Genotypic differences in wheat yield determinants within a NAM population based on elite parents

Priyanka A. Basavaraddi[1], Roxana Savin[1], Sivakumar Sukumaran[2], Matthew P. Reynolds[2], Simon Griffiths[3], Gustavo A. Slafer[1,4,*]

[1]Department of Crop and Forest Sciences, University of Lleida - AGROTECNIO Center, Av. R. Roure 191, 25198 Lleida, Spain

[2] International Maize and Wheat Improvement Center (CIMMYT), El Batán, Texcoco CP 56237, Mexico

[3] John Innes Centre, Norwich Research Park, Colney Ln, Norwich, NR4 7UH, United Kingdom

[4] ICREA, Catalonian Institution for Research and Advanced Studies, Spain

[*] Corresponding author. Email address: gustavo.slafer@udl.cat

Abstract

Future grain yield (GY) improvements require the identification of beneficial traits within the context of high yield potential and not just based on the pleiotropic effect of traits such as crop height and heading date. We evaluated 1937 lines from Nested Association Mapping (NAM) population derived from 13 bi-parental varietal crosses under field conditions. We selected 493 lines with similar time to anthesis to that of the two checks used in the study (across and within each family) which reduced the range of plant height in the selected lines. Yield components were measured in these 493 lines from which 231 lines were selected by excluding lines with lowest number of grains so excluded low yielding lines. Later the subset of 231 lines were evaluated in two field experiments (2016-17, CS1 and 2017-18, CS2). Numerical and physiological components of grain yield were measured. The two-step selection maximised GY within an acceptable range of variation for height and anthesis. GY in 231 lines showed very high G×E interaction. Taking both seasons together, we selected lines from upper and lower quartile GY groups to identify stable beneficial trait combinations for improved GY. Differences in GY were explained by grain number driven by increased spike dry weight at anthesis (SDWa) and fruiting efficiency (FE). Increased GY was accompanied by sink limitation. The data points towards increases in grain number as the route towards future GY increases in wheat breeding.

Key words: Pre-anthesis phases, fruiting efficiency, spike dry weight at anthesis, source-sink, yield

components

1. Introduction

Present rates of genetic gains in wheat grain yield (GY) are insufficient to satisfy future demands (Reynolds et al., 2012) which is estimated to increase 50% by 2050 from current level of demand (https://www.cimmyt.org/work/wheat-research/; accessed on 05.05.2020). In recent decades the rate of genetic gain has decreased (e.g. Aisawi et al., 2015; Flohr et al., 2018; Maeoka et al., 2020), in many cases to a standstill (e.g. Acreche et al., 2008; Chairi et al., 2018; de Oliveira Silva et al., 2020; Lo Valvo et al., 2018). To address this problem, we need to improve our understanding of physiological attributes likely to underpin future GY gains as well as to identify variation available within elite germplasm for these traits. Grain number per m^2 (GN) and average grain weight (AGW) are the two most important GY components (Slafer et al., 2014). Owing to larger plasticity it is GN that has delivered most GY improvements (Abbate et al., 1995; Calderini and Slafer, 1999; Fischer, 1985; Reynolds et al., 2009; Serrago et al., 2013; Siddique et al., 1989a; Slafer et al., 1990; Slafer and Andrade, 1989), even though it has much lower heritability than AGW (Sadras and Slafer, 2012).

Past improvements in wheat GN and GY came through the gradual accumulation of beneficial quantitative variation as well as a limited set of step changes such as the widespread deployment of semi-dwarf genes, chiefly Rht-1 (e.g. Calderini and Slafer, 1999; Flintham et al., 1997) and improving adaptation by changing time to anthesis to be more adequate for a specific region (e.g. Araus et al., 2002) particularly through changes in photoperiod and vernalisation sensitivity (González et al., 2005a; Griffiths et al., 2009; Shaw et al., 2012; Whitechurch and Snape, 2003). Reductions in plant height mediated by *R*ht-1 enhanced biomass partitioning to the juvenile spikes prior to anthesis (Brooking and Kirby, 1981; Fischer and Stockman, 1986; Miralles et al., 1998) which in turn allowed for an improved development of florets resulting in higher GN (Ferrante et al., 2013; Fischer and Stockman, 1986; Miralles et al., 1998; Siddique et al., 1989a). Introgression of Rht-1 alleles and homoeoalleles increased harvest index (HI) through increased GN and improved GY without major changes in biomass and a reduction in AGW, that naturally did not counteract the GN benefits (Bingham and Wellington., 1981; Calderini et al., 1995; Flintham et al., 1997; Miralles and Slafer, 1995; Shearman et al., 2005; Siddique et al., 1989b). Adjustments in time to anthesis have been critical to improve GY through improving adaptation mainly when the life cycle of the original genotypes exploited in a region did not allow maximum use of available resources or for stress avoidance (Araus et al., 2002). These two traits, that have been critical to

65 improve yields in the past, would be of limited importance in the future as they have already been
66 optimised in major wheat growing regions (e.g. Acreche et al., 2008; Maeoka et al., 2020; Slafer et
67 al., 2005).

