*Plant Genetic Resources: Characterization and Utilization*

**cambridge.org/pgr**

# Assessment of molecular diversity in doubled haploid lines of camelina (*Camelina sativa* (L.) Crantz), as a new emerging oil crop

Mozafar Sadeghikian[1], Abdollah Najaphy[1], Danial Kahrizi[1] and Hossein Rostami Amadvandi[2]

[1]Department of Plant Production and Genetics, Faculty of Agricultural Sciences and Engineering, Razi University, Kermanshah, Iran and [2]Dryland Agricultural Research Institute, Sararood Branch, Agricultural Research, Education and Extension Organization (AREEO), Kermanshah, Iran

## Research Article

**Cite this article:** Sadeghikian M, Najaphy A, Kahrizi D, Rostami Amadvandi H (2022). Assessment of molecular diversity in doubled haploid lines of camelina (*Camelina sativa* (L.) Crantz), as a new emerging oil crop. *Plant Genetic Resources: Characterization and Utilization* **20**, 249–254. https://doi.org/10.1017/S1479262123000114

Received: 10 June 2021
Revised: 9 February 2023
Accepted: 10 February 2023
First published online: 8 March 2023

**Keywords:**
Camelina; doubled haploid lines; genetic diversity; ISSR markers

**Author for correspondence:**
Abdollah Najaphy, E-mail: anajaphy@razi.ac.ir

**Abstract**

Camelina (*Camelina sativa* (L.) Crantz), an oilseed crop, belongs to the Brassicaceae family. Two unique features of camelina in comparison with the main oil crops are an adaptation to different environments and also its unique oil composition. The development of doubled haploid plants is one of the essential methods for crop improvement. The study of genetic diversity is an important step in planning crop breeding programmes. This research was conducted to evaluate the genetic variation of 81 camelina doubled haploid lines obtained from 15 crosses by inter simple sequence repeat (ISSR) markers. The total number of amplified bands was 243, of which 239 bands (98.3%) showed polymorphism. The percentage of polymorphic bands varied between 93.75 and 100. The size of the bands ranged from 50 to 1700 base pairs. The informative ISSRs were identified by estimating marker features: polymorphism information content, effective multiplex ratio, marker index and resolving power. Three markers had higher resolving power values (9.88, 8.5 and 7.46) and were the most informative markers to identify the lines. Cluster analysis based on the complete algorithm divided the lines into five groups, indicating relatively clear configuration from the geographic distribution patterns of the parents of the doubled haploid lines. Principal coordinate analysis classified the 81 camelina doubled haploid lines into six groups. The ISSR markers detected high polymorphism to reveal the genetic variation of camelina lines. The findings of this research, along with the characterization of biochemical traits of the lines, can improve breeding programmes achieve high-yielding camelina varieties with higher and better oil content.

## Introduction

Vegetable oils are mainly obtained from oilseeds such as soybean, sunflower, cottonseed, peanut and rapeseed, which are in high demand for water. However, these common oilseeds, despite their many benefits, have their limitations in terms of cultivation and climatic conditions (Shonnard *et al.*, 2010).

Camelina (*Camelina sativa*) is from the Brassicaceae family and has been shown in many experiments to have much lower water requirements and greater resistance to spring cold than the other oil crops, especially rapeseed (McVay, 2008). Regarding its wide adaptability, camelina has been listed among the most promising new crops for oil production in temperate and semi-arid regions (Vollmann and Eynck, 2015). Camelina is a crop native to Europe and South Asia and may grow like a weed in some farms. Its cultivation history dates back to 4000 years ago. The Former Soviet Union was the largest producer of this crop in the 20th century with 300,000 hectares under cultivation in 1950 (Gehringer, 2010). Public interest in camelina has been re-emerged due to its level of omega-3 fatty acids, favourable agronomic characteristics and low-input potential as a biofuel crop (Ghamkhar *et al.*, 2010). Camelina oil possesses high omega-3 fatty acids, so has many properties and uses in nutrition and health that may prevent many diseases and disorders. In industry, it is used as biofuel, resins, waxes, as well as for the production of cosmetics and pharmaceutical products. In oil refineries, industrial antioxidants, which are extremely dangerous to human health, are added to prevent oxidation as well as to increase the duration of the oil. However, due to the high content of alpha-tocopherol and high vitamin E in camelina oil, which are strong antioxidants, camelina does not need any additives to last (Kahrizi *et al.*, 2015).

