

## بررسی تنوع ژنتیکی گوسفند نژادهای بومی (افشاری، قزل و مغانی) و خارجی (دورپر، مریوس و رامنی) با استفاده از اطلاعات نشانگری متراکم

رقبه نبی‌لو<sup>۱</sup>، محمد باقر زندی باغچه مریم<sup>۲\*</sup> و محمد طاهر هرکی نژاد<sup>۳</sup>

۱، ۲ و ۳. دانشجوی سابق کارشناسی ارشد ژنتیک و اصلاح نژاد دام، استادیار و دانشیار، گروه علوم دامی، دانشکده کشاورزی، دانشگاه زنجان  
(تاریخ دریافت: ۱۳۹۷/۱/۲۹ - تاریخ پذیرش: ۱۳۹۷/۹/۳)

### چکیده

استفاده از نشانگرهای مولکولی در سال‌های اخیر جهت تعیین تنوع ژنتیکی بین جمعیت‌ها و گونه‌های مختلف جانوری کاربرد گسترده‌ای یافته است. مطالعات ژنتیک مولکولی، مقایسه تنوع ژنتیکی درون و بین نژادها و بازسازی از تاریخ نژاد و جمعیت اجدادی را امکان‌پذیر می‌نماید. این تحقیق با بررسی ساختار ژنتیکی در نژادهای بومی و خارجی با استفاده از ۵۰۰۰۰ نشانگر SNP با هدف شناسایی تنوع ژنتیکی انجام شد. اطلاعات ژنوتیپی مورد نیاز از پروژه Ovine HapMap شامل نژادهای بومی (نژاد افشاری، مغانی و قزل) و خارجی (نژاد دورپر، مریوس و رامنی) اخذ گردید. آنالیز آماری با استفاده از چندین روش لایه‌بندی جمعیت، بررسی ساختار جمعیت با استفاده از آمار چند متغیره بیزی و روش مبتنی بر مدل بررسی شد. نتایج به‌دست‌آمده از آنالیز مشخص مؤلفه‌های اصلی، تفکیک نژادهای ایرانی و خارجی را به‌خوبی نشان داد و هر دو تصویر آشکاری از ساختار ژنتیکی جمعیت‌های مورد بررسی را نشان داد. در روش DAPC، برای ارزیابی شمار بهینه خوشه با معیار ارزیابی BIC،  $K=4$  بهترین نتیجه را نشان داد. علی‌رغم این‌که در این مطالعه از روش‌های مختلف برای بررسی ساختار جمعیت استفاده شد، همه این روش‌ها توانستند ساختار جمعیت‌های بومی و خارجی را متمایز از هم نشان دهند که نژادهای ایرانی بررسی‌شده (افشاری، مغانی و قزل) در مقایسه با نژادهای خارجی دارای تشابه ژنتیکی بیشتری بوده و در یک گروه نژادی و نژادهای دورپر، مریوس و رامنی در گروه‌های مجزای ژنتیکی قرار گرفتند. البته زمانی که نژادهای داخلی به تنهایی مورد بررسی قرار گرفتند. این نژادها از لحاظ ژنتیکی کاملاً از هم متمایز بودند و میزان آماره تمایز جمعیتی  $F_{st}$  برای نژادهای افشاری، مغانی و قزل به ترتیب ۰/۰۳۸، ۰/۱۰۷ و ۰/۲۹۸ برآورد شد.

واژه‌های کلیدی: آنالیز لایه‌بندی جمعیت، تنوع ژنتیکی، نژاد بومی.

## Study of Genetic diversity of indigenous (Afshari, Moghani and Ghezel) and exotic (Romney, Merinos and Dorper) sheep breeds using high-density SNP markers

Roghayyeh Nabiloo<sup>1</sup>, Mohammad Bagher Zandi Baghcheh Maryam<sup>2\*</sup> and Mohammad Taher Harakinezhad<sup>3</sup>

1, 2. Former M. Sc. Student, Assistant Professor and Associate Professor, Department of Animal Science, Faculty of Agriculture, University of Zanjan, Iran

(Received: Apr. 18, 2018 - Accepted: Nov. 24, 2018)

### ABSTRACT

The use of molecular markers in recent years has been widely used to determine the genetic diversity between populations and animal species. Molecular genetics studies allow a comparison of genetic diversity within and across breeds and make a new insight to reconstruct the breed history and history of ancestral populations. The aim of this study was to investigate the genetic structure and genetic variation of indigenous and exotic sheep breeds using 50,000 SNP markers. Genotype data of indigenous breeds (Afshari, Ghezel and Moghani) and exotic breeds (Dorper, Merinos and Romney) were obtained from the Ovin HapMap project. Multiple population stratification analysis such as, population structure using multivariate statistics and model-based approach were applied. The result of all methods obviously showed the correct population differentiation. In DAPC  $K=4$  inferred best cluster results by BIC. Despite the use of different methods, all of these methods showed a distinct structure for indigenous and exotic populations and it can be concluded that Iranian sheep breeds has more genetic similarity rather than exotic breeds and Iranian sheep breeds grouped as one category and Romney, Merinos and Dorper breeds were separated as distant groups. However, when the indigenous breeds were studied alone, the breeds were genetically separated and population differentiation statistic ( $F_{st}$ ) for Afshari, Moghani and Qezel was 0.038, 0.107, and 0.298, respectively.

**Keywords:** Genetic diversity, indigenous breed, Population Stratification Analysis.

\* Corresponding author E-mail: mbzandi@znu.ac.ir

## مقدمه

گوسفند دام مهمی برای ایران محسوب می‌گردد و به لحاظ زیاد بودن جمعیت آن و تنوع نژاد و غیره بخش بزرگی از ذخایر ژنتیک دامی ایران را شامل می‌شود. بنابراین ضرورت شناسایی تنوع موجود در جمعیت‌های گوسفند بومی ایران و تعیین روابط تکاملی میان آن‌ها وظیفه‌ای است که بر عهده متخصصین ژنتیک و اصلاح دام کشور می‌باشد. نژادهای بومی در هر کشور به‌عنوان یک سرمایه ملی و محصولی استراتژیک محسوب می‌شوند و نیز به دلیل انتخاب پایین، تعداد زیاد پرورش‌دهندگان و محدودیت استفاده از تلقیح مصنوعی، معمولاً از تنوع ژنتیکی بالایی برخوردارند، از طرفی با گذشت زمان و کسب آگاهی بیشتر نسبت به اهمیت صفات مختلف نیازهای جدیدی مطرح می‌شود که متخصصین اصلاح نژاد را بر آن می‌دارد که از ذخیره ژنی دام‌های بومی استفاده نمایند. این مسئله بخصوص با افزایش تولید محصولات دامی و محصولات پیش‌بینی‌نشده در آینده لزوم حفظ تنوع ژنتیکی در دام‌های بومی را الزامی ساخته است (Frankham, 1994). چرا که یک گونه بدون حفظ تنوع ژنتیکی قادر به سازگاری با تغییرات محیطی و مبارزه با انگل‌ها نیست (Askari et al., 2011). در ایران حدود ۵۰ میلیون رأس گوسفند، شامل ۲۷ نژاد و اکوتیپ وجود دارد (Zamani et al., 2011) که لازم است تنوع ژنتیکی آن‌ها محاسبه و سپس در حفظ این نژادهای بومی تلاش شود. حفاظت باید براساس دانش عمیقی از منابع ژنتیکی نژادهای خاص باشد، لذا تلاش برای شناسایی و تعیین خصوصیات ژنتیکی نژادهای بومی و محلی بسیار اهمیت دارد (Shojaei et al., 2010; Zamani et al., 2011). امروزه نشانگرهای ژنتیکی نقش مهمی در مطالعات مرتبط با منشأ، تاریخچه و سیر تکاملی جوامع دامی ایفا می‌کنند (Fernández et al., 2013). دسترسی به تراشه‌های SNP فرصت مناسبی را جهت بررسی تنوع ژنتیکی و ساختار جمعیت‌های کوچک فاقد شجره را فراهم نموده است. نشانگرهای چندشکلی تک‌نوکلئوتیدی در بسیاری از حیوانات اهلی از جمله گاو (Mckay et al., 2007; Vonholdt et al., 2010)، سگ (Pertoldi et al., 2010)

(2010) و اسب (Petersen et al., 2013) به‌منظور شناسایی ساختار ژنتیکی جمعیت‌ها براساس روش بیزی به‌کار گرفته شده‌اند و به‌عنوان ابزاری استاندارد در این زمینه شناخته می‌شوند. تنوع ژنتیکی موجود در بین جوامع، حاصل عوامل بسیاری از جمله مهاجرت، جهش و اختلاط جمعیت‌هاست. این تنوع نشانه‌هایی از رانش تصادفی و انتخاب را به‌همراه دارد. آثار متقابل این عوامل می‌تواند سبب ایجاد جوامع و خرده جمعیت‌های مرتبط با آن شود. برای به‌دست آوردن اطلاعاتی درباره تاریخچه و سیر تکاملی جوامع، مشخص کردن خرده ساختارهای جمعیتی بسیار با اهمیت است، امروزه مطالعات متنوعی برای مشخص کردن خرده‌ساختارهای جمعیتی در دام‌ها صورت گرفته است (Kijas et al., 2009). تا به حال روش‌های آماری بسیاری برای بررسی ارتباط بین جوامع و مشخص کردن موقعیت افراد در جوامعی که از آن منشأ گرفته‌اند پیشنهاد شده است.

