



Integrating SCoT, CBDP, and ISSR molecular markers for genetic diversity assessment and taxonomic authentication of some Asteraceae species in Egypt

Mariam E. M. A. Elgohary · Ehab A. A. Salama · Hossam M. E. El-wakil ·
Nader R. Abdelsalam · Mohamed A. M. Atia

Received: 12 February 2025 / Accepted: 28 May 2025 / Published online: 19 June 2025
© The Author(s) 2025

Abstract Despite the few numbers of molecular investigations exploring the genetic diversity, very scarce information is available for supporting the taxonomic and anatomical investigations of Asteraceae species in Egypt. Thus, the adoption of modern DNA based marker techniques is an urgent need to dive into the taxonomic authentication, evolutionary relationship in addition to the calculation of genetic variations within and between Asteraceae species. Hence, the current study was conducted to assess the intraspecific genetic diversity and authenticate taxa at the species level among and within 20 different species of Asteraceae family. Using 19 primers combinations belonging to three different molecular marker systems namely, Start Codon Targeted (SCoT), CAAT box-derived Polymorphism (CBDP) and Inter Simple Sequence Repeats (ISSR). A total of 285

amplified DNA fragments were produced, including 273 polymorphic DNA bands, with an average of percentage of polymorphism (PPB) ranged from 91.10 to 100.00% and the average of polymorphism information content (PIC) for SCoT, CBDP, and ISSR markers were 0.349, 0.403, and 0.418 respectively. The population structure analysis suggested that the 20 Asteraceae species could be categorized into 2 or 7 or 8 genetic groups based on SCoT, CBDP, and ISSR markers datasets, respectively. Both factorial and phylogenetic analyses based on the presence or absence of the amplified bands/fragments were performed by Jaccard's similarity coefficient. Interestingly, Unweighted pair group method with arithmetic mean (UPGMA) cluster dendrograms for all markers system) two main clusters that all subsequently divided into different subgroups of species. Surprisingly, the UPGMA dendrogram generated based on CBDP marker data was divided into main clusters only one species of *Centaurea cristata* L. that collected from AGERI was clustered as a unique and main cluster separately while the remaining 19 species of Asteraceae family spreading over the first main cluster. Together, this study showed that the combination of these three molecular marker systems rather than independently could be useful tool for comprehensive evaluation of genetic diversity among and within Asteraceae species in Egypt.

Mariam E. M. A. Elgohary and Ehab A. A. Salama have contributed equally to this work.

M. E. M. A. Elgohary · E. A. A. Salama (✉) ·
H. M. E. El-wakil · N. R. Abdelsalam
Agricultural Botany Department (Genetics), Faculty
of Agriculture Saba Basha, Alexandria University,
Alexandria 21531, Egypt
e-mail: ehabsalama89@alexu.edu.eg

M. A. M. Atia (✉)
Genome Mapping Department, Agricultural Genetic
Engineering Research Institute (AGERI), Agricultural
Research Center (ARC), Giza 12619, Egypt
e-mail: matia@matia.org

Keywords Genetic variability · UPGMA · Evolutionary relationships · Asteraceae · DNA markers

Introduction

Globally, Asteraceae (Compositae, alternate name) which known as the sunflower family is categorized as the largest and most remarkably diverse family of angiosperms and flowering plants due to its wide distribution as well as growth suitability under various types of habitats mainly in tropical and subtropical regions all over the world (Rolnik and Olas 2021; Pantoja and Oliveira 2024). Approximately, it comprises about 1600–1700 genera and more than 25,000–33,000 worldwide species belonging to 13 subfamilies and 45 tribes which give it the wide adaptability across all continents except Antarctica (Marzouk et al. 2021; Shen et al. 2023). Asteraceae family members have a unique characteristics besides morphological features such as that it easy to grow and propagate, have small size individuals predominantly as well as have the capability to produce a variety of indigenous bioactive secondary metabolite compounds for empowering their survival and adaptation strategy (Liana 2020). All these previous advantageous factors of the Asteraceae family, let its species and bioproducts are frequently utilized as an important ornamental, food, cosmetics, and other daily life diverse purposes such as the therapeutic and medical applications to treat different diseases which eventually made it contributed to human health for more than 3000 years (Marković et al. 2022).

Despite the fact that Egypt country has a unique biodiversity pattern of containing many different plants, microorganisms and animals' species makes the country an active member of the Convention on Biological Diversity (CBD) since 1993 according to the Egyptian Biodiversity Strategy and Action Plan 2015–2030 (Temraz et al. 2016). At the same time, there is a few numbers of molecular marker investigations that explore the genetic diversity as well as support the taxonomic and anatomical investigations of Asteraceae species in Egypt. On the other hand, several works had done to determine the floristic composition of Asteraceae species in different geological protected areas across the country such as Saint Katherine Protectorate and other protected areas of Egypt

(Shaltout et al. 2004; Temraz et al. 2016; Khafagip et al. 2018; Adel et al. 2020). Some investigations were done using the morphological characters as a key data to confirm the classification and recognition of the Asteraceae species in Egypt (Adel et al. 2020; Elgazzar et al. 2020). Recently, a first comparative taxonomic study was conducted to revise the classification and identification as well as to estimate the inter and intraspecific variations among 156 specimens of Launaea family species collected from 63 sites in Egypt using macro- and micromorphological characters (Marzouk et al. 2021). It is important to state that, the reliability of these macro- and micromorphological characters for evaluating genetic diversity of germplasm, revising their classification and relationships is so low, not only because it influenced by many environment factors, and different plant development stages but also its very time consuming and tedious processes (Bidyananda et al. 2024). To avoid the challenges facing researchers to identify and classify the Asteraceae species in Egypt based on the floristic composition and other morphological features. To address these challenges, the application of different molecular marker systems acts as a best suitable and efficient tool for germplasm identification and characterization in addition to the molecular diversity analysis compared to the evaluation based on morphological characteristics (Özbek 2024).

In general, Egypt flora contains 98 genera and 227 species with 15 subspecies and 15 varieties of Asteraceae family in which all are widespread growing annually or perennially across various phytogeographical regions of the country as well as holding an important economic potentialities especially their utilization for so-called the folk medicine in Egypt (Bolous 2002; Amar et al. 2012). Thus, there is an urgent need to increase and expand the utilization of different molecular markers techniques to study the genetic diversity and calculate the genetic variation among these Asteraceae species in Egypt due to their nature intricate genetic background. Importantly, the advancement of genotyping techniques using both arbitrary and gene-targeted molecular marker systems allowed the extensive use and rapid development of molecular biology research to identify the germplasm and to assess the genetic diversity in different crops. In addition, these recent molecular marker techniques not only have the ability to estimate the diversity and evolutionary relationships within plant species

but also can be used for identifying and determining closely related genotypes within the collected species (Khodae et al. 2021).

Considering the molecular diversity research using different DNA based markers techniques in Asteraceae family away from using arbitrary molecular markers such Random Amplified Polymorphic DNA (RAPD), DNA amplification fingerprinting (DAF), and Arbitrarily primed polymerase chain reaction (AP-PCR) markers, there are other novel PCR based molecular markers have been extensively used to evaluate the genetic variation within Asteraceae species and individuals. For example, Conserved DNA-Derived Polymorphism (CDDP), Start Codon Targeted Polymorphism (SCoT), CAAT box -derived Polymorphism (CBDP), and Directed Amplification of Minisatellite-region DNA (DAMD) markers. Notably, these later gene-related target marker techniques were developed based on a short conserved DNA sequence regions of genome across different plant species (Hajibarat et al. 2015; Hamdan 2024) that make them all have a greater advantages over random molecular markers for precise conducting of the DNA fingerprinting assays as well as supporting the species identification and classification in Asteraceae family. Additionally, Inter Simple Sequence Repeats (ISSRs) are also another most efficient marker type that has the capacity to reveal highly reproducible and polymorphic bands from a single PCR amplification assay during the genetic diversity analysis from various plant breeding investigations (Pradeep Reddy et al. 2002). Interestingly, all these markers shared a common feature as it displays a high level of polymorphism and marker efficiency beside their higher reproducibility rank and generation of high polymorphic information content (PIC) compared with RAPD, DAF, and AP-PCR random markers. Taking together these advantages, several investigations have been reported the usefulness of these molecular markers together of individually to estimate the genetic diversity, phylogenetic relationships, and DNA fingerprinting in a wide range of Asteraceae species and other important crops species for instance; Chrysanthemum (Samarina et al. 2021), Milk thistle (Hamouda 2019), *Centaurea* species (Atia et al. 2021), Wheat (Heidari et al. 2017; Khodae et al. 2021; Shaban et al. 2022), Lesser burnet (Karataş et al. 2024), Okra (Puneeth

et al. 2023), Alfalfa (Alzahrani et al. 2023), Chickpea (Hajibarat et al. 2015), Common Bean (Hromádová et al. 2023), and Juniper plant (Al-Yasi and Al-Qathanin 2024).

More recently and interestingly, one comprehensive database called Asteraceae Genomics Database (AGD-<http://cbcb.cdutcm.edu.cn/AGD/>) has been developed and established by a Chinese research team (Wang et al. 2024) to integrate and update all genomic data related to Asteraceae family species or plants which makes it serves as a systematic and powerful resource in evolutionary biology, genetic molecular analysis, and other breeding fields. Furthermore, such kinds of databases and others related genomic repositories platforms would support in-depth study as well as analysis of the global genetic diversity among *Asteraceae* family species at their molecular level and resulting to empower many researchers/breeders to advance their various future diversity investigations. With this background, as each marker system has its pros and cons to generate more and valuable genetic information, the main objective of this current research study is to evaluate the efficiency of SCoT, CBDP, and ISSR molecular markers (either independently or in combination) to distinguishing and analyzing the intraspecific genetic diversity and authenticate taxa at the species level among and/or within some species of Asteraceae family. As per our knowledge and recent survey, this is the first work study in Egypt used these three molecular marker techniques together with this number of Asteraceae plants in Egypt. That's why we hopefully this study could provide an important step toward the better classification and conservation strategies of this family species in Egypt.

Materials and methods

The experimental investigations of this current research study were conducted at the Molecular Genetics and Genome Mapping Laboratory (MGGM), Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), Egypt as well as the Agricultural Biotechnology Laboratory (ABL), Agricultural Botany Department, Faculty of Agriculture Saba Basha, Alexandria University.

Collected plant materials and its geographical location

A total of twenty wild and endemic plant species of composite (Asteraceae) family were collected from their natural habitat in Borg EL Arab–Alexandria area, Alexandria desert road surrounding area, Abis Area, Alexandria governate-Egypt in addition to Agricultural Genetic Engineering Research Institute (AGERI), Cairo university-Egypt (Figs. 1 and 2). The geographic distribution map of Asteraceae species collected for this study was created by the help of Google Earth Pro software and ArcGIS 10.4 online base-map system (Fig. 1). The list of these collected plant species and their scientific names, date of collection and sites (Latitude and Longitude) are presented in Table 1.

