Mol Breeding (2008) 22:301–313
DOI 10.1007/s11032-008-9176-4

Utilization of SSR and AFLP markers for the assessment of distinctness in durum wheat

E. Noli · M. S. Teriaca · M. C. Sanguineti · S. Conti

Received: 28 October 2007 / Accepted: 7 March 2008 / Published online: 2 April 2008

Abstract Molecular techniques provide new possibilities to characterize advanced genetic materials for registration purposes and for the protection of breeders' rights. The objective of this work was to compare the simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP) markers with morphological descriptors, currently in use, for the assessment of variety distinctness in durum wheat. Fifty-six F_8 or F_9 lines or sublines at different levels of relatedness derived from four crosses, two BC_1- and 11 BC_3-derived lines, were characterized with 27 morphological traits, including 17 official Community Plant Variety Office (CPVO) descriptors, seven AFLP primer combinations, and 98 SSR primer pairs. The similarities based on all three marker classes reflected, on average, the degree of relatedness of the materials. However, molecular markers (MMs) and in particular SSRs produced classifications more closely representing the relationships among lines, allowing us to discriminate even tightly related genotypes that morphological descriptors failed to distinguish. The moderate correlations between similarities based on morphological and molecular data, and the wide range of MM differences corresponding to few or no morphological differences, imply that at present MMs should only be used as a complementary tool to assess distinctness. Based on these results, it seems that a set of 28 SSRs (one per chromosome arm) represents a useful prescreening tool to identify the entry pairs sufficiently different at MM level ($\geq$13 polymorphisms) for which a field evaluation could be avoided, with relevant savings in resources and optimization of the field trial design.

Keywords AFLPs · Durum wheat · DUS · Genetic similarity · Morphological descriptors · SSRs

Abbreviations

CPVO	Community Plant Variety Office
DUS	Distinctness uniformity stability
MM	Molecular marker
PBRs	Plant breeders' rights
SM	Simple matching
UPOV	Union pour la Protection des Obtentions Végétales

Introduction

The contribution of plant breeding to the spectacular progress in agriculture during the last half century, both in terms of yield increase and improvements of product quality, is evident and unquestionable (Fehr 1984; Royo et al. 2007). Variety development requires

E. Noli (✉) · M. S. Teriaca · M. C. Sanguineti · S. Conti
Seed Research and Testing Laboratory (LaRAS), Department of Agroenvironmental Science and Technology, University of Bologna, Viale Fanin 44, Bologna 40127, Italy
e-mail: enrico.noli@unibo.it

conspicuous investments that plant breeders, in particular those of private companies, should recoup in order to keep their activity profitable. This aspect has become even more crucial in recent years due to the prevailing involvement of the private sector in plant breeding and seed production. For this reason, a worldwide intellectual property protection system for plant breeders' rights (PBRs) has been set up and its principles established by the Union pour la Protection des Obtentions Végétales (UPOV). The system requires that a new plant variety, in order to be eligible for protection, must comply with specific prerequisites, namely distinctness from extant varieties, plus uniformity, and stability (DUS) of the same characteristics by which distinctness is ascertained. Commonly, candidate varieties are compared with existing varieties (or at least a sample of them) on the basis of a series of descriptors, mainly morphological, defined for each crop by experts and listed in the "UPOV Guidelines for DUS testing". For a few species, isozymes (maize and sunflower) and seed storage protein electrophoresis (wheat and barley) were included in the UPOV guidelines for DUS testing, but only in order to provide complementary information in those cases where distinctness could not be clearly established by means of traditional descriptors.

Several authors have highlighted the new possibilities and perspectives in varietal characterization opened up by the advent of molecular markers (MMs), emphasizing their advantages such as independence from the environment and rapidity of assessment (Soller and Beckman 1983; Smith and Smith 1992; Preston et al. 1999). Two main applications of MMs can be envisaged: varietal identification and varietal registration. In the first case, which consists of the verification of the identity of an already registered or protected variety, the adoption of molecular techniques is quite straightforward and does not generate special concerns (Camlin 2001). On the other hand, variety registration entails the use of MMs for a de novo, detailed description of new varieties for the assessment of DUS requisites, thus raising issues related to the levels of distinctness and uniformity that will be required. Molecular data, due to the large number of available loci and distinguishing power, could potentially reduce the current level of distance required to demonstrate distinctness, unless thresholds were established that conform to current standards of distinctness as determined by morphological data. Moreover, without suitable thresholds to determine acceptable levels of uniformity, the use of molecular data could lead to standards that are unnecessary and costly to monitor (ISF 2003, 2006). The opportunity of using MMs in the DUS examination has been debated within the UPOV; in this respect it was stressed that any innovative approach should be compared with the existing system established by the first UPOV Convention (UPOV 1961) that has been satisfactory so far in variety protection. Three options have been considered by the Working Group on Biochemical and Molecular Techniques (BMT) (UPOV 2004): (1) the use of MMs linked to the phenotypic characteristics adopted under the current system; (2) the use of MMs for the management of reference collections, i.e., of the sets of varieties of common knowledge with which candidate varieties should be compared; and (3) the use of MMs as an alternative to the current system.

No consensus has been reached regarding the acceptability of option 3 because it is felt that it would undermine the effectiveness of the protection under the UPOV system. Conversely, option 1 is difficult to apply on a large scale, at least for the time being, due to the high costs for the development of markers tightly linked to the traits of interest, even though it would be perfectly compatible with the current system. According to option 2, MMs could represent a useful means to improve the efficiency of the DUS assessment as a prescreening tool to identify well-distinct comparisons that would not need full phenotypic examination; nonetheless, its application requires the definition of thresholds for differences at the molecular level against those detected at the morphological level (van Eeuwijk and Baril 2001).

The objective of this work is to compare the morphological approach, presently in use, with a molecular approach based on SSR and AFLP markers in view of its possible application as an analytical tool to supplement the current procedure for the assessment of variety distinctness in durum wheat.

