ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/137544/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Hashemi, Mona, Rahimi Mianji, Godrat, Hossein Banabazi, Mohammad, Orozco ter-Wengel, Pablo and Biscarini, Filippo 2021. The Effect of microsatellite number and motif type on estimation of population parameters in genetic diversity studies i. Journal of Ruminant Research 8 (4) , pp. 39-54. 10.22069/EJRR.2021.17558.1732

Publishers page: https://doi.org/10.22069/EJRR.2021.17558.1732

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

***منا هاشمی[1]، قدرت رحیمی میانجی[2]، محمدحسین بنابازی[3]، پابلو اوروزکو ترونگل[4]، فیلیپو بیسکارینی[5]**
**[1]دانشجو دکترا و [2]استاد گروه ژنتیک و اصلاح نژاد دام، دانشکده علوم دامی وشیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری**
**[3]استادیار پژوهشی، بخش پژوهشهای بیوتکنولوژی، موسسه تحقیقات علوم دامی کشور،سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران**
[4]استاد ارشد علوم زیستی دانشگاه کاردیف، ولز
[5]محقق ارشد، شورای تحقیقات ملی (CNR) ایتالیا

***Mona Hashemi[1], Godrat Rahimi Mianji[2], Mohammad Hossein Banabazi[3], Pablo Orozro-ter-Wengel[4], Filippo Biscarini[5]**
[1]PhD student., and [2]Professor, Dept of genetics and animal breeding, Faculty of animal science and fisheries, Sari agriculture and natural resources university, [3]**Assistant Professor, Department of Biotechnology, Animal Science Research Institute of IRAN (ASRI), Agricultural Research, Education & Extension Organization (AREEO), Karaj, IRAN,** [4]Senior Lecturer, Cardiff University, Wales, [5]Senior Research Scientist, National Research Council (CNR), Italy

*نویسنده مسئول: mona.hashemi1012@gmail.com

# تاثیر تعداد و نوع نشانگرهای ریزماهواره بر برآورد پارامترهای جمعیتی درمطالعات تنوع ژنتیکی


**چکیده**

سابقه و هدف

ریزماهواره‌ها نواحی تکراری در دی.ان.آشامل 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 ریزماهواره‌ها در گونه‌های مختلف متفاوت است. حتی در گونه های بسیار نزدیک به هم مثل انسان و شامپانزه نیز تفاوت هایی وجود دارد. فراوانی موتیف‌های ریزماهواره‌ایی نیز در موجودات مختلف، متفاوت و دارای نرخ جهش متفاوتی می‌باشند. در ژنوم پستانداران، ریزماهواره‌های دی نوکلئوتیدی فراوانترین و پس از آنها ریزماهواره های مونو و تترا نوکلئوتیدی از وفور بالایی برخوردارند. تکرارهای تری نوکلئوتیدی معمولاً در گیاهان فراوان‌تر هستند. اما بررسی اثراین موتیف‌های ریزماهواره‌ای 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 ریزماهواره‌ای حاصل از 36 نمونه از گوسفندان اهلی و وحشی ایرانی از توالی کل ژنوم گوسفندان به روش توالی یابی نسل جدید، تعداد 163973 نشانگر ریزماهواره را شناسایی شد. پراکنش نمونه های اهلی مربوط به نواحی شمال غرب کشور و نمونه های وحشی مربوط به نواحی مرکزی و شمال غرب کشور می باشند. پس از طی مراحل مختلف جهت مرتب سازی 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 هرسه نشانگر دسته بندی شدند. سپس اسکریپتی در نرم افزار R جهت تهیه زیرمجموعه‌های مختلفی شامل10، 20، 30، 40، 50، 60، 70، 80، 90، 100، 150، 200، 250، 300، 400 و 500 نشانگر از هر یک از انواع موتیف‌ها تهیه نوشته شد و شش پارامتر معمول مورد استفاده در مطالعات تنوع ژنتیکی از جمله هتروزیگوسیته مشاهده شده، هتروزیگوسیته مورد انتظار، شاخص تنوع نئی، شاخص شانون، غنای آللی و ضریب همخونی با استفاده از هر یک از انواع موتیف‌ها و تعداد مختلفی از ریزماهواره‌ها در نرم افزار MSA (نسخه 4.05) محاسبه شد. برای هر پارامتر 10 تکرار در نظر گرفته شد. در نهایت میانگین و واریانس هر پارامتر در بین 10 تکرار محاسبه و نتایج به صورت نمودارهای باکس پلات در نرم افزار R (نسخه 3.3.3 ) گزارش شد. برای بررسی آماری اختلاف‌هایی که در مقدار برآورد پارامترهای مختلف با استفاده از تعداد و انواع مختلف موتیف‌های ریزماهواره‌ای مشاهده شد، از روش تحلیل واریانس برای آزمون فرض مقایسه میانگین زیرمجموعه‌های مختلف در نرم افزار R استفاده شد.

یافته‌ها

برآورد شش پارامتر با استفاده از تعداد و موتیف‌های مختلف نشانگرهای ریزماهواره‌ایی نتایج متفاوتی را نشان دادند. به طوریکه 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 40 نشانگر دی نوکلئوتیدی و کمترین آن با استفاده از 10 نشانگر تری و یا تترا نوکلئوتیدی به دست آمده است. برای بررسی وجود اختلاف معنی دار در نتایج به دست آمده از برآورد پارامترها با استفاده از تعداد و انواع مختلف موتیف-های ریزماهواره‌ای از تکنیک آنالیز واریانس برای مقایسه میانگین‌ها استفاده شد. در مواردی که اجرای مدل نتایج معنی داری نشان داد به منظور بررسی دوبه‌دوی زیرمجموعه‌هایی که با یکدیگر اختلاف معنی دار نشان دادند از روش مقایسه میانگین توکی استفاده شد.

نتیجه گیری

نتایج این پژوهش برتری استفاده از موتیف های نوع دی نوکلئوتیدی در مطالعات تنوع ژنتیکی بر جمعیت گوسفند را نشان داد. همینطور نتایج این پژوهش لزوم استفاده از حداقل 50 جایگاه ریزماهوره‌ای را برای داشتن برآوردهایی باثبات از پارامترهای جمعیتی در مطالعات تنوع ژنتیکی پیشنهاد میدهد.



