

# Genetic diversity and distinctiveness revealed by SSR markers among rapeseed (*Brassica napus* L.) genotypes from Australia, China and India

Sheng Chen<sup>1</sup>, Matthew N. Nelson<sup>1</sup>, Kioumars Ghamkhar<sup>2</sup>, Wallace A. Cowling<sup>1</sup>

<sup>1</sup>School of Plant Biology, <sup>2</sup>Centre for Legumes in Mediterranean Agriculture, Faculty of Natural and Agricultural Sciences, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

Email: chens@cyllene.uwa.edu.au

## Abstract

Rapeseed (*Brassica napus* L.) is one of the most important sources of vegetable oil. Its limited geographic range of origin and intensive selection in breeding programs have led to relatively limited genetic diversity in current breeding material. Consequently, many characters of commercial value show little genetic diversity in this important crop. We characterised the genetic diversity of 72 *B. napus* genotypes from Australia, China and India using 55 SSR markers that span the *B. napus* genome. The SSR markers amplified 365 alleles at 127 polymorphic loci and provided unique genetic fingerprints for 72 rapeseed genotypes, and genotypes from 5 other *Brassica* species. Hierarchical clustering and two-dimensional multidimensional scaling were used to identify three main groups: a group with all Indian genotypes, a group with 22 Chinese genotypes (China-1), and a mixed group. The mixed group could be divided into three subgroups: one with all Australian genotypes, one with Chinese genotypes (China-2), and one with genotypes from Europe and Canada. The subgroups from Australia, Europe/Canada and China-2 were more closely related to each other than they were with the Indian and China-1 groups. This suggested a relatively broader genetic diversity among the rapeseeds in China than among those in Europe, Canada and Australia. The Indian *B. napus* genotypes were rather distant to the other *B. napus* genotypes tested in this study, and India might be a good resource for expanding genetic diversity in *B. napus* breeding programs in Australia and China. 'Private alleles', a term primarily used to describe alleles that are unique to a particular endemic species, were most richly represented in the Indian rapeseed population (average 4.00 private alleles per genotype). The Chinese population had moderate richness (1.03), whereas both the Australian population (0.50) and the European and Canadian population (0.40) showed low private allelic richness. Private allelic richness supported the results obtained from hierarchical cluster analysis and suggested that the richness of private SSR alleles could be considered as a measure of genetic distinctiveness. Therefore they might be used as an important indicator when choosing parents in rapeseed breeding.

Key words: rapeseed, *Brassica napus*, genetic diversity, genetic distance, genetic distinctiveness, private allele, SSR.

## Introduction

Rapeseed (*Brassica napus*, AACC,  $2n = 38$ ) originated from a spontaneous hybridization between turnip rape (*B. rapa*, AA,  $2n = 20$ ) and a member of the cabbage family (*B. oleracea*, CC,  $2n = 18$ ) (UN, 1935; Kimber, 1995). Although this amphidiploid species was formed only in relatively recent agricultural times, it became the most important oilseed *Brassica* crop in the world (Liu, 2000). Since the release of canola-quality *B. napus* in the 1970s, Canada has contributed to the world's rapeseed production together with Europe. However, the narrow origin of *B. napus* and recent breeding practices have led to a relatively limited genetic diversity in modern *B. napus* cultivars (Becker, 1995). Consequently, many characters of value show little genetic diversity in this important crop for breeding purposes.

In China, *B. napus* rapeseed has displaced field rape (*B. rapa*), the traditional oilseed crop in China which has been cultivated for more than 6,000 years (Liu, 1985). Meanwhile, Chinese rapeseed cultivars have been genetically improved for adaptation to local environments mainly by introgressing genomic components from Chinese *B. rapa* (Qian, 2006). In Australia, a broad range of rapeseed varieties were introduced from Europe, Canada and Asia in the 1970s. During the next 30 years, approximately 5 breeding cycles were completed in a closed population and a number of blackleg resistant and high quality canola varieties were developed (Cowling, 2006). While the original parents were very diverse, the population inbreeding coefficient shows a loss of genetic diversity that will impact on future breeding progress (Cowling, 2007).

