






Microsatellite Markers: A Tool to Assess the Genetic Diversity of Yellow Mustard (*Sinapis alba* L.)

Eva Jozová ^{1,*} , Michael Rost ¹ , Andrea Rychlá ², Dagmar Stehlíková ¹ , Baveesh Pudhuvai ¹, Ondřej Hejna ¹, Pavel Beran ¹ , Vladislav Čurn ¹  and Miroslav Klíma ³

¹ Department of Genetics and Biotechnology, Faculty of Agriculture and Technology, University of South Bohemia in Ceske Budejovice, 370 05 Ceske Budejovice, Czech Republic; rost@fzt.jcu.cz (M.R.); stehld@fzt.jcu.cz (D.S.); bpodhuvai@fzt.jcu.cz (B.P.); hejna@fzt.jcu.cz (O.H.); beranp@fzt.jcu.cz (P.B.); curn@fzt.jcu.cz (V.Č.)

² OSEVA Development and Research, 746 01 Opava, Czech Republic; rychla@oseva.cz

³ Crop Research Institute, 161 06 Prague, Czech Republic; klima@vurv.cz

* Correspondence: jozova@fzt.jcu.cz

Abstract: Microsatellite markers were used for the assessment of genetic diversity and genetic structure in a germplasm collection of yellow mustard, *Sinapis alba* L. The comprehensive collection of genetic resources represented 187 registered varieties, landraces, and breeding materials. Microsatellites generated 44 polymorphic alleles in 15 loci. Eleven of them were medium to highly polymorphic, and the high levels of observed heterozygosity (0.12–0.83) and Nei's gene diversity index (0.11–0.68) indicated a high level of polymorphism. Based on PCoA and neighbor joining analyses, the genetic resources were divided into two groups. The range of genetic dissimilarity in the analysed collection was in the range of 0.00–1.00. The high level of dissimilarity between the accessions was documented by the high WAM value (33.82%). Bayesian clustering algorithms were performed in the STRUCTURE 2.3.4 software. The number of clusters was estimated at K = 2. The accessions were classified according to Q1/Q2 values. The low average values of the parameters Fst_1 (0.3482), Fst_2 (0.1916), and parameter alpha (0.0602) indicated substantial mating barriers between varieties and reproductive isolation due to the limited exchange of genetic resources between breeders. These results demonstrated the importance of extensive collections of genetic resources for the maintenance of genetic diversity and indicated considerable genetic differentiation among accessions.

Keywords: molecular markers; SSR; breeding; genetic structure; induced mutagenesis



Citation: Jozová, E.; Rost, M.; Rychlá, A.; Stehlíková, D.; Pudhuvai, B.; Hejna, O.; Beran, P.; Čurn, V.; Klíma, M. Microsatellite Markers: A Tool to Assess the Genetic Diversity of Yellow Mustard (*Sinapis alba* L.). *Plants* **2023**, *12*, 4026. <https://doi.org/10.3390/plants12234026>

Academic Editor: Shri Mohan Jain

Received: 19 October 2023

Revised: 23 November 2023

Accepted: 28 November 2023

Published: 29 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Yellow mustard (*Sinapis alba* L.) is a vital crop for Czech agriculture's food and condiment sectors. It originated in the Middle East and Mediterranean regions. Still, this crop has been produced worldwide for millennia as a condiment crop. It is extensively planted as a prominent speciality crop [1], and the seeds and volatile oils are used in traditional Chinese medicine [2]. This crop is drought-tolerant and resistant to major diseases and insect pests [3,4]. Yellow mustard also has phytosanitary properties and can potentially reduce harmful nematodes in soils, and the soil disinfestation effects of biofumigant compounds released from yellow mustard green manure have been reported [5]. Complete data from the literature and original data on *S. alba* exposed to elements such as metals from the thallium, arsenic, and platinum groups were collected, ensuring crop application for phytoremediation in contaminated areas [6]. The Czech Republic is at the forefront of this crop production, due to ideal climatic conditions and a long heritage in cultivation. Together with Canada and Ukraine, the Czech Republic is one of the world's largest producers and exporters of this crop. Yellow mustard is cultivated mainly for seeds and non-production purposes, such as an ameliorating catch crop for green manuring or as a significant component of biobelts. Mustard is also an important bee nectar and pollen-producing plant [7].

The range of genetic diversity, changes in genetic structure over the timespan of breeding in landraces, cultivated varieties, and genetic resources are still little understood, even though yellow mustard, like rapeseed, is a vital crop [4,8]. There is a widespread perception that the genetic variability is declining, due to breeders' emphasis on production and product/seed quality, compared to other intensively cultivated crops [9]. Morphological traits, previously the only tool for selection and cultivar identification, are no longer sufficient to identify emerging varieties and define their phenotypic variance, due to the influences of many environmental factors. This is also reflected in the often problematic use of DUS (Distinctness, Uniformity and Stability) tests for the description and legal protection of varieties [10].

The unremitting reduction in the number of cultivated species and varieties and their widespread loss has led to genetic erosion, the partial or total loss of landrace varieties and a narrowing of genetic diversity. The effects of a lack of genetic diversity can be seen in the loss of important genes and genetic vulnerability, as well as in the reduction in genetic gain in quantitative traits, which is difficult to overcome [9–11]. The breeders then have then sought to expand existing genetic diversity, e.g., by using landrace varieties, by distant hybridisation with taxonomically more distant species, and by applying in vitro techniques and targeted mutagenesis [12]. In practical breeding, the use of classical mutagenesis is limited, and only a few new varieties have been bred based on this approach. Because of the difficult detection of mutations and the not-quite-successful use of induced mutagenesis in plant breeding, a molecular or reverse genetics approach has been adopted since the beginning of the millennium through the TILLING (targeting induced local lesions in genome) method [13]. A more modern mutant detection procedure involves amplifying genes of interest, sequencing amplicons, and identifying point mutations [14].

