

A review of plant genetic diversity analysis using PCR-based markers banding pattern

Zahra Sadat Mousavi¹, Fatemeh Nasernakhaei¹(ORCID: 0000000277144331)

1. Department of Production Engineering and Plant Genetics, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

Correspondence:

Fatemeh Nasernakhaei

Email: f.nasernakhaei@scu.ac.ir

Received: 4/Aug/2024

Accepted: 4/Nov/2024

How to cite:

Mousavi, Z.S., & Nasernakhaei, F. (2025). A review of plant genetic diversity analysis using PCR-based markers banding pattern. *Crop Biotechnology*, 14 (2), 81-98. (DOI: [10.30473/cb.2024.71938.1980](https://doi.org/10.30473/cb.2024.71938.1980))

ABSTRACT

Genetic diversity is a crucial component of biodiversity, essential for preserving gene banks and enriching plant genetic resources. One way to assess genetic diversity is using band patterns (0 and 1) produced by PCR-based DNA markers (such as RAPD, SSR, ISSR, AFLP, SCoT etc.). After achieving reproducibility, the bands are selected, scored, and analyzed. The data is analyzed using multivariate statistical methods, including cluster and principal coordinate analysis. Genetic similarity or dissimilarity coefficients are used based on 0 and 1 data (for dominant and codominant markers), and allelic frequency coefficients (for codominant markers). Some algorithms such as UPGMA and Ward are employed to group the studied individuals and investigate their genetic relationships. Quantifying polymorphism and assessing genetic diversity within and between populations will depend on the type of marker (dominant and codominant) and reproduction mode. Parameters such as polymorphic information content (PIC), resolving power (Rp), marker index (MI), polymorphic loci/marker ratio, heterozygosity (H)/gene diversity, allelic diversity (A), the effective number of alleles (Ae), Shannon's index (I), Wright's F statistic (F_{IT} , F_{ST} , F_{IS}), G_{st} and analysis of molecular variance (AMOVA) are evaluated. Software tools (NTSYSpc, R, DARwin, PAST, Excel, PowerMarker, Popgen, and GenAlEx) can assist with grouping and estimating genetic diversity.

KEYWORDS

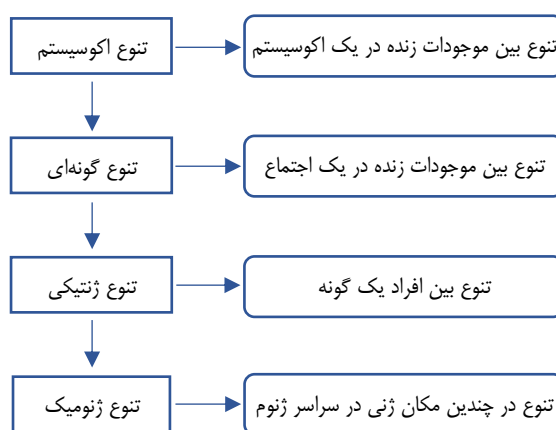
Dominant and Codominant markers, Gel scoring, Genetic diversity measurement, Multivariate statistical methods, Polymorphism measurement.

Figure 1. Types of diversity (Bhandari *et al.*, 2017)

به تغییرات ویژگی‌های وراثتی موجود در جمعیت یک گونه تنوع ژنتیکی می‌گویند (Swingland, 2001). تنوع در صفات وراثتی ممکن است خود را به شکل تغییر مورفولوژی، آناتومی، رفتار فیزیولوژیکی یا ویژگی‌های بیوشیمیایی نشان دهد (Bhandari *et al.*, 2017). تنوع ژنتیکی میان افراد، نشان‌دهنده وجود آلل‌های مختلف در مخزن ژنی و نیز ژنوتیپ‌های متفاوت در جمعیت‌ها می‌باشد (Caliskan, 2012). برخی منابع به گوناگونی‌های موجود در بین ژنوتیپ‌های مختلف با توجه به ترکیب ژنتیکی ژنوتیپ‌های مربوط به یک گونه یا بین گونه‌ها، تنوع ژنتیکی می‌گویند (Bhandari *et al.*, 2017). همچنین Salgotra and Chauhan (2023) تنوع ژنتیکی را در سه سطح تنوع بین گونه‌ها، تنوع بین جمعیت‌های یک گونه و تنوع بین افراد یک جمعیت تعریف می‌کنند. بنابراین تنوع ژنتیکی یک اصطلاح گسترده است که در تعبیر رایج، با گوناگونی ژنتیکی^{۱۰} به اشتباه، مترادف یکدیگر در نظر گرفته می‌شوند. گوناگونی^{۱۱} به تفاوت در یک یا چند صفت موجود زنده گفته می‌شود. به عبارتی گوناگونی ژنتیکی، تغییر در آلل‌های ژن یا تغییر در توالی‌های DNA/RNA در مخزن ژنی یک گونه یا جمعیت است که خود را در قالب فنوتیپ‌های متفاوت بیان می‌کند (Bhandari *et al.*, 2017). تنوع ژنتیکی از طریق شمارش تعداد ژن‌های مختلف در یک مخزن ژنی قابل اندازه‌گیری است، اما گوناگونی ژنتیکی^{۱۲} را تنها می‌توان تخمین زد و آن را می‌توان به عنوان بلوک‌های سازنده تنوع ژنتیکی در نظر گرفت (Bhandari *et al.*, 2017). جهش^{۱۳}، رانش ژنتیکی^{۱۴}، نوترکیبی^{۱۵} و جریان ژنی^{۱۶} برخی از منابع مختلف ایجادکننده تنوع ژنتیکی هستند (Salgotra and Chauhan, 2023).

مقدمه

تنوع زیستی^۱ به گوناگونی^۲ موجودات زنده اشاره دارد (Swingland, 2001) که ارزش آن برای بشر در دهه‌های اخیر به خوبی شناخته شده است و بسیاری استدلال می‌کنند که وجود آن برای توسعه پایدار فعالیت‌های مختلف انسانی، ضروری است (Shiva, 1994). به‌طور کلی تنوع در چهار سطح تعریف می‌گردد (شکل ۱). تنوع اکوسیستم^۳ نشان‌دهنده تغییرپذیری در میان جوامع مختلف گونه‌ها است. در سطح بعدی تنوع گونه‌ای^۴ است که نشان‌دهنده گونه‌های مختلف در یک اجتماع^۵ است که به آن غنای گونه‌ای^۶ نیز گفته می‌شود. تنوع ژنتیکی^۷ به تنوع موجود در ژنوتیپ‌های مختلف یک گونه اطلاق می‌گردد. به این دلیل، وجود آلل‌های متفاوت یک ژن در افراد مختلف، سبب ایجاد فنوتیپ‌های متفاوت می‌شود. تنوع ژنومیک^۸ را می‌توان به عنوان تنوع در چندین مکان ژنی^۹ در یک فرد تعریف کرد. در این میان، تنوع ژنتیکی بیش‌ترین توجه را در بین محققین بخش کشاورزی به خود اختصاص داده است (Bhandari *et al.*, 2017).



شکل ۱. انواع تنوع (Bhandari *et al.*, 2017)

10. Genetic variability
11. Variability
12. Genetic variation
13. Mutation
14. Genetic drift
15. Recombination
16. Gene flow

1. Biodiversity
2. Variety
3. Ecological diversity
4. Species diversity
5. Community
6. Species richness
7. Genetic diversity
8. Genomic diversity
9. Gene-loci

همچنین با استفاده از نشانگرها، شناسایی مولکولی^۴ منابع گیاهی در سطح DNA از طریق الگوی باندها یا انگشت نگاری^۵ امکان پذیر است. مقایسه اثر انگشت در مطالعات منابع ژنتیکی گیاهی برای برآورد تنوع ژنتیکی در یک مجموعه^۶ (درون یا بین گروهی)، و یا تخمین مفید بودن نشانگرهای DNA در شناسایی یک اکشن^۷ خاص استفاده می شود (Laurentin, 2009).

در این مطالعه مفاهیم مرتبط با ارزیابی تنوع ژنتیکی منابع گیاهی و چگونگی اندازه گیری آن با استفاده از نتایج حاصل از نشانگرهای مولکولی مبتنی بر PCR که شناسایی چندشکلی در آن‌ها بر اساس جداسازی روی ژل و امتیازدهی^۸ باندها (الگوی نوارها) است، مرور می گردد.

نشانگرهای مولکولی مبتنی بر PCR

نشانگرهای مولکولی مبتنی بر PCR (Kumar *et al.*, 2009; Kalendar and Schulman, 2014; Dar *et al.*, 2024; 2019; Amiteye, 2021; Bidyananda *et al.*, 2024) (جدول ۱) بازتاب تفاوت‌های موجود در ردیف DNA افراد می باشند و چندشکلی^۹ را در سطح DNA با استفاده از PCR شناسایی می کنند. این نشانگرها نواحی ژنومی رمزکننده و غیر رمزکننده خاصی هستند که به منظور تجزیه و تحلیل ژنوم مورد استفاده قرار می گیرند. همچنین به دلیل فراوانی فوق العاده آن‌ها، عدم تاثیرپذیری از شرایط محیطی، دقت و قابلیت مطلوب تفسیر نتایج، هم بارز^{۱۰} بودن بسیاری از این نشانگرها و سهولت تشخیص افراد هتروزیگوت^{۱۱} (ناخالص) از هموزیگوت^{۱۲} (خالص) اهمیت بیش تری نسبت به نشانگرهای ریخت شناسی پیدا کرده اند (Agarwal *et al.*, 2008; Naghavi *et al.*, 2009;) (Heidari, 2011; Darvish Zadeh and Azizi, 2015).

