

شناسایی تنوع در تعداد کپی قطعه ژنومی 15 نژاد گوسفند ایتالیایی با استفاده از چپ 50 کی گوسفندی

مریم نصرتی^{1*} - مجتبی طهمورث‌پور²

تاریخ دریافت: 1393/11/04

تاریخ پذیرش: 1394/05/26

چکیده

در سال‌های اخیر با پیشرفت تکنولوژی توالی‌یابی، مطالعات ژنومی به سمت شناسایی تنوع‌های جدید و نحوه تأثیر آن‌ها بر عملکرد صفات اقتصادی معطوف شده است. تنوع در تعداد کپی قطعه ژنومی با طول حداقل یک کیلو جفت‌باز و تشابه بیش از 90 درصد از جمله این تنوع‌ها است. در این پژوهش به هدف شناسایی این تنوع در ژنوم گوسفند از 360 رأس گوسفند نمونه خون تهیه شد و با استفاده از چپ 50 کی گوسفندی شرکت ایلومینا، 54241 جهش تک‌نوکلئوتیدی در سراسر ژنوم شناسایی شد. تعداد 2180 تنوع در تعداد کپی قطعه ژنومی با استفاده از نرم‌افزار پنسی‌ان‌وی شناسایی شد. کل طول این تنوع در ژنوم 8/5 مگا جفت‌باز بود (0/32 درصد). پس از ادغام نواحی مشترک 35 ناحیه تنوع در تعداد کپی قطعه ژنومی پُر تکرار در کروموزوم‌های اتوزوم با میانگین و میانه طول به ترتیب 178/61 و 135/25 کیلو جفت‌باز شناسایی شد. توزیع نواحی تنوع در ژنوم بیشتر در مناطق تلومری و سانتومری بود. از 35 ناحیه شناسایی شده 25 ناحیه به صورت کامل یا جزئی با 152 ژن مرجع در ژنوم انسان، گاو و گوسفند هم‌پوشانی داشت. این ژن‌ها در مسیرهای متابولیسمی ترانس‌دوکسیون بویای، سرطان، متابولیسم پورین و پریمیدین، بیوسنتز استروئیدها و بیماری‌های از قبیل استئوپوروسیز و کم‌خونی نقش داشتند. حضور این نواحی در مناطق ژنی این فرضیه را تقویت کرد که احتمالاً این تنوع قادر است مسیرهای متابولیسمی را تغییر دهد و منجر به تفاوت در عملکرد فنوتیپی بین افراد یک جمعیت شود. لذا به نظر می‌رسد شناسایی این تنوع می‌تواند افق‌های جدیدی را در برنامه‌های اصلاح‌نژادی باز کند.

واژه‌های کلیدی: پنسی‌ان‌وی، تنوع در تعداد کپی قطعه ژنومی، چپ 50 کی گوسفندی، گوسفند.

مقدمه

فنوتیپی بایستی بیش از جهش‌های تک‌نوکلئوتیدی باشد. در پستانداران این تنوع 25-4/0 درصد کل تنوع ژنوم را دربر می‌گیرد (1)، 9، 15 و 20).

اولین مطالعات بر روی این تنوع در انسان در سال 2006 صورت گرفت. تاکنون 29000 مورد تنوع در تعداد کپی قطعه ژنومی در ژنوم انسان شناسایی شده است که 9000 مورد آن، در نزدیکی یا داخل یک ژن قرار دارد. به‌طور کلی، 12 درصد تنوع ژنوم انسان ناشی از تنوع در قطعه کروموزومی است و تقریباً 40 درصد این ساختارها حداقل با یک ژن هم‌پوشانی دارد. این نوع تنوع با بیماری‌هایی از قبیل شیذوفرنی، دیابت نوع اول و سرطان و غیره در انسان مرتبط است (26، 4، 15، 11، 14، 22). مطالعات اخیر نشان می‌دهد تنوع در تعداد کپی قطعه ژنومی در حیوانات مزرعه‌ای با تنوع فنوتیپی صفات مورفولوژیکی مرتبط است. مثلاً تاج نخودی در طیور ناشی از تنوع در تعداد کپی قطعه ژنومی در اینترون یک ژن *SOX5* می‌باشد. در گوسفند دو برابر شدن در ژن *ASIP* منجر به تنظیم رنگدانه در پوشش بدن می‌شود. تاخیر در پردآوری جوجه‌ها به دلیل دوبره ناقص در ژن‌های *PRLR* و

اخیراً مطالعات ژنومی در حیوانات اهلی بر شناسایی تنوع‌های جدید در ژنوم و اثر آن در عملکرد فنوتیپی صفات اقتصادی متمرکز شده است. انواع تنوع ژنومی شامل حذف، اضافه، جایگزینی، بازترتیبی، معکوس شدن و تنوع در تعداد کپی قطعه ژنومی است. تنوع در تعداد کپی قطعه ژنومی شامل حذف، اضافه یا دو برابر شدن قطعه‌ای از DNA به طول یک کیلو جفت‌باز تا چندین مگا جفت‌باز است که شباهت میان این نواحی بیشتر از 90 درصد است. مطالعات نشان می‌دهد که این تنوع می‌تواند منجر به تغییر در دُر ژن، توالی‌های کُدکننده و تنظیم بیان ژن شود لذا، تأثیر آن بر تنوع

1- استادیار، گروه کشاورزی، دانشگاه پیام نور، صندوق پستی 3697-19395، تهران، ایران،

2- استاد گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد.

(*-نویسنده مسئول: (Email: m_nosrati@pnu.ac.ir)

DOI: 10.22067/ijasr.v0i0.43648

ایلو مینا³ برای تعیین ژنوتیپ 54241 جهش تکنوکلوتیدی در سراسر ژنوم استفاده شد. نمونه‌های DNA با غلظت 400-200 نانوگرم با محلول 0/1 نرمال سود دناتورده شده و سپس با محلول MA2 (کیت ایلومینا) خنثی شده تا برای تکثیر آماده شود. در مرحله بعد محلول MSM⁴ که حاوی مسترمیکس است به DNA اضافه شده و سپس پلیت در آن هیبریداسیون در دمای 37°C به مدت 24-20 ساعت انکوبه شد تا تکثیر صورت گیرد. این روش موجب تکثیر DNA به طور یکنواخت تا چندین برابر و بدون آریب (خطا) خواهد شد. DNA توسط سیستم آنزیمی تحت کنترل به قطعات کوچکتر بریده شد. DNA هضم شده با 2- پروپانول 100 درصد مخلوط و با سانتریفیوژ در دمای چهار درجه سانتی‌گراد رسوب داده شد و دوباره با محلول بافر هیبریداسیون سوسپانسیون شد. در مرحله بعد، نمونه‌های DNA بر روی چیپ قرار داده شد و به مدت 24-16 ساعت در دمای 48°C در آن هیبریداسیون نگهداری شد. در این مرحله از طریق سیستم مویرگی موجود در اطراف چیپ و تزریق محلول‌های هیبریداسیون، ابتدا DNA دناتورده شده و سپس هیبرید می‌شود. پس از هیبریداسیون، شستشوی چیپ برای حذف DNA هیبرید نشده انجام شد. نمونه‌ها برای بسط تک‌بازی و رنگ‌آمیزی آماده شد. به دلیل این که طول پرایمرهای (یک جفت) متصل به آرایه‌دانه⁵ تا یک نوکلئوتید قبل از محل جهش است، پس از اتصال DNA تک‌رشته (به‌عنوان الگو) به پروب 12 مری (به‌عنوان پرایمر)، بسط تک‌بازی در انتهای 3' پروب انجام می‌شود. بدین منظور محلول TEM در اطراف چیپ رها شده تا با استفاده از نوکلئوتیدهای نشان‌دار (فلورسنس) تکثیر صورت گیرد. در این محلول آدنین و تیمین به رنگ قرمز و سیتوزین و گوانین به رنگ سبز نشان‌دار شده است. پس از این مرحله DNA الگو توسط محلول فرمامید 95 درصد بر یک میکرومول EDTA حذف شد. در این حالت پرایمر بسط یافته متصل به آرایه‌دانه، تحت فرآیندهای مختلف قرار گرفت تا برای اسکن آماده شود. چیپ‌ها با استفاده از بید آرایه ریدر⁶ تحت لیزر قرار گرفت تا فلورسنس تک‌باز انتهای پرایمرهای بسط داده شده در آرایه‌دانه برانگیخته شود. اگر لوکاس موردنظر فاقد جهش (تیمین/آدنین یا سیتوزین/گوانین) باشد رنگ قرمز یا سبز ساطع می‌کند، اما اگر جهش صورت گرفته باشد، یعنی در مقابل آدنین مثلاً سیتوزین باشد رنگ زرد (ترکیب قرمز و سبز) ساطع می‌گردد. نور ساطع شده ثبت شد و داده‌های به‌دست آمده از این تصویر با بیواستدیو (نسخه 1)⁷ تصحیح شد. بدین معنی که گراف

SPEF2 می‌باشد. دو برابر شدن 4/6 کیلو جفت‌باز در اینترون شش ژن *stx17* منجر به خاکستری شدن رنگ بدن در اسب با افزایش سن می‌شود. حذف ناحیه بین ژنی به طول 11/7 کیلو جفت‌باز در ژنوم بز منجر به حذف شاخ می‌شود (24، 6، 27، 25).