از میان این روش‌ها می‌توان به تکنیک لایه‌بندی جمعیت با روش تحلیل مؤلفه‌های اصلی (PCA)<sup>۱</sup>، بررسی ساختار جمعیت با استفاده از روش چند متغیره (DAPC)<sup>۲</sup> و روش مبتنی بر مدل (STRUCTURE) اشاره کرد. در روش اول با استفاده از ماتریس واریانس کواریانس روابط خویشاوندی ژنومی مؤلفه‌های اصلی را تفکیک کرده و افراد را تنها بر اساس تعداد کمتر مؤلفه اصلی به گروه‌های مختلفی متمایز می‌کند. ابزار اصلی در PCA برای تجزیه و تحلیل ماتریس واریانس کواریانس و به‌عنوان روش کاهش خطی ابعاد برای تعیین ساختارهای جمعیتی استفاده می‌شود. تجزیه و تحلیل مؤلفه‌های اصلی ابزاری استاندارد در ژنتیک جمعیت است که در بررسی جمعیت‌های انسانی اروپائی و هندی (Lao et al., 2008) و جمعیت‌های گوسفند ایرانی (Moradi et al., 2012) استفاده شده است. در روش دوم از تکنیک کاهش ابعاد از قبیل تجزیه و تحلیل مؤلفه‌های اصلی (PCA) به‌همراه ابزارهای خوشه‌بندی همچون K-means

1. Principle Component Analysis  
2. Discriminant Analysis Principal Component

نژاد خارجی (رامنی تعداد ۲۴ رأس کشور انگلستان، مینوس تعداد ۲۵۹ رأس از سه کشور استرالیا، چین و آلمان و دورپر تعداد ۲۷ رأس از آفریقای جنوبی) پروژه Sheep HapMap کنسرسیوم بین‌المللی ژنومیکس گوسفند (ISCG) استفاده شد. نمونه‌ها با استفاده از تراشه ژنومی Illumina Ovine SNP 50K با بکارگیری پروتکل استاندارد شرکت Illumina تعیین ژنوتیپ شدند. این تراشه‌ها امکان تعیین ژنوتیپ همزمان حدود ۵۴۰۰۰ جایگاه مارکری را فراهم می‌کند.

**مراحل فیلتراسیون داده‌های حاصل از تعیین ژنوتیپ** برای اطمینان از کیفیت داده‌های حاصل از تعیین ژنوتیپ، در آنالیزهای نهایی مراحل، مختلف فیلتراسیون بر روی داده‌های اولیه با استفاده از نرم‌افزار Plink (Purcell, 2010) نسخه ۱٫۹ اعمال شد. بدین ترتیب در ابتدا حیوانات دارای بیش از ۵ درصد ژنوتایپ از دست رفته از آنالیزهای بعدی کنار گذاشته می‌شوند چون نمونه‌های با کیفیت پایین با احتمال بیشتری با داده‌های گمشده همراه هستند و منجر به افزایش خطای ژنوتایپ می‌شود (Barendse et al., 2009).

دو فاکتور  $MAF^1$  (حداقل فراوانی اللی) و  $Call\ rate^2$  (درصدی از نمونه‌ها که برای آن مارکر ژنوتایپ شده‌اند) برای هر SNP محاسبه شدند و SNPs که دارای  $MAF$  و  $Call\ rate$  به ترتیب کمتر از پنج درصد و ۹۵ درصد بودند حذف شدند. برای SNPs باقی‌مانده آنهایی که در هر کدام از نژادهای مورد مطالعه دارای تعادل هاردی-واینبرگ نبودند به‌عنوان معیاری از خطای تعیین ژنوتیپ کنار گذاشته شده، تعادل هاردی-واینبرگ در کلیه جایگاه‌ها مورد بررسی قرار گرفت و جایگاه‌های با  $P\text{-value}$  کمتر از  $10^{-6}$  از مجموعه داده‌ها کنار گذاشته شدند (Teo et al., 2007). در نهایت مارکریایی که موقعیت ژنومی آن‌ها نیز مشخص نبودند از مجموع مارکرها خارج شدند (جدول ۲). همچنین به‌منظور حذف SNPs که در حالت عدم تعادل پیوستگی با یکدیگر قرار داشتند تراکم مجموعه داده‌ها حاصل از کنترل کیفیت با

استفاده می‌شود. این روش به استخراج اطلاعات غنی از اطلاعات ژنتیکی، ارائه انتساب افراد به گروه، یک ارزیابی بصری از تمایز بین جمعیت و سهم الل‌های افراد در ساختار جمعیت کمک می‌کند (Jombart et al., 2010). DAPC بر مدل ژنتیکی خاصی وابسته نیست و در نتیجه فرض تعادل هاردی-واینبرگ و تعادل پیوستگی (لینکاژی) ندارد. به این ترتیب برای انواع موجودات صرف‌نظر از پلوئیدی بودن و نرخ نوترکیبی آن‌ها، کاربرد دارد (Jombart et al., 2010). در روش سوم از روش‌های استاندارد آماری بیزی برای برآورد پارامترهای جمعیتی استفاده می‌شود و معمولاً برای هر جامعه فرض تعادل هاردی-واینبرگ لحاظ می‌شود. در این روش اطلاعات پیشین مربوط به جمعیت‌ها و فراوانی اللی با اطلاعات موجود از ژنوتیپ افراد تلفیق شده و از آن‌ها جهت به‌دست آوردن توزیع پسین جمعیت‌ها و فراوانی اللی استفاده می‌شود. بنابراین امکان در نظر گرفتن اطلاعات اولیه در مورد جمعیت‌ها نظیر مکان‌های جغرافیایی آن‌ها نیز فراهم است (Holsinger et al., 2004). استنباط بیزی برای اولین بار به‌وسیله Symons (1981) جهت خوشه‌بندی مبتنی بر مدل مورد استفاده قرار گرفت و در حال حاضر به روشی استاندارد در تعیین ساختار ژنتیکی جمعیت‌ها تبدیل شده است. در این مطالعه ساختار جمعیت، تمایز و تنوع ژنتیکی در بین نژادهای بومی و خارجی با استفاده از چندین روش لایه‌بندی جمعیت (PCA)، بررسی ساختار جمعیت با استفاده از روش چندمتغیره (DAPC) و روش مبتنی بر مدل (STRUCTURE) جهت بررسی دقیق و نتیجه‌گیری کامل جمعیتی در مورد سه نژاد ایرانی (افشاری، مغانی و قزل) و نژادهای خارجی (دورپر، مینوس و رامنی) مورد بررسی قرار گرفت.

## مواد و روش‌ها

### نمونه‌گیری

در این پژوهش در مجموع اطلاعات تعیین ژنوتیپ تعداد ۹۹ رأس گوسفند بومی کشور (شامل ۳۷ رأس گوسفند افشاری استان زنجان، ۲۷ رأس گوسفند مغانی استان آذربایجان شرقی و ۳۵ رأس گوسفند قزل استان آذربایجان غربی) و ۳۱۰ رأس گوسفند از سه

1. Minor Allele Frequency

2. Consistency between animal replication

PCA روابط خویشاوندی ژنومی جهت تفکیک لایه‌بندی جمعیتی و شناسایی حیواناتی که خارج از گروه نژادی خود قرار گرفته‌اند با استفاده از نرم‌افزار Plink و ترسیم گراف‌ها در محیط R (<https://cran.r-project.org/bin/windows/base/>) انجام شد.

بررسی ساختار جمعیت با استفاده از روش آنالیز چند متغیره (DAPC)

به‌منظور تشخیص تمایز و تنوع ژنتیکی در بین و داخل گروه‌ها، آنالیز تشخیص مؤلفه‌های اصلی که یک روش پارامتری و غیر مبتنی بر مدل می‌باشد، با بسته آماری adegenet مربوط به نرم‌افزار R اجرا شد. شناسایی تعداد بهینه خوشه با اجرای پی‌درپی با افزایش K و راه‌حل‌های مختلف خوشه‌بندی در مقایسه با معیار اطلاعات بیزی (BIC) می‌باشد. در حالت ایده‌آل خوشه‌بندی بهینه باید با پایین‌ترین BIC مطابقت داشته باشد. در عمل بهترین BIC اغلب توسط نقطه حداقل در منحنی ارزش BIC به عنوان تابعی از K نشان داده شد. این تابع در آغاز داده‌ها را با PCA تبدیل می‌کند، سپس الگوریتم K-msans با افزایش مقادیر K اجرا می‌شود.

