



دانشگاه علوم کشاورزی و منابع طبیعی گorgan

نشریه پژوهش در نشخوارکنندگان

جلد نهم، شماره سوم، ۱۳۹۷

<http://ejrr.gau.ac.ir>

بررسی انحراف کدونی و تجزیه و تحلیل بیوانفورماتیکی توالی کد کننده

ژن کالپاستاتین در پستانداران

* پیام جلالی^۱، قدرت‌اله رحیمی میانجی^۲، محسن قلی‌زاده^۳، علی پاکدین پاریزی^۴

^۱دانش‌آموخته کارشناسی ارشد، ^۲استاد و استادیار گروه اصلاح دام دانشگاه علوم کشاورزی و منابع طبیعی ساری

^۳عضو هیات علمی پژوهشکده ژنتیک و زیست فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری

تاریخ دریافت: ۹۷/۲/۳۰؛ تاریخ پذیرش: ۹۷/۷/۲

چکیده

سابقه و هدف: کالپاستاتین از آنزیم‌های تشکیل‌دهنده سیستم پروتئولیتیک کالپاین است که این مجموعه پروتئینی پروتئولیتیک شامل پروتئازهای طبیعی وابسته به کلسیم^۲ است و در شکل‌گیری، تجزیه بافت‌های عضلانی و تردی گوشت پس از کشتار موثر بوده و به عنوان یک ژن شاخص در ارتباط با بازدهی رشد و کیفیت گوشت به حساب می‌آید. هدف از این مطالعه ارزیابی ساختار ژنی و پروتئینی کالپاستاتین با استفاده از ابزارهای بیوانفورماتیکی در چند گونه از پستانداران بوده است.

مواد و روش‌ها: در این پژوهش توالی ناحیه کد کننده ژن و پروتئین کالپاستاتین در شش گونه پستاندار (انسان، موش، گاو، کژگاو، گوسفند و بز) مورد بررسی قرار گرفت. توالی‌های ژنی و پروتئینی از بانک ژن بازیابی شد و مورد تجزیه و تحلیل قرار گرفتند. آنالیز همولوژی و همترازی، فیلوژنی، تنوع نوکلئوتیدی در طول ناحیه کد کننده این ژن و همچنین تنوع در کدون پایانی با استفاده از نرم افزارهای کلاستال امگا و مگا هفت انجام گرفت. برنامه‌های ساپما و پروتپارام برای بررسی همولوژی، تنوع اسیدهای آمینه و اسید آمینه انتهایی در ساختار پروتئینی، توالی‌های بازیابی شده از پایگاه داده ان سی بی آی، مورد استفاده قرار گرفتند. با استفاده از نرم‌افزار کدون دبلو (نسخه ۱.۴.۲) شاخص‌های ترجیح کدونی برآورد شد.

یافته‌ها: مقدار شاخص سازگاری کدون در گونه کژگاو بالاترین (۰/۲۵۶) و در گونه گوسفند کمترین مقدار (۰/۲۳۶) را داشت. آنالیز ساختار پروتئینی کالپاستاتین نشان داد که در کل توالی‌های مورد پژوهش، آمینواسید لیزین با ۶۲۳ جایگاه بیشترین و آمینواسید تریپتوفان با ۵ جایگاه کمترین تکرار را در ساختمان این پروتئین داشتند. نسبت آمینواسیدهای قطبی به آمینواسیدهای غیر قطبی در این پروتئین ۲ به ۱ بود. مقدار گرایش کدونی مترادف یا کارکرد نسبی کدون‌های مترادف برای آمینواسیدهای سرین و اسپارتیک اسید به‌عنوان اسید آمینه انتهایی در گونه‌های مختلف به ترتیب برابر با (AGC=۱/۳۸) و (GAU=۱/۰۱) محاسبه گردید. گونه گوسفند دارای بیشترین اسیدیته ایزوالکتریک و گونه بز در شاخص تعداد کدون‌های موثر دارای بیشترین مقدار بوده‌اند.

* نویسنده مسئول: payamjalali46@yahoo.com

نتیجه‌گیری: آنالیز توالی آمینواسیدی در زنجیره پروتئینی کالپاستاتین نشان داد که این پروتئین دارای بخش قابل توجهی از آمینواسیدهای آبگریز می‌باشد نقش بازدارندگی پروتئین کالپاستاتین روی فعالیت آنزیم کالپاین در عضلات و همچنین از آن‌جایی که بیشتر توالی کالپاین را آمینواسیدهای آبدوست شکل داده‌اند، نشان از نقش به‌سزای این آمینواسیدهای آبگریز برای فعالیت در مقابل بخش آبدوست کالپاین دارد. آنالیز انحراف کدونی نشان داد که در بین گونه‌های مورد پژوهش، در گونه کژگاو شاخص ترجیح کدونی بیشتر از سایر گونه‌ها بوده و کارکرد کدون‌های بهینه در این گونه کارآمدتر از سایر گونه‌های مورد مطالعه نشان داده شد.

واژه‌های کلیدی: پستانداران، کالپاستاتین، آنالیز بیوانفورماتیکی، گرایش کدونی، ساختار فضایی

مقدمه

با توجه به رشد روزافزون جمعیت انسانی و نیاز به منابع پروتئین حیوانی، رشد سریع حیوانات اهلی از اهمیت ویژه‌ای برخوردار است. تنظیم فیزیولوژی مصرف خوراک، رشد و انرژی در حیوانات تحت تاثیر ژن‌های متعددی قرار دارد (۸). تردی گوشت یکی از مهم‌ترین عوامل موثر بر کیفیت گوشت است که از طریق فرآیندهای بافت پیوندی و سیستم انقباضی ایجاد می‌شود. عوامل موثر بر تردی گوشت را پروتئازهایی می‌دانند که در اسیدپتت پس از کشتار فعال می‌شوند (۳۰). عواملی از قبیل دما، اسیدپتت، سرعت کاهش دما و ژنتیک حیوان نیز روند جمود نعشی را تحت تاثیر قرار می‌دهند. همچنین نحوه طبخ گوشت نیز بر تردی گوشت موثر است (۱۶). ژن کالپاستاتین روی کروموزوم شماره ۵ گوسفند (۱۴) و انسان و کروموزوم شماره ۷ گاو، کژگاو^۳ و بز قرار دارد که در شکل‌گیری و تجزیه بافت‌های عضلانیموثر است (۵ و ۱۸). کالپاستاتین در ارتباط با بازدهیرشد و کیفیت گوشت، به‌عنوان یکی از ژن‌های کاندیدا شناخته شده است (۲۲ و ۲۰). مطالعات بر روی نژادهای ایرانی نیز کاندیدا بودن این ژن در رابطه با تردی گوشت و صفات رشد را تایید کرده است (۱ و ۲۳). کالپاستاتین در واقع یکی از آنزیم‌های تشکیل‌دهنده سیستم پروتئولیتیک کالپاین است (۷). خانواده کالپاین

