تولیدات دامی
دوره ۱۹ ■ شماره ۱ ■ بهار ۱۳۹۶
صفحه‌های ۵۸-۴۷

مقایسه پروفایل تنوع CNV در ژنوم گوسفند با گاو هلشتاین

مریم نصرتی*

۱. استادیار، گروه علوم کشاورزی، دانشگاه پیام نور، تهران- ایران

تاریخ وصول مقاله: ۱۳۹۵/۰۲/۱۸ تاریخ پذیرش مقاله: ۱۳۹۵/۰۳/۲۳

چکیده

هدف از این پژوهش مقایسه CNV در تعدادی از نژادهای گوسفند ایتالیا با گاو هلشتاین بود. از ۵۸۰ راس گاو نر هلشتاین و ۳۶۰ راس گوسفند نمونه خون تهیه و DNA آن استخراج شد. ژنوتیپ چند شکلی‌های تک نوکلئوتیدی سراسر ژنوم با تراشه 50k گاوی و گوسفندی در هر دو گونه تعیین شد. با استفاده از نرم افزار PennCNV تعداد ۹۰۴ تنوع با میانگین طول ۱۵۴/۷ kb و میانه ۱۰۳/۷ kb در گوسفند و ۷۴۴ مورد با میانگین طول ۲۱۳/۷ kb و میانه ۱۲۴/۵ kb در گاو شناسایی شد. کل طول این تنوع در ژنوم گوسفند ۸/۵ Mb و در ژنوم گاو ۸۰ Mb بود. پس از ترکیب نواحی مشابه تعداد ۳۵ ناحیه تنوع پر تکرار در ژنوم گوسفند و ۱۴۱ مورد در ژنوم گاو شناسایی شد، فقط ۲۰ درصد نواحی CNV در ژنوم نژادهای گوسفند با ژنوم گاو هلشتاین همپوشانی داشت. علاوه بر این، نواحی این تنوع در ژنوم گوسفند با ۴۰ ژن مرجع و ۶۲ QTL در ژنوم گاو، و ۱۱۰ ژن مرجع در ژنوم انسان بطور کامل یا جزئی همپوشانی داشت. توزیع تنوع‌ها در هر دو گونه در نواحی ساب‌تلومریک و پری سانترومریک بیشتر بود. بر اساس این نتایج، شاید بتوان از روی پیوستگی این تنوع با نواحی ژنی در ژنوم گاو ژن‌ها و QTLهای ژنوم گوسفند را شناسایی نمود. برای کسب نتایج دقیق‌تر مطالعات بیشتری نیاز است.

کلیدواژه‌ها: پروفایل مقایسه‌ای ژنوم، تنوع ژنوم ، گاو هلشتاین، گوسفند، CNV

* نویسنده مسئول Email: m_nosrati@pnu.ac.ir

مقدمه

گاو و گوسفند در حدود ۱۰ هزار سال قبل اهلی شده‌اند [۲۳]. با پیشرفت روش‌های توالی‌یابی ژنوم، تعداد قابل توجهی نشانگر چند شکلی‌تک نوکلئوتیدی و ریزماهواره در ژنوم گاو شناسایی شد که منجر به افزایش سرعت شناسایی ژن‌ها و QTLهای ژنوم این گونه شد. در گوسفند، هنوز محل استقرار بسیاری از ژن‌ها و عملکرد آن‌ها معلوم نیست. این موضوع باعث شده تا شناسایی نشانگرهای نواحی رمزکننده و استفاده از آن‌ها در برنامه‌های اصلاح‌نژاد گوسفند نسبت به گونه گاو کمتر باشد [۱۲ و ۱۴]. یک راهکار عملی برای رفع این محدودیت مقایسه نواحی مختلف ژنوم در گونه‌های مشابه از جنبه‌های مختلف است. مقایسه ژنوم انسان و گوسفند [۵] و همچنین مقایسه نواحی غیر رمزکننده در ژنوم انسان و گاو [۲۱] از این نوع مطالعات است. بر همین اساس، مقایسه ژنوم گوسفند و گاو که شباهت زیادی با یکدیگر دارند می‌تواند از جنبه‌های مختلف مفید باشد [۴]. اولین نسخه نقشه پیوستگی ژنوم گوسفند در سال ۱۹۹۵ ارائه شد که شامل ۲۴۶ نشانگر ریزماهواره بود و در ۲۰۷۰ سانتی‌مورگان از ژنوم را در بر داشت [۴]. علاوه بر توسعه روند نقشه‌یابی نشانگرها، که به کندی در جریان بود، مطالعات نوع نشانگر از ریزماهواره به چند شکلی تک‌نوکلئوتیدی نیز تغییر نمود. این موضوع مقایسه ژنوم گوسفند را با گونه‌های مشابه آسانتر کرد. تنوع موجود بین ۱۲۰ ژن اورتولوگ در طول kb ۶۰ در ژنوم گاو و گوسفند با یکدیگر مقایسه و مشخص شد که نوکلئوتیدهای رمز کننده پروتئین در هر دو ژنوم فقط در سه درصد موارد با یکدیگر تفاوت دارند[۱۴]. بر این اساس، شناسایی میزان تشابه و یا تفاوت در تنوع‌های جدید می‌تواند برای شناسایی ژن‌ها و QTLهای بیشتر در ژنوم گوسفند بر اساس ژنوم گاو موثر باشد و پیشرفت برنامه‌های اصلاح‌نژاد در گوسفند تسریع شود. تنوع در تعداد کپی قطعه ژنومی (CNV) یکی از تنوع‌های جدید می‌باشد که اولین بار در سال ۲۰۰۶ در انسان شناسایی شد [۲۲ و ۱۵]. این تنوع در اثر حذف، اضافه یا دو برابر شدن قطعه‌ای از DNA به طول یک کیلو جفت باز تا چندین مگا جفت باز در ژنوم ایجاد می‌شود که توالی آن در حدود ۹۰ درصد با یکدیگر شباهت دارد [۲۲]. که باعث تغییر در دوز ژن، توالی‌های رمز کننده و تنظیم بیان ژن می‌شود [۹]. در انسان این نوع تنوع با بیماری‌هایی نظیر شیزوفرنی، دیابت نوع اول، سرطان در ارتباط است [۲۰ و ۹].

