# ردیابی مناطق تحت انتخاب مثبت در برخی از بزهای اهلی و وحشی ایران با استفاده از نشانگرهای مولکولی

صادق طاهری[۱]، سعید زره داران[۲]، علی جوادمنش[۲*]

۱. دانشجوی کارشناسی ارشد ژنتیک و اصلاح نژاد دام و طیور، گروه علوم دامی، دانشگاه فردوسی مشهد

۲. عضو هیأت علمی گروه علوم دامی، دانشگاه فردوسی مشهد

*ایمیل نویسنده مسئول: javadmanesh@um.ac.ir

**خلاصه**

**شناسایی این نشانه‌ها یکی از بحث‌های مورد علاقه متخصصان ژنتیک تکاملی است، زیرا می‌توانند اطلاعاتی از فرایندهای تکاملی تشکیل ژنوم تا اطلاعات عملکردی از نواحی ژنگانی را با شناسایی این نواحی به دست آورند. شناسایی نشانه‌های انتخاب ممکن است منجر به شناسایی مناطق مرتبط با صفات مهم که تحت انتخاب هستند شود. در این مطالعه، با هدف شناسایی مناطق تحت انتخاب مثبت بین گونه‌های اهلی و وحشی بز ایرانی، یک پویش گسترده در ژنوم با استفاده از چندشکلی‌های تک نوکلئوتیدی (SNP) انجام شد. داده‌های مطالعه حاضر از پروژه NEXTGENE گرفته شد. در این پژوهش ۲۰ راس بز اهلی و ۱۸ راس بز وحشی با تراشه‌های شرکت ایلومینا توالی یابی کل ژنوم شدند. برای مطالعه جهش‌ها در کل ژنوم برای هر حیوان، ابتدا جهش‌های معادل 50K شرکت ایلومینا استخراج شد. در این مطالعه برای شناسایی مناطق تحت انتخاب مثبت از آماره Fst استفاده شد. نتایج آزمون Fst بین دو جمعیت اهلی و وحشی نشان داد که ۴۸۰ منطقه ژنومی روی ۲۹ کروموزوم متأثر از انتخاب بودند. ژن‌های شناسایی شده با تنظیم اتوفاژی، تجزیه لیزین، چسبندگی سلولی، مسیر سیگنال دهی فسفولیپاز D، فعال‌سازی پلاکت، سیناپس گلوتاماتیک، سیناپس سروتونرژیک، مسیر سیگنال دهی گنادوتروپین و مسیر سیگنال دهی برای تمایز مناسب سلول‌های جنینی در ارتباط بودند. نتایج این مطالعه نشان داد که شناسایی امضاهای انتخاب ممکن است درک بهتری از فرایند اهلی سازی داشته باشد و ژن‌های نامزد می‌توانند در این فرایند سهیم باشند. با توجه به کوچک بودن جمعیت مورد مطالعه و کامل نبودن اطلاعات مربوط به حاشیه نویسی عملکردی ژن‌ها در گونه بز، مطمئنا مطالعات گسترده‌تر با تعداد نمونه‌های بیشتر، درک بهتری از ژن‌های کاندید برای صفات مهم اقتصادی در گونه‌های اهلی و وحشی و طراحی برنامه‌های اصلاح نژادی موفق، ایجاد خواهد نمود.**