یک مجموعه پروتئینی پروتئولیتیک هستند که شامل پروتئازهای طبیعی وابسته به کلسیم است (۳۲). این سیستم پروتئازی در فرآیندهای فیزیولوژیکی و پاتولوژیکی وابسته به کلسیم از قبیل توسعه و تجزیه عضلات، تردی گوشت بعد از کشتار، نقل و انتقالات و مرگ سلولی مؤثر است (۲۶). فعالیت مهارکنندگی کالپاستاتین بر μ کالپاین در مراحل اولیه تمایز عضله اتفاق می‌افتد (۲). وقوع جهش برای هر ژن در داخل یک گونه و همچنین در بین گونه‌های مختلف، یکسان نیست. این پدیده سبب ایجاد الگوهای غیر یکسان گرایش‌های کدونی در ژنوم جانداران، برای بیان ژن می‌شود که یک ژن را برای کد کردن آمینو اسید خاص در استفاده از کدون‌های مترادف به سمت استفاده از یک کدون خاص مجاب می‌کند. استفاده از کدون‌های مختلف، برای نمونه استفاده از کدون‌های ارجح که در سیر تکاملی با فراوانی تی آر ان ای مخصوص به خود در موجود زنده همراه هستند، برای بیان یک ژن، سبب دگرگونی در عملکرد و قابلیت بیان ژن می‌شود. در ضمن فشارهای ناشی از جهش و رانش ژنتیکی سبب دوام حیات کدون‌های جزئی که سهم بسیار کمی در تولید آمینو اسیدهای بیست گانه در پروتئین‌های مورد بررسی را دارند، می‌شوند. شاخص سازگاری کدون^۴ به‌عنوان معیارنسبت ریاضی از

4. Codon Adaptation index (CAI)

3. Yak

انحراف کدون‌های یک ژن به سمت کدون‌های ارجح در نظر گرفته می‌شود (۱۰). همچنین می‌توان از این شاخص برای پیش‌بینی میزان بیان ژن، ارزیابی میزان سازگاری ژن‌های ویروسی در دیگر میزبان‌ها و مقایسه ترجیح کدونی در ژن‌های مختلف و در گونه‌های متفاوت استفاده کرد. از سوی دیگر، اطلاعات به‌دست آمده از تحلیل داده‌های زیستی به وسیله علم بیوانفورماتیک، در به خط کردن توالی‌ها در بانک‌های اطلاعاتی برای یافتن شباهت‌ها و تفاوت‌های ژنی، پیش‌گویی ساختار و عملکرد محصولات ژن‌ها و یافتن ارتباط فیلوژنتیک میان ژن‌ها و توالی‌های پروتئینی کمک می‌کند (۱۱ و ۱۲). لذا، در تحقیق حاضر خصوصیات فیزیکی و شیمیایی ساختار پروتئین کالپاستاتین و نیز بررسی ساختار ناحیه کد کننده این جایگاه ژنی با استفاده از ابزارهای بیوانفورماتیک در گونه‌های مختلف پستانداران به منظور درک بهتر ماهیت تاثیر آن بر کیفیت گوشت مورد بررسی قرار گرفت.

مواد و روش‌ها

توالی‌های کدکننده ژن کالپاستاتین گاو، کژگاو، بز، گوسفند، انسان و موش از پایگاه داده ان سی بی آی بازیابی شد. همدردیفی توالی‌های ژنی با استفاده از ماژول کلاستال امگا (۱۳) انجام و کدون پایانی برای همه توالی‌ها در گونه‌های مختلف مشخص شد. با استفاده از نرم افزار مگا هفت (۷.۰.۱۸) ماتریس فاصله‌ای و میانگین اختلاف توالی‌ها مورد محاسبه قرار گرفت (۲۱). سپس آنالیز فیلوژنتیکی توالی‌های پروتئینی به روش اتصال مجاور^۹ انجام و درخت مربوطه بر مبنای شاخص شباهت رسم شد. بررسی مناطق غنی از بازهای گوانین و سیتوزین در ژنوم که به جزایر CpG معروف است به صورت آنالیز با

استفاده از برنامه CpGplot انجام شد. با استفاده از ماژولکدون یوزیچ در نرم‌افزار کدون دبلو (نسخه ۱.۴.۲) شاخص‌های انحراف کدونی برآورد شد (۲۷). از برنامه آنالیز ساپما برای آنالیز شاخص‌های مولکولی توالی‌های پروتئینی و تشخیص ساختار دوم پروتئین کالپاستاتیندر گونه‌های مورد پژوهش، استفاده شد. هم ترازوی چندگانه توالی‌های پروتئینی با نرم‌افزار مگا^{۱۸} (۷.۰.۱۸) هفت انجام گرفت و مقدار کارکرد نسبی کدون‌های مترادف برای اسیدهای آمینه انتهایی نیز محاسبه شد. از برنامه پروتپارام برای تعیین خصوصیات بیوشیمیایی و فیزیکی پروتئین از جمله: برآورد تعداد اسید آمینه‌های دارای بار الکتریکی مثبت (آرژنین+ لیزین)، تعداد اسیدهای آمینه دارای بار الکتریکی منفی (آسپاراژین+ گلوتامین)، وزن مولکولی، نقطه ایزوالکتریک، فرمول شیمیایی، شاخص ناپایداری، تعداد اتم‌های گوگرد، هیدروژن، نیتروژن، اکسیژن، کربن، تعداد کل اتم‌های موجود و شاخص گراویاستفاده شد. برنامه پروتپارام از الگوی مخفی مارکوف و شبکه‌های عصبی برای پیش‌بینی دقیق ساختار دوم پروتئین استفاده می‌کند. شاخص گرایش کدون‌های مترادف^۷ برای اسیدهای آمینه انتهایی زنجیره پروتئینی با استفاده از مدل زیر محاسبه شد (۳۴).

$$RSCU_i = \frac{X_i}{1/n \sum_{i=1}^n X_i} \quad \text{فرمول ۱}$$

n تعداد کدون‌های مترادف ($1 \leq n \leq 6$) برای اسید آمینه مورد نظر و X_i = تعداد دفعات وقوع کدون این شاخص برای تخمین میزان انحراف کدون‌های یک ژن به سمت کدون‌های برتر در ژن‌های با بیان بالا و همچنین تعیین کدون‌های برتر در مجموعه‌ای از ژن‌ها در یک گونه استفاده می‌شود. ارزش شاخص سازگاری کدون برای یک ژن، از

2. Mega7
3. Relative synonymous codon usage (RSCU)

1. Neighbour Joining (NJ)

فرمول ۵ $Fop=N$ (optimal codons)/N (synonymous codons)
 S میزان درصد CG در مکان سوم کدون می‌باشد. در مرحله بعد وفور کدون‌های بهینه^۱ نیز برای توالی‌ها با استفاده از مدل زیر محاسبه گردید (۲۶). این شاخص عبارت است از نسبت بین فراوانی کدون‌های بهینه و کل تعداد کدون‌های مترادف. این ارزش از ۰ (زمانی که ژن کدون بهینه ندارد) تا ۱ (زمانی که یک ژن ترکیبی از کدون‌های بهینه دارد) تغییر می‌کند (۳۳) و (۳۴).

درصد سیتوزین و گوانین موجود در توالی‌ها و GC3s (در جایگاه سوم) با نرم‌افزار کدون دبلو (نسخه ۱.۴.۲) محاسبه گردید. بررسی جزایر CpG با استفاده از برنامه آنالین CpGplot انجام گرفت.