DNA based molecular markers analysis and genetic diversity estimates

DNA extraction, quantification, and integrity check

DNA samples of twenty different Asteraceae family plant individuals were utilized for the molecular diversity analysis in this current study. For that, the fresh leaf samples were gently collected separately from each plant material and immediately kept in ice

box for further steps. The total genomic DNA was isolated from the youngest fresh leaves (100–250 mg) following the classical CTAB (2%) method protocol of Murray and Thompson (1980) and Doyle (1990). To quantify the isolated DNA concentration, 1ul of the extracted DNA samples was used to determine spectrophotometrically at 260 nm (A260) absorption using NanoDrop 1000 (Thermo Scientific). The purity of extracted DNA samples from protein and polysaccharide contamination was assessed by estimating the absorbance ratio at A260/A280 and A260/A230 respectively (Wilson and Walker 2005). Eventually, the DNA samples were tested for its integrity using 0.7% agarose gel electrophoresis then diluted to 100 ng/ul and kept under -30°C for further downstream processes.

PCR amplification and visualization of the amplified fragments by gel electrophoresis technique

To assess the genetic diversity of these selected plant materials (Table 1), the PCR amplification step was performed using the well extracted and purified genomic DNA sample of each collected plant species and screened by three different DNA based molecular marker systems viz., SCoT (7 primers), CBDP (5 primers) and ISSR (7 primers). The detailed information of these 19 different primers were presented

Fig. 1 Geographic distribution map of Asteraceae species assessed in this study

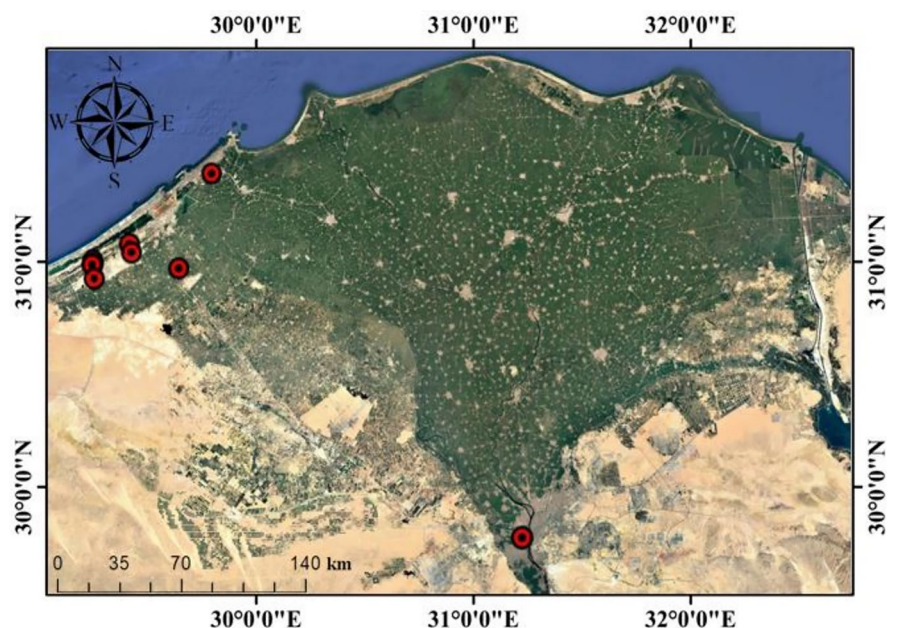


Fig. 2 The view of 20 plant species of Asteraceae family that found wild in the study areas (1; *Senecio Glaucus* L., 2; *Urospermum picroides* L., 3; *Glebionis coronaria* L., 4; *Launea nudicaulis* L., 5; *Sphagneticola trilobata* L., 6; *Conyza linifolia* L., 7; *Hedyponis rhogadioloides* L., 8; *Matricaria aurea* L., 9; *Leontodon tuberosus* L., 10; *Calendula arvensis* L., 11; *Echinops spinosus* L., 12; *Cardus getulus* L., 13; *Centaurea pallescens* L., 14; *Carduncellus mareoticus* Del., 15; *Centaurea alexandrina* Del., 16; *Centaurea calcitrapa* L., 17; *Scolymus hispanicus* L., 18; *Carthamus nitidus* L., 19. *Helianthus annuus* L., and 20; *Centaurea cristata* L.)



in Table 2. These primers were designed and chosen based on their efficacy in previous research such as Al-Qurainy et al. (2017) and Atia et al. (2021). Firstly, to amplify the seven SCoT primers, we followed the protocol described by (Abouseadaa et al. 2020). Whereas to conduct the CBDP analysis using five CAAT box—derived Polymorphism primers, a 7.5 volume reaction consisted of 3.75 ul 2×Taq PCR Master Mix (Green/ US Everbright In), 1.2 ul of primer, 1.5 ul of DNA, and completed by adding 1.05 ul of nuclease free water.

On the other hand, seven primers of ISSR were amplified into a 7.5 µl final reaction volume containing 2.3 ul 2×GoTaq® Master Mixes (Promega Corporation), 0.6 ul of primer, 2 ul of DNA template and the final volume was reached by add 2.6 ul of nuclease free water. The PCR condition for SCoT and CBDP were performed as follows: initial denaturation

at 94 °C for 5 min, followed by 40 cycles of DNA denaturation at 94 °C for 40 s, annealing temperatures from 49 to 50 °C for 50 s, extension at 72 °C for 1 min, and final extension at 72 °C for 10 min. While, in the ISSR analysis, the r PCR amplification condition was performed as follows: initial denaturation at 94 °C for 5 min, followed by 40 cycles of denaturation at 94 °C for 1 min, annealing temperature from 42 °C for 1 min, extension at 72 °C for 1 min, and final extension at 72 °C for 10 min.

The PCR amplification for Scot, CBDP, and ISSR molecular markers analysis was carried out using the Applied Biosystems 2720 Thermal cycler. The PCR products were separated by gel electrophoreses technique at 1.5% agarose gel stained with 10 ul ethidium bromide at a voltage of 100–120 V. Then the gel electrophoreses products were visualized using Gel Documentation Framework (Bio-Rad, USA).

Table 1 Information of the collected 20 plant species of Asteraceae family that evaluated in the present study

S. no.	Species	Collection site	GPS coordinates	Collector	Voucher number	Herbarium
1.	<i>Senecio glaucus</i> L.	Borg EL Arab	30° 53' 04" N, 29° 33' 21" E	Mariam Elgohary	AlexBA-C1	Alexandria University Herbarium
2.	<i>Urospermum picroides</i> L.	Borg EL Arab	30° 53' 04" N, 29° 33' 21" E	Mariam Elgohary	AlexBA-C2	Alexandria University Herbarium
3.	<i>Glebionis coronaria</i> L.	Borg EL Arab	30° 52' 53" N, 29° 33' 06" E	Mariam Elgohary	AlexBA-C3	Alexandria University Herbarium
4.	<i>Launea nudicaulis</i> L.	Borg EL Arab	30° 50' 08" N, 29° 33' 57" E	Mariam Elgohary	AlexBA-C4	Alexandria University Herbarium
5.	<i>Sphagnetocola trilobata</i> L.	AGERI	30° 01' 28" N, 31° 12' 09" E	Mariam Elgohary	AlexAG-C1	Alexandria University Herbarium
6.	<i>Conyza linifolia</i> L.	Borg EL Arab	30° 52' 53" N, 29° 33' 06" E	Mariam Elgohary	AlexBA-C5	Alexandria University Herbarium
7.	<i>Hedyponis rhogadioloides</i> L.	Borg EL Arab	30° 53' 46" N, 29° 33' 12" E	Mariam Elgohary	AlexBA-C6	Alexandria University Herbarium
8.	<i>Matricaria aurea</i> L.	Borg EL Arab	30° 53' 46" N, 29° 33' 12" E	Mariam Elgohary	AlexBA-C7	Alexandria University Herbarium
9.	<i>Lentodon tuberosus</i> L.	Borg EL Arab	30° 50' 08" N, 29° 33' 57" E	Mariam Elgohary	AlexBA-C8	Alexandria University Herbarium
10.	<i>Calendula arvensis</i> L.	Borg EL Arab	30° 56' 58" N, 29° 41' 28" E	Mariam Elgohary	AlexBA-C9	Alexandria University Herbarium
11.	<i>Echinops spinosus</i> L.	Borg EL Arab	30° 50' 08" N, 29° 33' 57" E	Mariam Elgohary	AlexBA-C10	Alexandria University Herbarium
12.	<i>Cardus getulus</i> L.	Borg EL Arab	30° 56' 58" N, 29° 41' 28" E	Mariam Elgohary	AlexBA-C11	Alexandria University Herbarium
13.	<i>Centaurea pallescens</i> L.	Alexandria Desert Road	30° 52' 02" N, 29° 53' 31" E	Mariam Elgohary	AlexDR-C1	Alexandria University Herbarium
14.	<i>Carduncellus maritoticus</i> Del.	Alexandria Desert Road	30° 55' 16" N, 29° 42' 08" E	Mariam Elgohary	AlexDR-C12	Alexandria University Herbarium
15.	<i>Centaurea alexandrina</i> Del.	Borg EL Arab	30° 50' 08" N, 29° 33' 57" E	Mariam Elgohary	AlexBA-C12	Alexandria University Herbarium
16.	<i>Centaurea calci-trapa</i> L.	Borg EL Arab	30° 50' 08" N, 29° 33' 57" E	Mariam Elgohary	AlexBA-C13	Alexandria University Herbarium
17.	<i>Scolymus hispanicus</i> L.	Abis	31° 11' 08" N, 29° 59' 40" E	Mariam Elgohary	AlexAS-C1	Alexandria University Herbarium

Table 1 (continued)

S. no.	Species	Collection site	GPS coordinates	Collector	Voucher number	Herbarium
18.	<i>Carthamus nitidus</i> L.	Abis	31° 11' 08" N, 29° 59' 40" E	Mariam Elgohary	AlexAS-C2	Alexandria University Herbarium
19.	<i>Helianthus annuus</i> L.	Borg EL Arab	30° 56' 58" N, 29° 41' 28" E	Mariam Elgohary	AlexBA-C14	Alexandria University Herbarium
20.	<i>Centaurea cristata</i> L.	AGERI	30° 01' 28" N, 31° 12' 09" E	Mariam Elgohary	AlexAG-C2	Alexandria University Herbarium

Table 2 Primers sequence information used to assess the genetic diversity

Marker type	Primer code	Sequence (5'–3')	Tm	References
SCoT marker	SCoT-7	CAACAATGGCTACCACGG	50 °C	Omar et al. (2020), Mohamed et al. (2024)
	SCoT-10	CAACAATGGCTACCAGCC	50 °C	
	SCoT-11	AAGCAATGGCTACCACCA	50 °C	
	SCoT-12	ACGACATGGCGACCAACG	50 °C	
	SCoT-13	ACGACATGGCGACCATCG	50 °C	
	SCoT-14	ACGACATGGCGACCACGC	50 °C	
	SCoT-15	ACGACATGGCGACCGCGA	50 °C	
CBDP marker	CBDP-4	TGAGCACGATCCAATAAG	50 °C	Atia et al. (2021)
	CBDP-5	TGAGCACGATCCAATCTA	50 °C	
	CBDP-7	TGAGCACGATCCAATCGA	50 °C	
	CBDP-9	TGAGCACGATCCAATGAT	50 °C	
	CBDP-15	TGAGCACGATCCAATTGA	50 °C	
ISSR marker	ISSR-4	ACACACACACACACACYG	42 °C	Tawfik and El-Mouhamady (2019)
	ISSR-9	GATAGATAGATAGATAGC	42 °C	
	ISSR-10	GACAGACAGACAGACAAT	42 °C	
	ISSR-11	ACACACACACACACACYA	42 °C	
	ISSR-12	ACACACACACACACACYC	42 °C	
	ISSR-13	AGAGAGAGAGAGAGAGYT	42 °C	
	ISSR-14	CTCCTCCTCCTCCTCTT	42 °C	

**T_m* melting temperature of primers

DNA bands scoring and data processing

In order to generate molecular data matrices following the electrophoresis step. Each DNA amplified band corresponded to different primer of SCoT, CBDP, and ISSR molecular markers respectively were scored visually for each species independently. The bands data were transformed into a binary matrix as follows; 1 (present of a fragment), 0 (absence of a fragment), and 0.999 (complete absence of band). To minimize the experimental errors, only distinct, clear, and reproducible bands were considered for scoring.

