https://ui.ac.ir/en

Journal of Taxonomy and Biosystematics
E-ISSN: 2322-2190
Document Type: Research Paper
Vol. 12, Issue 2, No.43, Summer 2020, P:2
Received: 30/09/2020 Accepted: 15/11/2020

Designing a Local Molecular Database to Facilitate the Identification of Mammal Species Habitats in Iran

Bahareh Zahedian
MSc, Department of Biodiversity and Ecosystem Management, Environmental Sciences Research Institute, Shahid Beheshti University, Tehran, Iran
baharehzahedian@gmail.com

Faraham Ahmadzadeh*
*Corresponding author: Associate Professor, Department of Biodiversity and Ecosystem Management, Environmental Sciences Research Institute, Shahid Beheshti University, Tehran, Iran
f_ahmadzade@sbu.ac.ir

Asghar Abdoli
Professor, Department of Biodiversity and Ecosystem Management, Environmental Sciences Research Institute, Shahid Beheshti University, Tehran, Iran
asabdoli@gmail.com

Mohammad Javidkar
Ph. D. School of Biological Sciences, The University of Adelaide, Adelaide, SA, Australia
m.javidkar@gmail.com

Abstract

Identification and protection of habitats are essential to protect wildlife. Identifying the habitats used by some mammal species with low population density requires the application of practical techniques in this field. Molecular studies on mammalian remains in nature and genome sequencing have facilitated the process of identifying habitats as well as monitoring their populations. Therefore, designing and launching a local molecular database to identify the sequences obtained from mammalian remains and ultimately facilitate the identification of their habitats is the main purpose of this study. For this purpose, some Iranian mammal species that were in conservation categories according to national laws or IUCN criteria, as well as the keystone species, were selected for this study. The non-invasive sampling of different mammalian species was performed from all over Iran and two gene regions including D-loop and COXI were chosen for sequencing. Additional sequences of the selected species belonging to D-loop, COXI, and Cyt b mitochondrial genes were also obtained from NCBI and a Local Molecular Database (LMD) was prepared. The performance of the LMD was determined using a known query sequence, and the database recognized the query sequence correctly. This database potentially enables the correct identification of the studied mammals' habitats by identifying sequences belonging to species that remain in their habitats and resolves the existing challenges in the field of habitat identification and conservation planning.

Key words: Species Identification, Mammals, Local Molecular Database, Mitochondrial Gene, Iran.

Copyright©2020, University of Isfahan. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/BY-NC-ND/4.0), which permits others to download this work and share it with others as long as they credit it, but they can't change it in any way or use it commercially.

Doi:10.22108/tbj.2020.125190.1124

تاکسونومی و بیوسیستماتیک، سال دوازدهم، شماره چهل و سوم ، تابستان ۱۳۹۹، صفحه ۱۷-۲۸
نوع مقاله: پژوهشی
دریافت مقاله: ۱۳۹۹/۰۷/۰۹ پذیرش نهایی: ۱۳۹۹/۰۸/۲۵

طراحی پایگاه اطلاعاتی مولکولی محلی برای تسهیل شناسایی زیستگاههای گونه‌های پستانداران شاخص در ایران

بهاره زاهدیان، کارشناسی ارشد گروه تنوع زیستی و مدیریت اکوسیستم‌ها، پژوهشکدهٔ علوم محیطی، دانشگاه شهید بهشتی، تهران، ایران
baharehzahedian@gmail.com

فراهم احمدزاده*، دانشیار گروه تنوع زیستی و مدیریت اکوسیستم‌ها، پژوهشکدهٔ علوم محیطی، دانشگاه شهید بهشتی، تهران، ایران
(مسوول مکاتبات)
f_ahmadzade@sbu.ac.ir

اصغر عبدلی، استاد گروه تنوع زیستی و مدیریت اکوسیستم‌ها، پژوهشکدهٔ علوم محیطی، دانشگاه شهید بهشتی، تهران، ایران
asabdoli@gmail.com

محمد جاویدکار، دکتری تخصصی دانشکدهٔ علوم زیستی، دانشگاه آدلاید، استرالیای جنوبی، استرالیا
m.javidkar@gmail.com

چکیده

شناسایی و حفاظت از زیستگاهها در راستای حفاظت از حیات وحش امری ضروری است. شناسایی زیستگاههای برخی از گونه‌ها نظیر برخی پستانداران که تراکم جمعیت کمی در طبیعت دارند، نیازمند به‌کارگیری تکنیکی کاربردی در این زمینه است. انجام مطالعات مولکولی روی نمایه‌های زیستی برجای‌مانده از پستانداران در طبیعت و توالی‌یابی ژنوم، فرآیند شناسایی زیستگاهها و همین‌طور پایش جمعیت آنها را تسهیل کرده است؛ بنابراین طراحی و راه‌اندازی پایگاه اطلاعاتی مولکولی محلی در راستای شناسایی توالی‌های به‌دست‌آمده از نمایه‌های پستانداران و درنهایت، تسهیل شناسایی زیستگاههای آنها هدف اصلی این پژوهش است؛ بدین منظور گونه‌هایی از پستانداران ایران که براساس قوانین ملی یا برپایهٔ IUCN در طبقات حفاظتی قرار داشتند و همچنین گونه‌های پستاندار پی‌سنگ برای مطالعه انتخاب شدند. نمونه‌برداری به روش غیرتهاجمی از گونه‌های مختلف پستاندار از سراسر ایران انجام گرفت. دو ناحیهٔ ژنی D-loop و COXI توالی‌یابی شدند؛ سپس توالی‌های گونه‌های منتخب، متعلق به ژن‌های میتوکندریایی D-loop، COXI و Cyt b نیز از NCBI اخذ و پایگاه اطلاعاتی مولکولی محلی با استفاده از نرم‌افزار BLAST تهیه و راه‌اندازی شد. عملکرد پایگاه اطلاعاتی مولکولی محلی طراحی‌شده با استفاده از یک توالی Query سنجیده شد و پایگاه، توالی Query را به‌طور صحیح بازشناخت. این پایگاه شناسایی زیستگاه پستانداران مطالعه‌شده را با استفاده از شناسایی توالی‌های متعلق به نمایه‌های زیستی باقی‌مانده از آنها در زیستگاهها مقدور می‌سازد و به رفع چالش‌های موجود درزمینهٔ شناسایی زیستگاهها برای برنامه‌ریزی‌های حفاظتی کمک شایانی خواهد کرد.