In total, 89,418 protein-coding genes have been identified in the camelina genome as a hexaploid plant ($2n = 40$, 750 Mb) (Kagale *et al.*, 2014). Unlike many crops, there was limited information about the genus *Camelina* and its biodiversity until 2005. The genome map of the camelina was drawn by Gehringer *et al.* (2006). Singh *et al.* (2015) developed a genomic map using 768 SNP markers and the Illumina GoldenGate SNP array method. Vollmann *et al.*

(2005) performed DNA analysis using RAPD markers to investigate the variation of camelina accessions in Australia and reported four distinct clusters. In that study, only 63% of the 30 markers studied were polymorphic. This indicates a relatively low diversity of the camelina accessions. Creating variation is essential for camelina improvement to consider positive characters and examine them in different environments (Vollmann *et al.*, 2005).

As mentioned, due to the polyploidy of the camelina genome, the classical breeding methods are partly complicated. On the other hand, cross-species hybridization in this plant has not been successful (Vollmann and Eynck, 2015). It seems that the production of double haploid plants isolated from the cultivation of anthers or microspores is a suitable method of creating diversity and improvement of this crop.

Analysis of genetic relationships in crops is a prerequisite for their breeding programmes, as it serves to provide information about genetic variation (Mohammadi and Prasanna, 2003). This research was conducted to assess the genetic diversity of 81 camelina doubled haploid (DH) lines using inter simple sequence repeat (ISSR) markers.

## Materials and methods

### Plant materials

In this study, 81 camelina DH lines were studied in terms of genetic diversity. Online supplementary Table S1 shows the list of these lines, their parents and origins. Camelina DH seeds were planted in October 2018 in the greenhouse of the College of Agricultural Sciences and Engineering, Razi University, Kermanshah, Iran.

### DNA purification

DNA extraction was performed in the laboratory of Zagros Bioidea Company located in Science and Technology Park, Kermanshah, Iran. Leaf samples of young leaves of potted seedlings grown in the greenhouse were prepared and immediately transferred to a freezer at −80°C in liquid nitrogen (−196°C). DNA extraction from leaf samples was performed using a kit made by Zagros Bioidea Company.

### ISSR amplification

Ten ISSR markers were used to detect genetic variation at the molecular level among all DH lines (Table 1). PCR amplification was done according to Najaphy *et al.* (2011) and Williams *et al.* (1990) in a BIOER XP CYCLER thermocycler. Agarose gel staining and visualization were conducted using ethidium bromide and a UV transilluminator, respectively.

### Statistical analysis

Amplified products were scored as (0) for the absence or (1) for the presence of each of the camelina DH lines. Jaccard's similarity coefficients were calculated based on the bivariate. Cluster analysis using the complete method and principal coordinate analysis (PCoA) were performed by NTSYS-pc software version 2.02.

The number of amplified and polymorphic bands, and the percentage of polymorphic bands (PPB) were measured for all of the markers. To determine the informative ISSRs to discriminate among camelina DH lines, polymorphism information content (PIC), effective multiplex ratio (EMR), marker index (MI) and resolving power (RP) were estimated. PIC was measured according to Mousapour Gorji *et al.* (2011), as $PIC = 1 - p^2 - q^2$, where $p$ is the frequency of the present band and $q$ is the frequency of the absent band. Diversity index (DI) was calculated as the average PIC value. EMR was recorded using the product of the number of polymorphic bands and the fraction of polymorphic bands (Kumar *et al.*, 2009). MI index was measured according to Powell *et al.* (1996) as the product of EMR and PIC. The formula $RP = \Sigma Ib$ was used to calculate RP, where Ib is band informativeness and $Ib = 1 - [2 \times (0.5 - p)]$, where $p$ is the proportion of the genotypes that contain the band (Altintas *et al.*, 2008).

