

مروری بر مفاهیم، روش‌ها و کاربرد نشانه‌های انتخاب در ژنوم دام‌ها

محمد مهدی کثیریان

دکترای ژنتیک و اصلاح دام، مرکز آموزش کشاورزی و منابع طبیعی رسول اکرم (ص) دامغان، ایران

چکیده

جمعیت دام‌ها از لحاظ عملکرد و بهره‌وری به‌طور مداوم در طی قرن‌ها از طریق پرورش انتخابی بهبود یافته است. استراتژی انتخاب، ردپایی در ژنوم می‌گذارد که به‌عنوان نشانه‌های انتخاب شناخته می‌شوند. با توجه به پیشرفت در فن‌آوری با بهره‌وری بالا در طول سال‌های اخیر شناسایی نشانه‌های انتخاب در ژنوم گسترده‌تر شده است. چنین مطالعاتی بینش‌های فراوانی مرتبط با اهلی سازی و فرآیندهای تکاملی ارائه می‌دهد که نتیجه آن منجر به تعداد زیادی نژاد دام که می‌توانند در محیط‌های و سیستم‌های تولیدی مختلف پرورش یابند. علاوه بر این، شناسایی ژن‌های تحت انتخاب مرتبط با صفات اقتصادی در جمعیت دام اهلی را تسهیل می‌نماید. روش‌های آماری مختلفی برای شناسایی نشانه‌ها انتخاب براساس توزیع فراوانی آللی، عدم تعادل پیوستگی (LD)، کاهش تنوع جایگاهی و تمایز جمعیت ایجاد شده است. هدف از این بررسی، مرور جامع بر مفهوم کلی، روش‌های مختلف شناسایی نشانه‌های انتخاب و ابزارهای بیوانفورماتیکی که در حال حاضر برای شناسایی نشانه‌های انتخاب در دسترس است و به‌طور خلاصه نتایج مطالعات اخیر مرتبط با نشانه‌های انتخاب که در گونه‌های مختلف دام انجام شده است.

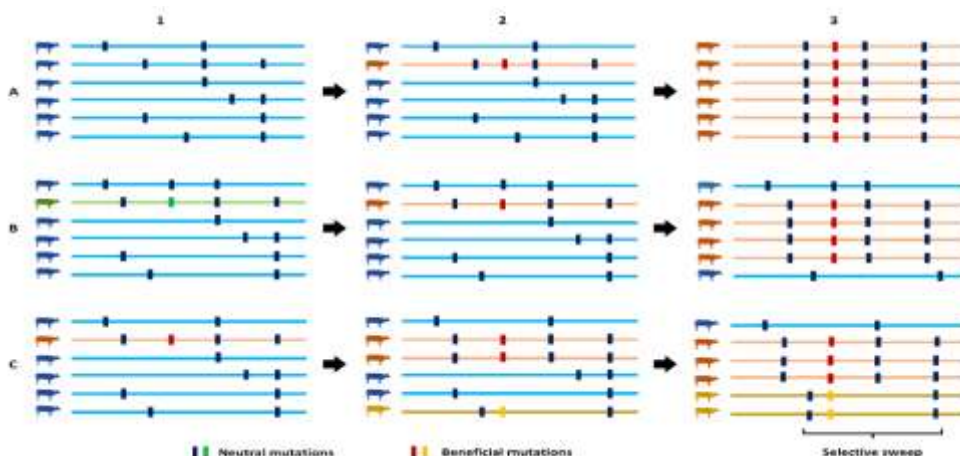
واژه های کلیدی: انتخاب مثبت، بیوانفورماتیک، ژنومیک، دام‌های اهلی، نشانه انتخاب

۱- مقدمه

اهلی شدن حیوانات نقش بسزایی در رشد جمعیت و فرهنگی انسان ها داشته است. اهلی شدن شامل فرآیندهای تکاملی مثل رانش ژنتیکی و هم خونی (غیر قابل کنترل)، انتخاب طبیعی (تا حدی قابل کنترل) و انتخاب مصنوعی (قابل کنترل) می باشد (میگنون و همکاران، ۲۰۰۵). این فرآیندهای تکاملی و جمعیتی منجر به گسترش طیف متنوعی از نژاد های دام اهلی با شرایط محیطی متنوع سازگار شده اند (گرونیولد و همکاران، ۲۰۱۰). در دوران باستان انتخاب ناخودگاه بدون اینکه به فکر تغییر نژاد باشند، بیشتر رایج بود. بعداً پرورش دهندگان مجرب، انتخاب آگاهانه یا علمی جهت دار برای یک هدف مشخص را دنبال می کردند. این استراتژی انتخاب باعث ایجاد تغییراتی در مناطق ژنومی خاص که ویژگی های خاصی از یک نژاد مانند مورفولوژی، ساختار بدن، عملکرد تولید، تولید مثل، رفتار، سازگاری با شرایط محیطی متفاوت و مقاومت در مقابل بیماری را کنترل می کند. الگو های ژنتیکی یا رد پای منحصر به فرد در نواحی ژنومیک ناشی از انتخاب را نشانه های انتخاب می نامند (جنسن و همکاران، ۲۰۱۶). کاهش تنوع ژنتیکی در بالا و پایین دست جهش مفید تحت انتخاب، تثبیت سریع آن جهش پس از چندین نسل ایجاد می شود. پدیده ای که بوسیله آن باعث افزایش فراوانی جهش جدید مفید در جمعیت و کاهش تنوع در جایگاه های خنثی می گردد را تحت عنوان جاروب انتخاب یا انتقال همراه می نامند.

انتخاب طبیعی یکی از نیروهای محرکه اصلی برای شکل دادن الگوی تنوع ژنتیکی تا در گونه ای، شانس بقاء و سازگاری تولید مثلی را فراهم نماید (کثیریان و همکاران، ۲۰۲۲). علاوه بر انتخاب طبیعی، داروین انتخاب مصنوعی را به دو دسته انتخاب ناآگاهانه و آگاهانه (انتخاب علمی) طبقه بندی کرد (گوری و همکاران، ۲۰۰۹).

جاروب انتخاب بسته به ناحیه، نوع و فراوانی جهش ممکن است سخت یا نرم و کامل یا جزئی باشد. در جاروب انتخاب سخت یا جاروب کلاسیک، فراوانی یک جهش نادر مفید به سرعت در زمان انتخاب افزایش می یابد و در مرحله نهایی، کل جمعیت برای آن جهش مفید ثابت می شوند (شکل-۱). بدین ترتیب با افزایش فراوانی بالا و پایین دست، نواحی هموزیگوتی طولانی و سطح بالا عدم تعادل پیوستگی (LD) تنوع ژنتیکی را در جمعیت کاهش می دهد (پریچارد و همکاران، ۲۰۱۰). در این نوع تمام آلل های مفید در یک جمعیت به یک منشاء جهش مشترک باز می گردد. در مقابل جاروب نرم دو نوع می باشد: جاروب تک منشاء و چند منشاء (هر میشنو همکاران، ۲۰۱۷). در جاروب نرم نوع اول، انتخاب روی تنوع ژنتیکی ثابت عمل می کند که قبلاً خنثی یا مضر بوده اما به دلیل تغییرات شرایط ژنتیکی و محیطی سازگار شده است (هرمیشن و همکاران، ۲۰۰۵). افزایش آلل سودمند، فراوانی را به سمت ثابت شدن و تغییرات ژنتیکی مختلفی را برجا می گذارد. در جاروب نرم نوع دوم با چند منشاء بیشتر در جمعیت های بزرگی که فراوانی جهش بالا باشد اتفاق می افتد.



