تولیدات دامی

دوره ۲۰ ■ شماره ۱ ■ بهار ۱۳۹۷

صفحه‌های ۱۲۰-۱۰۹

شناسایی و آنالیز حذف و درج‌ها در ژنوم شترهای تک‌کوهانهٔ ایرانی با استفاده از داده‌های توالی‌یابی کل ژنوم

رضا خلخالی‌ایوریق[۱]، سیدحسن حافظیان[۲*]، نعمت هدایت‌ایوریق[۳]، ایوب فرهادی[۴]، محمدرضا بختیاری‌زاده[۵]

۱. دانشجوی دکتری، گروه علوم دامی، دانشکدهٔ علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران.

۲. دانشیار، گروه علوم دامی، دانشکدهٔ علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران.

۳. استادیار، گروه علوم دامی، دانشکدهٔ کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران.

۴. استادیار، گروه علوم دامی، دانشکدهٔ علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران.

۵. استادیار، گروه علوم دام و طیور، پردیس ابوریحان، دانشگاه تهران، پاکدشت، ایران.

تاریخ وصول مقاله:۱۳۹۶/۰۷/۲۴ تاریخ پذیرش مقاله: ۱۳۹۶/۱۰/۱۰

چکیده

این مطالعه به‌منظور شناسایی حذف و درج‌های (ایندل‌ها) موجود در ژنوم شتر تک‌کوهانهٔ ایرانی و ارزیابی گروه‌های عملکردی مرتبط با آن‌ها، با استفاده از داده‌های توالی‌یابی شدهٔ ژنوم کامل دو نفر شتر تک‌کوهانهٔ ایرانی (شتر یزدی و شتر طرودی) انجام شد. در این تحقیق، برای افزایش دقت و صحت ایندل‌های شناسایی‌شده، از دو برنامهٔ قدرتمند شناسایی واریانت (GATK و SAMtools) استفاده شد. درنهایت، تنوع‌های شناسایی‌شدهٔ مشترک بین دو برنامه، بعد از پالایش کیفی، ایندل‌های نهایی در نظر گرفته شدند. مطالعهٔ حاضر به شناسایی تعداد ۳۵۱۴۲۹ ایندل برای شتر یزدی و ۳۴۱۴۷۹ ایندل برای شتر طرودی انجامید. برای انجام آنالیزهای بیشتر، از ایندل‌هایی استفاده شد که در حاشیه‌نویسی، به‌عنوان واریانت‌های با اثر بالا طبقه‌بندی‌شده بودند. تعداد ایندل‌های با اثر بالا، به‌ترتیب برای شتر یزدی و طرودی برابر ۳۴۲۴ و ۳۵۰۶ بودند. برای انجام مقایسه بین نمونه شترهای ایرانی و شترهای غیرایرانی، از داده‌های توالی‌یابی کل ژنوم نمونه شتری با منشاء آفریقایی استفاده شد. مقایسهٔ ایندل‌های با اثر بالا بین سه نمونه، نشان داد که تعداد ۱۵۹۵ ایندل، بین آن‌ها مشترک هستند. ارزیابی نتایج ژن آنتولوژی نشان داد که بسیاری از عبارات معنادار شده، مربوط به توانایی شترها برای تحمل شرایط سخت بیابانی است.

کلیدواژه‌ها: ایندل، ایلومینا، تنوع ژنتیکی، ژن آنتولوژی، گروه‌های عملکردی.

* نویسنده مسئول Email: Scientific66correspond@gmail.com

مقدمه

ایران در یکی از خشک‌ترین مناطق جهان واقع شده است. میانگین سهم سالانهٔ ایران از نزولات آسمانی، در حدود ۲۵۲ میلی‌متر است که این مقدار، یک‌سوم میانگین بارش در سطح جهان است. براساس برنامهٔ توسعهٔ سازمان ملل، سهم سرانهٔ ایرانیان از منابع آبی کشور، از ۲۰۲۵ مترمکعب در سال ۱۹۹۰، به ۸۱۶ مترمکعب در سال ۲۰۲۵ خواهد رسید [۱۷]. شرایط اقلیمی و کمبود منابع آبی دردسترس ازیک‌طرف و نیاز حیاتی صنعت دامپروری به منابع آب شیرین از طرفی دیگر سبب شد تا پرورش گونه‌هایی از دام‌ها که پیشتر چندان توجهی به آن نمی‌شد، کانون توجه قرار گیرد. گونه‌هایی که سازگاری بالاتری با شرایط اقلیمی کشور دارند و دارای راندمان بالاتری در مصرف آب و منابع غذایی هستند. در بین دام‌های اهلی، شترها را می‌توان گزینهٔ مناسبی برای این هدف در نظر گرفت.

شترهای تک‌کوهانه یکی از گونه‌های مختص اقلیم‌های بیابانی و گرم‌وخشک هستند که به‌صورت کاملاً مناسب، با این شرایط سازگار شده و روند تکاملی مناسب با این مناطق را طی کرده‌اند. پراکنش این‌گونه، بیشتر در آفریقا و آسیا است [۲۳]. شترها در مناطق ذکرشده، بیشتر به‌منظور تهیهٔ غذا و استفاده در حمل‌ونقل، بهره‌برداری می‌شوند. توانایی بالای شترها در تحمل شرایط سخت بیابانی[۱۲] در کنار کاربردهای پزشکی این حیوان[۱-۲] ، آن‌ها را به گونهٔ جذابی برای مطالعه و تحقیق تبدیل کرده است. اما متأسفانه کم‌توجهی به این‌گونه و محبوس کردن آن‌ها در بیابان‌ها باعث شده تا در مقایسه با دام‌های دیگر (گاو، گوسفند، مرغ و اسب)، توجه کمتری به اصلاح‌نژاد این‌گونهٔ ارزشمند شود. البته رشد تکنولوژی‌ها و ابزارهای کمک‌کننده به اصلاح نژاد (تکنولوژی توالی‌یابی کل ژنوم، استفاده از آرایه‌هایSNP chip، مطالعات پویش ژنومی، ژنومیکس مقایسه‌ای و انتخاب ژنومیک) می‌تواند عقب‌ماندگی شترها در بخش اصلاح‌نژاد را با سرعت مناسبی جبران کند. به نظر می‌رسد که توالی‌یابی کل ژنوم شتریان ازجمله شترهای تک‌کوهانه[۱۲ و ۲۳]، دوکوهانه[۱۶] و آلپاکا[۲۳] در سال‌های اخیر، باعث افزایش تمایل محققان برای ورود به مطالعات ژنومی و اصلاح‌نژاد شترها شده باشد.

