Genetic diversity and determination of the optimum number of RAPD markers in lettuce (*Lactuca sativa* L.)

Flávio Dessune Tardin[1], Antônio Teixeira do Amaral Júnior[1]*, Messias Gonzaga Pereira[1], Maria Celeste Gonçalves Vidigal[2], Rogério Figueiredo Daher[1] and Carlos Alberto Scapim[2]

[1]Laboratório de Melhoramento Genético Vegetal, Centro de Ciências e Tecnologias Agropecuárias, Universidade Estadual do Norte Fluminense, Av. Alberto Lamego, 2000, Horto, 28015-620, Campos dos Goytacazes, Rio de Janeiro, Brazil.
[2]Departamento de Agronomia, Universidade Estadual de Maringá, Av. Colombo, 5790, 87020-900, Maringá, Paraná, Brasil.
**Author for correspondence. e-mail amaraljr@uenf.br*

ABSTRACT. Twenty lettuce accessions from the UENF Germplasm Collection were assessed for genetic diversity using RAPD markers. Multivariate statistical techniques were used, such as grouping analyses, using the Tocher and Single Linkage methods. Analysis of 55 polymorphic markers, obtained with 25 primers, showed that there was sufficient variability in the material studied to be exploited in future breeding programs. The 'BGH 4060' and 'Grand Rapids' accessions were the most dissimilar, while 'BGH 4325' and 'BGH 4326' were the most similar. There was high agreement between ecogeographic origin and molecular similarity. It was further found by the stress statistic (Kruskal, 1964) and by the correlation between number of markers and increase in ideal formation of groups that 50 polymorphic markers is the optimum number for genetic diversity study in the assessed accessions.

Key words: *Lactuca sativa*, genetic diversity, RAPD markers.

RESUMO. Variabilidade genética em alface (*Lactuca sativa* L.) e determinação do número ótimo de marcas RAPD para estudos de diversidade molecular. Vinte acessos de alface da Coleção de Germoplasma da UENF foram avaliados quanto à diversidade genética, pelo uso de marcadores RAPD. Foram empregadas técnicas estatísticas multivariadas, como análises de agrupamento, utilizando-se os métodos de Tocher e 'Hierárquico do Vizinho Mais Próximo'. Pela análise de 55 fragmentos polimórficos obtidos com 25 'primers', os acessos BGH 4060 e Grand Rapids foram os mais dissimilares, enquanto Regina e Repolhuda Todo Ano foram os mais semelhantes. Constatou-se elevada concordância entre a origem ecogeográfica e a similaridade molecular. Verificou-se, ainda, que 50 fragmentos polimórficos foram o número ótimo de marcadores para uma análise segura da diversidade genética nos acessos avaliados.

Palavras-chave: *Lactuca sativa*, diversidade genética, marcadores RAPD.

Introduction

Lettuce, which originated in the Mediterranean basin, spread with the Roman conquests to France, England and the rest of Europe. With the discovery of the New World in 1492, it spread rapidly to the Americas and was introduced in Brazil in 1647 (Ryder, 1986).

Currently, it is a popular vegetable throughout the world and one of the most important in economic terms. It is part of the Brazilian diet, and it is the most important leafy vegetable (Alvarenga, 1999). It has a high pro-vitamin A content, 400 U.I. in 100 grams of green leaves, about four times the content found in tomatoes (Sonnenberg, 1985).

In spite of being the second most commercialized leafy vegetable in the State of Rio de Janeiro, Brazil, according to Ceasa (wholesale vegetable market) data of 1998, lettuce has not been investigated for genetic diversity; thus, the producers use a reduced number of genotypes which compete for the vulnerability of the varieties planted.

Thus, any research that can contribute to widen the knowledge on the species genetic diversity is necessary and justifiable. The DNA makers of the

RAPD type are a valuable option, because they consist of relatively simple procedures, with the advantages of not being subject to environmental influence or epistatic interactions and offering high sensitivity in detecting variability.

This study was carried out to investigate the genetic-molecular diversity among 20 *L. sativa* accessions from the Germplasm Collection at the Center of Agricultural Sciences and Technology (CCTA) at the Universidade Estadual do Norte Fluminense (UENF), Rio de Janeiro, Brazil, using the RAPD markers technique. Furthermore, by the RAPD technique, the optimum number of polymorphic markers recommendable for a safe analysis of the genetic diversity in lettuce was determined to cut costs in the development of future studies.

