Association analysis of agronomically important traits using SSR, SAMPL and AFLP markers in bread wheat

J. K. Roy[1,2], R. Bandopadhyay[1], S. Rustgi[1], H. S. Balyan[1] and P. K. Gupta[1,*]

[1]Molecular Biology Laboratory, Department of Genetics and Plant Breeding, Ch. Charan Singh University, Meerut 250 004, India
[2]Present address: Department of Plant Pathology, University of Minnesota, 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA

In bread wheat, marker-trait associations were studied for 14 agronomic traits using a set of 519 DNA-based molecular markers (derived from 20 SSR, 2 SAMPL and 8 AFLP primer pairs) with a set of 55 elite wheat genotypes. Using simple linear and multiple regressions, each of the 14 traits was regressed on all 519 available polymorphic markers (221 SSR markers, 43 SAMPL markers and 255 AFLP markers). A total of 131 SSR, 43 SAMPL and 166 AFLP markers gave significant associations with at least one of the 14 traits, either with linear or with multiple regression; 51 of these markers showed association using both approaches (linear and multiple regressions), and therefore may prove useful for marker-assisted breeding after necessary validation.

Keywords: AFLP, association analysis, regression, SAMPL, SSR, wheat.

DURING the last two decades, DNA-based molecular markers have been extensively used for a variety of purposes in many animal and plant systems, and bread wheat is no exception to this trend[1,2]. Using linkage-based association analysis (including QTL interval mapping) in bread wheat, a large number of genes for various traits (quality traits, resistance to biotic and abiotic stresses, etc.) have already been tagged with markers[2,3]. However, linkage-based studies conducted in the past allowed identification of genes/QTLs at distances as large as 10–30 cm from the closest markers, which is hardly suitable either for marker-assisted breeding or for identification/cloning of functional genes. Furthermore, in linkage-based analyses, only few genotypes that are used as parents of mapping populations could be screened for marker-trait associations, placing another limitation. In order to overcome these limitations of linkage-based analysis, in the recent past, association studies have been conducted, which not only allow mapping of genes/QTLs with higher level of confidence, but also allow detection of genes/QTLs, which would otherwise escape detection in linkage-based studies[4,5]. In humans, association mapping has already proved extremely useful in tagging a large number of Mendelian genes, particularly those involved in various diseases[6]. In plant systems also, a few association studies have been conducted[7], including two recent studies in wheat[8,9]. Availability of a large number of markers in bread wheat prompted us to utilize this approach in yet another study aiming at the detection of marker–trait associations in this crop. The present study involved regression analyses (both linear and multiple regressions) using as dependent variables the data on 14 individual phenotypic traits recorded in 55 elite exotic genotypes, and as independent variables the molecular marker data for the corresponding genotypes with each of a large number of DNA-based molecular markers (SSR, SAMPL and AFLP markers).

*For correspondence. (e-mail: pkgupta36@vsnl.com)

Materials and methods

Seed materials and field experiment

Seeds of 55 wheat accessions/genotypes were procured from the Directorate of Wheat Research, Karnal, India. These genotypes originated in 29 countries belonging to six continents. Information on the name/pedigree of the accessions is available elsewhere[10]. These 55 genotypes were grown in a randomized block design with three replications at the experimental farm, Ch. Charan Singh University, Meerut. Each genotype was evaluated in a single-row plot of 2.25 m with a plant-to-plant distance of 15 cm and row-to-row distance of 23 cm.

Recording of phenotypic data

Data from 15 plants (five plants per replication) of each genotype were recorded on the following 14 phenotypic traits: plant height (cm), days to flowering, days to maturity, tiller number, flag leaf area (cm^2), peduncle length (cm), spike length (cm), number of spikelets per spike, number of florets per spike, number of grains per spike, biological yield per plant (g), grain yield per plant (g), harvest index and 1000 grain weight (g).