در حیوانات مزرعه‌ای تنوع فنوتیپی تعدادی از صفات مورفولوژیکی، در ارتباط با CNVها است. بعنوان مثال تاج نخودی در طیور ناشی از دو برابر شدن در اینترون شماره یک ژن SOX5 می‌باشد. در گوسفند دو برابر شدن در ژن ASIP منجر به تنظیم رنگدانه در پوشش بدن می‌شود. تاخیر در پردرآوری جوجه‌ها به دلیل دوبله ناقص در ژن‌های PRLR و SPEF2 می‌باشد [۸]. بطور کلی این تنوع در پستانداران شامل ۲۵-۰/۴ درصد کل تنوع ژنوم می‌باشد [۲۰ و ۹]. اولین مطالعه برای شناسایی CNV در ژنوم حیوانات مزرعه‌ای در گاو صورت گرفت، که ۲۵ عدد CNV در سلول‌های جنسی شناسایی شد [۱۸]. با توسعه فناوری ساخت تراشه‌های برای تعیین ژنوتیپ هزاران چند شکلی تک‌نوکلئوتیدی موجود در ژنوم، شناخت این تنوع با سرعت بیشتری در سایر حیوانات مزرعه‌ای مانند جوجه گوشتی، بز و سگ انجام گرفت [۸ و۷]. اولین مطالعه برای شناسایی این تنوع در گوسفند با استفاده از هیبریداسیون مقایسه‌ای ژنومی و بکارگیری پروب گاوی در شش نژاد گوسفند انجام شد و ۱۳۵ ناحیه CNV بطول Mb ۱۰/۵ شناسایی شد [۷]. چندین مطالعه برای شناسایی CNVها در ژنوم گاو و تعداد محدودی نیز در ژنوم گوسفند صورت گرفته است. ولی برای بررسی تفاوت، تشابه، ماهیت و نحوه توزیع CNVها در دو گونه گاو و گوسفند مطالعه‌ای

انجام نشده است. لذا این پژوهش با هدف بررسی نحوه پراکندگی CNVها در سرتاسر ژنوم و مقایسه تفاوت‌ها و شباهت‌های CNVها در دو گونه گاو و گوسفند انجام شد.

مواد و روش‌ها

نمونه خون از تعدا ۱۵ نژاد گوسفند بومی ایتالیا و شامل Altamurana، Appenninica، Bagnolese، Bergamasca، Biellese، Delle Langhe، Fabrianes، Gent.le di Pugli، Istrian Pramenka، Laticauda، Massese، Pinzirita، Sambucana-Mnifest، Sopravissana و Valle del Belice در حدود ۲۴ نمونه از هر نژاد تهیه شد (۳۶۰ نمونه). همچنین نمونه خون از ۵۸۰ گاو نر هلشتاین دارای شجره از موسسه ANAFI تهیه شد. استخراج DNA از نمونه های خون با استفاده از کیت پرومگا در آزمایشگاه ژنتیک مرکز ژنومیک حیوانات اهلی دانشگاه بولونیا کشور ایتالیا انجام شد. تعدادی از نمونه‌ها قبل از تعیین ژنوتیپ با تراشه مورد نظر به دلائل مختلف نظیر کیفیت نامطلوب DNA حذف شدند. بدین ترتیب تعداد نمونه مورد آزمایش مربوط به گوسفند و گاو به ترتیب ۴۹۳ و ۲۴۴ نمونه بود. در این پژوهش از تراشه نشانگری 50 K گاوی نوع دو شرکت ایلومینا برای تعیین ژنوتیپ ۵۴۶۰۴ چند شکلی در ژنوم گاو استفاده شد، میزان باز خوانی این تراشه ۹۹ درصد است و میانگین فاصله پروب‌های آن در کل ژنوم ۴۹/۴ kb است. همچنین تراشه نشانگری50k گوسفندی شرکت ایلومینا برای تعیین ژنوتیپ ۵۴۲۴۱ چند شکلی در ژنوم گوسفند بر مبنای پروتکل Infinium HD Assay این شرکت استفاده شد، در این نوع تراشه میانگین فاصله بین چند شکلی‌های تک‌نوکلئوتیدی درسراسر ژنوم ۵۰/۹ kb است. پس ازطی مراحل مختلف برای خواندن ژنوتیپ هر چند شکلی، تراشه‌ها تحت لیزر قرار گرفت تا فلورسنس تک باز انتهای پرایمرهای بسط داده شده، برانگیخته شود. نور ساطع شده ثبت شد و داده‌های حاصل از این تصویر با نرم افزار GenomeStudio TM تصحیح شد، همچنین با استفاده از این نرم افزار، گراف تشکیل شده از چند شکلی‌ها برای کلیه نمونه‌ها کنترل شد و دسته بندی و نرمال کردن داده‌ها انجام گرفت. داده‌های حاصل از تعیین ژنوتیپ در هر دو گونه برای چند شکلی‌های تک‌نوکلئوتیدی و افراد با بیش از ۱۰ درصد ژنوتیپ از حذف شده از آنالیزها حذف شد و برای چند شکلی‌های با فراوانی آللی کوچک کمتر از ۰/۰۵ و چند شکلی‌های که در تعادل هاردی-وینبرگ نبودند، تصحیح شد. پس از انتخاب داده‌های مورد نیاز، فایل شدت سیگنال در فرمت متنی که همان فایل ورودی نرم افزار PennCNV بود تهیه شد. نرم افزارهای متعددی برای شناسایی CNV بر اساس داده‌های چند شکلی تک‌نوکلئوتیدی وجود دارد. متداول‌ترین آنها PennCNV می‌باشد که بر اساس مدل پنهان مارکوف با استفاده از برخی اطلاعات دیگر نظیر کل شدت سیگنال، نسبت شدت سیگنال هر چند شکلی تک‌نوکلئوتیدی، فاصله بین چند شکلی تک‌نوکلئوتیدی همسایه و فراوانی آللی چند شکلی تک‌نوکلئوتیدی، CNVها را شناسایی می‌کند. علاوه‌بر این با استفاده از روش‌های Bayesian امکان استفاده از اطلاعات شجره برای شناسایی CNVها را فراهم می‌کند و با تطبیق مدل‌های رگرسیون با محتوی GC بر امواج ژنومیکی غلبه کند و دقت شناسایی CNVها را افزایش دهد. لذا در اکثر مطالعات از این نرم‌افزار استفاده می‌شود و بطور رایگان در دسترس می‌باشد. فایل ورودی نرم‌افزار که همان فایل شدت سیگنال است، حاوی اطلاعات نام چند شکلی تک‌نوکلئوتیدی و موقعیت آن، پارامترهای لگاریتم نسبت شدت سیگنال (Log R Ratio(LRR)) و نسبت شدت سیگنال نسبی (Bllele Frequency (BAF)) دو آلل یک چند شکلی تک‌نوکلئوتیدی می‌باشد. با توجه به نحوه تعیین ژنوتیپ چند شکلی‌ها، از شدت سیگنال ساطع شده آلل A و B برای تهیه فایل‌های فوق استفاده شد. در شروع