Material and methods

Genetic materials

Twenty lettuce (*Lactuca sativa* L.) accessions were assessed from the Germplasm Collection at the Center of Agricultural Sciences and Technology (CCTA) at the Universidade Estadual do Norte Fluminense (UENF), state of Rio de Janeiro and are described in Table 1. Seedlings were produced in extruded polystyrene trays with drainage holes, previously filled with organic-vegetable substrate.

Table 1. Lettuce accessions with their respective identification and origin

Accession	Origin / Company	Classification[1/]
01. BGH 44	-	B
02. BGH 292	Sarzedo – MG, Brazil	B
03. BGH 303	-	B
04. BGH 410	Inhaúmas – GO, Brazil	C
05. BGH 502	-	B
06. BGH 2471	Copenhaguen, Denmark	R
07. BGH 2517	Copenhaguen, Denmark	L
08. BGH 2625	Valence, France	B
09. BGH 2629	Valence, France	R
10. BGH 3291	The Netherlands	L
11. BGH 4059	Santa Teresa – ES, Brazil	C
12. BGH 4060	Santa Teresa – ES, Brazil	B
13. BGH 4064	Colatina – ES, Brazil	B
14. BGH 4325	Ipameri – GO, Brazil	L
15. BGH 4326	Ipameri – GO, Brazil	L
16. Grand Rapids	Feltrin	L
17. Maravilha de Verão	Top Seed	B
18. Mimosa	Feltrin	L
19. Regina	Feltrin	L
20. Repolhuda Todo Ano	Feltrin	B

[1/] B = Butterhead; C= Crisphead; L = Looseleaf; and R = Roman.

DNA extraction

Healthy leaf samples were removed in bulk from twenty plants per accession, seven days after planting, while still in the extruded polystyrene tray, and immediately placed in plastic bags, properly identified by labels and submerged in liquid nitrogen. The samples thus obtained were taken to the Plant Genetic Breeding Laboratory (LMGV) of the CCTA/UENF and kept at –85°C in an ultrafreezer. The methodology, according to Doyle and Doyle (1990), was used to extract the DNA with modifications implemented by the LMGV.

DNA amplification, fragment separation by electrophoresis and band detection

One hundred and thirteen primers from Operon Technologies were tested, of which the following were used because they gave adequate resolution: OPC 05, OPC 11, OPH 12, OPI 11, OPN 12, OPO 04, OPR 04, OPR 13, OPT 08, OPV 01, OPV 06, OPV 10, OPV 15, OPV 19, OPW 04, OPW 05, OPW 09, OPY 02, OPAE 14, OPAE 15, OPAF 03, OPAF 10 and OPAF 11.

The amplification reactions were performed according to Willians *et al*. (1990), with modifications. For this, each reaction contained 12.05 μL distilled water, 2.50 μL buffer 10 X (Tris HCl a 10 mM, pH 8.3; KCl at 50 mM); 2.00 μL $MgCl_2$; 1.25 μL dNTPs; 0.20 μL of *Taq polimerase enzyme*; 2.0 μL of one primer; and 5.0 μL DNA. After amplification, the DNA fragments underwent electrophoresis in 1.4% agarose gel. The electrophoresis was performed by submerging the gel in TAE buffer (Tris base, sodium acetate, 0.5 M EDTA and distilled water) 0.5X at 100 volts for 3h, using as standard the size of the lambda bacteriophage DNA, from cleavages with the restriction enzymes *Bam Hl*, *Eco Rl* and *Hind III*.

After electrophoresis, the gels were stained for 30 minutes in ethyl bromide solution (75 μL of ethyl bromide / liter of distilled water) and photographed under ultraviolet light, using the Stratagene 'Eagle Eye' photodocumentation system.

Statistical analysis of RAPD markers

The polymorphic RAPD markers were used to make a matrix of binary data, by attributing **1** to the presence and **0** to the absence of a determined band. The distance among the pairs of accessions was calculated based on the Arithmetic Complement of the Jaccard Index (Amaral Júnior and Thiébaut, 1999) expressed by:

$$C_{ij} = 1 - \frac{a}{a+b+c}$$

where **a** represents the number of DNA fragments, codified with **1** (positive agreement) common to both the individuals and **b** and **c** registered the

number of DNA fragments, in which both the individuals disagree, respectively represented by **1-0** and **0-1**.

