



دانشگاه گوارکنندگان

نشریه پژوهش در نشخوارکنندگان

جلد هشتم، شماره چهارم، ۱۳۹۹

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

۳۹-۵۴

DOI: 10.22069/ejrr.2021.17558.1732

## تأثیر تعداد و نوع نشانگرهای ریزماهوره بر برآورد پارامترهای جمعیتی در مطالعات تنوع ژنتیکی

\* منا هاشمی<sup>۱</sup>، قدرت رحیمی میانجی<sup>۲</sup>، محمدحسین بنابازی<sup>۳</sup>، پابلو اوروزکو ترونگل<sup>۴</sup> و

فیلیپو بیسکارینی<sup>۵</sup>

<sup>۱</sup>دانشجو دکتری و <sup>۲</sup>استاد گروه ژنتیک و اصلاح نژاد دام، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری  
<sup>۳</sup>استادیار پژوهشی، بخش پژوهش‌های بیوتکنولوژی، موسسه تحقیقات علوم دامی کشور، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران،  
<sup>۴</sup>استاد ارشد علوم زیستی دانشگاه کاردیف، ولز، <sup>۵</sup>محقق ارشد، شورای تحقیقات ملی (CNR) ایتالیا

تاریخ دریافت: ۹۸/۱۱/۱۶؛ تاریخ پذیرش: ۹۹/۱۰/۸

### چکیده

**سابقه و هدف:** ریزماهوره‌ها نواحی تکراری در دی.ان.آ شامل آرایه‌های همگنی از موتیف‌های مونو، دی، تری، تترا، پنتا و هگزا نوکلئوتیدی با طول کمتر از یک کیلو جفت باز هستند که به صورت غیر تصادفی در سطح ژنوم پراکنده‌اند. تعداد و تراکم ریزماهوره‌ها در گونه‌های مختلف متفاوت است. حتی در گونه‌های بسیار نزدیک به هم مثل انسان و شامپانزه نیز تفاوت‌هایی وجود دارد. فراوانی موتیف‌های ریزماهوره‌ایی نیز در موجودات مختلف، متفاوت و دارای نرخ جهش متفاوتی می‌باشند. در ژنوم پستانداران، ریزماهوره‌های دی نوکلئوتیدی فراوانترین و پس از آنها ریزماهوره‌های مونو و تترا نوکلئوتیدی از وفور بالایی برخوردارند. تکرارهای تری نوکلئوتیدی معمولاً در گیاهان فراوان‌تر هستند. اما بررسی اثر این موتیف‌های ریزماهوره‌ای متفاوت روی برآورد پارامترهای تنوع ژنتیکی و یا ساختار ژنتیکی در جمعیت‌ها بحثی است که کمتر به آن توجه شده است.

**مواد و روش‌ها:** این پژوهش با استفاده از فایل نشانگرهای ریزماهوره‌ای حاصل از ۳۶ نمونه از گوسفندان اهلی و وحشی ایرانی از توالی کل ژنوم گوسفندان به روش توالی‌یابی نسل جدید، تعداد ۱۶۳۹۷۳ نشانگر ریزماهوره را شناسایی کرد. پراکنش نمونه‌های اهلی مربوط به نواحی شمال غرب کشور و نمونه‌های وحشی مربوط به نواحی مرکزی و شمال غرب کشور می‌باشند. پس از طی مراحل مختلف جهت مرتب‌سازی و فیلتراسیون جایگاه‌ها در نرم افزارهای Samtools و VCFtools نشانگرها در چهار فایل جداگانه شامل نشانگرهای دی، تری و تترا نوکلئوتیدی و نیز فایلی حاوی ترکیبی از هر سه نشانگر دسته‌بندی شدند. سپس اسکریپتی در نرم‌افزار R جهت تهیه زیرمجموعه‌های مختلفی شامل ۱۰، ۲۰، ۳۰، ۴۰، ۵۰، ۶۰، ۷۰، ۸۰، ۹۰، ۱۰۰، ۱۵۰، ۲۰۰، ۲۵۰، ۳۰۰، ۴۰۰ و ۵۰۰ نشانگر از هر یک از انواع موتیف‌ها تهیه نوشته شد و شش پارامتر معمول مورد استفاده در مطالعات تنوع ژنتیکی از جمله هتروزیگوسیت مشاهده شده، هتروزیگوسیت مورد انتظار، شاخص تنوع ثنی، شاخص شانون، غنای آللی و ضریب همخونی با استفاده از هر یک از انواع موتیف‌ها و تعداد مختلفی از ریزماهوره‌ها در نرم افزار MSA (نسخه ۴.۰۵) محاسبه شد. برای هر پارامتر ۱۰ تکرار در نظر گرفته شد. در نهایت میانگین و واریانس هر پارامتر در بین ۱۰ تکرار محاسبه و نتایج به صورت نمودارهای باکس پلات در نرم‌افزار R (نسخه ۳.۳.۳) گزارش شد. برای بررسی آماری اختلاف‌هایی

\* نویسنده مسئول: mona.hashemi1012@gmail.com

که در مقدار برآورد پارامترهای مختلف با استفاده از تعداد و انواع مختلف موتیف‌های ریزماهورهای مشاهده شد، از روش تحلیل واریانس برای آزمون فرض مقایسه میانگین زیرمجموعه‌های مختلف در نرم‌افزار R استفاده شد.

**یافته‌ها:** برآورد شش پارامتر با استفاده از تعداد و موتیف‌های مختلف نشانگرهای ریزماهورهایی نتایج متفاوتی را نشان دادند. به طوری که در هر زیر مجموعه، پارامتر برآورده شده با استفاده از موتیف‌های نوع دی مقدار عددی بالاتری را به خود اختصاص داد. همین‌طور در اکثر پارامترهای مورد بررسی بیشترین مقدار برآورده شده برای آن پارامتر با استفاده از ۴۰ نشانگر دی نوکلئوتیدی و کمترین آن با استفاده از ۱۰ نشانگر تری و یا تترا نوکلئوتیدی به دست آمده است. برای بررسی وجود اختلاف معنی‌دار در نتایج به دست آمده از برآورد پارامترها با استفاده از تعداد و انواع مختلف موتیف‌های ریزماهورهای از تکنیک آنالیز واریانس برای مقایسه میانگین‌ها استفاده شد. در مواردی که اجرای مدل نتایج معنی‌داری نشان داد به منظور بررسی دوبه‌دوی زیرمجموعه‌هایی که با یکدیگر اختلاف معنی‌دار نشان دادند از روش مقایسه میانگین توکی استفاده شد.

**نتیجه‌گیری:** نتایج این پژوهش برتری استفاده از موتیف‌های نوع دی نوکلئوتیدی در مطالعات تنوع ژنتیکی بر جمعیت گوسفند را نشان داد. هم‌منظور نتایج این پژوهش لزوم استفاده از حداقل ۵۰ جایگاه ریزماهورهای را برای داشتن برآوردهایی باثبات از پارامترهای جمعیتی در مطالعات تنوع ژنتیکی پیشنهاد می‌دهد.

**واژه‌های کلیدی:** تنوع ژنتیکی، گوسفند، موتیف‌های ریزماهورهایی، نشانگرهای ریزماهور

### مقدمه

آنزیمی دخیل نیست و علت آن به ناپایداری ذاتی دئوکسی نوکلئوتید اسید ریزماهورها ارتباط دارد. به دلیل ماهیت تکراری ریزماهورها، دو رشته کروموزوم می‌توانند در مقابل یکدیگر دچار سرخوردگی شوند. در حین سنتز دئوکسی نوکلئوتید اسید سرخوردگی منجر به افزایش یا کاهش یک یا چند واحد تکراری می‌شود. بسیاری از این حذف/ اضافه‌ها توسط سیستم تعمیر عدم تطابق<sup>۱</sup> اصلاح می‌شود و تنها بخش کوچکی که با این سیستم اصلاح نشده است به صورت جهش‌های ریزماهورهای باقی می‌ماند (۶). در ژنوم انسان ریزماهورها حدوداً ۳ درصد ژنوم را تشکیل می‌دهند (۱۳). تعداد و تراکم ریزماهورها در گونه‌های مختلف متفاوت است. حتی در گونه‌های بسیار نزدیک به هم مثل انسان و شامپانزه نیز تفاوت‌هایی وجود دارد (۲۷). در ژنوم پستانداران، ریزماهورهای دی نوکلئوتیدی فراوانترین

حفظ تنوع ژنتیکی جمعیت‌های دامی اساس بقا دراز مدت بیشتر گونه‌هاست که میبایست بر اساس اطلاعات جامعی از ساختار جمعیت‌ها و تنوع ژنتیکی درون و بین جمعیت‌ها مورد بررسی قرار گیرد (۲۰). در بررسی این تنوع، نشانگرهای ریزماهورها مدت زمان بسیاری است که به عنوان ابزاری سودمند در این مطالعات شناخته شده‌اند (۲۰). ریزماهورها دسته خاصی از دئوکسی نوکلئوتید اسید<sup>۱</sup> تکراری شامل آرایه همگنی از موتیف‌های مونو، دی، تری، تترا، پنتا و هگزا نوکلئوتیدی با طول کمتر از یک کیلو جفت باز هستند که به صورت غیر تصادفی در سطح ژنوم پراکنده‌اند (۱). ریزماهورها نسبت به جهش‌های نقطه‌ای نرخ جهش بالایی ( $10^{-2}$  تا  $10^{-6}$ ) دارند. ناپایداری دی.ان.آ ریزماهورها از یک مکانیسم جهش خاصی بنام سرخوردگی دی.ان.آ<sup>۲</sup> ایجاد می‌شوند (۲۱). در سرخوردگی دئوکسی نوکلئوتید اسید

1. DNA
2. DNA Slippage

3. Mismatch repairing system

و پس از آنها ریزماهوره‌های مونو و تترا نوکلئوتیدی از وفور بالایی برخوردارند. تکرارهای تری نوکلئوتیدی معمولاً در گیاهان فراوان‌تر هستند. دلیل این تفاوت را به مکانیسم‌های مختلف جهش و یا تعمیر موتیف‌های ریزماهوره‌ای در درون گونه‌ها نسبت داده‌اند. همچنین ممکن است در انتخاب موتیف‌های ریزماهوره‌ای مختلف در موجودات مختلف نیز تفاوت وجود داشته باشد (۷). اما بررسی اینکه آیا استفاده از موتیف‌های ریزماهوره‌ای متفاوت اثری بر برآورد نتایج پارامترهای ژنتیکی محاسبه شده در جمعیت‌ها و تعیین ساختار جمعیت‌ها دارد یا خیر، بحثی است که کمتر به آن پرداخته شده است.

