Plant Biotechnology Journal

*Plant Biotechnology Journal* (2016) **14**, pp. 735–748 doi: 10.1111/pbi.12424

# A multiparental cross population for mapping QTL for agronomic traits in durum wheat (*Triticum turgidum* ssp. *durum*)

Sara Giulia Milner[1], Marco Maccaferri[1], Bevan Emma Huang[2], Paola Mantovani[3], Andrea Massi[3], Elisabetta Frascaroli[1], Roberto Tuberosa[1] and Silvio Salvi[1,*]

[1]*Department of Agricultural Sciences, University of Bologna, Bologna, Italy*
[2]*Digital Productivity Flagship and Agriculture Flagship, CSIRO, St Lucia, Qld, Australia*
[3]*Società Produttori Sementi Bologna, Argelato, Bologna, Italy*

*Received 5 November 2014;*
*revised 8 April 2015;*
*accepted 29 April 2015.*
**Correspondence* (Tel +39 0512096645; fax +39 0512096241; email silvio.salvi@unibo.it)

**Keywords:** multiparental cross, QTL, *Triticum turgidum*, agronomic traits, SNP, founder haplotypes.

## Summary

Multiparental cross designs for mapping quantitative trait loci (QTL) provide an efficient alternative to biparental populations because of their broader genetic basis and potentially higher mapping resolution. We describe the development and deployment of a recombinant inbred line (RIL) population in durum wheat (*Triticum turgidum* ssp. *durum*) obtained by crossing four elite cultivars. A linkage map spanning 2664 cM and including 7594 single nucleotide polymorphisms (SNPs) was produced by genotyping 338 RILs. QTL analysis was carried out by both interval mapping on founder haplotype probabilities and SNP bi-allelic tests for heading date and maturity date, plant height and grain yield from four field experiments. Sixteen QTL were identified across environments and detection methods, including two yield QTL on chromosomes 2BL and 7AS, with the former mapped independently from the photoperiod response gene *Ppd-B1,* while the latter overlapped with the vernalization locus *VRN-A3*. Additionally, 21 QTL with environment-specific effects were found. Our results indicated a prevalence of environment-specific QTL with relatively small effect on the control of grain yield. For all traits, functionally different QTL alleles in terms of direction and size of genetic effect were distributed among parents. We showed that QTL results based on founder haplotypes closely matched functional alleles at known heading date loci. Despite the four founders, only 2.1 different functional haplotypes were estimated per QTL, on average. This durum wheat population provides a mapping resource for detailed genetic dissection of agronomic traits in an elite background typical of breeding programmes.

## Introduction

Quantitative trait locus (QTL) analysis on multiparental populations occupies an intermediate niche between conventional linkage mapping based on biparental populations and association mapping in germplasm collections. The idea of combining the genomes of multiple founders to ensure segregation for multiple QTL was first developed in mouse genetics with the name of heterogeneous stocks (Demarest *et al.*, 2001; McClearn *et al.*, 1970), which could be considered as an extension of the biparental advanced intercross lines (AILs) approach (Darvasi and Soller, 1995) to include broader genetic diversity. Subsequently, a large multiparent recombinant inbred line (RIL) panel from eight laboratory inbred strains was assembled (Churchill *et al.*, 2004; Threadgill *et al.*, 2002). Multiparental strategies were also adapted to plant genetics via multiparent advanced generation intercrosses (MAGIC, Cavanagh *et al.*, 2008; Mackay and Powell, 2007) and interconnected populations via di-allelic schemes or star designs (Huang *et al.*, 2011; Yu *et al.*, 2008). Linkage mapping in biparental populations often detects QTL with large support intervals because of limited recombination (Doerge, 2002; Holland, 2007). On the other hand, association mapping, which relies on both wider genetic diversity and cumulated historical recombinations, may suffer from inferential problems caused by residual hidden population structure (Flint-Garcia *et al.*, 2003). Mapping of QTL in multiparental populations is expected to combine the advantages of the two approaches: high number of informative crossovers, possibility to interrogate multiple alleles, limited population structure effects (Rebai and Goffinet, 2000). So far, only a few balanced multiparental populations have been developed in *Arabidopsis thaliana* (Huang *et al.*, 2011; Kover *et al.*, 2009), tomato (Pascual *et al.*, 2014), bread wheat (Huang *et al.*, 2012; Mackay *et al.*, 2014; Rebetzke *et al.*, 2014; Thépot *et al.*, 2015) and rice (Bandillo *et al.*, 2013).

In durum wheat, elite breeding germplasm has undergone progressively higher selection pressure towards strict grain industrial quality standards, resulting in very low genetic diversity (De Vita *et al.*, 2007; Maccaferri *et al.*, 2005), probably the lowest among cultivated wheats ($\pi = 0.0004 \times 10^{-3}$; Haudry *et al.*, 2007). In this context, the potentially higher informativeness of multiparental populations provides interesting opportunities for QTL discovery.

Here, we report on the QTL analysis of the first durum wheat multiparental RIL population obtained by crossing four elite cultivars from different origins to include additional genetic diversity. The population was used to assemble a linkage map and to map QTL for heading and maturity date, plant height and grain yield. Molecular investigation at well-known photoperiod (*Ppd*) and vernalization (*VRN*) genes, pivotal for seasonal and environ-

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd

mental adaptation in wheat (Cockram *et al.*, 2007), was carried out to verify QTL positions and effects.

The ultimate aim of this effort was to gain a better understanding about the genetic architecture of complex traits in an elite genetic background typical of commercial breeding programmes in durum wheat.

## Results

### Genetic diversity among founders and genetic map

Based on high-density SNP genotyping information, diversity in the four-way cross population from Neodur, Claudio, Colosseo and Rascon/2*Tarro (hereafter NCCR) was 0.25, which compares with 0.32 estimated from a reference collection of 253 cultivated tetraploid wheat accessions from worldwide. Additionally, NCCR captured 76% of the SNP alleles from the reference collection, with a SNP rate of 10.3% (19.7% in the reference collection). These results indicate that the genetic diversity represented by the NCCR founders represents well the overall range of diversity in durum wheat elite germplasm.

The NCCR linkage map included 7594 SNPs and spanned 2664 cM (1188 cM for the A genome and 1476 cM for the B genome, details reported in Table 1 and File S1), distributed in 14 linkage groups. The map included 1221 unique genomic positions (bins), varying from 43 on chr. 1A to 134 on chr. 1B (87.8 per chromosome on average). The mean distance between unique positions was 2.3 cM, and only 40 marker-pair intervals were ≥20 cM (Figure S1). When comparing the NCCR map to both the bread wheat consensus and the durum wheat consensus maps, positions were mostly monotonically related, as showed by Spearman's marker order rank correlation values >0.95 and >0.99, respectively, for all chromosome pairs (Figures S2 and S3). Gaps in chromosome coverage of the NCCR map compared to the consensus maps were mainly found for chrs. 1A, 3B, 4A and 5A, extending from 12.1 to 51.5% (Figures S2 and S4). These gaps were likely due to identity-by-descent (IBD) relationships among the founders, as commonly observed in other durum wheat crosses (Maccaferri *et al.*, 2005, 2014). SNP density (expressed as SNP/cM) along chromosomes showed differences between the NCCR and durum consensus maps as compared to the bread wheat consensus (examples are shown in Figure 1. All comparisons are shown in Figure S5). The high SNP density at peri-centromeric regions that was observed in bread wheat was less pronounced or absent in both durum maps. Such differences between bread and durum wheat can be due to both: (i) ascertainment bias due to the use of a SNP array prevalently developed from bread wheat and (ii) local differences in recombination rate between durum and bread wheat, mostly in favour of durum wheat (lower occurrence of historical introgressions from wheat-relative species, which are known to cause strong perturbation of recombination rate).

### Pattern of genomic information

Along each chromosome, we measured local genomic information as the percentage of RILs for which alleles could be assigned to the founders and the observed SNP-based haplotype richness (expressed as ratio over the theoretical maximum, i.e. four haplotypes), using a sliding window of 10 cM (Figure S7). Additionally, we compared this to the genotypic information content (GIC), computed at all positions as in Text S1, based on the trait grain yield and other traits, but observed few differences, which may reflect the presence of predominantly bi-allelic QTL detected in our analyses. In all cases, the different genomic information measures showed similar patterns along the chromosomes (Figure S7), indicating that a primary driver was the variation of similarity between founder genomes along chromosomes. In regions where the founders are more similar, the respective allelic assignment is less clear, and the number of founder haplotypes in the region will also drop; the GIC mirrors this trend.

**Table 1** Summary of main parameters of the NCCR map

| Chr. | SNPs (no.) | Length (cM) | Bins* (no.) | Recomb./chr. (no.) |
|---|---|---|---|---|
| 1A | 281 | 125 | 43 | 1.8 |
| 1B | 864 | 176 | 134 | 2.4 |
| 2A | 695 | 209 | 99 | 3.0 |
| 2B | 542 | 264 | 101 | 2.4 |
| 3A | 622 | 205 | 104 | 3.3 |
| 3B | 304 | 254 | 64 | 1.2 |
| 4A | 359 | 138 | 68 | 1.7 |
| 4B | 552 | 176 | 105 | 2.6 |
| 5A | 177 | 99 | 53 | 1.6 |
| 5B | 408 | 186 | 65 | 2.3 |
| 6A | 708 | 126 | 91 | 1.5 |
| 6B | 719 | 196 | 101 | 2.7 |
| 7A | 632 | 286 | 82 | 3.5 |
| 7B | 731 | 224 | 119 | 1.6 |
| Overall | 7594 | 2664 | 1229 | 31.6 |

*Bins are defined as groups of cosegregating markers.

### Genetic structure of NCCR population

Average assignment of each founder along the chromosomes for the 338 RILs ranged from 5.5% (both Claudio and Rascon/2*Tarro haplotypes, chr. 3B) to 30.4% (Neodur haplotype, chr. 5A). On the whole, 20.1% of the genome could not be assigned at the chosen probability threshold (0.7) to any founder. Detailed plots of founder assignment percentages for all chromosomes are reported in Figure 2. Across RILs, the percentage of regions where founders could not be assigned varied from 9.0% (chr. 5A) to 46.1% (chr. 3B). Percentage of founder assignment was typically lower at centromeric regions, reflecting the lower SNP density and SNP-based haplotype density among the four founders (Figures 1 and S7). The genome-wide average number of recombination events per line was 31.5, while it varied from 1.2 (chr. 3B) to 3.5 (chr. 7A) per morgan (Table 1).