68 Future gains in GN will provide the increased sink strength which many studies have pointed to
69 as required to increase GY, because of the frequent sink limitation for grain filling in wheat (Borrás
70 et al., 2004; Borrill et al., 2015; Reynolds et al., 2005; Serrago et al., 2013 and references quoted
71 there in). Final GN is a highly integrative trait (highly plastic and with low heritability; Sadras and
72 Slafer, 2012b) with many of the development processes that lead to contributing to the final
73 number. So, the identification of major genes or QTL directly and consistently controlling it is
74 unlikely. For these reasons it is important to understand which traits are responsible for differences
75 in GN within elite material and to show how they could be deployed by breeders aiming to improve
76 GY within elite × elite pedigrees by reducing sink limitation in their finished varieties.

77 While time to anthesis is tightly controlled in breeders selections around a local optimum, the
78 partitioning of the cycle into different duration of phases occurring before and after terminal
79 spikelet (TS) might still be improved (Slafer et al., 2001). Components of GN are formed from
80 sowing to a few days after anthesis (Slafer and Rawson, 1994) but the most sensitive phase is
81 demarcated by TS and anthesis, the late reproductive phase or LRP (Slafer, 2003; Fischer, 2011),
82 and in particular the last half of it (from flag leaf appearance to anthesis). Thus, it has been
83 hypothesised that lengthening the duration of the LRP, when floret development takes place,
84 would improve GN (Miralles and Slafer, 2007).

85 By the time anthesis is reached the stage is set for the realisation of GN, in fact the spike dry weight
86 at anthesis (SDWa) has been shown to be highly predictive of GN in a number of experiments
87 (Fischer, 2011; Ferrante et al., 2013). The physiological support for this mechanistic relationship
88 is that floret primordia survival is closely linked to SDWa and in wheat, being a cleistogamous
89 plant, most fertile florets become grains after anthesis. The number of fertile florets at anthesis
90 depends mainly on the balance between the initiation and mortality of floret primordia during the
91 LRP (Kirby, 1988; Prieto et al., 2018). Both floret mortality (Ferrante et al., 2013; González et al.,
92 2011) and survival (Ferrante et al., 2013, 2012; González et al., 2005b; Siddique et al., 1989a) seems
93 to depend on the availability of resources for spike growth from flag leaf appearance to anthesis.
94 The physiological dissection of this point in development has been taken further by Slafer *et al.*
95 (2015) using the concept of fruiting efficiency (FE, number of grains produced per unit SDWa)
96 and showing that FE can be useful towards genetically improving wheat GY (see also empirical
97 proofs in Acreche et al., 2008; Flohr et al., 2018; Lo Valvo et al., 2018b; Zhang et al., 2019).

For a proper identification of traits or trait combinations that are likely to be important and useful in modern wheat breeding programmes (i.e. beyond traits like plant height which are already optimised), it is important to study trait relationships within the context of elite germplasm. Although the analyses restricted to elite genotypes will naturally reduce substantially the degree of variation that could be expected from unselected lines of wider crosses (and would consequently yield less clear relationships). The advantage is that the materials used would resemble better what realistic breeding does (crosses of elite × elite) when aiming to improve yield, and therefore results and conclusions would be more likely truly applicable in actual breeding programmes. Therefore, in the present study we firstly grew a very large population of elite lines (1937 lines of a Nested Association Mapping, NAM, population produced by crossing elite parents) in the field at Ciudad Obregón, Mexico (Cd. Obregón) and from these initial results we further selected a relatively small, yet rather large, sub-set of 231 lines that were considered best performing (within germplasm that was already elite) to study them more in detail in field experiments carried out in Bell-lloc d'Urgell, Spain (Bell-lloc) over two cropping seasons.

2. Materials and Methods

2.1. Experimental field conditions

The first field evaluation of the whole NAM population was carried out in the 2015-16 cropping season at CIMMYT's experimental station (within the Norman E. Borlaug Experimental Field, CENEB) in Cd. Obregón, Sonora, North-West Mexico (lat. 27°23' N, 109°55'W). The experiment was sown on 10 December 2015 in small plots ("hills", 80 cm between hills, 30 cm long) at a density equivalent to 5 plants per plot.