ذخایر ژنتیکی گیاهی یکی از ارزشمندترین منابع طبیعی می باشند که آگاهی از تنوع موجود در آن از ملزومات ایجاد و غنی سازی بانک های ژنی و همچنین توسعه ی انواع محصولات زراعی جدید می باشد (Mohammadi, 2006; Darvish Zadeh and Azizi, 2015). بنابراین حفظ و استفاده از این ذخایر ژنتیکی نقش مهمی در کشاورزی، امنیت غذایی و جنگلداری خواهد داشت. موضوع تنوع ژنتیکی برای کشاورزان و به نژادگران به منظور به دست آوردن ارقام جدید با عملکرد بالا، کیفیت بهتر یا سازگارتر در برابر تنش های غیرزیستی (مانند گرما و سرمای شدید)، یا مقاوم تر در برابر آفات و عوامل بیماری زا و نیز مقاوم به آلاینده های مختلف هوا و خاک لازم است (Rao, 2004; Laurentin, 2009; Bhandari *et al.*, 2017; Shashidhara *et al.*, 2023). لذا وجود تنوع ژنتیکی بالا برای یک جمعیت به لحاظ بقا، سازگاری و تکامل آن جمعیت حیاتی است (Nonić and Šijačić-Nikolić, 2019).

با توجه به اهمیت تنوع ژنتیکی، ارزیابی آن با استفاده از نشانگرهای ژنتیکی^۱ می تواند اطلاعات ارزشمندی را در زمینه حفظ و نگهداری از گیاهان و همچنین استفاده از آن ها در برنامه های اصلاحی در اختیار پژوهشگران قرار دهد. نشانگرهای ژنتیکی، اشکال بیولوژیکی مانند ویژگی های فنوتیپی^۲ (صفات ریخت شناسی)، آنزیم ها/پروتئین ها و قطعاتی از DNA بوده که از والدین به نتاج^۳ به ارث می رسند و به دلیل استفاده از آن ها در ردیابی یک فرد یا یک ژن، نشانگر نامیده می شوند (Kadirvel *et al.*, 2015). نشانگرهای ژنتیکی به نشانگرهای ریخت شناسی و نشانگرهای مولکولی در سطح پروتئین و DNA (مبتنی و غیرمبتنی بر PCR) تقسیم می شوند. این نشانگرها در مطالعات فیلوژنتیک و تعیین میزان قرابت، انتقال ژن از گونه های خودرو، رده بندی گیاهان زراعی، تعیین محل ژن های کنترل کننده صفات کیفی و کمی، تهیه نقشه های ژنتیک و تسهیل انتخاب به کمک نشانگر کاربرد دارند (Naghavi *et al.*, 2009).

4. Molecular identity
5. Fingerprint
6. Collection
7. Accession
8. Scoring
9. Polymorphism
10. Co-dominant
11. Heterozygous
12. Homozygous

1. Genetic marker
2. Traits
3. Progeny

پلیمراز انجام می‌گیرد. جهت مشاهده قطعات تکثیر شده می‌توان از ژل آگارز یا پلی‌آکریل آمید استفاده نمود. سپس باندهای مشاهده شده بر روی ژل امتیازدهی شده و نهایتاً تجزیه و تحلیل داده‌ها انجام می‌گردد.

امتیازدهی

امتیازدهی باندها به صورت دستی یا با کمک برخی نرم‌افزارها (Mukhopadhyay and Bhattacharjee, 2016) نظیر TotalLab, PyElph, CLIQS, VisionWorks و Quantity One, ImageQuant انجام می‌شود. روش کار به این صورت خواهد بود که باندهای واضح انتخاب و به صورت وجود و عدم وجود باند (یک و صفر) برای نشانگرهای غالب و هم‌بارز و به صورت فراوانی آلی برای نشانگرهای هم‌بارز امتیازدهی می‌شود؛ البته در اکثر موارد نشانگرهای هم‌بارز مانند نشانگرهای غالب به صورت یک و صفر امتیازدهی می‌شوند (Mohammadi, 2006; Heras et al., 2016) (شکل ۳) و داده‌ها به صورت یک ماتریس دوتایی^۷ در اکسل آماده می‌شود (Sarwat, 2012). این ماتریس برای نشانگرهای غالب و هم‌بارز متفاوت است (شکل ۳). در صورت استفاده از نشانگر هم‌بارز، ماتریس دوتایی باید به ماتریسی از ژنوتیپ‌ها (هموزیگوت و هتروزیگوت) برای هر فرد در هر مکان^۸ تبدیل شود (Laurentin, 2009).

تجزیه و تحلیل اطلاعات

تجزیه و تحلیل تنوع ژنتیکی و ساختار جمعیت‌ها شامل کمی کردن تنوع درون جمعیت^۹ و بین جمعیت‌ها^{۱۰} و یا افراد^{۱۱}، کمی کردن روابط ژنتیکی^{۱۱} و نمایش روابط ژنتیکی است (De Vicente et al., 2004) که در ادامه به آن‌ها پرداخته می‌شود.

نشانگرهای مولکولی ممکن است به صورت نشانگرهای چند مکانی^۱ و یا تک مکانی^۲ باشند. در یک تجزیه و تحلیل با استفاده از نشانگرهای چند مکانی، امکان مشاهده باندهای زیادی وجود دارد، در حالی که برای نشانگرهای تک مکانی، حداکثر دو باند که مربوط به آلل‌های ژنوم دیپلوئید می‌باشد، رویت می‌گردد. بسته به اهداف مطالعه، محقق ممکن است از نشانگرهای تک یا چند مکانی و یا ترکیبی از هر دو استفاده کند. نشانگرهای چند مکانی اطلاعات گسترده‌ای از ژنوم افراد را ارائه می‌دهند، به همین دلیل در مطالعات ژنتیک جمعیت، با توجه به اینکه هدف تمایز میان افراد یک جمعیت بر اساس الگوی باندها (نوارها) است، نشانگرهای چند مکانی کارآمدتر هستند (Simpson, 1997). نشانگرهای مولکولی می‌توانند به صورت نشانگرهای چندشکل^۳ و یک شکل^۴ مشاهده شوند؛ بدین معنا که اگر نشانگر قادر به شناسایی تفاوت‌های میان افراد یک گونه یا گونه‌های متفاوت باشند، چندشکل و در صورتی که پتانسیل ایجاد تمایز میان ژنوتیپ‌ها را نداشته باشند، یک شکل تلقی می‌گردند (Kanouni and Khalily, 2010). نشانگرها را می‌توان تحت عنوان غالب^۵ (بارز) و هم‌بارز نیز تقسیم‌بندی کرد (Kadirvel et al., 2015). مزیت نشانگرهای هم‌بارز نسبت به نشانگرهای غالب، تمایز افراد هموزیگوت از هتروزیگوت است (Laurentin, 2009).

روند شناسایی گوناگونی‌های ژنتیکی افراد مورد مطالعه (ژنوتایپینگ)^۶ با استفاده از نشانگرهای

مبتنی بر PCR بر اساس امتیازدهی ژل

روش کلی ژنوتایپینگ با استفاده از نشانگرهای DNA مبتنی بر PCR به شرح شکل ۲ می‌باشد. ابتدا DNA گیاهی استخراج شده و پس از سنجش کمیت و کیفیت آن، با استفاده از نشانگرهای DNA واکنش زنجیره‌ای