Plant genetic resources should be characterised by employing molecular markers to reveal genotypic variations and unambiguously identify crop varieties. Molecular methods have been essential components of most research on genetic diversity over the last few decades, and molecular markers recognise DNA polymorphism between varieties and aid in gaining a better understanding of the genetic hierarchy of the parent components [15]. Among the most used types of markers to assess genetic diversity or seed purity are, for example, RAPDs (Random Amplified Polymorphic DNA) [16], which are no longer widely used due to poor repeatability of results, AFLP (Amplified Fragment Length Polymorphism) [17], ISSR (Inter Simple Sequence Repeats) [18], SSR (Simple Sequence Repeats) [4,19,20], or SNP (Single Nucleotide Polymorphism) [21]. However, it is essential to recognise that different markers have distinct attributes/properties that reflect various components of genetic variation and, therefore, can produce different results [22]. Although yellow mustard is an important crop, the number of studies devoted to this species is quite limited. A detailed study of yellow mustard accessions, maintained at Plant Gene Resources of Canada, was performed using AFLP markers [17]. The same collection of genetic resources was recently analysed using a modern approach to genotyping by sequencing, and new sets of SNPs were discovered and evaluated [21]. More attention was focused on coloured mustard, *B. juncea*, *B. carinata*, and *B. nigra*. An SSR marker was used in *B. juncea* [23] and *B. nigra* [24]. In *B. juncea*, the influence of natural outcrossing and cross-hybridization for the development of analysed varieties was demonstrated, based on the analysis of the genetic structure. Unique SSR alleles were used for specific DNA fingerprinting for the identification and legal protection of *B. juncea* varieties. Genome-specific SSR markers in *B. nigra* were used for assessment of the correct taxonomic classification of accessions. Based on multidimensional scaling analyses, genotypes were associated with country or region of origin. In *B. carinata* [25], the genetic diversity and population structure of genotypes of different origins was analysed using high-throughput SNP markers. The analysed genotypes were divided into two clusters that formed landraces/ecotypes, respectively breeding lines and varieties. This indicated a change in genetic structure due to breeding. The most comprehensive marker system used for the assessment of genetic diversity are microsatellites (SSRs, simple sequence repeats), as they have many desirable

properties, such as abundant genome distribution, the ability to form multiplexes, and codominance, i.e., the ability to distinguish between homozygotes and heterozygotes [26]. Molecular approaches improve germplasm conservation efforts and increase the use of plant genetic resources [27]. Furthermore, combining morphological and molecular markers is an excellent approach for selecting optimal genotypes for breeding, while maintaining adequate genetic diversity [28], and knowledge of genetic diversity is the basis for the effective management and utilisation of plant germplasms.

This study aimed to investigate the genetic structure of the yellow mustard core collection and to evaluate the genetic diversity among the yellow mustard accessions. A thorough investigation of the molecular properties of this diversity yields a group of the most diverse genotypes for use in breeding programmes. Marker-assisted selection using polymorphic markers among parents helps with the finding and choosing of suitable parents for hybridisation and backcross breeding for future use.

2. Results and Discussion

2.1. Microsatellite Polymorphism and Genetic Diversity of Genetic Resources

The wide collection of genetic resources for yellow mustard (*S. alba*) was represented by 198 unique accessions (registered varieties, landraces, and breeding materials). Microsatellites generated 113 alleles and, of the total alleles, 44 were polymorphic and detected in 15 loci, with an average of 2.93 alleles per locus and a maximum of six alleles in D3_A, P7, and P35_A loci. Five loci (BoREM1b_B, BoREM1b_C, BolAB19TF, P30_C, and P35_B) were classified as uninformative loci with MAF < 0.01. Only one allele was detected in these loci of one to three individuals/accessions. BoREM1b_B was successfully amplified in genotype 9 (Morocka, landrace) and BoREM1b_C in genotypes 51 (SIN 10/79, breeding material), 167, and 177 (Asta and Raduga, new registered varieties). BolAB19TF was amplified in genotype 130 (Sabon, new registered variety), P30_C in genotypes 169 and 186 (Rumba and Olga, new registered variety), and P35_B in genotype 174 (Bamberka, new registered variety). This study is part of a project focused on yellow and coloured mustard, and only the markers successfully amplified in yellow mustard are presented in Table 1. Samples with the occurrence of uninformative loci were also included in these results, because these genotypes with the presence of even this one allele differ from the set of other varieties. These unique/less common alleles are typical for a particular variety and lead to higher values of genetic diversity (see Figure S1 and WAM value). It is also noticeable that these less frequent alleles, with the exception of the Marocka landrace, occur in newly bred varieties and breeding materials. The size of the amplified alleles was in the range of 125–314 bp. Eleven loci were medium to highly polymorphic, with PIC values varying from 0.19 to 0.65. The highest PIC values were observed in loci D3_A (0.65) and P35_A (0.52), and also in loci BoPC34 (0.51) and P381 (0.50), characterised by a low number of alleles per locus. The level of observed heterozygosity was in the range of 0.12 to 0.83, with an average of 0.21. The highest observed heterozygosity was also recorded in loci with a high polymorphic index, namely D3_A (0.83), BoP34 (0.73), and P35_A (0.53). Nei's gene diversity index ranged between 0.11 and 0.68, with an average of 0.25. Such a high level of polymorphism, compared to other oil seed Brassicas, may be explained by a higher level of cross-pollination in yellow mustard [29]. Detailed results of the microsatellite analysis are given in Table 1.

Table 1. Detailed results of microsatellite analyses.

Microsatellite Marker	Number of Individuals	PIC	Na (Number of Alleles)	Number of Alleles in Population	Length of Alleles in Bp	Median and Max. Frequency of Alleles	Ho (Observed Heterozygosity)	He (Nei, 1973; Expected Heterozygosity)	Fis (Inbreeding Coefficient)	1-D (Simpson Index)	Hexp (Nei, 1978; Gene Diversity)	Evenness
BoREM1b_A	198	0.19	2	209	173–175	104.5; 187	0.12	0.11	−0.06	0.11	0.11	0.50
BoREM1b_B	198	0.00	1	3	181	3; 3	0	0	-	-	-	-
BoREM1b_C	198	0.00	1	3	187	3; 3	0	0	-	-	-	-
BoABI91TF	198	0.00	1	1	314	1; 1	0	0	-	-	-	-
BoPC34	198	0.51	3	323	149–151	148; 172	0.73	0.50	−0.45	0.50	0.50	0.94
P381	198	0.50	2	21	213–229	105; 11	0.05	0.50	0.90	0.50	0.51	1.00
D3_A	198	0.65	6	342	155–163	33.5; 149	0.83	0.67	−0.24	0.67	0.67	0.85
D3_B	198	0.44	5	101	167–171	10; 73	0.20	0.39	0.50	0.38	0.68	0.54
P7	198	0.27	6	119	141–156	3.5; 101	0.11	0.21	0.49	0.21	0.21	0.39
P9_B	198	0.38	3	184	125–138	32; 141	0.28	0.29	0.05	0.27	0.27	0.54
P9_C	198	0.35	4	231	145–160	25; 184	0.25	0.23	−0.06	0.23	0.24	0.49
P30_B	198	0.32	2	5	166–174	2.5; 4	0	0.32	1.00	0.32	0.36	0.72
P30_C	198	0.00	1	2	192	2; 2	0	0	-	-	-	-
P35_A	198	0.52	6	109	149–161	3.5; 68	0.53	0.47	−0.12	0.46	0.47	0.63
P35_B	198	0.00	1	1	173	1; 1	0	0	-	-	-	-

Allele sizes calculated from CE electropherograms were identical for the internal controls and stable in all repetitions. This uniformity and reproducibility for all microsatellites were indicators of correctly set protocols of molecular analyses. The microsatellites used in this study were originally designed for *Brassica oleracea* [30], *B. rapa* ssp. *chinensis*, *B. juncea*, and *B. nigra* genetic resources [31]. Our results showed very good cross-transferability and possible utilisation of these markers in yellow mustard. Successful amplification of specific fragments from the more distantly related *Brassica* species, including *S. alba* using primers designed for *B. napus*, were also reported [19]. Still, they did not analyse the diversity of microsatellites in yellow mustard genetic resources in more detail. Therefore, this study presents the first results of the analysis of microsatellites in a large set of genetic sources of yellow mustard and shows relatively high variability in the genetic resources, as well as the cross-transferability of markers and the occurrence of unique alleles.