In the following two seasons (2016-17, CS1 and 2017-18, CS2), field experiments were carried out near Bell-lloc d'Urgell, Lleida, North-East Spain (Lat. 41°38' N, 0°44' E in CS1 and Lat. 41°37' N, 0°47' E in CS2). Experiments were sown on 16 November 2016 and on 17 November 2017, both at the rate of 125 kg ha^{-1} aiming to attain an effective plant density of 250 plants per m^2.

The three experiments were carried out avoiding stresses: plots were always sown within the optimal dates to maximize yield, fully fertilized, irrigated, Weeds, pests and diseases were prevented or controlled. Soil nitrogen availability was determined in CS1 and CS2 at the beginning of the experiments. Eight samples from the soil surface to 0.9 m depth were randomly taken from the field were the experiments were sown and analysed for mineral N content. The average available N content of the experimental area was 133.1±9.3 and 115.4±8.8 KgN ha^{-1} in CS1 and CS2, respectively. This soil nitrogen availability was supplemented with 150 KgN ha^{-1} (as urea)

uniformly applied to each plot at the onset of tillering.

Meteorological data for the cropping periods were recorded from the Meteorological station located near CENEB for the first experiment and from the Meteorological station of Meteocat (Servei Meteorologic de Catalunya) close to the experimental fields in the last two experiments (Table 1).

In Bell-lloc (where the more detailed experiments were carried out), the average temperature for the whole cropping duration (November to July) of CS1 was 12.6 ºC whereas CS2 was 11.7 ºC. At the critical stage of anthesis both minimum and maximum temperatures were slightly higher during CS1 than CS2. In general, temperatures, both minimum and maximum were within the ranges normally occurring in the region during past 5 years. As mentioned above, the experimental fields were irrigated as needed: in Bell-lloc both experiments were irrigated around anthesis but in CS2 an additional irrigation was given at seedling emergence stage as late fall – early winter of 2017 was unusually dry (Table 1). Thus, there was only one irrigation in CS1 (on 19 April 2017) and two in CS2 (on 14 December 2017 and 5 May 2018). The average time to anthesis and physiological maturity for the lines was April and June for CS1 and CS2. Each irrigation was equivalent to 80 mm of rainfall.

Table. 1. Meteorological data for experiments in Ciudad Obregón 2015-16, and in Bell-lloc 2016-17 (CS 1) and 2017-18 (CS 2): monthly average of minimum (T min) and maximum (T max) temperatures (± standard error) as well as monthly cumulative precipitation. In all cases data are provided for the growing season (from the month of sowing to that of harvest).

		Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
T min (ºC)	**Cd. Obregón 2015-16**		8.08±0.54	5.79±0.31	8.64±0.56	9.84±0.40	11.24±0.39	14.05±0.57	22.12±0.60	
	Bell-lloc 2016-17	2.73±0.57	1.48±0.42	-1.23±0.66	2.70±0.57	3.50±0.44	4.95±0.45	9.60±0.46	15.10±0.51	16.42±0.41
	Bell-lloc 2017-18	-2.42±0.77	-1.67±0.64	1.80±0.67	-0.60±0.66	3.39±0.43	6.56±0.53	10.73±0.63	14.48±0.37	16.90±0.25
T max (ºC)	**Cd. Obregón 2015-16**		26.08±0.75	25.35±0.38	28.45±0.58	27.79±0.51	31.15±0.39	34.46±0.37	37.54±0.35	
	Bell-lloc 2016-17	13.90±0.50	6.96±0.84	8.45±0.81	13.69±0.48	18.49±0.65	21.78±0.67	26.80±0.70	32.09±0.91	32.73±0.63
	Bell-lloc 2017-18	13.28±0.64	8.37±0.61	12.52±0.75	10.86±0.78	15.93±0.52	20.43±0.86	24.02±0.49	29.48±0.70	33.75±0.39
Rainfall (mm)	**Cd. Obregón 2015-16**		0.2	2.1	2.1	8.4	1.0	0.3	1.7	
	Bell-lloc 2016-17	62.1	7.5	14.4	7.2	102.0	22.5	18.1	24.3	5.5
	Bell-lloc 2017-18	0.2	11.7	27.2	44.0	48.5	77.2	53.8	3.0	21.3