استفاده از نرم‌افزار Plink کاهش داده شد. این بخش ویرایش برای به‌دست آوردن میزان تنوع ژنتیکی پایه بین جمعیت‌ها انجام شد. این آنالیز نشانگرهای SNP که با یکدیگر LD دارند را از آنالیزهای بعدی کنار می‌گذارد تا نتایج بهتری از تنوع ژنتیکی بین افراد را نشان دهد (Petersen *et al.*, 2013)، که به این منظور در پنجره‌هایی که شامل ۵۰ SNP و با حرکت ۱۰ SNP رو به جلو در هر مرحله، SNPs دارای  $r^2$  (معیار عدم تعادل پیوستگی) بیش از ۰/۱ با یکدیگر از مجموعه داده‌ها حذف شدند.

### آنالیزهای آماری

آنالیز مؤلفه‌های اصلی (تکنیک لایه‌بندی جمعیت)

هدف از تجزیه به مؤلفه‌های اصلی آن است که واریانس موجود در داده‌های چند متغیره را به مؤلفه‌هایی تجزیه کند که اولین مؤلفه اصلی تا آنجا که ممکن است علت بیشترین واریانس موجود در داده‌ها را توجیه کند. دومین مؤلفه نیز به ترتیب بیشترین واریانس ممکن بعد از مؤلفه اول و الی آخر را توجیه نماید. به‌علاوه در این روش هر مؤلفه مستقل از مؤلفه‌های دیگر است، یعنی بین هر مؤلفه و مؤلفه‌های دیگر همبستگی وجود ندارد. آنالیز

جدول ۱. مناطق جغرافیایی و شمار حیوانات در هر منطقه

Table 1. The geographic regions and the numbers of animals in each region

Breed	Population size (approx)	Region sampled	Number sampled	Production Classification(s)
Afshari	2 Million	Iran, Zanjan	37	Meat
Moghani	1.5 Million	Iran, Ardabil	27	Meat and Milk
Ghezel	1 Million	Iran, Azerbaijan Sharghi	35	Meat and Milk
Romney	45 Million	England	24	Meat and Wool
Merino	72 Million	England, Australia and China	259	Wool
Dorper	20 Million	South African	27	Meat and Wool

جدول ۲. معیارهای مورد استفاده در کنترل کیفیت داده‌های حاصل از ژنوتایپ در سری اطلاعاتی نژادهای ایرانی و خارجی

Table 2. Measures of quality control from genotype in the Iranian and Exotic Breeds information series

Quality control thresholds	Iranian	Exotic	Information series
	Breeds	Breeds	Both breeds
Total number of animals	99	310	409
Total number of SNPs	49034	49034	49034
Remove SNPs with 5% > MAF in total samples	2892	1534	1207
Remove SNPs with 95% > Call Rate in total samples	71	14	10
Removing SNPs that have a large deviation from the H-W equilibrium (0.000001 > 0)	757	1315	1647
Remove unknown SNPs	379	372	374
LD Pruning	19479	19482	19618
Number of remained SNPs	26317	26317	26317
Relative remained SNPs	0.54	0.54	0.54

## 1. Bayesian Information Criterion

شده است) در اجزای مختلف میانگین گرفته شد. در مرحله دوم میزان تغییر تابع درست‌نمایی نسبت به  $K$  به دست آورده شد:

$$L'(K) = L(K) - L(K - 1)$$

و پس از این، مرتبه دوم تغییر  $L(K)$  نسبت به  $K$  نیز محاسبه شد:

$$|L''(K)| = |L'(K + 1) - L'(K)|$$

در نهایت مقدار  $\Delta K$  با تقسیم میانگین مقادیر مرتبه دوم تغییر  $L(K)$  نسبت به  $K$  بر انحراف استاندارد مقادیر درست‌نمایی به دست آورده شد:

$$\Delta K = m |L''(K)| / s[L(K)]$$

مقداری از  $K$  که به‌ازای آن مقدار  $\Delta K$  حداکثر شد، به‌عنوان بهترین مقدار  $K$  انتخاب شد. نتایج به‌دست‌آمده از این نرم‌افزار برای بررسی ساختار جمعیت‌ها مورد استفاده قرار گرفت.

## نتایج و بحث

### آنالیز لایه‌بندی جمعیت

برای ارزیابی اختلاف ژنتیکی میان جمعیت‌ها (۶ نژاد) و بررسی نحوه قرار گرفتن حیوانات در گروه‌های نژادی خود از آنالیز PCA با استفاده از اطلاعات روابط خویشاوندی ژنومی نمونه‌ها تعیین ژنوتیپ شده به‌عنوان داده مورد نیاز استفاده شد. نتایج این آنالیز نشان داد که PCA 1 باعث جدا شدن نژادها از یکدیگر شده و هر حیوان نیز بر روی PCA 2 از گروه نژادی خود فاصله می‌گیرد. اما همان‌طور که از شکل ۱ استنباط می‌شود، سه نژاد گوسفند ایرانی افشاری، مغانی و قزل نزدیک هم قرار گرفته‌اند که این بیانگر این نکته است که ژنوم نژادهای ایرانی به نسبت ژنوم نژادهای خارجی دارای تشابه ژنتیکی بیشتری می‌باشد، این نتایج نشان داده که با تکنیک لایه‌بندی جمعیتی می‌توان نژادهای ایرانی را در یک گروه کاملاً متمایز از نژادهای خارجی در نظر گرفت و نژادهای رامنی، مرینوس و دورپر که هر کدام دارای شاخصه‌های لایه‌بندی جمعیتی متمایزی می‌باشند گروه‌های نژادی متمایز در نظر گرفته شده‌اند که در نهایت مطابق شکل ۱، چهار گروه نژادی متمایز برای آنالیز تمایز جمعیتی تشکیل شد.

در نهایت با استفاده از آنالیز تمایز نژادی و بررسی گراف‌های به‌دست‌آمده از این روش در استخراج اطلاعات غنی از اطلاعات ژنتیکی، ارائه انتساب افراد به گروه، یک ارزیابی بصری از تمایز بین جمعیت و سهم الل‌های افراد در ساختار جمعیت کمک کرد.

بررسی ساختار جمعیت با استفاده از روش مبتنی بر مدل برنامه Structure (Porras-Hurtado *et al.*, 2013) پیاده‌سازی یک روش خوشه‌بندی مبتنی بر مدل برای استنتاج ساختار درست جمعیت با استفاده از داده‌های ژنوتایپ شامل نشانگرها (مارکرها) غیرمرتبط است. اختصاص مدل مناسب برای مشاهدات هر خوشه (جمعیت) اولین چالش در روش‌های خوشه‌بندی مبتنی بر مدل است. در اینجا فرض شده است که هر خوشه به‌وسیله یک مجموعه مشخص از فراوانی‌های اللی مدل‌بندی شده است. فرضیات اساسی مدل عبارتند از: تعادل هاردی-واینبرگ داخل جمعیت‌ها و تعادل پیوستگی کامل میان جایگاه‌ها. تحت این فرضیات هر الل در هر جایگاه و در هر ژنوتیپ یک نمونه مستقل از یک توزیع فراوانی خاص در نظر گرفته می‌شود.

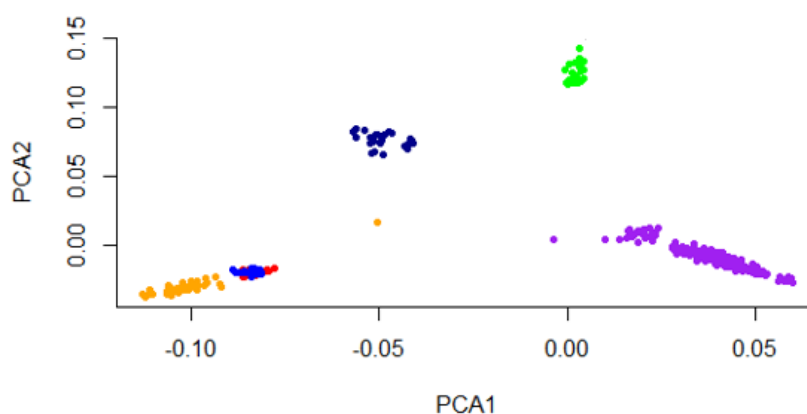
لذا در این مطالعه از نرم‌افزار Structure 2.3.4 (Pritchard *et al.*, 2010) جهت خوشه‌بندی جمعیت‌ها به گروه‌های ژنتیکی استفاده شد. مدل آمیختگی و فراوانی اللی تصحیح شده در نرم‌افزار Structure جهت آنالیز ساختار داده‌ها مورد استفاده قرار گرفتند. در هر بار اجرای برنامه ۵۰۰۰ دور گرم شدن به‌همراه تعداد ۵۰۰۰۰ تکرار برای مقادیر مختلف  $K$  (تعداد جمعیت‌های فرض‌شده) به‌کار گرفته شد. ارزیابی رسیدن به نقطه همگرایی با کمک بررسی مقادیر  $\alpha$  و مقادیر درست‌نمایی (که به‌وسیله نرم‌افزار در خروجی‌های هر تکرار ارائه می‌شوند) انجام شد. مناسب‌ترین تعداد جمعیت‌های فرض‌شده جهت تشریح گروه‌های ژنتیکی موجود در داده‌ها به‌کمک روش Evanno *et al.* (2005) انتخاب شد. بر اساس این روش از لگاریتم درست‌نمایی داده‌ها جهت تعیین مقدار حقیقی  $K$  (تعداد جمعیت‌های فرض‌شده) استفاده شد. در ابتدا از درست‌نمایی داده‌ها ( $L(K)$ ) (که در خروجی نرم‌افزار داده

می‌باشند. با توجه به مقادیر نوشته شده در زیر جمعیت شماره یک که مربوط به نژاد دورپر می‌باشد که در گروه استنباطی شماره دو قرار گرفته است و جمعیت شماره دو مربوط به نژاد مرینوس می‌باشد که در گروه های استنباطی شماره چهار، پنج و شش قرار گرفته است. در حقیقت نمونه‌های گوسفند نژاد مرینوس نمونه‌ها از سه کشور استرالیا، آلمان و چین اخذ شده است. جمعیت شماره سه مربوط به نژاد رامنی می‌باشد که در گروه استنباطی شماره سه قرار گرفته است و جمعیت‌های چهار، پنج و شش مربوط به سه نژاد بومی کشور (افشاری، مغانی و قزل) که آن‌ها را در گروه ژنتیکی استنباطی شماره یک طبقه‌بندی کرده است با توجه به نتایج به‌دست‌آمده می‌توان بیان کرد که نژادهای ایرانی را می‌توان در یک گروه ژنتیکی و نژادهای خارجی را در گروه‌های ژنتیکی مجزا از هم طبقه‌بندی کرد.