Materials and methods

Plant material

In this study, 69 experimental lines at an advanced inbreeding stage, derived from four crosses and two

Table 1 List of the examined lines, of the parents of the base crosses or backcrosses, and of their ancestors

Cross[a]	Lines	Cross[a]	Lines	Cross[a]	Lines
Sv × Ir	A655.1/.2	Ir × Me	C523	(Ma × Sv) × Sv^3	E731
	A657		C531		E733
	A658		C536		E734
	A661.1/.2_hl[b]		C543.1/.2_hy		E736
	A661_hy		C546.1/.2_hl		E737
	A662		C546_hy		
	A679		C554	(N224 × Sv) × Sv^3	$F^3$440.1/.2
	A680.1/.2		C568		$F^3$460
	A682.1/.2		C562.1/.2		$F^3$461
	A684		C563.1/.2		$F^3$203.1/.2
Ir × Ar	B340	Ar × Me	D424.1/.2_hl	(N224 × Sv) × Sv^1	$F^1$836
	B364		D424.1/.2_hy		$F^1$215
	B371		D436		
	B395_hl		D460		
	B395.1/.2_hy		D477		
	B400		D489_gl_hy		
	B402		D489_hl		
	B418.1/.2		D508		
	B420.1/.2		D445.1/.2		
	B375.1/.2		D466.1/.2		

[a] Sv: Svevo (Cymmit's selection/Zenit); Ir: Iride (Altar 84/Ionio); Ar: Arcobaleno (Altar 84/Chen); Me: Meridiano (Simeto/WB881//Duilio/F21); Ma: Manital (Marzotto/Mendos); N224 (Creso/A327//Lira/*T. dicoccoides*)

[b] hl: hairless; hy: hairy; gl: glaucous

backcrosses together with their parental lines, were considered (Table 1). In particular, the base crosses were made using three Italian varieties, i.e., Iride (Ir), Meridiano (Me), and Svevo (Sv), and the Spanish variety Arcobaleno (Ar). Moreover, Ir and Ar shared Altar 84 as a common parent. The breeding programs were carried out at Società Produttori Sementi Bologna and were aimed at obtaining high-yielding varieties with improved agronomic traits (earliness and resistance to the main diseases) and grain quality. For all crosses, starting from the F_4 generation, the selected materials were advanced by bulking the seed of the progeny, after an individual selection within the family aimed at removing off-type plants. With this procedure, F_9 lines from the Sv × Ir cross and F_8 lines from the other three crosses (Ir × Ar, Ar × Me, and Ir × Me) were developed. In the 2002 growing season, ten lines per cross, representative of the variability for some quality traits (protein content, yellow pigment colour, and gluten quality), were considered. As reported in Table 1, for the crosses Sv × Ir, Ir × Ar, and Ir × Me, two of the ten lines were related (as indicated by the same number), being derived from F_4 or F_5 plants of the same progeny differing for hairiness of the glume; for the cross Ar × Me, two pairs of related lines differing for the same morphological character were considered; overall, 35 of the tested lines can be considered as unrelated. For each cross, one plant at random for six lines and two plants for four lines (.1/.2) were chosen and bagged before anthesis to ensure selfing. At maturity, plants were harvested individually, seeds were cleaned manually to avoid contamination and treated with fungicide (Panoctine, Makhteshim Agan) prior to sowing. In this way, for each cross 14 genotypes were obtained. Due to the fact that seeds were bulked from F_4 onward, the sublines belonging to one pair (.1/.2) can be considered as having originated from the same F_3 plant or from the same plant in any subsequent generation preceding that in which the material was collected (F_8 in Sv × Ir or F_7 in the other crosses).

The remaining experimental lines were derived from two backcross programs with Svevo as recurrent parent: 5 BC_3F_6 lines from a cross with the bread wheat (*Triticum aestivum* L.) cv. Manital (Ma) as donor parent, 6 BC_3F_6 and 2 BC_1F_8 lines from a cross with the experimental line N224 derived from a double cross with a *Triticum dicoccoides* accession being one of the parents; in both cases the traits under transfer

were glutelin bands conferring high gluten quality. Two pairs of BC_3F_6 genotypes having N224 as donor were sublines derived from the same BC_3F_4 family.

The overall structure of these materials was such that the relationships among genotypes encompassed different levels of genetic similarity: lines obtained from crosses among different parents (Sv × Ir versus Ar × Me), with one (e.g. Sv × Ir versus Ir × Ar) or two common parents (same F_2 population); in the latter case the two parents were either related (Ar and Ir) or unrelated (the other crosses). Moreover, close genetic relationships were represented by lines deriving from the same backcross, or sublines derived from the same F_7–F_8 line.

Methods

Phenotypic characterization

Field trials were carried out in 2003 and 2004 at the farm of Produttori Sementi Bologna, Argelato in the Po Valley, Italy. Accessions were arranged in a randomized complete block design with three replicates. Plots spaced 0.72 m from each other, with double rows, 2.5 m long, were sowed with 200 seeds. Seventeen morphological traits of the 26 indicated in the CPVO protocol (CPVO 2003) were evaluated; the remaining nine were not included because they had not shown polymorphism in these materials. The CPVO traits considered were: frequency of plants with recurved flag leaves; glaucosity of neck; glaucosity of sheath; glaucosity and colour of the ear; awn colour; shape and width of the shoulder of the lower glume; length and shape of the beak of the lower glume; hairiness on external surface of the lower glume; grain shape; time of ear emergence; plant length; ear length, excluding awns; awns length; ear density. Ten additional traits were also considered, either because of discriminative value (straight or curved shape of neck; anthocyanin coloration of auricles; anthocyanin coloration of anthers; shape of the ear; length of ear extrusion), or because of agronomic interest (number of spikelets per ear; number of sterile basal spikelets per ear; number of grains per ear; number of grains per spikelet; grain mean weight). These additional traits were measured on ten ears per plot. Of the 27 characteristics scored, 11 were metric and 16 qualitative. The latter characteristics were assessed by visual observation and scored according to the CPVO protocol, except for shape of neck and shape of ear (on a 1–3 scale) and anthocyanin coloration of auricles and anthers (scored as presence/absence).

Genotypic characterization

For each genotype, 15 caryopses were planted in a pot and grown in the greenhouse in disease-free conditions until 3–4 leaves developed. An equal quantity of leaf material from 12 plants was harvested, bulked, lyophilized, and finely ground with a laboratory mill (Cyclotech PBI). DNA was extracted using a modified CTAB procedure (Saghai-Maroof et al. 1984). Molecular characterization was carried out with SSR and AFLP markers.

SSR analysis was conducted using 98 primer pairs (Table 2) directed to the amplification of di- and trinucleotide microsatellite loci, originally developed in bread wheat within five different research programs, WMS (Wheat Micro Satellite; Röder et al. 1998), WMC (Wheat Microsatellite Consortium; Gupta et al. 2002), BARC (USDA-ARS Beltsville Agricultural Research Center; Song et al. 2005), CFA and CFD (Sourdille et al. 2001; Guyomarc'h et al. 2002), and KSUM and CNL (Kansas State University and Cornell University; http://wheat.pw.usda.gov/ITMI/2002/EST-SSR; Yu et al. 2004). The loci were selected to provide an even coverage of the genome based on map information deduced from the above-cited papers, from the consensus map of bread wheat (Somers et al. 2004), and from the Triticeae database (GrainGenes 2004; http://wheat.pw.usda.gov).