نتایج و بحث

بررسی توالی‌های کد شونده در گونه‌های مورد پژوهش نشان داد که کدون پایانی برای همه توالی‌ها یکسان و TAA می‌باشد (جدول ۱). کدون خاتمه TAA دارای کارایی بیشتری در ترجمه است، زیرا هر دو، تری پپتیدهای فاکتورهای آزاد کننده ۱ و ۲ (فاکتور آزاد کننده ۱ رمز گشای TAA و TGA و فاکتور آزاد کننده ۲ رمز گشای TAA) می‌توانند کدون پایانی TAA را شناسایی کنند و به ترجمه خاتمه دهند (۳۳). تنوع در کدون پایانی سبب ایجاد یک توزیع فنوتیپی خاص در گونه‌ها می‌شود. ایجاد جهش در کدون پایانی، بروز وضعیت تازه‌ای از بیان پروتئینی را به دنبال خواهد داشت. به طور کلی دو روش وجود دارد که در آن یک ژن قادر است بر روند عملکرد یک صفت مؤثر باشد. رویه نخست، وقوع پدیده جهش در مناطق کد کننده ژن می‌باشد که ممکن است کیفیت پروتئین کد گذاری شده را تغییر دهد، در این مورد، جهش سبب

فراوانی همه کدون‌های استفاده شده در آن ژن محاسبه می‌شود. برای بدست آوردن این شاخص، ابتدا باید از ارزش‌های گرایش کدونی متفاوت برای ژن‌های با بیان بالا جدولی تهیه شود. از این جدول برای شناسایی کدون‌های با فراوانی بالا برای هر اسید آمینه استفاده می‌شود. سپس قدرت سازگاری یک کدون (wi) طبق فرمول زیر برآورد شد (۳۴).

$$w_i = \frac{RSCU_i}{RSCU_{max}} \quad \text{فرمول ۲}$$

$RSCU_{max}$ = بیشترین فراوانی ارزش کارکرد نسبی کدون‌های مترادف برای کدون‌های هر اسید آمینه. ارزش شاخص سازگاری کدون برای یک ژن به صورت میانگین هندسی ارزش‌های W_i برای همه کدون‌های مورد استفاده در آن ژن طبق مدل ذیل محاسبه شد (۳۴).

$$CAI = \left(\prod_{i=1}^L w_i \right)^{1/L} \quad \text{فرمول ۳}$$

L = تعداد کدون‌ها
 مقادیر نزدیک به ۱ نشان دهنده انحراف بیشتر به سمت ژن‌های با بیان بالا است. از این رو ژن‌های دارای عدد بزرگتر میزان بیان بیشتری دارند. همچنین مقدار شاخص تعداد کدون‌های مؤثر با نرم افزار کدون دبلو^۸ (نسخه ۱.۴.۲) مورد برآورد قرار گرفت. این شاخص برای تخمین میزان انحراف کدونی در یک ژن مورد استفاده قرار می‌گیرد (۳۳ و ۳۴). مقدار این شاخص در محدوده ۲۰ تا ۶۱ تغییر می‌کند. عدد ۲۰ نشان دهنده بیشترین انحراف کدونی و عدد ۶۱ بیانگر این است که هیچ انحرافی در کدون مترادف وجود نداشته و کدون‌ها برای هر اسید آمینه به طور یکسان مورد استفاده قرار گرفتند. شاخص تعداد کدون‌های مؤثر^۹ با استفاده از فرمول ذیل محاسبه شد (۳۳ و ۳۴).

$$ENC = 2 + S + (29/S^2) + (1 - S^2) \quad \text{فرمول ۴}$$

8. CodonW

9. Effective number of codons (ENC)

3. Frequency of optimal codon (Fop)

است که یک چهارم از محصول فنوتیپی پروتئین کالپاستاتین در گونه کژگاو حاصل بیان کدون‌های بهینه می‌باشد. این شاخص برای پیش‌بینی میزان بیان ژن، ارزیابی میزان سازگاری ژن‌های وکتورهای بیانی در میزبان‌های مختلف و مقایسه ترجیح کدونی استفاده می‌شود. نظریان و همکاران، توانستند با بهینه‌سازی ژن زیر واحد بتا کلراتوکسین شاخص سازگاری کدون که برای ژن طبیعی ۶۱ درصد بوده است را در ژن بهینه شده به مقدار ۹۲ درصد برسانند (۲۴).

تغییر اسید آمینه و در نتیجه تغییر در ساختار اولیه پروتئین شده که باعث نقص در عملکرد پروتئین می‌شود. رویه دوم، وقوع جهش در مناطق تنظیم کننده ژن است که می‌تواند میزان نسخه‌برداری ژن در سلول و نیز عملکرد پروتئین تولید شده را تحت تأثیر قرار دهد. بنابراین، انتظار می‌رود که این نوع جهش‌ها منجر به تغییرات فنوتیپی شوند (۲۵). آنالیز با نرم افزار کدون دبلو (نسخه ۱.۴.۲) نشان داد، مقدار شاخص سازگاری کدون در گونه کژگاو بالاترین و در گونه گوسفند کمترین مقدار بود. این بدین معنی

جدول ۱: شماره دسترسی، کدون پایانی، اندازه توالی و شاخص‌های انحراف کدونی برای توالی کدشونده ژن کالپاستاتین در گونه‌های مورد مطالعه

Table 1. Accession number, stop codon, CDS length and codon usage values for the coding sequences of the Calpastatin gene in studied species

گونه Species	شماره دسترسی Accession number	شاخص سازگاری کدون CAI	وفور کدون‌های بهینه Fop	تعداد کدون موثر ENC	تعداد کل کدون‌های مترادف L-sym	درصد آدنین و تیمین AT%	وزن مولکول دورشته‌ای MWDS	کدون پایانی Stop codon	اندازه‌ها CDS length
انسان Homo sapiens	D16217.1	0.247	0.418	51.94	698	54.54	54.54	TAA	2127
موش Mus musculus	AB026997.1	0.237	0.414	53.98	781	52.22	52.22	TAA	2187
گوسفند Ovisaries	U66320.1	0.236	0.424	53.99	717	50.98	50.98	TAA	2367
بز Capra hircus	GU944861.1	0.242	0.412	55.68	721	49.75	49.75	TAA	2361
گاو Bos Taurus	AF159246.1	0.255	0.430	52.05	776	50.70	50.70	TAA	2361
کژگاو Bos grunniens	EU009141.1	0.256	0.429	53.16	776	50.57	50.57	TAA	2172

L_Sym = تعداد کل کدون‌ها منهای کدون خاتمه، تریپتوفان و متیونین

L_Sym = Total number of codons except for stop codon, tryptophan and methionine

انتظار برای گاو و کژگاو بوده است و هر دو گونه با گونه بز به یک میزان اختلاف را نشان دادند. در ضمن بالاترین اختلاف بین موش با دام‌های اهلی مشاهده شد (جدول ۲).

اختلاف و تنوع در توالی‌های کد کننده ژن کالپاستاتین در قالب ماتریس فاصله توسط برنامه Pairwise Distance موجود در نرم‌افزار مگاهفت (۷.۰.۱۸) مورد بررسی قرار گرفت. نتایج نشان داد که بیشترین تشابه ژنتیکی در خروجی ماتریس مطابق

جدول ۲: ماتریس فاصله ژنتیکی بین گونه‌های مورد پژوهش با مقایسه توالی ژن کالپاستاتین

Table 2. Genetic matrix of distance among the studied species by comparing the Calpastatin gene sequence

گونه Species	گوسفند Ovis aries	بز Capra hircus	کز گاو Bos grunniens	گاو Bos Taurus	موش Mus musculus	انسان Homo sapiens
گوسفند Ovis aries	----	----	----	----	----	----
بز Capra hircus	0.038	----	----	----	----	----
کز گاو Bos grunniens	0.045	0.062	----	----	----	----
گاو Bos Taurus	0.046	0.062	0.008	----	----	----
موش Mus musculus	0.316	0.342	0.316	0.316	----	----
انسان Homo sapiens	0.195	0.217	0.201	0.198	0.264	----

