




# Unraveling diversity in wheat competitive ability traits can improve integrated weed management

Mariateresa Lazzaro<sup>1</sup> · Paolo Bàrberi<sup>1</sup>  · Matteo Dell'Acqua<sup>1</sup> · Mario Enrico Pè<sup>1</sup> · Margherita Limonta<sup>2</sup> · Delfina Barabaschi<sup>3</sup> · Luigi Cattivelli<sup>3</sup> · Paolo Laino<sup>4</sup> · Patrizia Vaccino<sup>5</sup>

Accepted: 29 November 2018

© INRA and Springer-Verlag France SAS, part of Springer Nature 2019

## Abstract

Weed pressure can be high in organic and low-input farming and reduce yield and produce quality. In these systems, integrated weed management includes different agronomic practices but rarely focuses on the use of more competitive cultivars, which would reduce reliance on direct weed control methods and their detrimental effects on soil and the environment. We characterized 160 common wheat (*Triticum aestivum* L.) accessions cultivated in Italy since the nineteenth century for four traits linked to competitive ability against weeds (above-ground biomass before stem elongation, tillering index, plant height, and flag leaf morphology) and for two production-related traits (grain yield and thousand-kernel weight). This approach aimed to identify the most suitable combinations of competitiveness and production traits, which often show trade-offs, and led to the identification of eight accessions with reduced grain yield to plant height trade-off. We genotyped the collection with SNP markers, revealing high molecular diversity and highlighting a trend of polymorphism loss passing from heritage to modern germplasm, with the presence of unique polymorphisms in both groups. These results underline the importance of studying both heritage and elite germplasm when focusing on traits that are not targeted by formal breeding, such as the competitive ability against weeds. Marker-trait associations (MTAs) with false discovery rates (FDR) < 5% were detected for all traits studied, while MTAs with FDR < 1% were detected for plant height, biomass, grain yield, and thousand-kernel weight. We identified MTAs confirming associations already reported in the literature as well as MTAs pinpointing new genomic regions that may disclose new breeding perspectives in common wheat. This study, for the first time, shows the high potential of interdisciplinary research bridging advanced genetic studies with agroecological approaches for selecting more competitive common wheat germplasm as additional tool in more sustainable integrated weed management systems.

**Keywords** Crop-weed interaction · Weed control · Landraces · Low-input breeding · Organic breeding · Genome-wide association · Marker-trait associations · Quantitative trait loci · *Triticum aestivum* L

✉ Paolo Bàrberi  
paolo.barberi@santannapisa.it

<sup>1</sup> Scuola Superiore Sant'Anna, Institute of Life Sciences, Piazza Martiri della Libertà 33, 56127 Pisa, Italy

<sup>2</sup> Atlas S.R.L., via Maestri del Lavoro 24/26, 26866 Sant'Angelo Lodigiano, Italy

<sup>3</sup> Consiglio per la Ricerca in agricoltura e l'analisi dell'economia Agraria, Research Centre for Genomics and Bioinformatics, Via San Protaso 302, 29017 Fiorenzuola d'Arda, Italy

<sup>4</sup> ICE SpA, via Sicilia, 10, 42122 Reggio Emilia, Italy

<sup>5</sup> Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Research Centre for Cereal and Industrial Crops, S.S. 11 for Torino km 2.5, 13100 Vercelli, Italy

## 1 Introduction

In organic and low-input wheat production, weed pressure can be high and control must be achieved without or with limited use of herbicides. Weed management is a key intervention for leveling the yield gap between organic and conventional wheat production, as weed control is one of the most significant agronomic problems faced by organic growers (Bàrberi 2002). In systems not using herbicides, weeds are controlled primarily through mechanical tools and agronomic practices, including the use of crop rotations, crop diversification, cover crops, and intercropping (Bàrberi 2002); however, in a weed management strategy fully based on agroecological principles, the use of more competitive cultivars should also be included. Traits such as tillering capacity, biomass accumulation before the stem

elongation phase, final plant height, and flag leaf morphology are functional in determining the competitiveness of wheat against weeds (Andrew et al. 2015). Crop breeding can be used in combination with agronomic approaches for obtaining a comprehensive integrated weed management strategy less reliant on direct weed control methods with potential negative effects on soil and the environment (Ahlgren 2004). Most of the traits related to competitive ability have a great potential for breeding towards improved weed control; however, they are time consuming to assess in breeding programs and are usually of little interest in conventional agriculture where weed control is mainly based on herbicide use. Also, the trade-off between productivity in weed-free situations and competitive ability is a main obstacle to the release of competitive cultivars (Andrew et al. 2015). Yield potential remains a primary criterion for cultivar choice by farmers hence it would be desirable to prioritize traits that confer higher competitive ability without incurring in considerable grain yield reduction in the absence of weed competition.

In the last decade, the availability of high resolution and cost-effective genotyping platforms have opened the way to genome-wide association studies (GWAS) aiming at identifying the genetic bases of complex traits. By exploiting linkage disequilibrium between markers and quantitative trait loci (QTL) controlling plant traits across all chromosomes, GWAS aims at identifying the genetic basis of complex phenotypes in natural or ad hoc-generated populations. GWAS has been widely adopted in different plant species to overcome some of the constraints inherent to bi-parental linkage mapping, such as the limited genetic diversity explored (Barabaschi et al. 2016). GWAS has been successfully used in both durum and common wheat, mainly focused to resistance to pathogens (e.g., Kidane et al. 2017), grain yield, heading date, plant height, number of fertile tillers, biomass, and number of seeds per spike (Zanke et al. 2014a, b, 2015; Mengistu et al. 2016).

In this study, we used a common wheat germplasm collection representing the outcome of wheat breeding in Italy across more than 100 years (Fig. 1).

Crop competitive ability against weeds was not a target for formal wheat breeding, initiated in Italy by Nazareno Strampelli, whose main objectives were to increase productivity, avoid heat-stress at grain filling, and reduce crop lodging (Salvi et al. 2013). Subsequent formal breeding programs aimed mainly at providing adequate cultivars for being grown in conventional farming systems where weeds are controlled by herbicides. The collection used in this work represents a unique resource for studying traits related to crop-weed interaction, as it comprises landraces, selected lines, and modern cultivars, which are expected to show different combinations of these traits.

In the current study, we screened the collection for four traits related to competitive ability (above-ground biomass before stem elongation phase, tillering index, plant height, and flag leaf morphology), and for two production-related traits (grain



**Fig. 1** View of the field experiment with the common wheat collection in the season 2013/14. Photo: P. Laino

yield and thousand-kernel weight). In addition, we conducted a GWAS on these traits for further advancing the knowledge about the genetic determination of competitive ability, and for increasing information about the trade-off between grain yield and the competitive ability of common wheat.