نشانی‌های ریزماهوره در بیشتر گونه‌ها از وفور بسیار بالایی برخوردار هستند. اما از آنجا که جداسازی آن‌ها به سرمایه گذاری قابل توجهی نیاز دارد، بسیاری از مطالعاتی که تاکنون از آنها جهت تحلیل تاریخچه تکاملی جمعیت‌ها و گونه‌ها استفاده کرده‌اند، تنها به استفاده از تعداد محدودی از این نشانگرها (معمولاً کمتر از ۵۰) بسنده کرده‌اند (۵). امروزه با پیشرفت‌هایی که در زمینه توالی‌یابی ژنوم صورت گرفته می‌توان به تعداد زیادی از این نشانگرها دسترسی داشت. این مطالعه با دسترسی گسترده به این نشانگرها در ژنوم گوسفند به بررسی تاثیر تعداد و نیز نوع موتیف نشانگرهای ریزماهوره‌ای، بر نتایج حاصل از مطالعه تنوع ژنتیکی و برآورد پارامترهای ژنتیکی می‌پردازد.

نتایج به دست آمده مکان و نوع موتیف را در تمام نمونه‌های مورد بررسی، صحیح نشان داد. تعداد ۱۶۳۹۷۳ جایگاه از انواع ریزماهوره‌های مونو، دی، تری و تترا نوکلئوتیدی در فایل موجود شناسایی شد. در فایلی جداگانه جایگاه‌های تک شکل شمارش و ذخیره شد. نوع موتیف و شمارش تعداد واحد تکراری آن باید در آلل مرجع<sup>۵</sup> و آلل (های) جایگزین<sup>۶</sup> یک شکل باشد.

آماده سازی داده‌ها: برای شروع، صحت موقعیت مکانی و نوع موتیف جایگاه‌ها در فایل داده موجود با تطبیق دادن تعدادی جایگاه به صورت نمونه روی ژنوم مرجع گوسفند (نسخه ۳.۱) "با شماره دسترسی GCA\_000298735.1 در سایت NCBI" مورد بررسی قرار گرفت. این کار در برنامه Samtools (۱۴ و ۱۵) صورت گرفت.

نتایج به دست آمده مکان و نوع موتیف را در تمام نمونه‌های مورد بررسی، صحیح نشان داد. تعداد ۱۶۳۹۷۳ جایگاه از انواع ریزماهوره‌های مونو، دی، تری و تترا نوکلئوتیدی در فایل موجود شناسایی شد. در فایلی جداگانه جایگاه‌های تک شکل شمارش و ذخیره شد. نوع موتیف و شمارش تعداد واحد تکراری آن باید در آلل مرجع<sup>۵</sup> و آلل (های) جایگزین<sup>۶</sup> یک شکل باشد.

نشانی‌های ریزماهوره در بیشتر گونه‌ها از وفور بسیار بالایی برخوردار هستند. اما از آنجا که جداسازی آن‌ها به سرمایه گذاری قابل توجهی نیاز دارد، بسیاری از مطالعاتی که تاکنون از آنها جهت تحلیل تاریخچه تکاملی جمعیت‌ها و گونه‌ها استفاده کرده‌اند، تنها به استفاده از تعداد محدودی از این نشانگرها (معمولاً کمتر از ۵۰) بسنده کرده‌اند (۵). امروزه با پیشرفت‌هایی که در زمینه توالی‌یابی ژنوم صورت گرفته می‌توان به تعداد زیادی از این نشانگرها دسترسی داشت. این مطالعه با دسترسی گسترده به این نشانگرها در ژنوم گوسفند به بررسی تاثیر تعداد و نیز نوع موتیف نشانگرهای ریزماهوره‌ای، بر نتایج حاصل از مطالعه تنوع ژنتیکی و برآورد پارامترهای ژنتیکی می‌پردازد.

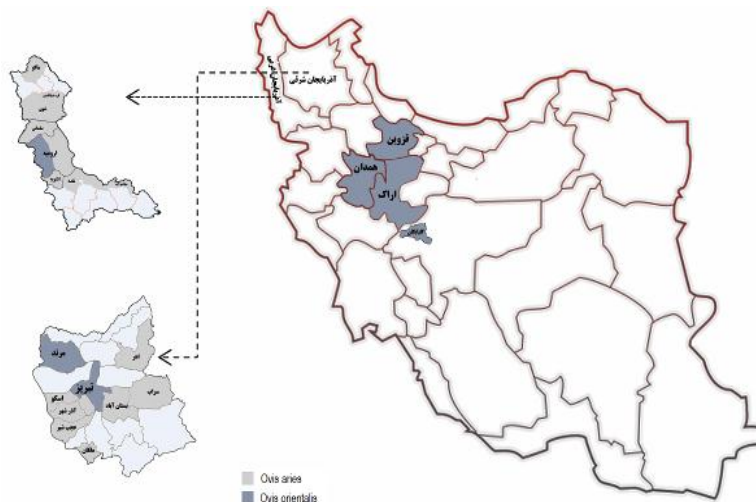
نشانی‌های ریزماهوره در بیشتر گونه‌ها از وفور بسیار بالایی برخوردار هستند. اما از آنجا که جداسازی آن‌ها به سرمایه گذاری قابل توجهی نیاز دارد، بسیاری از مطالعاتی که تاکنون از آنها جهت تحلیل تاریخچه تکاملی جمعیت‌ها و گونه‌ها استفاده کرده‌اند، تنها به استفاده از تعداد محدودی از این نشانگرها (معمولاً کمتر از ۵۰) بسنده کرده‌اند (۵). امروزه با پیشرفت‌هایی که در زمینه توالی‌یابی ژنوم صورت گرفته می‌توان به تعداد زیادی از این نشانگرها دسترسی داشت. این مطالعه با دسترسی گسترده به این نشانگرها در ژنوم گوسفند به بررسی تاثیر تعداد و نیز نوع موتیف نشانگرهای ریزماهوره‌ای، بر نتایج حاصل از مطالعه تنوع ژنتیکی و برآورد پارامترهای ژنتیکی می‌پردازد.

### مواد و روش‌ها

داده‌ها: این مطالعه روی تعداد ۳۶ نمونه از جمعیت گوسفندان ایرانی (۲۰ نمونه از جمعیت گوسفندان اهلی<sup>۱</sup> با شناسه رده بندی ۹۴۴۰ در سایت NCBI و ۱۶ نمونه از گونه وحشی<sup>۲</sup> با شناسه رده بندی

3. Next Generation Sequencing (NGS)
4. Variant Calling
5. Reference allele
6. Alternative allele

1. Ovis aries
2. Ovis orientalis



شکل ۱- موقعیت جغرافیایی مناطق تحت نمونه‌گیری

Figure 1. Geographical location of samples

است، از ادامه آنالیزهای تنوع ژنتیکی کنار گذاشته شدند. در ادامه برای بررسی اثر تعداد نشانگر زیرمجموعه‌های مختلفی از هر یک از انواع موتیف ایجاد شد. برای موتیف‌های نوع دی و mix-loci این زیرمجموعه‌ها به صورت ۱۰، ۲۰، ۳۰، ۴۰، ۵۰، ۶۰، ۷۰، ۸۰، ۹۰، ۱۰۰، ۱۵۰، ۲۰۰، ۲۵۰، ۳۰۰، ۴۰۰ و ۵۰۰ نشانگر و برای موتیف‌های نوع تری و تترا به صورت ۱۰ الی ۱۰۰ نشانگر بود. این زیرمجموعه‌ها با استفاده از اسکریپتی که در نرم افزار R به همین منظور نوشته شد، ایجاد شدند. این اسکریپت در داخل فایل موتیف‌های ریزماهوره‌ای مختلف به صورت تصادفی تعداد مورد نظر نشانگرها را انتخاب میکند. نمونه‌گیری با استفاده از این اسکریپت از داخل هر فایل و برای هر تعداد نشانگر ۱۰ بار تکرار شد.

