Theor Appl Genet (2013) 126:805–822
DOI 10.1007/s00122-012-2019-z

ORIGINAL PAPER

# Genomic regions associated with the nitrogen limitation response revealed in a global wheat core collection

**Jacques Bordes · C. Ravel · J. P. Jaubertie · B. Duperrier · O. Gardet · E. Heumez · A. L. Pissavy · G. Charmet · J. Le Gouis · F. Balfourier**

Received: 9 December 2011 / Accepted: 5 November 2012 / Published online: 29 November 2012


**Abstract** Modern wheat (*Triticum aestivum* L.) varieties in Western Europe have mainly been bred, and selected in conditions where high levels of nitrogen-rich fertilizer are applied. However, high input crop management has greatly increased the risk of nitrates leaching into groundwater with negative impacts on the environment. To investigate wheat nitrogen tolerance characteristics that could be adapted to low input crop management, we supplied 196 accessions of a wheat core collection of old and modern cultivars with high or moderate amounts of nitrogen fertilizer in an experimental network consisting of three sites and 2 years. The main breeding traits were assessed including grain yield and grain protein content. The response to nitrogen level was estimated for grain yield and grain number per $m^2$ using both the difference and the ratio between performance at the two input levels and the slope of joint regression. A large variability was observed for all the traits studied and the response to nitrogen level. Whole genome association mapping was carried out using 899 molecular markers taking into account the five ancestral group structure of the collection. We identified 54 main regions involving almost all chromosomes that influence yield and its components, plant height, heading date and grain protein concentration. Twenty-three regions, including several genes, spread over 16 chromosomes were involved in the response to nitrogen level. These chromosomal regions may be good candidates to be used in breeding programs to improve the performance of wheat varieties at moderate nitrogen input levels.

Communicated by A. Graner.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-012-2019-z) contains supplementary material, which is available to authorized users.

J. Bordes (✉) · C. Ravel · G. Charmet · J. Le Gouis · F. Balfourier
INRA, UMR 1095 Génétique, Diversité Et Ecophysiologie Des Céréales, 234 Avenue Du Brézet, 63100 Clermont-Ferrand, France
e-mail: jbordes@clermont.inra.fr

J. Bordes · C. Ravel · G. Charmet · J. Le Gouis · F. Balfourier
Université Blaise Pascal, UMR 1095 Génétique, Diversité Et Ecophysiologie Des Céréales, 63177 Aubière Cedex, France

J. P. Jaubertie · A. L. Pissavy
Agri Obtentions Chemin de La Petite, Minière, Guyancourt 78280, France

B. Duperrier
Limagrain Europe, BP3, 77390 Verneuil l'Etang, France

O. Gardet
INRA-UPS-CNRS-AgroParisTech, UMR de Génétique Végétale, Ferme Du Moulon, 91190 Gif-sur-Yvette, France

E. Heumez
INRA, UMR 1281 Stress Abiotiques Et Différenciation Des Végétaux Cultivés, Estrées-Mons, 80203 Péronne, France

## Introduction

Since the Green Revolution, breeding of modern wheat (*Triticum aestivum* L.) varieties in Western Europe has been mainly carried out using an optimal input management system for crop growth which includes applying high levels of nitrogen (N) fertilizer (Hirel et al. 2007). The use of N-rich fertilizers has greatly increased the quantity of nitrates leaching into groundwater, the emission of nitrous oxide derived from denitrification, and the consumption of fossil fuels necessary for their manufacture and application. The extent of these negative environmental impacts could

be reduced using less mineral N fertilizer and selecting crop varieties that are adapted to low input management systems. Such plants must be specifically N stress tolerant, i.e., they have to maintain good levels of both yield and grain protein content under moderate N deficiency (Witcombe et al. 2008). Although alleles specifically adapted to moderate input conditions may exist, it is unlikely they have been retained in most West European breeding programs, although they may have been maintained as genetic resources in germplasm banks.

Association genetics could be used to search for, in genetic resource collections, genes or QTL involved in the response to N level (RNL), i.e., alleles that limit the loss of productivity in conditions of moderate N availability. Contrasting genetic responses to N deficiency have often been reported either between varieties (e.g., Austin et al. 1977; Le Gouis et al. 2000; Barraclough et al. 2010; Gaju et al. 2011) or within mapping populations (An et al. 2006; Laperche et al. 2007, 2008; García-Suárez et al. 2010). The variability of the response can be quantified in different ways, for instance, by linear regression of each variety's performance on the mean environment grain yield as described by Finlay and Wilkinson (1963) or by global interaction variables such as the ratio and difference of values between N levels as used by Laperche et al. (2007).

The RNL value per se of a plant depends on its N use efficiency (NUE) and involves both N uptake and N utilization. N uptake efficiency is a genotype's ability to extract N from the soil and mainly depends on root attributes. N utilization is essentially how much grain a genotype can produce per unit of N taken up and this depends on both N assimilation and remobilization (Moll et al. 1982). Numerous traits summarized by Foulkes et al. (2009) have been identified as influencing NUE: root length density, stem capacity for N accumulation, low leaf lamina N concentration, post-anthesis remobilization of N from stems to grain, remobilization of N from leaves to grain, delayed senescence, and grain N concentration. Grain yield itself can be considered to be the sum total of the effects of all these traits, as observed by Barraclough et al. (2010) who found that it explained 77 % of the variation in NUE. Many key enzymes involved in N metabolism have been identified, e.g., nitrate reductase (NR), nitrite reductase (NiR), glutamine synthetase (GS), glutamate synthase (GOGAT), glutamate dehydrogenase (GDH), sucrose phosphate synthase (SPS), etc. Many genes or QTL that determine NUE have been mapped or positioned on genetic maps (An et al. 2006; Laperche et al. 2007; Habash et al. 2007; Fontaine et al. 2009) using quantitative traits.

To explore the diversity of wheat accessions conserved in the Clermont-Ferrand Genetic Resources Center, a core collection of 372 accessions (372CC) was selected based on passport and simple sequence repeat (SSR) marker data (Balfourier et al. 2007). The wide phenotypic variation of this core collection (Bordes et al. 2008) has been used to identify genetic associations for several quality traits (Bordes et al. 2011) and components of earliness (Bonnin et al. 2008; Rousset et al. 2011; Le Gouis et al. 2011) using molecular markers based on the diversity array technology (DArT), microsatellites (SSRs) and single nucleotide polymorphisms (SNPs). This core collection is therefore potentially a good resource for locating loci involved in the main agronomic traits and their response to N input levels. In bread wheat, several association mapping studies have highlighted loci acting on agronomic traits (Breseghello and Sorrells 2006; Roy et al. 2006; Crossa et al. 2007; Neuman et al. 2011) or quality traits (Ravel et al. 2009), but not to our knowledge responses to input level.

The objective of this work was to identify favorable alleles associated with agronomic traits, including grain yield (GY) and grain number per $m^2$ (GN), under moderate N levels from a subset of accessions selected from the global core collection of 372 accessions, by evaluating parameters during 2 years at three locations with two N input levels.

## Materials and methods

### Plant materials

A sample of 196 bread wheat accessions (196CC) from the INRA bread wheat core collection of 372 accessions (372CC) set up by Balfourier et al. (2007) was used. The accessions were sub-sampled based on field evaluation data obtained by Bordes et al. (2008) and passport data (geographic origin and registration date). To favor good grain production and limit plant lodging, the 196 accessions were chosen among the shortest accessions (50–130 cm tall) having a good grain weight (250–1,000 g $m^{-2}$). The accessions chosen originate from 38 different countries. About half are modern cultivars bred after 1960, the others are older accessions including landraces and cultivars from the nineteenth century.

### Experimental design

The experiment was conducted in 2007 and 2008 at three different locations in France: Joze (45.86′N, 3.30′E), Le Moulon (48.42′N, 2.08′E) and Mons (49.53′N, 3.00′E). Two input levels, called here high and moderate, were applied at each location (Table 1). The high input system corresponded to the usual farming practises used at each location for high yield objectives, which included dense planting and applying a high level

**Table 1** Description of experiments conducted at Mons, Le Moulon and Joze (France) for the 2006–2007 (2007) and 2007–2008 (2008) growing seasons on a collection of 196 bread wheats grown at two input levels (moderate = LN and high = HN)

| | Mons | | | | Le Moulon | | | | Joze | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2007 | | 2008 | | 2007 | | 2008 | | 2007 | | 2008 | |
| | LN | HN | LN | HN | LN | HN | LN | HN | LN | HN | LN | HN |
| Inorganic soil N (kg $ha^{-1}$) | 43 | 43 | 58 | 58 | 68 | 68 | 69 | 69 | 100 | 100 | 70 | 70 |
| N applied | | | | | | | | | | | | |
| Tillering | | 33 | | 40 | | 33 | | 40 | | 35 | | 50 |
| Jointing | 40 | 60 | 40 | 80 | 40 | 60 | 40 | 80 | 40 | | 50 | |
| Booting | 40 | 60 | | 40 | 40 | 60 | | 40 | | 40 | | 40 |
| Flowering | | | | | | | | | 50 | 50 | 40 | 40 |
| Total N supply | 123 | 196 | 98 | 218 | 148 | 221 | 109 | 229 | 190 | 225 | 160 | 200 |
| Difference HN–LN | 73 | | 120 | | 73 | | 120 | | 35 | | 40 | |
| Seeding density $m^{-2}$ | 220 | 220 | 220 | 220 | 220 | 220 | 220 | 220 | 280 | 170 | 280 | 170 |
| Fungicide treatments | No | Yes | Yes | Yes | No | Yes | Yes | Yes | No | Yes | Yes | Yes |
| Growth regulator | No | Yes | Yes | Yes | No | Yes | Yes | Yes | No | Yes | Yes | Yes |

Inorganic soil N was determined in February (growth stage 31) in the top soil layer (0–0.9 m)

of N fertilizer (HN) with appropriate pesticides. For the moderate input system the concentration of N applied with fungicide treatments was lowered by 35–120 kg N $ha^{-1}$ (LN). Seeding density was 40 % lower for both years at Joze only. Application of N was mainly reduced at the early tillering stage in 2007 and at the early tillering and booting stages in 2008. For HN in 2007, the growth regulator chloroethyltrimethylammonium chloride was applied at 9.6 kg a.i. $ha^{-1}$ to avoid lodging, with fungicide treatments. In 2008, the applications of growth regulator and fungicides were identical for LN and HN, so only N availability differed.

