



Association between genetic distances in wheat (*Triticum aestivum* L.) as estimated by AFLP and morphological markers

Eduardo Alano Vieira¹, Fernando Irajá Félix de Carvalho², Ivandro Bertan², Mauricio Marini Kopp², Paulo Dejalma Zimmer², Giovani Benin³, José Antônio Gonzalez da Silva², Irineu Hartwig², Gaspar Malone² and Antônio Costa de Oliveira²

¹*Empresa Brasileira de Pesquisa Agropecuária, Embrapa Cerrados, Planaltina, DF, Brazil.*

²*Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, Pelotas, RS, Brazil.*

³*Faculdade de Agronomia, Universidade Tecnológica Federal do Paraná, Pato Branco, PR, Brazil.*

Abstract

Molecular and morphological data analyses matrices are very informative tools for the estimation of genetic distances. We used AFLP markers, morphological traits and combined analysis to estimate the genetic distances between wheat genotypes and ascertain any associations between the two techniques. Nineteen wheat (*Triticum aestivum* L.) genotypes were analyzed using amplified fragment length polymorphism (AFLP) markers and field evaluated for two years. The matrices obtained by morphological and molecular marker data analyses revealed a significant but moderate correlation ($r = 0.47$), indicating that such techniques sample distinct genome regions. The combined analysis was found to be highly correlated with AFLP ($r = 0.97$) and moderately correlated with morphological ($r = 0.59$) markers. A possible explanation for such results is a bias caused by the much higher number of AFLP (229) than morphological (17) markers. Thus, it is evident that the combined analysis is not efficient when a very dissimilar number of markers are used in each isolated technique. Therefore, to obtain a better knowledge of the degree of divergence among genotypes it is necessary to consider each analysis separately.

Key words: reeding, genetic divergence, molecular markers, morphological traits.

Received: November 28, 2005; Accepted: October 3, 2006.

Introduction

The analysis of genetic distance is an auxiliary tool of great use in breeding programs and an important link between the conservation and use of available genetic resources (Mahammadi and Prasanna, 2003). A knowledge of genetic distance not only produces a better understanding of germplasm organization and higher efficiency during genotype sampling but also results in the biologically oriented choice of crosses and gene introgression from exotic germplasm, and can also be used to recommend cultivars when the goal is to increase the genetic basis of commercial cultivars for a given region.

In a breeding program, the genetic gain achieved through artificial selection is directly related to the amount of variability and gene quality present in the segregating population. Thus, the correct choice of parents employed in the development of the basic population can influence the final result of the artificial selection and promote a better al-

location of financial resources during the whole process of adjusting genotypes to a given environment (Nienhuis *et al.*, 1993; Bohn *et al.*, 1999). However, to confirm such expectations, it is necessary that the parents combine high means with an increase in variability for the characters under selection.

Molecular and morphological analyses are among the most used tools for the estimation of genetic distances within a group of genotypes. Molecular markers provide an excellent tool for obtaining genetic information and their use in the assessment of genetic divergence in wheat (*Triticum aestivum* L.) has increased in the last few years (Manifesto *et al.*, 2001; Corbellini *et al.*, 2002; Almanza-Pizón *et al.*, 2003; Máric *et al.*, 2004; Roy *et al.*, 2004). Also, because molecular markers are not subject to environmental influence they are considered superior to morphological markers (Máric *et al.*, 2004). Amplified fragment length polymorphism (AFLP) markers are the preferred type of molecular markers because of their multiplex power, their efficiency in detecting genetic variability and the robustness of AFLP assays (Vos *et al.*, 1995).

Morphological characters, in association with multivariate techniques, have been employed in quantifying genetic distance in wheat (Zeven and Schachl, 1989; Van Beuningen and Bush, 1997; Máric *et al.*, 2004; Roy *et al.*, 2004). However, phenotype expression is influenced by genotype *vs.* year and genotype *vs.* location interactions, leading to a low accuracy of quantitative genetic parameter estimates.

Nevertheless, the effectiveness of the simultaneous use of AFLP and morphological markers to measure genotype relationships, as well as the magnitude of association between these methods, is not yet fully understood. Thus, the objective of the work presented in this paper was to use AFLP, morphological and combined markers to estimate the genetic distance between wheat (*Triticum aestivum* L.) genotypes and to establish the degree of association between these techniques.