## Results

### ISSR polymorphism and marker parameters

In this experiment, 10 ISSR markers were used individually. Marker parameters are summarized in Table 1. The total number of amplified bands was 243, of which 239 bands (98.3%) showed polymorphism. The character ranged between 16 (P4) and 33

**Table 1.** ISSR marker parameters for genetic diversity evaluation of camelina DH lines

| Primer | Sequence 5′→3′ | PPB (%) | No. of polymorphic bands | Total amplified bands | PIC | MI | EMR | RP |
|---|---|---|---|---|---|---|---|---|
| P1 | GAG $(CAA)_5$ | 100 | 20 | 20 | 0.34 | 6.8 | 20 | 4.8 |
| P4 | CTG $(GT)_8$ | 93.75 | 15 | 16 | 0.33 | 4.64 | 14.06 | 3.36 |
| P6 | $(AG)_8$GCC | 100 | 29 | 29 | 0.25 | 7.25 | 29 | 4.15 |
| P7 | $(AG)_8$GTG | 100 | 28 | 28 | 0.29 | 8.12 | 28 | 4.87 |
| P8 | $(GA)_8$ACC | 100 | 25 | 25 | 0.25 | 6.25 | 25 | 3.68 |
| P9 | $(GA)_8$ATC | 100 | 31 | 31 | 0.30 | 9.30 | 31 | 5.62 |
| P10 | $(ATG)_5$ | 94.73 | 18 | 19 | 0.48 | 8.18 | 17.05 | 7.46 |
| MM14 | (CT)8RA | 100 | 23 | 23 | 0.33 | 7.59 | 23 | 4.72 |
| MM15 | (AC)8YT | 96.96 | 32 | 33 | 0.42 | 13.03 | 31.03 | 9.88 |
| MM16 | (AC)8YA | 94.73 | 18 | 19 | 0.49 | 8.35 | 17.05 | 8.5 |
| | Total | – | 239 | 243 | – | – | – | – |
| | Mean | 98.3 | 23.9 | 24.3 | 0.35 | 7.95 | 23.52 | 5.63 |

PPB, percentage of polymorphic bands; PIC, polymorphism information content; MI, marker index; EMR, effective multiplex ratio; RP, resolving power.

https://doi.org/10.1017/S1479262123000114 Published online by Cambridge University Press

(MM15). MM15 and P9 primers possessed the highest number of total amplified products as 32 and 31, respectively. The PPB differed from 93.75% for P4 to 100% for primers P1, P6, P7, P8, P9 and MM14. The size of the bands varied between 50 and 1700 base pairs. The highest PIC was observed for MM16 primer (0.49) and P10 (0.48). PIC ranged between 0.25 and 0.49 with an average of 0.35.

MI ranged from 4.64 to 13.03. The highest MI parameter was determined for ISSR primers P9 and MM16 which amplified more bands (32 and 31, respectively). EMR index depends on