For each primer pair, amplifications were performed according to the protocol provided by the authors with slight modifications; products were separated on 5% denaturing polyacrilamide gels and detected by silver staining.

PCR reactions were carried out in a final volume of 20 µl containing: PCR buffer 1×, $MgCl_2$ 1.5 mM, dNTPs 0.2 mM, 0.25 mM of each primer, 1 U *Taq* polymerase, and 50 ng template DNA. Amplifications were conducted using an MJ PTC-200 thermal cycler. For the primers WMC, BARC, and KSUM the amplification mix contained $MgCl_2$ 2.0 mM.

For the AFLP analysis, markers were generated by DNA digestion with *Sse*8387 (eight cutter) and *Mse*I (four cutter) enzymes followed by amplification with

Table 2 List of the SSR loci analyzed in this study

Locus[a]		Chromosome[b]	Locus[a]		Chromosome[b]
Xgwm136	+[c]	1AS	*Xwmc329*		1BS
Xwmc24		1AS	*Xgwm131*	+	1BS
Xgwm99	+	1AL	*Xwmc156*		1BL
Xcfa2263		2AS	*Xgwm268*	+	1BL
Xwmc177		2AS	*Xwmc44*		1BL
Xgwm448	+	2AS	*Xwmc25*		2BS
Xgwm95		2AS	*Xwmc154*		2BS
Xbarc5		2AL	*Xwmc243*	+	2BS
Xbarc220		2AL	*Xwmc257*		2BS
Xgwm294		2AL	*Xwmc27*		2BS
Xcfa2058		2AL	*Xgwm120*	+	2BL
Xwmc181		2AL	*Xgwm389*	+	3BS
Xcfa2086		2AL	*Xgwm1034*		3BS
Xgwm356		2AL	*Xgwm566*		3BS
Xbarc279	+	2AL	*Xgwm131*	+	3BL
Xgwm311		2AL	*Xwmc322*		3BL
Xbarc310		3AS	*Xgwm368*		4BS
Xbarc179		3AS	*Xwmc48*		4BS
Xgwm5	+	3AS	*Xwmc238*		4BS
Xwmc264	+	3AL	*Xgwm165b*	+	4BL
Xwmc215		3AL	*Xcfa2149*		4BL
Xwmc169		3AL	*Xwmc47*	+	4BL
Xgwm4		4AS	*Xcnl123*		4BL
Xgwm165a		4AS	*Xksum154*		4BL
Xbarc206	+	4AS	*Xgwm234*	+	5BS
Xbarc224		4AS	*Xgwm149*		5BS
Xgwm601		4AL	*Xgwm544*		5BS
Xgwm637	+	4AL	*Xgwm213*	+	5BL
Xwmc262		4AL	*Xgwm408*		5BL
Xwmc219		4AL	*Xgwm613*		6BS
Xgwm154		5AS	*Xgwc104*		6BS
Xgwm293		5AS	*Xcfd13*	+	6BS
Xgwm415		5AS	*Xcnl64*		6BS
Xbarc56		5AS	*Xgwm105*		6BL
Xgwm304	+	5AS	*Xgwm182*		6BL
Xgwm156		5AL	*Xbarc24*		6BL
Xgwm327		5AL	*Xgwm219*	+	6BL
Xwmc96-5A		5AL	*Xwmc323*		7BS
Xgwm291	+	5AL	*Xgwm569*		7BS
Xgwm459	+	6AS	*Xgwm537*	+	7BS
Xwmc163		6AL	*Xbarc65*		7BS
Xwmc256		6AL	*Xbarc85*		7BS
Xgwm570	+	6AL	*Xbarc72*		7BS
Xgwm169		6AL	*Xgwm577*		7BL
Xksum19	+	7AS	*Xwmc276*		7BL
Xbarc1088		7AS	*Xwmc273*		7BL
Xbarc64		7AS	*Xgwm344*	+	7BL
Xwmc83		7AS			
Xwmc488	+	7AL			
Xgwm282		7AL			
Xgwm63		7AL			
Xwmc525		7AL			
Total		Tested	99		
		Polymorphic	76		

[a] *Xgwm* (Gatersleben Wheat Microsatellite); *Xwmc* (Wheat Microsatellite Consortium); *Xcfa*, *Xcfd* (Clermont Ferrand, UMR Amélioration et Santé des plantes, INRA-UBP); *Xbarc* (Beltsville Agricultural Research Center); *Xcnl* (Cornell University); *Xksum* (Kansas State University Microsatellite)

[b] Chromosome (1–7), genome (A, B), chromosome arm (L, S)

[c] + Indicates a subset of 28 informative and independent markers used to assess distinctness

primers with three and two selective nucleotides, respectively. Seven primer combinations (S13/M51; S13/M59; S13/M60; S25/M49; S25/M55; S25/M56; S14/M61), previously utilized by Lotti et al. (2000) for the construction of a linkage map in the interspecific cross *T. durum* cv. Messapia × *T. dicoccoides* line MG4343, were tested. Amplification products were analyzed as for SSRs.

Data analysis

For all marker types, data were recorded in a spreadsheet arranging genotypes and traits in columns and rows, respectively. Moreover, in order to keep track of heterogeneity within a genotype, each trait/marker was scored over two rows.

Morphological data were standardized as

$$X'_{ki} = (X_{ki} - X_{\min})/(X_{\max} - X_{\min})$$

and phenotypic distances were calculated as the Manhattan coefficient

$$D_{ij} = 1/n \sum_{k=1}^{n} |X'_{ki} - X'_{kj}|,$$

where X_{ki} and X_{kj} are the observed values of the two lines i and j with respect to the kth variable and n is the

number of variables considered. The distances calculated according to this procedure are suitable when both metric and qualitative traits are considered. From phenotypic distances, phenotypic similarities were obtained as $1 - D_{ij}$ and used for cluster analysis. All calculations were performed using the subprograms STAND, SIMINT,and TRANSF of the statistical package NTSYS-pc, version 2.0 (Rohlf 1997).

For SSR data, a numerical code from 1 up to the total number of alleles was attributed to the different alleles at one locus, and genetic similarity was calculated using the simple matching (SM) coefficient (Sneath and Sokal 1973). AFLP data were scored as presence or absence of individual bands and recorded as 1 or 0, respectively; genetic similarity was obtained using the Jaccard coefficient. For both MM classes, similarity matrices were obtained using the subprogram SIMQUAL.

