Theor Appl Genet (2006) 112: 1195–1204
DOI 10.1007/s00122-006-0221-6

ORIGINAL PAPER

A. Blanco · R. Simeone · A. Gadaleta

Detection of QTLs for grain protein content in durum wheat

Received: 9 March 2005 / Accepted: 15 January 2006 / Published online: 2 February 2006

Abstract Grain protein content (GPC) of durum wheat (*Triticum turgidum* L. var. *durum*) is an important trait for the nutritional value of grain and for influencing the technological property of flour. Protein content is a quantitative trait negatively correlated with grain yield, thus increase in protein quantity usually results in yield reduction. This study was initiated to introgress alleles for high GPC from var. *dicoccoides* into durum wheat germplasm by the backcross inbred line (BIL) method and to identify molecular markers linked to high GPC alleles not associated with depressing effects on yield. The backcross line 3BIL-85 with high GPC and similar grain yield to the recurrent parent was backcrossed to Latino, and the generations F_2, F_3 and F_4 were evaluated for GPC and yield per spike (GYS) in three field trials. Three QTLs with major effects on GPC were detected on chromosome arms 2AS, 6AS and 7BL, identified by the markers *Xcfa2164*, *XP39M37*$_{(250)}$ and *Xgwm577*, respectively. Multiple regression analysis indicated that the three QTLs explained all the genetic variances of the trait. The high GPC parental line 3BIL-85 was not significantly different from the recurrent parent Latino for GYS, but the phenotypic correlation coefficient between GPC and GYS had negative values (from −0.02 to −0.28) in each trial, although it was statistically significant only in the F3 progeny trial. No co-located QTL for GYS was detected, excluding the hypothesis that the putative QTLs for GPC were indirect QTLs for low grain yield. The negative protein-yield response could be due to: (a) co-location of grain yield per spike QTLs with reduced phenotypic effects not detectable by the experimental design or statistical procedures, or to (b) opposite pleiotropic gene effects due to the major bio-energetic requirements for synthesis of protein then carbohydrates. Mapping loci by BILs should enable the production of near-isogenic lines in which the individual effects of each QTL can be examined in detail without confounding variations due to other putative QTLs.

Communicated by G. Wenzel

A. Blanco (✉) · R. Simeone · A. Gadaleta
Department of Environmental and Agro-Forestry Biology and Chemistry, University of Bari, Via Amendola, 165/A, 70126 Bari, Italy
E-mail: blanco@agr.uniba.it
Tel.: +39-80-5442992
Fax: +39-80-5442200

Introduction

Grain protein content (GPC) and gluten quality of durum wheat (*Triticum turgidum* L. var. *durum*) are the most important factors affecting pasta-making technology characteristics. In general, good cooking quality is related to a high level of protein and gluten content or an intermediate content of protein but high gluten quality. Considering the different pasta drying methods applied, D'Egidio et al. (1990) indicated that, at a low drying temperature, GPC and gluten quality were both important factors in determining pasta cooking quality, whereas at a high temperature only GPC was essential. Under normal cropping conditions, wheats show a low GPC, ranging between 10 and 14%. During the past 20 years, a rise in GPC has been achieved mainly through increased nitrogen fertilization. Any genetic improvement in GPC has been restricted by the negative correlation between productivity and GPC found in segregating populations in all cereals [see review by Simmonds (1995)]. As a result of this generally inverse relationship, high yielding cultivars have low GPC. Grain protein quantity is a typical quantitative trait controlled by a complex genetic system and influenced by environmental factors and management practices (nitrogen and water availability, temperature and light intensity). An extensive review by Konzak (1977) and more recent investigations (Levy and Feldman 1989; Stein et al. 1992; Snape et al. 1995; Blanco et al. 1996, 2002; Sourdille et al. 1996; Joppa et al. 1997; Perretant et al. 2000; Dholakia et al. 2001; Zanetti et al. 2001;

Borner et al. 2002; Groos et al. 2003; Prasad et al. 2003) have indicated that the factors influencing grain protein concentration in cultivated and wild wheats are located on all chromosomes.

The lack of sufficient genetic variation for useful traits within the cultivated wheats has limited the ability of plant breeders to improve grain yield and grain quality. A good source of genetic variation has been found in wild wheats and related species. The wild tetraploid wheat *T. turgidum* L. var. *dicoccoides* shows particular promise as a donor of useful genetic variation for several traits, including disease resistance, drought tolerance, and grain protein quality. However, wild germplasm contains both favorable traits and many commercially unacceptable characteristics. Backcrossing has been extensively used in self-pollinating crops to transfer simply inherited characteristics to cultivars which are deficient only in the characteristics being transferred. The application of backcrossing to the improvement of quantitative characteristics has been limited mainly because of the low heritability of these characteristics and the difficulty of transferring, simultaneously, relatively large numbers of genes. In contrast to more traditional procedures of wheat breeding, the backcross inbred line (BIL) method, first described by Wehrhahn and Allard (1965), produces BILs that can be tested in replicated trials over environments prior to selection. The development of molecular markers and molecular maps, and the recent strategies for molecular breeding, referred to as advanced backcross QTL analysis (Tanksley and Nelson 1996), have shown that the backcrossing method has considerable potential for the genetic analyses of quantitative traits and as a method of breeding higher yielding and quality crop varieties.

This study was initiated to: (a) introgress high GPC loci from var. *dicoccoides* into more adapted and agronomically acceptable durum wheat germplasm by means of the BIL method; (b) identify molecular markers associated to high protein content QTLs without depressing effects on yield and to be used in marker-assisted selection; and (c) study the relationship between QTLs for GPC and grain yield per spike (GYS).