Materials and Methods

The 19 wheat genotypes used in our study are listed in Table 1, two genotypes, the aluminum-tolerant BH1146 and the aluminum-sensitive Sonora64 genotypes (Bertan *et al.*, 2006) being included due to their well known responses to excess aluminum. To measure the morphological characters, experiments were conducted during the 2003 and 2004 cropping seasons in an experimental field belonging to the Genomic and Plant-improvement Center of the Federal University of Pelotas (Centro de Genômica e Fitomelhoramento/Universidade Federal de Pelotas – CGF/UFPEL),

in the municipality of Capão do Leão (31°52'00" S, 52°21'2" W, altitude = 13.24 m) in the southern Brazilian state of Rio Grande do Sul. The experimental design was completely randomized blocks with three replications, and the plots were composed of five 5 m rows spaced 0.2 m apart. It were disregarded the 2 outer rows and 0,5 m on the end of each of the 3 left inner rows, thus the useful area of each plot consisted of 4 m length of each three inner row. Each plot was base-fertilized with the equivalent of 300 kg ha⁻¹ of 5-20-20 NPK fertilizer and top-dressed with the equivalent of 60 kg ha⁻¹ N at tilling. Weeding was performed manually and ants were controlled using granulated baits, other pest and disease control measures being carried out according to the recommendations of the Southern Brazilian Wheat Research Committee (Comissão Sul Brasileira de Pesquisa de Trigo, 2002).

A total of 17 morphological characters were determined, according to procedures presented by Scheeren (1984): i) days from emergence to flowering (DEF); ii) days from emergence to maturation (DEM); iii) days from flowering to maturation (DFM); iv) plant stature in cm (PS); v) number of fertile tillers per linear meter (FTLM); vi) weight of a thousand grains (WTG); vii) hectoliter weight in kg hL⁻¹ (HW); viii) grain yield in kg ha⁻¹ (GY). The following morphological characters were measured on a sample of 25 plants per plot: ix) flag leaf blade width in cm (LBW); x) flag leaf blade length in cm (FLL); xi) leaf sheath length in cm (LSL); xii) peduncle length in cm (PL). In addition, the following characters were determined on a sample of 25 spikes per plot: xiii) spike length in cm (SL); xiv) number of spikelets per spike (NS); xv) spike weight in g (SW); xvi) number of grains per spike (NG) and xvii) number of grains per spikelet (NG/NS).

Morphological character data were subjected to analysis of variance (ANOVAR), considering the effects of genotypes and years as fixed. The character means, based on two years, were compared using the least significant differences (LSD) method (Steel and Torrie, 1980) at the 5% probability level ($p = 0.05$) and the generalized Mahalanobis distance (D^2) was obtained for all genotype pairs, based on two years of evaluation, using the Genes software (Cruz, 2001). From the genetic distance matrix, a dendrogram was constructed using the unweighted pair group method with arithmetic means (UPGMA) method. The adjustment between the distance matrix and the dendrogram was estimated by the cophenetic correlation coefficient (r ; Sokal and Rolf, 1962) using the NTSYS pc 2.1 software (Rolf, 2000).

The DNA used in the AFLP analyses was extracted according to the protocol described in Saghai-Marouf *et al.* (1984). The analyses of AFLP markers were performed according to the protocol described by Vos *et al.* (1995). The six primer combinations used were *M*-CTA/*E*-ACT (C_1), *M*-CAA/*E*-AGC (C_2), *M*-CAG/*E*-AGC (C_3), *M*-CAC/*E*-

Table 1 - Wheat genotypes used.

Genotype	Year of release*
BH1146	1955
BR18	1986
BR23	1987
BR35	1989
BRS49	1996
BRS119	1997
BRS120	1997
BRS177	1999
BRS194	2000
BRS208	2001
CEP24	1992
FUNDACEP29	1997
BRS192	2000
PF950354	-
RUBI	1998
SONORA64	1964
TB951	-
ICA1	1999
ICA2	1999

*(-)genotype not yet released commercially.

ACA (C₄), M-CAA/E-ACA (C₅) and M-CAG/E-ACT (C₆), where C_i represents the *i*th primer combination and *E* and *M* represent the *Eco*RI and *Mse*I restriction enzymes. The amplified fragments were electrophoresed in (w/v) 6% denaturing polyacrylamide gels and stained using silver nitrate (Creste *et al.*, 2001). Bands were scored as binary data (1 = presence and 0 = absence) and the average polymorphic information content (PIC) was calculated for each primer combination by applying the formula $PIC = 1 - \sum p_i^2$, where p_i is the frequency of the *i*th allele (Powell *et al.*, 1996). The marker index (MI) was calculated for each primer combination as $MI = PIC \times np_i$, where np_i is the number of polymorphic bands (Powell *et al.*, 1996).

A genetic similarity calculation was performed using the software NTSYS pc 2.1. The genetic similarity (S_{ij}) was measured using the Dice coefficient (Dice, 1945) according to the equation $S_{ij} = 2 N_{ij}/(N_i + N_j)$, where N_{ij} is the number of bands present in both genotypes *i* and *j*, N_i is the number of bands present in genotype *i*, and N_j is the number of bands present in genotype *j*. The genetic similarity was converted to genetic dissimilarity according to the equation $D_{ij} = 1 - S_{ij}$, in which D_{ij} is the genetic dissimilarity in each pair of *i* and *j* genotypes, and S_{ij} is the genetic similarity between each pair of *i* and *j* genotypes. The dissimilarity matrix produced was used to generate an UPGMA dendrogram, the adjustment between the dissimilarity matrix and the dendrogram being estimated from the cophenetic correlation coefficient (*r*) using the NTSYS pc 2.1 software. Cluster stability was measured by bootstrap analysis with 1,000 replications using the Winboot software (Yap and Nelson, 1996). The Genes software (Cruz, 2001) was used to estimate the minimum number of markers needed for the estimation of genetic distance with a correlation coefficient (*r*) of at least 0.95.

