Field Crops Research 123 (2011) 226–233

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Introgression of a major gene for high grain protein content in some Indian bread wheat cultivars

J. Kumar[a,c,1], V. Jaiswal[a,1], A. Kumar[a], N. Kumar[a], R.R. Mir[a,d], S. Kumar[a], R. Dhariwal[a], S. Tyagi[a], M. Khandelwal[e], K.V. Prabhu[b], R. Prasad[e], H.S. Balyan[a], P.K. Gupta[a,*]

[a] Molecular Biology Laboratory, Department of Genetics & Plant Breeding, Ch. Charan Singh University, Meerut 250004, India
[b] Division of Genetics, Indian Agriculture Research Institute, New Delhi 110012, India
[c] Division of Crop Improvement, Indian Institute of Pulses Research, Kanpur 208024, India
[d] International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Greater Hyderabad, India
[e] Indian Agricultural Statistics Research Institute, Pusa Campus, New Delhi 110012, India

ARTICLE INFO

Article history:
Received 3 January 2011
Received in revised form 14 May 2011
Accepted 14 May 2011

Keywords:
Bread wheat
Grain protein content
Leaf rust
MAS

ABSTRACT

In bread wheat, high grain protein content (GPC) determines nutritional value, processing properties and quality of the end-product. In view of this, marker-assisted selection (MAS) was performed for introgression of a major gene for high GPC (*Gpc-B1*) into 10 wheat genotypes. As a result, 124 BC_3F_5/F_6 progenies with *Gpc-B1* were developed and evaluated in multi-location field trials. Significant interaction of *Gpc-B1* with the recipient parent genotypes and the environment was noticed. However, a total of seven MAS-derived progenies with significantly higher GPC (14.83–17.85%) than their recipient parental genotypes and having no yield penalty were obtained. In these selected progenies, no significant negative correlation of grain yield with GPC (%) or protein yield was observed suggesting that GPC could be improved without yield penalty. This study thus suggested that MAS in combination with phenotypic selection is a useful strategy for development of wheat genotypes with high GPC associated with no loss in yield.

1. Introduction

In bread wheat, grain protein content (GPC) is an important grain quality trait, which determines nutritional value, processing properties, quality of the end products (bread and pasta) and market value of the grain. At 10% moisture content, the wheat grain is estimated to provide ~10 Mg (1 Mg = 10^6 g) of protein annually for human and livestock nutrition (see Brevis and Dubcovsky, 2010). However, despite proven value of higher GPC, only limited success has been achieved in breeding high GPC bread wheat genotypes using traditional methods. It has been observed that breeding efforts aimed at genetic improvement of grain yield resulted in lowering of GPC due to its negative association with grain yield (Simmonds, 1995; Brevis and Dubcovsky, 2010; references therein). Further, GPC is controlled by complex genetic system and is also influenced by the environment (Loffler and Busch, 1982; Simmonds, 1995; Lawlor, 2002). Therefore, genetic improvement of GPC without any yield penalty is still a challenge. It has also been argued that the improvement in GPC in modern wheat cultivars without associated yield penalties will require development of genotypes with higher N-use efficiency by increasing either the N-uptake or N-remobilization (see Brevis and Dubcovsky, 2010). However, in rare cases, it has been possible to develop some genotypes, which combine high yield with high level of GPC without the need to improve N-use efficiency (Cox et al., 1985).

In the past, a search for the genes/QTL for high GPC in wheat led to the discovery of a major QTL on chromosome arm 6BS explaining 66% of the variation in GPC in a population involving a tetraploid wheat [*Triticum turgidum* L. var. *dicoccoides* (Korn. In litt. in Schweinf.); accession FA15-3] with high GPC (Avivi, 1978; Joppa et al., 1997). The high GPC QTL, later designated as *Gpc-B1*, was introgressed into several hexaploid wheat cultivars, and a number of RFLP, SSR and CAPS markers closely linked with the gene were developed (Mesfin et al., 1999; Khan et al., 2000; Olmos et al., 2003; Distelfeld et al., 2004). Recently, construction of a complete physical map of a 250 kb region encompassing *Gpc-B1* allowed the development of an almost perfect marker (*Xuhw89*) that is tightly linked at a distance of 0.1 cM (Distelfeld et al., 2006). This was followed by positional cloning of *Gpc-B1*, thus facilitating the development of even "perfect" markers for this gene (Uauy et al., 2006).

Gpc-B1 has already been used for breeding wheat genotypes combining high GPC and high yield (see Brevis and Dubcovsky, 2010), although no such reports are available from India. Utilizing the available molecular markers for *Gpc-B1*, the present study was

* Corresponding author. Tel.: +91 1212768195; fax: +91 1212768195.
E-mail address: pkgupta36@gmail.com (P.K. Gupta).
[1] These authors contributed equally.

0378-4290/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2011.05.013

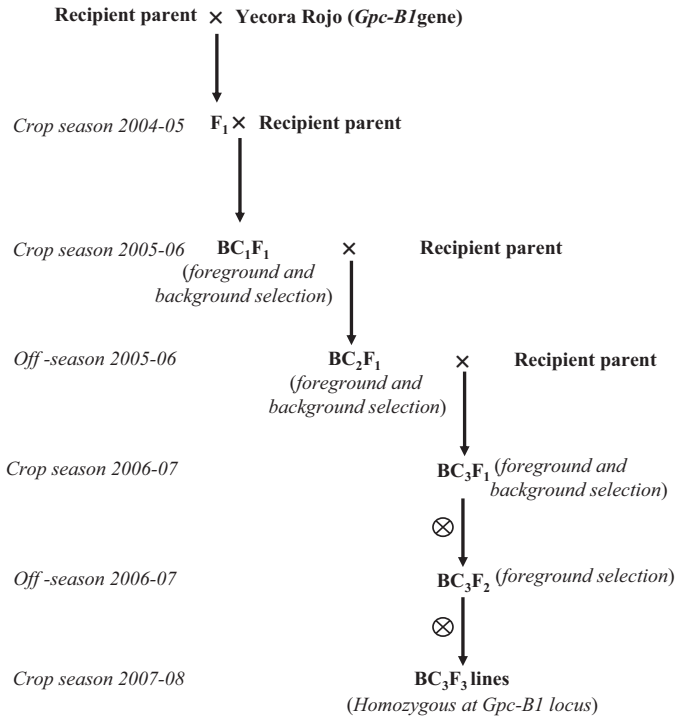

Fig. 1. Flow diagram showing steps involved in marker-assisted backcross breeding program.

aimed at developing high GPC lines through marker-assisted introgression of *Gpc-B1* in the backgrounds of 10 different Indian bread wheat genotypes, which had each low to moderate GPC. The impact of introgression of *Gpc-B1* on GPC and grain yield was examined.