Finally, from the three types of similarity matrices, clusters were built using the Unweighted Pair Group Method Using Arithmetic Average (UPGMA) procedure in the SAHN subprogram, considering all the descriptors and markers showing polymorphism among the six parents of the crosses. Cophenetic matrices (similarity values among accessions as defined by the output of cluster analysis) were calculated using the COPH subprogram, whereas correlation between matrices values were obtained with MXPLOT.

Distinctness was assessed through pairwise comparisons among all genotypes for all traits/markers considered. In the case of morphological traits, two genotypes were deemed distinct if they differed for at least one trait. For the traits assessed by visual observation, two genotypes were considered different if the expression of the respective characteristics differed by at least the span of one class of score; for the metric traits a *t*-test (P 0.01) was performed.

For MMs, two genotypes were considered distinct if they differed by at least one band or one locus for AFLP and SSR, respectively.

Results

Genetic variation among parental lines

A high level of polymorphism for morphological descriptors was detected between the durum wheat varieties on one side, and the bread wheat cv. Manital and the experimental line N224, on the other (17 and 15 polymorphic traits, respectively, when the latter two genotypes were compared to Svevo). Fewer differences were revealed among the durum varieties, ranging from 13 to 5 between the parents of the crosses Sv × Ir and Ir × Ar, respectively. Considering only the CPVO characters, clear differences were observed among the four durum varieties, namely for growth habit, for several descriptors of the glumes, for plant height, heading date, and length of ear extrusion. Regarding SSR analysis of the 98 primer pairs tested, 22 were monomorphic and 76 were polymorphic in at least one cross; one primer pair (WMS165) detected two loci. Fifty-two loci mapped on the A and 47 on the B genome. Even though our objective was to obtain an even coverage of the genome, the number of loci assayed per chromosome was quite variable, ranging from 3 to 13, and for three chromosome arms (1AL, 6AS, and 2BL) only one marker was analyzed. The level of polymorphism at SSR loci was similar in the two genomes (79 versus 77%, in A and B, respectively). Considering all the genotypes, chromosome regions completely monomorphic were observed in the telomeric part of 3AS and 4AS, in the centromeric part of 4BS, and in the middle part of 7AS and 7BS. Considering only the durum varieties, a low level of polymorphism was detected also in 2AL and in both arms of 2B. Nine nonparental alleles and four heterogeneous loci were detected in the cross-derived progenies, which represent 0.6% and 0.3%, respectively, of the data points, at loci that were polymorphic between parents of each cross. Higher frequencies of nonparental alleles and of heterogeneous loci (2.8% and 0.8%, respectively) were found in the BC-derived progenies.

The seven AFLP primer combinations produced a total of 206 bands, 56 of which were polymorphic in at least one cross. The total number of bands and the number of polymorphic bands per primer combination varied from 19 to 39 and from 4 to 11, respectively (data not shown). Because the cv. Messapia and the line MG4343, parents of the mapping population studied by Lotti et al. (2000), were also included in our analysis, the cross-reference to that genetic map allowed us to position 10 of the 56 AFLPs. In particular, six and four of the markers were attributed to the A and B genomes, respectively, mapping on 7 of the 28 chromosome

arms. In several cases, markers were located in regions poorly covered by SSRs (1AL, 6AS, 2BL, and 4BS). Nonparental alleles were detected in five lines at one AFLP locus, resulting in a frequency of 0.5% over the total data points scored.

Considering only the durum varieties, the cross Sv × Ir showed, similarly to morphological traits, the highest level of SSR and AFLP polymorphism (37% and 12%, respectively), whereas the Ir × Ar cross, whose parents are closely related, showed the lowest level (18% and 3%, respectively). As expected, higher levels of polymorphism on A and B genomes (69% and 47%, respectively) were found for SSR loci between Svevo, the recurrent parent of the backcrosses, and the donors Manital (bread wheat) and N224.

Genetic variation among progeny lines and with their parents

For all marker classes, similarities between lines belonging to different crosses were, as expected, lower than those between lines obtained from the same cross (Table 3). Likewise, similarities between lines derived from crosses with no common parent were lower than those between lines having one common parent, particularly when determined by MMs. Moreover, similarities between lines within crosses were like those between lines and their parents, but lower than between sublines. The latter was comparable, considering the SSR markers, to that observed between BC_3-derived lines, which in turn showed similarity values (0.93) identical to those with their recurrent parent. The level of similarity between the two BC_1 or with their recurrent parent was, as expected, lower than that observed for BC_3; nonetheless, the small sample size does not allow further comments.

Classifications based on morphological descriptors and marker data

The suitability of the three marker systems to describe the relationships among the 69 experimental lines can be evaluated by comparing the dendrograms shown in Fig. 1. Parental varieties were not included in this calculation in order to avoid biases in the cluster construction due to their occurrence in more than one base cross; in fact, they could be associated with only one of the groups of derived lines, preventing a faithful representation of the actual relationships with their descendants.

Table 3 Average similarities among genotypes of the crosses or backcrosses considered, based on morphological and molecular markers

Comparison	Similarity[a]		
	Morphological	SSR	AFLP
Lines of different crosses			
Without a common parent[b]	0.78	0.65	0.58
With a common parent	0.81	0.73	0.69
Lines of the same cross			
Between lines	0.83	0.82	0.77
Between sublines	0.95	0.96	0.91
Between lines and parental varieties	0.84	0.81	0.76
Lines of the same backcross			
Between BC_3 lines	0.86	0.93	0.93
Between BC_3 lines and recurrent parent	0.91	0.93	0.86
Between BC_1 lines	0.83	0.72	0.70
Between BC_1 lines with recurrent parent	0.82	0.84	0.78

[a] Calculated as (1 − Manhattan distance), simple matching, and Jaccard coefficients, for morphological, SSR, and AFLP data, respectively

[b] Sv × Ir and Ar × Me

The graphical representation based on the 27 morphological descriptors (Fig. 1) only partially reflects the genetic origin of the materials. The group in the upper part of the dendrogram is comprised of genotypes belonging either to the cross Sv × Ir or to the backcrosses in which Sv was used as recurrent parent, whereas the middle part is comprised of genotypes belonging to the three remaining crosses (Ir × Ar, Ar × Me, and Ir × Me) which, however, are rather dispersed in the cluster. It is worth noting that line $F^{1}836$, morphologically similar to the donor N224 for anthocyanin coloration of anthers and auricles and several quantitative characters, is outside of the group of the other backcross lines. The same line was correctly classified when only CPVO descriptors, not including the above-mentioned traits, were used (cluster not shown as it is similar to that presented, with a 0.82 correlation coefficient between cophenetic matrices). No cases of complete identity