### بررسی ساختار جمعیت با استفاده از روش آنالیز چند متغیره (DAPC)

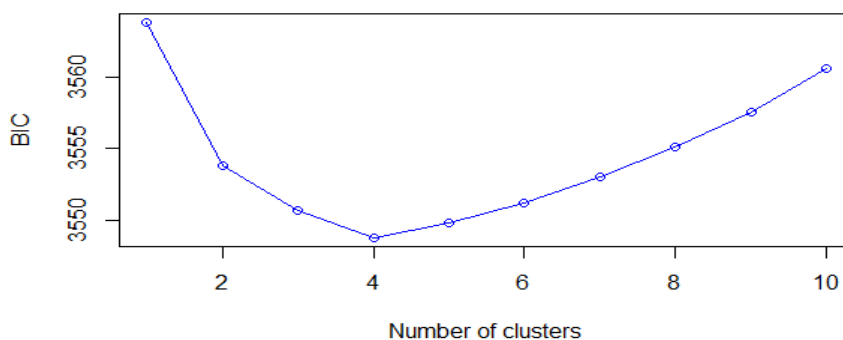
آنالیز مشخص کننده مؤلفه‌های اصلی توانایی تشخیص بهترین تمایز در بین نژادهای شناسایی و توصیف خوشه‌های ژنتیکی و همچنین احتمال عضویت هر فرد در جمعیت‌های مختلف در مطالعه را براساس توابع تفکیک حفظ شده فراهم می‌کند. شکل ۲ کاهش واضح BIC تا  $K=4$  خوشه و پس از آن افزایش BIC را نشان می‌دهد. با توجه به شکل در نتیجه تعداد خوشه‌های بدست آمده از نتایج، چهار خوشه ژنتیکی را توصیف می‌کند که شش نژاد مورد مطالعه می‌تواند در چهار خوشه ژنتیکی قرار داشته باشد.

قرار گرفتن جمعیت (نژاد) در گروه‌های واقعی و گروه‌های استنباط شده در شکل ۳ نشان داده که گروه‌های اصلی نژادی کاملاً با این روش قابل تفکیک



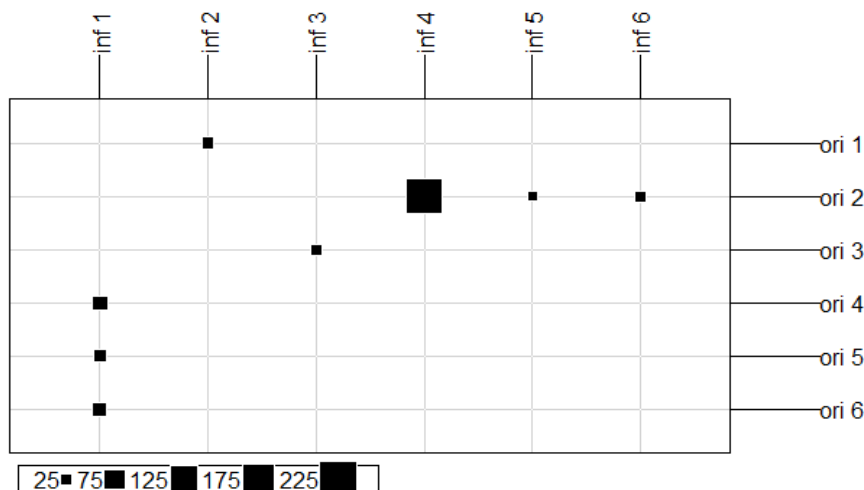
شکل ۱. خوشه‌بندی حیوانات براساس آنالیز PCA با استفاده از اطلاعات روابط خویشاوندی ژنومی. نمونه‌ها: نژادهای افشاری، مغانی، قزل، رامنی، مرینوس و دورپر به ترتیب با رنگ‌های نارنجی، قرمز، آبی، سبز، ارغوانی، آبی تیره نشان داده شده‌اند.

Figure 1. Animal clustering based on PCA analysis using genomic relationship information. Samples: Afshari, Moghani, Ghezel, Romney, Merinos and Dorper Breed's are shown in orange, red, blue, green, purple, dark blue, respectively.



شکل ۲. تعداد خوشه‌ها براساس معیار اطلاعات بیزین (BIC)

Figure 2. Number of clusters based on Bayesian Information Criterion (BIC)



شکل ۳. گروه‌های واقعی (ori) و گروه‌های استنباط شده (inf). (نژادهای افشاری، معانی و قزل در گروه استنباط شماره یک، نژاد دورپر در گروه استنباط شماره دو، نژاد مرینوس در گروه‌های استنباط شده (چهار، پنج و شش) و نژاد رامنی در گروه استنباط شماره سه قرار گرفته‌اند).

Figure 3. Actual groups (ori) and inferred groups (inf). (Afshari, Moghani and Ghezeli Breed's are located in inference group number one, Dorper Breed in inference group number 2, Merinos Breed in inference groups (four, five and six) and Romney Breed in the third inference group)

آنالیز تعلق نژادی مورد بررسی قرار گیرند که با این روش هر نمونه‌ای از هر جمعیتی با قرار دادن آنها در فایل‌های ورودی آنالیز نسبت تعلق هر نمونه به خوشه اصلی را تعیین کرد. همانطور که در این شکل ملاحظه می‌شود همه افراد در هر جمعیت به خود خوشه اصلی منتسب شده به جز این که چندین نمونه از نژاد مرینوس در خوشه چهار که از نظر ژنتیکی و آزمون انتساب به جمعیت مرینوس تعلق ندارند که ممکن است مربوط به نژادهای دیگری خارج از نژادهای مورد بررسی باشد و با توجه به اینکه اطلاعات گله‌ها و محل نمونه‌های نژادهای خارجی در دسترس نیست امکان بررسی پیشینه نمونه‌های مشخص شده وجود ندارد.

مزیت اصلی روش PCA همانند DAPC، توانایی آن برای شناسایی ساختارهای جمعیتی با حجم گسترده‌ی داده‌های نشانگری، زمان محاسباتی کمتر و نبود فرض درباره مدل ژنتیک جمعیتی است. با این حال PCA روش ناپارامتری بوده فاقد برخی ویژگی‌های ضروری بررسی ساختار جمعیت‌های زیستی (بیولوژیکی) است. PCA یک ارزیابی گروهی را فراهم نمی‌سازد و نیازمند پیش تعریف مناسبی از خوشه‌ها برای بررسی ساختار جمعیتی است و حتی برای به دست آوردن تصویر آشکار از تنوع بین جمعیت‌ها مناسب نیست (Jombart et al.,

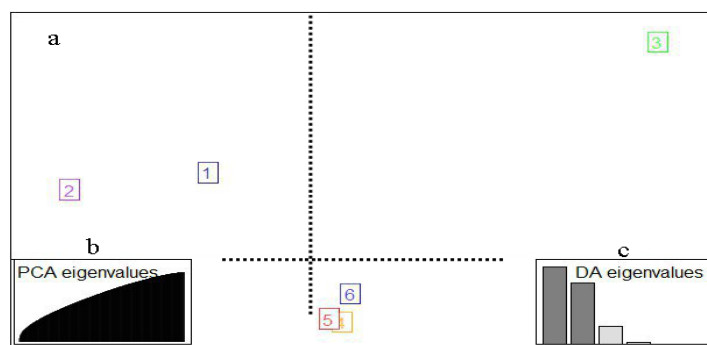
قرار گرفتن افراد در گروه‌های منتسب در شکل ۴ ارائه شده است. همان‌گونه که مشاهده می‌شود نژادهای ایرانی در یک گروه و نژادهای خارجی هر کدام برای خود گروه‌های مجزایی را تشکیل دادند همچنین نژادهای متعلق به مرینوس که از سه کشور مختلف بوده به خوبی تفکیک شده است. مقادیر ویژه مؤلفه‌های اصلی در تجزیه و تحلیل‌ها نمایش داده شده است که مقادیر ویژه مربوط به نسبت واریانس بین گروه‌ها بر روی واریانس درون گروه‌ها برای هر تابع تفکیک می‌باشد که بیانگر این نکته است که اطلاعات می‌تواند در داخل دو تابع تفکیک قرار بگیرد (شکل c-۴).