واژه‌های کلیدی: شناسایی گونه‌ها، پستانداران، پایگاه اطلاعاتی مولکولی محلی، ژن میتوکندریایی، ایران.

Copyright©2020, University of Isfahan. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/BY-NC-ND/4.0), which permits others to download this work and share it with others as long as they credit it, but they cannot change it in any way or use it commercially.
Doi:10.22108/tbj.2020.125190.1124

مقدمه

پایش گونه‌های حیات وحش و شناسایی زیستگاههای آنها در طبیعت، اطلاعات بسیار باارزشی به‌منظور برنامه‌ریزی‌های حفاظتی برای آنها در اختیار پژوهشگران قرار می‌دهد (Ishige *et al.*, 2017). شناسایی نقاط قوت و ضعف روش‌های معرفی‌شده در این باره نیز از اهمیت بسیار زیادی برخوردار است. پایش و همچنین شناسایی زیستگاههای برخی از گونه‌های جانوری نظیر برخی گونه‌های پستانداران که در وضعیت تهدیدشده قرار دارند و دارای اندازهٔ جمعیت محدود یا ازلحاظ رفتاری پنهان‌کار هستند، اغلب در طبیعت دشوار است (Battersby and Greenwood, 2004; Burton *et al.*, 2015). درحال حاضر روش‌های متعددی برای پایش، شناسایی زیستگاهها و نحوهٔ پراکندگی این گروه از جانوران به کار گرفته می‌شود که هریک دارای مزایا و معایبی هستند (Sadlier *et al.*, 2004; Barea-Azcón *et al.*, 2007). مشاهدهٔ مستقیم، بررسی و شناسایی نمایه‌های برجای‌مانده از گونه‌ها در زیستگاهها براساس ویژگی‌های ظاهری (Akrim *et al.*, 2018) و استفاده از دوربین‌های تله‌ای (Farhadinia *et al.*, 2017) ازجملهٔ این روش‌ها هستند. روش‌های نام‌برده باوجود سادگی، زمان‌بر و پرهزینه هستند و مهم‌تر از آن دقت لازم را هم ندارند؛ بنابراین به‌کارگیری روشی مؤثر، کاربردی و با دقت زیاد به‌منظور ردیابی گونه‌های پستانداران و به دنبال آن شناسایی زیستگاههای مهم آنها برای برنامه‌ریزی‌های حفاظتی در آینده امری ضروری به نظر می‌رسد.

رشد و توسعهٔ تکنیک‌های مولکولی فرصت‌های جدیدی در اختیار پژوهشگران قرار داده و نقشی بسیار کاربردی در حفاظت از تنوع زیستی ایفا کرده است (Lorenz *et al.*, 2005; Kress *et al.*, 2015). توالی‌یابی ژنوم در نمونه‌هایی نظیر پوست، مو، مخاط، بزاق، اسپرم، سرگین، ادرار و خون که همگی حاوی مولکول‌های DNA هستند، ردیابی گونه‌های جانوری به‌ویژه، گونه‌های پستانداران و شناسایی زیستگاههای آنها را تسهیل کرده است (Taberlet *et al.*, 2012; Wheat *et al.*, 2016; Ishige *et al.*, 2017)؛ همچنین تأثیرنپذیرفتن از محیط، غیرتهاجمی‌بودن، دقت زیاد و هزینهٔ معقول روش‌های مبتنی‌بر DNA از مزایای آن نسبت به روش‌های سنتی پایش حیات وحش و شناسایی زیستگاههای آنها به شمار می‌رود (Thomsen and Willerslev, 2015).

ژن‌های میتوکندریایی به‌دلیل نرخ جهش و فراوانی بیشتر نسبت به ژن‌های هسته‌ای، اغلب برای شناسایی گونه‌های جانوری استفاده می‌شوند (Nowak *et al.*, 2014). توالی‌های بسیاری از ژن‌های میتوکندریایی متعلق به گونه‌های جانوری در مرکز ملی اطلاعات زیست‌فناوری (National Center for Biotechnology Information (NCBI)) وجود دارد (Smith, 2016) و این ژن‌ها نشانگرهای مفیدی برای شناسایی گونه‌ها معرفی شده‌اند (Kefi *et al.*, 2009). پایگاه بارکد حیات (The Barcode of Life Data System (BOLD)) نیز کتابخانه‌ای آنلاین است که توالی‌های بارکد گونه‌ها را گردآوری می‌کند و امکان شناسایی و تجزیه و تحلیل توالی‌های متعلق به گونه‌ها را فراهم کرده است (Ratnasingham and Hebert, 2007). در این بین، ژن‌های میتوکندریایی سیتوکروم اکسیداز I (Cytochrome c oxidase I (COXI))، سیتوکروم b (Cytochrome *b* (Cyt b)) و ناحیهٔ

غیرکدشوندهٔ D-loop، نشانگرهایی مفید و کاربردی برای شناسایی گونه‌های جانوری معرفی شده‌اند و به‌منظور تهیهٔ بارکد DNA استفاده می‌شوند (Hebert *et al.*, 2003; Lorenz *et al.*, 2005; Hajibabaei *et al.*, 2006; Clare *et al.*, 2007; Robins *et al.*, 2007; Borisenko *et al.*, 2008; Yacoub *et al.*, 2013; Parkanyi *et al.*, 2014; Aksöyek *et al.*, 2016).