2.2. Genotypes and experimental design

In the experiment at Cd. Obregón we grew the whole NAM population while in the two field experiments at Bell-lloc we grew a selection of this population. The NAM population was generated from 13 bi-parental crosses where both parents in each cross were elite spring wheat varieties selected for having particular traits of interest to be included in the crosses. The parents of the crosses used were (i) Paragon, one of the best UK spring wheat cultivars considering yield potential and disease resistance, was the most common parent used; (ii) four CIMCOG (CIMMYT Core Germplasm: Orford et al., 2014) lines viz.: CIMCOG 49, 47, 3 and 32 characterised for their high values of biomass, grains per spike and harvest index; (iii) Weebill, a cultivar well known for having its high yield associated to superior average grain weight; (iv) MISR1, SUPER152, Pfau, Waxwing and Baj, all parents selected for their high yield related to earliness in time to anthesis; and (v) Wyalkatchem a high performing Australian variety. The 13 families were the lines derived of the following crosses: (1) Weebill × CIMCOG3, (2) Weebill × CIMCOG32, (3) Paragon × Pfau, (4) Paragon × Baj, (5) Paragon × Wyalkatchem, (6) Paragon × (Becard × Kachu), (7) Paragon × MISR1, (8) Paragon × Waxwing, (9) Paragon × Garcia, (10) Paragon × Super151, (11) Paragon × Synth type, (12) Paragon × CIMCOG47, and (13) Paragon × CIMCOG49 (please note that the order of the crosses mentioned here from 1 to 13 will be followed in the result section). Detailed descriptions of the populations, including Axiom 35K genotype files and genetic maps can be found at https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10996, accessed on 05.05.2020. All germplasm is deposited at the CIMMYT genebank.

In the experiment at Cd. Obregón, the original set of 1937 lines were grown in un-replicated hill-plots together with checks (the parents of the crosses and two well adapted genotypes, Reedling and Sokoll) replicated across the whole experiment. Plots were arranged as different families with embedded checks in an augmented design (considering the lines of the NAM, parents and replicated checks there were 2120 hill plots).

In the two experiments conducted in Bell-lloc, we grew 231 lines which is a sub-set from 1937 lines selected based on their field performances in the initial experiment at Cd. Obregón. Treatments consisted of 231 selected lines grown in unreplicated plots together with replicated check plots across the experiments using augmented design in a regular grid (for a scheme showing the distributions of the plots please see Supplementary Fig. S1), design which is commonly used to test large populations where it is not possible to have a complete replication of lines (Scott and Milliken, 1993). Plots in both

experiments were arranged in the field with random allocation of un-replicated 231 genotypes and replicated 3 checks (there were 26 check plots arranged in order to have two check plots in each of the 13 rows of plots arranged diagonally across rows of plots; Müller et al., 2010). The layout of the experiments had 13 rows and 20 columns of plots making it a total of 260 plots, of which 257 corresponded to lines and checks in which traits were measured (the other three plots were sown to complete the rectangular field layout but were not measured). In addition, the whole experiment had a set of 70 border plots that were not considered for measurements (were sown and maintained to avoid border effects on the plots allocated to rows 1 and 13 and to columns 1 and 20 of the measured plots). Each plot consisting of 6 rows was 0.2 m apart and 4 m long. The plot-to-plot distance was 0.5 m between rows and 1 m between columns and the entire experimental area had dimension of 57.5 x 98 m (Supplementary Fig. 1). Three cultivars viz., Paragon, Garcia and Paledor were the checks used both to quantify the spatial heterogeneity across field and as a reference for performance of well adapted cultivars. Paragon was used, as it was the most common parent of the studied NAM population while Garcia (http://www.genvce.org/variedades/trigo-blando/invierno/garcia/; accessed on 14.01.2020 or http://www.agrusa.com/Semillas.php?_b=&_un=1&_do=18&_tr=19; accessed on 05.05.2020) and Paledor (http://www.genvce.org/variedades/trigo-blando/invierno/paledor/; accessed on 14.01.2020 and http://www.agrusa.com/Semillas.php?_b=&_un=1&_do=18&_tr=22; accessed on 05.05.2020) were chosen to be two of the best performing local cultivars at the time we conducted the study. Paledor was indeed a check in the variety trials at least until the cropping season immediately before the CS1 (https://genvce.org/wp-content/uploads/2019/12/informe-genvce-cereal-de-invierno-2015-2016.pdf, accessed on 27.08.2020).