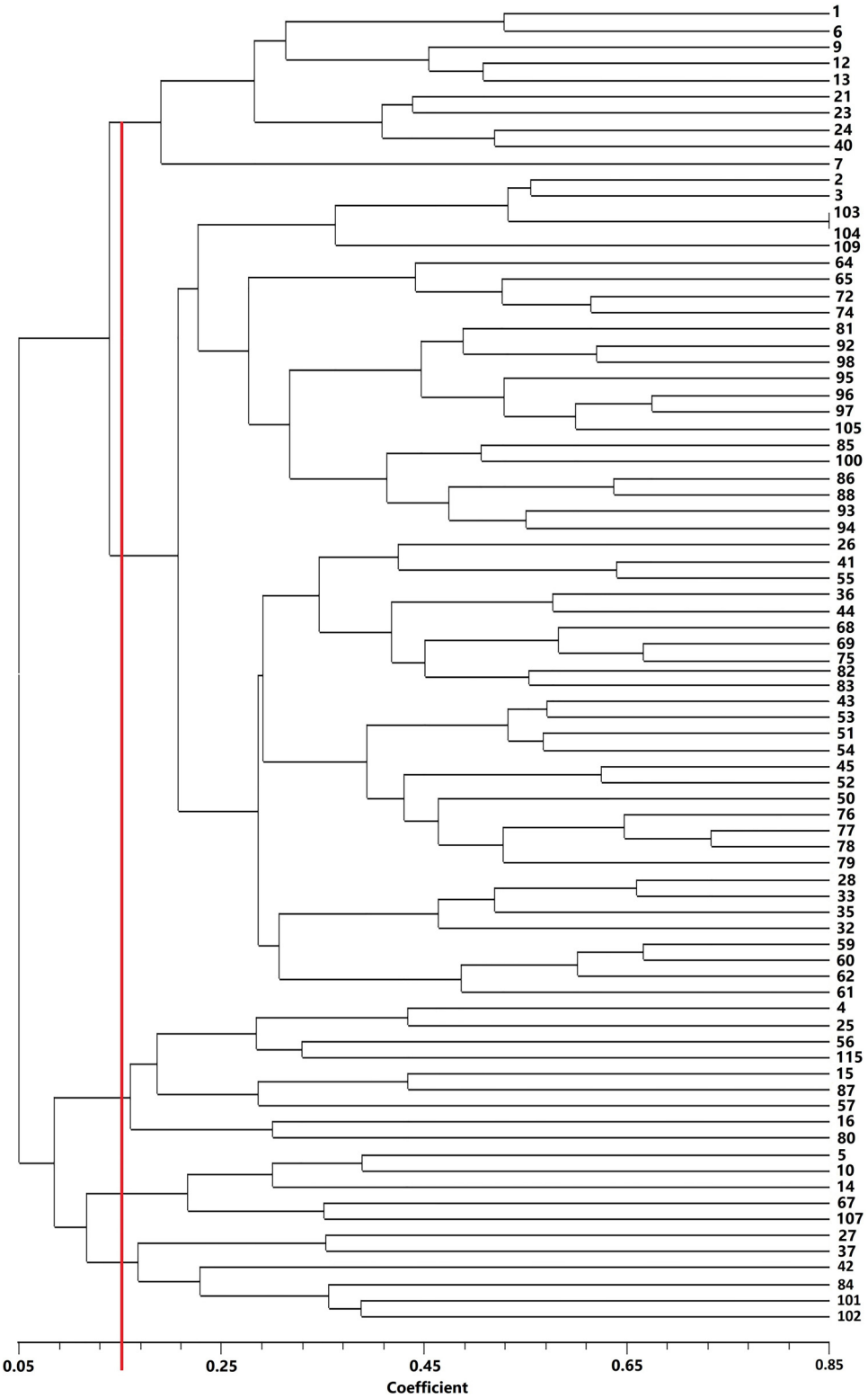


**Fig. 1.** Complete dendrogram of 81 camelina doubled haploid lines based on ISSR marker data.

https://doi.org/10.1017/S1479262123000114 Published online by Cambridge University Press

the PPB and the number of polymorphic bands. This feature marker varied from 14.06 to 31.03 with a mean of 23.52 (Table 1).

Three ISSR markers had higher RP values (9.88, 8.5 and 7.46 for MM15, MM16 and P10, respectively) and could be useful informative markers to identify DH lines (Table 1). This index varied from 3.36 to 9.88 with an average of 5.63 (Table 1).