**کلید واژه:** بز اهلی و وحشی، نشانه‌های انتخاب، Fst، SNP

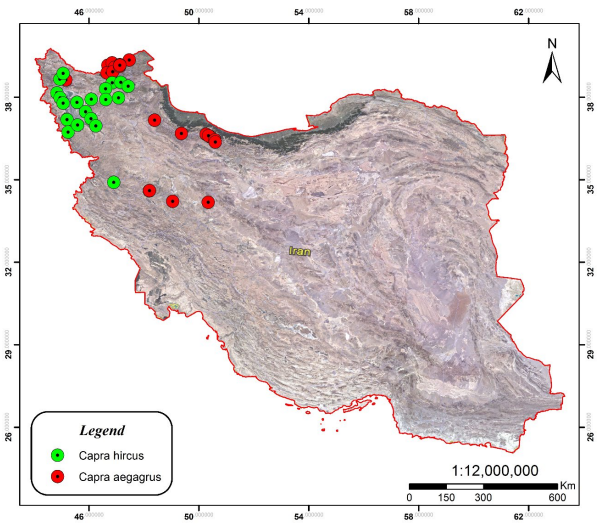
## ۱. مقدمه

بزهای وحشی از خانواده گاوسانان با نام فارسی کل و بز وحشی و اسم علمی *Capra aegagrus* است. این حیوانات از نظر جثه از بزهای اهلی امروزی بزرگتر هستند. به نرها کل و به ماده‌ها بز گفته می‌شود. کوتاه بودن دست و پا، سنگین بودن قسمت جلویی بدن و ساختمان خاص سم‌ها، آن‌ها را قادر ساخته که به آسانی در شیب‌های تند مناطق صخره‌ای تردد نمایند. بزها نیز جزو اولین حیواناتی بودند که اهلی شدند (بویازاغلو و همکاران، ۲۰۰۵). بسیاری از خصوصیات رفتاری

انواع وحشی در طی اهلی سازی تغییر یافته است (گراستیو و همکاران، ۲۰۰۵). بزها در طی اهلی‌سازی خصوصیات متفاوتی به دست آوردند که به آنها امکان زندگی و بقا در شرایط آب و هوایی متفاوت را می‌دهد (میراندا دلالاما و متیلو، ۲۰۱۰). این سازگاری‌ها، در نتیجه انتخاب‌های مصنوعی توسط پرورش دهندگان برای خصوصیات منحصر به فرد در خز، شیر و گوشت و همچنین انتخاب‌های طبیعی در شرایط متفاوت منجر به بزهای اهلی امروزی شده است. علاوه‌براین، در مقایسه با بزهای وحشی، بزهای اهلی از نظر کاهش اندازه بدن، رنگ بدن، مقاومت بدن و سازگاری با شرایط مختلف با بزهای وحشی تفاوت‌های زیادی دارند (مرادی و همکاران، ۲۰۱۲). انتخاب در طول زمان منجر به ایجاد نشانه‌هایی در سطح ژنوم می‌شود .این ردپاهای ژنومی که در نتیجه انتخاب بر روی ژنوم برجای می‌ماند، نشانه‌های انتخاب نامیده می‌شوند و در جهت شناسایی جایگاه‌هایی که تحت انتخاب بوده‌اند، مورد استفاده قرار می‌گیرند (کریتمن، ۲۰۰۰). با توجه به در دسترس بودن اطلاعات ژنگانی گونه‌های اهلی و وحشی و پیشرفت ابزارهای آماری، در سال‌های اخیر شناسایی این ردپاها در یکسری گونه‌های خاص امکان پذیر شده است (دانگ و همکاران،۲۰۱۳). شناسایی نشانه‌های انتخاب ممکن است منجر به شناسایی مناطق مرتبط با صفات مهم که تحت انتخاب هستند شود. بنابراین هدف از این مطالعه شناسایی نشانه‌های انتخاب در بزهای اهلی و وحشی ایران با استفاده از نشانگر تک نوکلئوتیدی می‌باشد.

