Theor Appl Genet (2003) 106:1032–1040
DOI 10.1007/s00122-002-1111-1

C. Groos · N. Robert · E. Bervas · G. Charmet

Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat

Received: 4 January 2002 / Accepted: 7 June 2002 / Published online: 3 October 2002

Abstract Grain yield and grain protein content are two very important traits in bread wheat. They are controlled by genetic factors, but environmental conditions considerably affect their expression. The aim of this study was to determine the genetic basis of these two traits by analysis of a segregating population of 194 F_7 recombinant inbred lines derived from a cross between two wheat varieties, grown at six locations in France in 1999. A genetic map of 254 loci was constructed, covering about 75% of the bread wheat genome. QTLs were detected for grain protein-content (GPC), yield and thousand-kernel weight (TKW). 'Stable' QTLs (i.e. detected in at least four of the six locations) were identified for grain protein-content on chromosomes 2A, 3A, 4D and 7D, each explaining about 10% of the phenotypic variation of GPC. For yield, only one important QTL was found on chromosome 7D, explaining up to 15.7% of the phenotypic variation. For TKW, three QTLs were detected on chromosomes 2B, 5B and 7A for all environments. No negative relationships between QTLs for yield and GPC were observed. Factorial Regression on G×E interaction allowed determination of some genetic regions involved in the differential reaction of genotypes to specific climatic factors, such as mean temperature and the number of days with a maximum temperature above 25 °C during grain filling.

Keywords *Triticum aestivum* L. · QTL · Molecular markers · Interaction · Factorial regression

Communicated by J.W. Snape

N. Robert · G. Charmet (✉)
INRA, Station d'Amélioration des Plantes,
234 avenue du Brézet 63039 Clermont-Ferrand Cedex 2, France
e-mail: charmet@clermont.inra.fr

E. Bervas
ULICE, ZAC "Les portes de Riom", 63200 RIOM France

C. Groos
ETH-IPW, Eschikon 33, 8315 Lindau Swizerland

Introduction

Yield and grain protein-content (GPC) are traits of primary importance in bread-wheat breeding programs. The first one is obviously a major determinant of farmer's incomes, while the second one is very important for bread quality. Accurate evaluation of these traits is made difficult by the importance of the genotype × environment (G×E) interactions (Robert et al. 2001) and, thus, determination of molecular markers linked to these traits would help plant breeders to develop cultivars that combine high yield with high grain protein-content.

For many years, a negative relationship between these two traits was observed but little was known whether this association was due to a close genetic relationship or to opposite environmental effects on the two traits. Oury (personal communication), combining results from 14 years of multi-site trials, has shown that this relationship was closer when average genetic values were considered, suggesting a genetic relationship between the two traits.

The genetic components of GPC have been extensively studied in bread wheat (Joppa et al. 1997; Prasad et al. 1999; Perretant et al. 2000; Zanetti et al. 2001). The greatest influence was detected by Joppa et al. (1997), who found a QTL explaining 66% of the phenotypic variation for GPC located on chromosome 6B. Fewer results are available for yield and, generally, the studies have focused only on a single chromosome (Hyne and Snape 1991; Araki et al. 1999; Shah et al. 1999). Some studies have examined the influence of the *Rht* dwarfing genes (Hyne and Snape 1991) on yield. More results are available on yield components, such as thousand-kernel weight (TKW) (Campbell et al. 1999; Prasad et al. 1999; Zanetti et al. 2001). To our knowledge, only one study has included results for GPC and yield components obtained on the same population (Zanetti et al. 2001).

Grain protein-content and yield are both largely influenced by environmental conditions such as soil fertility, rainfall or temperature. Many authors (e.g. Bhullar and Jenner 1985; Wardlaw and Wrigley 1994; Daniel and

Triboï 2000) have shown that temperature and nitrogen nutrition influence both grain weight and grain protein-content. Therefore, descriptors of environmental conditions, such as climate or soil traits are candidate covariates for interpreting G×E interactions for yield and GPC.

Different statistical models have been used in the analysis of G×E interactions, such as the Additive Main effects and Multiplicative Interaction (AMMI) model (Gauch 1988, 1992), Partial Least Squares Regression (Aastveit and Martens 1986) or Factorial Regression (Denis 1980, 1988). The AMMI model is useful in the dissection of G×E interaction, even if external covariates can not be introduced. Factorial Regression and Partial Least Squares Regression allow determination of environmental covariates influencing G×E, and Vargas et al. (1999), using these two models, have found similar results. Crossa et al. (1999) proposed the inclusion of molecular markers as genetic covariates in Factorial Regression to determine significant cross products between genetic and environmental covariables.

The objectives of this study were to:

(1) determine the genetic basis of the relationship between grain-protein content and yield or yield components through QTL analyses, and
(2) dissect G×E interaction for these traits using both environmental and genetic covariates.

Materials and methods

Plant material and technological analyses

The plant population studied has been described previously (Groos et al. 2002). It consisted of 194 F_7 RILs obtained by single-seed descent (SSD) from the cross between 'Renan' and 'Récital'. 'Récital' is more productive while 'Renan' has a higher grain protein-content (GPC) and thousand-kernel weight (TKW). The population was sown in autumn, 1998, and harvested in summer, 1999, in six locations in France, Châlons-en-Champagne (CHAL), Chartainvilliers (CVIL), Clermont-Ferrand (CF), Le Moulon (LM), Mons (MO) and Rennes (RN). The experimental field design consisted of a randomized trial with two replications, divided into blocks. The parental lines were replicated in every block to control field heterogeneity. Each plot measured between 5.4 m^2 and 7.5 m^2 depending on location, and plants were grown under normal field conditions.