**کلمات کلیدی:** تنوع ژنتیکی، موتیف های ریزماهواره‌ایی، نشانگرهای ریزماهواره، گوسفند


**مقدمه**

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 (20). در بررسی این تنوع، نشانگرهای ریزماهواره مدت زمان بسیاری است که به عنوان ابزاری سودمند در این مطالعات شناخته شده‌اند (20). ریزماهواره‌ها دسته خاصی از دی.ان.آ[1] 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(1). ریزماهواره‌ها نسبت به جهش‌های نقطه‌ای نرخ جهش بالایی($10^{-2}$ تا $10^{-6}$) دارند. ناپایداری دی.ان.آ ریزماهواره‌ها از یک مکانیسم جهش

1-DNA

خاصی بنام سرخوردگی دی.ان.آ[1] ایجاد می‌شوند(21). در سرخوردگی دی.ان.آ آنزیمی دخیل نیست و علت آن به ناپایداری ذاتی دی.ان.آ ریزماهواره‌ها ارتباط دارد. به دلیل ماهیت تکراری ریزماهواره ها، دو رشته کروموزوم می‌توانند در مقابل یکدیگر دچار سرخوردگی شوند. در حین سنتز دی.ان.آ سرخوردگی منجر به افزایش یا کاهش یک یا چند واحد تکراری می‌شود. بسیاری از این حذف/اضافه‌ها توسط سیستم تعمیری عدم تطابق[2] اصلاح می‌شود و تنها بخش کوچکی که با این سیستم اصلاح نشده است به صورت جهش‌های ریزماهواره‌ای باقی می‌مانند (6). در ژنوم انسان ریزماهواره‌ها حدوداً 3 درصد ژنوم را تشکیل میدهند (13). تعداد و تراکم ریزماهواره‌ها در گونه‌های مختلف متفاوت است. حتی در گونه های بسیار نزدیک به هم مثل انسان و شامپانزه نیز تفاوت هایی وجود دارد(27). در ژنوم پستانداران، ریزماهواره-های دی نوکلئوتیدی فراوانترین و پس از آنها ریزماهواره های مونو و تترا نوکلئوتیدی از وفور بالایی برخوردارند. تکرارهای تری نوکلئوتیدی معمولاً در گیاهان فراوان‌تر هستند. دلیل این تفاوت را به مکانیسم های مختلف جهش و یا تعمیر موتیف های ریزماهواره ای در درون گونه‌ها نسبت داده‌اند. همچنین ممکن است در انتخاب موتیف های ریزماهواره ای مختلف در موجودات مختلف نیز تفاوت وجود داشته باشد (7). اما بررسی اینکه آیا استفاده از موتیف‌های ریزماهواره-ایی 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.

نشانگرهای ریزماهواره 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 (معمولا کمتر از50) بسنده کرده‌اند (5). امروزه با پیشرفت‌هایی که در زمینه توالی یابی 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 ریزماهواره‌ای، بر نتایج حاصل از مطالعه تنوع ژنتیکی و برآورد پارامترهای ژنتیکی می‌پردازد.

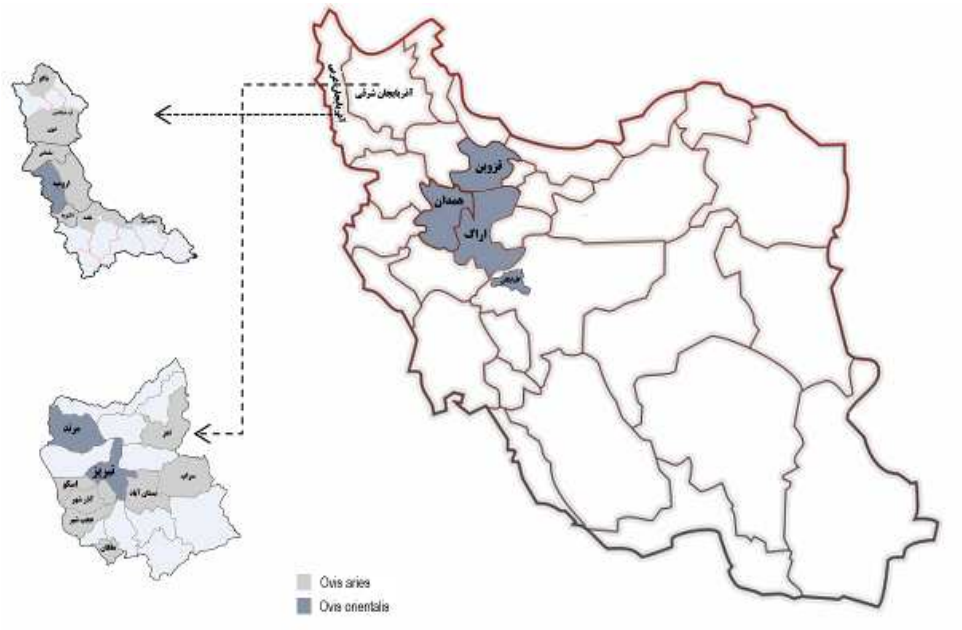
**مواد و روشها**

**داده ها**

این مطالعه روی تعداد 36 نمونه از جمعیت گوسفندان ایرانی (20 نمونه از جمعیت گوسفندان اهلی[3] با شناسه رده بندی 9440 در سایت NCBI و 16 نمونه از گونه وحشی[4] با شناسه رده بندی 469796 در سایت NCBI) که توالی کل ژنوم آنها به روش توالی یابی نسل جدید[5] موجود و فراخوانی واریانت ها[6] بر روی آنها انجام گرفته بود صورت گرفت. این مطالعه بر روی داده‌های ایجاد شده از پروژه NextGene به آدرس"https://projects.ensembl.org/nextgen/" در اختیار قرار گرفت. برنامه چهارچوب هفتم اتحادیه اروپا با کد FP7/2010-2014 سرمایه گذاری بر روی این پروژه را تحت قرارداد اعطای شماره 244356 بر عهده دارد. در داده‌های مورد مطالعه پراکنش نمونه های اهلی مربوط به نواحی شمال غرب کشور و نمونه های وحشی مربوط به نواحی مرکزی و شمال غرب کشور میباشند. شکل 1 موقعیت جغرافیایی محل های نمونه گیری را نشان می دهد.

---

[1]- DNA Slippage
[2] - mismatch repairing system
1- Ovis aries
2- Ovis orientalis
3- Next Generation Sequencing (NGS)
4- Variant Calling




شکل1- موقعیت جغرافیایی مناطق تحت نمونه گیری
Figure1- geographical location of samples

**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 3.1) "با شماره دسترسی GCA_000298735.1 در سایت NCBI" مورد بررسی قرار گرفت. این کار در برنامه Samtools (14 و 15) صورت گرفت. نتایج به دست آمده مکان و نوع موتیف را در تمام نمونه های مورد بررسی، صحیح نشان داد. تعداد 163973 جایگاه از انواع ریزماهواره‌های 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[1] و آلل (های) جایگزین[2] یک شکل باشد. در فایل داده اولیه مشاهده شد که در برخی جایگاه ها 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 284 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[3] از ریزماهواره های دی، تری و تترا نوکلئوتیدی تهیه شد. برای این کار از دستورات مختلف برنامه VCFtools استفاده شد(4). همچنین فایلی دیگر 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 ریزماهواره‌های مونو نوکلئوتیدی نیز از آنجا که روند تکاملی آنها تا به حال به خوبی بررسی نشده 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

1- Reference allele
2- Alternative allele
3- Variant Call Format (VCF)

نوع دی و mix-loci این زیرمجموعه‌ها به صورت 10، 20، 30، 40، 50، 60، 70، 80، 90، 100، 150، 200، 250، 300، 400 و 500 نشانگر و برای موتیف‌های نوع تری و تترا به صورت 10 الی 100 نشانگر بود. این زیرمجموعه ها با استفاده از اسکریپتی که در نرم افزار R به همین منظور نوشته شد، ایجاد شدند. این اسکریپت در داخل فایل موتیف‌های ریزماهواره‌ای مختلف به صورت تصادفی تعداد مورد نظر نشانگرها را انتخاب میکند. نمونه گیری با استفاده از این اسکریپت از داخل هر فایل و برای هر تعداد نشانگر 10 بار تکرار شد.