### Phenotypic data

In the NCCR population, heading date (HD), maturity date (MD), plant height (PH) and grain yield (GY) trait values were normally distributed and characterized by transgressive segregation with respect to parental lines (Figure 3). RILs ranged from 132.3 to 149.0 days for HD and from 180.4 to 193.3 days for MD. PH among RILs varied between 64.0 and 94.8 cm, while GY varied between 4.2 and 7.1 t/ha (Table 2). For all traits, highly significant differences among RILs were observed ($P < 10^{-3}$, Table 2). Heritability for the combined values over the four environments was 94.5% for HD, 86.6% for PH, 58.4% for MD and 47.7% for GY (Table 2). A significant correlation was observed between HD and MD ($r = 0.62$), GY and HD ($r = 0.26$) and GY and PH ($r = 0.33$). Combined and single-environment correlation values among traits are reported in Tables 3 and S1.

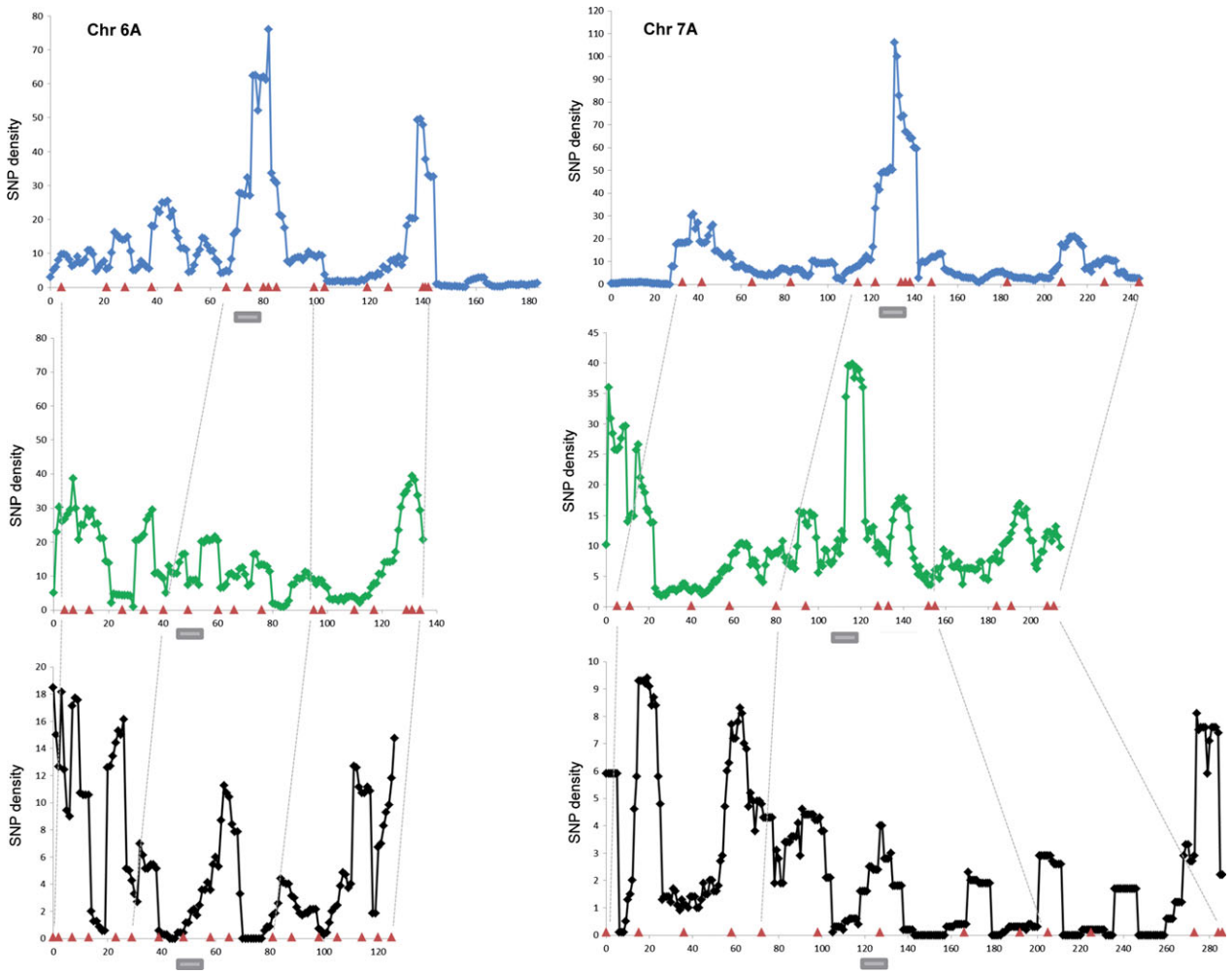


**Figure 1** Marker density along chr. 6A and chr. 7A. From top to bottom: bread wheat consensus map (blue-dotted plot. From Wang *et al.*, 2014), durum wheat consensus map (green-dotted plot. From Maccaferri *et al.*, 2014) and NCCR map (black-dotted plot; this study). *X*-axis: genetic map of the chromosome (cM); red triangles: position of anchor markers; *y*-axis: SNP density (1 cM-bin), sliding window of 7 cM. Centromere positions are shown as grey rectangles below the *x*-axis.

## Power simulation

A simulation study allowed us to define the minimum QTL size that can be detected in the NCCR population and the associated false-positive rate both in terms of detection power and precision across a range of explained percentage variance. Table 4 reports the results of this simulation study. In particular, reaching a power of 0.5 requires a minimum PVAR of 15%, which corresponds to a precision of 5 cM or less.

## Overview of QTL results

QTL analysis was carried out with two methods, IBD-CIM and IBS-SMA (see Experimental procedures). IBD-CIM combined across environments revealed a total of 18 QTL. The adjusted $R^2$ for the CIM full model was equal to 0.52 for HD, 0.40 for MD, 0.26 for PH and 0.12 for GY. IBS-SMA identified 29 QTL, including all but two of the 18 IBD-CIM QTL (Table 5; Figure 4), with cumulative $R^2$ values similar to those detected in the IBD-CIM analysis (+/− 0.01). Both methods identified the same four main chromosome regions (on chromosomes 2A, 2B distal, 2B proximal and 7A) enriched with QTL (QTL peaks within a 10-cM interval).

IBD-CIM on single-environment data identified 23, 15, 13 and nine QTL in Cad11, Cad12, Pr11 and Arg12, respectively (detailed results reported in Tables S2–S5), which could be grouped into 39 distinct QTL in total. Comparing the results of QTL analysis for the combined data with those obtained for single environments, we found that 14 of the 18 QTL detected with IBD-CIM across environments and 25 of 29 QTL identified with the IBS-SMA showed environmental specificity to some extent. Thereafter, we report and discuss QTL results from the combined analysis of data across environments only.

## QTL for phenology-related traits

Both IBD-CIM and IBS-SMA identified four HD QTL *(QHd.ubo-2A, QHd.ubo-2B, QHd.ubo-4B and QHd.ubo-7A.2*) with $R^2$ ranging from 2.9 to 24.7% (Table 5). IBS-SMA identified two additional HD QTL on chr. 6A ($R^2$ of 3.4%) and on chr. 7A ($R^2$ of 5.3%). All six QTL showed concomitant significant effects across environments for MD, with magnitude and ranking of effects similar to those observed for HD. For all but the QTL on chr. 4B, the same trend was observed for the corresponding MD QTL at target regions, with slight differences in relative effect sizes. QTL specific

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd, 

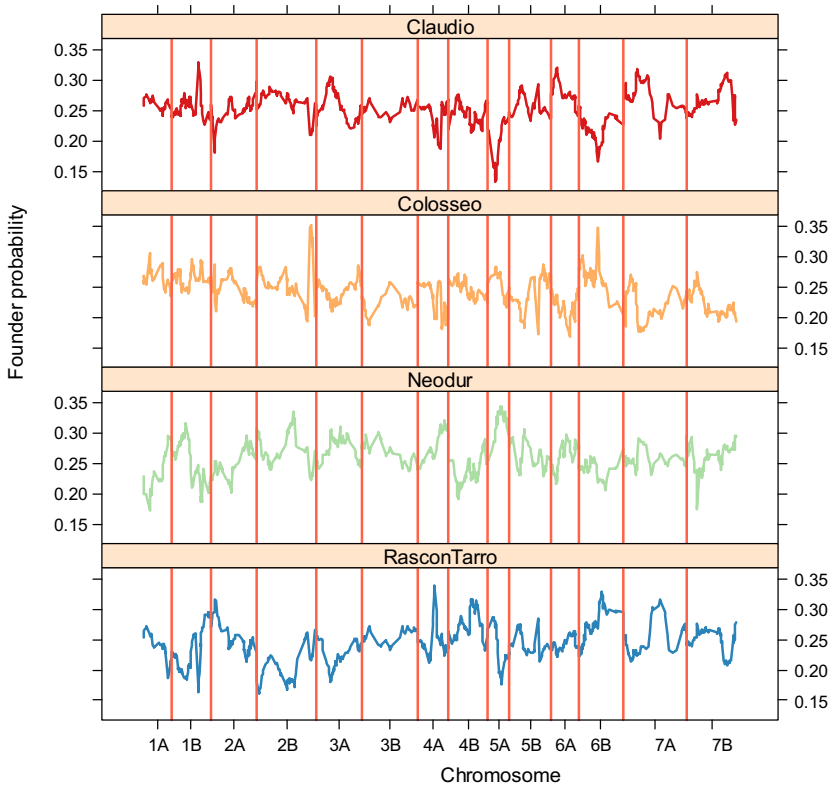


**Figure 2** Proportion of 338 RILs assigned to each of the four founders across the NCCR linkage map.

to MD and/or to MD and PH were also identified (*QMd.ubo-1B.1, QMd.ubo-1B.2, QMd.ubo-2A.2, QMd.ubo-2B.2, QMd.ubo-4B.2, QMd.ubo-5B* and *QMd.ubo-7B*). At these QTL, Rascon/2*Tarro was frequently associated with early maturity. Claudio contributed a strong late-maturity QTL (2a = 1.46 days. Claudio *vs.* Colosseo and Rascon/2*Tarro) at *QMd.ubo-7B,* while Neodur contributed late maturity at five of seven IBD-CIM MD QTL (Table 5).