Grain yield was measured for all plots. Grain protein-content (GPC) and thousand-kernel weight (TKW) of RILs were estimated only for one replication in each location. For the parental lines, GPC and TKW was measured for all replications in all locations, except at Rennes for TKW. Thousand-kernel weight was evaluated by weighing two samples of 100 kernels for each plot. Grain protein-content was determined on whole grains by NIR (Near Infrared Reflectance) spectroscopy using an Infratec 1221 Grain Analyser according to the AACC method 39-25 (AACC 1995).

The genetic linkage map was constructed on this population using microsatellite, RFLP and AFLP markers. This map was described by Groos et al. (2002). In order to saturate it, more microsatellites were added. The map used for the QTL analysis consisted of 254 loci on 38 linkage groups for a total length of 2,722 cM. It covered all the chromosomes, except chromosome 4D, with a genetic coverage of about 75% compared to reference maps (Röder et al. 1998; Guyomarc'h 2000). Some unlinked markers which did not deviate from the expected ratio (1:1) were also used in QTL analyses.

QTL detection

QTL analysis was performed using a Splus 'home made' program described by Groos et al. (2002). First, one-way analyses of variance (ANOVA) were used to detect significant differences between genotypic classes for each marker. Significant markers were used as candidates in a multiple regression model, in order to select a subset of non-redundant markers for further use as covariates. Then, on every linkage group in which at least one marker was found significant, the marker-regression method (Kearsey and Hyne 1994) was carried out to locate the QTLs more precisely and estimate their additive effects. This program allows the detection of two QTLs on the same chromosome using a two-dimensional scanning of the chromosome (Hyne and Kearsey 1995). The 95% confidence intervals of the QTL locations and additive effects were established by bootstrapping (Visscher et al. 1996) using 200 replicates for the one-QTL model and 400 for the two-QTLs model.

Statistical analyses of G×E

To assess the significance of the G×E interaction, we analyzed the replicated data available only for the parental lines through ANOVA using the GLM procedure in SAS software (1991).

For the RILs, the G×E interaction was first decomposed with the Additive Main effects and Multiplicative Interaction (AMMI) model (Gauch 1988, 1992) to determine whether the G×E interaction was structured or not, i.e. whether at least one multiplicative term accounted for a significant part of the G×E sum of squares. When no multiplicative term is significant, there is no need to go further in the analysis, as no covariate would explain a significant part of G×E interaction. In this study, we applied a two-term AMMI model on the data subset containing no missing values and, because we had no replicate for the RILs, we used the residual of this AMMI model (i.e. the rest of G×E interaction was not accounted for by the first two multiplicative terms) as the error term to compute pseudo-*F* statistics.

When the AMMI models revealed at least one significant multiplicative term, we further analysed the data by Factorial Regression (Denis 1980, 1988) using INTERA software (Decoux and Denis 1991). For factorial regression, we used environmental covariables. The study was primarily devoted to GPC, so we tested a set of covariates characterizing the grain-filling period, defined as the period between flowering and flowering plus 700 degree-days (Gate 1995). The covariates, which have been used, were Tm, the mean of the average daily temperature during the grain-filling period, T25, the sum of the maximum daily temperature in degree-days on a 25 °C basis, and NbD25, the number of days, with a maximum daily temperature above 25 °C during the grain-filling period. We tested these covariates for the three traits, and retained the covariates that accounted for most of the sum of squares, even when they were not significant. Then, instead of testing every marker in the same way as every climatic covariate, we tested by ANOVA whether some markers were significantly associated with the genotypic regression coefficient on environmental covariates. Finally, the markers, that significantly explained the genotypic regression coefficient, were introduced in a typical Factorial Regression model as proposed by Crossa et al. (1999). For a single pair of covariates, the model was:

$$E(X_{ij}) = \mu + \alpha_i + \beta_j + M_i \nu COV_j + \gamma_i COV_j + \delta_j M_i + \theta_{ij},$$

where μ is the overall mean, α_i is the genotype main effect, β_j is the environment main effect, ν is the coefficient of regression on the product of the covariables, γ_i is the specific response of genotype to the environmental covariate COV_j and δ_j indicates the weighting effect of environment j with respect to the M_i marker influence.

A multiplicative term was introduced after the covariates, in order to test whether the rest of the G×E interaction, not accounted for by Factorial Regression, still contained a structured fraction (Baril 1992). As in the AMMI model, the rest of the G×E interaction was used as an error term in computing *F* statistics.