2.3. Measurements and determinations

In the field experiment conducted at Cd. Obregón plant height and anthesis date were determined in all the 2120 hill plots. Based on these determinations, 493 lines, which had similar time to anthesis to that of the checks and discarding extremely short lines, were sampled at maturity and yield per hill as well as AGW were determined. Of these 493 lines, the 231 lines that exhibited best field performance were selected to be evaluated in the more detailed study carried out over the following two seasons in Bell-lloc. This selection was done maintaining the structure of the NAM population: we selected a similar number of lines within each of the 13 families of bi-parental crosses (i.e. the selection of the

213 231 lines have not favoured some of the RIL populations more than others; Table 2).

214

215 **Table. 2.** Number of lines of each of the 13 bi-parental crosses in the whole NAM population grown
216 initially in Cd. Obregón, and in the selected sub-set of lines analysed in more detail in the two field
217 experiments carried out in Bell lloc d'Urgell.

Bi-parental cross	Whole NAM population	Selected sub-set
Weebill x CIMCOG3	157	19
Weebill x CIMCOG32	88	22
Paragon x Pfau	86	15
Paragon x Baj	170	17
Paragon x Wayal	94	17
Paragon x Becard x Kachu	88	17
Paragon x MISR1	86	18
Paragon x Waxwing	77	17
Paragon x Garcia	158	17
Paragon x Super152	92	18
Paragon x Synth type	74	18
Paragon x CIMCOG47	409	19
Paragon x CIMCOG49	358	17
Total	**1937**	**231**

218

219 In the two field experiments in Bell-lloc we determined in each plot different stages of development
220 using the decimal code developed by Zadoks et al. (1974): seedling emergence (stage DC10), onset of
221 stem elongation (DC30), flag leaf emergence (DC39), heading (DC59), anthesis (DC65) and
222 physiological maturity (DC95). All the stages were recorded when 50% of the plot showed that stage
223 by monitoring each plot regularly (from once a week to thrice a week, depending on temperature).
224 The onset of stem elongation (OSE) was determined by touching the main shoot at the base just
225 above the ground to detect the first node and was repeated on several plants in each plot to record
226 the stage for that plot. Later, the OSE data from a parallel but smaller experiment conducted in the
227 same field, in which we also determined the stage of TS by periodic dissection of the apex, was used
228 to estimate the timing of TS from the OSE measurements. Length of phenological phases was
229 estimated in thermal time with base temperature of 0 ºC.

230 Plants were sampled at anthesis (stage DC65) and physiological maturity (DC95) from each individual
231 plot from 1 linear meter which was chosen randomly (from any of the 4 central rows and avoiding the

extreme 25 cm of the rows that were left as borders). Plants in that sampling area were manually pulled out to recover the whole above ground biomass and taken to the laboratory where they were processed to record number of plants, shoots, and productive shoots (shoots bearing spikes) and stem length from the soil level to the base of the spikes. Leaves (only leaf laminae), spikes and stems (including leaf sheaths) were separated and dried in a hot-air oven at 65 ºC for 72 h after which dry weights were recorded. At physiological maturity, spikes and shoots (stem + leaf lamina) were separated to measure biomass separately. The spikes were dried and weighed to attain spike dry weight which were then threshed to obtain grains. Later, the grains were counted and dried again for at least 24 h to measure the grain weights the grain weight was subtracted from spike dry weight to calculate chaff dry weight. Grain yield was calculated using grain number and grain weight then extrapolated to m^2 using total dry weight.

2.4. Analyses

For the data from field experiment in Cd. Obregón only descriptive statistics were performed. Data from the Augmented field experiments in Bell-lloc were analysed using Preliminary phenotypic analysis in GENSTAT (Version 19, VSN International, Hemel Hempstead, UK. Web page: Genstat.co.uk). As a first step in order to correct for spatial trends the replicated row and column entries was subjected to identify the model with most significant variables using the Bayesian information criterion (BIC; Schwarz, 1978) and Akaike information criterion (AIC; Akaike, 1974). Four models (*viz.* identity, autoregressive-AR 1, linear and random models) were compared among which model with the lowest BIC and AIC values was chosen (in all cases the same model was suggested by both criteria, as it most commonly happens; Burnham and Anderson, 2004). After the best correcting model was identified, genotypes were considered random to estimate variance components. These variance components were used in a subsequent analysis, fitting un-replicated entries as fixed terms to obtain un-shrunkenBest linear Unbiased Estimates (BLUEs; Gumedze and Dunne, 2011). The BLUEs were obtained for the directly measured variables using which derived traits were calculated. In the chosen model checks are run as extra genotypes that are replicated and effect of spatial heterogeneity on un-replicated genotypes (i.e. the genotypes being under evaluation) were accounted for using variation observed in checks where rows and column factors were considered random terms.