Our interdisciplinary work was conducted with the aim to: (i) identify germplasm with high competitive ability potential and with a combination of productivity and competitiveness traits adequate to low-input and organic farming; (ii) study the structure of the collection at molecular level; (iii) identify marker-trait associations (MTAs) for traits related to crop-weed interference.

## 2 Materials and methods

### 2.1 Plant material

We studied a collection of 160 wheat accessions (Table 1) that summarizes the typologies of common wheat cultivars grown in Italy from the nineteenth century until 2009 (Ormoli et al. 2015; Laino et al. 2015). Accessions were organized into six groups (A to F) based on the genealogical information available and the date of commercial release. Group A includes landraces in representation of local populations cultivated in Italy until the beginning of the twentieth century. Group B includes selections within landraces based on desirable types naturally occurring within existing populations. Group C includes the cultivars obtained by Strampelli and their derivatives. Group D includes varieties released before 1970 and directly derived from material developed by Strampelli, but with reduced height alleles other than the original ones, coming from cv. Akagomugi. Group E includes varieties developed since 1970 with pedigree deriving from material traditionally used in Italian breeding programs and group F includes varieties developed since 1970 with foreign germplasm (e.g., from CIMMYT) in their

**Table 1** List of accessions used with code, name, and group

Code	Name	Group	Code	Name	Group	Code	Name	Group
00040	Andriolo	A	00692	Marzuolo Toscano	B	00001	A 416	E
00132	Bianchetta	A	00693	Marzuolo Toscano	B	00256	Costante	E
00141	Bianco Nostrale	A	00050	Apulia B1	C	00964	Salgemma	E
00192	Carosella	A	00057	Ardito	C	00978	Sangiacomo	E
00289	Dente di cane	A	00701	Mentana	C	00431	Gemini	E
00440	Gentil Bianco	A	01200	Villa Glori 1	C	00212	Chiarano	F
00443	Gentil Rosso	A	00175	Cambio	C	00966	Salmone	F
00444	Gentil Rosso	A	00112	Balilla	C	00696	Mec	F
00657	Majorca	A	00028	Albimonte	C	00591	Leopardo	F
00676	Martinella	A	00029	Albimonte B1	C	00661	Manital	F
00680	Marzuolo	A	00196	Catria	C	00200	Centaurio	F
00721	Morru Canu	A	00198	Catria	C	00330	Etruria	F
00819	Precoce Bianco	A	00048	Apulia Precoce 1	C	00796	Pandas	F
00820	Precoce di Piemonte	A	00239	Comandante Baudi 1	C	00002	Abano	F
00881	Rieti 1	A	00596	Libero (Littorio)	C	00078	Arquà	F
00882	Rieti 2	A	00598	Libero (Littorio) 1	C	00159	Brasilia	F
00883	Rieti 3	A	00599	Libero (Littorio) 2	C	00628	Loreto	F
00885	Rieti 5	A	00600	Libero (Littorio) 3	C	00629	Loreto 2	F
00889	Rieti 10	A	00093	Ausonia	C	00769	Oderzo	F
00890	Rieti 11	A	00528	Impeto	C	00153	Bolero	F
00909	Rosso Olona	A	00905	Rosso di Salmour	C	00322	Eridano	F
00910	Rosso Piemonte	A	00267	Damiano	C	00477	Golia	F
01027	Solina	A	00268	Damiano	C	00118	Barra	F
01028	Solina 2	A	00946	S. Pastore	C	00708	Mieti	F
01029	Solina 3	A	00971	Salto 2	C	00715	Mol	F
01031	Solina 5	A	01221	Vivenza	C	01004	Serio	F
01032	Solina 6	A	00097	Autonomia B	C	00221	Colfiorito	F
01033	Solina A	A	00119	Bassi Lauro	C	00961	Sagittario	F
01134	Trigo Cossu	A	00052	Aquila	C	01301	Soissons	F
01136	Trigo Cossu	A	00411	Funo	C	00142	Bilancia	F
00139	Bianchetta A	A	00364	Fiorello	C	00401	Freccia	F
00140	Bianchetta B	A	00662	Mara A	C	00549	Isengrain	F
00535	Inallettabile	A	00663	Mara B	C	00044	Apache	F
00134	Bianchetta 83 1	B	00402	Freccia	C	00122	Belfiore	F
00135	Bianchetta 83 2	B	00490	Grifo	C	00165	Buon Pastor	F
00226	Cologna 21 A	B	00901	Rondine	C	00793	Palesio	F
00228	Cologna 80	B	00005	Abbondanza	C	00081	Artico	F
00229	Cologna 83	B	00006	Abbondanza	C	00145	Blasco	F
00448	Gentil Rosso 13	B	00679	Marzotto	C	01253	Bologna	F
00449	Gentil Rosso 160	B	00586	Leonardo	C	00816	PR22R58	F
00450	Gentil Rosso 4	B	00300	Dragone	C	01250	Aubusson	F
00452	Gentil Rosso 48	B	00301	Dragone	C	00158	Bramante	F
00493	Gua' 113	B	00304	Dusi	C	00464	Geronimo	F
00533	Inallettabile 8 A	B	00305	Dusi 3	C	00794	Palladio	F
00534	Inallettabile 8 B	B	00306	Dusi 3A	C	00222	Colledoro	F
00536	Inallettabile 96	B	00831	Produttore S. 6	D	00336	Exotic	F
00538	Inallettabile 3	B	00474	Glutinoso	D	01245	Antille	F
00541	Inallettabile Todaro	B	00064	Argelato	D	01251	Bandera	F
00658	Majorca 47	B	00627	Lontra	D	01304	Solehio	F
00659	Majorca 68	B	00545	Inerio	D	01241	Altamira	F
00660	Majorica 47	B	01069	Strampelli Nazareno	D	01248	Arrocco	F
00684	Marzuolo Aqui 3	B	00777	Orso	D	01319	Zanzibar	F
00685	Marzuolo Aqui 4A	B	00012	Adria	D			
00686	Marzuolo Aqui 4B	B	00055	Aquileja	D			

pedigree. Groups from B to E include at least one of the parents belonging to the preceding group.

