Euphytica (2005) 146: 109–116
DOI: 10.1007/s10681-005-0548-0

Efficiency of different PCR-based marker systems in assessing genetic diversity among winter rye (*Secale cereale* L.) inbred lines

Hanna Bolibok*, Monika Rakoczy-Trojanowska, Aneta Hromada & Robert Pietrzykowski
*Warsaw Agricultural University, Warsaw, Poland (*author for correspondence: e-mail: tapatik@yahoo.com)*

Received 30 November 2004; accepted 5 July 2005

Key words: genetic diversity, ISSR, rye, SAMPL, SSR

Summary

The aim of this study was to investigate the efficiency of ISSR, SSR, and SAMPL marker systems in detecting genetic polymorphism among 30 winter rye inbred lines and to compare the results of cluster analysis performed on data from these marker systems using different statistical methods and coefficients. Each marker system was able to discriminate among the materials analyzed with the lowest value of average genetic similarity (GS) obtained with ISSR markers (0.2888) and the highest with SAMPLs (0.5381). EST-derived SSRs turned out to be less efficient in detecting genetic diversity than those from genomic libraries (average GS values 0.3814 and 0.3221, respectively). The average GS value for combined SSR data was 0.3569. The lack of correlations between similarity and cophenetic matrices obtained with various methods systems suggests that different marker systems should be used simultaneously for a genetic diversity study to exploit as many sources of polymorphisms as possible.

Introduction

Different marker systems have been used routinely in assessing genetic diversity in plants. SSR markers are known to be an attractive tool for a number of approaches including genetic diversity analysis, due to its multiallelic nature, high reproducibility, and locus specificity. Previous sequence information is required for the development of SSR primers, which is a major drawback to their application in species studied less intensively. The number of SSR primers available for rye is still relatively small – 27 genomic (Saal & Wricke, 1999) and 196 EST-based (Hackauf & Wehling, 2002; Khlestkina et al., 2004). ISSR technique in which polymorphisms result from the length differences between inversely oriented closely spaced microsatellites is a cost-effective alternative to SSR. Amplified fragments can be resolved using agarose gels and the reproducibility is higher than in the other marker system using single arbitrary primer, namely RAPD. SAMPL is a relatively new molecular marker system basing on the AFLP methodology (Morgante & Vogel, 1994), technically more demanding than ISSR. Despite its many advantages, i.e., a high multiplex ratio, SAMPL has not been widely used for analysis of plant genomes. For the review of microsatellite sequence-based marker systems and their applications see Rakoczy-Trojanowska and Bolibok (2004).

Rye is an important crop due to its nutrient efficiency, winter hardiness, and tolerance to other environmental stresses such as drought. It is also a valuable genetic resource in wheat and triticale breeding. A better knowledge of genetic diversity and relationships between the existing rye inbred lines would aid in the development of breeding programs that efficiently utilize available rye germplasm. The objective of our work was to compare the efficiency of SSRs, ISSRs, and SAMPLs in detecting genetic diversity among 30 inbred lines of rye. The genetic relationships of these inbreds have not been determined before. To our knowledge it is also the first study comparing the efficiency of SAMPLs in detecting genetic diversity with those of other microsatellite sequence-based marker systems – ISSR and SSR. The second objective of the study was

to compare the results of cluster analysis performed on data from these marker systems using different statistical methods and coefficients.

Materials and methods

Plant material and DNA isolation

Thirty winter rye inbred lines, mainly of Polish origin, were used in the study. The lines were chosen to represent a wide genotypic range and include components of Polish mapping populations (L9, L318, Ot 1-3, 541, S21), male sterile forms differing in the resistance to powdery mildew (M1–M10), inbred lines differing in tissue culture ability and some morphological characters (P330, L9, L310, L318, H32, DW28, L29, L299, L324, L325, L329), as well as lines differing in the resistance to leaf rust (K3-K6, K9) and preharvest sprouting (Ot 1–3, 541). DNA was extracted from leaves of greenhouse-grown plants using CTAB method (Murray & Thompson, 1980).