Relationships between traits were analysed using the BLUEs obtained per plot for each season with linear regressions and we used Pearson's correlation coefficient. We made a raw estimate of the genetic component of the phenotypic correlation when considering all sources of variation together which was the average value for each genotype across both seasons. Genetic correlation as well as statistical significance between traits were calculated using META-R software (Alvarado et al., 2020) where the BLUEs of the two seasons were considered as replicates (since the management for the two experiments were similar and the selected lines lacked G×E interaction).

3. Results

3.1. Genetic variation in the whole NAM population and selection of a sub-set

Expectedly, the ranges of variation in phenology and in plant height were rather large when considering the whole NAM population of 1937 lines (Fig. 1).

Time to anthesis ranged from c. 75 to c. 120 d (Fig. 1a), equivalent to a thermal time range from c. 1150 to c. 2000 ºC d with a base temperature of 0 ºC. This large degree of variation was due to a few parents that had a considerably longer time to anthesis than most others in Cd Obregón as well as a large transgressive segregation particularly for longer periods to anthesis, as the longest times to anthesis in the lines analysed exceeded, by c. 10 d (c. 212 ºC d), the already large range of variation shown by the parents of the population (Fig. 1a). This was in part due to the inclusion of cultivars possessing valuable yield-determining traits beyond time to anthesis (such as Paragon) with a strong photoperiod sensitivity conferring late flowering and maladaptation in Cd. Obregón though many of the lines derived from Paragon would (c. three quarter of the parents differed in time to anthesis by less than a week in this growing condition and half of them flowered within the two days of difference shown by the two well adapted genotypes used as checks; Fig. 1a). As we aimed to identify traits of value beyond time to anthesis and plant height, the data from the first experiment was used to select against variation in time to anthesis that exceeded that of the best adapted local check varieties. Thus, the range of variation in time to anthesis in the selected 493 lines (which were then sampled at physiological maturity to measure hill-plot yield and AGW) was dramatically reduced (Fig. 1a), and could not be further reduced when selecting the 231 lines for later experiments (Fig. 1a).

Plant height in the whole NAM population also varied hugely, from c. 25 to almost 200 cm (Fig. 1b) and in this case mostly due to large transgressive segregation (likely due to segregation of Rht alleles

resulting in some lines being tall and others double dwarf), as parents of the 13 crosses ranged in height from c. 60 to 110 cm and most parents had a height very similar to that of the two well adapted checks (Fig. 1b). The selection of lines that had time to anthesis in the narrow range of best adapted checks reduced the range of variability in height to c. 50 to 120 0.cm and the final selection of lines to be further tested in later experiments reduced that variation further by discarding the shortest plants (<64 cm; Fig. 1b).

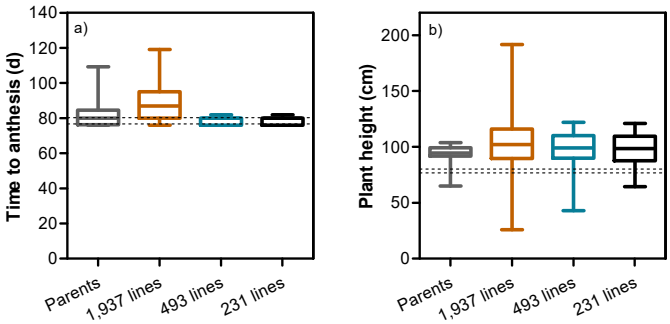


Figure 1. Boxplots for time from sowing to anthesis (a) and plant height (b) from the experiment carried out in Cd. Obregón (NW Mexico) in 2015-16 considering the variability within the parents of the 13 crosses, the whole original NAM population of 1,937 lines, the 493 lines that were sampled and for which yield components were determined, and the 231 lines that were finally selected to be further studied in later field experiments carried out in Bell-lloc (NE Spain). Dashed lines show the values corresponding to two well adapted genotypes used as checks in the experiment, viz. Reedling and Sokoll.

To produce the final selection of the subset of 231 lines to be analysed in more detail we considered the hill-plot yield and yield components of the 493 lines that were sampled in the experiment. The range of GY and its two major components were relatively large (Fig. 2), even though these lines displayed virtually no difference in time to anthesis and exhibited a range of plant height that is substantially reduced compared to the whole NAM population. Indeed, variation in time to anthesis or in plant height explained a negligible proportion of the genotypic variation in GY among the 493 lines (0.4% and 6.6%, respectively). As GY was more related to the number than to the weight of grains, and these major components were negatively related to each other (Supplementary Fig. S2), the selection was made eliminating the lines with lowest number of grains and lowest GY. Consequently, the selected sub-set of 231 lines (15-22 lines per each of the 13 crosses) reduced the

variability in GY and its two components shown in the set of 493 lines, through maintaining the lines with highest GY and GN per hill as well as with intermediate values of AGW (Fig. 2).