In each site-season (year × location combinations), two blocks were grown, one with HN and one with LN. Each block contained 234 plots (196 accessions plus height standard cultivars, six of which were repeated six times) sized 7.5 $m^2$. This experimental design was analyzed as a split-plot design, with nitrogen as the whole-plot factor, genotype as the split-plot factor and the six site-seasons as repetitions. To take into account possible heterogeneity of the soil, each block was divided into six sub-blocks. The 196 accessions plus the two cultivars Farandole and Recital were allocated to these six sub-blocks according to their height to limit plant competition. The six standard cultivars Apache, Caphorn, Soissons, Oratorio, Koreli and Renan were then added in each sub-block. Within a sub-block, the cultivars were randomized differently for each site-season-N level combination. Grain yield (GY, t $ha^{-1}$ at 0 % humidity), grain number (GN, $m^2$), thousand kernel weight (TKW, g), plant height (PH, cm), heading date (HD, days from the 1st January) and grain protein concentration (GPC, %) were determined for each plot.

## Statistical analysis

To control for intra-sub-block heterogeneity for each site-season and N level, subplot trait values were adjusted relative to the six cultivars repeated in each sub-block using the glm and lsmeans procedures in Statistical Analysis Software (SAS Institute Inc 1999) using the model:

$$Y_{ij} = \mu + G_i + B_j + e_{ij},$$

where $Y_{ij}$ represents the value of the trait under investigation for genotype $i$ in sub-block $j$, $\mu$ represents the general mean, $G$ represents the fixed genotypic effect, $B$ is the fixed sub-block effect, and $e_{ij}$ is the error term of the model.

Analysis of variance was then carried out with Statistica10 (StatSoft Inc, Tulsa, USA) for each trait using the subplot adjusted values and the following mixed model:

$$Y_{ikl} = \mu + G_i + \mathrm{Site}_k + N_l + G \times \mathrm{Site}_{ik} + G \times N_{il} + \mathrm{Site} \times N_{kl} + e_{ikl},$$

where $Y_{ikl}$ represents the value of the trait under investigation for genotype $i$ at site $k$ and $N$ level $l$, $G$ represents the random genotypic effect, $N$ the fixed nitrogen level effect, Site the fixed site-season effect, i.e., the combination of year and location, and $e_{ikl}$ is the error term of the model.

To take into account the split-plot structure, the Site and $N$ effects were both tested against the Site × $N$ interaction effect, the $G$ effect was tested against the $G \times$ Site

interaction effect, and the $G \times N$ interaction effect was tested against the $e_{ikl}$ error term.

The correlation coefficients between traits were computed independently for low and high N using the adjusted means of the first stage analysis averaged across the six site-season (2 years × 3 sites) trials.

### Variability of the response to N level

The RNL was computed in three ways: the difference between the value under HN and the value under LN (HN–LN), the ratio between HN and LN values (LN/HN), and the joint regression.

The difference is an estimate of the $G \times$ N interaction term, whereas the ratio approximates the deviation of the linear relationship between LN and HN values (Laperche et al. 2006). To estimate the variation of these two parameters, an analysis of variance was carried out with Statistica10 (StatSoft Inc, Tulsa, USA) for HN–LN and LN/HN using the following mixed model:

$$Y_{ik} = \mu + G_i + \text{Site}_k + e_{ik},$$

where $Y_{ik}$ represents the value of the trait under investigation for genotype $i$ and Site $k$, and $e_{ik}$ is the error term of the model.

The joint regression corresponds to the linear regressions of each variety's performance on the means of other varieties in the same environment (Finlay and Wilkinson 1963). Here, we used 12 environments, two per site-season (HN and LN), that were considered as independent to perform the regression analysis. The slope of this regression is then considered as representative of the sensitivity of a genotype to environmental factors varying in the network. We considered that the average yield of each environment is representative of N availability. The lower the slope coefficient, the more stable the genotype in response to moderate and high inputs. The most stable accessions are considered the most tolerant to different levels of input (Brancourt-Hulmel et al. 1994). The variability of this parameter, represented by the heterogeneity of the slopes, was estimated by the size of the genotype × environment interaction effect in an analysis of variance carried out with Statistica10 (StatSoft Inc, Tulsa, USA) using the following mixed model:

$$Y_{im} = \mu + G_i + Env_m + \beta_i \times \text{CovEnv}_m + e'_{im},$$

where $Y_{im}$ represents the value of the trait under investigation for genotype $i$ and environment $m$, $\beta$ represents the regression coefficient of genotype $i$ and $\text{CovEnv}_j$ represents the environmental covariate and $e'_{\text{im}}$ is the residual term of the model. Environment is the combination of Site-season × N level.

### Genetic map

The core collection was genotyped as previously described by Bordes et al. (2011) with 578 DArT markers generated by Triticarte Pty, Ltd. (Canberra, Australia; http://www.triticarte.com.au), 271 SSR and 109 SNP markers (Table S1). For association analyses, we considered rare alleles (i.e., those accounting for <2.5 % of the total number of alleles) as missing data. Finally, 899 markers (574 DArTs, 259 SSRs and 66 SNPs) were used to test association with phenotypic traits.

The map used as reference here was the consensus map described in Bordes et al. (2011) built with MetaQTL software (Veyrieras et al. 2007) using published data and the Somers et al. (2004) reference map. If markers had not been genetically mapped, they were placed close to the DArT markers with which the linkage disequilibrium (LD) was the highest. LD was calculated as $R^2$ (Weir 1996) with the Tassel v2.0.1 software (Bradbury et al. 2007).

### Genetic structure and association analysis

The population structure of the 372CC accessions was investigated using the STRUCTURE program (Pritchard et al. 2000) with 578 DArT markers distributed evenly across the genome (Bordes et al. 2011). Like the 372CC, the 196CC population was structured into five groups corresponding to the ancestral geographical groups (Horvath et al. 2009). To control for possible effects of this population structure, a matrix ($Q$-matrix) of the contribution of a genotype to each of the five inferred ancestor groups was computed.

Association between markers and traits was tested for each environment with the TASSEL version 2.0.1 software using the general linear model (GLM). In the whole core collection, the GLM-Q model, which controls the structure only, gave similar results to the more complex MLM-Q-K model, which controls for population structure ($Q$-matrix) and kinship between accessions ($K$-matrix) (Bordes et al. 2011).

The marker being tested was considered as a fixed effect. GLM was performed including the Q population structure coefficients as covariates and using 1,000 permutations to correct for multiple testing. The significance of associations between markers and traits was tested with an $F$ test and regarded as significant when the adjusted $p$ values were <0.05. The association of a marker with a trait is represented by its $r^2$ value, an estimate of the percentage of variance explained by the marker. It was previously shown that the most significant LD in 372CC were limited to 5 cM (Horvath et al. 2009), hence in the general description of associated regions, we have considered two adjacent markers located in an interval of <10 cM as

belonging to the same region. Maps and QTL were illustrated with the MapChart v2.1 software (Voorrips 2002).

## Results

### Characterization of the effects of the different environments

To investigate the genetic basis of wheat tolerance to moderate nitrogen levels, 196 accessions (196CC) from a larger worldwide core collection (Balfourier et al. 2007) were tested during 2 years at three locations (six site-seasons) and two input levels differing in the amount of N applied, pesticide treatment and seeding density. The ANOVA results of the different traits are presented in Table 2. The trait means at the two input levels HN and LN for each site-season are given in Table S2.

Mean GY ranged from 4.16 in JO07 at LN to 6.91 t ha$^{-1}$ in LM07 at HN (Table S2). The input level effect was significant at $p < 0.002$ for GY and GN but only significant at $p < 0.099$ and 0.18 for TKW and GPC, respectively (Table 2). The input level effect was not significant for PH and HD. GY and GN showed a large mean difference (1.31 t ha$^{-1}$ and 2,415 grains m$^{-2}$, respectively) between HN and LN (Table S2). For GY, the differences in the six site-seasons ranged from 0.2 t ha$^{-1}$ for JO08 to more than 2.0 t ha$^{-1}$ for JO07 and LM07. A strong year effect was observed at Joze. Despite similar reductions in N fertilizer input in both years (Table 1), we observed a large difference in GY in 2007 and a small one in 2008 and likewise the spike number per m$^2$ (Table S2). At JO07, the absence of N application at the tillering stage and use of a growth regulator had probably affected the number of tillers and consequently reduced the number of spikes. At JO08, the application of N at tillering at HN had little or no effect probably because of a severe drought at that time (rainfall records are given in Table S3) that must have strongly hindered N uptake. Consequently, the spike number at LN (387 spikes m$^{-2}$) was very similar to the spike number at HN (403 spikes m$^{-2}$) and the yields were similar. By contrast, the first N application for LN, on March 26, 2008 when soil moisture was good, was probably efficient. In conclusion, for JO08 the climatic conditions probably largely offset the difference in N supply.