**برآورد پارامترهای تنوع ژنتیکی**

شش پارامتر معمول تنوع ژنتیکی شامل هتروزیگوسیته مشاهده شده و مورد انتظار، شاخص اطلاعات شانون، تنوع ژنی نئی، غنای آللی[1] وضریب همخونی[2] در این پژوهش در نرم افزار MSA (نسخه 4.05) (5) محاسبه شدند. در بین نرم‌افزارهای موجود جهت انجام آنالیزهای تنوع ژنتیکی نرم افزار MSA به دلیل توانایی در آنالیز تعداد زیادی از داده های چند آللی انتخاب شد. برای تبدیل فرمت داده‌ها از فرمت خوانش واریانت‌ها به فرمت مورد قبول برنامه MSA اسکریپتی تحت زبان پایتون نوشته شد. این اسکریپ در فایل داده‌های موجود، فرمت خوانش آلل‌ها را از تعداد تکرار موتیف (در فایل‌های با فرمت خوانش واریانت‌ها) به اندازه محصولات واکنش زنجیره‌ای پلیمراز[3] که فرمت مورد نیاز برنامه MSA می‌باشد تغییر می‌دهد. این نرم افزار پارامترها را بر اساس هر جمعیت و نیز هر جایگاه گزارش می‌کند. محاسبه پارامترها برای هر یک از 10 تکرار در زیرمجموعه‌های مختلف نشانگری به صورت جداگانه انجام گرفت. سپس میانگین و واریانس هر پارامتر در بین 10تکرار محاسبه شد و در نهایت نتایج به صورت نمودارهای باکس پلات در نرم افزار R (نسخه 3.3.3 ) 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 (3.3.3) 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 iامین زیرمجموعه در jامین 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 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 ریزماهواره‌ایی 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 (2، 17، 19). طی

1- Allelic richness
2- $F_{IS}$
3- PCR

مطالعات صورت گرفته (7، 12، 13) فراوانی موتیف‌های ریزماهواره ایی و نرخ جهش در موجودات متفاوت است. برای مثال مونو ریزماهواره‌ها در ناحیه اینترونی پریمات‌ها و باکتری اشریشیا کلای فراوانتر از موتیف‌های دی و تری نوکلئوتیدی گزارش شده است. ریزماهواره‌های دی نوکلئوتیدی بیشتر در ژنوم پستاندارانی چون جوندگان گزارش شده است (20) و تری نوکلئوتیدها در مقایسه با سایر موتیف‌های ریزماهواره ای بیشتر در ژنوم گیاهان گزارش شده است (1). در این مطالعه در دو جمعیت گوسفند مورد بررسی از 163973 جایگاه ریزماهواره‌ای موجود، مونو ریزماهواره‌ها بیشترین تعداد (86854 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 9782 جایگاه شد. از تعداد 5739 جایگاه دارای اشکال نوکلئوتید اضافی تعداد 284 جایگاه به ادامه آنالیزها برگردانده شدند. درنهایت برآورد پارامترهای ژنتیکی با استفاده از تعداد 3997 ریزماهواره دی، 115 ریزماهواره تری و 140 ریزماهواره تترا نوکلئوتیدی، انجام گرفت. جدول 1 شمارش تعداد جایگاه ها در فایل داده اولیه و نیز در مراحل مختلف آماده سازی داده‌ها را نشان می‌دهد.

**جدول 1- تعداد جایگاه های ریزماهواره در حین مراحل آماده سازی نمونه‌ها**

**Table1- 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 جایگاه ها Initial number of loci |
| 18701 | 6664 | 41972 | 56812 | 124149 | جایگاه های تک شکل Monomorphic loci |
| 1078 | 397 | 8307 | 30042 | 39824 | جایگاه های چند شکل Polymorphic loci |
| 999 | 228 | 4512 | 0 | 5739 | جایگاه های چند شکل دارای مشکل آلل خوانی Polymorphic loci with allele calling problem |
| 79 | 169 | 3795 | 30042 | 34085 | جایگاه های چند شکل بدون مشکل آلل خوانی Polymorphic loci with no allele calling problem |
| 146 | 115 | 4066 | 30042 | 34369 | جایگاه های چند شکل پس از کاربرد اسکریپ Polymorphic loci with using script |
| 140 | 115 | 3997 | 0 | 4252 | جایگاه های نهایی Final loci |

## پارامترهای تنوع ژنتیکی

منطقه خاورمیانه به دلیل نزدیکی به ناحیه هلال بارور[1] 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 وهمکاران (2016) تنوع ژنتیکی و انشقاق 10 تیپ مختلف از گوسفندان دنبه دار ایرانی را با استفاده از 18 نشانگر ریزماهواره‌ای مورد بررسی قرار داده و سطح بالایی از تنوع را در بین تیپ‌های مورد بررسی نشان دادند(23). در اکثر مطالعات انجام گرفته بر روی تیپ‌های مختلف گوسفندان ایرانی از تعداد 4 الی 20جایگاه 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 (8، 16، 23، 24، 25، 27). در این مطالعه در آنالیز پارامترهای تنوع ژنتیکی در جمعیت‌های مورد بررسی تعداد 9 جایگاه دی نوکلئوتیدی خطای "آلل بیش از حد

