دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشهای تولیدات دامی

" مقاله پژوهشی"

پیش‌بینی جزایر CpG و متیلاسیون DNA در ژنوم گاو با استفاده از فراتحلیل ریزآرایه DNA و پویش ژنومی

زهرا بیرانوند[۱]، سید ضیاالدین میرحسینی[۲]، مصطفی قادری-زفره‌ای[۳]، سیدحسین حسینی‌مقدم[۴]
آرش فاضلی[۵] و کیانوش زرین کاویانی[۶]

۱- دانش‌آموخته دکتری ژنتیک و اصلاح دام، دانشگاه گیلان، رشت، ایران
۲ و ۴- استاد و استادیار، ژنتیک مولکولی، گروه علوم دامی، دانشگاه گیلان، رشت، ایران
۳- دانشیار بیوانفورماتیک و ژنتیک و اصلاح نژاد، گروه علوم دامی، دانشگاه یاسوج، یاسوج، ایران، (نویسنده مسوول: Mosmos741@yahoo.com , mghaderi@yu.ac.ir)
۵- دانشیار ژنتیک مولکولی، گروه زراعت و اصلاح نباتات، دانشگاه ایلام، ایلام، ایران
۶- عضو هیات علمی دانشگاه ایلام

تاریخ دریافت: ۱۳۹۷/۰۲/۲۵ تاریخ پذیرش: ۱۳۹۹/۰۱/۲۹
صفحه: ۹۵ تا ۱۰۶

چکیده

متیلاسیون DNA نوعی تغییر فراژنتیکی است که به‌طور مستقیم بر روی DNA تاثیر می‌گذارد. در پستانداران متیلاسیون DNA برای تکامل جنینی و تمایز سلول بنیادی ضروری است و این پدیده اساساً درون جزایرCpG اتفاق می‌افتد. در این پژوهش برای بررسی نیم‌رخ متیلاسیون DNA ژنوم گاو، از دو روش استفاده شد. در روش اول نیم‌رخ متیلاسیون DNA ژن‌های با بیان متفاوت حاصل از فراتحلیل داده‌های ریزآرایه DNA بیماری ورم پستان گاو شیری به‌دست آمد. برای انجام فراتحلیل داده‌های ریزآرایه DNA در این روش، از بسته metaDE در محیط R استفاده شد. سپس جهت پیش‌بینی جزایر CpG در ژن‌های با بیان متفاوت از ۵ الگوریتم شامل TJ, GF, CpG cluster, HMM ,GHMM استفاده شد. در روش دوم نیم‌رخ متیلاسیون DNA با استفاده از پویش کل ژنوم گاو انجام گرفت. همچنین جهت پیش‌بینی جزایر CpG متیله شده در کل ژنوم، ابتدا توسط الگوریتم HMM جزایر CpG در ژنوم گاو و برای هر کروموزوم برآورد شدند و سپس هم‌پوشانی CpG با Hypo/Hyper-Methylation توسط پایگاه برخط Galaxy محاسبه شد. نتایج روش اول نشان داد که از میان ۳۲ ژن با بیان متفاوت، ۱۴ ژن دارای جزایر CpG متیله شده بودند. این ژن‌ها، شامل ژن‌های LTF, APP, CCL5, CD40, CSNK1D, CX3CL1, DAPP1, NFKBIZ, S100A9, ISG15, MAP3K8, MX1, RDAD2, ZC3H12A بودند. نتایج روش دوم تعداد ۹۰۶۶۸ Hypo/Hyper-Methylation را در ژنوم گاو برآورد نمود که تعداد ۹۹۴۲ (۱۰.۹۶٪) جزیره CpG با Hypo/Hyper-Methylation دارای هم‌پوشانی بودند و به‌عنوان CpG متیله شده در نظر گرفته شدند. مقایسات ژنومی بین گونه‌ای متیلاسیون DNA نشان داد که نیم‌رخ کلی متیلاسیون DNA در اکثر گونه‌های مورد بررسی تقریبا مشابه هم بودند و به‌نظر می‌رسد که نیم‌رخ کلی متیلاسیون DNA بین گونه‌های مختلف به‌صورت محافظت شده باشد. حاصل این پژوهش نشان داد کنکاش متیلاسیون DNA در بیماری‌هایی که میزان وراثت‌پذیری پایینی دارند و بیشتر تحت تاثیر فرآیندهای فراژنتیک هستند، ضروری به‌نظر می‌رسد.

واژه‌های کلیدی: متیلاسیون DNA، جزایر CpG، فراژنتیک، گاو، فراتحلیل، الگوریتم

مقدمه

متیلاسیون DNA، از مهمترین ابزارهای فراژنتیک به‌شمار می‌رود و تنها تغییر فراژنتیکی است که به‌طور مستقیم روی DNA تاثیر می‌گذارد. متیلاسیون DNA مهم‌ترین نشانگر فراژنتیکی در بیماری‌هایی مانند ورم پستان گاو شیری می‌باشد که میزان وراثت پذیری در این گونه صفات کم است. نواحی کوچکی از DNA (۱ تا ۲٪) که جزایر CpG نامیده می‌شوند و عمدتاً در نواحی راه‌انداز ژن‌ها قرار دارند، شدیداً از متیلاسیون محافظت می‌شوند و با نقاط آغاز رونویسی در حدود ۷۰ درصد از ژن‌های انسانی مرتبط هستند (۸). توزیع CpG در ژنوم مهره‌داران همگن نبوده و بیشتر ژنوم خالی از CpG است. بیشتر جزایر CpG در ژنوم پستانداران و انسان غیرمتیله است تا مانعی در راه رونویسی از ژن ایجاد نکند. الگوی متیلاسیون از فردی به فرد دیگر و از بافتی به بافت دیگر متفاوت است (۱۳). متیلاسیون DNA شامل افزوده شدن گروه متیل روی باز سیتوزین، پس از ساخته شدن DNA می‌باشد و در نتیجه این انتقال آنزیمی، یک گروه متیل (-CH3) از S- آدنوزیل متیونین (دهنده متیل) به کربن موقعیت ۵ سیتوزین موجود در دی‌نوکلئوتید CpG منتقل می‌شود که این فرآیند برگشت‌پذیر است (۱۸). این واکنش آنزیمی توسط DNA متیل ترانسفرازها (DNA methyltransferases) DNMT انجام می‌گیرد. DNMT-که آنزیم‌های انتقال‌دهنده‌ی گروه متیل بر روی دی‌نوکلئوتیدهای CpG هستند هماهنگ با همانندسازی DNA عمل کرده تا الگوی متیلاسیون DNA را حفظ کنند. DNA متیل ترانسفرازها دی‌نوکلئوتیدهای CpG متیله شده بر روی رشته‌ی مادری را شناخته و دی‌نوکلئوتیدهای CpG مرتبط بر روی رشته‌ی دختری را متیله می‌کنند (۱۷) و این الگو را در طی فرآیند تمایز حفظ و به نسل بعدی سلول منتقل می‌کنند. در پستانداران، ۵ عضو از خانواده DNMT شناسایی شده است: DNMT1, DNMT2, DNMT3a, DNMT3b, DNMT3L که از بین این‌ها فقط DNMT1, DNMT3a, DNMT3b فعالیت متیل ترانسفرازی دارند. الگوی توارث متیلاسیون DNA همانند الگوهای مدیفیکاسیون‌های هیستونی (Modifications Histone)، توسط ماشین فراژنتیکی واسطه‌گری می‌شود. الگوی متیلاسیون طی تکامل

[Downloaded from rap.sanru.ac.ir on 2025-12-17]
[DOI: 10.52547/rap.11.29.95]