2. Materials and methods

2.1. Materials

A total of 10 bread wheat genotypes were used as recipient parents during the present study (Table 1). A hexaploid wheat genotype Yecora Rojo containing *Gpc-B1* responsible for high GPC (kindly provided by Jorge Dubcovsky, University of California, Davis, USA) was used as the donor parent.

2.2. Marker assisted selection (MAS)

2.2.1. DNA isolation and the markers

DNA from parental genotypes and backcross progenies was isolated from one-month-old plants using a modified CTAB method (Saghai-Maroof et al., 1984). Following markers were used for foreground selection in different segregating backcross populations: (i) flanking markers *Xgwm193* (SSR) and *XNor-B2* (CAPS), which are 7.5 cM apart (Khan et al., 2000), were used in BC_1F_1 and BC_2F_1; (ii) *Xuhw89* (SSR), 0.1 cM away from *Gpc-B1* (Distelfeld et al., 2006), was used in BC_3F_1.

2.2.2. Marker-assisted breeding

The scheme followed for marker-assisted breeding is presented in Fig. 1. F_1 plants were confirmed for their heterozygosity for the markers flanking *Gpc-B1* and were backcrossed with their respective recipient genotypes. In each backcross, foreground selection was carried out using the markers listed above and plants that were heterozygous for the parental alleles were selected. These selected plants were subjected to background selection (see below) and plants (up to 5) showing highest similarity with the recipient parent genome were selected. In BC_3F_1, after foreground and background selections, individual plants heterozygous for *Xuhw89* were selected and selfed to provide BC_3F_2, where plants homozygous at marker locus *Xuhw89* were selected and selfed to obtain BC_3F_3 seed. The BC_3F_3 seed was used to advance the MAS-derived progenies through selfing to obtain BC_3F_5/F_6 (BC_3F_5 and BC_3F_6) progenies before these were evaluated in multilocation trials for their GPC (%) and grain yield and other related traits.

2.2.3. Foreground selection

For foreground selection, PCR amplification was carried out in a reaction mixture of 20 μl containing 10 mM Tris–HCl (pH8.8), 50 mM KCl, 200 μM dNTPs (MBI Fermentas), 0.75 U Taq DNA polymerase (GeneScript), 0.2 μM primers and 50 ng template DNA. PCR cycle consisted of an initial denaturation for 5 min at 95 °C, followed by 39 cycles each with 1 min at 94 °C, 1 min at annealing temperature (which differs for different primers), with a final extension of 7 min at 72 °C. The amplification products due to SSR (*Xgwm193*) and allele specific perfect markers (*Xuhw89*) were resolved on 10% PAGE following silver staining. In case of CAPS (*XNor-B2*) marker, the amplified products were digested with 10 U of *Bam*H1 restriction enzyme overnight before resolving the products on 0.8% agarose gels following Khan et al. (2000). The molecular data were scored manually.

The BC_3F_5/F_6 progenies in the backgrounds of PBW343 (*Lr*24) and HD2329 (*Lr*24 + *Lr*28) were also screened with SCAR markers ($SCS73_{719}$ for *Lr*24 and $SCS421_{570}$ for *Lr*28) to confirm the retention of the two leaf rust resistance genes in the derived progenies. The PCR conditions that were used for amplification are available elsewhere (Prabhu et al., 2004; Kumar et al., 2010). The resistance of the positive lines containing *Lr*24 or *Lr*24 + *Lr*28 was also confirmed following leaf rust resistance tests carried out at the seedling stage (see below).

2.2.4. Background selection

For rapid recovery of the genome of each of the recipient parents, the background selection in each of the three backcross generations (BC_1F_1, BC_2F_1, and BC_3F_1) was carried out using a total of 92 SSRs that were polymorphic between each pair of the donor and the recipient genotypes (ESM 1). These SSRs were distributed throughout the wheat genome in a reference map (Somers et al., 2004). Information regarding chromosome location, primer sequences and PCR conditions used to amplify SSR markers are available elsewhere (Somers et al., 2004). The PCR products were resolved on 10% PAGE following silver staining; molecular data was scored manually.

2.2.5. Estimation of the recovery of the recipient genome

The recovery of recipient parent genome 'G' in the derived progenies, identified following foreground selection in each backcross and subsequent selfed generations was estimated using the following formula:

$$G = \left[\frac{X + 1/2Y}{N}\right] \times 100$$

where X = number of markers showing homozygosity for recurrent parent allele; Y = number of markers showing heterozygous state for the parental alleles; N = total number of parental polymorphic markers screened.

2.3. Evaluation of BC_3F_5/F_6 progenies for GPC (%) and grain yield

The BC_3F_5/F_6 seed obtained by selfing each of the BC_3F_4/F_5 progenies derived through MAS along with the seed of respective recipient genotypes was used for conducting field trials at three dif-

Table 1
Details of the recipient bread wheat genotypes used (after Kundu et al., 2006).