## QTL for plant height

Five QTL for PH were identified based on both IBD-CIM and IBS-SMA mapping methods. These were located on chrs. 1B, 2B, 5B, 7A and 7B, with one of the largest ($R^2$ = 7.8%. 2a = −3.91 cm. Colosseo *vs.* others) on 7A, corresponding to *VRN-A3* (Table 5; Figure 4). Four additional small-effect QTL were identified by IBS-SMA only (Table 5).

## QTL for grain yield

Both IBD-CIM and IBS-SMA identified two GY QTL, *QGy.ubo-2B* and *QGy.ubo-7A,* mapped on chrs. 2B and 7A, respectively (Table 5). For *QGy.ubo-7A*, the IBD-CIM-based QTL profile showed multiple local peaks; therefore, to be conservative in our estimates of this QTL position, a confidence interval based on the preliminary simple interval mapping analysis was reported in Table 5 and Figure 4. At both *QGy.ubo-2B* and *QGy.ubo-7A,* the Colosseo haplotype affected GY negatively (−0.35 and −0.39 t/ha, respectively) relatively to the other three founders (with indistinguishable effect). Interestingly, these two QTL showed contrasting associations with PH and phenology. Although *QGy.ubo-2B* comapped with *QPh.ubo-2B.1* (one of the main PH QTL detected here), its association with phenology was mild (only a concomitant small effect for MD was detected). The QTL is well distinct in position from the *Photoperiod-B1* (*Ppd-B1*) locus. The direction of allelic effects was consistent with the effects observed for PH. On the other hand, *QGy.ubo-7A* belonged to a QTL cluster strongly driven by the *VERNALIZATION-A3* (*VRN-A3* = *TaFT*) locus, with clear effects on development and phenology. In all cases, Colosseo haplotypes differed from the other founders in terms of genetic effect.

## Comparison of functional haplotypes and SNP-based haplotypes at QTL

At each QTL detected with IBD-CIM, we grouped different founder haplotypes to common 'functional haplotypes' whenever we could not statistically differentiate their genetic effects (in terms of value). We then compared functional haplotypes with 'SNP-based haplotypes' based on identity-by- state haplotype observation within the QTL confidence intervals. Despite the NCCR population being based on four founders, for the 18 QTL mapped across environments, the mean number of functional QTL haplotypes per QTL was 2.1 (Table 5), with no QTL showing four functionally different haplotypes. Two QTL showed three haplotypes and 16 QTL showed two haplotypes. For the same set of QTL, we identified on average 3.6 SNP-based haplotypes per QTL. Altogether, SNP-based haplotypes outnumbered functional haplotypes in all cases except two. At no QTL did the two grouping criteria provide the same results, although concordance was observed (e.g. extremely functionally different haplotypes were always classified as molecularly different except at one QTL, *QHd.ubo-2B*).

As a case study, we compared QTL effects (i.e. functional haplotypes), SNP-based haplotypes and allelic variants at causal

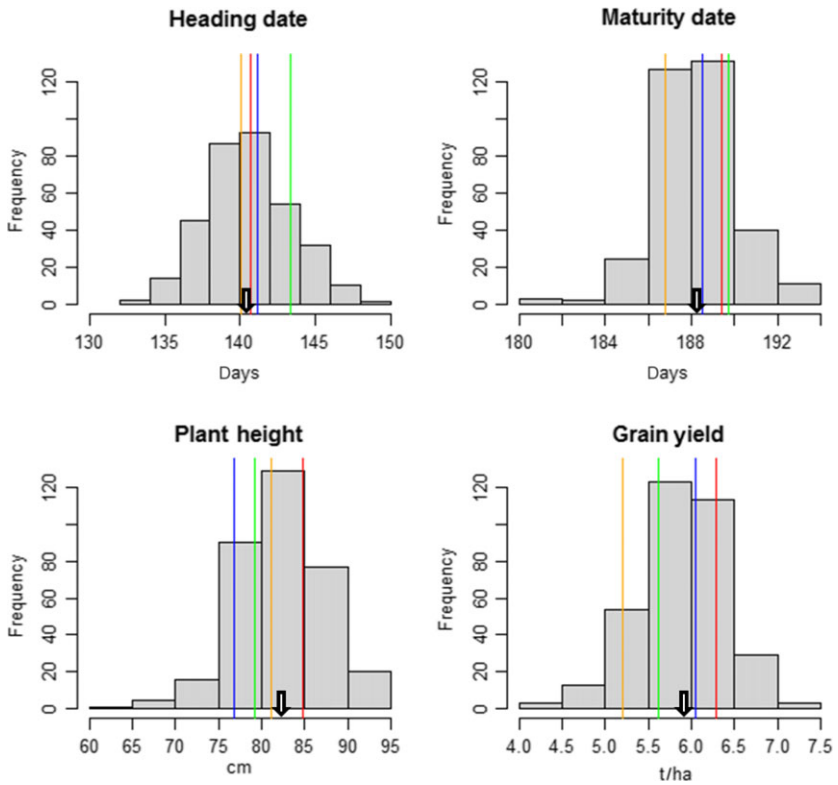


**Figure 3** RIL frequency distributions for the four traits investigated in this study. Mean values of the founder lines are identified by coloured bars: Neodur (green), Claudio (red), Colosseo (orange) and Rascon/2*Tarro (blue). For each trait, the population mean value is indicated by an arrow.

**Table 2** Summary of the analysis of phenotypic data, combined over years and environments

| Trait | Abbr. | Mean | Range | $MS_g$ | $MS_{g*e}$ | $h^2$ |
|---|---|---|---|---|---|---|
| Heading date | HD | 140.6 (days) | 132.3–149.0 | 61.9*** | 3.4*** | 94.4 |
| Maturity date | MD | 188.2 (days) | 180.4–193.3 | 22.8*** | 9.5*** | 58.4 |
| Plant height | PH | 82.1 (cm) | 64.0–94.7 | 153.1*** | 20.5*** | 86.6 |
| Grain yield | GY | 5.9 (t/ha) | 4.2–7.1 | 2.0*** | 1.0*** | 47.7 |

Abbr., used abbreviations; Mean, mean values; Range, value range; $MS_g$, mean square of RI lines; $MS_{g*e}$, mean square of the genotype by environment interaction; ***$P$-value <0.001; $h^2$, heritability of investigated traits.

**Table 3** Pearson correlation values and statistical significance calculated on the combined adjusted means of phenotypic data

| | HD | MD | PH |
|---|---|---|---|
| MD | 0.62*** | | |
| PH | 0.16** | 0.11* | |
| GY | 0.26** | 0.12* | 0.33** |

HD, heading date; MD, maturity date; PH, plant height; GY, grain yield; ***$P$-value <0.001; **$P$-value <0.01; *$P$-value <0.05.

genes for three main loci controlling HD (*QHd.ubo-2A*, *QHd.ubo-2B* and *QHd.ubo-7A.2)* that overlapped with *Ppd-A1*, *Ppd-B1* and *VRN-A3*, respectively (see Figure 5 with details reported in Text S2 and File S2). At *QHd.ubo-2A,* corresponding to *Ppd-A1*, allelic variation in the promoter region (Neodur and Colosseo *vs.* Claudio and Rascon/Tarro) concurred with the corresponding pattern of functional haplotypes. Additional significant differences for heading date between Neodur and Colosseo as captured by IBD-CIM were not detected by the allelic assay (Wilhelm *et al.*, 2009), while four haplotypes were identified based on SNP profiling on the confidence interval (Figure 5a). In this case, functional and SNP-based haplotypings are not necessarily conflicting as the four SNP-based haplotypes may only represent three functionally different alleles. Indeed, a shorter subportion of the confidence interval (as also outlined by the SNP-based QTL analysis—marker IWB67307. Figure 5a) indicates the

**Table 4** Power and precision in detection of QTL in the NCCR population for a range of percentage variance explained (PVAR). Power is measured as the proportion of replicates in which the most significant QTL was detected within 10 cM of the simulated location. The standard deviation (SD) of the distance from the simulated QTL is reported across all replicates where the most significant QTL was within 10 cM

| PVAR (%) | Power (ratio) | SD (cM) |
|---|---|---|
| 3 | 0.035 | 4.98 |
| 5 | 0.079 | 4.48 |
| 10 | 0.259 | 4.34 |
| 15 | 0.508 | 3.92 |
| 20 | 0.738 | 3.88 |

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd, 

**Table 5** Summary of QTL identified by IBD-CIM and IBS-SMA mapping methods for heading date (HD), maturity date (MD), plant height (PH) and grain yield (GY) across environments in the NCCR population