آنالیز تنوع ژنتیکی و چندشکلی‌ها در مقیاس بزرگ، درون یک گونه یا جمعیت، امکان شناسایی کاندیدهای جدید برای انتخاب را در اختیار ما قرار می‌دهند. همچنین شناخت دقیق‌تر ساختار ژنتیکی شترهای بومی ایران، امکان بهتری را برای حفظ ذخایر ژنتیکی این‌گونهٔ ارزشمند حیوانی در کشور فراهم می‌آورد. بسیاری از خصوصیات گونه‌های بومی، مختص همان مناطق بوده و حفظ و بهبود این خصوصیات منحصربه‌فرد منطقه‌ای، برای حفظ تنوع در سطح جهانی بسیار مهم است. شناسایی، حاشیه‌نویسی و تعیین اثر تنوع‌های ژنتیکی در سطح ژنوم می‌تواند کارایی بالایی در طراحی استراتژی‌های اصلاح‌نژادی شترها داشته باشد. ازاین‌رو در مطالعهٔ حاضر سعی شده است تا با استفاده از داده‌های مربوط به توالی‌یابی کل ژنوم دو نفر شتر تک‌کوهانهٔ ایرانی، به شناسایی، حاشیه‌نویسی و انجام آنالیزهای gene ontology (GO) مربوط به حذف و درج‌های (ایندل‌ها) ژنوم شترهای تک‌کوهانه پرداخته شود. شایان ذکر است که هدف این تحقیق، بیشتر بر نتایج مربوط به ایندل‌های با آثار بالا متمرکز شده است. به نظر می‌رسد که مطالعهٔ حاضر و مطالعاتی ازاین‌دست می‌توانند با انباشته‌کردن اطلاعات مربوط به تنوع‌های ژنتیکی شترها، به گسترش شناخت ما از این‌گونه و طراحی استراتژی‌های اصلاح‌نژادی مناسب کمک کنند.

مواد و روش‌ها

در این تحقیق، از نمونهٔ خون دو نفر شتر تک‌کوهانهٔ ایرانی به‌ترتیب از ایستگاه تحقیقاتی پرورش شتر استان یزد (شتر

یزدی) و ایستگاه پرورش و اصلاح‌نژاد شتر طرود در استان سمنان (شتر طرودی) برای توالی‌یابی کل ژنوم آن‌ها استفاده شد. نمونهٔ خون، از ورید گردنی حیوانات گرفته شد و تا زمان استخراج DNA در محیط آزمایشگاهی و در دمای ۲۰- درجهٔ سانتی‌گراد نگهداری شد. استخراج کل DNA با استفاده از کیت استخراج برای پستانداران (Real Biotech Corporation, RBC, South Korea) انجام پذیرفت. کنترل کمّی DNA استخراج‌شده با استفاده از دستگاه نانودراپ صورت گرفت و نسبت جذب طول موج‌های ۲۶۰ به ۲۸۰ به‌ترتیب برای شتر یزدی و شتر طرودی به‌عنوان معیار خلوص DNA گزارش شد. همچنین کنترل کیفی نمونه‌های استخراج‌شده DNA نیز با استفاده از ژل آگارز یک درصد انجام گرفت. توالی‌یابی کل ژنوم دو شتر مذکور با استفاده از تکنولوژی Hiseq 2000 شرکت ایلومینا (Illumina, San Diego, CA) در مرکز ژنوم دانشگاه آلبرتای کانادا انجام شد. توالی‌یابی به‌صورت Paired-end و در خوانش‌هایی با اندازه ۱۰۰ باز اجرا شد. برای تولید خوانش‌های زیاد و به‌منظور افزایش عمق توالی‌یابی، از دو کانال فلوسل برای توالی‌یابی کتابخانه ایجاد شده، استفاده شد. کنترل کیفی خوانش‌ها با استفاده از برنامهٔ FastQC [۱۰] صورت گرفت. به‌منظور حذف خوانش‌های با کیفیت پایین و بهبود کیفی داده‌ها برای انجام آنالیزهای بعدی، از برنامه Trimmomatic v0.36[۶] برای پالایش کیفی خوانش‌ها استفاده شد.

خوانش‌های پالایش‌شده، با استفاده از الگوریتم MEM برنامهٔ BWA v0.7.15 [۱۸] در ژنوم مرجع (شمار دسترسی Genbank: GCA-000767585.1) هم‌ردیف شدند. مرحلهٔ بعد از هم‌ردیف کردن خوانش‌ها در ژنوم مرجع، حذف خوانش‌های دو نسخه شده (Duplicated reads) است که این عملیات با استفاده از نرم‌افزار Picard صورت پذیرفت. در این مرحله، خوانش‌هایی حذف می‌شوند که در طی PCR، چند نسخه شده و در هنگام هم‌ردیفی، در جایگاه‌های یکسانی از ژنوم مرجع قرار می‌گیرند. در گام بعدی و برای کاهش خطای شناسایی ایندل‌ها، عملیات هم‌ترازی مجدد در اطراف ایندل‌ها با استفاده از برنامه GATK[۲۰] و در دو مرحله اجرا شد. در طی توالی‌یابی ژنوم، پلت‌فرم‌های توالی‌یابی دچار خطاهایی در برآورد نمره کیفی بازها می‌شوند که به نمره‌دهی بیش‌ازحد واقعی یا کمتر از آن به بازها می‌انجامد. مرحلهٔ BQSR با استفاده از الگوریتم‌های یادگیری ماشین، به شناسایی و تعدیل این‌گونه خطاها می‌پردازد. ازاین‌رو در آخرین گام قبل از شناسایی ایندل‌ها، کالیبراسیون مجدد نمرهٔ کیفی بازها (BQSR) نیز با استفاده از برنامهٔ GATK اجرا شد.