**Table 2** Results of split-plot analysis of variance for genotype, input level, genotype × input level interaction effects for 196 bread wheat genotypes grown in six site-seasons (2 years × 3 locations)

| Trait | Source of variation | Effect | SS | DF | MS | *F* | *p* |
|---|---|---|---|---|---|---|---|
| Grain yield | Genotype | Random | 446582 | 195 | 2290 | 26.9 | 0.001 |
| | Genot × input level | Random | 16610 | 195 | 85 | 1.6 | 0.001 |
| | Input level | Fixed | 126183 | 1 | 126183 | 32.4 | 0.002 |
| | Site-season | Fixed | 13380 | 5 | 2676 | 0.7 | 0.655 |
| Grain number per m$^2$ | Genotype | Random | 3.3E + 10 | 195 | 1.7E + 08 | 15.7 | 0.001 |
| | Genot × input level | Random | 2.1E + 09 | 195 | 1.1E + 07 | 1.4 | 0.001 |
| | Input level | Fixed | 6.1E + 09 | 1 | 6.1E + 09 | 37.1 | 0.002 |
| | Site-season | Fixed | 1.7E + 09 | 5 | 3.4E + 08 | 2.1 | 0.219 |
| Thousand kernel weight | Genotype | Random | 55548 | 195 | 285 | 19.4 | 0.001 |
| | Genot × input level | Random | 2857 | 195 | 15 | 1.0 | 0.597 |
| | Input level | Fixed | 2875 | 1 | 2875 | 4.1 | 0.099 |
| | Site-season | Fixed | 22031 | 5 | 4406 | 6.3 | 0.033 |
| Plant height | Genotype | Random | 1055409 | 195 | 5412 | 110.3 | 0.001 |
| | Genot × Input level | Random | 9569 | 195 | 49 | 0.7 | 1.000 |
| | Input level | Fixed | 146 | 1 | 146 | 0.0 | 0.857 |
| | Site-season | Fixed | 13873 | 5 | 2775 | 0.7 | 0.654 |
| Heading date | Genotype | Random | 111758 | 195 | 573 | 249.8 | 0.001 |
| | Genot × Input level | Random | 447 | 195 | 2 | 0.6 | 1.000 |
| | Input level | Fixed | 83 | 1 | 83 | 0.7 | 0.436 |
| | Site-season | Fixed | 123394 | 5 | 24679 | 211.9 | 0.001 |
| Grain protein concentration | Genotype | Random | 3144.2 | 195 | 16.1 | 40.9 | 0.001 |
| | Genot × Input level | Random | 76.9 | 195 | 0.4 | 0.7 | 1.000 |
| | Input level | Fixed | 291.6 | 1 | 291.6 | 2.4 | 0.180 |
| | Site-season | Fixed | 2820.3 | 5 | 564.1 | 4.7 | 0.058 |

Site-season: year × location combinations. Genotype was tested against genotype × input level, genotype × input level was tested against genotype × input level × site-season + genotype × site-season and input level and site-season were tested against input level × site-season

The less than favorable conditions during growth in 2007 resulted in a low level of disease, leading to a slight difference in the extent of attack by septoria (*Septoria tritici*), brown rust (*Puccinia recondita*) and powdery mildew (*Erysiphe graminis*) between the two input conditions. The absence of growth regulator for LN in 2007 had little effect on the root lodging difference between the two input levels. Thus, we may assume that the main factor differing between HN and LN was the N application and thus any effects observed on the traits were mainly caused by the difference in available N.

## Trait variation and correlations between traits

For all the traits, there was a highly significant ($p < 0.001$) genotypic effect (Table 2). A large variability was observed in the population for all the traits at the two input conditions. The lowest yielding cultivar was Nishikaze Komugi (Japan) producing 2.76 t ha$^{-1}$ and the highest was Cadenza (France) producing 10.14 t ha$^{-1}$. The highly significant genotype × input level interaction observed for GY and GN indicated significant variability in RNL. The variation of TKW, PH and HD was not affected by the genotype × N level interaction. Unlike GY, GPC showed no RNL. The Pearson's correlation coefficients between yield and its components were very high for GN at both input levels and low for TKW (Table 3). So the variation in GY is explained more by GN than by TKW. The other coefficients were very similar between HN and LN. We observed strong negative correlations between GY and GPC and between GY and PH and a positive correlation between GPC and PH. There was a moderate negative correlation between GY and HD such that late genotypes on average performed less well than early ones.

**Table 3** Correlation coefficients between traits measured on 196 bread wheat genotypes grown in six site-seasons (2 years × three sites)

| Trait | GY | GN | TKW | PH | HD | GPC |
|---|---|---|---|---|---|---|
| Grain yield | | **0.88** | **0.30** | **−0.76** | **−0.21** | **−0.80** |
| Grain number | **0.90** | | **−0.16** | **−0.72** | **−0.21** | **−0.77** |
| Thousand kernel weight | 0.11 | **−0.32** | | −0.11 | 0.03 | −0.09 |
| Plant height | **−0.80** | **−0.77** | 0.06 | | **0.51** | **0.61** |
| Heading date | **−0.24** | **−0.28** | 0.13 | **0.53** | | 0.04 |
| Grain protein content | **−0.84** | **−0.79** | −0.02 | **0.64** | **0.08** | |

Significant coefficients ($p < 0.05$) are in *bold*

Above and under the diagonal: low and high input levels, respectively

## High variability in the response to N level

The variability of the genotype response to different N input was estimated for GY by the three computed parameters HN–LN, LN/HN and the slope of the joint regression. For GN, as the sowing density was different between some environments we only used HN–LN and LN/HN. The variations in HN–LN and LN/HN were significant (Table 4). The HN–LN difference ranged from 0.05 to 2.94 t ha$^{-1}$ for GY and from −456 to 4,894 for GN. The range of the ratio LN/HN was from 0.65 to 0.99 for GY and from 0.68 to 1.05 for GN. For each genotype, we then computed the regression of its yield on the average yield for the 12 environments (site-season × input level combinations). The genotype × environment covariate (joint regression) explained 17.5 % of the genotype × environment interaction (Table 5). On average, the determination coefficients were relatively high ($r^2 = 0.5$) being maximal ($r^2 = 0.96$) for cultivar Gk Szoke (Hungary) and minimal ($r^2 = 0.04$) for Blanc Précoce (Switzerland). The slope for GY ranged from 0.16 to 2.16. Thus, the three parameters indicated a significant variability in the sensitivity to N limitation. The two parameters HN–LN and LN/HN correlated well (0.8) but were not significantly correlated with the slope of the joint regression, indicating that they describe different behaviors.

There was a positive relationship between the slope of the joint regression and the mean yield for the 196 accessions (Fig. 1). The lowest yielding third (cultivars yielding 3–5 t ha$^{-1}$) was composed mainly of old cultivars registered before 1960 and several recent cultivars mainly originating from Southeast Europe and probably not well suited to the French climate. In contrast, the highest yielding third (cultivars yielding 7–9 t ha$^{-1}$) was composed of modern varieties registered after 1960 with the exception of two old cultivars: Kid (France) and Hana (Czech Republic). Whichever GY graphic area was considered, the variability for the slope was similar (about 0.8 range). Among both old and recent cultivars were individuals with relatively shallow slopes indicating stable responses to contrasted environments and individuals with relatively steep slopes that respond much more. For instance, the recent and most productive cultivars with high yields and relatively low slope coefficients were cv. Arche (France) and Renan (France), whereas Bellovac (France) and CF3003-2 (France) with relatively high slope coefficients were less stable responding more to contrasting environments.

## Association analysis of phenotypic traits

As we found significant genotypic variations for all the traits (Table 2), we looked for genetic associations both in

**Table 4** Analysis of variance on two parameters for the response of grain yield and grain number per $m^2$ to the input level: the difference and ratio between the value under high level and the value under lower level

| Trait | Source of variation | DF | Grain yield | | | Grain number per $m^2$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | MS | *F* | *p* | MS | *F* | *p* |
| Difference HN–LN | Genotype | 195 | 170.4 | 7.3 | 0.001 | 2.2E + 07 | 1.8 | 0.001 |
| | Site-season | 5 | 7795.1 | 336.2 | 0.001 | 3.3E + 08 | 27.4 | 0.001 |
| Ratio LN/HN | Genotype | 195 | 0.022 | 4.0 | 0.001 | 0.029 | 2.8 | 0.001 |
| | Site-season | 5 | 1.765 | 321.3 | 0.001 | 0.972 | 96.1 | 0.001 |

Genotype and site are tested against genotype × site-season interaction effect

*HN* high input level, *LN* lower input level

**Table 5** Analysis of variance for grain yield where the genotype × environment interaction is decomposed by a joint regression analysis

| Source of variation | SS | DF | MS | *F* | *p* | Percent of genot × env |
|---|---|---|---|---|---|---|
| Genotype (Genot) | 446581.5 | 195 | 2290.2 | 39.8 | 0.001 | |
| Environment (Env) | 159050.8 | 11 | 14459.2 | 251.2 | 0.001 | |
| Genot × Env | 123478.7 | 2145 | 57.6 | | | |
| Genot × EnvCov | 21300.2 | 195 | 109.2 | | | 17.25 |
| Residual error | 102178.5 | 1950 | 52.4 | | | 82.75 |

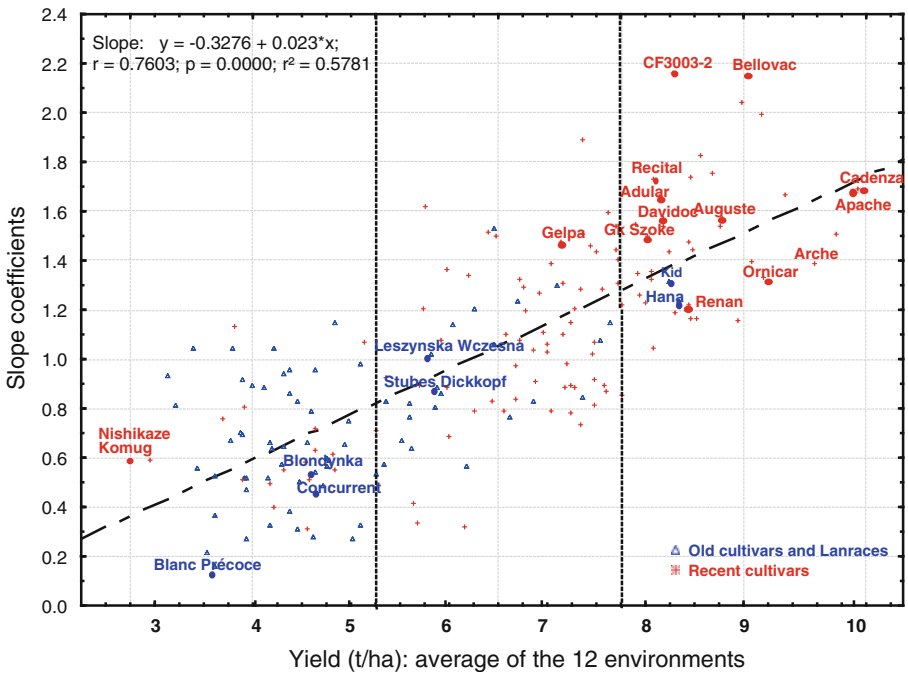


**Fig. 1** For 196 wheat accessions grown on 12 environments, slope coefficients of the joint regression against the average grain yield of the accessions. Old and recent cultivars: bred before and after 1960, respectively. Cultivars named are those mentioned in the text for a particular characteristic

each of the 12 environments and for the mean over the 12 environments. The underlying structure of the collection, i.e., five groups according to data from 578 SSR markers, explained 25 % of the phenotypic variance in GY. The markers found to be significantly associated with the mean over the 12 environments and in at least three environments were considered as stable markers. Only these stable markers were taken into account and reported in Fig. 2. When markers were significantly associated with at least five environments we considered them to be very stable. We also searched for associations with the three GY and GN response parameters previously described, RGY for GY and RGN for grain number per $m^2$. When one marker was associated with several response parameters only one association was reported in Fig. 2. All the details of the results for each parameter, the $r^2$ of each association, and the number of environments where the marker was associated are given in Table S4.