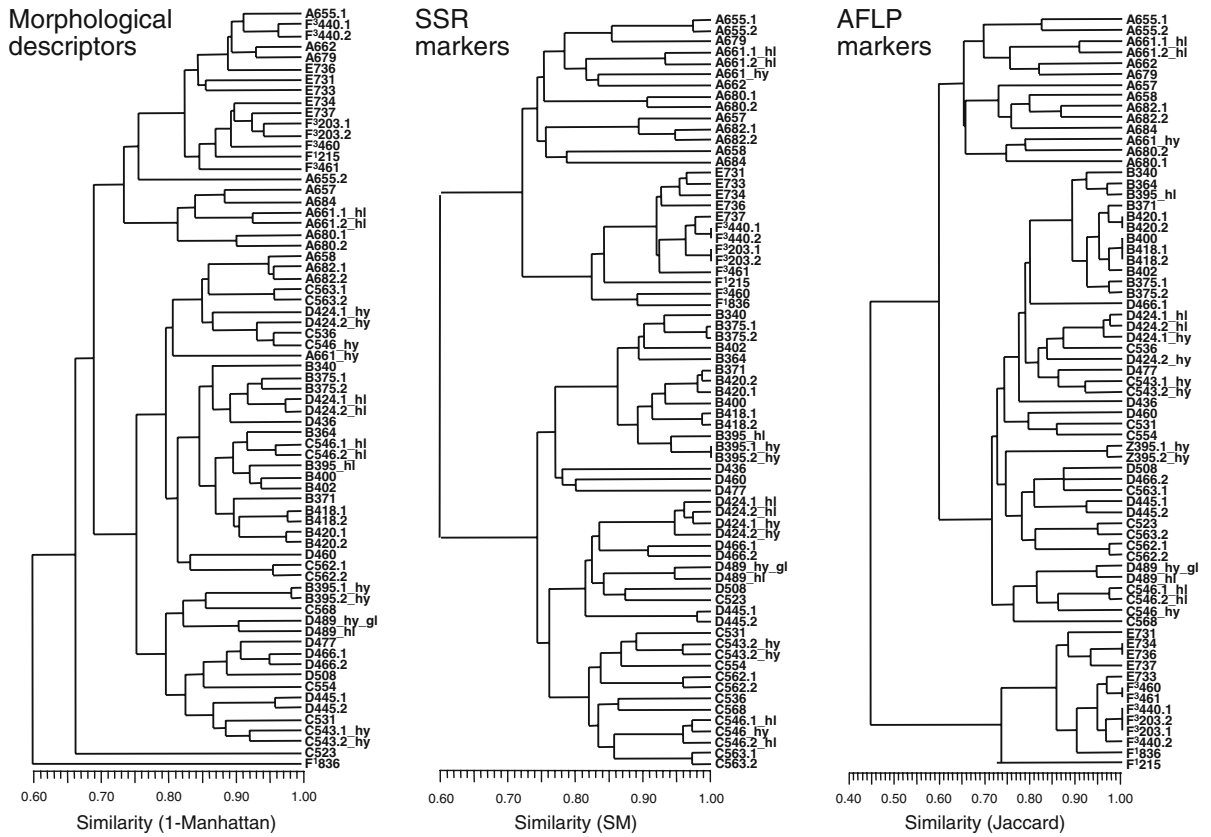

Fig. 1 Relationships among 69 lines of durum wheat revealed by UPGMA clustering on the basis of morphological descriptors, and SSR and AFLP markers. A: lines derived from Sv × Ir; B: from Ir × Ar; C: from Ir × Me; D: from Ar × Me; E: from (Ma × Sv) × Sv^3; F: from (N224 × Sv) × Sv^{1-3}

Table 4 Correlations between similarity matrices (below diagonal) and cophenetic matrices (above diagonal) based on different marker classes

	Morphological ($N = 27$)	SSR ($N = 76$)	AFLP ($N = 56$)
Morphological	**0.78****	0.63**	0.64**
SSR	0.66**	**0.80****	0.77**
AFLP	0.62**	0.89**	**0.89****

Cophenetic correlation values for the UPGMA dendrograms are shown in bold along the diagonal

** Statistically significant (P 0.01) according to Mantel test

can be detected in this dendrogram since metric variables were used for the calculation of similarities.

Although similar to that based on morphological descriptors, the analyses carried out using MMs produced dendrograms that were more in keeping with the pedigree of the materials. In particular, the dendrogram obtained from the 76 polymorphic SSRs clearly classified the genotypes into two main clusters, one including the progenies of the cross Sv × Ir and those of the backcrosses, the other including the remaining lines which, in this case, appeared almost perfectly grouped according to the base crosses; similar results (cluster not shown) were obtained when the analysis was based on a subset of 28 SSR markers evenly distributed in the genome (1 per chromosome arm) and highly informative (see those tagged with a plus symbol in Table 2). The few exceptions are the line C523 of the cross Ir × Me that clustered with the Ar × Me progenies, and the lines D436, D460, and D477 of this latter cross, which appeared slightly more similar to the Ir × Ar progenies.

Also the AFLP-based dendrogram grouped the progenies according to the crosses of origin, but with extensive mixing of the lines derived from crosses sharing Meridiano as common parent (Ar × Me and

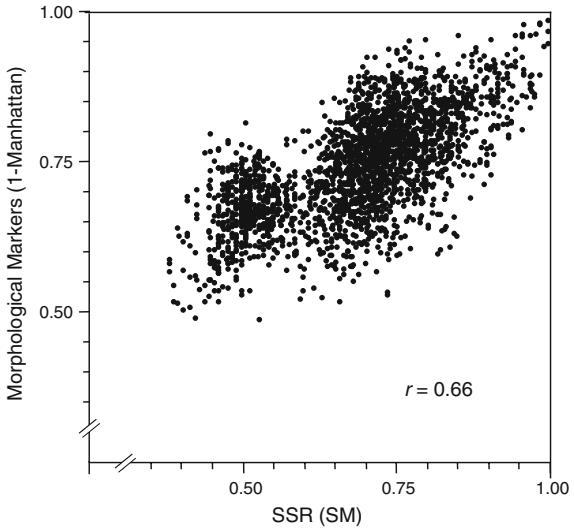

Fig. 2 Plot of genetic similarity estimates among 69 lines of durum wheat based on morphological descriptors and SSR markers

Ir × Me). In addition AFLP emphasized the high similarity among the Ir × Ar progenies, attributable to the narrow genetic base of the original cross. Moreover, these markers clustered the progenies of the Sv × Ir cross in a well-distinct group clearly separated from the others, whereas the SSR markers revealed their high similarity with the backcross lines having Svevo as a parent. As for the SSR markers, the AFLP dendrogram showed a few cases of identity among lines belonging to the same cross.

