Genetic Relationships Among *Brassica napus* Crops Based on SSR Markers

Pilar Soengas,1 Pablo Velasco, Guillermo Padilla, Amando Ordás, and María Elena Cartea

Department of Plant Genetics, Misión Biológica de Galicia (CSIC), Apartado 28, E-36080 Pontevedra, Spain

Additional index words. *Brassica napus*, genetic diversity, microsatellites, nabicol landraces, SSRs

Abstract. *Brassica napus* includes economically important crops such as oilseed rape, rutabaga, and leaf rape. Other vegetable forms of *Brassica napus*, namely nabicol and couve-nabiça, are grown in northwestern Spain and north of Portugal, respectively, and their leaves are used for human consumption and fodder. The relationship of nabicol with other *Brassica napus* leafy crops was studied before, but its origin remained unclear. The aims of this work were to study the genetic relationships among nabicol landraces and other *B. napus* crops based on microsatellites and to relate the genotypic differences with the use of the crop. The relationship among 35 *Brassica napus* populations representing different crops was studied based on 16 microsatellite markers. An analysis of molecular variance was performed partitioning the total variance into three components. The source of variation resulting from groups was defined considering the main use of the crop and accounted for a smaller percentage of variation than other sources of variation, proving that this division is not real. Populations clustered into seven different clusters using a similarity coefficient of 0.82. No clear association was evident between clusters and the main use of populations, suggesting genetic differences among populations could reflect differences in their origin/breeding or domestication. Spanish nabicol could have originated from a sample of couve-nabiças, and couve-nabiças could be used to improve nabicol landraces, because they have a narrow genetic basis that limits their potential for breeding.

Different hypothesis have been suggested for the origin of *B. napus*. Sinskaia and Schiemann (Gómez-Campo and Prakash, 1999) have proposed an origin in the Mediterranean region of southwest of Europe, where the two contributing parents (B. oleracea L. and B. rapa L.) overlapped in their natural distribution, probably in an agricultural environment. Song et al. (1988) and Song and Osborn (1992) proposed multiple origins of *B. napus* based on restriction fragment length polymorphism (RFLP) analyses. Repeated backcrosses of the interspecific hybrids to one or both parental populations have been suggested for the origin of some populations of *B. napus* (Palmer et al., 1983).

*Brassica napus* includes economically important crops such as oilseed rape (*B. napus* var. *pabularia* [DC.] Reichenb.). In Galicia (northwestern Spain), a cultivar of *B. napus var. pabularia*, namely ‘nabicol’, is a traditional crop grown by the farmers of the area from many years ago as an important horticultural product during the winter season (Cartea et al., 2005). The leaves are used both for human consumption and fodder. Nabicol landraces have been collected since the 1980s in northwestern Spain in the areas where the crop is important, mainly in the border with Portugal. Populations have been kept as an active collection at the ‘Misión Biológica de Galicia’ (MBG) until the present (Ordás and Baladron, 1985).

The origin of the local nabicol populations grown in Galicia is unclear. Rodríguez et al. (2005) found a low level of genetic diversity among them based on agronomic and morphologic characteristics, and they concluded that could be the result of a common genetic origin of the populations studied. Cartea et al. (2005) studied the relationships among 32 nabicol populations using 16 British populations of forage rape and rape kale based on RAPD markers and concluded that Spanish and British populations have different origin.

Simple sequence repeat (SSR; microsatellite) markers may provide a useful method to characterize, conserve, and use agricultural crop diversity. Identification and characterization of SSR markers have been done in different species of the *Brassica* genus (Lowe et al., 2004; Plieske and Strauss, 2001; Saal et al., 2001; Suwabe et al., 2002; Zwir-McFadden et al., 1996; Westman and Kresovich 1998). Microsatellite markers have been useful to study the genetic diversity of *Brassica* (Westman and Kresovich 1999), to characterize resynthesized rapeseeded lines (Seyis et al., 2003), to distinguish between spring/winter types or genotypes with high/low glucosinolate content (Charters et al., 1996), or to identify sources for improving heterotic potential in oilseed breeding (Hasan et al., 2005).