شکل ۱- شرح نمایی از انواع جاروب انتخاب. A: جاروب انتخاب کلاسیک یا سخت. B: جاروب انتخاب نرم ناشی از

تنوع ژنتیکی. C: جاروب نرم چند ناحیه ای (سروانان و همکاران، ۲۰۲۰)

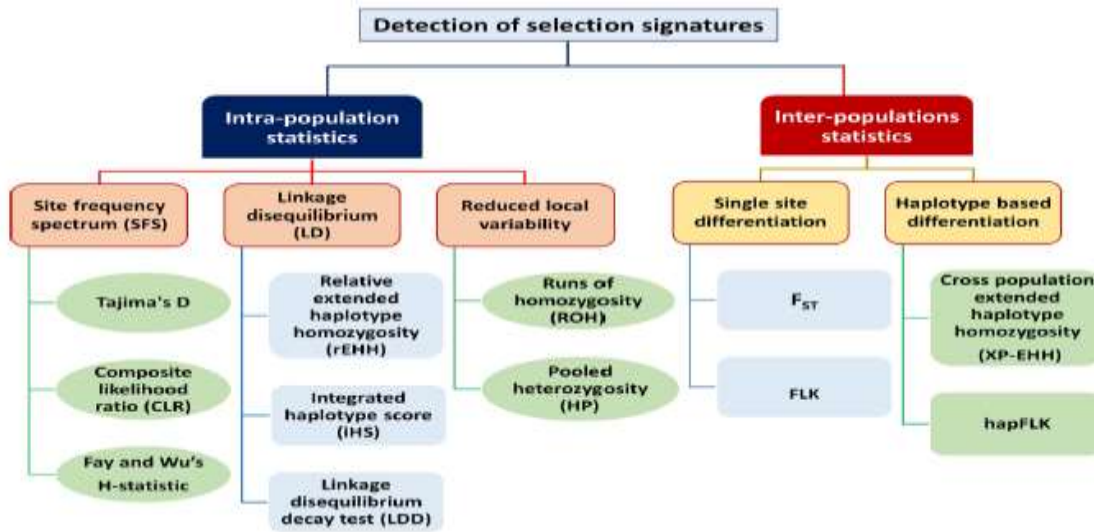
جهش‌ها به‌طور همزمان با افزایش فراوانی بالا می‌روند اما هیچ‌یک از جهش‌ها نمی‌توانند فیکس شوند (هرمیشن و همکاران، ۲۰۰۵). در اینجا همه آلل‌های مفید معمولاً با فراوانی‌های متوسط در یک جمعیت دیده می‌شوند که به منشاء جهش‌های مختلف بر می‌گردد. شناسایی نشانه‌های انتخاب در جاروب نرم کمتر مشخص می‌شود به دلیل اینکه تنوع ژنتیکی جمعیت به شدت کاسته نمی‌شود (پریچارد و همکاران، ۲۰۱۰). از این رو شناسایی جاروب نرم مشکل‌تر از شناسایی جاروب سخت است و احتمال تشخیص آن بستگی به اندازه نمونه دارد. جاروب انتخاب می‌تواند یک جاروب کامل (ثبیت کامل یک آلل مورد علاقه) و یک جاروب جزئی (آلل‌های مورد علاقه با هنوز با جایگاه‌های خنثی در جمعیت جدا می‌شوند) باشد (هرمیشن و همکاران، ۲۰۱۷). طبق تئوری خنثی در مبحث تکامل مولکول بیشتر تنوع ژنتیکی ناشی از جابجایی‌های خنثی با رانش ژنتیکی ارتباط دارد تا به تکامل سازگاری داروین (کیمورا و همکاران، ۱۹۶۸). این پیشنهاد بحث‌های زیادی بین انتخاب‌گرها و طبیعت دانان را دنبال داشته است. چندین مقاله در سال‌های اخیر در ارتباط با استفاده از روش‌های آمار مدرن که تعدادی از آن‌ها برای تمایز خنثی تغییرات انطباقی در ژنوم و تعدادی تکامل سازگار منتشر شده است (کرن و همکاران، ۲۰۱۸ و گریت و همکاران، ۲۰۱۷). با این حال فرضیات و مدل‌های رانش ژنتیکی بر مبنای تئوری خنثی بعنوان چهارچوب کلی برای توسعه روش‌های آماری جهت تشخیص جاروب انتخاب کمک می‌کند (کرن و همکاران، ۲۰۱۸). تحت مدل رانش، یک آلل مشترک بایستی هاپلوتایپ کوتاهی داشته باشد، اما اگر یک آلل مشترک دارای هاپلوتایپ طولانی داشته باشد بعنوان یک سیگنال خوب برای انتخاب در نظر گرفته می‌شود.

شناسایی نشانه‌های انتخاب در حال حاضر جزء علایق اصلی متخصصین ژنتیک حیوانی است. متخصصین ژنتیک حیوانی از نشانه‌های انتخاب برای شناسایی ژن‌ها و جهش‌های مفید که از مزیت انتخابی خاص برخوردار است در یک جمعیت اهلی خاص استفاده می‌کنند (کثیریان و همکاران، ۲۰۲۲). نشانه‌های انتخاب ممکن است تاریخچه تکاملی از توسعه نژادهای اهلی مختلف را فراهم سازد (اتو و همکاران، ۲۰۰۰). فناوری تعیین توالی (NGS)، آرایه‌های SNP با چگالی بالا، ابزارهای پیشرفته آماری و بیوانفورماتیکی توانایی‌های متخصصین ژنتیک حیوانی را در شناسایی مناطق ژنومی تحت انتخاب در گونه‌های مختلف دام را بهبود بخشیده است. رویکرد‌های مختلفی برای شناسایی نشانه‌های انتخاب وجود دارد که این روش‌ها شامل عدم تعادل پیوستگی (LD)، فراوانی آللی، کاهش تنوع جایگاه‌ها و ویژگی‌های هاپلوتایپ می‌باشد (قنبری و همکاران، ۲۰۱۴).

۲- روش‌های شناسایی نشانه‌های انتخاب

در ابتدای دهه ی ۱۹۹۰ چندین آزمون توسعه داده شده بود بعنوان مثال: آزمون d_n/d_s ، آزمون HKA، آزمون MKT که بر اساس مقایسه الگوی جایگزینی ژن و واگرایی انتخاب در گونه‌ها بود (نیلسون و همکاران، ۲۰۰۵). در آزمون d_n/d_s : نسبت میزان جابجایی در جایگاه‌های مشابه (d_s) (یعنی جایگاه‌های خنثی) به میزان جابجایی در جایگاه‌های غیر مشابه (d_n) (یعنی مکان‌های تحت انتخاب) (گلدمن و همکاران، ۱۹۹۴). در صورت انتخاب مثبت نسبت بیشتر از یک و در انتخاب منفی کمتر از یک خواهد بود. اگر $d_n/d_s=1$ باشد جایگاه‌ها تحت تاثیر تکامل خنثی قرار می‌گیرند (گلدمن و همکاران، ۱۹۹۴). معمولاً از این روش برای طبقه بندی سطوح انتخابی اتفاق افتاده در گذشته و نمایش الگوهای تکاملی حاصل از نتیجه انتخاب واگرا بین گونه‌ها استفاده می‌شود (مک ایچران و همکاران، ۲۰۰۹). از آزمون استاندارد، بعنوان مثال آزمون MKT در ابتدا برای شناسایی نشانه‌های انتخاب و سازگاری مولکول‌های تکاملی بین گونه‌ها بر اساس انتخاب واگرا نوکلئوتیدی استفاده می‌شد (مک دونالد و همکاران، ۱۹۹۱). اما بیشتر بعنوان MKT چند لوکوسی پیشرفته برای مقایسه تنوع تک نوکلئوتیدی درون گونه‌ای در مکان‌های مختلف DNA ایجاد شد (اگی و همکاران، ۲۰۰۸). بعداً مطالعات نشانه‌های انتخاب فقط به روش‌های شناسایی سطوح انتخاب تکاملی خرد در سطح جمعیت (درون گونه‌ها) بستگی داشت. در ژنتیک جمعیت سنتی فقط مکان‌های خاصی از ژنوم با یک یا چند مارکر به صورت آزمایشگاهی یا از مدل‌های آماری خاص برای مقایسات استفاده می‌شد. پیشرفت در توالی‌یابی با بالا و تکنولوژی فناوری SNP منجر به بررسی سطح وسیعی از ژنوم برای شناسایی نشانه‌های انتخاب جمعیت‌های درون یک گونه شده است (ماتوکومالی و همکاران، ۲۰۰۹).