Genotype	Year of release	Important features
RAJ3765	1996	Recommended for cultivation in north western and north eastern plain zones, medium sized amber grains, heat and drought tolerant, 12.50% protein content
K9107	1996	Recommended for cultivation in north eastern plain zones, bold and amber grains, 13.51% protein content, good cooking quality and market acceptability
PBW373	1996	Recommended for cultivation under late sown conditions in north western plain zone, bold amber grains, 12.95% protein content
PBW343	1996	Recommended for cultivation in north western plain zone, bold amber grains, 12% protein content
HD2687	1999	Recommended for cultivation in north western plain zone, medium sized amber grains, 11.81% protein content
HI977	1988	Recommended for cultivation in peninsular zone, medium sized amber grains, 12.59% protein content
PBW343 + *Lr*24 (three lines)	–	Lines produced through marker-assisted introgression of *Lr*24 in the background of cv. PBW343
HD2329 (*Lr*24 + *Lr*28)	–	Line produced through marker-assisted introgression of *Lr*24 + *Lr*28 in the background of cv. HD2329

ferent locations including Meerut, Ludhiana and Pantnagar during crop season 2009–2010. These three locations are situated in north western plain zone, the major wheat producing area of India. The trial was laid in a 12 × 12 simple lattice design at each of the three locations each with two replications. Each genotype was planted in a 3 m^2 plot with five rows each 3 m long, and with a row-to-row distance of 25 cm at a seed rate of 120 kg/ha. All recommended agronomic practices were followed. The data on each genotype in each replication were recorded on the following traits: (i) plant height (cm), (ii) tillers per m^2, (iii) grains per spike, (iv) 1000-grain weight (g), (v) plot yield (g), (vi) GPC (%) and (vii) protein yield. The data on plot yield (g) was converted into grain yield (t/ha) for further statistical analyses. The protein yield was also calculated in t/ha. The GPC (%) at 12% moisture content was estimated for each genotype in each replication using Infratech Grain Analyzer at Agharkar Research Institute, Pune.

2.4. Test for leaf rust resistance

For the evaluation of leaf rust resistance, the material was grown in growth chambers, under controlled environmental conditions, at the National Phytotron Facility, Indian Agricultural Research Institute, New Delhi, India. Ten-day-old (single-leaf stage) seedlings were inoculated with pathotype 77-5 (the most virulent and predominant pathotype of leaf rust in South East Asia) by spraying the inoculum suspended in water fortified with Tween-20 (0.75 μl/ml) at an average concentration of 20 urediospores/microscopic field (10×·10×). The inoculated seedlings were incubated for 36 h in humid glass chambers at a temperature of 23 ± 2 °C with a relative humidity of more than 95%. After incubation, plants were shifted to growth chambers with the same environmental conditions. Disease reaction was recorded 12 days after inoculation following Stakman et al. (1962).

2.5. Statistical analysis

The analysis of variance (ANOVA) was conducted for seven different traits including GPC using data for all the three locations and using the following as sources of variation: location, replication, block, genotype and genotype × location interaction. Arc-sine transformation was used to transform data on GPC (%) for the purpose of ANOVA. The ANOVA also included study of the contrast between lines carrying *Gpc-B1* gene and those lacking it. The background effect involving 10 recipient genotypes was also examined.

Significance of differences between means was tested using Tukey's test. In MAS-derived progenies, correlations between grain yield and GPC (%) and between grain yield and protein yield were worked out using data for individual locations and also data pooled over locations. In order to examine further the relationships between grain yield and GPC (%), two scatter diagrams for grain yield vs. GPC (%) were also prepared, one involving 124 MAS-derived progenies, and the other involving seven desirable selected progenies. Similar scatter plots for grain yield vs. protein yield were also prepared. Statistical analyses were conducted using the software available with PROC GLM in SAS (SAS 1996) and Microsoft Excel.

3. Results

The breeding scheme followed in the present study is presented in Fig. 1. On the basis of marker-assisted foreground and background selections, 124 progenies carrying the gene *Gpc-B1* were selected. The results of ANOVA for the three locations involving 10 recipient parents, a donor genotype and 124 MAS-derived BC_3F_5/F_6 progenies [29 (BC_3F_5) + 95 (BC_3F_6)] are presented in Table 2. The mean squares for genotypes were partitioned into the following two sources of variation, (i) parents and derived progenies, and (ii) fillers, although the mean squares due to fillers are not included in Table 2. It may also be seen that the mean squares due to locations, blocks, genotypes and parents/derived progenies were significant for all the seven traits including GPC (%) and protein yield with some exceptions. For the contrast, with and without *Gpc-B1*, the mean squares for GPC (%) were significant, but those for grain yield were not significant, suggesting that protein content in derived progenies with *Gpc-B1* was higher than in parents lacking *Gpc-B1*, and further suggesting that there is no yield penalty due to increase in protein content. Background effects due to the 10 recipient genotypes were also significant suggesting that the genetic background had some effect on the expression of *Gpc-B1*. However, mean squares for protein yield were not significant in contrast (with and without *Gpc-B1*) and for genetic background, suggesting that the presence of *Gpc-B1* and the associated genetic background did not affect protein yield. A comparison of mean values for grain yield, GPC (%) and protein yield between the recipient parents and the MAS-derived progenies is presented in Table 3.

3.1. Progenies with high GPC with no yield penalty

Means for GPC (%) and yield in the above 124 MAS-derived progenies carrying *Gpc-B1*, were also compared with those of the recipient parents. There were 71 progenies, which exhibited high

Table 2
Analysis of variance for grain yield, yield component traits, GPC (%) and protein yield of parental lines and BC_3F_5/F_6 MAS-derived progenies based on data of three locations.

Source	DF	Mean square						
		Plant height (cm)	Tiller per m^2	Grains per spike	1000-grain weight (g)	Grain yield (t/ha)	GPC (%) transformed	Protein yield (t/ha)
Location	2 (1)	14,546.00**	6,460,000**	13,165.00**	2807.60**	1829.30**	1.36**	39.85**
Replication	1 (1)	66.60	397,698**	57.51	77.49*	2.70	0.02*	0.08
Block (Rep.)	22 (22)	61.88**	62,518**	388.31	32.59**	4.74**	0.03**	0.07**
Genotype	143 (143)	162.24**	53,569**	418.82	51.85**	1.79	0.02**	0.05**
Parents and derived progenies	134 (134)	139.28**	45,351**	439.59	50.60**	1.77	0.02**	0.04**
Location × genotype	286 (143)	49.93**	43,982**	422.95	29.28**	1.58	0.01**	0.04**
With *Gpc-B1* vs. without *Gpc-B1*	1 (1)	0.28	68,700	20.48	284.58**	0.03	0.27**	0.03
Background	9 (9)	165.82**	37,926	52.03	59.89**	1.63	0.01**	0.04
Error	409 (265)	22.50	31,821	372.56	17.28	1.51	0.005	0.02

The values in parentheses indicate d.f. for plant height measured at two locations. *Note*: The values of mean squares due to fillers are not given.
* Significant at 5% level.
** Significant at 1% level.