شکل ۵ آنالیز کلی آزمون انتساب را نشان می‌دهد، به طوری که نسبت‌هایی از انتقال موفق افراد بر اساس توابع تفکیک‌کننده واریانس درون گروهی و بین گروهی به خوشه‌های اصلی خود را نشان می‌دهد که هر سطر مربوط به یک فرد و هر ستون به یک گروه می‌باشد. رنگ‌های آبی نشان‌دهنده مکان فرد در خوشه‌های قبلی ارائه شده آنالیز DAPC است. همان‌طور که مشاهده می‌شود، نتایج طبقه‌بندی DAPC سازگار به خوشه‌های اصلی است. همچنین این شکل زمانی مفید است که گروه‌های بیولوژیکی بدون داشتن اطلاعات پس‌زمینه و شجره‌ای آنها برای

درباره زمینه ژنتیکی جمعیت است ( Jombart *et al.*, 2010). در پژوهشی برای بررسی ساختار جمعیتی درختان آکاسیا با روش‌های مبتنی بر مدل و روش اکتسابی DAPC، این روش‌ها تداوم بالایی در برآورد ساختار جمعیت و استنباط احتمالات عضویت افراد به هر گروه را نشان دادند (Pometti *et al.*, 2014). روش PCA به‌عنوان جایگزینی برای روش‌های خوشه‌بندی بیزی پیشنهاد شده است ( Lee *et al.*, 2009; Liu & Zhao, 2006). همچنین تصحیح لایه‌بندی جمعیتی، روش DAPC به‌علت این‌که واریانس بین گروهی را افزایش و واریانس درون گروهی را کاهش می‌دهند بهتر از روش PCA است (Jombart *et al.*, 2010). در استنباط و تفسیر ساختار ژنتیکی از روش‌های پرشمار بررسی ساختار جمعیتی، روش DAPC با توجه به برتری‌هایی همچون تفسیر بهتر زیر جمعیت‌ها و انتساب خوشه‌ها بهتر از روش PCA عمل کرد (Sethuraman, 2013).

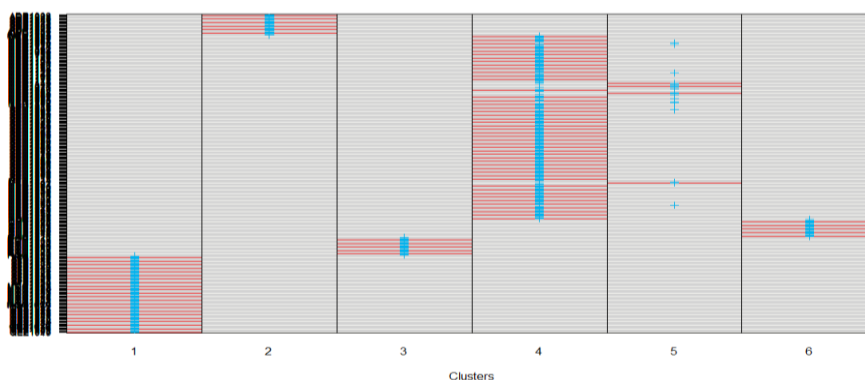
تجزیه و تحلیل مؤلفه‌های اصلی تنوع کلی میان افراد که شامل هر دو تنوع بین گروهی و تنوع درون گروهی می‌شود را خلاصه می‌کند و تصویر آشکاری از تفاوت بین گروه‌ها را نشان می‌دهد ( Jombart & Collins, 2015) که با نتایج به‌دست‌آمده از این پژوهش مطابقت دارد. برای ارزیابی ارتباط بین خوشه‌های مختلف، باید یک روش مناسب بر تنوع بین گروهی تمرکز کند و تنوع درون گروهی را نادیده بگیرد که تجزیه تشخیصی مؤلفه‌های اصلی این کار را انجام می‌دهد (Jombart & Collins, 2015).

روش DAPC نسبت به PCA در به تصویر کشیدن تفاوت بین گروهی برتری دارد که تفاوت گروهی را به‌طور مناسب‌تر نشان داد. مشخصه اصلی DAPC این است که یک روش پارامتری بوده و توانایی آن برای شناسایی ساختارهای ژنتیک در مجموعه داده‌های بزرگ در زمان محاسباتی ناچیز به نسبت PCA و بدون هیچ فرضی



شکل ۴. اختصاص افراد در گروه‌ها: آنالیز مؤلفه‌های اصلی (۴-ا) به‌همراه اطلاعات کل مقادیر ویژه (شکل ۴-ب) و مقادیر ویژه مورد نیاز (شکل ۴-ج)

Figure 4. Assigning individuals to groups: Principle Component Analysis (4-a) along with the total eigenvalues (4-b) and the required eigenvalues (4-c)



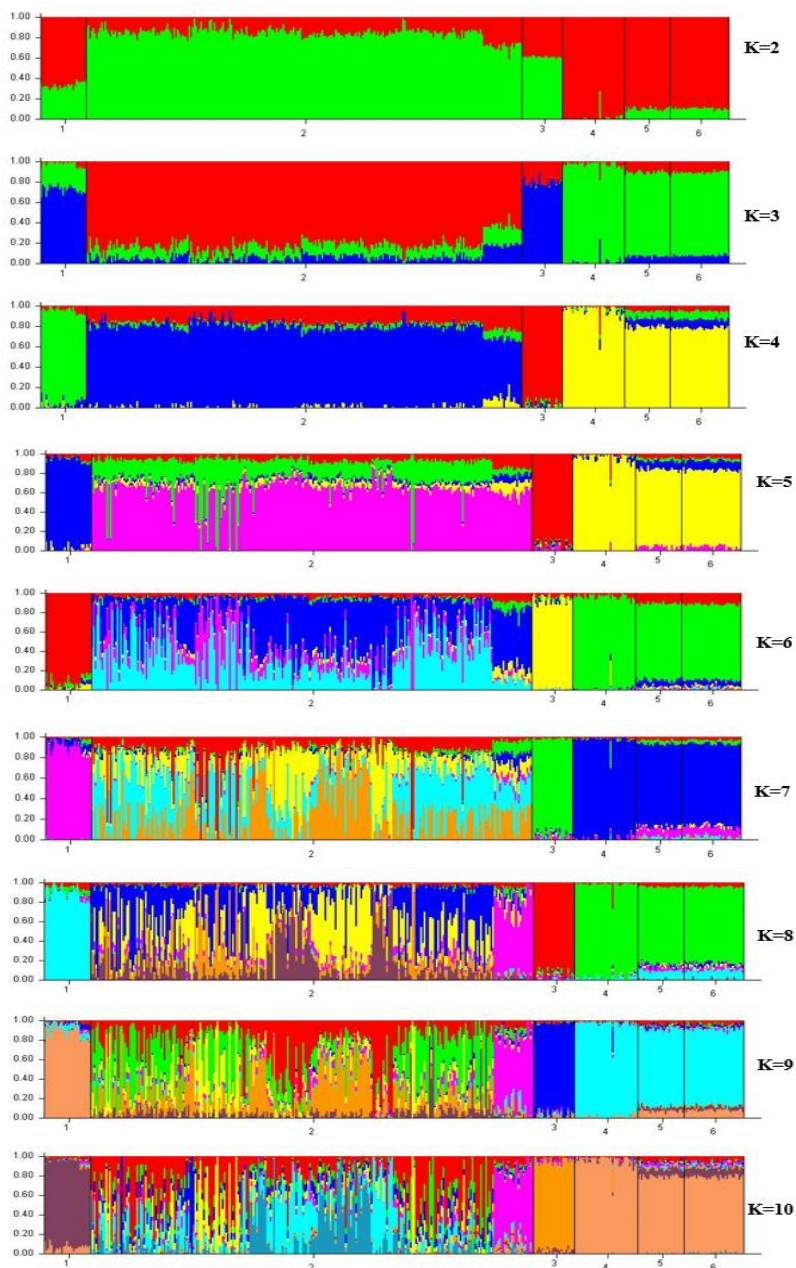
شکل ۵. عضویت افراد در گروه‌های مختلف. شماره‌های ۱ تا ۶ نشانگر جمعیت‌های استنباط شده شکل ۳ می‌باشد.  
Figure 5. Memberships in different groups. Number 1 to 6 represents inferred populations as in Figure 3



نژادها و همینطور خرده جوامع برای نمونه‌های مورد بررسی معین نیست. لذا دو پیمایش با افزوده شدن تعداد K از ۲ به ۱۰ جهت بررسی ساختار جمعیتی در گوسفندان ایرانی و خارجی مورد استفاده قرار گرفت (نتایج K=2 تا K=10 در شکل ۶ ارائه شده است).