ISSR assay

Fifty-five ISSR primers were tested, 14 primers showing clear and polymorphic banding patterns were chosen for genetic diversity analysis. Sequences of these primers are shown in Table 1. Amplifications were carried out in 15 μl reaction mixtures each containing 50 ng of template DNA, 0.4 μM ISSR primer, 0.2 mM each of the dNTPs, 2.5 mM $MgCl_2$, 1× PCR buffer, and 0.3 unit *Taq* polymerase (Invitrogen) using the following PCR profile: 30 s at 94 °C, 3 min at 50 or 55 °C (depending on the primer), 60 s at 72 °C, repeated 35 times. The amplification products were resolved by electrophoresis using 1% agarose gels containing ethidium bromide.

Table 1. List of ISSR primers used in experiments and their nucleotide sequences

Primer	Sequence 5′-3′
ISSR_2	GAGAGAGAGAGAGAGAGAC
ISSR_3	GAGAGAGAGAGAGAGAGAT
ISSR_14	GATCGTAGGTAGGTAGGTAGGTAGGTAG
ISSR_17	AGAGAGAGAGAGAGAGT
ISSR_20	CTCTCTCTCTCTCTCTT
ISSR_21	CTCTCTCTCTCTCTCTA
ISSR_27	ACACACACACACACACT
ISSR_32	CTCTCTCTCTCTCTCTG
ISSR_35	TCTCTCTCTCTCTCTCG
ISSR_36	ACACACACACACACACTC
ISSR_40	TGTGTGTGTGTGTGTATAT
ISSR_42	CACACACACAGAGAGAGA
ISSR_44	CCTCCTCCTCCTCCTA
ISSR_45	CCTCCTCCTCCTCCTG

SSR assay

A total of 38 SSR primers pairs were used for assessing genetic genetic diversity: 19 EST-derived SSR: SCM0019, SCM0021, SCM0041, SCM0045, SCM0047, SCM0077, SCM0083, SCM0091, SCM- 0095, SCM0098, SCM0135, SCM0139, SCM0140, SCM0141, SCM0159, SCM0172, SCM0183, SCM0150 (Hackauf & Wehling, 2002) and 19 genomic library-derived SSR: SCM2, SCM75, SCM69, SCM86, SCM101, SCM102, SCM104, SCM120, SCM138, SCM180, SCM243, SCM268, SCM304, SCM307 (Saal & Wricke, 1999) and WRM206, WRM216, WRM220, WRM225 (Rakoczy-Trojanowska et al., in preparation). The cycling parameters for SSR markers developed by Sall and Wricke (1999) followed the recommendations of the authors. The remaining SSRs were amplified using PCR profile No. 20 described by Pillen et al. (2000). Amplification products were electrophoresed in 6% denaturing polyacrylamide gels and silver stained according to Pillen et al. (2000).

SAMPL assay

The SAMPL procedure used in this study was based on the protocol of Singh et al. (2002). However, some modifications were introduced: 125 ng of genomic DNA were digested with restriction enzymes *Mse*I and *Hin*dIII. SAMPL primers used in the selective amplification included compound SSRs, 5′-anchored SSRs and 3′anchored SSRs. In total 222 combinations of 37 SAMPL primers (S) with 6 *Mse*I + 3 primers (M) were tested, of them 16 combinations of 10 SAMPL and 5 *Mse*I +3 primers (M5S1, M1S2, M6S2, M5S23, M3S28, M5S27, M1S29, M2S29, M3S29, M6S29, M5S12, M1S37, M3S37, M5S37, M5S38, M5S39) were chosen for the assessment of genetic diversity. Sequences of these primers are shown in Table 2. Amplification products were electrophoresed and visualized as described previously for SSRs.