Table 1. Details of 20 SSR primers used for PCR amplification of 55 elite wheat genotypes, their polymorphic information content (PIC) values, associated genes and associated characters

Primer designation	Chromosome assignment*	Number of alleles	PIC	Associated gene	Associated character**
WMC24	1AS	12	0.90	*Tri A1*	Plant height, spikelets/spike
WMC25	2BS, 2DS	11	0.77	*Lr-16*	Days to flowering, flag leaf area
WMC35	4B-	5	0.75	–	NA
WMC44	1BL	13	0.86	–	Harvest index
WMC47	4BS, 5AS, 5BS	4	0.73	–	NA
WMC76	7BL	7	0.73	–	NA
WMC83	7AS	9	0.85	–	Days to flowering, tiller no., grain yield, harvest index
WMC120	1A-	2	0.46	–	NA
WMC149	2AS, 2BL, 5BS	9	0.77	*Lr-16*	Spikelets/spike, days to flowering
WMC167	2DL	11	0.83	–	Days to flowering, tiller no., grain/spike, harvest index
WMC169	3AL	9	0.79	–	Days to flowering
WMC170	2A-	10	0.80	–	Days to flowering, biological yield, harvest index, days to maturity, florets/spike
WMC177	2AS	10	0.85	–	NA
WMC216	1BL, 1DL	13	0.87	*Glu-D1*	Grain/spike, days to flowering, days to maturity, flag leaf area, peduncle length
WMC221	7DL	2	0.30	–	NA
WMC243	2BS, 2DL, 6AS	3	0.64	*Lr-16*	Peduncle length, grain yield
WMC245	2BL, 2DL	5	0.74	*Lr-16*	NA
WMC254	4B-	12	0.85	–	Flag leaf area, spikelets/spike, florets/spike
WMC256	6A, 6D	4	0.21	–	NA
WMC267	5A-	3	0.65	–	Days to maturity

*Chromosomal arm assignment was not possible in all the cases.
**NA, Not associated.

Primers, DNA isolation and PCR amplification

The primers used included 20 SSR primers (Table 1), two SAMPL primer combinations ($S_6 \times M_{CAG}$ and $S_7 \times M_{CAG}$) and eight AFLP primer combinations ($E_{ACC} \times M_{CTA}$; $E_{ACC} \times M_{CTG}$; $E_{AAC} \times M_{CTC}$; $E_{AAC} \times M_{CAT}$; $E_{ACG} \times M_{CTT}$; $E_{ACG} \times M_{CAG}$; $E_{AGG} \times M_{CAA}$; $E_{AGG} \times M_{CAC}$). More details of the above three types of primers are available in published literature[10–12].

The protocols for DNA extraction, PCR amplification, polyacrylamide gel electrophoresis, gel drying, and autoradiography are also described elsewhere[10].

Chromosome and arm assignment

Twenty SSR primers were tried for chromosome and arm assignment using 42 nullisomic-tetrasomic (NT) and 24 ditelosomic (DT) lines, procured from B. S. Gill, Kansas State University, USA.

Search for candidate gene

Candidate genes were searched using cMAP (a map alignment tool) available at GrainGenes website (http://wheat.pw.usda.gov), utilizing map information for SSR markers used in the present study. Map information for the above SSR markers was determined from composite wheat genetic map (Appels' unpublished data) and location of genes was determined from gene map of wheat[13]. SAMPL and AFLP markers were not used for the above exercise due to lack of map (genetic and physical) information.

Data analysis

The bands on all the gels/autoradiograms were scored in a 1–0 binary format and used in regression analysis with phenotypic data of 14 different traits recorded on 55 wheat genotypes.

Analysis involving estimations of polymorphic information content (PIC), marker index (MI) and analysis of variance (ANOVA) was conducted using the above 1–0 data.

The association analysis was conducted using both simple linear regression and multiple regression. For simple linear regression, data on individual phenotypic trait were regressed on whole 1–0 binary marker data for each individual marker using Excel programme. Multiple regression analysis was conducted using SPSS software.

