

مقاله علمی - پژوهشی:

شناسایی و اعتبارسنجی نشانگرهای توالی ساده تکراری (SSR) برای ماهی گطان (*Luciobarbus xanthopterus*) بر پایه داده‌های RNA-Seq

غزال محمدی اهوازی^۱، محمدتقی بیگی نصیری^۱، محمود نظری^{۱*}، آیه سادات صدر^۲

*M.nazari@asnrukh.ac.ir

۱- گروه علوم دامی، دانشکده علوم دامی و صنایع غذایی، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان،

ملاثانی، ایران

۲- پژوهشکده آبی پروری جنوب کشور، موسسه تحقیقات علوم شیلاتی کشور، سازمان تحقیقات، آموزش و

ترویج کشاورزی، اهوان، ایران

تاریخ پذیرش: فروردین ۱۴۰۳

تاریخ دریافت: مهر ۱۴۰۲

چکیده

این پژوهش به منظور شناسایی و اعتبارسنجی نشانگرهای توالی ساده تکراری (SSR) مربوط به گونه ماهی گطان (*Luciobarbus xanthopterus*) با استفاده از داده‌های RNA-Seq انجام شد. پس از جداسازی بافت کبد، استخراج RNA از نمونه‌ها انجام گرفت. نمونه‌ها با استفاده از پلتفرم ایلومینا Novaseq 6000 توالی‌یابی شدند. برای بازسازی رونوشت‌ها، از نرم‌افزار Trinity (نسخه ۲.۱۵.۱) استفاده شد. شناسایی SSRs با استفاده از پایگاه MISA انجام شد. پرایمرهای مورد نیاز در واکنش زنجیره‌ای پلیمراز برای ۵ نشانگر SSR با نرم‌افزار PRIMER 3 طراحی و بر ۴۵ ماهی گطان مورد ارزیابی قرار گرفت. سرهم‌بندی رونوشت‌ها با برنامه Trinity، ۹۴۱،۸۹۴ یونی ژن و ۱،۷۶۸،۶۶۲ ترانسکرپت ایجاد کرد. در این مطالعه، بررسی ۱،۲۷۶،۸۲۵ توالی به‌وسیله نرم‌افزار MISA، منجر به شناسایی تعداد ۳۴۹،۸۷۱ نشانگر SSR شد. در میان SSRs، تکرارهای دو و سه نوکلئوتیدی به ترتیب با ۸۹/۳۷ درصد و ۸/۰۵ درصد، بیشترین تعداد تکرار را به‌خود اختصاص دادند. ۵ پرایمر مورد استفاده، ۲ جایگاه چندشکلی را نشان دادند. هتروزیگوسیتی مورد انتظار و مشاهده شده برای جایگاه MN1359 به ترتیب ۰/۷۸۵ و ۰/۱۴۷ و برای جایگاه GM1371 به ترتیب ۰/۷۷۰ و ۰/۰۹۳ محاسبه شد. به‌علاوه، آزمون کای مربع نشان داد که جمعیت برای هر دو جایگاه در تعادل هاردی-وینبرگ قرار ندارد. نتایج ارزیابی معیارهای مختلف تنوع ژنتیکی همگی گویای کاهش تنوع در بین جمعیت مورد مطالعه بود. به‌طور کلی، نتایج این تحقیق نشان داد که SSRs جدید می‌توانند برای اصلاح نژاد، مطالعات ژنومیک عملکردی و بررسی تنوع ژنتیکی ماهی گطان مفید باشند.

کلمات کلیدی: گطان، RNA-Seq، ترانسکرپتوم، توالی‌های ساده تکراری

*نویسنده مسئول



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

مقدمه

اگرچه کشور ایران در یک منطقه خشک و نیمه خشک قرار دارد، اما تنوع ماهیان آبهای داخلی ایران بسیار بالاست و حوضه‌های آبریز خزر و دجله دارای بیشترین تنوع گونه‌ای هستند که در این میان برخی از گونه‌های ماهیان آبهای داخلی ایران مانند جنس باربوس‌ها دارای ارزش صید اقتصادی هستند. اما با وجود اهمیت خاص این مناطق به عنوان متنوع‌ترین زیستگاه‌های ماهیان آبهای داخلی ایران، متأسفانه تهدیدهای زیادی نیز این تنوع ارزشمند را تهدید می‌کند که از جمله عوامل تهدید کننده تنوع زیستی ماهیان آبهای داخلی ایران می‌توان به خشک شدن آب رودخانه‌ها، ساخت سدها، بهره‌برداری بیش از حد از برخی گونه‌ها و عدم شناخت کافی از وضعیت دقیق گونه‌ها اشاره کرد (Abdoli, 2016).

ماهی گطان (*Luciobarbus xanthopterus* Heckel, 1843) از جمله ماهیان با ارزش و اقتصادی در مناطق جنوب غربی ایران است که در بخش‌هایی از منابع آبی استان خوزستان و مرزی آن زیست می‌کند (Hosseinzadeh, 2017). این ماهی در طبیعت از رشد بسیار خوبی برخوردار است و به خانواده کپور ماهیان (Cyprinidae) و جنس *Luciobarbus* (Fricke et al., 2007) تعلق دارد. علاوه بر این، به دلیل کوتاهی میزان طول نسبی روده و محتویات آن، می‌توان استنباط کرد که گطان یک ماهی همه‌چیزخوار است که به طور عمده از گیاهان و بی‌مهرگان تغذیه می‌کند (Eskandari et al., 2003). گونه گطان گستره وسیعی دارد و از گونه‌های ارزشمند در میان باربوس ماهیان منطقه به‌شمار می‌رود، ولی متأسفانه برخی از گزارش‌ها حاکی از آن است که جمعیت این گونه در آبهای ایران و عراق کاهش یافته است (Mortezavi Zadeh et al., 2011). همچنین این گونه تحت تاثیر سدها و برداشت آب و صید بی‌رویه، در معرض خطر انقراض قرار دارد (Freyhof, 2014).

کاهش تنوع ژنتیکی به‌ویژه برای جمعیت‌هایی که با شرایط محیطی استرس‌زا مواجه هستند، ممکن است منجر به کاهش زنده‌مانی جمعیت و افزایش احتمال انقراض شود

(Vandewoestijne et al., 2008; Markert et al., 2010).