The results of the AMOVA (Analysis of molecular variance) showed that the highest proportion of diversity was found within the analysed set of genetic resources (87.13%) and that the proportion of diversity and differentiation between types of genetic resources (4.6%) or between genotypes within types of genetic resources (8.27%), although significant, was always much lower (Table 2). Likewise, the null hypotheses were rejected (for parameters Φ_{CT} , Φ_{SC} — $p < 0.001$, and Φ_{ST} — $p < 0.005$). Thus, there are genetic differences between types of genetic resources, as well as between individual accessions of yellow mustard within the analysed set of genetic resources.

Table 2. Analysis of molecular variance (AMOVA) results for yellow mustard genetic resources (between and within types of genetic resource) with df, sums of squares (SS), mean squares (MS), genetic variance estimates (Sigma), proportion of variance (%) attributed to the different levels in the spatial hierarchy, Phi (Φ) statistics that are analogous to Wright's F-statistic, and p values (9999 permutations).

	df	SS	MS	Genetic Variance Estimates (Sigma)	Proportion of Variance (%)	Phi (Φ)	p -Values
Between types of genetic resources	2	19.41	9.705	0.143	4.602	$\Phi_{CT} = 0.0460$	<0.0001
Between genotypes within types of genetic resources	31	123.07	3.970	0.257	8.272	$\Phi_{SC} = 0.0867$	<0.0001
Within genetic resources	164	414.12	2.707	2.707	87.126	$\Phi_{ST} = 0.1287$	0.0028
Total	197	556.60	2.992	3.107	100		

Type of genetic resource = landrace/variety/breeding material.

The genetic diversity of the yellow mustard germplasms was calculated using DARwin 6 software. The dendrogram in Figure 1a shows the results of the weighted neighbor joining analysis, and, in Figure 1b, the results of the principal coordinates analysis. The extent of genetic variation in the analysed collection was in the range of 0.00–1.00 (Figure S1). The furthest distances were between accessions ‘47-635-YE 13’ and ‘Braco’. Additionally, the seven other most different/distant pairs of accessions always included the Chinese ‘47-635-YE 13’ genotype. Both the genetic diversity and dissimilarity range were much higher than the values reported by Fu et al. [17], given the more considerable extent of the analysed collection and the marker system used. The high level of dissimilarity between the accessions analysed was documented by the WAM value, which was 33.82%. A higher WAM value corresponded with the pattern of genotype distribution after factorial (PCoA) analysis (Figure 1b). In our previous study, focused on *Brassica napus* [32], we observed WAM values that were significantly lower (22.50% for SSR markers). These results showed the importance of extensive collections of genetic resources for preserving genetic diversity. The results of other studies conducted on mustard showed significantly lower values of genetic diversity [33], probably due to the low diversity of the collection and the small number of samples, which was incomparably lower than in this study.

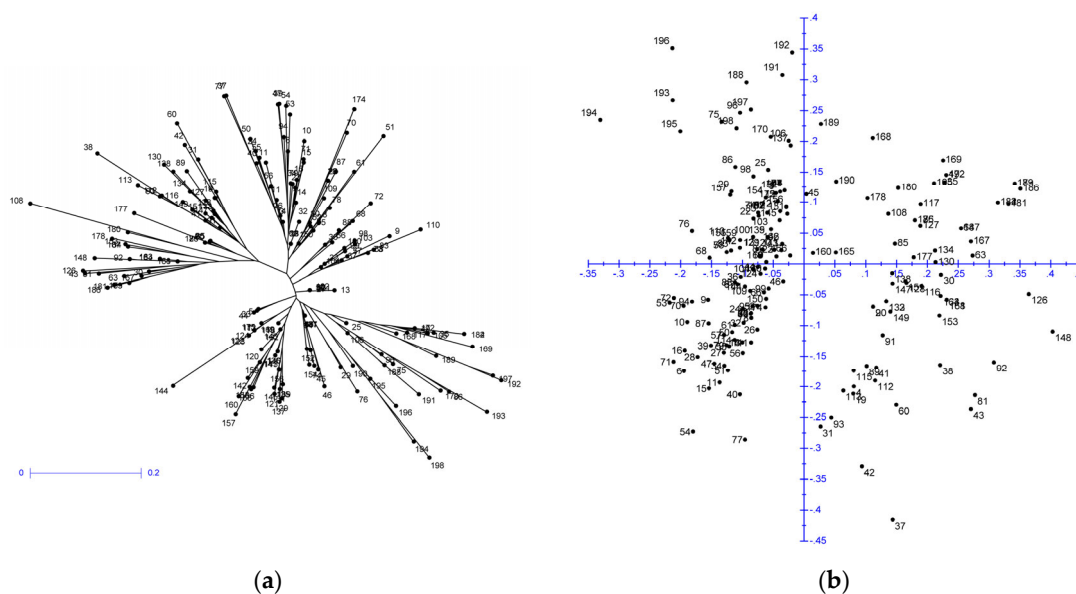


Figure 1. Outputs of statistical analyses of SSR markers in 198 *S. alba* accessions: (a) results of cluster analysis (weighted N J analysis); (b) results of principal coordinates analysis (PCoA).