بافت ایجاد می شود (۲۳) و زمانی‌که این الگو در یک سلول ایجاد شد، به شکل پایداری طی همانندسازی DNA در هر تقسیم سلولی، حفظ می‌شود. بنابراین وقتی سلول‌ها حاوی اطلاعات ژنتیکی یکسانی باشند، ماهیت متمایز خود را حفظ می‌کنند. طی تکامل و در قالب برنامه‌ریزی مجدد ژنومی، سیتوزین متیله‌ی یک دی‌نوکلئوتید CpG طی دآمیناسیون (Diamination) به تیمین تبدیل می شود. الگوهای متیلاسیون به‌طور نزدیکی با الگوهای بیان ژن مرتبط هستند. نواحی شدیداً متیله شده‌ی ژنوم با نوعی از سازمان‌بندی کروماتین که موجب مهار رونویسی می‌شود، مرتبط است. در انسان، چنین نواحی متیله شده‌ای اغلب حاوی توالی‌های تکراری است. همچنین متیلاسیون مانع از رونویسی توالی‌های مزاحم می شود و از جابجایی آنها در ژنوم جلوگیری می‌کند (۳۵). در مقابل، جزایر CpG غیرمتیله‌ی ژن‌ها با نواحی فعال کروماتین که شدیداً رونویسی می‌شوند، مرتبط است. معمولاً در تمامی بافت‌های طبیعی این جزایر غیرمتیله بوده و غالباً انتهای '۵ تعدادی از ژن‌ها را پوشش می‌دهند. زمانی‌که جزایر CpG ناحیه راه‌انداز فاقد متیلاسیون بوده و فاکتورهای نسخه‌برداری مناسب به آن دسترسی دارند، بیان ژن میسر می‌شود؛ در مقابل متیلاسیون جزایر CpG راه‌انداز با ساختار بسته کروماتین و در نتیجه عدم نسخه‌برداری ژن‌های مربوطه همراه خواهد بود. الگوی متیلاسیون در ژنوم پستانداران به این‌صورت است که دی‌نوکلئوتیدهای CpG در جزایر CpG به‌صورت غیرمتیله و در بدنه ژنوم متیله می‌باشند و تغییر این الگو یعنی هایپومتیلاسیون (Hypo-methylation) وسیع بدنه ژنومی منجر به ناپایداری کروموزومی و تغییر بیان ژن‌ها گردیده و یا هایپرمتیلاسیون (Hyper-methylation) جزایر CpG منجر به بروز ژن‌های دخیل درسرطان‌ها و یا بیماری‌های دیگر می‌شود. متیلاسیون DNA در فرآیندهای تمایز سلولی، نشانه‌گذاری ژنومی، مرگ برنامه‌ریزی شده سلول، سرطان‌ها، پیری، غیرفعال شدن تصادفی کروموزوم X و غیره نقش تعیین‌کننده‌ای دارد (۳۴).

در بررسی‌های صورت گرفته روی ژنوم گاو، نشان داده شده است که سطح متیلاسیون DNA در بدنه ژن‌ها نسبتا بالا و در راه‌انداز ژن‌ها پایین بود (۳۲،۱۷). اکثر جزایر CpG غیرمتیله بودند و جزایر CpG متیله در اینترون و بدنه ژن‌ها قرار داشتند. همچنین، همبستگی منفی بین سطح بیان ژن و سطح متیلاسیون در اطراف مکان‌های شروع رونویسی (Transcription Start Site) ژن‌ها مشاهده گردید، اما رابطه بین متیلاسیون DNA بدنه ژن و بیان ژن غیر یکنواخت بود. در نهایت نشان داده شد که ژن‌های با بیان متوسط دارای بالاترین سطح متیلاسیون در بدنه خود، اما ژن‌های دارای بیان بالا و پایین و ژن‌های خاموش دارای سطح متیلاسیون متوسط در بدنه ژن را داشتند و ژن‌های دارای بیشترین بیان، کمترین سطح متیلاسیون در بدنه ژن را داشتند (۳۲،۱۷). در پژوهشی بر روی ژنوم جوجه‌های گوشتی مشخص شد که کمترین تراکم متیلاسیون در جزایر CpG واقع در راه‌انداز ژن‌ها است و در بدنه ژن‌ها تراکم متیلاسیون جزایر واقع در اینترون بیشتر از اگزون و UTRs می‌باشد و جزایر CpG متیله شده در نواحی بین ژنی توزیع شده نیز قابل توجه هستند (۱۶). این نتایج با نیم‌رخ متیلاسیون به‌دست آمده در ژنوم خوک، مشابه بود (۲۱). کلدری و همکاران در پژوهش بر روی گوسفند نشان دادند که نواحی نزدیک TSS (Transcription Start Site) ژن‌ها دارای متیلاسیون پایین‌تری هستند (۴). لو و همکاران بیماری مارک در طیور را مورد بررسی قرار دادند. عامل بیماری مارک ویروسی به اسم آنکوویروس (Oncovirus) است که باعث تومورهای لنفاوی در جوجه‌های حساس می‌شود و ممکن است بر روی ثبات فراژنتیک در ژن CD4 تاثیر بگذارد. نشان داده شد که متیلاسیون راه‌انداز ژن CD4 در در جوجه‌های سویه L72 پس از آلودگی به ویروس مارک در سطح پایینی تنظیم می‌شود. متیلاسیون DNA که در سطح پایین راه‌انداز CD4 تنظیم می‌شود همبستگی منفی با تنظیم بیان ژن CD4 در طحال در سویه L72 را نشان داد (۲۳). آنچه از تمامی پژوهش‌ها این بخش قابل نتیجه‌گیری است این‌است که نیم‌رخ کلی متیلاسیون در اکثر پستانداران و حتی پرندگان مشابه هم می‌باشد و به‌نظر می‌رسد که بین گونه‌های مختلف به‌صورت محافظت شده است (۳۱،۱۷،۵).

مواد و روش‌ها

در این پژوهش برای درک بهتر نیم‌رخ متیلاسیون DNA ژنوم گاو، از دو روش بررسی متیلاسیون در ژن‌های متفاوت بیان‌شده حاصل از فراتحلیل داده‌های ریزآرایه DNA در بیماری ورم پستان گاو شیری و بررسی متیلاسیون DNA در کل ژنوم گاو استفاده می‌شود.

روش متکی به فراتحلیل داده‌های ریزآرایه DNA در بیماری ورم پستان گاو

در ابتدا ۵ آزمایش ریزآرآیه DNA با شماره دسترسی GSE15019, GSE24217, GSE24560, GSE25413, GSE50685 از پایگاه GEO استخراج شدند (جدول ۱) و با استفاده از بسته metaDE (https://rdrr.io /github /meta Omics/MetaDE/) متکی به زبان R فراتحلیل روی آن‌ها انجام شد. انجام فراتحلیل با روش فیشر و در سطح معنی‌داری یک درصد ($p<0.01$) انجام گرفت. هدف اصلی از انجام فراتحلیل در این پژوهش، بالا بردن میزان اعتماد به نتایج حاصل از ژن‌های با بیان متفاوت بود، از این‌رو فرا تحلیل روی آزمایش‌های مختلف انجام شد.

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

جدول ۱- اطلاعات مربوط به آزمایش‌های ریزآرایه DNA

Table 1. Information on DNA microarray experiments

آزمایش	نمونه سالم	نمونه بیمار	پلتفرم
	۸	۶	
GSE25413	۲۷	۶۱	Affymetrix
GSE24560	۲۳	۲۶	Affymetrix
GSE24217	۵	۵	Affymetrix
GSE15019	۵	۱۵	Affymetrix
GSE50685			

استخراج جزایر CpG بر اساس ژن‌های با بیان متفاوت

برای استخراج جزایر CpG بر اساس ژن‌های با بیان متفاوت، ابتدا توالی هر کدام از ژن‌ها از پایگاه‌های داده NCBI و ENSMBLE پیاده‌سازی شدند. سپس جهت پیش‌بینی جزایر CpG در ژن‌های با بیان متفاوت از ۵ الگوریتم مختلف شامل الگوریتم گاردینر گاردن و فرامر (۹)، الگوریتم تاکایی و جونز (۳۳)، الگوریتم CpGcluster (۱۵)، الگوریتم مدل مخفی مارکف (۳۷) و الگوریتم مدل مخفی مارکف تعمیم یافته (۳۹) استفاده شدند.

با استفاده از الگوریتم‌های یاد شده جزایر CpG مربوط به هر ژن به‌دست آمد. نتایج الگوریتم‌های مختلف با یکدیگر مقایسه شدند و نتایج جزایر CpG برای ژن‌های با بیان متفاوت در گاو با انسان و موش مقایسه شدند. در مرحله بعد، برای به‌دست آوردن نواحی دارای Hypo/Hyper-Methylation (HMRs) در ژن‌های متفاوت بیان‌شده، همپوشانی جزایر CpG با داده‌های استخراج شده از پژوهش سو و همکاران با استفاده از ابزارهای نرم‌افزار BEDTools (https://bedtools .readthedocs.io/en/latest/) برآورد شد (۳۲). سپس برای بررسی رابطه تراکم جزایر CpG و نرخ نوترکیبی، پس از استخراج مجموعه داده مربوط به نرخ نوترکیبی گاو (window size, 1Mb) از مطالعه ونگ و همکاران همبستگی بین تراکم جزایر CpG و نرخ نوترکیبی در سطح کروموزوم پیش‌بینی شد (۳۶).