نشانگرهای ژنتیکی مولکولی از ابزارهای قدرتمند برای تشخیص منحصر به فرد بودن ژنتیکی افراد، جمعیت‌ها یا گونه‌ها هستند (Doveri et al., 2008; Chenani et al., 2021). ریزماهورها به عنوان نشانگرهای ژنتیکی در سرتاسر ژنوم یوکاریوتی توزیع شده‌اند و برای تجزیه و تحلیل تنوع در بسیاری از گونه‌ها مورد استفاده قرار می‌گیرند (Wolfus et al., 1997; Ellis and Burke, 2007). توالی‌های ساده تکراری (SSRs)¹ که به عنوان ریزماهورها نیز شناخته می‌شوند، تکرارهای کوتاه پشت سر هم هستند که معمولاً در ژنوم‌ها یا رونوشت‌های موجودات مختلف یافت می‌شوند (Xia et al., 2016). نشانگرهای مولکولی SSR در مناطق کد کننده و غیر کد کننده ژنوم توزیع شده‌اند که امکان تمایز ژنوتیپ‌ها را فراهم می‌کنند. جایگاه‌های ریزماهورها چند شکلی، به طور گسترده برای تجزیه و تحلیل ساختار ژنتیکی جمعیت استفاده می‌شوند (Romero et al., 2019). به طور کلی، این نشانگرها به دلیل تکرارپذیری، فراوانی، ماهیت چند آللی، غلبه، پوشش ژنومی بالا و تنوع در کل ژنوم، بسیار مفید هستند (Jabeen et al., 2023). از دیگر مزایای این منابع مولکولی می‌توان به سهولت استفاده نسبی آنها برای مقایسه بین هم‌تایان ژنومی و کیفیت بالا اشاره کرد. همچنین مشخص شده است که تغییرات SSRs ممکن است موجب تغییر فعالیت ژن یا تنظیم بیان ژن شود. از این رو، فرآیندهای بیولوژیک و سلولی، از جمله ساختار پروتئین و رونویسی ژن را تحت تأثیر قرار می‌دهد (Dang et al., 2020).

توالی‌یابی RNA (RNA-Seq)² یک فن‌آوری مولکولی کارآمد برای مطالعه تکامل گونه‌ها، تعیین بیان ژن‌ها و نقشه‌برداری ژنتیکی و ... است (Rai et al., 2016). در واقع، این فناوری با فراهم کردن امکان کمی‌سازی سطح بیان همه ژن‌ها به طور همزمان، رویکرد جامعه تحقیقاتی را برای مطالعه بیان ژن متحول کرده است (Geraci et al., 2020). همچنین توالی‌یابی RNA برای انجام مطالعات ژنتیکی و

¹ Simple sequence repeats (SSRs)

² RNA sequencing

(Thermo Scientific Nanodrop. 2000. USA) استفاده شد. پس از استخراج RNA و اطمینان از کمیت و کیفیت نمونه‌های مورد نظر، شرکت توپازرن کاوش، نمونه‌ها را برای توالی‌یابی به خارج از کشور شرکت Novogene ارسال کردند. توالی‌یابی نمونه‌ها با استفاده از فناوری ایلومینا Novaseq 6000 صورت پذیرفت.

پس از دریافت داده‌های توالی‌یابی، قبل از تجزیه و تحلیل داده‌ها، به منظور بررسی و کنترل کیفیت داده‌های حاصل از توالی‌یابی از نرم‌افزار FastQC (نسخه 0.11.5) استفاده شد. سپس برای بازسازی رونوشت‌ها بر مبنای خوانش‌های توالی، از نرم‌افزار Trinity (نسخه 2.15.1) در حالت پیش‌فرض (K-mer ۲۵) استفاده شد. همچنین به منظور حذف توالی‌های تکراری و خوشه‌بندی خوانش‌ها از نرم‌افزار CD-HIT (نسخه 4.7) استفاده گردید. پس از کلاستر بندی خوانش‌ها برای شناسایی جایگاه‌های ریزماهواره موجود در رونوشت‌های مونتاژ شده حاصل از داده‌های RNA-seq گونه گطان از وب سرور MISA به آدرس <https://webblast.ipk-gatersleben.de/misa> با پارامترهای پیش‌فرض استفاده شد (Beier et al., 2017). در ادامه به منظور اعتبارسنجی نشانگرهای SSR شناسایی شده جدید، ۴۵ نمونه از باله دم‌ی ماهی گطان صید شده از رودخانه کارون تهیه شد. پس از استخراج DNA (طبق پروتکل کیت ستونی استخراج DNA ژنومیک شرکت دنا زیست آسیا)، کیفیت DNA های استخراجی به وسیله ژل آگارز ۱ درصد تعیین شد. در نهایت ۵ آغازگر، برای نشانگرهای SSR شناسایی شده جدید، با نرم‌افزار PRIMER 3 از پایگاه داده NCBI طراحی شد (جدول ۱). واکنش زنجیره‌ای پلیمرز با حجم نهایی ۲۵ میکرولیتر و ۳۵ چرخه بر نمونه‌های DNA استخراجی از باله دم‌ی ماهی گطان با استفاده از دستگاه ترموسایکلر Bio Rad و روش Touchdown صورت گرفت. برنامه حرارتی مورد استفاده به شرح ذیل است. 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 ۵۹-۶۴ درجه سانتی‌گراد، ۳۰ ثانیه در ۷۲ درجه سانتی‌گراد و ۱۰ دقیقه برای تکثیر نهایی در دمای ۷۲ درجه سانتی‌گراد

مولکولی به‌ویژه در موجودات غیر مدل که اطلاعات ژنتیکی بسیار اندکی از آنها موجود است، کاربرد زیادی دارد (Ghorbani et al., 2020). مطالعات متعددی در ارتباط با شناسایی مارکرهای SSR بر پایه داده‌های RNA-Seq در گونه‌های مختلف ماهی مانند *Brachymystax lenok* (Wen et al., 2020)، *Patagonian toothfish* (Touma et al., 2019)، *Betta splendens* (Chen et al., 2021)، *Amphiprion* (Fang et al., 2015) و *Coilia nasus ocellaris* (Yang et al., 2018) مورد بررسی قرار گرفته است.

پژوهش‌های متعددی در کشور ایران جهت بررسی تنوع ژنتیکی و ساختار جمعیت آبزیان در گونه‌های مختلف مانند سیاه کولی (Mohamadian et al., 2010)، مارماهی دهان گرد خزری (Bahmaninejad et al., 2018)، سگ ماهی آنگورا (Kashiri et al., 2015)، صیبتی (Ghasemi et al., 2019)، خیاطه (Hosseini et al., 2015) و شیربت (Karami Nasab et al., 2014) صورت گرفته است. اما به رغم این که ماهی گطان از ماهی‌های مرغوب منطقه به‌شمار می‌آید (Mortezavi Zadeh et al., 2011)، تاکنون هیچ‌گونه اطلاعات ژنتیکی برای این گونه گزارش نشده است. از این‌رو، با توجه به اهمیت ماهی گطان در منطقه جنوب غرب ایران و فقدان هر گونه اطلاعات در رابطه با ساختار ژنتیکی آن و اهمیت و نقش موثر تنوع ژنتیکی برای حفظ ذخایر تنوع گونه، مطالعه حاضر به منظور شناسایی و اعتبارسنجی نشانگرهای SSR مربوط به این گونه ماهی با استفاده از داده‌های RNA-Seq برای نخستین بار اجرا شد.