Genetic variation and marker efficiency analysis

The genetic diversity and molecular marker efficiency analysis were calculated following the polymorphism indices and parameters of iMEC (Amiryousefi et al. 2018), Online Marker Efficiency Calculator tool (<https://irscope.shinyapps.io/iMEC/>). The level of polymorphism of each marker was assessed to evaluate its effectiveness by certain criteria such as the polymorphism information content (Botstein et al. 1980), that measures the extent of genetic variation.

The discriminating power (Dp) and Effective multiplex ratio (EMR) of each marker were calculated

using the formula suggested by Tessier et al. (1999) and Powell et al. (1996), respectively. DARwin version 6.0.021. software (<https://darwin.cirad.fr/produit.php>) that developed by Perrier and Jacquemoud-Collet (2009) was employed to estimate the genetic diversity parameters among these different species of Asteraceae family. Similarities were estimated using Jaccard's coefficient (Bonham 1982). The genetic similarity coefficients and genetic distances were analyzed according to the method described by Nei (1973). To exhibit the genetic and phylogenetic relationships, a hierarchical clustering using UPGMA (Unweighted Pair Group Technique with Arithmetic Mean) was used to construct the dendrogram for each marker system based on the similarity coefficient (Schlee 1975).

To assess the population structure of the 20 species of Asteraceae family, The Bayesian algorithm-based STRUCTURE software version 2.3.4. (<https://web.stanford.edu/group/pritchardlab/structure.html>) that described by Hubisz et al. (2009) was used with all loci that generated from the three types of molecular markers. The sub-population number (K) was set from 1 to 10 by running 3 iterations with a burn-in time of 10,000 and a total of number of 1000 Markov Chain Monte Carlo (MCMC) repeats after burn-in process. To determine the optimal cluster number (K), the STRUCTURE results from previous step were packaged and uploaded to StructureSelector a web based online tool (<https://lmme.ac.cn/StructureSelector/>) to select and visualize the best genetic clusters based on DeltaK (ΔK) values (Evanno et al. 2005; Li and Liu 2018; Puneeth et al. 2023).

Results

DNA based molecular marker's efficiency and genetic polymorphism analysis

In the present study, 19 different DNA primers belonging to three modern molecular markers (SCoT, CBDP, and ISSR) were used for evaluating as well as providing better understanding of the genetic diversity in addition to the phylogenetic relations between 20 Asteraceae family species. As overall, all the nineteen primers were able to amplify the polymorphic DNA fragments among the twenty species. The gel electrophoresis profile of each marker type is shown

in Fig. 3A–C. Additionally, all important parameters and genetic information generated from the polymorphism calculation of SCoT, CBDP, and ISSR, respectively are outlined accordingly in Table 3. Considering the results obtained by seven SCoT primers, all these primers were tested for their ability to amplify the DNA samples of collected species and showed highly polymorphic amplification fingerprint patterns. All examined SCoT primers generated a total of 105 amplified DNA fragments (Table 3). Interestingly, all amplified fragments (105) were polymorphic and showed 100% percentage of polymorphism (PPB). While the polymorphic information content (PIC) for all SCoT primers were ranged from 0.328 (SCoT-13) to 0.444 (SCoT-7) with an average of 0.349 of PIC. Regarding the discriminating power (Dp) of SCoT marker primers, SCoT-14 primer showed the highest Dp value whereas SCoT-7 primer showed the lowest Dp value among all used SCoT primers. The effective multiplex ratio (EMR) for all SCoT primers varied from 1.000 (SCoT-7) to 12.737 (SCoT-13) and both SCoT-10 and SCoT-13 recorded the highest EMR of 12.421 and 12.737 values respectively with an average of 9.617 of EMR.

Regarding, the CBDP marker profiling, the five screened CBDP primers produced 78 amplified DNA fragments with a varied number of amplified bands per each primer (Table 3). For example, the CBDP-4 primer generated the lowest number of amplified bands (14) while the CBDP-2 primer showed the highest number of amplified bands (19). The PPB for all CBDP primers ranged from 86.67 to 100% with number of polymorphic bands ranged from 13 to 19, respectively. It is important to note that, the three CBDP primers (CBDP-4, CBDP-5, and CBDP-7) revealed 100% of PPB among all CBDP primers. Interestingly, the CBDP-15 primer recorded the highest value for PIC value (0.500) whilst the CBDP-9 primer displayed the lowest PIC value (0.340) among all CBDP primers. On the other side, CBDP-5 primer recorded high Dp (0.636) as well as EMR (11.474) values compared to other four CBDP primers with an average of 0.487 and 6.432 for both parameters, respectively. In our study, the obtained results from DNA fingerprinting profiles of ISSR marker system showed that an approximately of 102 amplified DNA fragments were produced with a varied number of amplified bands from each primer of 11–17 (Table 3).

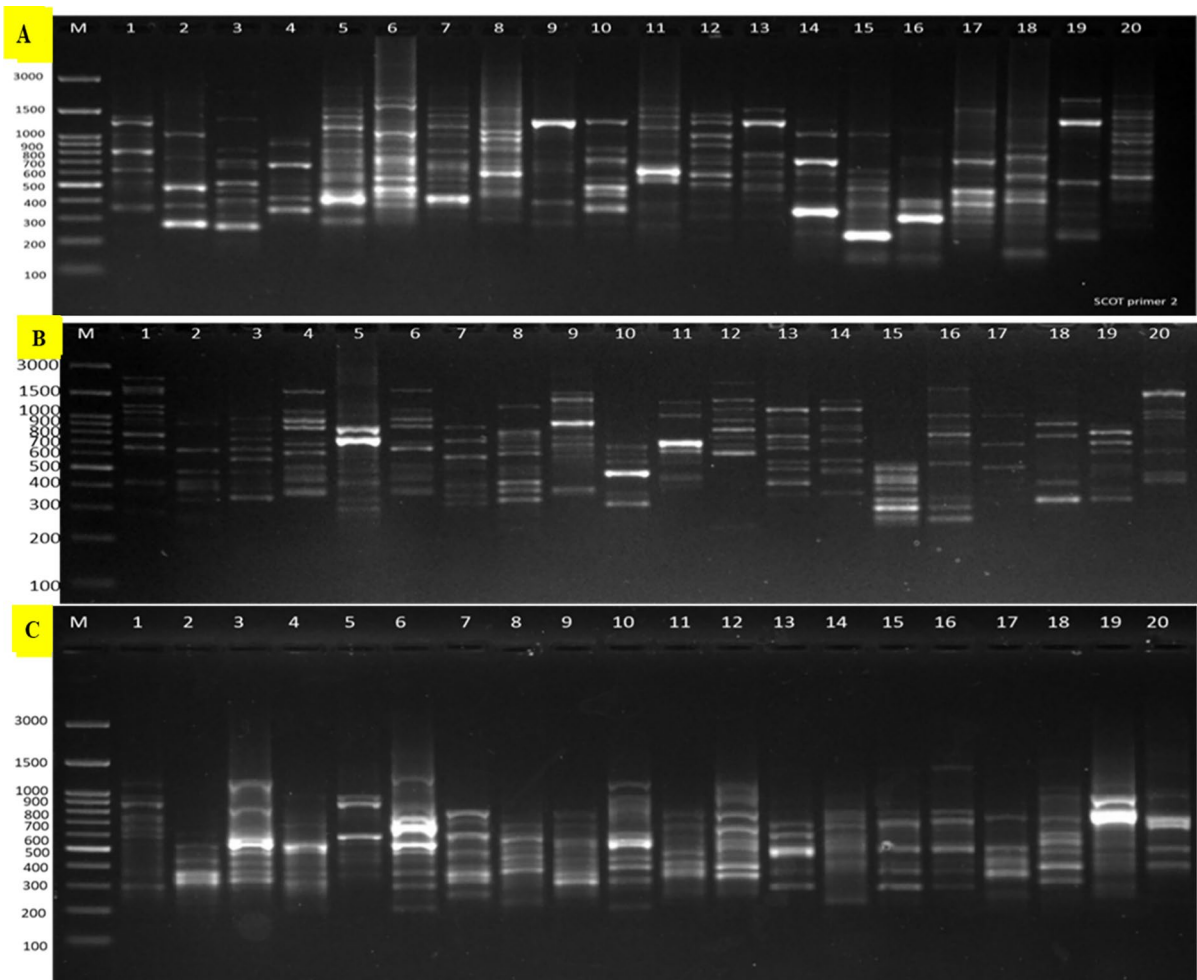


Fig. 3 PCR Amplification profiles using Start Codon Targeted (SCoT) (A), CAAT box-derived Polymorphism (CBDP) (B), and Inter Simple Sequence Repeat (ISSR) (C) primers among 20 species belonging to Asteraceae family

Notably, only ISSR-13 primer showed the highest number of polymorphic bands (17) as well as the highest percentage of polymorphism (100%) when compared to other remaining ISSR primers except (ISSR-9). The mean PPB percentage for all ISSR primers was 91.103%. While the PIC values of all ISSR primers varied from 0.327 (ISSR-10) to 0.585 (ISSR-9) whereas the discriminating power (D_p) values ranged from 0.212 (ISSR-9) to 0.643 (ISSR-11) with an average of 0.418 and 0.430 for PIC and D_p , respectively. On the other hand, three primers (ISSR-9, ISSR-12, and ISSR-14) recorded the lowest same values (1.000) as an effective multiplex ratio (EMR) when compared to other tested ISSR primers whilst the highest EMR value recorded by

two primers ISSR-4 and ISSR-13 of 11.526 and 11.053, respectively as displayed in Table 3.