با وجود مطالعات گسترده در زمینه شناسایی و توالی‌یابی این تنوع بر بیماری‌ها در انسان، مطالعه بر روی نحوه توزیع و توالی‌یابی آن بر صفات مهم اقتصادی در حیوانات اهلی در مراحل ابتدایی است. این امر ناشی از کمبود ابزارهای ژنومی در حیوانات اهلی بوده است. اولین مطالعه در این زمینه برای شناسایی تنوع در تعداد کپی قطعه ژنومی گاو توسط لیو و همکاران (18) با aCGH¹ (هیبریداسیون مقایسه‌ای ژنومی) صورت گرفت. آن‌ها توانستند 25 تنوع در تعداد کپی قطعه ژنومی را در سلول‌های جنسی شناسایی کنند. با توسعه تکنولوژی، ساخت چیپ‌های شناسایی جهش‌های تکنوکلوتیدی ژنوم، شناسایی این تنوع با دقت و سرعت بیشتری در سایر حیوانات مزرعه‌ای مانند جوجه گوشتی، بز و سگ انجام گرفته است (29، 9، 23). اولین مطالعه برای شناسایی این تنوع در گوسفند توسط فوتانزی و همکاران (8) صورت گرفت. این مطالعه با استفاده از aCGH با به‌کارگیری پروب گاو در شش نژاد گوسفند انجام شد و 135 ناحیه تنوع به طول 10/5 مگا جفت‌باز شناسایی شد. با توجه به پوشش قابل توجه این تنوع در ژنوم و اثر آن بر تنوع فنوتیپی صفات، اهمیت مطالعه بیشتر روی این نوع تنوع آشکارتر می‌شود. لذا در این پژوهش تنوع در تعداد کپی قطعه ژنومی در 360 نمونه از 15 نژاد مختلف گوسفند ایتالیایی با استفاده از چیپ 50 کی گوسفندی مورد بررسی قرار گرفت.

مواد و روش‌ها

نمونه‌گیری و تعیین ژنوتیپ

نمونه خون از نژادهای مختلف گوسفند شامل Altamura، Delle Biellese، Bergamasca، Bagnolese، Appenninica، Istrian Pramenka، Gent.le di Puglia، Fabrianese، Langhe، Sambucana-Mnifest، Pinzirita، Massese، Laticauda، Valle del Belice، Sopravissana، که بومی کشور ایتالیا بودند با متوسط 25 نمونه از هر نژاد تهیه شد. برخی نمونه‌ها قبل از تعیین ژنوتیپ با چیپ موردنظر به دلایل مختلف مانند کیفیت پایین DNA استخراج شد حذف شد و در پایان 360 نمونه باقی ماند. استخراج DNA از خون با استفاده از کیت استخراج DNA پرومگا² در آزمایشگاه ژنتیک مرکز ژنومیک حیوانات اهلی، دانشگاه بولونیا کشور ایتالیا انجام شد. در این پژوهش چیپ 50 کی گوسفندی شرکت

3-Illumina BovineSNP50 BeadChip (Illumina, Inc., USA)

4- Multi-Sample Amplification Master Mix

5- Bead Array

6- Bead Array reader

7- Illumina's BeadStudio v1.0

1-Array-comparative genomic hybridization

2-Wizard® Genomic DNA Purification kit (Promega Corporation, Madison, WI)

فایل شدت سیگنال ورودی به فایل‌های کوچکتر شکسته شد که هر فایل حاوی اطلاعات فوق برای هر فرد به‌طور مجزا است. فایل فراوانی آللی B در جمعیت (PFB) بر اساس بی.ا.ا.اف. هر نشانگر در جمعیت محاسبه شد. محل و موقعیت جهش تک‌نوکلئوتیدی‌ها روی کروموزوم از فایل PFB اخذ می‌شود، لذا در صورت تغییر در اسمبلی ژنوم تعیین محل جهش‌ها به‌راحتی امکان‌پذیر است. فایل جی-سی مدل بر اساس محتوی جی-سی یک مگا جفت‌باز از ژنوم ساخته شد. فایل مدل پنهان مارکوف (چ.ام.ام)⁷ مدل اچ.ام.ام را برای نرم‌افزار پنسی.ان.وی. فراهم می‌کند و مقدار شدت سیگنال مورد انتظار برای تنوع‌های مختلف را به برنامه القاء می‌کند. با توجه به این‌که مدل اچ.ام.ام برای همه ژنوم‌ها تقریباً یکسان است از فایل hh550.hmm که برگرفته از چپ انسان 550 کی است، استفاده شد. به این ترتیب تنوع در تعداد کپی قطعه ژنومی با استفاده از مدل زنجیره مارکوف در نرم‌افزار پنسی.ان.وی محاسبه شد (28). پارامتر سی.ان⁸ گزارش شده توسط نرم‌افزار پنسی.ان.وی در کروموزوم اتوزوم برای نواحی که فاقد تکرار هستند برابر 2 (موجود دیپلوئید) برای نواحی که حذف صورت گرفته برابر یک یا صفر، بسته به میزان حذف و برای نواحی دو یا سه برابر شده معادل به ترتیب سه و چهار است. پس از محاسبه تنوع در تعداد کپی قطعه ژنومی فیلترهای کنترل کیفیت نرم‌افزار پنسی.ان.وی اعمال شد. نمونه‌های با انحراف استاندارد ال.ا.ا.ر. کمتر از 0/03، بی.ا.ف.ی برابر با حداقل 0/01 و مقدار فاکتور امواج⁹ بین 0/05 تا 0/05- انتخاب شدند. تصحیح برای محتوی جی-سی با استفاده از فایل جی-سی مدل انجام شد. پس از اعمال فیلترهای موردنظر تعداد 184 نمونه باقی ماند. داده‌ها بر اساس اسمبلی Ovis_aris_3.0 ژنوم گوسفند برای بررسی تنوع ساختار ژنوم آنالیز شد.

مطالعات بیوانفورماتیکی

برای مقایسه نواحی تنوع در تعداد کپی قطعه ژنومی شناسایی شده در ژنوم گوسفند، با مطالعات قبلی در ژنوم گاو از ابزار liftOver وبسایت UCSC استفاده شد و نواحی موردنظر از ژنوم گوسفند¹⁰ به معادل آن در ژنوم گاو¹¹ تبدیل شد. به‌علاوه، برای مطالعه محتوی ژنی نواحی شناسایی شده، محدوده این نواحی در وبسایت UCSC با استفاده از ابزار table browser برای گونه‌های گوسفند، گاو و انسان مورد ارزیابی قرار گرفت. جهت بررسی عملکرد ژن‌های این نواحی، حدود آن وارد وبسایت دیوید¹² شد (10) و توالی‌یابی این ژن‌ها در

تشکیل شده از جهش‌ها برای کلیه نمونه‌ها کنترل شد و دسته بندی و نرمال سازی داده‌ها انجام گرفت. بر اساس مطالعات دیسکسن¹ و همکاران (5) گاهی در هنگام هیبریداسیون به دلیل نوسان در غلظت DNA نمونه‌های مختلف و محتوی جی-سی متفاوت، اتصال پروب‌ها به‌طور یکنواخت صورت نمی‌گیرد این پدیده را امواج ژنومیکی می‌نامند که بر بازخوانی ژنوتیپ‌ها اثر می‌گذارد. این پدیده احتمالاً بر صحت تعیین تنوع در تعداد کپی قطعه ژنومی نیز مؤثر است. بدین منظور ژن کال اسکور² نمونه‌ها بررسی شد. این پارامتر یک کمیت متریک برای نشان دادن قابلیت اطمینان ژنوتیپ‌های خوانده شده است که حد آستانه آن 0/15 است. ژنوتیپ‌ها با مقدار کمتر از 0/15 به دلیل عدم اطمینان در تعیین ژنوتیپ صحیح حذف شد. پس از انتخاب آیت‌های موردنیاز برای محاسبه تنوع، فایل شدت سیگنال در فرمت متنی که همان فایل ورودی نرم‌افزار پنسی.ان.وی است، تهیه شد. کلیه مراحل تعیین ژنوتیپ در وبسایت شرکت ایلومینا در دسترس می‌باشد.