جدول ۳: فراوانی نوکلئوتیدی در جایگاه‌های اول تا سوم کدون توالی ژن کالپاستاتین در گونه‌های مورد پژوهش

Table 3. Nucleotide dispersion from first to third codon sequences of the Calpastatin gene in studied species

گونه Species	تیمین جایگاه ۱ T-1	سیتوزین جایگاه ۱ C-1	آدنین جایگاه ۱ A-1	گوانین جایگاه ۱ G-1	موقعیت Pos #1	تیمین جایگاه ۲ T-2	سیتوزین جایگاه ۲ C-2	آدنین جایگاه ۲ A-2	گوانین جایگاه ۲ G2	موقعیت Pos #2
گوسفند Ovis aries	17	19.9	31.7	31.2	725.0	11	31.3	34.7	22.8	723.0
بز Capra hircus	17	20.5	30.3	31.8	732.0	12	31.8	33.6	22.7	726.0
کز گاو Bos grunniens	17	19.8	32.1	30.7	789.0	11	31.4	34.4	22.8	784.0
گاو Bos taurus	17	19.9	32.3	30.4	789.0	11	31.5	34.4	22.7	784.0
موش Mus musculus	18	19.5	32.4	30.5	790.0	13	29.7	34.1	23.0	787.0
انسان Homo sapiens	20	17.2	32.8	30.2	711.0	13	28.9	35.4	22.8	706.0
میانگین Avg.	18	19.5	31.9	30.8	756.0	12	30.8	34.4	22.8	751.7
گونه Species	تیمین جایگاه ۳ T-3	سیتوزین جایگاه ۳ C-3	آدنین جایگاه ۳ A-3	گوانین جایگاه ۳ G-3	موقعیت Pos #3	تیمین یوراسیل T(U)	سیتوزین C	آدنین A	گوانین G	کل Total
گوسفند Ovis aries	20	24.3	38.3	17.8	724.0	16.0	25.1	34.9	23.9	2172.0
بز Capra hircus	19	25.0	37.2	18.9	729.0	16.0	25.7	33.7	24.5	2187.0
کز گاو Bos grunniens	19	25.0	37.2	18.7	788.0	16.0	25.4	34.6	24.1	2361.0
گاو Bos taurus	20	25.0	36.9	18.4	788.0	16.1	25.5	34.6	23.8	2361.0
موش Mus musculus	23	20.9	36.3	19.7	790.0	18.0	23.4	34.3	24.4	2367.0
انسان Homo sapiens	21	21.0	41.3	16.3	710.0	18.1	22.3	36.5	23.1	2127.0
میانگین Avg.	20	23.5	37.8	18.3	754.8	16.7	24.6	34.7	24.0	2262.5

اسید آمینه انتهایی تنها در انسان اسید اسپارتیک است که دارای $pka=3.9$ بوده و در سایر گونه‌ها اسید آمینه سرین جای دارد که یک آمینو اسید الکل دار با فرمول $C_3H_7NO_3$ است که نسبت به آنزیم‌هایی مانند تریپسین، کیموتریپسین و الاستاز حساس می‌باشد. مقدار عددی گرایش کدون (میل کدون) در پروتئین برای هر کدام از کدون‌های مربوط به این آمینو اسیدها، کدون‌های ارجح را به ترتیب برای سرین ($AGC=1/38$) و برای اسید اسپارتیک ($GAU=1/01$) محاسبه کرد که در جدول ۴ آورده شد.

جایگاه هر نوکلئوتید در موقعیت‌های اول تا سوم تعیین و نتایج در جدول ۳ آورده شد. قرار گرفتن نوکلئوتید گوانین و سیتوزین در جایگاه سوم برای آنالیزهای انحرافات کدون در موجودات زنده بسیار مهم می‌باشد. قابل ذکر است که تعداد کل این نوکلئوتیدها در ژن سبب بروز رفتارهای ویژه خواهد شد چرا که محتوای این دو نوکلئوتید با انحراف کدون، طول ژن، مدت زمان همانند سازی و نرخ میتلاسیون همبستگی دارد (۱۵ و ۲۸).
بررسی توالی اسیدهای آمینه در پروتئین کالپاستاتین با نرم افزار بیوادیت (7.0.9.0) نشان داد

جدول ۴: مقدار RSCU برای کدون‌های جایگاه کالپستاتین در گونه‌های مورد مطالعه.

Table 4. The amount of RSCU for codons of the Calpestatin locus in studied species

کدون	تعداد	گرایش کدون	کدون	تعداد	گرایش کدون	کدون	تعداد	گرایش کدون	کدون	تعداد	گرایش کدون
Codon	Count	مترادف RSCU	Codon	Count	مترادف RSCU	Codon	Count	مترادف RSCU	Codon	Count	مترادف RSCU
UUU(F)	2.3	0.51	UCU(S)	15.3	1	UAU(Y)	3.5	1.08	UGU(C)	7.8	0.94
UUC(F)	6.8	1.49	UCC(S)	17.3	1.13	UAC(Y)	3	0.92	UGC(C)	8.8	1.06
UUA(L)	6.3	1.02	UCA(S)	19.8	1.3	UAA(*)	7.2	0.83	UGA(*)	17.5	2.03
UUG(L)	5.5	0.89	UCG(S)	5.8	0.38	UAG(*)	1.2	0.14	UGG(W)	5.8	1
CUU(L)	7.8	1.26	CCU(P)	16.2	1.05	CAU(H)	6.3	0.84	CGU(R)	3.3	0.44
CUC(L)	8.7	1.4	CCC(P)	16.3	1.06	CAC(H)	8.7	1.16	CGC(R)	1.3	0.18
CUA(L)	1.5	0.24	CCA(P)	23.7	1.54	CAA(Q)	19.2	1.25	CGA(R)	3.8	0.5
CUG(L)	7.3	1.18	CCG(P)	5.3	0.35	CAG(Q)	11.5	0.75	CGG(R)	5.8	0.77
AUU(I)	6.2	1.09	ACU(T)	14.3	1.09	AAU(N)	5.3	0.84	AGU(S)	12.3	0.81
AUC(I)	6.5	1.15	ACC(T)	16.5	1.25	AAC(N)	7.3	1.16	AGC(S)	21	1.38
AUA(I)	4.3	0.76	ACA(T)	20.2	1.53	AAA(K)	48.5	1.14	AGA(R)	29	3.81
AUG(M)	4.7	1	ACG(T)	1.7	0.13	AAG(K)	36.7	0.86	AGG(R)	2.3	0.31
GUU(V)	4.8	0.94	GCU(A)	19.8	1.39	GAU(D)	18	1.01	GGU(G)	8	0.61
GUC(V)	6.8	1.32	GCC(A)	21.3	1.5	GAC(D)	17.5	0.99	GGC(G)	8.2	0.63
GUA(V)	3.2	0.61	GCA(A)	14.3	1.01	GAA(E)	39.5	1.22	GGA(G)	26.2	2.01
GUG(V)	5.8	1.13	GCG(A)	1.5	0.11	GAG(E)	25.3	0.78	GGG(G)	9.8	0.75

جدول ۵: فراوانی نوکلئوتید GC در جایگاه اول تا سوم ژن کالپاستاتین در گونه‌های مورد مطالعه.