[1]- Fertile Crescent

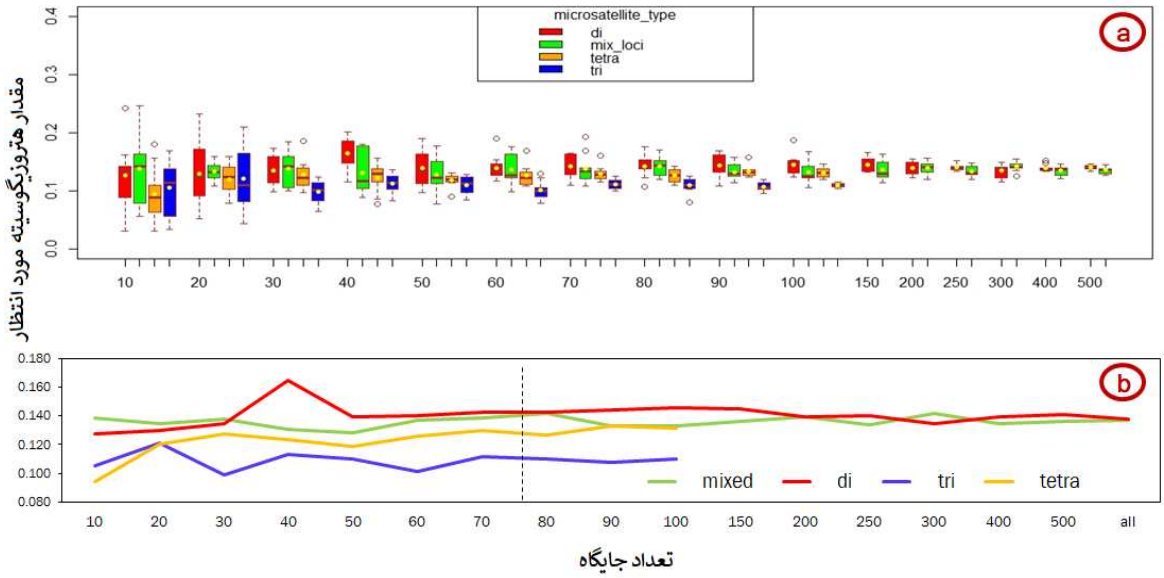
بزرگ" را نشان دادند. این جایگاه‌ها scaffold بودندکه از ادامه آنالیزها کنار گذاشته شدند و ادامه محاسبات با تعداد 4243 نشانگر (3988ریزماهواره دی نوکلئوتیدی) ادامه یافت. محاسبه شش پارامتر تنوع ژنتیکی با استفاده از زیرمجموعه‌های تعریف شده در جمعیت گوسفندان اهلی نتایج را بدین شرح ارائه داد: هتروزیگوسیته مشاهده شده (بیشترین مقدار 0/07 با 40 نشانگر دی، کمترین مقدار 0/031 با 10 نشانگر تترا)، هتروزیگوسیته مورد انتظار (بیشترین مقدار 0/165 با 40 نشانگر دی، کمترین مقدار 0/094 با 10 نشانگر تترا)، شاخص تنوع نئی (بیشترین مقدار 0/264 با 40 نشانگر دی، کمترین مقدار 0/174 با 30 نشانگر تری)، شاخص شانون (بیشترین مقدار 0/279 با 40 نشانگر دی، کمترین مقدار 0/155 با 10 نشانگر تترا)، غنای آللی (بیشترین مقدار 1/81با 40 نشانگر دی، کمترین مقدار 1/45 با 10 نشانگرتری). ضریب همخونی (بیشترین مقدار 0/504 با30 نشانگر mix-loci، کمترین مقدار 0/336 با 10 نشانگر دی).

در جمعیت گوسفندان وحشی نیز نتایج بدین صورت بود: هتروزیگوسیته مشاهده شده (بیشترین مقدار 0/074 با 40 نشانگر دی، کمترین مقدار 0/045 با 80 نشانگر تترا)، هتروزیگوسیته مورد انتظار (بیشترین مقدار 0/192 با 40 نشانگر دی، کمترین مقدار 0/114 با 60 نشانگر تری)، شاخص تنوع نئی (بیشترین مقدار0/267 با 40 نشانگر دی، کمترین مقدار0/183 با 60 نشانگر تری)، شاخص شانون (بیشترین مقدار 0/324 با 10 نشانگر mix-loci، کمترین مقدار 0/185 با 70 نشانگر تری)، غنای آللی (بیشترین مقدار 1/97 با 40 نشانگر mix-loci، کمترین مقدار1/61 با 20 نشانگر تری)، ضریب همخونی(بیشترین مقدار 0/484 با 70 نشانگر تترا، کمترین مقدار 0/331 با 60 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 (شکل های 2 تا 5، aو b).

در نمودارهای باکس پلات (شکل های 2-5،a) هر ستون میزان پراکندگی نتایج به دست آمده از 10 تکرار را نشان می‌دهد. این پراکندگی با افزایش تعداد نشانگر کاهش می‌یابد. مقایسه پلات‌ها نتایج متفاوتی از تخمین یک پارامتر را نشان می‌دهد به طوریکه 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 تترا 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 تترا نوکلئوتیدی در این فایل، نتایج قابل توضیح است. نقاط زرد رنگ روی هر ستون مقدار میانگین 10 تکرار را نشان میدهد. نمودارهای b بر اساس میانگین 10 تکرار ترسیم شده‌اند. بر اساس نتایج در استفاده از 10 الی 40 نشانگر، پراکنش زیادی بین تکرارهای مختلف مشاهده میشود. در اکثر پارامترهای مورد بررسی بیشترین مقدار برآورده شده برای آن پارامتر با استفاده از 40 نشانگر دی نوکلئوتیدی و کمترین آن با استفاده از 10 نشانگر تری و یا ترا نوکلئوتیدی به دست آمده است. با استفاده از50 نشانگر نمودارها (b) به سمت ثابت شدن پیش می‌روند و در استفاده از 80 نشانگر به بالا، واریانس‌ها به سمت صفر میل میکنند و تقریبا میتوان در برآورد هر پارامتر به یک مقدار عددی ثابت دست یافت.

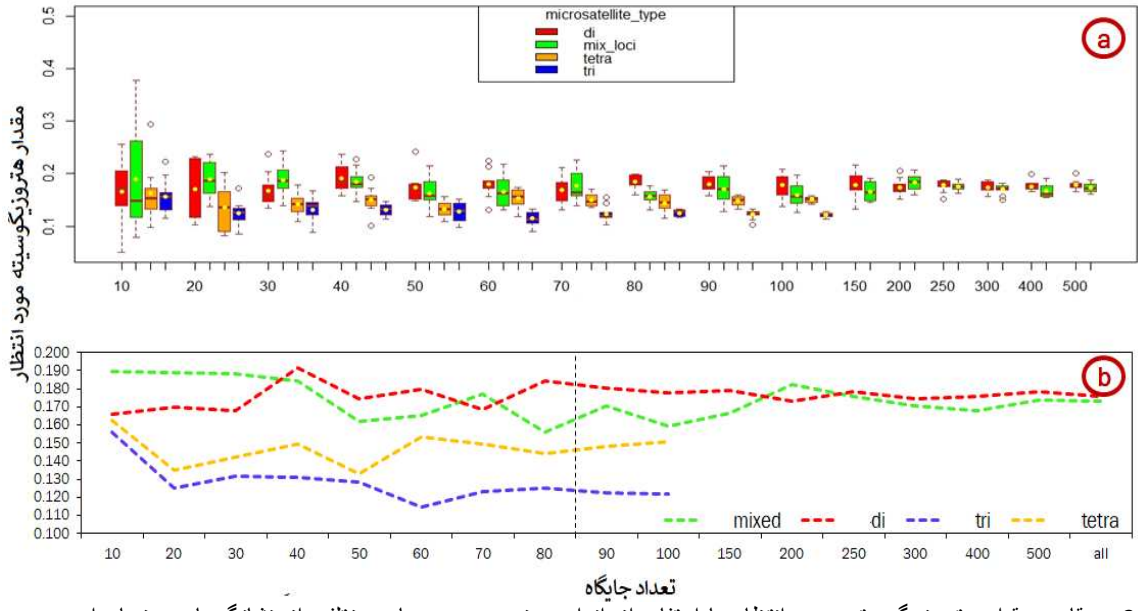
مقایسه پارامترهای تنوع ژنتیکی در دو جمعیت اهلی و وحشی در این مطالعه تنوع ژنتیکی بسیار پایینی در هر دو جمعیت را نشان داد که با مطالعات تنوع ژنتیکی پیشین صورت گرفته روی تیپ‌های مختلف گوسفندان ایرانی مغایرت دارد (8، 16، 23 24، 25، 27). از آنجا که این مطالعه به تعداد کافی نشانگر دسترسی داشته این میزان تنوع 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 ریزماهواره‌های تری) بود. وجود تنوع پایین 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.



