E

BMT/8/7
ORIGINAL: English
DATE: August 14, 2003

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA-PROFILING IN PARTICULAR

Eighth Session
Tsukuba, Japan, September 3 to 5, 2003

GENETIC VARIABILITY AND RELATIONSHIPS AMONG SUNFLOWER INBRED LINES FROM A FRENCH REFERENCE COLLECTION ASSESSED USING A SET OF WELL-CHARACTERIZED SSR MARKERS

Document prepared by experts from France

GENETIC VARIABILITY AND RELATIONSHIPS AMONG SUNFLOWER INBRED LINES FROM A FRENCH REFERENCE COLLECTION ASSESSED USING A SET OF WELL-CHARACTERIZED SSR MARKERS.

D Zhang[1], LS Zhang[1♦], V. Cellier[2], F. Blouet[2], J. Guiard[2]

1. BioGEVES. Le Magneraud. BP52. 17700 SURGERES, France
2. GEVES LA MINIERE, 78285 GUYANCOURT CEDEX, France
♦present address: Horticultural College, China Agricultural University
Beijing 100094, China.

Abstract:

The present paper presents at first the results on the assessment of molecular variability among 124 sunflower inbred lines, including 67 female maintainers (M lines) and 57 male restorers (R lines), using 78 SSR markers. A total of 281 alleles were obtained across the 124 elite inbred lines, with a mean of 3.5 alleles / SSR locus. The number of alleles per SSR ranged from 2 to 9. The Polymorphism Information Content (PIC) values varied from 0.06 to 0.81, with an average of 0.51. Relationships between the inbred lines were studied by the estimation of Rogers' distance. The general mean of distance values for M+R lines is 0.54, with a min of 0.013 and a max of 0.79. The great majority of the distance estimates is between 0.4 and 0.6. We will also show some results on intra-line variability study.

Introduction

The number of protected varieties and parent lines in sunflower increases year by year; more than 600 inbred lines are now presented in the French DUS reference collection. The genetic basis used for sunflower breeding tends to become narrower and narrower by the frequent use of same genetic resources for the common breeding objectives (e.g. grain yield and resistances). Furthermore, sunflower is a plant very sensitive to the genotype x place x year interaction; the phenotype of a same plant material may vary greatly according to the place and the year of growing. All these factors render more and more difficult the accurate description of the new inbred lines undergoing the DUS testing. The GEVES has been involved in the development of new descriptors in sunflower, in particular molecular markers (Gentzbittel et al. 1994, Teulat et al. 1994, Zhang et al. 1995), since number of years. Among the different markers systems available actually, microsatellite or Simple Sequence Repeat (SSR) is the marker system of choice for plant variety description and identification because of their simplicity for use, general high level of polymorphism, good coverage of genome, quality of information (single locus, co-dominance, reproducibility). Recently, SSR markers have been also developed and reported for genotyping inbred lines and genetic mapping in sunflower (Paniego et al. 2002, Yu et al. 2002, Tang et al. 2002, Tang and Knapp 2003). Moreover, the SSR primer pairs developed in these researches are publicly available.

The present study aimed (1) to select a set of SSR markers suitable for genotyping and (2) to assess the variability within and between sunflower inbred lines selected in the French DUS testing reference collection using this selected set of SSRs.

Materials and methods

SSR primer pairs used: SSR primer pairs used in our study have been developed under a collaborative research project "CARTISOL", involving private seed companies and public institutes (INRA and GEVES). The development work was ensured by the teams of S.J. Knapp (Oregon State University. USA) for the SSRs named "ORS-" and K. Edwards (University of Bristol, UK) for the SSRs named "SSL-". All the primers of the SSR markers named "ORS" have been published and publicly accessible (Tang et al. 2002). On the contrary, these named "SSL-" are property's markers which belong to "G.I.E. CARTISOL". All contacts for the access to these SSRs should be addressed to GIE-CARTISOL (20, rue Bachaumont 75002 Paris. France).

SSR analysis conditions: SSR analysis conditions have been previously described (Zhang and Becquet, 2002), using a LI-COR automated DNA analysis system.