هدف اصلی تهیهٔ بارکدهای DNA برای گونه‌های جانوری، گیاهی، آغازیان و ... شناسایی صحیح و دقیق گونه‌ها است و باوجود اینکه اغلب در گذشته تاکسونومیست‌ها از آن استفاده می‌کردند، امروزه توجه زیست‌شناسان حفاظت را نیز به خود جلب کرده است. بارکدهای DNA امروزه درزمینهٔ شناسایی فون و فلور یک محدودهٔ جغرافیایی، شناسایی گونه‌های مهاجم، شناسایی زیستگاههای حیات وحش و نحوهٔ پراکنش گونه‌ها در آن و همین‌طور شناسایی گونه‌های کشف و ضبط‌شده از شکارچیان به شکلی گسترده استفاده می‌شود (Kumar *et al.*, 2012; Ishige *et al.*, 2017). در این بین، پژوهشگران به کاربرد آن درزمینهٔ شناسایی توالی‌های به‌دست‌آمده از نمایه‌ها و پیروی آن، شناسایی زیستگاههای اشغال‌شده توسط حیات وحش بسیار توجه کرده‌اند.

امروزه پژوهشگرانی که در حیطهٔ حفاظت از حیات وحش مشغول به فعالیت‌اند با چالشی بسیار مهم و جدی درزمینهٔ شناسایی زیستگاههای برخی گونه‌های پستانداران روبه‌رو هستند؛ بنابراین به نظر می‌رسد تهیهٔ پایگاه اطلاعاتی مولکولی محلی که اطلاعات آن را توالی‌های گونه‌های پستانداران ایران تشکیل دهد، کمک شایانی به رفع چالش موجود می‌کند. پایگاه نام‌برده توالی‌های به‌دست‌آمده از نمایه‌های زیستی برجای مانده از گونه‌های پستانداران در طبیعت را تطبیق و شناسایی و به دنبال آن شناسایی زیستگاه آنها را به روشی دقیق و در کمترین زمان تسهیل خواهد کرد؛ بنابراین تهیهٔ پایگاه اطلاعاتی مولکولی محلی برای گونه‌های پستانداران ایران به‌منظور بهبود روند ردیابی و تسهیل شناسایی زیستگاههای آنها هدف اصلی پژوهش حاضر است.

مواد و روش‌ها

نمونه‌برداری

نمونه‌برداری و تهیهٔ بافت از گونه‌هایی از پستانداران ایران که براساس قوانین ملی یا برپایهٔ فهرست سرخ اتحادیهٔ جهانی حفاظت از محیط زیست (IUCN) در طبقات حفاظتی قرار داشتند و همین‌طور گونه‌های پی‌سنگ Keystone ، به‌صورت غیرتهاجمی از حیوانات تلف‌شده در اثر تصادفات جاده‌ای و لاشه‌های کشف‌شده از شکارچیان و موزه‌های تاریخ طبیعی در سراسر ایران، طی سال‌های ۱۳۹۶ الی ۱۳۹۷ تا حد امکان، انجام گرفت (جدول ۱). نمونه‌ها در تیوب‌های حاوی اتانول ۹۶ درصد و در دمای ۲۰- درجهٔ سانتی‌گراد تا پیش از استخراج DNA نگهداری شدند.

جدول ۱- فهرست نمونهٔ بافت‌های جمع‌آوری‌شدهٔ پستانداران از مناطق مختلف ایران

ردیف	نام فارسی	نام علمی	استان	نوع بافت
۱	روباه معمولی	*Vulpes vulpes*	کردستان	ماهیچه
۲	یوزپلنگ آسیایی	*Acinonyx jubatus venaticus*	سمنان	ماهیچه
۳	گربهٔ جنگلی	*Felis chaus*	ایلام	ماهیچه
۴	سیاه‌گوش	*Lynx lynx*	تهران	ماهیچه
۵	پلنگ ایرانی	*Panthera pardus saxicolor*	گیلان	ماهیچه
۶	فک خزر	*Pusa caspica*	گیلان	ماهیچه
۷	خرس قهوه‌ای سوری	*Ursus arctos syriacus*	قزوین	ماهیچه
۸	کفتار راه‌راه	*Hyaena hyaena*	قم	ماهیچه
۹	گراز	*Sus scrofa attila*	گیلان	ماهیچه
۱۰	مرال	*Cervus elaphus maral*	مازندران	ماهیچه
۱۱	گوزن زرد ایرانی	*Dama dama mesopotamica*	مازندران	ماهیچه
۱۲	شوکا	*Capreolus capreolus*	مازندران	ماهیچه
۱۳	آهوی گواتردار	*Gazella subgutturosa*	سمنان	ماهیچه
۱۴	قوچ و میش اوریال	*Ovis vignei*	گلستان	ماهیچه
۱۵	قوچ و میش ارمنی	*Ovis orientalis*	همدان	ماهیچه
۱۶	سنجاب ایرانی	*Sciurus anomalus*	کردستان	ماهیچه