Results

Phenotypic data

For each of the 14 phenotypic traits included in the present study, data on mean, range, mean squares and coefficients

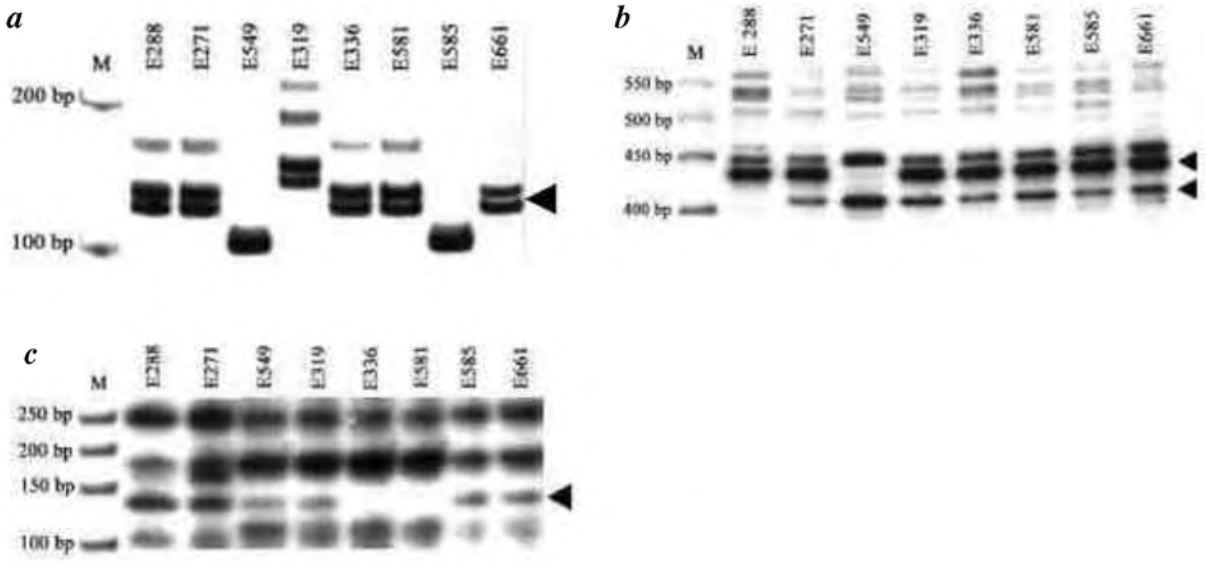

Figure 1. Polymorphism in a representative set of eight elite wheat genotypes detected using (*a*) SSR primer WMC24, M = 100 bp ladder; (*b*) SAMPL primer combination S7 × M_{CAG}, M = 50 bp ladder; and (*c*) AFLP primer combination E_{ACG} × M_{CAG}, M = 50 bp ladder. Arrowheads in each case indicate polymorphic fragments.

of phenotypic and genotypic variations are available with another study published elsewhere[10].

Information content of SSR, SAMPL and AFLP markers on 55 wheat genotypes

Detection of polymorphism: Twenty SSR primers gave 221 polymorphic bands, two SAMPL primers gave 43 polymorphic bands, and eight AFLP primers gave 255 polymorphic bands (Figure 1 *a–c*). For SSRs, the mean PIC value was 0.70, with a range from 0.21 (WMC256) to 0.90 (WMC24; Table 1); for SAMPL, the mean PIC value was 0.264 with a range from 0.035 to 0.499; and for AFLP, the mean PIC value for the different AFLP primer pairs was 0.21, with a range from 0.13 ($E_{AAC} \times M_{CAT}$) to 0.35 ($E_{ACG} \times M_{CAG}$). The MI value for SSR markers was 0.70, for SAMPL it was 9.61, and for AFLP it was 16.14.

Simple regression: Data on each of the 14 phenotypic traits were separately regressed on each of the polymorphic markers, including 221 SSR markers, 43 SAMPL markers and 255 AFLP markers. In each marker class, while there were markers, on each of which more than one trait regressed significantly, there were also individual traits which regressed significantly on more than one marker.