The genetic basis of nabicol landraces is narrow as was demonstrated by (Rodríguez et al., 2005) and (Cartea et al. 2005). Genetic improvement of nabicol could be done by introducing interesting genes from other *B. napus* crops. Information about genetic relationships of nabicol with other *B. napus* crops could be useful in designing breeding programs to obtain improved nabicol varieties.

The aims of this work are 1) to study the genetic relationships among nabicol landraces and other *B. napus* crops based on SSR markers, and 2) to relate the genotypic differences based on SSR markers with the use of the crop.

Material and Methods

Plant material. Thirty-five *B. napus* cultivars from different geographic origins and with different uses were evaluated in this study (Table 1). To facilitate the description of the material, populations were divided into four sets (leafy crops, oilseed crops, root vegetable crops, and unknown use). The first cluster was formed by nabicol landraces, couve-nabiças, and five cultivars of forage rape and rape kale from diverse origin. Nabicol landraces are from the northwest of Spain and were chosen based on previous morphologic and molecular characterizations (Cartea et al., 2005; Rodriguez et al., 2005) to represent the clusters obtained in those works. Portuguese populations of couve-nabiça are used for human consumption and were chosen to represent different geographic and climatic conditions. The second set included five oilseed *B. napus* populations, including spring and winter canola. Finally, the third and fourth sets consisted of rutabaga and wild populations, respectively (Table 1).

Simple sequence repeat analysis. Forty days after sowing, the 4 or 5 youngest leaves were collected. Leaf samples were taken from 5 plants of each accession and DNA was extracted following the method of Liu and Whittier (1994) with minor modifications. A total of 40 primer combinations were tested for their suitability. After prescreening, 16 primer pairs were chosen that gave clear, reproducible, and polymorphic amplification products at one or more loci in *B. napus*. The sequence of primers Ni2-C12, Ra2-E04, Na12-A02, O110-F11, Na12-E05, Na14-D07, Na14-E08, N12-F02, Na10-A08, O112-F02, O112-F11, Na10-D09, N14-D09, and Na10-F06 was taken from the public domain: www.ukcrop.net. The sequence of primers BRMS-030 and

---

Received for publication 9 Mar. 2006. Accepted for publication 11 Apr. 2006. Research supported by the Committee for Science and Technology of Spain (Project Cod. AGL2003-01366) and Excmu Diputación Provincial de Pontevedra, Spain. We thank David Ashley and Angela Pinigar from Horticulture Research International, (UK) and Noor Bas from the Centre for Genetic Resources (The Netherlands) for providing some populations. G. Padilla acknowledges a fellowship from the Cabildo Insular de La Palma. We thank the invaluable help of Elisa Santos and Rosaura Abiliera for all the laboratory work.

1To whom correspondence should be addressed. e-mail psoenagas@mbg.csrgas.es.
BRMS-037 was taken from Suwabe et al. (2002). Amplifications were performed by using a PTC-100 Thermal Cycler (MJ Research, Watertown, MA). The amplification consisted of a denaturing step at 95 °C for 5 minutes followed by 35 denaturing cycles at 95 °C for 30 seconds, annealing at 56 °C for 30 seconds, and elongation at 72 °C for 30 seconds. The program ended with an extra elongation period of 10 minutes followed by a continuous cycle at 4 °C. Polymerase chain reactions were carried out in a volume of 25 μL containing 50 ng of each primer, 0.6 unit of Taq polymerase (ECOGEN), 200 μM each dNTP, 1× reaction buffer, 2.0 mM MgCl₂, 50 ng DNA template, and distilled and autoclaved water. After amplification, SSR products were separated by electrophoresis on 6% nondenaturing acrylamide gels run in 1× TBE buffer. Gels were stained with ethidium bromide, run approximately for 2 hours at 250 V, and then were visualized under ultraviolet light.