مطالعات آزمایشگاهی

استخراج DNA از نمونهٔ بافت‌های جمع‌آوری‌شده براساس پروتکل Gentra (Qiagen, www.qiagen.com)، صورت گرفت و کمیت و کیفیت DNA استخراج‌شده با استفاده از روش اسپکتروفتومتری بررسی شد. تکثیر ناحیهٔ غیرکدشوندهٔ D-loop و ناحیهٔ ژنی COXI با استفاده از واکنش زنجیره‌ای پلی‌مراز (Polymerase Chain Reaction (PCR)) در حجم ۲۵ میکرولیتر انجام شد. چرخهٔ دمایی برای تکثیر نواحی نام‌برده عبارت بود از: پنج دقیقه در ۹۴ درجهٔ سانتی‌گراد و در ادامه ۳۵ چرخه، شامل ۳۰ ثانیه در ۹۴ درجهٔ سانتی‌گراد و ۳۰ ثانیه در شرایط دمایی متفاوت برای هر آغازگر منتخب (جدول ۲) و ۶۰ ثانیه مرحلهٔ طویل‌سازی در ۷۲ درجهٔ سانتی‌گراد و بسط نهایی با ۷۲ درجهٔ سانتی‌گراد در پنج دقیقه. برای حصول اطمینان از نتایج تحلیل‌های مبتنی‌بر PCR، همواره از کنترل‌های منفی و مثبت در آزمایش‌ها استفاده شد (Burkardt, 2000). محصولات حاصل از واکنش PCR، پس از بررسی موفقیت واکنش با الکتروفورز روی ژل آگارز یک درصد (Delidow *et al.*, 1993)، برای توالی‌یابی به شرکت Macrogen (Macrogen, Seoul, South Korea) ارسال شدند.

جدول ۲- توالی‌های آغازگرهای استفاده‌شده برای تکثیر نواحی ژنی منتخب

ناحیهٔ ژنی	توالی آغازگر	دمای اتصال	منابع
D-loop	L15995-F-5´-CTCCACTATCAGCACCCAAAG-3´ H16498-R-5´-CCTGAAGTAAGAACCAGATG-3´	۵۶ درجهٔ سانتی گراد	Taberlet and Bouvet, 1994; Fumagalli *et al.*, 1997
COXI	LCO1490-F-5´-GGTCAACAAATCATAAAGATATTGG-3´ HCO2198-R-5´-TAAACTTCAGGGTGACCAAAAAATCA-3´	۴۸ درجهٔ سانتی گراد	Folmer *et al.*, 1994

اخذ توالی‌های گونه‌های پستانداران تهدیدشدهٔ ایران از NCBI

طی پژوهش حاضر، توالی دو ناحیهٔ D-loop و COXI برای گونه‌های نمونه‌برداری‌شده (جدول ۱)، تهیه شد؛ سپس نام علمی کلیهٔ گونه‌های پستانداران ایران که در طبقات تهدیدشده قرار داشتند، در NCBI جستجو و توالی‌های متعلق به نشانگرهای D-loop، COXI و Cyt b اخذ شد؛ به‌عبارتی توالی‌های اخذشده، گونه‌های نمونه‌برداری‌شده و همچنین گونه‌هایی که طی این پژوهش نمونه‌برداری از آنها صورت نگرفت را شامل می‌شود.

راه‌اندازی پایگاه اطلاعاتی مولکولی محلی

توالی‌های تهیه‌شده طی پژوهش حاضر با نرم‌افزارهای Geneious pro v11.1.5 (Biomatters, www.geneious.com) و SeqScape v3 (Applied Biosystem, www.appliedbiosystem.com) ویرایش شد. فایلی با فرمت Fasta از توالی‌های DNA و پروتئین‌های حاصل از آن (توالی‌های تهیه‌شده طی این پژوهش و توالی‌های اخذشده از NCBI) به‌صورت مجزا تهیه شد؛ سپس از نرم‌افزار BLAST v.2.7.1+ به‌منظور ساخت پایگاه اطلاعاتی مولکولی محلی برای گونه‌های پستانداران تهدیدشدهٔ ایران استفاده شد. این نرم‌افزار این امکان را فراهم می‌آورد تا هر کاربری به‌صورت غیر برخط با استفاده از توالی‌های دلخواه و مدنظر خود یک پایگاه اطلاعاتی مولکولی محلی راه‌اندازی کند و جستجو و شناسایی توالی Query (توالی مجهول) را نیز در میان توالی‌های خود انجام دهد (Hu and Kurgan, 2019). امکان توسعهٔ پایگاه اطلاعاتی ساخته‌شده در آینده با استفاده از نرم‌افزار نام‌برده، نیز وجود دارد (Altschul *et al.*, 1990; Altschul *et al.*, 1997).

بررسی عملکرد پایگاه اطلاعاتی مولکولی در شناسایی توالی Query

به‌منظور بررسی عملکرد پایگاه اطلاعاتی مولکولی طراحی‌شده در شناسایی یک توالی Query، ابتدا نمونهٔ بافتی از لاشهٔ یک «مرال» (*Cervus elaphus maral*) از مرکز بازپروری در استان قم تهیه شد و توالی آن با استفاده از آغازگرهای استفاده‌شده در این پژوهش، به دست آمد. درواقع، برای حصول اطمینان از عملکرد پایگاه اطلاعاتی ساخته‌شده در شناسایی صحیح توالی Query، از نمونه‌ای با هویت معلوم استفاده شد؛ سپس فایلی با فرمت txt از توالی Query تهیه و از مجموعه‌ای از دستورهای blastn، blastp، blastx، tblastn و tblastx که خود زیرمجموعه‌ای از نرم‌افزار BLAST هستند، برای تطابق و شناسایی استفاده شد. blastp و blastx در پایگاه اطلاعاتی پروتئینی و blastn، tblastn و tblastx در پایگاه اطلاعاتی DNA جستجو می‌کنند. الگوریتم‌های مشابهی برای جستجو هنگام شناسایی به

کار گرفته می‌شود و تنها نوع توالی Query و پایگاه اطلاعاتی در هر جستجو متفاوت است (Altschul *et al.*, 1990).