"Private allele" is a term originally used in population biology to describe alleles that are unique to a particular endemic species. Slatkin termed an allele a "private allele" if it is found in only one of several populations in a larger collection (Slatkin, 1985) and he described several properties of private alleles in his simulations (Slatkin, 1985). The frequency of private alleles has since been used to assess the population structure and the rate of gene flow in animals, plants and plant pathogens. In recent years, it has also been used to analyse genetic diversity in grain crops such as maize (Choukan, 2006) and barley (Feng, 2006), but the implication to crop genetic breeding and the possible utilization strategies have not yet been discussed.

In this paper we investigated the genetic diversity of *B. napus* rapeseed in populations from Australia, China and India using mapped simple sequence repeat (SSR) markers that spanned the *B. napus* genome. Hierarchical clustering and two-dimensional multidimensional scaling were used to identify the main groups of genotypes. Private alleles, defined in this

study as SSR alleles that were present in only one geographic *B. napus* population, were determined for each of the main genotype groups. The value of private alleles as an indicator of distinctiveness of genetic germplasm, and as a guide for choosing parents in a breeding program, is discussed.

## Materials and methods

The plant materials for this study were obtained from 72 accessions of *B. napus* (rapeseed) from China, Australia, India, Europe and Canada, and 5 genotypes from other *Brassica* species. The genomic DNA was separately extracted from five individuals from each accession using the CTAB method. After pre-screening of 102 *Brassica* SSR markers, 55 primer pairs were chosen that gave clear, reproducible and polymorphic amplification products at one or more loci in *B. napus*. According to the primer dye types and the size range of the PCR products, the PCR products were pooled together for analysis on an AB3730xl capillary sequencer. The data were then evaluated using GeneMapper 3.7 (Applied Biosystems). On the basis of preliminary cluster analysis, one genotype was selected from each accession for further work. The selected genotypes were the most homozygous individuals from each accession.

*The assignment of alleles to loci within a SSR marker:* Polymorphic bands were firstly scored among genotypes with 1 and 0 for the presence and absence of a SSR band, respectively. The map locations of SSR markers were obtained from published maps of *B. napus* (Lowe, 2004; Piquemal, 2005) and from the originator of the SSR markers (Andrew Sharpe, personal communication). For primer pairs that amplified more than one locus, the following criteria were used to help assign each allele to a specific locus: firstly, alleles within a marker were regarded as belonging to the same locus if they showed an obvious codominant relationship from their segregation patterns among different lines; secondly, if their amplicon sizes were very close and followed a certain distinctive stuttering pattern, they were considered to belong to the same locus; thirdly, because most multi-locus markers detected two homoeologous loci in the A and C genomes of *B. napus*, it was often possible to infer from the *Brassica* species controls (including *B. rapa*, AA-genome and *B. oleracea*, CC-genome) whether a particular locus belonged to the A or C genomes. Furthermore, alleles present only in *B. nigra* (BB) and/or *B. juncea* (AABB) and/or *B. carinata* (BBCC) were assumed to be located on the B-genome and thus were regarded as novel loci.

*Data analysis:* Allele frequency was calculated using the NTSYSpc program (Rohlf, 2006). Genetic distances (GD) were calculated using the formula from Nei and Li (Nei, 1979). Data from the GD matrix among 72 *B. napus* genotypes and 5 other *Brassica* genotypes were subjected to hierarchical cluster analysis using the unweighted pair group method and arithmetic averages (UPGMA), and ordination by non-metric multidimensional scaling (MDS) from PRIMER 6 (Plymouth Routines In Multivariate Ecological Research) (Clarke, 2006).

