

Identification and validation of simple sequence repeats (SSRs) markers for *Androctonus crassicauda* scorpion using RNA-Seq data

Roshan Kaki 

Ph. D student in Department of Animal Science, Faculty of Animal science and Food Technology, Agricultural Science and Natural Resources University of Khuzestan, Mollasani, Iran. Email address: roashan.kaki@yahoo.com

Mahmood Nazari 

*Corresponding author. Associate Professor, Department of Animal Science, Faculty of Animal science and Food Technology, Agricultural Science and Natural Resources University of Khuzestan, Mollasani, Iran. Email address: M.Nazari@asnrukh.ac.ir

Hedaiat allah Rooshanfekr 

Professor, Department of Animal Science, Faculty of Animal science and Food Technology, Agricultural Science and Natural Resources University of Khuzestan, Mollasani, Iran. Email address: roshanfekr_hd@yahoo.com

Fatemeh Salabi 

Department of Venomous Animals and Anti-venom Production, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Ahvaz-Iran. E-mail: f.salabi@rvsri.ac.ir

Abstract

Objective

Androctonus crassicauda scorpion is known as one of the deadliest species of scorpion in the world for mammals. Simple sequence repeats (SSRs) markers are one of most complete molecular tools in the study of genetic diversity due to the appropriate genomic coverage and high repeatability. Simple sequence repeats markers for the scorpion *Androctonus crassicauda* have not been reported so far. Therefore, this study was conducted with the aim of identifying and validating simple repeat sequence markers of the scorpion *Androctonus crassicauda* using RNA-Seq data.

Materials and methods

The samples were evaluated qualitatively and quantitatively after RNA extraction from the scorpion venom gland. Then the samples were sequenced with the Illumina HiSeq 2000 platform.

Reconstruction of transcripts was done by De novo assembly method using Trinity software. Next, in order to identify new SSR markers, the transcriptome of the scorpion *Andrachtonus crassicauda* was analyzed using FullSSR software (version 1.5). To validate the identified SSR markers, 8 pairs of primers were evaluated by polymerase chain reaction method. Indices such as observed heterozygosity (Ho), expected heterozygosity (He) and fixation index (F) were calculated using GenAlEx software (version 6.5) for statistical analysis of polymorphic loci.

Results

Assembling transcripts by Trinity program generated a total of 744,804 transcripts and 563,526 unigenes. In this study, the examination of 952,725 sequences by FullSSR software led to the identification of 315,395 SSR markers. Among SSRs, two and three nucleotide repeats accounted for the highest number of repetitions with 71.85% and 22.36%, respectively. Among the 8 studied loci, only 2 loci showed polymorphism. The expected and observed heterozygosity for the RK1354 locus was calculated as 0.765 and 0.683 respectively, and for the NK1362 locus as 0.768 and 0.633 respectively. The fixation index statistic (F_{IS}) for both loci was almost 0.1, which indicates low inbreeding in the population.

Conclusions

The different criteria of genetic diversity all indicate that the population has high diversity, but there are factors that are reducing the diversity among the studied population. So that the population is out of balance and the number of observed heterozygotes is less than the expected heterozygotes. It is necessary to investigate the factors that have led to the reduction of diversity in order to prevent its further reduction. In general, the results of this research showed that new SSRs can be useful for understanding the population structure and investigating the genetic diversity of the scorpion *Andrachtonus crassicauda*.

Keywords: *Andrachtonus crassicauda*, simple sequence repeats, transcriptome

Paper Type: Research Paper.

Citation: Kaki, N., Nazari, M., Roshanfekar, H., & Salabi, F. (2025). Identification and validation of simple sequence repeats (SSRs) Markers for *Andrachtonus crassicauda* scorpion using RNA-Seq data. *Agricultural Biotechnology Journal*, 17 (2), 31-54.

Agricultural Biotechnology Journal 17 (2), 31-54.

DOI: 10.22103/jab.2025.24286.1625

Received: January 04, 2024.

Received in revised form: February 20, 2025.

Accepted: February 21, 2025.

Published online: March 31, 2025.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.



© the authors

Introduction

Scorpions have eight legs and are easily recognized by a pair of grasping pincers and a narrow, segmented tail, often carried in a characteristic forward curve over the back and always ending with a stinger. The evolutionary history of scorpions goes back 435 million years. Scorpions primarily prey on insects and other invertebrates, but some species hunt vertebrates. They use their pincers to restrain and kill prey, or to prevent their own predation. The scorpion of *Androctonus crassicauda* belongs to the Buthidae family, it has a strong and deadly venom for humans, and every year there are cases of death due to the sting of this scorpion. They are nocturnal and enjoys making nests where they hide in crevices during the day to stay moisturized. The Arabian fat-tailed scorpion (*Androctonus crassicauda*) is a species of dangerous scorpion usually found mainly in the Palaearctic region, in such countries as Turkey, Iran, and other southwestern Asian nations. This scorpion is the second most frequent causes of scorpion sting in south-west Iran. Despite the risks of keeping such a dangerously venomous species in captivity, *Androctonus* scorpions are frequently found in the exotic animal trade. The scorpion's main diet in captivity consists of cockroaches, grasshoppers, and crickets. However, the fat-tailed scorpion can go months without consuming food. Scorpions will generally try to kill and eat anything which moves and is smaller than themselves. This species is a generalist desert species. Adults can vary in colour from a light brown to reddish to blackish-brown, to black. The venom is mainly composed of neurotoxins, cardiotoxins, and possibly myotoxins. The venom is very powerful and considered among the most potent among scorpions, and is known to cause human death. Deaths mostly occur with respiratory arrest, heart failure, and shock. In the past, most studies conducted to classify scorpions and determine phylogenetic and evolutionary relationships were based on phenotype. Recently, molecular markers have been used in genetic diversity studies, paternity analysis, cultivar identification, phylogenetic relationships, and genetic mapping in scorpion genetics research. Various molecular markers are being developed including simple sequence repeats (SSRs), restriction fragment length polymorphisms (RFLPs), random amplification of polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), and single nucleotide polymorphisms (SNPs). Among these, SSRs, are very advantageous for a variety of applications, because they are multi-allele, co-dominant, highly informative, occur with high relative abundance and good coverage across the genome, and are experimentally reproducible. Simple sequence repeats or microsatellites are DNA stretches consisting of short, tandemly repeated di-, tri-, tetra- or penta-nucleotide motifs. Simple sequence repeats have been found in all eukaryotic species that were scrutinized for them. Simple sequence repeats are widely used in genetic fingerprinting, kinship analysis, and population genetics. Also, research show that they may play an important role in genome evolution and provide hotspots of recombination. SSRs

markers are one of most complete molecular tools in the study of genetic diversity due to the appropriate genomic coverage and high repeatability. Simple sequence repeats markers for the scorpion *Andractonus crassicauda* have not been reported so far. Therefore, this study was conducted with the aim of identifying and validating simple sequence repeats markers of the scorpion *Andractonus crassicauda* using RNA-Seq data by De Novo method. De novo assembly of RNA-seq data enables researchers to study transcriptomes without the need for a genome sequence; this approach can be usefully applied, for instance, in research on 'non-model organisms' of ecological and evolutionary importance, cancer samples or the microbiome.

Materials and methods

Scorpions were poisoned by electric shock method. After 72 hours of envenoming, the venom gland was isolated and RNA was extracted from the venom gland using the commercial RNeasy Animal Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. After confirming the quality of extracted RNA samples using Nanodrop and horizontal gel electrophoresis, the samples were sequenced with 150 base pair paired-end reads at Macrogen Company (Macrogen, Seoul, South Korea) using Illumina HiSeq 2000 sequencing platform (Illumina, San Diego, CA, USA) according to the instructions of the company. Checking the quality of raw sequences and clean data before and after editing was done using FASTQC software under the Linux operating system and removing low-quality reads using Trimmomatic v 0.36 software. After filtering, cleaning, and trimming the raw reads generated from the Illumina sequencing platform, clean reads were de novo assembled into contigs using Trinity software (v. 2.0.3) with optimized parameters. This approach is useful when a genome is unavailable, or when a reference-guided assembly is undesirable. In the absence of a sequenced genome to guide the reconstruction process, the transcriptome must be assembled *de novo* using only the information available in the RNA-seq reads. Also, CD-HIT software (version 4.7) was used to cluster the reads and remove duplicate sequences. Next, in order to identify new SSR markers, the transcriptome of the scorpion *Andractonus crassicauda* was analyzed using FullSSR software (version 1.5). To validate the identified SSR markers, 8 pairs of primers were designed using PRIMER 3 software and evaluated by polymerase chain reaction method. Indices such as observed heterozygosity (Ho), expected heterozygosity (he) and fixation index (F) were calculated using GenAEx software (version 6.5) for statistical analysis of polymorphic loci.

Results

To visualize the distribution of quality scores across all bases in a sequencing run, a box whisker-type diagram is often used. In this diagram, the yellow box represents the interquartile

range (25-75%), the upper and lower tentacles account for 10% and 90%, and the blue line indicates the average quality. The y-axis shows the quality score, with higher scores indicating better quality bases. To evaluate the overall sequencing quality of a run, the average sequencing quality per base is often used. A sequencing quality score of 30 or above is generally considered very good. The average data quality of this study was 36, which indicates good data quality. Assembling transcripts by Trinity program generated a total of 744,804 transcripts and 563,526 unigenes. In this study, the examination of 952,725 sequences by FullSSR software led to the identification of 315,395 SSR markers. The results showed that among these, 226,619 sequences contained one SSR and 88,779 sequences contained more than two SSRs. Among SSRs, two and three nucleotide repeats accounted for the highest number of repetitions with 71.85% and 22.36%, respectively. Among the 8 studied loci, only 2 loci showed polymorphism. The expected and observed heterozygosity for the RK1354 locus was calculated as 0.765 and 0.683 respectively, and for the NK1362 locus as 0.768 and 0.633 respectively. The fixation index statistic (F_{is}) for both loci was almost 0.1, which indicates low inbreeding in the population. The chi-square test performed on the population of the scorpion *Andrachtonus crassicoda* showed that Hardy-Weinberg equilibrium was not established for both loci. The chi-square values for loci RK1354 and NK1362 were 59.62 and 56.87, respectively, which were significant at the 1% level.