مواد و روش‌ها

در این پژوهش به منظور شناسایی نشانگرهای SSR در ماهی گطان (*Luciobarbus xanthopterus*) ابتدا دو نمونه ماهی گطان از رودخانه کارون صید و به آزمایشگاه منتقل گردید. پس از جداسازی بافت کبد، استخراج RNA از نمونه‌ها، با استفاده از کیت ستونی استخراج RNA (شرکت دنازی است آسیا) انجام گرفت. در نهایت به منظور بررسی RNA استخراجی و تعیین کیفیت و کمیت آن، از روش الکتروفورز ژل آگارز ۱ درصد و دستگاه نانودراپ

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 تنوع شانن، از نرم‌افزار GenAlex (نسخه 6.3) استفاده شد. توالی SSR چندشکل شناسایی شده با سایر توالی‌های موجود در پایگاه داده NCBI هم‌ردیف شد و در نهایت درخت فیلوژنتیک با استفاده از روش پیوند هم‌جواری Blast Tree View (Neighbor Joining)، به‌وسیله Widget از پایگاه داده NCBI ترسیم گردید.

صورت پذیرفت. مرحله اتصال آغازگرها طبق این روش به دو چرخه دمایی ۱۰ و ۲۵ تایی تقسیم شد که به‌ازاء هر چرخه ۰/۵ درجه سانتی‌گراد از دمای اتصال اولیه در ۱۰ چرخه اولیه کاهش یافت. پس از آن کیفیت محصولات حاصل از تکثیر به‌وسیله تکنیک PCR، با استفاده از الکتروفورز ژل آگارز ۳ درصد سنجیده شد. پس از تهیه تصویر ژل‌ها به‌وسیله دستگاه ژل‌داک، برای سنجش وزن مولکولی محصولات PCR بر حسب جفت باز (bp) و تعیین ژنوتیپ‌ها از نرم‌افزار GelAnalyzer (نسخه 19.1) استفاده گردید. در ادامه به منظور تجزیه و تحلیل‌های آماری برای جایگاه‌هایی

جدول ۱: خصوصیات آغازگرهای استفاده شده برای اعتبارسنجی SSR ها در این مطالعه
Table 2: Characterization of primers used for validation of SSRs in this study.

Number	SSR sequence	Primer sequence	Amplicon size (bp)	Temperature (°C)
1	(ATGAGC)13	F: CCAGCAGCCACCCTATATCC R: CTTGTGCGAGTTTGACGCTC	258	60
2	(GA)34	F: TCACCATCGCCATAATCCGC R: CGTATCGCCGGCATTAGCA	174	58
3	(AC)35	F: TICTTTTCGTCCGGTGCCTTA R: CAAAGTCTCTGTGAGGGCT	283	59
4	(AGAT)23	F: CTTGGTTGCAAGCAGCGTC R: GCAACCAGGGTCCAAAATGAA	225	60
5	(GA)34	F: TCGAGATCTTTGTCAACCAGCTT R: GTGGCGGATATGGAGGTTCC	363	60

کیفیت داده‌ها به‌وسیله آزمون‌های کنترل کیفیت نرم‌افزار FastQC نشان‌دهنده کیفیت بالای خوانش‌ها جهت آنالیزهای بعدی در ادامه روند کار بود (شکل ۲). بازسازی رونوشت‌ها به روش *De novo* با استفاده از نرم‌افزار Trinity در مجموع ۹۴۱۸۹۴ یونی ژن و ۱۷۶۸۶۶۲ ترانسکرپت ایجاد کرد. همچنین متوسط طول کانتیگ (contig) ۳۴۵ محاسبه شد. نتایج حاصل از نرم‌افزار Trinity در جدول ۲ ارائه شده است.

در ادامه به منظور شناسایی نشانگرهای SSR جدید، ترانسکرپتوم^۱ گطان با استفاده از وب سرور MISA مورد تجزیه و تحلیل قرار گرفت و در نهایت با بررسی تعداد ۱۲۷۶۸۲۵ توالی با اندازه کل ۷۱۳۱۷۶۹۸۵ جفت باز، در

نتایج

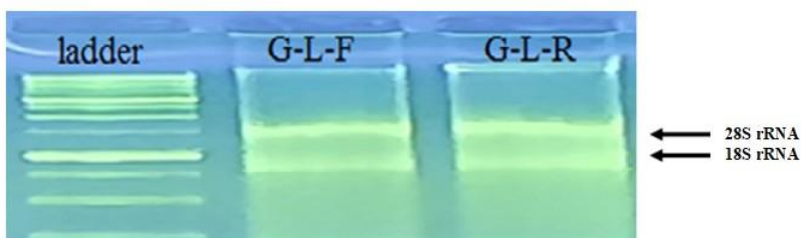
tRNAs تقریباً ۷۰ درصد از کل RNAs سلول را تشکیل می‌دهند، به همین دلیل مشاهده باندهای آن (۲۸S و ۱۸S) بر ژل آگارز به صورت شارپ می‌تواند بیانگر یک استخراج مطلوب و باندهای تخریب شده دال بر تخریب RNA و پایین بودن کیفیت باشد. استخراج RNA از بافت کبد، با موفقیت صورت پذیرفت و نتایج استخراج RNA بر الکتروفورز ژل آگارز ۱ درصد گویای فقدان هر گونه باند کاذب یا اسمیر و عدم تجزیه RNA طی فرایند استخراج است (شکل ۱).

توالی‌یابی RNA از نمونه‌های بافت کبد ماهی گطان، به‌ترتیب ۲۴۲۵۶۳۵۵ و ۴۸۳۰۹۸۲۱ خوانش خام برای نمونه‌های G-L-F و G-L-R تولید کرد. پس از دریافت داده‌های حاصل از توالی‌یابی در فرمت FastQ، بررسی

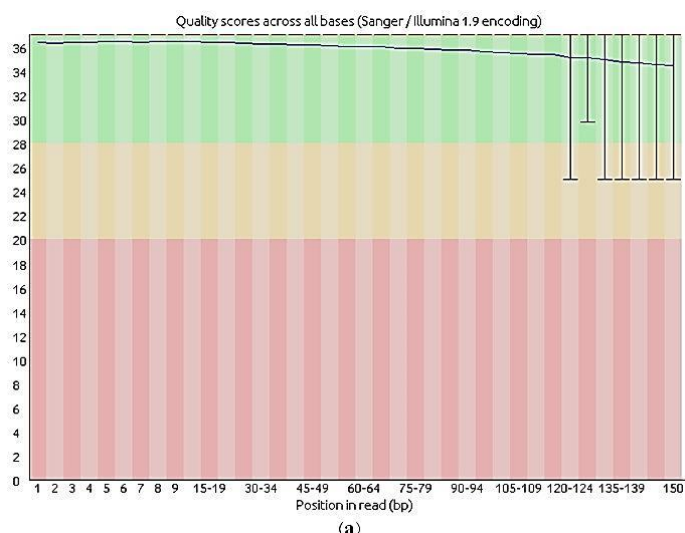
¹ Transcriptome

همچنین تعداد ۷۴۱۷۳ توالی‌های حاوی بیش از یک SSR شناسایی شدند (جدول ۳).