Table 3
Comparisons of mean values of grain yield, GPC (%), and protein yield in recipient parental genotypes and their corresponding MAS-derived progenies based on pooled data of all three locations.

Recipient parent	Grain yield (t/ha)		GPC (%)		Protein yield (t/ha)	
	Parental lines	Derived progenies	Parental lines	Derived progenies	Parental lines	Derived progenies
Raj3765	7.21	6.04	12.75	14.46	0.92	0.87
K9107	5.77	5.27	14.10	15.83	0.81	0.84
PBW373	6.32	6.06	13.50	14.56	0.85	0.88
PBW343	6.33	5.82	13.72	14.76	0.87	0.86
HD2687	6.43	4.95	14.63	15.24	0.94	0.75
HI977	6.43	6.23	13.83	14.96	0.89	0.93
PBW343 (*Lr*24)	6.87	6.19	13.85	14.55	0.95	0.90
PBW343 (*Lr*24)	4.46	6.23	13.98	14.58	0.62	0.91
PBW343 (*Lr*24)	6.91	6.87	13.40	14.07	0.93	0.97
HD2329 (*Lr*24 + *Lr*28)	6.10	5.12	13.98	15.32	0.85	0.78

GPC (%) at all the three locations with no yield penalty, although improvement in GPC (%) was not statistically significant. Only three progenies one at each location showed significantly higher GPC (%) without any yield penalty relative to their respective recipient parental genotypes, although similar significant change in protein yield was not observed in these three selected progenies (Table 4).

A perusal of pooled data from three locations, however, showed that five progenies involving three of the 10 recipient parents [two each belonging to genotypes HD2329 (*Lr*24 + *Lr*28) and Raj3765 and one belonging to HI977] had significantly higher GPC (%) with no significant reduction in yield (Table 5). The protein yield, plant height and the three yield component traits (tillers per m^2, grains per spike and 1000-grain weight) of the above five MAS-derived progenies also did not differ significantly from the corresponding values for recipient parents (Table 5).

One of the above five MAS derived progenies, also had significantly higher GPC at one of the three locations, so that altogether there were seven progenies, which either had higher GPC at one of the three locations or exhibited higher GPC in pooled data.

3.2. Recovery of the genome of recipient parent

Following background selection using 92 SSR markers that were distributed over all the 21 chromosomes, the recovery of the genome of recipient parent in the 124 MAS-derived progenies varied from 60.93% to 98.40%. However, in the seven desirable selected progenies showing high GPC without yield penalty, the recovery varied from 72.0% to 95.71%.

3.3. Correlations of grain yield with GPC (%) and protein yield

When data for 124 MAS-derived progenies was used separately, both for individual locations and the pooled data, grain yield had significant negative correlation with GPC (%) and significant positive correlation with protein yield (Fig. 2a and b). However, for the

Table 4
Mean values of plant height, yield component traits, grain yield, GPC (%), protein yield and the per cent recovery of the recipient parent genome (MAS-derived progenies) of the parent genotypes and the three MAS-derived progenies with significantly higher GPC (%) and no yield penalty based on data of individual locations.

Parent genotype/progeny number	Location	Plant height (cm)	Tillers per m^2	Grains per spike	1000-grain weight (g)	Grain yield (t/ha)	GPC (%)	Protein yield (t/ha)	Per cent recovery of the recipient parent genome
Raj3765[a]	Ludhiana	85.25	672.60	48.70	36.58	4.80	13.47	0.65	–
Raj3765-762	Ludhiana	88.10	713.60	53.60	27.69	4.50	17.46*	0.79	72.00
PBW343 (*Lr*24)[a]	Meerut	85.10	713.60	51.30	38.30	5.40	13.63	0.74	–
PBW343 (*Lr*24)-603	Meerut	100.30	1017.10	41.30	40.61	7.30	17.12*	1.39	95.71
HD2329 (*Lr*24 + *Lr*28)[a]	Pantnagar	–	1008.90	38.10	31.33	5.07	13.93	0.70	–
HD2329 (*Lr*24 + *Lr*28)-396	Pantnagar	–	1222.10	39.20	20.97	3.77	17.85*	0.67	87.03

[a] Parent genotype; in column 1, each parental genotype is followed by derived lines in row below.
* Significant at 5% level.

Table 5
Mean values of plant height, yield component traits, grain yield, GPC (%), protein yield and the per cent recovery of the recipient parent genome (MAS-derived progenies) of the parent genotypes and the five MAS-derived progenies with significantly higher GPC (%) and no yield penalty based on data of three locations.

Parent genotype/progeny number	Plant height (cm)	Tillers per m^2	Grains per spike	1000-grain weight (g)	Grain yield (t/ha)	GPC (%)	Protein yield (t/ha)	Per cent recovery of the recipient parent genome
Raj3765[a]	86.60	870.23	53.70	36.52	7.21	12.75	0.92	–
Raj3765-418	88.50	916.42	52.30	31.93	5.61	14.83*	0.82	73.08
Raj3765-762	84.25	1033.50	51.27	30.06	5.67	15.44*	0.80	72.00
HI977[a]	87.43	837.50	48.97	36.57	6.43	13.83	0.89	–
HI977-478	88.05	965.44	48.40	29.70	4.74	16.15*	0.76	87.03
HD2329 (*Lr*24 + *Lr*28)[a]	70.48	801.79	42.00	33.36	6.10	13.98	0.85	–
HD2329 (*Lr*24 + *Lr*28)-342	69.16	764.36	43.63	31.45	5.32	16.18*	0.84	88.89
HD2329 (*Lr*24 + *Lr*28)-367	72.00	743.70	43.20	27.85	4.95	15.79*	0.62	90.77

[a] Parent genotype; in column 1, each parental genotype is followed by derived lines in row below.
* Significant at 5% level.