استخراج جزایر CpG بر اساس پویش کل ژنوم

برای تشخیص جزایر CpG در قسمت‌های مختلف ژنوم که شامل ژن‌ها، نواحی بین ژنی، نواحی داخل ژنی و نواحی شروع رونویسی(TSS) بود، ابتدا اطلاعات مربوط به این نواحی ژنومی از پایگاه‌های (https://www.ncbi.nlm.nih.gov/) NCBI و UCSC (https://genome.ucsc.edu/) پیاده‌سازی شدند. سپس با استفاده از ابزارهای نرم‌افزار BEDTools همپوشانی جزایر CpG با نواحی مختلف ژنومی برآورد شدند. سپس جهت پیش‌بینی جزایر CpG متیله‌شده در کل ژنوم، ابتدا توسط الگوریتم HMM Hidden Markov (model) جزایر CpG در ژنوم گاو و برای هر کروموزوم برآورد شدند و سپس هم‌پوشانی جزایر CpG با HMRs توسط وب سرور Galaxy محاسبه شد. جزایری که با HMRs دارای همپوشانی بودند به‌عنوان جزایر CpG متیله‌شده در نظر گرفته شدند.

جزایر مرتبط با هر ژن به‌عنوان جزایر CpG مرتبط با ژن‌ها، جزایری که کل توالی آن‌ها در نواحی بین ژنی قرار داشت به‌عنوان جزایر CpG بین ژنی، جزایری که کل توالی آن‌ها در نواحی داخل ژنی قرار داشت به‌عنوان جزایر CpG داخل ژنی و جزایری که با TSS هم‌پوشانی داشتند را به‌عنوان جزایر مرتبط با TSS طبقه‌بندی شدند. سپس جهت پیش‌بینی جزایر CpG متیله‌شده در کل ژنوم، ابتدا توسط الگوریتم HMM Hidden Markov Model جزایر CpG در ژنوم گاو و برای هر کروموزوم برآورد شدند و سپس هم‌پوشانی جزایر CpG با HMRs توسط وب سرور Galaxy محاسبه شد. جزایری که با HMRs دارای هم‌پوشانی بودند به‌عنوان جزایر CpG متیله‌شده در نظر گرفته شدند. همچنین برای مقایسه جزایر CpG در ژنوم گاو و سایر نشخوارکنندگان اهلی، ژنوم سه گونه گوسفند و بز و شتر استخراج گردید و با استفاده از الگوریتم HMM جزایر CpG آنها برآورد گردید و با ژنوم گاو مقایسه شد.

نتایج و بحث

مقایسه الگوریتم‌های مختلف پیش‌بینی جزایر CpG

جزایر CpG برای ژن‌های با بیان متفاوت با استفاده از الگوریتم‌های گاردینر گاردن و فرامر GF، الگوریتم تاکایی و جونز TJ، الگوریتم CpGcluster، الگوریتم مدل مخفی مارکف HMM و الگوریتم مدل مخفی مارکف تعمیم یافته GHMM پیش‌بینی گردید. در جدول ۲ اطلاعات ژنومی و اطلاعات آماری جزایر CpG در ژن‌های با بیان متفاوت و با استفاده از الگوریتم‌های مختلف آورده شده است.

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

جدول ۲- مقایسه الگوریتم‌های مختلف پیش‌بینی جزایر CpG در ژن‌های با بیان متفاوت

Table 2. Comparison of different algorithms for CpG islands prediction in differentially expressed genes

ژن‌های با بیان متفاوت**	تاکایی و جونز	گاردین گاردن و فرامر	خوشه بندی جزایر CpG	مدل مخفی مارکف	مدل مخفی مارکف تعمیم یافته	عملکرد
LTF	۱۴	۶۷۱	۱۱۱۴	۹۲۱	۵۱۲	سیستم ایمنی ذاتی
APP	۲۵	۳۵۲۲	۵۲۲۳	۴۰۳۲	۳۱۴۷	گیرنده سطحی سلول
CD40	۶	۶۲۱	۱۰۲۱	۸۷۱	۴۳۶	سیستم ایمنی و پاسخ به التهاب
BCL2A1	۰	۶	۱۱	۹	۴	آپوپتوز و هموستازی
CCL2	۰	۰	۰	۰	۰	سیستم ایمنی و پاسخ به التهاب
CCL5	۵	۱۰۸	۲۱۵	۱۹۷	۸۷	سیستم ایمنی و پاسخ به التهاب
CCL20	۰	۰	۰	۰	۰	سیستم ایمنی و پاسخ به التهاب
CFB	۳	۲	۵	۳	۲	نقش در تکثیر لنفوسیت‌های B
CSNK1D	۱۰	۲۲۹۷	۳۲۴۱	۲۸۷۴	۱۹۶۴	نقش در تکثیر و ترمیم DNA
CX3CL1	۲	۰	۴	۲	۱	گیرنده کموکاین‌ها و سیستم ایمنی
CXCL2	۱	۳۲۱	۵۴۹	۴۱۴	۲۱۸	گیرنده کموکاین‌ها و سیستم ایمنی
CXCL3	۱	۲۶۳	۳۱۷	۳۰۱	۲۱۱	گیرنده کموکاین‌ها و سیستم ایمنی
CXCL5	۱	۴۱۹	۶۵۸	۵۳۶	۳۶۷	گیرنده کموکاین‌ها و سیتم ایمنی
CXCL8	۰	۰	۰	۰	۰	گیرنده کموکاین‌ها و سیستم ایمنی
DAPP1	۰	۱۵۵	۲۴۷	۱۹۶	۱۱۰	گیرنده فسفولیپیدها و پاسخ به التهاب
GRO1	۱	۲۴۸	۴۱۳	۳۱۶	۲۰۳	گیرنده کموکاین‌ها و سیتم ایمنی
IL1A	۰	۱۶	۲۴	۱۹	۱۴	گیرنده کموکاین‌ها و پاسخ به التهاب
IL1B	۲	۵۲	۵۴	۵۱	۳۸	گیرنده کموکاین‌ها و پاسخ به التهاب
IL1RN	۲	۰	۴	۳	۰	گیرنده کموکاین‌ها و پاسخ به التهاب
NFKBIA	۲	۸۴۷	۱۲۳۴	۸۶۵	۵۴۲	پاسخ به التهاب
NFKBIZ	۱	۶۱۹	۹۴۷	۶۷۳	۴۷۰	گیرنده لیپوپلی‌ساکاریدها
S100A8	۰	۰	۰	۰	۰	تکثیر و تمایز سلولی
S100A9	۱	۱۰	۲۱	۱۸	۸	تکثیر و تمایز سلولی
ISG15	۲	۰	۵	۴	۰	پاسخ به التهاب و فعالیت نوتروفیل‌ها
MAP3K8	۷	۷۷	۱۲۴	۱۰۲	۵۴	فعال‌کننده MAP کیناز و JNK کیناز
MARCKSL1	۱	۴۶۵	۷۵۴	۵۹۶	۳۷۲	تنظیم سایتواسکلتال
MX1	۷	۸۰۸	۱۴۵۴	۹۶۰	۶۳۱	پاسخ سلولی ضد ویروس
RND1	۱	۲۹	۴۹	۳۲	۲۳	پاسخ به فاکتورهای رشد بیرون سلولی
RSAD2	۴	۲۷۱	۴۷۰	۳۹۹	۲۱۷	سیستم ایمنی ذاتی
SAA3	۰	۰	۰	۰	۰	متابولیسم فولات
TLR2	۲	۱۱۹۷	۱۷۱۴	۱۲۳۱	۹۴۰	نقش در سیستم ایمنی
TNF	۱	۰	۳	۲	۰	اپوپتوز و تکثیر و تمایز سلولی
ZC3H12A	۸	۱۰۷۹	۱۷۸۴	۱۵۱۲	۹۲۰	آپوپتوز