آنالیز با پنسی.ان.وی و کنترل کیفیت

نرم‌افزارهای متعددی برای شناسایی تنوع در تعداد کپی قطعه ژنومی بر اساس داده‌های جهش تک‌نوکلئوتیدی وجود دارد. متداول‌ترین آن‌ها پنسی.ان.وی می‌باشد که از طریق وبسایت³ آن قابل دسترسی می‌باشد. فایل ورودی نرم‌افزار پنسی.ان.وی که همان فایل شدت سیگنال⁴ است، حاوی اطلاعات نام جهش تک‌نوکلئوتیدی تک‌نوکلئوتیدی و موقعیت آن، پارامترهای لگاریتم نسبت شدت سیگنال⁵ (ال.ا.ا.ر) و نسبت شدت سیگنال نسبی دو آلل جهش تک‌نوکلئوتیدی⁶ (بی.ای.اف) می‌باشد. با توجه به نحوه تعیین ژنوتیپ جهش تک‌نوکلئوتیدی‌ها، از شدت سیگنال ساطع شده آلل A و B (مقادیر X و Y) برای تهیه فایل‌های ال.ا.ا.ر. و بی.ای.اف استفاده می‌شود. به این ترتیب که ال.ا.ا.ر. کل شدت سیگنال نرمال شده برای هر دو آلل جهش تک‌نوکلئوتیدی می‌باشد که در زمان حذف، کاهش و در زمان اضافه شدن، افزایش می‌یابد. نحوه محاسبه آن به این صورت است اگر $R = Y + X$ باشد آنگاه $(R_{observed} / R_{expected})$ $LRR = \log_2$ است که $R_{expected}$ از طریق درون‌یابی خطی دسته‌های ژنتیکی جهش تک‌نوکلئوتیدی نمونه‌ها به دست می‌آید. بی.ای.اف، نرمال شده، نسبت شدت سیگنال نسبی آلل A و B است که از طریق محاسبه $\theta = \arctan(Y/X) / (\pi/2)$ به دست می‌آید. در شروع آنالیز،

7- Hidden Markov Model
8- Copy Number (cn)
9- waviness factor value
10- ISGC_Ovis_aris_1.0
11- Baylor 4.0/Bos Tau4
12- DAVID

1- Diskin
2- GenCall Score
3- <http://www.openbioinformatics.org/pennncnv/>
4- signal intensity file
5- Log R Ratio(LRR)
6-B Allele Frequency(BAF)

مسیرهای متابولیکی مختلف مورد ارزیابی قرار گرفت.

نتایج و بحث

شناسایی تنوع در تعداد کپی قطعه ژنومی

حذف تنوع‌ها با یک تکرار 904 تنوع در تعداد کپی قطعه ژنومی شامل 599 حذف، 111 اضافه و 194 حذف/اضافه باقی ماند. به‌طور متوسط 4/91 تنوع پس از تصحیح به‌ازاء هر رأس گوسفند شناسایی شد. میانگین و میانه این تنوع‌ها به‌ترتیب 170 و 123/9 کیلو جفت‌باز بود. 60 درصد تنوع‌ها طولی بین 100 تا 500 کیلو جفت‌باز داشتند. در این مطالعه تنوع با طول کمتر از 30 کیلو جفت‌باز مشاهده نشد. کل طول این تنوع 8/5 مگا جفت‌باز معادل 0/32 درصد از کروموزم‌های اتوزوم بود (جدول 1).

پس از تصحیح در مرحله خوانش جهش‌های تک‌نوکلئوتیدی، 184 نمونه باقی ماند. از این تعداد نمونه 2180 تنوع در تعداد کپی قطعه ژنومی یعنی 11/85 تنوع به‌ازاء هر رأس گوسفند شناسایی شد. کلیه تنوع‌ها با یک تکرار از ادامه آنالیزها حذف شد (58 درصد). پس از

جدول 1- خصوصیات تنوع در تعداد کپی قطعه ژنومی شناسایی شده در ژنوم گوسفند

حذف/اضافه	اضافه	حذف	کل	معیار اندازه‌گیری
Loss/gain	gain	loss	total	Item measurement
194	111	599	904	تنوع در تعداد کپی قطعه ژنومی Copy number variation
3.32	1.92	3.26	8.5	طول Length (Mbp)
302	240	105	170	متوسط طول Mean length (kb)
0.0	10(1.2%)	351 (38.8%)	361(40%)	Kb - <100 kb ≥ 30
194(21.4%)	101(11.2%)	248(27.4%)	543 (60%)	Kb - < 500 Kb ≥ 100

ناحیه تنوع در تعداد کپی قطعه ژنومی

پس از ادغام نواحی مشترک 88 ناحیه تنوع در تعداد کپی قطعه ژنومی در کروموزم‌های اتوزوم شناسایی شد که 53 مورد آن فقط در یک حیوان مشاهده شد و از ادامه آنالیزها حذف گردید. از 35 مورد پُر تکرار شش مورد اضافه، 21 مورد حذف و هشت مورد حذف/اضافه بود. این تعداد تنوع، 6/3 مگا جفت‌باز از ژنوم معادل 0/23 درصد از کل کروموزم‌های اتوزوم گوسفند را با میانگین 178/61 و میانه 135/25 کیلو جفت‌باز به‌خود اختصاص داد. طول ناحیه تنوع در تعداد کپی قطعه ژنومی از 34 تا 500 کیلو جفت‌باز متغیر بود. 31/4 درصد از نواحی تنوع طولی کمتر از 100 کیلو جفت‌باز داشتند (جدول 2).

در این مطالعه تعداد نواحی حذفی 3/5 برابر نواحی اضافی بود. 11 ناحیه شناسایی شده فراوانی کمتر از 2 درصد و 15 مورد فراوانی بالاتر از سه درصد داشتند. ناحیه 55441983-55475961 بر روی کروموزوم شماره 10 بیش‌ترین فراوانی (هشت درصد) را داشت. فونتازنی و همکاران (8) 135 ناحیه تنوع (75 حذف، 59 اضافه و یک مورد هر دو) با میانگین و میانه 77/6 و 55/8 کیلو جفت‌باز شناسایی کردند. علت تفاوت در طول شاید در روش شناسایی باشد. در این مطالعه تنوع‌ها با چیپ 50 کی شناسایی شد، اما در پژوهش قبلی از روش aCGH استفاده شد.