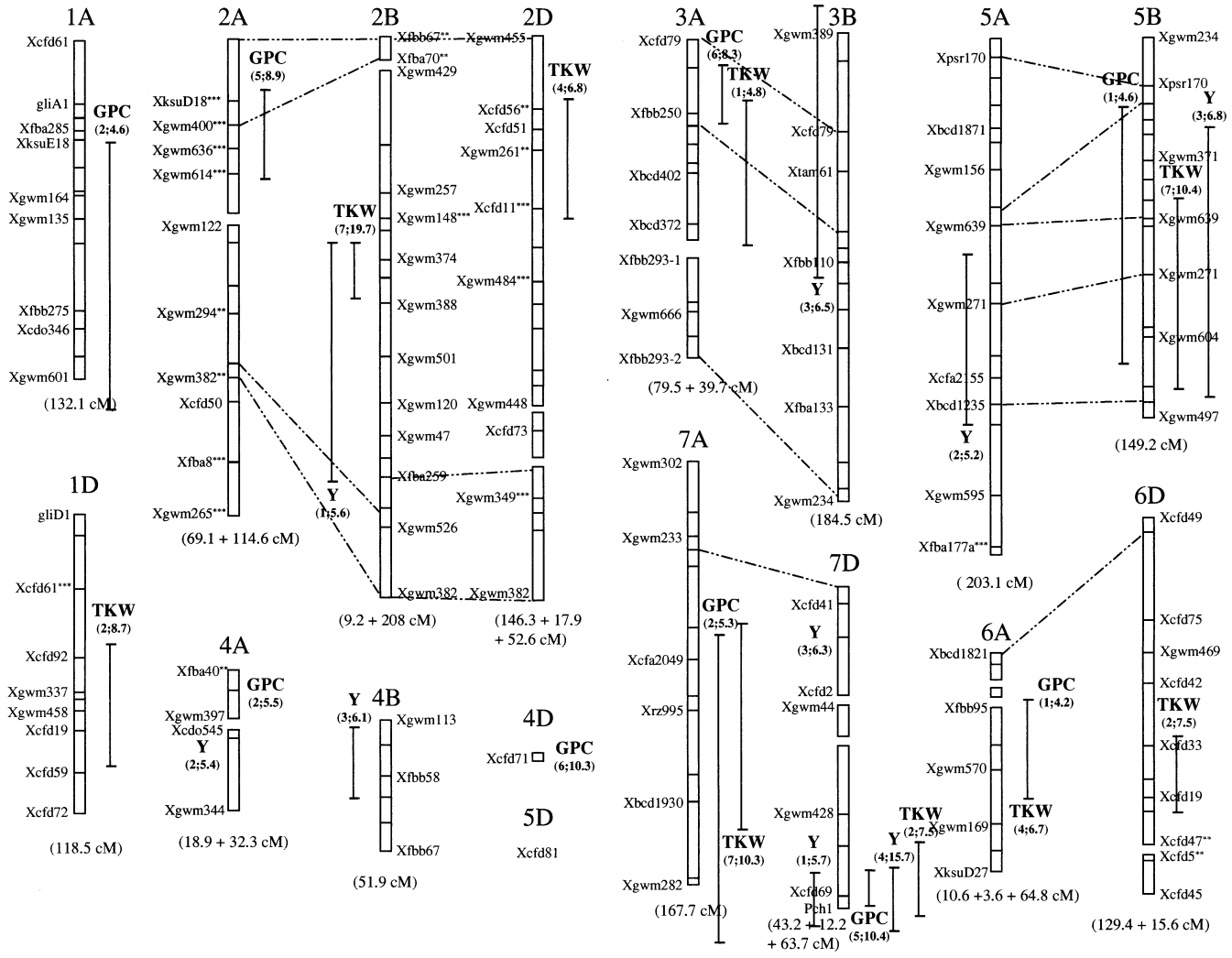

Fig. 1 Locations of the principal QTLs for grain protein-content (GPC), yield (Y) and thousand-kernel weight (TKW), and position of the markers involved in the G×E interaction for all these traits on the map 'Renan' × 'Récital'. □ distance on the chromosome between two anchor markers. ··· homoeologous relationship, И confidence interval of the QTL (when the QTL was detected in different environments, the confidence interval corresponds to the shortest). The favorable allele is by Renan when the line is on the right of the chromosome, and by Récital when it is on the left. (3;13.6) the first number corresponds to the number of environments, where the QTL was detected; the second is the highest value for R^2 for the character. (53 cM) length of the chromosome in centimorgans; when the chromosome is not in a singlepart, the lengths of the different parts are given

Results

Phenotypic variation

Table 1 summarizes the phenotypic data for the three traits observed in the six locations. Renan had a 2% to 3% higher protein content and kernels were 11 to 20 mg heavier than Recital, while grain yield was similar. For the first two traits, the mean RILs value was close to the parental mean, suggesting only weak epistatic effects. The pattern was less clear for yield, the mean RILs value being lower than the value of the poorest parental line in some locations. For all the three traits, the range of the RILs population was much larger than the range of parental lines suggesting that favourable alleles are present in both parents. For example, the difference between the lines with the highest and the lowest protein contents was nearly twice as high as the difference between Renan and Récital in most locations.

Pearson correlation coefficients were negatively significant ($P < 0.01$) between GPC and yield, with a mean of –0.40 but with large differences according to location (from $r = –0.66$ at Rennes to $r = –0.19$ at Chartainvilliers). Between yield and thousand-kernel weight, correlation coefficients were only positively significant ($P < 0.01$) at Clermont-Ferrand and Châlons-en-Champagne.

QTL detection

All putative QTLs are presented in Table 2 and the location of the most significant are shown on the genetic map in Fig. 1. The detection of QTLs was carried out for each location and for the mean over the six locations.