3.2. Genetic variation in, and relationships between, GY and phenology within the selected lines

When in the next two seasons these selected 231 lines were grown in Bell-lloc, the range of variation in time to anthesis was larger than for the same lines which had been selected in Cd. Obregón with the aim of constraining anthesis data. However, the timeframe of anthesis was much narrower than

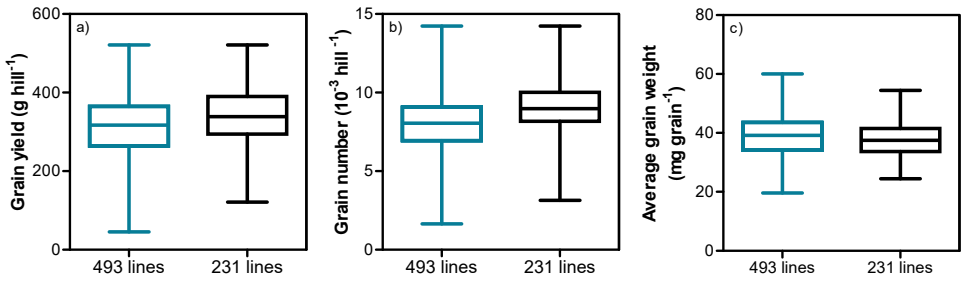


Figure 2. Boxplots of grain yield (a) and its two major components, grain number (b) and their average weight (c) in the experiment carried out in Cd. Obregón (NW Mexico) in 2015-16 considering the variability for the 493 lines that had similar time to anthesis to that of the well adapted checks, and the 231 lines that were further selected to vary less in plant height and which were finally selected to be further studied in later field experiments carried out in Bell-lloc (NE Spain).

would be expected from the whole NAM (n=1937) or a random selection of it. Even though the length of cropping cycle is longer in Spain than Mexico, the range in time to anthesis for the selected panel of 231 lines was much narrower than the original population of 1937 lines evaluated in Cd. Obregón (cf. Figs. 1a and 3a and b). It is also true that although lines were selected for having similar time to anthesis within the families, there was a noticeable variation, not only considering the whole population (1188-1481 ºC d in CS1 and 1209-1525 ºC d in CS2) but also within most families (Fig. 3a, b); implying that the variation across the whole population was not simply reflecting the variation across families but also within them. There was a reasonable degree of consistency for thermal time to anthesis between the two cropping seasons (Fig. 3c).

The same was true for plant height: lines differed in height both across and within families but genotypic differences in height were reasonably consistent across both seasons (Supplementary Fig.S3).