Data analysis. Each SSR band was scored as present (1) or absent (0) for all markers across all populations. The reason to treat codominant markers as if they were dominant is because *B. napus* is an amphidiploid and most of the primers amplified more than one locus, and it was not clear for some bands if they belong to one or the other locus. Only neat and polymorphic bands across populations were scored. Analysis of molecular variance (AMOVA) was performed by the software Arlequin version 2.0 (Schneider et al., 2000). The total variance was partitioned into three sources of variation: among groups, among populations within groups, and within populations. The source of variation among groups was defined considering the main use of the crop: leaf consumption (nabicol, couve-nabica, forage rape, and rape kale), oleiforous (oilseed rape), root consumption (rutabaga), and unknown use (wild populations). The significance level for variance components was tested with permutation tests.

The polymorphic information content for each population and SSR was computed as: \[ PI = 2f(1 - f) \]

where \( PI \) is the polymorphic information content, \( f \) is the frequency of the amplified SSR band in each population, and \( (1 - f) \) is the frequency of the absence of band following Roldan-Ruiz et al. (2000). PIC values were averaged across all loci.

A similarity matrix with the presence–absence data were constructed by the NTSYS-PC version 2.1 (Rohlf, 1998) based on the Dice coefficient, also known as the similarity coefficient of Nei and Li (1979). A band was considered as present in a population if at least was amplified in one individual of the population. The Dice coefficient is computed as: 

\[ D = 2a/(2a + b + c) \]

where \( a \) is the number of SSR bands shared by genotypes in each pairwise comparison and \( b \) and \( c \) are the number of SSR bands present in one genotype and not present in the other. Cluster analysis was performed using the unweighted pair group method with arithmetic averages (UPGMA) with the program NTSYS (Rohlf, 1998). A cophenetic correlation was calculated to test for the goodness-of-fit between similarity matrix obtained from the cluster
Results and Discussion

Fifty-eight SSR bands were scored across all populations. Analysis of molecular variance showed that all three sources of variation were significant (Table 2). The source variation among populations within groups accounted for 50.49% of the total variation, followed by within populations (34.01%) and among groups (15.50%). Larger differences among groups, defined by the main use of the crop, and lower differences within populations were expected because B. napus is considered an autogamous species and the populations studied represented different crops. B. napus has been reported to have a variable outcrossing rate that depends on the genotype and environmental conditions (Becker et al., 1992), and this could be partially responsible for the high level of variability found within populations. The fact that variation among groups is the smallest source of variation provides evidence that the division made according to the end use of the crop is not real and that differences among populations could be rather related to the geographic origin of populations or to their pedigree than to the use of the crop.

Part of the variability was present at the within-population level. The number of polymorphic bands varied between 27 for the wild accession BE062 and the rutabaga ‘Friese Gele’ and 40 for the couve-nabica BE013. PIC values varied between 0.01 for ‘Friese Gele’ and 0.21 for BE013 (Table 1). Couve-nabicas named BE013 and BE031 showed both a high number of polymorphic bands and PIC values (Table 1).

The cophenetic correlation between the similarity matrix based on Dice coefficient and the similarity matrix obtained from the cluster was 0.72. This is not a very high correlation (Rohlf, 1998); therefore, results from the cluster analysis should be taken cautiously. All populations, apart from BE030, grouped together for a similarity coefficient of 0.76. A relatively close relationship was revealed among B. napus cultivars that could be explained because of its recent origin compared with other Brassica species (Gómez-Campo and Prakash, 1999).