Table 2. Sequences of the primers used in SAMPL assays

Primer	Symbol	Sequence 5′-3′
*Mse*I+3_1	M1	GACGACCGACGAGTAACAA
*Mse*I+3_2	M2	GACGACCGACGAGTAACAC
*Mse*I+3_3	M3	GACGACCGACGAGTAACAG
*Mse*I+3_5	M5	GACGACCGACGAGTAACTC
*Mse*I+3_6	M6	GACGACCGACGAGTAACTG
SAMPL_1	S1	$(GA)_9A$
SAMPL_2	S2	$(GA)_9C$
SAMPL_12	S12	$GA(TCT)_{10}$
SAMPL_23	S23	$(CA)_8A$
SAMPL_27	S27	$(AC)_8T$
SAMPL_28	S28	$(TG)_8A$
SAMPL_29	S29	$(AG)_8C$
SAMPL_37	S37	$A(CA)_7(TA)_2T$
SAMPL_38	S38	$A(CA)_7(AT)_2A$
SAMPL_39	S39	$C(AC)_7(TA)_2T$

Data analysis

In order to obtain a measure of the usefulness of each marker system, the average polymorphic information content (PIC) and the marker index (MI) were calculated for each marker system applying the formulas given by Powell et al. (1996). For calculating the MI, it is assumed that all polymorphic bands detected with each SSR assay map to a single marker locus, whereas each polymorphic DNA fragment within an ISSR or a SAMPL assay unit is considered as a single dominant marker locus.

For calculating similarity coefficients, the presence or absence of bands was recorded 1 or 0, respectively for each marker system to generate a binary matrix. Genetic similarity (GS) was calculated on the basis of Jaccard's coefficient (Jaccard, 1908). A similarity matrix was used for cluster analysis to construct dendrograms using unweighted pair group method with arithmetic averages (UPGMA). Cophenetic values were computed for each dendrogram resulting in construction of cophenetic matrix for each marker type.

Additionally cluster analysis was done for each marker system using weighted pair-group method with arithmetic averages (WPGMA) and flexible clustering (FLEXI, with parameter $\beta = -0.25$) based on (1) simple matching coefficient and (2) Yule coefficient. Resulting dendrograms were arbitrarily divided into four clusters in order to investigate the consistency of assigning the inbreds to certain groups by the different statistical methods applied. All statistical analyses were performed by NTSYS-pc, Version 2.1 (Rohlf, 2001) for each marker type separately (ISSR, SAMPL, EST-derived SSRs, genomic SSRs) and for combined SSR data and combined data from all marker types.

Results and discussion

Levels of polymorphisms

All marker systems examined turned out to be useful tools for detection of polymorphism and assessing genetic diversity in winter rye, but the degree of resolution depended on the technique applied. The number of bands analyzed per assay unit varied from 2.74 for genomic SSR to 28.31 for SAMPLs with the proportion of polymorphic bands that were detected with different markers ranging from 0.71 for SAMPLs to 0.98 for EST SSRs (Table 3). PIC value was the highest in the case of ISSR markers – 0.88, the lowest in the case of genomic SSRs – 0.35. The number of loci (alleles) analyzed was 65 for ISSR, 101 for SSR (55 for EST-derived and 46

Table 3. Analyses of banding patterns generated by ISSR, SSR, and SAMPL assays for 30 winter rye inbred lines. The numbers in brackets refer to polymorphic assay units

Marker	Number of assay units	Number of bands	Number of bands per assay unit	Proportion of polymorphic bands	Number of polymorphic bands per assay unit	Number of loci per assay unit	PIC	MI
ISSRs	14 (14)	73 (73)	5.21 (5.21)	0.88	4.64	5.21	0.88	4.58
EST-SSRs	19 (18)	56 (55)	2.95 (3.06)	0.98	2.89	1	0.41	0.41
Genomic SSRs	19 (13)	52 (46)	2.74 (3.54)	0.88	2.42	1	0.35	0.35
All SSRs	38 (31)	107 (101)	2.82 (3.16)	0.94	2.66	1	0.38	0.38
SAMPLs	16	453	28.31	0.71	19.94	28.31	0.40	11.32

for genomic library-derived SSR), and 319 for SAMPL. Each marker system was able to discriminate among 30 inbred lines. However, when data from genomic SSRs were analyzed separately it was not possible to distinguish between two inbreds namely M9 and M10. Contrary to the results of Blair et al. (2003) and Cho at al. (2000) who found gene-based SSRs to be less polymorphic than anonymous genomic SSRs, in our study EST-SSRs appeared to be more efficient in detecting polymorphisms (PIC value 0.41 versus 0.35 for genomic SSRs). Results obtained with SAMPL markers (PIC = 0.40, MI = 11, 32) confirmed findings of other authors reporting a high usefulness of this system for revealing genomic differences (Tosti & Negri, 2002; Singh et al., 2002; Roy et al., 2002). In our study SAMPLs were shown to be more suitable for this purpose than SSR markers tested. Although the PIC values and MI values for SSRs were relatively low in comparison to other marker systems applied they have the big advantage of being locus specific, which