روش‌های آماری مختلفی برای شناسایی نشانه‌های انتخاب با استفاده از توالی DNA یا داده‌های ژنوتایپینگ SNP در جمعیت دام‌های اهلی ایجاد شده است (قنبری و همکاران، ۲۰۱۴). این روش‌ها را به‌طور کلی می‌توان به دو گروه تقسیم کرد: آماره درون جمعیتی و آماره بین جمعیتی (شکل-۲). آماره درون جمعیتی با مقایسه داده‌های ژنوتیپی در جمعیت‌ها، نشانه‌های انتخاب را جستجو می‌کند. این گروه شامل سه روش اصلی: فراوانی آللی، عدم تعادل پیوستگی (LD) و کاهش تنوع ژنتیکی است (ویگانند و همکاران، ۲۰۱۸). آماره بین جمعیتی عمدتاً به تمایز ناشی از فراوانی آللی اختصاصی مکان‌های خاص بین جمعیت را بررسی می‌کند (ژو و همکاران، ۲۰۱۵). این روش بر اساس single site و تمایز بین هاپلوتایپ‌ها تقسیم‌بندی می‌شوند.



شکل-۲: روش‌های مختلف شناسایی نشانه‌های انتخاب در جمعیت دام‌های اهلی (سروانان و همکاران، ۲۰۲۰)

۱-۲- آزمون‌های مبتنی بر توزیع فراوانی آللی

جاروب انتخاب باعث افزایش فراوانی آللی بالا و پایین و کاهش فراوانی آللی متوسط در جمعیت می‌گردد (اچازو و همکاران، ۲۰۰۹). روش SFS، یکی از روش‌های مبتنی بر توزیع فراوانی آللی است (رونن و همکاران، ۲۰۱۳). از دیگر آزمون‌های مبتنی بر توزیع فراوانی آللی را می‌توان آزمون تاجیما D (تاجیما، ۱۹۸۹)، آزمون فی و H (فی و همکاران، ۲۰۰۰)، و آزمون نسبت درست‌نمایی مرکب (لیندسی و همکاران، ۱۹۹۸) نام برد. آزمون تاجیما D، تفاوت بین میانگین تعداد تفاوت در نوکلئوتیدهای $(\theta\pi)$ و تعداد جایگاههای تفکیکی (θ_s) برآورد شده حاصل از داده‌های چند شکلی برآورد می‌گردد (کارلسون و همکاران، ۲۰۰۵). در صورت انتخاب مثبت یا جاروب انتخاب، جهش‌ها نادر باعث کاهش هتروزیگوتی و ارزش D منفی ($D < 0$) می‌شود. در مقابل، انتخاب متعادل منجر به افزایش فراوانی آلل‌های هدف و نهایتاً منجر به مثبت شدن ارزش D می‌شود.

آماره فی و H (فی و همکاران، ۲۰۰۰) براساس فراوانی آلل‌های اجدادی و مشتق شده انجام می‌شود با فرض اینکه آلل‌های اجدادی شناخته شده باشند. از این آماره برای شناسایی نشانه‌های انتخاب مثبت برای آلل‌های غیر اجدادی (مشتق شده) در یک فراوانی متوسط تا بالا استفاده می‌شود. این آماره به‌طور مکرر با آماره تاجیما D می‌تواند آل‌های مشتق شده در یک فراوانی پایین به متوسط را شناسایی کند (کادو و همکاران، ۲۰۱۴). آماره CLR، این آماره متفاوت از دو آماره بالا از طریق طیف فراوانی آلل‌ها در چند جایگاه و همچنین نرخ نوترکیبی نشانه‌های انتخاب را برآورد می‌نماید (وی و کیم، ۲۰۱۵). آماره CLR مبتنی بر فرضیه آماری است که مدل خنثی را براساس طیف فراوانی مکانی از پنجره ژنومی را با مدل جاروب انتخاب

مقایسه می کند (چن و همکاران، ۲۰۱۸). آماره CLR به نشانه‌های انتخاب مثبت در چندین جایگاه در یک جمعیت حساس است (ویلیامسون و همکاران، ۲۰۰۷). در حالت کلی این روش‌های مبتنی بر SFT برای جایگاه های SNP در کل ژنوم مناسب نیستند. پس از ظهور تکنولوژی ژنومیک، فاز هاپلوتایپ و مرحله سازی آن ها، روش‌های پیشرفته تر مبتنی بر عدم تعادل پیوستگی (LD) و هموزیگوسیتی هاپلوتایپ ها توسعه داده شده اند.

۳-۲- آزمون های مبتنی بر عدم تعادل پیوستگی^۱ (LD)

جاروب انتخاب، یک هاپلوتایپ فراوانی بالا با عدم تعادل پیوستگی (LD) را ایجاد می کند که در زمان افزایش سریع فراوانی هاپلوتایپ حامل جهش مفید اتفاق می افتد (کثیریان و همکاران، ۲۰۲۲). روش‌های مبتنی بر LD نواحی هموزیگوتی طولانی با هاپلوتایپ های فراوانی بالا ایجاد شده با جاروب انتخاب مورد هدف قرار می دهد. روش‌های مبتنی بر LD جهت شناسایی انواع جاروب انتخاب نرم و جزئی مفید هستند. هموزیگوسیتی هاپلوتایپ بسط یافته^۲ (EHH) مبتنی بر عدم تعادل پیوستگی جهت شناسایی نشانه‌های انتخاب در یک جمعیت را پیشنهاد دادند (ثابتی و همکاران، ۲۰۰۲). پس از شناسایی هاپلوتایپ مرکزی، سن هر هسته مرکزی با فرسایش LD با توجه به فاصله ارزیابی می شود. EHH این احتمال اینکه یک جفت کروموزوم حامل هاپلوتایپ هموزیگوتی مرکزی باشد. سپس هاپلوتایپ هموزیگوسیتی نسبی بسط یافته^۳ (rEHH) با مقایسه ارزش EHH از دو هاپلوتایپ مرکزی برآورد می گردد (ثابتی و همکاران، ۲۰۰۲). از روش هاپلوتایپ هموزیگوسیتی نسبی توسعه یافته برای شناسایی مناطقی که اخیراً مورد انتخاب مثبت و نیاز به آلل های اجدادی ندارند مورد استفاده قرار می گیرد. یک آماره جدیدی از EHH توسط (ویقیت و همکاران، ۲۰۰۶) تحت عنوان نمره هاپلوتایپ بسط یافته (iHS) که با فاصله نوترکیبی در آماره توسعه داده شده است. پس از توسعه تراشه های SNP این روش به طور گسترده ای در شناسایی نشانه‌های انتخاب استفاده شده است. نمره هاپلوتایپ بسط یافته از طریق هاپلوتایپ های اطراف یک SNP نسبت به کل ژنوم اندازه گیری می شود. در این روش هر SNP بعنوان یک SNP مرکزی و مقادیر EHH برای هر SNP مرکزی با توجه به آلل اجدادی و آلل مشتق شده از گروه خارجی استفاده می شود. این آماره، نقطه ای را نشان می دهد که در آن مقادیر ارزش EHH مشتق شده و اجدادی در زیر حد آستانه معین واقع شده اند. در نهایتاً ارزش iHS منفی ($iHS < -2$) نشان دهنده هاپلوتایپ بسطی بیشتری در آلل های مشتق شده در مقایسه با هاپلوتایپ های مرتبط با آلل های اجدادی است. در حالیکه مقادیر مثبت ارزش iHS ($iHS > 2$) نشان می دهد که آلل های اجدادی تحت جاروب انتخاب قرار گرفته اند (ویگانند و همکاران، ۲۰۱۸). بر خلاف روش‌های SFS یا روش‌های مبتنی بر توزیع فراوانی، روش‌های مبتنی بر عدم تعادل پیوستگی نیاز به فاز هاپلوتایپی، نقشه نوترکیبی، موقعیت ژنومی و اطلاعات اجدادی و مشتق شده هر SNP دارند. این روش برای شناسایی نشانه‌های انتخاب در فراوانی ها آلل های متوسط در جمعیت مناسب می باشد. در مقایسه با روش rEHH، روش iHS کمتر تحت تاثیر عوامل جمعیتی قرار می گیرد؛ بنابراین احتمال وقوع نتایج مثبت کاذب کمتر خواهد بود (ویقیت و همکاران، ۲۰۰۶).