SSR markers: Analyses of variances for regressions showed significant regression of each of 13 of the 14 traits (excluding 1000-grain weight) on one or more of a total of 99 of the 221 polymorphic SSR bands (Table 2). The associated markers each explained a maximum of 8.12 (tiller numbers) to 27.95% (for harvest index) of the total available variation for individual traits (Table 2).

SAMPL markers: Regression analyses showed significant regression of each of the 11 traits (excluding spike length, biological yield and 1000-grain weight) on a total of 19 of the 43 polymorphic SAMPL bands (Table 2). While on one extreme, flag leaf area was associated with five SAMPL markers (Table 2), on the other extreme, days to maturity and grain yield were both associated with a solitary SAMPL marker (Table 2). The associated markers each explained up to 8.11 (for days to maturity) to 20.20% (for harvest index) of the phenotypic variation for 11 different traits (Table 2).

AFLP markers: Significant regression was observed for each of 13 of the 14 traits (excluding days to maturity) involving 133 of the 255 polymorphic AFLP bands (Table 2). As extreme examples, while each of five different traits regressed on the same AFLP marker (*Xccs*$E_{AAC}M_{CTC}$*268*), peduncle length regressed only on two AFLP markers, and a solitary trait, florets per spike regressed on a maximum number of 63 AFLP markers (Table 2). Associated markers explained up to 11.22 (for grain yield) to 29.38% (for florets per spike) of the total variation available for different individual traits (Table 2).

Multiple regressions: SSR markers: Significant regressions of each of the 14 traits were observed on a variable number (47–55) of different SSR markers. Surprisingly, from these 55 associated markers, individual subsets (each having 45 to 52 SSR markers) were able to explain

Table 2. Details of simple linear regression analysis using SSR, AFLP and SAMPL markers involving 14 different phenotypic traits

	SSR		SAMPL		AFLP		
Phenotypic trait	T	R^2 (%)	T	R^2 (%)	T	R^2 (%)	Total markers
Plant height	7	16.69	2	9.21	12	16.95	21
Days to flowering	30	21.30	3	12.64	13	18.84	46
Days to maturity	12	16.74	1	8.11	–	–	13
Tiller number	3	8.12	4	9.14	9	20.65	16
Flag leaf area	7	13.68	5	9.75	8	15.04	20
Peduncle length	5	15.63	2	9.40	2	12.43	9
Spike length	8	10.03	–	–	12	15.23	20
Spikelets/spike	17	25.08	2	8.15	17	25.05	36
Florets/spike	15	18.14	2	11.44	63	29.38	80
Grains/spike	15	11.93	2	9.29	55	24.89	72
Biological yield	12	13.33	–	–	11	13.58	23
Grain yield	19	16.94	1	10.40	5	11.22	25
Harvest index	27	27.95	3	20.20	4	11.40	34
1000-grain weight	–	–	–	–	12	16.22	12

T, Total number of markers showing significant association with the trait.
R^2, Maximum variation of a trait explained by a marker out of total significantly associated markers.

Table 3. Details of analyses of variances (ANOVA) involving multiple regressions for 14 traits using 221 SSR, 43 SAMPL, and 255 AFLP polymorphic bands (only those having significant regression are included)