Plant material: Ten public inbred lines, HA89, H52, HA372, HA383, HA821, RHA274, RHA377, RHA801, PAC2 and RHA266, were used for evaluating the functionality as well as the quality of SSRs in our conditions. Variability between lines was assessed on 124 inbred lines selected from the reference collection maintained in GEVES for sunflower DUS testing in France. Among them, there are 67 female maintainer lines coded from M1 to M105 and 57 male lines (Restorer) coded from R1 to R80; the numbers after M or R are not continuous. Among these 124 samples, some of them have special relationships: (1) the form A (male sterile) and the form B (maintainer of male sterility) of a same female inbred line, (2) the normal version and the downy mildew resistant version of a same inbred line and (3) two seed lots of a same inbred line supplied as reference seed for DUS testing by seed companies in two different time periods. Leaves were harvested from 10 individual plants per inbred line in the fields at GEVES LE MAGNERAUD, in the summer of 2000 and 2001, and pooled for DNA extraction. The harvested leaves were stocked in a freezer at -20°C before use. For extracting the genomic DNA, leaves were lyophilized and grounded manually in a mortar. About 100 mg of the flour were submitted for DNA isolation, using the DNeasy Plant Mini kit (QIAGEN) for all the plant materials used in this study.

Data analysis: Gel analysis and data scoring were performed using the software RFLPscan^{TM} (LI-COR); data were manually validated. The molecular data generated were analyzed using the software LCDMV (Logiciel de Calcul de Distance Moléculaires entre Variétés), developed by the GEVES using SAS tools. This software allowed (1) the calculation of the value of Polymorphism Information Content (PIC) and of the genetic distance of Rogers (1972) and (2) the elaboration of UPGMA (un-weighted pair-group method average) dendrograms.

Results and discussion

Selection of a set of SSR markers for genotyping sunflower inbred lines: A total of 1111 primer pairs amplifying SSR loci were available for us in the end of CARTISOL project. Only a subset of 323 best primer pairs have been selected for the present study, according to the information supplied by the two developer laboratories; their quality was evaluated across 10 public inbred lines. Only 103 primer pairs out of the 323 corresponded to the criteria that

we fixed: clear amplification, simple profiling and single co-dominant locus. Then a validation trial has been organized to evaluate the robustness of these 103 SSRs. Four laboratories from 4 seed companies participated to this trial: EURALIS GENETIQUE, MONSANTO SAS, RAGT and SYNGENTA. The validation trial consisted of performing analysis of the 103 SSRs on the same 10 sunflower inbred lines with DNA supplied by our laboratory, each participating laboratory using its own working and analysis conditions. Among the four laboratories, two use a LI-COR automated DNA analysis system and the other two use gel electrophoresis system. After comparisons of results obtained by the 4 participating laboratories with those obtained by us, we fund

* 28 SSRs showing the same profiling in all participating laboratories,
* 5 SSR showing the same profiling in 4 of the 5 participating laboratories,
* 19 SSR showing the same profiling in 3 of the 5 participating laboratories.

It should be indicated there are more identical profiling in the 3 laboratories (Euralis génétique, RAGT et BioGEVES) where the same plate-form (LI-COR) has been used.

After the comparison of SSR profilings between laboratories, 78 SSRs have been retained and proposed for further genotyping use in sunflower (Table 1). All these retained SSR markers are co-dominant single locus and have a correct reproducibility.

Polymorphism level of the selected SSRs: The 78 SSRs generated a total of 276 alleles across the 124 sunflower inbred lines analyzed. The Table 1 gives information about the number of alleles as well as the PIC value for each SSR marker. The number of alleles per SSR locus varied from 2 to 9, with an average of 3.5. The PIC value ranged from 0.06 to 0.81, with a mean of 0.51. These results are very similar to those reported by Hundley et al. (2002), Paniego et al. (2002), Yu et al. (2002) and Tang et Knapp (2003), using different sets of sunflower inbred lines. This level of molecular polymorphism is comparable with that revealed using RFLPs among sunflower lines (Gentzbittel et al. 1994, Zhang et al. 1995).

All the 78 SSRs do not have the same utility for routine genotyping work and for variety identification in sunflower, based on the PIC value in Table 1. Generally, the higher the PIC value is and the more powerful the SSR marker is for discriminating the sunflower varieties. So it will be preferable to use those which show big value of PIC. On the contrary, it should avoid the use of those SSRs with low PIC values, in particular ORS617, ORS559, ORS779, ORS428 and ORS609 that have a PIC value less than 0.20.