**DGE: ژن‌های با بیان متفاوت در سطح ۱ درصد آماری ($p< 01$)

همان‌طور که در جدول ۲ آمده است، تعداد جزایر CpG و همچنین تراکم جزایر CpG پیش‌بینی شده توسط الگوریتم‌های مختلف بسیار متفاوت است. الگوریتم CpGcluster دارای بیشترین تعداد جزایر CpG و الگوریتم TJ دارای کمترین تعداد جزایر CpG در بین الگوریتم‌ها هستند. تعداد خوشه‌ها یا جزایر CpG در الگوریتم CpGcluster وابسته به ارزش p است و با تغییر آن تعداد جزایر پیش‌بینی شده تغییر خواهد کرد. نویسندگان این الگوریتم مقدار $p=10^{-5}$ را در نظر و بیان نموده‌اند که هرچند این الگوریتم تعداد بسیار زیادی جزایر CpG با طول کم را پیش‌بینی می‌کند، اما این جزایر کوچک عملکردی هستند (۱۱،۱۰). از طرفی الگوریتم CpGcluster به‌طور متوسط دارای محتوی GC و نسبت ObsCpG/ExpCpG بالاتر نسبت به سایر الگوریتم است که بیانگر آن است، الگوریتم CpGcluster گرایش به انتخاب جزایر کوچک‌تر، اما غنی از دی‌نوکلئوتیدهای CpG دارد. الگوریتم TJ جهت در نظر نگرفتن توالی‌های عناصر پراکنده کوتاه[1] (SINEs) و به‌خصوص توالی‌های تکراری[2] Alu در پیش‌بینی جزایر CpG در ژنوم انسان و موش طراحی گردیده است (۳۴،۳۳). بنابراین این می‌تواند از دلایل پیش‌بینی کمتر تعداد جزایر CpG توسط این الگوریتم باشد. همچنین اکثر الگوریتم‌ها میانگین طول جزایر CpG را برای این ژنوم‌ها زیر ۵۰۰ جفت نوکلئوتید برآورد کرده در حالی‌که کمینه طول برای الگوریتم TJ ۵۰۰ جفت نوکلئوتید است و این نیز می‌تواند دلیل دیگری برای پیش‌بینی کمتر تعداد جزایر CpG در این الگوریتم باشد. در پژوهش‌های دیگر محققین نیز چنین تفاوت‌های بزرگی در بین الگوریتم‌ها گزارش نموده‌اند (۱۵،۱۳،۱۱). یکی از دلایل اصلی بررسی و استخراج جزایر CpG نقش آن‌ها به‌عنوان نشانگر ژن‌ها می‌باشد (۲۲،۱۸،۳) و معمولاً در این زمینه الگوریتم‌هایی که تعداد جزایر CpG را نزدیک به تعداد ژن‌ها برآورد کنند بهتر می‌باشند. در پژوهش هان و ژائو الگوریتم TJ نسبت به الگوریتم CpGcluster تخمین دقیق‌تری از جزایر CpG داشت (۱۵). یکی از اصلی‌ترین تفاوت‌های بین این الگوریتم‌ها این بود که دو الگوریتم GF و TJ دارای محدودیتی به‌نام کمینه طول جزایر CpG بودند که به‌ترتیب برای این دو ۲۰۰ و ۵۰۰ جفت باز بود و سایر الگوریتم‌ها این محدودیت را نداشتند. کوچکترین جزایر CpG پیش‌بینی شده برای الگوریتم CpGcluster دارای طول ۸ جفت باز، برای

1- Short Interspersed Elements are short DNA sequences

2- The most common SINEs in primates

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

الگوریتم HMM دارای طول ۱۸ جفت باز و برای الگوریتم GHMM دارای طول ۲۲ جفت باز است. در الگوریتم GF حدود ۹۲ درصد جزایر CpG در دامنه ۲۰۰ تا ۵۰۰ جفت باز قرار داشتند و فقط ۸ درصد بیش از ۵۰۰ جفت باز بودند. دو الگوریتم HMM و GHMM تقریبا دارای توزیع طول مشابه بودند (۶۰ درصد زیر ۵۰۰ جفت باز و ۴۰ درصد بالای ۵۰۰ جفت باز). حدود ۸۵ درصد از جزایر CpG پیش‌بینی شده توسط الگوریتم TJ در دامنه ۵۰۰ تا ۱۰۰۰ جفت باز بودند. نتایج مشابه در سایر پژوهش‌ها مشاهده شد (۳۷،۱۵،۱۳). طول DNA یک فاکتور مهم در تشخیص جزایر CpG است و معمولا جزایر CpG با طول بیشتر از ۵۰۰ جفت باز بیشتر با راه‌انداز ژن‌ها مرتبط بوده و جزایر CpG دارای طول‌های کمتر در نواحی بین ژنی قرار دارند (۱۹،۱۵،۱۱). در این پژوهش الگوریتم TJ، HMM و GHMM اکثرا دارای طول بیش از ۵۰۰ جفت باز بودند و بنابراین ممکن است جهت تشخیص جزایر CpG مرتبط با راه‌انداز ژن‌ها مفید باشند. مهم‌ترین تفاوت بین الگوریتم CpGcluster و سایر الگوریتم‌ها طول پیش‌بینی شده توسط این الگوریتم است که منجر به تولید جزایر کوچک با طول کم می‌شود. ایجاد کنندگان این الگوریتم بیان کرده‌اند که این جزایر کوچک دارای وضعیت متیلاسیونی متفاوت بوده و عملکردی هستند (۱،۱۰). بنابراین، می‌توان گفت که این الگوریتم جهت کاوش و بررسی جزایر کوچک و عملکرد آن‌ها مناسب است. اکثر طراحان الگوریتم‌ها از ویژگی‌های زیستی جهت تایید و ارزیابی الگوریتم‌های پیش‌بینی جزایر CpG استفاده می‌کنند اما هیچکدام از الگوریتم‌های پیش‌بینی نمی‌توانند تمام ویژگی‌های زیستی مربوطه را با دقت و صحت مناسب پیش‌بینی کنند (۲۴).

مقایسه جزایر CpG ژن‌های با بیان متفاوت در گاو با موش و انسان

در این قسمت جزایر CpG در ژن‌های با بیان متفاوت در گاو با انسانو موش مقایسه شدند. به‌علت اینکه اطلاعات جزایر CpG موجود در این دو پستاندار براساس دو الگوریتم GF و TJ بوده است (۳۷،۱۲)، ما نیز نتایج حاصل از این دو الگوریتم را جهت مقایسه استفاده نمودیم. همان‌طور که در جدول ۳ مشخص است ویژگی‌های جزایر CpG در بین ژن‌های با بیان متفاوت و بر اساس هر یک از الگوریتم‌ها به‌طور قابل توجه متفاوت است.

جدول ۳- مقایسه تعداد جزایر CpG ژن‌های با بیان متفاوت در گاو با موش و انسان

Table 3. The comparison number of CpG islands of differentially expressed genes in cows with mice and humans

DGE**	گاو		انسان		موش	
	TJ	GF	TJ	GF	TJ	GF
LTF	۱۴	۶۷۱	۶	۳۷۳	۵	۱۶۲
APP	۲۵	۳۵۲۲	–	–	–	–
CD40	۶	۶۲۱	۲	۴۴۲	۴	۲۲۸
BCL2A1	۰	۶	۱	۱۲۹	۰	۰
CCL2	۰	۰	۰	۰	۰	۰
CCL5	۵	۱۰۸	۵	۴۶۰	۱	۹۰
CCL20	۰	۰	۰	۰	۰	۰
CFB	۳	۲	۱	۱۱۲	۱	۴۸
CSNK1D	۱۰	۲۲۹۷	۲۲	۵۳۳۵	۴	۱۱۷۰
CX3CL1	۲	۰	۶	۳۲۱	۵	۳۱۸
CXCL2	۱	۳۲۱	۱	۸۲۹	۱	۰
CXCL3	۱	۲۶۳	۱	۷۴۱	۱	۰
CXCL5	۱	۴۱۹	۱	۳۱۴	۱	۲۶۰
CXCL8	۰	۰	۰	۰	–	–
DAPP1	۰	۱۵۵	۴	۳۰۳	۴	۶۰۴
GRO1	۱	۲۴۸	۱	۶۳۲	۱	۱۰۷
IL1A	۰	۱۶	۲	۵۵۴	۰	۰
IL1B	۲	۵۲	۰	۰	۰	۰
IL1RN	۲	۰	۲	۵۰	۰	۱۰۹
NFKBIA	۲	۸۴۷	۱	۱۲۲۵	۲	۹۸۹
NFKBIZ	۱	۶۱۹	۵	۱۷۰۶	۷	۱۹۰۱
S100A8	۰	۰	۱	۰	۰	۰
S100A9	۱	۱۰	۱	۰	۰	۰
ISG15	۲	۰	۱۲	۲۵۷۱	۱	۰
MAP3K8	۷	۷۷	۳	۱۵۶۸	۱	۵۶۵
MARCKSL1	۱	۴۶۵	۱	۱۴۳۱	۱	۱۱۰۸
MX1	۷	۸۰۸	۱۱	۱۸۷۳	۰	۳۲
RND1	۱	۲۹	۳	۳۲۵	۷	۷۰۹
RSAD2	۴	۲۷۱	۱	۱۰۹۷	۱	۷
SAA3	۰	۰	۰	۰	۱	۳
TLR2	۲	۱۱۹۷	۰	۰	۲	۳۷۱
TNF	۱	۰	۲	۵	۲	۱۷
ZC3H12A	۸	۱۰۷۹	۷	۱۰۶۰	۵	۹۶۹