## Results

A total of 127 loci with 365 polymorphic alleles were amplified among 72 rapeseed genotypes and 5 genotypes from other *Brassica* species. In order to reveal the genetic relationships among these genotypes, an UPGMA dendrogram representing genetic dissimilarity among all these genotypes was constructed (Fig. 1). At a genetic dissimilarity level of about 0.57, all the rapeseed genotypes were clustered into one group and were clearly separated from the other five *Brassica* species. This result was confirmed by two-dimensional multidimensional scaling (2-D MDS) analysis (Fig. 2A). Among the 72 rapeseed genotypes, the 3 Indian genotypes were clearly clustered into an independent group at the genetic dissimilarity level of about 0.49, indicating that the Indian genotypes studied are genetically distant from other *B. napus* genotypes. At the genetic dissimilarity level of about 0.46, 22 Chinese genotypes clustered into a distinct group (China-1). The remaining 8 Chinese genotypes (China-2) were in a third mixed group with genotypes from Australia, Europe and Canada. In this mixed group, the Chinese and Australian genotypes were located in distinct subgroups at the genetic dissimilarity level of about 0.44, whilst the European and Canadian genotypes were combined in a third subgroup. At the similarity level of 0.4, the subset 2-D MDS analysis of the 72 *B. napus* genotypes (Fig. 2B) supported the division of genotypes into three major groups: the Indian group, the China-1 group, and the mixed group. At the similarity level of 0.38, the subset 2-D MDS analysis also supported the division of the mixed group into 3 subgroups (Fig. 2B). Figs. 1 and 2 indicate that the China-2 group shares more in common with the subgroups from AU and EU/CA than they do with the China-1 group.

A large number of private alleles were extremely rare, and occurred in only 1 genotype. Of a total of 59 private SSR alleles, 29 were located in one *B. napus* genotype and 13 were in two *B. napus* genotypes. A few private alleles were widely distributed within geographical populations. For example, one private allele from China was shared by 10 Chinese genotypes and one private allele from Australia was shared by 9 Australian genotypes. The number of private alleles per genotype varied significantly among the 4 populations with an average of 4.00 in IN, 1.03 in CN, 0.50 in AU and 0.40 in EU/CA (Table 1). The average richness of private alleles in Indian *B. napus* was much higher than the other three populations, although it is emphasised that only 3 genotypes were available from India. Some alleles in the 4 *B. napus* populations were also found in one or more of the genotypes from 5 other *Brassica* species. Up to 67% of the alleles in the IN population were found in other *Brassica* species, but none of the alleles from the EU/CA population were found in other *Brassica* species (Table 1).

A total of 287 alleles were detected with an average of 5.22 alleles per SSR marker in the 72 *B. napus* genotypes. The number of alleles per marker varied from only one allele with sN2552 to as many as 17 alleles with sN3514F. The alleles were distributed among the Australia (AU), China (CN), Europe/Canada (EU/CA) and India (IN) geographical populations as follows: 213 alleles for AU, 241 for CN, 158 for EU/CA and 191 for IN. The average number of alleles per marker and standard deviation (SD) were as follows: CN (4.38±1.97), AU (3.87±1.80), IN (3.47±1.69), and EU/CA (2.87±1.62). Of the 287 alleles, 120 (41.8%) were common to all four populations (common alleles) and the average number of common alleles

was  $2.18 \pm 1.42$ , with the highest number of common alleles (7) being detected with sN3514F while no common alleles were detected with 3 SSR markers Ni2-F02, sNRD71 and sORF73.