نتایج

ویژگی‌های پایگاه اطلاعاتی مولکولی محلی

درحال حاضر، توالی‌های ۲۴ گونهٔ پستاندار ایران از چهار راستهٔ گوشتخواران (Carnivora)، فردسمان (Perissodactyla)، زوج‌سمان (Artiodactyla) و جوندگان (Rodentia) و ۱۰ خانواده در پایگاه اطلاعاتی مولکولی محلی طراحی‌شده، گنجانده شده است (جدول ۳). توالی‌های این پایگاه را توالی‌های تهیه‌شده از گونه‌های نمونه‌برداری‌شده طی این پژوهش، به‌علاوهٔ توالی‌های اخذشده از NCBI تشکیل داده‌اند. این پایگاه در دو نسخه شامل توالی‌های DNA و توالی‌های پروتئین طراحی، نصب و راه‌اندازی شد. توالی‌های پایگاه DNA را توالی‌های ناحیه‌های D-loop، COXI و Cyt b و توالی‌های پایگاه پروتئین را، توالی‌های پروتئین‌های حاصل از توالی DNA ناحیه‌های COXI و Cyt b تشکیل داده‌اند.

جدول ۳- اطلاعات توالی‌های گونه‌های موجود در پایگاه اطلاعاتی مولکولی محلی طراحی‌شده

ردیف	خانواده	گونه
۱	سگ‌سانان (Canidae)	شاه‌روباه (*Vulpes cana*)، روباه ترکمنی (*Vulpes corsac*)، روباه شنی (*Vulpes rueppellii*) و روباه معمولی (*Vulpes vulpes*)
۲	گربه‌سانان (Felidae)	یوزپلنگ آسیایی (*Acinonyx jubatus venaticus*)، گربهٔ جنگلی (*Felis chaus*)، سیاه‌گوش (*Lynx lynx*) و پلنگ ایرانی (*Panthera pardus saxicolor*)
۳	کفتارها (Hyaenidae)	کفتار راه‌راه (*Hyaena hyaena*)
۴	فک‌ها (Phocidae)	فک خزر (*Pusa caspica*)
۵	خرس‌ها (Ursidae)	خرس قهوه‌ای سوری (*Ursus arctos syriacus*)
۶	اسب و الاغ (Equidae)	گورخر آسیایی (*Equus hemionus onager*)
۷	خوک‌ها (Suidae)	گراز (*Sus scrofa attila*)
۸	گوزن‌ها (Cervidae)	مرال (*Cervus elaphus maral*)، گوزن زرد ایرانی (*Dama dama mesopotamica*) و شوکا (*Capreolus capreolus*)
۹	گاوسانان (Bovidae)	جیر (*Gazella bennettii*)، آهوی کوهی (*Gazella gazella*)، آهوی گواتردار (*Gazella subgutturosa*)، کل و بز (*Capra aegagrus*)، قوچ و میش اوریال (*Ovis vignei*) و قوچ و میش ارمنی (*Ovis orientalis*)
۱۰	سنجاب‌ها (Sciunidae)	سنجاب ایرانی (*Sciurus anomalus*)

شناسایی توالی Query توسط پایگاه اطلاعاتی مولکولی محلی

توالی‌های تهیه‌شده از نمونهٔ بافت متعلق به «مرال»، نواحی ژنی D-loop و COXI، با کمک پایگاه اطلاعاتی مولکولی جستجو شدند. براساس نتایج، توالی متعلق به ناحیهٔ COXI با توالی موجود از این ناحیه و همین‌طور توالی D-loop نیز با توالی موجود از این ناحیه در پایگاه اطلاعاتی تطبیق یافتند و بنابراین شناسایی توالی با موفقیت و به‌طور صحیح توسط پایگاه اطلاعاتی طراحی‌شده انجام شد (جدول ۴ و ۵). این

پایگاه اطلاعاتی، خروجی آنالیزهای انجام‌شده توسط کاربر را در قالب فایل‌هایی با فرمت txt ارائه می‌دهد. اطلاعات مهم این فایل‌های خروجی را نام و توالی گونهٔ تطبیق‌یافته و میزان عددی پارامترهای E value و Score bits تشکیل می‌دهند. پارامتر E value بیانگر میزان دفعاتی است که ممکن است تطابق توالی Query با توالی‌های مرجع تطبیق‌یافته در پایگاه اطلاعاتی به‌صورت تصادفی انجام شده باشد. Score bits نیز بیانگر میزان شباهت توالی Query با توالی مرجع تطبیق‌یافته در پایگاه اطلاعاتی است. نتایج تطبیق توالی COXI «مرال» (توالی Query) با پایگاه اطلاعاتی مولکولی طراحی‌شدهٔ خود شامل پنج نوع BLAST است که همگی نتایج مشابهی را نشان داده‌اند و در تمامی آنالیزهای BLAST انجام‌شده، توالی Query با توالی مرجع گونهٔ «مرال» تطبیق یافته است (جدول ۴)؛ اما آنالیز BLAST توالی D-loop «مرال» (توالی Query) تنها blastn را شامل می‌شود (جدول ۵)؛ زیرا ناحیهٔ D-loop در پروتئین‌سازی شرکت نمی‌کند و نمی‌توان برای آن سایر انواع BLAST را انجام داد.

جدول ۴- نتایج BLAST توالی ناحیهٔ COXI با پایگاه اطلاعاتی مولکولی

ردیف	گونهٔ تطبیق‌یافته	ناحیهٔ ژنی	blastn	blastp	blastx	tbastn	tblastx
۱	مرال	COXI	E value= $9e^{-119}$, Score bits= 409	E value= $1e^{-051}$, Score bits= 147	E value= $2e^{-023}$, Score bits= 75.5	E value= $1e^{-025}$, Score bits= 81.3	E value= $2e^{-046}$, Score bits= 166

جدول ۵- نتایج BLAST توالی ناحیهٔ D-loop با پایگاه اطلاعاتی مولکولی

ردیف	گونهٔ تطبیق‌یافته	ناحیهٔ ژنی	blastn
۱	مرال	D-loop	E value= $8e^{-169}$, Score bits= 579

بحث

پایگاه اطلاعاتی طراحی‌شده مانند کتابخانهٔ مرجع کوچکی است که توالی‌های گنجانده‌شده در آن به منزلهٔ اطلاعات کتاب‌های این کتابخانه هستند. درواقع، این اطلاعات که ماهیتی ژنتیکی دارد، نخستین بانک اطلاعاتی ژنتیکی برخی از گونه‌های پستانداران ایران را در قالب پایگاه اطلاعاتی مولکولی محلی فراهم آورده است. نخستین نسخهٔ این پایگاه اطلاعاتی درحال حاضر روی رایانهٔ آزمایشگاه تنوع زیستی پژوهشکدهٔ علوم محیطی دانشگاه شهید بهشتی راه‌اندازی شده است و قابلیت راه‌اندازی روی رایانه‌های دیگر آزمایشگاهها و استفادهٔ سایر پژوهشگران را نیز دارد.