Population structure analysis

The Bayesian algorithm-based STRUCTURE software version 2.3.4 was used to determine and analysis the population structure of the twenty species of Asteraceae family evaluated in the current study. The DNA fingerprinting data that produced from screening these species by 19 different primers of SCoT, CBDP and ISSR molecular markers, respectively were subjected to the STRUCTURE software. Then the obtained results were used to estimate the population structure index by the available web

Table 3 Comparison of the efficiency of Start Codon Targeted (SCoT), CAAT box-derived Polymorphism (CBDP), and Inter Simple Sequence Repeat (ISSR) primers

Marker type	Primer code	TAB	NPB	PPB (%)	PIC	Dp	EMR
<i>Scot</i>	SCoT-7	15	15	100.00	0.444	0.385	1.000
	SCoT-10	15	15	100.00	0.350	0.573	12.421
	SCoT-11	15	15	100.00	0.337	0.528	9.632
	SCoT-12	15	15	100.00	0.359	0.609	11.895
	SCoT-13	15	15	100.00	0.328	0.500	12.737
	SCoT-14	15	15	100.00	0.370	0.673	9.158
	SCoT-15	15	15	100.00	0.350	0.572	10.474
	Total	105	105	–	2.538	3.840	67.317
	Mean	15	15	100.00	0.349	0.576	11.053
<i>CBDP</i>	CBDP-4	15	15	100.00	0.447	0.319	1.000
	CBDP-5	19	19	100.00	0.364	0.636	11.474
	CBDP-7	15	13	86.67	0.362	0.628	9.158
	CBDP-9	14	13	92.86	0.340	0.538	9.526
	CBDP-15	15	15	100.00	0.500	0.316	1.000
	Total	78	75	–	2.013	2.437	32.158
	Mean	15.60	15	95.906	0.403	0.487	6.432
<i>ISSR</i>	ISSR-4	17	15	88.24	0.341	0.541	11.526
	ISSR-9	13	13	100.00	0.585	0.212	1.000
	ISSR-10	14	11	78.57	0.327	0.496	9.947
	ISSR-11	11	10	90.91	0.365	0.643	6.579
	ISSR-12	15	14	93.33	0.527	0.282	1.000
	ISSR-13	17	17	100.00	0.351	0.578	11.053
	ISSR-14	15	13	86.67	0.430	0.259	1.000
	Total	102	93	–	2.926	3.011	42.105
	Mean	14.6	13.29	91.103	0.418	0.430	6.015

TAB total amplified bands, NPB number of polymorphic bands, PPB percentage of polymorphism, PIC polymorphism information content, Dp discriminating power, EMR effective multiplex ratio

based online tool called “StructureSelector” which performed based on DeltaK (ΔK) value and K ranging 1–10. Using the SCoT marker dataset for population structure analysis, the results of seven SCoT primers showed that K=2 was largest delta K (ΔK) value at as represented by a sharp peak (Fig. 4A) which led to classify the 20 species into two distinct populations; population 1 (red vertical bars) and population 2 (green vertical bars). It’s important to note that, some of these species showed as a pure and other species showed as a slight admixture genetic group which their vertical graph bar contains both red and green colors. Interestingly, the highest number of Asteraceae family species was genetically grouped to the population 1 (red vertical bars).

With regard to the CBPD marker results, by utilizing the obtained data recorded by its five primers, the population structure analysis divided the tested 20 species into many genetic groups (7) with different colors represented these different groups. However,

the most suitable K value for determining the genetic clusters of these evaluated Asteraceae family species based on CBPD marker system was at K=7 of ΔK value as shown in Fig. 4B.

On the other side, the obtained results from both STRUCTURE software and StructureSelector online tool to investigate the genetic diversity of Asteraceae family species using five primers of ISSR marker system revealed that, the ΔK value reached its peak at K=8 as represented by a sharp peak in Fig. 4C. Thus, the population structure analysis divided the 20 species into 8 genetic groups represented by 8 different colors in Fig. 4C.

Phylogenetic and cluster analysis through UPGMA method

The hierarchical clustering using UPGMA phylogenetic trees based on the genotypic data produced by 19 screened primers pairs of the three different DNA

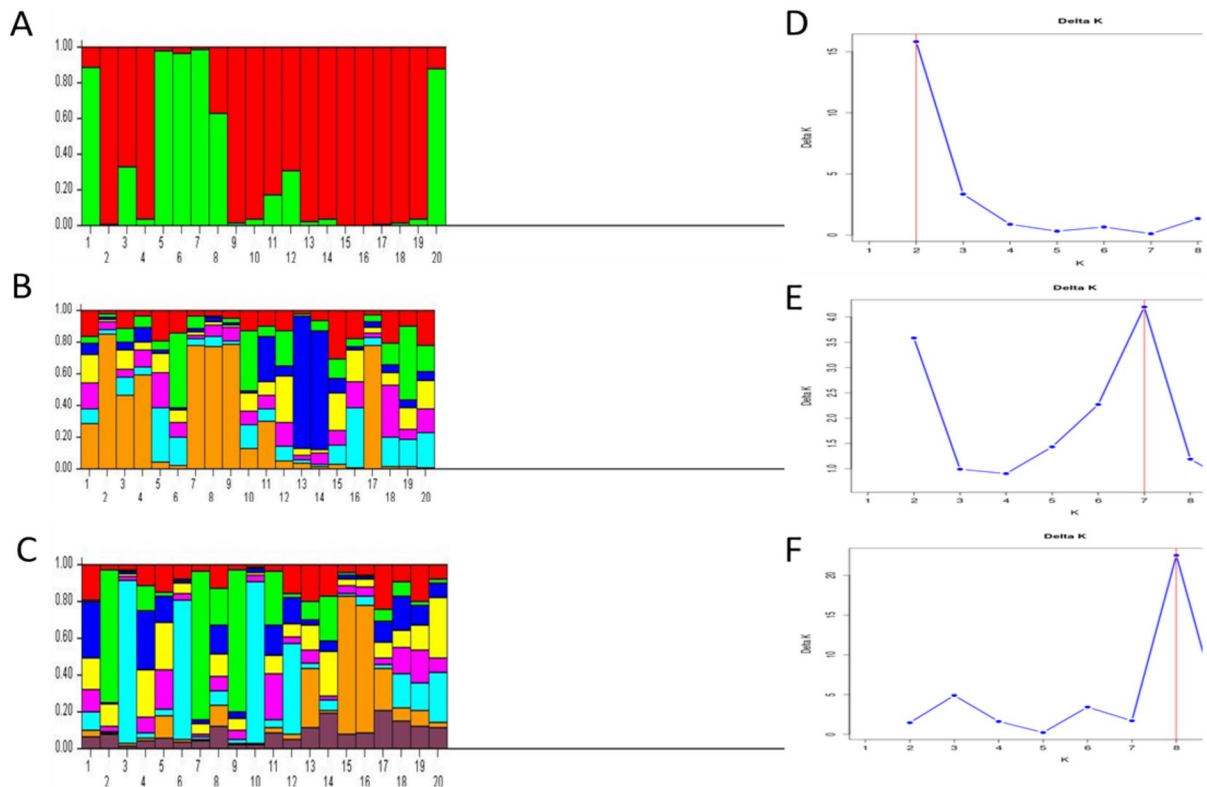
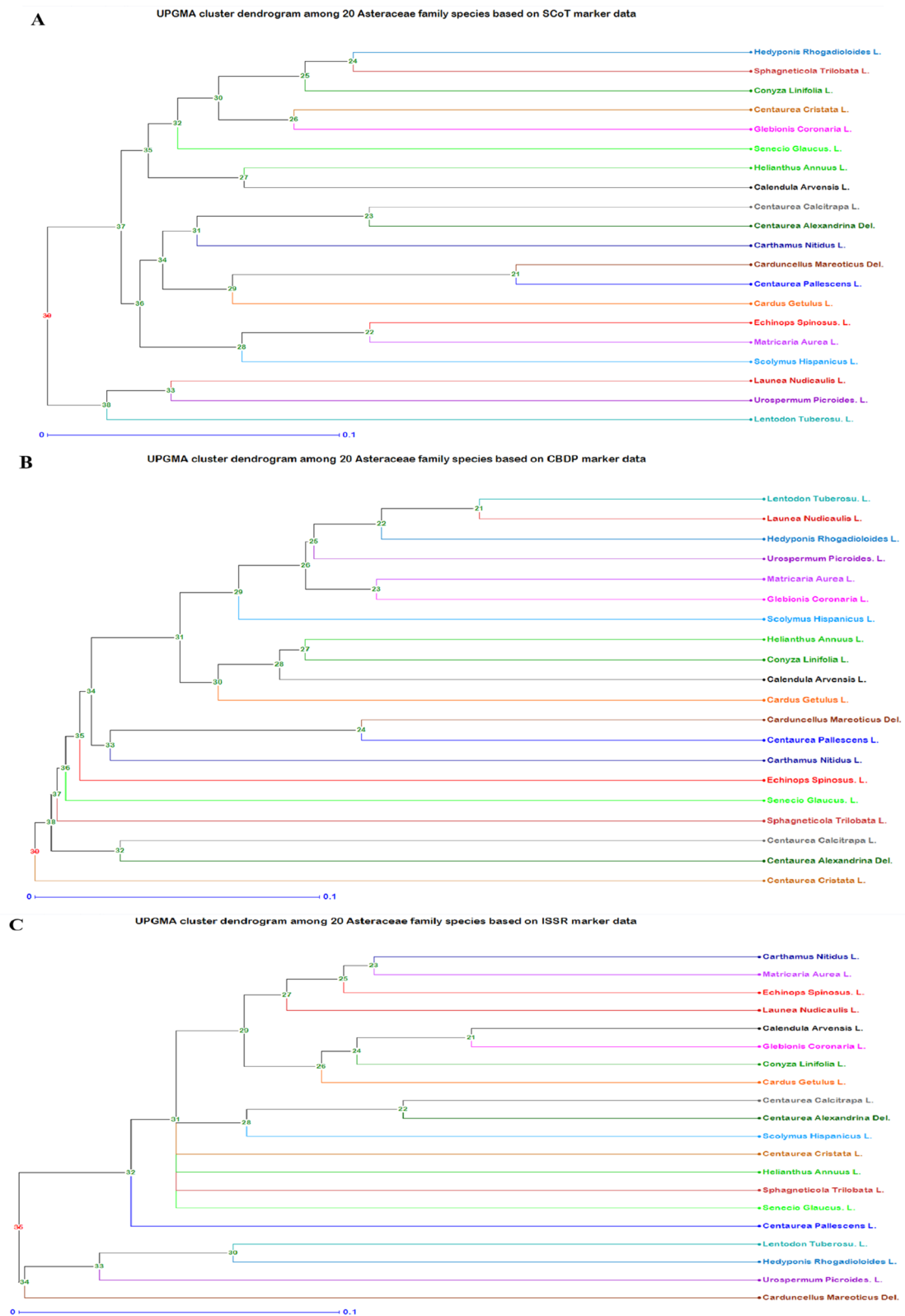