## 2.2 Phenotypic trait measurements

Phenotypic measurements on the collection were taken across five seasons (2011–2015) in four locations: S. Angelo Lodigiano (SAL, 45°14'17.7"N 9°24'21.6"E), Lodi (LO, 45°18'14.1"N 9°30'45.2"E), Fiorenzuola d'Arda (FIOR, 44°55'36.2"N 9°53'40.0"E), and Pisa (PI, 43°39'44.3"N 10°20'55.2"E) using field and greenhouse experiments, as detailed below. We investigated four traits related to wheat competitive ability against weeds:

- Above-ground biomass before stem elongation phase (BI), measured as dry weight (after oven-drying at 100 °C) of above-ground biomass per 1-m row (field trial at SAL in 2014/15, randomized complete block design with three replicates) and as dry weight (as above) of single plants (greenhouse trials at PI in 2013/14 and at FIOR in 2014/15, randomized complete block design with five replicates).
- Tillering capacity (TI), measured as number of culms divided by number of plants per 1-m row (field trial at SAL in 2014/15, randomized complete block design with three replicates) and as number of culms by individual plants (greenhouse trials at SAL and PI in 2013/14 and at FIOR in 2014/15, randomized complete block design with five replicates).
- Plant height at maturity (PH), measured as average height from the ground to the spike tip excluding the awns (field trials at SAL in 2013/14 and 2014/15, randomized complete block design with three replicates and at SAL and LO in 2010/11, 2011/12, and 2012/13, unreplicated randomized design experiment).
- Flag leaf (FG) morphology (area, length and average width) on five plants plot<sup>-1</sup> using a LI-COR Inc. area meter Li-3000 (field trials at SAL in 2013/14 and 2014/15, randomized complete block design with three replicates, and at LO in 2012/13, unreplicated randomized design experiment).

As indicators of yield potential, we measured:

- Grain yield (GY), as grain production on one 1-m<sup>2</sup> area plot<sup>-1</sup>, adjusted to 14% moisture content (field trials at SAL in 2013/14 and 2014/15, randomized complete block design with three replicates).
- Thousand-kernel weight (TKW), measured by weighing two 100-kernel samples (field trials at SAL and LO in 2010/11, 2011/12, and 2012/13, unreplicated randomized experiment).

All the experiments were conducted in weed-free conditions obtained by chemical control in field experiments. Nitrogen fertilization was applied at tillering (40 kg N ha<sup>-1</sup>, as ammonium nitrate) in all field experiments. Growth regulators, fungicides, and other pesticides were not applied in any of the experiments.

## 2.3 Genotypic evaluation

Genotyping was performed with the Infinium iSelect 90K array (Illumina Inc.), a high-density single nucleotide polymorphism (SNP) genotyping array containing 81,587 SNPs markers. SNPs were called using the hexaploid wheat pipeline in GenomeStudio V11 and mapped according to the consensus genetic linkage map for hexaploid wheat published by Wang et al. (2014).

## 2.4 Phenotypic data analysis

All data analyses were conducted using R environment for statistical computing, version 3.2.5. R/lme4 was used to calculate best linear unbiased predictors (BLUPs) of phenotypes with REML estimation. The model was run with a fixed intercept and random intercept for accession, environment (each experiment in one season was considered as one environment), block (when more than one replicate was available), and interaction between environment and block. The model was formulated as:

$$Y_{ijk} = \mu + a_i + e_j + e_j : b_k + \varepsilon_{ijk} \quad (1)$$

where  $Y_{ijk}$  is the phenotype for the accession  $i$  at environment  $j$  in block  $k$ ,  $\mu$  represents the grand mean,  $a$  represents the random effect of accession  $i$ ,  $e$  represents the random effect of the environment  $j$ ,  $e_i : b_j$  is the interaction between environment and block, and  $\varepsilon_{ijk}$  is the residual error. For continuous variables, we used a linear model and for count data a generalized model with Poisson distribution. The broad sense heritability ( $H^2$ ) of traits was calculated using the variance values from random effects model as:

$$H^2 = \delta^2_a / (\delta^2_a + \delta^2_f/n + \delta^2_b/m + \delta^2_\varepsilon/n * m) \quad (2)$$

where  $\delta^2_a$  is the total genetic variance,  $\delta^2_f/n$  is the total variance across environments divided by the number of environments ( $n$ ),  $\delta^2_b/m$  is the total variance across block divided by the number of blocks ( $m$ ) and  $\delta^2_\varepsilon/n * m$  is the residual variance divided by the number of environments×blocks. The BLUPs of all measured variables were used in a principal component analysis (PCA) performed with R/vegan.

R/lme4 was used to study the diversity between groups for competitive ability-related traits. The model was run with a fixed intercept for group and a random intercept for environment and block. The model with fixed effect for the group was



compared with a model with a fixed intercept for testing the significant effect of the grouping factor. R/multcomp was used for multiple comparisons of means (Tukey test).

Field experiments in 2013/14 and 2014/15, in which grain yield and final plant height were measured in the same trials, were used for exploring the trade-off between competitive ability and grain production. To do so, we ran a linear mixed model for grain yield with a fixed intercept for plant height and a random intercept for block. We studied the distribution of the model residuals in order to identify accessions for which the model significantly underestimated grain yield. These accessions are those for which yield is higher than expected from plant height measurements alone. In order to exclude plants with a favorable plant height to grain yield relationship, but with too little plant height values to effectively compete, we selected varieties with height > 90 cm. The average of residuals for the three replications was required to be > 1 to have a consistent empirical threshold of grain yield underestimation.

## 2.5 Genotypic data analysis

To investigate the diversity trend from landraces to modern lines, we compared the number of polymorphisms in groups A+B vs C+D+E. The number and chromosomal distribution of the polymorphisms specific of these two groups was calculated as well. A PCA was run on the Euclidean distance matrix of SNPs for all accessions. The genome-wide association study (GWAS) was performed with R/GAPIT. Minor allele frequencies were filtered at 5%, and markers that failed in more than 20% of the accessions were excluded. The VanRaden method (VanRaden 2008) was used to compute kinship and to account for familial relatedness in a compressed mixed linear model (CMLM). Principal components (PC) between 1 and 15 were iteratively used as covariates to include population structure as a fixed effect of the model. The SUPER method in R/GAPIT was used to calculate marker-trait associations. The best fit of the model was visually evaluated on quantile-quantile (Q-Q) plots, and number of PC covariates was chosen accordingly. Nominal *P* values obtained for each marker and trait were adjusted for false discovery rate (FDR) (Benjamini and Hochberg 1995). MTAs are reported and considered of interest when the FDR was < 1% (highly significant associations) and when FDR was between 1 and 5% (less stringent threshold).