همان‌طور که پایگاههای اطلاعاتی جهانی نظیر NCBI یا BOLD، توالی‌هایی با طول معین و کیفیت زیاد را شامل می‌شود، توالی‌های استفاده‌شده برای ساخت این پایگاه اطلاعاتی مولکولی نیز به‌منظور افزایش دقت در عملکرد آن، به دقت به‌طور کمی و کیفی بررسی و سنجش و از استفاده از توالی‌هایی با طول کوتاه یا کیفیت کم که خود منجر به ایجاد خطا

در فرآیند تطبیق و شناسایی توالی Query می‌شود، اجتناب شد.

ازجمله تفاوت‌های پایگاه اطلاعاتی مولکولی محلی طراحی‌شده در پژوهش حاضر با پایگاههای اطلاعاتی جهانی نظیر NCBI یا BOLD این بوده است که این پایگاه تنها توالی‌های برخی گونه‌های پستانداران ایران را شامل می‌شود و به‌صورت اختصاصی برای برخی از گونه‌های پستانداران ایران طراحی شده است. توالی‌های بارکد گنجانده‌شده در این پایگاه را گونه‌هایی از پستانداران ایران تشکیل داده‌اند که درحال حاضر اطلاعات بارکد اغلب آنها در پایگاههای اطلاعاتی جهانی موجود نیست که این امر آن را از بانک‌های اطلاعاتی ارائه‌پذیر به پژوهشگران توسط NCBI، متمایز ساخته است؛ علاوه‌بر آن این پایگاه برخلاف پایگاههای اطلاعاتی جهانی، آنالیزها را به‌صورت غیر برخط و بدون نیاز به اتصال به اینترنت انجام می‌دهد.

پایگاه اطلاعاتی مولکولی طراحی‌شده برخلاف سایر پایگاههای اطلاعاتی جهانی نام‌برده، امکان اخذ توالی مطابقت داده‌شده را به کاربر نمی‌دهد و اطلاعات خروجی آن نیز برخلاف سایر پایگاههای اطلاعاتی که شامل اطلاعات تکمیلی دربارهٔ تاکسون مطابقت داده‌شده است، به‌صورت فایلی با فرمت txt است که تنها اطلاعات توالی مطابقت داده‌شده و ارزش‌های عددی E value و Score bits را ارائه می‌دهد. درواقع، این پایگاه تنها برای تطابق و شناسایی توالی Query است.

در این پژوهش، طراحی و ایجاد پایگاه اطلاعاتی پروتئینی علاوه‌بر پایگاه اطلاعاتی DNA، امکان انجام انواع BLAST برای شناسایی توالی Query را میسر ساخته است؛ همان‌گونه که در پایگاههای اطلاعاتی جهانی نیز انجام انواع BLAST امکان‌پذیر است. آنالیز blastp که تنها با استفاده از توالی Query پروتئینی و پایگاه اطلاعاتی پروتئینی انجام‌پذیر است، منجر به حصول اطمینان بیشتر خواهد شد (Altschul *et al.*, 1997)؛ به‌عبارتی جستجو و تطابق توالی Query در پایگاه‌های اطلاعاتی پروتئینی نتایجی دقیق‌تر و مطمئن‌تر به همراه دارد (Yu *et al.*, 2007)؛ این هرگز بدان معنا نبوده است که نتایجی که تنها از آنالیز blastn به دست آمده‌اند مطمئن نیستند، بلکه حالت بهینه این است که برای شناسایی توالی Query از مجموع توالی‌های DNA و پروتئین آن و انواع BLAST در پایگاههای اطلاعاتی DNA و پروتئین استفاده شود.

استفاده از نشانگرهای مختلف میتوکندریایی در ایجاد پایگاه اطلاعاتی مولکولی از دیگر نقاط قوت این پایگاه به شمار می‌رود. همواره این عقیده وجود دارد که به‌کارگیری نشانگرهای متعدد میتوکندریایی برای انجام تطابق، منجر به کسب نتایجی دقیق‌تر و مطمئن‌تر خواهد شد (Mwale *et al.*, 2017). این امر میزان خطا در شناسایی توالی Query را به حداقل میزان ممکن کاهش می‌دهد. درواقع، کسب نتایج یکسان از تطابق‌های متعدد با نشانگرهای متعدد، این اطمینان را به وجود می‌آورد که هیچ‌گونه تطابقی تصادفی شکل نگرفته است.

پایگاه اطلاعاتی مولکولی طراحی‌شده طی پژوهش حاضر توالی متعلق به «مرال» را به‌طور صحیح بازشناخت؛ بنابراین این پایگاه شناسایی زیستگاههای پستانداران مطالعه‌شده را با استفاده از تطبیق توالی‌های متعلق به نمایه‌های زیستی موجود در زیستگاهها با اطلاعات توالی‌های گنجانده‌شده در آن، تسهیل می‌کند. روند این فرآیند بدین صورت خواهد بود: ۱.

