

Identification of Carnation Varieties Using Microsatellite Markers

P. Arens^{1, a}, D. Esselink¹, Y. Noordijk¹, L. Kodde¹, L. Hof², W. Wietsma² and B. Vosman¹

¹Plant Breeding, Wageningen UR, P.O. Box 16, 6700 AA Wageningen, The Netherlands

²Unit Varieties & Trials, Naktuinbouw, P.O. Box 40, 2370 AA Roelofarendsveen, The Netherlands

Keywords: SSR markers, mutants, ploidy level

Abstract

As in many ornamentals, also in carnation the number of varieties in common knowledge is large and identification throughout the chain from breeder to consumer using plant material from different stages and organs may be needed. Results in this study on the use of microsatellite markers from *Dianthus caryophyllus* L. for the characterization of carnation varieties as well as the construction and evaluation of a molecular database show that these markers show potential for identification purposes.

INTRODUCTION

As in many ornamentals, also in carnation the number of varieties in common knowledge is large. Variety registration and protection in carnation is based on morphological characteristics, described in the UPOV Test Guidelines (document TG/25/8). Identification throughout the chain from breeder to consumer depends on the availability of methods that can use plant material from different stages and organs. Furthermore, in situations where a suspicion of an infringement of a PBR has been raised, there may be a need for a quick and reliable comparison of varieties before a lengthy trial for comparison of DUS characteristics is considered. Therefore, complementary approaches, such as the use of microsatellite markers, are being evaluated and molecular databases are being constructed. Microsatellites are highly polymorphic and have the advantage of providing co-dominant markers based on PCR technology. In a sequenced tagged microsatellite site (STMS) approach, they produce simple banding patterns, especially suitable for automated and objective description of plant varieties which are easy to store in a database. New varieties or new markers can easily be added to an existing database.

In this paper, we present the results of our study on the use of microsatellite markers from *Dianthus caryophyllus* L. for the characterization of carnation varieties as well as the construction and evaluation of a molecular database.

MATERIALS AND METHODS

Plant Material

In total, 172 samples were analysed of which 133 samples were sent in by breeding companies and 12 samples were reference varieties from the carnation DUS test of 2005. Furthermore 27 samples (including 12 reference varieties) were analysed in duplex adding up to the total of 172 samples. The set of samples was chosen in such a way that it encompassed a selection of common varieties as well as some mutant groups.

Analysis of this set of samples was performed without prior knowledge of variety names and relationships.

Microsatellite Markers, Development of New Markers and Detection of Polymorphism

From previous studies (Smulders et al., 2000, 2003) 15 polymorphic microsatellite markers were available and these have been tested for their suitability in database

^a paul.arens@wur.nl

building. Furthermore, microsatellites were isolated from enriched small-insert genomic libraries essentially as described by Karagyozov et al. (1993) and modified by Van der Wiel et al. (1999). For initial screening of STMS markers, PCR amplification products were separated on a 6% acrylamide gel and visualized with silver staining according to the Promega Silver sequence DNA sequencing system as described by Wiel et al. (1999). STMS markers that showed well scorable peak patterns were chosen for characterisation of the collected set of varieties using an ABI 3700 DNA sequencer (Applied Biosystems). Only markers that showed no or very little stutter peaks and high quality patterns were selected into the final set of markers.

A reference allele system as previously described for tomato (Bredemeijer et al., 2002) was used i.e. the 12 reference varieties from the 2005 DUS test were present on both microtiter plates for consistency of allele scoring. The alleles of each sample were scored and scores transferred to a numeric database in an Excel spreadsheet for storage and further analyses.

Data Analysis

Co-dominant scoring of the markers in heterozygote samples to assess the complete genotype of a variety was complicated by the polyploid nature of carnation. Among carnation varieties ploidy levels can differ from diploid, triploid, tetraploid and hexaploid. Moreover aneuploidy is also common within carnation varieties.

Although differences were found in the amounts of product for different alleles, it turned out to be very difficult to use these differences to estimate whether a particular allele was present one, two or three times, and in this way deduce the actual genotype of a variety. Therefore only the allele composition was scored and recorded in a presence/absence matrix. We refer to this as the allelic phenotype (Becher et al., 2000).

Using the Jaccard similarity coefficient and the unweighted pair group method using arithmetic averages (UPGMA) module of NTSYS, version 2.1 the varieties were clustered and results analysed.