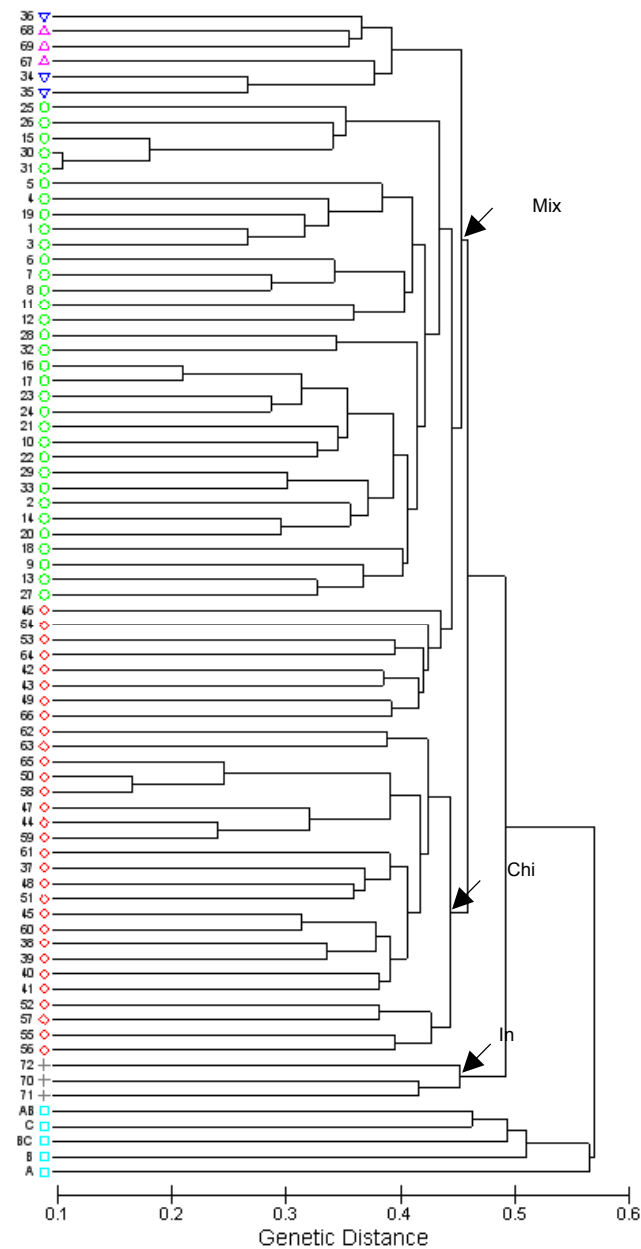


Fig. 1. Cluster analysis of Nei's matrix distances among 72 *B. napus* genotypes and a genotype from 5 other *Brassica* species (A, *B. rapa*, B, *B. nigra*, C, *B. oleracea*, AB, *B. juncea*, and BC, *B. carinata*) based on 365 alleles from 55 SSR markers. *B. napus* genotypes are from Australia (○), China (◇), India (+), Europe (Δ) and Canada (∇) and other *Brassica* species (□). The division points of the three major *B. napus* clusters (India, China-1 and "mixed") are indicated with arrows.