Materials and methods

Plant materials and field evaluation

A population of 92 BILs was developed using a procedure similar to that of Wehrhahn and Allard (1965). The semidwarf and high-yielding cultivar Latino of durum wheat was used as a recurrent parent, and the accession MG29896 of *T. turgidum* L. var. *dicoccoides* was used as donor parent because of its high GPC and the acceptable seed size, respectively. Three successive backcrosses were made to Latino followed by four generations of self-pollination. After the first backcross, 120 random BC_1 plants were chosen to initiate the backcross program. No selection was applied in any generation, and plants were chosen randomly for additional backcrossing or selfing. However, some plants were lost in the subsequent generations and a total of 92 BC_3F_5 lines were obtained and multiplied to have sufficient seeds for use in replicated trials. The complete series of BILs and the recurrent parent Latino were evaluated for grain quality and grain yield components in replicated trials in southern Italy at two locations over 2 years (Blanco et al. 2003). The line 3BIL-85 showed a significantly higher GPC and grain yield not significantly different than the recurrent parent in three of the four trials. To detect and map QTLs for high GPC, the line 3BIL-85 was backcrossed to Latino and the F_2 population grown with plants 10 cm apart per row and 30 cm between rows at Valenzano (Bari) in 2000. The parental lines and 142 F_3 or F_4 progenies were evaluated for GPC and GYS at Valenzano in 2001 (F_3) and Valenzano and Gaudiano in 2002 (F_4) by using a randomized complete block design with three replications. Seeds for F_4 progeny trials were obtained from all plants of each F_3 progeny. In all the experiments, each progeny was planted with 3 g of seeds in a single row 1 m long with 30 cm between rows. During the growing season, standard cultivation practices were used. Plots were hand harvested at maturity and GYS for each progeny was obtained by grain yield per row on the number of spikes per row (about 50–60 spikes). GPC, expressed as a percentage of protein on a dry weight basis, was directly obtained on a 2-g sample of whole-meal flour using near-infrared reflectance spectroscopy.

DNA extraction and molecular marker analysis

On the basis of the F_3 trial data, ten high and ten low GPC progenies with similar GYS were selected for making two separate DNA bulks with extreme phenotypes to be screened with microsatellite (SSR) and AFLP markers according to the bulked segregant analysis (Michelmore et al. 1991). Primers detecting polymorphisms between the bulks were then used to test each progeny of the segregant population.

Total DNA was extracted from 15 to 18 F_3 plants derived from each F_2 plant as described by Sharp et al. (1988). SSR markers representing chromosomes of the A and B genome of hexaploid wheat were chosen for analysis. The development of the SSR markers, primer sequences, chromosome location, and annealing temperature were reported by Röder et al. (1998) for GWM, Gupta et al. (2002) for WMC, Somers et al (2004) for BARC, and in the web site http://www.graingenes.gov for CFA. DNA amplifications were carried out in 25-µl reaction mixtures, each containing 100-ng template DNA, 2 μM of SSR primer, 200 μM of dNTP, 2.5 mM $MgCl_2$, 1× PCR buffer (10 mM Tris–HCl, pH 8.3, 10 mM KCl), and 1 unit of *Taq* DNA-polymerase. The following PCR profile in a Perkin Elmer DNA Thermal Cycler was used: initial denaturation at 95°C for 3 min, followed by 35 cycles of 95°C for 1 min, 50°C/65°C for

1 min, 72°C for 2 min with a final extension at 72°C for 10 min. The amplification products were resolved on 6% polyacrylamide gels and stained with ethidium bromide.

AFLP assays were performed essentially as described by Vos et al. (1995). Total DNA was digested with *Pst*I and *Mse*I, and the adaptors were ligated to the end of the restriction fragments. A total of 20 cycles of PCR were performed for pre-amplification. Selective amplification was performed using *Pst*I and *Mse*I primers that contained three selective nucleotides at the 3′ end. The amplification product was labeled using ^{33}P-labeled *Pst*I primer and visualized using Kodak X-OMAT film (exposure for 3–5 days at room temperature). Initially, 37 AFLP primer combinations (PCs) were tested on the two parental lines. From these, the best ten PCs were selected on the basis of the number of bands, clarity of pattern, and distribution on the gel. These were tested again on the parental lines and the high and low GPC bulks. The polymorphic PCs P39M37 and P39M40 between the bulks (*Pst*I primer selective bases: P39 AGA; *Mse* primer selective bases: M37 ACG, M40 AGC) were tested on the 142 progenies.

Statistical analysis and QTL detection

Standard procedures for analysis of variance of GPC and GYS were used for each trial. Least significant differences at $P = 0.05$ were used to compare means of GPC and GYS values of the different genotype groups. Genetic variance (σ^2_G) and broad-sense heritability (h^2_B) were obtained using the variance component estimates. Pearson phenotypic correlation coefficients were calculated between GPC and GYS in each experiment. The QTL detection was based on regression analysis of progeny mean values in individual experiments on polymorphic markers. The presence of a QTL near a marker locus was judged to be likely if (1) a significant effect was observed for a single marker–trait combination in at least one trial with $P \leq 0.001$ or (2) significant effects were observed, in the same direction, for a marker–trait combination at two or more trials with $P \leq 0.01$. The proportion of the phenotypic variance explained by segregation of each significant marker locus was determined by the R^2 value. Linkage of the molecular markers was analyzed by using MapMaker, version 3 for PC (Lander et al. 1987).

Results

Phenotypic variation

The analysis of variance for GPC and GYS revealed highly significant differences between F_3 and F_4 progenies in each of the field trials carried out at Valenzano in 2001 and at Valenzano and Gaudiano in 2002 (Table 1). When combined over environments, progenies, and progeny×environment effects were significant for both traits (data not shown). The means of parental lines and means and ranges of F_2, F_3 and F_4 populations, variance components, and broad-sense heritability estimates for each trait and for each experiment are presented in Table 2. The line 3BIL-85 was always significantly different from the recurrent parent Latino for GPC in each trial, while GYS was similar. Differences in mean values and variances of parental lines and of F_3 and F_4 progenies were observed among the trials carried out in different years and locations, very likely due to the different climatic conditions. The mean values for GPC of the F_3 and F_4 progeny populations ranged from 13.25 to 14.59% in different environments and were always close to the parental means, suggesting only weak epistatic effects. Both GPC and GYS of the F_2, F_3 and F_4 progenies showed normal distribution frequencies typical of quantitative traits. Broad-sense heritability, as estimated on progeny mean basis, ranged from 0.13 to 0.26 for GPC and from 0.14 to 0.29 for GYS among environments. GPC was significantly negatively correlated with GYS only at Valenzano 2001 ($r = -0.28$, $P \leq 0.001$). The phenotypic correlation coefficients for GPC between the different generations were always highly significant ($P \leq 0.01$ or $P \leq 0.001$) and ranging from 0.24 to 0.49.