RESULTS

Isolation and Characterization of Microsatellite Markers

The 15 polymorphic microsatellite markers available (Smulders et al., 2000, 2003) were supplemented with 28 newly developed markers. The set of 12 reference varieties was used for pattern quality assessment. Many markers, in particular the ones with a dinucleotide microsatellite, showed considerable to severe stuttering and were not amendable for use. Thirteen markers were selected on the basis of pattern quality (quality 1 or 2), using the standards described by Smulders et al. (1997). The selected markers were used for fingerprinting the set of 172 samples.

Marker Characteristics

The initial screening of the primer combinations was performed on a set of 12 varieties in which each marker was analyzed separately. To increase effectiveness of analyses, markers were subsequently tested in multiplexes of three or four markers to see which markers could be amplified together. Twelve markers could be amplified into four multiplexes of three markers. The thirteenth marker could only be amplified in simplex, but could be added to multiplex three afterwards for analysis. Table 1 shows some characteristics of the markers used in this study. For all STMS markers, allele sizes differed by one or multiple repeat units from each other, i.e. alleles were separated by steps of 2 bp for dinucleotide repeats or 3 bp for trinucleotide repeats. However, marker DINMADSBOX showed one allele that did not follow this relation. This suggests that another type of sequence variation is also involved in allelic diversity at this locus which complicated the allele designation of the three alleles involved of this dinucleotide microsatellite marker. Another important characteristic of the STMS markers observed is the occurrence of differences in allele intensity in heterozygous varieties. This was a

generally observed phenomenon among the STMS markers tested. Although differences exist between the selected loci, most alleles at each locus could be easily identified because the relative peak area of the less intense band was usually well over 15% of the main peak. This difference in amplification is partly due to differences in the number of copies of the alleles present, but may also depend on differences in amplification efficiency of alleles. Two STMS markers are probably located at chromosomes that are involved in aneuploidy. Marker DC09 showed five or six alleles in four tetraploid cultivars (all pot carnations) in repeated tests. Marker DC105 consistently showed three alleles in 17 diploid cultivars in separate tests.

Variety Identification Using STMS Markers

In order to calculate the discriminating power of the markers a database was built with the results of all 172 samples. Two to 10 alleles were generated by each STMS marker, and in total, 84 alleles were amplified (Table 1).

The numbers of allelic marker phenotypes varied from 3 to 20 (Table 2) and, in total, 163 marker phenotypes were found.

In total, 27 a priori known duplicate samples were present: in 16 duplicate samples, allelic phenotypes were identical; in 10 duplicate samples one allele was scored differently between duplicates; and in 1 duplicate sample, two alleles were scored differently. Almost all differences were due to selective amplification of alleles, stutter peaks and constant bands, combined with a threshold for peaks in allele designation. As a safeguard, a threshold of two alleles difference or less was used to group samples. This threshold avoids two varieties being considered to be different on the basis of technical errors.

Based on the threshold of two allele's difference, 19 groups of samples could be identified that can not be distinguished from each other. In this phase, sample variety information was obtained and compared with the results of the clustering analysis to study the consequences of the chosen threshold.

Out of the 19 identified groups, 6 groups consisted of samples from a variety which was put into the sample set twice without our knowledge. Furthermore, 10 groups could be identified which contained varieties that were color mutants. In all cases, duplicate samples and mutants from initial varieties were correctly grouped together.

Without the duplicate samples and mutants, 118 samples remained. Out of the 118 potentially different samples, 112 had unique profiles, taking into account the threshold of 2 alleles. The 7 remaining varieties fell into three groups for which the clustering could not be explained with the information available. Analyses of new plant material of these varieties gave exactly the same results. The first group (A) consisted of 3 varieties, of which one variety is a known mutant of one of the other two, that all showed an identical allelic phenotype (numbers 104, 137 and its mutant 133 in Figure 1). The second group consists of two varieties (sample 105, same variety as 47; put into the sample set twice without prior notification) from the same breeder that showed one allele difference (group C in Fig. 1; numbers 47 and 115). The third group consisted of five original varieties and one mutant that showed a difference of two alleles or less between them (Fig. 1 group B, numbers 130, 143, 225, 184, 219 and its mutant 218).

The product of the frequencies of the most common allelic phenotypes from all markers is the upper limit of the probability that two samples are identical by chance alone. This upper probability of an identical pattern in unrelated samples was 0.0000016. Average similarity between samples (0.46) was low, indicating a high level of genetic diversity between samples.