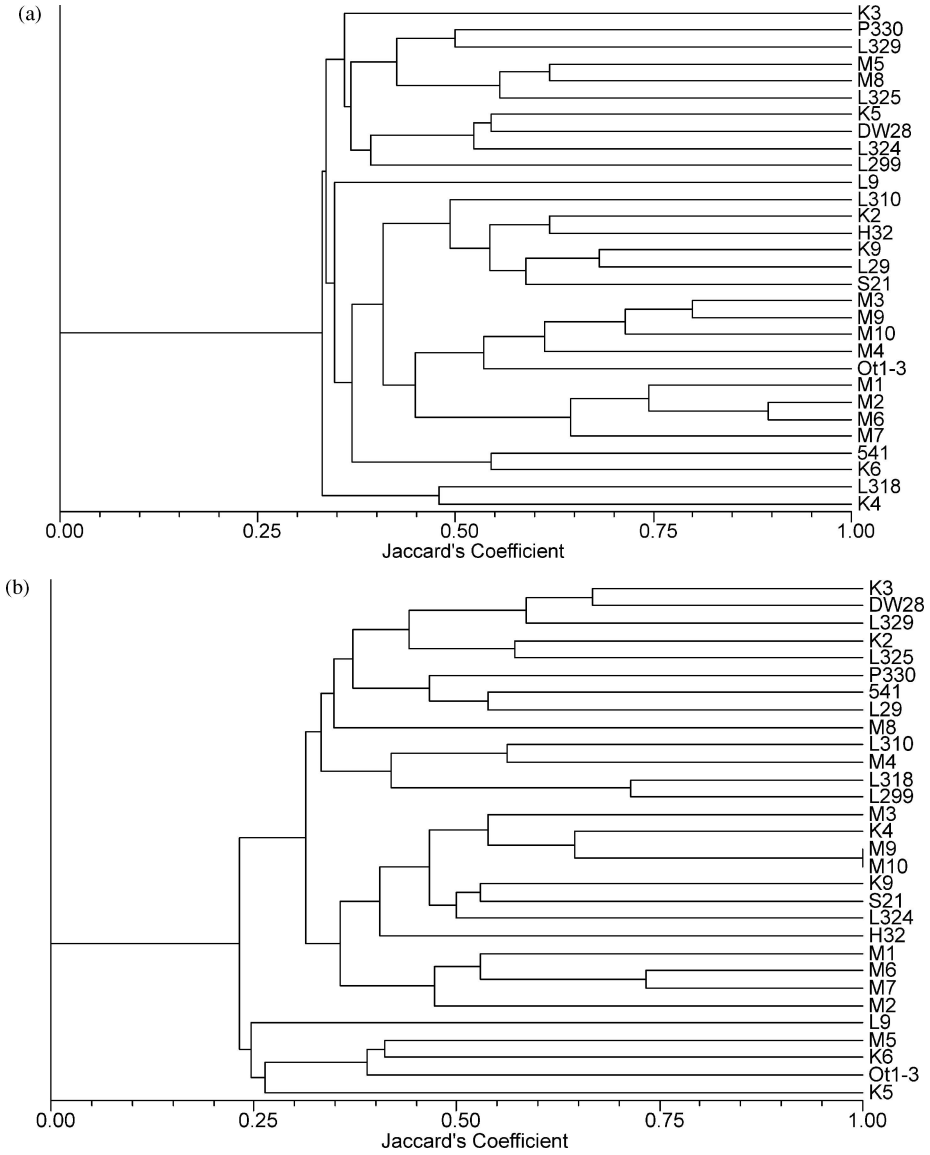

Figure 1. UPGMA dendrogram based on Jaccard's coefficient indicating genetic similarities among rye inbred lines derived from: (a) EST-SSR data, (b) genomic-SSR data.