آنالیز، فایل شدت سیگنال به فایل‌های کوچکتر تقسیم شد که هر فایل حاوی اطلاعات فوق برای یک فرد بود. فایل مدل GC بر اساس محتوی GC یک مگا جفت باز از ژنوم ساخته شد. فایل hmm مدل پنهان مارکوف را برای نرم افزار PennCNV فراهم می‌کند و مقدار شدت سیگنال مورد انتظار برای تنوع‌های مختلف را به برنامه القاء می‌کند. به این ترتیب کلیه CNVها با طول بیش از ۱ kb و بر اساس الگوریتم مدل پنهان مارکوف در نرم افزار PennCNV محاسبه شدند. پارامتر تعداد کپی گزارش شده توسط نرم‌افزار در کروموزوم اتوزوم برای نواحی که فاقد تکرار هستند برابر دو (موجود دیپلوئید) برای نواحی که حذف صورت گرفته برابر یک یا صفر، بر اساس میزان حذف و برای نواحی دو یا سه برابر شده معادل به ترتیب سه و چهار است. پس از محاسبه CNV فیلترهای کنترل کیفیت اعمال شد. نمونه‌های با انحراف استاندارد LRR کمتر از ۰/۰۳، BAF برابر با حداقل ۰/۰۱ و فاکتور امواج (waviness factor value) بین ۰/۰۵ تا ۰/۰۵- انتخاب شدند. تصحیح برای محتوی GC با استفاده از فایل مدل GC انجام شد. پس از اعمال فیلترهای مورد نظر تعداد ۱۸۴ نمونه برای گوسفند و ۴۷۳ نمونه برای گاو باقی ماند. داده‌ها بر اساس اسمبلی Ovis_aris_3.0 ژنوم گوسفند و اسمبلی UMD3.1 ژنوم گاو برای شناسایی CNVها آنالیز شد. کلیه آنالیزهای انجام شده فقط بر روی کروموزوم‌های اتوزوم بود و کروموزوم‌های جنسی از آنالیزها حذف شد.

با توجه به پوشش ضعیف ژن‌های شناسایی شده در ژنوم گوسفند، برای مقایسه نواحی CNV در ژنوم گوسفند، با توجه به مطالعات قبلی در ژنوم گاو، از ابزار liftOver وبگاه UCSC استفاده شد و نواحی مورد نظر به معادل آن در ژنوم گاو و انسان تبدیل شد. همچنین، برای مطالعه ژن‌های نواحی شناسایی شده، محدوده این نواحی در وبگاه UCSC با استفاده از ابزار table browser برای گونه‌های گوسفند، گاو و انسان ارزیابی شد. بمنظور بررسی عملکرد ژن‌های این نواحی، حدود آن در وبگاه DAVID ثبت و تاثیر این ژن‌ها در مسیرهای متابولیکی مختلف ارزیابی شد. همچنین برای بررسی وجود QTLها در محدود نواحی CNV، مناطق CNVهای پرتکرار در ژنوم گوسفند پس از تبدیل آن به نواحی معادل در ژنوم گاو بر اساس اسمبلی UMD3.1 در وبگاه QTLdb ثبت شد. بدلیل گستردگی حدود اطمینان بازه تشخیص QTLها، فقط موارد با حدود اطمینان کمتر از ۱۰ Mb در نظر گرفته شد.

نتایج و بحث

در نمونه‌های گوسفند پس از تصحیح در مرحله خواندن چند شکلی‌ها، تعداد ۱۸۴ نمونه باقی ماند. از این تعداد نمونه ۲۱۸۰، CNV یعنی ۱۱/۸۵ تنوع به ازای هر راس گوسفند در کروموزم‌های اتوزوم شناسایی شد. کلیه تنوع‌های با یک تکرار از ادامه آنالیزها حذف شد (۵۸ درصد). پس از حذف آن‌ها، ۹۰۴ CNV شامل ۶۰۶ حذف، ۲۹۸ اضافه باقی ماند. بطور متوسط تعداد ۴/۹، CNV پس از تصحیح به ازای هر راس گوسفند باقی ماند که میانگین و میانه این تنوع‌ها به ترتیب ۱۵۴/۷ kb و ۱۰۳/۷ بود و طول ۶۰ درصد تنوع‌ها بین ۱۰۰ تا ۵۰۰ kb بود. در این تحقیق تنوع با طول کمتر از ۳۰ kb مشاهده نشد. کل طول این تنوع ۸/۵ Mb معادل ۰/۳ درصد از کروموزم‌های اتوزوم بود (جدول ۱). در نمونه‌های گاو هلشتاین پس از تصحیح برای محتوی GC و کنترل کیفیت، در نهایت ۱۴۵۱ CNV برای کروموزوم‌های اتوزوم در ۴۷۳ راس گاو هلشتاین شناسایی شد. از کل تعداد CNVهای شناسایی شده ۷۰۷ مورد فقط در یک حیوان مشاهده شد (۴۹ درصد) که از آنالیزهای بعدی حذف شد. به ازای هر حیوان ۱/۶ CNV با میانگین ۲۱۳/۷ و میانه ۱۲۴/۵ کیلو جفت باز شناسایی شد. تعداد CNV شناسایی شده در ژنوم گوسفند تقریبا چهار برابر CNVهای شناسایی شده در ژنوم گاو بود (جدول ۱).