7. Binary matrix

8. Locus

9. Intrapopulation genetic diversity

10. Interpopulation genetic diversity

11. Genetic relationships

1. Multi-locus

2. Single-locus

3. Polymorph

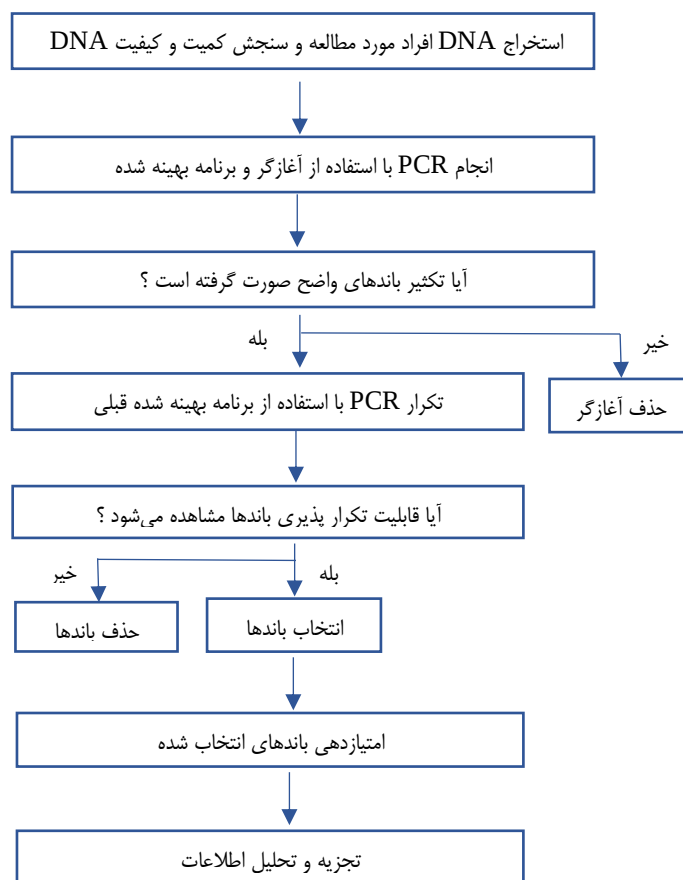
4. Monomorph

5. Dominant

6. Genotyping

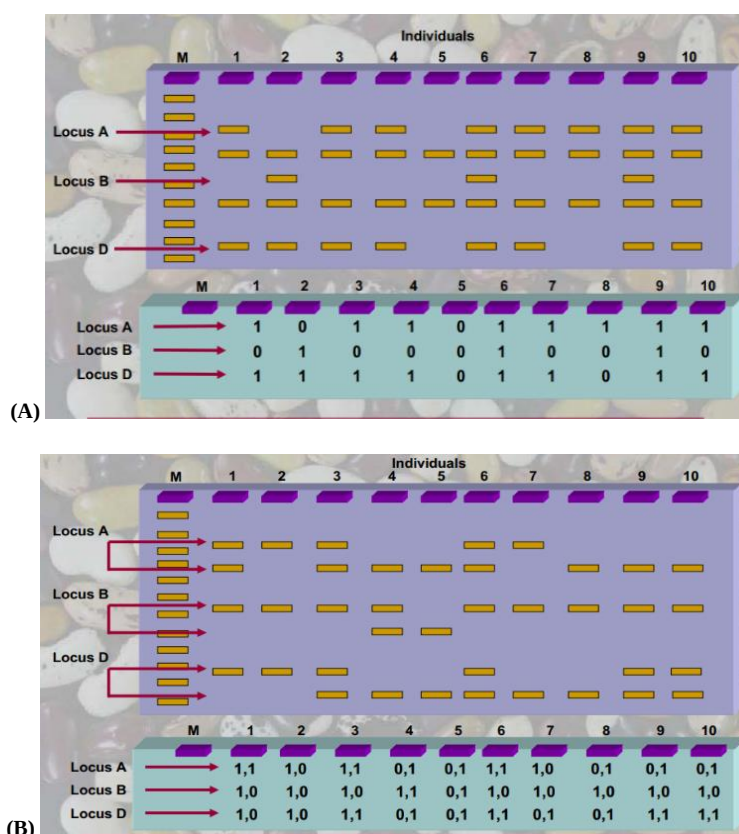
جدول ۱. برخی از نشانگرهای مبتنی بر PCR و استفاده شده در ارزیابی تنوع ژنتیکی
Table 1. Some of the PCR-based markers used in genetic diversity assessment

نام نشانگر	منبع	مزایا	معایب و محدودیت‌ها	مثال
دی‌ان‌ای چندشکل تکثیر شده تصادفی Amplified Polymorphic DNA	Williams <i>et al.</i> , (1990)	پوشش بهتر ژنوم	تکرارپذیری پایین، نشانگر غالب	<i>Abelmoschus esculentus</i> L. (Hamdan <i>et al.</i> , 2024)
ردیف‌های تکراری ساده یا ریزماهوره Repeats	Akkaya <i>et al.</i> , (1992)	تکرارپذیری و چندشکلی بالا، نشانگر هم‌پارز	نیاز به اطلاعات توالی	<i>Hordeum vulgare</i> L. (Esmaili <i>et al.</i> , 2014); <i>Solanum lycopersicum</i> L. (Hosseini <i>et al.</i> , 2023)
توالی‌های چندشکل تکثیر شده شکسته شده Amplified Polymorphic Sequence	Akopyanz <i>et al.</i> , (1992)	تکرارپذیری بالا، نشانگر هم‌پارز	نیاز به اطلاعات توالی برای ساخت آغازگر	<i>Solanum lycopersicum</i> L. (Pozharskiy <i>et al.</i> , 2023)
ناحیه تکثیر شونده با ردیف مشخص Characterized Amplified Region	Paran and Michelmore, (1993)	تکرارپذیری بالا، نشانگر هم‌پارز	نیازمند توالی ژن جهت ساخت آغازگر	<i>Amaranthus</i> (Ray and Roy, 2009)
تفاوت‌های ریزماهوره تکثیرشونده تصادفی Amplified Microsatellite Polymorphisms	Wu <i>et al.</i> , (1994)	عدم نیاز به اطلاعات توالی برای ساخت آغازگر، نشانگر هم‌پارز	مشکل تکرارپذیری	<i>Hordeum vulgare</i> (Yongcui <i>et al.</i> , 2005)
توالی‌های تکراری ساده میانی Inter-Simple Sequence Repeat	Zietkiewics <i>et al.</i> , (1994)	عدم نیاز به اطلاعات توالی برای ساخت آغازگر، نیاز به مقدار کم DNA	مشکل تکرارپذیری، نشانگر غالب	<i>Ricinus communis</i> (Sharifi Soltani <i>et al.</i> , 2023)
تفاوت طول قطعات حاصل از تکثیر Fragment Length Polymorphism	Vos <i>et al.</i> , (1995)	عدم نیاز به اطلاعات توالی	مشکل تکرارپذیری، نشانگر غالب	<i>Oryza sativa</i> L. (Alizadeh <i>et al.</i> , 2013); <i>Sorghum bicolor</i> L. Moench (Medraoui <i>et al.</i> , 2024)
چندشکلی در تکثیر بین رتروترانسپوزون‌ها-IRAP Retrotransposon Amplified Polymorphism	Kalendar <i>et al.</i> , (1999)	چندشکلی و پوشش ژنومی بالا	تشخیص آلل، نشانگر غالب	<i>Gladiolus</i> (Nazarbeigi <i>et al.</i> , 2024)
چندشکلی در تکثیر بین ریزماهوره و رتروترانسپوزون Retrotransposon-Microsatellite Amplified Polymorphism	Kalendar <i>et al.</i> , (1999)	چندشکلی و پوشش ژنومی بالا	از نظر تکنیکی دارای مشکل تشخیص آلل، نشانگر غالب	<i>Triticum turgidum</i> L. (Marzang <i>et al.</i> , 2020)
چندشکلی نواحی هدف کدون آغاز (SCoT) Start Codon Targeted	Collard and Mackill, (2009)	سادگی، مقرون به صرفه	گاها مشکل تکرارپذیری، نشانگر غالب	<i>Foeniculum vulgare</i> Mill. (Nikkerdar <i>et al.</i> , 2018); <i>Cressa cretica</i> L. (Jahangir, 2021)



شکل ۲. ژنوتایپینگ با استفاده از نشانگرهای DNA مبتنی بر PCR بر اساس الگوی باندها (Ng and Tan, 2015)

Figure 2. Genotyping using PCR-based DNA markers banding pattern (Ng and Tan, 2015)



شکل ۳. امتیازدهی ژل نشانگر (A) غالب و (B) هم‌بارز در ۱۰ فرد (De Vicente *et al.*, 2004)

Figure 3. Gel scoring A) dominant and B) codominant marker in 10 individuals (De Vicente *et al.*, 2004)

Vicente *et al.*, 2004; Nonić and Šijačić-Nikolić, 2019). همچنین در ارزیابی تنوع بین جمعیتی از تمایز بین جمعیتی برای یک مکان^۳ (g_{st})، تمایز بین جمعیتی برای چندین مکان^۴ (G_{st})، سهم جمعیت در تنوع ژنتیکی کل^۵، آماره F رایت^۶ و تجزیه واریانس مولکولی^۷ (AMOVA) استفاده می‌شود (De Vicente *et al.*, 2004; Nonić and Šijačić-Nikolić, 2019). برای برآورد پارامترهای مذکور می‌توان از نرم‌افزارهایی همچون GenAlEx و POPGENE، PowerMarker، Excel بهره برد. در ادامه به توضیح برخی از مفاهیم پارامترهای مورد استفاده برای بررسی میزان چندشکلی و تنوع ژنتیکی (جدول ۲) پرداخته می‌شود.

جمعیت به گروهی از افراد که علاوه بر مخزن ژنی مشترک، از توانایی تلاقی با یکدیگر نیز برخوردار هستند اطلاق می‌شود. ساختار هر جمعیت در سه سطح افراد، زیرجمعیت^۲ و جمعیت کل تقسیم می‌گردد. توصیف تنوع ممکن است در درون یک جمعیت یا بین جمعیت‌ها انجام شود (De Vicente *et al.*, 2004). زمانی که الگوی باندها امتیازدهی شد، می‌بایست بر اساس نوع نشانگر مورد استفاده (غالب یا هم‌بارز) و سیستم زادآوری گیاه، میزان چندشکلی و تنوع مشاهده شده در درون یا بین جمعیت‌ها از نظر کمی تعیین شود. از مهم‌ترین پارامترها در برآورد تنوع درون جمعیتی می‌توان به اندازه‌گیری میزان چندشکلی، نسبت مکان‌های چندشکل، میانگین تعداد آلل‌ها در هر مکان، تعداد آلل‌های موثر در هر مکان و میانگین هتروزیگوسیتی مورد انتظار اشاره کرد (De

3. Interpopulation differentiation for one locus (g_{st})
 4. Interpopulation differentiation for several loci (G_{st})
 5. Population's contribution to total genetic diversity
 6. Wright's F statistics
 7. Analysis of Molecular Variance (AMOVA)

1. Individuals
 2. Subpopulation

جدول ۲. برخی از پارامترهای مورد استفاده در برآورد میزان چندشکلی و تنوع ژنتیکی

Table 2. Some parameters used in the estimation of polymorphism and genetic diversity

پارامتر	محتوای اطلاعات چندشکلی	قدرت تفکیک	شاخص نشانگری	نسبت مکان های ژنی چندشکل / نسبت نشانگرهای چندشکل	هتروزیگوسیتی / تنوع ژنی	شاخص شانون	تنوع آلی	تعداد آلل های موثر	تجزیه و تحلیل واریانس مولکولی	آماره F رایت	پارامترهای نی
غالب	✓	✓	✓	✓**	✓**	✓			✓	✓§	✓
هم بارز	✓	✓	✓	✓*	✓*	✓	✓	✓	✓	✓	✓

§ آماره F_{ST} برای نشانگرهای غالب نیز قابل محاسبه است.محتوای اطلاعات چندشکلی^۱

محتوای اطلاعات چندشکلی نشان دهنده توانایی یک نشانگر در تشخیص چندشکلی میان افراد یک جمعیت است و به تعداد آلل های موجود و فراوانی توزیع آن ها بستگی دارد. لذا هر چه این پارامتر بالاتر باشد، ارزش آن نشانگر بیش تر است (Serrote *et al.*, 2020). این پارامتر برای نشانگرهای هم بارز از طریق رابطه $PIC = 1 - \sum_{i=1}^n p_i^2$ (Botstein 1980) و برای نشانگرهای غالب به روش $PIC = 2f_i(1 - f_i)$ محاسبه می شود که در آن ها منظور از p_i و f_i فراوانی آلل نام است. دامنه تغییرات این پارامتر برای نشانگرهای هم بارز از صفر تا یک و برای نشانگرهای غالب صفر تا نیم می باشد (Serrote *et al.*, 2020). برخی محققین همچون Chesnokov and Artemyeva (2015) آن را معادل تنوع ژنی^۲ در نظر گرفته اند.