Level of fixation of SSR loci and inbred lines: Table 2 shows that 51 out of 78 SSR loci (63.75%) are well fixed across all the 124 inbred lines. The rest of 29 SSR loci detected heterogeneity at least on one inbred line. Three SSRs, ssl13, ors151 and ors337, show highest level of no fixation; they detected heterogeneity in 8 and 9 lines respectively. It will be preferable to avoid the use of those SSRs which detect high level of heterogeneity among inbred lines, for genotyping and variety identification proposes. Among the 124 inbred lines studied, 79 (63.71%) are well fixed for all the 80 SSRs analysed; 45 sunflower lines are not fixed at least on one SSR locus. However, 32 lines of the 45 show heterogeneity only at one SSR locus. The highest level of heterogeneity concerns one line where 6 SSR loci are not fixed (7.5%).

Genetic relationships between the inbred lines: General information about the estimations of Rogers distance among the inbred lines is presented in Table 3. The general mean of distance values for M+R lines is 0.54, with a min of 0.013 and a max of 0.79. When the

estimations of genetic distances were made separately for M lines and R lines, the average of distance values is 0.5 for M lines and 0.52 for R lines. The great majority of the distance estimates is between 0.4 and 0.6. In fact, a lot of the inbred lines selected for this study are original lines and represent the whole range of morphophysiological variability observed in the French DUS reference collection. These results confirm our previous results obtained with RFLPs (Zhang et al. 1995).

The distance indices for the particular pairs of inbred lines are varied according to situations. These pairs of lines can be classed as 5 situations:

- Situation 1 - Form A and B of a same line : two pairs of lines are in this case. The distance varied from 0.03 to 0.06. Normally, the difference between the A and B lines is only the gene responsible for male sterility.

- Situation 2 – Two supplies of reference seed lot of a same line in two different time periods: We have 7 cases in this situation. The distances varied from 0.02 to 0.08, with a mean of 0.043. As the breeder continues to maintain his plant material after registration, small difference may occur between the seed lots of two supplies. Different allele forms have been already observed on sunflower inbred lines by isozyme analysis during the routine check of reference seed lots performed in our laboratory.

- Situation 3 – Two repetitions of a same line in the DUS reference collection: There is only one pair in this case. One difference has been observed out of 78 SSR loci analyzed. Some heterogeneity was observed in the fields on these two samples.

- Situation 4 – Two versions (normal VS modified) of a same line: There is only one pair of lines. it concerns the pair R8/R9. R8 is an original line and R9 is the modified version in which a downy mildew resistance gene has been introduced via traditional backcross. Their distance is 0.13.

- Situation 5 – Pairs of lines are more or less close phenotypically : We have 16 pairs of lines which have been declared close based on morphological investigations in the fields. For these pairs of lines, their molecular distances varied from 0.04 to 0.48, with a mean of 0.25. These results recall the triangular relationships between molecular distances and morphological distances already observed and described in maize (Bar-Hen and Charcosset. 1995) and in oilseed rape (Lombard, 2000) : paris of lines (or genotypes) having small morphological distances may have small or big molecular distances. The raisons leading to this situation have been explained by the authors.

Classification of sunflower inbred lines *:* Based on the Rogers distances, an UPGMA dendrogram showing the relationships among the 124 sunflower inbred lines analyzed was constructed (Figure 1). On the dendrogram, the R lines were grouped at the top and M lines at the bottom. Within each big group, there are several subgroups. However, two M lines (M 68 and M 105) have been classified among the R lines and four R lines (R35, R 53, R 25 and R 37) were mixed among the M lines. All these lines have original phenotype. This kind of mis-classifications for some inbred lines has already been reported in sunflower (Zhang et al. 1995, Paniego et al. 2002).