Genetic similarity was also estimated between all genotype pairs using the similarity index proposed by Gower (1971) which uses both binary and quantitative morphological data to estimate a unique similarity index ranging from 0 to 1, calculated as

$$S_{\alpha\beta} = \frac{\sum_{h=1}^p t_{\alpha\beta h} \delta_{\alpha\beta h}}{\sum_{j=1}^p \delta_{\alpha\beta j}}$$

where $t_{\alpha\beta h}$ is adjusted for the type of variable (*h*) and where if *h* is binary (as in AFLP analysis) then $t_{\alpha\beta h}$ equals 0 and $\delta_{\alpha\beta h} = 1$ if *h* for each genotypes is different while both $t_{\alpha\beta h}$ and $\delta_{\alpha\beta h}$ equal 1 if *h* is present in both genotypes or 0 if *h* is absent from both genotypes but if *h* is a quantitative variable, such as is the case for some morphological parameters, then

$$t_{\alpha\beta h} = 1 - \frac{|x_{\alpha h} - x_{\beta h}|}{(\max x_h - \min x_h)}$$

and $\delta_{\alpha\beta h}$ equals 1, where $\delta_{\alpha\beta h}$ is the variable value (*h*) for genotype α and $x_{\beta h}$ is the value for the same variable in genotype β . The genetic similarity was estimated using the Multiv v. 2.3 software (Pillar, 1997), was converted into genetic dissimilarity according to the equation $D_{\alpha\beta} = 1 - S_{ij}$ in which $D_{\alpha\beta}$ is genetic dissimilarity between each pair of α and β genotypes while $S_{\alpha\beta}$ is the genetic similarity between each pair of α and β genotypes. The dissimilarity matrix generated was used to construct an UPGMA dendrogram, the adjustment between the dissimilarity matrix and the dendrogram being estimated from the cophenetic correlation coefficient (*r*) using the NTSYS pc 2.1 software.

Mantel test (Mantel, 1967) with 1,000 permutations was used to estimate the correlation (association) significance between the distance matrices resulting from morphological, AFLP and combined analyses, the test being calculated using the NTSYS pc 2.1 software.

Results and Discussion

The analysis of variance indicated that for all the characters evaluated there were statistically significant differences ($p = 0.05$) between the genotypes studied and, for most of the characters evaluated, for years and the genotype times year interaction, justifying the need for evaluating genotypes for more than one year in order to obtain a reliable estimate of individual means for the majority of characters evaluated (Table 2). The coefficients of variation for the data shown in Table 2 were low (2 to 13%), indicating high experimental accuracy for the study.

Examination of maximum and minimum values for each of the characters measured showed that some genotypes presented means located at the top and bottom limits for a large number of characters. For example, the Sonora 64 genotype showed mean values at the bottom limit of the range for seven (DEF, DEM, LBW, LSL, GY, HW and NG/NS) out of the 17 measured characters (Table 2) while for genotype TB 951 a total of six characters had mean values located at the top (characters LBW, FLL, NE and SL) or bottom (characters PS and PL) limits (Table 2), indicating that these genotypes have very different morphological characters compared to the other genotypes. According to the LSD test, only the DEF characteristic showed no genotypes with means significantly below the overall mean while only the NG/NE characteristic showed no genotypes with means significantly above the overall mean (Table 2).

The overall Mahalanobis distance (D^2) estimated using morphological characters revealed that the most distant genotypes were TB 951 and BRS 194 while the closest were BRS 119 and BRS 208. The dendrogram showed that the most distinct genotypes were Sonora 64, TB 951 and BR 18 (Figure 1). The large distance found between the

Table 2 - Comparison of means and summary of variance of the characters: days from emergence to flowering (DEF), days from emergence to maturation (DEM), days from flowering to maturation (DFM), plant stature (PS), number of fertile tillers per linear meter (FTLM), flag leaf blade width (LBW), flag leaf blade length (FLL), leaf sheath length (LSL), peduncle length (PL), grain yield (GY) hectoliter weight (HW), number of spikelets per spike (NS), spike length (SL), spike weight (SW), number of grains per spikelet (NG/NS) and weight of a thousand grains (WTG) evaluated in 19 wheat genotypes.