برآورد پارامترهای تنوع ژنتیکی: شش پارامتر معمول تنوع ژنتیکی شامل هتروزیگوسیت مشاهده شده و مورد انتظار، شاخص اطلاعات شانون، تنوع ژنی نسبی، غنای آلی<sup>۲</sup> و ضریب همخونی<sup>۳</sup> در این پژوهش در

در فایل داده اولیه مشاهده شد که در برخی جایگاه‌ها علاوه بر تکرار موتیف، یک و یا دو نوکلئوتید اضافی دیگر نیز خوانده شده است که این امر منجر به شناسایی اشتباه یک آلل در آن جایگاه و نیز موجب افزایش تعداد آلل در جمعیت می‌شود. این جایگاه‌ها در فایل داده اصلی شناسایی و با تهیه اسکریپتی در پایتون سعی شد تا تعداد ۲۸۴ جایگاه را به جای حذف به ادامه آنالیزها برگردانیم. سپس با توجه به یکی از اهداف این پژوهش مبنی بر مقایسه اثر نوع موتیف بر نتایج حاصل از آنالیزهای تنوع ژنتیکی، سه فایل جداگانه در قالب فرمت خوانش واریانت‌ها<sup>۱</sup> از ریزماهوره‌های دی، تری و تترا نوکلئوتیدی تهیه شد. برای این کار از دستورات مختلف برنامه VCFtools استفاده شد (۴). همچنین فایل دیگری شامل هر سه نوع موتیف که از این به بعد از آن با عنوان mix-loci نام برده می‌شود، تهیه شد. در این بین از آنجا که جایگاه‌های تک شکل اطلاعاتی در زمینه میزان تنوع جمعیت در اختیار قرار نمی‌دهند و همچنین ریزماهوره‌های مونو نوکلئوتیدی نیز از آنجا که روند تکاملی آنها تا به حال به خوبی بررسی نشده

2. Allelic richness

3.  $F_{IS}$ 

1. Variant Call Format (VCF)

معنی دار در مقدار پارامترهای محاسبه شده با تعداد/موتیف‌های نشانگری مختلف را نشان دهد، جهت یافتن زیرمجموعه‌هایی که با یکدیگر اختلاف معنی دار داشتند از آزمون مقایسه میانگین به روش توکی استفاده شد.

### نتایج و بحث

**شمارش تعداد و نوع موتیف‌های ریزماهواره‌ای در جمعیت مورد مطالعه:** تاکنون در مطالعات بسیاری از نشانگرهای ریزماهواره (بدون تاکید بر نوع ریزماهواره استفاده شده) برای بررسی ساختار جمعیت‌ها و برآورد پارامترهای ژنتیکی در گونه‌های مختلفی از موجودات استفاده شده است (۲، ۱۷، ۱۹). طی مطالعات صورت گرفته (۷، ۱۲، ۱۳) فراوانی موتیف‌های ریزماهواره‌ای و نرخ جهش در موجودات متفاوت است. برای مثال مونو ریزماهواره‌ها در ناحیه اینترونی پرمات‌ها و باکتری اشیریشیا کلای فراوانتر از موتیف‌های دی و تری نوکلئوتیدی گزارش شده است. ریزماهواره‌های دی نوکلئوتیدی بیشتر در ژنوم پستاندارانی چون جوندگان گزارش شده است (۲۰) و تری نوکلئوتیدها در مقایسه با سایر موتیف‌های ریزماهواره‌ای بیشتر در ژنوم گیاهان گزارش شده است (۱). در این مطالعه در دو جمعیت گوسفند مورد بررسی از ۱۶۳۹۷۳ جایگاه ریز ماهواره‌ای موجود، مونو ریزماهواره‌ها بیشترین تعداد (۸۶۸۵۴ جایگاه) را به خود اختصاص دادند. این جایگاه‌ها به دلیل اینکه روند تکاملی آنها به خوبی مورد بررسی قرار نگرفته از ادامه آنالیزها کنار گذاشته شدند. حذف ریزماهواره‌های مونونوکلئوتیدی سبب کاهش تعداد نشانگرها به تعداد ۹۷۸۲ جایگاه شد. از تعداد ۵۷۳۹ جایگاه دارای اشکال نوکلئوتید اضافی تعداد ۲۸۴ جایگاه به ادامه آنالیزها برگردانده شدند. در نهایت برآورد پارامترهای ژنتیکی با استفاده از تعداد ۳۹۹۷ ریزماهواره دی، ۱۱۵ ریزماهواره تری و ۱۴۰

نرم افزار MSA (نسخه ۴.۰۵) (۵) محاسبه شدند. در بین نرم افزارهای موجود جهت انجام آنالیزهای تنوع ژنتیکی نرم افزار MSA به دلیل توانایی در آنالیز تعداد زیادی از داده‌های چند آلی انتخاب شد. برای تبدیل فرمت داده‌ها از فرمت خوانش واریانت‌ها به فرمت مورد قبول برنامه MSA اسکریپتی تحت زبان پایتون نوشته شد. این اسکریپت در فایل داده‌های موجود، فرمت خوانش آلل‌ها را از تعداد تکرار موتیف (در فایل‌های با فرمت خوانش واریانت‌ها) به اندازه محصولات واکنش زنجیره‌ای پلیمراز<sup>۱</sup> که فرمت مورد نیاز برنامه MSA می‌باشد تغییر می‌دهد. این نرم افزار پارامترها را براساس هر جمعیت و نیز هر جایگاه گزارش می‌کند. محاسبه پارامترها برای هر یک از ۱۰ تکرار در زیرمجموعه‌های مختلف نشانگری به صورت جداگانه انجام گرفت. سپس میانگین و واریانس هر پارامتر در بین ۱۰ تکرار محاسبه شد و در نهایت نتایج به صورت نمودارهای باکس پلات در نرم افزار R (نسخه ۳.۳.۳) گزارش شدند.

### آنالیز آماری

بررسی آماری اختلاف‌هایی که در مقدار برآورد پارامترهای مختلف با استفاده از تعداد و انواع مختلف موتیف‌های ریزماهواره‌ای به دست آمد به روش تحلیل واریانس تحت مدل ذکر شده در زیر در نرم افزار R نسخه (۳.۳.۳) مورد تجزیه و تحلیل قرار گرفت. این مدل برای هر یک از شش پارامتر مورد بررسی در دو جمعیت اهلی و وحشی به صورت جداگانه اجرا شد:

$$1) y_{ij} = \mu + \alpha_i + e_{ij}$$

که در آن  $y_{ij}$  پارامتر تنوع ژنتیکی محاسبه شده توسط  $i$ امین زیرمجموعه در  $j$ امین تکرار،  $\mu$  میانگین،  $\alpha_i$  تعداد جایگاه/نوع موتیف در زیرمجموعه  $i$  و  $e_{ij}$  اثر تصادفی باقیمانده است. در مواردی که مدل اجرا شده، تفاوت

ریزماهواره تترا نوکلئوتیدی، انجام گرفت. جدول ۱ شمارش تعداد جایگاه‌ها در فایل داده اولیه و نیز در

جدول ۱- تعداد جایگاه‌های ریزماهواره در حین مراحل آماده سازی نمونه‌ها

Table 1. Number of microsatellite loci during sample preparation steps

تترا ریزماهواره‌ها tetra-micros	تری ریزماهواره‌ها tri-micros	دی ریزماهواره‌ها di-micros	مونو ریزماهواره‌ها mono-micros	کل Total	
19779	7061	50279	86854	163973	تعداد اولیه جایگاه‌ها Initial number of loci
18701	6664	41972	56812	124149	جایگاه‌های تک شکل Monomorphic loci
1078	397	8307	30042	39824	جایگاه‌های چند شکل Polymorphic loci
999	228	4512	0	5739	جایگاه‌های چند شکل دارای مشکل آلل خوانی Polymorphic loci with allele calling problem
79	169	3795	30042	34085	جایگاه‌های چند شکل بدون مشکل آلل خوانی Polymorphic loci with no allele calling problem
146	115	4066	30042	34369	جایگاه‌های چند شکل پس از کاربرد اسکریپت Polymorphic loci with using script
140	115	3997	0	4252	جایگاه‌های نهایی Final loci

نوکلئوتیدی خطای "آلل بیش از حد بزرگ" را نشان دادند. این جایگاه‌ها scaffold بودند که از ادامه آنالیزها کنار گذاشته شدند و ادامه محاسبات با تعداد ۴۲۴۳ نشانگر (۳۹۸۸ ریزماهواره دی نوکلئوتیدی) ادامه یافت. محاسبه شش پارامتر تنوع ژنتیکی با استفاده از زیرمجموعه‌های تعریف شده در جمعیت گوسفندان اهلی نتایج را بدین شرح ارائه داد: هتروزیگوسیت مشاهده شده (بیشترین مقدار ۰/۰۷ با ۴۰ نشانگر دی، کمترین مقدار ۰/۰۳۱ با ۱۰ نشانگر تترا)، هتروزیگوسیت مورد انتظار (بیشترین مقدار ۰/۱۶۵ با ۴۰ نشانگر دی، کمترین مقدار ۰/۰۹۴ با ۱۰ نشانگر تترا)، شاخص تنوع نئی (بیشترین مقدار ۰/۲۶۴ با ۴۰ نشانگر دی، کمترین مقدار ۰/۱۷۴ با ۳۰ نشانگر تری)، شاخص شانون (بیشترین مقدار ۰/۲۷۹ با ۴۰ نشانگر دی، کمترین مقدار ۰/۱۵۵ با ۱۰ نشانگر تترا)، غنای آلی (بیشترین مقدار ۱/۸۱ با ۴۰ نشانگر دی، کمترین مقدار ۱/۴۵ با ۱۰ نشانگر تری). ضریب همخوانی (بیشترین مقدار ۰/۵۰۴ با ۳۰ نشانگر mix-loci، کمترین مقدار ۰/۳۳۶ با ۱۰ نشانگر دی).