Table 1 Analysis of variance for grain protein content (GPC) and grain yield per spike (GYS) of F_3 and F_4 progenies derived from crossing the cv. Latino of durum wheat and the backcross inbred line 3BIL-85 (= var. *dicoccoides* acc. MG29896×cv. Latino, BC_3F_5). Data from a randomized complete block design of F_3 progenies grown at Valenzano in 2001 and F_4 progenies grown at Valenzano and Gaudiano in 2002

Source of variation	DF	Mean square					
		F_3 progenies (Valenzano 2001)		F_4 progenies (Valenzano 2002)		F_4 progenies (Gaudiano 2002)	
		GPC	GYS	GPC	GYS	GPC	GYS
Replications	2	29.102***	0.161	0.274	0.147	2.896***	0.642***
Progenies	141	1.426***	0.134***	0.931**	0.101**	0.813***	0.090**
Error	279	0.695	0.060	0.641	0.069	0.406	0.062

**Significant differences at 0.01P
***Significant differences at 0.001P

Table 2 Means, ranges, genetic variances (σ^2_G), and heritability (h^2_B) for grain protein content (GPC) and grain yield per spike (GYS) in the F_2, F_3 and F_4 generations derived from crossing the cv. Latino of durum wheat and the backcross inbred line 3BIL-85 (=var. *dicoccoides* acc. MG29896×cv. Latino, BC_3F_5). The mean values of the selected ten progenies for low GPC (LP progenies) and for high GPC HP progenies) are reported. Data from F_2 plants grown at Valenzano in 2000 and from a randomized complete block design of F_3 and F_4 progenies grown at Valenzano in 2001 and at Valenzano and Gaudiano in 2002, respectively

	F_2 progenies (Valenzano 2000)		F_3 progenies (Valenzano 2001)		F_4 progenies (Valenzano 2002)		F_4 progenies (Gaudiano 2002)	
	GPC	GYS	GPC	GYS	GPC	GYS	GPC	GYS
Latino	14.05	3.05	12.65	2.23	13.83	1.60	12.14	1.80
3BIL-85	16.65	3.12	14.36	2.27	15.25	1.92	13.73	1.90
Progenies: means	15.55	3.07	13.59	2.31	14.59	1.77	13.25	1.86
(Ranges)	(11.3–20.8)	(1.1–4.4)	(11.2–16.4)	(1.3–2.9)	(13.2–16.0)	(1.0–2.3)	(11.7–14.9)	(1.3–2.3)
LP progenies	13.90	3.23	12.6	2.30	12.95	1.75	14.01	1.97
HP progenies	16.80	3.27	14.8	2.30	13.80	1.74	14.70	1.79
LSD (0.05*P*)			1.34	0.39	1.29	0.42	1.02	0.40
CV (%)			6.14	10.63	5.49	14.83	4.81	13.37
σ^2_G			0.244	0.025	0.097	0.011	0.136	0.009
σ^2_E			0.695	0.060	0.641	0.069	0.406	0.062
h^2_B			0.26	0.29	0.13	0.14	0.25	0.13

QTL detection

As many as 235 SSR primer pairs were used for detecting polymorphisms between the parents of the BIL population (cv. Latino and var. *dicoccoides* acc. MG29896). SSR markers were selected on the basis of published maps (Röder et al. 1998; Korzun et al. 1999) to identify from six to eight polymorphic markers for each of the 14 chromosomes of the A and B genomes of durum wheat. The analysis detected 138 polymorphic markers, with a polymorphism value (52.7%) higher than those usually observed among cultivated wheats (25–30%) because of the genetic distance between the parents. The marker analysis on Latino and 3BIL-85 identified 25 polymorphic markers and detected the introgression of several segments of the var. *dicoccoides* located on eight different chromosomes. The 3BIL-85 genome resulted constituted by 91.6% of the recurrent parent Latino and 8.4% of the wild donor parent var. *dicoccoides*, compared with the theoretically expected values of 93.75 and 6.25%, respectively.

On the basis of F_3 data, ten high and ten low GPC progenies with similar GYS (Table 2) were selected for making two DNA bulks of extreme phenotypes to be screened according to the bulked segregant analysis (Michelmore et al. 1991). The selected 25 polymorphic SSR primer pairs were used to amplify bulked and parental DNA simultaneously. Two SSR markers (*Xcfa2164-2A* and *Xgwm577-7B*) showed amplification profiles in the high and low GPC bulks similar to the high and low GPC parents, respectively. Such markers were considered potentially linked to QTLs for GPC and then screened on the 142 progenies. The linkage with protein content QTLs was tested by regression analysis of each marker locus on the F_2, F_3, and F_4 progeny data and on the mean over the F_3 and F_4 progenies. Each marker was also tested for linkage to QTLs for GYS in order to take into account the factors contributing to the negative correlation of GPC and GYS. The marker *Xcfa2164-2A* was found to be significantly associated at 0.01*P* or 0.001*P* with GPC in each trial and negatively associated at 0.05*P* with GYS only in the F_3 progeny trial. The marker *Xgwm577-7B* was significantly associated with GPC in both the F_4 progeny trials ($P \leq 0.01$) and on the F_3 and F_4 progeny mean ($P \leq 0.001$) and negatively associated with GYS in the F_3 progeny trial at $P \leq 0.01$ and over the F_3 and F_4 progeny mean at $P \leq 0.05$ (Table 3). In order to have a detailed map of the 2A chromosome segment(s) of 3BIL-85 carrying the interesting high GPC QTL(s), 19 SSR informative markers (i.e., polymorphic between Latino and the var. *dicoccoides*) located on chromosome 2A were analyzed. Ten markers were found to be polymorphic between Latino and 3BIL-85 and between the high and low protein bulks. Such markers were subsequently analyzed with the 142 progeny DNA and found to be significantly associated with GPC in each trial. The linkage analysis with MapMaker indicated that the ten SSR markers constituted a linkage group (LOD score > 3) spanning a 39.9 cM chromosome segment (Fig. 1).