Table 1 Value of grain-protein content (GPC), yield and thousand-kernel weight (TKW) for the RILs population and the parental lines in the different locations`

Location[a]	GPC (%)				Yield (qx/ha)				TKW (g)			
	Range[a]	Mean[b]	Rn[c]	Rc[c]	Range	Mean	Rn	Rc	Range	Mean	Rn	Rc
CHAL	9.7–14.2	11.6	13.4	10.3	55–106	84	80	98	37–60	48	55	44
CF	9.7–13.1	11.2	12.3	10.2	49–80	70	68	73	29–44	36	42	31
CVIL	9.0–13.8	11.1	12.3	9.8	61–141	104	106	109	36–58	47	56	39
LM	9.8–14.2	11.8	12.9	10.3	54–85	71	69	79	39–59	48	56	44
MO	9.2–14.5	11.7	12.9	10.5	52–125	84	81	92	35–57	46	54	43
RN	10.8–16.2	12.9	13.8	11.6	59–110	94	99	98	40–62	49	No[d]	No

[a] CHAL: Châlons en Champagne, CF: Clermont-Ferrand, CVIL: Chartainvilliers, LM: Le Moulon, MO: Mons, RN: Rennes
[b] Range and mean in the RILs population
[c] Rn and Rc corresponding to the two parental lines, respectively, Renan and Récital
[d] No: not obtained

Table 2 Chromosomal location of QTLs affecting grain-protein content (GPC %), yield (qx/ha) and thousand-kernel weight (TKW)g

Trait	Chromo-some	Trial[a]	R^{2b}	Pc[c]	Location on chromosome[d]	Additive value[e]	
GPC	1A	CF **CHAL**	4.3–4.6	0.59	39–94–144	0.13–0.29–0.50	Rn
	2A	CHAL **CVIL** LM MO Mn	4.4–8.9	0.61	20–45–55	0.16–0.32–0.45	Rn
	3A	**CF** CVIL LM MO RN Mn	4.1–8.3	0.85	10–21–33	0.11–0.21–0.32	Rn
	3B	**CVIL** LM	4.3–5.3	0.60	19–33–175	0.18–0.32–0.49	Rn
	4A	**RN** Mn	4.9–5.5	(a)		0.18	Rn
	4D	CF **CHAL** CVIL LM MO Mn	4.6–10.3	(a)		0.27	Rn
	5B	CHAL	4.6	0.83	27–43–128	0.14–0.25–0.39	Rc
	6A	**LM**	4.2	(a)		0.15	Rn
	7A	LM **RN**	4.5–5.3	0.60	68–110–189	0.12–0.27–0.44	Rn
	7Db[(b)]	CF CHAL CVIL **RN** Mn	6.4–10.4	0.77	49–61–63	0.17–0.28–0.43	Rn
Yield	2B	**CF**	5.6	0.82	68–158–162	0.9–1.5–2.3	Rc
	3B	CVIL **LM** Mn	4.4–6.5	0.84	–19–58–92	1.8–3.0–4.3	Rc
	4A	CHAL **Mn**	4.0–5.4	(a)		2.2	Rc
	4B	**CHAL** MO Mn	3.9–6.1	0.79	3–18–31	1.7–2.8–4.4	Rc
	5A	**CHAL** LM	4.3–5.2	0.69	85–114–152	2.1–3.3–4.9	Rc
	5B	CF **CHAL** Mn	3.9–6.8	0.69	35–38–141	1.3–2.5–3.9	Rn
	7Da[(b)]	**CHAL** MO Mn	4.4–6.3	(a)		2.5	Rc
	7Db[(b)]	**CHAL** CVIL MO Mn	3.7–15.7	0.94	46–58–71	2.9–4.8–7.1	Rn
	7Db[(b)]	RN	5.7	0.41	54–64–75	0.6–2.5–3.9	Rc
TKW	1X[(c)]	CF **CVIL** LM Mn	5.1–6.1	(a)		1.1	Rn
	1D	**LM** Mn	6.3–8.7	0.76	51–76–89	0.7–1.3–1.9	Rn
	2B	CF CHAL CVIL LM MO **RN** Mn	10.7–19.7	0.99	68–72–90	1.2–1.9–2.8	Rc
	2D	CF CHAL MO **RN**	4.7–6.8	0.65	25–37–72	0.8–1.4–2.2	Rn
	3A	RN	4.8	0.79	24–39–81	0.5–1.1–1.7	Rn
	5B	CF CHAL CVIL LM **MO** RN Mn	4.9–10.4	0.76	63–130–138	0.6–1.2–2.0	Rn
	6A	**CF** CHAL LM Mn	5.2–6.7	0.52	–6–19–36	0.6–1.1–1.6	Rn
	6D	RN **CVIL**	5.4–7.5	0.92	86–101–116	0.9–1.6–2.5	Rn
	7A	CF CHAL CVIL LM **MO** RN Mn	5.2–10.3	0.89	62–140–143	1.4–2.2–3.1	Rn
	7Db[(b)]	CHAL CVIL	4.8–7.5	0.49	42–62–71	0.6–1.4–2.2	Rn

[a] trial: CF = Clermont-Ferrand, CHAL = Châlons-en Champagne, CVIL = Chartainvilliers, LM = Le Moulon, MO = Mons, RN = Rennes, Mn =Mean; in bold, the trial for which the results of bootstrapping are given
[b] R^2: range of the determination coefficient calculated for the closest marker of the QTL on the chromosome determined by ANOVA ($P < 0.01$) in the different trials
[c] Pc: percentage of significant models using bootstrap resampling
[d] Location on chromosome: confidence interval and estimate of the position of the QTL on the chromosome in centimorgans determined by bootstrapping
[e] Additive value: confidence interval and estimate of the additive value determined by bootstrapping, and indicate the parent contributing to a higher-value allele, where Rn = Renan and Rc = Récital. (a): the linkage group is too small to allow 'regression marker', (b): the chromosome 7D is some linkage groups, (c): this linkage group is on a chromosome of group 1, but not determined

The trials where each QTL was detected are given, with the range of r^2 over sites, while chromosome location and additive effect, with their respective confidence intervals, are given only for the site where the QTL had the highest r^2.