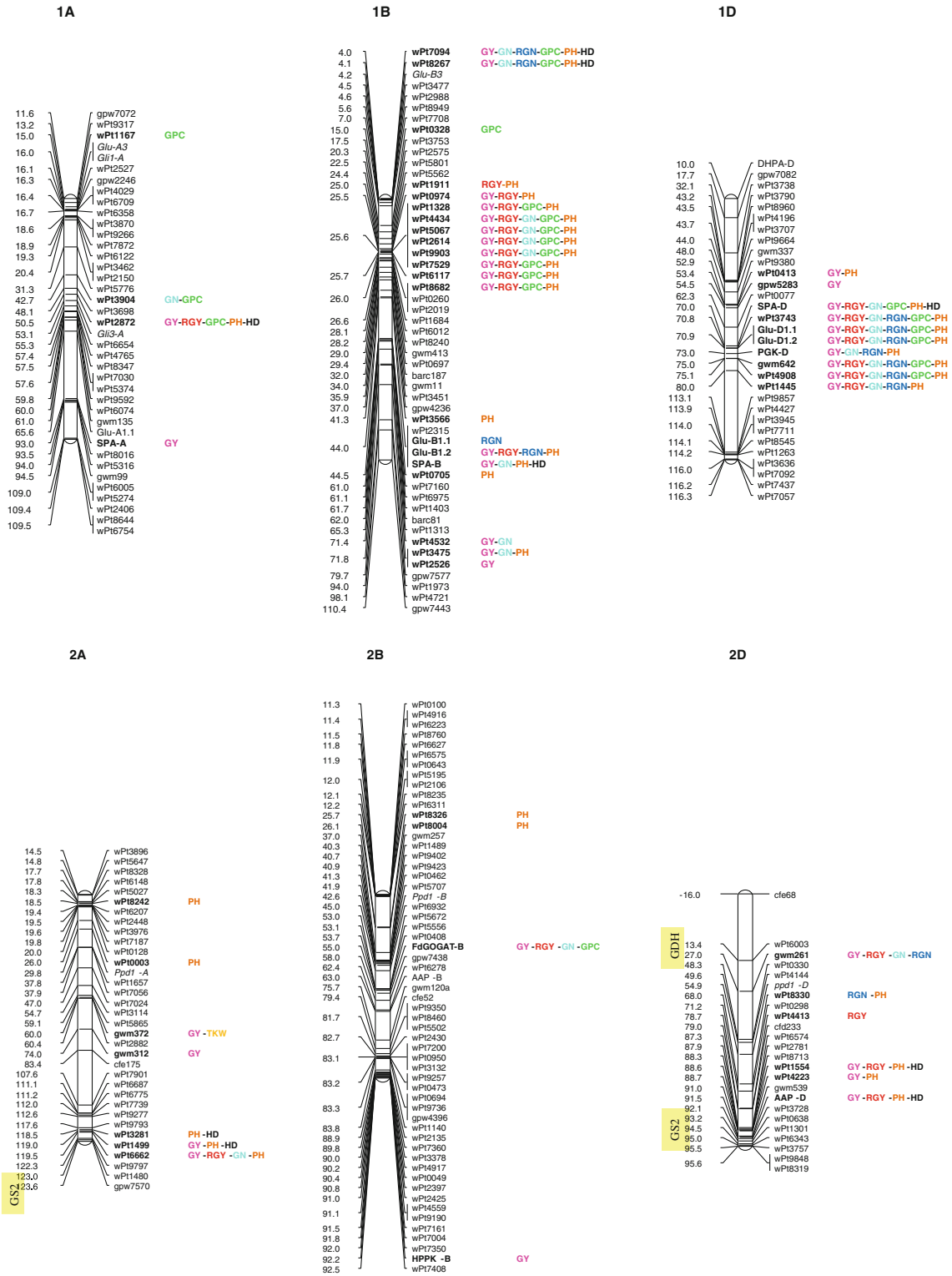


**Fig. 2** Consensus map showing DArT, SSR and SNP markers tested for association on a collection of 196 wheat accessions. Some genes not tested (*italic*) and some QTLs (*written vertically* and *highlighted in yellow*) found by Habash et al. (2007) and Fontaine et al. (2009) were also mapped. Significant markers are shown for grain yield (GY), response to nitrogen level assessed for GY (RGY), grain number per $m^2$ (GN), response to nitrogen level assessed for GN (RGN), grain protein concentration (GPC), plant height (PH), heading date (HD) and thousand kernel weight (TKW). Distances are in cM