پارامترهای تنوع ژنتیکی: منطقه خاورمیانه به دلیل نزدیکی به ناحیه هلال بارور<sup>۱</sup> که یکی از مراکز اصلی اهلی سازی میباشد همیشه مورد توجه محققین بوده است. ساختار جمعیت و تنوع ژنتیکی دامهای این منطقه در مطالعات بسیاری مورد تجزیه و تحلیل قرار گرفته است. برای مثال وحیدی و همکاران (۲۰۱۶) تنوع ژنتیکی و انشقاق ۱۰ تیپ مختلف از گوسفندان دنبه دار ایرانی را با استفاده از ۱۸ نشانگر ریزماهواره‌ای مورد بررسی قرار داده و سطح بالایی از تنوع را در بین تیپ‌های مورد بررسی نشان دادند (۲۳). در اکثر مطالعات انجام گرفته بر روی تیپ‌های مختلف گوسفندان ایرانی از تعداد ۴ الی ۲۰ جایگاه ریزماهواره‌ای برای آنالیز تنوع ژنتیکی در درون و بین جمعیت‌های مورد بررسی استفاده شده است که همگی تنوع ژنتیکی بالایی را در جمعیت‌های مورد بررسی خود گزارش کرده‌اند (۸، ۱۶، ۲۳، ۲۴، ۲۵، ۲۷). در این مطالعه در آنالیز پارامترهای تنوع ژنتیکی در جمعیت‌های مورد بررسی تعداد ۹ جایگاه دی

#### 1. Fertile Crescent

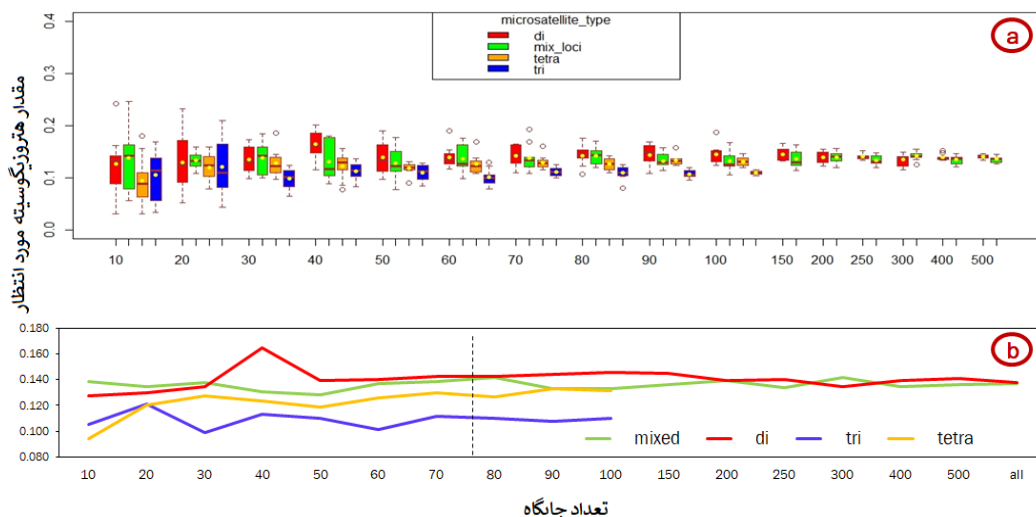
قابل توضیح است. نقاط زرد رنگ روی هر ستون مقدار میانگین ۱۰ تکرار را نشان می‌دهد. نمودارهای b بر اساس میانگین ۱۰ تکرار ترسیم شده‌اند. بر اساس نتایج در استفاده از ۱۰ الی ۴۰ نشانگر، پراکنش زیادی بین تکرارهای مختلف مشاهده می‌شود. در اکثر پارامترهای مورد بررسی بیشترین مقدار برآورده شده برای آن پارامتر با استفاده از ۴۰ نشانگر دی نوکلئوتیدی و کمترین آن با استفاده از ۱۰ نشانگر تری و یا ترا نوکلئوتیدی به دست آمده است. با استفاده از ۵۰ نشانگر نمودارها (b) به سمت ثابت شدن پیش می‌روند و در استفاده از ۸۰ نشانگر به بالا، واریانس‌ها به سمت صفر میل می‌کنند و تقریباً می‌توان در برآورد هر پارامتر به یک مقدار عددی ثابت دست یافت.

مقایسه پارامترهای تنوع ژنتیکی در دو جمعیت اهلی و وحشی در این مطالعه تنوع ژنتیکی بسیار پایینی در هر دو جمعیت را نشان داد که با مطالعات تنوع ژنتیکی پیشین صورت گرفته روی تیپ‌های مختلف گوسفندان ایرانی مغایرت دارد (۸، ۱۶، ۲۳، ۲۴، ۲۵، ۲۷). از آنجا که این مطالعه به تعداد کافی نشانگر دسترسی داشته این میزان تنوع پایین در جمعیت‌های مورد بررسی را می‌توان به تعداد کم نمونه و پراکندگی زیاد در بین نمونه‌ها نسبت داد. در این مطالعه پارامترهای تنوع ژنتیکی با اندکی تفاوت در جمعیت وحشی بالاتر بود که نشان دهنده فشار انتخابی کمتر بر روی این جمعیت می‌باشد. تنها مورد استثنا بالاتر بودن ضریب همخونی در جمعیت اهلی (با استفاده از ریزماهوره‌های تری) بود. وجود تنوع پایینی در جمعیت گوسفندان وحشی در این مطالعه باید زنگ خطری برای دست اندرکاران اصلاح نژاد کشور باشد. شناسایی آلل‌های اختصاصی در این جمعیت‌ها و اقدامات اصلاح نژادی در راستای حفظ تنوع باقیمانده در این جمعیت‌ها می‌تواند موضوع مطالعات آتی روی این جمعیت‌ها باشد.

در جمعیت گوسفندان وحشی نیز نتایج بدین صورت بود: هتروزیگوسیتة مشاهده شده (بیشترین مقدار ۰/۰۷۴ با ۴۰ نشانگر دی، کمترین مقدار ۰/۰۴۵ با ۸۰ نشانگر تری)، هتروزیگوسیتة مورد انتظار (بیشترین مقدار ۰/۱۹۲ با ۴۰ نشانگر دی، کمترین مقدار ۰/۱۱۴ با ۶۰ نشانگر تری)، شاخص تنوع نئی (بیشترین مقدار ۰/۲۶۷ با ۴۰ نشانگر دی، کمترین مقدار ۰/۱۸۳ با ۶۰ نشانگر تری)، شاخص شانون (بیشترین مقدار ۰/۳۲۴ با ۱۰ نشانگر mix-loci، کمترین مقدار ۰/۱۸۵ با ۷۰ نشانگر تری)، غنای آلی (بیشترین مقدار ۱/۹۷ با ۴۰ نشانگر mix-loci، کمترین مقدار ۱/۶۱ با ۲۰ نشانگر تری)، ضریب همخونی (بیشترین مقدار ۰/۴۸۴ با ۷۰ نشانگر تری)، کمترین مقدار ۰/۳۳۱ با ۶۰ نشانگر تری). این نتایج به شکل گرافیکی با استفاده از نمودارهای باکس پلات برای دو پارامتر در هر دو جمعیت به عنوان نمونه آورده شده است (شکل‌های ۲ تا ۵، a و b).

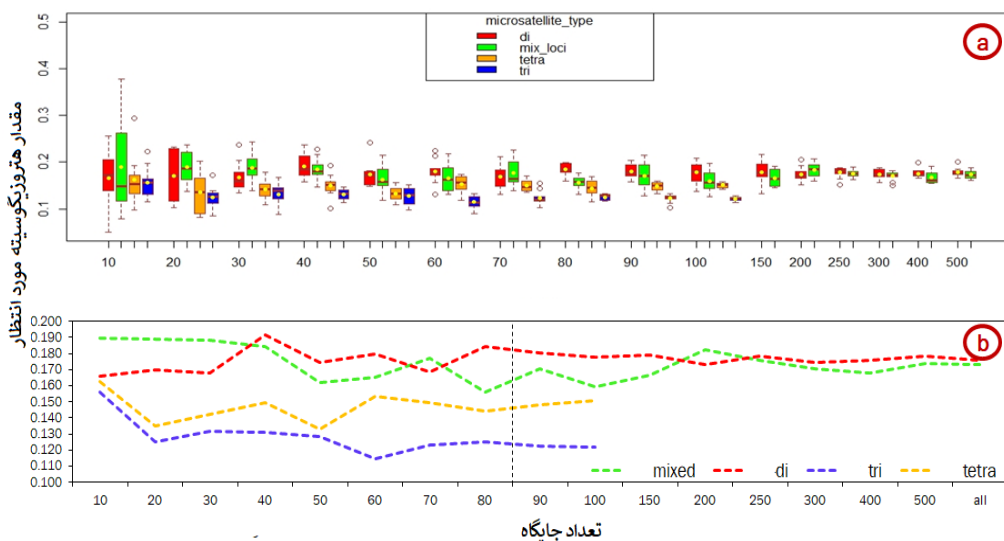
در نمودارهای باکس پلات (شکل‌های ۲-۵، a) هر ستون میزان پراکندگی نتایج به دست آمده از ۱۰ تکرار را نشان می‌دهد. این پراکندگی با افزایش تعداد نشانگر کاهش می‌یابد. مقایسه پلات‌ها نتایج متفاوتی از تخمین یک پارامتر را نشان می‌دهد به طوری که در هر زیر مجموعه، پارامتر برآورده شده با استفاده از موتیف‌های نوع دی مقدار عدد بالاتری را به خود اختصاص داده است (به جز پارامتر ضریب همخونی که با استفاده از موتیف‌های نوع ترا نوکلئوتیدی بالاترین مقدار به دست آمده است). پس از موتیف‌های دی نوکلئوتیدی، استفاده از mix-loci بالاترین مقدار برای هر پارامتر را نشان داد که با توجه به بالاتر بودن نسبت موتیف‌های دی نوکلئوتیدی در مقایسه با تری و ترا نوکلئوتیدی در این فایل، نتایج