**DGE: ژن‌های با بیان متفاوت در سطح ۱ درصد آماری (p< 01)

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

مقایسه جزایر CpG در ژنوم گاو و سایر نشخوارکنندگان اهلی

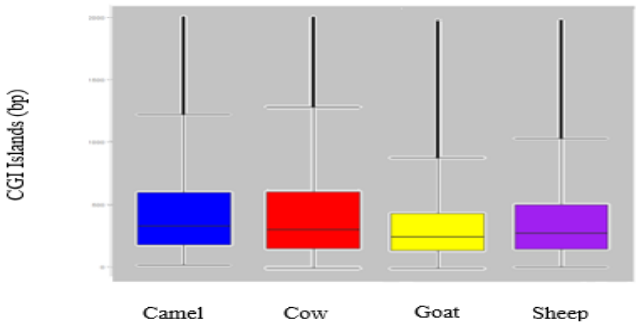
در این قسمت، ژنوم سه گونه گوسفند و بز و شتر استخراج شد و با استفاده از الگوریتم HMM جزایر CpG آنها برآورد و با ژنوم گاو مقایسه شد (۳۹). این پژوهش- بر اساس بررسی‌های ما- شاید اولین بررسی سیتماتیک جزایر CpG در بین نشخوارکنندگان اهلی باشد. در جدول ۴ اطلاعات آماری مقایسه‌ای ژنوم و جزایر CpG گاو با سه گونه گوسفند و بز و شتر آمده است.

جدول ۴- اطلاعات آماری ژنوم و جزایر CpG در ۴ گونه نشخوارکننده اهلی

Table 4. Statistical data on genome and CpG islands in 4 domestic ruminant species

گونه	ژنوم				جزایر CpG				
	سایز (Gb)	تعداد جفت کروموزوم	محتوی GC (٪)	نسبت Obs/ExpCpG	تعداد CGIs	تراکم CpGs(/Mb)	میانگین طول (bp)	محتوی GC	نسبت Obs/ExpCpG
گاو	۲/۶۳	۳۰	۴۱/۷	۰/۲۵۱	۹۰۶۶۸	۳۴/۴۷	۴۸۵	۶۳/۳	۰/۷۴۹
گوسفند	۲/۵۹	۲۷	۴۱/۸	۰/۲۵۳	۹۵۵۴۲	۳۶/۸۹	۴۲۰	۶۲/۲	۰/۷۲۹
بز	۲/۵۲	۳۰	۴۱/۸	۰/۲۵۱	۹۹۰۷۰	۳۹/۳۱	۳۶۱	۶۱/۲	۰/۷۳۰
شتر	۲/۰۵	۳۷	۴۱/۴۰	۰/۲۶۱	۸۸۹۴۹	۴۳/۳۹	۴۸۹	۶۰/۶	۰/۶۹۴

همانطور که در جدول ۴ مشخص است تعداد و تراکم جزایر CpG در بین ژنوم‌ها تغییر می‌کند. ژنوم بز دارای بیشترین تعداد جزایر CpG و ژنوم شتر دارای بیشترین تراکم جزایر CpG هستند. باکس پلات توزیع طول جزایر CpG در این چهار گونه بیانگر تفاوت کم میانگین و واریانس طول بین گونه‌ها می‌باشد. در این نمودار ژنوم بز دارای کمترین میانگین طول و واریانس است (شکل ۱).



شکل ۱- توزیع طول جزایر CpG ۴ گونه نشخوارکننده اهلی با استفاده از الگوریتم HMM

Figure 1. Distribution of CpG Island Lengths in 4 ruminant Species Using the HMM Algorithm

در پژوهش‌هایی که جهت توالی‌یابی ژنوم این چهار گونه انجام شده است، شباهت‌های بسیاری اعم از انشقاق اجدادی نزدیک، نقشه‌های مشابه و حتی ژن‌های ارتولوگ دارند (۱،۶،۸،۳۸). بنابراین، این دلایل می‌تواند علت نزدیکی جزایر CpG در این ژنوم‌ها باشد. در بین این ژنوم‌ها شتر بیشترین نزدیکی را با ژنوم گاو داشت. ژنوم گاو شباهت (Synteny) بالایی را با ژنوم شتر (نرخ پوشش بیش از ۸۵ درصد) به اشتراک گذاشته‌اند و همچنین انشقاق اجدادی بین گونه گاو و شتر حدود ۴۷/۲ میلیون سال قبل است که کمتر از سایر گونه‌ها است. این عوامل ممکن است علت نزدیکی بیشتر جزایر CpG در گاو و شتر باشد (۳۸). به‌دلیل کیفیت بالای حاشیه‌نویسی ژن‌ها در گاو نسبت به بقیه ژنوم‌ها داده‌های ژنی مربوط به گاو جهت جستجوی جزایر CpG در قسمت‌های مختلف ژنوم (ژن، نواحی بین ژنی، نواحی داخل ژنی و سایت‌های شروع رونویسی) استفاده شد. در جدول ۵ روابط بین تراکم جزایر CpG در قسمت‌های مختلف ژنوم با ویژگی‌های ژنومی در سطح کروموزوم (لگاریتم در مبنای ۱۰ سایز کروموزوم، محتوی GC ژنومی و نسبت ObsCpG/ExpCpG بررسی شده‌است. بر اساس انوتیشن ژن در پایگاه‌های داده NCBI و UCSC به‌ترتیب تعداد ۶۱۰۹۵، ۲۹۶۱۳، ۴۸۷۵۹ و ۱۲۳۱۵ جزیره CpG مرتبط با ژن‌ها، نواحی بین ژنی، نواحی داخل ژنی و جایگاه‌های شروع رونویسی هم‌پوشانی داشتند.

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

جدول ۵- همبستگی بین تراکم جزایر CpG در قسمت‌های مختلف ژنوم با ویژگی‌های ژنومی در سطح کروموزوم

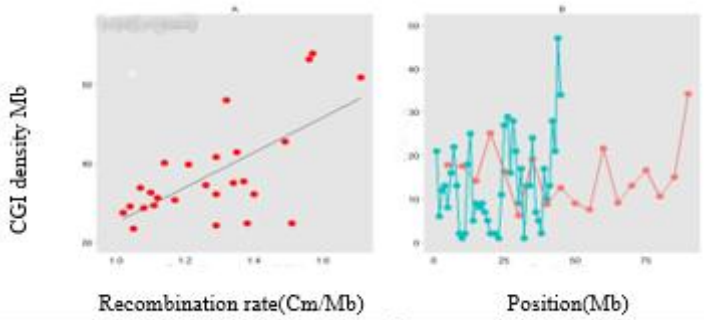
Table 5. Correlation between the density of CpG islands in different genome sections with genomic characteristics at the chromosome level

TSS CGIs (12315)		Intragenic CpGs(48795)		Intergenic CpGs(29613)		Gene-associated CGIs(61095)		
p	r	p	r	p	r	p	r	
۰/۰۰۳	−۰/۵۳	۰/۰۱	−۰/۴۸	۰/۰۳	−۰/۳۹	$۲/۹\times۱۰^{-۵}$	−۰/۶۱	$\log_{10}$ (chromosome size)
$۳/۲\times۱۰^{-۱۰}$	۰/۸۷	$۲/۳\times۱۰^{-۶}$	۰/۸۵	۰/۰۱	۰/۴۷	$۴/۱\times۱۰^{-۱۲}$	۰/۹۱	GC content
$۸/۶\times۱۰^{-۱۰}$	۰/۸۶	$۱/۸\times۱۰^{-۸}$	۰/۸۳	$۸/۵\times۱۰^{-۵}$	۰/۶۵	$۳/۱\times۱۰^{-۱۲}$	۰/۹۴	Obs_{CpG} /Exp_{CpG}

با بررسی تعداد جزایر CpG در نواحی داخل ژنی و بین‌ژنی مشخص می‌شود که تعداد جزایر CpG در نواحی بین ژنی بسیار کمتر از نواحی ژنی و داخل ژنی است و همچنین همبستگی کمتری با ویژگی‌های ژنومی دارند. این مشاهدات و همچنین تعداد بالای جزایر CpG در TSS بیانگر آن است که جزایر CpG از ویژگی‌های قابل توجه ژن در نشخوارکنندگان هستند. در بررسی سایر پژوهشگران روی قسمت‌های مختلف ژنوم سایر گونه‌ها نیز نتایج مشابهی گزارش شده است (۲۵،۱۴،۱۲).