Based on microsatellite analysis, it was possible to identify a large proportion of the analysed accessions unambiguously, but even when a relatively high number of markers were evaluated, several accessions of different origins had precisely the same patterns of microsatellite markers. There were two main groups of accessions with the same pattern of microsatellites. The first group included genotypes from Germany, Denmark, Netherlands, and Poland, but also genotypes from Australia and an old landrace from the Czech Republic/former Czechoslovakia (‘ABA’, ‘Borowska’, ‘BRSCHW 55705’, ‘Česká krajová’, ‘Perovska’, ‘Sito’, and ‘King’). The second block of genotypes with the same pattern of microsatellites mainly included Czech varieties (‘Andromeda’ and ‘Severka’), as well as varieties from Germany, Austria, and the Netherlands (‘Ascot’, ‘Emergo’, and ‘Mirly’). There were also other groups or pairs of indistinguishable accessions, such as ‘OP-2408’ and ‘OP-2415’, or ‘Olga’ and ‘OP-2291’, originating in one breeding station, but also cases have been reported where landraces have the same microsatellite profile as modern varieties, e.g., ‘Přerovská bílá’ vs. ‘Ludique’ and ‘Budakalászi sárga’, or ‘Dr. Schneider’ vs. ‘Concerta’ and ‘Espirit’. A relatively narrow genetic base should explain these results, and unrelated genotypes with an identical pattern of microsatellites are used in different breed-

ing programmes. These findings also reveal the need for better management of breeding populations based on genetic analysis, and further improvement is also increasingly dependent on the evaluation of genetic diversity resources and breeding populations. Breeding materials from the TILLING population showed significantly higher variability and distinct differences from other genetic resources. These materials were created precisely with the goal of extending genetic variability and their further utilization in breeding programmes (accessions 188–198 in the lower left part of the dendrogram in Figure 1a). Fu et al. [17] reported a limited range of genetic resources used in mustard breeding, and the importance of germplasm protection and narrowing genetic variation has been discussed for other cruciferous oilseeds, including mustards [8,33,34]. For this reason, it is appropriate to use molecular markers in selection programmes. Genetic resources (paternal components used for crossing) should be selected not only according to their phenotypic values but also in terms of genetic diversity or distance, and microsatellites can be used as effective practical tools to provide molecular data and evaluate the genetic relationship [35–37].

2.2. Genetic Structure of Yellow Mustard Varieties

Bayesian clustering algorithms were performed in the STRUCTURE 2.3.4 software to describe the genetic structure better and to explain the results from previous statistical analyses. Although this approach has been highly frequented in ecological studies [38] in the last period, it is also used for agricultural crops including oil seed brassicas, such as *B. napus* [32], *B. juncea* [23], and *B. carinata* [25]. Through the STRUCTURE software, we determined the likelihood of the data for several different K, with K ranging from one to seven, to eventually see the plateau pattern in the diagnostic plot. During these computations, we set the burnin period to 100,000, and 150,000 MCMC repeats after burnin. Such a setting seemed to be sufficient, as we attained good mixing and a stable process in equilibrium, with no excessive variations in the $\ln(\alpha)$ parameter estimates. The parameter alpha (the quantification of relative levels of admixture between populations) was allowed to vary for each population to obtain a more accurate estimate of the ancestry. The starting value for the Dirichlet parameter alpha was set to one with a uniform prior. We applied the model with admixture and correlated allele frequencies to attain greater power, in order to detect potential distinct populations that could be closely related. The distributional parameter lambda was computed in the initial pre-run (with lambda set to one), and the inferred value was used in sequential runs (the quantification of independence between markers in terms of allelic frequency distribution). Other hyperparameters were left at their default value. We did not use the LOCPRIOR parameter in our simulations. Consequently, for optimal K = two, we estimated the individual Q-matrix [38–40]. The detailed outputs of cluster number estimations are given in Table S2 and Figure 2. The results and outputs of the STRUCTURE software (Q1/Q2 values) are given in Figure 3 and Table S1 (Supplementary Data).

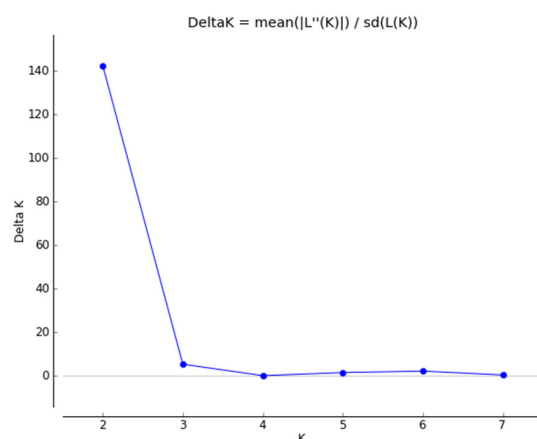


Figure 2. The magnitude of Delta K as a function of K.

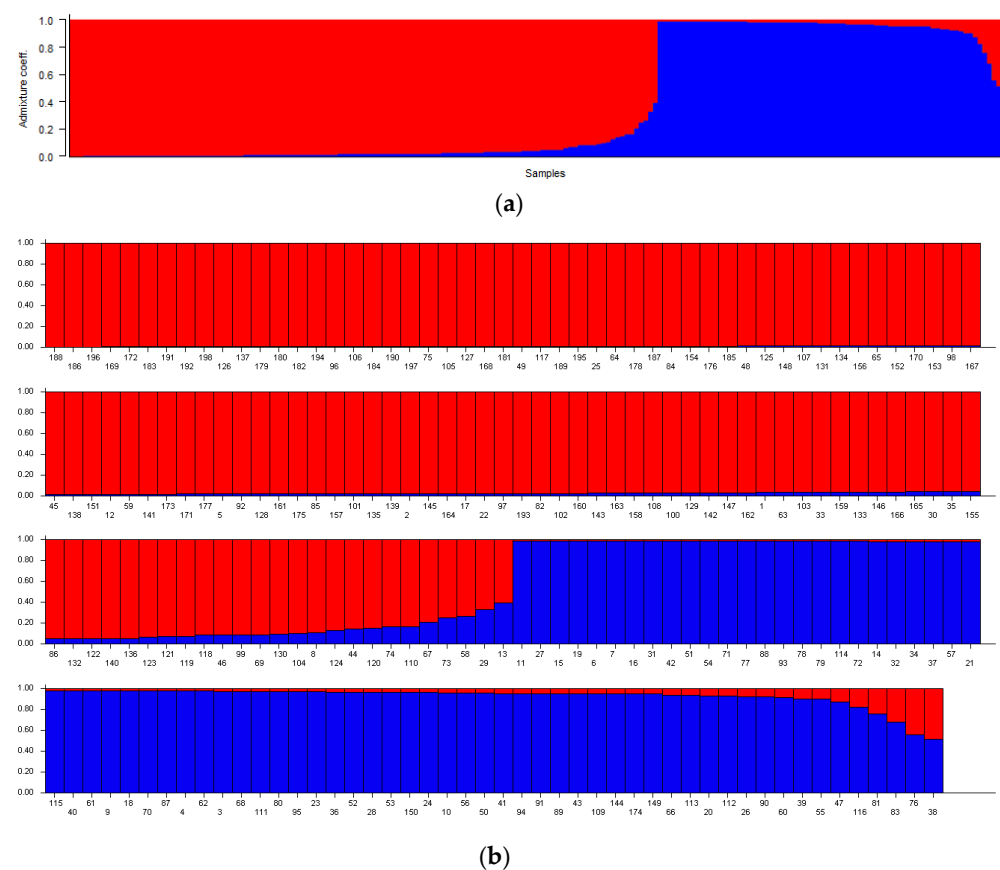


Figure 3. Detailed classification of accessions according to Q1 (blue)/Q2 (red) values: overall overview of the distribution of Q1/Q2 values in the analysed set of genetic resources (a); detailed analysis of genetic structure with marking of individual accessions (b).