Table 5. Frequency of GC nucleotide from first to third codon sequences of Calpestatin gene in studied species

گونه	درصد سیتوزین و گوانین جایگاه اول	درصد سیتوزین و گوانین جایگاه دوم	درصد سیتوزین و گوانین جایگاه سوم	درصد سیتوزین و گوانین کل
Species	GC1s	GC2s	GC3	GC
انسان	0.474	0.517	0.342	0.455
Homo sapiens				
موش	0.50	0.527	0.401	0.478
Mus musculus				
گوسفند	0.511	0.541	0.431	0.491
Ovisaries				
بز	0.523	0.545	0.445	0.503
Capra hircus				
گاو	0.503	0.527	0.436	0.494
Bos Taurus				
کوزگاو	0.505	0.517	0.437	0.495
Bos grunniens				

به مورد انتظار CpG، نرخ نوترکیبی و تراکم ژن مشاهده گردید. با افزایش سایز کروموزمها تراکم جزایر CpG کاهش یافت و همچنین تراکم جزایر CpG در نواحی تلومری کروموزمها بیشتر از سایر قسمت‌ها مشاهده گردید که می‌تواند علت همبستگی مثبت بین نرخ نوترکیبی و تراکم جزایر CpG باشد. جهت یافتن تفاوت‌های جزایر CpG بین ژنوم گاو و سایر مهره داران جزایر CpG در یازده مهره دار دیگر بررسی گردید. تغییرات تراکم جزایر CpG در بین حیوانات مورد مطالعه بسیار متفاوت بود. آنها نشان دادند که یافته‌های آنها می‌تواند در درک بهتر نقش اپی ژنتیکی و نقش تکامل مولکولی جزایر CpG در ژنوم گاو موثر باشد. در پژوهشی که برای مطالعه مقایسه ای این جزایر در شتر و گاوها صورت گرفت مانند نتایج به دست آمده از تحقیق، در آنجا نیز تعداد CpGI با نرخ مشابه برآورد شده است (۳ و ۴).

قرار گرفتن نوکلئوتید گوانین و سیتوزین در جایگاه سوم برای آنالیزهای انحرافات کدونی در موجودات زنده بسیار مهم می‌باشد. بررسی جزایر CpG به صورت آنلاین نشان داد که تمام توالی‌ها دارای این جزایر به طول بیش از ۲۰۰ نوکلئوتید می‌باشند. همچنین نسبت مشاهده شده به مورد انتظار که نسبت مشاهده شده این جزایر به مورد انتظار بوده بزرگتر از ۶۰ درصد می‌باشد. برآزنده و همکاران در پژوهشی به بررسی این جزایر در ژنوم گاو با استفاده از مدل‌های مخفی مارکوف پرداختند و در پژوهش آنها تعداد ۹۰۶۶۸ جزیره CpG در ژنوم گاو پیش بینی گردید. تغییرات تعداد و تراکم جزایر CpG در طول کروموزومها بسیار کم بود. کروموزوم ۲۵ دارای بیشترین تعداد (۴۵۵۶) و تراکم (۱۰۶۱۲۰ CGIs/Mb) جزایر CpG بود. همبستگی مثبت و معنی دار بین تراکم جزایر CpG با محتوی GC، نسبت مشاهده شده

جدول ۶: بررسی جزایر CpG در توالی کد شونده ژن کالپاستاتین در گونه‌های مورد مطالعه.

Table 6. Study of CpG islands in coding sequence of Calpastatin gene in studied species

گوسفند	انسان	گاو	کژگاو	موش	بز	نسبت مشاهده شده به مورد انتظار
Sheep	human	Cow	grunniens Bos	Mic	Goat	observed/ expected ratio
>60	>60	>60	>60	>60	>60	(O/E %)

بیشترین مشارکت در ساختار پروتئین کالپاستاتین را داشته‌اند. در پژوهشی بین بافت‌های سرطانی و سالم ریه، مشخص شد تغییر نسبت در آمینو اسیدها بگونه اتفاق افتاده که آمینو اسیدهای آرژنین، پرولین، سرین، گلايسين، لایزین و ترئونین بالاترین فراوانی را در بافت‌های سرطانی نسبت به بافت سالم پیدا کرده‌اند در حالی که در حالت سلامت بیشترین فراوانی آمینو اسیدی برای آمینو اسیدهای لوسین، سیستئین، فیل آلانین، هیستیدین و تریپتوفان بوده است (۱۸). توبستر و همکاران این گونه بیان کردند که در افراد سرطانی فراوانی آمینو اسیدهای سیستئین، تریپتوفان و هیستیدین در پروتئوم بالا بوده در حالی که آمینو اسید آرژنین کمترین فراوانی را داشته است (۲۴).

با بررسی جدول ۷ می‌توان دریافت که، ۳۳/۷۰٪ توالی آمینواسیدی این پروتئین را آمینواسیدهای غیرقطبی تشکیل داده‌اند که در واکنش‌های هیدروفوبیک شرکت دارند و آمینواسیدهای قطبی یونی نیز ۳۷/۶۶٪ کل توالی‌های موجود را شامل می‌شوند. نسبت آمینواسیدهای قطبی به غیرقطبی در این پروتئین ۲ به ۱ می‌باشد. آمینواسید تریپتوفان که در جذب نور بالای ۲۸۰ نانومتر پروتئین‌ها نقش اصلی را بازی می‌کند با ۵ تکرار در کل توالی‌ها و تنها در گونه گاو، کژگاو و گوسفند به ترتیب با ۲، ۲ و ۱ تکرار، کمترین مشارکت در ساختار این پروتئین را داشته در حالی که آمینواسیدهای لیزین، اسیدگلوتامیک و سرین به ترتیب با ۶۲۳، ۵۱۶ و ۴۹۵ تکرار در توالی‌ها

جدول ۷: فراوانی آمینو اسیدی در زنجیره پروتئین کالپاستاتین در گونه‌های مورد پژوهش

Table 7. Frequency of amino acid in Calpastatin protein chain in studied species

گونه species	آلانین Al	سیتئین Cy	آسپاراتیک Asp	گلو تامیک Glu	فیل آلانین Ph	گالینین Gly	هیستیدین Hi	ایزولوسین I	لیرین Ly	لوسین Le	کل Total
انسان Homo sapiens	61	6	67	83	5	37	7	18	100	45	429
موش Mus musculus	61	4	62	86	8	45	15	16	100	50	447
گوسفند Ovis aries	57	7	56	82	6	33	13	12	97	47	410
بز Capra hircus	71	5	62	89	6	31	6	13	96	44	423
کز گاو Bos grunniens	78	6	63	88	6	39	10	13	116	47	466
گاو Bos Taurus	77	5	62	88	6	38	10	13	114	47	460
جمع SUM	405	33	372	516	37	223	61	85	623	280	2635
میانگین Avg.	67.5	5.5	62	86	6.1	37.1	10.1	14.1	103.8	46.6	438.8
حداکثر Max	78	7	67	89	8	45	15	18	116	50	493
حداقل Min	57	4	56	82	5	31	6	12	96	44	393
انحراف معیار SD	7.8	0.8	2	2.3	0.6	3.5	2.5	1.8	7.4	1.4	

گونه‌ها species	متیونین M	آسپاراتین As	پروлін Pt	گلو تامین GI	آسپاراتین Ar	سرن Se	تیروزین Th	والین Va	تریوفان Tr	تیروزین Ty	کل Total
انسان Homo sapiens	10	10	69	24	18	71	46	25	0	6	279
موش Mus musculus	7	13	78	31	24	93	52	35	0	8	341
گوسفند Ovis aries	5	10	78	27	33	74	46	33	1	6	313
بز Capra hircus	7	9	82	20	24	76	49	32	0	6	305
کز گاو Bos grunniens	8	10	84	22	15	90	49	33	2	7	320
گاو Bos Taurus	8	12	84	22	16	91	51	33	2	7	32
جمع SUM	45	64	475	146	130	495	293	191	5	40	1884
میانگین Avg.	7.5	10.6	79.1	24.3	21.6	82.5	48.8	31.8	0.83	6.6	313.63
حداکثر Max	10	13	84	31	33	93	52	35	2	8	361
حداقل Min	5	9	69	20	15	71	46	25	0	6	266
انحراف معیار SD	1.1	1.2	4.1	3.1	5.3	8.8	1.9	2.28	0.83	0.6	