| | IBD-CIM | | | | | | | | | IBS-SMA | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | cM | c.i. | L Mrk–R Mrk | $R^2$ | LOD | 2a_Neo | 2a_Cla | 2a_Col | FH | cM | c.i. | Mrk | $R^2$ | LOD | a | Ref |
| HD | | | | | | | | | | | | | | | | |
| *QHd.ubo-2A* | 46.0 | 41.5–49.5 | IWB70098–IWB52303 | 24.7 | 31.6 | 1.94 | −0.45 | 3.05 | Col >N >Cla/RT | 42.0 | 7.3–51.5 | IWB67307 | 22.9 | >15.0 | −2.69 | Cla/RT |
| *QHd.ubo-2B* | 51.5 | 39.5–53.5 | IWB4604–IWA3868 | 14.1 | 19.7 | 1.60 | 1.73 | −1.28 | N/Cla >Col/RT | 57.0 | 39.5–57.0 | IWA8083 | 11.5 | 8.7 | −1.90 | Col/RT |
| *QHd.ubo-4B* | 59.1 | 49.1–70.1 | IWB6994–IWB2398 | 2.9 | 4.5 | 0.60 | −0.76 | 0.50 | N/Col/RT >Cla | 55.1 | 48.1–59.1 | IWA4916 | 4.3 | 3.2 | −1.17 | Cla/RT |
| *QHd.ubo-6A* | – | – | – | – | | – | – | – | – | 64.7 | 58.7–65.7 | IWA1235 | 3.4 | 2.5 | −1.12 | Cla/Col/RT |
| *QHd.ubo-7A.1* | – | – | – | – | | – | – | – | – | 19.7 | 0.0–48.7 | IWB54744 | 5.3 | 3.9 | 1.52 | N/Cla/RT |
| *QHd.ubo-7A.2* | 63.2 | 62.2–67.2 | IWA8390–IWB72200 | 15.8 | 21.7 | 0.15 | 0.45 | −2.55 | N/Cla/RT >Col | 65.7 | 59.7–98.8 | IWB56953 | 12.2 | 9.3 | 2.59 | N/Cla/RT |
| MD | | | | | | | | | | | | | | | | |
| *QMd.ubo-1B.1* | – | – | – | – | | – | – | – | – | 61.8 | 56.6–68.9 | IWB66172 | 5.7 | 4.2 | −0.96 | Cla/Col/RT |
| *QMd.ubo-1B.2* | – | – | – | – | | – | – | – | – | 130.5 | 122.7–134.5 | IWB6805 | 4.1 | 3.0 | −0.75 | Cla/RT |
| *QMd.ubo-2A.1* | 45.0 | 33.4–71.5 | IWB67517–IWB70098 | 17.7 | 19.2 | 0.70 | −0.30 | 1.72 | Col >N >Cla/RT | 42.0 | 7.3–45.0 | IWB67305 | 10.5 | 7.9 | 1.19 | Cla/RT |
| *QMd.ubo-2A.2* | – | – | – | – | | – | – | – | – | 83.6 | 73.0–91.1 | IWB1 | 5.8 | 4.2 | −1.04 | RT |
| *QMd.ubo-2B.1* | 51.5 | 39.5–63.1 | IWB4604–IWA3868 | 12.2 | 13.5 | 0.46 | 1.05 | −1.10 | N/Cla >Col/RT | 57.0 | 56.5–57.0 | IWA8083 | 9.3 | 7.0 | −1.13 | Col/RT |
| *QMd.ubo-2B.2* | – | – | – | – | | – | – | – | – | 97.9 | 87.3–97.9 | IWB5039 | 3.5 | 2.5 | 0.81 | N/Cla/RT |
| *QMd.ubo-4B.1* | – | – | – | – | | – | – | – | – | 49.6 | 48.1–59.1 | IWB72203 | 5.7 | 4.2 | −0.88 | Cla/RT |
| *QMd.ubo-4B.2* | 97.2 | 85.7–108.2 | IWB67166–IWB25207 | 4.8 | 6.1 | 0.49 | 0.64 | 1.16 | N/Cla. Col >RT | 101.2 | 92.2–101.2 | IWB66623 | 3.9 | 2.8 | −0.78 | RT |
| *QMd.ubo-5B* | 41.8 | 30.8–51.3 | IWA3870–IWB45033 | 4.5 | 5.4 | 0.81 | 0.97 | 0.36 | N/Cla >Col/RT | 42.3 | 39.8–42.3 | IWB37497 | 4.2 | 2.8 | −0.76 | Col/RT |
| *QMd.ubo-6A* | – | – | – | – | | – | – | – | – | 63.2 | 51.7–65.7 | IWA6724 | 3.9 | 2.9 | −0.81 | Cla/Col/RT |
| *QMd.ubo-7A.1* | – | – | – | – | | – | – | – | – | 48.7 | 35.7–49.7 | IWB71609 | 4.6 | 3.4 | −0.8 | Col/RT |
| *QMd.ubo-7A.2* | 63.2 | 58.7–67.2 | IWA8390–IWB72200 | 8.9 | 9.2 | 0.32 | 0.55 | −0.87 | N/Cla/RT >Col | 62.7 | 62.7–67.2 | IWB40574 | 5.3 | 3.9 | 1.05 | N/Cla/RT |
| *QMd.ubo-7A.3* | 110.9 | 98.9–163.1 | IWB319–IWB23989 | 3.1 | 4.1 | 1.00 | 0.32 | 1.23 | N/Col >Cla/RT | – | – | – | – | – | – | – |
| *QMd.ubo-7B* | 25.0 | 0–48.6 | IWA2568–IWA6901 | 2.5 | 3.3 | 0.27 | 1.46 | 0.00 | Cla >N/Co/RT | – | – | – | – | – | – | – |
| PH | | | | | | | | | | | | | | | | |
| *QPh.ubo-1B* | 41.4 | 31.3–51.9 | IWB66840–IWB8804 | 3.8 | 3.8 | 2.92 | 0.94 | 1.64 | N/Cla/Col >RT | 38.3 | 34.8–79.7 | IWB58817 | 3.8 | 2.7 | −2.24 | Cla/Col/RT |
| *QPh.ubo-2B.1* | 97.9 | 87.8–127.5 | IWB69270–IWB69796 | 8.1 | 8.0 | 1.05 | 0.12 | −2.91 | N/Cla/RT >Col | 96.9 | 87.8–97.9 | IWB69109 | 9.0 | 6.7 | 3.53 | N/Cla/RT |
| *QPh.ubo-2B.2* | – | – | – | – | | – | – | – | – | 127.5 | 127.5–151.0 | IWB20922 | 6.0 | 4.4 | 2.78 | N/Cla/RT |
| *QPh.ubo-4A* | – | – | – | – | | – | – | – | – | 120.0 | 116.5–131.5 | IWB23377 | 4.7 | 3.5 | −2.20 | N/RT |
| *QPh.ubo-5B* | 105.8 | 95.7–126.6 | IWB3693–IWB71533 | 4.6 | 4.7 | 0.14 | −2.48 | 0.15 | N/Col/RT >Cla | 105.3 | 96.7–113.8 | IWB11813 | 5.2 | 3.8 | 2.68 | N/Col/RT |
| *QPh.ubo-6A* | – | – | – | – | | – | – | – | – | 119.7 | 113.2–122.7 | IWA6537 | 5.3 | 3.8 | 2.79 | N/Cla/RT |
| *QPh.ubo-6B* | – | – | – | – | | – | – | – | – | 51.0 | 51.0–56.5 | IWA2451 | 3.5 | 2.4 | −2.30 | N/Cla/RT |
| *QPh.ubo-7A* | 67.2 | 60.7–76.8 | IWB67995–IWB33919 | 7.8 | 7.7 | −0.5 | 0.4 | −3.91 | N/Cla/RT >Col | 67.2 | 59.7–98.8 | IWA7301 | 8.4 | 6.0 | 3.89 | N/Cla/RT |
| *QPh.ubo-7B* | 138.4 | 104.8–180.5 | IWB10498–IWA7330 | 3.7 | 3.7 | −0.22 | 3.61 | 1.32 | Cla >N/Col/RT | 151.9 | 138.4–157.9 | IWB69205 | 3.5 | 2.6 | 2.24 | Cla/Col/RT |
| GY | | | | | | | | | | | | | | | | |
| *QGy.ubo-2B* | 103.4 | 87.8–127.5 | IWB69796–IWA5411 | 5.8 | 4.7 | −0.14 | −0.05 | −0.35 | N/Cla/RT >Col | 127.5 | 97.4–150.0 | IWB62437 | 5.3 | 3.8 | 0.23 | N/RT |
| *QGy.ubo-7A* | 110.9 | 98.8–123.4 | IWB319–IWB23989 | 8.8 | 7.1 | 0.16 | −0.11 | −0.39 | N/Cla/RT >Col | 60.7 | 59.7–87.8 | IWB30443 | 7.2 | 5.4 | 0.32 | N/Cla/RT |

HD, heading date; MD, maturity date; PH, plant height; GY, grain yield; cM, peak position in centimorgan (cM); c.i., QTL confidence interval (cM); L Mrk–R Mrk, left and right markers of the QTL peak position; $R^2$, proportion of phenotypic variance explained by individual QTL (%); LOD, $\log_{10}$ likelihood ratio of individual QTL; 2a_Neo, 2a_Cla, 2a_Col, founder genetic effect (2a) considering Rascon/2*Tarro as reference haplotype; FH, functional haplotypes at the QTL and ranking based on genetic effect (N, Neodur; Cla, Claudio; Col, Colosseo; RT, Rascon/2*Tarro); Mrk, marker with the highest significance; a, additive genetic effect (a); Ref, reference allele for computing the additive genetic effect (N, Neodur; Cla, Claudio; Col, Colosseo; RT, Rascon/2*Tarro).

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd. 


**Figure 4** NCCR linkage map reporting QTL positions for heading date (HD), maturity date (MD), plant height (PH) and grain yield (GY, data combined across environments. Green bars: QTL mapped by IBD-CIM. Red bars: QTL mapped by IBS-SMA (also denoted by '_t' suffix). For each QTL, bar and bar tick represent the LOD drop confidence interval and the LOD peak position, respectively.