With the objective of detecting any other introgressed *dicoccoides* segment influencing the GPC of 3BIL-85, 37 AFLP PCs were analyzed between Latino and 3BIL-85. A total of 212 AFLP polymorphic markers out of 2.590 markers were detected, indicating a polymorphism value (8.1%) similar to that found for SSR markers (10.6%). Ten PCs were used to amplify bulked and parental DNA simultaneously. Two AFLP loci (*XP39M37*$_{(250)}$ and *XP39M40*$_{(510)}$), polymorphic between the bulks, were screened on the 142 progenies. The marker *XP39M37*$_{(250)}$ was found to be significantly associated with GPC in the F_3 and F_4 (Gaudiano 2002) progeny data and over the F_3 and F_4 progeny mean ($P \leq 0.01$ or $P \leq 0.001$) and negatively associated at $P \leq 0.05$ with

Table 3 Regression coefficients (*b*) of molecular markers associated to QTLs for grain protein content (GPC) and grain yield per spike (GYS) in F_2 plants, in F_3 and F_4 progenies and mean F_3–F_4 derived from crossing the cv. Latino of durum wheat and the backcross inbred line 3BIL-85 (= var. *dicoccoides* acc. MG29896×cv. Latino, BC_3F_5)

Marker	Trait	F_2 progenies (Valenzano 2000)	F_3 progenies (Valenzano 2001)	F_4 progenies (Valenzano 2002)	F_4 progenies (Gaudiano 2002	Mean F_3–F_4 progenies
Xcfa2164–2A	GPC	1.50**	0.51***	0.39**	0.57***	0.50***
	GYS	0.04	−0.10*	−0.05	0.05	-0.04
XP39M37$_{(250)}$-6A	GPC	0.41	0.84***	0.23	0.32**	0.46***
	GYS	0.26	−0.07	−0.02	−0.10*	−0.06*
Xgwm577-7B	GPC	0.77	0.45*	0.42**	0.40**	0.42***
	GYS	0.40	−0.15**	−0.02	−0.07	−0.10*

*Significantly different at 0.05*P*
**Significantly different at 0.01*P*
***Significantly different at 0.001*P*

GYS at Gaudiano 2002 and over the F_3 and F_4 progeny mean (Table 3). The chromosome location of *XP39M37$_{(250)}$* on the short arm of chromosome 6A was obtained by using a mapping population developed by Blanco et al. (2004).

The GPC frequency distributions of F_3 and F_4 progenies from the three genotype groups for each marker showed continuous distribution and covered the entire range of GPC, but the distribution patterns varied somewhat with genotypes (see Fig. 1 for *Xcfa2164-2A*). '1' refers to the marker allele associated with the 2AS chromosome segment from the low GPC parent (Latino), and '2' refers to the marker allele associated with the 2AS chromosome segment from the high GPC parent (3BIL-85). Variation within a genotypic class results from the genetic effects of other QTLs, environmental variation, interactions between the 2A QTL and other QTLs, and variability in the phenotypic evaluations (experimental error). The *t* test applied to homozygote genotypes based on the SSR markers confirmed the regression analysis, i.e., the substitution of the low GPC parent alleles with the high GPC parent alleles resulted in a significant increase of the GPC (Table 4).

The amount of phenotypic variation for GPC explained by each QTL, estimated by the coefficient of genetic determination (R^2), ranged from 4.2 to 21.7% in individual trials. The R^2 value was 16.8% for *Xcfa2164-2A*, 14.4% for *XP39M37$_{(250)}$-6A* and 9.1% for *Xgwm577-7B*, respectively, over the F_3 and F_4 progeny mean data.

Discussion

QTLs for grain protein content

The present study was initiated to introgress high GPC loci from var. *dicoccoides* into more adapted and agronomically acceptable wheat germplasm and to identify novel alleles for high GPC not associated with depressing effects on yield. In order to investigate QTLs as single Mendelian factors, we used a set of BILs according to the advanced backcross QTL analysis (Tanksley and Nelson 1996). BILs are characterized by a low proportion of the donor parent genome in each of the population members and therefore are ideally suited for mapping intervarietal and interspecific variation. QTLs can be broken down in different lines so that each line carries only one or few QTLs that can be easily analyzed. Recent studies have indicated that a QTL can be a single locus as demonstrated with some QTLs mapped or cloned in tomato (Frary et al. 2000; Fridman et al. 2000), rice (Takahashi et al. 2001), *Arabidopsis* (El-Assal et al. 2001), and durum wheat (Olmos et al. 2003). We evaluated the Latino-*dicoccoides* BILs for grain quality and grain yield components in four replicated trials (Blanco et al. 2003). Six superior BILs for increased GPC were selected, while retaining the desirable agronomic characteristics of the recurrent parent and without other undesirable characteristics of the donor parent, such as fragile spikes, hard glumes, late maturity, and tallness. The marker analysis on the six high GPC lines detected the introgression of several segments of the var. *dicoccoides* located on eight different chromosomes, but no clear evidence of marker-QTL association was obtained. One of the high GPC lines (3BIL-85) was then backcrossed to Latino, and the generations F_2, F_3 and F_4 were used for QTL detection. Three QTLs with major effects on GPC were detected on chromosome arms 2AS, 6AS and 7BL, identified by the markers *Xcfa2164-2A*, *XP39M37$_{(250)}$-6A*, and *Xgwm577-7B*, respectively. The favorable effects on GPC were due to alleles from the 3BIL-85 parent with high GPC. The phenotypic effect of each QTL was variable across locations, thus indicating that individual QTLs seem to be sensitive to environmental factors. The multiple regression analysis indicated that the three putative QTLs explained 27.3, 12.1, 25.3% of the phenotypic variance of GPC in F_3, F_4 (Valenzano 2001), and F_4 (Gaudiano 2001), respectively. Taking into account the heritability values observed in each trial (Table 2) and the genotypic coefficient of determination ($R^2_G = R^2_P/h^2$), the three QTLs explained all the genetic variance of the segregant population analyzed.