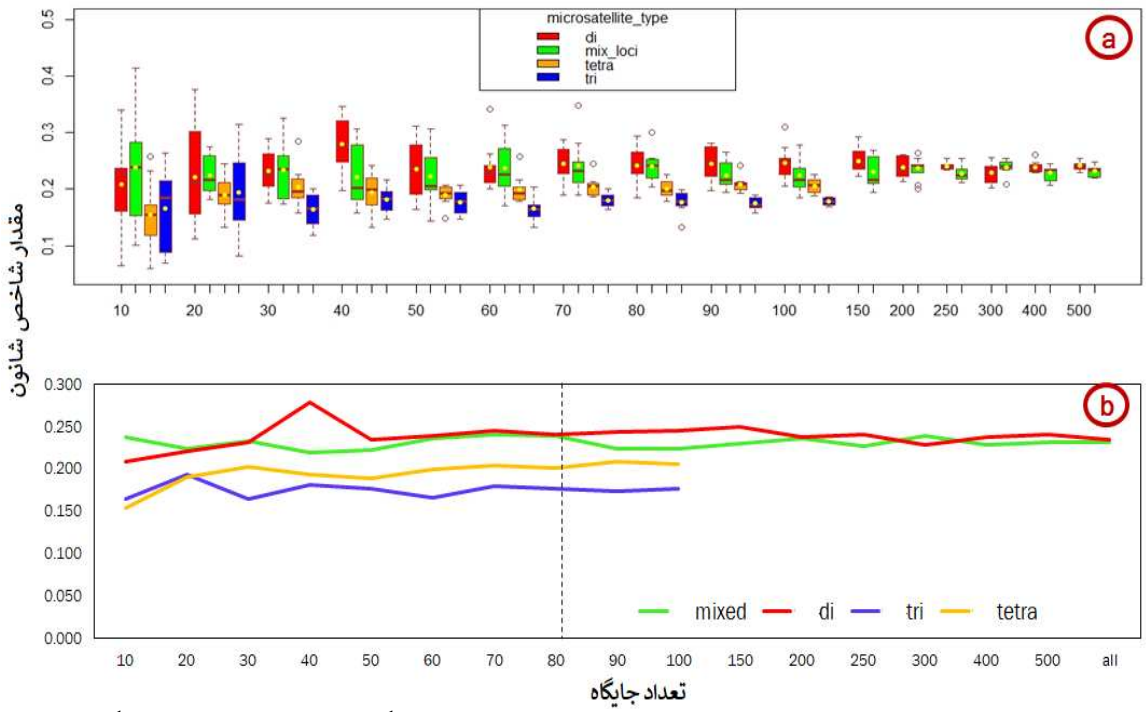

شکل 2- مقایسه مقدار هتروزیگوسیته مورد انتظار با استفاده از انواع و زیرمجموعه های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان اهلی (a: پراکنش داده ها در 10 تکرار برای هر زیرمجموعه از نشانگرها، b: مقدار میانگین به دست آمده از 10 تکرار در هر زیرمجموعه از نشانگرها)

Figure2-Comparing expected heterozygosity values with different subset of microsatellite types in Ovis aries



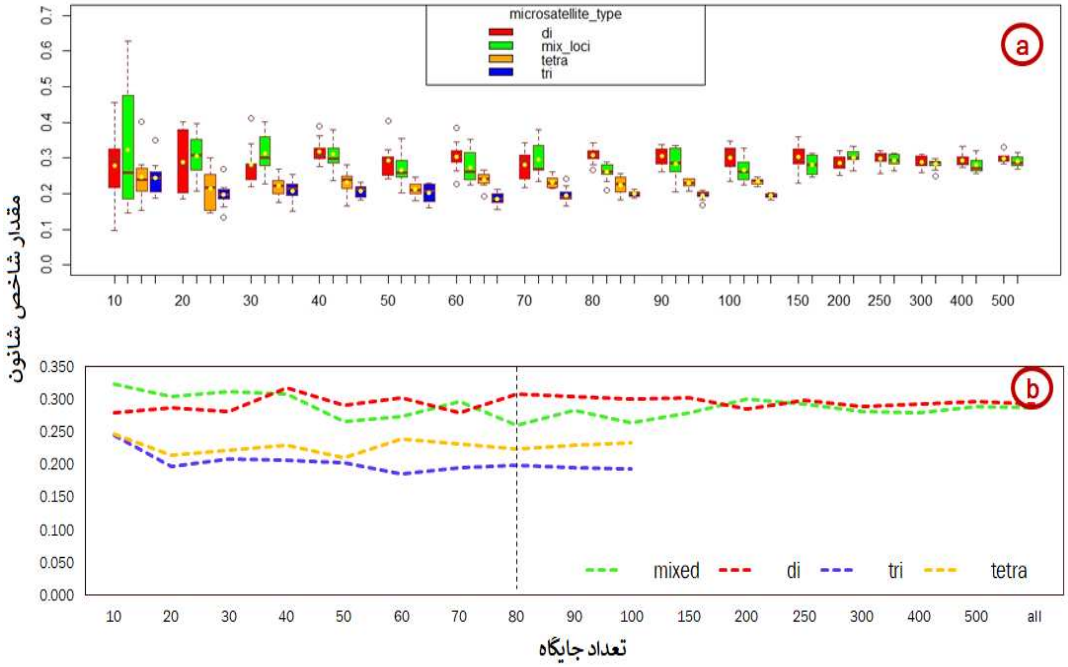

شکل3 - مقایسه مقدار هتروزیگوسیته مورد انتظار با استفاده از انواع و زیرمجموعه های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان وحشی(a: پراکنش داده ها در 10 تکرار برای هر زیرمجموعه از نشانگرها، b: مقدار میانگین به دست آمده از 10 تکرار در هر زیرمجموعه از نشانگرها)

Figure4-Comparing expected heterozygosity values with different subset of microsatellite types in Ovis orientalis




شکل4- مقایسه مقدار شاخص شانون با استفاده از انواع و زیرمجموعه های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان اهلی(a: پراکنش داده ها در 10 تکرار برای هر زیرمجموعه از نشانگرها، b: مقدار میانگین به دست آمده از 10 تکرار در هر زیرمجموعه از نشانگرها)

Figure3-Comparing Shanon values with different subset of microsatellite types in Ovis aries