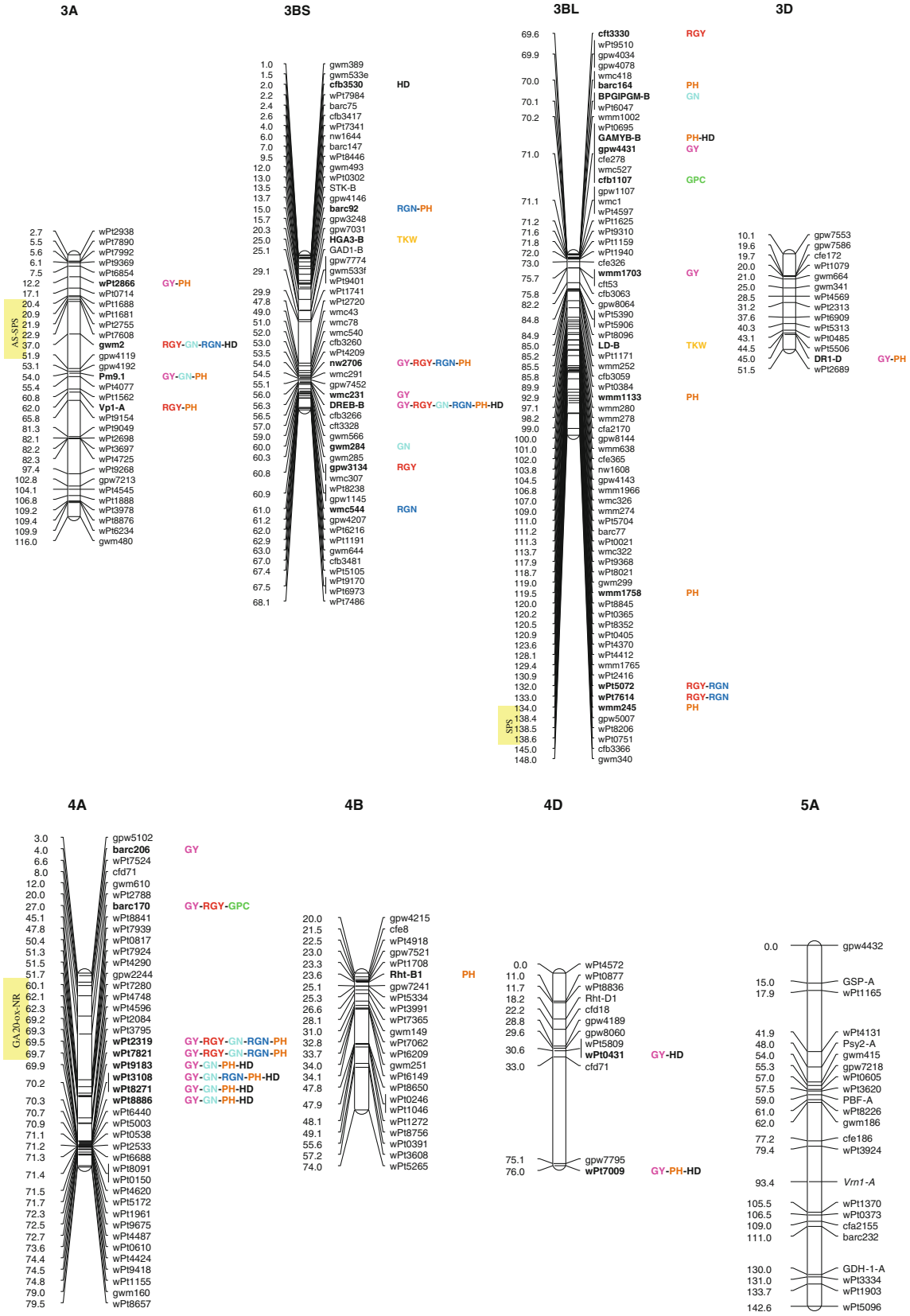


**Fig. 2** continued

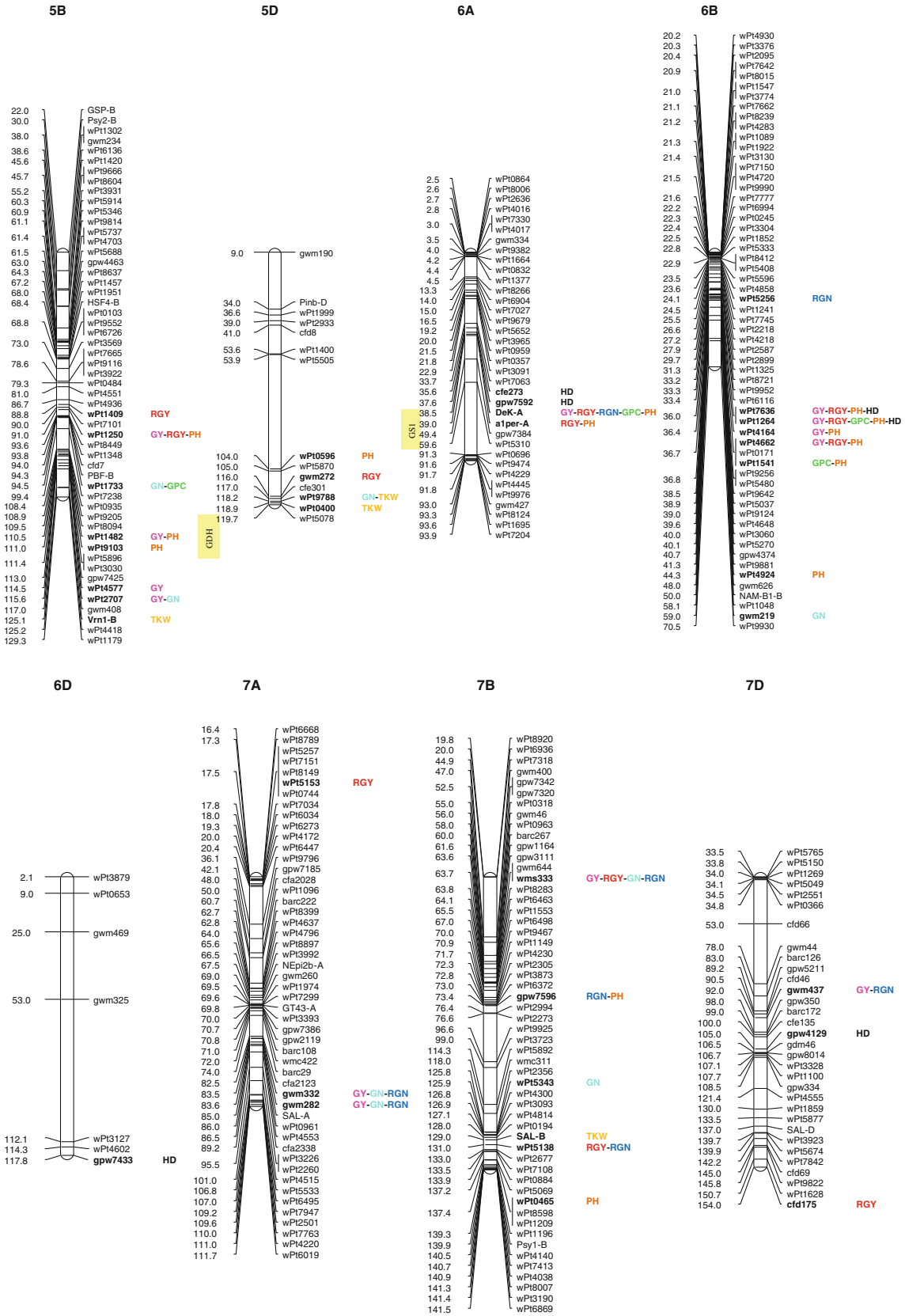


**Fig. 2** continued

We identified 54 main associated regions spread along all chromosomes except chromosome 5A. Some regions were represented by only one marker (DArT, SSR or SNP) while others were represented by clusters of two to seven markers. The length of segments covered by clusters varied from 0.1 cM for two DArT markers on chromosome 1B to 10 cM for seven DArT, SSR and SNP markers on chromosome 1D.

The association of cluster wPt-1911/wPt-8682 on chromosome 1BS with almost all the traits was due to the 1BL/1RS translocation carried by 19 genotypes. When these 19 genotypes were not included in the analysis, only the distal cluster wPt-7094/wPt-8267 remained associated. In the following presentation of results, the wPt-1911/wPt-8682 cluster will be treated separately from the other regions.

*Chromosome regions associated with grain yield*

Thirty-one regions spread over 16 chromosomes were involved in the control of GY variation. Twenty-one of these regions were detected in more than five environments. For the mean over the 12 environments, the $r^2$ of significant markers ranged from 4.4 to 25.3 % (gwm282 on 7A). With SNP 13 genes were identified that are possibly involved in the variation in GY.

For RGY, 24 regions spread over 16 chromosomes were involved. The $r^2$ of significant markers of the three parameters (HN–LN, LN/HN and the slope) ranged from 4.7 to 22.9 % (for nw2706 on 3B). Of the 24 regions involved in the response to input level, 15 were involved in GY, including homoeologous genes *Glu-B1* (1B) and *Glu-D1* (1D), *SPA-D* (1D), *FdGOGAT-B* (2B), *AAP-D* (2D), *DREB-B* (3BS) and *DeK1-A* (6A). Seven regions were independent of regions involved in GY including genes *Vp1-A* (3A) and *A1per-A* (6A). Two markers, wPt-4413 (2D) and gwm272 (5D), each representing one of these independent regions, were significantly associated with all three N parameters. Twelve regions were significantly associated with two parameters and nine with only one parameter.

*Chromosome regions associated with grain number per m$^2$*

Twenty regions spread over 14 chromosomes, including 8 SNP were involved in the control of GN. For the mean over the 12 environments the $r^2$ of the significant markers ranged from 4.1 to 22.5 % (for both gwm332 and gwm282 on 7A).

For the two parameters estimating the response to the input level (HN–LN and LN/HN), 17 regions spread over 11 chromosomes were involved including genes *Glu-B1, Glu-D1*, *PGK-D*, *DREB-B* and *DeK1-A*. Only nine regions coincided with the 20 regions directly associated with GN, hence 11 were independent. For HN–LN and LN/HN the $r^2$ of the significant markers ranged from 4.2 to 26.0 % (for gwm332 on 7A).

*Coincidences between loci associated with GY and grain number per m$^2$*

Of the 31 and 20 regions involved in GY and GN, respectively, 13 coincided, including the genes *SPA-B* (1B), *SPA-D*, *Glu-D1*, *PGK-D* (1D), *FdGOGAT-B, PM9.1* (3A) and *DREB-B*. For the response to input level, of the 23 and 17 regions involved, respectively, in the GY and GN parameters, 11 coincided including genes *Glu-B1*, *Glu-D1*, *DREB-B* and *DeK1-A*. The most explanatory region was represented by SSR nw2706 on chromosome 3B, which was significantly associated with mean GY in five environments with an $r^2$ value of 21.1 %. It was also associated with two RGY parameters with $r^2$ values of 18.7 and 22.9 % and with one RGN parameter with an $r^2$ value of 23.2 %. Two very close SSR markers on chromosome 7A, gwm332 and gwm282, were also highly involved in GY and GN but not in RGY. For GY and across the environments, the response of the five most frequent alleles of nw2706 was different than the response of the four most frequent alleles of gwm282 (Fig. 3). For gwm282 the decrease in yield was almost the same for each allele. For example, a difference of 1 t ha$^{-1}$ was observed between the extreme alleles (249 vs. 210) in all the environments. The decrease was different between alleles for nw2706. For example, the theoretical difference estimated from the slope of the regression line between the two extreme alleles 202 and 174 was 2.5 t ha$^{-1}$ in high yielding environments and only 1.0 t ha$^{-1}$ in low yielding environments. Most of the 47 accessions carrying allele 202 were recent (1978–1998) or relatively recent (1960–1975) French varieties, only 12 accessions were old cultivars, whereas most of the 17 accessions carrying allele 174 were landraces and very old cultivars (1900–1939) only seven were recent (1974–1975) or relatively recent (1962–1967).

*Associations with other traits*

TKW, the second yield component, was associated with six regions including the genes *Hga3-B*, *LD-B* (3B) and *Vrn-B1* (5B). With an $r^2$ value of 22.4 % the most explanatory marker was SSR gwm372 on chromosome 2A, the same marker that was associated with GY.

PH was associated with 30 regions including 14 SNP markers spread over 17 chromosomes. Twenty of these coincided with the regions involved in GY. HD was associated with 17 regions including the genes *SPA-B* (1B), *SPA-D*, *AAP-D*, *DREB-B* and *GAMYB-B* (3B). Eleven of these regions coincided with the regions found to be

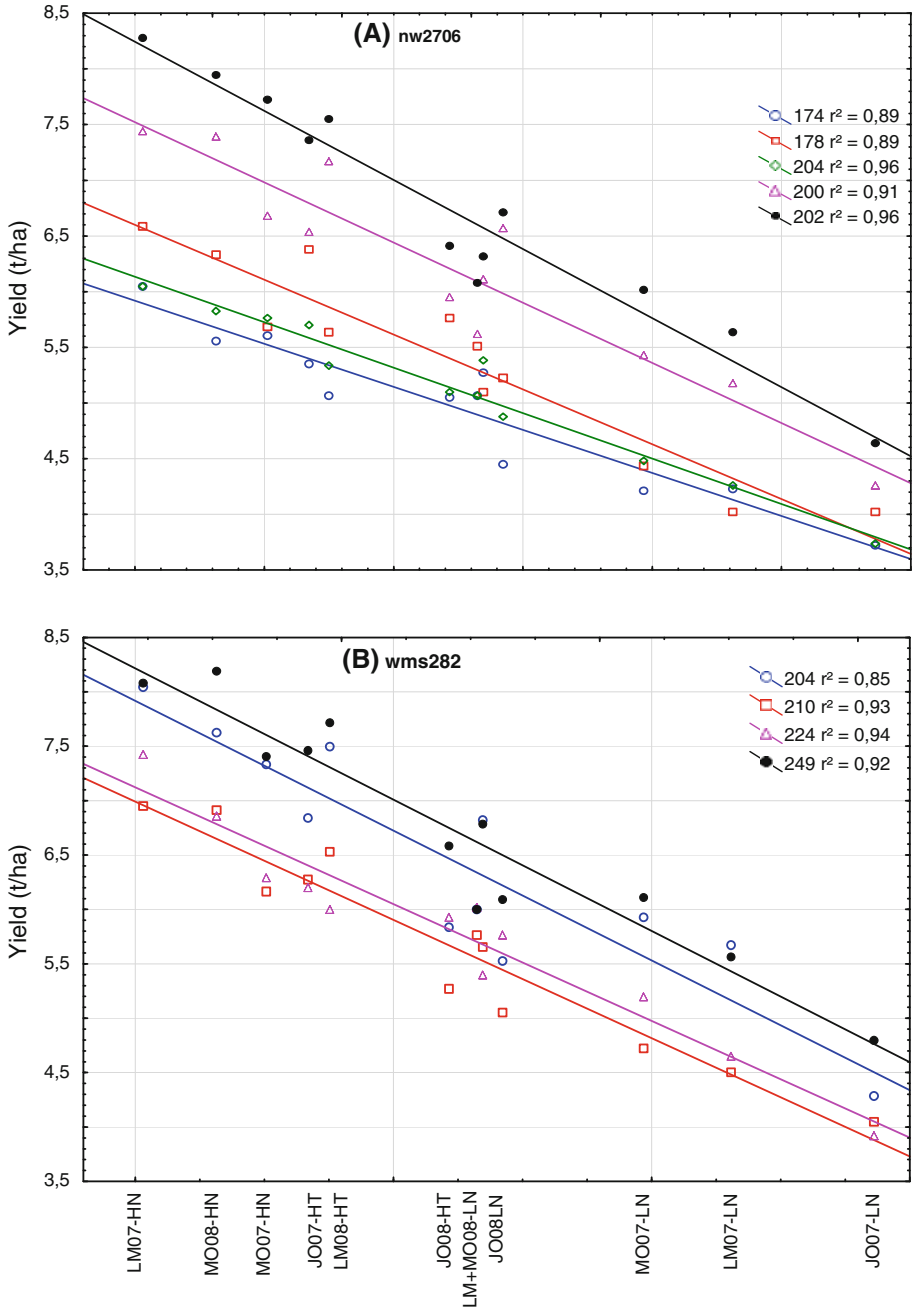


**Fig. 3** Response of the most represented alleles of the SSR markers nw2706 (**a**) and gwm282 (**b**) for grain yield in the 12 environments average of the 196 accessions ranked from the largest to the smallest. Allele values are calculated as the average for all the *lines* that carry it. Environments are coded as LM, MO and JO for Joze, Mons and Le Moulon, respectively, 07 and 08 for 2007 and 2008, respectively, and HN and LN for high input and low input, respectively. SSR alleles are named according to their migration size

involved in PH and ten with regions involved in GY. Finally, GPC was associated with 11 regions including the genes *SPA-D*, *Glu-D1*, *FdGOGAT-B* and *DeK1-A*. Seven of these regions corresponded to regions involved in GY.