Conclusions

The different criteria of genetic diversity all indicate that the population has high diversity, but there are factors that are reducing the diversity among the studied population. So that the population is out of balance and the number of observed heterozygotes is less than the expected heterozygotes. It is necessary to investigate the factors that have led to the reduction of diversity in order to prevent its further reduction. In general, the results of this research showed that new SSRs can be useful for understanding the population structure and investigating the genetic diversity of the scorpion *Andrachtonus crassicouda*.

Author Contributions

Conceptualization, M.N. and H.R.; methodology, R.K.; software, F.S.; formal analysis, R.K.; writing—original draft preparation, M.N. and R.K.; writing—review and editing, H.R.; project administration, M.N.; funding acquisition, M.N. All authors have read and agreed to the published version of the manuscript. All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

The RNA-seq reads are available in National Center for Biotechnology Information database under the accessions: Repository name: Raw data of male's venom glands of *A. crassicauda* Data identification number: PRJNA1040487. Direct link: <https://www.ncbi.nlm.nih.gov/sra/PRJNA1040487>

Acknowledgements

We acknowledge the Research deputy of Agricultural Sciences and the Natural Resources University of Khuzestan for all their support.

Ethical Considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

The study was funded by the Agricultural Sciences and the Natural Resources University of Khuzestan, Iran. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.

شناسایی و اعتبارسنجی نشانگرهای توالی ساده تکراری (SSRs) عقرب آندراکتونوس کراسیکودا (*Androctonus crassicauda*) با استفاده از داده‌های RNA-seq

روشن کاکي

دانشجوی دکتری ژنتیک و اصلاح نژاد دام، گروه علوم دامی - دانشکده علوم دامی و صنایع غذایی - دانشگاه علوم کشاورزی و منابع طبیعی خوزستان - ملاتانی - ایران. رایانامه: roashan.kaki@yahoo.com

محمود نظری

*نویسنده مسئول: دانشیار گروه علوم دامی - دانشکده علوم دامی و صنایع غذایی - دانشگاه علوم کشاورزی و منابع طبیعی خوزستان - ملاتانی - ایران. رایانامه: M.Nazari@asnrukh.ac.ir

هدایت اله روشنفکر

استاد گروه علوم دامی - دانشکده علوم دامی و صنایع غذایی - دانشگاه علوم کشاورزی و منابع طبیعی خوزستان - ملاتانی - ایران. رایانامه: roshanfekr_hd@yahoo.com

فاطمه ثعلبی

سازمان تحقیقات آموزش و ترویج کشاورزی - موسسه تحقیقات واکسن و سرم سازی رازی شعبه اهواز، گروه جانوران سمی و تولید پادزهر - اهواز، ایران. رایانامه: f.salabi@rvsri.ac.ir

تاریخ دریافت: ۱۴۰۳/۱۰/۱۵ تاریخ دریافت فایل اصلاح شده نهایی: ۱۴۰۳/۱۲/۰۲ تاریخ پذیرش: ۱۴۰۳/۱۲/۰۳

چکیده

هدف: عقرب آندراکتونوس کراسیکودا (*Androctonus crassicauda*) به عنوان یکی از کشنده‌ترین گونه عقرب در جهان برای پستانداران شناخته شده است. نشانگرهای توالی ساده تکراری (SSR) به دلیل پوشش مناسب ژنومی و تکرارپذیری بالا یکی از کاملترین ابزارهای مولکولی در بررسی تنوع ژنتیکی به شمار می‌آیند. نشانگرهای توالی ساده تکراری برای عقرب آندراکتونوس کراسیکودا تاکنون گزارش نشده‌اند. لذا این مطالعه با هدف شناسایی نشانگرهای توالی ساده تکراری مربوط به عقرب آندراکتونوس کراسیکودا با استفاده از داده‌های RNA-Seq انجام شد.

مواد و روش‌ها: پس از استخراج RNA از غده زهر عقرب، ارزیابی کیفی و کمی انجام گرفت. سپس نمونه‌ها با پلتفورم Illumina HiSeq 2000 توالی‌یابی شدند. برای ویرایش داده‌های خام و رسیدن به دقت و صحت بالاتر از نرم‌افزار Trimmomatic نسخه (0.36) استفاده شد. بازسازی رونوشت‌ها به روش De novo با استفاده از نرم‌افزار Trinity انجام شد. در ادامه به منظور شناسایی نشانگرهای SSR جدید، ترانسکریپتوم عقرب آندراکتونوس کراسیکودا با استفاده از نرم‌افزار FullSSR (نسخه 1.5) مورد آنالیز قرار گرفت. به منظور اعتبارسنجی نشانگرهای SSR شناسایی شده، ۸ جفت پرایمر طراحی و با روش واکنش زنجیره‌ای پلیمرز مورد ارزیابی قرار گرفت.

نتایج: سرهم‌بندی رونوشت‌ها توسط برنامه Trinity، در مجموع ۷۴۴۸۰۴ ترانسکریپت و ۵۶۳۵۲۶ یونی ژن ایجاد کرد. در این مطالعه، بررسی ۹۵۲۷۲۵ توالی به وسیله نرم‌افزار FullSSR، منجر به شناسایی تعداد ۳۱۵۳۹۵ نشانگر SSR شد. در میان SSRها، تکرارهای دو و سه نوکلئوتیدی به ترتیب با ۷۱/۸۵ درصد و ۲۲/۳۶ درصد بیشترین تعداد تکرار را به خود اختصاص دادند. از ۸ جایگاه مورد بررسی ۲ جایگاه چندشکلی را نشان داد. هتروزیگوسیتی مورد انتظار و مشاهده شده برای جایگاه RK1354 به ترتیب ۰/۷۶۵ و ۰/۶۸۳ و همچنین برای جایگاه NK1362 به ترتیب ۰/۷۶۸ و ۰/۶۳۳ محاسبه شد. آماره شاخص تثبیت (F_{is}) برای هر دو جایگاه تقریباً ۰/۱ بود که نشان‌دهنده همخونی کم در جمعیت است.

نتیجه‌گیری: نتایج ارزیابی معیارهای مختلف تنوع ژنتیکی همگی گویای این است که جمعیت از تنوع بالایی برخوردار است، اما عواملی در حال کاهش تنوع در بین جمعیت مورد مطالعه هستند. به طوری که جمعیت از تعادل خارج شده است و تعداد هتروزیگوت‌های مشاهده شده کمتر از هتروزیگوت‌های مورد انتظار شده است. لازم است عواملی که منجر به کاهش تنوع شده است مورد بررسی قرار گیرد تا از کاهش بیشتر آن جلوگیری گردد. به طور کلی، نتایج این تحقیق نشان داد که SSRهای جدید می‌توانند برای شناخت ساختار جمعیت و بررسی تنوع ژنتیکی عقرب اندراکتونوس کراسیکودا مفید باشند.

کلیدواژه‌ها: اندراکتونوس کراسیکودا، ترانسکریپتوم، نشانگرهای توالی ساده تکراری

نوع مقاله: پژوهشی.

استناد: کاکي روشن، نظری محمود، روشنفکر هدایت اله، ثعلبی فاطمه (۱۴۰۴) شناسایی و اعتبارسنجی نشانگرهای توالی ساده تکراری (SSRs) عقرب آندراکتونوس کراسیکودا (*Androctonus crassicauda*) با استفاده از داده‌های RNA-seq. مجله بیوتکنولوژی کشاورزی، ۱۷(۲)، ۳۱-۵۴.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.



© the authors

عقرب‌ها در تمام قاره‌ها پراکنده شده‌اند که این پراکندگی در مناطق گرمسیری و نیمه گرمسیری بیشتر است. حدود ۵۰ گونه از عقرب‌ها اهمیت پزشکی و بهداشتی ویژه‌ای دارند. کشور ایران با توجه به نوع اقلیم و آب و هوا از نظر وجود بندپایان بسیار غنی است. ایران در زمره‌ی کشورهایی است که گونه‌های زیادی از عقرب‌ها به‌ویژه گونه‌های خطرناک مانند همی‌اسکورپیوس لپتوروس و آندروکتونوس کراسیکودا گزارش شده است (Dehghani et al., 2023). موارد مرگ ناشی از گزش در تمام مناطق کشور رخ می‌دهد ولی اکثر مرگ‌ها در استان‌های خوزستان، سیستان و بلوچستان، کرمان، هرمزگان و ایلام اتفاق می‌افتد. استان‌های خوزستان و ایلام به دلیل قرار گرفتن در مرز با کشور عراق و نیز دارا بودن آب و هوای گرم و نیمه گرمسیری از استان‌های مستعد تنوع عقرب هستند (Ghorbani et al., 2021). عقرب‌ها از معدود "فسیل‌های زنده" هستند که در مورفولوژی بسیار محافظه کار هستند و توزیع گسترده‌ای دارند به نحوی که تقریباً در همه انواع زیستگاه‌ها جای دارند. با چنین ویژگی‌های جذابی، عقرب‌ها به عنوان ارگانیسم‌های مدل عالی برای مطالعات فیلوژنیک و تکاملی واجد شرایط هستند.