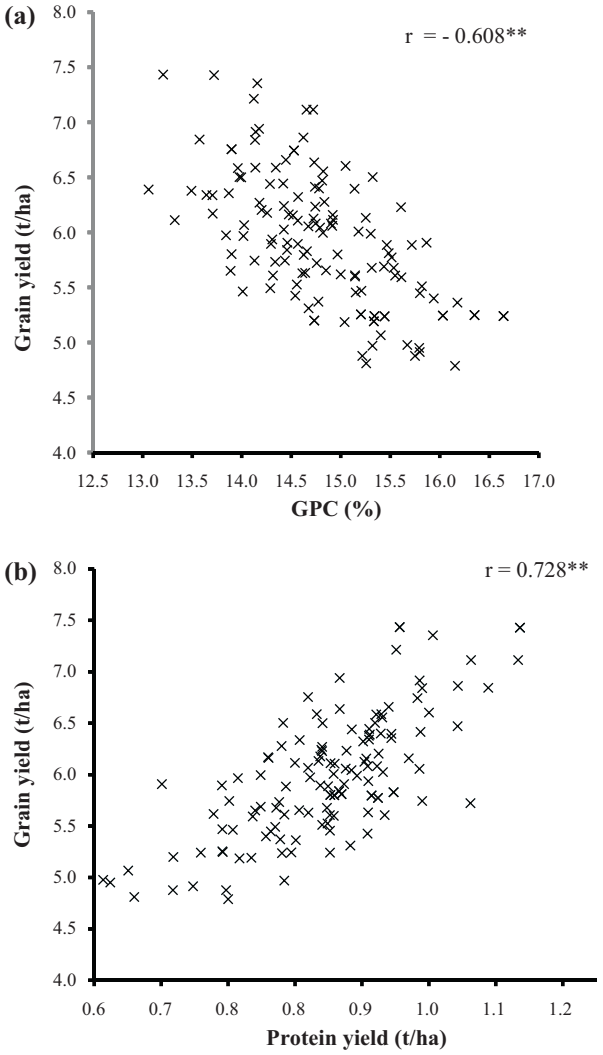

Fig. 2. Scatter plots of 124 MAS-derived progenies using data pooled over three locations. (a) Grain yield vs. GPC (%) and (b) grain yield vs. protein yield.

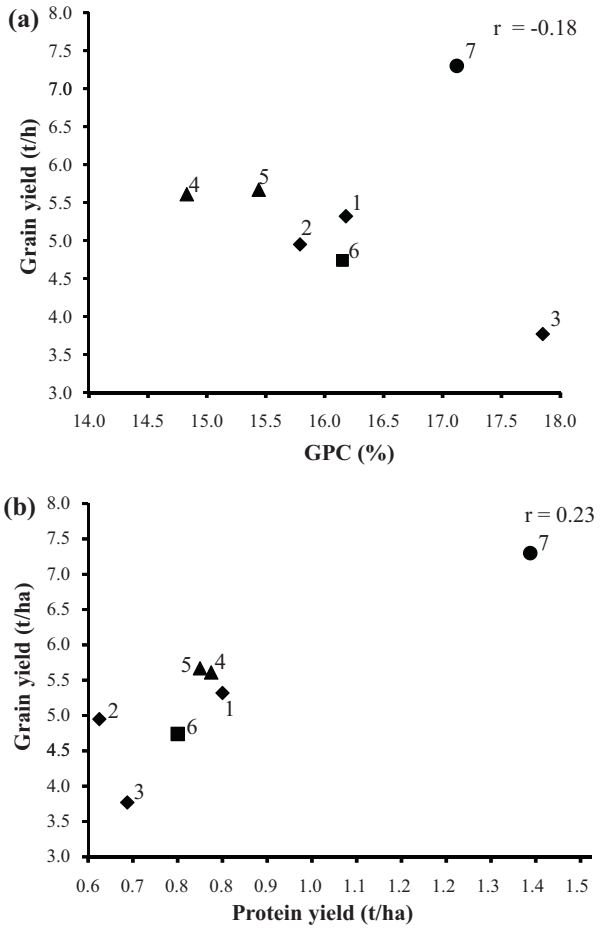

Fig. 3. Scatter plots of seven selected high GPC MAS-derived progenies based on pooled data of the three locations (progenies: 1, 2, 4, 5 and 6) and data of individual locations (progenies: 3 and 7). (a) Grain yield vs. GPC (%) and (b) grain yield vs. protein yield. 1, HD2329 (*Lr*24 + *Lr*28)-342; 2, HD2329 (*Lr*24 + *Lr*28)-367; 3, HD2329 (*Lr*24 + *Lr*28)-396; 4, Raj3765-418; 5, Raj3765-762; 6, HI977-478; 7, PBW343 (*Lr*24)-603.

selected seven progenies, no significant correlations were observed ($r = -0.18$ for GPC (%); $r = 0.23$ for protein yield). The scatter plots of grain yield vs. GPC (%) and grain yield vs. protein yield for the above selected seven MAS-derived progenies showed dispersed distribution confirming no correlation (Fig. 3a and b).

3.4. Pyramiding of Gpc-B1 and leaf rust resistance genes (Lr24 or Lr24 + Lr28)

Efforts were also made to pyramid *Gpc-B1* on to the leaf rust resistance genes (*Lr*; *Lr*24 + *Lr*28) that were introgressed earlier into to two recipients (PBW343 and HD2329). Four derived lines, three belonging to PBW343 and carrying *Lr24*, and one belonging

to HD2329 carrying *Lr*24 + *Lr*28 were used as recipients. Following introgression of *Gpc-B1* involving MAS, sixty (60) of the 124 BC_3F_5/F_6 progenies (with *Gpc-B1* gene) involved the above two recipients [PBW343 (*Lr*24); HD2329 (*Lr*24 + *Lr*28)] and were therefore also tested for the leaf rust resistance genes using SCAR markers. The analysis showed that all the above progenies carried corresponding *Lr* gene(s). These progenies also exhibited hypersensitive reaction to leaf rust confirming successful pyramiding of *Gpc-B1* in combination with leaf rust resistance genes.