Genotypes [†]	Characters [‡]																
	DEF (days)	DEM (days)	DFM (days)	EP (cm)	FTLM (unit.)	LBW (cm)	FLL (cm)	LSL (cm)	PL (cm)	GY (kg ha ⁻¹)	HW (kg hL ⁻¹)	NS (unit.)	SL (cm)	SW (g)	NG (unit.)	NG/NS (unit.)	WTG (g)
BH1146	71.83	128.67 S	56.83 S	103.33 S	84.33	1.26	19.06 I	19.80 S	20.07	2418.52 I	72.43 I	14.97	8.99	1.74	32.97	2.20	41.43
BR18	72.00	112.83	40.83	79.17 I	75.67	1.45	24.03 S	17.20	16.09	2311.11 I	70.90 I	14.76 I	9.01	2.00	31.95 I	2.18	44.06 S
BR23	75.83 S	122.33 S	46.50	95.83	82.50	1.33	21.95	17.34	21.02 S	3682.41 S	76.67	17.13	10.15 S	2.15	39.90	2.33	40.21
BR35	79.67 S	116.67 S	37.00	96.67	81.00	1.39	20.80	20.16 S	16.22	3347.22	74.53	16.54	8.65	2.09	39.11	2.35	40.74
BRS49	72.33	113.67	41.33	103.33 S	87.67	1.45	19.82	18.39	17.92	2911.11	73.00	17.65 S	8.43	2.05	41.33	2.36	36.39
BRS119	70.83	114.17	43.33	90.83	83.17	1.24	24.85 S	18.77	18.21	3072.22	76.20	17.20	9.03	2.23	44.50	2.59	38.13
BRS120	72.83	116.33 S	43.50	96.67	114.33 S	1.46	24.33 S	18.02	18.99	3324.08	76.73	16.00	8.90	1.95	34.70	2.16	41.03
BRS177	73.17 S	118.67 S	45.50	95.00	109.33 S	1.30	23.13	18.32	18.24	4159.19 S	77.77 S	15.77	9.25	1.88	40.67	2.58	35.37
BRS194	68.50	114.67	46.17	100.00	104.17	1.24	16.28 I	15.82 I	20.04	3690.67 S	72.70 I	17.27	7.83 I	2.09	39.30	2.28	40.21
BRS208	66.67	110.17	43.50	90.83	86.50	1.20	22.22	18.14	18.17	3271.30	75.47	15.93	8.18	2.08	39.57	2.48	40.65
CEP24	65.17	107.83	42.67	109.17 S	91.83	1.37	24.46 S	21.17 S	24.21 S	3230.56	76.60	15.48	8.98	2.01	33.80	2.18	46.64 S
CEP29	59.83	108.33	48.50	85.83 I	82.83	1.50 S	20.51	16.83	22.79 S	3166.67	74.97	16.03	9.18	2.30	42.83	2.67	40.67
BRS192	65.17	111.00	45.83	87.50	74.67	1.47	22.73	17.74	14.86 I	3223.87	77.90 S	15.87	8.62	2.26	41.93	2.64	35.20
PF950354	63.00	100.00 I	37.00	107.50 S	85.50	1.26	19.78	17.43	20.05	2944.45	76.00	17.50	10.68 S	2.25	41.47	2.38	39.88
RUBI	63.17	102.33 I	39.17	92.50	115.83 S	1.25	19.92	16.18 I	15.48 I	3938.89 S	79.50 S	14.93	8.11	1.85	35.23	2.36	37.14
SONORA64	44.83 I	86.50 I	41.67	82.50 I	80.83	1.01 I	19.08 I	15.39 I	15.78 I	1697.22 I	69.90 I	14.90	8.30	1.48 I	29.41 I	1.98 I	36.23
TB951	58.50 I	98.33 I	39.83	69.17 I	77.17	1.71 S	24.78 S	16.18 I	12.39 I	2404.54 I	75.80	18.73 S	11.70 S	2.48 S	49.57 S	2.64	35.70
ICAI	56.17 I	100.50 I	44.33	105.83 S	97.17	1.31	23.48 S	18.61	25.92 S	2956.48	76.10	15.77	8.14	1.81	37.22	2.35	34.63
ICAI2	53.33 I	108.50	55.17 S	84.17 I	67.67 I	1.27	15.84 I	15.41 I	13.51 I	2557.41 I	73.70	16.18	8.64	1.93	40.38	2.49	33.69 I
SV (DF)																	
MS genotypes (18)	447*	556*	163*	658*	1114*	0.13*	48*	15*	74*	2195747*	36*	7*	5*	0.32*	141*	0.22*	70*
MS years (1)	26163*	22709*	122	3427*	10827*	0.09	50*	86*	272*	18101066*	24*	13	3	1*	597*	0.88*	4
MS G x Y (18)	407*	699*	159*	56*	252*	0.03*	7*	2*	10*	604834*	15*	1	0.28	0.16*	52*	0.17*	27*
MS residue (72)	19	11	22	17	142	0.01	2	0.50	2	84168	2	0.70	0.33	0.07	17	0.05	8
CV (%)	7	3	11	4	13	7	6	4	8	9	2	5	6	13	11	9	7
Mean	65.94	110.10	44.14	93.46	88.53	1.34	21.42	17.73	18.41	3068.94	75.10	16.24	8.99	2.03	38.73	2.38	38.84
Amplitude	34.84	42.17	19.83	40	48.16	0.7	8.94	5.78	13.53	2461.97	9.6	3.97	3.87	1	20.16	0.69	10.37
LSD (5%)	7.00	5.30	7.55	6.73	19.33	0.15	2.04	1.12	2.39	471.40	2.37	1.36	0.93	0.43	6.67	0.36	4.49

[†]Abbreviations: CV = coefficient of variation; DF = degrees of freedom; MS = mean square; LSD = least square difference; SV = source of variation.