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd, 

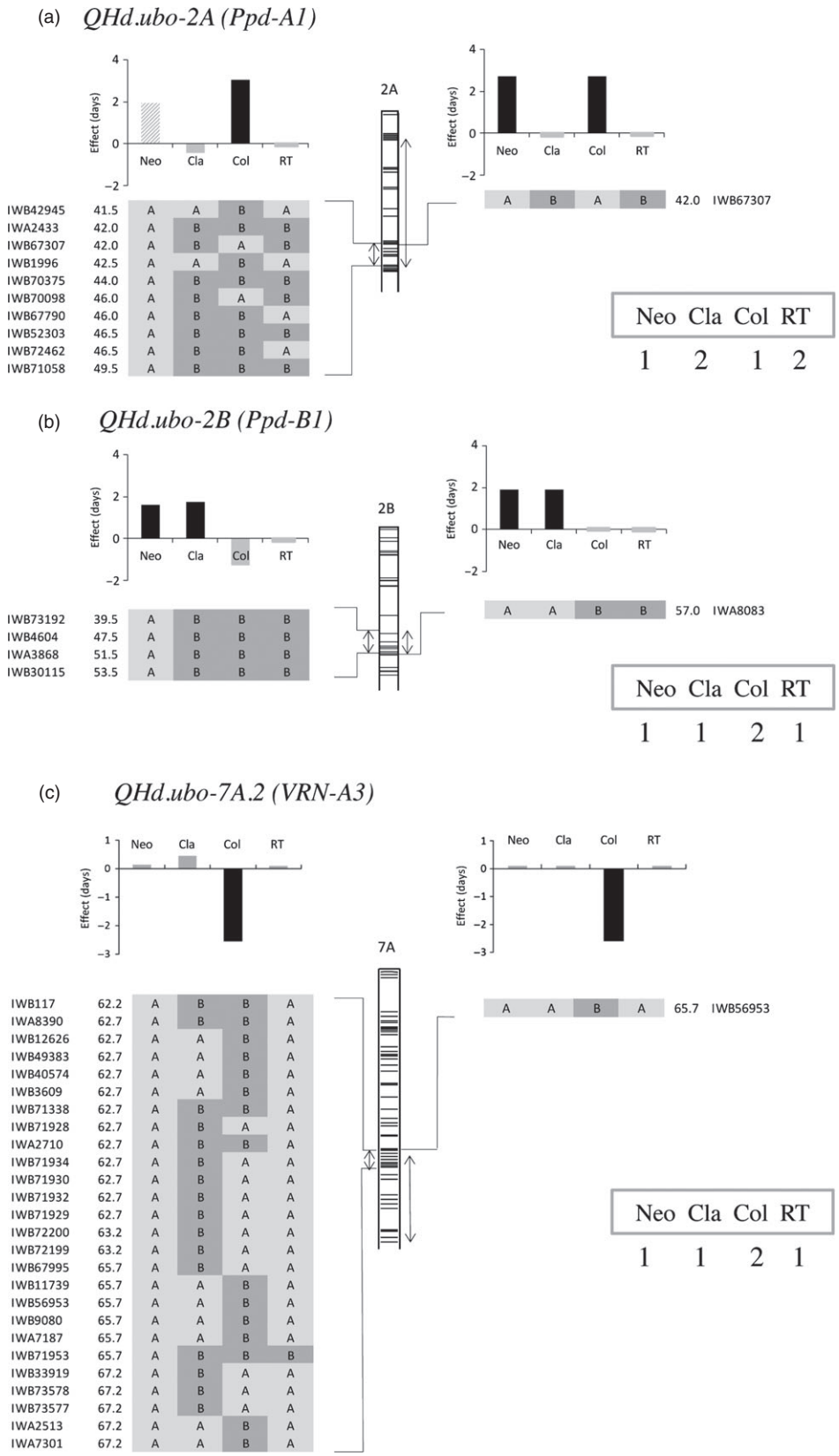


**Figure 5** Functional effects, SNP-based haplotypes and allelic causal variation at *QHd.ubo-2A* (a), *QHd.ubo-2B* (b) and *QHd.ubo-7A.2* (c). For each QTL, the functional (phenotypic) QTL effects as estimated by IBD-CIM and IBS-SMA are reported on the top left and top right, respectively (QTL effects were always estimated taking Rascon/2*Tarro (RT), as reference haplotype); SNP-based haplotypes in the QTL confidence interval are reported on the bottom left; parental alleles at the most significant SNP marker from the IBS-SMA are reported on central right; allelic variation observed at the causal gene is reported as number codes on the right-bottom corner of each subsection (different alleles are represented by different numbers).

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd, 

presence of SNP-based haplotypes fully concordant with functional haplotypes.

At *QHd.ubo-2B,* corresponding to *Ppd-B1*, the earliness effect of the Colosseo private allelic was associated with a gene copy number polymorphism (four copies *vs.* one copy for the other three parents, Figure S8), similarly to what reported in bread wheat (Díaz *et al.*, 2012). Notably, this was clearly captured only by the IBD-CIM method. In fact, the IBS-SIM method based on single bi-allelic SNP was unable to significantly detect the earliness effect of the Colosseo allele. Further functional variation at *Ppd-B1* other than copy number variation could explain the observed difference in QTL effect between the haplotypes of Neodur/Claudio and Rascon.

At *QHd.ubo-7A.2,* corresponding to *VRN-A3,* the two functional alleles observed at the phenotypic level (Colosseo *vs.* Neodur, Claudio and Rascon/Tarro) were completely characterized both by allelic variation in the promoter region (Figures 5 and S8) and by SNP haplotype variation.

## Discussion

### NCCR genetic map and population structure

As expected, our four-way cross showed higher diversity than that of biparental populations while capturing 76% of the diversity present in the elite durum wheat collection considered in this study. As a comparison, typical biparental durum wheat populations represented 46–61% diversity of the collection (see Maccaferri *et al.*, 2014). These results confirmed the high informativeness of the NCCR population for QTL mapping. The diversity of NCCR population was similar to that captured in an eight-way cross bread wheat population (Mackay *et al.*, 2014), whose parents represented 74% of the diversity captured by a 64-accession panel.

The NCCR genetic map spanned 2664 cM, similarly to the SNP-based durum wheat consensus map (2631 cM, Maccaferri *et al.*, 2014). The map was considerably longer than the single maps obtained from elite × elite biparental populations (Maccaferri *et al.*, 2014), and it is slightly longer (7.1%) than the A and B genomes of the bread wheat consensus map (Wang *et al.*, 2014). Overall, the quality of the NCCR map appeared satisfactory because the marker order matched closely with both durum and bread wheat consensus maps. This notwithstanding, our results confirmed that assembling a linkage map in a multiparental cross still presents difficulties. Ahfock *et al.* (2014) highlighted marker density and founder distribution patterns as potential factors affecting multiparental map construction. One particular issue affecting our study was the genealogical relationships between the four founders (i.e. the presence of shared chromosome regions, or IBD, see Maccaferri *et al.*, 2007), which inherently prevented perfect assignment of founder haplotypes and reduced accuracy in estimating crossover positions. As a consequence, fewer recombination events were observed than expected by simulation, as seen in other multiparental crosses (Huang *et al.*, 2012). Simulations conducted on hexaploid bread wheat and based on a consensus map of total length 2500 cM (Huang *et al.*, 2012) highlighted that a mean of 51 and 77 recombination events per biparental or four-parent RIL population, respectively, was expected, but the average observed number was lower. Such expected values, adjusted for the 14 durum chromosomes, approximate to 34.0 and 51.3 genome-wide recombination events for a biparental and a four-parental population, respectively.

### IBS *vs.* IBD approaches for mapping QTL in NCCR

Performing a QTL analysis with both IBD-CIM and IBS-SMA enabled us to compare the effectiveness of two contrasting QTL mapping methods. The main difference between the two approaches was in how marker information at each putative QTL position was modelled. The IBS-SMA analysis relied on independent tests at each bi-allelic marker position regardless of common-ancestor origin. The main drawback with this approach is that plants with identical marker alleles (IBS) are grouped together at tested SNP positions, irrespectively of possible low/incomplete linkage disequilibrium between markers and causal allelic variation (Würschum, 2012). In contrast, the IBD-CIM analysis relied on IBD estimates of founder haplotype probabilities, as it attempts to trace the parental origin of SNP alleles in population lines. This approach should have higher power in cases where QTL variation is multi-allelic, but loses power due to the additional degrees of freedom in the test when the QTL is in fact bi-allelic and/or in the presence of IBD relationships among founders.

Overall, the two QTL mapping methods mostly agreed in terms of QTL numbers, map positions, effects and proportion of explained variance. When we analysed the IBD-CIM results in terms of functional haplotypes, we identified on average 2.1 haplotypes per locus, a context which provides both IBD-CIM and IBS-SMA similar QTL detection power. This partially explains the similar results obtained with the two mapping methods. Nevertheless, a more detailed analysis of three QTL with known causal genes evidenced interesting differences between the two QTL analysis methods. In the case of *QHd.ubo-2B*, only IBD-CIM allowed tracing parents' molecular variation at *Ppd-B1*, unlike with the SNP-based founder haplotype pattern in the surrounding region. This was not the case for *QHd.ubo-7A.2*, where the alleles at the candidate causal locus (*VRN-A3*) appeared in complete linkage with the surrounding SNP-based haplotype.

### Analysis of QTL results

In NCCR population, a relatively low number of QTL showed significant and strong effects across environments, while most QTL showed relatively small and environment-specific effects.

This was not unexpected based on the low genetic diversity present in durum wheat cultivated germplasm, where only a limited number of major effect QTL are supposed to be still segregating, depending on the elite materials and target environments considered (reviewed in Cavanagh *et al.*, 2013; Collins *et al.*, 2008; Würschum, 2012). We compared our QTL results with a similar four-way study conducted in a bread wheat population of larger size (1100 RILs. Huang *et al.*, 2012). This population yielded nine PH QTL as compared to six in NCCR. The number of QTL identified in the two populations appears to be comparable based on the relative difference in population size.

To further assess the informativeness of QTL analysis in our population, we investigated the outcome of QTL mapping using both simulated and real data. In both cases, the population performed relatively well to detect larger effect QTL, but did not have high power to detect QTL accounting <10% of the phenotypic variation, most likely due to sample size. While Valdar *et al.* (2006) concluded that 500 RILs were sufficient to have high power to detect QTL explaining 10% of the phenotypic variance, our findings have shown different outcomes most likely reflecting number and choice of founders, background genetic effects, marker density, sample size and, particularly, the distinct features of elite germplasm in self-fertile crop species. The high levels of

similarity between founder genomes may make it difficult to detect QTL with private effects. Additionally, power is also expected to be reduced under more complex scenarios (i.e. QTL interactions), while in case of effects which are shared among founders, the power is expected to improve.

Considering GY, 2.8 QTL on average were identified in single environments, while only two were significant across environments with relatively minor effects (0.35–0.39 t/ha); additionally, they collectively explained a small percentage (12%) of the total phenotypic variation. Our results confirmed that GY and phenology QTL often overlap and that GY QTL not associated with phenology effects have small genetic effect in elite backgrounds (Maphosa *et al.*, 2014; Reynolds and Tuberosa, 2008; Simmonds *et al.*, 2014) with some notable exceptions (Maccaferri *et al.*, 2008). In our study, at least one GY *per se* (i.e. independent from phenology) QTL (*QGy.ubo-2B*) was identified and appears worthy of further characterization. Altogether, our results also suggest that genetic variation for GY in elite durum wheat could be exploited in a genomic selection framework (Desta and Ortiz, 2014) where statistically subsignificant marker–trait associations can be conveniently used in marker-assisted selection-based breeding programmes.

The importance of determining the actual QTL haplotypes segregating in relevant gene pools was first underlined in plant genomics-assisted breeding by Ching *et al.* (2002) in maize, and more in general by Peleman and Rouppe van der Voort (2003), who proposed the 'breeding by design' concept. Efforts in the fine-scale definition of haplotypes segregating in breeding germplasm, greatly facilitated by high-density SNP-based maps, have been carried out in maize (Gore *et al.*, 2009), rice (Yamamoto *et al.*, 2010) and soya bean (Zhang *et al.*, 2014). In our case, the founder haplotypes have been identified and reported for all detected QTL, including two major QTL for phenology and, importantly, one QTL for grain yield *per se*. These SNP-based haplotypes could be deployed to enhance genomics-assisted breeding of durum wheat.