اندراکتونوس کراسیکودا (*Androctonus crassicauda*) عقرب غیر حفاری است که اندازه آن تا ۱۲ سانتی‌متر می‌رسد. این گونه که عقرب سیاه نیز نامیده می‌شود آندروکتونوس کراسیکودا گونه‌ای است به رنگ قهوه‌ای، خرمایی، زیتونی تا سیاه، شانه‌ها، کلیسرها، انتهای پاها و سطح شکمی عقرب به رنگ زرد روشن می‌باشد. این عقرب دارای نیش بزرگی است که هنگام گزش موجب درد شدیدی می‌شود (Salabi et al., 2024). بنابر مطالعات این عقرب عامل ۲۸/۷ درصد از موارد عقرب‌گزیدگی در استان خوزستان و اولین عامل گزش انسان بوده است (Dehghani et al., 2009) و تقریباً همین سهم از گزش‌های کاشان (۳۰ درصد) را نیز دارد مرگ و میر ناشی از گزش این جانور بیشتر در کودکان مشاهده شده است (Dehghani et al., 2018). این کژدم پراکندگی وسیع در کشور داشته و از نواحی شمال کشور تا استان‌های جنوبی خلیج فارس و تمام استان‌های غرب کشور، کرمان، بوشهر، سمنان و خوزستان (اهواز، امیدیه، بستان، سوسنگرد، ماهشهر، خرمشهر و آبادان)، ایلام (دهلران، مهران و ایوان)، آذربایجان غربی (چالدران، خوی، ماکو، اشنویه و ارومیه)، کردستان (مریوان)، خراسان بزرگ (تایباد، خواف، قائن، بیرجند و نهبندان)، کرمانشاه (جوانرود، سرپل ذهاب و قصر شیرین)، کرمان (بافت، بردسیر، بم، جیرفت، راور، رفسنجان، زرنده، سیرجان، شهر بابک، عنبرآباد، کرمان، کهنوج و منوجان) صید شده است (Ebrahimi et al., 2015). تنوع ژنتیکی یکی از عوامل مهم برای اصلاح‌گران دام است و کمبود یا فقدان این تنوع قدرت انتخاب‌های آینده را کاهش می‌دهد، بنابراین آگاهی از تنوع ژنتیکی و مدیریت منابع ژنتیکی به عنوان اجزای مهم برنامه‌های اصلاح‌دام تلقی می‌شوند (Rohipoor et al., 2019; Rohipoor et al., 2021). در سال‌های اخیر نشانگرهای ژنتیکی به ابزاری قابل اعتماد و مناسب جهت مطالعات ژنتیکی و تنوع جمعیت‌ها تبدیل شده‌اند (Mohammadabadi et al., 2021). در حاضر حال نشانگرهای ریزماهواره به دلیل پوشش مناسب ژنومی و تکرارپذیری بالا یکی از بهترین و کاملترین ابزارهای مولکولی در بررسی تنوع ژنتیکی به شمار می‌آیند (Zane et al., 2002). این توالی‌ها عموماً از چندشکلی بالایی برخوردار هستند و در نواحی کدکننده و غیر کدکننده ژنوم موجودات مختلف حضور دارند

(Zane et al., 2002). ریزماهورها یا توالی‌های ساده تکراری (SSRs)، تکرارهای پشت سر هم متشکل از یک تا شش جفت نوکلئوتید هستند که می‌توانند در هر نقطه از کل ژنوم یافت شوند (Toth et al., 2000). SSRها خنثی در نظر گرفته می‌شدند، اما در سال‌های اخیر، تحقیقات عملکردهای بیولوژیکی آنها را تایید کرده است (Contente et al., 2002; Hefferon et al., 2004). ساختار تکراری SSRها باعث ایجاد رویدادهای لغزش DNA پلیمراز می‌شود که واحدهای تکراری را اضافه یا حذف می‌کند، و نرخ جهش را افزایش می‌دهد و ممکن است باعث بیماری شود (Mirkin, 2007; Ellegren, 2004). بیش از ۳۰ اختلال مندل به جهش‌های SSR نسبت داده می‌شود که بیشتر آنها به دلیل گسترش زیاد تکرارهای کدکننده سه نوکلئوتیدی، مانند بیماری‌های هانتینگتون و ماچادو- جوزف ایجاد می‌شوند (Mirkin, 2007). گسترش یا انقباضات SSR در مناطق کدکننده پروتئین می‌تواند عملکرد ژن را به طور بالقوه تغییر دهد، فنوتیپ‌ها را تغییر دهد و ممکن است با سازگاری سریع با محیط‌های جدید همراه باشد (Trifonov, 2004; Li et al., 2004). ویژگی‌ها و عملکرد توزیع ریزماهور ممکن است در ژنوم‌های منحصر به فرد متفاوت باشد (Li et al., 2004; Wang et al., 2015). علاوه بر این، اپی‌ژنوم شامل مکانیسم‌های مختلفی است، به عنوان مثال متیلاسیون DNA، بازسازی مجدد، اصلاحات دم هیستون، microRNAهای کروماتین و RNAهای بلند غیرکدکننده، با عوامل محیطی مانند تغذیه، عوامل بیماری‌زا، و آب و هوا برای تأثیر بر مشخصات بیان ژن‌ها و ظهور فنوتیپ‌های خاص در تعامل هستند (Barazandeh et al., 2016). تعاملات چند سطحی بین ژنوم، اپی‌ژنوم و عوامل محیطی ممکن است رخ دهد (Amiri Roudbar et al., 2020). همچنین، شواهد متعددی حاکی از تأثیر تنوع اپی‌ژنوم بر سلامت و تولید است (Alavi et al., 2022). بیان ژن‌های یوکاریوتی به طور موقت و چند بعدی کنترل می‌شود. تنها مجموعه نسبتاً کوچکی از کل ژنوم در هر نوع بافت بیان می‌شود و بیان ژن‌ها به مرحله رشد بستگی دارد. بنابراین، بیان ژن در یوکاریوت‌ها برای هر بافت خاص است (Rabieh et al., 2020; Sabahi et al., 2020; Koohgivi et al., 2019; Safaei et al., 2023). همچنین میزان فراورده‌های ژنی که در همان بافت و همچنین سایر بافت‌های سازنده آن محصول ساخته می‌شود، بیان آن ژن را تنظیم می‌کند. یکی از فعالیت‌های اساسی، مطالعه ژن‌ها و پروتئین‌های مرتبط با صفات و بررسی آنها در سطح سلولی یا کروموزومی است (Bordbar et al., 2022; Mohammadabadi et al., 2024). استفاده از داده‌های ترانسکریپتومی بطور قابل توجهی بررسی تنوع ژنتیکی گونه‌ها در سال‌های اخیر مورد توجه قرار گرفته است. آنالیز داده‌های ترانسکریپتومی بطور قابل توجهی بررسی تنوع ژنتیکی بین گونه‌ها را تسهیل بخشیده و اطلاعات جامعی از شناسایی گونه‌ها و ارتباط بین آنها را در اختیار ما قرار می‌دهد. مطالعه ترانسکریپتوم برای درک عناصر کاربردی ژنوم و ارزیابی ساختار مولکولی سلول‌ها و بافت‌ها ضروری است (Mohamadi et al., 2024).

روش توالی‌یابی RNA یا RNA-Seq یک روش قدرتمند بر پایه‌ی توالی‌یابی است که محققان را قادر می‌سازد تا توالی‌یابی و کمیت رونوشت‌های RNA را در یک محصول کامل بیان ژن (ترانسکریپتوم) تعیین کنند. این تکنیک برای بررسی ژنوم تازه توالی‌یابی شده و شناسایی ژن‌های جدید یا ایزوفرم‌های آن در موجوداتی که تفسیرات ژنی هنوز ناکامل مانده بسیار ارزشمند است.

ضمناً به ما امکان می‌دهد تا فعالیت‌های ژنی را در بافت‌های مختلف ارگانیسم‌ها، مراحل و یا تحت شرایط گوناگون بررسی کنیم (Haas et al., 2013). اساس روش توالی‌یابی RNA (RNA-Seq) بر برش کل mRNA رونویسی شده به قطعات کوچکتر و توالی‌یابی این قطعات با استفاده از پلتفرم‌های متنوع می‌باشد. انتخاب نمونه و بافت مناسب با توجه به اهداف بیولوژیکی برای دستیابی به ترانسکرپتوم بسیار مهم است همچنین تکرار بیولوژیکی به منظور افزایش صحت نتایج حائز اهمیت است. (Corney et al., 2016).

توالی‌یابی از نو یا دنوو (De novo) یکی از روش‌های آنالیز داده‌های حاصل از RNA seq است. در این روش، بدون در دسترس داشتن داشتن ژنوم اولیه موجود، قطعات کوتاه ("کانتیگ") می‌تواند قطعات بزرگ را تشکیل داده و توالی ژنوم یا ترانسکرپتوم را بسازد. با استفاده از توالی یاب دنوو می‌توان ژنوم‌های طولانی‌تر و پیچیده‌تر را توالی‌یابی کرد، زیرا این توالی‌یاب‌ها این امکان را می‌دهند که نواحی تکراری در ژنوم شناخته شوند. همچنین با استفاده از این نوع توالی‌یاب‌ها می‌توان ژنوم گونه‌های جدید را توالی‌یابی کرد که با استفاده از آن‌ها می‌توان تغییرات ساختاری مانند اضافه شدن و حذف شدن نوکلئوتیدها را در ژنوم یک گونه شناسایی کرد. این روش در ماهی‌ها (Luca et al., Wen et al., 2024; Mohamadi Ahvazi et al., 2024) و گیاهان (Nemati et al., 2020; Li et al., 2021) به دفعات زیاد استفاده شده است. در یک مطالعه توزیع ریزماهورها (SSR) در ژنوم دو گونه عقرب *Mesobuthus martensii* و *Centruroides exilicauda* مورد بررسی قرار گرفت (Wang et al., 2016). مشخص شد که نشانگرهای توالی ساده تکراری در ابتدا و انتهای اینترون و اگزون ژنها می‌توانند حضور داشته باشند.