بررسی ساختار جمعیت با استفاده از روش مبتنی بر مدل

برنامه Structure (مبتنی بر استنباط بیزی) جهت بررسی ساختار ژنتیکی و آمیختگی در مجموعه داده‌های گوسفندان بومی و خارجی شامل ۲۶۳۱۷ جایگاه به کار گرفته شد. نظر به این‌که دقیقاً تعداد



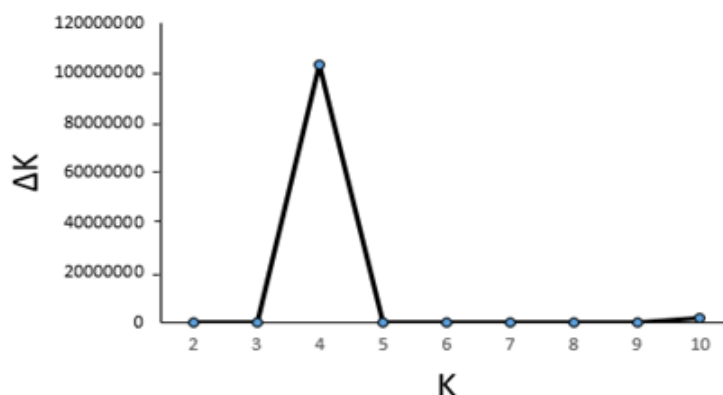
شکل ۶. میزان انتساب به خوشه‌های مختلف در ۱۰ مقدار مختلف بررسی شده از K (نمودارهای مربوط به برآوردهای مقادیر Q توسط نرم‌افزار STRUCTURE هر حیوان با یک خط عمودی که به K قطعه رنگی تقسیم شده، نشان داده شده که طول آن نشانگر سهم هر کدام از K خوشه استنتاجی است. شماره‌های ۱ تا ۶ نشانگر جمعیت‌های از پیش تعیین شده می‌باشند).

Figure 6. Cluster assignment for ten values of K. (investigated graphs for estimates of Q values by STRUCTURE software. Each animal is represented by one vertical line divided into K colored segments, whose length indicates The contribution of each K is the inferential cluster. Number 1 to 6 represent outline populations).

در  $K=4$  به دست آمد، که نشان می‌دهد جمعیت‌ها در چهار گروه (زیر جمعیت) قرار دارند (شکل ۷). نسبت انتساب هریک از نژادها به خوشه‌های مختلف در  $K=4$  در جدول ۳ نمایش داده شده است. همچنین مقادیر مختلف فراوانی اللی به اشتراک گذاشته شده بین جمعیت‌های فرض شده به تفکیک خوشه‌ها در جدول ۴ نشان داده شده است، داده‌های این جدول نشان می‌دهد که فاصله ژنتیکی بین اولین و دومین خوشه استنباط شده (نژادهای ایرانی و نژاد دورپر) در مقایسه با فاصله ژنتیکی روابط خویشاوندی خوشه‌های اول و دوم و با خوشه سوم (نژاد رامنی) و خوشه چهارم (نژاد مرینوس) کمتر بوده است. همچنین فاصله ژنتیکی نژادهای ایرانی با نژاد دورپر در مقایسه با نژادهای رامنی و مرینوس بیشتر بوده است. کمترین فاصله ژنتیکی بین نژادهای ایرانی و نژاد دورپر احتمالاً به دلیل آمیختگی این نژاد از گوسفند پرشین بلک با دورست شاخدار می‌باشد (Cloete & Olivier, 2010; Cloete *et al.*, 2007). نتایج حاصل از این آنالیز مشابه و تأییدکننده نتایج‌های به دست آمده از آنالیز PCA و DAPC می‌باشد. جهت بررسی دقیق نژادهای ایرانی پارامترهای مذکور به صوت مجزا بر روی داده‌های ژنوتایپ گوسفندان ایرانی قرار گرفت و نژادهای گوسفند افشاری، مغانی و قزل به تفکیک تمایز جمعیتی بر اساس فراوانی اللی تمایز قابل قبولی را دارا می‌باشند.

در  $K=2$  (فرض شده که وجود تنها دو جمعیت ایرانی و خارجی کل تغییرات موجود در داده‌ها را توجیه می‌کند) نژادهای ایرانی افشاری (۹۸/۶ درصد)، مغانی (۸۹/۸ درصد) و قزل (۸۹/۴ درصد) در یک خوشه سهیم بودند. نتایج حاصل از این مطالعه نشان‌دهنده مطابقت خوشه‌بندی ژنتیکی جمعیت‌ها با فواصل جغرافیایی و محیط‌های پرورش آنهاست. نژادهای ایرانی که در منطقه سرد و کوهستانی شمال غرب کشور نگهداری می‌شوند در اولین سطح خوشه‌بندی ( $K=2$ ) در خوشه یکسانی طبقه‌بندی شده‌اند و به لحاظ ژنتیکی نیز فاصله کمتری را در سطوح بالای خوشه‌بندی نشان داده‌اند. در  $K=3$  دوباره نژادهای ایرانی خوشه ژنتیکی متمایزی را تشکیل دادند، همچنین نژاد مرینوس خوشه مستقلی را برای خود تشکیل داد و نژادهای رامنی و دورپر تقریباً ترکیب ژنتیکی مشابهی را نشان دادند.

در  $K=4$  سه نژاد ایرانی (رنگ زرد) همچنان در یک خوشه مستقل ظاهر شده و متمایز از نژادهای خارجی بوده و نژادهای دورپر (رنگ سبز)، مرینوس (آبی) و رامنی (قرمز) نیز هریک برای خود خوشه‌های متمایزی را تشکیل دادند. در مقادیر  $K$  بالاتر از چهار نشانه‌های آمیختگی بیشتر در نژادهای مرینوس ظاهر شد. بررسی ساختار ژنتیکی جمعیت‌ها بر اساس داده‌های SNP و با استفاده از روش Evanno *et al.* (2005) نشان داد که بیشترین میزان  $\Delta K$  در مقابل تغییرات  $K$



شکل ۷. نمایش تغییرات مقادیر  $\Delta K$  به ازای مقادیر مختلف  $K$  (تعداد جمعیت‌های فرض شده). حداکثر مقدار  $\Delta K$  نشان‌دهنده مناسب‌ترین مقدار  $K$  است.

Figure 7. Representation of changes in  $\Delta K$  values per different values of  $K$  (number of assumed populations). Maximum point of  $\Delta K$  depicts the optimal  $K$  value.

جدول ۳. میزان عضویت هر یک از جمعیت‌ها در چهار خوشه استنباط شده (k=4)

Table 3. Proportion of membership of each population in each of the four Inferred clusters (K= 4)

Population	F <sub>ST</sub> Value	Inferred Clusters				Number of individuals
		1	2	3	4	
Dorper	0.1111	0.012	0.911	0.043	0.033	27
Merinos	0.0864	0.017	0.035	0.166	0.783	259
Romney	0.0726	0.006	0.029	0.951	0.014	24
Afshari	0.1449	0.955	0.015	0.017	0.014	37
Moghani	0.2521	0.796	0.072	0.049	0.082	27
Ghezel	0.1759	0.784	0.082	0.050	0.083	35

جدول ۴. اختلاف مقادیر فراوانی الی میان هر یک از جفت خوشه‌های استنباط شده در (k=4)

Table 4. Allele frequency divergence between Inferred clusters in K=4

Cluster number	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	-	0.0643	0.0530	0.0540
Cluster 2	0.0643	-	0.0488	0.0474
Cluster 3	0.0530	0.0488	-	0.0353
Cluster 4	0.0540	0.0474	0.0353	-

جدول ۵- میزان عضویت هر یک از نژادهای ایرانی در سه خوشه استنباط شده (K = 3)

Table 5. Proportion of membership of each Iranian sheep breeds in each of the four Inferred clusters (K= 3)

Population	F <sub>ST</sub> Value	Inferred Cluster			Number of individuals
		1	2	3	
Afshari	0.038	0.188	0.631	0.181	37
Moghani	0.107	0.910	0.079	0.012	27
Ghezel	0.298	0.739	0.240	0.020	35

مکانیسم‌های بیولوژیکی که این نژاد قادر به بقای در محیط‌های محلی است را فراهم می‌کند. توسعه سیستم‌های کشاورزی فشرده براساس تعداد کمی از نژادهای تجارتي به تدریج به کاهش تنوع در برخی از جمعیت‌های دام منجر شده است (Wei, 2015).