Ten chromosomal regions were detected for GPC in our population, with individual r^2 ranging from 4.2 to 10.4%, while eight were detected for grain yield (r^2 from 3.9 to 15.4) and nine for TKW (from 4.7 to 19.7).

Table 3 Analysis of variance for grain-protein content (GPC), yield and thousand-kernel weight (TKW) on the parental lines

Item	GPC			Yield			TKW		
	SS	*df*	*F*-value	SS	*df*	*F*-value	SS	*Df*	*F*-value
Location	41.16	5	40.50***	34206.69	5	328.03***	4,087.43	4	412.52***
Replication	6.49	5	6.38***	1,417.10	5	13.08***	5.69	5	0.46 ns
Block (replication × location)	32.35	94	1.69**	2,756.30	94	1.41 ns	343.90	85	1.63 ns
Genotype	291.26	1	1,432.77***	2,557.76	1	122.64***	7,094.24	1	2,863.92***
Genotype × location	5.74	5	5.65***	1,557.37	5	14.93***	175.08	4	17.67***
Error	19.72	97		2,064.74	99		217.99	88	
R-square of model	0.95			0.95			0.98		

** and ***, significant respectively at the 0.01 and 0.001 probability level
ns, not significant

Table 4 Model for G×E interaction for grain-protein content in the population Renan × Récital

Model	Source	Sum of squares	*df*	Mean square	Prob > *F*
AMMI	First term	54.55	148	0.37	0.0002
	Second term	44.10	146	0.30	0.0148
	Residual	96.71	426	0.23	
Factorial regression	Tm	39.27	144	0.27	0.32
	With Xgwm469	2.59	1	2.59	0.002
	With Xgwm156	0.78	1	0.78	0.077
	With Xgwm257	1.84	1	1.84	0.0077
	NbD25	34.55	144	0.24	0.66
	With Xgwm469	0.52	1	0.52	0.15
	With Xgwm156	1.54	1	1.54	0.01
	With Xgwm257	0.12	1	0.12	ns
	Multiplicative term	47.13	143	0.33	0.036
	Residual	71.49	280	0.26	

For GPC, four QTLs were detected in at least four of the six locations. These "stable" QTLs were located on chromosomes 2A, 3A, 4D and 7D, with 'Renan' having the favourable alleles. On chromosome 4D, the QTL was associated with an unlinked marker. Thus it was not possible to determine precisely the effect of this QTL and it could be underestimated. Unfortunately, it has been difficult to find polymorphic markers on chromosome 4D for our population. When combined in a multiple regression model, the different QTLs altogether explained between 12 and 30% of the phenotypic variation of grain-protein content, depending on the location considered.

Of the nine QTLs detected for TKW, three were significant for all six environments and the mean (Table 2). They were located on chromosomes 2B, 5B and 7A. The strongest QTL on chromosome 2B explained up to 20% of the variation of the trait and the positive allele for this QTL was from Récital, the parental line with the lowest TKW. For the other QTLs, the favourable allele came from Renan. On chromosomes 5B and 7A, the two-QTLs model was more significant than the one-QTL model for some locations. Thus, these two chromosomes seemed to carry two different regions influencing TKW. The part of phenotypic variation explained by these different QTLs in a multiple regression model ranged from 30.5% at Le Moulon to 35% at Chartainvilliers.

No consistent QTL was detected for yield in our population, except on chromosome 7D, with a QTL explaining nearly 15% of the phenotypic variation. On this chromosome, the favourable parent depended on the location: at Rennes, the favourable effect on yield was due to the 'Récital' allele, while at Châlons-en Champagne, Chartainvilliers and for the mean, the effect was due to the allele from 'Renan' (Table 2 and Fig. 1). The part of the variation explained by the QTLs in the multiple regression models ranged from 6% at Rennes to 30% at Châlons en Champagne.

Some co-location of QTLs occurred for the different traits on chromosomes 2B, 3A, 5B, 7A and 7D (Fig. 1). In most cases, the QTL was 'stable' only for one of the traits. On chromosomes 5B and 7D, QTLs were detected for the three traits. On chromosome 5B, the favourable parent for GPC was Récital *versus* Renan for TKW and yield. On chromosome 7D, the favourable allele was from Renan for the three traits, except for the yield at Rennes.

G×E interaction analyses

Analyses of variance using the balanced subset of parental lines data indicated that location, genotype and G×E interaction effects were highly significant for all traits (Table 3). Indeed, only two genotypes were used in this analysis. However, we could expect a similar result for the whole data set. Moreover, using a balanced subset of data for 40 lines replicated twice in three locations, we

Table 5 Model for G×E interaction for yield in the population Renan × Récital

Model	Source	Sum of squares	*df*	Mean square	Prob > *F*
AMMI	First term	20,460.91	181	113.04	$<1 \times 10^{-5}$
	Second term	14,802.04	179	82.69	$<1 \times 10^{-5}$
	Residual	17,313.46	525	32.98	
Factorial regression	Tm	9,405.16	177	53.14	0.015
	With Pch1	1,089.71	1	1,089.71	0.0001
	With Xfba285	420.49	1	420.49	0.0017
	T25	12,977.90	177	73.32	$<1 \times 10^{-5}$
	With Pch1	1,710.02	1	1,710.02	$<1 \times 10^{-5}$
	With Xfba285	55.70	1	55.70	0.24
	Multiplicative term	14,972.55	177	84.59	$<1 \times 10^{-5}$
	Residual	14,007.40	348	40.25	