[‡]Means followed by S and I in the columns, are, respectively, superior or inferior to the character mean at p = 0.05.

*significant at p = 0.05 by the F test.

Sonora 64 and TB 951 genotypes and the other genotypes in the dendrogram generated by morphological characters was expected, since their means were located at the top or bottom limits for a large number of phenotypic characters. The cophenetic correlation coefficient ($r = 0.80$) showed a fair degree of agreement between the graphical representation of distances and the original matrix, supporting the visual inferences drawn from Figure 1.

The six AFLP primer combinations used generated a total of 262 bands, of which 239 (91.2%) were polymorphic among the 19 genotypes studied. Analysis of the minimum number of informative markers revealed that among all the markers obtained at least 200 were required for the combined analysis to have a correlation coefficient of 0.95. Since the minimum number of informative markers was very close to the number of polymorphic markers evaluated (239) subsequent analyses were performed with all markers.

The PIC values were very close for the primer combinations used, so that the higher MI values were detected for combinations presenting the higher number of polymorphic bands (Table 3). The combinations producing the highest number of polymorphic bands (48) were *M-CTA/E-ACT* and *M-CAC/E-ACA* while the combination producing the

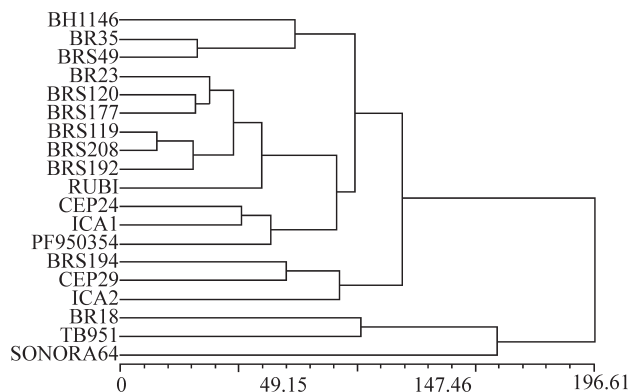


Figure 1 - Morphological character unweighted pair group method with arithmetic means (UPGMA) dendrogram resulting from the analysis of 19 wheat genotypes using the Mahalanobis distance based on 17 phenotypic characters as a measure of genetic distance. The value of the cophenetic correlation coefficient (r) was 0.80.

lowest number of polymorphic bands (31) was *M-CAA/E-ACA*. The low number of monomorphic bands (23) obtained by using all primer combinations demonstrated that AFLP analysis has a high potential for detecting the genetic variability present in these wheat genotypes. A similar scenario was reported by Corbellini *et al.* (2002) who analyzed 40 wheat genotypes from Central and Southern Europe using five AFLP primer combinations and obtained an average of 40 polymorphic bands per primer combination and a total of 200 polymorphic bands. Slightly lower levels of polymorphism have also been detected in wheat, such as the 59% reported by Almanzá-Pizon *et al.* (2003) and the 47% reported by Roy *et al.* (2004). Taken together, these results indicate that AFLP markers are efficient in detecting genetic variability in wheat.

The genetic dissimilarity estimated using AFLP markers showed that the most similar genotypes were BR 35 and BRS 120 and the most dissimilar were TB 951 and ICA 1. Clustering percentage values above 30% for 1,000 bootstrap cycles occurred in only four groups (BR 18 and TB 951 (57%), BRS 119 and BRS 192 (40%), BRS 35 and BRS 120 (37%) and BR 23 and BRS 177 (30%)), revealing that such clusters are the most consistent (Figure 2). The cophenetic correlation coefficient for the dendrogram ($r = 0.85$) indicated good agreement between the graphical display of distances and the original matrix, supporting the visual inferences suggested in Figure 2.

Of the four most consistent AFLP analysis clusters, three (BR 18 and TB 951; BR 23 and BRS 177; BRS 119 and BRS 192) were consistent with the distance estimated using morphological characters (Figures 1 and 2). A certain agreement between the distances estimated through these two techniques was found, as evidenced by a moderate but significant correlation ($r = 0.47$) between the morphological genetic distance matrix and the AFLP marker matrix (Table 4). Previous studies have reported that small genetic distances as estimated by molecular markers were consistently associated with small phenotypic distances (Dillmann *et al.*, 1997; Lefebvre *et al.*, 2001) while large molecular distances can either be associated with large or small phenotypic distances (Dillmann *et al.*, 1997; Lefebvre *et al.*, 2001).