برخلاف گاو، گوسفند دارای سطح نسبتاً بالایی از تنوع ژنتیکی است (Beynon *et al.*, 2015). مطالعه اخیر روی ۲۸۱۹ فرد از ۷۴ نژاد گوسفند مختلف که از آسیا، آفریقا، کارائیب، آمریکای شمالی و جنوبی، اروپا، استرالیا و نیوزیلند جمع آوری شده است، نشان داد که نژادهای گوسفند اهلی دارای اندازه مؤثر جمعیت (Ne) بالایی از ۱۰۰ تا ۱۳۱۷ هستند درحالی که در نژادهای گاو معاصر مقدار Ne از ۹۹ برای هلشتاین تا ۹۷ برای هر فورد متغیر است (Kijas *et al.*, 2012). درجه بالایی از اشتراک هاپلوتایپ در میان نژادهای گوسفند نیز منشأ مشترک از نژادها را پیشنهاد می‌کند. این در مقایسه با داده های گاو (*Bos indicus*) و خوک (*Sus scrofa*)، که شواهد روشن از رویداد اهلی شدن را نشان می‌دهد (Elsik *et al.*, 2012; Groenen *et al.*, 2009). نژاد مرینوس دارای مقادیر زیادی اختلاط ژنی با بیشتر نژادهای

عوامل مختلف مانند شرایط آب و هوایی متفاوت، تغییرات پوشش گیاهی، نوسانات جوی، شیوع بیماری‌های مختلف، موتاسیون، انتخاب طبیعی و مصنوعی و غیره باعث شده است جمعیت گوسفند به شدت تحت تأثیر قرار گیرد و حیواناتی با تنوع ژنوتیپی و فنوتیپی بالا از قبیل اندازه بدن، شکل، رنگ و تولیدات متفاوت حاصل گردد. بنابراین فشار انتخاب باعث افزایش فراوانی الی برخی جهش‌ها در نواحی خاص ژنوم شده است (Pan *et al.*, 2013). لذا تجزیه و تحلیل توزیع الی‌ها در جمعیت‌ها برای روشن نمودن تنوع ژنتیکی و روابط جوامع اهمیت دارد (Kalinowski, 2004).

مطالعات تاریخی جمعیت شناختی در داخل و بین گونه‌ها همچنین برآورد تعداد نشانگرهای ژنتیکی مورد نیاز برای به‌طور مؤثر پیش‌بینی ارزش‌های پرورش فردی (یعنی انتخاب ژنوم) را امکان‌پذیر می‌سازد و نشان می‌دهد که چگونه این پیش‌بینی‌ها قابل انتقال به سایر جمعیت‌ها خواهد بود (Goddard *et al.*, 2007; Dekkers *et al.*, 2004). درک ساختار و منشأ دام (نژاد) برای موفقیت انتخاب ژنومی، انتخاب به کمک نشانگر و تشخیص QTLها از طریق مطالعات ژنومی، بسیار مهم است. مطالعه نژادهای نادر فرصت شناسایی تنوع و درک

تجزیه و تحلیل DAPC و Structure یکسان و در چهار گروه ژنتیکی طبقه‌بندی شدند. نتایج حاصل از این تحقیق با نتایج به دست آمده توسط Mehri (2006) که تنوع ژنتیکی درون و بین نژادهای گوسفند ایرانی (افشاری، مغانی، قزل، کردی کردستان، مهربان، شال و ماکویی) را با پانزده جفت آغازگر ریزماهورهای بررسی نمود کرده بودند، مطابقت دارد. این محقق گزارش کرده بود که کم‌ترین فاصله ژنتیکی بین افشاری و قزل و بیش‌ترین فاصله ژنتیکی بین ماکویی و مهربان وجود دارد. در تحقیق دیگری توسط Pirkhezraeeian *et al.* (2015) از طریق تجزیه و تحلیل ژنتیکی و فیلوژنتیکی ژنوم میتوکندری نشان دادند که گوسفند نژاد افشاری در بین گونه‌های مربوط به هاپلوتاایپ A که همه آنها مربوط به گوسفندان آسیا می‌باشد بیش‌ترین قرابت را با گوسفندان بلوچی و مغانی ایران دارد. Vahidi *et al.* (2016) تنوع ژنتیکی بین ۱۰ نژاد گوسفند ایرانی (افشاری، مغانی، ماکویی، شال، کلکوئی، زندی، نائینی، تالشی، ترکی- قشقائی و کبودی شیرازی) را مورد مطالعه قرار دادند. کمترین مقدار فاصله بین جمعیت‌های افشاری و مغانی و بیش‌ترین بین شال و زندی بود. با توجه به نتایج بدست آمده می‌توان گفت که نژادهای ایرانی قرابت ژنتیکی بالایی با یکدیگر دارند.

#### نتیجه‌گیری کلی

این تحقیق با بررسی تفاوت فراوانی SNPs در سطح ژنوم گوسفند با استفاده از Ovine SNP 50K و سپس شناسایی ساختار دقیق نژادهای بومی و خارجی انجام شد. آنالیزها با استفاده از دوسری اطلاعات شامل نژادهای بومی و خارجی و با استفاده از داده‌های ژنوتایپ پروژه Sheep Hapmap و با به‌کارگیری آنالیزهای آماری مختلف انجام شد و در نهایت تمایز و تنوع ژنتیکی در بین نژادهای بومی و خارجی شناسایی شدند. در این مطالعه با توجه به این که روش‌های مختلف برای بررسی ساختار جمعیت به‌کار برده شد و همه روش‌های ذکر شده توانست ساختار جمعیت‌های بومی و خارجی را نشان دهد نتایج به دست آمده از این تحقیق و همچنین جدول‌های ۳، ۴ و ۵ حاکی از این

دیگر است. این نتایج بیانگر استفاده گسترده از پدرهای مریئوس در سراسر اروپا پس از قرون وسطی است (Kijass *et al.*, 2012).

در مطالعه‌ای روی هفت نژاد بز ایرانی با ۱۴ نشانگر ریز ماهوره PC اول ۲۵/۳۹ درصد واریانس PC دوم ۱۹/۷۰ درصد واریانس را توجیه کرد، آماره  $F_{ST}$  نیز محاسبه گردید و آنالیز Structure که جزء روش‌های مبتنی بر مدل است، نژادهای بز ایرانی را به سه خوشه مرکزی تقسیم کرد و اختلاطی بین خوشه‌های مرکزی و شمالی مشهود بود در حالی که خوشه غربی استخر ژنی کاملاً مجزایی تشکیل داد (Vahidi *et al.*, 2014). در بررسی ساختار ژنتیکی گاوهای بومی ایران براساس نشانگرهای چندشکلی تک نوکلئوتیدی، روش مبتنی بر مدل به‌کار گرفته شد و نتایج سه خوشه را توجیه کرد (Karimi *et al.*, 2015). در مطالعه دیگری بر روی گاوهای ایران روش‌های تجزیه و تحلیل مؤلفه‌های اصلی (PCA) و تجزیه و تحلیل جداسازی مؤلفه‌های اصلی (DAPC) اجرا شد. روش DAPC توانست با معیار ارزیابی BIC،  $K=3$  بهترین نتیجه را نشان دهد. روش DAPC در بررسی ساختار جمعیتی نسبت به روش PCA به دلیل افزایش واریانس بین جمعیت‌ها و کاهش واریانس درون گروهی و همچنین ارائه تصویر روشنی از ساختار جمعیتی کارآمد بوده و می‌تواند در کنترل کیفیت و تصحیح لایه‌بندی جمعیتی در بررسی‌های ارتباطی جایگزینی برای PCA باشد (Azizi *et al.*, 2017).

نژادهای محلی می‌تواند به‌عنوان یک مجموعه مهم ژن برای صفات سازگاری باشد و می‌تواند برای اطمینان از کارایی برنامه‌های پرورش حیات در آینده جغرافیایی مورد استفاده قرار گیرد که در این راستا مطالعه دقیق تنوع گوسفند و ساختار ژنتیکی در بین سه نژاد ایرانی و مقایسه آن با سه نژاد خارجی انجام شد. این شش جمعیت با استفاده از آنالیز PCA، تجزیه و تحلیل DAPC و Structure مورد مقایسه قرار گرفتند. نتایج به دست آمده، یک هماهنگی خوب بین DAPC و Structure زمانی که همه افراد مورد بررسی به صورت ناشناخته مطالعه شدند هر دو روش نتایج مشابهی را ارائه دادند. خوشه‌های محاسبه شده توسط

نامگذاری شود، همیشه اختلاف نظر وجود داشت و هم‌اکنون هم وجود دارد. اما براساس نتایج به‌دست‌آمده می‌توان گفت حتی با وجود این‌که اگر منشأ مشترکی هم برای نژادهای موجود قائل باشیم ولی با توجه به پرورش این دام به‌مدت طولانی در مناطق مختلف به‌تدریج ساختار ژنتیکی و فنوتیپ آنها از هم فاصله گرفته است، لذا با توجه به خصوصیات فنوتیپی ثابت آنها می‌توان آنها را نژاد نامید.