نسبت مکان های ژنی چندشکل /نسبت نشانگرهای چندشکل^۳

یکی از پارامترهایی که برای بررسی تنوع ژنتیکی با استفاده از نشانگرهای غالب و هم بارز قابل استفاده است محاسبه چندشکلی می باشد. از آن جا که Ott (1992)

سیستمی که فراوانی رایج ترین مرحله^۴ آن کم تر از ۰/۹۵ باشد را به عنوان چندشکل معرفی می کند، در نشانگرهای غالب که فرض می شود هر باند یک مکان را نشان می دهد مرحله ای که او به آن اشاره می کند همان توجه به «حضور و عدم حضور باند در هر فرد» در میان افراد تشکیل دهنده گروه مورد مطالعه است. به عنوان مثال اگر یک باند در ۹۴ درصد از افراد دیده شود، رایج ترین مرحله «حضور باند» خواهد بود، بنابراین به عنوان چندشکل در نظر گرفته می شود. همچنین اگر باندی فقط در ۶ درصد افراد مشاهده شود، شایع ترین مرحله «عدم حضور باند» خواهد بود که در ۹۴ درصد افراد دیده می شود، بنابراین باز هم، باند چندشکل خواهد بود (Laurentin, 2009). این پارامتر برای نشانگرهای غالب به جای مکان های چندشکل، نشانگر چندشکل نامیده می شود. در نشانگرهای هم بارز یک مکان با تعداد باندهایی که آلل می نامند نشان داده می شود؛ بنابراین، اگر رایج ترین مرحله یعنی حضور یا عدم حضور باند (آلل) در کم تر از ۹۵ درصد از افراد تشکیل دهنده گروه مورد مطالعه دیده شود، مکان چندشکل است. این پارامتر به صورت نسبت تعداد مکان های ژنی چندشکل بر تعداد کل مکان ها برای نشانگرهای هم بارز یا نسبت نشانگرهای چندشکل بر تعداد کل نشانگرها برای نشانگرهای غالب قابل محاسبه است (Laurentin, 2009).

1. Polymorphic Information Content (PIC)

2. Gene diversity

3. Polymorphic loci/marker ratio

4. Stage

اندازه نمونه کم تر حساس است. هنگامی که H_E و H_O مشابه هستند (تفاوت قابل توجهی نداشته باشند)، تلاقی در جمعیت تقریباً تصادفی است (Chesnokov and Artemyeva, 2015). هتروزیگوسیتی برای هر مکان به روش $h = 1 - \sum p_i^2$ محاسبه می گردد که در آن p_i فراوانی آلل نام است (De Vicente et al., 2004). مقادیر این پارامتر از صفر تا یک متغیر است و زمانی که تعداد زیادی آلل در فراوانی های مساوی وجود داشته باشد هتروزیگوسیتی مورد انتظار به حداکثر می رسد. میانگین هتروزیگوسیتی مورد انتظار در تمام مکان ها تخمینی از گوناگونی ژنتیکی^۶ در جمعیت است که با کم کردن فراوانی های مورد انتظار هموزیگوت ها در یک مکان از ۱ محاسبه می شود، این عمل برای همه مکان ها تکرار و سپس میانگین آن ها محاسبه می گردد (De Vicente et al., 2004). به گفته Laurentin (2009) رابطه فوق جهت محاسبه فراوانی آللی است و هتروزیگوسیتی یا تنوع ژنی برای هر مکان طبق رابطه $h = 2n(1 - \sum x_i^2)/(2n - 1)$ برآورد می گردد؛ که در این رابطه منظور x_i فراوانی آلل نام (یا نشانگر) و n تعداد افراد است. برای به دست آوردن یک دید کلی از تنوع ژنتیکی در یک جمعیت (از نظر مفهومی) هتروزیگوسیتی/تنوع ژنی مفهومی مناسب تر از PIC است (Laurentin, 2009).

تنوع آللی^۸ و تعداد آلل های موثر^۹

ساده ترین راه برای اندازه گیری تنوع ژنتیکی، کمی کردن تعداد آلل های موجود است. تنوع آللی (A) میانگین تعداد آلل ها در هر مکان است و برای توصیف تنوع ژنتیکی استفاده می شود و از طریق $A = n_i/n_1$ برآورد می گردد که در آن n_i تعداد کل آلل های موجود در مکان ها و n_1 تعداد مکان ها است (Govindaraj et al., 2015; Mukhopadhyay and Bhattacharjee, 2016). این پارامتر نسبت به اندازه نمونه حساس بوده و با افزایش

قدرت تفکیک^۱ و شاخص نشانگری^۲

قدرت تفکیک پارامتری است که برای مشخص کردن توانایی نشانگر در تشخیص تفاوت میان ژنوتیپ ها استفاده می شود و به روش $R_p = \sum I_b$ محاسبه می گردد (Kumar and Agrawal, 2019). در این رابطه، منظور از I_b اطلاعات باندها^۳ است که میزان عددی آن از صفر تا یک متغیر و حاصل از رابطه $I_b = 1 - (2|0.5 - p|)$ می باشد. در رابطه مذکور p نسبت افراد دارای باند به کل افراد است (Chesnokov and Artemyeva, 2015; Govindaraj et al., 2015).

شاخص نشانگری معیاری برای تعیین کارایی نشانگر در شناسایی مکان های چندشکل در بین ژنوتیپ های مختلف است که از حاصل ضرب محتوای اطلاعات چندشکلی (PIC) و نسبت چندگانه موثر^۴ (EMR) از طریق رابطه $MI = EMR \times PIC$ برآورد می گردد. نسبت چندگانه موثر از طریق رابطه $EMR = n_p(n_p/n)$ محاسبه می گردد که در آن منظور از n_p تعداد مکان های چندشکل و n تعداد کل مکان ها است (Nagaraju et al., 2001; Kumar and Agrawal, 2019).

هتروزیگوسیتی/تنوع ژنی^۵

هتروزیگوسیتی (H) احتمال متفاوت بودن دو آلل در یک مکان (که به طور تصادفی از جمعیت انتخاب شده اند) است که تنها در نشانگرهای هم بارز (به دلیل شناسایی افراد هموزیگوت از هتروزیگوت) قابل استفاده است. اصطلاح تنوع ژنی همین کاربرد را دارد با این تفاوت که برای نشانگرهای غالب به کار می رود زیرا افراد هموزیگوت و هتروزیگوت به طور مستقیم متمایز نمی گردند (Laurentin, 2009). هتروزیگوسیتی در دو سطح هتروزیگوسیتی مورد انتظار (H_E) و هتروزیگوسیتی مشاهده شده (H_O) تعریف می گردد که معمولاً از H_E به منظور توصیف تنوع ژنتیکی استفاده می گردد زیرا نسبت به

6. Frequencies

7. Genetic variability

8. Allelic diversity (A)9. Effective number of alleles (A_e)

1. Resolving Power (RP)

2. Marker Index (MI)

3. Informative bands (I_b)

4. Effective Multiplex Ratio (EMR)

5. Heterozygosity (H)/Gene diversity

بیش‌تری نسبت به غنای آلی یا تعداد آل‌ها می‌باشد (Sherwin *et al.*, 2017). این پارامتر بر اساس رابطه p_i فراوانی حضور یا عدم حضور یک باند در یک مکان است (Shannon and Weaver, 1949). از نرم‌افزاری نظیر GenAlEx می‌توان برای اندازه‌گیری این شاخص استفاده نمود.