Table 3 - Amplified fragment length polymorphism (AFLP) primer combinations, total number of bands, number and percentage of polymorphic bands, average polymorphic information (PIC) and marker index (MI) per primer combination for 19 wheat genotypes.

AFLP primer combination	Total number of bands	Number of polymorphic bands (%)	PIC	MI
<i>M-CTA/E-ACT</i>	51	48 (94.1%)	0.32	15.36
<i>M-CAA/E-AGC</i>	41	38 (92.7%)	0.36	13.68
<i>M-CAG/E-AGC</i>	38	34 (89.5%)	0.37	12.58
<i>M-CAC/E-ACA</i>	52	48 (92.3%)	0.35	16.80
<i>M-CAA/E-ACA</i>	36	31 (86.1%)	0.37	11.47
<i>M-CAG/E-ACT</i>	44	40 (90.9%)	0.39	15.60
Average	44	40 (91.2%)	0.36	14.25

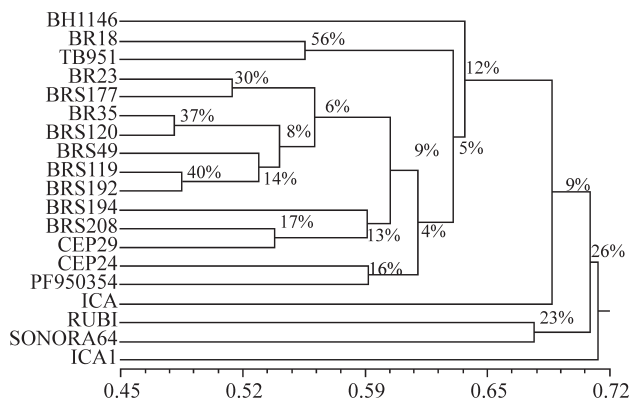


Figure 2 - Amplified fragment length polymorphism (AFLP) marker unweighted pair group method with arithmetic means (UPGMA) dendrogram resulting from the analysis of 19 wheat genotypes using the complement of the similarity index of Dice (1945) based on AFLP markers as a measure of genetic distance. The values found in the groups indicate the percentage number of times that the genotypes grouped together out of 1000 bootstrap cycles using the Winboot program. The value of the cophenetic correlation coefficient (r) was 0.85.

The moderate association between genetic distances estimated using molecular and phenotypic markers can be explained by a range of factors. Molecular analysis provides a wider genome sampling than the morphological analysis, since a study comparing both techniques rarely evaluates the same, or even a similar, number of morphological and molecular markers. The association between estimates is also influenced by the fact that a large portion of the variation detected by molecular markers is non adaptive and, therefore, not subject to either natural or artificial selection. On the other hand, the phenotypic characters are subject to both natural and artificial selection, aside from their high environmental dependence. Moreover, it is not always the case that two identical phenotypes are determined by the same genes, *i.e.*, distinct genes may lead to similar phenotypes. Thus, it is clear that such estimates are closer when there is an association between the loci controlling the targeted morphological traits (quantitative trait loci, or QTLs) and the evaluated bands and when a large number of morphological traits are evaluated (Dillmann *et al.*, 1997; Schut *et al.*, 1997; Lefebvre *et al.*, 2001; Máric *et al.*, 2004; Roy *et al.*, 2004).

Máric *et al.* (2004) investigated wheat and reported an $r = 0.12$ correlation between distances estimated using random amplified polymorphic DNA (RAPD) markers and 12 morphological characters. This was lower than the $r = 0.47$ correlation detected by us, possibly because we used a larger number of morphological characters (17) and hence a larger sample of the wheat genome. In another study of wheat, Roy *et al.* (2004) reported that the correlation between genetic distances estimated in wheat using AFLP markers and 14 morphological characters was $r = 0.072$, indicating an association close to null. However, although the morphological measurements conducted by Roy *et al.* (2004) were also performed in two years the ge-

notypes were evaluated under spaced-plant conditions, unlike in our study in which the genotypes were evaluated under full-row conditions. Taken together, these studies may indicate that under full-row conditions genetic distances tend to be closer than under spaced-plant conditions in which genotypes can express their phenotype without facing competition for space and nutrients.

Dreisigacker *et al.* (2004) analyzed 68 advanced wheat lines using 99 simple sequence repeat (microsatellite) (SSR) markers but did not differentiate the lines regarding the five distinct mega-environments in which they were selected. According to Dreisigacker *et al.* (2004) there may be three possible reasons for unsuccessful separation: i) selection based on mega-environment adaptation has not been practiced long enough to differentiate the germplasm; ii) genes conferring fitness to one mega-environment are not unique to that mega-environment and may confer fitness to several other mega-environments; and iii) adaptation to a mega-environments is not based on accretion of random genes but rather on a limited set of specific genes.

In our study, genetic dissimilarity as estimated through the combined analysis of AFLP and morphological characters showed that the most similar genotypes were BR 35 and BRS 120 and that the most dissimilar were TB 951 and ICA 1. The most distinct genotypes in the collection used in this study were ICA 2 and Sonora 64 (Figure 3). The cophenetic correlation coefficient of the dendrogram ($r = 0.81$) showed good agreement between the graphical representation of the distances and the original matrices, which enabled more accurate visual inferences to be drawn using Figure 3.