4. Discussion

In the past, marker-assisted selection (MAS) has been successfully implemented to introgress and pyramid major genes/QTL for different traits in wheat (Gao et al., 2005; Miedaner et al., 2009; Gupta et al., 2008, 2010; Barloy et al., 2007; Kumar et al., 2010). In the present study, we successfully introgressed high GPC gene *Gpc-B1* through MAS into 10 Indian bread wheat genotypes for the first time and also developed progenies having high GPC without any yield penalty. In the past also, transfer of *Gpc-B1* was achieved for improvement of GPC without any yield penalty (Kade et al., 2005; Brevis and Dubcovsky, 2010). Development of two commercial bread wheat cultivars, namely 'Lassik' (University of California, Davis, USA) and 'Farnum' (Washington State University, Pullman, USA) and a durum wheat cultivar 'Westmore' (Arizona Plant Breeders, AZ, USA) carrying the gene *Gpc-B1* was also reported (see Brevis and Dubcovsky, 2010; http://variety.wsu.edu/extensionpubs/Farnum_trifold.pdf,http://uvdavis.edu/files2/57360.pdf). The variety 'Lassik', in particular, showed highly significant increase in GPC without any yield penalty relative to the recipient parent 'Anza'. As in the above studies, the results of the present study involving improvement of GPC without yield penalty in four of the ten Indian bread wheat genotypes confirmed that in wheat breeding programs, GPC can be improved without loss in grain yield.

4.1. Performance of MAS-derived progenies: GPC and grain yield

Variation in the magnitude of both GPC (%) and protein yield was noted among the progenies derived from a common recipient parent as well as among the progenies involving different recipients. Variation was also observed when the same progeny was evaluated at three different locations. This was substantiated by significant differences among genotypes and locations as well as significant genotype-by-location interactions for GPC as revealed by ANOVA. Background mean squares involving 10 recipient genotypes were also significant for GPC (%) (Table 2). This suggested significant interaction of *Gpc-B1* with the recipient genotype and also with the environment. Significant background effects also suggest that interactions with different recipient genotypes differ. Such significant ($P < 0.05$) interactions were also reported by Brevis and Dubcovsky (2010), while evaluating the NILs for *Gpc-B1* in the backgrounds of six hexaploid (Anza, Yecora Rojo, Attila, RS15, UC1037 and UC1041) and three tetraploid (Kofa, Kronos and UC1113). Being a quantitative trait, the expression of *Gpc-B1* as a whole is also influenced by epistatic and environmental interactions (Kuspira and Unrau, 1957; Law et al., 1978; Morris et al., 1978; Stein et al., 1992; Snape et al., 1995; Kulwal et al., 2005). As a result, a large effect of *Gpc-B1*, can be somewhat modified depending on the genetic background into which the gene is introgressed (Davies et al., 2006).

A significant negative correlation was observed between the grain yield and GPC (%) when data for all the 124 MAS-derived lines, recorded at individual locations was used. The situation did not change when data pooled across locations was used. However, similar examination involving grain yield and protein yield showed significant positive correlation between the two traits. Apparently, the increase in GPC (%) in MAS-derived progenies was not always associated with proportional decline in grain yield in all the progenies, thus making it possible to identify some progenies which will have high GPC (%) without having any adverse effect on protein yield and grain yield. It may be noted that when correlations of grain yield with GPC (%) and protein yield were examined using seven progenies that were selected for significantly higher GPC, no significant correlations were noticed. Therefore, we feel that it is important to exercise phenotypic selection for GPC and grain yield in lines carrying *Gpc-B1* to identify progenies with significantly higher GPC (%) with no yield penalty.

Based on the analysis of pooled data of three locations, we were able to recover five progenies (in three genetic backgrounds) having significantly higher GPC (%) with no yield penalty, although on the basis of data from individual locations, only three progenies, one at each location had this attribute. However, when pooled data and data from individual locations were compared, there was one progeny (Raj3765-762), which exhibited higher GPC (%) with no yield penalty, not only in the pooled data, but also in the data from one of the three locations.

As described in the results earlier, there were 71 progenies, which exhibited high GPC (%) at all the three locations with no yield penalty, although improvement in GPC (%) was marginal (increment 0.14–9.81%) and not statistically significant. Together the results of the present study indicated the role of genotype-by-environment interaction in determining GPC (%) and grain yield in wheat and the role of phenotypic selection in identification of progenies combining high GPC with high grain yield. The GPC (%) of the seven improved progenies varied from 14.83% to 17.85% representing an increment of 12.93% to 29.62% over the GPC of their respective recipient genotypes. The scatter plots of the values of grain yield vs. GPC (%) and grain yield vs. protein yield also suggested that it is possible to combine high GPC due to the *Gpc-B1* with high yield. This is consistent with earlier reports that *Gpc-B1* has limited negative impact, if any, on wheat yield (Kade et al., 2005; Brevis and Dubcovsky, 2010). It may be speculated that these progenies with high GPC and no yield penalty may have an efficient nitrogen uptake and/or nitrogen re-mobilization from leaf and stem tissues contributing to grain development leading to breakage of the known negative correlation between grain yield and GPC.

The grain yield of the MAS-derived progenies was in the range of 5.32–6.10 t/ha, which is within the range of the wheat yields recovered in the experimental fields in India. The mean values of the plant height and yield contributing traits of the above seven MAS-derived progenies also did not differ from their respective recipient genotypes and together contributed to observed comparable grain yield in the MAS-derived progenies.