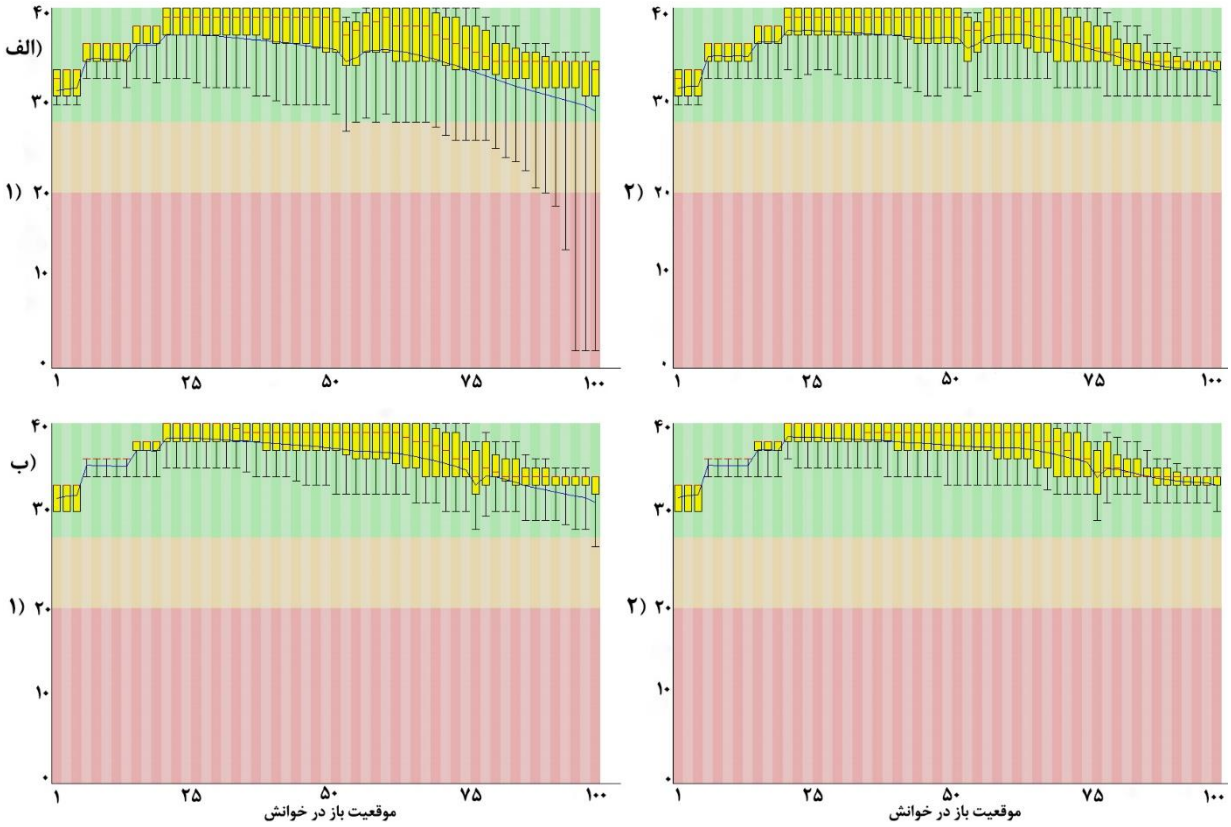
برای شناسایی ایندل‌های با دقت و صحت بالا، از دو نرم‌افزار قدرتمند در زمینهٔ شناخت تنوع‌های ژنتیکی، GATK و SAMtools [۱۹] استفاده شد. خروجی نرم‌افزارهای مذکور در قالب دو فایل VCF بود. به‌منظور پالایش و غربال کیفی ایندل‌های به‌دست‌آمده، از برنامه‌های BCFtools و VCFfilter استفاده شد. درنهایت دو فایل VCF به‌عنوان فایل‌های نهایی حاوی ایندل‌های شناسایی‌شده با صحت بالا در نظر گرفته شدند.

حاشیه‌نویسی ایندل‌های به‌دست‌آمده، توسط برنامه SnpEff [۸] انجام شد. درنهایت، آنالیز GO مربوط به ایندل‌های به‌دست‌آمده از مراحل ذکرشده، توسط نرم‌افزار برخط DAVID v6.8[۱۵] صورت پذیرفت. برای انجام این مرحله، نخست ژن‌های حاوی ایندل‌های با اثر بالا برای هر کدام از نمونه‌ها و همچنین برای ایندل‌های مشترک بین سه نمونه و مشترک بین دو نمونه ایرانی استخراج‌شده و سپس با استفاده از نرم‌افزار DAVID به انجام حاشیه‌نویسی عملکردی آن‌ها اقدام شد. مقادیر P-value تصحیح‌شده (تصحیح بنجامینی) کمتر از ۰/۱ به‌عنوان مقادیر معنادار برای نتایج حاصل از ژن آنتالوژی در نظر گرفته شد. شایان

ذکر است، به‌منظور مقایسه شترهای ایرانی با شترهای خارجی، داده‌های مربوط به توالی‌یابی شتری با منشاء آفریقایی از پایگاه اطلاعاتی NCBI (شماره دسترسی: SRX1013838) استخراج‌شده و تمامی مراحل صورت گرفته دربارهٔ نمونه‌های ایرانی، در مورد این شتر نیز انجام پذیرفت. درنهایت، نتایج به‌دست‌آمده برای مقایسه بین نمونه‌ها به کار گرفته شد.

نتایج و بحث

تعداد خوانش‌های حاصل از توالی‌یابی کل ژنوم به‌ترتیب برای شتر یزدی و طرودی برابر ۹۲۰۳۶۶۹۵۴ و ۸۴۳۴۵۵۱۴۴ به‌دست آمد که این تعداد، بعد از پالایش کیفی به‌ترتیب به ۸۹۹۷۱۴۱۰۲ (۹۷/۷۶ درصد خوانش‌های اولیه) و ۸۲۶۲۲۹۴۸۴ (۹۷/۹۶ درصد خوانش‌های اولیه) خوانش کاهش پیدا کرده و درنهایت باعث افزایش کیفیت میانگین خوانش‌ها به‌ویژه در انتهای خوانش‌ها شد (شکل ۱). بررسی هم‌ردیفی خوانش‌ها با ژنوم مرجع نشان داد که درصد خوانش‌هایی که با موفقیت در ژنوم مرجع هم‌ردیف شده‌اند، برای شتر یزدی ۹۷/۸ درصد و برای شتر طرودی ۹۷/۴ درصد هستند.