## Conclusions

A large multiparental RIL population produced by crossing four elite durum wheat cultivars was assembled and utilized for identifying QTL for phenology and grain yield. Notably, the NCCR map confirmed the value of the multiparental cross to maximize the survey of genome diversity for QTL discovery in elite backgrounds. Detailed analysis of QTL effects evidenced 2.1 functional haplotypes per QTL (to a maximum of three). Despite this scenario, analysis based on founder haplotype probabilities proved to be more efficient in QTL mapping in our multiparental population as compared to a conventional bi-allelic assay. Our results have clearly confirmed the prevalence of relatively minor and environment-specific grain yield QTL in elite germplasm typically utilized in commercial breeding programmes. Although the relationship between phenology and yield was confirmed, it is noteworthy that a major QTL for grain yield *per se* across environments was identified on chromosome 2B. Our results indicate that this multiparental population provides valuable opportunities for the genetic dissection of agronomic important traits of breeding relevance in durum wheat.

## Experimental procedures

### Plant material and crossing design

A balanced, four-parental cross, hereafter identified as NCCR, was derived from four durum wheat cultivars (Neodur, Claudio, Colosseo and Rascon/2*Tarro) representing breeding lineages which differed as to both origin and phenotypes for traits of agronomic interest. Neodur (184.7/Valdur/Edmore) is a French photoperiod-sensitive late cultivar showing high number of spikelets per ear; Claudio (Sel. CIMMYT 35/Durango/IS1938/Grazia) shows wide adaptability to Southern Europe and resistance to drought and powdery mildew; Colosseo (Creso derived with Italian landrace introgressions) presents high-yielding ears (with balanced yield components); Rascon/2*Tarro (Rascon-37/2*Tarro-2, CIMMYT germplasm) is a photoperiod-insensitive cultivar with high yield potential due to an high number of grains per spikelet. The cultivars were crossed according to a pairwise scheme (Neodur/Claudio//Colosseo/(Rascon/2*Tarro)). Biparental $F_1$s were subsequently crossed to produce approximately 400 four-way $F_1$ hybrids which were advanced by single-seed descent to $F_7$, and seeds were finally bulked in the $F_8$ generation. A final population of 338 RILs was utilized in this study.

### Sample preparation and NCCR genotyping

For each NCCR RIL, DNA was extracted from four plants' leaf tissue (Qiagen DNeasy kit standard protocol by Qiagen, Valencia, CA) and normalized to 50 ng/μL. Genotyping was carried out by means of a wheat-dedicated SNP Illumina 90K array comprising 81 587 SNPs (Wang *et al.*, 2014). As automated Illumina allele calling was not feasible due to the tetraploidy of durum wheat, a specific cluster file was developed by manual adjustment of the Illumina Genome Studio cluster calling for each marker to correctly capture variation in data. The cluster file is available upon request to Dr. Martin Ganal (TraitGenetics GmbH, Gatersleben, Germany). Such a dedicated cluster file (Wang *et al.*, 2014) allowed us to score 25 631 SNP markers. After eliminating 6633 markers showing >10% missing data, we examined NCCR for potential segregation distortion. In NCCR, two types of segregation patterns were present due to the founder allele distributions, namely 1:1 (e.g. 0-1-0-1) and 1:3 (e.g. 0-1-1-1). Markers were therefore filtered allowing a distortion of +/− 30%.

### Estimation of founder genetic diversity

Diversity index (DI) among the four NCCR parents, the parents of previously established biparental RIL populations and the 253 accessions from the whole durum panel, was calculated as following: $DI = 1 - \sum_{i=1}^{k} p_i^2$ (Weir, 1990), where $p_i$ is the frequency of the *i*th allele.

The number of captured alleles was calculated as the number of observed alleles in both NCCR and biparental groups with regard to the whole durum panel. The polymorphic SNPs were counted as ratio over the 76 102 durum functional SNPs included among the wheat Illumina 90K array (Wang *et al.*, 2014).

### Map construction

The NCCR linkage map was assembled using mpMap (Huang and George, 2011). Maximum likelihood estimates of recombination fraction (*r*) between each SNP pair were obtained using the *mpestrf* function with default parameter settings. The markers were then assigned to linkage groups when the pairwise LOD score exceeded 5.0, producing 50 initial groups with the *mpgroup* function. The LOD and recombination fraction matrices were explored with interactive heatmaps in R (R Core Team, 2013), and the resulting 50 marker clusters were iteratively collapsed to 22 based on the closest linkage between clusters. The 22 linkage groups were then combined into 14 groups

corresponding to the durum wheat chromosomes based on markers in common with a recently released bread wheat (*T. aestivum*) consensus 90K SNP-based map (Wang *et al.*, 2014). Markers were ordered using the *mporder* function. Map positions were estimated using the function *computemap* (Haldane map function, maxOffset = 1) that allows for the imputation of missing *r* values based on the value at the most highly correlated marker. Centromeres' map positions were deduced from the bread wheat consensus. Based on the projection of the NCCR-mapped markers on the bread and durum wheat reference maps, the observed genome coverage of the NCCR map was estimated for each chromosome. Gaps were identified when relative distance between two adjacent common markers was higher than 5%. For each chromosome, the total sum of gap relative distances was estimated. Additionally, Spearman's rank correlation values between the two considered maps have been calculated for all chromosome pairs. The same analysis was carried out using a recently published durum wheat consensus map (Maccaferri *et al.*, 2014).

### Founder haplotype reconstruction

Founder haplotype probabilities were computed with the *mpprob* function in mpMap using a Hidden Markov model implemented in R/qtl (Broman *et al.*, 2003). Haplotypes were then reconstructed by imputing known founder alleles whenever the corresponding probabilities at a genomic location exceeded 0.7 for each RIL. The original bi-allelic SNP data were thus recoded on the basis of the most likely founder haplotype at each particular linkage block. Recombination events were declared when the founder allele changed along the RIL genome.

### Estimation of genomic information

Three indexes were used to describe the genomic information patterns along the chromosomes. To locally estimate haplotype diversity among founders (i.e. SNP-based haplotype density), we counted the number of observed haplotypes along each chromosome using a 10-cM sliding window; we referred this number to the theoretical maximum (equal to four haplotypes). The second index was the percentage of RILs (of 338) with a founder haplotype assigned based on an assignment threshold equal to 0.7. Finally, a genotypic information content (see GIC, MapQTL, Van Ooijen, 2004) was calculated along the chromosomes as reported in Text S1.

### Experimental design and field-data evaluation

The phenotypic evaluation was conducted in four field trials during two growing seasons (2010–2011 and 2011–2012) in three locations in the Po Valley: Cadriano (44°33′N lat., 11°24′E long.) in 2010–2011 (Cad11) and 2011–2012 (Cad12) with plot sizes of 2.4 $m^2$; Poggio Renatico (44°45′N, 11°25′E) in 2010–2011 (Pr11) with plots of 4 $m^2$; and Argelato (44°39′N, 11°20′E) in 2011–2012 (Arg12) with plots of 4 $m^2$. The 338 RILs, the four parents (two entries each) and five control cultivars Levante, Meridiano, Orobel, Saragolla and Svevo (three entries each) were evaluated in each environment according to an incomplete block design. A 19 × 19 α-lattice (Patterson and Williams, 1976) with two replications was considered for each environment. Blocking was designed by means of FACTEX procedure (SAS Institute Inc., 2011). The following traits were measured: heading date (HD, reported as number of days from sowing to heading), maturity date (MD, reported as number of days from sowing to maturity), plant height (PH, reported in cm) and grain yield (GY, reported in t/ha at 13% moisture). HD was recorded when 50% of the ears in a plot were fully emerged from the flag leaf, while MD was determined when 50% of the ear peduncles in a plot turned yellow. PH was measured at maturity from ground level to the top of the terminal spikelet (excluding awns) on four main culms per plot. Data for MD in Poggio Renatico and PH in Argelato were not available.

### Statistical analysis of phenotypic data

Analysis of variance per environment was performed with the LATTICE procedure (SAS Institute Inc., 2011) according to Cochran and Cox (1960). Mean genotype values were adjusted for the spatial design if the relative efficiency was greater than 105% when compared to that of a randomized complete block design. The ANOVA over environments was performed with the SAS GLM procedure based on the least square means of each environment *via* the SAS GLM procedure (SAS Institute Inc., 2011). After weighting the ANOVA for replications in each location, the residual pooled over environments was considered as the error term. Heritability on entry mean basis ($h^2$) for each trait was estimated across the four environments (*n*) and the two replications (*r*) per environment, as:

$$h^2 = \sigma_g^2/(\sigma_g^2 + \sigma_{g*e}^2/n + \sigma^2/nr)$$

and for the single environment, across the two replications (*r*) as:

$$h^2 = \sigma_g^2/(\sigma_g^2 + \sigma^2/r)$$

where $\sigma_g^2$ is the genotypic variance, $\sigma_{ge}^2$ is the genotype by environment interaction variance, and $\sigma^2$ is the residual error variance for each model. The descriptive statistics and the correlation values (Pearson's correlation coefficients) were calculated on the adjusted means.

### QTL mapping

QTL analysis over environments was conducted with two distinct approaches: (i) single marker analysis using the SNP bi-allelic classes (identity-by-state single marker analysis, IBS-SMA) and (ii) composite interval mapping exploiting the four founders' identity-by-descent haplotype probabilities (identity-by-descent composite interval mapping, IBD-CIM). IBS-SMA was performed at each SNP position using Tassel by fitting a general linear model (GLM) and an *F*-test (Bradbury *et al.*, 2007). IBD-CIM was performed in mpMap software (Huang and George, 2011). In IBS-SMA, to correct for multiple testing, a genome-wide significance level was calculated based on the effective number of independent tests, $M_{eff}$ (Cheverud, 2001; Nyholt, 2004). This procedure has been previously shown to significantly lessen the overly conservative Bonferroni's adjustment (Gao *et al.*, 2010). The effective number of independent tests (linkage disequilibrium blocks) was inferred from Haploview (Barrett *et al.*, 2005), with the four-gamete rule (threshold for haplotype frequency in the population as 20%). Genome-wide, we estimated 500 blocks, corresponding to a markerwise *P*-value $<10^{-4}$ and an experiment-wise *P*-value <0.05. QTL confidence intervals were computed using a sliding window (±four markers) applied to IBS-SMA LOD values, with the interval boundaries defined by a LOD drop below the arbitrary value of 1.44.