## مواد و روش

داده‌های مطالعه حاضر از پروژه NextGENE گرفته شد. نمونه‌گیری از حیوانات وحشی به طور عمده از حیوانات شکار شده و از بافت‌های ذخیره شده در مراکز حفاظت شده در ایران انجام شد. نمونه‌گیری از نژادهای محلی در مناطق دور افتاده به منظور جلوگیری از ورود احتمالی نژادهای صنعتی صورت گرفت (شکل ۱). از مجموع داده‌های مربوط به بز ایرانی موجود در پروژه NextGENE داده‌های توالی‌یابی کل ژنوم مربوط به ۲۰ راس بز اهلی و ۱۸ راس بز وحشی در این مطالعه مورد استفاده قرار گرفت. برای بررسی جهش‌های موجود در کل ژنوم هر حیوان، ابتدا جهش‌های معادل با 50K شرکت ایلومینا با استفاده از نرم‌افزار Plink v1.9 استخراج شدند.




**شکل ۱: توزیع جغرافیایی گونه‌های بز ایرانی. نمونه‌ها از مناطق مختلف ایران جمع آوری شد. دایره‌های سبز و قرمز به ترتیب نشان دهنده بزهای اهلی و بزهای وحشی است.**

برای بررسی کنترل کیفیت داده‌های ژنوتایپی از نرم‌افزار Plink v1.9 استفاده شد. شاخص‌های کنترل کیفیت شامل نرخ خوانش حیوان (mind)، نرخ خوانش SNP (geno)، فراوانی آلل کمیاب (maf) و تعادل هاردی – واینبرگ بودند. حیوانات با نرخ تعیین ژنوتیپ کمتراز ۹۹ درصد، SNPها با نرخ تعیین ژنوتیپ کمتراز ۹۹/۹ درصد، فراوانی آلل کمیاب کمتر از ۵ درصد و انحراف از تعادل هاردی واینبرگ با شدت $P < 10^{-6}$ از داده‌ها کنار گذاشته شدند. همچنین برای اینکه یک دیدگاه کلی در زمینه ساختار جمعیتی وجود داشته باشد، نمونه‌ها مورد تجزیه‌ی مؤلفه‌های اصلی (PCA) که بر پایه ماتریس روابط خویشاوندی ژنگانی است، قرار گرفتند. از PCA برای شناسایی دام‌هایی که خارج از گروه‌های نژادی خود قرار دارند استفاده می‌شود. این آنالیز در محیط نرم‌افزار R i386 3.6.2 انجام شد. نشانه‌های انتخاب را می‌توان با مقایسه تفاوت بین جمعیتی در نژادهای مختلف با روش‌های آماری شناسایی کرد. شاخص تثبیت رایت یکی از اصلی‌ترین آزمون‌های آماری مبتنی بر تمایز جمعیتی است که برای شناسایی جایگاه‌های تحت انتخاب مثبت که اهمیت بالایی برای صفات اقتصادی دارند مورد استفاده قرار می‌گیرد (اکی و همکاران، ۲۰۰۲).

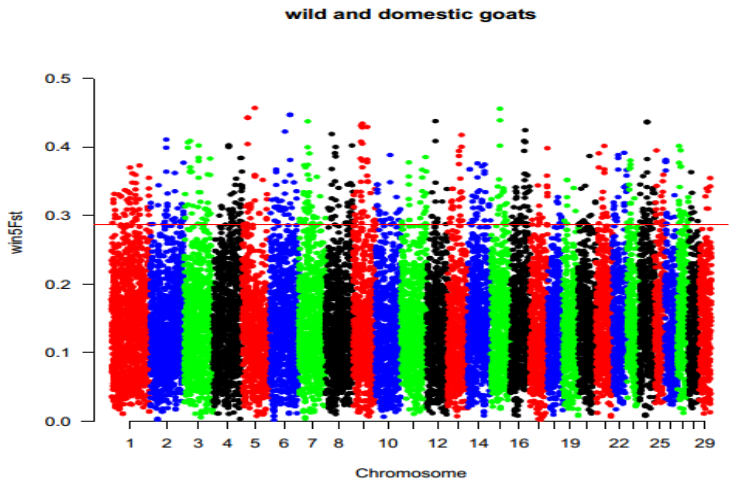
$$F_{ST} = \frac{\sigma_S^2}{\sigma_T^2} = \frac{\sigma_S^2}{\bar{\rho}\,(1-\bar{\rho})}$$