The accessions were classified into particular groups, according to Q1/Q2 values (Figure 3a). The mean alpha value was estimated at 0.0602, indicating substantial mating barriers between breeding populations (varieties) and reproductive isolation, due to the limited exchange of genetic resources between individual breeders. The mean values of F_{st_1} and F_{st_2} were 0.3482 and 0.1916, respectively. Also, these values indicated considerable genetic differentiation among the accessions, as was reported previously [41].

The detailed analysis of the raw data and outputs from the STRUCTURE software indicated the presence of a transition zone between the two main clusters estimated, according to parameter Delta K = two. Therefore, detailed classification of accessions according to Q1/Q2 values (Figure 3b) led to the formation of three blocks: a/the first block included accessions with Q1 values in the range of 0.008–0.095; b/the second block formed a transition zone with Q1 values in the range of 0.107–0.396; c/the third block included accessions with Q1 values in the range of 0.515–0.991. Most accessions (113; 57.1%) were assigned to the first block; the third block comprised 73 accessions (36.8%) and 12 accessions (6.1%) formed the transition group. In the first most diverse groups, there were both earlier varieties and landraces, as well as modern varieties and breeding materials. In the second and third groups were landraces and bred varieties, but not new breeding materials. These results from the analysis of the genetic structure were supported by the outputs from the AMOVA, where, although genetic differentiation between types of genetic resources and between individual accessions of yellow mustard within the collection of genetic resources was statistically significant, the proportion of variance between types of genetic resources (landrace, variety, and breeding material) was very low (4.6%). On the contrary, the proportion of variance between genotypes within the analysed set of genetic resources was the highest (87.13%). A more detailed analysis of genetic structure results shows

somewhat more interesting outputs than in the case of the NJ analyses. It was evident that varieties originating from some countries or breeding programmes have similar or the same genetic structures. However, it is more common that, according to the genetic structure, the varieties from one country of origin are classified into two different clusters, and the year of origin of the variety is significantly reflected, because the main factor that distinguishes these clusters is the parameter “year of origin/registration”. It also often reflects changes in the breeding programme, the use of different genetic resources, or the focus on the breeding of new traits (Figure 4).

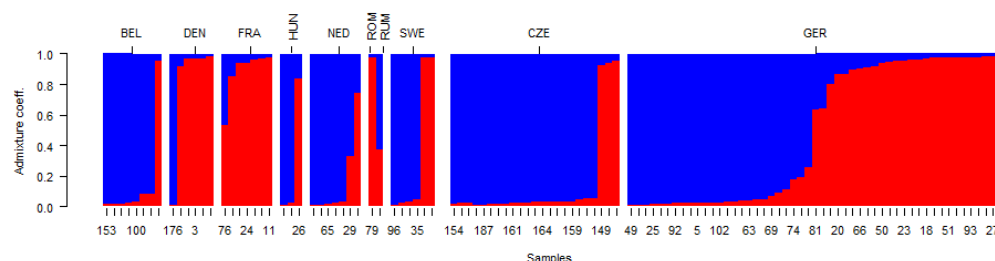


Figure 4. Classification of accessions from selected countries of origin, according to Q1 (red)/Q2 (blue) values and the year of registration.

This distribution of varieties is evident, for example, in Hungarian varieties, where the most recent and older varieties differ significantly in genetic structure. A similar characteristic can be seen in varieties from the Netherlands, Belgium, Denmark, France, Sweden, and Romania. Different characters and explanations may be given for grouping German and Czech varieties. Four German varieties (‘Seco’, ‘Semper’, ‘Setoria’, and ‘Sigri’) belonged to the same group with closely similar Q1 values. They were bred at one breeding station, and their genetic structure reflects the significant similarity of the genetic resources used in breeding programmes. In the case of Czech varieties, both landraces (‘Přerovská bílá’, ‘Česká krajová’), the older variety ‘Zlata’, and modern varieties from the breeding stations Opava and Krukanice belong to the same group. The second cluster comprises current modern varieties from breeding programmes in Selton and BOR (‘Agent’ and ‘Veronika’). In these cases, there is a strong influence of a limited set of genetic resources, and molecular marker methods and the evaluation of genetic diversity can help to appropriate the management of genetic resources and to spread genetic variability and shifts in genetic structures. In this study, some landraces and older varieties were not significantly different from modern varieties. However, it is still evident that the origin of the variety, country of origin, or the breeding programme significantly influences the genetic structure. Similar results were presented in oil seed brassicas [32] or in *B. juncea* [42] and *B. carinata* [34] genetic resources. In contrast, in *B. rapa*, no significant loss of genetic diversity was recorded, although most total variation was attributed to within-cultivar variation in cross-pollinated cultivars [43]. The large geographical variation of genetic resources is a significant source of genetic variation, and high variation among subpopulations of genetic sources has been observed in *B. juncea* [44].

An essential problem in current breeding is the insufficient range of genetic diversity and the related limited availability of interesting genetic resources, such as donors of required traits and characteristics [45]. The depletion and narrowing of genetic diversity occurred in the context of the development of modern breeding in the early 20th century, due to intensive breeding focused on a limited range of traits and the influence of the cultivation of high-yielding and high-quality, but often relatively closely related, varieties of agricultural crops [46,47]. One possibility to expand the range of genetic variability is TILLING technology [48]. This modern approach, combining classical mutagenesis with NGS techniques and the detection of genes carrying mutations, seems to be a very promising breeding method. Plants carrying a change/mutation in the gene of interest are potential donors of a new characteristic and are incorporated into breeding programmes [49]. The

results of the microsatellite analysis show, in this study, a significant spread of genetic variability in the mutant population. All plants from the mutant population were incorporated into the yellow mustard breeding programme and, in the following years, qualitative traits and their possible changes due to mutation will be assessed.

3. Materials and Methods

Plant material and DNA extraction. For the evaluation of genetic diversity, a pilot set of 198 high-value accessions was chosen from the collection of yellow mustard kept at the OSEVA Development and Research in Opava, Czech Republic. These accessions were selected based on their value for current breeding programmes, including modern and old registered varieties, landraces, breeding material, and accessions from the TILLING population. The accessions came from various countries, representing a wide range of spring-type cultivars with a wide range of seed colours. The list of genetic resources, their accession numbers in the gene bank, and their country of origin are given in Table S1, in the Supplementary Data. The seeds were seeded on a horticultural peat substrate and grown under laboratory conditions at a temperature of 21 °C and 16/8 h for five days. The bulk of 36 cotyledons was sampled and gently dried in silica gel. The dried material was homogenised using glass beads in a Retsch® Mixer Mill MM400 (Haan, Germany) for 30 s at a maximal frequency (30 Hz). The DNA was extracted by a modified CTAB PVP method [50], which is more time-consuming, but the obtained DNA is very pure and highly concentrated.