		Mean square					
Trait	Source of variance	*df*	SSR	*df*	SAMPL	*df*	AFLP
Plant height	*x*	47	(291.670**)	16	(496.838**)	53	(258.944**)
	y	7	(2.218)	38	(151.965)	1	(0.001)
Days to flowering	*x*	47	(291.670**)	16	(496.833**)	53	(258.944**)
	y	7	(2.218)	38	(151.965)	1	(0.001)
Days to maturity	*x*	47	(291.670**)	16	(496.838**)	53	(258.944**)
	y	7	(2.218)	38	(151.965)	1	(0.001)
Tiller number	*x*	47	(291.670**)	16	(496.838**)	53	(258.944**)
	y	7	(2.218)	38	(151.965)	1	(0.001)
Flag leaf area	*x*	49	(7.077**)	16	(49.657**)	46	(27.339**)
	y	5	(0.014)	38	(12.271)	8	(0.399)
Peduncle length	*x*	52	(37.309**)	11	(16.853**)	51	(6.799**)
	y	2	(0.023)	43	(3.755)	3	(0.024)
Spike length	*x*	49	(68.661**)	9	(105.787**)	53	(36.606**)
	y	5	(0.165)	45	(21.956)	1	(0.003)
Spikelets/spike	*x*	51	(3.957*)	16	(159.713**)	49	(68.666**)
	y	3	(0.001)	38	(21.311)	5	(0.121)
Florets/spike	*x*	51	(17.713**)	5	(46.218**)	52	(17.373**)
	y	3	(0.011)	49	(13.721)	2	(0.001)
Grains/spike	*x*	51	(4.261*)	4	(7.532**)	53	(3.808**)
	y	3	(0.008)	50	(1.705)	1	(0.001)
Biological yield	*x*	51	(3.957*)	9	(12.174**)	45	(4.559*)
	y	3	(0.001)	45	(2.05)	9	(0.022)
Grain yield	*x*	50	(124.283**)	18	(242.890**)	53	(117.345**)
	y	4	(1.281)	36	(51.313)	1	(0.001)
Harvesting index	*x*	45	(71.068**)	1	(25927.583**)	53	(60.550**)
	y	9	(1.231)	53	(6724.543)	1	(0.001)
1000-grain weight	*x*	50	(7645.864**)	–	–	48	(7955.024**)
	y	4	(8.795)	–	–	6	(81.206)

*,**Significant at 5 and 1%, respectively; *x*, Regression; *y*, Residual; *df*, degree of freedom.