در این رابطه $\sigma_S^2$، $\sigma_T^2$ و $\bar{\rho}$ به ترتیب واریانس فرآوانی آللی بین جمعیت‌های متفاوت، واریانس فرآوانی آللی در کل جمعیت و میانگین فرآوانی آللی در کل جمعیت می‌باشند. آماره $F_{ST}$ در هر جایگاه با استفاده از دو روش رایت و نااریب تتا (ویر و کوکرهام، ۱۹۸۴) در محیط نرم‌افزار R i386 3.6.2 به کمک پکیج FSThet محاسبه شد. برتری روش $F_{ST}$ نااریب تتا نسبت به $F_{ST}$ رایت در نظر گرفتن ناهمگنی اندازه جمعیت‌ها است (بوهمانوا و همکاران، ۲۰۱۰). $F_{ST}$ میزان تفاوت بین دو جمعیت را در دامنه ۰ (بدون تفاوت) و ۱ (کاملا متفاوت) در هر جایگاه ژنی نشان می‌دهد. علاوه براین با توجه به روش تتا که یک روش نااریب است احتمال دارد ارزش‌های منفی نیز به دست آید (اکی و همکاران، ۲۰۰۲). در این مطالعه برای شناسایی بهتر نواحی تحت انتخاب به جای استفاده از ارزش $F_{ST}$ هر SNP از میانگین ارزش $F_{ST}$ هر 5SNP مجاور (Win5) استفاده شد. نشانگرهایی که ارزش $F_{ST}$ آن‌ها در صدک ۹۵ بالایی بودند به عنوان مناطق متأثر از انتخاب معرفی شدند (وی و همکاران، ۲۰۱۶). ژن‌ها با استفاده از SNPهایی که در بازه‌ی ۵٪ بالای $F_{ST}$ و XP-EHH قرار داشتند، توسط نرم‌افزار Plink v1.9 با استفاده از gene list شرکت ایلومینا شناسایی شدند. از ClueGo v2.5.8 برای شناسایی مسیرهای قابل توجه KEGG استفاده شد. این برنامه یک افزونه Cytoscape است و به طور قابل توجهی تفسیرهای بیولوژیکی ژن ها را توسعه می دهد. مسیرهایی که میزان P_Value آن‌ها کمتر از ۰/۰۵ بود به عنوان مسیرهای تحت انتخاب معرفی شدند. برای تجزیه و تحلیل چرخه‌های بیولوژیکی، از پایگاه‌های اطلاعات ژنی آنلاین KEGG استفاده شد.