است. مقدار شاخص آلیفاتیک در توالی زنجیره پروتئینی کالپستاتین در گونهانساندر مقایسه با سایر گونه‌ها بیشتر بوده و این بدان معنی است که پروتئین کالپستاتین در انسان در برابر گرما مقاوم‌تر از سایر گونه‌هاست. قاضی و حدادی در بررسی پروتئین گیرنده $IFN-\alpha/\beta$ برای مقابله با اثرات آنتاگونیستی پروتئین VP35 ویروس ابولا، مقدار شاخص آلیفاتیک در این پروتئین را ۸۵/۵۰ و در پروتئین نو ترکیب را ۸۷/۶۶ اعلام کردند و بیان داشتند که عامل این افزایش در مقدار شاخص، جان‌شینی بیشتر آمینواسیدهای آلانین، والین، ایزولوسین و لوسین بوده است که مقاومت دمایی بیشتری در زمان دناتوره شدن را نشان می‌دهند (۹).

سرین موجود در توالی آمینو اسیدی کالپستاتین از دسته الکل نوع اول بوده که در کاتالیز آنزیمی به عنوان گروه نوکلئوفیل نقش دارد و در فعالیت آنزیمی پروتئین کالپستاتین موثر است. در مجموع ۳۲/۹۹٪ توالی‌های پروتئینی در گونه‌های مورد پژوهش را آمینو اسیدهای ضروری تشکیل داده اند که بیانگر حجم وابستگی این پروتئین به مواد خوراکی دریافت شده می‌باشد. آمینواسیدهای گوگرددار تنها ۰/۰۱۷۲٪ توالی‌ها را در بر می‌گیرند. برنامه ساپما نشان داد که در هر گونه، توالی پروتئینی دارای چه تعداد ماریپچ آلفا، صفحہ بتا، پیچ بتا و ماریپچ تصادفی باشد. گونه موش دارای بیشترین ماریپچ آلفا و همچنین گونه کژ گاو به خاطر داشتن صفحات بتای فراوان در ساختار پروتئینی خود، قالب زیگزاک مانند را به خود گرفته

جدول ۸: بررسی ساختار فضایی پروتئین CAST در گونه‌های مورد مطالعه

Table 8. Study of spatial structure of CAST protein in studied species

گونه Species	شماره دسترسی Accession number	شاخص آلفا هلیکس Aliphatic index	ماریپچ آلفا Hh	صفحات بتا Ee	پیچ بتا Tt	ماریپچ تصادفی Cc
انسان Homo sapiens	BAA03747.1	53.56	253(35.73%)	29(4.10%)	34(4.80%)	392(55.37%)
موش Mus musculus	BAA84768.1	53.29	283(35.91%)	23(2.92%)	34(4.31%)	448(56.85%)
گوسفند Ovisaries	AAB07483.1	52.95	216(29.88%)	35(4.84%)	33(4.56%)	439(60.72%)
بز Capra hircus	ADE58451.1	53.04	234(32.17%)	31(4.26%)	33(4.53%)	430(59.07%)
گاو Bos Taurus	AAD43339.1	51.74	241(30.66%)	33(4.20%)	32(4.07%)	480(61.07%)
کژگاو Bos grunniens	ABS19944.1	51.87	247(31.42%)	36(4.58%)	36(4.58%)	467(59.41%)

پروتئینی دارای تعداد اتم‌های اکسیژن، هیدروژن، نیتروژن و کربن بیشتر اما تعداد اتم گوگرد آن کمتر در مقایسه با سایر گونه‌ها بوده است. مطابق جدول شماره ۸، گونه گوسفند با شماره دستیابی AAB07483.1 با فرمول $C_{3384}H_{5549}N_{983}O_{1163}S_{12}$ دارای بیشترین اسیدیته ایزوالکتریک بوده است. این وضعیت باعث می‌شود که پروتئین کالپستاتین در

خصوصیات بیوشیمیایی و فیزیکی پروتئین کالپستاتین در گونه‌های مورد مطالعه در جداول ۱۰ و ۹ ارایه شده است. نیمه عمر پروتئین کالپستاتین در شرایط برون تنی (لوله آزمایشگاهی) در همه گونه‌ها ۳۰ ساعت، در محیط کشت مخمر کمتر از ۲۰ ساعت و در محیط کشت اشرشیاکلی کمتر از ۱۰ ساعت برآورد شد. گونه موش با بلندترین طول توالی زنجیره

شده که پایداری پروتئینی را افزایش می‌دهد. بیشترین بار الکتریکی مثبت (آرژنین+لیزین) در توالی گونه کژگاو و کمترین مقدار در گونه انسان مشاهده گردید.

ساختار خود در این گونه مقدار زیادی از صفحات بتا داشته باشد. بالا بودن اتم گوگرد در توالی زنجیره پروتئینی کالپاستاتین در گونه انسان نیز سبب تشکیل آمینو اسید سیستین و پیوندهای دی سولفیدی بیشتر

جدول ۹: خصوصیات بیوشیمیایی توالی‌های آمینواسیدی پروتئین کالپاستاتین در گونه‌های مورد پژوهش

Table 9. Biochemical characteristics of amino acid sequences of Calpastatin protein in specimens.

شماره دسترسی پروتئین Protein number	وزن مولکولی MW	اسیدیته ایزوالکتریک PI	تعداد آمینو اسید Number of amino acids	مقدار بار مثبت +charge	مقدار بار منفی -charge	شاخص پایداری II	شاخص گراوی GRAVY
BAA03747.1	76484.7	4.99	708	118	150	49.83	-1.193
BAA84768.1	84921.7	5.37	788	124	148	50.09	-1.164
AAB07483.1	78998.9	6.08	723	130	138	52.87	-1.247
ADE58451.1	78379.7	5.01	728	120	151	51.64	-1.164
AAD43339.1	84089.26	5.38	786	130	150	48.76	-1.149
ABS19944.1	84018.29	5.38	786	131	151	48.55	-1.141

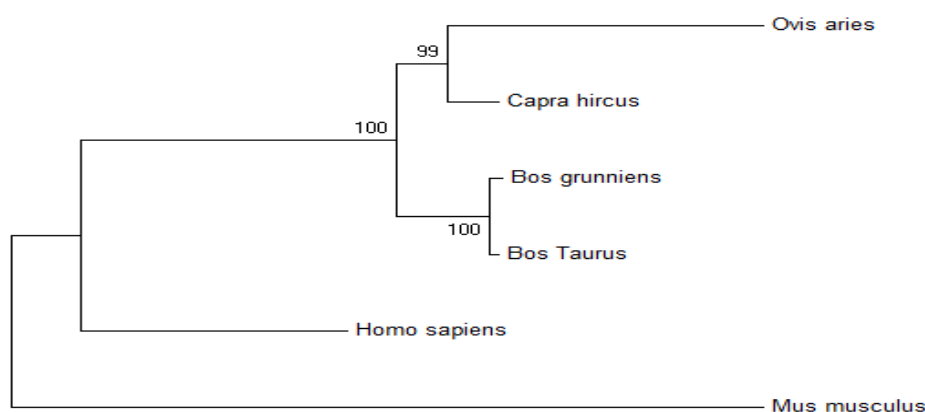
جدول ۱۰: خصوصیات فیزیکی توالی‌های آمینواسیدی پروتئین کالپاستاتین در گونه‌های مورد پژوهش