### *Genetic relationships and cluster analysis*

Jaccard's similarity coefficients using ISSR data for camelina DH lines were calculated. Genetic similarity ranged between 0.0.05 and 0.85 (data not shown). The minimum similarity (0.05) was recorded between lines 27 and 45. These two DH lines were obtained from separate crosses with different parents (online Supplementary Table S1). DH lines 103 and 104 had the most similarity coefficient (0.85).

To investigate genetic relationships among camelina DH lines and to classify the lines, cluster analysis was performed using Jaccard's similarity coefficients matrix. Various clustering methods were tested and finally, complete method was selected as the most appropriate due to the lowest rate of chaining and a considerable cophenetic correlation coefficient of $r = 0.89$. The dendrogram grouped the camelina DH lines into five clusters (Fig. 1):

Group I: DH lines 1, 6, 9, 12, 13, 21, 23, 24, 40, 7.
Group II: DH lines 2, 3, 103, 104, 109, 64, 65, 72, 74, 81, 92, 98, 95, 96, 97, 105, 85, 100, 86, 88, 93, 94, 26, 41, 55, 36, 44, 68, 69, 75, 82, 83, 43, 53, 51, 54, 45, 52, 50, 76, 77, 78, 79, 28, 33, 35, 32, 59, 60, 62, 61.
Group III: DH lines 4, 25, 56, 115, 15, 87, 57, 16, 80.
Group IV: DH lines 5, 10, 14, 67, 107.
Group V: DH lines 27, 37, 42, 84, 101, 102.

### *Principal coordinate analysis based on ISSR markers*

PCoA was carried out using the marker system (Fig. 2). The Latent roots (Eigen values) calculated by ISSRs data showed that the first five principal coordinates accounted for 25.27, 3.61, 2.74, 2.34 and 2.10% of total variability, respectively (36.06% cumulatively). PCoA classified the 81 camelina DH lines into six distinct groups. Most of the lines that came together in the scatter plot were in the same cluster dendrogram (Fig. 2).

## Discussion

In the presented research, the genetic variation of camelina DH lines was assessed by ISSR markers. The set of 81 DH lines, obtained from 15 different crosses, possessed the parents from The Former Soviet Union and some parts of Europe (online Supplementary Table S1).

The efficiency of the used markers was assessed by calculating their PIC parameter. The high average polymorphism (98.3%) among camelina DH lines and mean PIC value of 0.35 (DI = 0.35) in our study indicates considerable genetic variation