We compared the genomic regions involved in the quantitative expression of GPC found in the Latino-3BIL-85 backcross population with the map location of QTLs found in different genetic materials. The protein

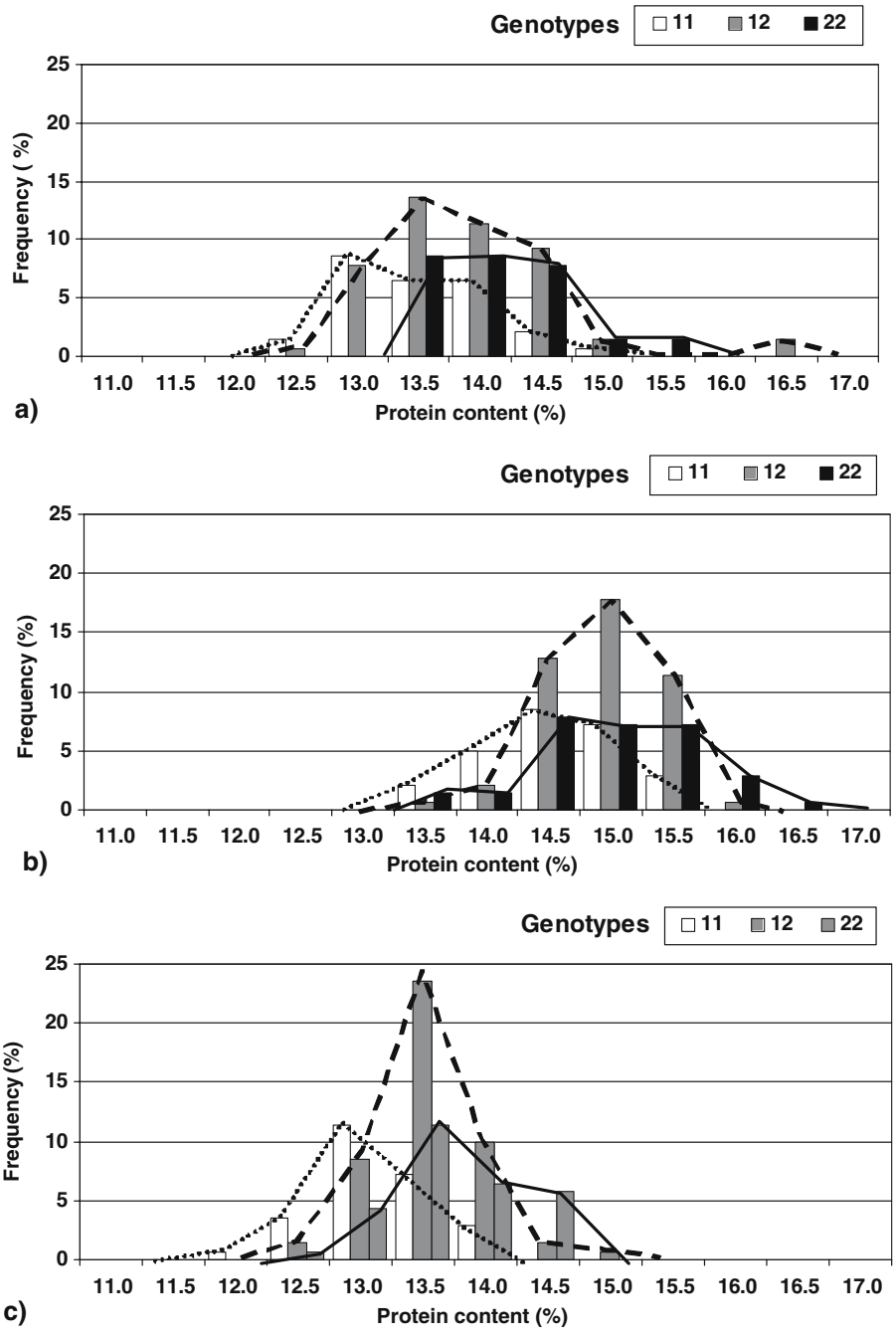

Fig. 1 Frequency distributions of grain protein content (GPC) values of three 2 genotypes based on Xcfa2164-2A in a F2 population of Latino x 3BIL-85. GPC data 3 from (**a**) F2:F3 progenies grown at Valenzano in 2001 and (**b**) and (**c**) F2:F4 progenies 4 grown at Valenzano and Gaudiano in 2002, respectively. 1 = marker allele from cv. 5

content QTL detected on chromosome arm 2AS (Fig. 2), which explained about 16.8% of phenotypic variance, appeared to be the most promising because of lack of association with low GYS. The influence of group-2 chromosomes on GPC control was already reported by Joppa and Cantrell (1990) using durum wheat—var. *dicoccoides* chromosome substitution lines. More recently Groos et al. (2003) and Prasad et al. (2003) reported a protein content QTL on chromosome arm 2AS. Based on some markers in common on the reported maps, the QTL detected in the present study should be considered different as it is localized on 2AS near the centromere, while Groos et al. (2003) and Prasad et al. (2003) mapped the QTL in the distal region of 2AS more then 50 cm distant from the common markers *Xgwm122–2A* and *Xgwm515–2A*.