لیو و همکاران (21) در بررسی تنوع در تعداد کپی قطعه ژنومی ژنوم سه نژاد گوسفند با استفاده از چیپ 50 کی با 256 نمونه موفق شدند 3624 تنوع در تعداد کپی قطعه ژنومی (3416 حذف و 208 اضافه) با میانگین و میانه طول به‌ترتیب 144/6 و 122/9 کیلو جفت‌باز شناسایی کنند. البته آن‌ها تنوع‌ها با یک تکرار را از آنالیزهای خود حذف نکردند. نتایج آن‌ها به‌لحاظ تعداد با توجه به اندازه نمونه و میانه با نتایج این پژوهش مطابقت دارد. در مطالعه حاضر 58 درصد تنوع‌ها یک تکرار داشتند که با نتایج مطالعه فونتازنی و همکاران (8) مطابقت داشت (60 درصد). تعداد تنوع به‌ازاء هر رأس گوسفند (11/85) با نتایج مشاهده شده در بُز (17/9)، گوسفند (16/9) و (14/16) و جوجه گوشتی (9/6) تشابه داشت (8، 9، 21، 29). تعداد و طول تنوع شناسایی شده بستگی به تعداد نمونه، روش شناسایی، نرم‌افزار مورد استفاده، گونه دام و تراکم نشانگر مورد استفاده برای شناسایی، اسمبلی ژنوم و غیره دارد. لذا، اختلاف در نتایج مطالعات مختلف می‌تواند دلایل متعددی داشته باشد. بر اساس مطالعه بی و همکاران (2) با افزایش تعداد نمونه تعداد تنوع‌های شناسایی شده در سطح ژنوم افزایش می‌یابد. تعداد حذف شدن‌ها در این مطالعه 5/4 برابر تعداد اضافه‌ها بود. در اکثر مطالعات تعداد حذف‌ها چندین برابر اضافه شدن است که این امر می‌تواند به‌دلیل عوامل بیولوژیکی باشد و این‌که نوترکیبی همولوگوس غیر آلی بیشتر تمایل به ایجاد حذف دارد تا اضافه، باشد (3).

جدول 2- مشخصات و فراوانی نواحی تنوع در تعداد کپی قطعه ژنومی در ژنوم گوسفند
Table 2- Characterization & copy number variation regions in sheep genome

معیار اندازه گیری (measurement item)	کل تعداد ناحیه تنوع Total No. CNV	تعداد حذف No. loss	تعداد اضافه No. gain	حذف/اضافه Loss/gain
ناحیه سی.ان.وی Copy number variation region(CNVR)	35	21	6	8
کل طول (مگا جفت‌باز) Total length	6.25	3.03	1.41	1.81
متوسط طول تنوع (کیلو جفت‌باز) Mean CNV length	178.61	144.09	236.59	225.76
30 kb ≤ x < 100 Kb	11	8	0.0	3
100 Kb ≤ x < 500 Kb	24	13	6	5

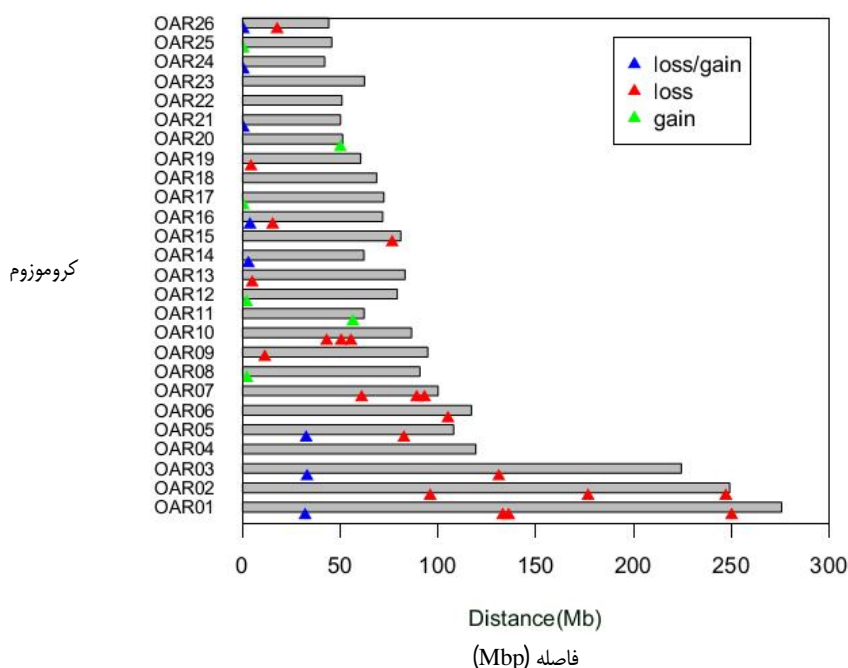
نقشه کروموزومی نواحی تنوع

برای بررسی نحوه توزیع نواحی تنوع بر روی کروموزوم‌ها، پروفایل آن با استفاده از اسمبلی Ovis_aris_3.0 ژنوم گوسفند ترسیم شد. نقشه به دست آمده نشان داد که توزیع نواحی تنوع به صورت کاملاً تصادفی نبوده و بسته به کروموزوم و محل کروموزومی تغییر می‌کند. تعداد زیادی ناحیه تنوع در تعداد کپی قطعه ژنومی در نواحی ساب‌تلومریک و پری‌سانترومریک قرار دارند. تلومر به ناحیه‌ای در انتهای کروموزوم اطلاق می‌شود که مانع از تجزیه انتهای کروموزوم و امتزاج کروموزوم‌ها با یکدیگر می‌شود و در مهره‌داران حاوی توالی‌های تکراری بلند (TTAGGG)_n به طول 20-3 کیلو جفت‌باز است (کلاهک). ناحیه ساب‌تلومریک ناحیه‌ای بین کلاهک تلومر و کروماتین است. هر کروموزوم حاوی دو ناحیه ساب‌تلومریک است. این ناحیه بیشتر حاوی نواحی دوبرابر شده (اس.دی)، ساتلیت‌ها و توالی اتصال به صورت (TTAGGG)_n است. نواحی پری‌سانترومریک و ساب‌تلومریک 3/4 درصد توالی ژنوم گاو را شامل می‌شوند که 2/5- 1/5 برابر تنوع در تعداد کپی قطعه ژنومی بیشتری نسبت به سایر نواحی دارند و 6/9 درصد - 6/7 درصد نواحی پلی‌مورف ژنوم را شامل می‌شوند (17). شاید دلیل دیگر، تراکم بیشتر نواحی تنوع در نواحی تلومری و سانترومی چپ مورد استفاده باشد. زیرا مطالعات هو و همکاران (12) نشان داد که نواحی تنوع شناسایی شده با چپ متراکم‌تر توزیع یکنواخت‌تری در ژنوم دارد، در حالی که تنوع شناسایی شده با چپ 50 کی بیشتر در نواحی تلومری و سانترومی قرار دارد. کروموزم یک دارای بیش‌ترین تعداد ناحیه تنوع اما کمترین پوشش بود (شکل 1). در مطالعه هو و همکاران (13) بیش‌ترین تعداد ناحیه تنوع بر روی کروموزوم یک و شش گاو قرار داشت. در حالی که کروموزوم یک گاوی بیش‌ترین منطقه برای دو برابر شدن³ را ندارد (19).

روش aCGH دقت و قدرت بیشتری در شناسایی این نوع تنوع دارد، زیرا توزیع پروب‌های در سراسر ژنوم یکنواخت‌تر است و حدود نواحی تنوع را دقیق‌تر تشخیص می‌دهد (17). از این رو تعداد نواحی تنوع شناسایی شده بیشتر و طول آن کوتاه‌تر است. علاوه‌براین در مقایسه با روش aCGH، روش چپ آرایه حاوی پروب‌های کمتری برای نواحی غیرپلی‌مورف است، لذا فقط تنوع‌های بزرگ را شناسایی می‌کنند (12). اما، شاید دلایل این اختلاف در نتایج فقط این موضوع نباشد، چرا که تعداد نمونه در مطالعه فوتنانزی و همکاران (8) بسیار کم‌تر (11 نمونه) از این مطالعه بود و آن‌ها از پروب‌های ژنوم گاو برای شناسایی تنوع موجود در ژنوم گوسفند استفاده کرده‌اند و ممکن است به‌خاطر اختلاف در ژنوم گوسفند و گاو برخی پروب‌ها به‌درستی با ژنوم گوسفند هیبرید نشده باشد. در مطالعه دیگر بر روی ژنوم بز با همین روش 127 ناحیه تنوع با پوشش 11/2 مگا جفت‌باز از ژنوم شناسایی شد (9). در مطالعه لیو و همکاران (21) تعداد 238 ناحیه تنوع با میانگین و میانه 253 و 186 کیلو جفت‌باز و طول کلی 60 مگا جفت‌باز شناسایی شد. این اختلاف در نتایج می‌تواند به دلیل عدم حذف نواحی تنوع با یک تکرار در مطالعه لیو و همکاران (21) و تفاوت در اسمبلی ژنوم مورد استفاده در دو پژوهش و تعداد نمونه باشد. در مطالعه هو و همکاران (12) هنگامی که نتایج چپ 777 کی در دو اسمبلی بی.تی.ای.یو. 4/0¹ و یو.ام.دی. 3/1² ژنوم گاو با یکدیگر مقایسه شد تنوع‌های شناسایی شده در ژنوم گاو با اسمبلی بی.تی.ای.یو. 4/0 طول کوتاه‌تری در مقایسه با نتایج حاصله از اسمبلی یو.ام.دی. 3/1 داشتند، اما به لحاظ تعداد و طول کلی تقریباً یکسان بودند. آن‌ها دلیل این موضوع را در جابجایی درست بسیاری از کانتیگ‌های بی‌محل و دوبرابر شدن‌های ساختگی موجود اسمبلی بی.تی.ای.یو. 4/0 را در اسمبلی جدید دانستند.