Correlations between similarity matrices

The correlation between the similarity estimates among the 69 experimental lines (Table 4) based on the two types of MMs was high ($r = 0.89$), whereas those between similarities calculated using morphological descriptors and SSRs or AFLPs were moderate ($r = 0.66$ and 0.62, respectively). The scatterplot describing the relationships among similarity estimates based on SSRs and morphological descriptors is shown in Fig. 2. The correlations between cophenetic matrices were comparable to those between similarity matrices, indicating a similar degree of topological resemblance between dendrograms. Lastly, for each marker system, the value of the cophenetic correlation coefficient indicated that the graphical representations given by the clusters shown in Fig. 1 quite closely reflected the relationships among genotypes expressed in terms of similarity.

Assessment of distinctness

An evaluation of the discriminative ability of the three types of markers is presented in Table 5, which shows the number and the proportion of pairwise

Table 5 Number of indistinguishable pairs of entries (lines and parents) on the basis of the CPVO descriptors and of the molecular markers considered

Comparison	Indistinguishable entry pairs (*n*)											
	CPVO[a]		AFLP[b]		SSR[c]							
					All		2 per chrom. arm		1 per chrom. arm		1 per chrom.	
Within crosses or backcrosses												
Sv × Ir	6	(2)	0	(0)	0	(0)	0	(0)	0	(0)	2	(2)
Ir × Ar	6	(4)	4	(2)	1	(1)	1	(1)	3	(3)	3	(3)
Ar × Me	6	(2)	0	(0)	0	(0)	0	(0)	0	(0)	2	(1)
Ir × Me	4	(3)	0	(0)	0	(0)	0	(0)	1	(1)	3	(3)
(N224 × Sv) × Sv^{1-3}	5		4		2		2		2		6	
(Ma × Sv) × Sv^{3}	0		1		0		0		6		6	
Among all entries	27	(0.97%)	9	(0.32%)	3	(0.11%)	3	(0.11%)	12	(0.43%)	22	(0.79%)

Within each cross numbers in brackets indicate the pair of sublines that were not distinct

[a] Based on 17 descriptors

[b] Based on 56 polymorphic bands

[c] Numbers of loci considered: all = 99 markers; 2 per chromosome arm = 53 (for three chromosome arms only one locus was tested); 1 per chromosome arm = 28; 1 per chromosome = 14

comparisons between entries (lines and parents) that were not distinct. If difference in only one CPVO descriptor were adopted as a general criterion to establish distinctness (as currently is the practice for DUS testing), the number of indistinguishable entry pairs was 27, from a total of 2,775 comparisons among the 75 entries (0.97%). Quite surprisingly, five cross-derived lines belonging to three different populations did not differ from one of the parents, whereas all the backcross-derived lines were always distinguishable from the recurrent parent, even if for just one character. On the other hand, the five indistinguishable pairs showed evident differences at the molecular level, with 11-18 SSR and 4-16 AFLP polymorphisms, respectively. Overall, considering the difference at only one locus as a criterion of distinctness for MMs, only three (0.11%) and nine (0.32%) cases of indistinguishable pairs were observed for SSRs and AFLPs, respectively. The higher discriminative capacity of SSRs was not only due to their abundance compared to morphological characters but also to their higher informativeness. Based on 14 SSRs (1 per chromosome), the number of indistinguishable line pairs was slightly, although not significantly (χ^2 test; P 0.05), lower (22, 0.79%) than that obtained with the 17 CPVO descriptors. As regards AFLPs, the level of discrimination achieved with the 56 polymorphic bands was similar to that obtained with 28 SSR markers (1 per chromosome arm), but significantly higher than that achieved with morphological characters.

As expected, all the indistinguishable comparisons were found within crosses. Moreover, considering all the marker types, they referred mainly to pairs of sublines or to comparisons between lines and their parents. In the backcross having N224 as donor parent, a better discrimination than that attained with the CPVO characters was reached when at least two SSR markers per chromosome (1 per chromosome arm) were considered, although a complete separation of the six BC lines was not reached even with 99 loci. It should be pointed out that the only two indistinguishable comparisons were those between sublines originating from the same BC_3F_4 family. In the backcross Sv × Ma, the 17 CPVO descriptors were sufficient to distinguish all genotypes, whereas a higher number of SSRs (2 per chromosome arm) was necessary to reach the same level of discrimination.

When considering the capacity to distinguish sublines, both MM types proved to be more efficient than CPVO descriptors: in fact in 11 out of 16 cases differences were not detected with CPVO markers, whereas SSRs and AFLPs failed only in one and two cases, respectively; the three latter pairs were also found to be morphologically indistinguishable. In particular, in one case (not shown), in spite of no morphological difference, up to seven SSR and ten AFLP polymorphisms were observed.

Table 6 shows, for all the pairwise comparisons between entries that were either not distinct or that were distinct for one, two, or three CPVO descriptors, the corresponding number of polymorphic MMs. In

Table 6 Number of AFLP and SSR loci polymorphic between genotypes either morphologically not distinct or distinct for one or more CPVO descriptors

CPVO[a]		Polymorphic loci (*n*)								
Distinctive descriptors (*n*)	Comparisons (*n*)	AFLP[b]			All SSR[c]			1 SSR per chrom. arm[d]		
		Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max
0	27	5.5 ± 5.2	0.0	16.0	8.1 ± 6.5	0.0	18.5[e]	4.2 ± 3.5	0.0	11.0
1	64	8.7 ± 5.2	1.0	18.0	12.4 ± 5.9	1.5	24.0	6.2 ± 3.5	0.0	13.5
2	186	9.0 ± 5.6	0.0	31.0	14.4 ± 7.4	2.0	41.0	7.6 ± 4.2	0.0	20.5
3	306	11.4 ± 6.1	0.0	34.0	17.9 ± 7.2	1.0	42.5	9.5 ± 4.2	0.0	22.5

[a] Based on 17 polymorphic descriptors

[b] Based on 56 polymorphic bands

[c] Based on 76 polymorphic loci

[d] Based on 28 polymorphic loci (1 per chromosome arm)

[e] For SSR markers decimals in minimum and maximum values are due to the scoring of heterozygous loci

the 27 comparisons in which no differences were detected at the morphological level, on average 5.5 and 8.1 polymorphic loci were revealed when considering the AFLP set and all the SSR markers, respectively. For both MM types, polymorphism increased as the number of CPVO distinctive characters rose; moreover a considerable widening in the range was observed following an increase of the number of distinctive characters, with minimum values remaining close to zero. The number of differences detected using 28 highly informative SSR markers (1 per chromosome arm) was about half that observed with the whole SSR set (76). When considering only these markers, on average 4.2 differences were observed in morphologically indistinguishable comparisons, with a minimum of 0, a maximum of 11, and a standard deviation of 3.5; it should be pointed out that an equal standard deviation was detected for the 64 comparisons differing for one CPVO character. Using this estimate of the standard deviation, the upper limit of the frequency distribution of the marker differences corresponding to indistinguishable morphological comparisons can be calculated for a given level of probability; with a mean of 4.2 this upper limit is equivalent to 10.1 and 12.5 differences at SSR loci, for $P = 0.05$ and $P = 0.01$, respectively. Therefore a number of SSR differences lower than 11 and 13 is expected in indistinguishable morphological comparisons in 95% and 99% of the cases.