Table 4. Markers common in both simple linear and multiple regression analyses

Trait	Marker type	Marker designation
Plant height	SSR	*Xwmc24*$_{124}$
	SAMPL	–
	AFLP	*XccsE*$_{ACC}$*M*$_{CTG}$*148*, *XccsE*$_{ACC}$*M*$_{CTG}$*456*
Days to flowering	SSR	*Xwmc83*$_{81}$, *Xwmc149*$_{162}$, *Xwmc167*$_{163}$, *Xwmc169*$_{160}$, *Xwmc170*$_{207}$, *Xwmc216*$_{142}$
	SAMPL	*XccsS*$_{6}$*M*$_{CAG}$*250*, *XccsS*$_{6}$*M*$_{CAG}$*370*, *XccsS*$_{6}$*M*$_{CAG}$*420*
	AFLP	*XccsE*$_{ACC}$*M*$_{CTA}$*205*, *XccsE*$_{ACC}$*M*$_{CTA}$*350*, *XccsE*$_{ACC}$*M*$_{CTG}$*340*, *XccsE*$_{ACG}$*M*$_{CAG}$*230*
Days to maturity	SSR	*Xwmc170*$_{207}$, *Xwmc216*$_{142}$, *Xwmc267*$_{212}$
	SAMPL	*XccsS*$_{6}$*M*$_{CAG}$*420*
	AFLP	–
Tiller number	SSR	*Xwmc83*$_{154}$
	SAMPL	*XccsS*$_{6}$*M*$_{CAG}$*365*, *XccsS*$_{7}$*M*$_{CAG}$*230*
	AFLP	*XccsE*$_{ACC}$*M*$_{CTA}$*115*, *XccsE*$_{ACC}$*M*$_{CTA}$*390*, *XccsE*$_{ACG}$*M*$_{CAG}$*245*
Flag leaf area	SSR	*Xwmc25*$_{147}$, *Xwmc25*$_{162}$, *Xwmc216*$_{220}$, *Xwmc254*$_{220}$
	SAMPL	*XccsS*$_{6}$*M*$_{CAG}$*365*, *XccsS*$_{7}$*M*$_{CAG}$*135*, *XccsS*$_{7}$*M*$_{CAG}$*240*
	AFLP	*XccsE*$_{ACC}$*M*$_{CTA}$*350*
Peduncle length	SSR	*Xwmc216*$_{178}$, *Xwmc243*$_{187}$
	SAMPL	–
	AFLP	*XccsE*$_{ACC}$*M*$_{CTG}$*148*
Spike length	SSR	–
	SAMPL	*XccsS*$_{7}$*M*$_{CAG}$*240*
	AFLP	*XccsE*$_{ACC}$*M*$_{CTG}$*148*, *XccsE*$_{AGG}$ *M*$_{CAC}$*245*
Spikelets/spike	SSR	*Xwmc24*$_{116}$, *Xwmc24*$_{124}$, *Xwmc149*$_{225}$, *Xwmc167*$_{171}$, *Xwmc254*$_{156}$
	SAMPL	–
	AFLP	*XccsE*$_{AAC}$*M*$_{CTC}$*128*, *XccsE*$_{ACC}$*M*$_{CTA}$*290*, *XccsE*$_{ACC}$*M*$_{CTA}$*390*, *XccsE*$_{ACC}$*M*$_{CTG}$*456*, *XccsE*$_{ACG}$*M*$_{CAG}$*230*
Florets/spike	SSR	*Xwmc170*$_{191}$, *Xwmc254*$_{156}$
	SAMPL	*XccsS*$_{6}$*M*$_{CAG}$*370*
	AFLP	*XccsE*$_{AAC}$*M*$_{CTC}$*160*, *XccsE*$_{ACC}$*M*$_{CTA}$*295*, *XccsE*$_{ACC}$*M*$_{CTA}$*390*, *XccsE*$_{ACC}$*M*$_{CTG}$*456*, *XccsE*$_{ACG}$*M*$_{CAG}$*230*, *XccsE*$_{AGG}$*M*$_{CAA}$*425*, *XccsE*$_{AGG}$*M*$_{CAC}$*245*
Grain/spike	SSR	*Xwmc24*$_{124}$, *Xwmc167*$_{171}$
	SAMPL	*XccsS*$_{7}$*M*$_{CAG}$*135*
	AFLP	*XccsE*$_{AAC}$*M*$_{CTC}$*168*, *XccsE*$_{AAC}$*M*$_{CTC}$*386*, *XccsE*$_{ACC}$*M*$_{CTG}$*456*, *XccsE*$_{ACG}$*M*$_{CAG}$*160*, *XccsE*$_{ACG}$*M*$_{CAG}$*230*, *XccsE*$_{AGG}$*M*$_{CAA}$*425*
Biological yield	SSR	*Xwmc170*$_{186}$
	SAMPL	–
	AFLP	*XccsE*$_{ACC}$*M*$_{CTA}$*290*, *XccsE*$_{AGG}$*M*$_{CAA}$*200*, *XccsE*$_{AGG}$*M*$_{CAC}$*245*
Grain yield	SSR	*Xwmc83*$_{81}$, *Xwmc216*$_{178}$, *Xwmc243*$_{187}$
	SAMPL	–
	AFLP	*XccsE*$_{ACG}$*M*$_{CAG}$*245*
Harvest index	SSR	*Xwmc44*$_{291}$, *Xwmc83*$_{81}$, *Xwmc167*$_{171}$, *Xwmc170*$_{174}$
	SAMPL	*XccsS*$_{6}$*M*$_{CAG}$*250*
	AFLP	*XccsE*$_{ACC}$*M*$_{CTG}$*276*
1000-grain weight	SSR	–
	SAMPL	–
	AFLP	*XccsE*$_{AAC}$*M*$_{CTC}$*386*, *XccsE*$_{ACC}$*M*$_{CTA}$*115*

In SSR marker designation, the first three letters followed by 2–3 digits denote the name of SSR primer and subscripts of 2–3 digits in lower case denote size of the band in each case. In AFLP marker designation, ccs represents Ch. Charan Singh University, E_{xxx} and M_{xxx} are names of AFLP *Eco*RI and *Mse*I primers, followed by three digits indicating the size of AFLP bands.

97 to 100% of the total available variation for each of the 14 phenotypic traits examined.