The protein content QTL identified by the AFLP marker $XP39M37_{(250)}$-*6A*, localized on chromosome arm 6AS close to the centromere, explained about 14.4% of phenotypic variance of the trait. Blanco et al. (2002) found a significant protein content QTL marked by the same $XP39M37_{(250)}$ locus in a recombinant inbred line (RIL) population derived from crossing the cv. Messapia and the acc. MG4343 of the var. *dicoccoides*. Perretant et al. (2000) were able to identify a major QTL on chromosome arm 6AS, close to the AFLP locus

Table 4 Mean values of grain protein content (GPC) and grain yield per spike (GYS) of homozygote genotypes based on microsatellite markers in F_3 and F_4 progenies and mean F_3–F_4 derived from crossing the cv. Latino of durum wheat and the backcross inbred line 3BIL-85 (= var. *dicoccoides* acc. MG29896×cv. Latino, BC_3F_5). 11 = marker alleles from cv. Latino (low grain protein content) and 22 = marker alleles from 3BIL-85 (high grain protein content)

Marker		F_3 progenies (Valenzano 2001)		F_4 progenies (Valenzano 2002)		F_4 progenies (Gaudiano 2002)		Mean F_3–F_4 progenies	
	Genotypes trait	11	22	11	22	11	22	11	22
Xcfa2164-2A	GPC	13.27	13.76#	14.31	14.74***	12.94	13.48#	13.51	14.00#
	GYS	2.37	2.28*	1.78	1.72	1.89	1.83*	2.14	2.11
XP39M37$_{(250)}$-6A	GPC	13.20	14.01#	14.52	14.64	13.04	13.42***	13.59	14.05#
	GYS	2.36	2.25**	1.74	1.73	1.90	1.79**	2.16	2.06**
Xgwm577-7B	GPC	13.43	13.95***	14.43	14.84***	13.07	13.445***	13.64	14.08#
	GYS	2.37	2.22***	1.77	1.76	1.88	1.80	2.18	2.07***

#Significantly different at 0.10*P*
*Significantly different at 0.05*P*
**Significantly different at 0.01*P*
***Significantly different at 0.001*P*

XE38S60$_{(200)}$, which explained about 17% of phenotypic variance of GPC. These QTLs could be allelic as they appear to be at similar positions on the proximal region of 6AS between the centromere and the *Gli-A2* locus.

The third QTL for GPC detected in the present study was located on chromosome arm 7BL close to the SSR marker *Xgwm577* that explained 9.1% of the phenotypic variance of the trait. Blanco et al. (1996, 2002) and Sourdille et al. (1999) reported QTLs for GPC on 7AS and Stein et al. (1992) and Zanetti et al. (2001) on 7A without arm information.

The present study confirms that some of the protein content QTLs appear to be at similar positions in different genetic backgrounds, indicating a similar genetic control in different germplasms. In most cases population sizes used for mapping QTLs are between 100 and 200 genotypes (the present study analyzed a population of 142 progenies), therefore analytical approaches to QTL analysis provide poor precision about QTL location. The process of first identifying linkage between markers and traits in a mapping population and subsequently testing the effects of markers in near-isogenic lines (NILs) seems to provide strong evidence for QTL position and effect. We are developing a set of substitution lines, recombinant for the 2AS region of 3BIL-85, in which the location of the 2A QTL can be determined more accurately. A combination of QTL mapping and substitution mapping provides the basis for a realistic approach of identifying and locating a QTL to a smaller region, a prerequisite for marker-assisted selection and for isolating genes via positional cloning (Lecomte et al. 2004).

Relationship between QTLs for grain protein content and grain yield

One of the major objectives of wheat breeding programs has been to increase GPC while maintaining or increasing grain yield of lines released for commercial production. Simultaneous improvement in both traits has been limited by the generally negative relationship between GPC and grain yield components in wheats, as well as in barley, maize, oat and sorghum, with reported correlations ranging from −0.2 to −0.8 (see review by Simmonds 1995). As a result of these inverse relationships, high yielding wheat cultivars have generally low increase GPC, especially when grown under fertile areas or under irrigation and where nitrogen is not a limiting factor. The negative yield–protein correlation can be attributed to environmental factors, genetic components, dilution of grain nitrogen with a much larger grain biomass accumulation, or to bio-energetic requirements for synthesis of carbohydrates and proteins (Bathia and Rabson 1987). The genetic basis of trait correlations may include single genes with pleiotropic gene effects or the tight linkage of several genes controlling different traits. The extensive review by Konzak (1977) and more recent investigations (Levy and Feldman 1989; Stein et al. 1992; Snape et al. 1995; Blanco et al. 1996, 2002; Sourdille et al. 1996; Joppa et al. 1997; Perretant et al. 2000; Dholakia et al. 2001; Zanetti et al. 2001; Borner et al. 2002; Groos et al. 2003; Prasad et al. 2003) have indicated that factors influencing protein concentration in cultivated and wild wheats are located on all chromosomes. However, only a few studies have included results for GPC and yield components obtained on the same population (Zanetti et al. 2001; Blanco et al. 2002; Borner et al. 2002; Groos et al. 2003) to determine whether the GPC-yield negative relationship resulted from opposite pleiotropic gene effects and/or co-location of different QTLs. Zanetti et al. (2001) determined QTLs for GPC and thousand-kernel weight and found both positive and negative genetic relations, depending on the QTL. However, yield was not evaluated in their cross between wheat and spelta. Blanco et al. (2002) detected seven independent genomic regions involved in the expression of high GPC in a tetraploid RIL population, six of which were associated