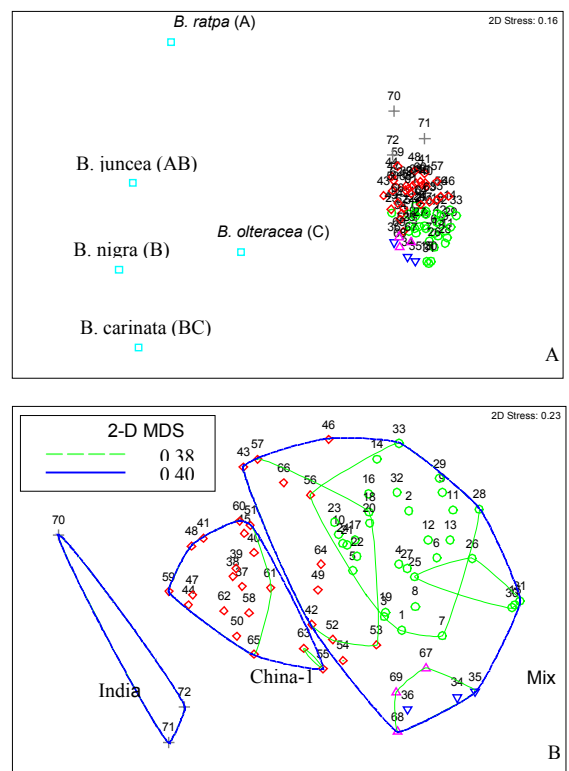


Fig. 2. Associations among 72 rapeseed *B. napus* genotypes and a genotype from 5 other *Brassica* species revealed by 2-D MDS analysis, based on 365 alleles from 55 SSR markers. The genotypes include *B. napus* from Australia (○), China (◇), India (+), Europe (Δ) and Canada (∇) and other *Brassica* species (□). The three major *B. napus* clusters (India, China-1 and "mixed") are indicated. **A.** The rapeseed *B. napus* genotypes were clustered into one group which is clearly separated from the other five *Brassica* species. **B.** the subset 2-D MDS analysis of the 72 *B. napus* genotypes showing the division into three major groups: the Indian group, the China-1 group and the mixed group at the similarity level of 0.4 (dark line), and the subgroups at the similarity level of 0.38 (pale line).

**Table 1. Summary of the private SSR alleles detected in 4 different *B. napus* geographic populations. From a total of 72 *B. napus* genotypes, 6 were doubled haploid derivatives of varieties present in the study and were excluded from this analysis. Some private alleles in *B. napus* genotypes were also found in one or more of the genotypes from 5 other *Brassica* species, and these alleles are also listed in the table**

Population	Number of genotypes	Private alleles		Private alleles found in 5 other <i>Brassica</i> species	
		Number	Richness	Number	Percentage
Australia	28	14	0.50	6	42.86
China	30	31	1.03	7	22.58
India	3	12	4.00	8	66.67
Europe/Canada	5	2	0.40	0	0.00
Total	66	59	0.89	21	35.59

## Discussion

In this study, Australian *B. napus* genotypes belonged to a “mixed” group with European and Canadian genotypes and the China-2 sub-group at the genetic dissimilarity level of about 0.46. However, they remained as an independent sub-group at the genetic dissimilarity level of about 0.44. It is known that the Australian *B. napus* breeding pool originated in 1970 from lines introduced from Europe, Canada and Asia. Of the canola varieties released in Australia from 1995 to 2002, approximately half of the ancestral contribution was from Asian *B. napus* or *B. juncea*, and half from European or Canadian *B. napus*. Australian *B. napus* has undergone approximately 5 breeding cycles in a closed population since 1970 (Cowling, 2006). Based on the genome-wide SSR markers used in this study, Australian genotypes have formed their own distinctive SSR allelic pattern during 30 years of breeding in genetic isolation.

The Chinese *B. napus* were allocated into two different groups, China-1 and China-2. These two groups were not associated with a particular oil quality (double high and double low), or to a particular growth habit (spring type and winter/semi-winter type). Of the 31 private SSR alleles in the Chinese *B. napus* population, 7 alleles (23%) were found in the other *Brassica* species tested (5 alleles in the A genome of *B. rapa* and 2 alleles in the C-genome of *B. carinata*). This indicated that *B. rapa* and *B. carinata* might have been involved in the breeding of Chinese *B. napus*. This is consistent with the breeding history of *B. napus* in China. Since *B. napus* was introduced to China in the 1930's to 1940's, the genomic components from *B. rapa* have been frequently introgressed into *B. napus*. More than 50% of the *B. napus* cultivars released in China were derived from *B. napus* × *B. rapa* crosses (Liu, 1985).

The *B. napus* from India was the most distant group and had the highest allelic richness of private alleles (as high as 4.0 private alleles per genotype) (Table 1). Eight of the 12 Indian private SSR alleles were found in one or more of the 5 *Brassica* species tested, indicating that these Indian *B. napus* genotypes contained novel alleles introgressed from other *Brassica* species. Therefore the Indian *B. napus* might be very valuable in broadening the *B. napus* gene pool of other populations.