## 3 Results and discussion

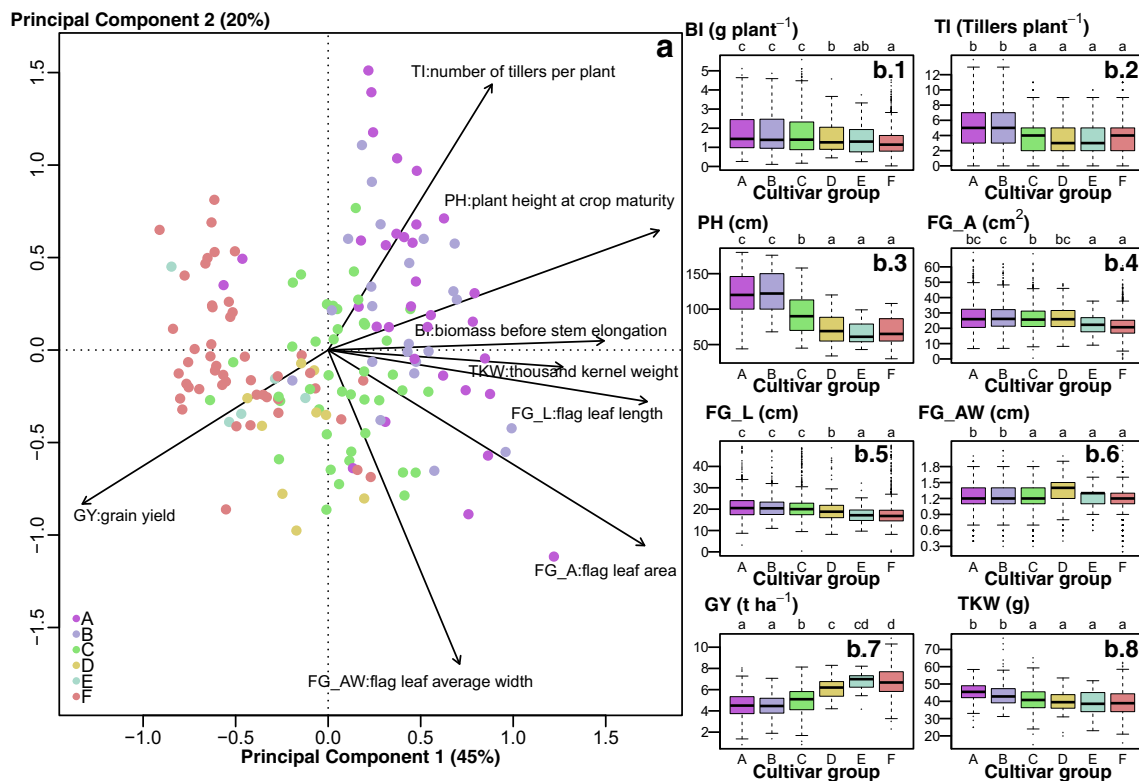
### 3.1 Competitive ability potential and trade-off with grain yield

The wheat germplasm collection was highly diverse for the competitive ability and production traits studied ( $P < 0.001$  for

all traits). An overall view of the phenotypic variability assessed in the experimental trials is depicted in the PCA plot (Fig. 2(a)). The first two PCs on phenotypic measurements explained 65% of total variance. The first PC (representing 45% of the variance) separates groups A and B from the others, especially from E and F. We can observe a temporal trajectory in the distribution of the six groups. By looking at the plot from top right to bottom left, the first encountered are the oldest accessions (groups A and B), then the accessions developed in the first decades of formal breeding (C and D) and those released after 1970 (E and F). Groups C and D are positioned around the origin of the axes, indicating high phenotypic variability. The accessions in groups C and D are not homogeneous for the traits studied, showing that the breeding material obtained by Strampelli (C) and immediately after (D) maintains a broad range of diversity for competitive ability traits. A positive correlation of biomass and flag leaf length with thousand-kernel weight was detected (0.49 and 0.31 respectively). Several studies, especially in rice, have shown a positive correlation between flag leaf traits and thousand-kernel weight (e.g., Khaliq et al. 2008; Cui et al. 2014). In our collection, flag leaf length was positively correlated with the other competitive traits such, i.e., biomass and plant height (0.44 and 0.55 respectively). Instead, flag leaf width was not significantly correlated with grain yield. In the PCA plot, the trade-off between grain yield and all the other traits is evident.

The six groups of wheat accessions showed significant variability in the average values for all the traits measured (Fig. 2(b1–8)). Regarding early season biomass accumulation, germplasm with higher values can be found not only in groups A and B but also in group C, which comprises the cultivars obtained by Strampelli and derivatives. The number of tillers per plant and plant height is higher in groups A and B. Considering the measured components of the flag leaf morphology, area is higher in group B, length is higher in groups A to C whereas average width is higher in group D. Grain yield was higher in modern cultivars (groups D, E, F) whereas thousand-kernel weight was higher in groups A and B. Expectedly, groups A to C showed higher average values for all competitive ability traits but flag leaf width. It is recognized that older genotypes are often more competitive than recent varieties, especially when competitive ability is associated with high early biomass accumulation, number of tillers, and height (Murphy et al. 2008; Andrew et al. 2015). This result confirms that heritage germplasm and the pool of varieties released at the beginning of formal breeding are an important source from which cultivars with high potential competitive ability against weeds can be selected.

Nevertheless, the trade-off between grain yield and competitive ability in weed-free situations should be carefully considered when choosing germplasm suitable for low-input and organic farming. We identified eight accessions with plant height > 90 cm (an empirical threshold of a tall canopy) and



**Fig. 2** **a** Ordination plot from a principal component analysis representing an overall view of the phenotypic variability within the wheat germplasm collection assessed in the experimental trials. The color of the points indicates accessions belonging to different cultivar groups (A to F), while arrows represent measured traits. We can observe a temporal trajectory in the distribution of the six groups with groups A and B (landraces and selections within landraces) separated on the first principal component from the others (modern varieties), especially from E and F (varieties developed since 1970). **b** The sets of predicted values of each competitive ability trait in each group are represented with box-plots. In each plot, groups that do not share the same letters are significantly different

from each other for the considered trait. Group A: landraces; Group B: selections within landraces; Group C: cultivars obtained by Strampelli and their derivatives; Group D: varieties released before 1970 from material by Strampelli, with reduced height alleles other than from cv. Akagomugi; Group E: varieties developed since 1970 with pedigree deriving from material traditionally used in Italian breeding programs; F: varieties developed since 1970 with foreign germplasm in the pedigree. BI: total plant above-ground biomass before stem elongation phase; TI: number of tillers per plant; PH: plant height at crop maturity; FG\_A: flag leaf area, FG\_L: flag leaf length, FG\_AW: flag leaf average width; GY: grain yield; TKW: thousand-kernel weight

average residuals of the model used to estimate grain yield as a function of height > 1. This latter value means that the model underestimated grain yield of at least 1 t/ha, which can be considered a reasonable empirical threshold of mismatch between the predicted and actual trade-off between grain yield and plant height. Two of these eight accessions belong to the group of Rieti populations (codes 00883 and 00885) and another is the landrace Bianco Nostrale (code 00141), all belonging to group A (Table 1). Gentil Rosso (code 00452) and Guà (code 00493) belong to group B, whereas cv. Albimonte (code 00028), Aquileja (code 00055), and Loreto (code 00628) belong respectively to groups C, D, and F (Table 1).