1- Btau_4.0
 2- UMD3.1

3- Segmental duplication(SD)



شکل 1- پروفایل توزیع نواحی تنوع در تعداد کپی قطعه ژنومی بر روی کروموزوم‌های اتوزوم گوسفند

Figure 1- the distribution profile of copy number variation regions in sheep autosomal chromosome

بزرگی کروموزوم یک مرتبط باشد. کروموزوم 26، 25 و 24 با پوشش به ترتیب 2 درصد، 0/8 درصد و 0/9 درصد طول کروموزوم، بیش‌ترین طول ناحیه تنوع در تعداد کپی قطعه ژنومی را داشتند. در کروموزوم های چهار، 18، 22 و 23 هیچ ناحیه تنوعی شناسایی نشد. 9 کروموزوم (2، 6، 7، 9، 10، 13، 15 و 19) فقط حاوی ناحیه حذف به‌طول کلی 3/82 مگا جفت‌باز بودند (جدول 3).

بر اساس مطالعات تفاوت آشکار و واضحی بین نواحی دو برابر شده و تنوع در تعداد کپی قطعه ژنومی وجود ندارد. دوبرابر شدن قطعه ای، به ناحیه‌ای به‌طول بیش از یک کیلو جفت‌باز اتلاق می‌شود که بیش از 90 درصد تشابه توالی داشته باشد و به‌نظر می‌رسد، این نواحی همان تنوع در تعداد کپی قطعه ژنومی اجدادی هستند که در جمعیت تثبیت شده‌اند (19). لذا شاید کمترین پوشش تنوع در کروموزوم یک (0/06 درصد) به کمترین میزان دوبرابر شدن در این کروموزوم و یا

جدول 3- مشخصات کروموزوم‌های اتوزوم گوسفند به لحاظ پوشش تنوع در تعداد کپی قطعه ژنومی

Table 3- the sheep autosomal chromosome characterization for CNVR coverage

کروموزوم Chr.	طول (bp) length	کل طول ناحیه تنوع Total length of CNVR(bp)	تعداد ناحیه تنوع No. CNVR	درصد پوشش کروموزوم Chr. Coverage %	کروموزوم Che.	طول (bp) length	کل طول ناحیه تنوع Total length of CNVR (bp)	تعداد	درصد پوشش
								ناحیه تنوع No. CNVR	کروموزوم Chr. Coverage %
OAR1	275612895	167855	4	0.06	OAR13	83079144	193525	1	0.2
OAR2	248993846	204301	3	0.08	OAR14	62722625	228607	1	0.3
OAR3	224283230	148704	2	0.07	OAR15	80923592	268156	1	0.3
OAR5	107901688	178524	2	0.2	OAR16	71719816	573372	2	0.8
OAR6	117031472	101639	1	0.09	OAR17	72286588	300793	1	0.4
OAR7	100079507	323648	3	0.3	OAR19	60464314	304446	1	0.5
OAR8	90695168	120357	1	0.1	OAR20	51176841	306982	1	0.6
OAR9	94726778	127868	1	0.1	OAR21	50073674	356052	1	0.7
OAR10	86447213	422158	3	0.5	OAR24	42034648	358678	1	0.9
OAR11	62248096	157132	1	0.3	OAR25	45367442	363424	1	0.8
OAR12	79100223	170845	1	0.2	OAR26	44077779	874327	2	2

مقایسه نواحی شناسایی شده در این مطالعه با سایر نتایج مطالعه سایر نشخوارکنندگان

در مقایسه نتایج تحقیق حاضر با نتایج مطالعه فونتازنی و همکاران (8) هیچ کدام از محل‌های نواحی شناسایی شده در این مطالعه در نتایج آن‌ها یافت نشد. این امر دلایل متعددی می‌تواند داشته باشد، اما به نظر می‌رسد شاید یکی از مهم‌ترین دلایل آن تفاوت در اسمبلی مورد استفاده در دو پژوهش باشد. بسیاری از کانتینگ‌های به محل در اسمبلی Ovis_aris_1.0 ژنوم گوسفند که در مطالعه فونتازنی و همکاران (8) استفاده شده بود در اسمبلی جدید (Ovis_aris_3.0) که در این پژوهش استفاده شده است، جایابی شده است. لذا احتمالاً محل نواحی تنوع در تعداد کپی قطعه ژنومی تغییر کرده است. نواحی شناسایی شده در 2 مورد به صورت جزئی با نتایج مطالعه لیو و همکاران (21) بر روی ژنوم گوسفند، مطابقت داشت. نتایج این پژوهش با مطالعات مشابه در ژنوم گاو با چپ 50 کی‌گای 777 کی‌گای نیز مقایسه شد. از 35 ناحیه تنها 26 ناحیه تبدیل به معادل آن در ژنوم گاو شد که دو مورد از این 26 مورد بر روی کروموزوم ناشناخته در بی.تی.ای.یو. 4/0 بود، که از آنالیزها حذف شد. ناحیه‌ای به طول 193 کیلو جفت‌باز بر روی کروموزوم 9 گاو در مطالعه بی و همکاران (2) با ناحیه‌ای مشابه بر روی کروموزوم 9 گوسفندی در این پژوهش مطابقت داشت. علاوه بر این، نتایج حاضر با نتایج مطالعه هو و همکاران (13) با چپ 50 کی‌گای در چهار مورد و با چپ 777 کی‌گای (12) در 10 مورد مطابقت جزئی و کامل داشت (جدول 4). با توجه به این که با افزایش تراکم نشانگر مورد استفاده، قدرت شناسایی ناحیه تنوع بیشتر می‌شود، لذا مطابقت با نتایج چپ 777 کی‌گای بیشتر بود. در اکثر مطالعاتی که تاکنون بر روی ژنوم گاو صورت گرفته است، میزان هم‌پوشانی بین نتایج پایین است. به طور کلی نواحی تنوع در تعداد کپی قطعه ژنومی در مطالعه هو و همکاران (13) در حدود 18/9 درصد با مطالعه بی و همکاران (2) و 7/84 درصد با مطالعه فدیستا و همکاران (7) و 8/36 درصد با مطالعه لیو و همکاران (17) در گاو هم‌پوشانی داشت. به عبارتی 70 درصد نواحی شناسایی شده در مطالعه آن‌ها در سایر منابع وجود نداشت. لذا با توجه به درصدهای پایین هم‌پوشانی در مطالعات مختلف به نظر می‌رسد، هنوز بسیاری از تنوع‌های ساختار کروموزومی در ژنوم گاو شناسایی نشده است. شرایط مشابه با ژنوم گاو برای ژنوم گوسفند برقرار است به طوری که نتایج این مطالعه هم‌پوشانی بسیار پایینی با دو پژوهش انجام شده بر روی ژنوم گوسفند و همین‌طور گاو داشت. لذا با توجه به این که این مطالعه دومین مطالعه در سطح کل ژنوم گوسفند است و به نظر می‌رسد بسیاری از این تنوع‌ها هنوز شناسایی نشده است.