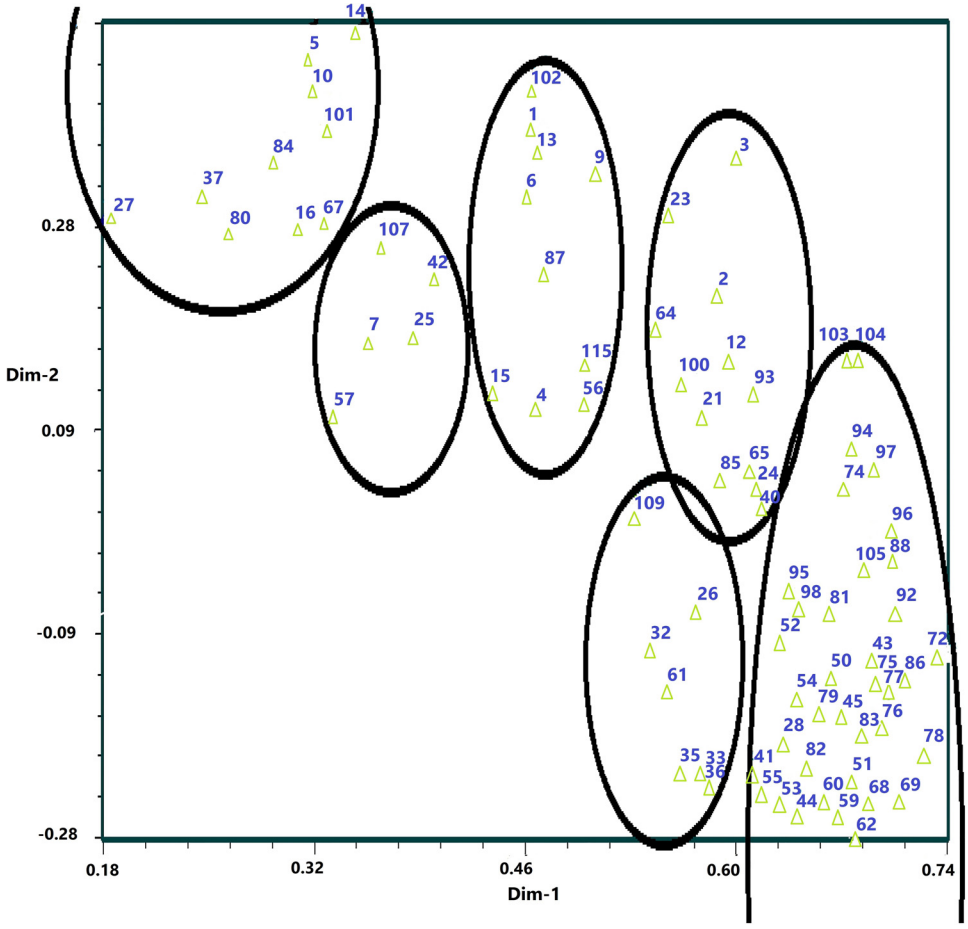


**Fig. 2.** Plot of camelina doubled haploid lines by principal coordinate analysis using the ISSR data.

https://doi.org/10.1017/S1479262123000114 Published online by Cambridge University Press

among the lines and high efficiency of the ISSR markers to distinguish the variation. An average PIC, i.e. DI = 0.29 indicating moderate genetic variation of *C. sativa* spring accessions using SNPs was reported by Luo *et al.* (2019). Najaphy *et al.* (2011) detected 80.2% polymorphism among 30 wheat accessions and reported an average PIC value of 0.22 using ISSR markers. Rostami-Ahmadvandi *et al.* (2013) recorded a mean PIC value of 0.39 for ISSR markers in cumin.

EMR based on the product of the number of polymorphic bands and the fraction of polymorphic bands, and MI index as the product of EMR and PIC were the two other parameters recorded for the marker system. The primers that produced more polymorphic bands had a higher EMR record. On the other hand, most of the primers with higher PIC showed higher MI values (Table 1). Different reports have been used on MI and EMR parameters to assess the discriminatory characteristic of molecular markers in some plants such as wheat (Najaphy *et al.*, 2011), thyme (Yousefi *et al.*, 2015) and apricot (Kumar *et al.*, 2009).

The ability of primers to identify the DH lines was also assessed by determining RP. This index ranged between 3.36 and 9.88 with a mean of 5.63 (Table 1). Three ISSR markers: M15, MM16 and P10 had higher RP values (9.88, 8.5 and 7.46, respectively) and could be useful informative markers to distinguish DH lines (Table 1). RP parameter presents a partly precise estimate of the number of accessions that can be distinguished using a marker. The index may provide no knowledge about the ability of a marker to show the phylogenetic relationships of the accessions (Prevost and Wilkinson, 1999). Some reports while calculating the ISSR parameters in the study of different crops genetic diversity introduced the resolution power and MI as indicator parameters for informative primers (Najaphy *et al.*, 2011; Yousefi *et al.*, 2015).

The genetic similarity matrix of camelina DH lines based on Jaccard's similarity coefficients using ISSR data was calculated (data not shown). The highest similarity coefficient (0.85) was found between lines 103 and 104. Interestingly, the two DH lines were also originated from different crosses by the parents from different countries. Line 103 had been obtained from the cross between Svalöf (originated from Sweden) and Ukrajinskij (from the Former Soviet Union). DH line 104 originated from a cross between Przybrodzka (from Poland) and Hoga (from Denmark).