SAMPL markers: Significant regression was observed for each of 13 of the 14 traits (excluding 1000-grain weight) on a total of 38 markers; individual traits regressed on 1 to 18 markers. Details of regression analysis are available in Table 3. Among these 13 traits, harvest index showed significant regression on only one SAMPL marker, while grain yield showed significant regression on 18 SAMPL

markers. The associated markers explained 5 (harvest index) to 66% (spikelets per spike) of the variation for individual traits.

AFLP markers: Significant association was observed for 54 of the 255 polymorphic markers with at least one of the 14 traits, the number of markers associated with individual traits ranging from 45 to 53 markers (Table 3). Only a solitary trait (biological yield) regressed on as many as 45 AFLP markers; eight other traits (plant height, days to maturity, days to flowering, tiller number, spike length, grains per spike, grain yield and harvest index) were associated each with as many as 53 AFLP markers. Each of the remaining five traits showed association with 46 to 52 AFLP markers. These associated markers together explained 98 to 100% variations of the 14 phenotypic traits examined.

Altogether, out of 519 DNA markers (SSR = 221, SAMPL = 43 and AFLP = 255), a total of 340 markers (SSR = 131, SAMPL = 43 and AFLP = 166) were identified to show association with at least one of the 14 agronomic traits with either of the two approaches; 51 DNA markers (SSR = 23, SAMPL = 7 and AFLP = 21) exhibited association, each with at least one of the 14 traits following both the above approaches. Markers that were found to be associated with an individual trait both in linear and multiple regressions were considered to be relatively more reliable (Table 4).

Chromosome/arm assignment of markers: A number of SSR markers could be assigned to individual chromosomes and to their arms using nullisomic–tetrasomic lines and the corresponding ditelocentrics. The results are included in Table 1.

Candidate genes: A search was made for genes lying close to the markers showing association with one or more traits. Six of the trait-associated markers had proximity to the following genes: *Tri A1*, *Lr-16* and *Glu-D1* (see Table 1 for details). Apparently, these genes may not be candidate genes for the associated traits, but may have either a pleiotropic effect or an indirect effect on the concerned trait through other related traits.

Discussion

Molecular markers linked with QTL/major genes for traits of interest are being routinely developed in several crops using materials derived from planned crosses such as F_2, RIL, DH populations, etc. Hopefully, some of these markers will be used for MAS in future wheat breeding programmes. However, non-availability of mapping populations and substantial time needed to develop such populations are sometimes major limitations in the identification of molecular markers for specific traits. Another limitation is the absence of tight linkage observed in these studies. To overcome these limitations, and as an alternative to planned populations, molecular markers for traits of interest have been identified through association studies conducted using germplasm collections[7]. Such association studies involving the use of germplasm collections for the identification of molecular markers have so far been only sparingly conducted in crop plants, including cereal species such as rice[14], barley[15], wheat[8,9] and ryegrass[16].

The present study involved a set of 55 exotic wheat genotypes, which constitute an important and diverse international elite germplasm of wheat, exhibiting moderate to high genetic variability for the 14 phenotypic traits examined during the present study[10]. It is obvious that for any specific trait, an international germplasm like the one used in the present study (relative to the Indian wheat germplasm having a narrow genetic base) would allow us to sample a much larger number of QTLs/genes and their corresponding alleles for identification of marker–trait associations.

Using both simple linear regression and multiple regression methods, a total of 340 molecular markers (SSR = 131; SAMPL = 43; AFLP = 166) were identified, each of which showed significant association with at least one of the 14 traits. Out of the above 340 markers, 51 were identified using both the regression methods (SSR = 23, SAMPL = 7 and AFLP = 21; Table 4). In both analyses, a maximum of 11 common markers were identified for days to flowering, followed by 10 markers each for spikelets per spike and florets per spike. For the remaining traits, the number of common markers varied from 2 to 9 (Table 4). It can be seen that each of a number of individual SSR, SAMPL and AFLP markers showed association with two to eight phenotypic traits. Such associations and/or prediction of 97–100% phenotypic variation by markers may be attributed to the presence of false positives and/or the use of structured populations in the present study. The AFLP/SAMPL markers used in the present study are not mapped; therefore, it is difficult to study population structure at this stage. Also, it is difficult to eliminate false positives with available methods. Therefore, markers identified during the present study need to be subjected to validation and/or functional analysis of respective traits, which is beyond the scope of the present study. However, we believe that at least some of the markers identified during the present study would be validated and used for MAS involving Indian germplasm of bread wheat.