جدول۱. مشخصات CNV و نواحی CNV در ژنوم گاو هلشتاین و گوسفندان ایتالیایی

معیار اندازه گیری	نوع دام	تعداد دام(راس)	کل تعداد تنوع	متوسط تنوع /هر نمونه	تنوع پر تکرار	متوسط طول (kb)	میانه طول (kb)	loss	gain	Loss/gain	کل CNV (Mb)
CNV	گاو	۴۷۳	۱۴۵۱	۱/۶	۷۴۴	۲۱۳/۷	۱۲۴/۵	۳۸۹	۳۵۵	۱/۱	۸۰
	گوسفند	۱۸۴	۲۱۸۰	۱۱/۸۵	۹۰۴	۱۵۴/۷	۱۰۳/۷	۶۰۶	۲۹۸	۲/۰۳	۸/۵
ناحیه CNV	گاو	۴۷۳	۳۸۸	۰/۸	۱۴۱	۲۴۳/۹	۱۳۳/۹	–	–	–	۹۴/۶
	گوسفند	۱۸۴	۸۸	۰/۴۸	۳۵	۱۴۷/۷	۱۲۰/۲				۱۳
ناحیه CNV پر تکرار	گاو	۴۷۳	۱۴۱	۰/۳	۱۴۱	۱۹۶/۸	۱۳۲/۱	۵۷	۳۴	۵۰	۲۷/۷
	گوسفند	۱۸۴	۳۵	۰/۲	۳۵	۱۷۳/۶	۱۳۱/۶	۲۱	۶	۸	۶/۳

انواع CNVها در گونه گوسفند با ۱۸۴ نمونه از۱۵ نژاد مختلف حاصل شد، ولی در نمونه‌های گاوی این تعداد مربوط ۴۷۳ نمونه از نژاد هلشتاین بود. لذا با وجودی که با افزایش اندازه نمونه تعداد CNVهای شناسایی شده افزایش می‌یابد [۷] اما به نظر می‌رسد تاثیر تنوع نژادی در شناسایی CNVها بیشتر است. علاوه بر این افزایش تنوع نژادی، بیشتر منجر به شناسایی CNV با یک تکرار شد بطوری که ۵۸ درصد CNVها در گوسفند فقط دارای یک تکرار بود (در مقایسه با ۴۹ درصد در گاو). از این رو برای شناسایی بهتر و دقیق‌تر جایگاه CNVهای ژنوم هر گونه افزایش تعداد نمونه همراه با تنوع نژادی نتایج مطلوب‌تری خواهد داشت.

کل طولCNV شناسایی شده پرتکرار (CNVها بیش از یک تکرار) در ۴۷۳ راس گاو معادل Mb ۱۶۰ و در ۱۸۴ راس گوسفند Mb ۱۴۸ بود که پس از حذف دام‌هایی با CNV مشابه در گوسفند Mb ۸/۵ و در گاو Mb ۸۰ از کروموزوم‌های اتوزوم شامل CNVها بود که معادل ۰/۳ درصد از کل ژنوم گوسفند و ۲/۷ درصد از کل ژنوم گاو بود. این موضوع تایید کرد که قسمت‌های بیشتری از ژنوم گاو توسط CNV پوشیده شده است. بنابراین با وجودی که تعداد CNVها در ژنوم گوسفند بیشتر از گاو بود اما کل ناحیه‌ای که در ژنوم گاو توسط CNV ها پوشیده شده بیش از نه برابر ژنوم گوسفند بود. شاید مقداری از این فزونی بدلیل بزرگی ژنوم گاو و همچنین مقداری از کاستی بدلیل اندازه نمونه کم گوسفند در این تحقیق باشد، بطور کلی به نظر می‌رسد که در گاو هلشتاین بدلیل انتخاب مداوم و طولانی بسیاری از CNVها در ژنوم تثبیت شده‌اند. مطالعات عدم تعادل پیوستگی بر روی ژنوم گاو هلشتاین ایتالیا نشان داد که انتخاب مداوم در نژاد هلشتاین باعث شده قطعاتی از ژنوم که عدم تعادل پیوستگی آن‌ها زیاد است، در این نژاد بزرگتر از سایر نژادهای شیری در ایتالیا باشد[۱]. به نظر می‌رسد در مورد CNVها نیز این امر صادق باشد. یعنی انتخاب سبب شده تا CNVهای بزرگتری در ژنوم گاو در مقایسه با ژنوم گوسفند ایجاد شود.

میانگین و میانه طول CNVها نیز این موضوع را تایید کرد. میانه CNV در هر دو گونه گاو و گوسفند تقریبا مشابه بود ولی میانگین آن در گاو هلشتاین بیشتر از گوسفند بود. این به مفهوم وجود CNV با طول بلندتر در ژنوم گاو است. طول CNVها در گوسفند بین kb ۳۰ تا

۰/۵ Mb بود ولی این مقدار در گاو هلشتاین بین ۴۴ kb تا ۱/۵ Mb بود. اگر این موضوع پذیرفته شود که انتخاب منجر به تثبیت CNVهای بلند در ژنوم گاو شده است، پس این CNVها تاثیر مثبت بر عملکرد ژن‌ها داشته‌اند چون در اثر انتخاب در ژنوم باقی مانده، طویل‌تر شده اند و تثبیت گردیده‌اند. سپس CNVهای تثبیت شده بصورت یک تنوع جدید به نام دو برابر شدن قطعه‌ای (Segmental duplication (SD)) در ژنوم تثبیت می‌شوند. بر اساس مطالعات انجام شده تفاوت آشکار و واضحی بین SD و CNV وجود ندارد [۱۹]. بطور کلی SD به ناحیه‌ای بطول بیش از ۱ kb گفته می‌شود که حدود ۹۰ درصد توالی مشابه داشته و به نظر می‌رسد این نواحی همان CNV اجدادی هستند که در جمعیت تثبیت شده‌اند [۱۸]. در مطالعات مختلف میزان همپوشانی CNVها با SD بین ۴۰ تا ۶۰ درصد متغیر بوده است [۱۰و۱۷].