شکل 5- مقایسه مقدار شاخص شانون با استفاده از انواع و زیرمجموعه های مختلفی از نشانگرهای ریزماهواره در جمعیت گوسفندان وحشی(a: پراکنش داده ها در 10 تکرار برای هر زیرمجموعه از نشانگرها، b: مقدار میانگین به دست آمده از 10 تکرار در هر زیرمجموعه از نشانگرها)

Figure5-Comparing Shanon values with different subset of microsatellite types in Ovis orientalis

**بررسی معنی داری تعداد و نوع موتیف بر برآورد پارامترها**

در مطالعه‌ای جهت اندازه گیری فاصله ژنتیکی بین چهار جمعیت انسانی با استفاده از 213 نشانگر ریزماهواره، کوپر و همکاران (1999) 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 (3). در مطالعه دیگری بر روی دروزوفیلا ملانوگاستر مشخص شد استفاده از نشانگرهای ریزماهواره در موقعیت‌های کروموزومی مختلف می‌تواند باعث تفاوت‌های معنی دار بین ساختار جمعیت ها شود و جالب اینکه هریک از ساختارهای جمعیتی که با استفاده از ترکیبات مختلف نشانگری به دست می آیند با دارا بودن احتمالات پسین بالا (P≥0/95) از نظر آماری دارای پشتوانه بالایی هستند. در این مطالعه مشخص گردید هر اندازه تعداد 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 120 جایگاه نشانگری نیاز است (17). 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 معنی دار 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 2 و 4 آورده شده است. اعداد داخل جدول مقدارP به دست آمده از این بررسی را نشان میدهد.

**جدول2- ارزیابی اثر تعداد نشانگر در برآورد پارامترهای ژنتیکی با مقایسه مقدار P به دست آمده**

**Table2- Evaluating the effect of number of loci on genetic parameter estimation comparing P values**

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

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

نتایج نشان میدهد با استفاده از موتیف‌های دی (و همینطور mix-loci) تفاوت معنی داری بین زیرمجموعه‌های نشانگری مختلف در برآورد شش پارامتر مورد نظر مشاهده نمیشود اما استفاده از موتیف‌های تری (در جمعیت گوسفندان وحشی) و تترا نوکلئوتیدی (در جمعیت گوسفندان اهلی) در برآورد برخی پارامترهای ژنتیکی معنی دار است. به منظور بررسی دقیق‌تر و یافتن اینکه دقیقاً کدام دسته از نشانگرها با یکدیگر اختلاف معنی دار دارند، مقایسه جفتی میانگین‌ها به روش توکی بر روی مواردی که دارای مقدار P معنی دار بودند صورت گرفت که نتایج آن در جدول 3 گزارش شده است.

**جدول 3- آزمون مقایسه میانگین به روش توکی بین تعداد نشانگرهای مختلف در دو جمعیت گوسفند اهلی و وحشی**

**Table3- Tukey significant between different numbers of loci in Ovis aries and Ovis orentalis population**

| جمعیت(population) | گوسفند اهلی(Ovis aries) | | | گوسفند وحشی(Ovis orientalis) | | | |
|---|---|---|---|---|---|---|---|
| نوع موتیف (motif type) | tetra | | | tri | | | |
| پارامتر (Parameter) | هتروزیگوسیته مشاهده شده (Het-obs) | هتروزیگوسیته مورد انتظار (Het-exp) | شاخص شانون (Shanon index) | هتروزیگوسیته مورد انتظار (Het-exp) | شاخص شانون (Shanon index) | ضریب همخونی ($F_{IS}$) | غنای آللی (Allelic richness) |
| **10 loci** | $0.031^{b}$ | $0.094^{b}$ | $0.155^{b}$ | $0.156^{a}$ | $0.245^{a}$ | $0.449^{a}$ | $1.718^{a}$ |
| **20 loci** | $0.047^{a}$ | $0.120^{ab}$ | $0.190^{ab}$ | $0.125^{b}$ | $0.197^{b}$ | $0.362^{ab}$ | $1.607^{b}$ |
| **30loci** | $0.044^{ab}$ | $0.128^{a}$ | $0.203^{a}$ | $0.132^{ab}$ | $0.208^{ab}$ | $0.355^{b}$ | $1.650^{ab}$ |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1.632[ab] | 0.354[b] | 0.206[b] | 0.131[ab] | 0.194[ab] | 0.123[ab] | 0.041[ab] | **40loci** |
| 1.631[ab] | 0.335[b] | 0.203[b] | 0.128[b] | 0.189[ab] | 0.118[ab] | 0.043[ab] | **50loci** |
| 1.619[b] | 0.331[b] | 0.185[b] | 0.114[b] | 0.199[a] | 0.126[ab] | 0.045[a] | **60loci** |
| 1.619[b] | 0.341[b] | 0.196[b] | 0.123[b] | 0.204[a] | 0.130[a] | 0.045[a] | **70loci** |
| 1.631[ab] | 0.345[b] | 0.199[b] | 0.125[b] | 0.201[a] | 0.127[a] | 0.047[a] | **80loci** |
| 1.622[ab] | 0.346[b] | 0.196[b] | 0.122[b] | 0.208[a] | 0.133[a] | 0.047[a] | **90loci** |
| 1.624[ab] | 0.334[b] | 0.194[b] | 0.122[b] | 0.206[a] | 0.131[a] | 0.047[a] | **100loci** |

*میانگین‌هایی که دارای حروف مشترک هستند از نظر آماری دارای اختلاف معنی دار نیستند (در سطح اطمینان 95درصد).

اعداد داخل جدول مقدار پارامترهای محاسبه شده با تعداد مختلفی از نشانگرها را نشان میدهد. بررسی نتایج در جمعیت گوسفندان اهلی وجود اختلاف معنی دار در برآورد پارامترهای نشان داده شده با استفاده از 10 و 70 نشانگر را نشان میدهد. در جمعیت گوسفندان وحشی نیز متفاوت بودن نتایج به دست آمده از تخمین پارامترها در مقایسه 10 و 60 نشانگر قابل مشاهده است. هرچند استثناهایی در مورد برخی پارامترها وجود دارد اما به طور قطع میتوان گفت در استفاده از 10 و 70 نشانگر تفاوت‌های معنی دار در برآورد پارامترها وجود دارد. در واقع میتوان گفت هرچه تعداد نشانگرها بیشتر میشود تفاوت در میانگین‌ها معنی دارتر میشود (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 زیرمجموعه‌های مختلف موتیف‌ها استفاده شد. این بررسی با تعداد 10، 20، 50، 70 و 100 نشانگر در هر دو جمعیت صورت گرفت.