پوشش نواحی تنوع در تعداد کپی قطعه ژنومی با نواحی ژنی از 35 ناحیه شناسایی شده تنها 25 ناحیه به صورت کامل یا جزئی با یک ژن هم‌پوشانی داشتند. در انسان تعداد 110 ژن رفرنس در 16 ناحیه تنوع در تعداد کپی قطعه ژنومی و در گاو 40 ژن رفرنس در 10 ناحیه شناسایی شد. ژن رفرنس توالی ژنومی است که به عنوان رفرنس استاندارد برای ژن‌های شناخته شده استفاده می‌شود و بخشی از پروژه ال.ا.ر.جی¹ است. در گوسفند تنها دو ناحیه تنوع با ژن‌های رفرنس هم‌پوشانی داشتند. به دلیل گستردگی نتایج حاصله از ارائه لیست ژن‌های شناسایی شده این بخش خودداری شد. بررسی مسیرهای متابولیکی ژن‌های موجود در نواحی تنوع در تعداد کپی قطعه ژنومی نشان داد که برخی از این ژن‌ها در مسیرهای متابولیکی ترانس دوکسیون بویای، سرطان در انسان و متابولیسم پورین و پریمیدین در گونه گاو و بیوسنتز استروئیدها هم در انسان و هم در گونه گاو نقش دارند. علاوه بر این برخی از ژن‌های شناسایی شده با بیماری‌های از قبیل استئوپوروسیز، کم خونی، اریتروسیتوزیس و مطالعات ارتباطی مقاومت به انسولین در انسان مرتبط بودند.

دو ناحیه تنوع در تعداد کپی قطعه ژنومی 481088-180295 در کروموزوم 12 و 385305-156698 در کروموزوم 16 به ترتیب ژن‌های رفرنس MYOG و Mir103 را در ژنوم گوسفند به طور کامل احاطه کرده بود. ژن MYOG از خانواده MYOD می‌باشد که در مایوژنز یعنی سنتز پروتئین میوفیبریل در عضلات اسکلتی و تنظیم تعداد میوفیبریل‌ها نقش دارد (31). مطالعات نشان داد که بیان مضاعف ژن Mir103 در غدد پستانی بز منجر به افزایش نسخه‌برداری از ژن‌های مرتبط با سنتز چربی تری‌گلیسیرید و اسیدهای چرب غیر اشباع در سلول‌های اپی‌تلیال غدد پستانی می‌شود (16). در مطالعه لیو و همکاران (21) 127 ناحیه تنوع با 1043 ژن مرجع از ژنوم انسان و 106 ناحیه تنوع با 270 ژن مرجع در ژنوم گاو هم‌پوشانی داشتند. علت اختلاف در تعداد ژن شناسایی شده، تفاوت در تعداد نواحی تنوع شناسایی شده در دو مطالعه بود (238 در مقابل 35 ناحیه). اما ژن‌ها شناسایی شده در هر دو مطالعه در مسیرهای متابولیکی مشابه فعالیت داشتند. ژن‌های به دست آمده در مطالعه فونتازنی و همکاران (8) با بیماری ژنتیک مندلی و طول عمر در انسان ارتباط داشت که مشابه نتایج حاضر بود.

در مطالعه کیجیس و همکاران (14) در حدود 141 ژن بر اساس اسمبلی بی.تی.ای.یو. 4/0 به طور جزئی یا کامل با 42 ناحیه تنوع از 51 ناحیه شناسایی شده هم‌پوشانی داشتند. هم‌چنین در مطالعه فدیستا و همکاران (2010) در حدود 348 ژن با 90 ناحیه هم‌پوشانی داشتند.

جدول 4- مقایسه همپوشانی نواری شناسایی شده با مطالعات مشابه در گوسفند و گاو
 Table 4- comparison of CNVR coverage with similar studied in sheep and cattle

کروموزم Chr	کروموزم chr	ناحیه تنوع در تعداد کمی قطعه ژنومی در گوسفند CNVR in Sheep	مقابل ناحیه در ژنوم گاو یا تبدیل در UCSC CNVR in cattle converted by UCSC	بی و همکاران (۲) Bae <i>et al.</i> (2)	هو و همکاران (۳) Hou <i>et al.</i> (3)	هو و همکاران (۱۲) (Biau4.0) Hou <i>et al.</i> (12)(Biau4.0)	لیو و همکاران (۲۱) Liu <i>et al.</i> (21)
1	1	133504096-133431841	4164033-4125261			4160400-4076896	
1	1	136423681-136305985	6972136-6855383			6975308-6969184	
1	1	249833893-249760850	116038720-115972051				
2	2	176607026-176503646	57001953-56898963				
2	3	489045-7495	126543539-126025529		126380127-126097657		
2	8	95923117-95820546	23496725-23441832				
8	9	2396681-2245347	14870609-14719674				
9	9	11688408-11494856	1408671-1223924				
7	10	61291011-60986565	60471824-60304880	1399219-1053431	1463199-1203864	1339470-1287617	12923703-11627386
7	10	89723172-89566040	87913856-87757476		87724129-87481635	877633321-87758707, 87855464-8782655, 87878635-87873996	
7	10	93688220-93560351	91759540-916392258				
3	11	13225301-13149640	97477207-97407786				
10	12	4353825-43417893	42865579-42758928				
10	12	50482176-50346927	49467686-49333866			42784397-42752600	
10	12	55475961-55441983	54355735-54322451				
13	13	4989997-4889447	4775863-4676755				
15	15	76549178-76486721	72470935-72408585				
17	17	496828-140776	483757-138526			445444-442724	
16	20	385305-156698	2773071-2548393			2754521-2695880	
16	20	15482312-15425949	15091121-15035002				
19	22	4319831-4184256	4109630-3975553				
20	23	50605513-50503874	48453256-48352345				
21	-	499813-231657	-				5220352-4744399
26	27	451489-166516	462308-182063			318690-238583	
26	27	17757292-17719912	17452961-17416230		17547965-17397248		

به‌خصوص که در مقایسه با جهش تک‌نوکلئوتیدی هر واحد از این تنوع حاوی اطلاعات بیشتری است. علاوه‌براین، شناسایی حدود 152 ژن رفرنس در نواحی شناسایی شده، این فرضیه را تقویت می‌کند که این نوع تنوع قادر است، طیفی از عملکردهای مولکولی را تحت تاثیر قرار دهد و بیان ژن و مسیرهای متابولیکی را تغییر دهد و حتی منجر به تفاوت در عملکرد فنوتیپی بین افراد یک جمعیت و نژادهای مختلف شود. حفاظت تنوع در تعداد کپی قطعه ژنومی در بین پستانداران می‌تواند منجر به ابقاء و اکتساب یک دژ خاص از یک ژن شود. از این رو این ژن یا خانواده آن به‌صورت تکراری در پستانداران یافت می‌شود (19). لذا به‌نظر می‌رسد شناسایی این تنوع ممکن است افق‌های جدیدی را در برنامه‌های اصلاح نژادی باز کند. مثلاً استفاده از تنوع‌های تثبیت شده در ژنوم، در مطالعات انتخاب بر مبنای نشانگر و یا مطالعات ارتباطی از جنبه‌های مختلف این اُفق‌ها می‌باشد. اما هم‌اکنون شناسایی این تنوع به‌خصوص در حیوانات اهلی در مرحله بسیار ابتدایی قرار دارد، لذا مطالعات بیشتر در این زمینه برای بررسی پتانسیل استفاده از آن در برنامه‌های اصلاح نژادی ضروری به‌نظر می‌رسد.

در مطالعه هو و همکاران (12) 939 ژن مرجع و 2153 پپتید در وب‌سایت انسمبل¹ که معادل 1855 ژن واحد در این وب‌سایت بود به‌دست آمد که با 3432 ناحیه تنوع هم‌پوشانی داشت. لیو و همکاران (20) هم‌پوشانی بین جایگاه صفات کمی و ناحیه تنوع را بررسی کردند. در مطالعه آن‌ها بسیاری از تنوع‌ها با جایگاه صفات کمی ورم پستان، شمار سلول‌های سوماتیکی²، اسکور سلول‌های سوماتیکی³ و مقاومت به انگل، درجه ماربلینگ، سخت‌زایی، نرخ باروری، طول دوره بارداری و تعداد تلقیح به‌زاء هر بارداری هم‌پوشانی داشت.