رابطه تراکم CpGs با نرخ نوترکیبی

همبستگی بین تراکم جزایر CpG و نرخ نوترکیبی در سطح کروموزوم پیش‌بینی شد. همبستگی مثبت و معنی‌دار (r=0.61, p=0.00031) بین این دو ویژگی مشاهده شد (شکل ۲ (A)). در پژوهش‌های پیشین رابطه بین نوترکیبی و ویژگی‌های مرتبط با GC (محتوی GC، دی‌نوکلئوتیدهای CpG) مثبت و معنی‌دار گزارش شده است (۳۵،۳۰،۲۸). در این پژوهش نیز همبستگی‌های بین تراکم جزایر CpG و ویژگی‌های مرتبط با GC مثبت و معنی‌دار بود که این می‌تواند از دلایل همبستگی بین تراکم جزایر CpG و نوترکیبی باشد. در پژوهش هان و همکاران (۱۲) نیز همبستگی بین تراکم جزایر CpG و نرخ نوترکیبی مثبت و معنی‌دار گزارش شده که نتایج تحقیق حاضر با آن مطابقت دارد. به‌دلیل اینکه نرخ نوترکیبی وابسته به فاصله از سانترومر است (۳۶،۲۶) و معمولا از سانترومر به سمت تلومر نرخ نوترکیبی افزایش می‌یابد (۳۶،۲۷،۲۰،۱۲). ما روند تراکم جزایر CpG را در طول کروموزم‌ها بررسی کردیم و مشخص شد که بیشترین تراکم جزایر CpG در نواحی تلومری کروموزوم‌ها است (شکل ۲ (B)).



شکل ۲- (A) همبستگی بین تراکم CpGs و نرخ نوترکیبی (cM/Mb). (B) توزیع تراکم CpGs در طول کروموزوم های گاو، خط قرمز (کروموزوم ۱۰) و خط آبی (کروموزوم ۲۸)، در سایر کروموزوم ها نیز همین روند مشاهده شد.

Figure 2. (A) Correlation between CpGsdensity and recombination rate (cM / Mb). (B) Distribution of CpGs density over cattle chromosomes, red line (chromosome 10) and blue line (chromosome 28), also observed in other chromosomes

این روند نیز می‌تواند یکی از دلایل همبستگی مثبت بین تراکم جزایر CpG و نرخ نوترکیبی باشد. البته باید توجه داشت که کشف علیت رابطه مشاهده شده بین تراکم جزایر CpG و نرخ نوترکیبی مشکل است مثلا اینکه کدام یک از آنها علت دیگری است، و بنابراین نیاز به بررسی و تجزیه و تحلیل مکانسیم‌های مؤثر بر نوترکیبی جهت تشخیص سازوکار مولکولی این رابطه است.

بررسی وضعیت متیلاسیون ژن‌های با بیان متفاوت

در این مطالعه پس از مشخص کردن جزایر CpG برای هر یک از ژن‌های متفاوت بیان شده، به بررسی وضعیت متیلاسیون جزایر CpG در ژن‌های با بیان متفاوت پرداخته شد. جدول ۶ نشان‌دهنده وضعیت متیلاسیون ژن‌های با بیان متفاوت می‌باشد.

نتایج بررسی متیلاسیون در ژن‌های با بیان متفاوت نشان داد که ژن‌های LTF, APP, CCL5, CD40, CSNK1D, CX3CL1, DAPP1, NFKBIZ, S100A9, ISG15, MAP3K8, MX1, RDAD2, ZC3H12A دارای جزایر CpG متیله شده می‌باشند. متیلاسیون سیتوزین در نواحی پالیندرومیک CpG در این ژن‌ها نقش مهمی را در خاموشی ژن‌های ذکر شده در توسعه و پیشرفت بیماری ورم پستان بازی می‌کنند.

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

بررسی وضعیت متیلاسیون جزایر CpG در کل ژنوم

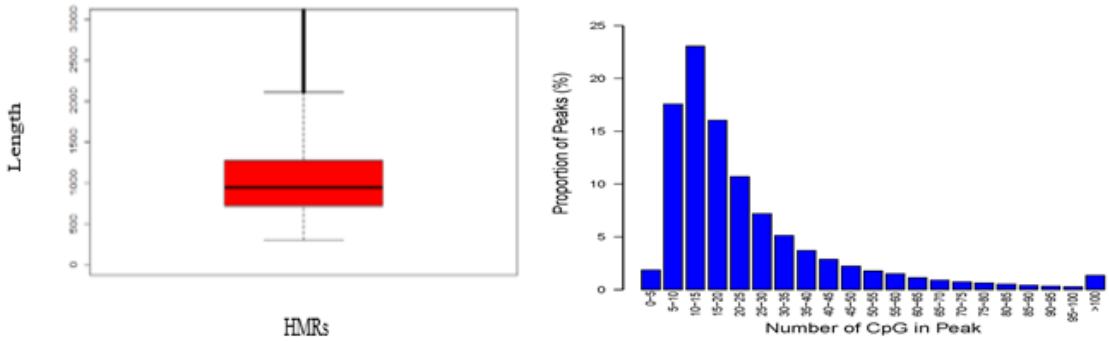
همانطور که شکل ۳ نشان می‌دهد، اکثر HMRs دارای تعداد ۵ تا ۲۵ دی نوکلئوتید CpG هستند. در این پژوهش، تعداد ۹۰۶۶۸ جزایر CpG در ژنوم گاو پیش بینی شد، که تعداد ۹۹۴۲ جزیره با HMRs دارای هم پوشانی بودند و به عنوان جزایر CpG متیله شده در نظر گرفته شدند. به لحاظ نسبی ۱۱ درصد از جزایر CpG متیله بودند که بیانگر آن است اکثر جزایر CpG در بافت جفت غیرمتیله هستند. در پژوهش سو و همکاران با استفاده از روش MeDIP-seq تعداد ۱۴۵۲۱۸ ناحیه دارای متیلاسیون بالا HMRs در ژنوم گاو تشخیص داده و استخراج شد (۳۲). توزیع طول HMRs در شکل ۳ نشان داده شده است. میانگین طول HMRs برابر با ۱۰۸۰/۱۶ است. حدود ۵/۸۹ درصد از ژنوم توسط HMRs پوشش داده می‌شود.

جدول ۶- بررسی وضعیت متیلاسیون جزایر CpG در ژن‌های با بیان متفاوت

Table 6. Examination of the methylation status of CpG islands in differential expressed genes

HMRs	GF	TJ	DGE**
۴	۶۷۱	۱۴	LTF
۲	۳۵۲۲	۲۵	APP
۱	۶۲۱	۶	CD40
۰	۶	۰	BCL2A1
۰	۰	۰	CCL2
۱	۱۰۸	۵	CCL5
۰	۰	۰	CCL20
۰	۲	۳	CFB
۵	۲۲۹۷	۱۰	CSNK1D
۱	۰	۲	CX3CL1
۰	۳۲۱	۱	CXCL2
۰	۲۶۳	۱	CXCL3
۰	۴۱۹	۱	CXCL5
۰	۰	۰	CXCL8
۴	۱۵۵	۰	DAPP1
۰	۲۴۸	۱	GRO1
۰	۱۶	۰	IL1A
۰	۵۲	۲	IL1B
۰	۰	۲	IL1RN
۰	۸۴۷	۲	NFKBIA
۱	۶۱۹	۱	NFKBIZ
۰	۰	۰	S100A8
۱	۱۰	۱	S100A9
۱	۰	۲	ISG15
۴	۷۷	۷	MAP3K8
۰	۴۶۵	۱	MARCKSL1
۱	۸۰۸	۷	MX1
۰	۲۹	۱	RND1
۲	۲۷۱	۴	RSAD2
۰	۰	۰	SAA3
۰	۱۱۹۷	۲	TLR2
۰	۰	۱	TNF
۱	۱۰۷۹	۸	ZC3H12A

**DGE: ژن‌های با بیان متفاوت در سطح ۱ درصد آماری (p< 01)



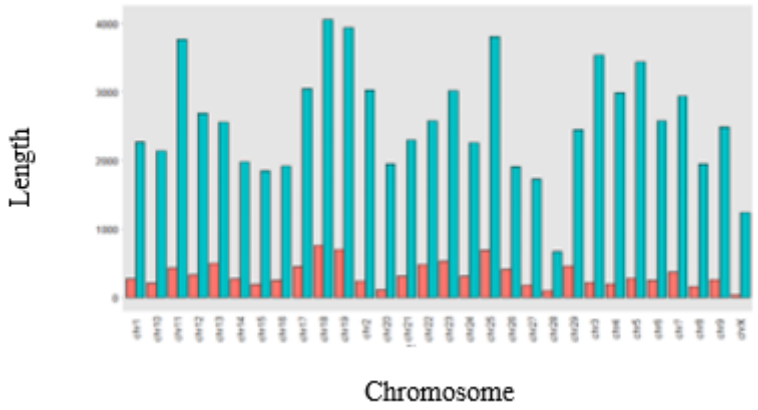
شکل ۳- سمت چپ: توزیع دی نوکلئوتیدهای CpG در HMRs، سمت راست: اعداد دی نوکلئوتیدهای CpG در HMRs