We also found a higher correlation ($r = 0.97$) between the combined and molecular matrices than between the combined and morphological matrices ($r = 0.59$), with both correlations being significant at $p = 0.05$ by the Mantel test

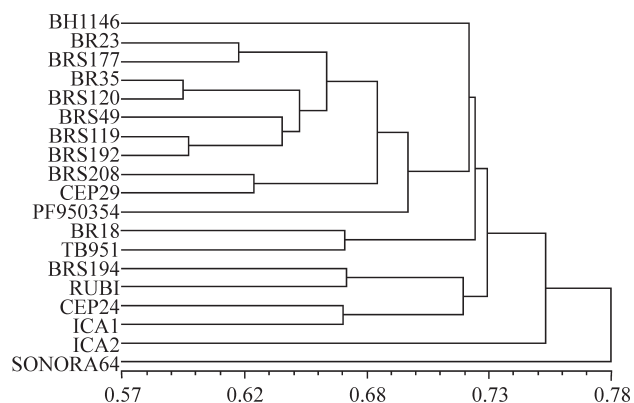


Figure 3 - Combined morphological character and amplified fragment length polymorphism (AFLP) marker unweighted pair group method with arithmetic means (UPGMA) dendrogram resulting from the analysis of 19 wheat genotypes using the complement of the similarity index of Gower (1971) obtained from the combined analysis of phenotypic and AFLP markers as a measure of genetic dissimilarity. The value of the cophenetic correlation coefficient (r) was 0.81.

with 1,000 permutations (Table 4). Such difference in association between matrices could be based on the different number of data points for AFLP markers (229) and phenotypic characters (17). The results obtained in our study suggests that to obtain a more complete understanding of the degree of genotype divergence it is necessary to consider the molecular and morphological data separately. Franco *et al.* (2001) have suggested that better genotype discrimination is obtained by the combined molecular and morphological data when one determines, *a priori*, the minimum number of markers that will lead to the same results as the combination of all markers. However, Warburton *et al.* (2002) has pointed out that this strategy may lead to errors because the presence of translocated segments on chromosomes within the studied genotypes can lead to overestimation of genetic distances, especially when there is no information available about the presence or absence of the translocated segments concerned. Warburton *et al.* (2002) also showed that only four AFLP markers, located on translocated wheat segments, out of 40 polymorphic markers were sufficient to mistakenly cluster 101 sister lines only as function of the presence of translocated segments and that separation of the 101 sister lines as a function of their pedigrees was achieved only when these four markers were eliminated from the analyses.

In our study, the morphological and molecular marker analyses showed that the most divergent genotype is Sonora 64 (Figures 1 and 2). However, this genotype should not be prioritized in crosses, because it presents the lowest grain yield value (Table 2). Among the characters most targeted in plant breeding are those related to productivity (GY and HW) which determine the ultimate genotype performance. For such characters, the choice of cross aiming to increase the productivity plateau should include genotypes with high GY, however divergent for the pool of evaluated characters. Genotypes that are high yielding and also divergent probably have different QTLs controlling GY and these QTLs could be combined into a new genotype superior to both parents (transgressive segregant). In this sense, BR 23 and/or RUBI are promising genotypes, since they fulfill the requirements of being high yielding (Table 2) and divergent as indicated by both morphological and molecular data (Figures 1 and 2). Among the possible

crossing combinations, special attention should be drawn to the promising cross between BRS 177 and Rubi, which was high yielding, divergent and demonstrated high HW (Table 2, Figures 1 and 2).

Acknowledgments

The authors are grateful to the Brazilian agencies FAPERGS, CNPq and CAPES for grants and graduate research and research productivity fellowships. The authors also thank Dr. Ana Pontaroli and Dr. Shavanor Smith from the Bennetzen's laboratory at the University of Georgia for kindly revising this manuscript.