۳-۲- آزمون های مبتنی بر کاهش مکانی تنوع ژنتیکی

این روش‌ها نسبت کاهش تنوع ژنتیکی نسبی مناطق از ژنومی به میانگین ژنوم را متمایز می کند. این روش‌ها شامل ROH^۴ (مک کوئیلان و همکاران، ۲۰۰۸) و HP^۵ (روبین و همکاران، ۲۰۱۰) می باشد. روش ROH، طول های پیوسته از ژنوتیپ های هموزیگوسیتی در یک فرد زمانیکه دو تا هاپلوتایپ دارای یک جد مشترک همسان (IBD^۶) را شناسایی می

^۱ - Linkage disequilibrium

^۲ - extended haplotype homozygosity

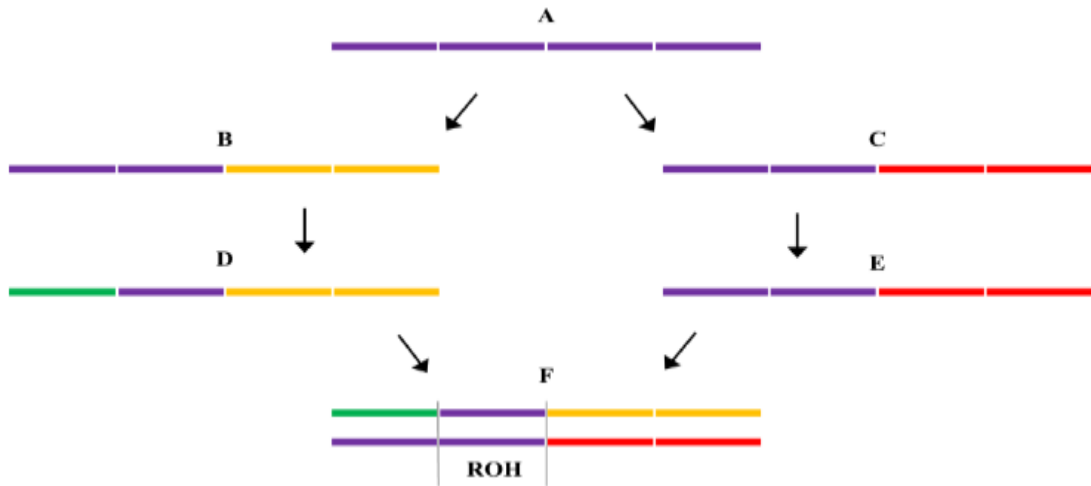
^۳ - relative extended haplotype homozygosity

^۴ - runs of homozygosity

^۵ - pooled heterozygosity

^۶ - identical by descent

کند(ریبيلاتو و همکاران، ۲۰۱۸)(شکل-۳). ROH ها به طور گسترده برای ارزیابی سطوح اصلاح ژنومیک، ساختار جمعیت و پیشینه جمعیت در دام‌های اهلی استفاده می شود (کوریک و همکاران، ۲۰۱۴). بر اساس تئوری انتقال همراه، یک جاروب انتخاب باید نقاطی از مکان های هموزیگوتی باشند که هموزیگوسیتی بالاتری بالاتر از میانگین ژنوم را نشان دهد(آلمیدا و همکاران، ۲۰۱۹). بنابراین از روش ROH می توان برای شناسایی نشانه‌های انتخاب استفاده کرد زیرا افراد تحت انتخاب طولانی مدت باعث هموزیگوسیتی شدن اطراف جایگاه یا جایگاه هدف می‌شوند(فوراتان و همکاران، ۲۰۱۸، ریبيلاتو و همکاران، ۲۰۱۸ و زی و همکاران، ۲۰۱۹).



شکل-۳: شرح نمایی از تشکیل ROH. F: نشان دهنده ROH (رنگ بنفش) است که از جفت شدن بخش های همولوگ حاصل از جد مشترک A تشکیل شده است (ریبيلاتو و همکاران، ۲۰۱۸).

۵-۲ - آزمون های مبتنی بر تمایز جمعیتی تک سایتی

روش های مبتنی بر تمایز جمعیتی تک سایتی شامل FST (رایت و همکاران، ۱۹۴۹) و FLK (بونهومن و همکاران، ۲۰۱۰) می باشد. شاخص تثبیت (FST) (رایت و همکاران، ۱۹۴۹) براساس تفاوت فراوانی آللی بین جمعیت ها محاسبه می شود. ارزش FST از ۰ (بدون تمایز) و ۱ (تفاوت ثابت بین جمعیت ها) است. تفاوت فراوانی آللی بالا بین جمعیت ها در هر مکان مشخص (بعنوان مثال: مقادیر بالا FST) نشان دهنده انتخاب مثبت است در حالیکه مقادیر کم FST نشان دهنده انتخاب منفی است (ژو و همکاران، ۲۰۱۵).