This type of selection minimizes grain yield reduction in weed-free conditions when high weed competitive ability potential is prioritized for cultivar choice. Although our selection criterion is, strictly speaking, valid only within our wheat collection, a similar approach could be applied to any wheat germplasm collection where an estimation of the trade-off between grain yield and competitive ability is sought. It would

be good to include more than one competitive ability trait in the model for increasing the generalization potential of the trade-off relationship.

### 3.2 Molecular characterization

Among the set of 81,587 SNPs on the 90K Illumina array, 36,397 were polymorphic and 23,116 had a minor allele frequency (MAF) above 5% and failed in less than 20% of the accessions. Of these markers, 19,432 had a univocal position on the hexaploid wheat consensus map (Wang et al. 2014). Wheat sub-genome D is the least represented in our dataset with 1808 markers against 7890 on sub-genome A and 9734 on sub-genome B. The underrepresentation of the D sub-genome is due to the natural lack of polymorphism and consequently to a bias on the array for this sub-genome (Wang et al. 2014; Rimbart et al. 2018).

The accessions belonging to groups A to E were used for studying the trend of genetic diversity from landraces to

modern lines. We excluded group F from this analysis in order to have information about Italian breeding before the introduction of parental lines from other breeding programs. The number of polymorphisms was 28,556 in groups A+B and 26,875 in groups C+D+E. Altogether, populations and mass-selected lines had 1681 more polymorphisms than the modern germplasm in our collection (4.6% of the polymorphisms in the overall collection). The collection showed a trend of reduction of the polymorphism number passing from heritage to modern germplasm. Heritage germplasm was shown to be more diverse than modern accessions also by Riaz et al. (2016), who performed a genetic characterization of common wheat accessions from N. I. Vavilov Institute of Plant Genetic Resources and compared them to a group of modern cultivars and elite breeding lines from Australia and the International Maize and Wheat Improvement Center (CIMMYT).

Roussel et al. (2005) suggested that modern lines might have reduced genetic diversity compared to landraces because of the use of a small number of genetically similar varieties within the different modern breeding programs. The decrease of diversity due to the breeding bottleneck as proposed by Roussel et al. (2005) might not be generalized because, over time and geographical area, breeding has produced variable outcomes regarding genetic diversity conservation (Winfield et al. 2018). Nevertheless, it is important to note that from previous investigations it also emerged that heritage germplasm may keep variability not present in modern germplasm (Lopes et al. 2015; Riaz et al. 2016; Winfield et al. 2018). Winfield et al. (2018) characterized a collection of landraces from 32 countries around the world using an 820K SNP array, and detected a high number of polymorphisms unique to landraces or to modern lines. Winfield et al. (2018) suggested that the polymorphisms unique to the landraces might indicate the presence of a valuable genetic variability that is not being incorporated in modern breeding programs.

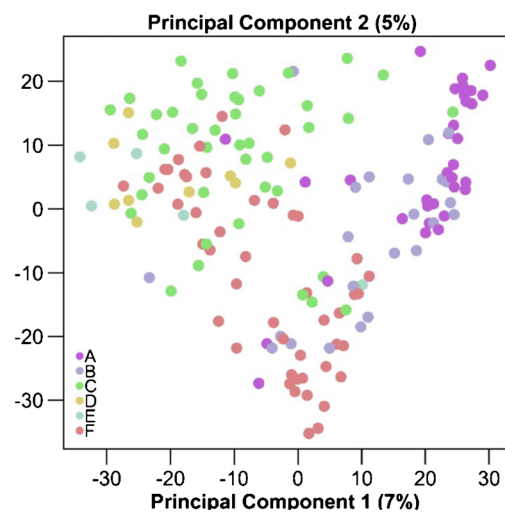
Among the detected polymorphisms in our study, 3761 SNP were specific of groups A+B and 2080 of groups C+D+E. This result highlights that, even if we detected higher diversity in heritage germplasm, both heritage and modern germplasm showed unique polymorphisms, indicating that it is useful to have both of them represented for the genetic dissection of a wide range of traits, especially for those not targeted during conventional breeding programs.

The collection presented a low genetic structure. To reach 50% of the original genetic variance, it was necessary to consider the first 21 SNP-based PCs. The first two PCs (Fig. 3) accounted for 7% and 5% of the molecular variance, respectively. The accessions in group A (landraces) and group B (mass selection from landraces) clustered on positive values of PC1. Even if the first PC explained only 7% of the molecular diversity, it is worth noting that it separated the landraces and mass-selected lines from all other germplasm in the collection. This happened even though part of the germplasm

comprised in this group was used in downstream breeding (in groups C, D, and E). Groups C to E seemed admixed on the PCA plot (Fig. 3). This result is in line with the design of the collection, that includes for each group at least one of the parents belonging to the preceding group. Group F was the most recognizable on the PCA plot among the groups of modern germplasm. Despite the weak genetic structure of the collection, the separation of groups A+B and C+D+E is of particular interest especially in regard to traits not targeted by the Green Revolution, and as such not retained in the breeding process. This diversity may be particularly useful when searching for material adapted to peculiar conditions such as cultivation in organic and low-input systems.