4.2. Gene pyramiding

The molecular marker-assisted selection (MAS) during the present study was also successful in pyramiding *Gpc-B1* over important leaf rust resistance gene(s) *Lr24* or *Lr*24 + *Lr*28 in a number of progenies in the backgrounds of two important wheat cvs. HD2329 and PBW343. Four such progenies [HD2329 (*Lr*24 + *Lr*28)-342, HD2329 (*Lr*24 + *Lr*28)-367, HD2329 (*Lr*24 + *Lr*28)-396 and PBW343 (*Lr*24)-603] had higher GPC with no yield penalty either in the pooled data or at individual locations. We earlier demonstrated successful pyramiding of two leaf rust resistance genes *Lr*24 and *Lr*28 and a QTL for pre-harvest sprouting tolerance in wheat (Kumar et al., 2010). In the past, similar pyramiding of genes was reported in wheat for leaf rust resistance genes *Lr*13, *Lr*34 and *Lr*37 (Kloppers and Pretorius, 1997) and powdery mildew resistance genes *Pm*3, *Pm*4a and *Pm*21 (Liu et al., 2000).

4.3. Recovery of recipient parent genome

It may be recalled that the recovery of the genome of recipient parent in the seven selected lines was not as high as one would expect after three backcross generations. This is not surprising in view of the limited population size that was used as a trade-off due to limited resources, and the 10 backcross populations that we handled simultaneously. Further, each of the seven progenies having high GPC without any yield penalty had 72.00–95.71% of the genome of recipient parent suggesting that full restoration of the recipient genotype may not always be necessary, and that a restricted backcross breeding program may be followed for the selection of the superior genotypes.

5. Conclusions

The introgression of the gene *Gpc-B1* through marker-assisted backcrossing in combination with phenotypic selection is a useful strategy for developing wheat genotypes with high GPC (%) without adverse effect on grain yield. Also, markers linked to other economically important traits such as leaf rust resistance could help in pyramiding genes for more than one trait following rapid multi-trait selection through MAS in wheat.

Acknowledgements

Thanks are due to DBT, New Delhi for providing financial support to carry out this study, to GBPAU&T, Pantnagar and PAU, Ludihana for providing field facilities and to ARI, Pune for help in estimation of grain protein content. Financial support received from DST (FIST-program), New Delhi and UGC (SAP-DRS Program), New Delhi is duly acknowledged. CSIR, New Delhi awarded SRFs to RRM and AK. PKG was awarded the position of INSA Honorary Scientist and NASI-Senior Scientist.

Appendix A.

ESM 1. List of 92 SSRs used for background selection during MAS in wheat.

1, barc111	2, barc119	3, barc124	4, barc125	5, barc126	6, barc134	7, barc146	8, barc172
9, barc183	10, barc204	11, barc228	12, barc24	13, barc243	14, barc98	15, cfa2129	16, cfa2193
17, gdm113	18, gdm126	19, gdm136	20, gdm145	21, gdm153	22, gdm63	23, gdm72	24, gdm88
25, gwm113	26, gwm149	27, gwm179	28, gwm191	29, gwm193	30, gwm251	31, gwm261	32, gwm274
33, gwm30	34, gwm301	35, gwm333	36, gwm341	37, gwm382	38, gwm448	39, gwm473	40, gwm513
41, gwm540	42, gwm550	43, gwm608	44, gwm636	45, gwm654	46, wmc150	47, wmc160	48, wmc169
49, wmc17	50, wmc175	51, wmc177	52, wmc179	53, wmc181	54, wmc201	55, wmc219	56, wmc235
57, wmc241	58, wmc273	59, wmc291	60, wmc323	61, wmc331	62, wmc335	63, wmc336	64, wmc397
65, wmc405	66, wmc413	67, wmc418	68, wmc419	69, wmc420	70, wmc438	71, wmc475	72, wmc48
73, wmc486	74, wmc487	75, wmc488	76, wmc491	77, wmc508	78, wmc517	79, wmc525	80, wmc532
81, wmc533	82, wmc59	83, wmc593	84, wmc608	85, wmc627	86, wmc640	87, wmc651	88, wmc664
89, wmc667	90, wmc773	91, wmc797	92, wmc825				

References

Avivi, L., 1978. High grain protein content in wild tetraploid wheat Korn. In: Ramanujam, S. (Ed.), 5th Wheat Genetics Symposium. Indian Soc Genet Plant Breed, New Delhi, pp. 372–380.

Barloy, D., Lemoine, J., Abelard, P., Tanguy, A.M., Rivoal, R., Jahier, J., 2007. Marker assisted pyramiding of two cereal cyst nematode resistance genes from *Aegilops variabilis* in wheat. Mol. Breed. 20, 31–40.

Brevis, J.C., Dubcovsky, J., 2010. Effect of chromosome region including the *Gpc-B1* locus on wheat protein and protein yield. Crop Sci. 50, 93–104.

Cox, M.C., Qualset, C.O., Rains, D.W., 1985. Genetic variation for nitrogen assimilation and translocation in wheat. III. Nitrogen translocation in relation to grain yield and protein. Crop Sci. 26, 737–740.

Davies, J., Berzonsky, W.A., Leach, G.D., 2006. A comparison of marker-assisted and phenotypic selection for high grain protein content in spring wheat. Euphytica 152, 117–134.

Distelfeld, A., Uauy, C., Fahima, T., Dubcovsky, J., 2006. Physical map of the wheat high-grain protein content gene *Gpc-B1* and development of a high-throughput molecular marker. New Phytol. 169, 753–763.

Distelfeld, A., Uauy, C., Olmos, S., Schlatter, A.R., Dubcovsky, J., Fahima, T., 2004. Microcolinearity between a 2-cM region encompassing the grain protein content locus *GPC-6B1* on wheat chromosome 6B and a 350-kb region on rice chromosome 2. Funct. Integr. Genomics 4, 59–66.

Gao, A.L., He, H.G., Chen, Q.Z., 2005. Pyramiding wheat powdery mildew resistance genes *Pm2*, *Pm4a* and *Pm21* by molecular marker-assisted selection. Acta Agron. Sin. 31, 1400–1405.