Populations grouped into seven different clusters using a similarity coefficient of 0.82 (Fig. 1). The first cluster consisted on 23 populations. Most of the couve-nabica and nabical populations were together in cluster I, but they were differentiated in two subclusters (Fig. 1): subcluster Ia comprised 13 populations and was formed by most of the nabical populations from Spain, together with one Spanish commercial forage rape, one Spanish commercial rutabaga (‘Colinabo Rocalba’), 2 Portuguese commercial couve-nabicas (‘César Santos’ and ‘Couve-Nabica’) and one winter oilseed rape (Fig. 1). Subcluster Ib consisted on four accessions of couve-nabica, besides one nabical, three oilseed raps, and two forage raps. Cluster II was composed of two forage rape populations and a wild accession. A single population of nabical (MBG-BRS0054) appeared on cluster III. Cluster IV was composed of 2 couve-nabicas. In cluster V, one couve-nabica plus a wild population and an oilseed rape were grouped together. Two rutabagas were classified into cluster VI. Finally, one single accession of couve-nabica comprised VII.

Most nabical populations were similar, because all of them except MBG-BRS0054 were grouped together, confirming previous studies. A low level of genetic diversity among the nabical landraces was found at the molecular (Cartea et al., 2005) and morphologic (Rodriguez et al., 2005) levels. These authors concluded that the origin of these landraces must be in common. Population MBG-BRS0054 classified far away from the rest of the nabical populations in this work. This population showed some alleles that were not present in the rest of nabical populations but were in couve-nabicas.

Taking into account the main use of the populations, most the nabical and couve-nabica were in cluster I. All the bands present in nabical populations were in the couve-nabica ones, but 7 bands present in couve-nabicas were not in nabical populations (data not shown), suggesting that nabical from the northwest of Spain could have originated from a subgroup of couve-nabicas. Both crops are very similar morphologically and the regions where they are grown are quite close. Commercial nabical sold by seed companies are landraces that have not been selected, except for selection carried out by growers. The variability present in couve-nabicas could be used in breeding to improve nabical landraces, because they have a narrow genetic basis that limits its potential for breeding (Cartea et al., 2005; Rodriguez et al., 2005).

Forage rape ‘Colza Rocalba’ was in cluster Ia very close to the commercial population named ‘Nabical’. ‘Colza Rocalba’, sold as ‘forage rape’, could be a nabical population. The origin of the commercial seed is unknown, but it could be possible that they were collected on the same or very close areas and sold under different names. The rest of forage rape populations (‘Vysokopolskij’ and ‘Mara’) were classified into cluster Ib; rape kale populations (‘Ragged Jack kale’ and ‘Russian kale’) were grouped apart on cluster II. They are also morphologically distinct, because rape kales present wavy leaf margins. They are used for human consumption in salads or boiled. Forage raps are harvested in the autumn to supplement the grass feed for cattle or for sheep. Sometimes both groups are not clearly distinguished because rape kales can be also denominated forage rape. Forage rape and rape kales were also classified separately in previous works based on molecular markers (Cartea et al., 2005; Hasan et al., 2005). Rape kales, represented by the variety ‘Asparagus kale’, could have an independent origin from the rest of rapeseed and rutabagas studied by Song et al. (1988) and Song and Osborn (1992), and it would be more related to B. rapa than other B. napus populations. Based on this work and previous ones, rape kale populations could have a different origin than the rest of B. napus crops, including nabical and couve-nabica.

Winter oilseed rape ‘Belinda’ was in cluster Ia very close to ‘Couve-Nabica’. ‘Couve-Nabica’ could be obtained from an oilseed rape variety, because a previous work (E. Cartea, unpublished data) showed a very low level of erucic acid content in seeds, similar to commercial populations of oilseed rape. Three oilseed rape populations were in cluster Ib, ‘Valle del oro’ and ‘Express’ winter crops and ‘Petranova’, a spring crop. Spring and winter populations usually classified separately, because they constitute two genetically different groups (Diers and Osborn, 1994; Hasan et al., 2005; Lombard et al., 2000; Seys et al., 2003). The fact that ‘Valle del oro’, ‘Express’, and ‘Petranova’ classify together could be the result of a common genetic background.