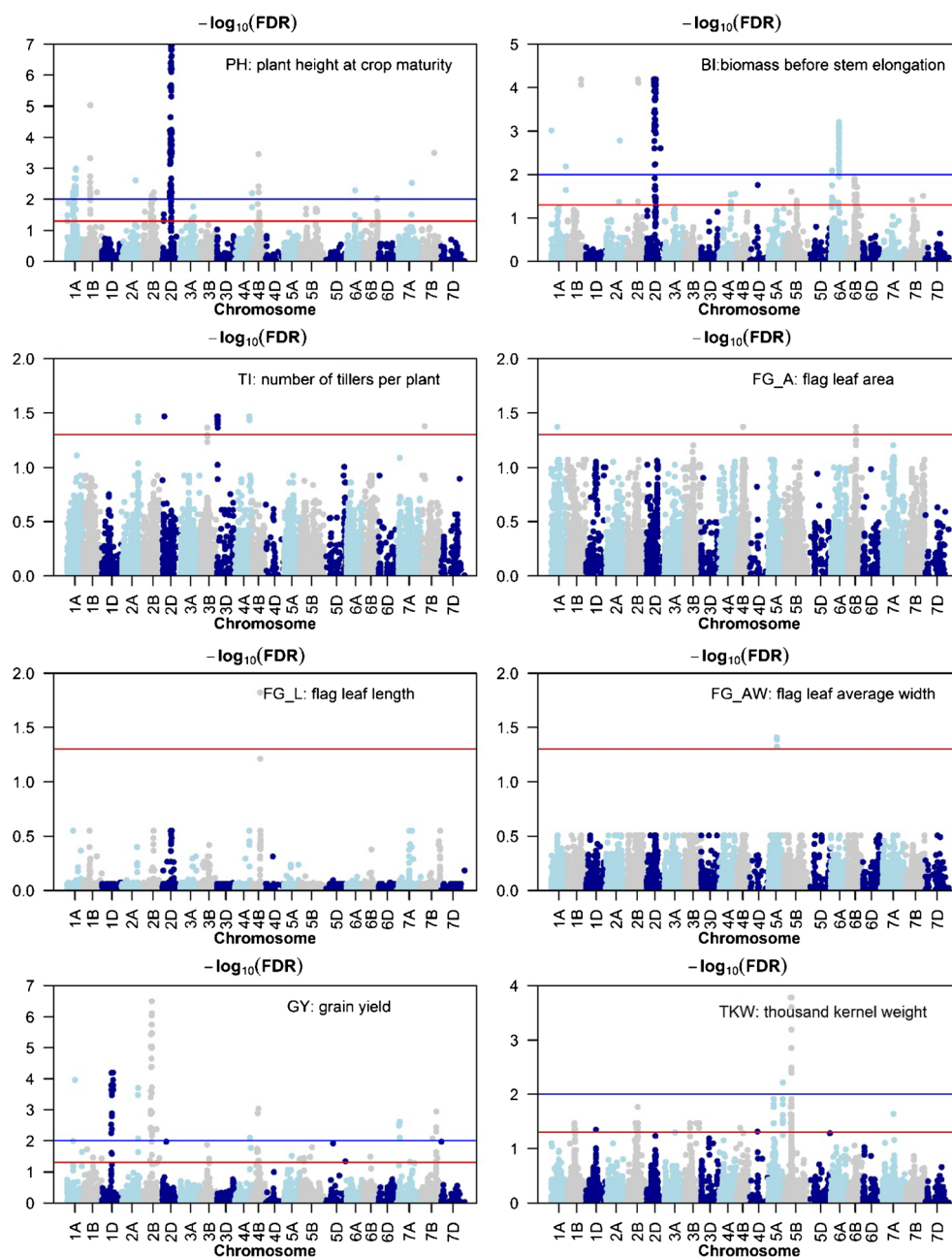
### 3.3 Marker-trait associations

The significant phenotypic diversity for competitive ability-related traits was a good base for their genetic dissection. Nevertheless, broad sense heritability ( $H^2$ ) varied markedly among the traits and the measurements for above-ground biomass at an early stage and for tillering index were the least repeatable ( $H^2 = 0.30$  and  $0.57$  respectively). We used the BLUP values for each trait as phenotype for the association study as already done by several other studies. In the association study, we could detect MTAs with FDR < 5% for all studied traits. MTAs with FDR < 1% were detected for plant height, biomass, grain yield, and thousand-kernel weight (Fig. 4).



**Fig. 3** Ordination plot from a principal component analysis representing the genetic diversity within the wheat germplasm collection. Principal component 1 separated landraces and mass-selected lines (groups A and B) from all other germplasm in the collection. The color of the points indicates accessions belonging to different cultivar groups (A to F). Group A: landraces; Group B: selections within landraces; Group C: cultivars obtained by Strampelli and their derivatives; Group D: varieties released before 1970 from material by Strampelli, with reduced height alleles other than from cv. Akagomugi; Group E: varieties developed since 1970 with pedigree deriving from material traditionally used in Italian breeding programs; F: varieties developed since 1970 with foreign germplasm in the pedigree

**Fig. 4** Manhattan plots for genome-wide association study (GWAS) of the phenotypic traits measured. Markers on each chromosome are plotted in light blue (A genome), gray (B genome), and dark blue (D genome). The indication of the chromosome is given on the x-axis. The y-axis represents the negative base 10 logarithm of the false discovery rate (FDR) values. In this graph, the strongest associations with the smallest  $P$  values, exhibit the greatest negative logarithms. The blue and red lines indicate the threshold of FDR = 1% (highly significant associations) and of FDR = 5% (indicative associations), respectively



Plant height is a key morphological trait in wheat because it has implications on harvest index, lodging, and competitive ability against weeds, with a composite effect on grain yield (Hoad et al. 2012; Griffiths et al. 2012). The presence of the dwarfing gene *Rht8* on chromosome 2D, derived from cv. Akagomugi, was expected in our cultivar pool due to the presence of several cultivars, belonging to Group C, obtained by Nazareno Strampelli through hybridization of cv. Akagomugi with Italian germplasm, mainly cv. Rieti. Indeed, we detected two groups of closely linked SNPs on chromosome 2D associated with plant height. Besides the signal on chromosome 2D, we could detect associations with plant height on other chromosomes. Determinants of plant

height other than *Rht8* were expected as well, as groups D and F contain varieties that, according to their genealogy, do not derive from cv. Akagomugi. The MTAs on chromosomes 1A (95.20 cM), 2B (82.43 cM), 2D (12.38 cM), 4B (57.5, 58.1 cM), 5B (60 cM), and 6A (123.48 cM) are close to putative quantitative trait loci (QTL) for plant height reported in Zanke et al. (2014b). The association on chromosomes 1B, 2A, 3A, 4A, 6B, 7A, and 7B did not match previously reported positions. The many associations for plant height that we identified are indicative of the diversity of alleles implicated in the determination of this trait. Plant height heritability in our experiment was 0.95. Even if the heritability of biomass was 0.30 in our phenotyping experiment, we could identify



associations on seven chromosomes with  $FDR < 1\%$ . Many of the associations for above-ground biomass overlapped with those for plant height. This happened in particular for the associations on chromosomes 1A, 1B, 2A, 2B, 2D, and 6A. This result can be explained by the positive correlation between plant height and biomass accumulation (Austin et al. 1989). Our data shows that this relationship holds true not only for straw biomass at harvest, but also for the biomass accumulation during the tillering phase.