Table 6 Model for G×E interaction for thousand kernel weight in the population Renan × Récital

Model	Source	Sum of squares	*df*	Mean square	Prob > *F*
AMMI	First term	919.98	125	7.36	$<1 \times 10^{-5}$
	Second term	786.46	123	6.39	$<1 \times 10^{-5}$
	Residual	1,201.09	357	3.36	
Factorial regression	Tm	644.54	121	5.33	0.017
	With Xcfd81	64.24	1	64.24	0.0002
	With Xgwm257	55.27	1	55.27	0.0004
	With Xgwm639	18.72	1	18.72	0.027
	Multiplicative term	833.00	121	6.88	0.0001
	Residual	1,378.48	351	3.93	

verified that the error variances obtained were of same magnitude than those from the AMMI model (data not shown), justifying their use in *F* statistics.

Results of ANOVA using AMMI or Factorial Regression models are given in Tables 4, 5 and 6, according to the trait. In AMMI models, the two multiplicative terms were highly significant ($P < 0.001$) for yield and TKW, but only the first one for GPC. In the Factorial Regression model, one covariate (Tm) was retained in the model for TKW, and two covariates, Tm and T25, for yield. Three markers, *Xcfd81*, *Xgwm257* and *Xgwm639*, linked to the genotypic regression coefficient on Tm were found to give significant cross products with Tm for explaining G×E interaction for TKW. These markers are located on 2B, 5B and 5D (Fig. 1). Similarly, two markers, *Pch1* (resistance gene) on chromosome 7D and *Xfba285* on chromosome 1A, gave significant cross products with both Tm and T25 for modelling G×E for yield. For GPC, none of the climatic covariates used alone explained a significant part of G×E; however, several cross products with markers *Xgwm257*, *Xgwm156* and *Xgwm469* appeared to be highly significant. These markers are located on chromosomes 2B, 5A and 6B, respectively (Fig. 1).

Discussion

Detection of 'stable' QTLs

Among the four 'stable' QTLs detected for GPC, none co-located with any of the storage protein loci, located on chromosomes group 1 and 6, or with previously published QTLs for GPC using other populations (Blanco et al. 1996; Joppa et al. 1997; Prasad et al. 1999; Perretant et al. 2000; Zanetti et al. 2001). For example, Zanetti et al. (2001) found a quite strong QTL ($r^2 = 25\%$) on chromosome 5A using a bread-wheat × spelt population. On this chromosome, we found only a small and inconsistent QTL, with a very large confidence interval, in two locations. Some 'stable' QTLs detected in our population could correspond to homoeologous QTLs reported in previous studies. We detected a QTL on 3AS, which could be due to a gene homoeologous with one explaining the QTL detected by Zanetti et al. (2001) on 3BS. The strongest of our 'stable' QTLs for GPC was located on chromosome 7D, close to the *Pch1* locus for resistance to eyespot. This locus is located in an alien fragment inherited from *Aegilops ventricosa* through interspecific hybridization (Maia et al. 1967). Thus, this higher GPC may originate from an *Ae. ventricosa* allele, although the length of the introgressed fragment is unknown.

The QTLs detected for TKW appeared different from those detected in previous studies on wheat (Campbell et al. 1999; Varshney et al. 2000; Zanetti et al. 2001), even when we compared homoeologous chromosomes. The QTL on chromosome 2B co-located with a QTL for heading date detected in the same population (data not shown). Thus, we can not reject the hypothesis that this QTL for TKW resulted from an indirect effect of earliness, from Récital, and more-favorable weather conditions during grain-filling for early lines. A gene for re-

sponse to photoperiod was reported on this chromosome (*Ppd-B1* Welsh et al. 1973).

For yield, only one QTL can be considered as 'stable'. This QTL, detected in four locations, was located on chromosome 7DL, close to *Pch1*. The effect of this QTL depends on the trial location. In most cases where this QTL was detected, the favorable allele came from Renan, whereas at Rennes it came from Récital. Because the field trials were made under normal conditions with fungicide treatment, it may be postulated that this QTL did not result from a direct effect of *Pch1*. Moreover, few or no eyespot symptoms were observed in the trials. This suggests that other important genes influencing grain yield may have been introgressed. Our results contradict those of Worland et al. (1990), who found high yield losses associated with the introgression of *Ae. ventricosa* segments, including the eyespot resistance gene.

Co-locations between QTLs for GPC, yield and yield components

In our population, there was a significant negative correlation between GPC and yield, in agreement with many previous studies (for example, Löffler et al. 1983; Cox et al. 1985; Gauer et al. 1992). It is important to consider co-locations between QTLs for the different traits, to determine whether this negative relationship resulted from opposite pleïotropic gene effects. Such a result would make the simultaneous improvement of these two traits difficult or impossible.

Such colocations occurred on chromosome 3A (GPC-TKW), 5B (GPC-TKW-Y), 7A (GPC-TKW) and 7D (GPC-TKW-Y). Surprisingly, in most cases, the favourable alleles for GPC and yield or TKW came from the same parental line, thus contradicting the hypothesis that QTLs for GPC are indirect QTLs for low yield. On two chromosomes, the effects on the different traits were opposite. On chromosome 5B, the Renan allele increased yield and TKW while it decreased GPC. However, this latter effect was low and was only detected in one location. The case of chromosome 7D is more complicated. The Renan allele had favourable effects for all traits in most locations, except for the yield at Rennes. On this chromosome, it appears that the favourable effect of the Renan allele for GPC was not the result of a negative effect for yield due to a gene in the introgression from *Ae. ventricosa*. Clearly this QTL was associated with the G×E interaction for yield, and this will be discussed below.