Figure 3. Expansion of CpG DiNucleotides in HMRs, right: Di Nucleotides CpG in HMRs (Peak)

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

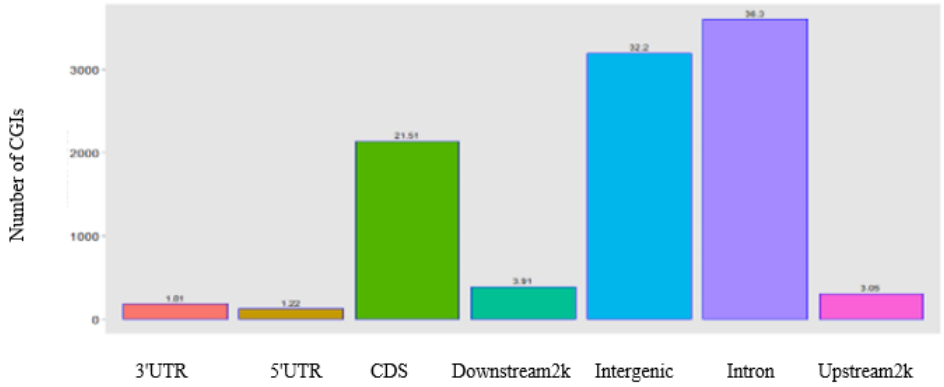
[DOI: 10.52547/rap.11.29.95]

در مطالعه سو و همکاران ۲۰/۱۲ درصد از جزایر متیله بودند (۳۲). که تفاوت با مطالعه حاضر می‌تواند به‌علت الگوریتم‌های مختلف استفاده شده جهت پیش‌بینی جزایر CpG باشد. تعداد جزایر متیله و غیرمتیله برای هر کروموزوم محاسبه گردید و نمودار مربوطه رسم شد (شکل ۴). نمودار بیانگر این مطلب است که در هر کروموزوم نیز اکثر جزایر CpG غیر متیله هستند.



شکل ۴- نمودار وضعیت متیلاسیون جزایر CpG در کروموزم‌ها. نمودار قرمز بیانگر CpGs متیله و نمودار آبی بیانگر CpGs غیرمتیله
Figure 4. Status of methylation of CpG islands in chromosomes. The red chart represents the Methyl CpGs and the blue graph represents non-Methyl CpGs

در مرحله بعد وضعیت جزایر CpG متیله شده در قسمت‌های مختلف ژنوم محاسبه شد (شکل ۵). همانگونه که در شکل مشخص است بیشترین تعداد جزایر CpG متیله‌شده در ناحیه اینترون (۳۶/۶ درصد) قرار دارند و سپس نواحی بین ژنی (۳۲/۲ درصد) و پس از آن نواحی کدکننده ژنی (۲۱/۵۱ درصد) بودند. نواحی راه‌انداز دارای کمترین تعداد جزایر CpG متیله هستند.



شکل ۵- توزیع وضعیت CpGs متیله شده در قسمت‌های مختلف ژنوم. درصد CpGs در هر ناحیه بر روی ستون‌ها درج شده است.
Figure 5. Distribution of methylation of CpGs in different parts of the genome. The percentage of CpGs per area is plotted on the columns

در بررسی سو و همکاران نتایج مشابهی گرفته شده است اما در تعداد و درصد تفاوت‌هایی وجود دارد که احتمالا می‌تواند
به‌علت الگوریتم‌های مختلف استفاده شده جهت پیش‌بینی جزایر CpG باشد (۳۲). در پژوهش هوو و همکاران بر روی ژنوم جوجه‌های گوشتی مشخص‌شد که کمترین تراکم متیلاسیون در جزایر CpG واقع در پروموتور ژن‌ها است و در بدنه ژن‌ها تراکم متیلاسیون جزایر واقع در اینترون بیش‌تر از اگزون و UTRs می‌باشد و جزایر CpG متیله شده در نواحی بین ژنی توزیع شده نیز قابل توجه هستند که نتایج تحقیق حاضر با آن مطابقت دارد (۱۶). همچنین کوواک و همکاران در پژوهشی پروفایل متیلاسیون را در ژنوم خوک بررسی نمودند که نتایج آنها مشابه نتایج به‌دست آمده در ژنوم جوجه و در مطابقت با پژوهش ما بود (۲۱). در پژوهش‌های بال و همکاران و دو و همکاران نیز پروفایل متیلاسیون جزایر CpG مشابه آنچه ذکر گردید گزارش شده‌است (۷،۲). کلدری و همکاران در پژوهش بر روی گوسفند نشان دادند که نواحی نزدیک TSS ژن‌ها دارای متیلاسیون پایین‌تری هستند (۴).

[Downloaded from rap.sanru.ac.ir on 2025-12-17]
[DOI: 10.52547/rap.11.29.95]

آنچه از بررسی پروفایل متیلاسیون CpGs در این مطالعه و سایر مطالعات بررسی‌شده حاصل می‌شود این است که پروفایل کلی متیلاسیون در اکثر پستانداران و حتی پرندگان مشابه هم می‌باشد و به‌نظر می‌رسد که بین گونه‌های مختلف به‌صورت محافظت شده است (۳۱،۱۷،۵).

مقایسات ژنومی بین گونه‌ای متیلاسیون DNA نشان داد که نیم‌رخ کلی متیلاسیون DNA در اکثر گونه‌های مورد بررسی تقریبا مشابه هم بودند و به‌نظر می‌رسد که احتمالا نیم‌رخ کلی متیلاسیون DNA بین گونه‌های مختلف به‌صورت محافظت شده باشد. حاصل این پژوهش نشان داد کنکاش متیلاسیون DNA در بیماری‌هایی که میزان وراثت پذیری پایینی دارند و بیشتر تحت تاثیر فرآیندهای فراژنتیک هستند، ضروری به‌نظر می‌رسد. در نهایت این یافته‌ها می‌تواند در چشم‌انداز تکامل ژنتیک گاو و سایر نشخوارکنندگان موثر بوده و نقش جزایر CpG در عملکردهای ژنتیکی، فراژنتیکی و مولکولی را بیان کند.

منابع

1. Archibald, A., N. Cockett, B. Dalrymple, T. Faraut and J. Kijas. 2010. The sheep genome reference sequence: a work in progress. Animal Genetics, 41: 449-453.
2. Ball, MP., J.B. Li, Y. Gao, J.H. Lee and E.M. LeProust. 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nature Biotechnology, 27: 361-368.
3. Bird, A.P. 1987. CpG islands as gene markers in the vertebrate nucleus. Trends in Genetics, 3: 342-347.
4. Couldrey, C., R. Brauning, J. Bracegirdle, P. Maclean, H.V. Henderson and J.C. McEwan. 2014. Genome-wide DNA methylation patterns and transcription analysis in sheep muscle. PloS One, 9(7): e101853.
5. Couldrey, C., R. Brauning, H. Henderson and J. McEwan. 2015. Genome- wide DNA methylation analysis: no evidence for stable hemimethylation in the sheep muscle genome. Animal Genetics, 46: 185-189.
6. Dong, Y., M. Xie, Y. Jiang, N. Xiao and X. Du. 2013. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). Nature Biotechnology, 31: 135-141.
7. Du, X., L. Han, A.Y. Guo and Z. Zhao. 2012. Features of methylation and gene expression in the promoter-associated CpG islands using human methylome data. Journal of Molecular Biology, 142: 461-469.
8. Elsik, C.G., R.L. Tellam and K.C. Worley. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science, 324: 522-528.
9. Gardiner-Garden, M and M. Frommer. 1987. CpG islands in vertebrate genomes. Journal of Molecular Biology, 196: 261-282.
10. Hackenberg, M., G. Barturen, P, Carpena, P.L. Luque-Escamilla, C. Previti and J.L. Oliver. 2010 .Prediction of CpG-island function: CpG clustering vs. sliding-window methods. BMC Genomics, 11: 327-334.
11. Hackenberg, M., C. Previti, P.L. Luque-Escamilla, P. Carpena, J. Martínez-Aroza and J.L. Oliver. 2006. CpGcluster: a distance-based algorithm for CpG-island detection. BMC Bioinformatics, 7: 446-452.
12. Han, L., B. Su, W.H. Li and Z. Zhao. 2008. CpG island density and its correlations with genomic features in mammalian genomes. Genome Biology, 9(5): R79.
13. Han, L. and Z. Zhao. 2008. Comparative analysis of CpG islands in four fish genomes. Genomics, 85: 301-309.
14. Han, L. and Z. Zhao. 2009. Contrast features of CpG islands in the promoter and other regions in the dog genome. Genomics, 94: 117-124.
15. Han, L. and Z. Zhao. 2009. CpG islands or CpG clusters: how to identify functional GC-rich regions in a genome? BMC Bioinformatics, 10: 65-73.
16. Hu, Y., H. Xu, Z. Li, X. Zheng and X. Jia. 2013. Comparison of the genome-wide DNA methylation profiles between fast-growing and slow-growing broilers. PloS One, 8(3): e56411.
17. Huang, Y.Z., J.J. Sun, L.Z. Zhang, C.J. Li and J.E. Womack. 2014. Genome-wide DNA methylation profiles and their relationships with mRNA and the microRNA transcriptome in bovine muscle tissue (Bos taurine). PloS One, 10(11): e0140467.
18. Illingworth, R.S. and A.P. Bird. 2009. CpG islands–‘a rough guide’. FEBS letters, 583: 1713-1720.

