Industrial Crops & Products 113 (2018) 111–120

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

Genetic structure identification and assessment of interrelationships between *Brassica* and allied genera using newly developed genic-SSRs of Indian Mustard (*Brassica juncea* L.)

Binay K. Singh[a,*], Shashi Bhushan Choudhary[b], Sushma Yadav[c], Era Vaidya Malhotra[d], Reema Rani[c], Supriya Ambawat[c], Priyamedha[c], Avinash Pandey[a], Rishikesh Kumar[a], Sudhir Kumar[a], Hariom Kumar Sharma[c], Devendra K. Singh[a], P.K. Rai[c]

[a] ICAR – Indian Institute of Agricultural Biotechnology, Ranchi 834 010, Jharkhand, India
[b] ICAR – National Bureau of Plant Genetic Resources, Regional Station, Ranchi 834 010, Jharkhand, India
[c] ICAR – Directorate of Rapeseed-Mustard Research, Bharatpur 321 303, Rajasthan, India
[d] ICAR – National Bureau of Plant Genetic Resources, New Delhi 110 012, India

ARTICLE INFO

Keywords:
Brassica
Cross-transferability
genic-SSR
Genetic variability
Genetic relationships

ABSTRACT

Brassica and allied genera are valuable sources of edible oils, vegetables, condiments and other products. They also find utility for specific applications as renewable industrial or fuel oils. Wild and weedy species of these genera are precious genetic resources of agronomic and economic traits. Scientists have developed a large number of genomic resources for cultivated species of *Brassica.* However, members of allied genera including wild and weedy species have remained mainly neglected. In this study, we tested the cross-transferability of 460 genic-SSRs, out of the 4108 genic-SSRs developed in our earlier study in *Brassica juncea*, to 21 accessions representing 19 species from eight different genera of the family *Brassicaceae*. A total of 200 genic-SSRs yielded single amplicons of which 150 exhibited cross-transferability. Cross-taxa allelic distribution analysis of the cross-transferable genic-SSRs highlighted preferential accumulation of 15 private alleles in cultivated Brassicas with distinct niche-specific roles. One hundred twenty-one genic-SSRs were found to be polymorphic with polymorphism information content values ranging from 0.09 to 0.66. The cross-transferable genic-SSRs were employed to identify population structure, Nei's genetic distance-based clustering, and principal coordinate analysis. The results convincingly demonstrated the efficacy of cross-transferable genic-SSRs in delineating taxonomic status of species of *Brassica* and allied genera and thus could be efficiently used in more extensive marker-based studies.

1. Introduction

The family *Brassicaceae* includes 338 genera and 3709 species (Al-Shehbaz et al., 2006). Among them, the genus *Brassica* is economically very important. It comprises of 37 species (Gómez-Campo, 1980) out of which humans have extensively cultivated *Brassica napus*, *Brassica juncea*, *Brassica carinata*, *Brassica rapa*, *Brassica nigra* and *Brassica oleracea*. *Brassica napus* and *Brassica rapa* are cultivated mainly for canola oil while *Brassica juncea*, *Brassica carinata*, and *Brassica nigra* are grown for the production of mustard oil and commercial spice (Labana and Gupta, 1993; Priyamedha et al., 2015). *Brassica oleracea* which is the most diverse species of the genus *Brassica* is cultivated for the production of six distinct vegetables, collectively known as cole crops (Balkaya et al., 2005). Besides *Brassica*, some of the other important genera belonging to the family *Brassicaceae* are *Camelina*, *Crambe*, *Lepidium*, and *Sinapis*. Several of the species belonging to these genera have already established their potential as new edible oil/protein crops, medicinal crops, biodiesel fuel crops, or platforms for bio-products or molecular farming (Gugel and Falk, 2006; Warwick et al., 2007; Singh et al., 2014).