Microsatellite amplification. A total of 15 microsatellite markers were used: BoREM1b_A, BoREM1b_B, BoREM1b_C, BolAB19TF, BoPC34 [51], P381, D3_A, D3_B, P7, P9_B, P9_C, P30_B, P30_C, P35_A, and P35_B [31]. Three primers were used in each PCR reaction, one universal FAM-labelled primer (Ba02 6-FAM or Ba03 6-FAM), a forward primer extended by the sequence of the labelled primer, and a reverse primer [52]. The PCR reaction was carried out in a 10 µL, in 1× reaction buffer (75 mM Tris—HCl, pH = 8.8, 20 mM (NH₄)₂ SO₄, 0.01% Tween 20, 2.5 mM MgCl₂, 200 µM dNTPs), 10 pmol of each primer, 0.5 U of Taq Purple DNA polymerase PPP Master Mix (Top-Bio, Vestec, Czech Republic) or 2× Master Mix with standard buffer (NEB, Ipswich, MA, USA), 2× BSA, and 50 ng of template DNA. The amplification was carried out on a Biometra TProfessional thermocycler (Göttingen, Germany) in the following temperature profile: initial denaturation of 5 min at 94 °C, 31 cycles of 60 s at 95 °C, 60 s at 52–59 °C (depending on the primer used), and 60 s at 72 °C/68 °C, with a final elongation of 15 min at 72 °C/68 °C.

Data Analysis. The results of the PCR analysis were evaluated by fragment analysis on an ABI 3500 genetic analyser (Applied Biosystems, Waltham, MA, USA). Molecular data were analysed using GeneMapper version 3.0 software. The weighted arithmetic mean of genetic distances (WAM) was calculated based on frequency values in the appropriate class of genetic diversity [32]. Diversity measures were calculated in adegenet for the R software version 4.2.2 [53,54]. They consisted of the number of alleles (NA), major allele frequencies (MAF), observed heterozygosity (Ho), expected heterozygosity (He), and the polymorphism information content (PIC). To determine genetic diversity among and within genetic resources, AMOVAs [55] were performed with the package “pegas” in R [56]. The null hypotheses associated with differentiation parameters (Φ) were as follows: there is no genetic difference between types of genetic resources ($\Phi_{CT} = 0$), between genotypes within types of genetic resources ($\Phi_{SC} = 0$), or between genotypes within collection of genetic resources ($\Phi_{ST} = 0$) [55]. The data were statistically evaluated using traditional genetic diversity metrics in MVSP 3 (Kovach Computing Services, Pentraeth, UK) and DARwin 6 (CIRAD, F) software. Intending to estimate the population/cluster number, we used the Bayesian clustering approach, involving MCMC estimation, and followed Evanno et al. [40] and Kolář et al. [38]. STRUCTURE software was used to assess the genetic structure of analysed germplasms [39].

4. Conclusions

The issue of genetic diversity, its conservation, protection, and exploitation are the key points in plant breeding. In this study, we focused on the detailed description and molecular analysis of the genetic variability of a large collection of yellow mustard genotypes. High values of observed heterozygosity, diversity index, and WAM index indicated a high level of polymorphism between individual accessions. Also, outputs from AMOVA supported these findings and showed a low proportion of variance (4.6%) between types of genetic resources (landrace, variety, and breeding material) and a high proportion of variance between genotypes within the analysed set of genetic resources (87.13%). In addition to the standard descriptive characteristics, two other procedures were applied, for an overall assessment of genetic variability and the degree of diversity of the genetic resources: PCoA analysis and Bayesian clustering algorithms. The results show a more valuable interpretation and highlight another way of processing the genetic diversity data and indicated mating barriers between varieties and reproductive isolation, due to the limited exchange of genetic resources between breeders. The range of genetic diversity was much higher than reported in other studies, concerning the large and variable collection of genetic resources. These results showed the importance of the proper management of genetic resources for the preservation of genetic diversity. Analysis of genetic structure may lead (as shown in our example) to a clarification of the grouping of accessions (genetic resources and varieties) and recommendations for expanding the genetic background in breeding programmes. The choice of parental components based on phenotypic characteristics and genetic diversity facilitates the selection of genetic resources for breeding programmes and can contribute to the success of breeding.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/plants12234026/s1>, Table S1. List of analysed genetic resources of yellow mustard and outputs of the STRUCTURE software (Q1/Q2 values). Table S2. Detailed output of cluster number estimation and statistics for different values of K. Figure S1. The graphic representation of the degree of dissimilarity between genetic resources—output from DARwin 6.0 software.

Author Contributions: Conceptualization, V.Č., M.K. and E.J.; methodology, E.J.; software, M.R.; validation, E.J. and O.H.; formal analysis, E.J. and O.H.; investigation, V.Č.; resources, A.R.; data curation, V.Č.; writing—original draft preparation, V.Č.; writing—review and editing, P.B., D.S., B.P., E.J. and M.K.; visualisation, M.R.; supervision, V.Č.; project administration, M.K.; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Agriculture of the Czech Republic, grant number QK1910225, and the Grant Agency of the University of South Bohemia in České Budějovice, grant number GAJU 080/2022/Z.

Data Availability Statement: Additional data files and project outputs are available at <http://biocentrum.zf.jcu.cz/laborator.php?sub=metodiky> (accessed on 20 November 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hemingway, J.S. The mustard species: Condiment and food ingredient use and potential as oilseed crops. In *Brassica Oilseeds Production and Utilization*; Kimber, D., McGregor, D.I., Eds.; CAB International: Wallingford, UK, 1995; pp. 373–383.
2. Ruan, S.F.; Wang, Z.X.; Xiang, S.J.; Chen, H.J.; Shen, Q.; Liu, L.; Wu, W.F.; Cao, S.W.; Wang, Z.W.; Yang, Z.J.; et al. Mechanisms of white mustard seed (*Sinapis alba* L.) volatile oils as transdermal penetration enhancers. *Fitoterapia* **2019**, *138*, 104195. [[CrossRef](#)] [[PubMed](#)]
3. Bodnaryk, R.P.; Lamb, R.J. Mechanisms of resistance to the flea beetle, *Phyllotreta cruciferae* (Goeze), in mustard seedlings, *Sinapis alba* L. *Can. J. Plant Sci.* **1991**, *71*, 13–20. [[CrossRef](#)]
4. Kumari, P.; Singh, K.P.; Rai, P.K. Draft genome of multiple resistance donor plant *Sinapis alba*: An insight into SSRs, annotations and phylogenetics. *PLoS ONE* **2020**, *15*, e0231002. [[CrossRef](#)] [[PubMed](#)]
5. Valdes, Y.; Viaene, N.; Moens, M. Effects of yellow mustard amendments on the soil nematode community in a potato field with focus on *Globodera rostochiensis*. *Appl. Soil Ecol.* **2012**, *59*, 39–47. [[CrossRef](#)]