Variability within line: Four lines and 30 individual plants per line were analysed using 50 SSR markers (the even numbers from 1 to 100 out of the 103 SSRs preselected). This study was realised by the four seed companies (EURALIS Génétique, RAGT, MONSANTO SAS

and SYNGENTA); each company analysed only one of the four lines. The results obtained are the following:

	Line 1	**Line 2**	**Line 3**	**Line 4**
Number of SSRs analysed	43	41	40	47
Number of SSRs non fixed	1	4	1	1
% of SSRs fixed	97.7	90.2	95.1	97.9

Heterogeneity has been detected on all the four inbred lines : 1 SSR locus on 3 lines and 4 SSR loci on one line. This represents a level of homogeneity, in term of fixed SSR loci, of 90.2% to 97.9%. It should be indicated that there is no documented heterogeneity, based on field investigations. These results show that the four inbred lines studied have been correctly fixed at molecular level, even if the breeding and varietal selection had been based on phenotypic characteristics.

The present study permitted the establishment of a very efficient SSR analysis conditions in sunflower, using a LI-COR automated DNA analysis system. A set of 80 SSR loci were screened and selected of more than one thousand of SSRs available. All these SSR markers are single simple co-dominant markers. This is extremely important for genotyping work, especially for variety identification. The robustness of these SSRs was assessed by performing a ring trial involving 4 other laboratories. This set of SSR markers has been characterized across a set of 124 sunflower inbred lines representing the morphophysiological variability observed in the French reference collection used for DUS testing. Our study showed that the great majority of the selected SSRs have a very good polymorphism level among the sunflower inbred lines analysed. Moreover, the SSRs demonstrated also their utility in the study of genetic relationships among sunflower lines. Among these 80 SSR markers, those, well fixed and with a big PIC value, will be very useful for variety identification in sunflower, for example variety description, purity testing, hybrid formula verification, variety identity checking in the process of seed certification.

Acknowledgements
This research program was funded by a grant of three years (1999 to 2002). The French inter-professional association AMSOL is our partner in this project. Four seed companies, EURALIS GENETIQUE. MONSANTO SAS. RAGT and SYNGENTA, participated and contributed to the validation work of the selected SSR markers as well as the intra-line variability study.

References

Bar-Hen A, Charcosset A (1995) Relationship between molecular and morphological distances in a maize inbred lines collection. Application for breeders' rights protection. In Biometrics in plant breeding: application of molecular markers. pp57-66.

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Research 19:4008.

Gentzbittel L, Zhang YX, Vear F, Griveau B, Nicolas P (1994) RFLP studies of genetic relationships among inbred lines of the cultivated sunflower. Helianthus annuus L.: evidence for distinct restorer and maintainer germplasm pools. Theor Appl Genet 89: 419-425.

Hundley M, Hu XY, Scott T, Ernst C, Cui C, Gupta M, Thompson S (2002) Utilization of SSR markers to determine genetic similarity and heterotic association in *Helianthus annuus*. Plant, Animal & Microbe Genomes X. post 201.

Leclercq P (1966) Une stérilité mâle utilisable pour la production d'hybride simples de tournesol. Ann Amélior Plant 16 :134-144.

Lombard V (2002) Estimation de la proximité génétique de variétés de colza sur la base de marqueurs moléculaires: consequences pour l'inscription et la protection variétale. Thèse doctorale de l'Institut National Agronomoque Paris-Grignon. pp136.

Paniego N, Echaide M, Munoz M, Fernandez L, Torales S, Faccio P, Fuxan I, Carrera M, Zandomeni R, Suarez EY, Hopp HE (2002) Microsatellite isolation and charaterization in sunflower (*Helianthus annuus* L.). Genome 45 : 34-43.

Rogers JS (1972) Measures of similarity and genetic distance. Studies in genetics VII. U. Texas Publ. 7216:145-153.

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2003) Simple sequence repeat map of the sunflower genome. Theor Appl Genet 105: 1124-1136.

Tang S, Knapp SJ (2003) Microsatellites uncover extraordinary molecular genetic diversity in native American land races and wild population of cultivated sunflower. Theor Appl Genet (in press).

Teulat B, Zhang YX, P Nicolas (1994) Characteristics of random amplified polymorphic DNA markers discriminating *Helianthus annuus* inbred lines. Agronomie, 14: 497-502.