## Discussion

Using a large experimental network, we identified chromosome regions associated with the main wheat breeding traits: grain yield, grain number per $m^2$, thousand kernel weight, plant height, heading date and grain protein concentration. We also identified regions involved in the response to input level, mainly differing in the amount of N applied, and found specific alleles that may mitigate reduction in yields when the crop is managed at moderate input levels. The phenotyping study by Bordes et al. (2008) was used to sample from the 372CC genotypes maximizing the variability for agronomic traits while favoring shorter cultivars to limit root lodging. The highly significant genotypic variance observed for all the traits demonstrated the large variability of this sample of 196 accessions which

facilitated a genome wide association analysis to locate regions involved in determining these traits.

The significant effect of input level confirmed the view that N availability has a large impact on most cereal yield traits but not PH. The absence of a significant effect on PH was probably due to the use of the growth regulator to avoid lodging. The Pearson's correlation coefficient confirmed the strong influence of GN and the low influence of TKW on GY, mainly at HN. The correlations found were similar, mainly at HN, to correlation coefficients between yield and TKW and GN, approximately 0.3 and 0.8, respectively, reported by Kuchel et al. (2007). The negative correlation between GY and both GPC and PH could be explained by the inclusion of many old accessions that are very tall with a low grain yield, yet rich in grain proteins. In modern cultivars, the negative correlation between GY and GPC is well known (reviewed by Oury et al. 2003) and the value we report here is similar to the average observed in French registration trials (Oury and Godin 2007).

Assuming that the observed effects were mainly caused by the N fertilizer level, the genotype × input level interaction can be interpreted as a response to available N. The three parameters used to characterize the response of GY and GN to the input level were significant (HN–LN and LN/HN) or explained more than 17 % of the heterogeneity of the slopes in the joint regression analysis, close to the 22 % value obtained by Brancourt-Hulmel et al. (1994). In the following discussion, while we focus in the mapped regions associated with RNL, some associations of the main chromosome regions with raw traits are also discussed.

## Associations spread over the whole genome for raw and response traits

Associated markers were fairly well distributed throughout the genome. They were generally very stable and were even significant in all environments, particularly for PH. All chromosomes, except chromosome 5A, were involved in determining at least one trait.

Some regions, mainly on group 1 chromosomes, were particularly rich in associations; the cluster including the 1BL/1RS translocation, the cluster around the gene *Glu-D1* including the *SPA-D* gene, and some other regions including the *Glu-B1* and *SPA-B* genes acted on almost all the traits. The respective positive and negative impact of the 1BL/1RS translocation on GY and GPC is well established (Crossa et al. 2007). The structural *Glu1* genes encode high molecular weight subunits of glutenin (HMW-GS), a storage protein (Payne 1987). They influence the quantity of protein fractions (Zhang et al. 2011) and the quantity and quality of HMW-GS (Branlard et al. 2003). Variation in N application also influences protein components and the ratio of gliadin to glutenin subunits (Johansson et al. 2001). In the same region encompassing SSR gwm642, Habash et al. (2007) detected a QTL for PH, grain filling, plant N content and grain weight per ear. In the *Glu-D1* region, including DArT wPt-3743 as in our results, Crossa et al. (2007) found significant associations with GY. The transcription factor *Spa1* plays an important role in plant development (Ravel et al. 2009) and a strong association of this region with HD has already observed (Le Gouis et al. 2011).

On the contrary, some regions expected to be involved in the determination of specific traits showed no or few significant associations here. The absence of associated markers on chromosome 5A, noted for bearing the dwarfing gene *Rht-12* (Peng et al. 1999), the vernalization requirement gene *Vrn1-A* (Law et al. 1976) and the awn inhibitor B1 gene (McIntosh et al. 1998), does not accord with several QTL and association or linkage studies in which chromosome 5A is demonstrated to influence grain yield, thousand kernel weight, response to input level and flowering time (Quarrie et al. 2005, Laperche et al. 2007, Rousset et al. 2011; Worland et al. 1994). Our sample of 196 accessions was largely divided into short and tall cultivars, reflecting the fact that about half were old cultivars and about half were recent varieties possibly carrying the *Rht-12* dwarfing gene, a mutant allele induced in 1968 by gamma ray treatment of Karcagi 522 (Viglasi 1968). The PH means of old and recent varieties differed greatly in all 12 environments ($p < 0.000$, data not shown). The absence of association observed for PH and other traits was very likely due to an absence of LD between markers and *Rht-12*, and possibly *Vrn1-A* and *B1*. So coverage of 5A chromosomes was probably not sufficient to analyze the effects of some genes known to affect the studied traits. The Rht-B1 gene on chromosome 4B was only associated with PH in six environments, and although not associated with GY means, it was significantly associated with GY (mean, $p = 0.05$) in three environments. Unlike Laperche et al. (2007) who observed a significant influence of *Rht-B1* on the sensitivity of a mapping population to N limitation, *Rht-B1* was not associated with RNL in our study. It is possible that the inclusion of recent short accessions giving very low yields in our conditions, like Nishikaze Komug (Japan), the least productive and the shortest (GY 3.17 t $ha^{-1}$; PH 75 cm) of the 196 accessions (and possibly lacking the dwarfing gene), could disrupt the association test for the 4B and 5A chromosomes.

Complex traits such as GY, PH, GN, HD and GPC are under the control of numerous genomic regions (e.g., Börner et al. 2002; Groos et al. 2003; Prasad et al. 2003; Charmet et al. 2005; Blanco et al. 2006; Kuchel et al. 2007; Crossa et al. 2007). The pleiotropic effects of several loci for different traits are also consistent with significant

correlations being observed between these traits. The numerous regions involved in determining PH were very often involved in GY and also GPC and HD. Unlike other traits, the few associations found for TKW may be partly explained by the strong influence of environmental conditions during the grain filling period when the crop is more susceptible to drought and heat stress. Donmez et al. (2001) observed that this trait was the most variable of the yield components using a population of 14 old (1873–1945) and recent (1964–1995) cultivars. Marker gwm372 on chromosome 2A was the most explanatory for TKW variation ($r^2 = 22.2$ %) but with moderate stability (in four environments of the 12 tested). Interestingly, this region was associated with GY but not with GN. For the other five regions associated with TKW, four corresponded to SNP in genes, including *Hga3* and *LD* on chromosome 3B, *Vrn* on 5B and *SAL1* on 7B. The *Hga* gene codes for an enzyme which synthesizes the pectic polysaccharide homogalacturonan (HGA) that accounts for approximately 60 % of total plant pectin, a major component of primary cell walls (Willats et al. 2001). The LuminiDependens (*LD*) gene and the *Vrn-B1* gene are involved in the control of wheat development and were shown to be associated with vernalization requirement in an association study on a subsample of 372CC (Rousset et al. 2011). The sub-aleurone layer (*SAL*) gene, which plays an important role in determining the number of aleurone layers and affects vacuolar sorting proteins (Sabelli and Larkins 2009), could be involved in TKW. Thus, these four genes are candidates as acting on TKW.

### Several genes and markers implicated in RNL

The response to the input level appeared to be explained by associations with all three homoeologous groups. Several enzymes involved in plant nitrogen assimilation may explain NRL differences observed in wheat, mainly NR (EC 1.6.6.1), NiR (EC 1.7.7.1), GS (EC 6.3.1.2), glutamate synthases NADH-GOGAT (EC 1.4.1.14) and Fd-GOGAT (EC 1.4.7.1) and GDH (EC 1.4.1.2) (Habash et al. 2007; Hirel et al. 2007; Fontaine et al. 2009). The genes encoding these enzymes and QTL determining their activities are spread over the whole genome. For instance, 26 QTL for GDH activity are spread over 13 chromosomes and 25 QTL for GS activity are spread over 12 chromosomes (Fontaine et al. 2009) while loci involved in N metabolism are found throughout the genome (Habash et al. 2007). According to our consensus map using common markers, several of the regions we identified co-localized with these N metabolism-related regions, thus bearing out our initial assumption that most of the response effects observed between input levels are due to differences in available nitrogen.

The large number of regions observed to be involved in RNL included several genes or markers in LD with genes that encode enzymes involved in N metabolism, e.g., the ferredoxin-dependent (*Fd*)-*Gogat* gene on chromosome 2B and markers wPt-6662 and wPt-1554 possibly in LD with GS2 genes about 5 cM away on chromosomes 2A and 2D, respectively (Habash et al. 2007). SSR nw2706 on chromosome 3B is very informative for RNL in terms of GY and GN and maps in the vicinity of the NADH-*Gogat* gene, a candidate gene for an NUE QTL (Masood Quraishi et al. 2011). GOGAT catalyzes the conversion of glutamine and 2-oxoglutarate to two molecules of glutamate, thus providing glutamate for ammonium assimilation while GS plays a major role in fixing ammonium to form the amino acid glutamine, hence these enzymes have a central role in nitrogen assimilation and recycling (Krapp et al. 2005).