6. Krasnodębska-Ostrega, B.; Sadowska, M.; Biaduń, E.; Mazur, R.; Kowalska, J. *Sinapis alba* as a useful plant in bioremediation—Studies of defense mechanisms and accumulation of As, Tl and PGEs. *Int. J. Phytoremediat.* **2022**, *24*, 1475–1490. [[CrossRef](#)] [[PubMed](#)]
7. Masierowska, M. Floral nectaries and nectar production in brown mustard (*Brassica juncea*) and white mustard (*Sinapis alba*) (Brassicaceae). *Plant Syst. Evol.* **2003**, *238*, 97–107. [[CrossRef](#)]
8. Katepa-Mupondwa, F.; Gugel, R.K.; Raney, J.P. Genetic diversity for agronomic, morphological and seed quality traits in *Sinapis alba* L. (yellow mustard). *Can. J. Plant Sci.* **2006**, *86*, 1015–1025. [[CrossRef](#)]
9. Khoury, C.K.; Brush, S.; Costich, D.E.; Curry, H.A.; de Haan, S.; Engels, J.M.M.; Guarino, L.; Hoban, S.; Mercer, K.L.; Miller, A.J.; et al. Crop genetic erosion: Understanding and responding to loss of crop diversity. *New Phytol.* **2022**, *233*, 84–118. [[CrossRef](#)]
10. Yu, J.K.; Chung, Y.S. Plant Variety Protection: Current Practices and Insights. *Genes* **2021**, *12*, 1127. [[CrossRef](#)]
11. Govindaraj, M.; Vetriventhan, M.; Srinivasan, M. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genet. Res. Int.* **2015**, *2015*, 431487. [[CrossRef](#)]
12. Garland, S.; Curry, H.A. Turning promise into practice: Crop biotechnology for increasing genetic diversity and climate resilience. *PLoS Biol.* **2022**, *20*, e3001716. [[CrossRef](#)] [[PubMed](#)]
13. McCallum, C.M.; Comai, L.; Greene, E.A.; Henikoff, S. Targeted screening for induced mutations. *Nat. Biotechnol.* **2000**, *18*, 455–457. [[CrossRef](#)] [[PubMed](#)]
14. Tsai, H.; Howell, T.; Nitcher, R.; Missirian, V.; Watson, B.; Ngo, K.J.; Lieberman, M.; Fass, J.; Uauy, C.; Tran, R.K.; et al. Discovery of Rare Mutations in Populations: TILLING by Sequencing. *Plant Physiol.* **2011**, *156*, 1257–1268. [[CrossRef](#)] [[PubMed](#)]
15. El-Esawi, M.A. Genetic diversity and evolution of Brassica genetic resources: From morphology to novel genomic technologies—A review. *Plant Genet. Resour.* **2017**, *15*, 388–399. [[CrossRef](#)]
16. Granot, D.; Shabelsky, E.; Schaffermann, D.; Yaniv, Z. Analysis of genetic variability between populations of *Sinapis alba* and the effect of cultivation on the variability. *Acta Hort.* **1996**, *407*, 67–74. [[CrossRef](#)]
17. Fu, Y.B.; Gugel, R.K.; Katepa-Mupondwa, F. Genetic diversity of *Sinapis alba* germplasm as revealed by AFLP markers. *Plant Genet. Resour.* **2006**, *4*, 87–95. [[CrossRef](#)]
18. Yadav, M.; Rana, J.S. ISSR markers assisted studies on genetic diversity in *Brassica juncea*. *Int. J. Pharma Bio Sci.* **2012**, *3*, 402–411.
19. Plieske, J.; Struss, D. Microsatellite markers for genome analysis in *Brassica*. I. Development in *Brassica napus* and abundance in Brassicaceae species. *Theor. Appl. Genet.* **2001**, *102*, 689–694. [[CrossRef](#)]
20. Javidfar, F.; Cheng, B. Construction of a genetic linkage map and QTL analysis of erucic acid content and glucosinolate components in yellow mustard (*Sinapis alba* L.). *BMC Plant Biol.* **2013**, *13*, 142. [[CrossRef](#)]
21. Fu, Y.B.; Cheng, B.; Peterson, G.W. Genetic diversity analysis of yellow mustard (*Sinapis alba* L.) germplasm based on genotyping by sequencing. *Genet. Resour. Crop Evol.* **2013**, *61*, 579–594. [[CrossRef](#)]
22. Karp, A.; Edwards, K.J. *Molecular Techniques in the Analysis of the Extent and Distribution of Genetic Diversity*; International Plant Genetic Resources Institute: Rome, Italy, 1997; pp. 11–38.
23. Singh, K.H.; Singh, L.; Parmar, N.; Kumar, S.; Nanjundan, J.; Singh, G.; Thakur, A.K. Molecular characterization and genetic diversity analysis in Indian mustard (*Brassica juncea* L. Czern & Coss.) varieties using SSR markers. *PLoS ONE* **2022**, *17*, e0272914. [[CrossRef](#)]
24. Pradhan, A.; Nelson, M.N.; Plummer, J.A.; Cowling, W.A.; Yan, G. Characterization of *Brassica nigra* collections using simple sequence repeat markers reveals distinct groups associated with geographical location, and frequent mislabelling of species identity. *Genome* **2011**, *54*, 50–63. [[CrossRef](#)] [[PubMed](#)]
25. Tesfaye, M.; Feyissa, T.; Hailesilassie, T.; Kanagarajan, S.; Zhu, L.-H. Genetic diversity and population structure in ethiopian mustard (*Brassica carinata* A. Braun) as revealed by single nucleotide polymorphism markers. *Genes* **2023**, *14*, 1757. [[CrossRef](#)] [[PubMed](#)]
26. Guichoux, E.; Lagashe, L.; Wagner, S.; Chaumeil, P.; Léger, P.; Lepais, O.; Lepoitteiven, C.; Malausa, T.; Revardel, E.; Salin, F.; et al. Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* **2011**, *11*, 591–611. [[CrossRef](#)] [[PubMed](#)]
27. Mondini, L.; Noorani, A.; Pagnotta, M. Assessing plant genetic diversity by molecular tools. *Diversity* **2009**, *1*, 19–35. [[CrossRef](#)]
28. Varshney, R.K. Designing future crops: Genomics-assisted breeding comes of age. *Trends Plant. Sci.* **2021**, *26*, 631–649. [[CrossRef](#)]
29. Olsson, G. Self-incompatibility and outcrossing in rape and white mustard. *Hereditas* **1960**, *46*, 241–252. [[CrossRef](#)]
30. Treccarichi, S.; Ben Ammar, H.; Amari, M.; Cali, R.; Tribulato, A.; Branca, F. Molecular markers for detecting inflorescence size of *Brassica oleracea* L. crops and *B. oleracea* complex species (n = 9) useful for breeding of broccoli (*B. oleracea* var. italica) and cauliflower (*B. oleracea* var. botrytis). *Plants* **2023**, *12*, 407. [[CrossRef](#)]
31. Cui, X.; Dong, Y.; Hou, X.; Cheng, Y.; Zhang, J.; Jin, M. Development and characterization of microsatellite markers in *Brassica rapa* ssp. chinensis and transferability among related species. *Agric. Sci. China* **2008**, *7*, 19–31. [[CrossRef](#)]
32. Havlickova, L.; Jozova, E.; Rychla, A.; Klima, M.; Kucera, V.; Curn, V. Genetic diversity assessment in winter oilseed rape (*Brassica napus* L.) collection using AFLP, ISSR and SSR markers. *Czech J. Genet. Plant Breed.* **2014**, *50*, 216–225. [[CrossRef](#)]
33. Vazhakkannadi, V.; Singh, N.; Vasudev, S.; Yadava, D.; Kumar, S.; Naresh, S.; Bhat, S.; Prabhu, K. Assessment of genetic diversity in *Brassica juncea* (Brassicaceae) genotypes using phenotypic differences and SSR markers. *Rev. Biol. Trop.* **2013**, *61*, 1919–1934.
34. Khedikar, Y.; Clarke, W.; Chen, L.; Higgins, E.; Kagale, S.; Koh, C.; Bennett, R.; Parkin, I. Narrow genetic base shapes population structure and linkage disequilibrium in an industrial oilseed crop, *Brassica carinata* A. Braun. *Sci. Rep.* **2020**, *10*, 12629. [[CrossRef](#)] [[PubMed](#)]