Discussion

In the present work, materials were chosen as being representative of those that can be currently found in applied durum wheat breeding and, as such, characterized by different levels of genetic similarity from closely related to unrelated. As expected on the basis of the high level of inbreeding, the cross-derived lines showed a good degree of uniformity as indicated by the low frequency of heterogeneous loci revealed by SSRs. Moreover, in these materials the occurrence of nonparental alleles was also quite low, whereas in BC-derived lines a higher level was reached. In the latter case, the presence of unexpected alleles can likely be attributed to heterogeneity within the recurrent parent Svevo since in both backcrosses they were detected at the same loci.

An overall consistency in the ranking of the similarity values based on the three marker systems was observed considering the parental lines as well as the progenies; this ranking reflects the degree of relatedness among the considered genetic materials. In fact, on average, the similarity between sublines was higher than between lines; the similarities between BC_3-derived lines as well as between these and the corresponding recurrent parent were higher than between cross-derived lines; within-cross similarity was always higher than between-cross and, in the latter case, it was higher when crosses with a common parent were considered.

Cluster analysis based on SSR and AFLP data produced representations that, compared to those obtained with morphological descriptors, more faithfully reflected the genetic structure of the materials confirming that MMs represent a more suitable tool to study genetic relationships (Graner et al. 1994; Powell et al. 1996; Noli et al. 1997; Beyene et al. 2005). In particular, SSRs showed the highest resolution power, grouping the materials almost perfectly according to their origin. The less accurate classification obtained with AFLPs, even compared to that based on a lower number of SSR markers (28), could be attributed to an uneven and perhaps insufficient sampling of the genome. In fact, the scant information on the map position of the AFLPs studied herein combined with the well-known tendency of this class of markers to cluster in pericentromeric regions (Moore et al. 1993) suggest that a larger number of AFLP markers could be necessary to obtain more accurate estimates of genetic similarity. However, it should be pointed out that we used a methylation-sensitive enzyme (*Sse*8387) to generate restriction fragments, thus the AFLP clustering should, at least partially, have been reduced. In this respect, Powell et al. (1997) in barley, and Law et al. (1998) in bread wheat, observed a more uniform AFLP genomic distribution when using methylation-sensitive enzymes.

The high correlation ($r = 0.89$) between similarity estimates obtained with SSR and AFLP markers observed in this study is quite close to that ($r = 0.81$) reported by Maccaferri et al. (2007) who considered a set of durum wheat varieties representative of different international breeding programs. The lower correlation between similarity values based on molecular markers (either SSR or AFLP) and

morphological descriptors is also in keeping with the findings of the same authors. A moderate relationship between MMs (RFLPs) and morphological data was also reported by Autrique et al. (1996) in durum wheat, whereas in bread wheat low to nonsignificant correlations were observed using different PCR-based markers (Marić et al. 2004; Roy et al. 2004; Fufa et al. 2005). Such poor correlations have been attributed to the limited number of the traits and of the genes underlying their expression (Shut et al. 1997). Moreover, the limited variation often observed for morphological traits, the inaccuracy in the estimate of the similarities therein obtained, and possible epistatic interactions among their determinants could be further reasons for the limited correlations. The lack of a high correlation between molecular and morphological similarities that we observed suggests that, in order to assess the distinctness among accessions, the information provided by the former should be considered as complementary rather than alternative to that obtained by the latter.

MMs have been shown to be a powerful tool for genotype characterization (Law et al. 1998; Lefebvre et al. 2001), and in the present study they allowed for the detection of differences even between highly related genotypes. On the other hand, the difficulty in assessing the distinction through this innovative approach relates to the wide range of variation for MM differences associated to none or very few (1–2) differences in morphological descriptors. Thus, the use of a marker-based threshold of similarity, below which two genotypes are declared as distinct without the need of further evidence of morphological differences, should be restricted to particular situations. In fact, the use of SSRs could be proposed to improve the efficiency of the DUS assessment serving as a prescreening tool to identify well-distinct comparisons that would not need full phenotypic examination (Nuel et al. 2001); nonetheless the application of a set of defined, representative markers, requires the definition of a threshold value of differences at the molecular level calibrated against the level of distinction attained at the morphological level. In this study, using one informative SSR per chromosome arm, we found that 13 or more polymorphisms corresponded to at least one morphological difference, with a probability level of 0.01. Therefore, field examination to assess distinctness with respect to reference varieties could be restricted to comparisons showing 12 or fewer SSR differences. In the materials considered herein, characterized by a rather high level of similarity, the application of this threshold would allow expensive morphological comparisons to be limited to 55.4% of the possible ones. Furthermore, the knowledge of genetic similarities could be helpful to optimize the design of field trials by assigning to neighboring plots the most similar entry pairs.

Prediction of distinctness could be further improved when MMs tightly linked to the genes controlling the expression of morphological descriptors are available. In contrast to this view, considering the advantages of MMs not only for distinctness purposes but also for describing varieties in relation to their value for cultivation and use, other authors (Cook and Reeves 2003) have suggested a wider utilization of molecular tools. In this perspective, MMs linked to relevant agronomic traits, such as disease resistance, adaptation to specific environments or input systems, etc., or to quality traits, could be considered in order to overcome the limitations of the current registration system based only on morphological descriptors.

Acknowledgements The authors wish to thank the Società Produttori Sementi-Bologna for supplying the plant materials and for hosting the field trials. They would also like to thank Dr. A. De Montis, who contributed pedigree information on the materials used in this study, Dr. M. Maccaferri for technical assistance for molecular analysis, and Dr. J.S.C. Smith for critical reading of the manuscript.