شکل ۱. کیفیت خوانش‌های مربوط به شتر یزدی قبل (الف ۱) و بعد از (الف ۲) پالایش کیفی و همچنین کیفیت خوانش‌های مربوط به شتر طرودی قبل (ب ۱) و بعد از (ب ۲) پالایش کیفی

جدول ۱. خلاصه اطلاعات مربوط به ایندل‌های شناسایی‌شده برای شترهای تک‌کوهانهٔ یزدی، طرودی و آفریقایی

	شتر یزدی	شتر طرودی	شتر آفریقایی
تعداد کل ایندل‌ها	۳۵۱۴۲۹	۳۴۱۴۷۹	۳۳۴۲۷۴
کل ایندل‌های هموزیگوت	۱۷۹۹۲۴	۱۷۴۸۶۴	۱۷۳۴۹۰
کل ایندل‌های هتروزیگوت	۱۷۱۵۰۵	۱۶۶۶۱۵	۱۶۰۷۸۴
درج هتروزیگوت/درج هموزیگوت	۱۰۴۸۷۱/۸۱۳۱۴	۱۰۲۵۴۲/۷۸۷۶۴	۱۰۱۰۱۳/۷۸۷۸۸
حذف هتروزیگوت/حذف هموزیگوت	۷۵۰۵۳/۹۰۱۹۱	۷۲۳۲۲/۸۷۸۵۱	۷۲۴۷۷/۸۱۹۹۶
نرخ درج/حذف	۱/۱۳ (۱۸۶۱۸۵/۱۶۵۲۴۴)	۱/۱۳ (۱۸۱۳۰۶/۱۶۰۱۷۳)	۱/۱۶ (۱۷۹۸۰۱/۱۵۴۴۷۳)

در مطالعهٔ حاضر، تعداد ایندل‌های شناسایی‌شده در ژنوم شتر یزدی ۳۵۱۴۲۹، در ژنوم شتر طرودی ۳۴۱۴۷۹ و برای شتر آفریقایی برابر با ۳۳۴۲۷۴ بود. بررسی میزان هتروزیگوسیتی ایندل‌های شناسایی‌شده در سه نمونهٔ ژنوم شتر، نشان داد که درصد ایندل‌های هتروزیگوت نمونهٔ آفریقایی (۴۸/۱۰ درصد) در مقایسه با نمونه‌های ایرانی (۴۸/۸۰ درصد برای یزدی و ۴۸/۷۹ درصد برای طرودی) اندکی کمتر است (جدول ۱). هرچند این تفاوت چندان بالا نیست، ولی تا حدودی می‌تواند بازتاب‌دهندهٔ نحوهٔ پرورش و اصلاح‌نژاد شترها در دو منطقهٔ جغرافیایی متفاوت باشد. سبک پرورش شتر در ایران به‌گونه‌ای است که اختلاط‌های ژنتیکی و جریان ژنی بیشتری در بین شترهای مناطق مختلف اتفاق می‌افتد. ازاین‌رو انتظار بر این است که میزان هتروزیگوسیتی در نژادهای ایرانی تا حدودی بالاتر باشد. نکتهٔ جالب توجه این است که نسبت ایندل‌های هتروزیگوت به هموزیگوت در حذف‌ها بیشتر از درج‌ها بود، به‌طوری‌که این نسبت در درج‌ها به ترتیب برای شترهای یزدی، طرودی و آفریقایی برابر با ۰/۷۷، ۰/۷۷ و ۰/۷۸ درصد به دست آمد درحالی‌که نسبت مذکور در حذف‌ها به‌ترتیب برابر با ۱/۲۰، ۱/۲۱ و ۱/۱۳ بود. احتمالاً این مسئله می‌تواند نشان‌دهندهٔ این امر باشد که ایندل‌های درج بیشتر از حذف‌ها تثبیت‌شده و به حالت هموزیگوت بروز کرده‌اند، هرچند اثبات این ادعا به مطالعات بیشتری نیاز دارد.

حاشیه‌نویسی ایندل‌های شناسایی‌شده نشان داد که این واریانت‌ها از نظر میزان تأثیر بر ژن‌ها و پروتیئن‌های حاصله، در سه طبقه‌بندی شامل ایندل‌های با تأثیر بالا، تأثیر متوسط و تأثیر پیراینده (Modifier) قرار دارند. ایندل‌های با تأثیر بالا، ایندل‌هایی هستند که با آثاری مانند تولید ناقص پروتئین، منجر به تغییر عملکرد آن می‌شوند. درحالی‌که ایندل‌های با آثار متوسط و پیراینده، چنین تأثیرات بالایی در پروتئین‌ها نمی‌گذارند. ممکن است این ایندل‌ها در قسمت‌های غیرکدکننده بروز کرده و یا تغییرات حاصل از این ایندل‌ها هم‌معنی باشند و در نتیجه به ساختار پروتئین‌ها آسیب وارد نکنند. هرچند ایندل‌ها مسئول ۲۴ درصد از بیماری‌های قابل توارث در انسان هستند [۲۴]؛ اما آثار آن‌ها همیشه جزء آثار منفی و مخرب دسته‌بندی نمی‌شود. ایندل‌ها در بعضی از مواقع آثار مثبت و سودمند نیز از خود نشان می‌دهند[۹]. مثالی از ایندل‌های سودمند را می‌توان در میکروب‌ها مشاهده کرد که تغییر قالب (Frame shift) باعث افزایش توانایی آن‌ها در پاسخ به آثار نامطلوب محیطی می‌شود[۲۲]. یکی دیگر از آثار مثبت ایندل‌ها ایجاد ماهیچه‌های دوگانه در گاو است که صفتی سودآور به‌شمار می‌رود[۱۴].