is invaluable in some genetic investigations. The ISSRs, technically least demanding among the marker systems tested, turned out to be very effective in detecting polymorphism in winter rye (PIC = 0.88, MI = 4.58). ISSRs are also the most cost-effective, as they can be resolved on agarose, no previous sequence information is required for design of the primers, and there is no need for using restriction enzymes. All the facts make them a very attractive tool for analyzing winter rye genome.

Genetic similarity and cluster analysis

The average GS value based on Jaccard coefficient for combined SSR data was 0.3569 which was intermediate to the values derived from ISSR and SAMPL datasets (0.2888 and 0.5381, respectively). The average value of similarity coefficient was higher for EST-derived SSRs (0.3814) than for those from genomic libraries (0.3221). The average GS value for combined data from all marker systems was 0.4902.

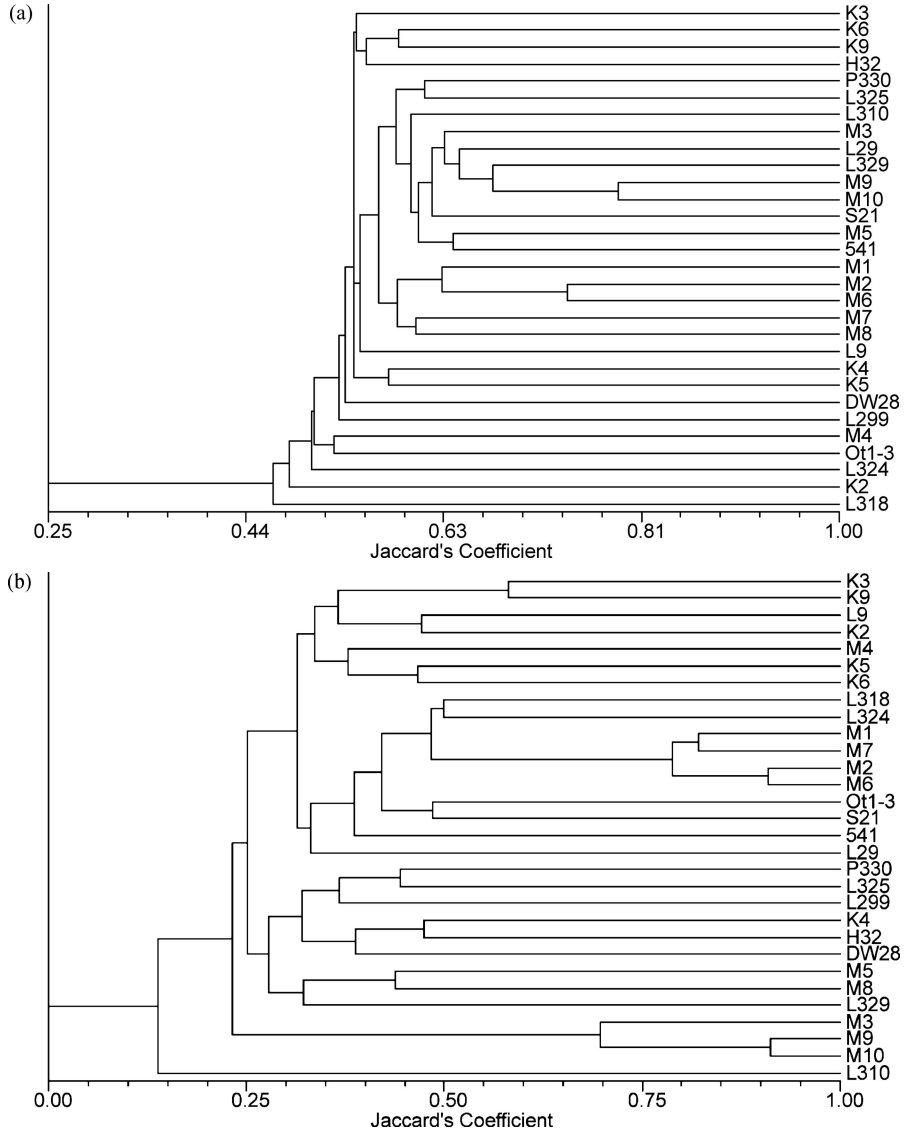

Figure 2. UPGMA dendrogram based on Jaccard's coefficient indicating genetic similarities among rye inbred lines derived from: (a) SAMPL data, (b) ISSR data.