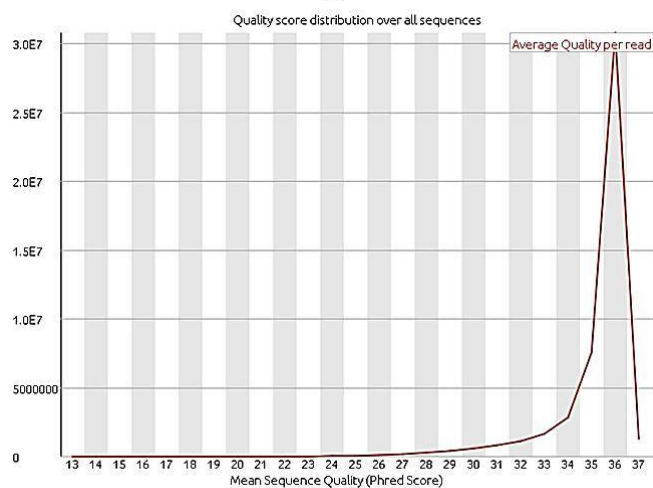
نهایت منجر به شناسایی تعداد ۳۴۹۸۷۱ نشانگر SSR جدید گردید که از این میان ۲۲۹۹۹۳ توالی، حاوی SSR بودند.



شکل ۱: کیفیت RNA استخراجی از بافت کبد ماهی گطان (*Luciobarbus Xanthopterus*) بر ژل آگارز ۱ درصد
Figure 1: The qualification of RNA extraction from liver tissue of yellowfin barbel (*Luciobarbus Xanthopterus*) on 1% agarose gel



(a)



(b)

شکل ۲: نمودار کیفیت توالی‌یابی به ازای هر باز (a) و امتیاز کیفیت به ازای توالی (b)
Figure 2. Quality assessment tests. (a) per base and (b) per sequence quality scores

جدول ۲: نتایج حاصل از رونوشت‌های اسمبل شده در نرم‌افزار

Trinity
Table 4: Results of assembled transcripts in Trinity software

Trinity	Statistics
941894	Total trinity genes
1768662	Total trinity transcripts
35.97	% GC
821	Contig N50
609.24	Average contig
345	Median contig length
1077536230	Total assembled bases

با توجه به نتایج گزارش شده از بررسی ۱۲۷۶۸۲۵ توالی، تکرارهای دو و سه نوکلئوتیدی به ترتیب با ۸۹/۳۷ درصد و

۸/۰۵ درصد، بیشترین تعداد تکرار را به خود اختصاص دادند و توالی‌های ۶ نوکلئوتیدی دارای کمترین تعداد بودند (جدول ۴).

شایان ذکر است، برای تشخیص جایگاه SSR ترانسکریپتوم در توالی‌های مونتاژ شده حداقل از ۶، ۵، ۵ و ۵ واحد تکرار پیوسته به ترتیب برای شناسایی موتیف‌های SSR کامل دی، تری، تترا، پنتا و هگزا نوکلئوتیدی استفاده شد. از مجموع SSRs دی نوکلئوتید شناسایی شده (AC/GT، AG/CT، AT/AT و CG/CG)، موتیف AG/CT با فراوانی ۶۵/۲۱ درصد، بیشترین فراوانی را به خود اختصاص داد در حالی که موتیف CG/CG با درصد فراوانی ۰/۰۷ بسیار نادر است (شکل ۳).

جدول ۳: نشانگرهای SSR شناسایی شده با استفاده از وب سرور MISA

Table 3: SSR markers detected using the MISA web server

Number	SSR markers
1276825	Total number of sequences examined
713176985	Total size of examined sequences (bp)
349871	Total number of identified SSRs
229993	Number of SSR containing sequences
74173	Number of sequences containing more than 1 SSR
61192	Number of SSRs present in compound formation

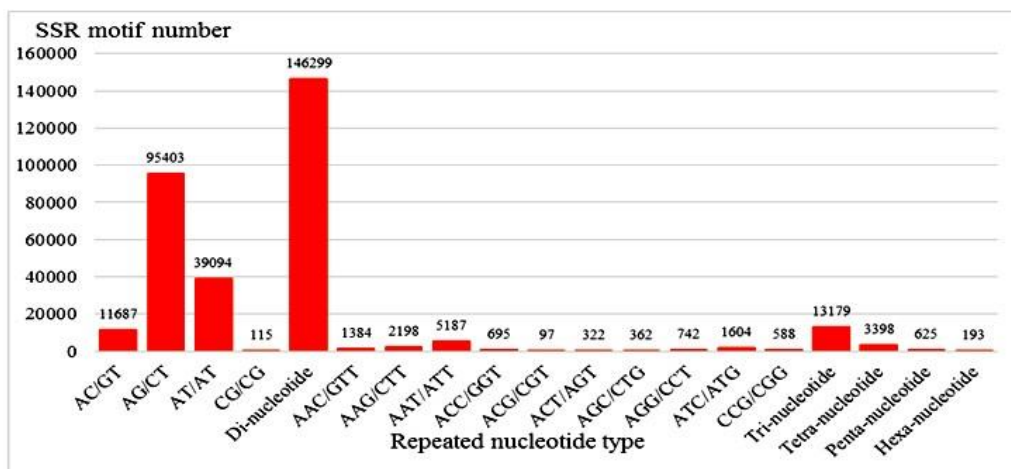
جهت بررسی میزان خلوص، اندازه‌گیری شد. این نسبت در نمونه‌های استخراجی در این پژوهش ۲-۱/۸ بود. پس از بررسی کمیت و کیفیت DNAs استخراجی، واکنش PCR بر نمونه‌های استخراجی اجرا شد و محصولات حاصل آن با استفاده از ژل آگارز ۳ درصد، مورد بررسی قرار گرفتند. از ۵ پرایمر مورد استفاده در این تحقیق، ۲ جایگاه (پرایمر شماره ۳ و ۴) چندشکلی بوده و سایر جایگاه‌ها تک باند بودند. در این تحقیق، این جایگاه به صورت MN1359 و GM1371 نامگذاری شد. دامنه آلی مشاهده شده برای جایگاه MN1359، ۱۵۷-۲۲۵ جفت نوکلئوتید و برای جایگاه GM1371، ۲۶۰-۳۰۰ جفت نوکلئوتید بود. در جدول ۵ پارامترهای آماری برآورد شده برای محاسبه تنوع ژنتیکی از جمله مقدار هتروزیگوسیتی مشاهده شده، میزان هتروزیگوسیتی مورد انتظار و تعداد آلل‌های واقعی، اندازه آلل‌های موثر و تعداد نمونه برای جایگاه MN1359 و GM1371 جمعیت گونه گطان استان خوزستان ارائه شده است.