FST رایت مفروض به اندازه بی نهایت جمعیت است. بنابراین ممکن است FST بیش از حد بزرگ برای جمعیت های کوچک رخ دهد. تعداد SNP های بیشتر (>500) باعث افزایش تمایز زیر جمعیت ها حتی در صورت کوچک بودن جمعیت نشان می دهد. علاوه بر روش FST رایت چندین روش برای برآورد آماره FST وجود دارد. FST نی - GST (نی و همکاران، ۱۹۷۳)، FST ویر و کوکرهام (ویر و کوکرهام، ۱۹۸۴)، FST هادسون (هادسون و همکاران، ۱۹۹۲)، FST هولیسنگر (هولیسنگر و همکاران، ۲۰۰۴)، FST مخصوص جمعیت (ویر و هیل، ۲۰۰۲)، FST بر اساس مدل دریافت (نیلچولسون و همکاران، ۲۰۰۲) و FST براساس مدل بی‌زین (گیانولا و همکاران، ۲۰۱۰) وجود دارد. نرم افزار های متعددی برای برآورد آماره های فوق وجود دارد (بهاتیا و همکاران، ۲۰۱۳). یکی از مزیت های FST نسبت به روش های مبتنی بر عدم تعادل پیوستگی (LD) و توزیع فراوانی آللی (SFS)، FST برای یک SNP خاص است و می تواند تنوع ژنتیکی واقعی تحت انتخاب را شناسایی کند. در آماره FST به جای آنالیز هر SNP به طور جداگانه و خاص بهتر است از میانگین چند SNP

تحت عنوان win FST مقایسه شوند. هنگام استفاده از داده های طبقه یا کلاس بندی شده SNP ها، آماره FST می تواند نتایج مثبت و منفی کاذب را تشخیص دهد (فاریلو و همکاران، ۲۰۱۴).
FLK یا TFLK یک آماره بسط داده شده از آماره LK اصلی (لیونتین و کراکانر، ۱۹۷۳) است. از طریق نسبت واریانس های مشاهده شده و مورد انتظار داده ها FST به واریانس تخمینی خنثی مقایسه می شود. FLK فاصله درخت فیلوژنیک را از طریق ماتریس خویشاوندی F تخمین می زند. بنابر این تغییرات ایجاد شده در اندازه های موثر جمعیت در طولانی مدت را با زیر شاخه های مرتبط جمعیت در نظر می گیرد (بون هومی و همکاران، ۲۰۱۰). ماتریس F رانش ژنتیکی مورد انتظار در هر جمعیت و کواریانس مورد انتظار بین آن است. آماره FLK یک آماره پارامتریک قدرتمند است که براحتی نشانه های انتخاب را در بین جمعیت ها با استفاده از یک دسته داده های ژنوتیپی بزرگ و عظیم شناسایی می کند. در مقایسه با روش FST، کارایی روش FLK در شناسایی نشانه های انتخاب خطای نوع ۱ (مثبت کاذب) را کاهش می دهد (بون هومی و همکاران، ۲۰۱۰).

۲-۶- آزمون های مبتنی بر هاپلوتایپ

روش های تمایز مبتنی بر هاپلوتایپ شامل هموزیگوسیتی هاپلوتایپ بسطی^y (XP-EHH) (ثابتی و همکاران، ۲۰۰۷) و hapFLK یا همان فرمت آماری FLK مبتنی بر هاپلوتایپ (فاریلو و همکاران، ۲۰۱۳) می باشد. در این روش ها از اطلاعات هاپلوتایپی در چندین جمعیت برای شناسایی نشانه های انتخاب استفاده می شود. روش XP-EHH یک روش تمایز مبتنی بر هاپلوتایپ که توسط ثابتی و همکاران (۲۰۰۷) توسعه داده شده اند. برای محاسبه XP-EHH بین جمعیت های A و B، ابتدا ارزش iHH برای هر جمعیت به طور جداگانه و سپس با ادغام EHH کل نمونه آماره XP-EHH محاسبه می شود. ارزش XP-EHH جهت دار است و ارزش های مثبت و منفی را نشان می دهد که انتخاب به ترتیب در جهت B و A رخ داده است. فاریلو و همکاران (۲۰۱۳) یک روش FLK مبتنی بر هاپلوتایپ به نام hapFLK را پیشنهاد دادند. آماره hapFLK بر اساس اطلاعات هاپلوتایپی و ساختار جمعیت استوار است که منجر به قدرت بالای این آماره جهت شناسایی نشانه های انتخاب بکار می رود. این آماره بر خلاف آماره FST اندازه موثر جمعیت های مختلف را در بر می گیرد. ماتریس خویشاوندی جمعیت در hapFLK مشابه روش FLK، اما به جای محاسبه فراوانی آلی از فراوانی هاپلوتایپ ها تعیین می شود (فاریلو و همکاران، ۲۰۱۳).

۳- ابزار های بیوانفورماتیکی آنالیز نشانه های انتخاب

شناسایی نشانه های انتخاب با ایجاد داده های ژنوتیپی از ریزآرایه SNP و تکنولوژی توالی یابی بالا شروع شد. این داده های ژنوتیپی خام باید مطابق روش تجزیه و ابزار های بیوانفورماتیکی پردازش شوند زیرا هر یک از آن ها به فرمت خاصی نیاز دارند (کاد زو و همکاران، ۲۰۱۴). اگر چه روش تشخیص نشانه های انتخاب با این ابزار پیچیده نیست اما از مرحله ژنوتیپ تا نشانه های انتخاب می تواند برای محققانی که با ابزار بیوانفورماتیکی مناسب فرمت داده ها آشنا نیستند یک مرحله چالش برانگیز باشد. چندین برنامه نرم افزاری خاصی برای تبدیل داده ها توسعه داده شده اند. PDG Spider (لیسچر و همکاران، ۲۰۱۰) و FcGENE (روش یارا و همکاران، ۲۰۱۴) توسعه یافته اند که می توانند طیف وسیع تری از فرمت های ورودی و خروجی را مدیریت کنند. برای کنترل کیفیت داده ها و فیلتراسیون SNP ها می توان از برنامه های نرم افزاری مانند PLINK، CSNPQC، JMP Genome، SVS و VCF Tools استفاده کرد. در روش PLINK (پورسل و همکاران، ۲۰۰۷) به دلیل سرعت و قابلیت اطمینان، پرکاربردترین منبع باز برای مدیریت داده های SNP بکار می رود. همچنین PLINK می تواند ابزار مناسبی برای شناسایی نشانه های انتخاب از طریق روش های ROH و FST بکار می رود. در کنار PLINK، بسته نرم افزاری R (تیم آرکو، ۲۰۲۰) به طور عمومی برای مدیریت، تجزیه و تحلیل داده ها بکار می رود. جدا از

⁷ - Cross population extended haplotype homozygosity

این ها JMP Genomics (SAS Institute Inc , Cary,Ne,USA) و تغییرات SNP ها (SVS,Golden) (Helix,Bozeman, MT,USA) برجسته ترین ابزارهای نرم افزاری مجاز برای مدیریت داده ها، تجزیه و تحلیل بکار می روند. پس از کنترل کیفیت داده ها برخی از روش های شناسایی نشانه های انتخاب مانند XP-EHH, LHR و iHS نیاز به محاسبه ژنوتیپ از دست رفته و فاز بندی هاپلوتایپ ها دارند. این کار را می توان با استفاده از ابزار های ایمپوت کردن مانند: Beagle, PedImpute, MACH, IMPUTE2, fastPHASE و ALPHAPHASE (نیکولازی و همکاران، ۲۰۱۵) انجام داد. در بین ابزارهای فوق، ابزار Beagle (هووی و همکاران، ۲۰۰۹) و IMPUTE2 (براونینگ و همکاران، ۲۰۱۸) جزء ابزارهای مهم برای ایمپوت و فاز بندی در ژنومیک دامها استفاده می شود. نرم افزارهای متفاوتی برای تشخیص نشانه های انتخاب با استفاده از روش های متفاوتی در دسترس است. انتخاب برنامه های نرم افزاری عمدتاً با روش های آنالیز، نیاز به داده ها و سیستم عامل و سرعت محاسبه تعیین می شود. اکثر نرم افزارهای مورد استفاده در شناسایی نشانه های انتخاب در پلت فرم های Linux و برخی هم در محیط ویندوز مناسب هستند. ابزارهای بیوانفورماتیکی رایج برای تشخیص نشانه های انتخاب در جمعیت دامها در جدول (۱) ذکر شده است.