Fig. 4 Graphical summary of STRUCTURE results of 20 species based on the generated markers datasets. **A–C** Population structure of 20 Asteraceae family species inferred by STRUCTURE analysis based on Start Codon Targeted (SCoT), CAAT box-derived Polymorphism (CBDP), and Inter Simple Sequence Repeat (ISSR) markers data respectively. Each solid bar represents a single species, while each color denotes a

specific population subset. The horizontal axis numbers 1–20 represents 20 studied individuals as presented in Table 1. **D–F** The estimated membership fraction using LnP(D)-derived delta K (ΔK) with cluster number (K) ranged from 1 to 10 for K=2, 7, and 8 respectively calculated by StructureSelector web based online tool

based molecular marker techniques were able to explain the genetic relationships among the 20 species belonging to Asteraceae family. The UPGMA dendrogram trees of SCoT, CBSP, and ISSR molecular markers are presented in Fig. 5A–C. According to the SCoT marker dendrogram, the evaluated species were distinguished into two main clusters with only three species viz., *Launea nudicaulis* L., *Urospermum picroides* L., and *Leontodon tuberosus* L. in the second main cluster and other 17 species spreading over the first main cluster (Fig. 5A). The second main cluster subdivided also into two main sub-clusters; the first sub-cluster consisted of eight species while the second sub-cluster consisted of nine species. The UPGMA dendrogram based similarity coefficients from CBDP marker profiles revealed that Asteraceae family species have been grouped into two

main groups similar to the SCoT primer dendrogram (Fig. 5A). Surprisingly, the second group consisted of only one species of *Centaurea cristata* L. that collected from AGERI while the remaining 19 species of Asteraceae family spreading over the first main cluster. The first group that consisted of major studied species of Asteraceae family have been subdivided into two sub-clusters. The first sub-cluster contained 17 species whilst the second sub-cluster contained only two species of *Centaurea calcitrapa* L., and *Centaurea alexandrina* Del. that haven collected from the same study area of Borg EL Arab (Fig. 5B). The cluster tree based on UPGMA of ISSR marker DNA fingerprinting results also categorized the study species into two main groups similar two other SCoT and CBDP cluster trees. Accordingly, four species formed the cluster II whereas the remaining 16 Asteraceae



◀**Fig. 5** Unweighted pair Group method with arithmetic mean (UPGMA) cluster dendrogram illustrating the genetic relationships among 20 Asteraceae family species. **A** Based on seven Start Codon Targeted (SCoT) primers. **B** Based on five CAAT box-derived Polymorphism (CBDP) primers and **C** Based on seven Inter Simple Sequence Repeat (ISSR) primers

family species spreading over the cluster I (Fig. 5C). The cluster I was divided into two sub-sub clusters and the sub-sub cluster II consisted only one Asteraceae species of *Centaurea pallescens* L., that collected from Alexandria Desert Road study area.

Interestingly, the sub-sub cluster I consisted of four species of Asteraceae family viz., *Centaurea cristata* L., *Helianthus annuus* L., *Sphagneticola trilobata* L., and *Senecio glaucus* L. showed highly genetic similarity among other 15 species that spread over sub-sub cluster II. Surprisingly, these previous mentioned four species of Asteraceae family have been collected from two different study areas of Borg EL Arab and AGERI respectively. The UPGMA dendrogram based similarity coefficients which integrated all the three markers' data generated from SCoT, CBDP and ISSR profiles revealed that Asteraceae family species have been grouped into two main groups (Fig. 8A). Surprisingly, the second cluster consisted of only three species of Asteraceae family viz., *Lentodon tuberosu* L., *Launea nudicaulis* L., and *Urospermum picroides* L. that collected from the same study location of Borg EL Arab. Interestingly, the cluster I consisted of that consisted major species Asteraceae family was divided into two main sub-subgroups. The second sub-subgroup consisted of two species of Asteraceae family while the first sub-subgroup consisted of remaining 15 species (Fig. 8A).

Factorial analysis to determine the species' genetic relationships

The factorial analysis represented as a diagram in Fig. 5, is a kind of principal coordinate analysis (PCoA) which facilitate the identification of genetic relations among studied species/individuals that showing a close genetic similarity group based on their genetic similarity coefficient values. It resulted into a two main dimensional scatter plots of species/genotypes and the species belonging to a single plot represent a set of genetically similar species. The essential genetic base in this analysis that, the closer

the distance between studied species the closer the genetic relationship between them, and vice versa. Therefore, in our investigations, the obtained results from the factorial analysis based on SCoT marker data using DARWin software revealed that the studied species grouping was in accordance with the clustering analysis carried out by UPGMA similarity dendrogram (Fig. 6A). Interestingly, the *Centaurea cristata* L. species was clustered separately from the other species in addition to some other species not only were clustered separately but also were separated independently from one another such that occurred between the two species of *Matricaria aurea* L., and *Calendula arvensis*. Regarding the factorial analysis based one five CBDP primers, the obtained results observed that three species, *Launea nudicaulis* L., *Calendula arvensis* L., and *Centaruea alexandrina* Del. have a slight deviation in which they were clustered separately and independently from the other evaluated Asteraceae species in the group which they all belong to. These results were also in a line with the UPGMA dendrogram of similarity clustering based on CBDP marker DNA fingerprinting results (Fig. 6B). In case of the factorial analysis results obtained from the ISSR marker profiles, the results were in an agreement with the results illustrated in the clustering dendrogram in Fig. 6C. Interestingly, out of all 20 species, *Carduncellus mareoticus*, *Scolymus hispanicus*, and *Urospermum picroides* L. were clustered separately and independently from the other species. Considering the factorial analysis based on the 19 primers of all three used molecular markers techniques (Fig. 8B), the obtained results observed that the Asteraceae family species PCoA were in an agreement with the results illustrated in the clustering dendrogram in Fig. 8A with a slight difference such as *Scolymus hispanicus* L. that clustered separately and independently from *Launea nudicaulis* L.

The dissimilarity matrix developed using DARwin software from three molecular markers

The possibility to assess the genetic diversity in the twenty species belonging to Asteraceae family in current analysis can be observed in Fig. 7A–C as regards the matrix dissimilarity generated from the Nei genetic dissimilarity coefficient. The results generated from SCoT marker system (Fig. 7A) revealed that the lowest genetic dissimilarity value

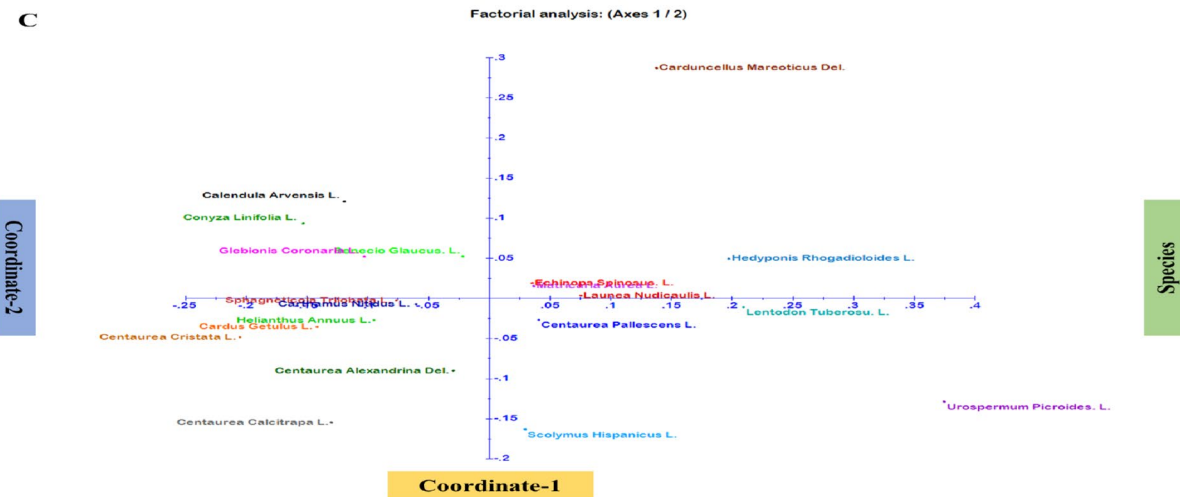
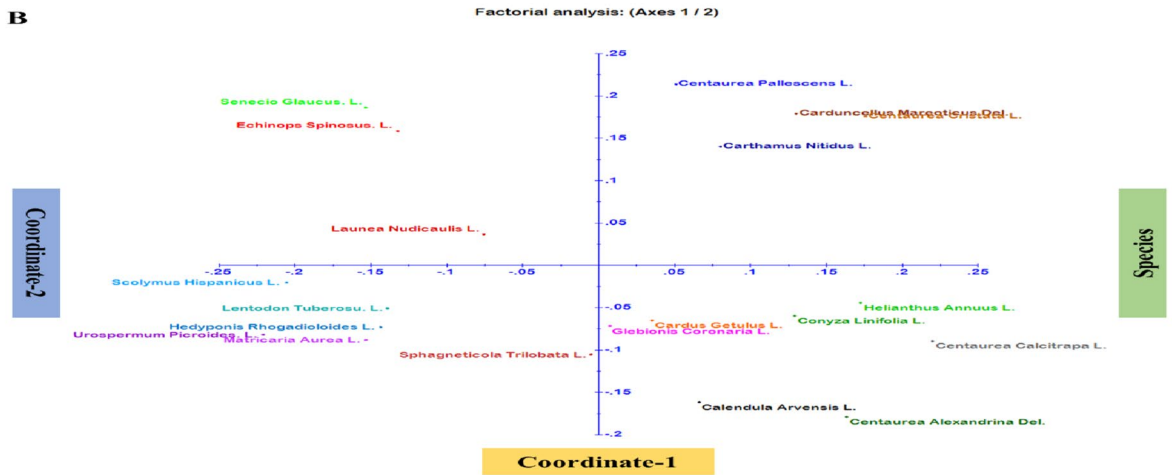
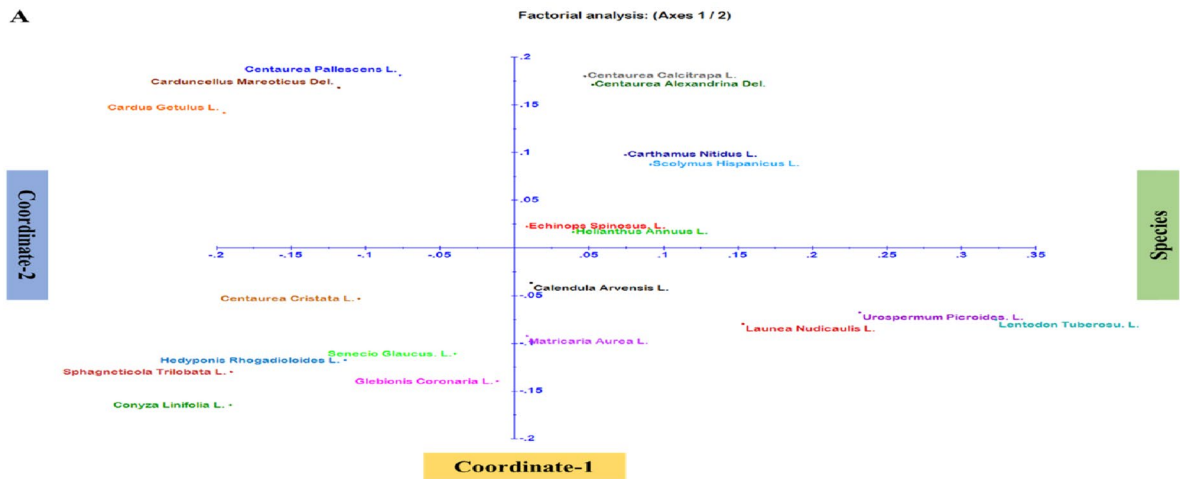