Rutabaga populations, ‘Friese Gele’ and ‘Lollo’, were in cluster IIa that was very close to the rest of the groups and the third rutabaga accession, ‘Colinabo Rocalba’, was genetically closer to other nabical populations (cluster I) than to the rest of rutabagas. Other authors have found that rutabagas cluster together (Diers and Osborn, 1994; Song and Osborn, 1992). As mentioned before, ‘Friese Gele’ and ‘Lollo’ had a low level of variability. ‘Colinabo Rocalba’ had a high level of variability and probably it includes in its pedigree local germplasm from the northwest of Spain.

Wild BE061 clustered with the rape kales and BE062 with oilseed rape populations and a couve-nabica. Following Gómez-Campo and Prakash (1999), there are not wild forms of B. napus in nature, although it often occurs as an escape. Wild populations BE061 and BE062 could be escapes from cultivated populations instead of truly wild genotypes, because at least BE062 has a low level of genetic diversity and both populations classified with other morphotypes.

Table 2. Analysis of molecular variance (AMOVA) for 35 populations of Brassica napus.

| Source of variation | df | Sum of squares | Variance components | Percentage of variation | P value1 |
|---------------------|----|----------------|---------------------|------------------------|---------|
| Among groups        | 3  | 61.2           | 0.45                | 15.5                   | <0.0001 |
| Among populations within groups | 31 | 257.2       | 1.47                | 50.5                   | <0.0001 |
| Within populations  | 139| 176.6         | 0.99                | 34.0                   | 0.00098 |
| Total               | 173| 496.0          | 2.91                |                        |         |

1Sum of squares, variance components, percentage of variation, and P value for the sources of variation. Probability to obtain a random value ≥ observed value using 1023 permutations.
No clear association was evident between clusters formed and the main use of the populations, because different crops are mixed in several clusters. To classify populations, differences between types of cultivars are not as important as differences in the rest of the genome that does not determine the use of the crop. This could be because differences in the use of the plant are determined by few genes and the number of SSR markers used in this work is not enough to detect their effect. Zhao et al. (2005) in a molecular characterization of B. rapa populations found that different morphotypes are often more related to other morphotypes from the same region than similar morphotypes from different regions, suggesting an independent origin or separate breeding/domestication.

In conclusion, genetic differences based on SSR markers among the populations analyzed are not related with their use as different crops. These differences probably reflect their origin breeding/domestication instead. Spanish nabicol is closely related to the Portuguese couve-nabicás and it could have originated from a sample of these. Couve-nabicas could be used to improve nabicol landraces, because nabicol has a narrow genetic basis that limits its potential for breeding.

**Literature Cited**

Becker, H.C., C. Damgaard, and B. Karlsson. 1992. Environmental variation for outcrossing rate in rapeseed (Brassica napus). Theor. Appl. Genet. 84:303–306.

Cartea, M.E., P. Soengas, A. Picoaga, and A. Ordás. 2005. Relationships among Brassica napus (L.) germplasm from Spain and Great Britain as determined by RAPD markers. Genet. Resour. Crop Evol. 52:655–662.

Charters, Y.M., A. Robertson, M.J. Wilkinson, and G. Ramsay. 1996. PCR analysis of oilseed rape cultivars (Brassica napus L. ssp. oleifera) using 5’-anchored simple sequence repeat (SSR) primers. Theor. Appl. Genet. 92:442–447.

Diers, B.W. and T.C. Osborn. 1994. Genetic diversity of oilseed Brassica napus germplasm based on restriction fragment length polymorphism. Theor. Appl. Genet. 88:662–668.

Gómez-Campo, C. and S. Prakash. 1999. Origin and domestication. Pp. 33–58. In: C. Gómez-Campo, (eds). Biology of Brassica cospecies. Elsevier Science B.V. Amsterdam. The Netherlands.

Hasan, M., F. Seys, A.G. Badani, J. Pons-Kühnemann, W. Friedt, W. Lühs, and R.J. Snowden. 2005. Analysis of genetic diversity in the Brassica napus L. gene pool using SSR markers. Genet. Resour. Crop Evol. in press.

Liu, Y.G. and R.F. Whitter. 1994. Rapid preparation of megabase plant DNA from nuclei in agarose plugs and microbeads. Nucleic Acids Res. 22:2168–2169.