شکل ۲- مقایسه مقدار هتروزیگوسیت مورد انتظار با استفاده از انواع و زیرمجموعه‌های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان اهلی (a): پراکنش داده‌ها در ۱۰ تکرار برای هر زیرمجموعه از نشانگرها، (b): مقدار میانگین به دست آمده از ۱۰ تکرار در هر زیرمجموعه از نشانگرها)

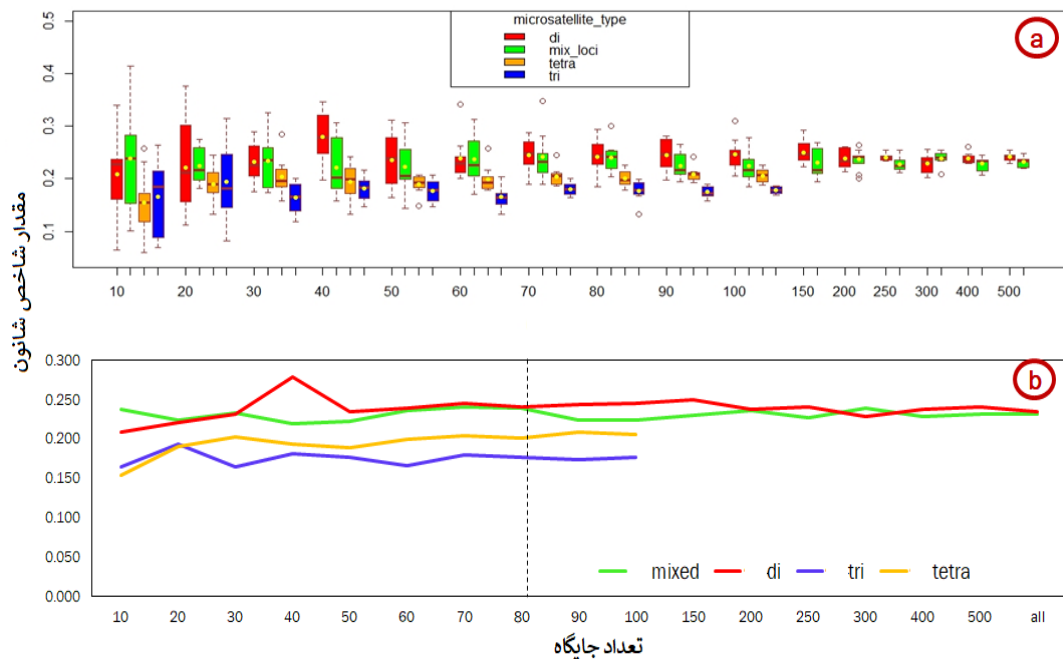
Figure 2. Comparing expected heterozygosity values with different subset of microsatellite types in *Ovis aries*



شکل ۳- مقایسه مقدار هتروزیگوسیت مورد انتظار با استفاده از انواع و زیرمجموعه‌های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان وحشی (a): پراکنش داده‌ها در ۱۰ تکرار برای هر زیرمجموعه از نشانگرها، (b): مقدار میانگین به دست آمده از ۱۰ تکرار در هر زیرمجموعه از نشانگرها)

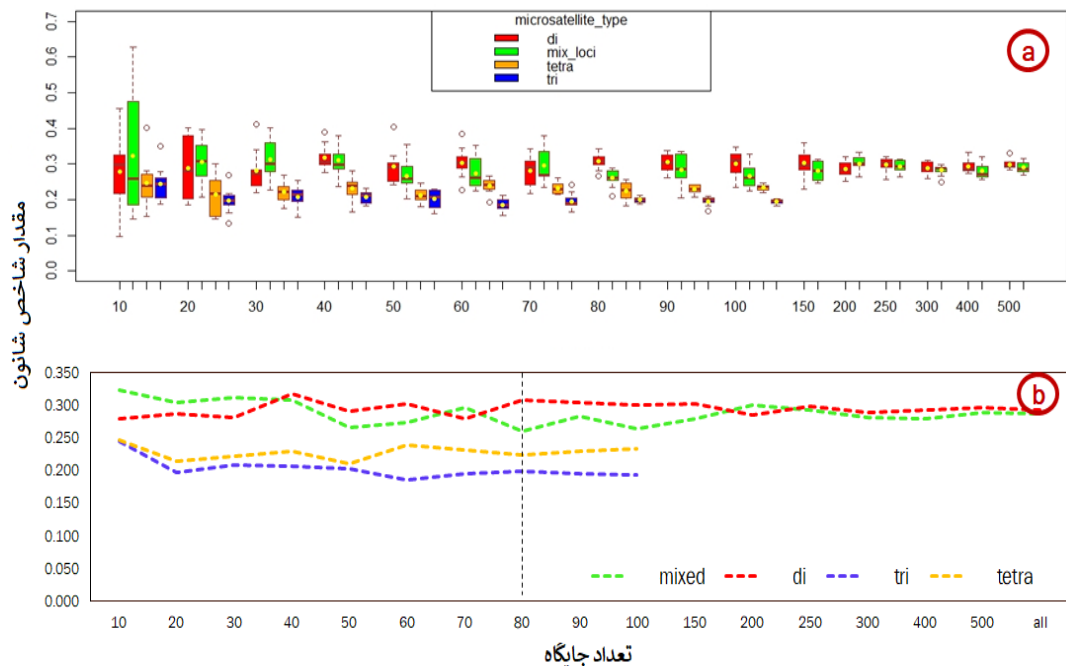
Figure 4. Comparing expected heterozygosity values with different subset of microsatellite types in *Ovis orientalis*





شکل ۴- مقایسه مقدار شاخص شانون با استفاده از انواع و زیرمجموعه‌های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان اهلی (a); پراکنش داده‌ها در ۱۰ تکرار برای هر زیرمجموعه از نشانگرها، (b): مقدار میانگین به دست آمده از ۱۰ تکرار در هر زیرمجموعه از نشانگرها

Figure 3. Comparing Shannon values with different subset of microsatellite types in *Ovis aries*



شکل ۵- مقایسه مقدار شاخص شانون با استفاده از انواع و زیرمجموعه‌های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان وحشی (a); پراکنش داده‌ها در ۱۰ تکرار برای هر زیرمجموعه از نشانگرها، (b): مقدار میانگین به دست آمده از ۱۰ تکرار در هر زیرمجموعه از نشانگرها

Figure 5. Comparing Shannon values with different subset of microsatellite types in *Ovis orientalis*

نشانگرها کمتر باشد این اثر در آنها شدیدتر است. در این مطالعه مشخص شد برای داشتن ساختارهای جمعیتی یکسان از نظر ژنتیکی، به تعداد حداقل ۱۲۰ جایگاه نشانگری نیاز است (۱۷). این دو مطالعه تنها مواردی بودند که گریزی به اثر تعداد نشانگر و ترکیبات مختلف آنها بر روی نتایج پارامترهای جمعیتی داشتند. از اینرو این مطالعه برای اولین بار به بررسی جامع تعداد و نوع موتیف نشانگرهای ریزماهواره بر نتایج پارامترهای تنوع ژنتیکی میپردازد. در این مطالعه پس از محاسبه شش پارامتر مختلف تنوع با استفاده از زیرمجموعه‌های مختلف نشانگری برای بررسی معنی‌دار بودن یا نبودن استفاده از تعداد نشانگر و نوع موتیف از تکنیک آنالیز واریانس برای مقایسه میانگین زیرمجموعه‌های مختلف استفاده شد. نتایج این بررسی به ترتیب در جدول ۲ و ۴ آورده شده است. اعداد داخل جدول مقدار P به دست آمده از این بررسی را نشان می‌دهد.