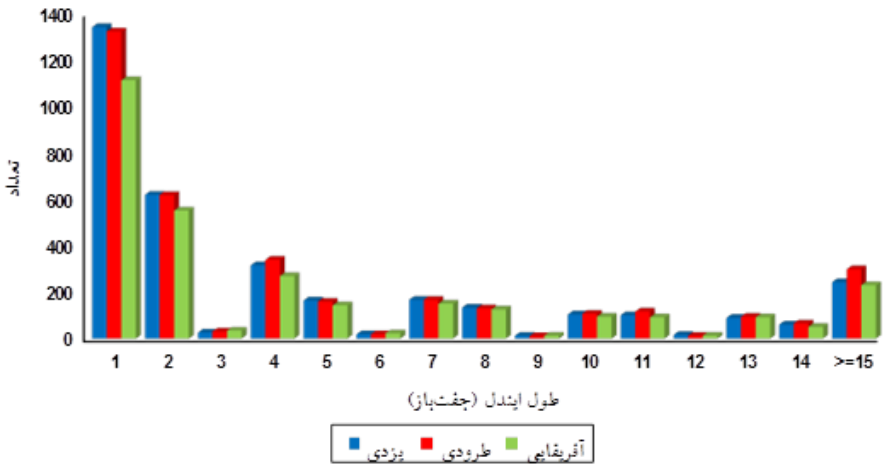
ایندل‌های دگرقالب، ایندل‌های ایجاد کنندهٔ ژن پیوندی (Gene fusion) و ژن پیوندی معکوس (Gene fusion reverse)، ایندل‌های ایجاد شده در ناحیه Splice site (دهنده و گیرنده) و ایندل‌های تغییردهنده کدون‌های آغازین و پایانی (Start_lost, Stop_gained و Stop_lost) در طبقه‌بندی ایندل‌های با تأثیر بالا قرار می‌گیرند. تقسیم‌بندی تعداد ایندل‌ها براساس آثارشان در جدول ۲ آمده است.

درصد اندکی از ایندل‌ها، دارای آثار بالایی هستند، به‌طوری‌که فقط ۰/۹۷ درصد از ایندل‌های شناسایی‌شده در ژنوم شتر یزدی، ۱/۰۳ درصد در شتر طرودی و ۰/۹۰ درصد در شتر آفریقایی در این دسته‌بندی قرار می‌گیرند. در میان ایندل‌های با تأثیر بالا، ایندل‌های دگرقالب دارای بیشترین سهم هستند و تعداد آن‌ها به‌ترتیب برای شتر یزدی، شتر طرودی و شتر آفریقایی برابر با ۳۲۵۷، ۳۳۲۷ و ۲۸۲۱ است. در هر سه نمونه مورد مطالعه، بیش از ۹۸ درصد از ایندل‌ها جزء ایندل‌های پیراینده بودند که آثار چشمگیری بر ژن‌ها و پروتئین‌ها ندارند، بنابراین ردیابی آثار آن‌ها بسیار دشوار به نظر می‌رسد.

بررسی نمودار مربوط به‌اندازهٔ ایندل‌های با آثار بالا (شکل ۲) نشان می‌دهد که اکثر ایندل‌ها، دارای طولی برابر یا کمتر از سه باز هستند، به‌طوری‌که ایندل‌های با اندازه‌های مذکور، به‌ترتیب ۵۷/۸۹، ۵۶/۳۶ و ۵۶/۶۹ درصد از کل ایندل‌های با آثار بالا در شترهای یزدی، طرودی و آفریقایی را به خود اختصاص دادند.

جدول ۲. تقسیم‌بندی تعداد ایندل‌های شناسایی‌شده براساس میزان تأثیر بر ژن‌ها و پروتئین‌ها

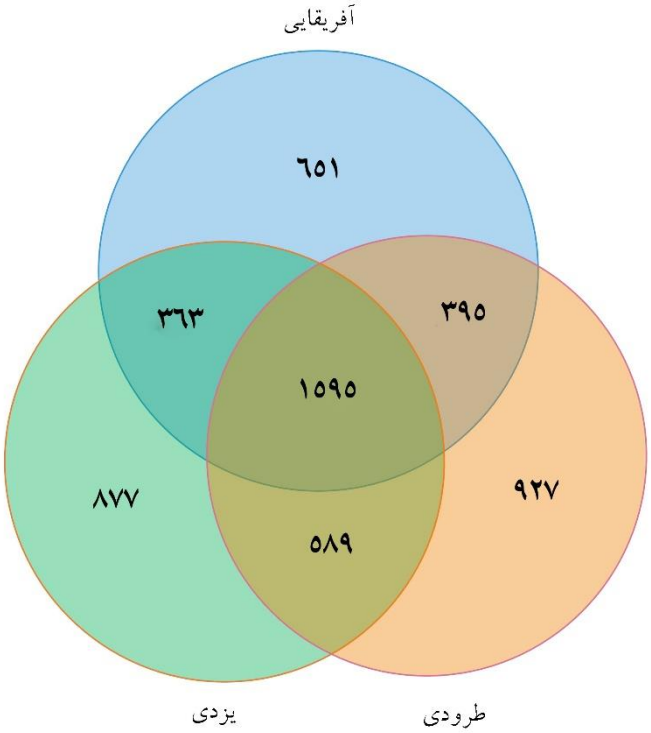
	یزدی	طرودی	آفریقایی
ایندل‌های با آثار بالا	۳۴۲۴ (۰/۹۷ درصد)	۳۵۰۶ (۱/۰۳ درصد)	۳۰۰۴ (۰/۹۰ درصد)
ایندل‌های با آثار متوسط	۱۰۴۸ (۰/۳۰ درصد)	۱۰۴۷ (۰/۳۱ درصد)	۹۴۱ (۰/۲۸ درصد)
ایندل‌های با آثار پیراینده	۳۴۶۹۵۷ (۹۸/۷۳ درصد)	۳۳۶۹۲۶ (۹۸/۶۶ درصد)	۳۳۰۳۲۹ (۹۸/۸۲ درصد)



شکل ۲. نمودار پراکنش اندازهٔ ایندل‌های با اثر بالا در ژنوم شترهای یزدی، طرودی و آفریقایی

نکتهٔ جالب توجه دیگری که در این نمودار به‌وضوح مشاهده می‌شود، فراوانی پایین ایندل‌هایی است که اندازهٔ آن‌ها ضریبی از عدد سه است. این امر نشان می‌دهد که ایندل‌هایی که اندازهٔ آن‌ها ضریبی از تعداد بازهای یک کدون (سه باز) نباشند، می‌توانند آثار بالاتری بر ساختار یک ژن و پروتئین حاصل از آن بگذارند. این مسئله به‌خصوص دربارهٔ ایندل‌های دگرقالب صدق می‌کند و ایندل‌هایی که اندازهٔ آن‌ها ضریبی از عدد سه نباشد، به‌راحتی قالب خوانش ژن را تغییر داده و درنهایت بر محصول آن اثر می‌گذارند.