Yu JK, Mangor J, Thompson L, Edwards KJ, Slabaugh MB, Knapp SJ. 2002 Allelic diversity of simple sequence repeat markers among elite inbred lines in cultivated sunflower. Genome 45: 652-660.

Zhang D, Becquet V (2002) Towards the optimisation of SSR analysis conditions in sunflower using a LI-COR automated DNA analysis system. Plant, Animal & Microbe Genomes X, post 202.

Zhang YX, Gentzbittel L, Vear F, Nicolas P (1995) Assessment of inter- and intra-inbred line variability in sunflower (Helianthus annuus) by RFLPs. Genome 38: 1040-1048.

Table 1 Number of alleles and the value of PIC of the 78 SSRs obtained by genotyping the 124 sunflower inbred lines

N°	SSR	Number of alleles	PIC	N°	SSR	Number of alleles	PIC
1	Ors10	2	0.34	41	Ors613	5	0.62
2	Ors146	3	0.52	42	Ors617	3	0.06
3	Ors151	3	0.66	43	Ors621	5	0.63
4	Ors170	2	0.26	44	Ors656	5	0.51
5	Ors176	9	0.81	45	Ors674	4	0.60
6	Ors203	4	0.69	46	Ors675	2	0.46
7	Ors303	2	0.29	47	Ors677	3	0.21
8	Ors307	3	0.53	48	Ors691	4	0.62
9	Ors309	2	0.48	49	Ors716	3	0.33
10	Ors310	4	0.71	50	Ors727	5	0.47
11	Ors317	4	0.62	51	Ors779	2	0.12
12	Ors329	2	0.41	52	Ors781	2	0.30
13	Ors337	2	0.32	53	Ors788	4	0.74
14	Ors338	2	0.19	54	Ors810	2	0.48
15	Ors342	4	0.42	55	Ors811	4	0.62
16	Ors380	3	0.62	56	ssl13	4	0.22
17	Ors407	3	0.43	57	ssl15	5	0.65
18	Ors428	2	0.13	58	ssl171	4	0.62
19	Ors432	3	0.52	59	ssl216	4	0.46
20	Ors437	4	0.72	60	ssl231	4	0.66
21	Ors442	3	0.55	61	ssl241	4	0.64
22	Ors486	3	0.63	62	ssl255	4	0.74
23	Ors502	3	0.38	63	ssl283	5	0.76
24	Ors509	4	0.43	64	ssl29	4	0.68
25	Ors510	3	0.37	65	ssl3	4	0.70
26	Ors513	2	0.44	66	ssl30	8	0.81
27	Ors53	3	0.61	67	ssl38	4	0.72
28	Ors533	4	0.58	68	ssl51	2	0.38
29	Ors543	5	0.71	69	ssl54	7	0.45
30	Ors546	3	0.57	70	ssl63	5	0.68
31	Ors547	4	0.68	71	ssl81	5	0.60
32	Ors559	2	0.11	72	ssl84	5	0.73
33	Ors57	2	0.48	73	ssl85	3	0.57
34	Ors591	2	0.30	74	ssl89	2	0.48
35	Ors595	4	0.71	75	ssl9	5	0.74
36	Ors605	4	0.66	76	Ors121	3	0.60
37	Ors609	2	0.16	77	Ors159	2	0.44
38	Ors610	5	0.47	78	Ors215	3	0.32
39	Ors7	3	0.51				
40	Ors78	4	0.63				
	Mean					**3.5**	**0.51**