**جدول4-ارزیابی اثر نوع موتیف در برآورد پارامترهای ژنتیکی با استفاده از زیرمجموعه‌های مختلف با مقایسه مقدار P**
**Table4-Evaluating the effect of motif type on genetic parameter estimation using different subsets comparing P values**

| **جمعیت (population)** | **گوسفند اهلی (Ovis aries)** | | | | | **گوسفند وحشی(Ovis orentalis)** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **تعداد جایگاه (Number of loci)** | 10 | 20 | 50 | 70 | 100 | 10 | 20 | 50 | 70 | 100 |
| **پارامتر (Parameter)** | | | | | | | | | | |
| **هتروزیگوسیته مشاهده شده (Het-obs)** | 0.18 | 0.172 | 0.081 | 0.000*** | 0.00*** | 0.186 | 0.051 | 0.001*** | 0.000*** | 0.00*** |
| **هتروزیگوسیته مورد انتظار (Het_exp)** | 0.263 | 0.831 | 0.041 | 0.003** | 0.00*** | 0.685 | 0.002** | 0.000*** | 0.000*** | 0.00*** |
| **شاخص شانون (Shanon index)** | 0.737 | 0.468 | 0.004 | 0.000*** | 0.00*** | 0.302 | 0.0006*** | 0.000*** | 0.000*** | 0.00*** |
| **شاخص نئی (Nei)** | 0.566 | 0.281 | 0.012* | 0.003** | 0.00*** | 0.748 | 0.010* | 0.002** | 0.0002*** | 0.00*** |
| **غنای آللی(Ar)** | 0.004** | 0.017* | 0.000*** | 0.000*** | 0.00*** | 0.030* | 0.00*** | 0.000*** | 0.000*** | 0.00*** |
| **ضریب همخونی ($F_{IS}$)** | 0.295 | 0.494 | 0.138 | 0.1 | 0.02* | 0.592 | 0.226 | 0.057 | 0.000*** | 0.00*** |

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

در دو جمعیت مورد بررسی، با استفاده از 10 نشانگر اختلاف معنی دار بین موتیف‌های ریزماهواره‌ای 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 نمایان میشود به طوریکه با استفاده از 100 نشانگر، در برآورد هر شش پارامتر در هر دو جمعیت اختلافات معنی دار مشاهده میشود. بررسی میانگین‌ها برای یافتن موتیف‌هایی که اختلاف معنی دار نشان داده اند با آزمون مقایسه میانگین توکی صورت گرفت که نتایج آن در جدول 5 آورده شده است.

**جدول5-** آزمون مقایسه میانگین به روش توکی بین موتیف‌های مختلف و مقایسه نتایج با استفاده از تعداد نشانگرهای مختلف در جمعیت گوسفندان اهلی و وحشی

**Table5- Tukey significant test between different motif types by comparing results using different number of loci in Ovis aries and Ovis Orientalis populations**

| جمعیت (population) | | گوسفند اهلی(Ovis.aries) | | | | | | گوسفند وحشی(Ovis.orientalis) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| تعداد جایگاه( Number of loci) | پارامتر (Parameter) / نوع موتیف (Motif type) | هتروزیگوسیته مشاهده شده (Het-obs) | هتروزیگوسیته مورد انتظار (Het_exp) | شاخص شانون (Shanon) | شاخص نئی (Nei) | غنای آللی (Ar) | ضریب همخونی ($F_{IS}$) | هتروزیگوسیته مشاهده شده (Het-obs) | هتروزیگوسیته مورد انتظار (Het_exp) | شاخص شانون (Shanon) | شاخص نئی (Nei) | غنای آللی (Ar) | ضریب همخونی ($F_{IS}$) |
| 10 | mixloci | - | - | - | - | $1.705^{a}$ | - | - | - | - | - | $1.973^{a}$ | - |
| | di | - | - | - | - | $1.625^{ab}$ | - | - | - | - | - | $1.910^{a}$ | - |
| | tri | - | - | - | - | $1.447^{b}$ | - | - | - | - | - | $1.718^{a}$ | - |
| | tetra | - | - | - | - | $1.484^{b}$ | - | - | - | - | - | $1.663^{a}$ | - |
| 20 | mixloci | - | - | - | - | $1.675^{ab}$ | - | - | $0.189^{a}$ | $0.305^{a}$ | $0.255^{a}$ | $1.893^{a}$ | - |
| | di | - | - | - | - | $1.706^{a}$ | - | - | $0.170^{ab}$ | $0.288^{ab}$ | $0.231^{ab}$ | $1.962^{a}$ | - |
| | tri | - | - | - | - | $1.545^{ab}$ | - | - | $0.125^{b}$ | $0.197^{c}$ | $0.204^{b}$ | $1.607^{b}$ | - |
| | tetra | - | - | - | - | $1.530^{b}$ | - | - | $0.135^{b}$ | $0.214^{bc}$ | $0.203^{b}$ | $1.638^{b}$ | - |
| 50 | mixloci | - | - | $0.222^{ab}$ | $0.210^{ab}$ | $1.709^{a}$ | - | $0.071^{a}$ | $0.174^{a}$ | $0.292^{a}$ | $0.244^{a}$ | $1.954^{a}$ | - |
| | di | - | - | $0.236^{a}$ | $0.240^{a}$ | $1.717^{a}$ | - | $0.067^{a}$ | $0.162^{a}$ | $0.267^{a}$ | $0.240^{a}$ | $1.833^{a}$ | - |
| | tri | - | - | $0.177^{b}$ | $0.195^{b}$ | $1.507^{b}$ | - | $0.057^{ab}$ | $0.128^{b}$ | $0.203^{b}$ | $0.201^{b}$ | $1.631^{b}$ | - |
| | tetra | - | - | $0.189^{b}$ | $0.200^{b}$ | $1.542^{b}$ | - | $0.047^{b}$ | $0.133^{b}$ | $0.212^{b}$ | $0.205^{b}$ | $1.636^{b}$ | - |
| 70 | mixloci | - | $0.139^{a}$ | $0.241^{a}$ | $0.226^{ab}$ | $1.743^{a}$ | - | $0.069^{a}$ | $0.177^{a}$ | $0.296^{a}$ | $0.250^{a}$ | $1.917^{a}$ | $0.450^{a}$ |
| | di | - | $0.143^{a}$ | $0.245^{a}$ | $0.245^{a}$ | $1.738^{a}$ | - | $0.066^{a}$ | $0.168^{ab}$ | $0.281^{a}$ | $0.236^{a}$ | $1.895^{a}$ | $0.433^{a}$ |
| | tri | - | $0.112^{b}$ | $0.180^{b}$ | $0.196^{b}$ | $1.514^{b}$ | - | $0.052^{b}$ | $0.123^{c}$ | $0.196^{b}$ | $0.196^{a}$ | $1.619^{b}$ | $0.341^{b}$ |
| | tetra | - | $0.130^{ab}$ | $0.204^{b}$ | $0.213^{ab}$ | $1.560^{b}$ | - | $0.045^{b}$ | $0.150^{bc}$ | $0.232^{b}$ | $0.228^{a}$ | $1.651^{b}$ | $0.484^{a}$ |
| 100 | mixloci | $0.051^{b}$ | $0.133^{a}$ | $0.224^{ab}$ | $0.219^{ab}$ | $1.666^{a}$ | $0.427^{ab}$ | $0.063^{b}$ | $0.159^{ab}$ | $0.265^{b}$ | $0.230^{b}$ | $1.853^{b}$ | $0.394^{bc}$ |
| | di | $0.058^{a}$ | $0.146^{a}$ | $0.246^{a}$ | $0.237^{a}$ | $1.736^{a}$ | $0.422^{ab}$ | $0.073^{a}$ | $0.178^{a}$ | $0.301^{a}$ | $0.258^{a}$ | $1.958^{a}$ | $0.421^{ab}$ |
| | tri | $0.044^{c}$ | $0.110^{b}$ | $0.178^{c}$ | $0.190^{c}$ | $1.517^{b}$ | $0.402^{b}$ | $0.052^{c}$ | $0.122^{c}$ | $0.194^{d}$ | $0.192^{c}$ | $1.624^{c}$ | $0.334^{c}$ |
| | tetra | $0.047^{bc}$ | $0.131^{a}$ | $0.206^{b}$ | $0.214^{b}$ | $1.558^{b}$ | $0.470^{a}$ | $0.051^{c}$ | $0.151^{b}$ | $0.234^{c}$ | $0.229^{b}$ | $1.652^{c}$ | $0.470^{a}$ |