Cluster analysis using the complete method depicted genetic relationships and classified the lines into five groups (Fig. 1). The clustering method indicated a relatively clear configuration from the geographic distribution patterns of the parents of the DH lines. Nevertheless, there is a mixture of DH lines with different parents, especially in group II (Fig. 1, online Supplementary Table S1). The phenomenon may be related to genetic interchange between populations cultivated in close proximity to one another in the Former Soviet Union and some parts of Europe. The relationships between geographic distribution and genetic variation have been shown in some other studies (Fracaro and Echeverrigaray, 2006; Agostini *et al.*, 2008; Najaphy *et al.*, 2011; Yousefi *et al.*, 2015; Luo *et al.*, 2019; Li *et al.*, 2021). Ghamkhar *et al.* (2010) reported a high genetic variation of 53 camelina genotypes from Ukraine.

By the analysis, PCoA revealed six distinct groups. The output of grouping DH lines obtained from PCoA was partly comparable and consistent with the cluster analysis (Fig. 2). It is noteworthy that most of the camelina DH lines that occupied positions together in the scatter plot of PCoA were classified in the same grouping dendrogram. Li *et al.* (2021) observed a clear pattern of genetic divergence in population structure of a worldwide collection of 222 camelina accessions and principal component analysis.

## Conclusion

DH is one of the applied mapping populations that can facilitate QTL mapping and molecular breeding programmes. The identification of genetic diversity and relationships among *C. sativa* DH lines is useful in breeding efforts. In our investigation, ISSR markers were employed to explore the genetic variation of *C. sativa* DH lines. The ISSRs generated high polymorphism for assessing genetic diversity and can be used in mapping QTLs and distinctive fingerprinting investigations of camelina accessions. Our results showed that the *C. sativa* DH lines were genetically diverse, and clustering of the lines according to the geographic distribution patterns of the parents of the DH lines was depicted. PIC, MI and RP are suggested as marker features to identify informative ISSRs. The molecular diversity evaluated in this research, along with the characterization of biochemical traits of the lines, can improve classical and molecular breeding programmes of the crop.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S1479262123000114.

## References

**Agostini G, Echeverrigaray S and Souza-Chies TT** (2008) Genetic relationships among South American species of Cunila D. Royenex L. based on ISSR. *Plant Systematics and Evolution* **274**, 135–141.

**Altıntas S, Toklu F, Kafkas S, Kilian B, Brandolini A and OZkan H** (2008) Estimating genetic diversity in durum and bread wheat cultivars from Turkey using AFLP and SAMPL markers. *Plant Breeding* **127**, 9–14.

**Fracaro F and Echeverrigaray S** (2006) Genetic variability in *Hesperozygis ringens* Benth. (Lamiaceae), an endangered aromatic and medicinal plant of Southern Brazil. *Biochemical Genetics* **44**, 479–490.

**Gehringer A** (2010) *Development of Camelina (Camelina sativa Crtz.) Genotypes and Winter Rapeseed (Brassica napus L.) Hybrids for Marginal Locations* (Doctoral dissertation). Justus Liebig University, Giessen.

**Gehringer A, Friedt W, Luhs W and Snowdon RJ** (2006) Genetic mapping of agronomic traits in false flax (*Camelina sativa* subsp. sativa). *Genome* **49**, 1555–1563.

**Ghamkhar K, Croser J, Aryamanesh N, Campbell M, Kon'kova N and Francis C** (2010) Camelina (*Camelina sativa* L.) Crantz as an alternative oilseed: molecular and ecogeographic analyses. *Genome* **53**, 558–567.

**Kagale S, Koh C, Nixon J, Bollina V, Clarke WE, Tuteja R, Spillane C, Robinson SJ, Links MG, Clarke C, Higgins EE, Huebert T, Sharpe AG and Parkin IA** (2014) The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nature Communications* **5**, 1–11.