The matrix of binary distances was used to group the genotypes by the Tocher and Single Linkage methods using the GENES computer program (Cruz, 1997). In the Tocher grouping, the first group was formed by the pair with the lowest distance value (C_{ij}). From then on, new groups were formed adopting the criterion of the mean intragroup distance being lower than any intergroup distance. The group identification process by the Single Linkage Method was performed by successive identifications of the closest genotypes, starting from the most similar pair until a dendrogram was established (Cruz and Regazzi, 1997).

The optimum number of RAPD markers was also determined for studying the dissimilarity of the genotypes assessed by correlation among the number of markers and the increase in the groups ideal formation, and also by the stress statistic (Kruskal, 1964), based on the mean from 20 markers combinations, sampled 20 times, from the fifteenth polymorphic. The GQMOL computer program was used, developed by the Genetic Sector at the Universidade Federal de Viçosa, Minas Gerais, Brazil.

Results and discussion

Polymorphic fragments

Table 2 shows that there was 67.03% polymorphism among the amplified fragments; in a total of 82 developed bands, 55 were polymorphic.

Grouping analysis

Six genotypes groups were formed in the Tocher grouping (Table 3). The group I, the largest, contained twelve genotypes. Groups II, III and IV contained two genotypes each, while groups V and VI contained only one genotype each. The greatest genetic similarity was observed between genotypes 19 (Regina) and 20 (Repolhuda Todo Ano), which were in group I together with the genotypes 1 (BGH 44), 3 (BGH 303), 7 (BGH 2517), 9 (BGH 2629), 6 (BGH 2471), 2 (BGH 292), 18 (Mimosa), 15 (BGH 4326), 14 (BGH 4325) and 8 (BGH 2625).

Genotype 11 (BGH 4059) was one of the most divergent and formed group VI alone. Also, genotype 5 (BGH 502) alone made up group V.

It is interesting to point out that there was high agreement in this grouping between the ecogeographic origin and molecular similarity. The latter consideration is supported by the fact that the great majority of the genotype pairs from the same locality was not disassociated in distinct groups. For example: introduction 2 (BGH 292) and 3 (BGH 303), 6 (BGH 2471) and 7 (BGH 2517), 8 (BGH 2625) and 9 (BGH 2629), and 14 (BGH 4325) and 15 (BGH 4326), derived, respectively from Sarzedo (MG, Brazil), Copenhagen (Denmark), Valence (France) and Ipameri (GO, Brazil), which were present in group I.

Table 2. Oligonucleotids primers used and their respective base sequences and number of the fragments associated to them

Primers	Sequence 5' - 3'	Number of Polymorphic Fragments	Number of Monomorphic Fragments
OPC 05	GATGACCGCC	1	3
OPC 11	AAAGCTGCGG	4	0
OPH 12	ACGCGCATGT	2	0
OPI 11	ACATGCCGTG	3	1
OPN 12	CACAGACACC	2	2
OPO 04	AAGTCCGCTC	2	0
OPR 04	CCCGTAGCAC	2	2
OPR 13	GGACGACAAG	4	0
OPT 08	AACGGCGACA	0	1
OPV 01	TGACGCATGG	2	2
OPV 06	ACGCCCAGGT	1	0
OPV 10	GGACCTGCTG	1	2
OPV 15	CAGTGCCGGT	2	0
OPV 19	GGGTGTGCAG	2	0
OPW 04	CAGAAGCGGA	4	3
OPW 05	GGCGGATAAG	1	1
OPW 09	GTGACCGAGT	5	0
OPY 02	CATCGCCGCA	3	1
OPAE 01	TGAGGGCCGT	2	1
OPAE 07	GTGTCAGTGG	2	2
OPAE 14	GAGAGGCTCC	0	2
OPAE 15	TGCCTGGACC	2	0
OPAF 03	GAAGGAGGCA	3	0
OPAF 10	GGTTGGAGAC	1	1
OPAF 11	ACTGGGCCTC	4	3
TOTAL		55	27

Table 3. Grouping by the Tocher method of 20 *Lactuca sativa* genotypes based on 55 polymorphic RAPD fragments