بررسی معنی‌داری تعداد و نوع موتیف بر برآورد پارامترها: در مطالعه‌ای جهت اندازه‌گیری فاصله ژنتیکی بین چهار جمعیت انسانی با استفاده از ۲۱۳ نشانگر ریزماهواره، کوپر و همکاران (۱۹۹۹) برای اولین بار تفاوت در میانگین طول نشانگرهای ریزماهواره را گزارش کرد. در این مطالعه اثر این تفاوت بر روی معیار فاصله ژنتیکی  $(\delta\mu)^2$  مورد بررسی قرار گرفت که نتایج نهایی تفاوت در مقادیر این معیار با استفاده از طول‌های مختلف ریزماهواره‌ای را نشان داد (۳). در مطالعه دیگری بر روی دروزوفیلا ملانوگاستر مشخص شد استفاده از نشانگرهای ریزماهواره در موقعیت‌های کروموزومی مختلف می‌تواند باعث تفاوت‌های معنی‌دار بین ساختار جمعیت‌ها شود و جالب اینکه هریک از ساختارهای جمعیتی که با استفاده از ترکیبات مختلف نشانگری به دست می‌آیند با دارا بودن احتمالات پسین بالا ( $P \geq 0.95$ ) از نظر آماری دارای پشتوانه بالایی هستند. در این مطالعه مشخص گردید هر اندازه تعداد

جدول ۲- ارزیابی اثر تعداد نشانگر در برآورد پارامترهای ژنتیکی با مقایسه مقدار P به دست آمده

**Table 2. Evaluating the effect of number of loci on genetic parameter estimation comparing P values**

گوسفند وحشی (Ovisorientalis)				گوسفند اهلی (Ovisaries)				جمعیت (population)
mixloci	tetra	tri	di	mixloci	tetra	tri	di	نوع موتیف (Motif type)
								پارامتر (Parameter)
0.771	0.928	0.676	0.945	0.482	0.0074*	0.874	0.766	هتروزیگوسیت مشاهده شده (Het_obs)
0.382	0.371	0.0004***	0.894	0.999	0.0146*	0.808	0.458	هتروزیگوسیت مورد انتظار (Het_exp)
0.365	0.426	0.0002***	0.909	0.996	0.0086***	0.748	0.192	شاخص شانون (Shanon index)
0.902	0.148	0.235	0.433	0.493	0.364	0.682	0.105	شاخص نی (Nei)
0.589	0.594	0.0356*	0.933	0.875	0.094	0.221	0.267	غناي آلی (Allelic richness)
0.449	0.443	0.0027**	0.986	0.172	0.996	0.906	0.291	ضریب همخونی (Fis)

\*، \*\*، \*\*\* به ترتیب نشان دهنده سطح معنی‌داری ( $P < 0.05$ )، ( $P < 0.01$ )، ( $P < 0.001$ )

برآورد برخی پارامترهای ژنتیکی معنی‌دار است. به منظور بررسی دقیق‌تر و یافتن اینکه دقیقاً کدام دسته از نشانگرها با یکدیگر اختلاف معنی‌دار دارند، مقایسه جفتی میانگین‌ها به روش توکی بر روی مواردی که دارای مقدار P معنی‌دار بودند صورت گرفت که نتایج آن در جدول ۳ گزارش شده است.

نتایج نشان می‌دهد با استفاده از موتیف‌های دی (و همین‌طور mix-loci) تفاوت معنی‌داری بین زیرمجموعه‌های نشانگری مختلف در برآورد شش پارامتر مورد نظر مشاهده نمیشود اما استفاده از موتیف‌های تری (در جمعیت گوسفندان وحشی) و ترا نوکلئوتیدی (در جمعیت گوسفندان اهلی) در

جدول ۳- آزمون مقایسه میانگین به روش توکی بین تعداد نشانگرهای مختلف در دو جمعیت گوسفند اهلی و وحشی

**Table 3. Tukey significant between different numbers of loci in Ovis aries and Ovis orientalis population**

گوسفند وحشی (Ovis orientalis)			گوسفند اهلی (Ovis aries)			جمعیت (population)	
tri			tetra			نوع موتیف (motif type)	
غناي آلي (Allelic richness)	ضرب همخونی (F <sub>IS</sub> )	شاخص شانون (Shanon index)	هتروزیگوسیتة مورد انتظار (Het-exp)	شاخص شانون (Shanon index)	هتروزیگوسیتة مورد انتظار (Het-exp)	مشاهده شده (Het-obs)	پارامتر (Parameter)
1.718 <sup>a</sup>	0.449 <sup>a</sup>	0.245 <sup>a</sup>	0.156 <sup>a</sup>	0.155 <sup>b</sup>	0.094 <sup>b</sup>	0.031 <sup>b</sup>	10 loci
1.607 <sup>b</sup>	0.362 <sup>ab</sup>	0.197 <sup>b</sup>	0.125 <sup>b</sup>	0.190 <sup>ab</sup>	0.120 <sup>ab</sup>	0.047 <sup>a</sup>	20 loci
1.650 <sup>ab</sup>	0.355 <sup>b</sup>	0.208 <sup>ab</sup>	0.132 <sup>ab</sup>	0.203 <sup>a</sup>	0.128 <sup>a</sup>	0.044 <sup>ab</sup>	30loci
1.632 <sup>ab</sup>	0.354 <sup>b</sup>	0.206 <sup>b</sup>	0.131 <sup>ab</sup>	0.194 <sup>ab</sup>	0.123 <sup>ab</sup>	0.041 <sup>ab</sup>	40loci
1.631 <sup>ab</sup>	0.335 <sup>b</sup>	0.203 <sup>b</sup>	0.128 <sup>b</sup>	0.189 <sup>ab</sup>	0.118 <sup>ab</sup>	0.043 <sup>ab</sup>	50loci
1.619 <sup>b</sup>	0.331 <sup>b</sup>	0.185 <sup>b</sup>	0.114 <sup>b</sup>	0.199 <sup>a</sup>	0.126 <sup>ab</sup>	0.045 <sup>a</sup>	60loci
1.619 <sup>b</sup>	0.341 <sup>b</sup>	0.196 <sup>b</sup>	0.123 <sup>b</sup>	0.204 <sup>a</sup>	0.130 <sup>a</sup>	0.045 <sup>a</sup>	70loci
1.631 <sup>ab</sup>	0.345 <sup>b</sup>	0.199 <sup>b</sup>	0.125 <sup>b</sup>	0.201 <sup>a</sup>	0.127 <sup>a</sup>	0.047 <sup>a</sup>	80loci
1.622 <sup>ab</sup>	0.346 <sup>b</sup>	0.196 <sup>b</sup>	0.122 <sup>b</sup>	0.208 <sup>a</sup>	0.133 <sup>a</sup>	0.047 <sup>a</sup>	90loci
1.624 <sup>ab</sup>	0.334 <sup>b</sup>	0.194 <sup>b</sup>	0.122 <sup>b</sup>	0.206 <sup>a</sup>	0.131 <sup>a</sup>	0.047 <sup>a</sup>	100loci

\*، \*\*، \*\*\* به ترتیب نشان دهنده سطح معنی داری (P < 0.05)، (P < 0.01)، (P < 0.001)

معنی دار در برآورد پارامترها وجود دارد. در واقع میتوان گفت هرچه تعداد نشانگرها بیشتر می شود تفاوت در میانگینها معنی دارتر می شود (نتایج آورده نشده). همانند ارزیابی اثر تعداد نشانگر، برای بررسی معنی دار نوع موتیف بر برآورد پارامترهای تنوع ژنتیکی نیز از تکنیک آنالیز واریانس برای مقایسه میانگین زیرمجموعه های مختلف موتیفها استفاده شد. این بررسی با تعداد ۱۰، ۲۰، ۵۰، ۷۰ و ۱۰۰ نشانگر در هر دو جمعیت صورت گرفت.

اعداد داخل جدول مقدار پارامترهای محاسبه شده با تعداد مختلفی از نشانگرها را نشان می دهد. بررسی نتایج در جمعیت گوسفندان اهلی وجود اختلاف معنی دار در برآورد پارامترهای نشان داده شده با استفاده از ۱۰ و ۷۰ نشانگر را نشان می دهد. در جمعیت گوسفندان وحشی نیز متفاوت بودن نتایج به دست آمده از تخمین پارامترها در مقایسه ۱۰ و ۶۰ نشانگر قابل مشاهده است. هرچند استثنایی در مورد برخی پارامترها وجود دارد اما به طور قطع میتوان گفت در استفاده از ۱۰ و ۷۰ نشانگر تفاوت های

جدول ۴- ارزیابی اثر نوع موتیف در برآورد پارامترهای ژنتیکی با استفاده از زیرمجموعه های مختلف با مقایسه مقدار P

**Table 4. Evaluating the effect of motif type on genetic parameter estimation using different subsets comparing P values**

گوسفند وحشی (Ovis orientalis)					گوسفند اهلی (Ovis aries)					جمعیت (Population)
100	70	50	20	10	100	70	50	20	10	تعداد جایگاه (Number of loci)
										پارامتر (Parameter)
0.00***	0.000***	0.001***	0.051	0.186	0.00***	0.000***	0.081	0.172	0.18	هتروزیگوسیتة مشاهده شده (Het-obs)
0.00***	0.000***	0.000***	0.002**	0.685	0.00***	0.003**	0.041	0.831	0.263	هتروزیگوسیتة مورد انتظار (Het_exp)
0.00***	0.000***	0.000***	0.0006***	0.302	0.00***	0.000***	0.004	0.468	0.737	شاخص شانون (Shanon index)
0.00***	0.0002***	0.002**	0.010*	0.748	0.00***	0.003**	0.012*	0.281	0.566	شاخص نی (Nei)
0.00***	0.000***	0.000***	0.00***	0.030*	0.00***	0.000***	0.000***	0.017*	0.004**	غناي آلي (Ar)
0.00***	0.000***	0.057	0.226	0.592	0.02*	0.1	0.138	0.494	0.295	ضرب همخونی (F <sub>IS</sub> )

\*، \*\*، \*\*\* به ترتیب نشان دهنده سطح معنی داری (P < 0.05)، (P < 0.01)، (P < 0.001)

نمایان میشود به طوری که با استفاده از ۱۰۰ نشانگر، در برآورد هر شش پارامتر در هر دو جمعیت اختلافات معنی دار مشاهده می شود. بررسی میانگین ها برای یافتن موتیف هایی که اختلاف معنی دار نشان داده اند با آزمون مقایسه میانگین توکی صورت گرفت که نتایج آن در جدول ۵ آورده شده است.