همانطور که در بالا اشاره شد نشانگرهای توالی ساده تکراری (SSR) یا ریزماهورها به دلیل پوشش مناسب ژنومی و تکرارپذیری بالا یکی از بهترین و کاملترین ابزارهای مولکولی در بررسی تنوع ژنتیکی به شمار می‌آیند. نشانگرهای توالی ساده تکراری برای عقرب اندراکتونوس کراسیکودا تاکنون گزارش نشده‌اند. لذا این مطالعه با هدف شناسایی و اعتبار سنجی نشانگرهای توالی ساده تکراری مربوط به عقرب اندراکتونوس کراسیکودا با استفاده از داده‌های RNA-Seq به روش اسمبلی دنوو (De novo) انجام شد.

مواد و روش‌ها

در این مطالعه ابتدا به منظور استخراج RNA و بررسی ترانسکرپتوم، نمونه‌گیری از بافت دم سه عقرب اندراکتونوس کراسیکودا در موسسه تحقیقات واکسن و سرم سازی رازی- شعبه جنوب غرب کشور (اهواز) تهیه گردید. استخراج RNA به وسیله محلول TRIZOL انجام شد. پس از استخراج RNA، به منظور بررسی یکپارچگی RNA استخراج شده از الکتروفورز ژل آگارز ۲ درصد با ولتاژ ۱۰۰ ولت به مدت ۳۰ دقیقه استفاده شد. همچنین ارزیابی کیفی RNA با استفاده از دستگاه نانودراپ و خوانش طول موج‌های

۲۶۰ و ۲۸۰ نانومتر و ۲۶۰ به ۲۳۰ نانومتر انجام شد. در نهایت پس از تایید کمیت و کیفیت RNA استخراج شده، مقدار مناسبی از نمونه‌ها به منظور توالی‌یابی به شرکت ماکروژن کره ارسال گردید. فرایند توالی‌یابی به وسیله شرکت ماکروژن و با استفاده از پلتفرم Illumina HiSeq 2000 و بر مبنای دستورالعمل سازنده، برای تولید خوانش ۱۵۰ جفت باز با استفاده از Paired – end انجام شد. بررسی کیفیت داده‌های توالی‌یابی شده قبل و بعد از ویرایش با استفاده از نرم افزار FASTQC تحت سیستم عامل لینوکس انجام شد. این نرم‌افزار علاوه بر فهرست تعداد خوانش‌ها و رمزگذاری کیفیت آنها، اطلاعات مربوط به کیفیت و محتوای بازها، طول خوانش و همچنین حضور بازهای مبهم و مضاعف‌شدگی‌ها را گزارش و مصورسازی نمود. برای ویرایش داده‌های خام RNA-seq و رسیدن به دقت و صحت بالاتر از نرم‌افزار Trimmomatic (نسخه 0.36) استفاده شد. این مرحله شامل حذف آداپتورها و حذف یا ویرایش (کوتاه نمودن) خوانش‌های بی‌کیفیت است. خوانش‌های تمیز پس از تایید کیفیت مجدد با نرم‌افزار FASTQC برای سرهمبندی رونوشت‌ها مورد استفاده قرار گرفتند. در این مطالعه مونتاژ داده‌های تمیز حاصل از توالی‌یابی RNA با استفاده از استراتژی سرهمبندی به روش دنوو و با بکارگیری نرم‌افزار Trinity (ورژن 2.0.3) انجام شد. در یکپارچه‌سازی صورت گرفته حداقل طول رونوشت‌ها ۲۰۰ جفت باز در نظر گرفته شد. برای ارزیابی کیفیت کانتیگ‌های تولید شده و انتخاب بهترین کایمر از شاخص‌های آماری هر اسمبلی از جمله طول N50 (حداقل طولی از کانتیگ‌ها که ۵۰ درصد از کل کانتیگ‌ها دارای این اندازه هستند)، تعداد کل جفت بازها در هر اسمبلی استفاده شد. در نهایت برای کلاستر بندی کانتیگ‌هایی که ۹۵ درصد هم‌پوشانی داشتند از نرم‌افزار CD-HIT (نسخه 4.7) استفاده گردید.

در ادامه به منظور شناسایی نشانگرهای SSR جدید، رونوشت‌های مونتاژ شده با استفاده از نرم افزار FullSSR (نسخه 1.5) مورد آنالیز قرار گرفت. به منظور اعتبارسنجی و تجزیه و تحلیل SSRهای شناسایی شده جدید، تعداد ۶۰ عدد عقب از مناطق مختلف جمع آوری شد. از میان SSRهای بدست آمده، ۸ نشانگر SSR به صورت تصادفی انتخاب و ۸ پرایمر توسط نرم افزار PRIMER 3 از سایت NCBI برای این نشانگرها طراحی شد (جدول ۱).

واکنش زنجیره‌ای پلیمرز با حجم نهایی ۲۵ میکرولیتر و ۳۵ چرخه بر روی نمونه‌های DNA استخراج شده از عقب اندراکتونوس کراسیکودا با استفاده از دستگاه ترموسایکلر Bio Rad و روش Touchdown صورت گرفت. برنامه حرارتی مورد استفاده به شرح زیر می‌باشد. دمای واسرشته‌سازی اولیه ۹۵ درجه سانتی‌گراد به مدت ۴ دقیقه و به دنبال آن ۳۵ چرخه هر کدام شامل ۳۰ ثانیه در ۹۵ درجه سانتی‌گراد، ۳۰ ثانیه در ۶۴-۵۹ درجه سانتی‌گراد، ۳۰ ثانیه در ۷۲ درجه سانتی‌گراد و ۱۰ دقیقه برای تکثیر نهایی در دمای ۷۲ درجه سانتی‌گراد صورت پذیرفت. پس از آن کیفیت محصولات حاصل از تکثیر به وسیله PCR، با استفاده از الکتروفورز ژل آگارز ۳ درصد سنجیده شد. پس از تهیه تصویر ژل‌ها به وسیله دستگاه ژل‌داک، برای سنجش وزن مولکولی محصولات PCR بر حسب جفت باز و تعیین ژنوتیپ‌ها از نرم‌افزار GelAnalyzer (نسخه 19.1) استفاده گردید. در ادامه به منظور تجزیه و تحلیل‌های آماری برای جایگاه‌هایی که چندشکلی در آنها مشاهده شده از جمله تعداد آلل‌های واقعی (Na)، تعداد آلل‌های موثر در هر

لوکوس (Ne)، هتروزیگوسیتی مشاهده شده (Ho) و مورد انتظار (He) و شاخص تثبیت (F_{IS}) از نرم افزار GenAlex (نسخه 6.5) استفاده شد.

جدول ۱. خصوصیات آغازگرهای استفاده شده برای اعتبار سنجی SSRها در این مطالعه

Table 1. Characterization of primers used for validation of SSRs in this study

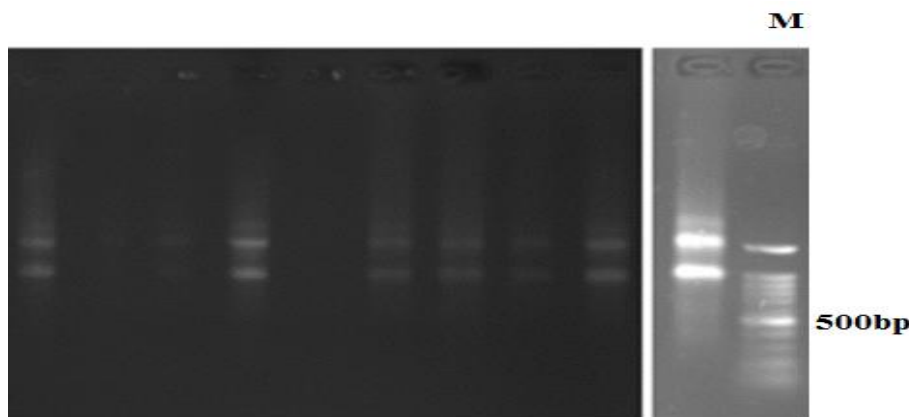
Row	Repetitive sequence	Primer sequence	length piece
1	(ACTC)17	F: TGCTTGAAATTGTGGGGTGC R: AACAGCTCCAAAAACAGTGCC	232
2	(AGAA)14	F: CACCCACTTACGAAGGGATACA R: CTCCTGAGCAAGGATCGGT	295
3	(TCCT)13	F: GCTGAGGACAATAGGGGAGG R: TACACCCTGGCTAAGGCTCT	236
4	(TG)54	F: AACGAGTTCAATGTGTGCGC R: ACACACACACACACACACGT	186
5	(TG)44	F: GTCACGTCAGGTGGTTGTGA R: ACACACAGACAGACACACACA	150
6	(AC)47	F: TGAAGATTGCTTGGCTGTCA R: TGGCTTCGTGGTCATTCTATG	246
7	(AC)56	F: TGCTGTCGTCCTATAATTGGT R: TCCAATCTGTGATAATCAGGAGT	195
8	(AT)97	F: TGTGACTGTTTATTCTGTCAA R: TCTTCGGATCACTCCTCTCA	250

نتایج

الکتروفورز RNA استخراج شده روی ژل آگارز ۲ درصد (شکل ۱) نشان داد که کیفیت RNA های استخراج شده خوب بود. باندهای RNA ریبوزومی ۱۸S و ۲۸S قابل تشخیص بوده و اسمیر میان این دو باند مربوط به mRNA ها بود که نشاندهنده کیفیت مناسب RNA استخراج شده بود. همچنین نتایج نانودراپ نشان داد که نمونه های RNA استخراج شده فاقد آلودگی بودند، زیرا نسبت ۲۶۰ نانومتر: ۲۸۰ نانومتر از ۱/۷۷ تا ۱/۹۰ متغیر بود.