بود که با وجود افراد خالص در این جمعیت‌ها نژادهای ایرانی از اکوتیپ‌های مختلف و نژادهای مختلف می‌باشند که در مقایسه با نژادهای خارجی مانند مرینوس، رامنی و دورپر، اشتراک ژنتیکی زیادی دارند. اما به‌هرحال در بررسی نژادهای داخلی به‌صورت مجزا مشخص شده که این نژادها نیز با هم متفاوت بوده و در خوشه‌های جداگانه‌ای قرار می‌گیرند. در گذشته در مورد اینکه گوسفند موجود را اکوتیپ، نژاد یا توده

## REFERENCES

1. Askari, N., Abadi, M. M. & Baghizadeh, A. (2011). ISSR markers for assessing DNA polymorphism and genetic characterization of cattle, goat and sheep populations. *Iranian Journal of Biotechnology*, 9(3), 222-9.
2. Azizi, Z., Moradi Shahrababak, H. & Shahrababa, M. (2017). Comparison of PCA and DAPC methods for analysis of Iranian Buffalo population structure using SNPchip90k data. *Iranian Journal of Animal Science*, 48(2), (153-161). (in Farsi)
3. Barendse, W., Harrison, B. E., Bunch, R. J., Thomas, M. B. & Turner, L. B. (2009). Genome wide signatures of positive selection. The comparison of independent samples and the identification of regions associated to traits. *BMC Genomics*, 10, 178.
4. Beynon, S. E., Slavov, G. T., Farré, M., Sunduimijid, B., Waddams, K., Davies, B., Haresign, W., Kijas, J., Macleod, L. M., Jamie Newblod, C., Davies, L. & Larkin, D.M. (2015). Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. *BMC Genetics*, 16(1), 65.
5. Cloete, J. J. E., Cloete, S. W. P., Olivier, J. J. & Hoffman, L. C. (2007). Terminal crossbreeding of Dorper ewes to Ile de France, Merino Landsheep and SA Mutton Merino sires: Ewe production and lamb performance. *Small Ruminant Research*, 69(1-3), 28-35
6. Cloete, S. W. P. & Olivier, J. J. (2010). South African sheep and wool industries. *International Sheep and Wool Handbook*, 95-104.
7. Dekkers, J. C. (2004). Commercial application of marker-and gene-assisted selection in livestock: strategies and lessons. *Journal of animal science*, 82(13-suppl), E313-E328.
8. Elsik, C. G., Tellam, R. L. & Worley, K. C. (2009). The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 324(5926), 522-528.
9. Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, 14(8), 2611-2620.
10. Frankham, R. (1994). Conservation genetics. *Annual Review of Genetics*, 29(1), 305-327.
11. Fernández, M. E., Goszczynski, D. E., Lirón, J. P., Villegas-Castagnasso, E. E., Carino, M. H., Ripoli, M. V., Rogberg-Muñoz, A., Posik, D. M., Peral-García, P. & Giovambattista, G. (2013). Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. *Genetics and Molecular Biology*, 36(2), 185-191.
12. Kalinowski, S. T. (2004). Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics*, 5(4), 539-543.
13. McKay, S. D., Schnabel, R. D., Murdoch, B. M., Matukumalli, L. K., Aerts, J., Coppieters, W. & Mannen, H. (2007). Whole genome linkage disequilibrium maps in cattle. *BMC Genetics*, 8(1), 74.
14. Glaubitz, J. C. (2004). Convert: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes*, 4(2), 309-310.
15. Goddard, M. E. & Hayes, B. J. (2007). Genomic selection. *Journal of Animal breeding and Genetics*, 124(6), 323-330.
16. Groenen, M. A., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F. & Li, S. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, 491(7424), 393-398
17. Holsinger, K. E. & Wallace, L. E. (2004). Bayesian approaches for the analysis of population genetic structure: an example from *Platanthera leucophaea* (Orchidaceae). *Molecular Ecology*, 13(4), 887-894.
18. Jombart, T. (2008). Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403-1405.

19. Jombart, T., Devillard, S. & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11(1), 94.
20. Jombart, T. & Collins, C. (2015). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. *London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling*.
21. Karimi, K., Koshkoiyeh, A. E. & Fuzi, M. A. (2015). Analysis of genetic structure of Iranian indigenous cattle populations using dense single nucleotide polymorphism markers. *Animal Production Research*, 4(3).
22. Kijas, J. W., Townley, D., Dalrymple, B. P., Heaton, M. P., Maddox, J. F., McGrath, A. & Tang, D. (2009). A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLoSOne*, 4(3), e4668.
23. Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Neto, L. R. P., San Cristobal, M. & Paiva, S. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, 10(2), e1001258.
24. Lao, O., Lu, T. T., Nothnagel, M., Junge, O., Freitag-Wolf, S., Caliebe, A. & Holmlund, G. (2008). Correlation between genetic and geographic structure in Europe. *Current Biology*, 18(16), 1241-1248.
25. Lee, C., Abdool, A. & Huang, C.-H. (2009). PCA-based population structure inference with generic clustering algorithms. *BMC bioinformatics*, 10(Suppl 1), S73.
26. Liu, N. & Zhao, H. (2006). A non-parametric approach to population structure inference using multilocus genotypes. *Human Genomics*, 2(6), 1.
27. Mehri, H. (2006). *Study of genetic structure and evolutionary relationships of six sheep breeds of Iran using microsatellite markers*. Master's thesis on livestock breeding, Faculty of Agriculture, Zanjan University. 100 p. (in Farsi)
28. Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G. & McEwan, J. C. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*, 13(1), 10.
29. Pan, D., Zhang, S., Jiang, J., Jiang, L., Zhang, Q & Liu, J. F. (2013). Genome-Wide detection of Selective Signature in Chinese Holstein. *Public Library of Science*, 8, 2-9.
30. Patterson, N., Price, A. L. & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet*, 2(12), e190.
31. Pertoldi, C., Tokarska, M., Wójcik, J. M., Kawalko, A., Randi, E., Kristensen, T. N. & Bendixen, C. (2010). Phylogenetic relationships among the European and American bison and seven cattle breeds reconstructed using the BovineSNP50 Illumina Genotyping BeadChip. *Acta Theriologica*, 55(2), 97-108.
32. Petersen, J. L., Mickelson, J. R., Cothran, E. G., Andersson, L. S., Axelsson, J., Bailey, E. & da Câmara Machado, A. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS One*, 8(1), e54997.
33. Pirkhezraeeian, Z., Tahmorespour, M., Mohammadhashemi, A., Pirani, N. & Azghandi, M. (2015). Genetic and phylogenetic analyses of HVR-I region of mtDNA in Afshari sheep breed. *Biological Biotechnology in Agriculture (Scientific-Research)*, 6 (1), 65-71. (in Farsi)
34. Pometti, C. L., Bessega, C. F., Saidman, B. O. & Vilardi, J. C. (2014). Analysis of genetic population structure in *Acacia caven* (Leguminosae, Mimosoideae), comparing one exploratory and two Bayesian-model-based methods. *Genetics and Molecular Biology*, 37(1),
35. Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, A. & Lareu, M. (2013). An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in Genetics*, 4, 98.
36. Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.
37. Pritchard, J. K., Wen, X. & Falush, D. (2010). Documentation for STRUCTURE software, version 2.3. University of Chicago, Chicago, IL.
38. Purcell, S. (2010). PLINK, version 1.07. URL: <http://pngu.mgh.harvard.edu/~purcell/plink>.
39. Sethuraman A. (2013). *On inferring and interpreting genetic population structure-applications to conservation, and the estimation of pairwise genetic relatedness*. Ph.D. dissertation, Iowa State University, Iowa State.
40. Shojaei, M., Abadi, M. M., Fozi, M. A., Dayani, O., Khezri, A. & Akhondi, M. (2010). Association of growth trait and Leptin gene polymorphism in Kermani sheep. *In Journal of Cell and Molecular Research*, 2 (1), 67-73
41. Teo, Y. Y., Fry, A. E., Clark, T. G., Tai, E. S. & Seielstad, M. (2007). On the usage of HWE for identifying genotyping errors. *Annals of Human Genetics*, 71(5), 701-703.
42. Vahidi, S. M. F., Tarang, A. R., Anbaran, M. F., Boettcher, P., Joost, S., Colli, L., Garcia, J. F. & Ajmone-Marsan, P. (2014). Investigation of the genetic diversity of domestic *Capra hircus* breeds reared within an early goat domestication area in Iran. *Genetics Selection Evolution*, 46(1), 27.

43. Vahidi, S. M. F., Faruque, M. O., Falahati Anbaran, M., Afraz, F., Mousavi, S. M., Boettcher, P. & Negrini, R. (2016). Multilocus genotypic data reveal high genetic diversity and low population genetic structure of Iranian indigenous sheep. *Animal Genetics*, 47(4), 463-470.
44. Vonholdt B. M., Pollinger J. P., Lohmueller K. E., Han E., Parker H. G., Quignon P., Degenhardt J. D., Boyko A. R., Earl D. A., Auton A., Reynolds A., Bryc K., Brisbin A., Knowles J. C., Mosher D. S., Spady T. C., Elkahouloun A., Geffen E., Pilot M., Jedrzejewski W., Greco C., Randi E., Bannasch D., Wilton A., Shearman J., Musiani M., Cargill M., Jones P. G., Qian Z., Huang W., Ding Z. L., Zhang Y. P., Bustamante C. D., Ostrander E. A., Novembre J. & Wayne R. K. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature*, 464, 898-902.
45. Wei, C., Wang, H., Liu, G., Wu, M., Cao, J., Liu, Z. and Liu, C. (2015). Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *BMC Genomics*, 16(1), 194.
46. Zamani, P., Akhondi, M., Mohammadabadi, M. R., Saki, A. A., Ershadi, A., Banabazi, M. H. & Abdolmohammadi, A. R. (2011). Genetic variation of Mehraban sheep using two intersimple sequence repeat (ISSR) markers. *African Journal of Biotechnology*, 10(10), 1812-1817.