پس از ترکیب نواحی مشترک، ۸۸ ناحیه CNV در گوسفند شناسایی شد که ۵۳ مورد آنها فقط در یک حیوان مشاهده شد. در گاو نیز ۳۸۸ ناحیهCNV شناسایی شد که ۲۴۷ مورد آن منحصر به گونه گاو بودند. در هر دو آزمایش نواحی خاص گونه‌ای از ادامه آنالیزها حذف شد و نسبت نواحی CNV پر تکرار به کل تعداد نواحی CNV در هر دو گونه تقریبا برابر بود (در گاو ۳۶/۳ درصد در مقایسه با ۳۹/۷ درصد در گوسفند). تعداد کل نواحی CNV در گاو تقریبا ۴/۴ برابر گوسفند بود. میانگین و میانه طول نواحی CNV پرتکرار در دو گونه مشابه بود (جدول ۱).

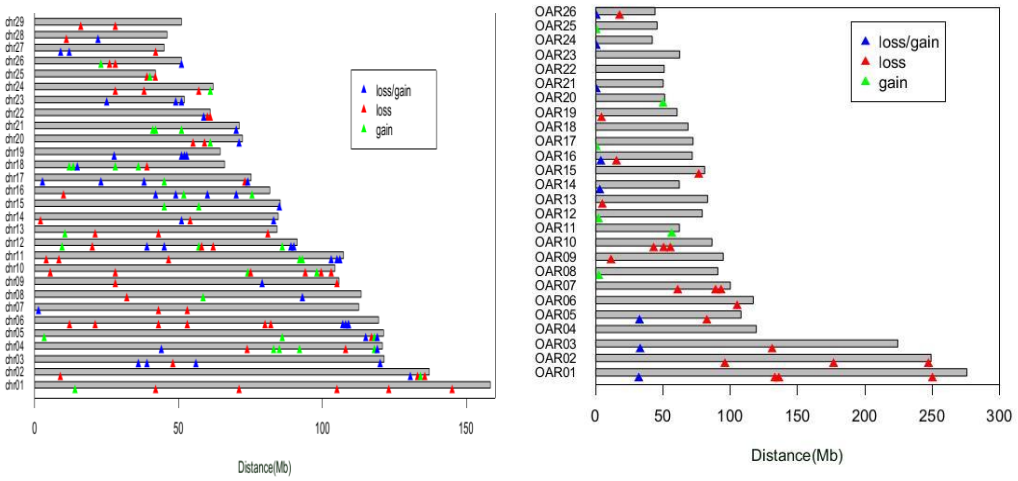
مقایسه حدود نواحی CNV ژنوم گوسفند پس از معادل سازی با ژنوم گاو نشان داد که فقط در هشت مورد با نواحی CNV ژنوم گاو بر روی کروموزم‌های یک، دو، سه، هفت، هشت، ۱۱، ۲۰، ۲۴ همپوشانی داشتند (جدول۲). در سایر کروموزم‌ها بسیاری از نواحی در دو گونه با تفاوت چند صد کیلو جفت باز در نزدیکی یکدیگر قرار داشتند. با توجه به تشابه زیاد ژنوم گاو و گوسفند، تشابه قابل توجهی بین نواحی CNV مورد انتظار بود. از نتایج حاصل دو مطلب استنباط می‌شود. (۱) با توجه به برابر نبودن طول کروموزوم‌ها در دو گونه، تنوع‌های نزدیک به هم احتمالا مشابه هستند، که شاید این نواحی همان نواحی پر CNV بین‌گونه‌ای باشند. (۲) به نظر می‌رسد با توجه به همپوشانی ضعیف نواحی CNV شناسایی شده در مطالعات مختلف، تعداد زیادی از این تنوع در ژنوم هر دو گونه هنوز ناشناخته است، که نیاز به مطالعات بیشتری دارد.

جدول ۲. همپوشانی نواحی CNV ژنوم گوسفند با نواحی این تنوع در ژنوم گاو

کروموزوم	ناحیه CNV در ژنوم گوسفند (ISGC_Ovis_aris_3.0)	کروموزوم	معادل ناحیه CNV در ژنوم گاو (UMD3.1)	نواحی مشابه در ژنوم گاو هلشتاین در تحقیق حاضر
۱	۴۸۹۰۴۵-۷۴۹۵	۳	۱۲۱۳۸۲۴۴۵-۱۲۰۹۰۷۴۴۰	۱۲۱۳۷۴۸۲۵-۱۲۰۱۲۲۱۷۶
۲	۲۴۷۲۹۰۶۴۹-۲۴۷۲۹۰۵۱۵	۲	۱۳۵۱۱۱۶۱۱-۱۳۵۰۷۱۶۴۶	۱۳۵۵۴۰۱۲۶-۱۳۵۰۹۶۰۶۰
۳	۴۴۲۵۲۳-۸۳۸۴۵	۱۱	۱۰۵۵۹۴۷۳۸-۱۰۵۲۷۰۱۹۸	۱۰۶۴۳۲۸۳۲-۱۰۵۵۵۷۳۰۴
۷	۸۹۷۲۳۱۷۲-۸۹۵۶۶۰۴۰	۱۰	۹۳۸۶۳۹۱۱-۹۳۷۰۰۲۱۹	۹۴۰۸۳۵۲۵-۹۳۹۸۸۹۳۰
۸	۲۳۹۶۶۸۱-۲۲۴۵۳۴۷	۹	۱۵۵۲۵۳۲۳-۱۵۳۶۵۴۵۴	۱۵۵۵۳۸۷۳-۱۵۳۴۰۵۵۸
۱۱	۵۶۳۵۶۵۱۸-۵۶۲۸۶۹۲۹	۱۹	۵۸۰۴۳۱۱۷-۵۷۹۷۲۷۸۷	۵۷۹۸۰۶۹۷-۵۷۷۴۸۵۵۲
۲۰	۵۰۶۰۵۵۱۳-۵۰۵۰۳۸۷۴	۲۳	۵۱۶۸۱۷۳۱-۵۱۵۵۷۱۲۱	۵۱۷۴۶۳۹۴-۵۱۲۱۶۷۴۴
۲۴	۴۸۱۷۹۴-۱۷۴۸۱۲	۲۵	۵۷۶۹۶۸-۲۳۵۶۶۴	۱۸۳۷۲۱۳-۱۳۸۳۸۲