Group	Genotypes [1/]
I	19, 20, 1, 3, 7, 9, 6, 2, 18, 15, 14, 8
II	10, 12
III	4, 16
IV	13, 17
V	5
VI	11

[1/] 1 = BGH 44, 2 = BGH 292, 3 = BGH 303, 4 = BGH 410, 5 = BGH 502, 6 = BGH 2471, 7 = BGH 2517, 8 = BGH 2625, 9 = BGH 2629, 10 = BGH 3291, 11 = BGH 4059, 12 = BGH 4060,13 = BGH 4064, 14 = BGH 4325, 15 = BGH 4326, 16 = Grand Rapids, 17 = Maravilha de Verão, 18 = Mimosa, 19 = Regina, 20 = Repolhuda Todo Ano

Five groups were established by the Single Linkage method, shown in Figure 1, taking as base the discrepant changes of level in the dendrogram, where group I united the genotypes belonging to groups I and II by the Tocher method. Genotype 5 (BGH 502) formed the second dendrogramic group, while in Tocher it was present on group V. Genotypes 4 (BGH 410) and 16 (Grand Rapids) formed the third group in both methods. The fourth group by the Single Linkage method was composed

by genotype 11 (BGH 4059) and the fifth group, made up of the genotypes 13 (BGH 4064) and 17 (Maravilha de Verão). These groups were equivalent, respectively, to groups VI and IV by the Tocher method. Thus, the results obtained by the Single Linkage method were very similar to those by the Tocher method.

The perfect genetic similarity between genotypes 19 (Regina) and 20 (Repolhuda Todo Ano) was not expected as these materials are phenotypically divergent in the literature, and it can be quoted for example, that Regina is classified as a classic example of a looseleaf variety while, as the name indicates, the 'Repolhuda Todo Ano' is a classic example of the butterhead type. However, during the course of this study, the 'Regina' variety plants morphoagronomically assessed showed the cabbage formation in the field, which called attention at the time. The most plausible explanation for this is a possible error made in the introduction of the seeds of this material in the UENF Germplasm Collection, raising the hypothesis that the genotypes 19 (Regina) and 20 (Repolhuda Todo Ano) may constitute identical genomes.

Six of the 25 primers used in the present study, OPC 05, OPC 11, OPI 11, OPN 12, OPO 04 and OPW 09 were also used by Kesseli *et al*. (1994) to construct a genetic linking map in lettuce. These primers, in the study by Kesseli *et al*. (1994) amplified seven distinct regions of the genome, of which five were present in different linkage groups. Considering that the authors identified eight large linkage groups (>70 cM) it may be supposed comparatively that the present study covered a large part of the lettuce genome.

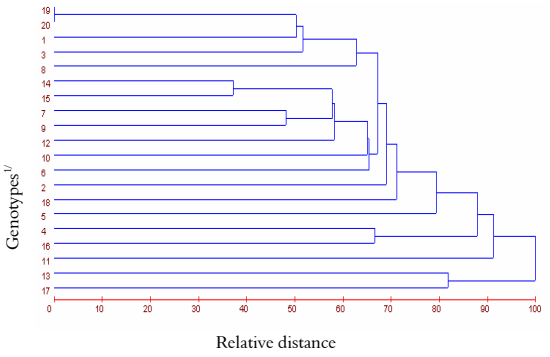

Figure 1. Genetic dissimilarity among 20 *Lactuca sativa* genotypes, by the Single Linkage method, based on 55 polymorphic fragments and 27 monomorphic fragments using the Arithmetic Complement of the Jaccard Index. 1/ 1 = BGH 44, 2 = BGH 292, 3 = BGH 303, 4 = BGH 410, 5 = BGH 502, 6 = BGH 2471, 7 = BGH 2517, 8 = BGH 2625, 9 = BGH 2629, 10 = BGH 3291, 11 = BGH 4059, 12 = BGH 4060,13 = BGH 4064, 14 = BGH 4325, 15 = BGH 4326, 16 = Grand Rapids, 17 = Maravilha de Verão, 18 = Mimosa, 19 = Regina, 20 = Repolhuda Todo Ano

Table 4. Estimates of the mean values of the correlation among the number of markers and the increase in the ideal formation of groups and mean of the stress coefficient estimate