جدول ۴: خلاصه‌ای از نوع و تعداد موتیف‌های تکراری

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

Unit size	Number of SSRs	Frequency (%)
Di-nucleotide	146299	89.37
Tri-nucleotide	13179	8.05
Tetra-nucleotide	3398	2.07
Penta-nucleotide	625	0.38
Hexa-nucleotide	193	0.13

به دلیل فقدان هر گونه اطلاعات در ارتباط با ژنوم گونه ماهی گطان، ابتدا SSRs شناسایی و سپس به منظور اعتبارسنجی نشانگرهای SSR شناسایی شده جدید، ۵ جفت پرایمر طراحی و با کمک روش واکنش زنجیره‌ای پلیمرز تکثیر شدند. استخراج DNA از نمونه‌های بافت باله دمی، با موفقیت صورت گرفت. تک باند مشاهده شده در مورد تک نمونه‌ها بیانگر صحت انجام آزمایش استخراج DNA از نمونه‌های مورد نظر بود. پس از اطمینان از کیفیت باندها بر ژل آگارز، نسبت طول موج‌های ۲۶۰ نانومتر به ۲۸۰ نانومتر



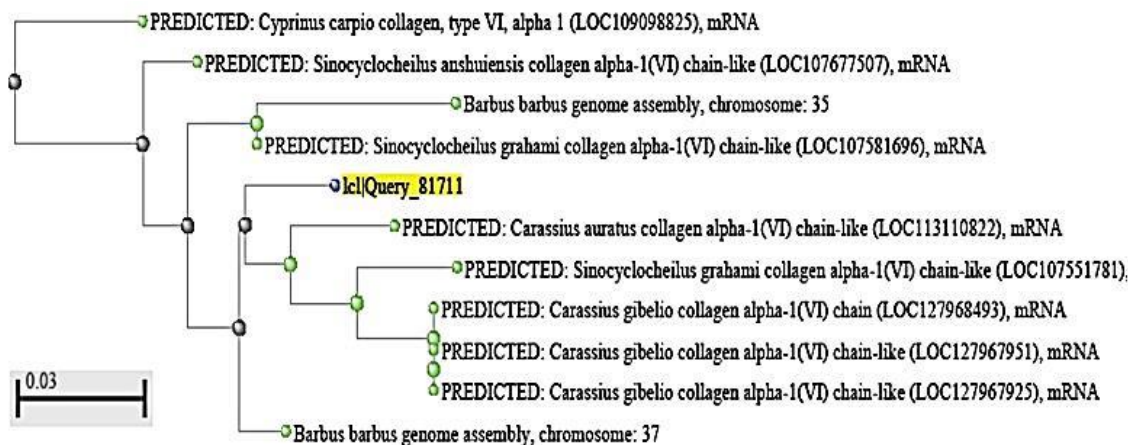
شکل ۳: توزیع فراوانی توالی‌های کوتاه تکراری (SSR) بر پایه نوع موتیف
Figure 3: Frequency distribution of simple sequence repeats (SSRs) based on motif type

جدول ۵: پارامترهای آماری برآورد شده برای محاسبه تنوع ژنتیکی جایگاه MN1359 و GM1371 در جمعیت گونه گطان
Table 5: Statistical parameters estimated to calculate the genetic diversity of the locus MN1359 and GM1371 in yellowfin barbel population

Locus	number of samples	Different Alleles (N _a)	The number of effective alleles (N _e)	Shannon index (I)	Observed heterozygosity (H _o)	Expected heterozygosity (H _e)
MN1359	45	5	4.651	1.568	0.147	0.785
GM1371	45	5	4.342	1.532	0.093	0.770

به‌دست آمد که در سطح ۱ درصد، معنی‌دار بود. با استفاده از توالی جدید شناسایی شده در این تحقیق (جایگاه MN1359)، درخت فیلوژنتیک برای این توالی و سایر توالی‌های موجود در پایگاه NCBI نیز رسم شد (شکل ۴).

تست کای اسکور انجام شده بر جمعیت ماهی گطان رودخانه کارون عدم برقراری تعادل هاردی-وینبرگ را برای هر دو جایگاه نشان داد. مقدار کای مربع برای جایگاه‌های MN1359 و GM1371 به ترتیب ۱۹۸/۸۱ و ۱۹۵/۷۲



شکل ۴: رابطه فیلوژنتیک جایگاه MN1359 با سایر توالی‌های موجود در پایگاه داده NCBI
Figure 4: Phylogenetic relationship of the MN1359 locus with other sequences in the NCBI database

بحث

با وجود اهمیت اقتصادی ماهی گطان، تعداد پژوهش‌های صورت گرفته در ارتباط با ارزیابی جنبه‌های مختلف این گونه، اندک است. بررسی ژنوم، ترانسکریپتوم و ساختار جمعیت ماهی گطان (*Luciobarbus xanthopterus*) به عنوان یک گونه در معرض انقراض (Freyhof, 2014) ضروری است. هیچ‌گونه اطلاعاتی در زمینه ژنوم ماهی گطان که در خطر انقراض است، در دسترس نیست. یک ابزار سریع و مقرون به صرفه برای شناسایی نشانگرهای مولکولی چندشکلی به‌ویژه SSRs، در مقایسه با روش‌های جداسازی سنتی، توالی‌یابی ترانسکریپتوم است (Zane et al., 2002). لذا، در این تحقیق نشانگرهای SSR مربوط به گونه ماهی گطان با استفاده از داده‌های RNA-Seq شناسایی شدند و جهت صحت کار مورد ارزیابی قرار گرفتند. امروزه نشانگرهای SSR رویکردی موثر برای شناسایی ارتباطات فنوتیپی و منبع مناسبی برای ایجاد نقشه‌های ژنتیکی، تجزیه و تحلیل تنوع ژنتیکی و انتخاب به کمک نشانگر مولکولی است (Zhang et al., 2014). این نشانگرها به دلیل سطوح بالای چندشکلی، پرکاربردترین نشانگرهای مولکولی در مطالعات ژنتیک جمعیت به‌شمار می‌روند (Li et al., 2002). در پژوهش حاضر، تعداد ۳۴۹۸۷۱ نشانگر SSR بر اساس یونی ژن‌های اسمبل شده حاصل از توالی‌یابی ترانسکریپتوم گونه گطان شناسایی شد که حدود ۱۸/۰۱ درصد از کل توالی‌های مورد بررسی حاوی SSR بودند. این تعداد کمی بیشتر از موارد گزارش شده در مطالعات مربوط به گونه‌های مختلف ماهی بود (Yang et al., 2018; Wen et al., 2020). از مجموع SSRs شناسایی شده، ۷۴۱۷۳ توالی حاوی بیش از یک SSR، تعداد ۶۱۱۹۲ SSR در فرم‌های ترکیبی نیز وجود دارند. علاوه بر این، فراوان ترین نوع SSR تکرار دی نوکلئوتید و به دنبال آن تری نوکلئوتید به‌دست آمد که با نتایج گزارش شده در رابطه با سایر گونه‌های ماهی از جمله *Brachymystax lenok* (Wen et al., 2020) *Patagonian toothfish* (Touma et al., 2019) و *Amphiprion ocellaris* (Yang et al., 2018) مطابقت دارد که این امر گویای صحت و دقت بالای یافته‌های این تحقیق است. همچنین از مجموع SSRs