مقایسه حدود نواحی CNV این پژوهش با حدود نواحی CNV شناسایی شده در یک مطالعه دیگر که با تراشه 50k گوسفندی انجام شد همپوشانی نداشت[۱۹]. همچنین در یک تحقیق دیگر که با روش متفاوتی انجام شده بود همپوشانی یافت نشد[۷]. این موضوع می‌تواند دلائل متعددی داشته باشد. استفاده از اسمبلی ISGC_Ovis_aris_1.0 ژنوم گوسفند برای معادل سازی نواحی CNV با اسمبلی Btau.4.0 ژنوم گاو یک عامل است[۱۹] هر دو اسمبلی نسبت به اسمبلی‌های استفاده شده در این تحقیق دارای کانتیگ‌های بی‌محل و یا بامحل اشتباه بیشتری هستند که می‌تواند منجر به عدم همپوشانی شود. همچنین تفاوت در روش استفاده می‌تواند دلیل این اختلاف باشد[۷]. در این تحقیق [۷] علاوه بر متفاوت بودن اسمبلی از روش هیبدریداسیون مقایسه‌ای ژنوم استفاده شد که دقت و قدرت بیشتری در شناسایی این نوع تنوع دارد زیرا توزیع کاوشگرهای آن در سراسر ژنوم یکنواخت‌تر از تراشه 50k است و تعداد کاوشگر بیشتری برای نواحی غیر پلی‌مورف دارد [۱۷] بنابراین تعداد نواحی تنوع شناسایی شده بیشتر و طول آن کوتاه‌تر بود. لذا حدود نواحی CNV جابجا شده وهمپوشانی ضعیف نواحی در دو مطالعه قابل انتظار است. مشابه این حالت در مطالعه نواحی CNV در ژنوم گاو نیز مشاهده شده است[۱۱]، بطوری که نواحی شناسایی شده بر روی ژنوم گاو با نواحی CNV در سایر مطالعات بین ۱۸/۹– ۷/۸۴ درصد همپوشانی داشت[۲، ۶، ۱۷]. این مطالعات نشان داد تشابه نواحی شناسایی شده در گاو هلشتاین کمتر از ۲۰ درصد است. این موضوع علاوه بر تایید همپوشانی ضعیف بین نواحی CNV به این نکته نیز اشاره می‌کند که هنوز تعداد قابل توجهی از نواحی حاوی این تنوع در ژنوم گاو شناسایی نشده است، با توجه به مطالعه کمتر بر روی CNVهای ژنوم گوسفند به نظر می‌رسد این موضوع در مورد ژنوم گوسفند بیشتر باشد. از این رو احتمال همپوشانی نواحی در بین دو گونه کمتر می‌شود.

برای بررسی نحوه توزیع نواحی CNV بر روی کروموزم‌ها، پروفایل آن با استفاده از اسمبلی Ovis_aris_3.0 ژنوم گوسفند (شکل ۱–راست) واسمبلی UMD3.1 ژنوم گاو (شکل ۱–چپ) ترسیم شد. نقشه‌های حاصل نشان داد که تراکم این نواحی در ژنوم گاو بیشتر از ژنوم گوسفند بود. با توجه به حجم ژنوم گاو و تعداد نمونه و همچنین خلوص نژادی آن این مطلب قابل پیش بینی بود.



شکل ۱. پروفایل توزیع نواحی CNV در ژنوم گوسفند(راست) و گاو(چپ)

علاوه بر این در تراشه 50k گوسفندی در مقایسه با نوع گاوی بسیاری از چند شکلی‌های تک نوکلئوتیدی به درستی جایابی نشده اند و یا بی‌محل هستند، این موضوع باعث کاهش در تعداد نواحی CNV شناسایی شده می‌شود. در هر دو گونه توزیع نواحی بصورت تصادفی نبوده و بسته به کروموزوم و محل کروموزومی تغییر می‌کند بطوری که تعداد زیادی از آنها در نواحی ساب تلومریک و پری‌سانترومریک قرار داشتند. شاید دلیل دیگر تجمع بیشتر در این نواحی این باشد که کلیه ۲۹ جفت کروموزوم اتوزوم گاو اکروسانترومریک هستند. اکروسانترومریک یعنی سانترومر در یکی از دو انتهای کروموزوم واقع شده است و بازوی کوتاه بسیار کوچک است [۳]. نواحی پری-سانترومریک و ساب‌تلومریک ۳/۴ درصد توالی ژنوم را شامل می‌شوند که ۲/۵-۱/۵ برابر CNV بیشتری نسبت به سایر نواحی دارند و ۶/۹-۶/۷ درصد نواحی پلیمورف ژنوم می‌باشند [۱۷]. البته ممکن است دلیل دیگر تراکم بیشتر نواحی CNV در نواحی تلومری و سانترومی، تراشه مورد استفاده باشد. در یک تحقیق مشخص شد که نواحی تنوع شناسایی شده با تراشه متراکم‌تر توزیع یکنواخت‌تری در ژنوم دارد ولی تنوع‌های شناسایی شده با تراشه 50k بیشتر در نواحی تلومری و سانترومی قرار دارد[۱۰].