Ironically, modern agricultural methods and technologies have drastically compromised the genetic diversity of crops leading to their enhanced vulnerability to ever-expanding biotic and abiotic threats. Growing threats of climate change have further accentuated the

Abbreviations: SSR, simple sequence repeat; CTAB, cetyl trimethyl ammonium bromide; PCR, polymerase chain reaction; PIC, polymorphism information content; PCoA, principal coordinate analysis; EST, expressed sequence tag

* Corresponding author.
E-mail addresses: binaybio@gmail.com, binay_kumar123@rediffmail.com (B.K. Singh).

https://doi.org/10.1016/j.indcrop.2018.01.023
Received 8 September 2017; Received in revised form 4 January 2018; Accepted 10 January 2018

problem (Cowling, 2007; Singh et al., 2017). Nevertheless, the genus *Brassica* has been bestowed with nearly 100 species and genera of wild and weedy relatives which are invaluable genetic resources of agronomic and economic traits (Warwick et al., 2009). To reap out benefits from these genetic resources, understanding the structure and genetic diversity spectrum are a prerequisite. In this context, molecular markers are useful tools, and these have greatly enhanced the genetic analysis of crop plants in recent years (Varshney et al., 2005). Among the variety of molecular marker techniques available, plant breeders widely use SSR markers for molecular-based breeding. Gene-based SSR markers are particularly useful for this purpose since SSRs located in transcribed regions have the potential to become functional markers in genic regions (Varshney et al., 2005). In an earlier study, we tested the transferability of 161 genomic-SSR markers derived from *Brassica napus*, *Brassica nigra*, *Brassica rapa* and *Brassica oleracea* to eleven different species of *Brassica* and allied genera (Singh et al., 2012a). The results were not very encouraging, as, among the 161 genomic-SSR markers, only 70 (43.5%) showed transferability to at least one of the eleven species included in the study. Genic-SSR markers are likely to show a higher level of transferability among species and genera since they are derived from the transcribed region of the genome that is usually conserved between related species (Varshney et al., 2005).

Keeping these in mind, we developed a set of 4108 genic-SSRs in *Brassica juncea* in one of our earlier studies (Singh et al., 2016). In the present study, we used 339 genic-SSRs, out of the 4108 genic-SSRs, for cross-species amplification in *Brassica rapa*, *Brassica nigra*, *Brassica napus*, *Brassica carinata*, *Brassica tournefortii*, *Brassica fruticulosa*, *Brassica spinescens*, *Sinapis alba*, *Crambe abyssinica*, *Diplotaxis assurgens*, *Diplotaxis siettiana*, *Diplotaxis tenuisiliqua*, *Diplotaxis muralis*, *Diplotaxis gomez-campoi*, *Capsella bursa-pastoris*, *Enarthrocarpus lyratus*, *Camelina sativa* and *Lepidium sativum,* representing 18 species from eight different genera of the family *Brassicaceae*. In addition to these, a Chinese cultivar of *Brassica juncea* was also included in the study. To assess the efficacy of cross-transferable genic-SSRs, we used them to evaluate the genetic diversity and establish inter-specific as well as inter-generic genetic relationships among the studied taxa. The cross-transferable genic-SSRs successfully distinguished the studied taxa and grouped them as per their recognized taxonomy. The markers identified in the study could be further used for cross-genome comparisons and subsequently map-based prediction of the location of candidate genes linked with agriculturally essential traits within different species of the family *Brassicaceae*. Besides, it would also pave the way for more extensive studies on transferability of genic-SSR markers for wild or lesser known species of *Brassica* and allied genera.

2. Materials and methods

2.1. Plant materials

Twenty-one accessions belonging to 19 species from eight different genera of the family *Brassicaceae* were analyzed in this study. We obtained the seed samples of cultivated *Brassica* species from Germplasm Unit, ICAR – Directorate of Rapeseed-Mustard Research, Rajasthan (India), and the seeds of wild species from ICAR – National Research Centre on Plant Biotechnology, New Delhi (India). Fig. 1 depicts the photographs of selected taxa. Supplementary Table 1 summarizes the details of the plant materials used in the study.