Although certain patterns of grouping revealed by cluster analyses were consistent across dendrograms obtained using the data from each marker system (i.e., lines M1, M2, M6, M7 and M3, M9, M10 were always within the same clusters) the overall display of the inbreds was not identical. The UPGMA dendrograms based on Jaccard's coefficient derived from data from each marker system are shown in Figures 1 and 2. Such a result could be expected since the polymorphic fragments in the marker systems examined have different origin. It should be however noted that the results might be influenced by a relatively small number of loci analyzed with ISSR and SSR markers, especially in comparison with a fairly large number of SAMPL loci. The comparison of dendrograms obtained using different statistical methods and coefficients also demonstrated differences in clustering (Figures 3 and 4). Dissimilar patterns of clustering obtained depending on marker system, statistical method, and coefficient lead us to the conclusion that in order to take the advantage of different sources of polymorphism and to obtain more reliable information on

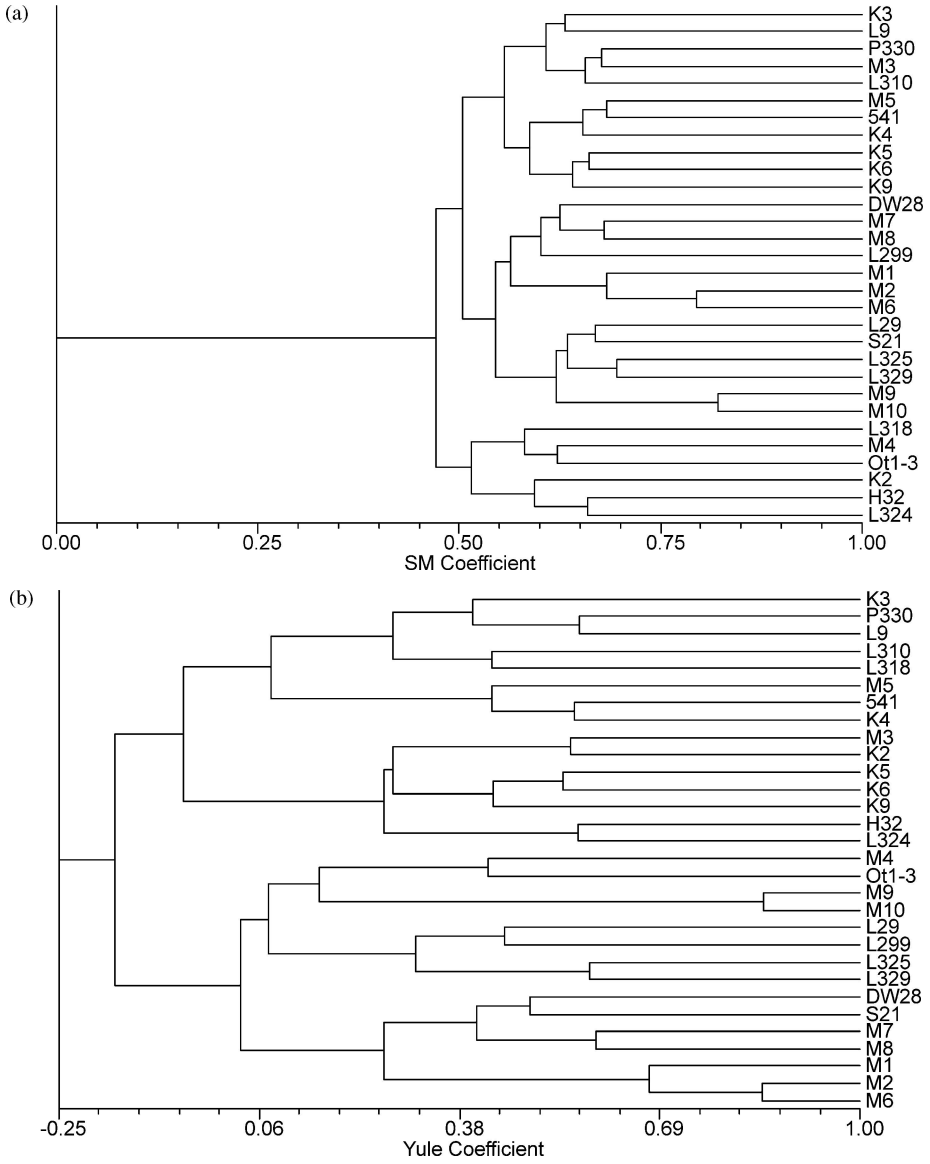

Figure 3. FLEXI dendrograms derived from SAMPL data based on: (a) SM coefficient, (b) Yule coefficient.