Table 10. Physical characteristics of amino acid sequences of Calpastatin protein in studied species

شماره توالی پروتئین protein seq	کربن C	هیدروژن H	نیتروژن N	اکسیژن O	گوگرد S	فرمول Formul	کل اتم‌ها Total atoms
BAA03747.1	3262	5348	910	1166	16	C ₃₂₆₂ H ₅₃₄₈ N ₉₁₀ O ₁₁₆₆ S ₁₆	10702
BAA84768.1	3631	5916	1034	1282	11	C ₃₆₃₁ H ₅₉₁₆ N ₁₀₃₄ O ₁₂₈₂ S ₁₁	11874
AAB07483.1	3384	5549	983	1163	12	C ₃₃₈₄ H ₅₅₄₉ N ₉₈₃ O ₁₁₆₃ S ₁₂	11091
ADE58451.1	3354	5487	937	1191	12	C ₃₃₅₄ H ₅₄₈₇ N ₉₃₇ O ₁₁₉₁ S ₁₂	10981
AAD43339.1	3609	5896	1004	1270	13	C ₃₆₀₉ H ₅₈₉₆ N ₁₀₀₄ O ₁₂₇₀ S ₁₃	11792
ABS19944.1	3608	5895	1001	1267	14	C ₃₆₀₈ H ₅₈₉₅ N ₁₀₀₁ O ₁₂₆₇ S ₁₄	11785

سبب دور شدن این گونه به صورت نسبی از سایر نشخوارکنندگان شده است. گونه‌های گاو و کژگاو همچنین بز و گوسفند در شاخه‌های نزدیک‌تر به هم جای گرفتند. نتیجه آنالیز نشان داد گونه گوسفند با گونه موش بیشترین فاصله تکاملی و به عبارت دیگر بیشترین ناهمسانی در توالی پروتئینی را نشان می‌دهند.

از داده‌های حاصل از توالی‌های مولکولی برای بررسی روابط تکاملی بین گونه‌های مختلف و برآورد زمان واگرایی ارگانسیم‌ها در تبارشناسی مولکولی استفاده می‌شود (۲۹). رابطه تکاملی پروتئین کالپاستاتین در شکل ۱ ارایه شد. همانگونه که در شکل مشاهده می‌شود نشخوارکنندگان بیشترین نزدیکی و یک پارچگی را در توالی آمینواسیدی حفظ کرده‌اند. البته جهش‌های ایجاد شده در توالی پروتئینی گوسفند



شکل ۱: درخت فیلوژنتیک میان پستانداران بر اساس توالی اسیدآمینۀ ژن CAST

در ارتباط با صفات رشد، صفات تولیدی و صفات مربوط به مقاومت بیماری‌ها هستند اهمیت واقعی خود را نشان می‌دهد.

منابع

1. Bahrapour, V., Mohammadabadi, M.R., Mirzaei, H.R., Baghizadeh, A., Dashab, G.R., Mohammadi, A., Alinaghizadeh, R.A., Soflaei, M. and Khesali, A. 2008. Molecular analysis of Calpastatin gene in Kermani sheep herds. *J. Agri. Sci and Natural Resources*. 15:124-131. (In Persian)
2. Balcerzak, D., Cottin, P., Poussard S., Cucuron, A., Brustisj, J. and Ducastaing, A. 1998. Calpastatin-modulation of m-calpain activity is required for myoblast fusion. *Europ. J. Cell. Bio*. 75: 247-253.
3. Barazandeh, A., Mohammadabadi, M.R., Ghaderi, M. and Nezamabadipour, H. 2016. Predicting CpG islands and their relationship with genomic feature in cattle by hidden Markov model algorithm. *J. App. Anim. Sci*. 6:571-579. (In Persian)
4. Barazandeh, A., Mohammadabadi, M.R., Ghaderi, M. and Nezamabadipour, H. 2016. Genome-wide analysis of CpG islands in some livestock genomes and their relationship with genomic features. *Czech. J. Anim. Sci*. 61: 487-495.
5. Bishop, M.D., Koohmaraie, M., Killefer, J. and Kappes, S. 1993. Restriction fragment length polymorphisms of the

نتیجه‌گیری نهایی

مطالعه و بررسی ساختار ژنتیکی افراد در سطح DNA، فرصت‌های بیشماری از پیشرفت ژنتیکی مرتبط به صفات اقتصادی مهم در دام‌های اهلی را فراهم کرده است (۶). با استفاده از ابزارهای بیوانفورماتیکی می‌توان روی داده‌های مولکولی تحلیل‌های دقیق، سریع و گسترده‌ای را انجام داد. در ژنتیک مولکولی جهت بررسی تکامل ژنتیکی، توالی ژنوتیپی افراد به منظور بررسی مجموعه‌ای از جایگاه‌های ژنی خاص بررسی می‌گردد. استفاده از این اطلاعات در برنامه‌های اصلاح نژادی، پتانسیل کافی به منظور افزایش دقت در انتخاب را فراهم می‌کند (۱۹). بررسی بیوانفورماتیکی امکان قدم زدن در سطح پروتئین، RNA و DNA را ایجاد کرده تا از طریق آن برنامه ریزان اصلاح دام بتوانند تصمیمات راحت تری درباره نقشه‌های اصلاحی داشته باشند. مقایسات بین گونه‌ای نیز این امکان را فراهم کرده تا با توجه به صفت مورد نظر در بین گونه‌ها، توالی که هدف اصلاح گران را بهتر تامین می‌کند شناخته شود و در مرحله بعد با مهندسی ژنتیک، جایگزین توالی ژن مربوط در دام‌های هدف گردد. این رویه خصوصاً در ژن‌هایی که