1. Langridge, P., Lagudah, E. S., Holton, T. A., Appels, R., Sharp, P. J. and Chalmers, K. J., Trends in genetic and genome analyses in wheat: a review. *Aust. J. Agric. Res.*, 2001, **52**, 1043–1077.
2. Varshney, R. K., Balyan, H. S. and Langridge, P., Wheat. In *The Genome: Cereals and Millets* (ed. Kole, C.), Science Publishers, Inc., Enfield (NH), USA, 2005, pp. 121–219.
3. Gupta, P. K., Varshney, R. K., Sharma, P. C. and Ramesh, B., Molecular markers and their applications in wheat breeding. *Plant Breed.*, 1999, **118**, 369–390.
4. Darvasi, A., Weintreb, A., Minke, V., Weller, S. and Soller, M., Detecting marker QTL linkage and estimating QTL gene effect

and map location using a saturated genetic map. *Genetics*, 1993, **134**, 943–951.
5. Neale, D. B. and Savolainen, O., Association genetics of complex traits in conifers. *Trends Plant Sci.*, 2004, **9**, 325–330.
6. Botstein, D. and Risch, N. J., Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nature Genet.*, 2003, **33**, 228–237.
7. Gupta, P. K., Rustgi, S. and Kulwal, P. L., Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.*, 2005, **57**, 461–485.
8. Maccaferri, M., Sanguineti, M. C., Noli, E. and Tuberosa, R., Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breed.*, 2005, **15**, 271–289.
9. Breseghello, F. and Sorrells, M. E., Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*, 2005, DOI 10.1534/genetics.105.044586.
10. Roy, J. K., Lakshmikumaran, M. S., Balyan, H. S. and Gupta, P. K., AFLP-based genetic diversity and its comparison with diversity based on SSRs, SAMPL and phenotypic traits in bread wheat. *Biochem. Genet.*, 2004, **42**, 43–59.
11. Prasad, M., Varshney, R. K., Roy, J. K., Balyan, H. S. and Gupta, P. K., The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. *Theor. Appl. Genet.*, 2000, **100**, 584–592.
12. Roy, J. K., Balyan, H. S., Prasad, M. and Gupta, P. K., Use of SAMPL for a study of DNA polymorphism, genetic diversity and possible gene tagging in bread wheat. *Theor. Appl. Genet.*, 2002, **104**, 465–472.
13. Hart, G. E., Gale, M. D. and McIntosh, R. A., Linkage maps of *Tritium aestivum* (hexaploid wheat, $2n = 42$, genomes A, B, and D) and *T. tauschii* ($2n = 14$, genome D). In *Genetic Maps: Locus Maps of Complex Genomes* (ed. O'Brien, S. J.), Cold Spring Harbor Laboratory Press, 1993, pp. 6204–6219.
14. Virk, P. S., Ford-Lloyd, B. V., Jackson, M. T., Pooni, H. S., Clemeno, T. P. and Newbury, H. J., Predicting quantitative variation within rice germplasm using molecular markers. *Heredity*, 1996, **76**, 296–304.
15. Pakniyat, H. *et al.*, AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. *Genome*, 1997, **40**, 332–344.
16. Skøt, L. *et al.*, An association mapping approach to identify flowering time genes in natural populations of *Lolium perenne* (L.). *Mol. Breed.*, 2005, **15**, 233–245.

ACKNOWLEDGEMENTS. Financial assistance provided by the Department of Biotechnology, New Delhi to carry out this study is acknowledged. During the period of the study, P.K.G. was supported as a Senior Scientist by the Indian National Science Academy, New Delhi.

Received 16 May 2005; revised accepted 24 November 2005