دو ناحیه CNV با حدود ۱۸۰۲۹۵-۴۸۱۰۸۸ در کروموزوم ۱۲ و ۱۵۶۶۹۸-۳۸۵۳۰۵ در کروموزوم ۱۶ به ترتیب ژن‌های مایوگ (MYOG) و MiR103 را در ژنوم گوسفند بطور کامل احاطه کرده بودند. ژن مایوگ از خانواده مایود (MYoD) می‌باشد که در سنتز پروتئین میوفیبریل در عضلات اسکلتی و تنظیم تعداد میوفیبریل‌ها نقش دارد [۲۴]. همچنین بیان مضاعف ژن MiR103در غدد پستانی بز منجر به افزایش نسخه‌برداری از ژن‌های مرتبط با سنتز چربی شیر می‌شود که این خود باعث افزایش اندازه قطرات چربی، انباشتگی تری‌گلیسیرید و اسیدهای چرب غیراشباع در سلول‌های اپی‌تلیال غدد پستانی می‌شود [۱۶]. با توجه به پوشش ضعیف ژن‌های شناسایی شده در ژنوم گوسفند، برای شناسایی ژن‌های بیشتر، نواحی CNV ژنوم گوسفند (ISGC_Ovis_aris_1.0) به معادل آن در ژنوم گاو(Baylor 4.0/Bos Tau4) و انسان (GRCh38/hg38) تبدیل شد. نتایج نشان داد که ۱۶ ناحیه CNV شناسایی شده در ژنوم گوسفند پس از تبدیل به معادل آن در ژنوم انسان با ۱۱۰ ژن رفرنس و ۱۰ ناحیه CNV با ۴۰ ژن رفرنس در ژنوم گاو همپوشانی داشت. چون بسیاری از ژن‌های ژنوم گوسفند شناسایی نشده است و تشابه قابل توجهی در نواحی رمز کننده در ژنوم هر دو گونه گوسفند و گاو وجود دارد[۱۴] و از طرفی حفاظت CNVها در بین پستانداران می‌تواند منجر به ابقا و اکتساب یک دز خاص از یک ژن شود که این ژن یا خانواده آن در پستاندارن تکرار می‌شود[۱۳ و۱۱]، لذا اگر ژن‌های مهمی در ژنوم گاو با نواحی CNV همپوشانی داشته یا در عدم تعادل پیوستگی باشند شاید بتوان از روی محل نواحی CNV ژنوم گاو اقدام به شناسایی آن ژن‌ها و یا QTLها در ژنوم گوسفند کرد. به هر حال برای نتیجه‌گیری دقیقتر مطالعات وسیع‌تر و جامع‌تری نیاز است.

بررسی مسیرهای متابولیکی ژن‌های موجود در نواحی CNV فوق نشان داد که تعدادی از این ژن‌ها در مسیرهای متابولیکی ترانسدوکسیون بویای، سرطان در انسان و متابولیسم پورین و پریمیدین در گونه گاو و بیوسنتز استروئیدها در انسان و گاو نقش دارند. علاوه بر این تعدادی از ژن‌های شناسایی شده با بیماری‌های نظیر استئوپروسیز، کم خونی، اریتروسیتوزیس و مطالعات ارتباطی مقاومت به انسولین در انسان مرتبط بود. ژن‌های حاصل در یک پژوهش با بیماری ژنتیک مندلی و طول عمر در انسان ارتباط داشت[۷] که مشابه نتایج این پژوهش بود. پس از بررسی حدود نواحی CNV در وب سایت

QTLdb از ۳۵ ناحیه CNV در ژنوم گوسفند ۳۲ ناحیه دارای معادل در ژنوم گاو بود. از این تعداد فقط ۲۴ ناحیه بطور کامل یا جزئی در محدود ۶۲ QTL بودند. این QTLها با صفاتی از قبیل، تولید شیر، درصد چربی شیر، درصد پروتئین شیر، میزان تولید چربی شیر، آسان زایی، وزن لاشه، وزن بدن، ماربلینگ و ... در ارتباط بودند. بررسی CNV ژنوم گوسفند برای شناسایی QTLها فقط وجود چهار QTL را نشان داد که با صفات تولید مثل، سطح ایمنوگلوبولین A و اسید های چرب غیر اشباع شیر مرتبط بودند. برای بررسی بیشتر در خصوص ارتباط CNVها با ژن‌ها و QTLها در مرحله اول تایید CNVهای شناسایی شده با qPCR ضروری است.

پس از حذف تنوع‌ها با یک تکرار CNV ۹۰۴ برای گوسفند و ۷۴۴ مورد برای گاو شناسایی شد. تعداد CNV شناسایی شده در ژنوم گوسفند تقریبا چهار برابر CNV شناسایی شده در ژنوم گاو بود، ولی کل طول CNV شناسایی شده پرتکرار پس از حذف دام‌های با CNV مشترک در گوسفند ۸/۵ Mb و در گاو ۸۰ Mb بود. پس از ادغام نواحی مشترک، ۵۳ ناحیه CNV پر تکرار در گوسفند و ۲۴۷ مورد در گاو شناسایی شد و مقایسه حدود نواحی CNV ژنوم گوسفند با ژنوم گاو فقط هشت مورد همپوشانی نشان داد و مقایسه با نواحی CNV شناسایی شده در سایر مطالعات همپوشانی موجود ضعیف بود. در هر دو گونه توزیع نواحی CNV در مناطق ساب‌تلومریک و پری‌سانترومریک بیشتر بود. بررسی نواحی CNV گوسفند در ژنوم انسان و گاو نشان داد که ۱۶ ناحیه CNV با ۱۱۰ ژن رفرنس در ژنوم انسان و ۱۰ ناحیه CNV با ۴۰ ژن رفرنس در ژنوم گاو همپوشانی داشت. به علاوه ۲۴ ناحیه CNV بطور کامل یا جزئی با ۶۲ QTL همپوشانی داشتند، که با صفاتی از قبیل، تولید شیر، درصد چربی شیر، درصد پروتئین شیر، میزان تولید چربی شیر، آسان‌زایی، وزن لاشه، وزن بدن، ماربلینگ و ... ارتباط دارند.

تشکر و قدردانی

بدینوسیله از لوکا فونتانزی استاد دانشگاه بلونیا- ایتالیا بخاطر ارائه داده‌های این پژوهش و راهنمایی در طی مراحل تجزیه و تحلیل این داده‌ها ، همچنین از مسئولین موسسه ANAFI بابت تامین اعتبار این پروژه تشکر و قدردانی می‌شود.

منابع

۱. نصرتی م، طهمورث پور م، فونتانزی ل و نصیری م ر (۱۳۹۵) بررسی فاز عدم تعادل لینکاژی بر روی کرموزوم شماره ۶ گاوهای نژاد هلشتاین. بیوتکنولوژی کشاورزی کرمان. ۹۸-۸۳:(۲) ۸

2. Bae JS, Cheong HS, Kim LH, NamGung S, Park TJ, Chun JY, et al. (2010) Identification of copy number variations and common deletion polymorphisms in cattle. BMC Genomics 11: 232.

3. Bohmanova J, Sargolzaei M, and Schenkel FS (2010) Characteristics of linkage disequilibrium in North American Holsteins. BMC Genomics 11.421.

4. Crawford AM, Dodds KG, Ede AJ, Pierson CA, Montgomery GW,Garmonsway HG, Beattie AE, Davies K, Maddox JF, Kappes SW, et al. (1995) An autosomal genetic linkage map of the sheep genome. Genetics 140:703-724.