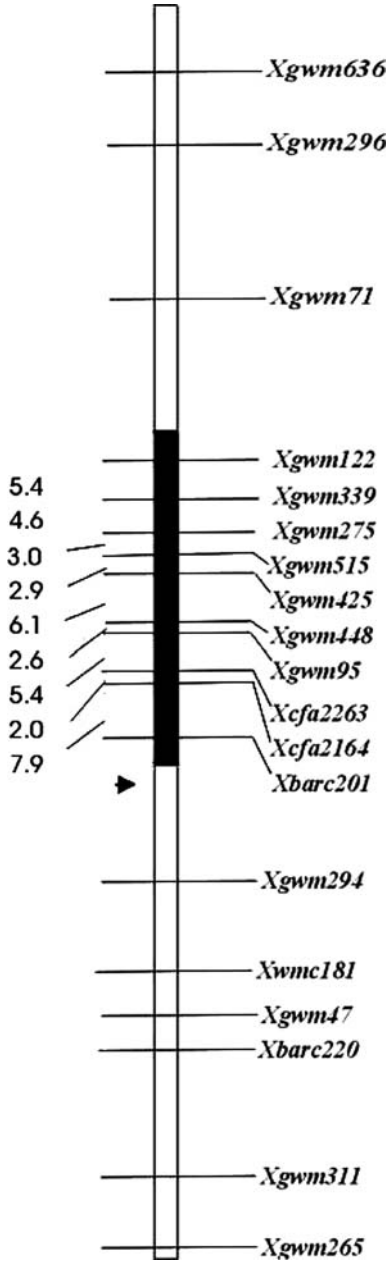

Fig. 2 Linkage map of the 2A chromosome segment introgressed from var. *dicoccoides* 21 in the backcross inbred line 3BIL-85 (*black region*). Map distances (cM) and 22 microsatellite markers significantly associated to grain protein content are shown on the 23 *left* and *right sides*. The centromere (*filled triangle*) and the introgressed segment were positioned 24 according to previously published maps (Röder et al. 1998; Blanco et al. 2004)

with QTLs for low grain yield. Borner et al. (2002) found two QTLs for GPC co-located with QTLs for low GYS on chromosome arms 2DS and 7AS in the ITMI mapping population. Groos et al. (2003) found co-location of QTLs for GPC, 1,000-kernel weight and grain yield on five different chromosomes, but in some cases the favorable alleles came from the same parental line. Only one important QTL for yield was however found in their bread wheat RIL population.

The primary components of grain yield per area are GYS and number of spike per area. The latter is strongly affected by environmental factors and by agrotechnique practices. The components of GYS are the number of kernels per spike and the kernel weight which are negatively correlated each other and should not be considered separately for detecting grain yield QTLs. In the present study, we considered GYS taken on plants grown at normal field density. Although the high GPC parental line 3BIL-85 was not significantly different from the recurrent parent Latino for GYS, the phenotypic correlation coefficient between GPC e GYS had negative values (from -0.02 to -0.28) in the F_2, F_3, and F_4 progeny trials, but it was statistically significant only in the F_3 progeny trial. The presence of a QTL near a marker locus was judged to be likely if a significant effect was observed for a single marker–trait combination in at least one environment with $P \leq 0.001$ or at least two environments with $P \leq 0.01$. The regression analysis detected three putative QTLs for GPC explaining all the genotypic variation of the trait. No co-located QTL for GYS was detected by using the same or less stringent criteria. These evidences excluded the hypothesis that the detected QTLs for GPC were indirect QTLs for low grain yield due to dilution of grain nitrogen with a much smaller grain biomass accumulation. The correlated response on GYS was a decrease of the trait not significant or close to statistical significance. Such a negative protein–yield response could be due to: (a) co-location of GYS QTLs with reduced phenotypic effects not detectable by the experimental design or statistical procedures; or to (b) opposite pleiotropic gene effect due to the major bioenergetic requirements for synthesis of protein then carbohydrates. Under the same soil nitrogen supply, a GPC increase should produce a GYS decrease. If the 'opposite pleiotropic effects' occur, the breeder should simply accept the need for compromises between high grain yield and high GPC which will be determined by economic factors. In all cases, a more detailed genetic analysis of linkage groups on large populations is needed to distinguish between clusters of tightly linked genes and single genes with pleiotropic effects.

Conclusions

The relatively low heritability of quantitative traits makes it difficult to select for useful alleles from wild or cultivated germplasm based on phenotypic evaluation of single plants and/or in single environments. The production of backcrossing lines is a laborious and sometimes expensive procedure, but the availability of BILs for quantitative traits facilitates identification of the genes involved in the inheritance of continuous variation, as BILs can be evaluated in replicated trials to identify those which deviate significantly from the recurrent parent. Mapping loci by BILs should enable the production of NILs in which the individual effects of each

QTL can be examined in detail without confounding variations due to other putative QTLs. The NILs for GPC could also be useful to conduct physiological studies aimed at investigating the potential mechanisms leading to high protein content and to elucidate if the high protein concentration of low-yielding genotypes is primarily due to genes for high GPC and not to loci for low grain yield. Combining QTLs for GPC with those for high yield and other agronomically important traits will require information on the QTLs controlling all of these traits in the genetic material under study. The implication for marker-assisted selection in breeding programs is that in the QTL analysis aiming to determine which regions are to be transferred using markers, it is of utmost importance to examine whether those regions contain QTLs for other traits that will affect the total performance of the genotypes.

Acknowledgements The research project was supported by grants from Ministero della Ricerca Scientifica e Tecnologica, project 'PRIN 2003' and from Ministero delle Politiche Agricole, Alimentari e Forestali, project 'FRUMISIS'.