جدول ۱- ابزارهای بیوانفورماتیکی برای شناسایی نشانه های انتخاب در حیوانات اهلی

ابزار/ نرم افزار	آخرین نسخه ^۱	روش	پلت فرم ^۲	مجوز ^۳	لینک منبع
Arlequin	3.5.2.2 (2015)	Tajima's D	W	F	http://cmpg.unibe.ch/software/arlequin35/
BayeScan	2.1 (2012)	FST (Bayesian approach)	L/M	F	http://cmpg.unibe.ch/software/BayeScan/index.html
cgaTOH	(۲۰۱۳)	ROH	W/L/M	F	http://www.cs.kent.edu/~zhao/TOH/
DetectRUNS R-package	0.9.6 (2019)	ROH	W/L/M	F	https://CRAN.R-project.org/package=detectRUNS
DnaSP	6.12.03 (2019)	Tajima's D, Fay and Wu's H, Fu and Li's test, MKT	W/L	F	http://www.ub.edu/dnasp
Hapbin	(2015)	EHH, iHS, XP-EHH	W/L/M	F	https://github.com/evotools/hapbin
HapFLK	1.4 (2017)	HapFLK, FLK	W/L/M	F	https://forge-dga.jouy.inra.fr/projects/hapflk/files
HGDP Selection	(2009)	iHS, XP-EHH, FST, CLR	W/L/M	F	http://hgdp.uchicago.edu/Software/
HierFstat R-package	0.04-22	Fst	W/L/M	F	https://CRAN.R-project.org/package=hierfstat
OmegaPlus	2.2.2 (2013)	LD based "ω"	L	F	http://www.exelixis-lab.org/software.html
Pcadapt R-package	4.3.1 (2020)	PCAdapt	W/L/M	F	https://CRAN.R-project.org/package=pcadapt
PLINK	1.90 (2015)	ROH, FST	W/L/M	F	http://zzz.bwh.harvard.edu/plink/
PoPoolation	1.2.2 (2011)	Tajima's Pi, Watterson's Theta and Tajima's D	W/L/M	F	https://sourceforge.net/projects/popoolation/
PopGenome	2.7.5 (2020)	Tajima's D, FST (Bayesian approach)	W/L/M	F	https://CRAN.R-project.org/package=PopGenome
Rehh R-package	3.1.0 (2020)	EHH, iHS, Rsb, XP-EHH	W/L/M	F	https://CRAN.R-project.org/package=rehh
Selscan	1.2.0 (2017)	EHH, iHS, XP-EHH, nSL	W/L/M	F	https://github.com/szpiech/selscan
SNP Variation Suite	8.4.0 (2015)	Fst	W/L/M	L	https://www.goldenhelix.com/forms/SNP_Variation
Sweed	3.2.1 (2013)	CLR	L	OS	https://cme.h-its.org/exelixis/web/software/sweed/index.html
Sweep	1.1 (2010)	LRH, EHH	W/L	F	https://software.broadinstitute.org/mpg/sweep/
Sweepfinder2	1.0 (2016)	CLR	L	F	http://degiorgiogrupo.fau.edu/sf2.html
TRES	1.0 (2015)	Wright's Fst	W/L/M	F	http://mlkd.csd.auth.gr/bio/tres
VariScan	2.0.3 (2012)	Tajima's D, Fu and Li's tests	W/L/M	F	http://www.ub.edu/softevol/variscan
VCftools	0.1.13 (2015)	Tajima's D, FST	L/M	F	https://github.com/vcftools/vcftools
XP-CLR	1.1.2 (2019)	XP-CLR	L	F	https://reich.hms.harvard.edu/software
ZANARDI	(2015)	ROH	L/M	OS	https://github.com/bioinformatics-ptp/Zanardi

۱- براساس سال ۲۰۲۰؛ ۲- پلت فرم W: ویندوز، L: لینوکس، M: Mac؛ ۳- مجوز: L: قابلیت خرید، F: رایگان، OS: دسترسی آزاد

۴- مطالعات نشانه‌های انتخاب در دام‌های اهلی

شناسایی نشانه‌های انتخاب به دلیل اهمیت آن اصلاح نژاد دام به شدت مورد مطالعه قرار گرفته است. مطالعات تنوع ژنتیکی فقط درجه یا میزان تمایز بین جمعیت‌ها یا نژادها را بررسی می‌کند اما نشانه‌های انتخاب فراتر از این است و دلیل واقعی تنوع ژنتیکی را پیدا می‌کند. شناسایی نشانه‌های انتخاب برای توصیف منابع ژنتیکی دام و تعیین ژن‌های تاثیرگذار در صفات اقتصادی و اکولوژیکی مهم هستند (سزارانی و همکاران، ۲۰۱۸). این مطالعات به شناسایی جهش‌های مفید دارای یک مزیت انتخابی در یک جمعیت یا نژاد خاص کمک می‌کند. شناسایی نشانه‌های انتخاب برای درک عمیق منشاء جمعیت‌ها و فرآیند‌های ژنتیکی تاثیرگذار تمایز جمعیت‌ها در دام‌های اهلی ضروری است. نشانه‌های انتخاب بینش بهتری از انتخاب مصنوعی و امکان بهبود ژنتیکی بیشتری در دام‌ها فراهم می‌کند. درک اینکه چگونه انتخاب بروی یک جمعیت خاص عمل می‌کند برای اصلاح صفات مهم اقتصادی و برنامه‌های اصلاح نژادی کمک می‌کند.

شناسایی نشانه‌های انتخاب بعنوان یک روش مکمل (کامل کننده) برای مطالعات ژنومی (GWAS) برای ارتباط ژن‌های کاندید تحت انتخاب با فنوتیپ که متعاقباً می‌تواند در انتخاب ژنومی استفاده شود مورد استفاده قرار داد (چن و همکاران، ۲۰۱۶). مطالعات نشانه‌ها انتخاب بهتر است برای صفات که گران، پیچیده و گاهی اوقات مطالعه آن با GWAS غیر ممکن است انجام شود. بعنوان مثال تحمل به تلورانس دما یا سازگاری حیوانات با آب و هوای خاص، مقاومت در برابر بیماری‌ها، پاسخ به استراتژی‌های مختلف اصلاح دام، پیشرفت ژنتیکی انتخاب و... که برای یک ویژگی خاص در یک جمعیت یا نژاد انتخاب شده اند (مایورانو و همکاران، ۲۰۱۸).