Lombard, V., C.P. Baril, P. Dubreuil, F. Blouet, and D. Zhang. 2000. Genetic relationships and fingerprinting of rapeseed cultivars by AFLP: consequences for varietal registration. Crop Sci. 40:1417–1425.

Lowe, A.J., C. Moule, M. Trick, and K.J. Edwards. 2004. Efficient large-scale development of microsatellites for marker and mapping applications in Brassica crop species. Theor. Appl. Genet. 108:1103–1112.

Nei, M. and W.H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5260–5273.

Ordás, A. and J.J. Baladrón. 1985. Collecting Brassicas in northwestern Spain. Crucifer. Newsl. 10:14.

Palmer, J.D., C.R. Shields, D.B. Cohen, and T.J. Osborn. 1983 Chloroplast DNA evolution and the origin of amphidiploid Brassica species. Theor. Appl. Genet. 65:181–189.

Plieske, J. and D. Strauss. 2001. Microsatellite markers for genome analysis in Brassica. I. Development in Brassica napus and abundance in Brassicaceae species. Theor. Appl. Genet. 102:689–694.

Roldán-Ruiz, I., E. Calisy, T.J. Gillilan, R. Coll, M.J.T. van Eijk, and M. De Loose. 2000. Estimating genetic conformity between related ryegrass (Lolium) populations. 2. AFLP characterization. Mol. Breeding. 6:593–602.

Rodriguez, V.M., M.E. Cartea, G. Padilla, P. Velasco, and A. Ordás. 2005. The nabicol: a horticultural crop in northwestern Spain. Euphytica 142:237–246.

Rohlf, F.J. 1998. NTYSYS-pc Numerical Taxonomy and Multivariate Analysis System. Version 2.1. EXETER software. Setauket, New York.

Saal, B., J. Plieske, J. Hu, C.F. Quros, and D. Strauss. 2001. Microsatellite markers for genome analysis in Brassica. II Assignment of rapeseed microsatellites to the A and C genomes and genetic mapping in Brassica oleracea L. Theor. Appl. Genet. 102:695–699.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin ver 2.0: a software for population
genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. Available at: http://anthro.unige.ch/arlequin. Accessed June 2, 2006.

Seyis, F., R.J. Snowdon, W. Luhs, and W. Friedt. 2003. Molecular characterization of novel re-synthesized rapeseed (Brassica napus) lines and analysis of their genetic diversity in comparison with spring rapeseed cultivars. Plant Breed. 122:473–478.

Song, K.M., T.C. Osborn, and P.H. Williams. 1988. Brassica taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). Theor. Appl. Genet. 75:784–794.

Song, K. and T.C. Osborn. 1992. Polyphyletic origins of Brassica napus: new evidence based on organelle and nuclear RFLP analyses. Genome 35:992–1001.

Suwabe, K., H. Iketani, T. Nunome, T. Kage, and M. Hirai. 2002. Isolations and characterization of microsatellites in Brassica rapa L. Theor. Appl. Genet. 104:1092–1098.

Szewe-McFadden, A.K., S. Kresovich, S.M. Bliek, S.E. Mitchell, and J.R. McFerson. 1996. Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated Brassica species. Theor. Appl. Genet. 93:534–538.

Westman, A.L. and S. Kresovich. 1998. The potential for cross-taxon simple-sequence repeat (SSR) amplification between Arabidopsis thaliana L. and crop brassicas. Theor. Appl. Genet. 96:272–281.

Westman, A.L. and S. Kresovich. 1999. Simple sequence repeat (SSR)-based marker variation in Brassica nigra genebank populations and weed populations. Euphytica 109:85–92.

Zhao, J., X. Wang, B. Deng, P. Lou, J. Wu, R. Sun, Z. Xu, J. Vromans, M. Koornneef, and G. Bonnema. 2005. Genetic relationships within Brassica rapa as inferred from AFLP fingerprints. Theor. Appl. Genet. 110:1301–1314.