## نتایج و بحث

پس از استخراج جهش‌های معادل با 50K شرکت ایلومینا از فایل‌های توالی یابی کل ژنوم، مرحله کنترل کیفیت داده‌ها با استفاده از نرم‌افزار Plink v1.9 انجام شد. کنترل کیفیت داده‌ها نشان داد که در نهایت ۴۳۱۷۰ و ۲۶۳۶۵ نشانگر SNPمربوط به ۲ گونه (۲۰ بز اهلی، ۱۸ بز وحشی) برای تجزیه وتحلیل‌های بعدی انتخاب شدند. هدف از تجزیه داده‌ها به مولفه‌های اصلی توجیه واریانس موجود در بین داده‌ها است. PCA نشان داد که مولفه PC1 باعث جدا شدن گونه‌های اهلی و وحشی از یکدیگر و مولفه PC2 باعث قرار گرفتن افراد در گونه‌های نژادی خود شد. به عبارت دیگر نژادهای اهلی و وحشی بز در گروه‌های جدا از هم قرار گرفتند. به منظور پیدا کردن نواحی تحت انتخاب از روش‌های مختلفی استفاده شد. همبستگی بین $F_{ST}$ رایت و تتا در گونه بز نشان دهنده همبستگی بالا (۰/۹۹۹) بین این دو آماره در جمعیت اهلی و وحشی است. با توجه به اینکه روش $F_{ST}$ نااریب تتا نسبت به $F_{ST}$ رایت ناهمگنی اندازه جمعیت‌ها را در نظر می‌گیرد، از همبستگی بالا این دو آماره در این مطالعه می‌توان نتیجه گرفت که جمعیت‌های مورد مقایسه از نظر اندازه جمعیت تفاوت چندانی با هم ندارند و $F_{ST}$ محاسبه شده از دقت خوبی برخوردار است. در این مطالعه، برای شناسایی مناطق متأثر از انتخاب مثبت در کل ژنوم برای بزهای اهلی و وحشی آماره $F_{ST}$ در محیط نرم‌افزار R x64 4.0.4 محاسبه شد. توزیع ژنومیک $F_{ST}$ به روش win5 برای تمام SNPها بر روی ژنوم انجام شد. از SNPهایی که در بازه‌ی ۵٪ بالایی $F_{ST}$ قرار داشتند، به منظور شناسایی ژن‌ها استفاده شد. نگاره منهتن بین هر دو جمعیت اهلی و وحشی بز ترسیم شد (شکل ۲). بخش‌هایی از ژنوم که دارای $F_{ST}$ بالایی هستند نشان‌دهنده تمایز بین دو گونه اهلی و وحشی می‌باشند که بر اثر انتخاب‌های طبیعی و مصنوعی در طی نسل‌های قبل به وجود آمده است. نتایج $F_{ST}$ نشان داد در گونه بز ۴۸۰ منطقه ژنومی روی ۲۹ کروموزوم بین دو گونه اهلی و وحشی بز شناسایی شد. مناطق و ژن‌های شناسایی شده با مسیرهای تنظیم اتوفاژی، تجزیه لیزین، چسبندگی سلولی، مسیر سیگنال دهی فسفولیپاز D، فعال‌سازی پلاکت، سیناپس گلوتاماتیک، سیناپس سروتونرژیک، مسیر سیگنال دهی گنادوتروپین و مسیر سیگنال دهی برای تمایز مناسب سلول‌های جنینی در ارتباط بودند (جدول ۱). نتایج $F_{ST}$ نشان داد که برخی از ژن‌ها از جمله MYCBP2، EDIL3، JARID2، RABGAP1L، ASTN2، CDKAL1، GLIS3، MAP6 و ROCK2 که در مناطق تحت انتخاب شناسایی شده‌اند، در سایر مطالعات نشانه‌های انتخاب که با روش $F_{ST}$ روی بز انجام شده است نیز گزارش شده است. در مطالعه‌ای در سال ۲۰۱۹ با استفاده از توالی ژنوم سه نژاد بز چینی و پنج نژاد بز دیگر انجام شد. نتایج نشان داد که ژن‌های MYCBP2، EDIL3، JARID2، RABGAP1L و ASTN2 بعنوان یک ژن کاندید در مناطق تحت انتخاب شناسایی شدند (گوو و همکاران، ۲۰۱۹). علاوه‌براین در یک مطالعه دیگر که از طریق تجزیه و تحلیل