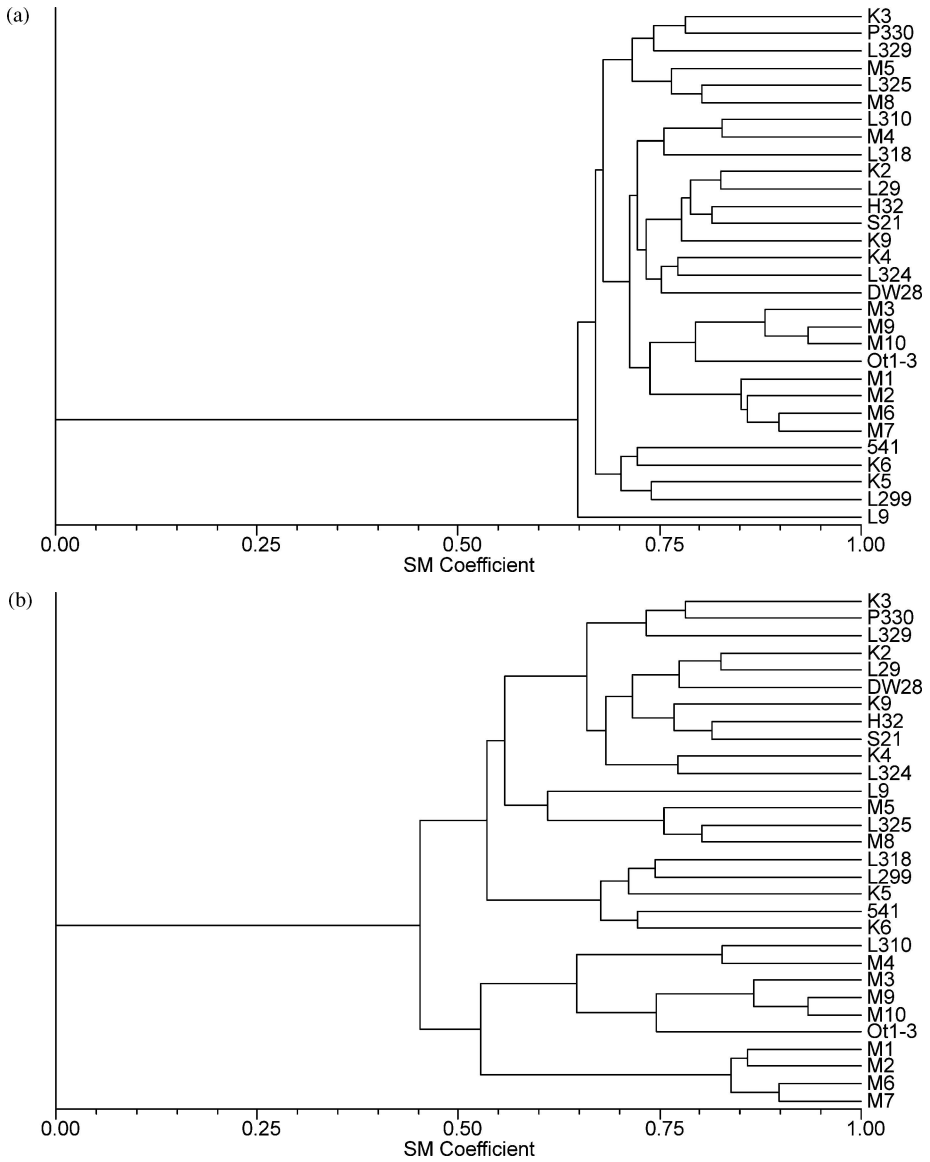

Figure 4. Dendrograms based on SM coefficient derived from combined SSR data obtained using (a) UPGMA, (b) FLEXI.

genetic relationships of the materials studied different marker systems and statistical methods should be used simultaneously in genetic diversity studies. Unfortunately, numerous studies concerning genetic similarity are based only on single marker systems which might lead to false conclusions.

Acknowledgments

We are thankful to Prof. P. Masojc from Agricultural University of Szczecin and Dr. L. Madej from Plant Breeding and Acclimatization Institute in Radzikow for providing seeds of some inbred lines. This research was partially supported by the State Committee of Scientific Research grant No. 3 P06A 025 25.

References

Blair, M.W., F. Pedraza, H.F. Buenida, E. Gaitán-Solís, S.E. Beebe, P. Gepts & J. Tohme, 2003. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 107: 1362–1374.

Cho, Y.G., S. Temnykh, X. Chen, L. Lipovich, S.R. McCouch, N. Ayres & S. Cartinhour, 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice. Theor Appl Genet 100: 713–722.

Hackauf, B. & P. Wehling, 2002. Identification of microsatellite polymorphisms in an expressed portion of the rye genome. Plant Breed 121: 17–25.

Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. Bull Soc Vaudoise Sci Nat 44: 223–270.

Khlestkina, E.K., M.H.M. Than, E.G. Pestsova, M.S. Röder, S.V. Malyshev, V. Korzun & A. Börner, 2004. Mapping of 99 new microsatellite-derived loci in rye (*Secale cereale* L.) including 39 expressed sequence tags. Theor Appl Genet 19: 725–732.