مقایسهٔ ایندل‌های با آثار بالا در بین سه نمونهٔ مورد مطالعه، نشان داد که تعداد بسیاری از این نوع ایندل‌ها در هر سه نمونه مشترک هستند که احتمالاً به دلیل تثبیت این واریانت‌ها در بین شترهای جهان باشد. البته، در این زمینه برای اثبات این فرضیه به مطالعات بسیاری نیازمندیم. تعداد ۱۵۹۵ ایندل با اثر بالا در بین هر سه نمونه مشترک بودند که این تعداد در برگیرندهٔ ۴۶/۵۸ درصد، ۴۵/۴۹ درصد و ۵۳/۰۹ درصد از کل ایندل‌های با اثر بالا در شترهای یزدی، طرودی و آفریقایی بود. همان‌گونه که انتظار می‌رفت، تعداد ایندل‌های مشترک بین شترهای یزدی و طرودی (۵۸۹) بیشتر از تعداد ایندل‌های مشترک بین شترهای یزدی و آفریقایی (۳۶۳) و همچنین طرودی و آفریقایی (۳۹۵) بود. در این مطالعه، ۸۷۷ ایندل برای شتر یزدی، ۹۲۷ ایندل برای شتر طرودی و ۶۵۱ ایندل برای شتر آفریقایی به‌صورت اختصاصی بودند و هیچ اشتراکی با نژادهای دیگر نشان ندادند (شکل ۳).



شکل ۳. نمودار وِن مربوط به ایندل‌های با آثار بالا در شترهای یزدی، طرودی و آفریقایی

حاشیه‌نویسی عملکردی برای ایندل‌های شناسایی‌شده در ژنوم سه نمونه و همچنین دربارهٔ ایندل‌های مشترک بین آن‌ها و ایندل‌های مشترک بین دو نمونهٔ ایرانی اجرا شد. حاشیه‌نویسی عملکردی، ژن‌های ورودی را براساس فرایندهای زیستی، عملکردهای مولکولی و اجزای سلولی طبقه‌بندی می‌کند [۳]. نتایج نشان دادند که به‌ترتیب، ۲۱۳۱، ۲۱۵۱ و ۱۹۵۵ ژن در ژنوم شترهای یزدی، طرودی و آفریقایی دارای ایندل‌های با اثر بالا هستند. تعداد ۱۱۰۵ ژن از بین ژن‌های شناسایی‌شده در بین هر سه نمونه مشترک بودند. تعداد ژن‌های مشترک و اختصاصی بین دو نمونهٔ ایرانی برابر با ۳۶۲ ژن به دست آمد که در واقع این ژن‌ها در نمونهٔ آفریقایی دچار ایندل نشدند. ازاین‌بین، تعداد ۵۲۸ ژن برای شتر یزدی، ۵۵۲ ژن برای شتر طرودی و ۴۳۷ ژن برای شتر آفریقایی اختصاصی بودند.

تعداد عبارت GO معنادار (P-value تصحیح‌شده کمتر از ۰/۱) برای فرایندهای زیستی در شتر یزدی ۳۳، در شتر طرودی ۴۵ و در شتر آفریقایی ۱۰۰ بود. دربارهٔ شتر یزدی، ۱۱ عبارت از ۳۳ عبارت معنادار شده (۳۳/۳ درصد) به‌طور مستقیم مربوط به فرایندهای توسعه‌ای در مراحل مختلف زندگی حیوان می‌شدند. شش عبارت (۱۸/۲ درصد) نیز مربوط به توسعه و شکل‌گیری نورون‌ها و سیستم عصبی در شتر یزدی بودند. عبارات معنادار جالب توجه دربارهٔ شتر یزدی شامل تنظیم انتقال یون فلزی (GO:0010959)، انتقال یون فلزی (GO:0030001)، سازمان‌دهی سیتواسکلتون (GO:0007010) و تنظیم انتقال غشایی کاتیونی (GO:1904062) بودند که در رابطه با توانایی کلیه‌ها برای حفظ هر چه بیشتر آب و کنترل غلظت یون‌هایی از قبیل مس (Cu^{2+})، روی (Zn^{2+}) و آهن (Fe^{2+}) در بدن است. سیتواسکلتون‌ها در توسعه و شکل‌گیری کلیه‌ها دخیل هستند [۴]. کلیه‌ها به دلیل درگیری مستقیم در تنظیم غلظت و تعادل یون‌ها و الکترولیت‌های مهم و همچنین تنظیم میزان خروجی آب از طریق ادرار، برای شترها که در شرایط کم آب بیابان زندگی می‌کنند اندام بسیار مهمی به شمار می‌رود.

شتر طرودی نیز نتایجی مشابه با شتر یزدی را نشان داد؛ اما سهم عبارات GO مرتبط با اندام کلیه در این شتر بیشتر از شتر یزدی بود. همچنین عبارتی در رابطه با پروتئین‌های Ras (GO:0007265) نیز در نتایج حاصل از حاشیه‌نویسی عملکردی این نمونه مشاهده شد. پروتئین‌های Ras دست‌هایی از پروتئین‌های کوچک هستند (وزن مولکولی: تقریباً ۲۱۰۰۰ دالتون) که در حضور ترکیب GDP غیرفعال بوده ولی با اتصال به GTP به حالت فعال در می‌آیند. مطالعات نشان می‌دهد که این پروتئین‌ها جزء مهمی در سیستم علامت‌دهی سلول‌های ایمنی هستند و در تنظیم تمایز سلولی، بلوغ سلولی و فعالیت انواع سلول‌های دخیل در ایمنی مشارکت دارند [۲۱].