DISCUSSION

As for many crops, hundreds of carnation varieties are registered. Identification of a variety from all others depends on the ability of the molecular markers used to produce a unique fingerprint. For this, it is necessary that the markers used are highly discriminative. When analyzing data obtained from marker analysis, it is essential that a

threshold is used to prevent two varieties from being considered different on the basis of technical errors. In addition, in future this threshold may be used as a way to introduce a minimum distance. Without the mutants and duplicate samples, 118 potentially different varieties were present in the sample set. From these 118 varieties, 112 unique patterns were identified: in pair-wise comparisons, 99.9% of the original varieties were discriminated. Only three groups, involving the clustering of nine varieties, could not be explained with the available information. A similar result was reported by Smulders et al. (2003), who also found a clustering of 2 varieties to a group of mutant varieties that could not be explained based on knowledge of the varieties origin. In view of the low average similarity between varieties in general, these findings warrant further study. It is possible that these high genetic similarities between seemingly different varieties are due to the fact that, with the polyploid nature of carnation, differences in the allelic composition of varieties are not distinguished because only allelic phenotypes can be recorded. Differences resulting from differences in allele dosage are not taken into account. Alternatively, the clear morphological difference that exists may be due to only very few genetic differences and varieties may actually be very closely related. Finally, it is also possible that varieties are extremely similar to each other and can not be identified uniquely with the markers used. Other markers are available (Smulders et al., 2000, 2003), although they were not considered suitable for database construction. With the set of 13 markers that were selected in this study, all known mutants and their original varieties were correctly grouped together. Since mutants may result from a single point mutation, it is extremely unlikely that they can be discriminated from their original variety because the chance that such a mutation would involve one of the microsatellite markers used in this study is remotely small. Average similarity between all samples was 0.46, which is low, and indicates a high level of genetic diversity in carnation. Therefore, the set of microsatellite markers described in this study can be used for genotyping carnation varieties for identification purposes. In addition, the markers may be used to spot putative mutants among candidate varieties. For practical applications, it is recommended to analyze all samples in duplicate and reanalyze all duplicates that show differences in alleles before their addition to the database. In this way, the threshold value can be lowered and accurate grouping of samples can be achieved.

ACKNOWLEDGEMENTS

This research was funded by the Ministry of Agriculture, Nature and Food quality of the Netherlands under the method development scheme for plant breeders' rights research.

Literature Cited

- Becher, S.A., Steinmetz, K., Weising, K., Boury, S., Peltier, D., Renou, J.-P., Kahl, G. and Wolff, K. 2000. Microsatellites for cultivar identification in *Pelargonium*. *Theor. Appl. Genet.* 101:643-651.
- Bredemeijer, G.M.M., Cooke, R.J., Ganal, M.W., Peeters, R., Isaac, P., Noordijk, Y., Rendell, S., Jackson, J., Röder, M.S., Wendehake, K., Dijcks, M., Amelaine, M., Wickaert, V., Bertrand, L. and Vosman, B. 2002. Construction and testing of a microsatellite database containing more than 500 tomato varieties. *Theor. Appl. Genet.* 105:1019-1026.
- Karagyozov, L., Kalcheva, I.D. and Chapman, V.M. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acids Res.* 21:3911-3912.
- Smulders, M.J.M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P. and Vosman, B. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theor. Appl. Genet.* 97:264-272.
- Smulders, M.J.M., Rus-Kortekaas, W. and Vosman, B. 2000. Microsatellite markers useful throughout the genus *Dianthus*. *Genome* 43:208-210.

- Smulders, M.J.M., Noordijk, Y., Rus-Kortekaas, W., Bredemeijer, G.M.M. and Vosman, B. 2003. Microsatellite genotyping of carnation varieties. *Theor. Appl. Genet.* 106:1191-1195.
- Wiel, C., Arens, P. and Vosman, B. 1999. Microsatellite retrieval in lettuce (*Lactuca sativa* L.). *Genome* 42:139-149.

Tables

Table 1. Characteristics of the 13 selected microsatellite markers for carnation.

Multiplex	Locus	Repeat motif	No. of alleles
1	DCD224 (Fam)	(CTT) ₁₀	7
	DCD105 (Hex)	(TCT) ₂₆	9
	DC14 (Ned)	(GAT) ₇ (GCC) ₇	5
2	DC16 (Fam)	(TGA) ₁₉	8
	DCF005 (Hex)	(TGTTTGT) ₅	8
	DINMADSBOX (Ned)	(TA) ₇	6 ¹
3	DC12 (Fam)	(AACCT) ₃ (CGG) ₆	2
	DINCARACC (Hex)	(TA) ₈	7
	DC09 (Ned)	(TGA) ₁₁ (GAT) ₅	10
	DC22 (Fam)*	(TCT) ₁₅	3
4	DC06 (Fam)	(AAC) ₆	5
	DC10 (Hex)	(GCG) ₁₃₋₂	7
	DC27 (Ned)	(CGG) ₅	7