در دو جمعیت مورد بررسی، با استفاده از ۱۰ نشانگر اختلاف معنی دار بین موتیف های ریز ماهواره ای مختلف در برآورد پارامترها مشاهده نشد. تنها برای پارامتر غنای آللی اختلاف های معنی دار مشاهده شد. با افزایش تعداد نشانگرها این اختلاف های معنی دار در مورد پارامترهای بیشتری

جدول ۵- آزمون مقایسه میانگین به روش توکی بین موتیف های مختلف و مقایسه نتایج با استفاده از تعداد نشانگرهای مختلف در جمعیت گوسفندان اهلی و وحشی

Table 5. Tukey significant test between different motif types by comparing results using different number of loci in *Ovis aries* and *Ovis Orientalis* populations

تعداد جایگاه (of loci)	پارامتر (Parameter)	جمعیت (population)									
		گوسفند اهلی ( <i>Ovis.aries</i> )					گوسفند وحشی ( <i>Ovis.orientalis</i> )				
	نوع موتیف (Motif type)	هر روز گوسپسته مورد انتظار (Het_exp)	شخص شانون (Shannon)	شخص نی (Nei)	غنای آللی (AI)	هر روز گوسپسته مشاهده شده (Het_obs)	هر روز گوسپسته مورد انتظار (Het_exp)	شخص شانون (Shannon)	شخص نی (Nei)	غنای آللی (AI)	هر روز گوسپسته مشاهده شده (Het_obs)
10	mixloca	-	-	-	1.705 <sup>a</sup>	-	-	-	-	1.973 <sup>a</sup>	-
	di	-	-	-	1.625 <sup>ab</sup>	-	-	-	-	1.910 <sup>b</sup>	-
	tri	-	-	-	1.447 <sup>b</sup>	-	-	-	-	1.718 <sup>a</sup>	-
	tetra	-	-	-	1.484 <sup>b</sup>	-	-	-	-	1.663 <sup>a</sup>	-
20	mixloca	-	-	-	1.675 <sup>ab</sup>	-	-	-	-	1.893 <sup>a</sup>	0.255 <sup>a</sup>
	di	-	-	-	1.706 <sup>a</sup>	-	-	-	-	1.962 <sup>a</sup>	0.231 <sup>ab</sup>
	tri	-	-	-	1.545 <sup>ab</sup>	-	-	-	-	1.607 <sup>b</sup>	0.204 <sup>b</sup>
	tetra	-	-	-	1.530 <sup>b</sup>	-	-	-	-	1.638 <sup>b</sup>	0.203 <sup>b</sup>
50	mixloca	-	-	-	1.709 <sup>a</sup>	0.210 <sup>ab</sup>	0.222 <sup>ab</sup>	-	-	1.954 <sup>a</sup>	0.244 <sup>a</sup>
	di	-	-	-	1.717 <sup>a</sup>	0.240 <sup>a</sup>	0.236 <sup>a</sup>	-	-	1.833 <sup>a</sup>	0.240 <sup>a</sup>
	tri	-	-	-	1.507 <sup>b</sup>	0.195 <sup>b</sup>	0.177 <sup>b</sup>	-	-	1.631 <sup>b</sup>	0.201 <sup>b</sup>
	tetra	-	-	-	1.542 <sup>b</sup>	0.200 <sup>b</sup>	0.189 <sup>b</sup>	-	-	1.636 <sup>b</sup>	0.205 <sup>b</sup>
70	mixloca	-	-	-	1.743 <sup>a</sup>	0.226 <sup>ab</sup>	0.241 <sup>a</sup>	0.139 <sup>a</sup>	-	1.917 <sup>a</sup>	0.250 <sup>a</sup>
	di	-	-	-	1.738 <sup>a</sup>	0.245 <sup>a</sup>	0.245 <sup>a</sup>	0.143 <sup>a</sup>	-	1.895 <sup>a</sup>	0.236 <sup>a</sup>
	tri	-	-	-	1.514 <sup>b</sup>	0.196 <sup>b</sup>	0.180 <sup>b</sup>	0.112 <sup>b</sup>	-	1.619 <sup>b</sup>	0.196 <sup>a</sup>
	tetra	-	-	-	1.560 <sup>b</sup>	0.213 <sup>ab</sup>	0.204 <sup>b</sup>	0.130 <sup>ab</sup>	-	1.651 <sup>b</sup>	0.228 <sup>a</sup>
100	mixloca	0.394 <sup>bc</sup>	0.230 <sup>b</sup>	0.265 <sup>b</sup>	1.666 <sup>a</sup>	0.219 <sup>ab</sup>	0.224 <sup>ab</sup>	0.133 <sup>a</sup>	0.051 <sup>b</sup>	1.853 <sup>b</sup>	0.230 <sup>b</sup>
	di	0.421 <sup>ab</sup>	0.258 <sup>a</sup>	0.301 <sup>a</sup>	1.736 <sup>a</sup>	0.237 <sup>a</sup>	0.246 <sup>a</sup>	0.146 <sup>a</sup>	0.058 <sup>a</sup>	1.958 <sup>a</sup>	0.258 <sup>a</sup>
	tri	0.334 <sup>c</sup>	0.192 <sup>c</sup>	0.194 <sup>d</sup>	1.517 <sup>b</sup>	0.190 <sup>c</sup>	0.178 <sup>c</sup>	0.110 <sup>b</sup>	0.044 <sup>c</sup>	1.624 <sup>c</sup>	0.192 <sup>c</sup>
	tetra	0.470 <sup>a</sup>	0.229 <sup>b</sup>	0.234 <sup>c</sup>	1.558 <sup>b</sup>	0.214 <sup>b</sup>	0.206 <sup>b</sup>	0.131 <sup>a</sup>	0.047 <sup>bc</sup>	1.652 <sup>c</sup>	0.229 <sup>b</sup>

\*موردی که با (-) نشان داده شده میانگین هایی هستند که ارزش P معنی دار در مدل برای آنها گزارش نشده و بنابراین تحت آزمون مقایسه میانگین قرار نگرفته اند. \*\*میانگین هایی که دارای حروف مشابه هستند از نظر آماری دارای اختلاف معنی دار نیستند (در سطح اطمینان ۹۵ درصد).

نشانگرها را نشانگرهای دی نوکلئوتیدی تشکیل می دهند. همین طور در مورد پارامترهای برآورد شده با استفاده از موتیف های نوع تری و تترا نیز اختلافات معنی دار کمی مشاهده شد. بیشترین میزان اختلاف مربوط به مقایسه میانگین پارامترها با استفاده از

در اکثر مقایسات جفتی بین میانگین ها، پارامترهای برآورد شده با استفاده از موتیف های نوع دی و mix-loci با یکدیگر اختلاف معنی دار ندارند که دلیل آن را می توان به یکسان بودن نسبی فایل داده آنها نسبت داد چرا که در فایل داده mix-loci اکثر

استفاده از حداقل ۵۰ جایگاه ریزماهوره‌ای را برای داشتن برآوردهایی باثبات از پارامترهای جمعیت در مطالعات تنوع ژنتیکی پیشنهاد می‌دهد. تعداد و نوع موتیف ریزماهوره‌ها احتمالاً در برآورد دیگر پارامترها از جمله پیش‌بینی وقایع دموگرافیک رخ داده در جمعیت‌ها تاثیر گذار خواهد بود که می‌تواند موضوع مطالعات آتی باشد. نتایج این پژوهش لزوم بررسی و اعمال دقت بیشتر روی نتایج حاصل از نرم افزارهایی که واریانت‌های ژنتیکی را خوانش می‌کنند (مانند نرم‌افزار lobSTR) نشان می‌دهد. خوانش اشتباه توالی ریزماهوره‌ها در چنین نرم‌افزارهایی می‌تواند سبب تشخیص نادرست آلل‌ها و در نتیجه اعلام نادرست تعداد آلل در محاسبات آتی شود که به نوبه خود بر نتایج حاصل از مطالعات ژنتیکی در جمعیت‌ها تاثیر بسزایی دارد. نتایج به دست آمده از این پژوهش سطح پائینی از تنوع ژنتیکی در دو جمعیت از گوسفندان اهلی و وحشی مورد مطالعه را نشان می‌دهد که باید هشدار برای مسئولین طرح‌های حفاظت ژنتیکی در گوسفندان وحشی و هم چنین طرح‌های اصلاح نژادی در گوسفندان اهلی در داخل کشور باشد.

#### منابع

1. Abdurakhmonov, I.Y. 2016. Introduction to microsatellites: basics, trends and highlights. Abdurakhmonov, I.Y (eds), Microsatellite Markers, Intech open. Uzbekistan. 4-16.
2. Charoensook, R., Gatphayak, K., Brenig, B. and Knorr, C. 2019. Genetic diversity analysis of Thai indigenous pig population using microsatellite markers. Asian-Australas Journal of Animal Science. 32(10): 1491-1500.
3. Cooper, G., Amos, W., Bellamy, R., Siddiqui, M.R., Frodsham, A., Hill, A. V. and Rubinsztein, D.C. 1999. An empirical exploration of the  $(\delta\mu)^2$  genetic distance for 213 human microsatellite markers. American Journal of Human Genetics. 65(4): 1125-1133.