داده‌های کل ژنوم ۴۴ بز بومی انجام شد، نتایج $F_{ST}$ نشان داد که ASTN2 و CDKAL1 به عنوان ژن‌های کاندید در مناطق تحت انتخاب معرفی گردیدند (بریولای و همکاران، ۲۰۱۹). همچنین در سال ۲۰۱۸ در مطالعه‌ای که روی نژادهای بز اوگاندا انجام گرفت، نتایج $F_{ST}$ نشان داد که دو ژن MAP6 و ROCK2 در مناطق تحت انتخاب شناسایی شده اند (اونزیما و همکاران، ۲۰۱۸). در مطالعه دیگری که برای شناسایی مناطق تحت انتخاب در هشت نژاد بز اهلی چینی انجام شد، ژن GLIS3 به عنوان ژن کاندید در مناطق انتخابی گزارش شد. این ژن مربوط به رنگ بدن در بز سیاه تایهانگ چینی است (ونگ و همکاران، ۲۰۱۶). مسیر سیگنال دهی برای تمایز مناسب سلول‌های جنینی نقش‌های بیشماری در کنترل تکثیر سلولی، الگوی بافت، نگهداری و توسعه سلول‌های بنیادی دارد (ژاویر و همکاران، ۲۰۱۶). طبق نتایج این دو ژن GLI2 و CSNK1G1 در مسیر سیگنال دهی برای تمایز مناسب سلول‌های جنینی شناسایی شدند. این مسیر در مطالعه ای که بر روی تشخیص ژن‌های متفاوت بیان شده در طول سیکل فولیکول مو سالانه در بز انجام شد، نیز گزارش شده است (ژانگ و همکاران، ۲۰۲۰). مولکول‌های CAM پروتئین هایی هستند که روی سطح سلول بیان می‌شوند و نقش مهمی در طیف وسیعی از فرآیندهای بیولوژیکی شامل هموستاز، پاسخ ایمنی، التهاب، جنین زایی و توسعه بافت عصبی ایفا می‌کنند (بارکلی، ۲۰۰۳).




**شکل ۲: توزیع ارزشهای WIN5 در سطح ژنوم بز اهلی در مقایسه با بز وحشی: موقعیت ژنومی SNPها بر روی محور X و ارزش $F_{ST}$ آنها بر روی محور Y نمایش داده شده است. خط ترسیم شده نشان‌دهنده آستانه معنی-داری ۹۵ است.**

**جدول ۱: چرخه‌های شناسایی شده با ژن‌های مرتبط با آماره $F_{ST}$ در بز**

| Pathway Name | P-Value | Gene Name |
|---|---|---|
| Glutamatergic synapse | 0.001 | *GRIK2, GRM4, GRM8, ITPR1, PLCB1, PLD1, PRKCA* |
| Cell adhesion molecules (CAMs) | 0.002 | *CD80, CDH4, CLDN11, NCAM2, NRXN1, NRXN3, NTNG1, SDC1* |
| Regulation of autophagy | 0.014 | *ATG5, ATG7, PIK3C3* |
| Phospholipase D signaling pathway | 0.014 | *GRM4, GRM8, PLCB1, PLD1, PRKCA, SYK* |
| Serotonergic synapse | 0.021 | *GABRB2, HTR4, ITPR1, PLCB1, PRKCA* |
| Platelet activation | 0.025 | *ITPR1, PLCB1, PRKCI, ROCK2, SYK* |
| Hedgehog signaling pathway | 0.027 | *CSNK1G1, EVC2, GLI2* |

| | | |
|---|---|---|
| GnRH signaling pathway | 0.039 | *ITPR1, PLCB1, PLD1, PRKCA* |
| Lysine degradation | 0.044 | *ALDH2, KMT2E, NSD1* |

## نتیجه گیری

**نتایج حاصل از این مطالعه با مشخص نمودن ژن‌های کاندیدا احتمالی مرتبط با صفات اقتصادی و مکان-های ژنومی که در اثر انتخاب تغییر پیدا کرده‌اند، می‌تواند در برنامه‌های اصلاح نژادی بز در کشور مورد استفاده قرار گیرد. البته با توجه به کامل نبودن اطلاعات مربوط به حاشیه نویسی عملکردی ژن‌ها در گونه بز، تراکم پایین SNP و کوچک بودن جمعیت مورد بررسی، مطالعات گسترده‌تر با تعداد نمونه‌های بیشتر، درک بهتری از ژن‌های کاندیدا برای صفات مهم اقتصادی در گونه‌های اهلی و وحشی ایجاد خواهد نمود. افزایش اطلاعات در این زمینه مطمئنا می‌تواند تاثیر مناسبی بر طراحی برنامه‌های اصلاح نژاد در بزهای کشور بگذارد.**

## منابع

Boyazoglu, J., Hatziminaoglou, I., & Morand-Fehr, P. J. S. R. R. (2005). The role of the goat in society: Past, present and perspectives for the future. *60*(1-2), 13-23.