*مواردی که با (-) نشان داده شده میانگین هایی هستندکه ارزش P معنی دار در مدل برای آنها گزارش نشده و بنابراین تحت آزمون مقایسه میانگین قرار نگرفته اند.**میانگین‌هایی که دارای حروف مشابه هستند از نظر آماری دارای اختلاف معنی دار نیستند (در سطح اطمینان 95درصد).

در اکثر مقایسات جفتی بین میانگین‌ها، پارامترهای برآورد شده با استفاده از موتیف‌های نوع دی و mix-loci با یکدیگر اختلاف معنی دار ندارند که دلیل آن را میتوان به یکسان بودن نسبی فایل داده آنها نسبت داد چرا که در فایل داده mix-loci اکثر نشانگرها را نشانگرهای دی نوکلئوتیدی تشکیل میدهند. همینطور در مورد پارامترهای برآورد شده با استفاده از موتیف‌های نوع تری و تترا نیز اختلافات معنی دار کمی مشاهده شد. بیشترین میزان اختلاف مربوط به مقایسه میانگین پارامترها با استفاده از موتیف‌های نوع دی با تترا و دی با تری است و این به این معناست که در برآورد پارامترهای تنوع ژنتیکی استفاده از ریزماهواره‌های تری یا تترا تفاوت آنچنانی بر روی نتایج نهایی نخواهد داشت اما استفاده از ریزماهواره‌های دی نوکلئوتیدی قطعا 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 معنی دار بیشتری بین داده ها قابل رویت است.

**نتیجه گیری نهایی**

نتایج مطالعات تنوع ژنتیکی در جمعیت‌ها می تواند تحت تاثیر تعداد جایگاه‌های مختلف و نیز موتیف‌های مختلف نشانگرهای ریزماهواره قرار گیرد. نتایج به دست آمده از این پژوهش تفاوت معنی دار در استفاده از موتیف‌های نوع دی در برآورد پارامترهای ژنتیکی را نشان داد. از آنجاییکه این ریزماهواره‌ها در سطح ژنوم فراوانی بالاتری دارند میتوانند در مطالعات بررسی ساختار جمعیتی روی گونه‌های دامی مناسب‌تر باشند. نتایج این پژوهش همچنین لزوم

استفاده از حداقل 50 جایگاه ریزماهواره‌ای را برای داشتن برآوردهایی باثبات از پارامترهای جمعیت در مطالعات تنوع ژنتیکی پیشنهاد میدهد. تعداد و نوع موتیف ریزماهواره‌ها احتمالا در برآورد دیگر پارامترها از جمله پیش بینی 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 نرم افزار lobSTR) نشان می‌دهد. خوانش اشتباه توالی ریزماهواره‌ها 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 هشداری برای مسئولین طرح‌های حفاظت ژنتیکی در گوسفندان وحشی و هم چنین طرح های اصلاح نژادی در گوسفندان اهلی در داخل کشور باشد.

**منابع:**


1. Abdurakhmonov, I. Y. 2016. Introduction to microsatellites: basics, trends and highlights. P4-16, In: Abdurakhmonov, I. Y (eds), Microsatellite Markers, Intech open. Uzbekistan.
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., & 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.
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., Durbin, R. 2011. The variant call format and VCFtools. Bioinformatics. 27(15), 2156-2158.
5. Dieringer, D., & 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., & 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. Esmaeilkhanian 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., Collins, F., et al. 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., & 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., & Gupta, V. S. 2001. Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol. Biol. Evol., 18(7), 1161-1167.
13. Lagercrantz, U., Ellegren, H., & Andersson, L. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res, 21(5),1111-1115.
14. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Szustakowki, J ., et al .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 SH, A. C., 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., & Schlötterer, C.2011. Genealogical lineage sorting leads to significant, but incorrect Bayesian multilocus inference of population structure. Molecular ecology, 20(6), 1108-1121.
18. Schlötterer, C., & 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., & Jeon, G. J. 2016. Microsatellite Analysis of the Genetic Diversity and Population Structure in Dairy Goats in Thailand. Asian-Australas J. Anim. Sci., 29(3), 327-332.
20. Sheriff, O., & Alemayehu, K. 2018. Genetic diversity studies using microsatellite markers and their contribution in supporting sustainable sheep breeding programs: a review. Cogent food and agriculture, 4(1).DOI: 10.1080/23311932.2018.1459062
21. Trivedi, S. 2004. Microsatellites (SSRs): puzzles with puzzle. Indian Journal of Biotechnology. 3, 331-347.
22. Tautz, D., & Schlötterer, C.1994. Simple sequences. Current Opinion in Genetics & 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. Anim Genet, 47: 463-470. doi:10.1111/age.12429
24. Vajed Ebrahimi, M. T., Mohammad Abadi, M. R., & Esmaeilzadeh, 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., & Esmailizadeh, A. 2017. Using microsatellite markers to analyze genetic diversity in 14 sheep types in Iran. Arch. Anim. Breed., 60(3), 183-189.
26. Webster, M. T., Smith, N. G., & Ellegren, H. 2002. Microsatellite evolution inferred from human-chimpanzee genomic sequence alignments. Proc. Natl. Acad. Sci. U S A, 99(13), 8748-8753.
27. Zahedi, Z., Esmaeelkhanian, S., & Torshizi, R. V. 2007. Microsatellite variation in one breed of Iranian sheep with 12 markers. Pak. J. Biol. Sci., 10(24), 4455-4460.

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

**Abstract**

**Background and objective:**

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

**Material and methods**

In the present study with using the 36 VCf file of microsatellite markers extracted from 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 north-west part of Iran and ovis orientalis samples are belongs to central part and north-west of Iran. After rearrangement and filtration on data file using Samtools and VCFtools softwares, we classified the whole markers on four different motif types including di- tri- and tetra-nucleotide microsatellites and a file includes 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 a 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 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 value for most parameters were obtained by 40-di and 10-tri/tetra microsatellites respectively. 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 the case with significant results, Tukey Honestly Significant Difference test was used to test pairwise contrasts between different subsets.

**Conclusion**

Our result proposes 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 motifs, microsatellite marker, sheep