35. Sobotka, R.; Dolanská, L.; Čurn, V.; Ovesná, J. Fluorescence-based AFLPs occur as the most suitable marker system for oilseed rape cultivar identification. *J. Appl. Genet.* **2004**, *45*, 161–173. [PubMed]
36. Cowling, W.A. Genetic diversity in Australian canola and implications for crop breeding for changing future environments. *Field Crops Res.* **2007**, *104*, 103–111. [CrossRef]
37. Chikkaputtaiah, C.; Debbarma, J.; Baruah, I.; Havlickova, L.; Deka Boruah, H.P.; Curn, V. Molecular genetics and functional genomics of abiotic stress-responsive genes in oilseed rape (*Brassica napus* L.): A review of recent advances and future. *Plant Biotechnol. Rep.* **2017**, *11*, 365–384. [CrossRef]
38. Kolář, F.; Fér, T.; Štěch, M.; Trávníček, P.; Dušková, E.; Schönschwetter, P.; Suda, J. Bringing together evolution on serpentine and polyploidy: Spatiotemporal history of the diploid-tetraploid complex of *Knautia arvensis* (Dipsacaceae). *PLoS ONE* **2012**, *7*, e39988. [CrossRef] [PubMed]
39. Pritchard, J.K.; Stehens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef]
40. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [CrossRef]
41. Bird, C.; Karl, S.A.; Smouse, P.E.; Toonen, R.J. Detecting and measuring genetic differentiation. In *Phylogeography and Population Genetics in Crustacea*; Koenemann, S., Held, C., Schubart, C., Eds.; CRC Press: Boca Raton, FL, USA, 2011; pp. 31–55. [CrossRef]
42. Yao, Q.L.; Chen, F.B.; Fang, P.; Zhou, G.F.; Fan, Y.H.; Zhang, Z.R. Genetic diversity of Chinese vegetable mustard (*Brassica juncea* Coss) landraces based on SSR data. *Biochem. Syst. Ecol.* **2012**, *45*, 41–48. [CrossRef]
43. Ofori, A.; Becker, H.C.; Kopisch-Obuch, F.J. Effect of crop improvement on genetic diversity in oilseed *Brassica rapa* (turnip-rape) cultivars, detected by SSR markers. *J. Appl. Genet.* **2008**, *49*, 207–212. [CrossRef]
44. Abdel-haleem, H.; Luo, Z.; Szczepanek, A. Genetic diversity and population structure of the USDA collection of *Brassica juncea* L. *Ind. Crops Prod.* **2022**, *187*, 115379. [CrossRef]
45. Sanchez, D.; Sadoun, S.B.; Mary-Huard, T.; Allier, A.; Moreau, L.; Charcosset, A. Improving the use of plant genetic resources to sustain breeding programs' efficiency. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2205780119. [CrossRef] [PubMed]
46. Bhandari, H.; Bhanu, A.; Srivastava, K.; Singh, M.; Shreya; Hemantaranjan, A. Assessment of genetic diversity in crop plants: An overview. *Adv. Plants Agric. Res.* **2017**, *7*, 279–286. [CrossRef]
47. Begna, T. Role and economic importance of crop genetic diversity in food security. *Int. J. Agric. Sci. Food Technol.* **2021**, *7*, 164–169. [CrossRef]
48. Kurowska, M.; Daszkowska-Golec, A.; Gruszka, D.; Marzec, M.; Szurman, M.; Szarejko, I.; Maluszynski, M. TILLING: A shortcut in functional genomics. *J. Appl. Genet.* **2011**, *52*, 371–390. [CrossRef] [PubMed]
49. Mba, C. Induced mutations unleash the potentials of plant genetic resources for food and agriculture. *Agronomy* **2013**, *3*, 200–231. [CrossRef]
50. Doyle, J. DNA Protocols for Plants. In *Molecular Techniques in Taxonomy*; Hewitt, G.M., Johnston, A.W.B., Young, J.P.W., Eds.; Springer: Berlin/Heidelberg, Germany, 1991; pp. 283–293. [CrossRef]
51. Louarn, S.; Torp, A.; Holme, I.; Andersen, S.B.; Jensen, B. Database derived microsatellite markers (SSRs) for cultivar differentiation in *Brassica oleracea*. *Genet. Resour. Crop Evol.* **2007**, *54*, 1717–1725. [CrossRef]
52. Jozová, E.; Čurn, V. Microsatellite genotyping with 4-color fluorescent detection using multiple-tailed primers in oilseed rape. *Úroda* **2014**, *12*, 191–194.
53. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020; Available online: <https://www.R-project.org/> (accessed on 30 October 2022).
54. Jombart, T.; Ahmed, I. Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **2011**, *27*, 3070–3071. [CrossRef]
55. Excoffier, L.; Smouse, P.E.; Quattro, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **1992**, *131*, 479–491. [CrossRef]
56. Paradis, E. *pegas*: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **2010**, *26*, 419–420. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.