The richness rate of private SSR alleles in each group was in the order of IN > CN > AU ≥ EU/CA. This was congruent with the hierarchical clustering order based on the genetic dissimilarity level, suggesting that the presence of private alleles could be considered as a measure of both genetic diversity and genetic distinctiveness. Populations such as EU/CA may have provided the base breeding stock for most of the world's *B. napus* breeding populations; hence EU/CA has very low private allelic richness. Australian germplasm was derived from EU/CA and Chinese ancestors, and also has low private allelic richness. Chinese and Indian *B. napus* germplasm has most likely undergone introgression with other *Brassica* species, which has increased the frequency of private alleles.

The richness rate of private SSR alleles represents the distinctiveness and novelty of certain population, and may provide breeders with a useful indicator for evaluating and utilizing novel germplasm. The relationship between the richness rate of private SSR alleles and breeding potential for agriculturally-important traits has yet to be determined. The enhancement of the richness of private alleles through SSR markers could be a valuable role of molecular markers in plant breeding. We anticipate that the richness of private SSR alleles will be correlated with the abundance of distinctive agriculturally-important alleles in the new germplasm. Private SSR alleles might be used as an important indicator when choosing parents for a breeding program.

## References

- Becker HC, Engqvist, G.M., Karlsson, B. (1995) Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. *Theor Appl Genet* 91: 62-67
- Choukan R, Hossainzadeh, A., Ghannadha, M. R., Warburton, M. L., Talei, A. R., Mohammadi, S. A (2006) Use of SSR data to determine relationships and potential heterotic groupings within medium to late maturing Iranian maize inbred lines. *Field Crops Research* 95: 212-222
- Clarke KR, Gorley, R. N (2006) PRIMER v6: user manual/tutorial. PRIMER-E Ltd, Plymouth, UK
- Cowling WA (2006) Genetic diversity in Australian canola and implications for future breeding. *In* ASA, ed, Proceedings of the 13th Australian Agronomy Conference, Perth, Western Australia
- Cowling WA (2007) Genetic diversity in canola for changing environments. *In* 12IRC, ed, 12th International Rapeseed Congress, Wuhan, China
- Feng Z-Y, Zhang, L.-L., Zhang, Y.-Z., Ling, H.-Q. (2006) Genetic diversity and geographical differentiation of cultivated six-rowed naked barley landraces from the Qinghai-Tibet plateau of China detected by SSR analysis. *Genetics and Molecular Biology* 29: 330-338

- Kimber DS, McGregor, D. I. (1995) Brassica oilseeds: production and utilization. CABI Publishing, Wallingford, UK
- Liu H (1985) Rapeseed genetics and breeding. Shanghai Science and Technology Press, Shanghai
- Liu H (2000) Genetics and breeding in rapeseed. China Agricultural University Press, Beijing
- Lowe AJ, Moule, C., Trick, M., Edwards, K. J (2004) Efficient large scale development of microsatellites for marker and mapping applications in Brassica crop species. *Theor Appl Genet* 108: 1103–1112
- Nei M, Li, W. H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76: 5269–5273
- Piquemal J, Cinquin, E., Couton, F., Rondeau, C., Seignoret, E., Doucet, I., Perret, D., Villeger M. J., Vincourt, P., Blanchard, P (2005) Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. *Theor Appl Genet* 111: 1514–1523
- Qian W, Li, M., Meng, J., Frauen, M., Sass, O., Noack, J., Jung, J. (2006) Introgression of genomic components from Chinese Brassica rapa contributes to widening the genetic diversity in rapeseed (*B. napus* L.), with emphasis on the evolution of Chinese rapeseed. *Theor Appl Genet* 113: 49-54
- Rohlf FJ (2006) NTSYS-PC 2.20f. Numerical taxonomy and multivariate analysis system, Exeter Software, Setauket, NY
- Slatkin M (1985) Rare alleles as indicators of gene flow. *Evolution* 39: 53-65
- Slatkin M, Takahata, N. (1985) The average frequency of private alleles in a partially isolated population. *Theoretical Population Biology* 28: 314-331
- UN (1935) Genomic analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn J Bot* 7: 389–452