14. Hediger, R., Ansari, H.A. and Stranzinger, G.F. 1991. Chromosome banding and gene localizations support extensive conservation of chromosome structure between cattle and sheep. *J. Cytogen. Geno. Res.* 57: 127- 134.
15. Ikemura, T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *J. Mol. Biol. Evol.* 2: 13-34.
16. Jiang, S.T. 1998. Contribution of muscle proteinases to meat tenderization. *Proceedings of the National science Council, Republic of China. Part B, Life. Sci.* 22: 97-107.
17. Jones, P.A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Natu. Rev. Genet.* 13: 484-492.
18. Kappes, S.M., Keele, J.W., Stone, R.T., Sonstegard, T.S., Smith, T.P, L., McGraw, R.A., Lopezcorrales, N.L. and Beattie, C.W. 1997. A second-generation linkage map of the bovine genome. *J. Geno. Res.* 7: 235.
19. Khalid, R. 2010. Application of data mining in bioinformatics. *Indian. J. Comp. Sci. Eng.* 1: 114-118.
20. Khosravi, M., Fakhr Kazemi, M., Mohamadi, A. and Nasiri, M.R. 1389. Study on the Genetic polymorphisms of candidate genes (Calpastatin and *BoLA*) in Sistani cattle using PCR-RFLP. *J. Agri. Bio. technol.* 9: 35-42.
21. Kumar, S., Stecher, G. and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *J. Mol. Biol. Evol.* 33: 1870-4.
22. Mahdavi Mamghani, A., Shoja, J., Pirany, N. and Elyasi, G. 1388. PCR-SSCP comparison of the Calpastatin gene polymorphism in Iranian Ghezel sheep and Sarabi cows. *J. Anim. Sci. Res.* 19: 1-8.
23. Mohammadabadi, M.R. 2016. Allelic Diversity of Calpastatin Gene in Sanjabi Sheep. *J. Molecular And Cellular Research.* 28-3: 395-402. (In Persian).
24. Nazarian, Sh., Arefpur, M.A., Bagheripur, M.J. and Oladi, Gh. 2013. Bioinformatics study and review of gene expression optimized for subunit B bovine calpastatin gene. *J. Anim. Sci.* 71:2277.
6. Dodds, K.G., McEwan, J.C. and Davis, G.H. 2007. Integration of molecular and quantitative information in sheep and goat industry breeding programmes. *Small Ruminant. Res.* 70:32-41.
7. Elyasi Zarin Ghobaii, Gh., Shojae Ghias, J., Nasiri, M.R., Pirahri, A. and Javanmard, A. 1384. Polymorphism in Calpastatin Gene by MSPI/RFLP. Fourth Biotechnology Conference of the Islamic Republic of Iran - Tehran . 2-1. (In Persian).
8. Esmaily, S., Mohamad Talee, A. and Alamzadeh, A. 2011. computer bioinformatic and identification of motifs in Na⁺/H⁺ antiporter plant promoters. The 7 National Biotechnology Congress of Iran - Tehran . (In Persian).
9. Ghazi, Y. and Hadadi, F. 2018. Bioinformatics design of the recombinant IFN- α / β receptor protein to counteract the antagonistic effects of the VP35 protein Ebola virus. *J. Biosafety.* 2: 29.0430.
10. Gupta, S.K., Bhattacharyya, T.K. and Ghosh, T.C. 2004. Synonymous codon usage in *Lactococcus lactis*: mutational bias versus translational selection. *J. Bio. mol. Struct. Dyn.* 21: 527-535.
11. Hadizadeh, M., MohammadAbadi, M., Niazi, A., Esmailzadeh Koshkoiyeh, A., Mehdizadeh Gazooei, Y. and MolaeiMoghbel, S. 2013. Use of bioinformatics tools to study exon 2 of GDF9 gene in Tali and Beetal goats. *J. Modern Genetics.* 8: 283-288. (In Persian).
12. Hadizadeh, M., Niazi, A., Mohammad Abadi, M., Esmailzadeh, A. and Mehdizadeh Gazooei, Y. 2014. Bioinformatics analysis of the BMP15 exon 2 in Tali and Beetal goats. *J. Modern Genetics.* 9-1: 117-127. (In Persian).
13. Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic Acids Symposium Series.* 41: 95-98.

30. Taylor, R., Christiansen, J. and Goll, D. 1991. Immunolocalization of the calpains and calpastatin in human and bovine platelets. *J. Biom. Bioch. Acta*, 50: 491-498.
31. Bergund, U.W. and Helleday, T.h. 2017. Mutations in cancer cause gain of Cysteine, Histidine, and Tryptophan the expense of a net loss of Arginine on the proteome level. *J. Bio. Mole.* 7: 49.
32. Usmanova, N.M. and Tomilin, N.V. 2008. Bioinformatic analysis of retroelement associated sequences in human and mouse promoter. *J. Wor. Acad. Sci. Eng. Tech.* 2: 146-154.
33. Wei, Y., Wang, J. and Xia, X. 2016. Coevolution between stop codon usage and release factors in bacterial species. *J. Mol. Biol. Evol.* 33-9: 2357-2367.
34. Wright, F. 1990. The 'effective number of codons' used in a gene. *Gene*. 87: 23-29.
- chloratoxin as a candidate for vaccine. *J. Isfahan. Medi. School.* 32: 270-279.
25. Nei, M. 2007. The new mutation theory of phenotypic evolution. *J. Procee. Nati. Acad. Sci. USA.* 104: 12235-12242.
26. Palmer, B.R., Hickford, J.G.H. and Bickerstaffe, R. 1997. A candidate gene approach to animal quality traits. *Proceedings of the New Zealand Society of Animal Production.* 57: 294-296.
27. Peden, J.F. 2000. Analysis of codon usage. Doctoral dissertation, University of Nottingham.
28. Rao, Y.S., Chai, X.W., Wang, Z.H.F., Nie, Q.H. and Zhang, X.Q. 2013. Impact of GC content on gene expression pattern in chicken. *J. Gene. Selec. Evol.* 45:9.
29. Stamatakis, A. 2005. Phylogenetics: applications, software and challenges. *J. Canc. Geno. Prot.* 2: 301-306.



Codon usage deviations and bioinformatics analysis of encoding sequence of Calpastatin gene in some mammalian species

*P. Jalali¹, GH. Rahimi Mianji², M. Gholizadeh³, A. Pakdin Parizi⁴

¹M.Sc. Graduated, ²Professor, ³Assistant Prof., Dept. of Animal Sciences, Sari Agricultural Sciences and Natural Resources University, Sari, Iran, ⁴Assistant Prof, Genetics and agricultural Biotechnology Institute of Tabarestan, Agricultural Sciences and Natural Resources University of Sari

Received: 20/05/2018; Accepted: 24/09/2018

Abstract

Background and Object: Calpastatin (CAST) is one of the enzymes of Calpine proteolytic system. The proteolytic protein complex contains Ca⁺² dependent proteases which contributes in construction, degradation and muscle tissue compression after slaughter, and is also regarded a candidate gene associated with growth efficacy and quality of meat.

Materials and Methods: In this study, gene and protein coding sequences region of CAST in six mammalian species (human, rat, cattle, yak, sheep and goat) were examined. Gene and protein sequences were retrieved from gene bank and then analyzed. Homology analysis and alignment, phylogenetic and nucleotide diversity and variation in coding region and stop codon were carried out using the Clustal Ω and Mega7 softwares. SOPMA and Protparam programs were used for homology and alignment analyses, and to investigate the atoms diversity in protein structure, amino acid and terminal amino acid diversity in sequences retrieved from the NCBI database. Preferred codon sequences were obtained using CodonW software to explore the codon usage locus.

Results: Codon adaptation index (CAI) had the highest value for yak (0.256) and lowest value for sheep (0.236). Analysis by bioinformatic software for better understanding of protein structure of CAST showed that, in all sequences, lysine amino acid was the most frequent by 623 observations and tryptophan was the least with 5 repeats in the structure of the protein. The ration of Polar amino acids to non-polar amino acids in the protein was 2. The relative efficiency of synonymous codons (RSCU) for the amino acids serine and aspartic acid as the terminal amino acid in different species were, respectively, (AGC =1/38) and (GAU =1/01). *Ovisaries* species showed the maximum PI and *Capra hircus* species had the highest value of effective number of codons index (ENC).

Conclusion: hydrophobic amino acids constitute the main part of the amino acid sequence of Calpastatin protein. Given the role of inhibition of Calpastatin protein for the activity of the enzyme Calpain in muscle and considering that the most sequences of Calpain are captured by hydrophilic amino acids, the explore of amino acid sequence in Calpastatin and the role of these hydrophobic amino acids against the hydrophilic amino acids in Calpain is important. Calpastatin protein is much more tolerant in humans than ruminants. The codon bias analysis of the studied species showed that, in the evolution, *Bos grunniens* protein species had higher phenotype appearance for preferred codons than other species and function of the optimal codons were shown to be stronger than others.

Keywords: Mammals, Calpastatin, Bioinformatics analysis, Codon bias, spatial structure

*Corresponding author; payamjalali46@yahoo.com