Gupta, P.K., Balyan, H.S., Kumar, J., Kulwal, P.K., Kumar, N., Mir, R.R., Kumar, A., Prabhu, K.V., 2008. QTL analysis and marker assisted selection for improvement in grain protein content and pre-harvest sprouting tolerance in bread wheat. In: 11th Wheat Genetics Symposium (IWGS) , Brisbane Australia, August 24–29, pp. 1–3.

Gupta, P.K., Langridge, P., Mir, R.R., 2010. Marker-assisted wheat breeding: present status and future possibilities. Mol. Breed. 26, 145–161.

Joppa, L.R., Du, C., Hart, G.E., Hareland, G.A., 1997. Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosomal lines. Crop Sci. 37, 1586–1589.

Kade, M., Barneix, A.J., Olmos, S., Dubcovsky, J., 2005. Nitrogen uptake and remobilization in tetraploid 'Langdon' durum wheat and a recombinant substitution line with the high grain protein gene *Gpc-B1*. Plant Breed. 124, 343–349.

Khan, I.A., Procunier, J.D., Humphreys, D.G., Tranquilli, G., Schlatter, A.R., Marcucci-Poltri, S., Frohberg, R., Dubcovsky, J., 2000. Development of PCR-based markers for high grain protein content gene from *Triticum turgidum* ssp. *dicoccoides* transferred to bread wheat. Crop Sci. 40, 518–524.

Kloppers, F.J., Pretorius, Z.A., 1997. Effects of combinations amongst genes *Lr13*, *Lr34* and *Lr37* on components of resistance in wheat to leaf rust. Plant Pathol. 46, 737–750.

Kulwal, P., Kumar, N., Kumar, A., Gupta, R.K., Balyan, H.S., Gupta, P.K., 2005. Gene networks in hexaploid wheat: interacting quantitative trait loci for grain protein content. Funct. Integr. Genomics 5, 254–259.

Kumar, J., Mir, R.R., Kumar, N., Kumar, A., Mohan, A., Prabhu, K.V., Balyan, H.S., Gupta, P.K., 2010. Marker-assisted selection for pre-harvest sprouting tolerance and leaf rust resistance in bread wheat. Plant Breed..

Kundu, S., Shoran, J., Mishra, B., Gupta, P.K., 2006. Indian Wheat Varieties at a Glance, first print. Directorate of Wheat Research, Karnal-132001, India. Research Bulletin No. 21.

Kuspira, J., Unrau, J., 1957. Genetic analysis of certain characters in common wheat using whole chromosome substitution lines. Can. J. Plant Sci. 37, 300–326.

Law, C.N., Young, C.F., Brown, J.W.S., Snape, J.W., Worland, A.J., 1978. The study of grain protein control in wheat using whole-chromosome substitution lines. In: IAEA (Ed.), Seed Protein Improvement by Nuclear Techniques. IAEA, Vienna, pp. 483–502.

Lawlor, D.W., 2002. Carbon and nitrogen assimilation in relation to yield: mechanisms are the key to understanding production systems. J. Exp. Bot. 53, 773–787.

Liu, J., Liu, D., Tao, W., Li, W., Wang, S., Chen, P., Cheng, S., Gao, D., 2000. Molecular marker-facilitated pyramiding of different genes for powdery mildew resistance in wheat. Plant Breed. 119, 21–24.

Loffler, C.M., Busch, R.H., 1982. Selection for grain protein, grain yield, and nitrogen partitioning efficiency in hard red spring wheat. Crop Sci. 22, 591–595.

Mesfin, A., Frohberg, R.C., Anderson, J.A., 1999. RFLP markers associated with high grain protein from *Triticum turgidum* L. var. *dicoccoides* introgressed into hard red spring wheat. Crop Sci. 39, 508–513.

Miedaner, T., Wilde, F., Korzun, V., Ebmeyer, E., Schmolke, M., Hart, L., Schon, C.C., 2009. Marker selection for *Fusarium* head blight resistance based on quantitative trait loci (QTL) from two European sources compared to phenotypic selection in winter wheat. Euphytica 166, 219–227.

Morris, R., Mattern, P.J., Schmidt, J.W., Johnson, V.A., 1978. Studies on protein, lysine and leaf rust reactions in the wheat cultivar 'Atlas 66' using chromosome substitutions. In: Ramanujam, S. (Ed.), Proc. 5th Int. Wheat Genet Symp. Indian Soc Genet Plant Breed, IARI, New Delhi, pp. 447–454.

Olmos, S., Distelfeld, A., Chicaiza, O., Schlatter, A.R., Fahima, T., Echenique, V., Dubcovsky, J., 2003. Precise mapping of a locus affecting grain protein content in durum wheat. Theor. Appl. Genet. 107, 1243–1251.

Prabhu, K.V., Gupta, S.K., Charpe, A., Koul, S., 2004. SCAR marker tagged to the alien leaf rust resistance gene *Lr19* uniquely marking the *Agropyron elongatum*-derived gene *Lr24* in wheat: a revision. Plant Breed. 123, 417–420.

Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A., Allard, R.W., 1984. Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. U.S.A. 81, 8014–8018.

Simmonds, N.W., 1995. The relation between yield and protein in cereal grain. J. Sci. Food Agric. 67, 309.

Snape, J.W., Hyne, V., Aitken, K., 1995. Targeting genes in wheat using marker-mediated approaches. In: Li, Z.S., Xin, Z.Y. (Eds.), Proc. 8th Int. Wheat Genet Symp. China Agric Scientech Press, Beijing, pp. 749–759.

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

Stakman, E.C., Stewart, D.M., Loegering, W.Q., 1962. Identification of Races of *Puccinia graminis* var. *tritici*. U.S. Dept. of Agric. Publ. E617, USDA, Washington, DC, USA.

Stein, I.S., Sears, R.G., Gill, B.S., Hoseney, R.C., Cox, T.S., 1992. Heterogeneity of the 'Wichita' wheat monosomic set for grain quality and agronomic traits. Crop Sci. 32, 581–584.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., Dubcovsky, J., 2006. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science 314, 1298–1301.