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

19. Ioshikhes, I.P. and M.Q. Zhang. 2000. Large-scale human promoter mapping using CpG islands. Nature Genetics, 26: 61-63.
20. Jensen-Seaman, M.I., T.S. Furey, B.A. Payseur, Y. Lu and K.M. Roskin. 2004. Comparative recombination rates in the rat, mouse, and human genomes. Genome Research, 14: 528-538.
21. Kwak, W., J.n. Kim, D. Kim, J.S. Hong and J.H. Jeong. 2014. Genome-wide DNA methylation profiles of small intestine and liver in fast-growing and slow-growing weaning piglets. Asian-Australasian Journal of Animal Sciences, 27: 1532-1538.
22. Lai, W.K. and M.J. Buck. 2010. ArchAlign: coordinate-free chromatin alignment reveals novel architectures. Genome Biology, 11(12): R126.
23. Luo, J., Y. Yu, H. Zhang, F. Tian and S. Chang. 2011. Down-regulation of promoter methylation level of CD4 gene after MDV infection in MD-susceptible chicken line. Proc. BMC Proceedings. BioMed Central, 5(4): S7.
24. Medvedeva, Y.A. 2011. Algorithms for CpG Islands Search: New Advantages and Old Problems. In Bioinformatics-Trends and Methodologies, 9: 448-481.
25. Medvedeva, Y.A., M.V. Fridman, N.J. Oparina, D.B. Malko and E.O. Ermakova. 2010. Intergenic, gene terminal and intragenic CpG islands in the human genome. BMC Genomics, 11: 48-55.
26. Paape, T., P. Zhou, A. Branca, R. Briskine, N. Young and P. Tiffin. 2012 .Fine-scale population recombination rates, hotspots, and correlates of recombination in the Medicago truncatula genome. Genome Biology and Evolution, 4: 726-737.
27. Poissant, J., J.T. Hogg, C.S. Davis, J.M. Miller, J.F. Maddox amd D.W. Coltman. 2010. Genetic linkage map of a wild genome: genomic structure, recombination and sexual dimorphism in bighorn sheep. BMC Genomics, 11: 524-532.
28. Rao, Y.S., Z.F. Wang, X.W. Chai, Q.H. Nie and X. Q. Zhang. 2013. Relationship between 5′ UTR length and gene expression pattern in chicken. Genetica, 141: 311-318.
29. Rice, P., I. Longden and A. Bleasby. 2000. EMBOSS: the European molecular biology open software suite. Elsevier Current Trends, 16(6): 276-277.
30. Romiguier, J., V. Ranwez, E.J. Douzery and N. Galtier. 2010. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. Genome Research, 20: 1001-1009.
31. Sati, S., V.S. Tanwar, K.A. Kumar, A. Patowary and V. Jain. 2012. High resolution methylome map of rat indicates role of intragenic DNA methylation in identification of coding region. PloS One, 7(2): e31621.
32. Su, J., Y. Wang, X. Xing, J. Liu and Y. Zhang. 2014. Genome-wide analysis of DNA methylation in bovine placentas. BMC Genomics, 15: 12-20.
33. Takai, D. and P.A. Jones. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proceedings of the National Academy of Sciences, 99: 3740-3745.
34. Takai, D. and P.A. Jones. 2003. The CpG island searcher: a new www resource. In Silico Biology, 3: 235-240.
35. Tortereau, F., B. Servin, L. Frantz, H.J. Megens and D. Milan. 2012. A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. BMC Genomics, 13: 586-594.
36. Weng, Z.Q., M. Saatchi, R.D. Schnabel, J.F. Taylor and D.J. Garrick. 2014. Recombination locations and rates in beef cattle assessed from parent-offspring pairs. Genetics Selection Evolution, 46: 34-42.
37. Wu, H., B. Caffo, H.A. Jaffee, R.A. Irizarry and A.P. Feinberg. 2010. Redefining CpG islands using hidden Markov models. Biostatistics, 11: 499-514.
38. Wu, H., X. Guang, M.B. Al-Fageeh, J. Cao and S. Pan. 2014. Camelid genomes reveal evolution and adaptation to desert environments. Nature Communications, 5: 5188-5199.
39. Xi, L., Y. Fondufe-Mittendorf, L. Xia, J. Flatow, J. Widom and J.P. Wang. 2010. Predicting nucleosome positioning using a duration Hidden Markov Model. BMC bioinformatics, 11: 346-357.

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]

Predicting CpG Islands and DNA Methlation in the Cow Genome Using DNA Microarray Meta-Analysis and Genome Wide Scanning

Zahra Biranvand[1], Seyed Zia-ud-Din Mir Hosseini[2], Mostafa Ghaderi-Zafra[3], Seyed Hossein Hosseini Moghaddam[4], Arash Fazeli[5] and Kianoosh Zarrin Kaviani[6]

1- Graduate of PhD in Genetics and Animal Breeding, University of Guilan, Rasht, Iran
2 and 4- Professor of Molecular Genetics, Department of Animal Sciences, University of Guilan, Rasht, Iran
3- Associate Professor of Genetics and Breeding, Department of Animal Sciences, Yasouj University, Yasouj, Iran
(Corresponding Author: Mosmos741@yahoo.com, mghaderi@yu.ac.ir)
5- Associate Professor of Molecular Genetics, Department of Agriculture and Plant Breeding, Ilam University, Ilam, Iran
6- Faculty member of Ilam University
Received: May 15, 2018 Accepted: April 17, 2020

Abstract

DNA methylation is a type of epigenetic changes that directly affects DNA. In mammals, DNA methylation is essential for fetal development and stem cell differentiation and this phenomenon essentially occurs within the CpG islands. In this study, two methods were used to study the DNA methylation profile of cow genome. In the first method, the DNA methylation profile of the differentially expressed genes from meta-analysis of DNA microarray data on mastitis were obtained. In order to perform the meta-analysis in the first method, the metaDE package in R environment, was used. Then five algorithms including TJ, GF, CpG cluster, HMM and GHMM were used to predict CpG islands in different genes. In the second method, DNA methylation profiling was performed using whole cow genome scanning. Also, for prediction of methylated CpG islands in whole genome, HMM algorithm was first estimated in bovine genome for each chromosome and then CpG overlap with Hypo / Hyper-Methylation was calculated by Galaxy Online database. The results of the first method showed that among 32 differentially expressed genes, 14 genes involved methylated CpG islands. These genes included LTF, APP, CCL5, CD40, CSNK1D, CX3CL1, DAPP1, NFKBIZ, S100A9, ISG15, MAP3K8, MX1, RDAD2, ZC3H12A. Results of the second method identified a total 90668 Hypo / Hyper-Methylation in the bovine genome, among which 9942 (10.96%) CpG islands overlapped with Hypo / Hyper-Methylation and were considered as methylated CpG. Genomic comparisons were also made between species for DNA methylation. The results showed that the overall DNA methylation profile was almost similar for majority of studied species and it seems that the overall profile of DNA methylation is likely to be conserved between different species. The results of this study showed that DNA methylation seems necessary in diseases with low heritability and which are more influenced by epigenetic processes.

Keywords: DNA Methylation, Cpg Islands, Epigenetic, Cow, Meta-Analysis, Algorithm

[Downloaded from rap.sanru.ac.ir on 2025-12-17]

[DOI: 10.52547/rap.11.29.95]