Fig. 6 Factorial analysis based on the genetic similarity coefficients using 19 primers for 20 Asteraceae species. A–C for Start Codon Targeted (SCoT), CAAT box-derived Polymorphism (CBDP), and Inter Simple Sequence Repeats (ISSR) principal component analysis (PCoA), respectively

was between *Centaurea pallescens*, and *Cardunculus mareoticus* Del., species (0.161) while highest genetic dissimilarity value was between *Leontodon tuberosus* L., and *Cardus getulus* L. (0.660). It is importantly to note that, the obtained dissimilarity matrix results of SCoT marker profiles were in line with the results recorded in Figs. 5A and 6A of similarity clustering analysis. As regard to dissimilarity matrix resulted from the CBDP mark DNA fingerprinting data, the genetic dissimilarity values ranged from 0.200 to 0.667 (Fig. 7B). Interestingly, the genetic dissimilarity between *Launea nudicaulis* L., and *Leontodon tuberosus* species was lower value among all analyzed Asteraceae family species and these results were also in agree with the UPGMA cluster dendrogram of CBDP marker type (Fig. 5B). On the other hand, the obtained results in Fig. 7B were not agree with the results recorded the factorial analysis based on the genetic similarity coefficients between *Launea nudicaulis* L., and *Leontodon tuberosus* species (Fig. 6B). Regarding the genetic dissimilarity matrix of ISSR marker system, the obtained results as shown in Fig. 7C observed that the genetic dissimilarity between *Glebionis coronaria* L., and *Calendula arvensis* species was the lower value of 0.180 (Fig. 7C)

While the highest genetic dissimilarity value of 0.640 was observed between *Urospermum picroides* L., and *Centaurea cristata* L. species. These obtained results were in a line with the ISSR marker cluster dendrogram (Fig. 5C) and not agree with ISSR factorial analysis results (Fig. 6C). Considering the obtained dissimilarity matrix results from integrating all molecular marker profiles data (Fig. 8C). The obtained results revealed that, the lowest genetic dissimilarity value was between *Glebionis coronaria* L., and *Calendula arvensis* L., species (0.2836) while highest genetic dissimilarity value was between *Urospermum picroides* L., and *Centaurea pallescens* L., species of Asteraceae family (0.5104). Noticeably, these obtained results of dissimilarity matrix generated from all molecular markers profiles data (Fig. 8C) were agree with the

results recorded in Fig. 8A and Fig. 8A of similarity clustering analysis.

Discussion

The urgent need to study and diving deeply into the genetic diversity of any genetic resource materials, is the essential step toward better understanding of evolutionary relationships as well as the adaptation of many plant species of this planet. Additionally, with the nature intricate genetic background of Asteraceae family species in Egypt. Hence, the utilization as well as the effective combination of different molecular markers techniques playing a pivotal role to studying its genetic diversity patterns of Asteraceae family species in Egypt. In addition to that, it's able calculating accurately the genetic variation of this family due to their widely growing habitats and adaption of these species Egypt. Therefore, the current research work was primarily aimed to explore the genetic diversity among 20 species belonging to different species of *Asteraceae* family collected from four different geographical areas in northern part of Egypt. In addition, our research study also aimed to examine the efficiency of combining three DNA based marker systems to provide more realistic information related to the genetic diversity among these studied species. In general, our results suggest that the efficiency of SCoT, CBDP, and ISSR molecular markers techniques was strong in DNA fingerprinting assessment of the evaluated individuals of Asteraceae family species. However, our results observed that the highest number of polymorphic DNA fragments/bands were amplified by SCoT marker primers (105) followed by ISSR marker primers (93), and CBDP marker primers (75), respectively. Additionally, out of 19 designed primers used in this study, the primers: CBDP-15 (0.500), ISSR-12 (0.527), and ISSR-9 (0.585) recorded the highest values for PIC when compared to other 16 screened primers. As PIC parameter is an important index to estimate the genetic status beside its ability to emphasize the discrimination power of various DNA primers has proven to study the genetic relationships among many plant species (Ravindran et al. 2023; Salama et al. 2024). It was reported in earlier research that, PIC values smaller than 0.25 indicates low levels of polymorphism associated to a locus, whereas PIC values between 0.25 and 0.5

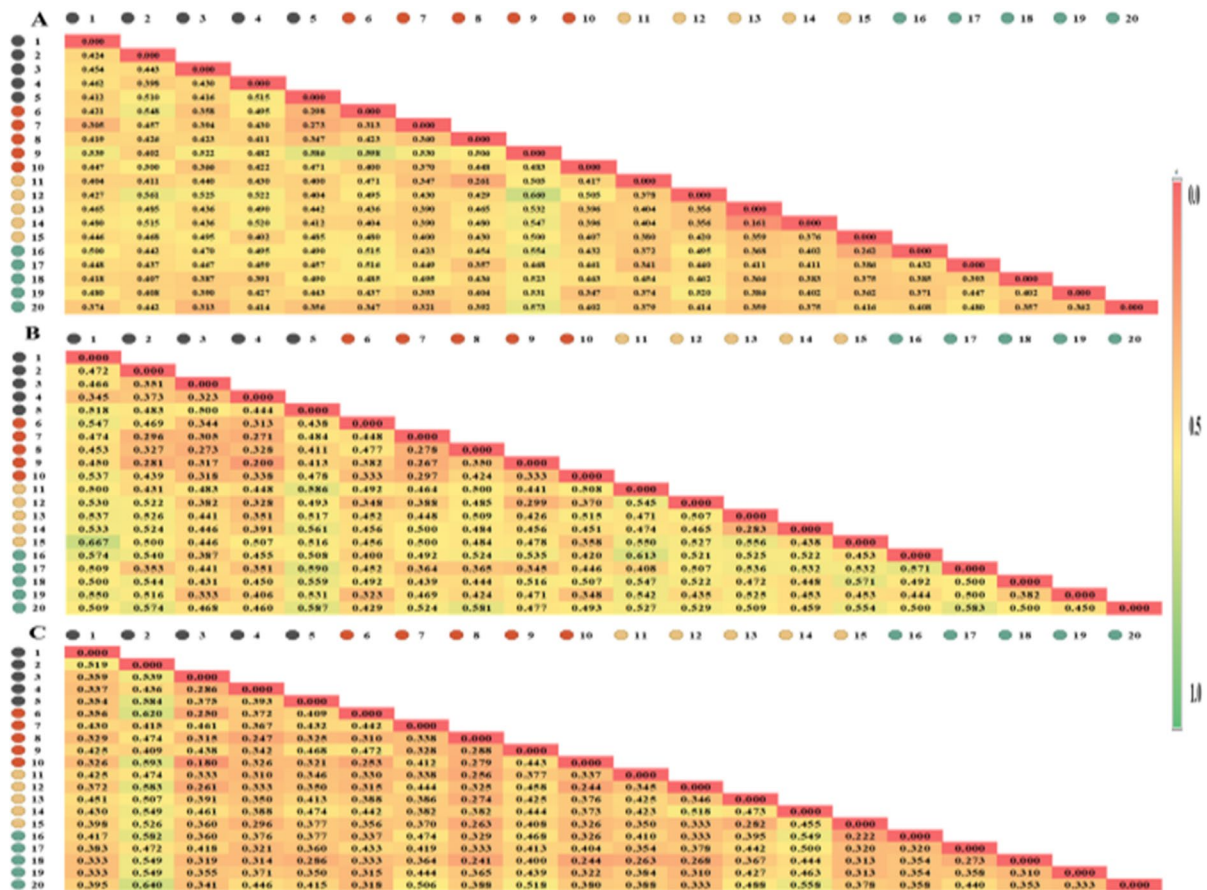


Fig. 7 Dissimilarity matrix among 20 Asteraceae species. **A–C** for Start Codon Targeted (SCoT), CAAT box-derived Polymorphism (CBDP), and Inter Simple Sequence Repeats (ISSR), respectively

imply moderate levels of polymorphism, while PIC values greater than 0.5 indicate high levels of polymorphism (Zhang et al. 2022; Yirgu et al. 2023). In this regard, the primers namely, CBDP-15, ISSR-12, and ISSR-9 that observed the PIC values equal or greater than to 0.5, these primers suggested to be more efficient to calculate the existing genetic diversity among studied species belonging to of Asteraceae family. On the other hand, SCoT-7, CBDP-4, and ISSR-14 observed a PIC value greater than to 0.4 could be also useful primers to obtain this goal along with the above-mentioned primers (PIC value greater than or equal to 0.5). Importantly, the utilization of PIC values of SCoT, CBDP, and ISSR molecular markers to be a strong discriminatory power tool in accumulating the accurate variability knowledge of different Asteraceae family species has been reported in many molecular diversity investigations such as the

research done by Atia et al. (2021), Samarina et al. (2021) and Jamshidi et al. (2024). More recently, one molecular study combined five SCoT primers along with five ISSR primers to estimate the genetic variability among two cultivated plant species and their three wild relatives from Asteraceae family. The results of this investigation demonstrated that, even with the lower number of ISSR, and SCoT primers still their ability to inspecting the inter-species diversity among the Asteraceae family species was highly effective and advantageous. Together, our results supporting the usefulness of combining three molecular techniques to understand the genetic diversity as well as the phylogenetic relationships between Asteraceae family species in Egypt.