توالی یابی از نمونه های بافت دم عقرب تعداد ۵۱۷۷۹۹۷۰۴ خوانش خام تولید کرد. پس از دریافت داده های حاصل از توالی یابی در فرمت FastQ، برای کنترل کیفیت اولیه از نرم افزار FASTQC استفاده شد. این نرم افزار علاوه بر فهرست تعداد خوانش ها و

رمزگذاری کیفیت آنها، اطلاعات مربوط به کیفیت و محتوای بازها، طول خوانش و همچنین حضور بازهای مبهم و مضاعف‌شدگی‌ها را گزارش و مصورسازی کرد. نتایج تعیین کیفیت داده‌ها در نرم‌افزار FASTQC نشان داد که تمام داده‌های این مطالعه در وضعیت قبول قرار داشتند و میانگین کیفیت داده‌ها ۳۶ بود که بیانگر کیفیت خوب داده‌ها بود (شکل ۲).



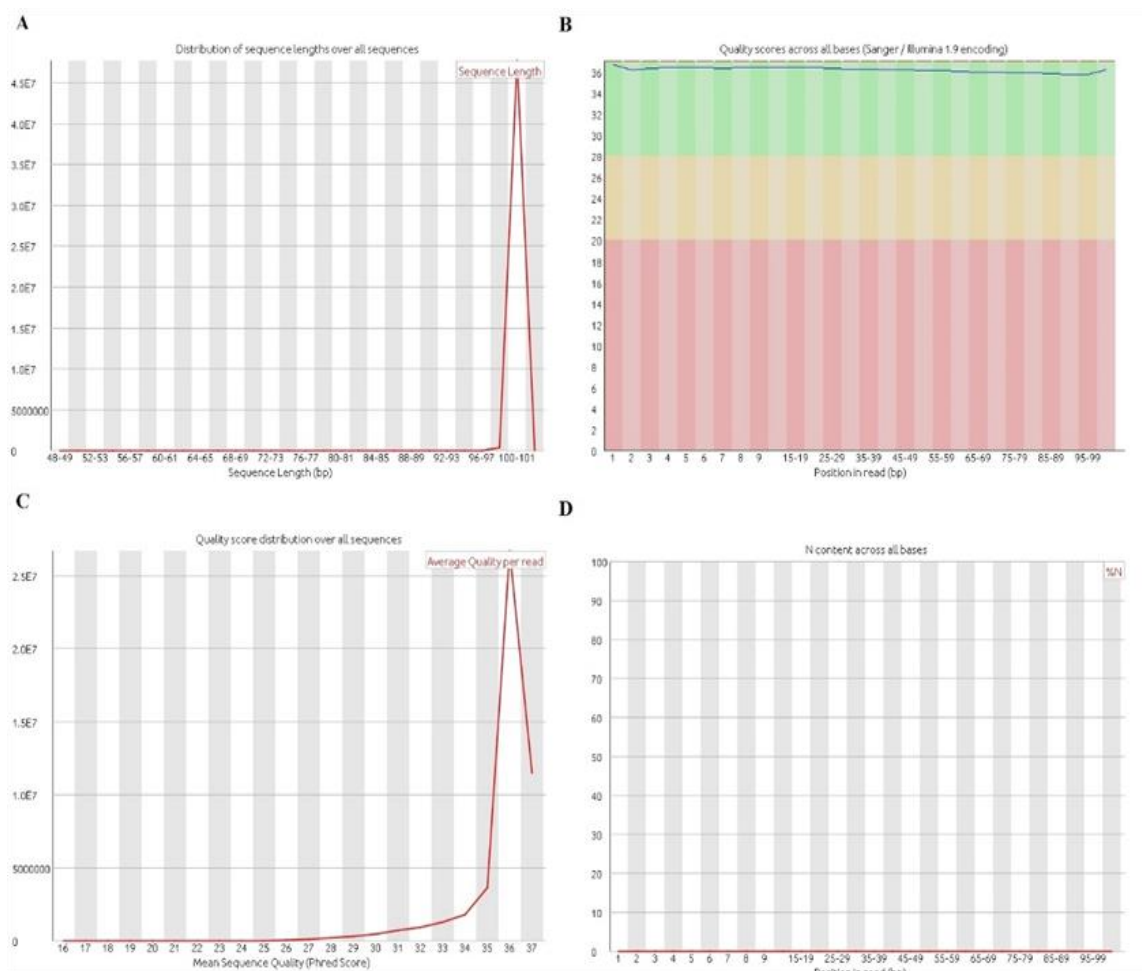
شکل ۱. کیفیت RNA استخراج شده از بافت دم عقرب اندراکتونوس کراسیکودا بر روی ژل آگارز ۲ درصد. M: نشانگر مولکولی ۱۰۰ جفت باز

Figure 1. The Quality of RNA extracted from the tail tissue of the scorpion *Andrachtonus crassicauda*. M: Molecular marker 100 base pair

بازسازی رونوشت‌ها به روش دنوو با استفاده از نرم افزار Trinity انجام شد و نتایج پیکربندی ترانسکریپتوم نشان داد که در مجموع ۷۴۴۸۰۴ کانتیگ و ۵۶۳۵۲۶ یونی ژن ایجاد کرد. میانگین میزان GC آنها ۳۴/۰۱ درصد بود (جدول ۲). در ادامه به منظور شناسایی نشانگرهای SSR جدید، ترانسکریپتوم عقرب اندراکتونوس کراسیکودا با استفاده از نرم افزار FullSSR (نسخه ۱.۵) مورد آنالیز قرار گرفت. در نهایت با بررسی تعداد ۹۵۲۷۲۵ توالی تعداد ۳۱۵۳۹۵ نشانگر SSR جدید شناسایی گردید. نتایج نشان داد از این میان ۲۲۶۶۱۹ توالی حاوی یک SSR و تعداد ۸۸۷۷۹ توالی دارای بیش از دو SSR بود (جدول ۳). با توجه به نتایج گزارش شده از بررسی ۹۵۲۷۲۵ توالی تکرارهای دو و سه نوکلئوتیدی به ترتیب با ۷۱/۸۵ درصد و ۲۲/۳۶ درصد بیشترین تعداد تکرار را به خود اختصاص دادند و توالی شش نوکلئوتیدی به میزان ۰/۰۲ درصد کمترین تعداد را داشتند (جدول ۴).

به منظور اعتبارسنجی نشانگرهای SSR شناسایی شده در عقرب اندراکتونوس کراسیکودا ۸ جفت پرایمر طراحی و با روش واکنش زنجیره‌ای پلیمرز تکثیر شدند. استخراج DNA از نمونه بافت دم عقرب با موفقیت انجام گرفت تک باند مشاهده شده در ژل استخراج DNA نمونه‌ها نشانگر صحت و کیفیت استخراج DNA از نمونه‌های مورد نظر بود. پس از انجام واکنش PCR بر روی نمونه‌های استخراج شده، محصولات حاصل از آن با استفاده از ژل آگارز ۳ درصد مورد بررسی قرار گرفتند. از ۸ پرایمر استفاده شده در این تحقیق ۲ جایگاه (پرایمر شماره یک و پرایمر شماره سه) چند شکلی نشان دادند و بقیه جایگاه‌ها تک باند بودند. در این

تحقیق، این جایگاه‌ها برای پرایمر شماره ۱ به صورت RK1354 و برای پرایمر شماره ۳ به صورت NK1362 نامگذاری شد. دامنه آلی مشاهده شده برای جایگاه RK1354 بین ۲۳۲ تا ۴۳۵ جفت نوکلئوتید و برای جایگاه NK1362 بین ۳۲۶ تا ۴۲۴ جفت نوکلئوتید بود.



شکل ۲. کیفیت توالی RNA-seq از مجموعه داده‌های *Androctonus crassicauda* توسط نرم افزار FASTQC (A) توزیع طول خوانش‌ها (B) کیفیت نوکلئوتیدهای خوانده شده (C) کیفیت توالی‌های خوانده شده (D) محتوای نوکلئوتیدهای خوانش نشده

Figure 2. The Quality of RNA-seq sequences from *Androctonus crassicauda* data by Fastqc software. A: Read length distribution B) Quality of read nucleotides C) Quality of read sequences D) Content of unread nucleotides

پارامترهای برآورد شده برای محاسبه تنوع ژنتیکی از جمله میزان هتروزیگوسیتی مشاهده شده، هتروزیگوسیتی مورد انتظار، شاخص تثبیت، تعداد آلل‌های واقعی، اندازه آلل‌های موثر و تعداد نمونه برای جایگاه دارای چندشکلی در جدول ۵ ارائه شده است.

جدول ۲. نتایج حاصل از رونوشت‌های اسمبل شده در نرم‌افزار Trinity

Table 2. The results of assembled transcripts in Trinity software

statistics	Trinity
Number of genes	563526
Number of transcripts	744804
GC percentage	34.10
Average coating	471.18
Total assembled bases	350938383

جدول ۳. نشانگرهای SSR شناسایی شده با استفاده از نرم‌افزار FullSSR

Table 3. SSR markers detected using the Full SSR software

Checked items	Number
Total examined sequences	952725
Total detected SSR	315395
Number of sequences with one SSR	226619
Number of sequences with more than one SSR	88779

جدول ۴. خلاصه‌ای از نوع و تعداد موتیف‌های تکراری

Table 4. Summary of the type and number of repetitive motifs

Unit size	Number of SSR	Percentage
Di-nucleotide	226619	71.85
Tri-nucleotide	70523	22.36
Tetra-nucleotide	16694	5.29
Penta-nucleotide	1509	0.48
Hexa-nucleotide	50	0.02

مقدار هتروزیگوسیتی مشاهده شده برای جایگاه RK1354 برابر با ۰/۶۸۳ و برای جایگاه NK1362 برابر با ۰/۶۳۳ بود و

مقدار هتروزیگوسیتی مورد انتظار برای جایگاه RK1354 برابر با ۰/۷۶۵ و برای جایگاه NK1362 برابر با ۰/۷۶۸ بود. مقدار آماره

شاخص تبیت (Fis) برای جایگاه‌های RK1354 و NK1362 به ترتیب ۰/۱۰۷ و ۰/۱۷۵ محاسبه شد.