Morgante, M. & J. Vogel, 1994. Compound Microsatellite Primers for the Detection of Genetic Polymorphism, U.S. Patent Appl. 08/326456.

Murray, M.G. & W.F. Thompson, 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acid Res 8: 4321–4325.

Pillen, K., A. Binder, A.B. Kreuzkam, L. Ramsay, R. Waugh, J. Förster & J. Léon, 2000. Mapping new EMBL-derived barley microsatellites and their use in differentiating German barley cultivars. Theor Appl Genet 101: 652–660.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey & A. Rafalski, 1996. The comparison of RFLP, RAPD, AFLP, and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2: 225–238.

Rakoczy-Trojanowska, M. & H. Bolibok, 2004. Characteristics and comparison of three classes of microsatellite-based markers and their application in plants. Cell Mol Biol Lett 9: 221–238.

Rohlf, F.J., 2001. NTSYS-pc Numerical Taxonomy and Multivariate Analysis System, Version 5.1. Exeter, Setauket, NY.

Roy, J.K., H.S. Balyan, M. Prasad & P.K. Gupta, 2002. Use of SAMPL for a study of DNA polymorphism, genetic diversity, and possible gene tagging in bread wheat. Theor Appl Genet 104: 465–472.

Saal, B. & G. Wricke, 1999. Development of simple sequence repeat markers in rye (*Secale cereale* L.). Genome 42: 964–972.

Singh, A., A. Chaudhury, P.S. Srivastava & M. Lakshmikumaran, 2002. Comparison of AFLP and SAMPL markers for assessment of intra-population genetic variation in *Azadirachta indica* A. Juss Plant Sci 162: 17–25.

Tosti, N. & V. Negri, 2002. Efficiency of three PCR-based markers in assessing genetic variation among cowpea (*Vigna unguiculata* subsp. *unguiculata*) landraces. Genome 45: 268–275.