آماره F رایت^۴ و G_{st}^۵

یکی از راه‌های مطالعه ساختار ژنتیکی جمعیت‌ها استفاده از آماره‌های F (F_{IT}, F_{ST}, F_{IS}) است که امکان آنالیز ساختار جمعیت‌های در حال تقسیم^۶ را فراهم می‌کند (De Vicente *et al.*, 2004). آماره F رایت یا شاخص تثبیت^۷ به منظور برآورد تفاوت هتروزیگوسیتی مشاهده شده و هتروزیگوسیتی مورد انتظار وقتی که تعادل هاردی واینبرگ برقرار باشد به کار می‌رود (Laurentin, 2009). تمایز در سه سطح کل گروه ($F_{IT} = 1 - (H_I/H_T)$), بین گروه‌های فرعی ($F_{ST} = 1 - (H_S/H_T)$) و درون گروه‌های فرعی ($F_{IS} = 1 - (H_I/H_S)$) ارزیابی می‌شود؛ که در این روابط منظور از H_I میانگین هتروزیگوسیتی مشاهده شده در تمام مکان‌ها، H_T میانگین هتروزیگوسیتی مورد انتظار در کل گروه و H_S میانگین هتروزیگوسیتی مورد انتظار در گروه‌های فرعی است. آماره‌های مذکور برای نشانگرهای هم‌باز قابل محاسبه‌اند زیرا نیازمند برآورد هتروزیگوسیتی مشاهده شده می‌باشند (Laurentin, 2009). آماره F_{ST} برای نشانگرهای غالب نیز قابل اجرا است. دامنه تغییرات آن از صفر (عدم تمایز ژنتیکی) تا یک (حداکثر تمایز) متغیر می‌باشد؛ در این میان بازه ۰/۰۵-۰/۰۵ نشان‌دهنده تمایز ژنتیکی کم، ۰/۰۵-۰/۱۵ متوسط، ۰/۱۵-۰/۲۵ زیاد و بالاتر از ۰/۲۵ بسیار زیاد می‌باشد (De Vicente *et al.*, 2004; Laurentin, 2009).

اندازه نمونه جمعیت احتمال آشکار شدن یک آل جدید (نادر) افزایش می‌یابد (Laurentin, 2009).

اصطلاح تعداد آل‌های موثر (A_e) که توسط Hartl and Clark (1997) مطرح شد (Laurentin, 2009) عبارت از اندازه‌گیری تعداد آل‌های با فراوانی برابر است و جهت دستیابی به سطح معینی از تنوع ژنی لازم می‌باشد. از آن‌جا که جمعیتی با سه آل با فراوانی مشابه نسبت به جمعیتی که یک آل با فراوانی بالا و دو آل با فراوانی کم، چندشکلی بیش‌تری نشان می‌دهد، لذا A_e معیار بهتری نسبت به A است (Bagherieh Najjar *et al.*, 2014). تعداد آل‌های موثر از طریق $A_e = 1/(1 - h) = 1/\sum p_i^2$ محاسبه می‌شود که p_i فراوانی آل iام در یک مکان و $h = 1 - \sum p_i^2$ میزان هتروزیگوسیتی در یک مکان است (De Vicente *et al.*, 2004). بنابراین دستیابی به بیش‌ترین تعداد آل‌های موثر زمانی رخ می‌دهد که هتروزیگوسیتی بالا باشد و این مسئله مستلزم برابر بودن فراوانی همه آل‌ها است. البته این محاسبه ممکن است تحت تاثیر اندازه نمونه قرار گیرد (De Vicente *et al.*, 2004).

شاخص شانون^۱

شاخص تنوع شانون برای ارزیابی تنوع در سطوح مختلف از ژن‌ها و جمعیت‌ها گرفته تا همه گونه‌ها و اکوسیستم‌ها پیشنهاد شده است. اگرچه برخی از محققین استفاده از این شاخص به عنوان معیار تنوع ژنتیکی را رد می‌کنند (Hennink and Zeven, 1990; Konopiński, 2020). در مطالعات ژنتیک جمعیت، وقتی در برخی از جمعیت‌ها چند آل مشترک غالب باشد و در جمعیت‌های دیگر تغییرات^۲ به‌طور یکنواخت توسط همه آل‌ها ایجاد شود امکان تشخیص سطح تغییرات بین جمعیت‌ها (با تعداد آل یکسان) توسط شاخص شانون فراهم خواهد شد. شاخص شانون در مقایسه با هتروزیگوسیتی، به از دست دادن واریانت‌های نادر (مثلاً به دلیل تنگناهای ژنتیکی)^۳ حساس‌تر است. همچنین این شاخص حامل اطلاعات

4. Wright's F-statistics

5. Interpopulation differentiation for several loci

6. Subdivided populations

7. Fixation index

1. Shannon's index

2. Variation

3. Genetic bottlenecks

استفاده می‌گردد (Kosman and Leonard, 2005). در واقع ژنوتیپ‌هایی با ژن‌های مشابه، فاصله ژنتیکی کم‌تری با یکدیگر دارند (Bhandari et al., 2017). جهت برآورد عدم تشابه ژنتیکی می‌توان از ضرایب مبتنی بر فراوانی آلی^۳ نظیر اقلیدسی^۴، راجر^۵، راجر تغییر یافته^۶، رینولد^۷، نی^۸ ۷۲^۹ (فاصله ژنتیکی استاندارد نی)^۹ و نی و همکاران^{۱۰} ۸۳^{۱۱} استفاده نمود (Mohammadi, 2006).

دسته دیگر ضرایب مورد استفاده جهت برآورد تشابه ژنتیکی یا عدم تشابه افراد، مبتنی بر داده‌های صفر و یک^{۱۱} است که متداول‌ترین آن‌ها شامل ضرایب جاکارد^{۱۲} (J)، دایس^{۱۳} (D) یا نی و لی^{۱۴} (NL) و تطابق ساده^{۱۵} (SM) می‌باشد (Mohammadi, 2006). در انتخاب ضرایب مذکور، Laurentin (2009) به نوع نشانگر (غالب یا هم‌بارز) و همچنین به نوع سیستم زادآوری گیاه (خودگشن یا دگرگشن) توجه کرده است. او ضرایب جاکارد و دایس را در بررسی افراد غیرخویشاوند و ضریب تطابق ساده را در بررسی افراد خویشاوند با استفاده از نشانگرهای غالب پیشنهاد کرده است. همچنین او در بررسی روابط ژنتیکی با استفاده از نشانگرهای هم‌بارز به‌طور عمده ضریب دایس را برای گیاهان خودگشن پیشنهاد می‌کند (Laurentin, 2009). از آن‌جا که در ضریب دایس وزن تعداد باندهای مشترک بین دو فرد، دو برابر ضریب جاکارد در نظر گرفته می‌شود، در صورت هموزیگوت بودن افراد مورد بررسی، برآورد شباهت ژنتیکی آن‌ها با استفاده از این دو ضریب (دایس و جاکارد) یکسان است و در صورت هتروزیگوت بودن آن‌ها، برآورد این دو ضریب با استفاده از نشانگرهای هم‌بارز که قدرت تفکیک افراد هموزیگوت و

برخی منابع G_{ST} را معادل F_{ST} گزارش کرده‌اند (Nei, 1973) که تنوع ژنی بین جمعیت‌ها را اندازه‌گیری می‌کند؛ لذا برای برآورد آن باید تعداد زیادی از مکان‌ها را مورد مطالعه قرار داد (De Vicente et al., 2004). این پارامتر جهت مطالعه ساختار ژنتیکی یک جمعیت تقسیم شده بدون در نظر گرفتن تعداد آلل‌ها در هر مکان یا نیروهای تکاملی کاربردی می‌باشد. تمایز ژنی (G_{ST}) طبق رابطه $G_{ST} = D_{ST}/H_T$ برآورد می‌گردد که در آن G_{ST} تمایز ژنی مرتبط با کل جمعیت، D_{ST} تمایز ژنی (میانگین تنوع ژنی) میان زیرجمعیت‌ها و H_T تنوع ژنی در کل جمعیت است (Laurentin, 2009).

تجزیه و تحلیل واریانس مولکولی

از روش‌های رایج برای محاسبه آماره‌های F می‌توان به تجزیه و تحلیل واریانس مولکولی (AMOVA) اشاره کرد که آزمایش وجود ساختار جمعیت سلسله‌مراتبی را آسان می‌کند (Excoffier et al., 1992). با این حال، این تکنیک مستلزم آن است که ساختار سلسله‌مراتبی جمعیت از قبل شناخته شده باشد. بنابراین، یک تجزیه و تحلیل خوشه‌بندی به عنوان مثال، Structure (Pritchard et al., 2000) ممکن است بر روی داده‌ها انجام شود، که نتایج آن سپس به عنوان مبنایی برای AMOVA استفاده می‌شود (Meirmans, 2012). همچنین AMOVA می‌تواند اجزای واریانس بین و درون گروه‌ها را تخمین زند و سهم هر یک از تنوع درون و بین جمعیت‌ها را نسبت به تنوع کل مشخص نماید (Excoffier et al., 1992).

اندازه‌گیری تشابه^۱ و عدم تشابه^۲ ژنتیکی (فاصله ژنتیکی) و ضرایب مورد استفاده

در ارزیابی روابط ژنتیکی میان افراد می‌توان از ماتریس عدم تشابه ژنتیکی استفاده کرد که میزان تفاوت میان جفت افراد، محاسبه و از آن برای توصیف ساختار جمعیت بر اساس قرابت نسبی هر فرد با سایر افراد مورد مطالعه،

3. Genetic dissimilarity coefficients for allelic informative marker data
4. Euclidean
5. Rogers
6. Modified Rogers
7. Reynolds et al., (1983)
8. Nei, (1972)
9. Nei's standard genetic distance
10. Nei et al., (1983)
11. Similarity coefficients for allelic noninformative marker data
12. Jaccard (J)
13. Dice (D)
14. Nei and Li (NL)
15. Simple Matching (SM)

1. Similarity
2. Dissimilarity (distance) matrix

هدف آن ایجاد نمودار گرافیکی بر اساس فاصله‌ی میان افراد است؛ به گونه‌ای که فاصله بین نقاط در نمودار، نزدیک به تفاوت‌های اصلی باشد (Mohammadi and Prasanna, 2003).