ابتدا، نمایه‌های زیستی باقی‌مانده از گونه‌های پستانداران از زیستگاهها جمع‌آوری و مختصات جغرافیایی آنها ثبت می‌شود؛ ۲. از نمونه‌های جمع‌آوری‌شده DNA استخراج و با استفاده از نشانگرهای پیشنهادی در این پژوهش، توالی از آنها تهیه می‌شود؛ ۳. توالی‌های تهیه‌شده در جایگاه توالی Query با استفاده از پایگاه اطلاعاتی مولکولی محلی تهیه‌شده جستجو و هویت آنها مشخص می‌شود؛ ۴. درانتها، با معلوم‌شدن هویت نمایه‌ها، مشخص می‌شود هر ناحیه از زیستگاه که نمایه‌ها از آن برداشت شده‌اند، توسط چه گونه‌هایی استفاده می‌شوند. پایش گونه‌های پستانداران و همین‌طور شناسایی زیستگاههای آنها از طریق توالی‌یابی ژنوم نمایه‌های زیستی برجای‌مانده در زیستگاه و شناسایی آنها توسط پایگاه اطلاعاتی مولکولی، امروزه ازجمله روش‌هایی است که پژوهشگران به‌صورت گسترده استفاده می‌کنند و روشی دقیق برای مطالعهٔ پستانداران معرفی شده است (Laguardia *et al.*, 2015; Wheat *et al.*, 2016). بازیابی اطلاعات DNA متعلق به نمونه‌های جمع‌آوری‌شده از محیط، پتانسیل رفع چالش‌های موجود درزمینهٔ شناسایی زیستگاههای حیات وحش را دارد (Bohmann *et al.*, 2014; Thomsen and Willerslev, 2015; Ishige *et al*, 2017).

پایگاه اطلاعاتی مولکولی طراحی‌شده، هدف اصلی پژوهش حاضر مبنی‌بر کمک به رفع چالش‌های موجود درزمینهٔ شناسایی زیستگاههای پستانداران را میسر ساخته است. این پایگاه اطلاعاتی با تهیه و گردآوری اطلاعات ژنتیکی بخشی از تنوع زیستی این کشور، به شکلی غیرمستقیم نقشی مهم درزمینهٔ حفاظت از تنوع زیستی ارزشمند این مرز و بوم ایفا خواهد کرد. درحال حاضر تنها محدودیت این پایگاه در کوچک‌بودن آن است. هنگامی که این پایگاه اطلاعاتی هیچ‌گونه اطلاعاتی از توالی Query دردسترس نداشته باشد، این احتمال وجود دارد که شناسایی را با دقت کمتر و در سطوح فراتر از گونه یعنی جنس یا خانواده انجام دهد (Whitworth *et al.*, 2007; Khedkar *et al.*, 2014)؛ به‌عبارتی میزان دقت پایگاه اطلاعاتی در شناسایی توالی Query رابطهٔ مستقیمی با حجم اطلاعات پایگاه ازلحاظ طیف گونه‌های گنجانده‌شده در آن دارد (Kumar *et al.*, 2012)؛ اما همان‌طور که پیش از این نیز اشاره شد، امکان توسعه و گسترش آن در آینده وجود دارد و هرچه توالی‌های تعداد گونه‌های بیشتری در آن گنجانده شود، امکان شناسایی گونه‌های بیشتری میسر و شناسایی نیز با دقت بیشتری انجام خواهد شد. درواقع، این پایگاه پتانسیل تبدیل‌شدن به یک پایگاه اطلاعاتی مولکولی برای کل گونه‌های پستانداران ایران را دارد.

سپاسگزاری

نویسندگان مقاله مراتب قدردانی خود را از تمامی اشخاصی که در نمونه‌برداری پژوهش حاضر مساعدت کردند و همین‌طور از داوران محترم برای ارائهٔ نظرات بسیار ارزشمندشان، ابراز می‌کنند.

منابع

Akrim, F., Mahmood, T., Max, T., Nadeem, M.S., Qasim, S. & Andleeb, S. (2018). Assessment of bias in morphological identification of carnivore scats confirmed with molecular scatology in the north-eastern Himalayan region of Pakistan. *Peer Journal*, 6, e5262.

Aksöyek, E., Ibis, O., Özcan, S., Moradi, M. & Tez, C., (2016). DNA barcoding of three species (*Canis aureus, Canis lupus,* and *Vulpes vulpes*) of Canidae. *Mitochondrial DNA part A,* 28(5), 1-9.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology,* 215(3), 403-410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search program. *Nucleic Acid Research,* 25(17), 3389-3402.

Barea-Azcón, J. M., Virgós, E., Ballesteros-Duperón E., Moleón, M. & Chirosa, M. (2007). Surveying carnivores at large spatial scales: a comparison of four broad-applied methods. *Biodiversity and Conservation,* 16(4), 1213-1230.

Battersby, J. E. & Greenwood, J. J. D. (2004). Monitoring terrestrial mammals in the U.K: past, present, and future, using lessons from the bird world. *Mammal Review,* 34(1-2), 3-29.

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W. & Bruyn, M. D. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution,* 29(6), 358-367.

Borisenko, A. V., Lim, B. K., Ivanova, N. V., Hanner, R. H. & Hebert, P. D. N. (2008). DNA barcoding in surveys of small mammal communities: a field study in Suriname. *Molecular Ecology Resource,* 8(3), 471-479.

Burkardt, H. J. (2000). Standardization and Quality Control of PCR Analyses. *Clinical Chemistry and Laboratory Medicine,* 38(2), 87-91.

Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E. & Boutin, S. (2015). Review: wildlife camera trapping: a review and recommendation for linking surveys to ecological processes. *Journal of Applied Ecology,* 52(3), 675-685.

Clare, E. L., Lim, B. K., Engstrom, M. D., Eger, J. L. & Hebert, P. D. N. (2007). DNA barcoding of Neotropical bats: species identification and discovery within Guyana. *Molecular Ecology Notes,* 7(2), 184-190.