References

Autrique E, Nachit MM, Monneveux P, Tanksley SD, Sorrells ME (1996) Genetic diversity in durum wheat based on RFLPs, morphophysiological traits and coefficient of parentage. Crop Sci 36:735–742

Beyene Y, Botha AM, Myburg AA (2005) A comparative study of molecular and morphological methods of describing relationships in traditional Ethiopian highland maize. African J Biotech 4:586–595

Camlin MS (2001) Possible future roles for molecular techniques in the identification and registration of new plant cultivars. Acta Hortic 546:289–296

Cooke RJ, Reeves JC (2003) Plant genetic resources and molecular markers: variety registration in a new era. Plant Genet Res 1:81–87

CPVO (2003) European Union – Community plant variety office. Protocol for distinctness, uniformity and stability tests: durum wheat (*Triticum durum* Desf.) http://www.cpvo.eu.int/documents/TP/agricoles/TPdurumwheat120.pdf

Fehr WR (1984) Genetic contributions to yield grains of five major crop plants. Spec. Publ. 7. CSSA, Madison, WI

Fufa H, Baenziger PS, Beecher BS, Dweikat I, Graybosh RA, Eskridge KM (2005) Comparison of phenotypic and molecular marker-based classifications of hard red winter wheat cultivars. Euphytica 145:133–146

Graner A, Ludwig WF, Melchinger AE (1994) Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. Crop Sci 34:1199–1205

Gupta P, Balyan H, Edwards K, Isaac P, Korzun V, Röder M, Gautier MF, Joudrier P, Schlatter A, Dubcovsky J, De la Pena R, Khairallah M, Penner G, Hayden M, Sharp P, Keller B, Wang R, Hardouin J, Jack P, Leroy P (2002) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. Theor Appl Genet 105:413–422

Guyomarc'h H, Sourdille P, Charmet G, Edwards KJ, Bernard M (2002) Characterisation of polymorphic markers from *T. tauschii* and transferability to the D-genome of bread wheat. Theor Appl Genet 104:1164–1172

ISF (2003) ISF view on intellectual property http://www.worldseed.org/pdf/ISF_View_on_Intellectual_Property.pdf

ISF (2006) Use of DNA markers for DUS testing, essential derivation and identification http://www.worldseed.org/Position_papers/Use_DNA_Markers.htm

Law JR, Donini P, Koebner RMD, Reeves JC, Cooke RJ (1998) DNA profiling and plant variety registration. III: the statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. Euphytica 102:335–342

Lefebvre V, Goffinet B, Chauvet JC, Caromel B, Signoret P, Brand R, Palloix A (2001) Evaluation of genetic distances between pepper inbred lines for cultivar protection purposes: comparison of AFLP, RAPD and phenotypic data. Theor Appl Genet 102:741–750

Lotti C, Salvi S, Pasqualone A, Tuberosa R, Blanco A (2000) Integration of AFLP markers into an RFLP-based map of durum wheat. Plant Breed 119:393–401

Maccaferri M, Stefanelli S, Rotondo F, Tuberosa R, Sanguineti MC (2007) Relationships among durum wheat accessions. I. Comparative analysis of SSR, AFLP, and phenotypic data. Genome 50:373–384

Marić S, Bolarić S, Martinčić J, Pejić I, Kozumplik V (2004) Genetic diversity of hexaploid wheat cultivars estimated by RAPD markers, morphological traits and coefficients of parentage. Plant Breed 123:366–369

Moore G, Abbo S, Cheung W, Foote T, Gale MD, Koebner RMD, Leitch A, Leitch I, Money T, Stancombe P, Yano MM, Flavell RB (1993) Key features of cereal genome organization as revealed by the use of cytosine methylation-sensitive restriction endonucleases. Genomics 15:472–482

Noli E, Salvi S, Tuberosa R (1997) Comparative analysis of genetic relationships in barley based on RFLP and RAPD markers. Genome 40:607–616

Nuel G, Baril C, Robin S (2001) Varietal distinctness assisted by molecular markers: a methodological approach. Acta Hortic 546:65–71

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

Powell W, Thomas WTB, Baird E, Lawrence P, Booth A, Harrower B, Mc Nicol JW, Waugh R (1997) Analysis of quantitative traits in barley by the use of amplified fragment length polymorphism. Heredity 78:48–59

Preston LR, Harker N, Holton T, Morell MK (1999) Plant cultivar identification using DNA analysis. Plant Var Seeds 12:191–205

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. Genetics 149:2007–2023

Rohlf FJ (1997) NTSYS-pc version 2.0. Exter, Setauket, NY

Roy JK, Lakshmikumaran MS, Balyan HS, Gupta PK (2004) AFLP-based genetic diversity and its comparison with diversity based on SSR, SAMPL, and phenotypic traits in bread wheat. Biochem Genet 42:43–59

Royo C, Álvaro F, Martos V, Ramdani A, Isidro J, Villegas D, García del Moral LF (2007) Genetic changes in durum wheat yield components and associated traits in Italian and Spanish varieties during the 20th century. Euphytica 155:259–270

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 83:1757–1761

Shut JW, Qi X, Stam P (1997) Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor Appl Genet 95:1161–1168

Smith JSC, Smith OS (1992) Fingerprinting crop varieties. Adv Agron 47:5–140

Sneath PHA, Sokal RR (1973) Numerical taxonomy. WH Freeman, San Francisco, CA, USA

Soller M, Beckman JS (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 67:25–33

Somers DJ, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor Appl Genet 109:1105–1114

Song QJ, Shi JR, Singh S, Fickus EW, Costa JM, Lewis J, Gill BS, Ward R, Cregan PB (2005) Development and mapping of microsatellite (SSR) markers in wheat. Theor Appl Genet 110:550–560

Sourdille P, Guyomarc'h H, Baron C, Gandon B, Chiquet V, Artiguenave F, Edwards K, Foisset N, Dufour P, Bernard M (2001) Improvement of the genetic maps of wheat using new microsatellite markers. Proceedings of the 9th plant and animal genome. Final Abstract Guide. San Diego January 13–17 2001 P425 p167

UPOV (1961) International convention for the protection of new varieties of plants. http://www.upov.int/en/publications/conventions/index.html

UPOV (2004) Administrative and legal committee – molecular techniques. UPOV CAJ/50/4

van Eeuwijk FA, Baril CP (2001) Conceptual and statistical issues related to the use of molecular markers for distinctness and essential derivation. Acta Hortic 546:35–53

Yu J-K, La Rota M, Kantety RV, Sorrells ME (2004) EST derived SSR markers for comparative mapping in wheat and rice. Mol Genet Genomics 271:742–751