تست کای مربع انجام شده بر روی جمعیت عقرب اندراکتونوس کراسیکودا عدم برقراری تعادل هاردی-وینبرگ را برای هر دو جایگاه نشان داد. مقدار کای مربع برای جایگاه‌های RK1354 و NK1362 به ترتیب $59/62$ و $56/87$ بدست آمد که در سطح ۱ درصد معنی‌دار بود.

جدول ۵. پارامترهای آماری برآورد شده برای محاسبه تنوع ژنتیکی جایگاه RK1354 و NK1362 در جمعیت عقرب اندراکتونوس کراسیکودا

Table 5. Statistical parameters estimated to calculate the genetic diversity of the locus RK1354 and NK1362 in the *Androctonus crassicauda* scorpion population

Loci	N	No. of Different Alleles	No. of Effective Alleles	Observed Heterozygosity	Expected Heterozygosity	Fixation Index (F_{IS})
RK1354	60	5	4.256	0.683	0.765	0.107
NK1362	60	5	4.301	0.633	0.768	0.175

بحث

عقرب آندروکتونوس کراسیکودا یک گونه عقرب مهم در زمینه عقرب‌زدگی است و مرگ و میر ناشی از آن در استان خوزستان می‌باشند. شناسایی و دسته‌بندی عقرب‌های استان خوزستان به دلیل فراوانی میزان تنوع گونه‌های عقرب در این استان از اهمیت ویژه‌ای برخوردار است. پیش از این، بررسی تنوع گونه‌های عقرب بیشتر بر اساس خصوصیات مورفولوژیکی انجام می‌شده است. اما با توجه به خطاهای بسیار این روش محققین به بررسی تنوع گونه‌های عقرب با استفاده از ژن‌ها پرداختند. نشانگرهای زیستی مختلف برای بررسی روابط فیلوژنتیک گونه‌های مختلف ابداع و مورد استفاده قرار گرفته است. به عنوان مثال نشانگرهای میکروستلایت (SSR)، ژن‌های هسته، ژنهای میتوکندری و غیره است. DNA میتوکندری (mtDNA) به طور گسترده در مطالعات فیلوژوگرافی و فیلوژنتیک استفاده شده است. در مطالعات بسیاری شناسایی ژنتیکی عقرب و تنوع ژنتیکی عقرب آندراکتونوس کراسیکودا، با استفاده از ژن‌های میتوکندری انجام شده است (Ozkan et al., 2010).

یک ابزار سریع و مقرون به صرفه برای شناسایی نشانگرهای مولکولی چندشکلی، به ویژه SSRها، در مقایسه با روش‌های جداسازی سنتی، توالی‌یابی ترانسکریپتوم است (Zane et al., 2002; Mohamadi ahvazi et al., 2024). بر خلاف ژنوم که در تمامی سلول‌های یک موجود زنده ثابت است (به استثنای جهش یافته‌ها)، تحت شرایط مختلف الگوی بیان ژن‌ها تغییر خواهد

کرد (Dunlop, 2010). داده‌های ترانسکریپتومی که مجموعه‌ای از مولکول‌های (mRNA) هستند به دلیل تکرار پذیری بالا و میزان خطای کم، بهترین نتایج را در تمایز گونه‌ها داشته‌اند و در این تحقیق به عنوان یک نشانگر نوآورانه و جدید مورد بررسی قرار گرفته‌اند. لذا در این تحقیق نشانگرهای SSR مربوط به گونه عقرب اندراکتونوس کراسیکودا با استفاده از داده‌های RNA-Seq شناسایی شده و جهت صحت کار مورد ارزیابی قرار گرفتند. امروزه نشانگرهای SSR یک رویکرد موثر برای شناسایی ارتباطات فنوتیپی و همچنین منبع مناسبی برای ایجاد نقشه‌های ژنتیکی، تجزیه و تحلیل تنوع ژنتیکی و انتخاب به کمک نشانگر مولکولی است (Mohamadi Ahvazi et al., 2024). این نشانگرها به دلیل سطوح بالای چندشکلی، پرکاربردترین نشانگرهای مولکولی در مطالعات ژنتیک جمعیت بشمار می‌روند (Li et al., 2004).

در این پژوهش تعداد ۳۱۵۳۹۵ نشانگر SSR بر اساس یونی ژن‌های اسمبل شده حاصل از توالی‌یابی ترانسکریپتوم عقرب اندراکتونوس کراسیکودا شناسایی شد. که حدود ۳۳/۱ درصد از توالی‌های مورد بررسی حاوی SSR بودند. از بررسی ۹۵۲۷۲۵ SSRهای دو و سه نوکلئوتیدی به ترتیب با ۷۱/۸۵ درصد و ۲۲/۳۶ درصد بیشترین تعداد تکرار را به خود اختصاص دادند. در مطالعه عقرب‌های *Centruroides exilicauda* و *Mesobuthus martensii*، نشانگر SSRهای سه نوکلئوتیدی فراوان‌ترین انواع SSR بودند و SSRهای پنج نوکلئوتیدی به میزان ۰/۰۲ درصد کمترین تعداد را داشتند (Wang et al., 2016).

در این پژوهش از بین SSRهای شناسایی شده در عقرب اندراکتونوس کراسیکودا تنوع ژنتیکی در ۸ جایگاه مورد بررسی و ارزیابی قرار گرفت. نتایج ژل الکتروفورز حضور هشت SSR را تایید کرد. از میان SSRهای مورد بررسی، ۶ جایگاه فاقد چند شکلی بودند و فقط ۲ جایگاه چند شکلی داشتند. بر طبق نتایج ارائه شده در جدول ۵، هر دو جایگاه از تعادل هاردی-وینبرگ انحراف دارد و هتروزیگوسیتی مورد انتظار و مشاهده شده برای جایگاه RK1354 به ترتیب ۰/۷۶۵ و ۰/۶۸۳ و همچنین برای جایگاه NK1362 به ترتیب ۰/۷۶۸ و ۰/۶۳۳ بود. مطابق با این نتایج، در مطالعه ای که بر روی هشت جایگاه ریز ماهواره دو و سه نوکلئوتیدی در عقرب چینی *Mesobuthus martensii* انجام شد میزان هتروزیگوسیتی مورد انتظار در این جایگاه‌ها از ۰/۱۹ تا ۰/۸۶۰ و هتروزیگوسیتی مشاهده شده از ۰/۱۹ تا ۰/۴۹۲ متغیر بود و تعدادی از جایگاه از تعادل هاردی-وینبرگ انحراف داشتند. این محققین کمبود آشکار هتروزیگوت‌ها را به این دلیل دانستند که نمونه‌ها از چندین جمعیت که یک منطقه جغرافیایی بزرگ را پوشش می‌دادند جمع‌آوری شده بودند (Ji et al., 2008). در مطالعه حاضر، بالا بودن میزان هتروزیگوسیتی مشاهده شده در این جایگاه‌ها بیانگر سطح بالای تنوع در جمعیت است. بالا بودن میزان هتروزیگوسیتی می‌تواند به علت نبودن انتخاب در گزینش والدین، وجود جهش در آلل‌ها و همچنین پراکندگی در محیط زندگی طبیعی که باعث کم شدن میزان همخونی و افزایش میزان هتروزیگوسیتی می‌شود. اگر تمام اعضای یک جمعیت الگوی ژنتیکی یکسانی را برای یک صفت خاص داشته باشند آن جمعیت از نظر صفت یاد شده هموزیگوت است و اگر برای یک ویژگی خاص بیش از یک الگوی ژنتیکی در جمعیتی مشهود باشد، آن جمعیت را از نظر صفت مذکور، هتروزیگوت به شمار می‌آورند. با افزایش هتروزیگوتی، تنوع ژنتیکی زیاد می‌شود. بررسی تنوع آلی در یک جمعیت مبحثی به نام تنوع ژنومی را مطرح می‌کند که در سطح مولکولی در توالی‌ها منعکس می‌شود (Howard et al., 2019).

یکی از شاخص‌های مهم در بررسی‌های تنوع ژنتیکی شاخص تثبیت (Fis) است. این شاخص بیانگر متوسط تفاوت میزان هتروزیگوسیتی مشاهده بر اثر عدم آمیزش تصادفی در یک زیر جمعیت است که مقدار همبستگی بین موقعیت‌های دو آلل که به طور تصادفی از هریک از زیر جمعیت‌ها انتخاب شوند را بیان می‌دارد (Shakeri et al., 2021). شاخص تثبیت برای جمعیت حدود ۰/۱ محاسبه گردیده است. این به معنای کم بودن همخونی در جمعیت است و نشان‌دهنده بالا بودن تنوع در جامعه است. در این پژوهش تعداد آلل‌های واقعی ۵ عدد به دست آمد. تعداد آلل‌های واقعی در آزمایش‌های مختلف تغییر می‌کند و به عنوان معیاری برای بررسی تنوع، به شدت و اندازه نمونه بستگی دارد و در آزمایشات با تعداد نمونه‌های متفاوت، تعداد آلل‌های مختلف به دست می‌آید (Mohamadi ahvazi et al., 2019). نتایج به دست‌آمده از برآورد معیارهای تنوع درون جمعیتی مانند شاخص تثبیت، تعداد آلل مؤثر، هتروزیگوسیتی مورد انتظار و هتروزیگوسیتی مشاهده شده نشان‌دهنده تنوع مناسب در جمعیت‌های مورد مطالعه بود که می‌توان آن‌ها را به عنوان یک ذخیره ژنتیکی مناسب برای اهداف مختلف پرورشی و پزشکی در کشور به حساب آورد. کمتر شدن مقدار هتروزیگوسیتی مشاهده نشده از هتروزیگوسیتی مورد انتظار نشان‌دهنده این است که عواملی در حال کاهش تنوع در بین جمعیت مورد مطالعه هستند. به طوری که جمعیت از تعادل خارج شده است و تعداد هتروزیگوت‌های مشاهده شده کمتر از هتروزیگوت‌های مورد انتظار شده است.