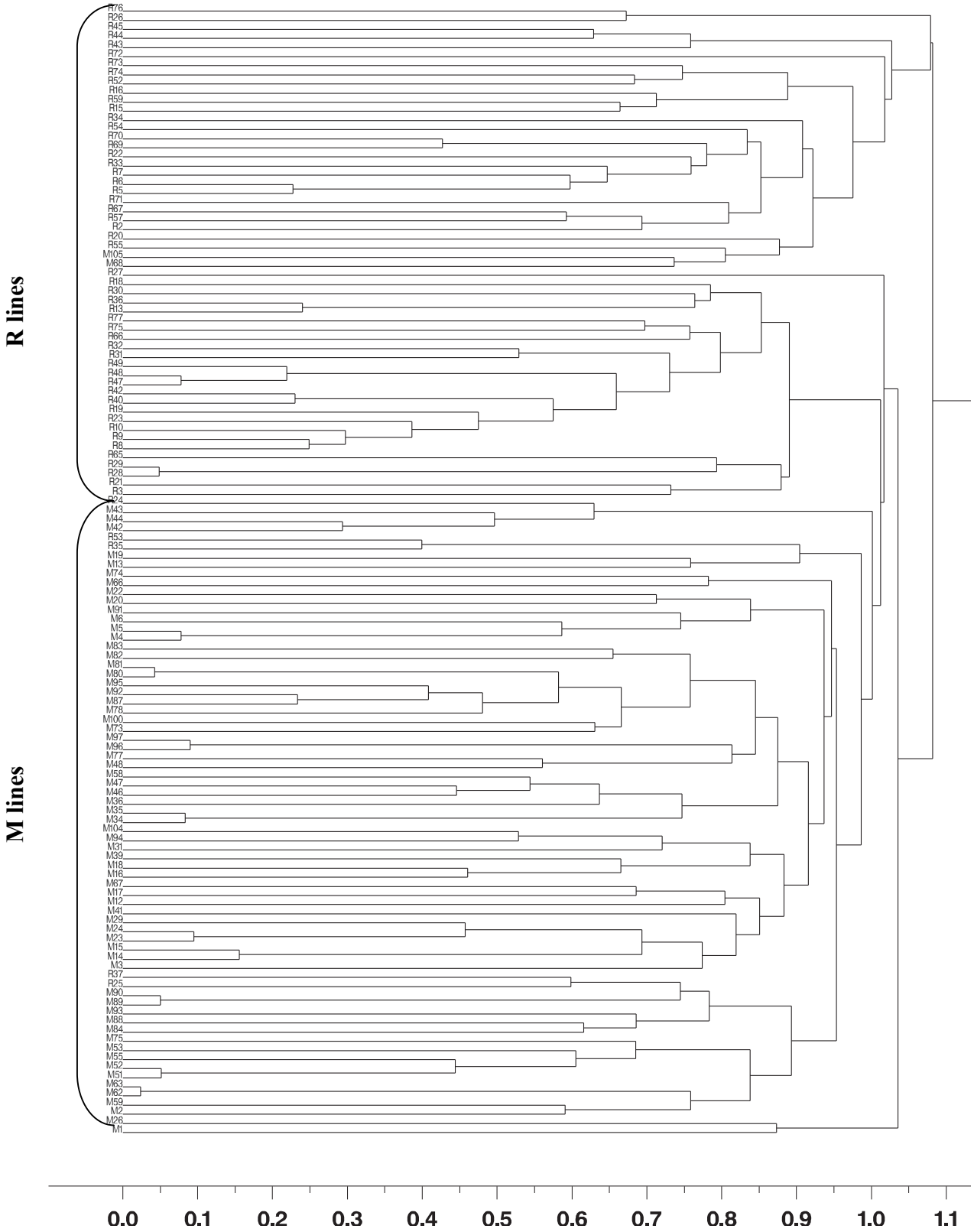
Table 2 Level of fixation of SSR loci across the inbred lines analyzed

Level of fixation	Number of SSRs	%	SSRs marker
Fixed for all lines	51	63.75	
Not fixed on 1 line	13	16.25	
Not fixed on 2 lines	8	10	ors170, ors176, ors329, ors486, ors547, ors610, ors159, ors121
Not fixed on 3 lines	3	3.75	ors502. ssl9. ssl54
Not fixed on 4 lines	2	2.5	ssl30. ssl63
Not fixed on 5 lines	1	1.25	ssl13
Not fixed on 8 lines	1	1.25	ors151
Non fixed on 9 lines	1	1.25	ors337

Table 3 General Information of the estimations of Rogers distances

	Number of lines	Distance of Rogers			
Group of lines		Moyen	SD	Min	Max
M lines	67	0.5	0.0597	0.013	0.74
R lines	57	0.52	0.0575	0.026	0.79
M+R lines	124	0.54	0.0587	0.013	0.79

Figure 1 UPGMA Dendrogram showing groupings of the 124 sunflower inbred lines constructed according to the Rogers' distances estimated using data generated using 78 SSR markers



[End of document]