For interval mapping, a linear model was fit by estimating separated fixed effects for each of the four founders at each putative QTL position (Rascon/2*Tarro was arbitrarily chosen as

© 2015 Society for Experimental Biology, Association of Applied Biologists and John Wiley & Sons Ltd, 

the reference haplotype), using the *mpIM* function in mpMap (program 'qtl'). We first ran simple interval mapping (SIM) and subsequently included cofactors corresponding to the number of SIM-detected QTL in a composite interval mapping (CIM) approach (Jansen, 1993; Zeng, 1993) to account for background variation. Testing was performed on a reduced version of the NCCR map representing only unique genomic positions; at each location, we computed a Wald statistic for the overall significance of all founder effects. QTL were identified if the Wald statistic *P*-value was $<10^{-3}$. All QTL were simultaneously fit in a full model. The percentage of phenotypic variation accounted for by each individual QTL ($R^2$) was determined as the square of the semi-partial correlation coefficient, fitting the multiple regression model. This can be interpreted as the percentage of the variance due uniquely to that QTL, accounting for all other QTL in the model. *Post hoc t*-test on the founder effects and a Bonferroni correction ($\alpha = 0.05$) were used to estimate the number of statistically different classes of founder haplotypes at each QTL, which will be hereafter named functional haplotypes as their identification was based only on effect on trait and not on molecular genotype status. The 2-LOD confidence interval (for combined data) and the 1-LOD confidence interval (for single-environment data) were calculated based on transformed *P*-value [−log10(*P*)] profiles. SNP-based haplotypes at QTL regions were defined upon the pattern of SNP markers encompassing the QTL confidence intervals.

### Molecular characterization at *Ppd* and *VRN* target loci

*Ppd-A1* has been targeted following the assay by Wilhelm *et al.* (2009), while the *Ppd-B1* locus was characterized using a copy number variation (CNV) TaqMan® assay (Díaz *et al.*, 2012). CNV analysis was performed using a 7500 Fast Real-time PCR and a TaqMan® Universal PCR Master Mix, No AmpErase UNG (A. Biosystems, Foster City, CA). Raw data have been analysed by CopyCaller v2.0 (A. Biosystem). *VRN-A3* (*TaFT,* Yan *et al.*, 2006) has been characterized by means of an assay targeted to Colosseo's 1.475-bp deletion in the upstream of 5′-UTR region (Ricci *et al.*, unpublished data). Amplifications were undertaken using primers as follows:

wild type-specific Fwd1 (5′-CTCCGGAAATGCTCGCCAGG T-3′),

Colosseo-specific Fwd2 (5′-GAAGGCTCCCCATCAGTAACACT GGAA-3′) and

Rev (5′-CGAAACGTTAGCTTACACGGGACGCT-3′).

PCR protocol was as follows: 95 °C for 2 min and 35 cycles at 94 °C for 30 s, 67 °C for 20 s and 72 °C for 30 s, with a final extension at 72 °C for 5 min. PCR products were separated by 2%—agarose gel electrophoresis stained with ethidium bromide.

### Power simulation

We generated 2500 simulated data sets based on the true NCCR genotype data in the founders and progeny. For each replicate, we generated a phenotype by randomly selecting a SNP to be a QTL. This QTL was assigned a private single founder effect representing 3, 5, 10, 15 or 20% of the phenotypic variation; founder values at the SNP were based on the haplotype probabilities estimated from real data. For each simulated data set, we performed IBD QTL mapping using R/mpMap and recorded whether a QTL was detected within 10 cM of the true location, with a significance threshold of *P*-value <0.001. We estimated the power as the proportion of replicates in which such a peak occurred, and estimated the precision of QTL mapping in these replicates by considering the mean and standard deviation of the distance of the QTL mapping peak from the true QTL location.

## Acknowledgements

This research was supported by the EU FP6 BIOEXPLOIT project and the Ager Agroalimentare e Ricerca—Project 'From Seed to Pasta—Multidisciplinary approaches for a more sustainable and high quality durum wheat production'.

## Conflict of interest

The authors have no conflict of interest to declare.

## References

Ahfock, D., Wood, I., Stephen, S., Cavanagh, C.R. and Huang, B.E. (2014) Characterizing uncertainty in high-density maps from multiparental populations. *Genetics*, **198**, 117–128.

Bandillo, N., Raghavan, C., Muyco, P.A., Sevilla, M.A.L., Lobina, I.T., Dilla-Ermita, C.J., Tung, C., McCouch, S., Thomson, M., Mauleon, R., Singh, R.K., Gregorio, G., Redoña, E. and Leung, H. (2013) Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. *Rice*, **6**, 11.

Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633–2635.

Broman, K.W., Wu, H., Sen, S. and Churchill, G.A. (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, **19**, 889–890.

Cavanagh, C., Morell, M., Mackay, I. and Powell, W. (2008) From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* **11**, 215–221.

Cavanagh, C.R., Chao, S., Wang, S., Huang, B.E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G.L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., Lopez da Silva, M., Bockelman, H., Talbert, L., Anderson, J.A., Dreisigacker, S., Baenziger, S., Carter, A., Korzun, V., Laurent Morrell, P., Dubcovsky, J., Morell, M.K., Sorrells, M.E., Hayden, M.J. and Akhunov, E. (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl Acad. Sci. USA*, **110**, 8057–8062.

Cheverud, J.M. (2001) A simple correction for multiple comparisons in interval mapping genome scans. *Heredity*, **87**, 52–58.

Ching, A., Caldwell, K.S., Jung, M., Dolan, M., Smith, O.S., Tingey, S., Morgante, M. and Rafalski, A.J. (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* **3**, 19.

Churchill, G.A., Airey, D.C., Allayee, H., Angel, J.M., Attie, A.D., Beatty, J., Beavis, W.D., Belknap, J.K., Bennett, B., Berrettini, W., Bleich, A., Bogue, M., Broman, K.W., Buck, K.J., Buckler, E., Burmeister, M., Chesler, E.J., Cheverud, J.M., Clapcote, S., Cook, M.N., Cox, R.D., Crabbe, J.C., Crusio, W.E., Darvasi, A., Deschepper, C.F., Doerge, R.W., Farber, C.R., Forejt, J., Gaile, D., Garlow, S.J., Geiger, H., Gershenfeld, H., Gordon, T., Gu, J., Gu, W., de Haan, G., Hayes, N.L., Heller, C., Himmelbauer, H., Hitzemann, R., Hunter, K., Hsu, H.C., Iraqi, F.A., Ivandic, B., Jacob, H.J., Jansen, R.C., Jepsen, K.J., Johnson, D.K., Johnson, T.E., Kempermann, G., Kendziorski, C., Kotb, M., Kooy, R.F., Llamas, B., Lammert, F., Lassalle, J.M., Lowenstein, P.R., Lu, L., Lusis, A., Manly, K.F., Marcucio, R., Matthews, D., Medrano, J.F., Miller, D.R., Mittleman, G., Mock, B.A., Mogil, J.S., Montagutelli, X., Morahan, G., Morris, D.G., Mott, R., Nadeau, J.H., Nagase, H., Nowakowski, R.S., O'Hara, B.F., Osadchuk, A.V., Page, G.P., Paigen, B., Paigen, K., Palmer, A.A., Pan, H.J., Peltonen-Palotie, L., Peirce, J., Pomp, D., Pravenec, M., Prows, D.R., Qi,

Z., Reeves, R.H., Roder, J., Rosen, G.D., Schadt, E.E., Schalkwyk, L.C., Seltzer, Z., Shimomura, K., Shou, S., Sillanpää, M.J., Siracusa, L.D., Snoeck, H.W., Spearow, J.L., Svenson, K., Tarantino, L.M., Threadgill, D., Toth, L.A., Valdar, W., de Villena, F.P., Warden, C., Whatley, S., Williams, R.W., Wiltshire, T., Yi, N., Zhang, D., Zhang, M. and Zou, F.; Complex Trait Consortium. (2004) The collaborative cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* **36**, 1133–1137.

Cochran, W.G. and Cox, G.M. (1960) *Experimental Designs*. New York: John Wiley and Sons.

Cockram, J., Jones, H., Leigh, F.J., O'Sullivan, D., Powell, W., Laurie, D.A. and Greenland, A.J. (2007) Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *J. Exp. Bot.* **58**, 1231–1244.

Collins, N.C., Tardieu, F. and Tuberosa, R. (2008) Quantitative trait loci and crop performance under abiotic stress: where do we stand? *Plant Physiol.* **147**, 469–486.

Darvasi, A. and Soller, M. (1995) Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics*, **141**, 1199–1207.

De Vita, P., Li Destri Nicosia, O., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N. and Cattivelli, L. (2007) Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *Eur. J. Agron.* **26**, 39–53.

Demarest, K., Koyner, J., McCaughran, J. Jr, Cipp, L. and Hitzeman, R. (2001) Further characterization and high-resolution mapping of quantitative trait loci for ethanol induced locomotor activity. *Behav. Genet.* **31**, 79–91.

Desta, Z.A. and Ortiz, R. (2014) Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* **19**, 592–601.

Díaz, A., Zikhali, M., Turner, A.S., Isaac, P. and Laurie, D.A. (2012) Copy number variation affecting the *Photoperiod-B1* and *Vernalization-A1* genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLoS ONE*, **7**, e33234.

Doerge, R.W. (2002) Multifactorial genetics: mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.* **3**, 43–52.

Flint-Garcia, S.A., Thornsberry, J.M. and Buckler, E.S. (2003) Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* **54**, 357–374.

Gao, X., Becker, L.C., Becker, D.M., Starmer, J.D. and Province, M.A. (2010) Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet. Epidemiol.* **34**, 100–105.

Gore, M.A., Chia, J.M., Elshire, R.J., Sun, Q., Ersoz, E.S., McMullen, M.D., Grills, G.S., Ross-Ibarra, J., Ware, D.H. and Buckler, E.S. (2009) A first-generation haplotype map of maize. *Science*, **326**, 1115–1117.

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S. and David, J. (2007) Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Mol. Biol. Evol.* **24**, 1506–1517.

Holland, J.B. (2007) Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* **10**, 156–161.