To our knowledge, no previous study determined the genetic control of GPC and grain yield in the same population. For their cross between wheat and spelta, Zanetti et al. (2001) determined QTLs for GPC and TKW. They found both positive and negative genetic relations, depending on the QTL. However yield was not evaluated.

Although the relationship between yield and TKW was not close, some co-locations between QTLs for yield and TKW were found, on chromosomes 2B, 5B and 7D. In most cases, QTLs for yield had weak effects and were not detected in most locations. Our results suggest that QTLs for TKW, which are more-precisely detected, could be used in breeding programs for yield improvement, because in most cases these QTLs for TKW collocated with QTLs for yield.

QTLs for G×E interactions

Crossa et al. (1999) first suggested the use of molecular markers in Factorial Regression (FR) together with environmental covariates as a means of modelling and interpreting G×E interactions. In a dataset from maize yield trials, they found more than 30 markers to be significant in individual FR and 13 significant cross products with nine environmental covariates. In our study, we found very few markers to be significant when used alone in FR, with the exception of *Pch1* for yield. Therefore, we used an alternative approach.

The results of AMMI model analyses (Table 4, 5 and 6) of our population data showed that G×E interaction was less structured for GPC than for the other traits, because only the first term was highly significant. This could explain why no environmental covariate was found to be significant in FR for GPC. However, three significant ($P < 0.01$) cross-products were found for GPC (Table 4), two with the environmental covariate Tm, the average daily temperature during grain filling, and one with NbD25, the number of days with a maximum daily temperature above 25 °C. Indeed, high temperatures during grain filling are known to affect the relative rates of accumulation of carbon and nitrogen in grain, and therefore GPC (Grayboch et al. 1996; Daniel and Triboï 2000). Perhaps because temperatures were not sufficiently different between the locations of our study, we did not detect any direct effect of these covariates on G×E interaction for GPC, but we detected genetic regions that could be related to differential responses of genotypes to temperature during grain filling. These regions were located on chromosomes 2D, 5A and 6D, which showed no QTL for GPC in any location.

For TKW, the effect of mean temperature (Tm) was found to be significant ($P < 0.05$) in the FR model, suggesting that accumulation of carbon in the grain is more influenced by temperature than that of nitrogen. Three cross-products were found significant with Tm. The genetic regions accounting for a part of the G×E interaction were located on 2B, 5B and 5D. On chromosome 5B, the marker detected was located in the confidence interval for a 'stable' QTL for TKW, detected in all locations. It was not possible to determine whether these two genetic effects resulted from the same gene or whether this region contains different genes involved in the control of either the main effect or the interactive part of the trait. On chromosome 2B, the marker detected was close to the QTL for TKW with the strongest effect, although it was not included in the confidence interval. As previously mentioned, because of the presence of a QTL for

heading date on the same chromosome, the question remained whether earliness had an indirect effect on the QTL for TKW, on G×E interaction, on both or on neither of these. The last marker was located on chromosome 5D, where no QTL was detected.

Two environmental covariates, Tm and T25, were found to be significant for grain yield. Two markers gave significant cross products with Tm and one with T25. The two markers were *Pch1* on chromosome 7D and *Xfba285* on chromosome 1A. Obviously, yield is the product of two components, the number of grains per m^2 determined before flowering, and the kernel weight, mostly determined after flowering. This was not a comprehensive study of G×E interaction for yield, because only covariates associated with environmental conditions during grain filling were considered. However the genomic regions that interact with Tm are different for TKW and yield.

No co-location was observed between interacting genetic regions for the three traits. Much work is still needed to obtain better understanding of G×E interactions. These results have identified some candidate genomic regions that are likely to be involved.

Conclusion

In this study, we have reported some strong and stable QTLs for grain protein-content and yield, and one of its components, TKW. No strong negative pleïotropic effect was detected for grain protein-content and yield, suggesting that it will be possible to use these results for the improvement of these two economically important traits in the same breeding scheme. QTLs for TKW could thus be used efficiently in yield breeding because of the existence of co-location between QTLs controlling the two traits. Finally in a first analysis of G×E interactions, some specific regions involved in wheat for GPC and yield and interacting with temperature conditions during grain filling, were identified.

Acknowledgements This research was supported by the French ministers of Research and of Agriculture and by private companies in the framework of a national plant genomics program, 'Génoplante'. The authors gratefully acknowledge M. Trottet for providing the studied RILs population; the technical staff in the different experimental locations for the field trials; the technical staff of the society ULICE for providing the technological data; N. Marty, M.R. Perretant, V. Bodusseau, M. Bilous, S. Négre, B. Charef and L. Gervais for their contributions to the molecular mapping. They also thank F. Vear, P. Sourdille and the two anonymous referees for their helpful comments on the manuscript. The experiments comply with current laws of France where the experiments were performed.