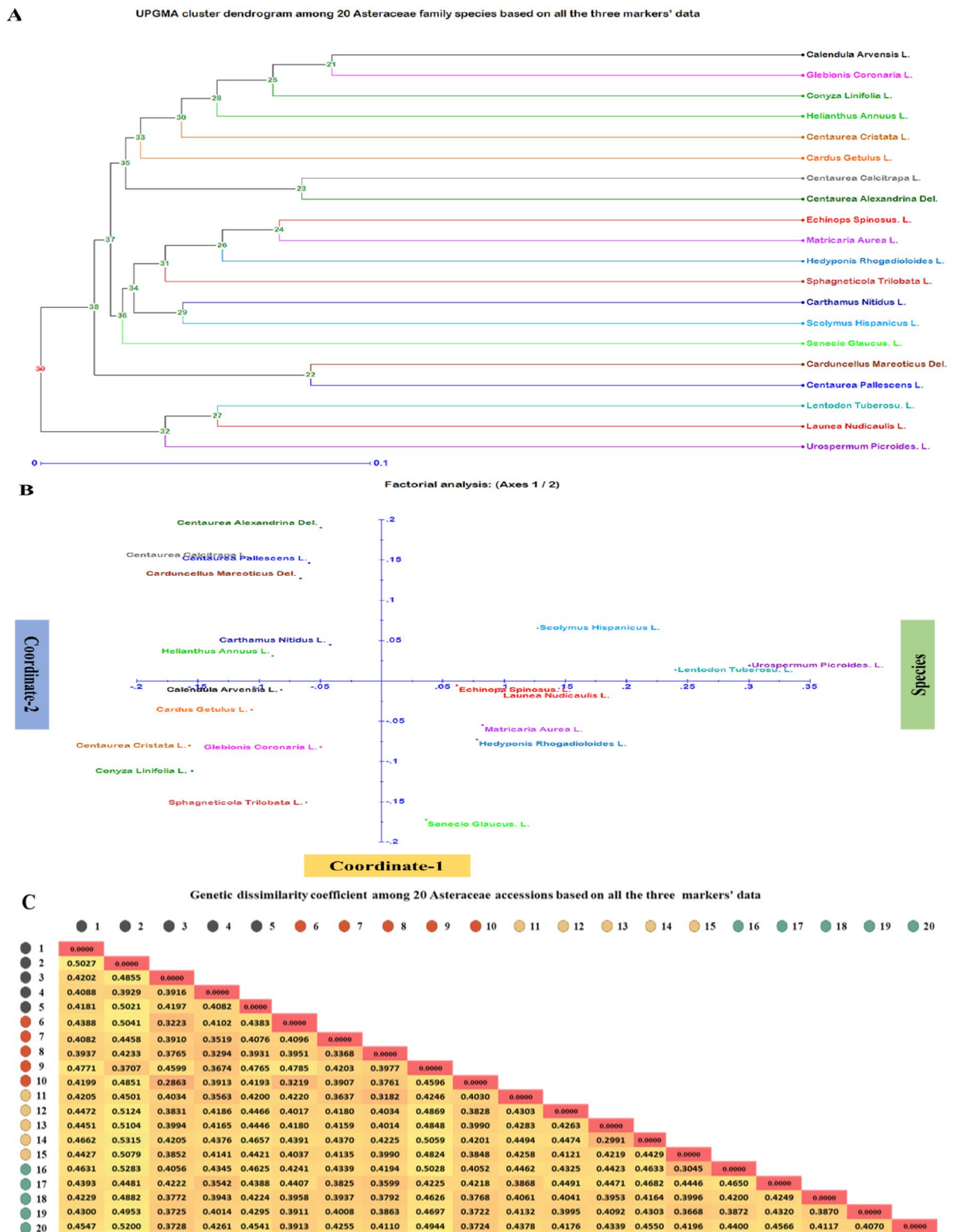
The strong and/or weak correlation between different geographic distribution factors and genetic diversity as well as phylogenetic relationships within

plant populations has been reported previously in many investigations based on clustering and population structure analyses. Thus, in our investigation, the population structure estimated by ISSR marker primers grouped the 20 species into different non-uniform genetic groups (eight different colors) as found in CBDP marker population structure analysis results that separated the studied species into seven subpopulations. However, despite being pure as showed in the population structure results of SCoT marker type, On the contrary, both CBDP and ISSR marker profiles observed two other categories of slight admixture and much admixture species population structure. It is very important to state that type of population structure might be due to the various membership fractionation of each species among all evaluated Asteraceae family species. Here, our report showed that the results of phylogenetic and cluster analysis through UPGMA method and the results of factorial analysis (Principal Coordinate Analysis) based on three molecular marker data were highly similar, and the genetic relationships among Asteraceae family species could be seen more intuitively. It was reported that the combination of these two types of genetic analysis could be more effective in analyzing the genetic relationships among different plant species (Kongjaimun et al. 2022; Li et al. 2022; PENG et al. 2023). Interestingly, the results obtained from all genetic tests in our current study such as population structure, UPGMA clustering, PCoA and dissimilarity analyses were found to be consistent with each other. Moreover, in CBDP UPGMA cluster *Centaurea Cristata* L. that collected from AGERI was grouped alone in one main cluster while in both SCoT and ISSR UPGMA cluster was grouped in sub-sub cluster. With regard to the geographical distribution of Asteraceae species, it was reported that the geographical distribution affect genetic diversity and population structure leading to clear genetic differences due to the variable habitat conditions as well as accumulated gene flow among studied species (Xiong et al. 2020; Shooshtari et al. 2024). In agree with this fact, there were three species of Asteraceae family *Lentodon tuberosu* L., *Launea nudicaulis* L., and *Urospermum picroides* have been grouped in a separate cluster of the dendrogram generated by integrating all the three molecular markers data (Fig. 8A) and all the three species were collected the same study location of Borg EL Arab.

In Summary, our current investigation confirmed the potential of Start Codon Targeted (SCoT), and CAAT box-derived Polymorphism (CBDP) as a modern as well as informative functional gene targeted molecular markers beside Inter Simple Sequence Repeats (ISSR) in estimating the genetic variation among and within Asteraceae family species. Additionally, these molecular marker systems would be more useful for the molecular breeding programs focusing mainly on wild species (hidden treasure) as well as Asteraceae family crops for instance construction of linkage and QTL maps, genetic diversity assessment, germplasm identification and characterization as well as biotic and a biotic stress tolerance. Taking together, the results of our study, for better conservation and sustainable utilization of the Asteraceae germplasm it is not advisable to adopt the taxonomic characterization alone because it led to unambiguous identification of species and varieties. That's why its critically important to combine the macro- and micromorphological characters along with more novel and informative DNA based molecular markers to avoid the challenges facing researchers to identify and classify the Asteraceae species in Egypt based on the floristic composition. Considering the weakness points of each used molecular markers tested in this study we suggested the using of advanced sequencing-based approaches (e.g., chloroplast genes, ITS) to validate or refine the patterns observed in the present study, particularly for higher-level taxonomic relationships among and within the studied species. To the best of our knowledge, this is the first report on genetic diversity analysis among 20 species of Asteraceae family using three molecular markers, i.e., SCoT and CBDP, along with ISSR markers.

Conclusion

In essence, the results and information obtained from this current study may help more researchers and scientists not only to deep understanding of the genetic diversity between and within the unique Asteraceae family genetic resources but also could contribute to the broader utilization of its stand-out potentials for achieving sustainable agriculture and food security goals especially to Egypt country. In summary, our results revealed that the effective combination of different molecular marker



◀**Fig. 8** Genetic diversity analysis of 20 Asteraceae species. **A** Unweighted pair group method with arithmetic mean (UPGMA) cluster dendrogram illustrating the genetic relationships among 20 Asteraceae family species. **B** Factorial analysis based on the genetic similarity coefficients using all the three markers' data of 20 Asteraceae species. **C** Dissimilarity matrix among 20 Asteraceae species for all the three markers' data

techniques such as SCoT, CBDP, and ISSR markers were able to estimate as well as provided comprehensive overview of the genetic diversity present in the studied Asteraceae family species. Our results indicated that the targeted approach of combining multiple functional and non-functional marker systems could provide a valuable insight for long-term DNA fingerprinting and comprehensive perspectives on the genomic landscape of different Asteraceae family species which subsequently helping and supporting the future conservation and management efforts for wild plants in Egypt. Additionally, the obtained results suggested that more focus must be directed towards harnessing of the available genetic diversity of Asteraceae family as its utilization could possess substantial implications for future molecular breeding and genetic improvement programs of different wild and cultivated species in Egypt. We also proposing in future that more investigations have to be carried out using the sequencing-based approaches (e.g., chloroplast genes, ITS) to validate or refine the patterns observed in the present study, particularly for higher-level taxonomic relationships among and within the studied species.

Acknowledgements The authors would like to express their deepest gratitude to Professor Dr. Hossam M. E. El-wakil for his sincerely contributions, dedication, and expertise which significantly advanced this work study. We honor his memory with this manuscript. We also acknowledge tireless efforts and support made by the Head of Agricultural Botany Department, Faculty of Agriculture Saba Basha, Alexandria University, and the Director of the Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center (ARC), Cairo University for their support while conducting the various experiments of this research work.

Author contributions "H.M.E.E., M.A.M.A., E.A.A.S., and M.E.M.A.E., conceived and designed the study experiments; H.M.E.E., M.A.M.A., and E.A.A.S. supervised the experiments; M.E.M.A.E., and E.A.A.S., performed the experiments; N.R.A., and M.A.M.A., supported the experiments conduction. E.A.A.S., and M.E.M.A.E., carried out the full statistical analyses, and prepared the original draft of manuscript. M.A.M.A., E.A.A.S., N.R.A., and M.E.M.A.E., contributed to the editing and revisions of the manuscript; M.A.M.A., and H.M.E.E.,

contributed to the funding acquisition and resources. This work was carried out in collaboration among all authors. "

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). This research work was supported by the Documentation of Elite Natural Resources in Egypt through DNA-Barcoding Technology Project funded by the Academy of Scientific Research & Technology (ASRT).

Data availability All data supporting the findings of this study are available within the paper and its Supplementary Information.

Declarations

Competing interests The authors declare no competing interests.

Ethical approval and consent to participate Not applicable as the present study doesn't involve any human subjects or animal research.