دربارهٔ شتر آفریقایی علاوه بر معنادار شدن عباراتی در زمینهٔ توسعه و رشد اندام‌ها و همچنین توسعه و عملکرد کلیه‌ها، عباراتی دربارهٔ سیستم ایمنی و پاسخ به محرکات بیماری‌زا نیز مشاهده شد که نتایج جالبی به نظر می‌رسند. ازجمله می‌توان به عباراتی مرتبط با تولید و تنظیم مثبت تولید سیتوکین (GO:0001819، GO:0001817 و GO:0001816) اشاره کرد. سیتوکین‌ها علاوه بر نقشی که در سیستم ایمنی بدن دارند [۱۳]، در پاسخ به استرس اکسیداتیو نیز ایفای نقش می‌کنند. استرس اکسیداتیو در شرایطی بروز می‌کند که عدم تعادلی بین مقدار گونه‌های واکنش‌پذیر اکسیژن (ROS) و توانایی بدن برای داکسیده کردن این ترکیبات به وجود آید[۵]. شایان ذکر است که ترکیبات ROS در شرایط پاسخ به انواع استرس‌ها ازجمله شوک گرمایی افزایش پیدا می‌کنند و افزایش آن‌ها تا حد خاصی می‌تواند به تحریک تولید پروتئین‌های شوک حرارتی کمک‌کننده باشد اما تجمع بیش‌ازحد آن‌ها می‌تواند

برای سلول‌ها، مرگ‌آور باشد[۱۱]. شوک گرمایی در شترها به دلیل زندگی در شرایط بیابانی می‌تواند مشکلی جدی به شمار بیاید و ازاین‌رو مشاهدهٔ عبارات معنادار در زمینهٔ تولید سیتوکین، حاکی از توسعهٔ ژن‌هایی است که برای تحمل این شرایط به شترها کمک می‌کنند. در عبارات معنادار شده برای فرایندهای زیستی شتر آفریقایی، چندین عبارت در ارتباط با قلب و سیستم خون‌رسانی (GO:0060047، GO:0003015، GO:0072359 و GO:0072358) نیز وجود دارد که به نظر می‌رسد تفسیر این عبارات به مطالعات فیزیولوژیکی بیشتر، دربارهٔ شترها نیاز داشته باشد.

در مطالعه‌ای [۲۳]، به گروهی از شترها، ۲۴ روز محدودیت دسترسی به آب داده شد و بعد از این مدت، ژن‌های بیان‌شده در بخش کورتکس و مدولای کلیه این حیوانات بررسی شد و در مقایسه با گروه شاهد، ژن‌های با افزایش بیان و کاهش بیان مشخص شدند. سپس آنالیز GO برای این ژن‌ها صورت گرفت. مقایسهٔ عبارات GO به‌دست‌آمده در مطالعهٔ حاضر با تحقیق مذکور نتایج جالبی را مشخص کرد. نتایج مقایسه‌ها نشان داد که ۲۳ عبارت از ۳۳ عبارت به‌دست‌آمده (۶۹/۷ درصد) در شتر یزدی با مطالعهٔ ذکرشده دارای اشتراک هستند. این تعداد برای شتر طرودی برابر با ۳۱ عبارت از ۴۵ عبارت به‌دست‌آمده (۶۸/۹ درصد) و برای شتر آفریقایی نیز ۵۸ عبارت از ۱۰۰ عبارت به‌دست‌آمده (۵۸ درصد) بود. این شواهد به‌احتمال فراوان می‌تواند نشان‌دهندهٔ این امر باشد که ژن‌های با بیان بالا در شرایط کم‌آبی در مسیر تکامل خود دچار تغییرات شده‌اند. مطالعهٔ ژن‌های شناسایی‌شده در تحقیق کنونی که با مطالعهٔ مذکور[۲۳]مشترک بودند، می‌توانند به درک بهتر ما از روند تکاملی شترها در مسیر عادت کردن به محیط بیابانی، کمک‌کننده باشند.

نتایج نشان داد که هیچ عبارت GO معناداری برای ژن‌های مشترک بین دو نمونهٔ ایرانی وجود ندارد. اما در آنالیز ژن‌های مشترک بین سه نمونهٔ مورد مطالعه، در بخش فرایندهای زیستی، سه عبارت GO (GO:0032502، GO:0044707 و GO:0007275) معنادار شدند. عبارات معنادار شده برای عملکردهای مولکولی (برای مثال: GO:0001882، GO:0032549، GO:0032555، GO:0005524 و GO:0017076) دربارهٔ ژن‌های مشترک، حاکی از اهمیت سوخت‌وساز مرتبط با تولید مادهٔ ژنتیکی (DNA یا RNA) در این حیوانات است. زندگی در شرایط نامساعد و پراسترس بیابان و قرار گرفتن در معرض عواملی مانند اشعهٔ فرابنفش خورشید و میزان نمک بالا (در گیاهانی که غذای شترها هستند) می‌تواند باعث افزایش تخریب DNA شده [۷] و در نتیجه، نیاز به تولید و تعمیر این مواد حیاتی را بالا ببرد. ازاین‌رو شناسایی عباراتی که مرتبط با این مسئله هستند می‌تواند تأییدی بر این ادعا باشد. در میان ژن‌های مشترک بین سه نمونه، ژن‌هایی ازجمله ATF6، STOX1، ATF6B، TMBIM6 و ANKRD2 وجود دارند که در پاسخ به استرس‌های محیطی نقش داشته و با ایجاد تغییراتی در میزان رونویسی از قسمت‌هایی از DNA به تعیین پاسخ مناسب به محرک‌ها و استرس‌ها کمک می‌کنند[۳].

نتایج به‌دست‌آمده در این مطالعه نشان داد که وقوع ایندل‌ها، بیشتر در ژن‌هایی رخ داده است که در عادت‌پذیری شترها به شرایط بیابانی دخیل بوده‌اند. ژن‌های دخیل در توسعه و عملکرد کلیه‌ها، سیستم ایمنی، سیستم همانندسازی و تعمیر DNA و همچنین سیستم‌های پاسخ به محرک‌های خارجی ازجمله بخش‌هایی هستند که توسط ایندل‌ها دست‌خوش تغییرات شده‌اند. نبود اطلاعات فنوتیپی قابل اعتماد در شترها باعث شد تا مطالعهٔ شناسایی ارتباط بین این ایندل‌ها و صفات فنوتیپی انجام نپذیرد. ازاین‌رو در مطالعات پیش‌رو، نیاز به اطلاعات فنوتیپی

احساس می‌شود و وجود داده‌هایی ازاین‌دست می‌تواند بـه پیشبرد دانسته‌های محققان از شترها کمک شایانی کند. این مطالعه در جهان جزء نخستین قدم‌هـا در مسـیر اسـتفاده از داده‌های توالی‌یابی کل ژنوم شتر بـرای شناسـایی تنوعـات ژنتیکی است. ورود محققان ایرانی به این زمینه می‌تواند بـه شناسـایی قابلیـت‌هـای ژنتیکـی شـترهای ایرانـی و برنامه‌ریزی‌های دقیق اصلاح نژادی بیانجامد.