The gene *AAP* on chromosome 2D encodes an amino acid permease. Members of the amino acid permease family are good candidates for proteins involved in the phloem loading process (Koch et al. 2003). One member is mainly expressed in root tissue (Okumoto et al. 2004), suggesting a potential role in the uptake and distribution of amino acids and hence in nitrogen metabolism. The *AAP* gene could also be in LD with the GS2 gene about 3 cM away (Habash et al. 2007). On chromosome 6A, the *Dek* gene encoding a calpain cysteine proteinase involved in seed development (Lid et al. 2005) and the *A1per* gene encoding xanthine/uracil/vitamin C permease located at about 1 cM from GS1 genes could be in LD with this gene. SSR gwm261 and gwm272 on chromosomes 2D and 5D, respectively, coincide with a *GDH* gene positioned by Habash et al. (2007). The GDH enzyme plays important roles in controlling tissue ammonium concentration and N assimilation and recycling (Dubois et al. 2003; Fontaine et al. 2009). SSR gwm2 on chromosome 3A coincides with several QTL reported to act on yield components and N content of different organs of plants and co-localizes with the asparagine synthetase (*AS2*) and *SPS* genes (Habash et al. 2007). DArT markers wPt-5072 and wPt-7614 involved in the RNL of GY and GN map to a distal region of chromosome 3BL and may be in LD with an *SPS* gene localized about 6 cM away (Habash et al. 2007). In plants, a significant portion of assimilated N is channeled into the relatively inert amino acid asparagine, serving to store N and/or transport it from sources to sinks. SPS, a key regulatory enzyme in the pathway of sucrose biosynthesis, has been shown to be linked to QTL controlling plant growth and yield, mainly in maize (Castleden et al. 2004).

In contrast to chromosome 4B, a cluster on chromosome 4A from markers wPt-2319 to wPt-8886 was strongly associated with PH ($r^2$, 9.1–13.1) with very stable markers (in all 12 environments) so it could be in strong LD with a gene involved in this trait. This locus also acts on earliness

and yield (Crossa et al. 2007) and was involved in RNL of both GY and GN, mapping in the vicinity of the *NR* and *GA-20Ox* genes (Habash et al. 2007). NR activity correlates with N absorption (Kichey et al. 2006, 2007). *GA20-ox* codes for gibberellin GA 20-oxidase which catalyzes the final steps in the synthesis of bioactive GAs (Yamaguchi 2008). It is possible that GA modulates lateral root proliferation associated with regulation of plant allometry during the stress response that could have an effect on N absorption (Gou et al. 2010).

To a lesser degree, it could be argued that the cluster between wPt-7636 and wPt-4662 on chromosome 6B influences PH even though Crossa et al. 2007 found no significant association in this region. In a nearby region, however, one QTL involved in grain number per ear, grain weight per ear, grain N per ear, N per grain and percentage peduncle N (Habash et al. 2007) could explain the involvement of this cluster in RNL.

Other genes such as Vip1-3A, known to contribute to susceptibility of modern hexaploid wheat varieties to PHS (Groos et al. 2002), or DREB-B, known to control the expression of stress-responsive genes via ABA-independent pathways in Arabidopsis (Pradeep et al. 2006), and other markers significantly associated with RNL were in regions not yet established as acting in nitrogen assimilation. These regions are now candidate loci for RNL in wheat.

Finally, several favorable alleles have now been identified that minimize the loss of yield at low input level compared to high input level which can be used in breeding programs where the adaptation to moderate N level is important. For instance, the SSR nw2706 marker explains most of the response to input level. The cultivars carrying the 174 and 204 alleles of this marker were more stable between environments than cultivars carrying the 202 allele. Allele 174 was found in landraces or old cultivars such as Concurrent (Netherlands, 1905) and Blondynka (Poland, 1920) but also in some varieties such as the relatively recent Gelpa (France, 1967). Allele 204, found in several old cultivars, was also found in cv Adular (Germany) and cv Davidoc (France) registered in 1986 whose average yields were of 8.4 and 8.2 t $ha^{-1}$, respectively. Therefore, it is now possible to incorporate these cultivars carrying favorable alleles into breeding programs as they are less demanding in nitrogen while productivity levels are compatible with current demand.

### Influence of the 1BL/1RS translocation

The 1BL/1RS translocation had an unfavorable effect on RNL in terms of the chosen criteria giving greater HN–LN differences, lower LN/HN ratios and higher slope coefficients (data not shown). However, in low input conditions, the average yield performance of the 19 accessions carrying the translocation was significantly better than the average performance of the other individuals (data not shown). Greater root biomass has been reported for spring wheat lines with the 1BL-1RS translocation (Ehdaie and Waines 2003), an effect confirmed in primary synthetic spring wheats grown in Australia (Dreccer et al. 2004) in conditions where nitrate rapidly leaches down the soil profile. Rooting depth is therefore an important attribute for soil N capture. The 1BL/1RS translocated lines may be of interest for increasing yield in low N situations. However, the 1BL/1RS translocation also has a deleterious effect on the rheological properties of bread wheat flour (Wieser et al. 2000) although there is considerable variation in the magnitude of such effects (Pena et al. 1990) that could be highly dependent on the genetic background (Lee et al. 1995). Some sister lines from the CIMMYT nursery bearing the 1BL/1RS translocation have good bread-making properties (Pena et al. 1990). The 1BL/1RS translocation has the advantage of carrying several disease resistances (Kim et al. 2004) and potentially confers beneficial health effects since the flour has an increased content in dietary fiber (Bordes et al. 2011).

## Conclusion

The 196CC contains a wide genetic variability that can be used to dissect the genetic determinism of the main agronomic traits targeted in wheat breeding. The genetic variability includes that for RNL and several loci involved in this trait have been newly identified or confirmed at the gene or marker level. Denser high-throughput genotyping can be used to identify new regions, to map some regions more precisely and to detect new candidate genes.

Favorable alleles of associated markers could now be the focus of breeding programs to reduce N input without unacceptable decreases in grain yield. The presence of favorable alleles in both old and recent cultivars suggests that they have not been eliminated by the breeding process mostly conducted at an optimal nitrogen level. To increase the range of the available genetic variability in the confirmed chromosome regions, new alleles will be sought in genetic resources available in germplasm banks. Finally, the ongoing sequencing of wheat chromosome 3B (Paux et al. 2008) will enable a more precise study of regions specifically involved in the response to available nitrogen.

**Acknowledgments** The experimental work was partly supported by the Enterprise Competitiveness Fund Project "Semences de Demain".

## References


Agarwal PK, Agarwal P, Reddy MK, Sopory SK (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Rep 25:1263–1274

An DG, Su JY, Liu QY, Zhu YG, Tong YP, Li JM, Jing RL, Li B, Li ZS (2006) Mapping QTLs for nitrogen uptake in relation to the early growth of wheat (*Triticum aestivum* L.). Plant Soil 284:73–84

Austin RB, Ford MA, Edrich JA, Blackwell RD (1977) The nitrogen economy of winter wheat. J Agric Sci 88:159–167

Balfourier F, Roussel V, Strelchenko P, Exbrayat-Vinson F, Sourdille P, Boutet G, Koenig J, Ravel C, Mitrofanova O, Beckert M, Charmet G (2007) A worldwide bread wheat core collection arrayed in a 384-well plate. Theor Appl Genet 114:1265–1275

Barraclough PB, Howarth JR, Jones J, Lopez-Bellido R, Parmar S, Caroline E, Shepherd CE, Hawkesford MJ (2010) Nitrogen efficiency of wheat: genotypic and environmental variation and prospects for improvement. Eur J Agron 33:1–11

Blanco A, Simeone R, Gadaleta A (2006) Detection of QTLs for grain protein content in durum wheat. Theor Appl Genet 113:563–565

Bonnin I, Rousset M, Madur D, Sourdille P, Dupuits C, Brunel D, Goldringer I (2008) FT genome A and D polymorphisms are associated with the variation of earliness components in hexaploid wheat. Theor Appl Genet 116:383–394

Bordes J, Branlard G, Oury FX, Charmet G, Balfourier F (2008) Agronomic characteristics, grain quality and flour rheology of 372 bread wheats in a worldwide core collection. J Cereal Sci 48:569–579

Bordes J, Ravel C, Le Gouis J, Charmet G, Balfourier F (2011) Use of global wheat core collection for association analysis of flour and dough quality traits. J Cereal Sci 54:137–147

Börner A, Schumann E, Fürste A, Cöster 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

Bradbury PJ, Zhang ZC, Kroon D, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633–2635

Brancourt-Hulmel M, Lecomte C-M, Bérard P, Galic N, Trouvé B, Sausseau C (1994) Sélection et stabilité du rendement chez le blé tendre d'hiver. Agronomie 14:611–625

Branlard G, Dardevet M, Amiour N, Igrejas G (2003) Allelic diversity of HMW and LMW glutenin subunits and omega gliadins in French bread wheat (*Triticum aestivum* L.). Genet Resour Crop Evol 50:669–679

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177

Castleden CK, Aoki N, Gillespie VJ, MacRae EA, Quick WP, Buchner P, Foyer CH, Furbank RT, Lunn JE (2004) Evolution and function of the sucrose-phosphate synthase gene families in wheat and other grasses. Plant Physiol 135:1753–1764

Charmet G, Robert N, Branlard G, Linossier P, Martre P, Triboi (2005) Genetic analysis of dry matter and nitrogen accumulation and protein composition in wheat kernels. Theor Appl Genet 111:540–550

Crossa J, Burgueno J, Dreisickacker S, Vargas M, Herrera-Foessel SA, Lillemo M, Singh RP, Trethowan R, Warburton M, Franco J, Reynolds M, Crouch JH, Ortiz R (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177:1889–1913

Donmez E, Sears RG, Shroyer JP, Paulsen GM (2001) Genetic gain in yield attributes of winter wheat in the Great Plains. Crop Sci 41:1412–1419

Dreccer F, Ogbonnaya G, Borgognone J, Wilson J (2004) Variation in shoot and root growth in primary synthetic wheats—implications for overcoming water deficits in marginal environments. In: Proceedings of the 4th international crop science congress, Brisbane, 26 Sept–1 Oct 2004 (http://www.cropscience.org.au/icsc2004/poster/1/3/2/637_dreccer.htm)

Dubois F, Tercé-laforgue T, Gonzalez Moro MB, Estavillo JM, Sangwan R, Gallais A, Hirel B (2003) Glutamate dehydrogenase in plants: is there a new story for an old enzyme? Plant Physiol Biochem 41:565–576

Ehdaie B, Waines JG (2003) 1RS translocation increases root biomass in Very-type wheat isogenic lines and associates with grain yield. In: Pogna NE, Romano M, Pogna EA, Galterio G, eds. Proceedings of the 10th International Wheat Genetics Symposium, Vol. 2. ISC: Paestum, Italy, p 693–695