5. Dalrymple BP, Kirkness EF, Nefedov M, McWilliam S, Ratnakumar A, Barris W, Zhao S, Shetty J, Maddox JF, O'Grady M, Nicholas F, Crawford AM, Smith T, de Jong PJ, McEwan J, Oddy VH, Cockett NE (2007) Using comparative genomics to reorder the human genome sequence into a virtual sheep genome. Genome Biology 8 R152.

6. Fadista J, Thomsen B, Holm L, and Bendixen C (2010) Copy number variation in the bovine genome. BMC Genomics 11:284.

7. Fontanesi L, Beretti F, Martelli PL, Colombo M, Dall'Olio S, Occidente M, Portolano B, Casadio R, Matassino D, and RussoV, (2011) A first comparative map of copy number variations in the sheep genome. Genomics 97: 158-165.

8. Fontanesi L, Martelli PL, Beretti F, Riggio V, Dall'Olio S, Colombo M, Casadio R, Russo V and Portolano B (2010) An initial comparative map of copy number variations in the goat (Capra hircus) genome. BMC genomics 11: 639.

9. Hastings PJ, Lupski JR, Rosenberg SM, and Ira G (2009) Mechanisms of change in gene copy number. Nature Reviews Genetics 10(8): 551–564.

10. Hou Y, Bickhart DM, Hvinden ML, Li C, Song J, Boichard AD, Fritz S, Eggen A, DeNise S, Wiggans GR, Sonstegard TS, Van Tassell CP, and Liu, GE (2012) Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. BMC Genomics 13: 1471-2164.

11. Hou Y, Liu GE, Bickhart DM, Cardone MF, Wang K, Kim ES, Matukumalli LK, Ventura M, Song J, Vanradan PM, Sonstegard TS, and Van Tassell CP. (2011) Genomic characteristics of cattle copy number variations. BMC genomics 12:127.

12. Ihara N, Takasuga A, Mizoshita K et al. (2002) Mapping of over 1100 bovine polymorphic microsatellite markers to the USDA-MARC cattle linkage map. Proceeding International Society of Animal Genetics, D-160.

13. Kijas JW, Barendse W, Barris W, Harrison B, McCulloch R, McWilliam S,and Whan V (2011) Analysis of copy number variants in the cattle genome. Gene 482: 73–77.

14. Kijas JW, Menzies M, Ingham A (2006) Sequence diversity and rates of molecular evolution between sheep and cattle genes. Animal Genetic 37, 171–174.

15. Li X, Tan L, Liu X, Lei S, Yang T, Chen X, Zhang F, Fang Y, Guo Y, Zhang L, Yan H, Pan F, Zhang Z, Peng Y, Zhou Q, He L, Zhu L, Cheng J, Zhang L and Liu L (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, Luo J, Zhang L, Wang W, Gou D (2013) MiR-103 Controls Milk Fat Accumulation in Goat (Capra hircus) Mammary Gland during Lactation. PLoS ONE 8(11): e79258. doi: 10.1371/journal.pone.0079258.

17. Liu GE, Hou Y, Zhu B, Cardone MF, Jiang J, et el. (2010) Analysis of copy number variations among diverse cattle breeds. Genome Research Journal 20: 693-703.

18. Liu GE, Van Tassel CP, Sonstegard TS, Li RW, Alexander LJ, Keele JW, Matukumalli LK, Smith TP, and Gasbarre LC (2008) Detection of germ line and somatic copy number variations in cattle. Journal of Development Biology 132: 231-237.

19. Liu J, Zhang L, Xu L, Ren H, Lu J, Zhang X, Zhang S, Zhou S, Wei C, Zhao F, and Du L (2013) Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. BMC Genomics 14:229.

20. Mills RE, Walter K, Stewart C, Handsaker RE et al. (2011) Mapping copy number variation by population-scale genome sequencing. Nature 470: 59-65.

21. Miziara MN, Riggs PK, Amaral ME (2004) Comparative analysis of noncoding sequences of orthologous bovine and human gene pairs. Genetic Molecular Research 3, 465–473.

22. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH. Carson AR, Chen W et al (2006) Global variation incopy number in the human genome. Nature 444: 444-454.

23. Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, Ajmone-Marsan P (2008) Are cattle, sheep, and goats' endangered species? Molecular Ecology 17(1):275-84.

24. Zhang Z, Xu F, Zhang Y, Li W, Yin Y, Zhu C, Du L, Elsayed AK, Li B, (2014) Cloning and expression of MyoG gene from Hu sheep and identification of its myogenic specificity. Molecular Biology Report.41(2):1003-13.

Journal of
Animal Production
(College of Abouraihan – University of Tehran)
Vol. 19 ■ No. 1 ■ Spring 2017

Comparison of copy number variation profile in sheep with Holstein breed

Maryam Nosrati *

Assistant Professor, Department of Agricuture, Payame Noor University, Tehran, IRAN

Received: May 7, 2016 **Accepted: June 12, 2016**

Abstract

The aim of this study was to compare copy number variation (CNV) in some Italian sheep breeds with Holstein cows. Blood samples were collected form 580 Holstein bulls and 360 different Italian sheep breeds and then DNA was extracted. The SNPs genotypes across the genome were determined by ovine and bovine 50K BeadChip in both species. By PennCNV, the 904 CNV with mean and median size of 154.7 kb and 103.7 kb in sheep and 744 CNV with mean and median of 213.7 kb and 124.5 kb in Holstein were detected, respectively. A total length of this variation were 8.5 Mb in sheep and was 80 Mb in cattle. After merging similar regions, the 35 and 141 non- unique copy number variation regions (CNVR) were detected in sheep and cows, respectively which 20% of sheep's CNVRs overlapped with cattle's CNVRs. In addition to, these regions in Sheep genome were partially or completely overlapped with 40 RefGen & 62 QTL in cattle and 110 RefGen in Human. CNVRs distribution in both species was more in subtelomeric and pericentromeric regions. According to these results, it could be possible to use bovine CNVRs which were closely linked to genes for characterizing functional genes and QTLs in sheep, however more studies are need in this regards.

Keywords: CNV, Comparative genome profile, Genome variation, Holstein cows, Sheep.

Email: m_nosrati@pnu.ac.ir