دی‌نوکلئوتید شناسایی شده موتیف AG/CT با فراوانی ۶۵/۲۱ درصد، بیشترین فراوانی را به‌خود اختصاص داد درحالی‌که موتیف CG/CG با درصد فراوانی ۰/۰۷ بسیار نادر است که با نتایج سایر محققین مطابقت دارد (Fang et al., 2015). این مورد در مهره‌داران نیز به‌خوبی نشان داده شده است (Toth et al., 2000). در پژوهش حاضر از بین SSRs شناسایی شده در گونه گطان، تنوع ژنتیکی در ۵ جایگاه مورد بررسی و ارزیابی قرار گرفت. نتایج ژل الکتروفورز حضور ۵ SSR را تایید کرد و نشان داد که SSRs جدید شناسایی شده، مورد تایید است. از میان این ۵ جایگاه، ۳ جایگاه فاقد چند شکلی بوده و فقط ۲ جایگاه چندشکل بود. میزان هتروزیگوسیتی به عنوان یکی از معیارهای متداول بررسی تنوع درون یک جمعیت است (Frankham, 2008; Rohipoor et al., 2021). با توجه به نتایج به‌دست آمده در جدول ۵، مقدار هتروزیگوسیتی مشاهده شده برای جایگاه MN1359 و GM1371 به ترتیب برابر با ۰/۱۴۷ و ۰/۰۹۳ و هتروزیگوسیتی مورد انتظار برای جایگاه MN1359 و GM1371 به ترتیب برابر ۰/۷۷۰ و ۰/۷۷۰ محاسبه شد. پایین بودن میزان هتروزیگوسیتی مشاهده شده در این جایگاه‌ها را می‌توان بیانگر سطح پایین تنوع آن درون جمعیت دانست. کاهش هتروزیگوسیتی در ذخائر می‌تواند به دلیل ایجاد تنگناهای ژنتیکی از جمله صید بی‌رویه، آمیزش‌های خویشاوندی و تخریب زیستگاه‌ها باشد (Norris et al., 1999; Rohipoor et al., 2019). طبق نتایج حاصله در پژوهش حاضر، تعداد آلل‌های واقعی ۵ به‌دست آمد. تعداد آلل واقعی به عنوان معیاری برای بررسی تنوع، به‌شدت به اندازه نمونه بستگی دارد و در آزمایش‌های با تعداد نمونه‌های متفاوت، تعداد آلل‌های مختلفی ایجاد می‌شود (Mohamadi Ahwazi et al., 2019). همچنین تعداد آلل موثر به عنوان معیاری برای بیان آللهایی که هتروزیگوسیتی یکسان ایجاد می‌کنند، از تعداد آلل واقعی کمتر بود. باتوجه به این امر که تفاوت زیاد بین این دو عامل دلیلی بر فراوانی بالا در هر جایگاه است، پس نتایج حاصله در این تحقیق، بر پایین بودن میزان تنوع دلالت دارد. همچنین شاخص شانن، به عنوان معیار مفیدی برای اندازه‌گیری تنوع در جایگاه‌ها برای این جمعیت برای جایگاه

نژاد دیگر را اندازه‌گیری می‌نماید. معمولاً برای نشان دادن میزان انشقاق گروه‌ها، طول شاخه‌هایی که آنها را به هم متصل می‌کنند، مورد قیاس قرار می‌گیرد. تجزیه و تحلیل فیلوژنتیک بر مبنای رسم درخت فیلوژنی^۱، ابزاری برای جدا کردن جمعیت‌های خاص و مشخص کردن گونه‌های نیازمند به حفاظت است (Zhang and Hewitt, 2003). نتایج درخت فیلوژنتیک تأیید کرد که ماهی گطان به خانواده کپور ماهیان (Cyprinidae) و جنس *Luciobarbus* تعلق دارد که با نتایج گزارش‌های قبلی، همخوانی دارد (Fricke *et al.*, 2007).

نتیجه‌گیری

این پژوهش اولین تحقیقی است که بر پایه داده‌های RNA-Seq، نشانگرهای SSR گونه گطان را شناسایی، معرفی و ارزیابی کرده است. نتایج ارزیابی معیارهای مختلف تنوع ژنتیکی همگی گویای کاهش تنوع در بین جمعیت مورد مطالعه بود. علاوه بر این، نتایج حاصله ابزار قدرتمندی برای مطالعات ژنتیکی آینده برای این گونه به ویژه در مورد اساس ژنتیکی باربوس ماهیان و به طور کلی، ماهی گطان ارائه می‌دهد. همچنین می‌تواند برای مطالعات بیولوژیک، ساختار جمعیت و بررسی تنوع ژنتیکی مفید باشند.

تشکر و قدردانی

نگارندگان از مسئولین پژوهشی دانشگاه علوم کشاورزی و منابع طبیعی خوزستان برای حمایت مالی و از پژوهشکده آبی‌پروری جنوب کشور به دلیل فراهم نمودن امکانات آزمایشگاهی و همکاری صمیمانه برای اجرای تحقیق، تشکر و قدردانی می‌نمایند.

منابع

Abdoli, A., 2016. Field guide of fishes of inland waters of Iran. Iranshenasi Publications, Tehran, Iran. pp. 16-41. (in Persian)

MN1359 و GM1371 به ترتیب برابر ۱/۵۶۸ و ۱/۵۲۳ محاسبه شد که هرچه مقدار عددی این شاخص به صفر نزدیک‌تر باشد، دلالت بر پایین بودن تنوع دارد (Nazari and Mohamadi Ahwazi, 2022). با توجه به مقدار عددی محاسبه شده بر جمعیت مورد مطالعه، می‌توان استنباط کرد که تنوع گونه‌ای درون جمعیت این ماهی نسبتاً کم است. از دلایل این پدیده می‌توان به ورود آلودگی‌های شهری، صنعتی، کشاورزی و برداشت شن و ماسه بی‌رویه از بستر رودخانه اشاره کرد که باید از لحاظ زیست محیطی بیشتر به آن توجه شود. این گزارش نشان داد که می‌توان از نشانگرهای ریزماهوره (SSR) جهت بررسی تنوع ژنتیکی و ساختار جمعیت ماهی گطان استفاده کرد. این موضوع در ماهیانی همچون سیاه‌کولی (Mohamadian *et al.*, 2010)، مارماهی دهان گرد خزری (Bahmaninejad *et al.*, 2018)، سگ ماهی آنگورا (Kashiri *et al.*, 2015)، صیبتی (Ghasemi *et al.*, 2019) و خیاطه (Hosseini *et al.*, 2015) به اثبات رسیده است. تنوع ژنتیکی ماهی شیربت (گونه دیگری از باربوس ماهیان در رودخانه کارون استان خوزستان)، با استفاده از ریزماهوره‌ها نشان داد که تنوع ژنتیکی بالایی در این گونه وجود دارد (Karami Nasab *et al.*, 2014). این محققین از ۷ ریزماهوره برای بررسی ساختار جمعیت ماهی شیربت استفاده کردند که ۷ ریزماهوره مورد استفاده همگی چندشکلی را نشان دادند. فیلوژنتیک^۱، علمی است که با در نظر داشتن شباهت‌ها و تفاوت‌های بین گونه‌ها و نژادها، به توضیح روابط تکاملی موجودات می‌پردازد (Nazari and Mohamadi Ahwazi, 2022). علم طبقه‌بندی^۲، انواع موجودات زنده را به گروه‌های بسیار زیادی تقسیم می‌نماید به نحوی که اعضاء هر یک از این گروه‌ها، ویژگی‌های یکسان یا مشابهی دارند. بدین ترتیب، امکان طبقه‌بندی گونه‌ها فراهم می‌شود. در مطالعات تکامل، طبقه‌بندی شامل تشکیل فیلوژنی نیز می‌گردد. به دلیل نشان دادن مسیرهای تکاملی، از درختان فیلوژنتیک برای درک روابط تکاملی می‌توان استفاده کرد. در واقع، شاخه‌ها، فاصله تکاملی بین یک نژاد با