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant breeding programme. Aust J Agric Res 14:742–754

Fontaine JX, Ravel C, Pageau K, Heumez E, Dubois F, Hirel B, Le Gouis J (2009) A quantitative genetic study for elucidating the contribution of glutamine synthetase, glutamate dehydrogenase and other nitrogen-related physiological traits to the agronomic performance of common wheat. Theor Appl Genet 119:645–662

Foulkes MJ, Hawkesford MJ, Barraclough PB, Holdsworth MJ, Kerr S, Kightley S, Shewry PR (2009) Identifying traits to improve the nitrogen economy of wheat: recent advances and future prospects. Field Crops Res 114:329–342

Gaju O, Allard V, Martre P, Snape J, Heumez E, Le Gouis J, Moreau D, Bogard M, Griffiths S, Orford S, Hubbart S, Foulkes J (2011) Identification of traits to improve the N-use efficiency of wheat genotypes. Field Crops Res 123:139–152

García-Suárez JV, Röder MS, Díaz De León JL (2010) Identification of QTLs and associated molecular markers of agronomic traits in wheat (*Triticum aestivum* L.) under two conditions of nitrogen fertilization. Cereal Res Commun 38:459–470

Gou J, Strauss SH, Tsai CJ, Fang K, Chen Y, Jiang X, Busov VB (2010) Gibberellins regulate lateral root formation in Populus through interactions with auxin and other hormones. Plant Cell 22:623–639

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

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

Habash D, Bernard S, Schondelmaier J, Weyen J, Quarrie S (2007) The genetics of nitrogen use in hexaploid wheat: n utilisation, development and yield. Theor Appl Genet 114:403–419

Hirel B, Le Gouis J, Ney B, Gallais A (2007) The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. J Exp Bot 58:2369–2387

Horvath A, Didier A, Koenig J, Exbrayat F, Charmet G, Balfourier F (2009) Diversity and linkage disequilibrium analysis along the chromosome 3B of bread wheat (*Triticum aestivum* L.) in contrasted plant materials. Theor Appl Genet 119:1523–1537

Johansson E, Prieto-Linde ML, Jönsson JÖ (2001) Effects of wheat cultivar and nitrogen application on storage protein composition and bread making quality. Cereal Chem 78:19–25

Kichey T, Heumez E, Pocholle P, Pageau K, Vanacker H, Dubois F, Le Gouis J, Hirel B (2006) Combined agronomic and

physiological aspects of nitrogen management in wheat (*Triticum aestivum* L.). Dynamic and integrated views highlighting the central role for the enzyme glutamine synthetase. New Phytol 169:265–278

Kichey T, Hirel B, Heumez E, Dubois F, Le Gouis J (2007) In winter wheat (*Triticum aestivum* L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. Field Crop Res 102:22–32

Kim W, Johnson JW, Baenziger PS, Lukaszewski AJ, Gaines CS (2004) Agronomic effect of wheat–rye translocation carrying rye chromatin (1R) from different sources. Crop Sci 44:1254–1258

Koch W, Kwart M, Laubner M, Heineke D, Stransky H, Frommer WB, Tegeder M (2003) Reduced amino acid content in transgenic potato tubers due to antisense inhibition of the leaf $H^+$/amino acid symporter StAAP1. Plant J 33:211–220

Krapp A, Saliba-Colombani V, Daniel-Vedele F (2005) Analysis of C and N metabolisms and of C/N interactions using quantitative genetics. Photosynth Res 83:251–263

Kuchel H, Williams K, Langridge P, Eagles HA, Jefferies SP (2007) Genetic dissection of grain yield in bread wheat II. QTL-by-environment interaction. Theor Appl Genet 115:1015–1027

Laperche A, Brancourt-Hulmel M, Heumez E, Gardet O, Le Gouis J (2006) Estimation of genetic parameters of a DH wheat population grown at different N stress levels characterized by probe genotypes. Theor Appl Genet 112:797–807

Laperche A, Brancourt-Hulmel M, Heumez E, Gardet O, Hanocq E, Devienne-Barret F, Le Gouis J (2007) Using genotype × nitrogen interaction variables to evaluate the QTL involved in wheat tolerance to nitrogen constraints. Theor Appl Genet 115: 399–415

Laperche A, Le Gouis J, Hanocq E, Brancourt-Hulmel M (2008) Modelling nitrogen stress with probe genotypes to assess genetic parameters and genetic determinism of winter wheat tolerance to nitrogen constraint. Euphytica 16:259–271

Law CN, Worland AJ, Giorgi B (1976) Genetic-control of ear-emergence time by chromosomes-5a and chromosomes-5D of wheat. Heredity 36:49–58

Le Gouis J, Beghin D, Heumez E, Pluchard P (2000) Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. Eur J Agron 12:163–173

Le Gouis J, Bordes J, Ravel C, Heumez E, Faure S, Praud S, Galic N, Remoué C, Balfourier F, Allard V, Rousset M (2011) Genome wide association analysis to identify chromosomal regions determining components of earliness in wheat. Theor Appl Genet. doi:10.1007/s00122-011-1732-3

Lee JH, Graybosch RA, Peterson CJ (1995) Quality and biochemical effects of 1BL/1RS wheat–rye translocation in wheat. Theor Appl Genet 90:105–112

Lid SE, Olsen L, Nestestog R, Aukerman M, Brown RC, Lemmon B, Mucha M, Opsahl-Sorteberg HG, Olsen OA (2005) Mutation in the *Arabidopisis thaliana* DEK1 calpain gene perturbs endosperm and embryo development while over expression affects organ development globally. Planta 221:339–351

Masood Quraishi U, Abrouk M, Murat F, Pont C, Foucrier S, Desmaizieres G, Confolent C, Rivière N, Charmet G, Paux E, Murigneux A, Guerreiro L, Lafarge S, Le Gouis J, Feuillet C, Salse J (2011) Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution. Plant J 65:745–756

McIntosh RA, Hart GE, Devos KM Gale MD, Rogers WJ (1998) Catalogue of gene symbols for wheat. In: Proceedings of the 9th international wheat genetics symposium, Saska toon, Canada, pp 1–235 (1998)

Neumann K, Kobiljski B, Dencie S, Varshney RK, Borner A (2011) Genome-wide association mapping: a case study in bread wheat (*Triticum aestivum* L.). Mol Breeding 27:37–58

Okumoto S, Koch W, Tegeder M, Fischer WN, Biehl A, Leister D, Stierhof YD, Frommer WB (2004) Root phloem-specific expression of the plasma membrane amino acid proton co-transporter AAP3. J Exp Bot 55:2155–2168

Oury FX, Godin C (2007) Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favorable genotypes? Euphytica 157:45–57

Oury FX, Bérard P, Brancourt-Hulmel M, Depatureaux C, Doussinault G, Galic N, Giraud A, Heumez E, Lecomte C, Pluchard P, Rolland B, Rousset M, Trottet M (2003) Yield and grain protein concentration in bread wheat: a review and a study of multi-annual data from a French breeding program. J Genet Breed 57:59–68

Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeyer W, Lagudah E, Somers D, Kilian A, Alaux M, Vautrin S, Bergès H, Eversole K, Appels R, Safar J, Simkova H, Dolezel J, Bernard M, Feuillet C (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. Science 322:101–104

Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on breadmaking quality. Annu Rev Plant Physiol 38:141–153

Pena RJ, Amaya A, Rajaram S, Mujeeb-Kazi A (1990) Variation in quality characteristics associated with some spring 1B/1R translocation wheats. J Cereal Sci 12:105–112

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256–261

Prasad M, Kumar N, Kulwal PL, Röder MS, Balyan HS, Dhaliwal HS, 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

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusic D, Waterman E, Weyen J, Schondelmaier J, Habash DZ, Farmer P, Saker L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti MC, Hollington PA, Aragues R, Royo A, Dodig D (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. Theor Appl Genet 110:865–880

Ravel C, Martre P, Romeuf I, Dardevet M, El-Malki R, Bordes J, Duchateau N, Brunel D, Balfourier F, Charmet G (2009) Nucleotide polymorphism in the wheat transcriptional activator Spa influences its pattern of expression and has pleiotropic effects on grain protein composition, dough viscoelasticity and grain hardness. Plant Physiol 151:33–44

Rousset M, Bonnin I, Remoué C, Falque M, Rhoné B, Veyrieiras J-B, Madur D, Balfourier F, Le Gouis J, Santoni S, Goldringer I (2011) Deciphering the genetics of growth habit by an association study on candidate genes in bread wheat (*Triticum aestivum* L.). Theor Appl Genet. doi:10.1007/s00122-011-1636-2

Roy JK, Bandopadhyay R, Rustgi S, Balyan HS, Gupta PK (2006) Association analysis of agronomically important traits using SSR, SAMPL and AFLP markers in bread wheat. Curr Sci 90:683–689

Sabelli PA, Larkins BA (2009) The development of endosperm in grasses. Plant Physiol 149:14–26

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

Veyrieras J-B, Goffinet B, Charcosset A (2007) MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. BMC Bioinformatics 8:1–16

Viglasi P (1968) Short strawed mutants of Karcagi 522 winter wheat induced by gamma rays. Acta Agron Hung 17:205–214

Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Weir BS (1996) Genetic data analysis II Methods for discrete population genetic data. Sinauer Associates Inc, Sunderlan, p 376

Wieser H, Kieffer R, Lelley T (2000) The influence of 1B/1R chromosome translocation on gluten protein composition and technological properties of bread wheat. J Sci Food Agric 80: 1640–1647

Willats WGT, McCartney L, Mackie W, Knox JP (2001) Pectin: cell biology and prospects for functional analysis. Plant Mol Biol 47:9–27

Witcombe JR, Hollington PR, Howarth CJ, Reader S, Steele KA (2008) Breeding for abiotic stresses for sustainable agriculture. Phil Trans R Soc Biol 363:703–716

Worland AJ, Sayer EJ, Borner A (1994) The genetics and breeding potential of Rht12, a dominant dwarfing gene in wheat. Plant Breed 113:187–196

Yamaguchi S (2008) Gibberellin metabolism and its regulation. Annu Rev Plant Biol 59:225–251

Zhang Y, Tang J, Zhang Y, Yan J, Xiao Y, Zhang Y, Xia X, He Z (2011) QTL mapping for quantities of protein fractions in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 122:971–987