Mignon-Grasteau, S., Boissy, A., Bouix, J., Faure, J.-M., Fisher, A. D., Hinch, G. N., . . . Prunet, P. J. L. P. S. (2005). Genetics of adaptation and domestication in livestock. *93*(1), 3-14.

Miranda-de La Lama, G., & Mattiello, S. J. S. R. R. (2010). The importance of social behaviour for goat welfare in livestock farming. *90*(1-3), 1-10.

Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G., & McEwan, J. C. J. B. g. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. 13(1), 1-15.

Kreitman, M. J. A. r. o. g., & genetics, h. (2000). Methods to detect selection in populations with applications to the human. *1*(1), 539-559.

Dong, Y., Xie, M., Jiang, Y., Xiao, N., Du, X., Zhang, W., . . . Liang, J. J. N. b. (2013). Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). *31*(2), 135-141.

Akey, J. M., Zhang, G., Zhang, K., Jin, L., & Shriver, M. D. J. G. r. (2002). Interrogating a high-density SNP map for signatures of natural selection. *12*(12), 1805-1814.

Weir, B. J. E. F.-s. f. t. a. o. p. s. E. et Cockerham, CC,(1984). *38*, 1358-1370.

Bohmanova, J., Sargolzaei, M., & Schenkel, F. S. J. B. g. (2010). Characteristics of linkage disequilibrium in North American Holsteins. *11*(1), 1-11.

Wei, C., Wang, H., Liu, G., Zhao, F., Kijas, J. W., Ma, Y., . . . Wu, M. J. S. r. (2016). Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *6*(1), 1-11.

Guo, J., Zhong, J., Li, L., Zhong, T., Wang, L., Song, T., & Zhang, H. J. G. S. E. (2019). Comparative genome analyses reveal the unique genetic composition and selection signals underlying the phenotypic characteristics of three Chinese domestic goat breeds. 51(1), 1-18.

Berihulay, H., Li, Y., Gebrekidan, B., Gebreselassie, G., Liu, X., Jiang, L., & Ma, Y. J. F. i. g. (2019). Whole Genome Resequencing Reveals Selection Signatures Associated With Important Traits in Ethiopian Indigenous Goat Populations. 10, 1190.

Onzima, R., Upadhyay, M., Mukiibi, R., Kanis, E., Groenen, M., & Crooijmans, R. J. A. g. (2018). Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. 49(1), 59-70.

Wang, X., Liu, J., Zhou, G., Guo, J., Yan, H., Niu, Y., . . . Lan, X. J. S. r. (2016). Whole-genome sequencing of eight goat populations for the detection of selection signatures underlying production and adaptive traits. 6(1), 1-10.

Xavier, G. M., Seppala, M., Barrell, W., Birjandi, A. A., Geoghegan, F., & Cobourne, M. T. (2016). Hedgehog receptor function during craniofacial development. Developmental biology, 415(2), 198-215.

Zhang, Y., Wu, K., Wang, L., Wang, Z., Han, W., Chen, D., . . . Liu, Z. J. G. (2020). Comparative study on seasonal hair follicle cycling by analysis of the transcriptomes from cashmere and milk goats. 112(1), 332-345.

Barclay, A. N. (2003, August). Membrane proteins with immunoglobulin-like domains—a master superfamily of interaction molecules. In Seminars in immunology (Vol. 15, No. 4, pp. 215-223). Academic Press.