Delidow, B. C., Lynch, J. P., Peluso, J. J. & White, B. A. (1993). Polymerase Chain Reaction: Basic Protocols. *Methods in Molecular Biology,* 15, 1-29.

Farhadinia, M. S., Hunter, L. T. B., Jourabchian, A., Hosseini-Zavarei, F., Aakbari, H., Ziaie, H., Schaller, G. B. & Jowkar, H. (2017). The critically endangered Asiatic cheetah *Acinonyx jubatus venaticus* in Iran: a review of the recent distribution, and conservation status. *Biodiversity and Conservation,* 26(5), 1027-1046.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoej, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology,* 3(5), 294-299.

Fumagalli, L., Pope, L. C., Taberlet, P. & Moritz, C. (1997). Versatile primers for the amplification of the mitochondrial DNA control region in marsupials. *Molecular Ecology,* 6(12), 1199-1201.

Hajibabaei, M., Singer, G. A. & Hickey, D. A. (2006). Benchmarking DNA barcodes: an assessment using available primate sequences. *Genome,* 49(7), 851-854.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Journal of Proceedings of the Royal Society,* 270(1512), 313-322.

Hu, G. & Kurgan, L. (2019) Sequence similarity searching. *Current Protocols in Protein Science,* 95(1), e71.

Ishige, T., Miya, M., Ushio, M., Sado, T., Ushioda, M., Maebashi, K., Yonechi, R., Lagan, P. & Matsubayashi, H. (2017). Tropical-forest mammals as detected by environmental DNA at natural saltlicks in Borneo. *Biological Conservation*, 210, 281-285.

Kefi, R., Hsouna, S., Beraud-Colomb, E. & Abdelhak, S. (2009). Mitochondrial DNA: properties and applications. *Archieves de I'Institut Pasteur de Tunis*, 86(1-4), 3-14.

Khedkar, G. D., Abhayankar, S. B., Nalage, D., Ahmed, S. N. & Khedkar, C. D. (2014). DNA barcode-based wildlife forensics for resolving the origin of claw samples using a novel primer cocktail. *Mitochondrial DNA Part A*, 27(6), 3932-3935.

Kress, W. J., García-Robledo, C., Uriarte, M. & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution*, 30(1), 25-35.

Kumar, U. S., Ratheesh, R. V., Thomas, G. & George, S. (2012). The Use of DNA barcoding in wildlife forensics: a study of sambar deer (*Rusa unicolor*). *Forest Science and Technology*, 8(4), 224-226.

Laguardia, A., Wang, J., Shi, F. L., Shi, K. & Riordan, P. (2015). Species identification refined by molecular scatology in a community of sympatric carnivores in Xinjiang, China. *Zoological Research*, 36(2), 72-78.

Lorenz, J. G., Jackson, W. E., Beck, J. C. & Hanner, R. (2005). The problems and promise of DNA barcodes for species diagnosis of primate biomaterials. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462), 1869-1878.

Meiklejohn, C., Montooth, K. & Rand, D. (2007). Positive and negative selection on the mitochondrial genome. *Trends in Genetics*, 23(6), 259-263.

Mwale, M., Dalton, D. L., Jansen, R., Roelofse, M., Pietersen, D., Mokgokong, P. S. & Kotze, A. (2017). Forensic application of DNA barcoding for identification of illegally traded African pangolin Scales. *Genome*, 60(3), 272-284.

Nowak, C., Büntjen, M., Steyer, K. & Frosch, C. (2014). Testing mitochondrial markers for noninvasive genetic species identification in European mammals. *Conservation Genetics Resources*, 6(1), 41-44.

Parkanyi, V., Ondruska, L., Vasicek, D. & Slamecka, J. (2014). Multilevel D-loop PCR identification of hunting game. *Applied Translational Genomics*, 3(1), 1-7.

Ratnasingham, S. & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes*, 7(3), 355-364.

Robins, J., Hingston, M., Matisoo-Smith, E. & Ross, H. (2007) Identifying *Rattus* species using mitochondrial DNA. *Molecular Ecology Notes*, 7(5), 717-729.

Sadlier, L. M. J., Webbon, C. C., Baker, P. J. & Harris, S. (2004) Methods for monitoring red foxes *Vulpes vulpes* and badgers *Meles meles*: are field signs the answer? *Mammal Review*, 34(1–2), 75-98.

Smith, D. R. (2016). The past, present, and future of mitochondrial genomics: have we sequenced enough mtDNA? *Briefings in Functional Genomics*, 15(1), 47-54.

Taberlet, P. & Bouvet, J. (1994). Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proceeding of the Royal Society B: Biological Sciences*, 225(1344), 195-200.

Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, 21(8), 1789-1793.

Thomsen, P. F. & Willerslev, E. (2015). Environmental DNA, an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, 4-18.

Wheat, R. E., Allen, J. M., Miller, S. D. L., Wilmers, C. & Levi, T. (2016). Environmental DNA from residual saliva for efficient noninvasive genetic monitoring of Brown Bears (*Ursus arctos*). *PLoS One*, 11(11), e0165259.

Whitworth, T. L., Dawson, R. D., Magalon, H. & Baudry, E. (2007) DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). *Proceeding of The Royal Society B: Biological Sciences*, 274(1619), 1731-1739.

Yacoub, H. A., Fathi, M. M. & Mahmoud, W. M. (2013). DNA barcode through cytochrome *b* gene information of mtDNA in native chicken strains. *Mitochondrial DNA*, 24(5), 528-537.

Yu, Y., Gertz, E. M., Agarwala, R., Schäffer, A. A. & Altschul, S. F. (2006) Retrieval accuracy, statistical significance, and compositional similarity in protein sequence database searches. *Nucleic Acids Research*, 34(20), 5966-5973.