Tillering capacity had a heritability of 0.57. For this trait, a high genotype by environment interaction was already observed (Mengistu et al. 2016); nevertheless, we could identify associations on six chromosomes. Wang et al. (2016) reported associations for tillering on chromosomes 2D, 2B, and 5A. Our association on chromosome 2D, at 4 cM, is on the same linkage group but at a totally different position from that reported in Wang et al. (2016) (at above 200 cM). We did not identify the previously described tiller inhibitor genes *Tin* (Kuraparthy et al. 2006) or *Gli-A2*, reportedly having an effect on tiller number (Li et al. 2002). It is important to note that the genetic architecture of this rather elusive trait has not yet been fully described; further characterization of our newly detected MTAs may provide additional information in this respect.

The heritability of flag leaf traits was 0.88, 0.84, and 0.71 for area, length, and average width, respectively. We identified three associations for flag leaf area (chromosomes 1A, 4B, 6B), and different ones for flag leaf length (chromosome 4B) and flag leaf width (chromosome 5A). Fan et al. (2015) reported associations for flag leaf area on chromosome 4B and 6B. We detected MTA at 61.84 cM and 71.76 cM, respectively. For flag leaf length, Fan et al. (2015) detected an association on chromosome 4B. In our study, the association was detected at 73.84 cM of chromosome 4B. Nevertheless, the associations in the two studies cannot be compared due to the fact that the positions reported by Fan et al. (2015) are based on the 591 markers map by Cui et al. (2014), while the association on our study has instead a map position based on 40,267 SNP markers, with no common markers between the two maps.

An association for flag leaf width on chromosome 5A is reported by Yang et al. (2016) at 67.1 cM. The association on chromosome 5A for flag leaf width in our study was detected at 84.12 cM. As the previous case, different positions may be due to the use of different genetic maps and cannot be directly compared. To our knowledge, the association for flag leaf area on chromosome 1A is a new finding.

Grain yield and thousand-kernel weight are both very complex traits (Zanke et al. 2015), as confirmed by the 26 and 13 associations detected in our study. Among the associations for thousand-kernel weight on chromosome 5B we could detect an overlap with an association for above-ground biomass (68.36 and 68.93 cM respectively). Similarly, the associations

for grain yield on chromosomes 2D and 6B were detected also for biomass accumulation. The pleiotropy of the determinants of plant height, grain yield, and thousand-kernel weight to biomass accumulation is a further confirmation of the complexity of the relationship between production-related traits and competition-related traits.

The study of this collection was useful to get a picture in terms of the genetic variability in the germplasm cultivated by Italian farmers from the end of nineteenth century to nowadays. The genetic diversification of the collection gives support to the usefulness of this collection for investigating the genetic determination of traits of interest for peculiar conditions not targeted by conventional wheat breeding. At the same time, new regions associated with the target traits were highlighted by GWAS study, giving proof of the usefulness of this collection in a novel breeding perspective. This knowledge can be used as a starting point for further investigations dissecting the molecular bases of phenotypes of interest.

## 4 Conclusion

In this interdisciplinary study, the integration of agroecological principles and current molecular breeding technologies aimed to develop a methodology for selecting wheat germplasm characterized by higher competitive ability against weeds to be used in more sustainable integrated weed management systems for organic and low-input production. The study combined the characterization of the molecular and phenotypic diversity of a collection of 160 accessions harnessing heritage and modern germplasm representative of common wheat cultivated in Italy from the end of the nineteenth century to 2009. This collection represents an important resource for Italian common wheat breeding programs (private, public, or participatory) that aim to develop germplasm adapt to organic and low-input agriculture where the inclusion of traits for higher crop competitive ability against weeds is a desirable selection goal. Further studies and use of this material may allow to select and develop cultivars more competitive against weeds as well as accessions in which the trade-off between competitive ability and grain yield is acceptable, a key asset for wheat cultivation in Italy and beyond. This approach, if applied to other wheat germplasm collections or to other crop species, can empower integrated weed management by making available more competitive cultivars selected by applying agroecological principles and exploiting the most advanced breeding technologies.

**Funding information** The authors acknowledge the Italian Ministry of Agriculture for partially funding this research in the framework of the Project RGV-FAO (DM 11746, 10/04/17) and the International PhD program in Agrobiodiversity of Scuola Superiore Sant'Anna for providing the scholarship to the first author.