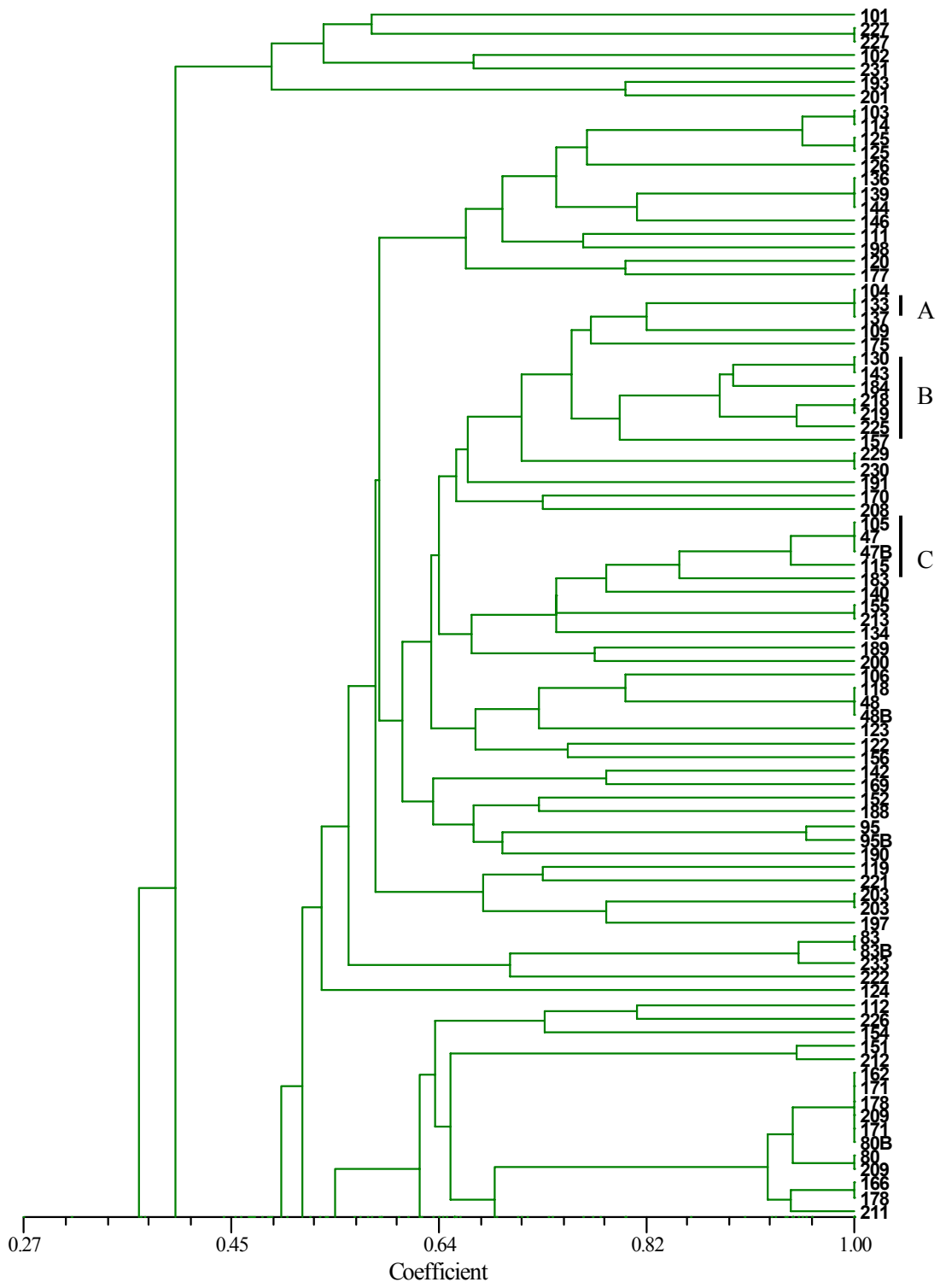
*Locus DC22 is separately amplified, and after mixing, analyzed together with other loci.

¹Presence of one allele that through a 1bp insertion/deletion does not fit with the other alleles in a dinucleotide array.

Table 2. Total number of detected allelic phenotypes for each locus.