نتیجه گیری: در این پژوهش بر اساس داده‌های RNA-Seq، نشانگرهای SSR عقرب اندراکتونوس را شناسایی، معرفی و ارزیابی شد. نتایج ارزیابی معیارهای مختلف تنوع ژنتیکی همگی گویای این است که جمعیت از تنوع بالایی برخوردار است اما عواملی در حال کاهش تنوع در بین جمعیت مورد مطالعه هستند. به طوری که جمعیت از تعادل خارج شده است و تعداد هتروزیگوت‌های مشاهده شده کمتر از هتروزیگوت‌های مورد انتظار شده است. لازم است عواملی که منجر به کاهش تنوع شده است مورد بررسی قرار گیرد تا از کاهش بیشتر آن جلوگیری گردد. به طور کلی، نتایج این تحقیق نشان داد که نشانگرهای SSR جدید می‌توانند برای شناخت ساختار جمعیت و بررسی تنوع ژنتیکی عقرب اندراکتونوس کراسیکودا مفید باشند.

سپاسگزاری: از معاونت پژوهشی دانشگاه علوم کشاورزی و منابع طبیعی خوزستان به دلیل حمایت‌های مالی برای انجام

این پژوهش تشکر بعمل می‌آید.

منابع

- شاکری رئوف، جوانمرد آرش، حسن پور کریم، عباسی مختار علی و همکاران (۱۴۰۰) بررسی شاخص‌های تنوع ژنتیکی در یک جمعیت از گاو هلشتاین با استفاده از نشانگرهای متراکم اسنپ پژوهشهای تولیدات دامی ۱۲(۳۲)، ۱۴۹-۱۴۰.
- صباحی راضیه، نظری محمود، بیگی نصیری محمدتقی، قربانی محمدرضا (۱۳۹۹) اثر پودر میوه گیاه پنج‌انگشت بر بیان ژن GnRH هیپوتالاموس مرغان تخمگذار. پژوهشهای تولیدات دامی ۱۱(۳۰)، ۱۰۰-۹۲.

- کوه گیوی کی آرام، روشنفکر هدایت الله، نظری محمود، طاطار احمد (۱۳۹۸) تاثیر جایگزینی DL- متیونین با L- متیونین و سطوح مختلف پروتئین جیره بر بیان ژن میوستاتین در بلدرچین ژاپنی. مجله بیوتکنولوژی کشاورزی ۱۱(۲)، ۳۶-۲۳. 1
- ربیع موله، روشنفکر هدایت الله، نظری محمود، قربانی محمدرضا (۱۴۰۰) بررسی بیان ژن آنزیمهای آنتی اکسیدانی در جوجههای گوشتی تغذیه شده با ویتامین E، عصاره گیاه خرفه و پسته وحشی تحت استرس گرمایی. نشریه دامپزشکی ایران ۱۷(۲)، ۵۱-۶۰.
- روحی پور مرضیه، نظری محمود، بیگی نصیری محمدتقی (۱۳۹۸) آنالیز ژنتیکی و فیلوژنیک جمعیت بز عدنی بر اساس سیتوکروم B. پژوهشهای تولیدات دامی ۱۰(۲۶)، ۸۴-۸۹.
- روحی پور مرضیه، نظری محمود، بیگی نصیری محمدتقی (۱۳۹۹) تعیین ساختار، روابط فیلوژنتیک و تنوع ژنتیکی ناحیه کنترل میتوکندریایی بز عدنی. مجله ژنتیک نوین ۱۵(۴)، ۲۹۷-۳۰۴.
- محمدی اهوازی غزال، بیگی نصیری محمدتقی، نظری محمود، صدر آیه سادات (۱۴۰۳) شناسایی و اعتبارسنجی نشانگرهای توالی ساده تکراری (SSR) برای ماهی گطان (*Luciobarbus xanthopterus*) بر پایه دادههای RNA-Seq. مجله علمی شیلات ایران ۳۳ (۱)، ۲۷-۱۵.
- محمدی اهوازی غزال، نظری محمود، محمدآبادی محمدرضا، حیدری راضیه (۱۳۹۸) آنالیز ژنتیکی و فیلوژنتیکی ناحیه HVR1 میتوکندری در سه نژاد گوسفند ایرانی. مجله ژنتیک نوین ۱۴(۳)، ۲۱۱-۲۱۹.
- محمدآبادی محمدرضا، قاسمی میمنی مهرداد، منتظری مهدیه (۱۴۰۰) بررسی تنوع ژنتیکی شترهای بومی شمال استان کرمان با استفاده از آماره های F. مجله اصلاح و به نژادی دام ۱(۲)، ۱۷-۵.
- محمدآبادی محمدرضا، کرد محبوبه، نظری محمود (۱۳۹۷) مطالعه بیان ژن لپتین در بافت های مختلف گوسفند کرمانی با استفاده از Real Time PCR. مجله بیوتکنولوژی کشاورزی ۱۰(۳)، ۱۲۳-۱۱۱.
- نعمتی ایمان، صدقی محمد، حسینی سالکده قاسم، توکل افشاری رضا، تقوی محمدرضا (۱۳۹۸) توالی یابی de novo ترانسکریپتوم و آنالیز عملکردی ژن های با بیان متفاوت بذرهای بزرگ و کوچک گیاه توق خاردار (*Xanthium strumarium L.*) در دوره نمو. علوم و فناوری بذر ایران ۸(۲)، ۲۲۷-۲۱۱.

References

- Alavi, M., Mozafari, M.R., Ghaemi, S., Ashengroph M., Hasanzadeh Davarani, F., & Mohammadabadi, M. (2022). Interaction of Epigallocatechin Gallate and Quercetin with Spike Glycoprotein (S-Glycoprotein) of SARS-CoV-2: In Silico Study. *Biomedicines*, 10(12), e3074. <https://doi.org/10.3390/biomedicines10123074>.
- Amiri Roudbar, M., Mohammadabadi, M.R., Ayatollahi Mehrgardi, A., Rostami, A. A. Momen, M., Morota, G., Brito Lopes, F., Gianola, D., & Rosa, G. (2020). Integration of single nucleotide variants and whole-genome DNA methylation profiles for classification of

- rheumatoid arthritis cases from controls. *Heredity*, 124, 658-674. <https://doi.org/10.1038/s41437-020-0301-4>.
- Barazandeh, A., Mohammadabadi, M.R., Ghaderi-Zefrehei M., & Nezamabadi-Pour, H. (2016). Genome-wide analysis of CpG islands in some livestock genomes and their relationship with genomic features. *Czech Journal of Animal Science*, 61(11), 487–495. <https://doi.org/10.17221/78/2015-CJAS>
- Bordbar, F., Mohammadabadi, M., Jensen, J., Xu, L., Li, J., & Zhang, L. (2022). Identification of candidate genes regulating carcass depth and hind leg circumference in simmental beef cattle using Illumina Bovine Beadchip and next-generation sequencing. *Animals*, 12(9), e1103. <https://doi.org/10.3390/ani12091103>.
- Contente, A., Dittmer, A., Koch, M.C., Roth, J., & Dobbelsstein, M. A. (2002). A polymorphic microsatellite that mediates induction of PIG3 by p53. *Nature Genetics*, 30(3), 315–320. <https://doi.org/10.1038/ng836>.
- Corney, D.C., & Basturea, G.N. (2016). RNA-seq using next generation sequencing. *Materials and Methods*, 3, 203. <https://doi.org/10.13070/mm.en.3.203>
- Dehghani, R., Djadid, N.D., Shahbazzadeh, D., & Bigdelli, S. (2009). Introducing *Compsobuthus matthiesseni* (Birula, 1905) scorpion as one of the major stinging scorpions in Khuzestan, Iran. *Toxicon*, 54 (3), 272-5. <https://doi.org/10.1016/j.toxicon.2009.04.011>.
- Dehghani, R., Ghorbani, A., Varzandeh, M., & Karami-Robati, F. (2023). Toxicity Mechanism of Dangerous Scorpion Stings in Iran. *Journal of Arthropod-Borne Disease*, 17(2), 105-119. <https://doi.org/10.18502/jad.v17i2.13616>
- Dehghani, R., Khobdel, M., & Sabati, H. (2018). Control of scorpions in military environments: a review study. *Journal of Military Medicine*, 20(1), 3-13. <http://eprints.bmsu.ac.ir/id/eprint/703>
- Dunlop, J.A. (2010). Geological history and phylogeny of Chelicerata. *Arthropod Structure & Development*, 39, 124–142. <https://doi.org/10.1016/j.asd.2010.01.003>
- Ebrahimi, M., Azizi, K., Moemenbellah-Fard, M.D., Fakoorziba, M.R., & Soltani, A. (2015). Morphometry Indices of the Black Fat-tailed Scorpion *Androctonus crassicauda* (Scorpiones Buthidae), from Fars Province, Southern Iran. *Journal of Entomology*, 12(1), 39-47. <https://doi.org/10.3923/je.2015.39.47>
- Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, 5, 435–445. <https://doi.org/10.1038/nrg1348>
- Ghorbani, A., Mansouri, B., & Baradaran, M. (2021). Effects of climate variables on the incidence of scorpion stings in Iran for five years. *Journal of Venomous Animals and Toxins including Tropical Diseases*, 27, e20200110. <https://doi.org/10.1590/1678-9199-JVATITD-2020-0110>.

- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M. MacManes, M.D. Ott, M.m Orvis, J., & Pochet, N. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, 8, 1494-1512. <https://doi.org/10.1038/nprot.2013.084>
- Hefferon, T.W., Groman, J.D., Yurk, C.E., & Cutting, G.R. (2004). A variable dinucleotide repeat in the CFTR gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing. *Proceedings of the National Academy of Sciences of USA*, 101(10), 3504–35. <https://doi.org/10.1073/pnas.0400182101>.
- Howard, R.J., Edgecombe, G.D., Legg, D.A., Pisani, D., & Lozano-Fernandez, J. (2019) Exploring the evolution and terrestrialization of scorpions (Arachnida: Scorpiones) with rocks and clocks. *Organisms Diversity Evolution*, 19, 71–86. <https://doi.org/10.1007/s13127-019-00390-7>
- Koohgivi, K. A., Rooshanfekr, H. A., Nazari, M., & Tatar, A. (2019). Effect of DL- Methionine Replacement with L- Methionine and Different Dietary Protein Levels on Myostatin Gene Expression in Japanese Quails. *Agricultural Biotechnology Journal*, 11(2), 23-36. (In Persian). <https://doi.org/10.22103/jab.2019.13274.1099>
- Li, Y.C., Korol, A.B., Fahima, T., & Nevo, E. (2004). Microsatellites within genes: structure, function, and evolution. *Molecular Biology and Evolution*, 21(6), 991–1007. <https://doi.org/10.1093/molbev/msh073>
- Luca, M., Danilo, B., Giorgia, G., Adriana, C., & Oliana, C. (2021). De novo transcriptome assembly, functional annotation and characterization of the Atlantic bluefin tuna (*Thunnus thynnus*) larval stage. *Marine Genomics*, 58, 100834. <https://doi.org/10.1016/j.margen.2020.100834>
- Mirkin, S. (2007). Expandable DNA repeats and human disease. *Nature*, 447, 932–940. <https://doi.org/10.1038/nature05977>
- Mohamadi Ahvazi, G., Beigi Nassiri, M.T., Nazari, M., & Sadr. A.S. (2024). Identification and validation of simple sequence repeats (SSR) markers for yellowfin barbel (*Luciobarbus xanthopterus*) using RNA-Seq data. *Iranian Scientific Fisheries Journal*, 33 (1), 15-27 (In Persian). <https://doi.org/10.22092/ISFJ.2024.131190>
- Mohamadi ahvazi, G., Nazari, M., Mohamadabadi, M.R., & Heidari R. (2019). Genetic and phylogenetic analysis of mitochondrial HVR1 region in three breeds of Iranian sheep. *Modern Genetic Journal*, 14(3), 211-219 (In Persian). <http://mg.genetics.ir/article-1-110-en.html>
- Mohammadabadi, M., Babenko Ivanivna, O., Borshch, O., Kalashnyk, O., Ievstafiieva, Y., & Buchkovska, V. (2024). Measuring the relative expression pattern of the UCP2 gene in

- different tissues of the Raini Cashmere goat. *Agricultural Biotechnology Journal*, 16(3), 317-332. <https://doi.org/10.22103/jab.2024.24337.1627>
- Mohammadabadi, M., Kord, M., & Nazari, M. (2018). Studying expression of leptin gene in different tissues of Kermani Sheep using Real Time PCR. *Agricultural Biotechnology Journal*, 10(3), 111-123 (In Persian). https://jab.uk.ac.ir/article_2211.html?lang=en
- Mohammadabadi, M., Ghasemi Meymandi, M., & Montazeri, M. (2021). The Study of genetic diversity of camels in north of Kerman province using F statistics. *Breeding and Improvement of Livestock*, 1(2), 5-17 (In Persian). <https://doi.org/10.22034/bilj.2021.144024>
- Nemati, I., Sedghi, M., Hoseini Salekdeh, G., Tavakkol afshari, R., & Naghavi, M. (2020). De novo transcriptome sequencing and functional annotation of differentially expressed genes in large and small seeds of common cocklebur (*Xanthium strumarium* L.) during seed development. *Iranian Journal of Seed Science and Technology*, 8(2), 211-227 (In Persian). <https://doi.org/10.22034/ijssst.2019.122712.1220>
- Ozkan, O., Ahmet, C., & Zafer, K. (2010). A study on the genetic diversity of *Androctonus Crassicauda* from Turkey. *Journal of Venomous Animals and Toxins including Tropical Diseases*, 16(4), 599-606. <https://doi.org/10.1590/S1678-91992010000400010>
- Li, Q., Su, X., Ma, H., Du, K., Yang, M., Chen, B., Fu, S., Fu, T., Xiang C., Zhao, Q., & Xu, L. (2021). Development of genic SSR marker resources from RNA-seq data in *Camellia japonica* and their application in the genus *Camellia*. *Scientific Reports*, 11, 9919. <https://doi.org/10.1038/s41598-021-89350-w>
- Rabieh, M.H., Rooshanfekr, M., Nazari, M., & Ghorbani, M.R. (2020). Gene expression of antioxidant enzymes fed wild pistachio (*Pistachio atlantica*), purslane (*portulaca oleracea*) extract and vitamin E under in broiler chickens under heat stress condition. *Iranian Veterinary Journal*, 17(2), 51-60 (In Persian). <https://doi.org/10.22055/IVJ.2019.154392.2084>
- Rohipoor, M., Nazari, M., & Beigi Nassiri, M.T. (2019). Genetic and Phylogenetic Analysis of Adani Goat Population Based on Cytochrome B Gene. *Research on animal production*, 10(26), 84-89 (In Persian). <https://doi.org/10.29252/rap.10.26.84>
- Rohipoor, M., Nazari, M., & Beigi Nassiri, M.T. (2021). Population structure, Genetic diversity and phylogenetic analysis of control region of mtDNA in Adani goat breed. *Modern Genetic Journal*, 15(4), 297-304 (In Persian). <http://mg.genetics.ir/article-1-94-en.html>
- Sabahi, R., Nazari, M., Beigi Nassiri, M.T., & ghorbani, M.R. (2020). The Effect of Vitex Agnuse Castus Fruit Powder on Hypothalamic GnRH Gene Expression in Laying Hens. *Research on Animal Production*, 11(30), 92-100 (In Persian). <https://doi.org/10.52547/rap.11.30.92>

- Safaei, S.M.H., Dadpasand, M., Mohammadabadi, M., Atashi, H., Stavetska, R., Klopenko, N., & Kalashnyk, O. (2022). An origanum majorana leaf diet influences myogenin gene expression, performance, and carcass characteristics in Lambs. *Animals (Basel)*, 13(1), 14. <https://doi.org/10.3390/ani13010014>
- Salabi, F., Nemati, M., & Lari Baghal, M. (2024). Fractionation and enzymes activity measurement of the scorpion *Androctonus crassicauda* (Scorpionida: Buthidae) venom. *Iranian Veterinary Journal*, 20(1), 109-120. <https://doi.org/10.22055/IVJ.2023.422895.2646>
- Shakeri, R., Javanmard, A., Hasanpur, K., Abbasi, M., Mazlom, S.M., Khansefid, M., & Rahimi Varposhti, M. (2021). Assessment of genetic diversity within Holstein population using bovine SNP chip data. *Research on Animal Production*, 12(32), 140-149 (In Persian). <https://doi.org/10.52547/rap.12.32.140>
- Toth, G., Gaspari, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, 10(7), 967–981. <https://doi.org/10.1101/gr.10.7.967>.
- Trifonov, E.N. (2004). Tuning function of tandemly repeating sequences: A molecular device for fast adaptation. Springer Nature, Netherlands, pp 115-138.
- Wang, C., Du, L.M., Li, P., Yang, M., Li, W., & Shen, Y. (2015). Distribution patterns of microsatellites in the genome of the german cockroach (*Blattella germanica*). *Acta Entomologica Sinica*, 58(10), 1037–1045.
- Wang, C., Kubiak, L., Du, L., Li, W., Jian, Z., Tang, C., Fan, Z., Zhang, X., & Yue, B. (2016). Comparison of microsatellite distribution in genomes of *Centruroides exilicauda* and *Mesobuthus martensii*. *Gene*, 594(1), 41-46. <https://doi.org/10.1016/j.gene.2016.08.047>
- Wen, S., Li, P., Wang, F., Li, J., Liu, H., & Li, N. (2020). De novo assembly and microsatellite marker development of the transcriptome of the endangered *Brachymystax lenok* tsinlingensis. *Genes & Genomics*, 42(7), 727-734. <https://doi.org/10.1007/s13258-020-00939-3>.
- Ji, Y.J., Leng, L., Shi, C.M., Hua, Y.P., & Zhang, D.X. (2008). Eight polymorphic microsatellite markers developed in the Chinese scorpion, *Mesobuthus martensii* (Scorpiones: Buthidae). *Molecular Ecology Resources*, 8(6), 1454–1456. <https://doi.org/10.1111/j.1755-0998.2008.02215.x>
- Zane, L., Bargelloni, L., & Patarnello, T. (2002). Strategies for microsatellite isolation. *Molecular Ecology*, 11(1), 1–16. <https://doi.org/10.1046/j.0962-1083.2001.01418.x>.