References

- Almanza-Pinzón MI, Hairallah M, Fox PN and Walburton ML (2003) Comparison of molecular markers and coefficients of parentage for the analysis of genetic diversity among spring bread wheat accessions. *Euphytica* 130:77-86.
- Bertan I, Carvalho FIF, Oliveira AC, Silva JAG, Benin G, Vieira EA, Silva GO, Hartwig I, Valério IP and Finatto T (2006) Dissimilaridade genética entre genótipos de trigo avaliados em cultivo hidropônico sob estresse por alumínio. *Bragantia* 65:55-63 (abstract in English).
- Bohn M, Friedrich H and Melchinger AE (1999) Genetic similarities among Winter wheat cultivars determined on basis of RFLPs, AFLPs, SSRs and their use for predicting variance. *Crop Sci* 39:228-237.
- Comissão Sul Brasileira de Pesquisa de Trigo (2002) Indicações Técnicas da Comissão Sul-Brasileira de Pesquisa de Trigo. Embrapa Trigo, Passo Fundo, 74 pp.
- Corbellini M, Perenzin M, Accerbi M, Vaccino P and Borghi B (2002) Genetic diversity in bread wheat, as revealed by coefficient of parentage and molecular markers, and its relationship to hybrid performance. *Euphytica* 123:273-285.
- Creste S, Tulmann-Neto A and Figueira A (2001) Detection of single sequence repeat polymorphism in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol Biol Rep* 4:1-8.
- Cruz CD (2001) Programa Genes: Aplicativo Computacional em Genética e Estatística. UFV, Viçosa, 648 pp.
- Dice LR (1945) Measures of the amount of ecological association between species. *Ecology* 26:297-307.
- Dillmann C, Bar-Hen A, Guérin D, Charcosset A and Murigneux A (1997) Comparison of RFLP and morphological distances between maize *Zea mays* L. inbred lines. Consequences for germplasm protection purposes. *Theor Appl Genet* 95:92-102.
- Dreisigacker S, Zhang P, Warburton ML, Van Ginkel M, Hoisington D, Bohn M and Melchinger AE (2004) SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to defferent megaenvironments. *Crop Sci* 44:381-388.
- Franco J, Crossa J, Ribaut JM, Bertran J, Warburton ML and Khairallah M (2001) A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. *Theor Appl Genet* 103:944-952.
- Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* 27:857-874.

Table 4 - Correlation between genetic distance matrices estimated through the analyses of phenotypic characters (MD_P), amplified fragment length polymorphism (AFLP) markers (MD_{AFLP}) and combined analysis (MD_{P+AFLP}) for 19 wheat genotypes.

	MD_P	MD_{AFLP}	MD_{P+AFLP}
MD_P	1	0.47*	0.59*
MD_{AFLP}		1	0.97*
MD_{P+AFLP}			1

*Significant correlation at $p = 0.01$ by the Mantel test with 1000 permutations.

- Lefebvre V, Goffinet B, Chauvet JC, Caromel B, Signoret P, Brand R and 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.
- Manifesto MM, Schlatter AR, Hopp HE, Suárez EY and Dubcovsky J (2001) Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers. *Crop Sci* 41:682-690.
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209-220.
- Máric S, Laric S, Artincic J, Pejic I and Kozumplink V (2004) Genetic diversity of hexaploid wheat cultivars estimated by RAPD markers, morphological traits and coefficients of parentage. *Plant Breed* 123:366-369.
- Mohammadi SA and Prasanna BM (2003) Analyses of genetic diversity in crop plants – Salient statistics tools and considerations. *Crop Sci* 43:1235-1248.
- Nienhuis J, Slocum MK, Devos DA and Muren R (1993) Genetic similarity among *Brassica oleracea* genotypes as measured by restriction fragment length polymorphisms. *J Am Hort Sci* 118:298-303.
- Pillar VP (1997) Multivariate exploratory analysis and randomization testing using Multiv. *Coenoses* 12:145-148.
- Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2:225-238.
- Rohlf FJ (2000) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, version 2.1. Exeter Software, New York, 83 pp.
- Roy JK, Lakshmikumaran MS, Balyan HS and Gupta PK (2004) AFLP-based genetic diversity and its comparison with diversity based on SSR, SAMPL, and phenotypic traits in bread wheat. *Bioc Gen* 42:43-59.
- Saghai-Marouf MA, Soliman KM, Jorgensen RA and Allard RW (1984) Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosome location and population dynamics. *Proc Nat Acad Sci USA* 89:1477-1481.
- Scheeren PL (1984) Instruções para Utilização de Descritores de Trigo (*Triticum* sp.) e Triticale (*Triticum* sp.). Embrapa Trigo, Passo Fundo, 32 pp.
- Schut JW, Qi X and Stam P (1997) Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theor Appl Genet* 95:1161-1168.
- Sokal RR and Rohlf FJ (1962) The comparison of dendrograms by objective methods. *Taxon* 11:30-40.
- Steel RGD and Torrie JH (1980) Principles and Procedures of Statistics: A Biometrical Approach. McGraw-Hill Book Company, New York, 633 pp.
- Van Beuningen LT and Bush RH (1997) Genetic diversity among North American spring wheat cultivars: III. Cluster analysis based on quantitative morphological traits. *Crop Sci* 37:981-988.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: A new technique for DNA fingerprinting. *Nuc Acids Res* 23:4407-4414.
- Warburton ML, Skovmand B and Mujeeb-Kazi (2002) The molecular genetic characterization of the 'Bobwhite' bread wheat family using AFLPs and the effect of the T1BL.1RS translocation. *Theor Appl Genet* 104:868-873.
- Yap IV and Nelson RJ (1996) Winboot: A program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI, Manila, 22 pp.
- Zeven AC and Schachl R (1989) Groups of bread wheat landraces in the Australian Alps. *Euphytica* 41:235-246.

Associate Editor: Márcio de Castro Silva Filho