Number of markers	Mean of the estimate of	
	Correlation	Stress coefficient
15	0.6907	0.1600
16	0.6575	0.1616
17	0.6789	0.1505
18	0.7345	0.1410
19	0.7147	0.1380
20	0.7460	0.1263
21	0.7191	0.1303
22	0.7365	0.1321
23	0.7921	0.1987
24	0.7836	0.1176
25	0.8226	0.1046
26	0.7994	0.1075
27	0.8141	0.1040
28	0.8093	0.0987
29	0.8243	0.0933
30	0.8393	0.0884
31	0.8428	0.0921
32	0.8640	0.0813
33	0.8760	0.0839
34	0.8650	0.0821
35	0.8898	0.0723
36	0.8801	0.0749
37	0.9065	0.0680
38	0.8942	0.0695
39	0.9119	0.0636
40	0.9115	0.0635
41	0.9271	0.0592
42	0.9284	0.0585
43	0.9381	0.0556
44	0.9359	0.0509
45	0.9552	0.0461
46	0.9540	0.0450
47	0.9631	0.0406
48	0.9644	0.0388
49	0.9730	0.0353
50	0.9762	0.0318
51	0.9807	0.0282
52	0.9856	0.0248
53	0.9917	0.0195
54	0.9948	0.0146
55	1.0000	0.0000

Recommendable number of polymorphic markers

The mean of the correlation among the number of markers and the increase in the ideal formation of groups (Table 4) show mean values varying from 0.6575 and 1.000, which occurred, respectively, when 15 and 55 markers were sampled. When 0.05 was adopted as the cutting point for the stress statistic, it was initially concluded that, in the present study, the use of 45 to 51 primers would be sufficient to express the variability present. However, for the samples of 45, 46, 47, 48 and 49 primers, at least one sample presented stress values, for the estimate of the mean over 0.05, and therefore it was not considered as a potential optimum number of markers to detect the variability in the studied population. But 50 was the number of markers which did not present any sample values higher than 0.05 for the composition of the estimated mean of statistical stress, and therefore it

was considered the optimum number of markers to investigate the genetic diversity in the studied material. For this magnitude of markers, the mean correlation was 0.9762, which, because it was close to 1.0000, portrays the good precision of the estimates obtained with the primers used, which will cut costs in future researches.

References

ALVARENGA, M.A.R. *Crescimento, teor e acúmulo de nutrientes em alface americana (Lactuca sativa L.) sob doses de nitrogênio aplicadas no solo e de níveis de cálcio aplicados via foliar.* Lavras, 1999. Tese ((Doutorado em Agronomia) - Universidade Federal de Lavras, Lavras, 1999.

AMARAL JÚNIOR, A.T.; THIÉBAUT, J.T.L. *Análise multivariada na avaliação da diversidade em recursos genéticos vegetais*. Campos dos Goytacazes: UENF. 1999.

Ceasa. Centrais de Abastecimento do Estado do Rio de Janeiro: Desempenho da comercialização nas unidades de Rio de Janeiro: Ceasa-RJ. 1998.

CRUZ, C.D. *Programa Genes*: aplicativo computacional em genética e estatística. Viçosa: Imprensa Universitária, 1997.

CRUZ, C.D.; REGAZZI, A.J. *Modelos biométricos aplicados ao melhoramento genético*. Viçosa: UFV/Imprensa Universitária. 1997.

DOYLE, J.J.; DOYLE, J.L. Isolation of plant DNA from fresh tissue. *Focus*, v. 12, p. 13-15, 1990.

KESSELI, R.V. *et al*. Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. *Genetics,* Bethesda, v.136, p.1435-1446, 1994.

KRUSKAL, J.B. Multidimensional scaling by optimizing goodness of fit to on non-metric hypothesis. *Psycometrika*, v.29, p.1-127, 1964.

RYDER, E.J. Lettuce breeding. In: BASSET, M.J. (Ed.). *Breeding vegetable crops*. Gainesville, Florida: AVI Publishers. 1986. p.436-476.

SONNENBERG, P.E. *Olericultura especial*: alface, alho, cebola, cenoura, batata e tomate. Goiânia: UFG, 1985.

WILLIAMS, J.G.K. *et al*. DNA polymorfism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res., Oxford,* v.18, p.6531-6535, 1990.

Received on May 17, 2002.
Accepted on July 14, 2002.