۵- چالش‌ها و مشکلات نشانه‌های انتخاب

در بررسی‌های مورد بحث، از استراتژی‌های مختلف پیشرفته برای شناسایی نشانه‌های انتخاب در جمعیت‌های مختلف استفاده کرده‌اند. با این حال محدودیت عمده در این مطالعات وجود نتایج مثبت کاذب و منفی کاذب است. عوامل مختلفی، سطوح دموگرافی (مانند افزایش موثر جمعیت و مهاجرت)، نوع و قدرت انتخاب، عدم تجانس جهش‌ها و نوترکیبی نشانه‌های انتخاب را تغییر می‌دهد (ویگاند و لسی، ۲۰۱۸). در اندازه جمعیت کوچک، تغییرات دموگرافیک مانند تنگه ژنتیکی یا اثرات بنیان‌گذار اغلب باعث کاهش تنوع ناشی از طریق دریافت ژنتیکی که باعث تغییر فراوانی آللی و هاپلوتایپی مشابه نشانه‌های انتخاب می‌شود (نیلسون، ۲۰۰۵). جدا از عوامل دموگرافیک، نشانه‌های انتخاب ممکن است تحت تاثیر نرخ متغیر نوترکیبی در ژنوم قرار بگیرد. مناطق ژنومی با نرخ پایین از نوترکیبی ممکن است از نشانه‌های انتخاب تقلید کنند (پایسر و همکاران، ۲۰۱۵). قدرت و نوع انتخاب بروی نشانه‌های انتخاب تاثیر می‌گذارد. بعنوان مثال در انتخاب منفی، حذف آلل‌های مضر منجر به کاهش تنوع ژنتیکی در مکان‌های مجاور شود و این ممکن است از نشانه‌های انتخاب تقلید کند (ویتی و همکاران، ۲۰۱۳). بنابراین برای جلوگیری از نتایج مثبت کاذب و منفی تمام عوامل مشکل‌ساز را در نظر گرفت. ترکیبی از رویکرد‌های جایگزینی و روش‌های ترکیبی مانند CSS (سیگنال‌های انتخاب کامپوزیت) و DCMS (ترکیبی از همبستگی سیگنال‌های چندگانه) یک روش موثر برای پرداختن انواع نشانه‌های انتخاب است (یورچینکو و همکاران، ۲۰۱۸).

یکی دیگر از مشکلات تکراری محققان در نظر گرفتن ماهیت داده‌های ژنومی در بررسی نشان‌های انتخاب است (ویتی و همکاران، ۲۰۱۳). تا به امروز بیشتر مطالعات نشانه‌های انتخاب از داده‌های SNP بدست آمده از آرایه‌های ژنوتایپینگ برای شناسایی چند شکلی‌ها استفاده می‌شد. اشکال عمده و اساسی کشف نشانگر SNP، آلل‌های با فراوانی پایین را در نظر نمی‌گیرد که در نتیجه باعث حذف چنین آرایه‌هایی می‌شود (لاچنس و همکاران، ۲۰۱۳) که این امر باعث نتایج گمراه‌کننده خواهد شد. استفاده از اندازه جمعیت بزرگ، تعیین توالی یابی کل ژنوم، حداکثر احتمال، روش‌های بی‌زین و مبتنی بر آپلوتیپ ممکن است مشکلات مربوط به استفاده از داده‌های SNP جلوگیری شود (مامولانس و همکاران، ۲۰۱۸). با این حال هزینه توالی یابی کل ژنوم بالا است و تجزیه و تحلیل آماری برای داده‌های عظیم در دام‌های اهلی با چالش مواجه است.

۶- بحث و نتیجه‌گیری

در بررسی مروری یک نمای کلی از مفاهیم، روش‌ها و کاربردهای نشانه‌های انتخاب در دام‌ها اهلی ارائه شد. شناسایی نشانه‌های انتخاب می‌تواند به درک اهلی سازی، توسعه نژادها، ساختار جمعیت، تکامل سازگار و پاسخ به انتخاب کمک نماید. علاوه بر این به مطالعات GWAS، انتخاب ژنومیک و اجرای برنامه‌های اصلاح نژادی کمک می‌کند. در حال حاضر بیشتر مطالعات نشانه‌های انتخاب در دام‌ها بر روی تغییرات ژنتیکی در سطح نوکلئوتید به دلیل دسترسی آسان با استفاده از فن‌آوری پیشرفته ژنومی متمرکز بود اما چندین تغییر ژنتیکی دیگر در سطوح مختلف ممکن است در اثر انتخاب مانند انواع تعداد کپی (CNV)، تغییرات کروموزومی (مانند: حذف، دو برابر شدن و وارونگی) و تغییرات اپی ژنتیک قرار گیرد. روش‌های آماری و ابزارهای بیوانفورماتیکی بیشتری در سال‌های اخیر گسترش یافته‌اند تا همه این تغییرات ژنتیکی با قالب داده‌های انعطاف پذیر کنترل کنند تا داده‌های عظیم و پیچیده را به‌طور کارآمد مدیریت و دقت شناسایی نشانه‌های انتخاب را بهبود ببخشند.

منابع

- 1- Ablondi, M., Viklund, Å, Lindgren, G., Eriksson, S., Mikko, S., 2019. Signatures of selection in the genome of Swedish warmblood horses selected for sport performance. *BMC Genomics* 20, 717. <https://doi.org/10.1186/s12864-019-6079-1>. Achaz, G., 2009.
- 2- Achaz, G., 2009. Frequency Spectrum Neutrality Tests: One for All and All for One. *Genetics* 183, 249–258.
- 3- Bahbahani, H., Clifford, H., Wragg, D., Mbole-Kariuki, M.N., Van Tassell, C., Sonstegard, T., Woolhouse, M., Hanotte, O., 2015. Signatures of positive selection in East African Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. *Sci. Rep.* 5, 11729.
- 4- Bahbahani, H., Musa, H.H., Wragg, D., Shuiep, E.S., Almathen, F., Hanotte, O., 2019. Genome diversity and signatures of selection for production and performance traits in dromedary camels. *Front. Genet* 10, 893.
- 5- Bhatia, G., Patterson, N., Sankararaman, S., Price, A.L., 2013. Estimating and interpreting FST: The impact of rare variants. *Genome Res* 23, 1514–1521. <https://doi.org/10.1101/gr.154831.113>.
- 6- Bhatia, G., Patterson, N., Sankararaman, S., Price, A.L., 2013. Estimating and interpreting FST: The impact of rare variants. *Genome Res* 23, 1514–1521. <https://doi.org/10.1101/gr.154831.113>.
- 7- Cadzow, M., Boocock, J., Nguyen, H.T., Wilcox, P., Merriman, T.R., Black, M.A., 2014. A bioinformatics workflow for detecting signatures of selection in genomic data. *Front. Genet.* 5, 293
- 8- Caramelli, D., 2006. The Origins of Domesticated Cattle. *Hum. Evol.* 21, 107–122.
- 9- Carlson, C.S., 2005. Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Res* 15, 1553–1565.
- 10- Cesarani, A., Sorbolini, S., Criscione, A., Bordonaro, S., Pulina, G., Battacone, G., Marletta, D., Gaspa, G., Macciotta, N.P.P., 2018. Genome-wide variability and selection signatures in Italian island cattle breeds. *Anim. Genet.* 49, 371–383.
- 11- Diao, S., Luo, Y., Ma, Y., Deng, X., He, Y., Gao, N., Zhang, H., Li, J., Chen, Z., Zhang, Z., 2018. Genome-wide detection of selective signatures in a Duroc pig population. *J. Integr. Agric.* 17, 2528–2535.
- 12- Fariello, M.I., Boitard, S., Naya, H., SanCristobal, M., Servin, B., 2013. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics* 193, 929–941.
- 13- Fariello, M.I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., Cristobal, M.S., Boitard, S., 2014. Selection signatures in worldwide sheep populations. *PLoS One* 9, e103813.
- 14- Fay, J.C., 2011. Weighing the evidence for adaptation at the molecular level. *Trends Genet* 27, 343–349.
- 15- Gibson, J., Morton, N.E., Collins, A., 2006. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Gen.* 15, 789–795.