**Data availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author upon request.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Ahlgren S (2004) Environmental impact of chemical and mechanical weed control in agriculture: a comparing study using life cycle assessment (LCA) methodology. SIK rapport Nr 719 2004. The Sdewish Institute for Food and Biotechnology, Gothenburg, Sweden
- Andrew IKS, Storkey J, Sparkes DL (2015) A review of the potential for competitive cereal cultivars as a tool in integrated weed management. *Weed Res* 55:239–248. <https://doi.org/10.1111/wre.12137>
- Austin RB, Ford MA, Morgan CL (1989) Genetic improvement in the yield of winter wheat: a further evaluation. *J Agric Sci* 112:295–301. <https://doi.org/10.1017/S0021859600085749>
- Barabaschi D, Tondelli A, Desiderio F, Volante A, Vaccino P, Valè G, Cattivelli L (2016) Next generation breeding. *Plant Sci* 242:3–13. <https://doi.org/10.1016/j.plantsci.2015.07.010>
- Bärberi P (2002) Weed management in organic agriculture: are we addressing the right issues? *Weed Res* 42:177–193. <https://doi.org/10.1046/j.1365-3180.2002.00277.x>
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 57:289–300
- Cui F, Fan X, Zhao C, Zhang W, Chen M, Ji J, Li J (2014) A novel genetic map of wheat: utility for mapping QTL for yield under different nitrogen treatments. *BMC Genet* 15:57. <https://doi.org/10.1186/1471-2156-15-57>
- Fan X, Cui F, Zhao C, Zhang W, Yang L, Zhao X, Han J, Su Q, Ji J, Zhao Z, Tong Y, Li J (2015) QTLs for flag leaf size and their influence on yield-related traits in wheat (*Triticum aestivum* L.). *Mol Breed* 35: 24. <https://doi.org/10.1007/s11032-015-0205-9>
- Griffiths S, Simmonds J, Leverington M, Wang Y, Fish L, Sayers L, Alibert L, Orford S, Wingen L, Snape J (2012) Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. *Mol Breed* 29:159–171. <https://doi.org/10.1007/s11032-010-9534-x>
- Hoad SP, Bertholdsson N-Ø, Neuhoff D, Köpke U (2012) Approaches to breed for improved weed suppression in organically grown cereals. In: Lammerts van Bueren ET, Myers JR (eds) *Organic crop breeding*. Wiley-Blackwell, Chichester, pp 61–76
- Khaliq I, Irshad A, Ahsan M (2008) Awns and flag leaf contribution towards grain yield in spring wheat (*Triticum aestivum* L.). *Cereal Res Commun* 36:65–76. <https://doi.org/10.1556/CRC.36.2008.1.7>
- Kidane YG, Hailemariam BN, Mengistu DK, Fadda C, Pè ME, Dell'Acqua M (2017) Genome-wide association study of *Septoria tritici blotch* resistance in Ethiopian durum wheat landraces. *Front Plant Sci* 8:1586. <https://doi.org/10.3389/fpls.2017.01586>
- Kuraparthy V, Sood S, Dhaliwal HS, Chhuneja P, Gill BS (2006) Identification and mapping of a tiller inhibition gene (*tin3*) in wheat. *Theor Appl Genet* 114:285–294. <https://doi.org/10.1007/s00122-006-0431-y>
- Laino P, Limonta M, Gerna D, Vaccino P (2015) Morpho-physiological and qualitative traits of a bread wheat collection spanning a century of breeding in Italy. *Biodivers Data J* 3:e4760. <https://doi.org/10.3897/BDJ.3.e4760>
- Li WL, Nelson JC, Chu CY et al (2002) Chromosomal locations and genetic relationships of tiller and spike characters in wheat. *Euphytica* 125:357–366. <https://doi.org/10.1023/A:1016069809977>
- Lopes MS, El-Basyoni I, Baenziger PS et al (2015) Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *J Exp Bot* 66:3477–3486. <https://doi.org/10.1093/jxb/erv122>
- Mengistu DK, Kidane YG, Catellani M, Frascaroli E, Fadda C, Pè ME, Dell'Acqua M (2016) High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. *Plant Biotechnol J* 14:1800–1812. <https://doi.org/10.1111/pbi.12538>
- Murphy KM, Dawson JC, Jones SS (2008) Relationship among phenotypic growth traits, yield and weed suppression in spring wheat landraces and modern cultivars. *Field Crop Res* 105:107–115. <https://doi.org/10.1016/j.fcr.2007.08.004>
- Ornoli L, Costa C, Negri S, Perenzin M, Vaccino P (2015) Diversity trends in bread wheat in Italy during the 20th century assessed by traditional and multivariate approaches. *Sci Rep* 5:1–7. <https://doi.org/10.1038/srep08574>
- Riaz A, Hathorn A, Dinglasan E, Ziemis L, Richard C, Singh D, Mitrofanova O, Afanasenko O, Aitken E, Godwin I, Hickey L (2016) Into the vault of the Vavilov wheats: old diversity for new alleles. *Genet Resour Crop Evol* 64:531–544. <https://doi.org/10.1007/s10722-016-0380-5>
- Rimbert H, Darrier B, Navarro J, Kitt J, Choulet F, Leveugle M, Duarte J, Rivière N, Eversole K, on behalf of The International Wheat Genome Sequencing Consortium, le Gouis J, on behalf The BreedWheat Consortium, Davassi A, Balfourier F, le Paslier MC, Berard A, Brunel D, Feuillet C, Poncet C, Sourdille P, Paux E (2018) High throughput SNP discovery and genotyping in hexaploid wheat. *PLoS One* 13:e0186329. <https://doi.org/10.1371/journal.pone.0186329>
- Roussel V, Leisova L, Exbrayat F, Stehno Z, Balfourier F (2005) SSR allelic diversity changes in 480 European bread wheat varieties released from 1840 to 2000. *Theor Appl Genet* 111:162–170. <https://doi.org/10.1007/s00122-005-2014-8>
- Salvi S, Porfiri O, Ceccarelli S (2013) Nazareno Strampelli, the ‘prophet’ of the green revolution. *J Agric Sci* 151:1–5. <https://doi.org/10.1017/S0021859612000214>
- VanRaden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, 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 AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo MC, Dvorak J, Morell M, Dubcovsky J, Ganai M, Tuberosa R, Lawley C, Mikoulitch I, Cavanagh C, Edwards KJ, Hayden M, 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. <https://doi.org/10.1111/pbi.12183>
- Wang Z, Liu Y, Shi H, Mo H, Wu F, Lin Y, Gao S, Wang J, Wei Y, Liu C, Zheng Y (2016) Identification and validation of novel low-tiller number QTL in common wheat. *Theor Appl Genet* 129:603–612. <https://doi.org/10.1007/s00122-015-2652-4>
- Winfield MO, Allen AM, Wilkinson PA, Burrridge AJ, Barker GLA, Coghill J, Waterfall C, Wingen LU, Griffiths S, Edwards KJ (2018) High-density genotyping of the A.E. Watkins collection of hexaploid landraces identifies a large molecular diversity compared to elite bread wheat. *Plant Biotechnol J* 16:165–175. <https://doi.org/10.1111/pbi.12757>

- Yang D, Liu Y, Cheng H, Chang L, Chen J, Chai S, Li M (2016) Genetic dissection of flag leaf morphology in wheat (*Triticum aestivum* L.) under diverse water regimes. *BMC Genet* 17:94. <https://doi.org/10.1186/s12863-016-0399-9>
- Zanke CD, Ling J, Plieske J et al (2014a) Genetic architecture of main effect QTL for heading date in European winter wheat. *Front Plant Sci* 5:217. <https://doi.org/10.3389/fpls.2014.00217>
- Zanke CD, Ling J, Plieske J, Kollers S, Ebmeyer E, Korzun V, Argillier O, Stiewe G, Hinze M, Neumann K, Ganal MW, Röder MS (2014b) Whole genome association mapping of plant height in winter wheat (*Triticum aestivum* L.). *PLoS One* 9:e113287. <https://doi.org/10.1371/journal.pone.0113287>
- Zanke CD, Ling J, Plieske J, Kollers S, Ebmeyer E, Korzun V, Argillier O, Stiewe G, Hinze M, Neumann F, Eichhorn A, Polley A, Jaenecke C, Ganal MW, Röder MS (2015) Analysis of main effect QTL for thousand grain weight in European winter wheat (*Triticum aestivum* L.) by genome-wide association mapping. *Front Plant Sci* 6:644. <https://doi.org/10.3389/fpls.2015.00644>