2.2. Identification of genic-SSRs for cross-transferability study

Out of the total of 4108 genic-SSRs detected in the unigene contigs developed from the expressed sequences of *Brassica juncea* cv. CS-52 (Singh et al., 2016), this study evaluated the transferability of 460 genic-SSR loci. The study included only those SSRs which contained simple sequences of 2–6 nucleotides reiterated at least five times. We did not consider complex SSRs for the study. Primer sets for all the 460 genic-SSRs were custom synthesized (Eurofins Genomics, India) by using sequence information reported by Singh et al. (2016).

2.3. DNA extraction and cross-species PCR amplification of genic-SSRs

Total genomic DNA was extracted from fresh and healthy leaves from young plants using the CTAB method of Murray and Thompson (1980). The quality of DNA was checked by electrophoresis on 1.0% agarose gel. The primer sets of the 460 genic-SSRs were tested for PCR amplification using genomic DNA extracted from leaf tissues of *Brassica juncea* cv. NRCDR-2. The primer pairs which yielded clear and unambiguous amplicons were examined for cross-transferability and polymorphism among other cultivated and wild species of *Brassica* and allied genera. We performed all the PCR reactions in 10-µl reaction mixtures consisting of 1 × PCR buffer, 1.5 µM $MgCl_2$, 200 µM dNTPs each, 250 nM of each primer, 0.25 U of *Taq* DNA polymerase and 20 ng of genomic DNA. All the PCR reagents used in the study were procured from MBI, Fermentas, USA. The PCR thermal profile was set as: initial denaturation at 94 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 45 s, primer annealing at 60 °C for 30 s, extension at 72 °C for 30 s and a final extension at 72 °C for 7 min. The PCR products were visualized in 3.5% MetaPhor (FMC BioProducts, Rockland, ME, USA) agarose gels containing 0.5 ng/ml of ethidium bromide (Sigma-Aldrich, St. Louis, USA).

2.4. Scoring and data analysis

The amplicons of genic-SSRs obtained from each accession were resolved on the MetaPhor agarose gel system. The molecular weights of DNA bands were estimated by comparing with 100 bp DNA ladder. The PIC for each genic-SSR was calculated using PowerMarker software (Liu, 2003). GenAlEx v. 6.1 software was used to calculate pair-wise Nei's genetic distance and private alleles. The neighbour-joining dendrogram (cluster analysis) based on Nei's genetic distance for genic-SSRs was generated using MEGA 6 software.

Bayesian model-based clustering algorithm of STRUCTURE 2.3 software was employed for the molecular marker based population structure analysis (Pritchard et al., 2000, 2010). The software was run using putative population origin with parameters as followed by Sharma et al. (2016) and Choudhary et al. (2017). Based on ΔK method of Evanno et al. (2005), the optimum value of K was determined in Structure Harvester online software (Earl and vonHoldt, 2012; http://taylor0.biology.ucla.edu/structureHarvester/) by using STRUCTURE 2.3 software output as input. Further, STRUCTURE 2.3 software was run at desired K values for 100,000 burn-in periods, and 100,000 MCMC replicates. The structure diagram was prepared using the same software. Principal coordinate analysis based on allele frequencies was conducted by using XLSTAT statistical software (Addinsoft, 2007).

3. Results and discussion

3.1. PCR amplification, cross-taxa transferability, and polymorphism analysis

Genic-SSRs represent expressed regions of the genome. These SSRs are highly useful for analyzing the functional diversity in germplasm collections. Besides, they can be useful in marker-assisted breeding, comparative mapping and evolutionary studies. Genic-SSR markers are transferable to related taxa. Typically, cross-taxa transferability shows an inverse relationship with the evolutionary distance (Dube et al., 2017). In this study, we evaluated a set of 460 genic-SSRs identified in *Brassica juncea* cv. CS-52 for its utility across other cultivated and wild relatives of *Brassica* and allied genera. Of the 460 primer pairs, preliminarily tested for amplification of genomic DNA prepared from *Brassica juncea* cv. NRCDR-2, 339 generated scorable amplicons. However, only 200 of these primer pairs yielded single fragments. We