**Kahrizi D, Rostami-Ahmadvandi H and Akbarabadi A** (2015) Feasibility cultivation of camelina (*Camelina sativa*) as medicinal-oil plant in rain fed conditions in Kermanshah-Iran's first report. *Journal of Medicinal Plants and By-products* **2**, 215–218.

**Kumar M, Mishra GP, Singh R, Kumar J, Naik PK and Singh SB** (2009) Correspondence of ISSR and RAPD markers for comparative analysis of genetic diversity among different apricot genotypes from cold arid deserts of trans-Himalayas. *Physiology and Molecular Biology of Plants* **15**, 225–236.

**Li H, Hu X, Lovell JT, Grabowski PP, Mamidi S, Chen C, Amirebrahimi M, Kahanda I, Mumey B, Barry K, Kudrna D, Schmutz J, Lachowiec J and Lu C** (2021) Genetic dissection of natural variation in oilseed traits of camelina by whole-genome resequencing and QTL mapping. *The Plant Genome* **14**, e20110.

https://doi.org/10.1017/S1479262123000114 Published online by Cambridge University Press

**Luo Z, Brock J, Dyer JM, Kutchan T, Schachtman D, Augustin M, Ge Y, Fahlgren N and Abdel-Haleem H** (2019) Genetic diversity and population structure of a *Camelina sativa* spring panel. *Frontiers in Plant Sciences* **10**, 1–12.

**McVay KA** (2008) Camelina production in Montana. http://agresearch.montana.edu› carc› Articles.

**Mohammadi SA and Prasanna BM** (2003) Analysis of genetic diversity in crop plants-salient statistical tools and considerations. *Crop Science* **43**, 1236–1248.

**Mousapour Gorji A, Poczai P, Polgar Z and Taller J** (2011) Efficiency of arbitrarily amplified dominant markers (SCOT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. *American Journal of Potato Research* **88**, 226–237.

**Najaphy A, Ashrafi Parchin R and Farshadfar E** (2011) Evaluation of genetic diversity in wheat cultivars and breeding lines using inter simple sequence repeat markers. *Biotechnology and Biotechnological Equipment* **4**, 2634–2638.

**Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalsky A** (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* **2**, 225–238.

**Prevost A and Wilkinson MJ** (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics* **98**, 107–112.

**Rostami-Ahmadvandi H, Cheghamirza K, Kahrizi D and Bahraminejad S** (2013) Comparison of morpho-agronomic traits versus RAPD and ISSR markers in order to evaluate genetic diversity among *Cuminum cyminum* L. accessions. *Australian Journal of Crop Sciences* **7**, 361–367.

**Shonnard DR, Williams L and Kalnes TN** (2010) Camelina-derived jet fuel and diesel: sustainable advanced biofuels. *Environmental Progress and Sustainable Energy* **29**, 382–392.

**Singh R, Bollina V, Higgins EE, Clarke WE, Eynck C, Sidebottom C, Gugel R, Snowdon R and Parkin IA** (2015) Single nucleotide polymorphism identification and genotyping in *Camelina sativa*. *Molecular Breeding* **35**, 1–13.

**Vollmann J and Eynck C** (2015) Camelina as a sustainable oilseed crop: contributions of plant breeding and genetic engineering. *Biotechnology Journal* **10**, 525–535.

**Vollmann J, Grausgruber H, Stift G, Dryzhyruk V and Lelley T** (2005) Genetic diversity in camelina germplasm as revealed by seed quality characteristics and RAPD polymorphism. *Plant Breeding* **124**, 446–453.

**Williams JG, Kubelik KAR, Livak KJ, Rafalski JA and Tingey SV** (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**, 6531–6553.

**Yousefi V, Najaphy A, Zebarjadi A and Safari H** (2015) Molecular characterization of thymus species using ISSR markers. *Journal of Animal and Plant Sciences* **25**, 1087–1094.

https://doi.org/10.1017/S1479262123000114 Published online by Cambridge University Press