نتیجه‌گیری کلی

به‌طور کلی در این پژوهش 2180 تنوع و 88 ناحیه تنوع در 184 حیوان شناسایی شد. مقایسه این نتایج با نتایج به‌دست آمده از مطالعات دیگر بر روی ژنوم گوسفند و گاو نشان داد که هم‌پوشانی بین نواحی شناسایی شده در مطالعات مختلف بسیار کم و ضعیف است. این بدین معنی می‌تواند باشد که احتمالاً تعداد زیادی از این تنوع هنوز شناسایی نشده است. لذا به‌نظر می‌رسد این تنوع می‌تواند درصد قابل‌توجهی از تنوع موجود در ژنوم را به‌خود اختصاص دهد.

منابع

- 1- Aboura, A., C. Dupas, G. Tachdjian, M. Portnoi, N. Bourcigaux, D. Dewailly, R. Frydman, B. Fauser, N. Ronci-Chaix, B. Donadille, P. Bouchard, S. Christin-Maitre. 2009. Array Comparative Genomic Hybridization Profiling Analysis Reveals Deoxyribonucleic Acid Copy Number Variations Associated with Premature Ovarian Failure. *Journal of Clinical Endocrinology & Metabolism*, 94: 4540-4546.
- 2- Bae, J. S., H. S. Cheong, L. H. Kim, S. NamGung, T. J. Park, J. Y. Chun, J. Y. Kim, C. F. Pasaje, J. S. Lee and H. D. Shin. 2010. Identification of copy number variations and common deletion polymorphisms in cattle. *BMC Genomics*, 11: 232.
- 3- Clop, A., O. Vidal and M. Amills. 2012. Copy number variation in the genomes of domestic animals. *Animal Genetics*, 43: 503-517
- 4- Conrad, D. F., D. Pinto, R. Redon, L. Feuk, O. Gokcumen, Y. Zhang, J. Aerts, T. D. Andrews, C. Barnes, P. Campbell, T. Fitzgerald, M. Hu, C. H. Ihm, K. Kristiansson, D. G. MacArthur, J. R. MacDonald, I. Onyiah, A. W. C. Pang, S. Robson, K. Stirrups, A. Valsesia, K. Walter, J. Wei, The Wellcome Trust Case Control Consortium, C. Tyler-Smith, N. P. Carter, C. Lee, S. W. Scherer and M. E. Hurles. 2009. Origins and functional impact of copy number variation in the human genome. *Nature*. 464: 704-712.
- 5- Diskin, S. J., M. Li, J. Hou, S. Yang, J. Glessner, M. Hakonarson, M. Bucan, J. M. Maris and K. Wang. 2008. Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Research*. 36: e126.
- 6- Elferink, M. G., A. A. Vallée, A. P. Jungerius, R. P. Croijmans and M. A. Groenen. 2008. Partial duplication of the PRLR and SPEF2 genes at the late feathering locus in chicken. *BMC Genomics* 9: 391.
- 7- Fadista, J., B. Thomsen, L. Holm and C. Bendixen. 2010. Copy number variation in the bovine genome. *BMC Genomics*. 11:284
- 8- Fontanesi, L., F. Beretti, P. L. Martelli, M. Colombo, S. Dall'Olio, M. Occidente, B. Portolano, R. Casadio, D. Matassino and V. Russo. 2011. A first comparative map of copy number variations in the sheep genome. *Genomics*. 97: 158-165

1- www.ensembl.org

2- somatic cell score

3- somatic cell count

- 9- Fontanesi, L., P. L. Martelli, F. Beretti, V. Riggio, S. Dall'Olio, M. Colombo, R. Casadio, V. Russo and B. Portolano. 2010. An initial comparative map of copy number variations in the goat (*Capra hircus*) genome. *BMC genomics*, 11: 639.
- 10- Huang, D.W., B. T. Sherman and R. A. Lempicki. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1):1-13.
- 11- Hastings, P. J., J. R. Lupski, S. M. Rosenberg and G. Ira. 2009. Mechanisms of change in gene copy number. *Nature Review Genetic*, 10: 551–564.
- 12- Hou, Y., D. M. Bickhart, M. L. Hvinden, C. Li, J. Song, D. A. Boichard, S. Fritz, A. Eggen, S. DeNise, G. R. Wiggans, T. S. Sonstegard, C. P. Van Tassell and G. E. Liu. 2012. Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. *BMC Genomics*, 13: 1471-2164
- 13- Hou, Y., G. E. Liu, D. M. Bickhart, M. F. Cardone, K. Wang, E. S. Kim, L. K. Matukumalli, M. Ventura, J. Song, P. M. Vanradan, T. S. Sonstegard and C. P. Van Tassell. 2011. Genomic characteristics of cattle copy number variations. *BMC genomics*, 12:127
- 14- Kijas, J. W., W. Barendse, W. Barris, B. Harrison, R. McCulloch, S. McWilliam and V. Whan. 2011. Analysis of copy number variants in the cattle genome. *Gene*, 482: 73–77.
- 15- Li, X., L. Tan., X. Liu., S. Lei., T. Yang., X. Chen., F. Zhang., Y. Fang, Y. Guo, L. Zhang, H. Yan, F. Pan, Z. Zhang, Y. Peng, Q. Zhou, L. He, X. Zhu, J. Cheng, L. Zhang and Y. Liu. 2010. A genome wide association study between copy number variation (CNV) and human height in Chinese population. *Journal of Genetic & Genomics*. 37: 779-785
- 16- Lin, X., J. Luo, L. Zhang, W. Wang and D. Gou. 2013. MiR-103 Controls Milk Fat Accumulation in Goat (*Capra hircus*) Mammary Gland during Lactation. *Plos One*, 8(11): e79258
- 17- Liu, G. E., Y. Hou, B. Zhu, M. F. Cardone, L. Jiang, A. Cellamare, A. Mitra, L. J. Alexander, L. L. Coutinho, M. E. Dell'Aquila, L. C. Gasbarre, G. Lacalandra, R. W. Li, L. K. Matukumalli, D. Nonneman, L. C. Regitano, T. P. Smith, J. Song, T. S. Sonstegard, C. P. Van Tassell, M. Ventura, E. E. Eichler, T. G. McDanel and J. W. Keele. 2010. Analysis of copy number variations among diverse cattle breeds. *Genome Research Journal*, 20: 693-703
- 18- Liu, G. E., C. P. Van Tassel, T. S. Sonstegard, R. W. Li, L. J. Alexander, J. W. Keele, L. K. Matukumalli, T. P. Smith and L. C. Gasbarre. 2008. Detection of germ line and somatic copy number variations in cattle. *Journal of Developmental Biology*, 132: 231-237.
- 19- Liu, G. E., M. Ventura, A. Cellamare, L. Chen, Z. Cheng, B. Zhu, C. Li, J. Song and E. E. Eichler. 2009. Analysis of recent segmental duplications in the bovine genome. *BMC Genomics*, 10:571
- 20- Liu, Z. J. 2011. *Next Generation Sequencing and Whole Genome Selection in Aquaculture*. Wiley-Blackwell press. 1st Edition. pp. 87-90
- 21- Liu, J., L. Zhang, L. Xu, H. Ren, J. Lu, X. Zhang, S. Zhang, X. Zhou, C. Wei, F. Zhao and L. Du. 2013. Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. *BMC Genomics*, 14:229
- 22- Mills, R. E., K. Walter, C. Stewart, R. E. Handsaker, K. Chen, C. Alkan, A. Abyzov, S. C. Yoon, K. Ye, R. K. Cheetham, A. Chinwalla, D. F. Conrad, Y. Fu, F. Grubert, I. Hajirasouliha, F. Hormozdiari, L. M. Iakoucheva, Z. Iqbal, S. Kang, M. T. Kidd, M. K. Konkel, J. Korn, E. Khurana, D. Kural, H. Y. Lam, J. Leng, R. Li, Y. Li, C. Y. Lin, R. Luo, X. J. Mu, J. Nemesh, H. E. Peckham, T. Rausch, A. Scally, X. Shi, M. P. Stromberg, A. M. Stütz, A. E. Urban, J. A. Walker, J. Wu, Y. Zhang, Z. D. Zhang, M. A. Batzer, L. Ding, G. T. Marth, G. McVean, J. Sebat, M. Snyder, J. Wang, K. Ye, E. E. Eichler, M. B. Gerstein, M. E. Hurles, C. Lee, S. A. McCarroll, J. O. Korbelt. 1000 Genomes Project. 2011. Mapping copy number variation by population-scale genome sequencing. *Nature*. 470: 59-65
- 23- Nicholas, T. J., Z. Cheng, M. Ventura, K. Mealey, E. E. Eichler and J. M. Akey. 2009. The genomic architecture of segmental duplications and associated copy number variants in dogs. *Genome Research Journal*, 19: 491–499
- 24- Norris, B. J. and V. A. Whan. 2008. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Research Journal*, 188: 1282–1293.
- 25- Pailhoux, E., B. Vigier, S. Chaffaux, N. Servel, S. Taourit, J. P. Furet, M. Fellous, F. Grosclaude, E. P. Cribiu, C. Cotinot and D. Vaiman. 2001. A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nature Genetics*. 29: 453–458.
- 26- Redon, R., S. Ishikawa, K. R. Fitch, L. Feuk, G. H. Perry, T. D. Andrews, H. Fiegler, M. H. Shapero, A. R. Carson, W. Chen, E. K. Cho, S. Dallaire, J. L. Freeman, J. R. González, M. Gratacòs, J. Huang, D. Kalaitzopoulos, D. Komura, J. R. MacDonald, C. R. Marshall, R. Mei, L. Montgomery, K. Nishimura, K. Okamura, F. Shen, M. J. Somerville, J. Tchinda, A. Valsesia, C. Woodward, F. Yang, J. Zhang, T. Zerjal, J. Zhang, L. Armengol, D. F. Conrad, X. Estivill, C. Tyler-Smith, N. P. Carter, H. Aburatani, C. Lee, K. W. Jones, S. W. Scherer and M. E. Hurles. 2006. Global variation in copy number in the human genome. *Nature*. 444: 444-454.
- 27- Rosengren, P. G., A. Golovko, E. Sundstrom, I. Curik, J. Lennartsson, P. G. Rosengren, A. Golovko, E. Sundström, I. Curik, J. Lennartsson, M. H. Seltenhammer, T. Druml, M. Binns, C. Fitzsimmons, G. Lindgren, K. Sandberg, R. Baumung, M. Vetterlein, S. Strömberg, M. Grabherr, C. Wade, K. Lindblad-Toh, F. Pontén, C. H.