- 16- Grivet, D., Avia, K., Vaattovaara, A., Eckert, A.J., Neale, D.B., Savolainen, O., GonzálezMartínez, S.C., 2017. High rate of adaptive evolution in two widespread European pines. *Mol. Ecol.* 26, 6857–6870.
- 17- Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Finlay, E.K., Jianlin, H., Groeneveld, E., Weigend, S., 2010. Genetic diversity in farm animals - a review. *Anim. Genet.* 41, 6–31.
- 18- Hermisson, J., Pennings, P.S., 2005. Soft Sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169, 2335–2352.
- 19- Hermisson, J., Pennings, P.S., 2017. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol. Evol.* 8, 700–716.
- 20- Holsinger, K.E., 2004. Analysis of Genetic diversity in geographically structured populations: A Bayesian perspective. *Hereditas* 130, 245–255.
- 21- Howie, B.N., Donnelly, P., Marchini, J., 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 5, e1000529.
- 22- Jensen, J.D., Foll, M., Bernatchez, L., 2016. The past, present and future of genomic scans for selection. *Mol. Ecol.* 25, 1–4
- 23- Kasiriyani, M.M., GHolizadeh, M., Rahimimianji, GH and Moradi, 2022. MH. Selection signatures associated with the number of lambs per lambing in Baluchi ewes. *Animal Production Research*. Vol. 11, No. 4, 2022 (47-60)
- 23- Kijas, J.W., Lenstra, J.A., Hayes, B., Boitard, S., Porto Neto, L.R., San Cristobal, M., Servin, B., McCulloch, R., Whan, V., Gietzen, K., Paiva, S., Barendse, W., Ciani, E., Raadsma, H., McEwan, J., Dalrymple, B., 2012. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol* 10, e1001258.
- 24- Kimura, M., 1968. Evolutionary rate at the molecular level. *Nature* 217, 624–626.
- 25- Kryazhimskiy, S., Plotkin, J.B., 2008. The population genetics of dN/dS. *PLoS Genet* 4, e1000304.
- 26- Lachance, J., Tishkoff, S.A., 2013. SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. *BioEssays* 35, 780–786.
- 27- Lischer, H.E.L., Excoffier, L., 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299.
- 28- Malomane, D.K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A.R., Simianer, H., 2018. Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics* 19, 22.
- 29- Manunza, A., Cardoso, T.F., Noce, A., Martínez, A., Pons, A., Bermejo, L.A., Landi, V., Sánchez, A., Jordana, J., Delgado, J.V., Adán, S., Capote, J., Vidal, O., Ugarte, E., Arranz, J.J., Calvo, J.H., Casellas, J., Amills, M., 2016. Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. *Sci. Rep.* 6, 27296.
- 30-McDonald, J.H., Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351, 652–654.
- 31-Nicholson, G., Smith, A.V., Jonsson, F., Gustafsson, O., Stefansson, K., Donnelly, P., 2002. Assessing population differentiation and isolation from single-nucleotide polymorphism data. *J. Royal Stat. Soc. Series B* 64, 695–715.
- 32-Nielsen, R., 2005. Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218.
- 33-Otto, S.P., 2000. Detecting the form of selection from DNA sequence data. *Trends Genet* 16, 526–529.
- 34-Pritchard, J.K., Pickrell, J.K., Coop, G., 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* 20, R208–R215.
- 35-Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for wholegenome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.
- 36-Purfield, D.C., McParland, S., Wall, E., Berry, D.P., 2017. The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS One* 12, e0176780.

- 37- Qanbari, S., Pimentel, E.C.G., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A.R., Simianer, H., 2010. A genome-wide scan for signatures of recent selection in Holstein cattle. *Anim. Genet.* 41, 377–389.
- 38-Qanbari, S., Simianer, H., 2014. Mapping signatures of positive selection in the genome of livestock. *Livest. Sci.* 166, 133–143.
- 39-R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- 40- Rebelato, A.B., Caetano, A.R., 2018. Runs of homozygosity for autozygosity estimation and genomic analysis in production animals. *Pesqui. Agropecu. Bras.* 53, 975–984.
- 41-Rubin, C.J., Zody, M.C., Eriksson, J., Meadows, J.R.S., Sherwood, E., Webster, M.T., Jiang, L., Ingman, M., Sharpe, T., Ka, S., Hallböök, F., Besnier, F., Carlborg, Ö., Bed'hom, B., Tixier-Boichard, M., Jensen, P., Siegel, P., Lindblad-Toh, K., Andersson, L., 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464, 587–591.
- 42-Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., Campbell, S.J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., Lander, E.S., 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419, 832–837.
- 43-Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., Schaffner, S.F., Lander, E.S., 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918.
- 44-Signer-Hasler, H., Burren, A., Neuditschko, M., Frischknecht, M., Garrick, D., Stricker, C., Gredler, B., Bapst, B., Flury, C., 2017. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genet. Sel. Evol.* 49, 83.
- 45-Taye, M., Kim, J., Yoon, S.H., Lee, W., Hanotte, O., Dessie, T., Kemp, S., Mwai, O.A., Caetano-Anolles, K., Cho, S., Oh, S.J., Lee, H.-K., Kim, H., 2017. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. *BMC Genet* 18, 11.
- 46-Vitti, J.J., Grossman, S.R., Sabeti, P.C., 2013. Detecting natural selection in genomic data. *Annu. Rev. Genet.* 47, 97–120.
- 47-Voight, B.F., Kudravalli, S., Wen, X., Pritchard, J.K., 2006. A map of recent positive selection in the human genome. *PLoS Biol* 4, e72.
- 48-Vy, H.M.T., Kim, Y., 2015. A composite-likelihood method for detecting incomplete selective sweep from population genomic data. *Genetics* 200, 633–649.
- 49-Weigand, H., Leese, F., 2018. Detecting signatures of positive selection in non-model species using genomic data. *Zool. J. Linn. Soc.* 184, 528–583.
- 50-Weir, B.S., Hill, W.G., 2002. Estimating F-statistics. *Annu. Rev. Genet* 36, 721–750.
- 51-Williamson, S.H., Hubisz, M.J., Clark, A.G., Payseur, B.A., Bustamante, C.D., Nielsen, R., 2007. Localizing recent adaptive evolution in the human genome. *PLoS Genet* 3, e90.
- 52-Yurchenko, A.A., Daetwyler, H.D., Yudin, N., Schnabel, R.D., Vander Jagt, C.J., Soloshenko, V., Lhasaranov, B., Popov, R., Taylor, J.F., Larkin, D.M., 2018. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Sci. Rep* 8, 12984.
- 53-Yang, L., Xu, L., Zhu, B., Niu, H., Zhang, W., Miao, J., Shi, X., Zhang, M., Chen, Y., Zhang, L., Gao, X., Gao, H., Li, L., Liu, G.E., Li, J., 2017. Genome-wide analysis reveals differential selection involved with copy number variation in diverse Chinese Cattle. *Sci. Rep.* 7, 14299.
- 54-Zhao, F., McParland, S., Kearney, F., Du, L., Berry, D.P., 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet. Sel. Evol.* 47, 49.

An overview of the concepts, methods, and applications of selection signatures in the animal genome

Mohammad Mehdi Kasiriyani

Phd of Genetics and animal Breeding Center for Agriculture and Natural Resources Rasoul Akram Damghan, Iran

Abstract

Livestock populations have been continuously improved in terms of performance and productivity over the centuries through selective breeding. The selection strategy leaves traces in the genome known as selection signatures. Due to the progress in high-throughput technology, the identification of selection signatures in the genome has become more widespread in recent years. Such studies provide many insights related to domestication and evolutionary processes that result in a large number of livestock breeds that can be bred in different environments and production systems. In addition, it facilitates the identification of genes under selection related to economic traits in the livestock population. Various statistical methods have been developed to identify the selection signatures based on allelic frequency distribution, linkage disequilibrium (LD), reduction of locus diversity and population differentiation. Bioinformatics tools currently available to identify selection cues and briefly the results of recent studies related to selection signatures conducted in different livestock species.

Keywords: Positive selection, bioinformatics, genomics, domestic animals, selection signatures