References

Aastveit H, Martens H (1986) ANOVA interactions interpreted by partial least squares regression. Biometrics 42:829–844

Araki E, Miura H, Sawada S (1999) Identification of genetic loci affecting amylose content and agronomic traits on chromosome 4A of wheat. Theor Appl genet 98:977–984

American Association of Cereals Chemists (1995) Approved methods of the American Association of Cereal Chemists, 9th edn. The American Association of Cereals Chemists, Saint Paul, Minnesota

Baril C (1992) Factor regression for interpreting genotype-environment interaction in bread-wheat trials. Theor Appl Genet 83:1022–1026

Bhullar SS, Jenner CF (1985) Differential responses to high temperatures of starch and nitrogen accumulation in the grain of four cultivars of wheat. Aust J Plant Physiol 12:363–375

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

Campbell KG, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft × hard wheat cross. Crop Sci 39:1184–1195

Cox M, Qualset C, Rains D (1985) Genetic variation for nitrogen assimilation and translocation in wheat. I. Dry matter and nitrogen accumulation. Crop Sci 25:430–435

Crossa J, Vargas M, van Eeuwijk FA, Jiang C, Edmeades GO, Hoisington D (1999) Interpreting genotype × environment interaction on tropical maize using linked molecular markers and environmental covariables. Theor Appl Genet 99:611–625

Daniel C, Triboï E (2000) Effects of temperature and nitrogen nutrition on the grain composition of winter wheat: effects on gliadin content and composition. J Cereal Sci 32:45–56

Decoux G, Denis J-B (1991) INTERA Logiciels pour l'interprétation statistique de l'interaction entre deux facteurs. Laboratoire de Biométrie, Versailles, INRA

Denis J-B (1980) Analyse de régression factorielle. Biométrie-Praximétrie 20:1–34

Denis J-B (1988) Two-way analysis using covariates. Statistics 19:123–132

Gate P (1995) Ecophysiologie du blé, de la plante à la culture. Technique et Documentation, Lavoisier. Paris, France

Gauch HG (1988) Model selection and validation for yield trials with interaction. Biometrics 44:705–715

Gauch HG (1992) Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elvesier, Amsterdam

Gauer L, Grant C, Gehl D, Bailey L (1992) Effects of nitrogen fertilization on grain protein content, nitrogen uptake, and nitrogen use efficiency of six spring wheat (*Triticum aestivum* L.) cultivars, in relation to estimated moisture supply. Can J Plant Sci 72:235–241

Graybosch R, Peterson C, Shelton D, Baenziger P (1996) Genotypic and environmental modification of wheat flour protein composition in relation to end-use quality. Crop Sci 36:296–300

Groos C, Gay G, Perretant MR, Gervais L, Bernard M, Dedryver F, Charmet G (2002) Study of the relationship between preharvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread wheat cross. Theor Appl Genet 104:39–47

Guyomarc'h H (2000) Développement de marqueurs microsatellites à partir de *Triticum tauschii*. Application à l'évolution des génomes du genre *Triticum* et à la recherche de QTL de qualité technologique en panification chez le blé tendre (*Triticum aestivum* L.). Thèse de l'Université Blaise Pascal

Hyne G, Snape JW (1991) Mapping quantitative trait loci for yield in wheat. In: Pesek J, Herman M, Hartmann J (eds) Biometrics in plant breeding. Proc 8th Meeting Eucarpia Section Biometrics on Plant Breeding. Brno, Czechoslovakia, pp 47–56

Hyne V, Kearsey MJ (1995) QTL analysis: further uses of 'marker regression'. Theor Appl Genet 91:471–476

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

Kearsey MJ, Hyne V (1994) QTL analysis: a simple 'marker-regression' approach. Theor Appl Genet 89:698–702

Löffler C, Bush R, Wiersma J (1983) Recurrent selection for grain protein percentage in hard red spring wheat. Crop Sci 23:1097–1101

Maia N (1967) Obtention de blé tendre résistants au piétin-verse (*Cercosporella herpotrichoides*) par croisements interspécifiques. C. R. Acad Agric France 52:149–154

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, Varshney RK, Kumar A, Balyan HS, Sharma PC, Edwards KJ, Singh H, Dhaliwal HS, Roy JK, Gupta PK (1999) A microsatellite marker associated with a QTL for grain protein content on chromosome arm 2DL of bread wheat. Theor Appl Genet 99:341–345

Robert N, Hennequet C, Bérard P (2001) Dry matter and nitrogen accumulation in wheat kernel: genetic variation in rate and duration of grain filling. J Genet Breed 55:297–306

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

SAS Institute Inc (1991) SAS/STAT^{R} user's guide: release 6.03 edn. SAS Institute, Cary, North Carolina

Shah MM, Gill KS, Baesinger PS, Yen PS, Kaeppler SM, Ariyarathne HM (1999) Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. Crop Sci 39:1728–1732

Vargas M, Crossa J, van Eeuwijk FA, Ramirez ME, Sayre K (1999) Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype × environment interaction. Crop Sci 39:955–967

Varshney RK, Prasad M, Roy JK, Kumar N, Harjit-Singh, Dhaliwal HS, Balyan HS, Gupta PK (2000) Identification of eight chromosomes and a microsatellite marker on 1AS associated with QTLs for grain weight in bread wheat. Theor Appl Genet 100:1290–1294

Visscher PM, Thompson R, Haley CS (1996) Confidence intervals in QTL mapping by bootstrapping. Genetics 143:1013–1020

Wardlaw IF, Wrigley CW (1994) Heat tolerance in temperate cereals: an overview. Aust J Plant Physiol 21:695–703

Welsh JR, Keim DL, Pirasteh B, Richards RD (1973) Genetic control of photoperiod response in wheat. In: Sears ER, Sears LMS (eds) Proc 4th Int Wheat Genet Symposium, Colombia, Missouri, USA, pp 879–884

Worland AJ, Law CN, Li WM (1990) The location of genes depressing yield associated with the transfer of eyespot resistance from *Aegilops ventricosa*. Annu Rep, Inst Plant Sci Res:7–8

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