منابع

[1]. Agrawal RP, Jain S, Shah S, Chopra A and Agarwal V (2011) Effect of camel milk on glycemic control and insulin requirement in patients with type 1 diabetes: 2-years randomized controlled trial. European journal of clinical nutrition. 65(9): 1048-1052.

[2]. Al-Yousef N, Gaafar A, Al-Otaibi B, Al-Jammaz I, Al-Hussein K and Aboussekhra A (2012) Camel urine components display anti-cancer properties in vitro. Journal of ethnopharmacology. 143(3): 819-825.

[3]. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT and Harris MA (2000) Gene Ontology: tool for the unification of biology. Nature genetics. 25(1): 25-29.

[4]. Bacallao R (1995) The role of the cytoskeleton in renal development. In Seminars in nephrology. 15: 285-290.

[5]. Bhattacharya S (2015) Reactive oxygen species and cellular defense system. Free Radicals in Human Health and Disease. 17-29.

[6]. Bolger AM, Lohse M and Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15): 2114-2120.

[7]. Burg MB, Ferraris JD and Dmitrieva NI (2007) Cellular response to hyperosmotic stresses. Physiological Reviews. 87(4): 1441-1474.

[8]. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X and Ruden DM (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(2): 80-92

[9]. Das S, Upadhyaya HD, Srivastava R, Bajaj D, Gowda CLL, Sharma S, Singh S, Tyagi AK and Parida SK (2015) Genome-wide insertion–deletion (InDel) marker discovery and genotyping for genomics-assisted breeding applications in chickpea. DNA Research. 22(5): 377-386.

[10]. FastQC: a quality control tool for high throughout sequence data (2016). http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/

[11]. Fedyaeva AV, Stepanov AV, Lyubushkina IV, Pobezhimova TP and Rikhvanov EG (2014) Heat shock induces production of reactive oxygen species and increases inner mitochondrial membrane potential in winter wheat cells. Biochemistry (Moscow). 79(11): 1202-1210.

[12]. Fitak RR, Mohandesan E, Corander J and Burger PA (2016) The de novo genome assembly and annotation of a female domestic dromedary of North African origin. Molecular ecology resources. 16(1): 314-324.

[13]. Ganesh BB, Bhattacharya P, Gopisetty A and Prabhakar BS (2011) Role of cytokines in the pathogenesis and suppression of thyroid autoimmunity. Journal of Interferon and Cytokine Research. 31(10): 721-731.

[14]. Grobet L, Martin LJR, Poncelet D, Pirottin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Ménissier F, Massabanda J and Fries R (1997) A deletion in the bovine myostatin gene causes the double–muscled phenotype in cattle. Nature genetics. 17(1): 71-74.

[15]. Huang DW, Sherman BT and Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 4(1): 44-57.

[16]. Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z, Zhang H, Wang L, Hasi S, Zhang Y, Li J and Shi Y (2012) Genome sequences of wild and domestic bactrian camels. Nature Communications. 3: 1202-1210.

[17]. Lehane S (2014) The Iranian water crisis. Strategic Analysis Paper. Future Directions International.

[18]. Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 25(14): 1754-1760.

[19]. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics. 25(16): 2078-2079.

[20]. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M and DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research. 20(9): 1297-1303.

[21]. Olson MF and Marais R (2000) Ras protein signalling. Seminars in immunology. 12: 63-73.

[22]. Van-Der Woude MW and Bäumler AJ (2004) Phase and antigenic variation in bacteria. Clinical Microbiology Reviews. 17(3): 581-611.

[23]. Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, Zhang L, Abutarboush MH, Xing Y, Xie Z and Alshanqeeti AS (2014) Camelid genomes reveal evolution and adaptation to desert environments. Nature communications. 5: 5188.

[24]. Yan Y, Yi G, Sun C, Qu L and Yang N (2014) Genome-wide characterization of insertion and deletion variation in chicken using next generation sequencing. PloS one. 9(8): p.e104652.

Journal of
Animal Production
(College of Abouraihan – University of Tehran)
Vol. 20■ No. 1 ■ Spring 2018

Detection and analysis of deletions and insertions in genome of Iranian dromedary camels using whole genome sequencing data

Reza Khalkhali Ivriq[1] , Seyed Hasan Hafezian[2*], Nemat Hedayat Ivriq[3] , Ayoub Farhadi[4] , Mohammad Reza Bakhtiarizadeh[5]

1. Ph.D. Student, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran
2. Associate Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran
3. Assistant Professor, Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardebil, Iran
4. Assistant Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran
5. Assistant Professor, Department of Animal and Poultry Science, College of Aburaihan, University of Tehran, Pakdasht, Iran

Received: October 16, 2017 **Accepted: December 31, 2017**

Abstract

This study was carried out for identification of deletions, insertions (INDELs) and assessment of their related functional groups in two Iranian dromedary camels (Yazdi camel and Trodi camel) using whole genome sequencing data. In this study, two powerful variant callers (GATK and SAMtools) were used to increase precision and accuracy of detected INDELs. Finally, common identified variants between two programs, after quality filtration, were considered as final INDELs. The present study led to identification of 351429 INDELs for Yazdi camel and 341479 INDELs for Trodi camel. Annotated INDELs that classified as high impact INDELs were used for further analysis. The numbers of high impact INDELs were 3424 and 3506 for Yazdi and Trodi camels, respectively. To compare Iranian camels with non-Iranian camels, we used whole genome sequencing data of one African origin camel. Comparison of high impact INDELs between three samples showed that 1595 INDELs were common between them. Assessment of gene ontology's results showed that many of significant terms are related to the ability of camels to withstand serve desert conditions..

Keywords: functional groups, gene ontology, genetic diversity, illumina, INDEL..

Email: Scientific66correspond@gmail.com