Locus	No. of allelic phenotypes	PIC value based on allelic phenotypes	Frequency of most common allelic phenotype	No. of different alleles in most common allelic phenotype
DCD224	20	0.77	0.42	2
DCD105	17	0.79	0.32	2
DC14	10	0.79	0.35	1
DC16	12	0.52	0.67	1
DCF005	15	0.70	0.45	1
DINMADSBOX	15	0.79	0.35	2
DC12	3	0.63	0.44	2
DC9	17	0.78	0.33	1
DINCARACC	14	0.71	0.49	1
DC22	5	0.42	0.74	1
DC6	9	0.64	0.54	1
DC27	16	0.61	0.55	1
DC10	10	0.34	0.81	1

Figures



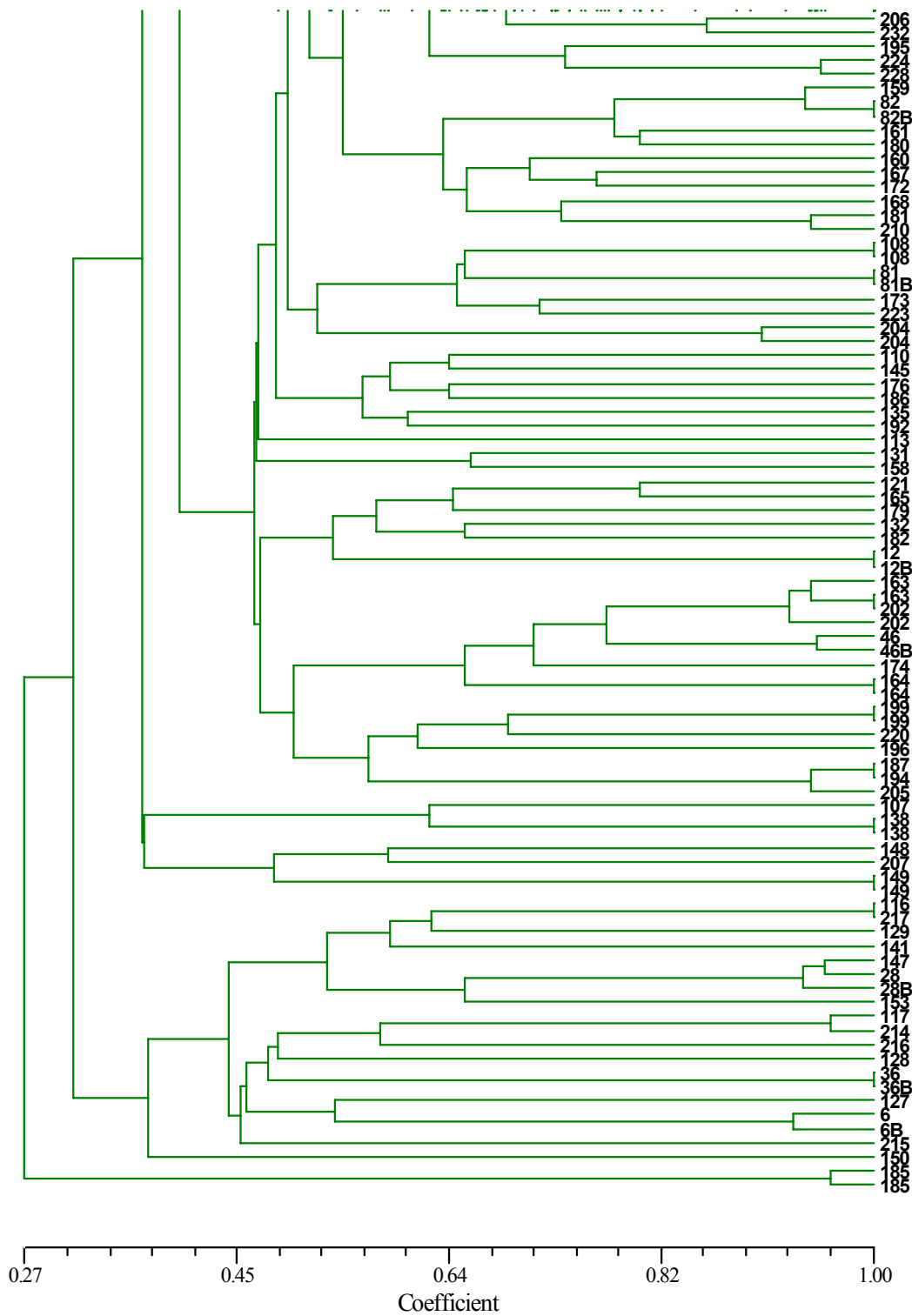


Fig. 1. Cluster analysis (UPGMA) of the carnation varieties. Numbers represent different samples, replicate samples have identical numbers. Duplicate samples from the reference varieties are indicated with a B after the number. Groups which clustering could not be explained with the information available are indicated (A, B, C).