تجزیه خوشه‌ای و تجزیه به مختصات اصلی با استفاده از برخی نرم‌افزارها مانند NTSYSpc، DARwin، PAST و R بر اساس ضرایب و الگوریتم‌های متفاوتی ترسیم می‌گردند.

انواع الگوریتم‌های گروه‌بندی

در گروه‌بندی سلسله‌مراتبی بر اساس روش اندازه‌گیری فاصله بین افراد، از الگوریتم‌هایی نظیر نزدیک‌ترین همسایگی^۶، دورترین همسایگی^۷، جفت گروه بدون وزن با میانگین حسابی^۸ (UPGMA) و نیز روش Ward استفاده می‌گردد. مبنای اندازه‌گیری در این چهار روش به ترتیب کم‌ترین، بیش‌ترین، میانگین و مجموع مربعات فاصله جفت افراد می‌باشد (De Vicente et al., 2004; Maulik et al., 2011). UPGMA و Ward از متداول‌ترین الگوریتم‌های مورد استفاده هستند (Mohammadi and Prasanna, 2003) که اهمیت آن‌ها را می‌توان به سهولت دسترسی به آن‌ها در برنامه‌های آماری، درک آسان‌تر نتایج حاصل از آن‌ها و کوتاه بودن زمان مورد نیاز برای تجزیه و تحلیل نسبت داد (Ebrahimi & Zienalabedini, 2013). در روش UPGMA وجود اثر زنجیره‌ای^۹، تفسیر داده‌ها را مشکل می‌کند که از معایب اصلی آن شمرده شده است در صورتی که در روش Ward اثر زنجیره‌ای به‌ندرت دیده می‌شود (Mohammadi and Prasanna, 2003). روش دیگری به نام اتصال همسایگی^۱ (NJ) نیز وجود دارد که ابتدا با یک درخت با توپولوژی ستاره مانند شروع و سپس با انتخاب متوالی جفت تاکسون‌ها بر اساس فاصله آن‌ها ادامه یافته تا درخت نهایی ترسیم گردد (Yang and Rannala, 2012). در برخی منابع توصیه به استفاده از روش NJ و Ward شده است (Ebrahimi & Zienalabedini, 2013).

هتروزیگوت را دارند، متفاوت خواهد بود (Link et al., 1995; Mohammadi, 2006). در ضریب تطابق ساده، عدم وجود باند برای هر دو فرد مورد مقایسه به عنوان تشابه در نظر گرفته می‌شود، در حالی که ضرایب جاکارد و دایس آن را به عنوان معیاری برای تشابه در نظر نمی‌گیرند (Kosman and Leonard, 2005).

تجزیه خوشه‌ای و تجزیه به مختصات اصلی

روش‌های آماری چند متغیره راه مهمی برای نمایش روابط ژنتیکی و گروه‌بندی افراد می‌باشد که رایج‌ترین آن‌ها شامل تجزیه خوشه‌ای^۱ و تجزیه به مختصات اصلی^۲ می‌باشد (Mohammadi and Prasanna, 2003). در تجزیه خوشه‌ای، افراد بر اساس اطلاعات توصیف‌کننده آن‌ها و روابطشان گروه‌بندی می‌گردند. در واقع افراد درون یک گروه، مشابه یا مرتبط با یکدیگر بوده و از افراد گروه‌های دیگر متمایز می‌شوند. هر چه شباهت در یک گروه و تفاوت بین گروه‌ها بیش‌تر باشد، خوشه‌بندی بهتر یا متمایزتر است (Tan et al., 2006). یکی از روش‌های خوشه‌بندی، سلسله‌مراتبی^۳ است که با ساختار درخت‌مانند مشخص می‌شود. این روش دارای دو رویکرد بالا به پایین (تجزیه‌ای)^۴ و پایین به بالا (تجمعی)^۵ می‌باشد. در حالت تجمعی هر فرد در یک خوشه مجزا قرار گرفته و سپس خوشه‌ها با توجه به شباهت‌شان با یکدیگر ادغام می‌شوند. ابتدا، دو خوشه مشابه (آن‌هایی که کم‌ترین فاصله را دارند) ادغام شده تا یک خوشه جدید در پایین سلسله‌مراتب تشکیل شود؛ سپس یک جفت دیگر از خوشه‌ها ادغام شده و به سطح بالاتری از سلسله‌مراتب پیوند داده می‌شود و به همین ترتیب سلسله‌مراتبی از خوشه‌ها از پایین به بالا ایجاد می‌گردد. در صورتی که در رویکرد بالا به پایین ابتدا همه افراد در یک خوشه قرار گرفته، سپس به‌تدریج تقسیم می‌شوند (Sarstedt and Mooi, 2011). تجزیه به مختصات اصلی (PCoA) روشی است که با ماتریسی از شباهت‌ها یا عدم تشابه بین مجموعه‌ای از افراد شروع شده و

6. Single linkage
7. Complete linkage
8. Unweighted Pair Group Method with Arithmetic Mean (UPGMA)
9. Chaining effect

1. Cluster analysis
2. Principal Coordinate Analysis (PCoA)
3. Hierarchical clustering
4. Divisive clustering
5. Agglomerative clustering

نمونه‌ها یا زمان و بودجه محدود است مناسب می‌باشد (De Vicente *et al.*, 2004).

برنامه‌های نرم‌افزاری برای بررسی تنوع ژنتیکی
تجزیه و تحلیل بر پایه ابزار آماری سبب تفسیر بهتر داده‌ها می‌گردد. انواع مختلفی از نرم‌افزارها (Aremu, 2012; Tanavar *et al.*, 2014; Pagnotta, 2018) برای تجزیه و تحلیل تنوع ژنتیکی وجود دارد که اکثر آن‌ها بر اساس تحلیل آماری چند متغیره هستند. اغلب نرم‌افزارها به صورت رایگان در دسترس و برای رایانه‌های شخصی مناسب هستند (Bhandari *et al.*, 2017). در این مطالعه به برخی از پرکاربردترین نرم‌افزارها در جدول ۳ اشاره شده است.

یکی از رایج‌ترین روش‌های سنجش کارایی هر یک از الگوریتم‌های مورد استفاده در تجزیه خوشه‌ای، ضریب همبستگی کوئتیک (r) است که به منظور تطابق شباهت نشان داده شده توسط درخت و ماتریس تشابه به کار می‌رود. مقادیر مساوی ۰/۹ و یا بیش‌تر، کارایی الگوریتم مورد استفاده را نشان می‌دهد اما لازم به ذکر است که پایین بودن مقدار ضریب همبستگی مبنی بر عدم کارایی درخت نبوده و ممکن است مرتبط به داده‌های مولکولی گمشده باشد (Mohammadi, 2006; Darvish Zadeh, 2015 and Azizi). همچنین روش دیگری به نام نمونه برداری مجدد با جایگزینی با همان ماتریس داده (بوتسترپ) وجود دارد که امکان محاسبه انحراف معیارها و واریانس‌ها را فراهم می‌کند و برای شرایطی که تعداد

جدول ۳. برخی از رایج‌ترین برنامه‌های نرم‌افزاری برای تجزیه و تحلیل تنوع ژنتیکی (Govindaraj *et al.*, 2015; Mukhopadhyay and Bhattacharjee, 2016; Bhandari *et al.*, 2017)

Table 3. Some of the most common software programs for analyzing genetic diversity (Govindaraj *et al.*, 2015; Mukhopadhyay and Bhattacharjee, 2016; Bhandari *et al.*, 2017)

نرم‌افزار	برخی از ویژگی‌های اصلی	منبع	آدرس سایت
NTSYSpc (Numerical Taxonomy System for personal computer)	مبتنی بر ضرایب تشابه و عدم تشابه، تجزیه خوشه‌ای، تجزیه به مختصات اصلی و تجزیه به مولفه‌های اصلی	Rohlf, (2002)	https://www.appliedbiostat.com/ntsyspc/ntsyspc.html
DARwin (Dissimilarity Analysis and Representation for windows)	تخمین فاصله و عدم تشابه، تجزیه خوشه‌ای و تجزیه به مختصات اصلی	Perrier and Jacquemoud-Collet, (2006)	https://darwin.cirad.fr/product.php
PAST (PAleontological STatistics software for analysis and education)	پارسیمونی، تجزیه به مولفه‌های اصلی و تجزیه به مختصات اصلی	Hammer <i>et al.</i> , (2001)	https://www.nhm.uio.no/english/research/resources/past/
PowerMarker (A Powerful software for Marker data analysis)	مناسب داده‌های SSR در بررسی ژنتیک جمعیت، تعداد آلل‌ها، تنوع ژنی، ضریب خودگشنی، ساختار جمعیت، آزمون منتل، هتروزیگوسیتی، محتوای اطلاعات چندشکلی، فراوانی آللی، تعادل هاردی-واینبرگ	Liu and Muse, (2005)	https://brcwebportal.cos.ncsu.edu/powermarker/
POPGENE (Microsoft Window-based Freeware for Population Genetic Analysis)	مناسب برای نشانگرهای غالب و هم‌باز، فراوانی ژنی، تعداد آلل‌ها، تعداد آلل‌های موثر، مکان‌های چندشکل، تنوع ژنی، شاخص شانون، آماره‌های F، جریان ژنی، فاصله ژنتیکی (مبتنی بر ضریب نی) و تجزیه خوشه‌ای (بر اساس Neighbor Joining و UPGMA)	Yeh <i>et al.</i> , (1999)	https://sites.ualberta.ca/~fyeh/popgene.html
GenAlEx (Genetic Analysis in Excel)	مناسب برای نشانگرهای غالب و هم‌باز، هتروزیگوسیتی مشاهده شده و مورد انتظار، شاخص نشانگری، شاخص تنبیت، Fst، فاصله ژنتیکی نی، شاخص شانون و تجزیه به مختصات اصلی	Peakall and Smouse, (2006)	http://biology-assets.anu.edu.au/GenAlEx/Welcome.html
Arlequin	تخمین فراوانی آللی، تجزیه و تحلیل واریانس مولکولی	Excoffier <i>et al.</i> , (2010)	https://cmpg.unibe.ch/software/arlequin35
FSTAT	تنوع ژنی، بوتسترپ	Goudet, (2002)	https://www2.unil.ch/popgen/softwares/fstat.htm