¹ Phylogenetics

² Taxonomy

³ Phylogenetic tree

- Bahmaninejad, A., Farahmand, H., Eagderi, S., Elmdoust, A., 2018.** Study of separation between fall and spring run populations of Caspian lamprey, *Caspiomyzon wagneri*, in the southern Caspian Sea (Shirud River) due to different migration time using microsatellite markers. *Journal of Aquaculture Sciences*, 6:73-81. (in Persian)
- Beier, S., Thiel, T., Münch, T., Schols, U. and Mascher, M., 2017.** MISA-web: a web server for microsatellite prediction. *Bioinformatics*, 33, 2583-2585. DOI:10.1093/bioinformatics/btx198
- Chen, H., Li, X., Wang, Y., Zhu, Ch., Huang, H., Yang, W. and Li, G., 2021.** De Novo Transcriptomic Characterization Enables Novel Microsatellite Identification and Marker Development in *Betta splendens*. *Life*, 11:803. DOI:10.3390/life11080803
- Chenani, H., Nazari, M., Beigi Nassiri, M. T., Roshanfekr, H. Aghaie, A., 2021.** Exonic SNP in MHC-DMB2 is associated with gene expression and humoral immunity in Japanese quails. *Veterinary Immunology and Immunopathology*, 239:110302. DOI:10.1016/j.vetimm.2021.110302
- Dang, Z., Huang, L., Jia, Y., Peter J. Lockhart, P.J., Fong, Y. and Tian, Y., 2020.** Identification of genic SSRs provide a perspective for studying environmental adaptation in the endemic shrub *Tetraena mongolica*. *Genes*, 11:322. DOI:10.3390/genes11030322
- Doveri, S., Lee, D., Maheswaran, M. and Powell, W., 2008.** Molecular markers: History, features and applications. In: Kole, C. and Abbott, A.G. (Eds). Principles and Practices of Plant Genomics. *The Science Publishers*, USA. pp. 23-68.
- Ellis, J.R. and Burke, J.M., 2007.** EST-SSRs as a resource for population genetic analyses. *Heredity*, 99(2):125-132. DOI:10.1038/sj.hdy.6801001
- Eskandari, G., Safi Khani, H., Dehghan, S. and Esmaili, F., 2003.** Gathering and feeding of Gattan fish (*barbus xanthopterus*) in Karkheh and Horulazim rivers. *Iranian Scientific Fisheries Journal*, 21-42. DOI:10.22092/ISFJ.2003.113604 (in Persian)
- Fang, D.A., Zhou, Y.F., Duan, J.R., Zhang, M.Y., Xu, D.P., Liu, K., Xu, P. and Wei, Q., 2015.** Screening potential SSR markers of the anadromous fish *Coilia nasus* by de novo transcriptome analysis using Illumina sequencing. *Genetics and Molecular Research*, 14(4):14181-14188. DOI:10.4238/2015.November.13.1
- Frankham, R., 2008.** Genetic adaptation to captivity in species conservation programs. *Molecular Ecology*, 17:325-333. DOI:10.1111/j.1365-294X.2007.03399.x
- Freyhof, J., 2014.** *Luciobarbus xanthopterus*. *The IUCN Red List of Threatened Species*, DOI:10.2305/IUCN.UK.2014-1.RLTS.T19383627A19849886.en
- Fricke, R., Bilecenoğlu, M. and Sari, H.M., 2007.** Annotated checklist of fish and lamprey species (*Gnathostoma* and *Petromyzontomorphi*) of Turkey, including a Red List of threatened and declining species. *Stuttgarter Beiträge zur Naturkunde*, 706:1-172.

- Geraci, F., Saha, I. and Bianchini, M., 2020.** RNA-Seq Analysis: Methods, Applications and Challenges. *Frontiers in Genetics*, 11:1-3. DOI:10.3389/fgene.2020.00220
- Ghasemi, S.A., Faghih, A. and Fakhri, A., 2019.** Genetic diversity Investigation of Silver Seabream *Sparidentex hasta* (Day, 1878) in Persian Gulf and Oman Sea using Microsatellite marker. *Journal of Animal Researches*, 32:3. (in Persian)
- Ghorbani, A., Rostami, M., Afshari Far, A. and Azarbaijani, R., 2020.** RNA sequencing (RNA Seq) and its data analysis. Marja Alam Publications. Shiraz, IRAN. 12 P. (in Persian)
- Hosseini, S.M., Shabani, A., Kalangi Miandre, H. and Safari, R., 2015.** Genetic structure of *Alburnoides eichwadi* (De Filippi, 1863) in Gilan (Roodsar Paulrood River) using microsatellite markers. *Journal of Applied Ichthyological Research*, 3(4):49-60. (in Persian)
- Hosseinzadeh sahafi, H., 2017.** Strategic roadmap for development of warmwater fish culture, Iranian fisheries science research institute, Tehran, IRAN. 11 P. (in Persian)
- Jabeen, S., Saif, R., Distefano, G., Haq, R., Haider, W., Hayat, A. and Naz, Sh., 2023.** *De novo* Genome Assembly, Functional Annotation and SSR Mining of *Citrus reticulata* "Kinnow" from Pakistan. *BioRxiv*, DOI:10.1101/2023.03.27.534305
- Karami Nasab, M., Shabani, A., Kalangi Miandreh, H. and Sharbati, S., 2014.** Genetic diversity of Shirbut (*Barbus grypus* Heckel, 1843) fish in Karun and Dez rivers of Khuzestan province using microsatellite markers. *Journal of Applied Ichthyological Research*, 2(1):63-74. (in Persian)
- Kashiri, H., Shabani, A. and Qudsi, V., 2015.** Population Structure of *Oxynoemacheilus sangorae* (Steindachner, 1897) in Gheshlagh (Kurdistan Province), Sefidbarg and Gamasiab (Kermanshah Province) rivers using microsatellite markers. *Journal of Aquatic Ecology*, 5(3):1-11. (in Persian)
- Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E., 2002.** Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, 11:2453-2465. DOI:10.1046/j.1365-294x.2002.01643.x
- Markert, J.A., Champlin, D.M., Gutjahr-Gobell, R., Grear, J.S. and Kuhn, A., McGreevy Jr, T.J., Roth, A., Bagley, M.J. and Nacci, D.E., 2010.** Population genetic diversity and fitness in multiple environments. *BMC Ecology and Evolution*, 10:205:1-13. DOI:10.1186/1471-2148-10-205
- Mohamadi ahwazi, Gh., Nazari, M., Mohamadabadi, M.R. and Heidari, R., 2019.** Genetic and phylogenetic analysis of mitochondrial HVR1 region in three breeds of Iranian sheep. *Modern genetics journal*, 14(3):209-217. (in Persian)
- Mohamadian, S., Rezvani Gilkolaei, S., Kazemian, M., Kamali, A., Tagvi, M.J., Rooh Elhi, Sh. and Nirani, M., 2010.** The study of genetic diversity and population structure of *Vimba vimba persa* (Pallas, 1814) populations in the eastern and western