Consent for publication We affirm that all authors have reviewed and approved the content of this manuscript and agree to its submission for publication.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abouseadaa HH, Atia MA, Younis IY, Issa MY, Ashour HA, Saleh I, Osman GH, Arif IA, Mohsen E (2020) Gene-targeted molecular phylogeny, phytochemical profiling, and antioxidant activity of nine species belonging to family Cactaceae. *Saudi J Biol Sci* 27:1649–1658
- Adel EG, Nahed EH, Khafagi AA, Mostafa NA (2020) Two additions to the Asteraceae of Egypt. *Not Sci Biol* 12:14–21
- Al-Qurainy F, Khan S, Nadeem M, Tarroum M, Gaafar ARZ (2017) Antioxidant system response and cDNA-SCoT marker profiling in *Phoenix dactylifera* L. plant under salinity stress. *Int J Genom* 2017:1537538

- Al-Yasi HM, Al-Qthanin R (2024) Comparing genetic differentiation and variation using ISSR and SCoT among Juniper plant markers in Saudi Arabia. *Front Plant Sci* 15:1356917
- Alzahrani OR, Alshehri MA, Alasmari A, Ibrahim S, Oyouni AA, Siddiqui ZH (2023) Evaluation of genetic diversity among Saudi Arabian and Egyptian cultivars of alfalfa (*Medicago sativa* L.) using ISSR and SCoT markers. *J Taibah Univ Sci* 17:2194187
- Amar M, Hassan A, El Sherbeny EA (2012) Assessment of genetic diversity in some wild plants of asteraceae family by ribosomal DNA sequence. *Egypt J Genet Cytolpl* 41:195–208
- Amiryousefi A, Hyvönen J, Pocai P (2018) iMEC: online marker efficiency calculator. *Appl Plant Sci* 6:e01159
- Atia MAM, El-Moneim DA, Abdelmoneim TK, Reda EH, Shakour ZTA, El-Halawany AM, El-Kashoury E-SA, Shams KA, Abdel-Azim NS, Hegazy M-EF (2021) Evaluation of genetic variability and relatedness among eight Centaurea species through CAAT-box derived polymorphism (CBDP) and start codon targeted polymorphism (SCoT) markers. *Biotechnol Biotechnol Equip* 35:1230–1237
- Bidyananda N, Jamir I, Nowakowska K, Varte V, Vendrame WA, Devi RS, Nongdam P (2024) Plant genetic diversity studies: Insights from DNA marker analyses. *Int J Plant Biol* 15:607–640
- Bolous, L., 2002. Flora of Egypt Vol. (3) printed al hadara publishing Cairo, Egypt.
- Bonham CD (1982) A dissimilarity coefficient and its use. *Rangel Ecol Manag* 35:90–92
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314
- Doyle JJ (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Elgazzar A, El-Husseini N, Khafagi A, Mostafa NAA (2020) Computer-generated keys to the flora of Egypt. 10. The spineless taxa of Asteraceae. *Egypt J Bot* 60:97–131
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Hajibarat Z, Saidi A, Hajibarat Z, Talebi R (2015) Characterization of genetic diversity in chickpea using SSR markers, start codon targeted polymorphism (SCoT) and conserved DNA-derived polymorphism (CDDP). *Physiol Mol Biol Plants* 21:365–373
- Hamdan YA (2024) Genetic diversity assessment of Palestinian safflower (*Carthamus tinctorius* L.) utilizing DAMD molecular markers. *Biotechnol Lett* 46:1293–1302
- Hamouda M (2019) Molecular analysis of genetic diversity in population of *Silybum marianum* (L.) Gaertn in Egypt. *J Genet Eng Biotechnol* 17:12
- Heidari P, Etminan A, Azizinezhad R, Khosroshahi M (2017) Genomic variation studies in durum wheat (*Triticum turgidum* ssp. durum) using CBDP, SCoT and ISSR markers. *Indian J Genet Plant Breed* 77:379–386
- Hromadová Z, Gálová Z, Mikolášová L, Balážová Ž, Vivodík M, Chňápek M (2023) Efficiency of RAPD and SCoT markers in the genetic diversity assessment of the common bean. *Plants* 12:2763
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9:1322–1332
- Jamshidi B, Etminan A, Mehrabi AM, Pour-Aboughadareh A, Shooshtari L, Ghorbanpour M (2024) Genetic diversity of Artemisia species based on CAAT-box derived polymorphism (CBDP) and start codon targeted (SCoT) markers. *Genet Resour Crop Evol* 71:1–11
- Karataş R, Kasapoğlu AG, Güllap MK, İlhan E, Aydın M (2024) Genetic diversity analysis of lesser burnet (*Sanguisorba minor* Scop.) genotypes collected from Erzurum Province, Türkiye Using inter-simple sequence repeat (ISSR) markers. *Genet Resour Crop Evol* 72:1–11
- Khafagip O, Elsaiedp A, Hegazyp R, Moursy M (2018) Flora and plant communities of the eastern sector of Saint Katherine Protectorate, South Sinai, Egypt, pp 2348–7968
- Khodaei L, Azizinezhad R, Etminan AR, Khosroshahi M (2021) Assessment of genetic diversity among Iranian *Aegilops triuncialis* accessions using ISSR, SCoT, and CBDP markers. *J Genet Eng Biotechnol* 19:5
- Kongjaimun A, Takahashi Y, Yoshioka Y, Tomooka N, Mongkol R, Somta P (2022) Molecular analysis of genetic diversity and structure of the lablab (*Lablab purpureus* (L.) sweet) gene pool reveals two independent routes of domestication. *Plants* 12:57
- Li YL, Liu JX (2018) StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour* 18:176–177
- Li L, Ou W, Wang Y, Peng J, Wang D, Xu S (2022) Comparison of genetic diversity between ancient and common populations of *Docynia delavayi* (Franch.) Schneid. *Gene* 829:146498
- Liana D (2020) Phytochemical screening, in vitro antimalarial activity and genetic diversity of selected Asteraceae medicinal plants indigenous to Thailand. Thesis. Chulalongkorn Digital Collections. 0.58837/CHULA.THE.2020.425
- Marković M, Pljevljakušić D, Matejić J, Nikolić B, Smiljić M, Đelić G, Papović O, Đokić M, Stankov-Jovanović V (2022) The plants traditionally used for the treatment of respiratory infections in the Balkan Peninsula (Southeast Europe). *Lek Sirov* 42:68–88
- Marzouk RI, El-Darier SM, Kamal SA, Nour IH (2021) Comparative taxonomic study of Launaea Cass. (Asteraceae, Cichorioideae) in Egypt. *Taxonomy* 1:192–209
- Mohamed HM, El-Assal SE-DSM, Gamal El-Din A-KY, El-Khishin DA (2024) Analysis of genetic diversity and population structure in some Egyptian Berseem (*Trifolium alexandrinum*) accessions based on ISSR, SCoT and SRAP markers. *Genet Resour Crop Evol* 71:1–15
- Murray M, Thompson W (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8:4321–4326
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci* 70:3321–3323
- Omar MA, Saleh M, Mohamed GG, El Demerdash M (2020) Utilizing Start Codon Targeted (SCoT) markers for DNA fingerprinting and diversity analysis of *Artemisia herba-alba* plant species. *J Basic Environ Sci* 7:66–71

- Özbek Ö (2024) Molecular markers used to reveal genetic diversity and phylogenetic relationships in crop plants. *OBM Genet* 8:1–25
- Pantoja TR, Oliveira AM (2024) Asteraceae Family species used as medicinal plants in the treatment of respiratory diseases in the state of Amapá, Amazon–Brazil: a systematic review. *Rev Gestão Soc Ambient* 18:e03273
- Peng Y, Hayat F, Li J, Xiao J, Tu P, Chen J, Li C, Song W, Xuan W (2023) Genetic diversity analysis of 20 Abiu Germplasm based on SCoT molecular markers. *Appl Ecol Environ Res* 21:4295–4311
- Perrier X, Jacquemoud-Collet J (2009) DARwin software: dissimilarity analysis and representation for windows (version 5.0. 157).
- Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2:225–238
- Pradeep Reddy M, Sarla N, Siddiq E (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* 128:9–17
- Puneeth P, Lata S, Yadav R, Wankhede DP, Tomar B, Choudary H, Tomer A, Bidaramali V, Talukdar A (2023) Exploring the genetic diversity using CAAT box-derived polymorphism (CBDP) and start codon targeted (SCoT) markers in cultivated and wild species of okra (*Abelmoschus esculentus* (L.) Moench). *Genet Resour Crop Evol* 70:749–761
- Ravindran S, Ramar A, Balaraman SS, Sankar MP, Ahmed MA, Salama EA (2023) Testing the genetic variability of turmeric (*Curcuma longa* L.) genotypes through simple sequence repeats genomic-based microsatellite markers. *ACS Agric Sci Technol* 4:63–71
- Rolnik A, Olas B (2021) The plants of the Asteraceae family as agents in the protection of human health. *Int J Mol Sci* 22:3009
- Salama EA, Farid MA, El-Mahalawy YA, El-Akheder A, Abo-shosha AA, Fayed AM, Yehia W, Lamlom SF (2024) Exploring agro-morphological and fiber traits diversity in cotton (*G. barbadense* L.). *BMC Plant Biol* 24:403
- Samarina LS, Malyarovskaya VI, Reim S, Yakushina LG, Koninskaya NG, Klemeshova KV, Shkhalakhova RM, Mat-skiv AO, Shurkina ES, Gabueva TY (2021) Transferability of ISSR, SCoT and SSR markers for *Chrysanthemum* × *Morifolium* Ramat and genetic relationships among commercial Russian cultivars. *Plants* 10:1302
- Schlee D (1975) Numerical taxonomy. The principles and practice of numerical classification. *J Storage* 24(2):263–268
- Shaban AS, Arab S, Basuoni MM, Abozahra MS, Abdelkawy AM, Mohamed M (2022) SCoT, ISSR, and SDS-PAGE investigation of genetic diversity in several Egyptian wheat genotypes under normal and drought conditions. *Int J Agron* 2022:7024028
- Shaltout K, Heneidy S, Al-Sodany Y, Eid E, Hatim M, EE-G (2004) Evaluation of botanical conservation measures in St. Katherine Protectorate, Medicinal Plants Conservation Project. Egypt
- Shen F, Qin Y, Wang R, Huang X, Wang Y, Gao T, He J, Zhou Y, Jiao Y, Wei J (2023) Comparative genomics reveals a unique nitrogen-carbon balance system in Asteraceae. *Nat Commun* 14:4334
- Shooshtari L, Pour-Aboughadareh A, Etminan A, Ghorbanpour M, Bocianowski J (2024) Genetic diversity and population structure of Iranian oak (*Quercus* spp.) accessions based on ISSR and CBDP markers. *Genet Resour Crop Evol* 71:1–13
- Tawfik RS, El-Mouhamady ABA (2019) Molecular genetic studies on abiotic stress resistance in sorghum entries through using half diallel analysis and inter-simple sequence repeat (ISSR) markers. *Bull Natl Res Centre* 43:1–17
- Temraz T, Zedan H, Fouda M, Saber M, Salama W, Harhash K (2016) Egyptian biodiversity strategy and action plan (2015–2030). *Minist Environ* 1–83
- Tessier C, David J, This P, Boursiquot J-M, Charrier A (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *Theor Appl Genet* 98:171–177
- Wang L, Yang H, Xu G, Liu Z, Meng F, Shi L, Liu X, Zheng Y, Zhang G, Yang X (2024) Asteraceae genome database: a comprehensive platform for Asteraceae genomics. *Front Plant Sci* 15:1445365
- Wilson K, Walker J (2005) Principles and techniques of biochemistry and molecular biology, 6th edn. Cambridge University Press, Cambridge
- Xiong S, Zhao Y, Chen Y, Gao M, Wu L, Wang Y (2020) Genetic diversity and population structure of *Quercus fabri* Hance in China revealed by genotyping-by-sequencing. *Ecol Evol* 10:8949–8958
- Yirgu M, Kebede M, Feyissa T, Lakew B, Woldeyohannes AB, Fikere M (2023) Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare* L.) germplasm. *BMC Genom Data* 24:7
- Zhang Z, Lin Z, Wei M, Chen Z, Shen M, Cao G, Wang Y, Zhang Z, Zhang D (2022) Development of single nucleotide polymorphism and association analysis with growth traits for black porgy (*Acanthopagrus schlegelii*). *Genes* 13:1992

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.