Huang, B.E. and George, A.W. (2011) R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines. *Bioinformatics*, **27**, 727–729.

Huang, X., Paulo, M.J., Boer, M., Effgen, S., Keizer, P., Koornneef, M. and Eeuwijk, F.V. (2011) Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. *Proc. Natl Acad. Sci. USA*, **108**, 4488–4493.

Huang, B.E., George, A.W., Forrest, K.L., Kilian, A., Hayden, M.J., Morell, M. and Cavanagh, C. (2012) A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnol. J.* **10**, 826–839.

Jansen, R.C. (1993) Interval mapping of multiple quantitative trait loci. *Genetics*, **135**, 205–211.

Kover, P.X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I.M., Purugganan, M.D., Durrant, C. and Mott, R. (2009) A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* **5**, e1000551.

Maccaferri, M., Sanguineti, M.C., Noli, E. and Tuberosa, R. (2005) Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breeding*, **15**, 271–290.

Maccaferri, M., Sanguineti, M.C., Xie, C., Smith, J.S.C. and Tuberosa, R. (2007) Relationships among durum wheat accessions. II. A comparison of molecular and pedigree information. *Genome*, **50**, 385–399.

Maccaferri, M., Sanguineti, M.C., Corneti, S., Araus Ortega, J.L., Ben Salem, M., Bort, J., DeAmbrogio, E., del Moral, L.F.G., Demontis, A., El-Ahmed, A., Maalouf, F., Machlab, H., Martos, V., Moragues, M., Motawaj, J., Nachit, M., Nserallah, N., Ouabbou, H., Royo, C., Slama, A. and Tuberosa, R. (2008) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics*, **178**, 489–511.

Maccaferri, M., Ricci, A., Salvi, S., Milner, S.G., Noli, E., Martelli, P., Casadio, R., Akhunov, E., Scalabrin, S., Vendramin, V., Ammar, K., Blanco, A., Desiderio, F., Distelfeld, A., Dubcovsky, J., Fahima, T., Faris, J., Korol, A., Massi, A., Mastrangelo, A.M., Morgante, M., Pozniak, C., N'Diaye, A., Xu, S. and Tuberosa, R. (2014) A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* **13**, 648–663.

Mackay, I. and Powell, W. (2007) Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* **12**, 57–63.

Mackay, I.J., Bansept-Basler, P., Barber, T., Bentley, A.R., Cockram, J., Gosman, N., Greenland, A.J., Horsnell, R., Howells, R., O'Sullivan, D.M., Rose, G.A. and Howell, P.J. (2014) An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties and validation. *G3*, **4**, 1603–1610.

Maphosa, L., Langridge, P., Taylor, H., Parent, B., Emebiri, L.C., Kuchel, H., Reynolds, M.P., Chalmers, K.J., Okada, A., Edwards, J. and Mather, D.E. (2014) Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. *Theor. Appl. Genet.* **127**, 1607–1624.

McClearn, G.E., Wilson, J.R. and Meredith, W. (1970) The use of isogenic and heterogenic mouse stocks in behavioral research. In *Contributions to Behavior-Genetic Analysis: The Mouse as Prototype* (Lindzey, G. and Thiessen, D.D., eds), pp. 3–22. New York: Appleton Century Crofts.

Nyholt, D.R. (2004) A simple correction for multiple testing for single nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* **74**, 765–769.

Pascual, L., Desplat, N., Huang, B.E., Desgroux, A., Bruguier, L., Bouchet, J.-P., Le, Q.H., Chauchard, B., Verschave, P. and Causse, M. (2014) Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnol. J.* **13**, 565–577.

Patterson, H.D. and Williams, E.R. (1976) A new class of resolvable incomplete block designs. *Biometrika*, **63**, 83–92.

R Core Team. (2013) *R: A Language and Environment for Statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Peleman, J.D. and Rouppe van der Voort, J. (2003) Breeding by design. *Trends Plant Sci.* **8**, 330–334.

Rebai, A. and Goffinet, B. (2000) More about quantitative trait locus mapping with diallel designs. *Genet. Res.* **75**, 243–247.

Rebetzke, G.J., Verbyla, A.P., Verbyla, K.L., Morell, M.K. and Cavanagh, C.R. (2014) Use of a large multiparent wheat mapping population in genomic dissection of coleoptile and seedling growth. *Plant Biotechnol. J.* **12**, 219–230.

Reynolds, M. and Tuberosa, R. (2008) Translational research impacting on crop productivity in drought-prone environments. *Curr. Opin. Plant Biol.* **11**, 171–179.

Ricci, A., Maccaferri, M., Salvi, S., Sanguineti, M.C., Stefanelli, S., Corneti, S., Cattivelli, L. and Tuberosa, R. Genetic analysis of heading time, grain weight, grain yield and grain yield components in colosseo lloyd durum wheat (*Triticum turgidum* ssp. *durum*) population. Unpublished.

SAS Institute Inc. (2011) Base SAS® 9.3 Procedures Guide. Cary, NC: SAS Institute Inc.

Simmonds, J., Scott, P., Leverington-Waite, M., Turner, A.S., Brinton, J., Korzun, V., Snape, J. and Uauy, C. (2014) Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biol.* **14**, 191.

Thépot, S., Restoux, G., Goldringer, I., Hospital, F., Gouache, D., Mackay, I. and Enjalbert, J. (2015) Efficiently tracking selection in a multiparental population: the case of earliness in wheat. *Genetics*, **199**, 609–623.

Threadgill, D.W., Hunter, K.W. and Williams, R.W. (2002) Genetic dissection of complex and quantitative traits: from fantasy to reality via a community effort. *Mamm. Genome*, **13**, 175–178.

Valdar, W., Flint, J. and Mott, R. (2006) Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics*, **172**, 1783–1797.

Van Ooijen, J.W. (2004) *MapQTL® 5, Software for the Mapping of Quantitative Trait Loci in Experimental Populations*. Wageningen: Kyazma B.V.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M., Salvi, S., Milner, S.G., Cattivelli, L., Mastrangelo, A.M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., International Wheat Genome Sequencing Consortium, Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A.R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K.J., Hayden, M. and Akhunov, E. (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **12**, 787–796.

Weir, B.S. (1990) *Genetic Data Analysis*. Sunderland, MA: Sinauer Associates.

Wilhelm, E.P., Turner, A.S. and Laurie, D.A. (2009) Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (*Triticum durum* Desf.). *Theor. Appl. Genet.* **118**, 285–294.

Würschum, T. (2012) Mapping QTL for agronomic traits in breeding populations. *Theor. Appl. Genet.* **125**, 201–210.

Yamamoto, T., Nagasaki, H., Yonemaru, J., Ebana, K., Nakajima, M., Shibaya, T. and Yano, M. (2010) Fine definition of the pedigree haplotypes of closely related rice cultivars by means of genome-wide discovery of single-nucleotide polymorphisms. *BMC Genom.* **11**, 267.

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M. and Dubcovsky, J. (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of FT. *Proc. Natl Acad. Sci. USA*, **103**, 19581–19586.

Yu, J., Holland, J.B., Mc Mullen, M.D. and Buckler, E.S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics*, **178**, 539–551.

Zeng, Z. (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc. Natl Acad. Sci. USA*, **90**, 10972–10976.

Zhang, D., Kan, G., Hu, Z., Cheng, H., Zhang, Y., Wang, Q., Wang, H., Yang, Y., Li, H., Hao, D. and Yu, D. (2014) Use of single nucleotide polymorphisms and haplotypes to identify genomic regions associated with protein content and water-soluble protein content in soybean. *Theor. Appl. Genet.* **127**, 1905–1915.

## Supporting information

Additional Supporting information may be found in the online version of this article:

**Figure S1** Classes of inter-marker distances (in cM), NCCR map.
**Figure S2** Scatterplots representing the relationship between the NCCR map (x) and the bread wheat consensus map (y) for the 14 chromosomes.
**Figure S3** Scatterplots representing the relationship between the NCCR map (x) and the durum wheat consensus map, (y) for the 14 chromosomes.
**Figure S4** Gaps in chromosome coverage of the NCCR map. Gaps were calculated projecting NCCR mapped markers on the bread wheat reference map (a) and durum wheat reference map (b), considering intervals of 5 cM.
**Figure S5** Marker density of the bread wheat consensus map (blue-dotted plot), durum wheat consensus map (green-dotted plot) and the NCCR map (black-dotted plot) for the 14 chromosomes. *x*-axis: genetic map of the chromosome (cM); *y*-axis: SNP density (1 cM-bin), sliding window of 7 cM. Anchor markers (monotonically increasing) are indicated by red triangles. Centromere positions are shown as grey rectangles below the *x*-axis.
**Figure S6** NCCR linkage map reporting QTL positions for heading date (HD), maturity date (MD), plant height (PH) and grain yield (GY). Green bars: QTL mapped by IBD-CIM, data combined across environments. Red bars: QTL mapped by IBS-SMA (also denoted by '_t' suffix), data combined across environments. Single-environment QTL mapping results are represented as following: brown bars, Cad11; blue bars, Cad12; pink bars, Pr11 and pale green bars, Arg12. For each QTL, bar and bar tick represent the LOD drop confidence interval and the LOD peak position, respectively.
**Figure S7** Map informativeness along the chromosomes. Black line: SNP-based haplotype density (number of haplotypes/number of founders, 10-cM sliding window). Red line: percentage of 338 RILs with a founder haplotype assigned. Blue line: Genomic Information Content estimated by the multi-allelic-GIC formulation based on the GY full model. Centromere positions are shown as grey rectangles.
**Figure S8** Allelic variation at *QHd.ubo-2A* (1), *QHd.ubo-2B* (2) and *QHd.ubo-7A.2* (3).
**File S1** NCCR genetic map.
**File S2** Co-linearity analysis including wheat-barley-*Brachypodium*-rice for three NCCR heading date QTL based on the gene-associated SNPs mapped in the QTL confidence intervals.
**Table S1** Pearson correlation values and statistical significance calculated on single-environment data.
**Table S2** Single-environment Cad11 QTL detected by IBD-CIM.
**Table S3** Single-environment Cad12 QTL detected by IBD-CIM.
**Table S4** Single-environment Pr11 QTL detected by IBD-CIM.
**Table S5** Single-environment Arg12 QTL detected by IBD-CIM.
**Text S1** Estimation of the genotypic information content (GIC).
**Text S2** Co-location and co-linearity analysis of three NCCR heading date QTL.