1. Neighbor-joining (NJ)
2. Bootstrapping

نتیجه‌گیری

تنوع ژنتیکی برای سازگاری گیاهان با تغییرات محیطی لازم و ضروری است. واکنش گیاهان و توانایی‌های سازگاری آن‌ها به تغییرات آب و هوایی به سطح تنوع ژنتیکی آن‌ها بستگی دارد. نشانگرهای مولکولی به‌طور گسترده برای مطالعات تنوع ژنتیکی منابع گیاهی و ژنتیک جمعیت کاربرد دارند. استفاده از این نشانگرها به منظور اصلاح و بهبود محصولات، حفاظت و معرفی گیاهان در معرض خطر و با ارزش ضروری است. همچنین شناسایی واریته‌های گیاهی جدید و تشخیص تغییرات ژنتیکی انواع گیاهان را امکان‌پذیر می‌سازند و بینش‌های ارزشمندی را در مورد تغییرات ژنتیکی موجود در درون و بین جمعیت‌های گیاهی ارائه می‌دهند. لذا شناسایی و برآورد مناسب تنوع ژنتیکی برای اهداف مذکور حائز اهمیت است. بنابراین ابزارهای مولکولی و روش‌های آماری می‌توانند دریچه جدیدی را برای درک بهتر چندشکلی و روابط ژنتیکی بین افراد باز کنند و به محققین اجازه دهند تا به سؤالات مهم در این زمینه‌ها پاسخ دهند.

لذا در این مطالعه مروری بر مفاهیم پایه مرتبط به کمی کردن تنوعات درون جمعیت‌های گیاهی (تنوع آلی، تعداد آل‌های موثر و هتروزیگوسیتی/تنوع ژنی)، تنوعات بین جمعیت‌های گیاهی (و یا افراد) (آماره F رایت، G_{st} و تجزیه و تحلیل واریانس مولکولی)، روابط ژنتیکی بین آنها (ضرایب مبتنی بر تشابه و فاصله ژنتیکی) و نیز نمایش این روابط (خوشه‌بندی و تجزیه به مختصات اصلی) بر اساس داده‌های مولکولی مبتنی بر امتیازدهی ژل صورت گرفته است. همچنین به برخی از رایج‌ترین برنامه‌های نرم‌افزاری در این زمینه اشاره می‌گردد.

تشکر و قدردانی

از حمایت مالی معاونت پژوهش و فناوری دانشگاه شهید چمران اهواز در قالب پژوهانه به شماره SCU.AA1402.165 در انجام این تحقیق، تشکر و قدردانی می‌گردد.

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

References

- Agarwal, M., Shrivastava, N., & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports*, 27, 617-631.
- Akkaya, M. S., Bhagwat, A. A., & Cregan, P. B. (1992). Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*, 132 (4), 1131-1139.
- Alizadeh, M., Nematzadeh, G. A., Ebrahimi, M.A., & Hashemi, H. (2013). Fingerprinting of some rice (*Oryza sativa* L.) germplasm via AFLP markers. *Crop Biotechnology*, 3 (4), 53-60. (In Persian)
- Akopyanz, N., Bukanov, N. O., Westblom, T. U., & Berg, D. E. (1992). PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. *Nucleic Acids Research*, 20 (23), 6221-6225.
- Amiteye, S. (2021). Basic concepts and methodologies of DNA marker systems in plant molecular breeding. *Heliyon*, 7 (10), 1-20.
- Aremu, C. O. (2012). Exploring statistical tools in measuring genetic diversity for crop improvement. *Genetic Diversity in Plants*, 340-348.
- Bagherieh Najjar, M. B., Kohan Baghkheirati, E., & Abasi, Z. (2014). *DNA fingerprinting in plants principles, methods and applications*. Gorgan, Golestan University. (In Persian)
- Bhandari, H. R., Bhanu, A. N., Srivastava, K., Singh, M. N., & Shreya, H. A. (2017). Assessment of genetic diversity in crop plants - an overview. *Advances in Plants & Agriculture Research*, 7 (3), 279-286.
- Bidyananda, N., Jamir, I., Nowakowska, K., Varte, V., Vendrame, W. A., Devi, R. S., & Nongdam, P. (2024). Plant genetic diversity studies, insights from DNA marker analyses. *International Journal of Plant Biology*, 15 (3), 607-640.
- Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32 (3), 314.
- Caliskan, M. (Ed.). (2012). *Genetic diversity in plants*. BoD-Books on Demand.

- Chesnokov, Y. V., & Artemyeva, A. M. (2015). Evaluation of the measure of polymorphism information of genetic diversity. *Agricultural Biology*, 50 (5), 571-578.
- Collard, B. C., & Mackill, D. J. (2009). Start codon targeted (SCoT) polymorphism, a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology Reporter*, 27, 86-93.
- Dar, A. A., Mahajan, R., & Sharma, S. (2019). Molecular markers for characterization and conservation of plant genetic resources. *The Indian Journal of Agricultural Sciences*, 89 (11), 1764-1772.
- Darvish Zadeh, R., & Azizi, H. (2015). *Molecular markers and their application in the analysis of genetic diversity*. Urmia, University of Urmia. (In Persian)
- De Vicente, M. C., Lopez, C., & Fulton, T. (2004). Genetic diversity analysis with molecular marker data, learning module. International Plant Genetic Resources Institute (IPGRI).
- Ebrahimi, M. A., & Zienalabedini, M. (2013). Application of genomic and unigene-based microsatellite markers in conservation and management of genetic resources of some Iranian crops. *Crop Biotechnology*, 4, 133-148.
- Esmaili, A., Talebi, B., Drikvand, R., Ebrahimi, M. A., & Hossienpour, T. (2014). Evaluation of genetic diversity among hull-less and hulled rain-fed barley genotypes using SSR molecular marker. *Crop Biotechnology*, 3 (5), 103-111. (In Persian)
- Excoffier, L., Smouse, P. E., & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes, application to human mitochondrial DNA restriction data. *Genetics*, 131 (2), 479-491.
- Excoffier, L., Lischer, H. E. L. (2010) Arlequin suite ver 3.5, a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567.
- Goudet, J. (2002). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3.2). <http://www.unil.ch/izea/software/fstat.html>.
- Govindaraj, M., Vetriventhan, M., & Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances, an overview of its analytical perspectives. *Genetics Research International*, 2015, 1-14.
- Hamdan, Y. A., Hawamda, A. I., Basheer-Salimia, R., & Salman, M. (2024). Genetic diversity assessment of Palestinian okra landraces (*Abelmoschus esculentus* L.) through RAPD marker. *Genetic Resources and Crop Evolution*, 1-8.
- Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST, paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, 4 (1), 1-9.
- Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics*. Sunderland, Massachusetts, Sinauer Associates.
- Heidari, B. (2011). *Plant molecular breeding*. Shiraz, University of Shiraz. (In Persian)
- Hennink, S., & Zeven, A. C. (1990). The interpretation of Nei and Shannon-Weaver within population variation indices. *Euphytica*, 51, 235-240.
- Heras, J., Domínguez, C., Mata, E., Pascual, V., Lozano, C., Torres, C., & Zarazaga, M. (2016). A survey of tools for analysing DNA fingerprints. *Briefings in Bioinformatics*, 17 (6), 903-911.
- Hosseini, F., Niknejad, A., Sorkhilaleloo, B., Kohpalekani, J. A., & Khameneh, M. M. (2023). Valuation of genetic diversity of tomato (*Solanum lycopersicum* L.) based on the comparison of ISSR and SSR marker efficiency. *Agricultural Biotechnology Journal*, 15 (2), 63-82. (In Persian)
- Jaccoud, D., Peng, K., Feinstein, D., & Kilian, A. (2001). Diversity arrays, a solid state technology for sequence information independent genotyping. *Nucleic Acids Research*, 29 (4), e25.
- Jahangir, M. (2021). Genetic diversity and population structure of *Cressa cretica* L. in Ahvaz using morphological and molecular markers (SCoT and ISSR). The thesis was submitted in partial fulfillment of the requirements for the Master of Science degree (M.Sc.). Shahid Chamran University of Ahvaz. (In Persian)
- Kadirvel, P., Senthilvel, S., Geethanjali, S., Sujatha, M., & Varaprasad, K. S. (2015). Genetic markers, trait mapping and marker-assisted selection in plant breeding. *Plant Biology and Biotechnology, Volume II, Plant Genomics and Biotechnology*, 65-88.
- Kalendar, R., Grob, T., Regina, M., Suoniemi, A., & Schulman, A. (1999). IRAP and REMAP, two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics*, 98, 704-711.
- Kalendar, R., & Schulman, A. H. (2014). Transposon-based tagging, IRAP, REMAP, and iPBS. Molecular plant taxonomy, methods and protocols, 233-255.