References

Bathia CR, Rabson R (1987) Relationship of grain yield and nutritional quality. In: Nutritional quality of cereal grains: genetic and agronomic improvement. Agronomy Monograph no. 28, ASA-CSSA-SSSA, Madison, WI 53711, USA, pp 11–43

Blanco A, De Giovanni C, Laddomada B, Sciancalepore A, Simeone R, Devos KM, Gale MD (1996) Quantitative trait loci influencing grain protein content in tetraploid wheats. Plant Breed 115:310–316

Blanco A, Pasqualone A, Troccoli A, Di Fonzo N, Simeone R (2002) Detection of grain protein content QTLs across environments in tetraploid wheats. Plant Mol Biol 48:615–623

Blanco A, Gadaleta A, Simeone R (2003) Variation for yield and quality components in durum wheat backcross inbred lines derived from ssp. *dicoccoides*. Aust J Agric Res 54:163–170

Blanco A, Simeone R, Cenci A, Gadaleta A, Tanzarella OA, Porceddu E, Salvi S, Tuberosa R, Figliuolo G, Spagnoletti P, Roder MS, Korzun V (2004) Extention of the "Messapia x *dicoccoides*" linkage map of *Triticum turgidum* (L.) Thell. Cell Mol Biol Lett 9:529–541

Borner A, Schumann E, Furste A, Coster H, Leithold B, Röder MS, Weber WE (2002) Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). Theor Appl Genet 105:921–936

D'Egidio MG, Mariani BM, Nardi S, Novaro P, Cubadda R (1990) Chemical and technological variables and their relationships: a predictive value equation for pasta cooking quality. Cereal Chem 67:275–281

Dholakia BB, Ammiraju JSS, Sandra DK, Singh H, Katti MV, Lagu MD, Tamhankar SA, Rao VS, Gupta VS, Dhaliwal HS, Ranjekar PK (2001) Molecular marker analysis of protein content using PCR-based marker in wheat. Biochem Genet 39:325–338

El-Assal SE, Alonso-Blanco C, Peeters AJM, Raz V, Koorneef M (2001) A QTL for flowering time in *Arabidopsis* reveals a novel allele for CRY2. Nat Genet 29:435–440

Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpter KB, Tanksley SD (2000) *Fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88

Fridman E, Pleban T, Zamir D (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484bp within an invertase gene. Proc Natl Acad Sci USA 97:4718–4723

Groos C, Robert N, Bervas E, Charmet G (2003) Genetic analysis of grain protein content, grain yield and thousand-kernel weight in bread wheat. Theor Appl Genet 106:1032–1040

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

Joppa LR, Cantrell RG (1990) Chromosomal location of genes for grain protein content of wild tetraploid wheat. Crop Sci 30:1059–1064

Joppa LR, Du C, Hart GE, Hareland GA (1997) Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. Crop Sci 37:1586–1589

Konzak CF (1977) Genetic control of the content, aminoacid composition and processing properties of proteins in wheat. Adv Genet 19:407–582

Korzun V, Röder MS, Wandekake K, Pasqualone A, Lotti C, Ganal MW, Blanco A (1999) Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. Theor Appl Genet 98:1202–1207

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lecomte L, Saliba-Colombani V, Gautier A, Gomez-Jimenez MC, Duffé P, Buret M, Causse M (2004) Fine mapping of QTLs on cromosome 2 affecting the fruit architecture and composition of tomato. Mol Breed 13:1–14

Levy AA, Feldman M (1989) Location of genes for high grain protein percentage and other quantitative traits in wild wheat, *T. turgidum* var. *dicoccoides*. Euphytica 41:113–122

Michelmore RW, Paran I, Kesselli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Olmos S, Diestelfeld A, Chicaiza O, Schatter AR, 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

Perretant MR, Cadalen T, Charmet G, Sourdille P, Nicolas P, Boeuf C, Tixier MH, Branlard G, Bernard S, Bernard M (2000) QTL analysis of bread-making quality in wheat using a doubled haploid population. Theor Appl Genet 100:1167–1175

Prasad M, Kumar N, Kulwal PL, Roder MS, Balyan HS, Dhaliwal HS, Roy JK, Gupta PK (2003) QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. Theor Appl Genet 106:659–667

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

Simmonds NW (1995) The relation between yield and protein in cereal grain. J Sci Food Agric 67:309–315

Sharp PJ, Kreis M, Shewry PR, Gale MD (1988) Location of b-amylase sequences in wheat and its relatives. Theor Appl Genet 75:286–290

Snape JW, Hyne V, Aitken K (1995) Targeting genes in wheat using marker-mediated approaches. In: Li ZS, Xin ZY (eds) Proc 8th Int Wheat Genetics Symp, Beijing, 20–25 July 1993, China Agric Scientech Press, Beijing, China pp 749–759

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

Sourdille P, Perretant MR, Charmet G, Leroy P, Gautier MF, Joudrier P, Nelson JC, Sorrells ME, Bernard M (1996) Linkage between RFLP markers and gene affecting kernel hardness in wheat. Theor Appl Genet 93:580–586

Sourdille P, Perretant MR, Charmet G, Cadalen T, Tixier MH, Joudrier P, Gautier MF, Branlard G, Bernard S, Boeuf C, Bernard M (1999) Detection of QTL for bread making quality in wheat using molecular markers. In: Scarascia Mugnozza GT, Porceddu E, Pagnotta MA (eds) Genetics and breeding for crop quality and resistance. Kluwer , Netherlands, pp 361–366

Stein IS, Sears RG, Hoseney RC, Cox TS, Gill BS (1992) Chromosomal location of genes influencing grain protein concentration and mixogram properties in Plainsman V winter wheat. Crop Sci 32:573–580

Takahashi Y, Shomura A, Sasaki T, Yano M (2001) *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity encodes the α-subunity of protein CK2. Proc Natl Acad Sci USA 98:7922–7927

Tanksley SD, Nelson JC, (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191–203

Vos P, Herges R, Blecker M, Reijans M, Van De Lee T, Hores M, Frijters A, Pop J, Pelman J, Kuiper M, Zabeau M (1995) AFLP a new technique for DNA fingerprinting. Nucl Acid Res 23: 4407–4414

Wehrhahn C, Allard RW (1965) The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. Genetics 51:109–119

Zanetti S, Winzeler M, Feulillet C, Keller B, Messmer M (2001) Genetic analysis of bread-making quality in wheat and spelt. Plant Breed 120:13–19