- coastline of the Caspian Sea (Havigh river and GorganRoud river) using microsatellite markers. *Taxonomy and Biosystematics*, 5:29-38. (in Persian)
- Mortezavi Zadeh, S., Moazedi, J., Yooneszadeh Feshalami, M. Jorfi, E., 2011.** Determination of artificial propagation biotechnic of *Barbus xanthopterus*. *Iranian Scientific Fisheries Journal*, 19:137-142. DOI:10.22092/ISFJ.2017.109967 (in Persian)
- Nazari, M. and Mohamadi Ahwazi, Gh., 2022.** Genetic and phylogenetic analysis of mitochondrial D-loop HVR I region in three breeds of native sheep Iran (Taleshi, Shal and Makui). *Veterinary Research & Biological Products*, 35(1):31-39. DOI:10.22092/VJ.2021.352317.178 (in Persian)
- Norris, A.T., Bready, D.G. and Cuningham, E.D., 1999.** Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) population. *Aquaculture*, 247-264. DOI:10.1016/S0044-8486(99)00212-4
- Rai, R., Chauhan, S.K., Singh, V.V., Rai, M. and Rai, G., 2016.** RNA-seq analysis reveals unique transcriptome signatures in systemic lupus erythematosus patients with distinct autoantibody specificities. *PLoS One*, 11:1-35. DOI:10.1371/journal.pone.0166312
- Rohipoor, M., Nazari, M. and 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. DOI:10.29252/rap.10.26.84 (in Persian)
- Rohipoor, M., Nazari, M. and Beigi nassiri, M.T., 2021.** Population structure, Genetic diversity and phylogenetic analysis of control region of mtDNA in Adani goat breed. *Modern genetics journal*, 15(4):297-304. (in Persian)
- Romero, M., Mujica, A., Pineda, E., Ccamapaza, Y. and Zavalla, N., 2019.** Genetic identity based on simple sequence repeat (SSR) markers for Quinoa (*Chenopodium quinoa willd.*). *Ciencia e Investigacion Agraria*, 46(2):166-178. DOI:10.7764/rcia.v45i2.2144
- Toth, G., Gaspari, Z. and Jurka, J., 2000.** Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, 10:967-981. DOI:10.1101/gr.10.7.967
- Touma, J., García, K.K., Bravo, S., Leiva, F., Moya, J., Vargas-Chacoff, L., Reyes, A. and Vidal, R. 2019.** De novo Assembly and Characterization of *Patagonian toothfish* Transcriptome and Develop of EST-SSR Markers for Population Genetics. *Frontiers in Marine Science*, 6:1-13. DOI:10.3389/fmars.2019.00720
- Vandewoestijne, S., Schtickzelle, N. and Baguette, M., 2008.** Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biology*, 6:1-11. DOI:10.1186/1741-7007-6-46.
- Wen, S., Li, P., Wang, F., Li, J., Liu, H. and Li, N., 2020.** De novo assembly and microsatellite marker development of the transcriptome of the endangered

Brachymystax lenok tsinlingensis. *Genes and Genomics*, 42:727–734.
DOI:10.1007/s13258-020-00939-3

from transcriptome of yellow catfish. *Molecules*, 19(10):16402-16415.
DOI:10.3390/molecules191016402

Wolfus, G.M., Garcia, D.K., Alcivar-Warren, A., 1997. Application of the microsatellite technique for analyzing genetic diversity in shrimp breeding programs. *Aquaculture*, 152:35–47. DOI:10.1016/S0044-8486(96)01527-X

Xia, E.H., Yao, Q.Y., Zhang, H.B., Jiang, J.J., Zhang, L.P. and Gao, L.Z., 2016. CandiSSR: An Efficient pipeline used for identifying candidate polymorphic SSRs based on multiple assembled sequences. *front Plant Science*, 6:1-9.
DOI:10.3389/fpls.2015.01171

Yang, W., Lin, B., Li, G., Chen, H. and Liu, M., 2018. Sequencing and Transcriptome Analysis for Reproduction-Related Genes Identification and SSRs Discovery in Sequential Hermaphrodite *Amphiprion ocellaris*. *Turkish Journal of Fisheries and Aquatic Sciences*, 19:1049-1059.
DOI:10.4194/1303-2712-v19_12_07

Zane, L., Bargelloni, L. and Patarnello, T., 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology*, 11:1–16.
DOI:10.1046/j.0962-1083.2001.01418.x

Zhang, D.X. and Hewitt, G.M., 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, 12:563-584.
DOI:10.1046/j.1365-294X.2003.01773.x

Zhang, J., Ma, W., Song, X., Lin, Q., Gui, J.F. and Mei, J., 2014. Characterization and development of EST SSR markers derived

Identification and validation of simple sequence repeats (SSR) markers for yellowfin barbel (*Luciobarbus xanthopterus*) using RNA-Seq data

Mohamadi Ahvazi G.¹; Beigh Nassiri M.T.¹; Nazari M.^{1*}; Sadr A.S.²

*M.nazari@asnrukh.ac.ir

1- Department of Animal Science, Faculty of Animal Science and Food Technology, Agricultural Science and Natural Resources University of Khuzestan, Mollasani, Iran

2- Aquaculture Research Center -South of Iran, Iranian Fisheries Science Research Institute (IFSRI), Agricultural Research, Education and Extension Organization (AREEO), Ahvaz, Iran

Abstract

This research was conducted to identify and validate the simple sequence repeats (SSR) markers related to yellowfin barbel (*Luciobarbus xanthopterus*) using RNA-Seq data. After collecting the liver tissue, RNA was extracted from the samples. The samples were sequenced using the Illumina Novaseq 6000 platform. Trinity software (version 2.15.1) was used to reconstruct transcripts. Identification of SSR was done using the MISA web server. The primers needed in the polymerase chain reaction for 5 SSR markers were designed by PRIMER 3 software and validated using 45 yellowfin barbel. Transcript assembly by the Trinity program generated 941,894 unigenes and 1,768,662 transcripts. In this study, the examination of 1,276,825 sequences by MISA software led to the identification of 349,871 potential SSR markers. Among the SSRs, di- and trinucleotide repeats had the highest number of repetitions with 89.37% and 8.05%, respectively. From the 5 primers used, only two of them showed polymorphism. The expected and observed heterozygosity for the MN1359 locus were calculated as 0.785 and 0.147, and for the GM1371 locus given 0.770 and 0.093, respectively. Moreover, the chi-square test showed that population was not in Hardy-Weinberg equilibrium. The results of different genetic diversity criteria showed a decrease in diversity in yellowfin barbel. Generally, this research showed that the new SSRs can be helpful for molecular breeding, functional genomics studies, and the genetic diversity analysis of yellowfin barbel.

Keywords: Yellowfin barbel, RNA-Seq, Transcriptome, Simple sequence repeats

*Corresponding author