- Kanouni, H., & Khalily, M. (2010). *Advances in crop breeding, laboratory protocols of molecular markers*. Urmia, Academic Jihad. (In Persian)
- Konopiński, M. K. (2020). Shannon diversity index, a call to replace the original Shannon's formula with unbiased estimator in the population genetics studies. *PeerJ*, 8, e9391.
- Kosman, E., & Leonard, K. J. (2005). Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology*, 14 (2), 415-424.
- Kumar, J., & Agrawal, V. (2019). Assessment of genetic diversity, population structure and sex identification in dioecious crop, *Trichosanthes dioica* employing ISSR, SCoT and SRAP markers. *Heliyon*, 5 (3), 1-26.
- Kumar, P., Gupta, V. K., Misra, A. K., Modi, D. R., & Pandey, B. K. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics*, 2 (4), 141-162.
- Laurentin, H. (2009). Data analysis for molecular characterization of plant genetic resources. *Genetic Resources and Crop Evolution*, 56, 277-292.
- Link, W., Dixkens, C., Singh, M., Schwall, M., & Melchinger, A. E. (1995). Genetic diversity in European and Mediterranean faba bean germ plasm revealed by RAPD markers. *Theoretical and Applied Genetics*, 90, 27-32.
- Liu, K., & Muse, S. V. (2005). PowerMarker, an integrated analysis environment for genetic marker analysis. *Bioinformatics*, 21 (9), 2128-2129.
- Marzang, N., Abdollahi Mandoulakani, B., Shaaf, S., Ghadmizadeh, M., Bernousi, I., Abbasi Holasou, H., & Sadeghzadeh, B. (2020). IRAP and REMAP -based genetic diversity among Iranian, Turkish, and international durum wheat (*Triticum turgidum* L.) cultivars. *Journal of Agricultural Science and Technology*, 22 (1), 271-285.
- Maulik, U., Bandyopadhyay, S., & Mukhopadhyay, A. (2011). *Multiobjective genetic algorithms for clustering, applications in data mining and bioinformatics*. Germany, Springer Berlin Heidelberg.
- Medraoui, L., Rabeh, K., Ater, M., & Filali-Maltouf, A. (2024). Genetic diversity analysis of sorghum (*Sorghum bicolor* L. Moench) landraces from northwestern Morocco using ISSR and AFLP markers. *Genetic Resources and Crop Evolution*, 71 (2), 835-850.
- Meirmans, P.G. (2012). AMOVA-based clustering of population genetic data. *Journal of Heredity*, 103 (5), 744-750.
- Mohammadi, A. (2006). *Analysis of molecular data from the point of view of genetic diversity*. The 9th Congress of Agricultural Sciences and Plant Breeding of Iran, 96-119. (In Persian)
- Mohammadi, S. A., & Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop Science*, 43(4), 1235-1248.
- Mooi, E., & Sarstedt, M. (2011). *A concise guide to market research*. Springer-Verlag Berlin Heidelberg.
- Mukhopadhyay, T., & Bhattacharjee, S. (2016). Genetic diversity, importance and measurements. conserving biological diversity, a multiscaled approach. *Research India Publications, New Delhi*, 251-295.
- Nagaraju, J., Reddy, K. D., Nagaraja, G. M., & Sethuraman, B. N. (2001). Comparison of multilocus RFLPs and PCR-based marker systems for genetic analysis of the silkworm, *Bombyx mori*. *Heredity*, 86 (5), 588-597.
- Naghavi, R., Ghariazi, B., & Hosseini Salekdeh, G. H. (2009). *Molecular markers*. Tehran, University of Tehran. (In Persian)
- Nazarbeigi, M., Roein, Z., & Sabouri, A. (2024). Investigating the genetic diversity of native Iranian *Gladiolus* with ISSR and IRAP molecular markers. *Journal of Plant Production Research*, 31 (1), 89-110. (In Persian)
- Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, 106 (949), 283-292.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, 70 (12), 3321-3323.
- Ng, W. L., & Tan, S. G. (2015). Inter-simple sequence repeat (ISSR) markers, are we doing it right. *ASM Science Journal*, 9 (1), 30-39.
- Nikkerdar, F., Farshadfar, M., Ebrahimi, M. A., & Shirvani, H. (2018). Genetic diversity among fennel (*Foeniculum Vulgare* Mill.) landrace using SCoT Markers. *Journal of Crop Breeding*, 9 (24), 95-102. (In Persian)
- Nonić, M., & Šijačić-Nikolić, M. (2019). Genetic diversity, sources, threats, and conservation. *Life on Land*, 1-15.
- Orita, M., Suzuki, Y., Sekiya, T., & Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics*, 5 (4), 874-879.
- Ott, J. (1992). Strategies for characterizing highly polymorphic markers in human gene mapping. *American Journal of Human Genetics*, 51 (2), 283-290.

- Pagnotta, M. A. (2018). Comparison among methods and statistical software packages to analyze germplasm genetic diversity by means of codominant markers. *Multidisciplinary Science Journal*, 1 (1), 197-215.
- Paran, I., & Michelmore, R. W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics*, 85, 985-993.
- Peakall, R., & Smouse, P.E., (2006). GenAlEx 6, genetic analysis in Excel. population genetic software for teaching and research. *Molecular Ecology*, 6 (1), 288-295.
- Perrier, X., & Jacquemoud-Collet, J. P. (2006). DARwin software, <http://darwin.cirad.fr/darwin>.
- Pozharskiy, A., Kostyukova, V., Khusnitdinova, M., Adilbayeva, K., Nizamdinova, G., Kapytina, A., Kerimbek, N., Taskuzhina, A., Kolchenko, M., Abdrakhmanova, A., Kisselyova, N., Kalendar, R., & Gritsenko, D. (2023). Genetic diversity of the breeding collection of tomato varieties in Kazakhstan assessed using SSR, SCAR and CAPS markers. *PeerJ*, 11, 1-19.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155 (2), 945-959.
- Rani, V., Yadav, M. K., Singh, R., & Yadav, D. (2024). Genetic diversity assessment in cereal crops. in sustainable utilization and conservation of plant genetic diversity (pp. 363-398). Singapore, Springer Nature Singapore.
- Rao, N. K. (2004). Plant genetic resources, advancing conservation and use through biotechnology. *African Journal of Biotechnology*, 3 (2), 136-145.
- Ray, T., & Roy, S. C. (2009). Genetic diversity of *Amaranthus* species from the Indo-Gangetic plains revealed by RAPD analysis leading to the development of ecotype-specific SCAR marker. *Journal of Heredity*, 100 (3), 338-347.
- Rohlf, F. J. (2002). NTSYSpc, numerical taxonomy system, version 2.1, Exeter Publishing, Setauket, NY, USA.
- Salgotra, R. K., & Chauhan, B. S. (2023). Genetic diversity, conservation, and utilization of plant genetic resources. *Genes*, 14 (1), 174.
- Sarwat, M. (2012). ISSR, a reliable and cost-effective technique for detection of DNA polymorphism. *Plant DNA Fingerprinting and Barcoding, Methods and Protocols*, 103-121.
- Serrote, C. M. L., Reiniger, L. R. S., Silva, K. B., dos Santos Rabaiolli, S. M., & Stefanel, C. M. (2020). Determining the polymorphism information content of a molecular marker. *Gene*, 726, 1-15.
- Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication, by CE Shannon (and recent contributions to the mathematical theory of communication), W. Weaver. University of Illinois Press.
- Shashidhara, N., Dhanalakshmi, T. N., & Santosh, D. T. (2023). *Plant genetic resources in sustainable agriculture, current state, future perspectives, and challenges*. New Delhi.
- Sharifi Soltani, S., Ranjbar, G. A., Kazemitabar, S. K., Pakdin-Parizi, A., & Najafi Zarrini, H. (2023). Investigation of genetic diversity in castor (*Ricinus communis*) ecotypes using ISSR markers. *Journal of Plant Research (Iranian Journal of Biology)*, 36 (3), 294-308. (In Persian)
- Sherwin, W. B., Chao, A., Jost, L., & Smouse, P. E. (2017). Information theory broadens the spectrum of molecular ecology and evolution. *Trends in Ecology & Evolution*, 32 (12), 948-963.
- Shiva, V. (1994). Agriculture and food production. *Environmental Education Dossiers*, 9, 2-3.
- Simpson, J. (1997). Molecular markers. *Boletín de la Sociedad Botánica de México*, 60, 73-76.
- Swingland, I. R. (2001). Biodiversity, definition of. *Encyclopedia of Biodiversity*, 1, 377-391.
- Tan, P. N., Steinbach, M., & Kumar, V. (2006). Cluster analysis, basic concepts and algorithms. *Introduction to Data Mining*. (Boston), 525-612.
- Tanavar, M., Kelestanie, A. R. A., & Hoseni, S. A. (2014). Software programs for analyzing genetic diversity. *International Journal of Farming and Allied Sciences*, 3 (5), 462-466.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Homes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. (1995). AFLP, a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23 (21), 4407-4414.
- Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalsky, J. A., & Tingey, S. V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18, 6531-6535.
- Wu, K. S., Jones, R., Danneberger, L., & Scolnik, P. A. (1994). Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research*, 22 (15), 3257-3258.
- Yang, Z., & Rannala, B. (2012). Molecular phylogenetics, principles and practice. *Nature Reviews Genetics*, 13 (5), 303-314.
- Yeh, F. C., Yang, R. C., & Boyle, T. (1999). POPGENE. microsoft windows based freeware for population genetic analysis. release 1.31. University of Alberta, Edmonton. (<http://www.ualberta.ca/s-fyeh/fyeh>).

- Yongcui, H., Zehong, Y., & Xiujin, L. (2005). Genetic diversity among barley germplasm with known origins based on the RAMP and ISSR markers. *Scientia Agricultura Sinica*, 38 (12), 2555-2565.
- Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, 20 (2), 176-183.