موتیف‌های نوع دی با تترا و دی با تری است و این به این معناست که در برآورد پارامترهای تنوع ژنتیکی استفاده از ریزماهوره‌های تری یا تترا تفاوت آنچنانی بر روی نتایج نهایی نخواهد داشت اما استفاده از ریزماهوره‌های دی نوکلئوتیدی قطعاً نتایج حاصله را تحت تاثیر قرار داده و برآوردهای دقیق تری ارائه می‌دهد. همینطور باید به این نکته دقت کرد که با به کارگیری تعداد بیشتری از نشانگرها، اختلاف معنی‌دار بیشتری بین داده‌ها قابل رویت است.

#### نتیجه‌گیری نهایی

نتایج مطالعات تنوع ژنتیکی در جمعیت‌ها می‌تواند تحت تاثیر تعداد جایگاه‌های مختلف و نیز موتیف‌های مختلف نشانگرهای ریزماهوره قرار گیرد. نتایج به دست آمده از این پژوهش تفاوت معنی‌دار در استفاده از موتیف‌های نوع دی در برآورد پارامترهای ژنتیکی را نشان داد. از آنجایی که این ریزماهوره‌ها در سطح ژنوم فراوانی بالاتری دارند می‌توانند در مطالعات بررسی ساختار جمعیتی روی گونه‌های دامی مناسب‌تر باشند. نتایج این پژوهش همچنین لزوم

4. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, Handsaker, R.E., Lunter, G., T.Marth, G., T., Sherry, S., McVean, G. and Durbin, R. 2011. The variant call format and VCFtools. Bioinformatics. 27(15): 2156-2158.
5. Dieringer, D. and Schlötterer, C. 2003. microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes. 3(1): 167-169.
6. Eisen, J.A. and Hanawalt, P.C. 1999. A phylogenomic study of DNA repair genes. proteins and processes. Mutation research. 435(3): 171-213.
7. Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics. 5(6): 435-445.

8. Esmailkhanian S.B.M. 2006. Genetic variation within and between five Iranian sheep populations using microsatellite markers. *Pakistan Journal of Biological Sciences*. 9: 2488-2492.
9. Gibbs, R.A., Weinstock, G.M., Metzker, M.L., Muzny, D.M., Sodergren, E.J., Scherer, S. and Collins, F. 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*. 428(6982): 493-521.
10. Gymrek, M., Golan, D., Rosset, S. and Erlich, Y. 2012. lobSTR: A short tandem repeat profiler for personal genomes. *Genome Research*. 22(6): 1154-1162.
11. Hentschel, C.C. 1982. Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. *Nature*. 295(5851): 714-716.
12. Katti, M.V., Ranjekar, P.K. and Gupta, V.S. 2001. Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution*. 18(7): 1161-1167.
13. Lagercrantz, U. and Andersson, L. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research*. 21(5): 1111-1115.
14. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M. and Szustakowki, J. 2001. Initial sequencing and analysis of the human genome. *Nature*. 409(6822): 860-921.
15. Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 27(21): 2987-2993.
16. Nanekarani, S.H., A.C. and Amirmozafari, N. 2011. Genetic analysis of Karakul sheep breed using microsatellite markers. *African Journal of Microbiology Research*. 5: 703-707.
17. Orozco-terWengel, P., Corander, J. and Schlötterer, C. 2011. Genealogical lineage sorting leads to significant, but incorrect Bayesian multilocus inference of population structure. *Journal of Molecular Ecology*. 20(6): 1108-1121.
18. Schlötterer, C. and Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*. 20(2): 211-215.
19. Seilsuth, S., Seo, J.H., Kong, H.S. and Jeon, G.J. 2016. Microsatellite analysis of the genetic diversity and population structure in dairy goats in Thailand. *Asian-Australas Journal of Animal Science*. 29(3): 327-332.
20. Sheriff, O. and Alemayehu, K. 2018. Genetic diversity studies using microsatellite markers and their contribution in supporting sustainable sheep breeding programs: Review: *Cogent Food and Agriculture*. 4(1).
21. Trivedi, S. 2004. Microsatellites (SSRs): puzzles with puzzle. *Indian Journal of Biotechnology*. 3: 331-347.
22. Tautz, D. and Schlötterer, C. 1994. Simple sequences. *Current Opinion in Genetics and Development*. 4(6): 832-837.
23. Vahidi, S.M.F., Faruque, M.O., Falahati Anbaran, M., Afraz, F., Mousavi, S.M., Boettcher, P., Joost, S., Han, J.L., Colli, L., Periasamy, K., Negrini, R. and Ajmone-Marsan, P. 2016. Multilocus genotypic data reveal high genetic diversity and low population genetic structure of Iranian indigenous sheep. *Journal of Animal Genetic*. 47: 463-470.
24. Vajed Ebrahimi, M.T., Mohammad Abadi, M.R. and Esmailzadeh, A.K. 2016. Analysis of genetic diversity in five Iranian sheep population using microsatellite markers. *Journal of Agriculture Biotechnology*. 7 (4).
25. Vajed Ebrahimi, M.T., Mohammadabadi, M. and Esmailzadeh, A. 2017. Using microsatellite markers to analyze genetic diversity in 14 sheep types in Iran. *Arch. Journal of Animal Breeding*. 60(3): 183-189.
26. Webster, M.T., Smith, N.G. and Ellegren, H. 2002. Microsatellite evolution inferred from human-chimpanzee genomic sequence alignments. *Proc. . Proceedings of the National Academy of Sciences* 99(13): 8748-8753.
27. Zahedi, Z., Esmaelkhanian, S. and Torshizi, R.V. 2007. Microsatellite variation in one breed of Iranian sheep with 12 markers. *Pakistan journal of biological sciences: PJBS*. 10(24): 4455-4460.



## The Effect of microsatellite number and motif type on estimation of population parameters in genetic diversity studies

\*M. Hashemi<sup>1</sup>, G. Rahimi Mianji<sup>2</sup>, M.H. Banabazi<sup>3</sup>, P. Orozro-ter-Wengel<sup>4</sup> and F. Biscarini<sup>5</sup>

<sup>1</sup>Ph.D. Student, and <sup>2</sup>Professor, Dept. of Genetics and animal breeding, Faculty of Animal Science and Fisheries, Sari Agriculture and Natural Resources university, <sup>3</sup>Assistant Prof., Dept. of Biotechnology, Animal Science Research Institute of Iran (ASRI), Agricultural Research, Education & Extension Organization (AREEO), Karaj, Iran, <sup>4</sup>Senior Lecturer, Cardiff University, Wales,

<sup>5</sup>Senior Research Scientist, National Research Council (CNR), Italy

Received: 02/05/2020; Accepted: 12/28/2020

### Abstract

**Background and objective:** Microsatellites are repetitive regions in DNA including a homogeneous array of mono, di-, tri-, tetra-, penta-, and hexa-nucleotides with a length of less than 1 Kbp which are non-randomly distributed in the genome. The number and density of microsatellites are varying within species even in very close species such as humans and chimpanzees. The frequency of microsatellite motifs and their mutation rate are reported differently in various organisms. Di-nucleotide microsatellites are the most abundant motifs followed by mono and tetra microsatellite motifs in mammalian genomes. Tri-nucleotide microsatellites are more frequent in plants. However, the effect of different microsatellite motifs on genetic diversity or population structural parameters is a topic that has received less attention.

**Material and methods:** In the present study, using the 36 VCf file of microsatellite markers extracted from the whole genome of Iranian sheep “*Ovis aries* and *Ovis orientalis*” by NGS, the total number of 163973 microsatellite markers were detected. The distribution of *Ovis aries* samples is from the north-west part of Iran and *Ovis orientalis* samples belongs to the central part and north-west of Iran. After rearrangement and filtration on the data file using Samtools and VCFtools software, we classified the whole markers on four different motif types including di- tri- and tetra-nucleotide microsatellites and a file include all three microsatellite types. Several subsets including 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 400 and 500 markers were generated from each microsatellite motif types using an R script. Six common genetic diversity parameters including observed and expected heterozygosity, Nei diversity index, Shanon index, Allelic richness, and  $F_{IS}$  were calculated for each different subset of number and motif type of microsatellites in MSA (V.4.05) software. 10 replications were considered for each parameter. The mean and variance were calculated among 10 replications and results were represented by boxplots using R (v.3.3.3). The statistical investigation of parameter estimation differences using different microsatellite number and motif types were analyzed using ANOVA for testing the hypothesis of equality of means in R (v.3.3.3).

**Results:** Estimation of all six parameters revealed various results using a different number of loci as well as motif types. Additionally, the results revealed higher values for parameters estimated with di microsatellite motifs compared to others. In addition, the highest and lowest values for most parameters were obtained by 40-di and 10-tri/tetra microsatellites respectively.

\*Corresponding author; mona.hashemi1012@gmail.com



The statistical significance on findings of parameter values using different number/motif of microsatellite markers were analyzed using ANOVA in R (v.3.3.3). In cases where the implication of the model showed significant results, Tukey Honestly Significant Difference test was used to test pairwise contrasts between different subsets.

**Conclusion:** Our results propose a better application of di microsatellites for genetic diversity studies in sheep populations. Moreover, results showed that for stable estimation of population parameters in genetic diversity studies a minimum of 50 microsatellite loci are needed.

**Keywords:** Genetic diversity, Microsatellite marker, Microsatellite motifs, Sheep