- Heldin, J. Sölkner and L. Andersson. 2008. A cis-acting regulatory mutation causes premature hair greying and susceptibility to melanoma in the horse. *Nature Genetics*, 40: 1004–1009.
- 28- Wang, K., M. Li, D. Hadley, R. Liu, J. Glessner, S. F. Grant, H. Hakonarson and M. Bucan. 2007. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research Journal*, 17:1665-1674.
- 29- Wang, X., S. Nahashon, T. K. Feaster, A. Bohannon-Stewart and N. Adefope. 2010. An initial map of chromosomal segmental copy number variations in the chicken. *BMC Genomics*, 11: 351.
- 30- Wright, D., H. Boije, J. R. Meadows, B. Bed'hom, D. Gourichon, A. Vieaud, M. Tixier-Boichard, C. J. Rubin, F. Imsland, F. Hallbook and L. Andersson., 2009. Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. *PLoS Genetics*, 5: e1000512.
- 31- Zhang, Z., F. Xu, Y. Zhang, W. Li, Y. Yin, C. Zhu, L. Du, A. K. and B. Li. 2014. Cloning and expression of MyoG gene from Hu sheep and identification of its myogenic specificity. *Molecular Biology Reports*, 41(2):1003-13.

Copy number variation detection in sheep genome by using ovine BeadChip 50k

Maryam Nosrati^{1*}, Mojtaba Tahmoorespur²

Received: 24-01-2015

Accepted: 17-08-2015

Introduction Recently, genomic research in livestock is focused on genomic variation and its effect on phenotypic performance in economic traits. Copy number variation (CNV) is one of these variations in genome including insertion, deletion and duplication of 1 kb to 1 Mb segment with more than 90% similarity. CNVs can change gene structure and dosage, can regulate gene expression and function and (1, 4). In mammals, it is important source of variability in genomes and it contains 0.4-25% of whole genome variation. Some researches carried out in livestock have been demonstrated that CNV affecting genes or gene regions are associated with several phenotypic traits. For example, CNV in intron 1 of the SOX5 gene causes the pea-comb phenotype in chicken. CNV affects also the Agouti locus in sheep and goats and contributes to the variability of coat color in these two species. The late feathering locus in this avian species includes a partial duplication of the PRLR and SPEF2 genes and Dominant white locus in pigs includes alleles determined by duplications of the KIT gene (2, 5, 6). In spite of, many researches carried out in human represent association between CNV with both complex genetic diseases and traits; however, far too little attention has been paid to CNV in farm animal. This paper will focused on detecting of CNV in sheep genome.

Materials and method The sheep genomic DNA was extracted from blood of 360 Italian ewes using DNA Purification kit (Promega Corporation, Madison, WI). Markers were genotyped by Illumina ovineSNP50 BeadChip according to instructions. It is containing 54,241 markers that uniformly span the entire ovine genome (Illumina, Inc., USA). After completion of the assay, the BeadChips were scanned with a two-color, confocal Bead Array reader. Scanned image intensities were loaded directly into Illumina's BeadStudio 1.2 software. When normalization was completed, the clustering process was performed to assess cluster position for each marker and to determine individual genotypes. LRR and BAF of sample were reported. The PFB file was calculated based on the BAF of each marker in these populations. The sheep GC model file was generated by calculating the GC content of the 1 Mb genomic region surrounding each marker (500 Kb each side). CNVs were inferred using a PennCNV (<http://www.openbioinformatics.org/penncnv/>). Penn CNV quality filters were applied after CNV detection. High quality samples with a standard deviation (SD) of LRR < 0.30 and with the default set: BAF drift as 0.01 and waviness factor value between - 0.05 and 0.05, were used respectively. In addition, the program argument: the "lastchr 26" in the "detect" argument were used for specific CNVs. CNVRs were determined by aggregating overlapping CNVs identified in different animals. The UCSC table browser tool was used to identify the gene content located within or partially overlapping with the CNVRs and DAVID Bioinformatics Resources (<http://david.abcc.ncifcrf.gov>) was used for further GO functional analysis, including Gene Ontology.

Results and Discussion After all filtration 184 samples were remained. All CNVs and CNVRs found in one sample were omitted from further analysis. Finally 904 CNVs (599 losses, 111 gains and 194 losses/gains) were detected. The average number of CNVs per sample was 4.91, with an average length and median size of 170 kb and 123.9 kb, respectively. 60% of all CNVs had length between 100 kb to 500 kb. This result was similar to other research (2, and 3). After all CNVs were aggregated for the CNV region (CNVR), 88 CNVRs were identified that 55 event were found just in one sample and were omitted. The average and median size of CNVR were 178.61 kb and 135.25 kb. The profile of CNVRs location on Ovis_aris_3.0 genome were shown distribution of CNVRs was not randomly. The highest percentages covering of CNVRs located on chromosomes 16, 24, 25 and 26 (0.8%, 0.9%, 0.8% and 2%, respectively). It's similar to result of Liu et al (21). Gene ontology (GO) analysis can provide insight into the functional enrichment of CNVs. For this reason, we ran GO analysis using DAVID <http://david.abcc.ncifcrf.gov>. Two CNVRs (chr12:481088-180295, chr16: 385305-156698) entirely encompassing MYOG RefSeq gene and Mi103 respectively. The gene content of the 25 CNVR s, we used a BLASTN search for homologous human and cattle sequences using the UCSC table browser tool. There were 110 RefSeq homologous human genes located within or partially overlapping with 16 CNVRs and similarly, there were 40 RefSeq homologous cattle genes located within or partially overlapping with 10 CNVRs.

1 - Assistant Professor in Faculty of Agriculture, Payame Noor University, PO BOX. 19395-3697, Tehran, Iran,

2 - Professor in Department of Animal Science, Faculty of agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.

(*Corresponding Author Email: m_nosrati@pnu.ac.ir)

Conclusion Comparing CNVs and CNVRs identified in sheep genome with CNVRs reported in cattle showed demonstrated low level of similarity, so this genomic variation had great potential detection and using in breeding scheme in sheep industry.

Keyword: CNV, Ovine Illumina 50k BeadChip, Penncnv, Sheep.