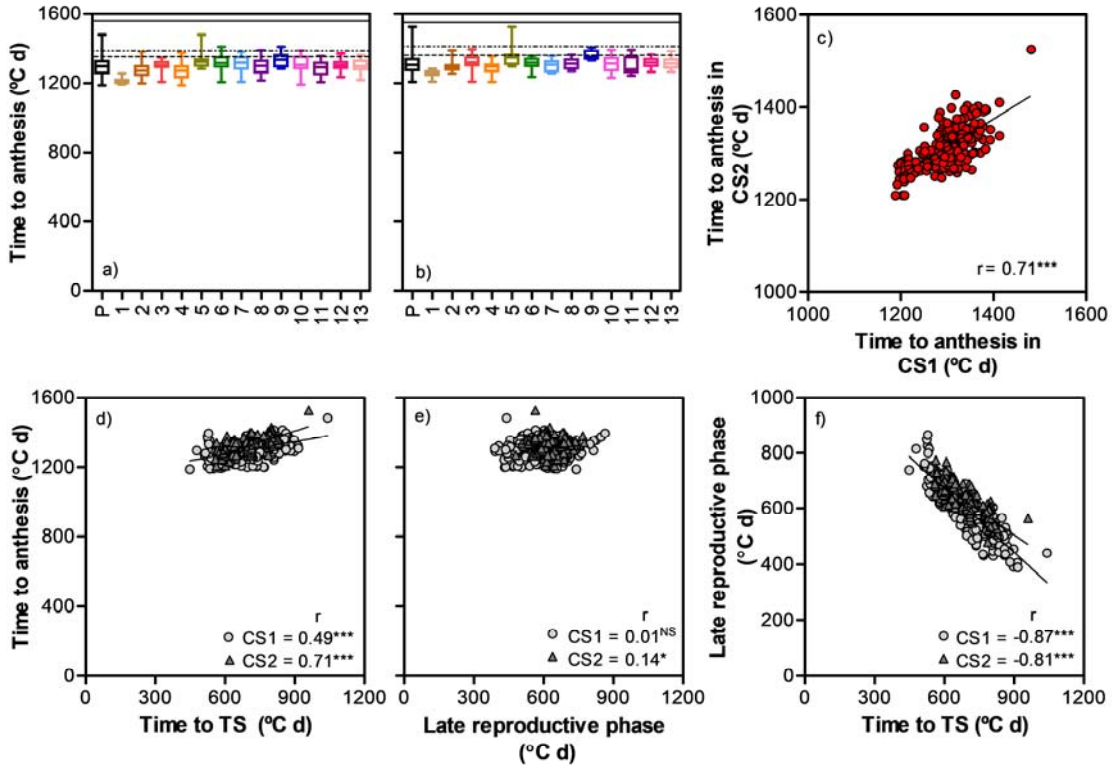


Figure 3. Upper panels: Boxplots showing variability in time from sowing to anthesis within the whole population (P) and within families (13 bi-parental crosses) along with three checks Paragon (solid line), Paledor (dotted line) and Garcia (dashed line) in the first (CS1, a) and second cropping season (CS2, b), and consistency of time to anthesis over the two cropping seasons (c). Bottom panels: Relationships between time to anthesis and its component phases: time from sowing to terminal spikelet (TS, d) and time from then to anthesis, i.e. the late reproductive phase (e); as well as between the two component phases (f) for the 231 lines grown in the first (CS1) and second (CS2) cropping seasons. Note: The crosses corresponding to serial number 1-13 is given in materials and methods; graphs c, d, e and f do not include checks; origin of the graph c does not begin at 0. Coefficients of correlations and their significance level (* $p<0.05$; *** $p<0.001$; NS=non-significant) are shown for the relationships.

Genetic variation in thermal time to anthesis was related to variation for each of the two component phases considered: time from sowing to TS (Fig. 3d) and time since then to anthesis (Fig. 3e), though

the correlations were stronger with the initial phase to TS, embracing the vegetative and early reproductive phases, than with the LRP, suggesting that variation in time to anthesis was mainly controlled by the duration of the first phase in this panel. Furthermore, there was a clear trend for compensation between duration of these two phases both the seasons (Fig. 3f), that was naturally only partial (otherwise there would have been no variation in thermal time to anthesis), as revealed by the slopes that were higher (i.e. less negative) than -1 (-0.76 and -0.57 in CS1 and CS2, respectively). This means that in both cropping seasons it was possible to identify lines with the same time to anthesis but differing oppositely in the duration of the phases of leaf and spikelet initiation and of floret development.

Genotypic variation in GY was large (440 to 1181 g m^{-2} and 459 to 1067 g m^{-2} in CS1 and CS2, respectively). And once again, the variation across the subset of the NAM population reflected more the variation within families than differences across families (Fig. 4a, b). For most of the families, and therefore for the whole population, there were several lines with greater GY than the local checks, Paledor and Garcia, which were modern commercial high-yielding cultivars. However, unlike with time to anthesis and plant height, there was a very large G×E interaction for GY, evident from the absence of significant relationship between GY of the lines across the two cropping seasons (Fig. 4c). This lack of consistency between seasons was also evident for the yield difference between the two well adapted cultivars: while their difference was not significant in CS1 (7.82±0.40 and 7.12±0.42 Mg ha-1) it was larger and highly significant in CS2 (9.19±0.46 and 7.70±0.17).

Although the differences in time to anthesis and plant height could potentially interfere with the relationships between GY and traits other than these two, such potential interference would not be critical in this study as there were no clear relationships between GY and either time to anthesis (Fig. 4d) or plant height (Fig. 4e). Although time to anthesis was significantly related to GY in CS2, it explained only 7% of the GY variation, whilst time to anthesis in CS1 and plant height in both seasons explained less than 1% of the variation in GY (Fig. 4d, e).

Even though the total time to anthesis did not explain differences in GY within the whole population, the partitioning of that time into phases occurring before or after TS seemed to have some significance: across lines and within each of the two seasons GY tended to be negatively related to the duration of the first phase, time from sowing to TS (Supplementary Fig. S4a) and positively related to the length of the following phase, the LRP (Supplementary Fig. S4b). However, even when statistically significant, these relations were weak.

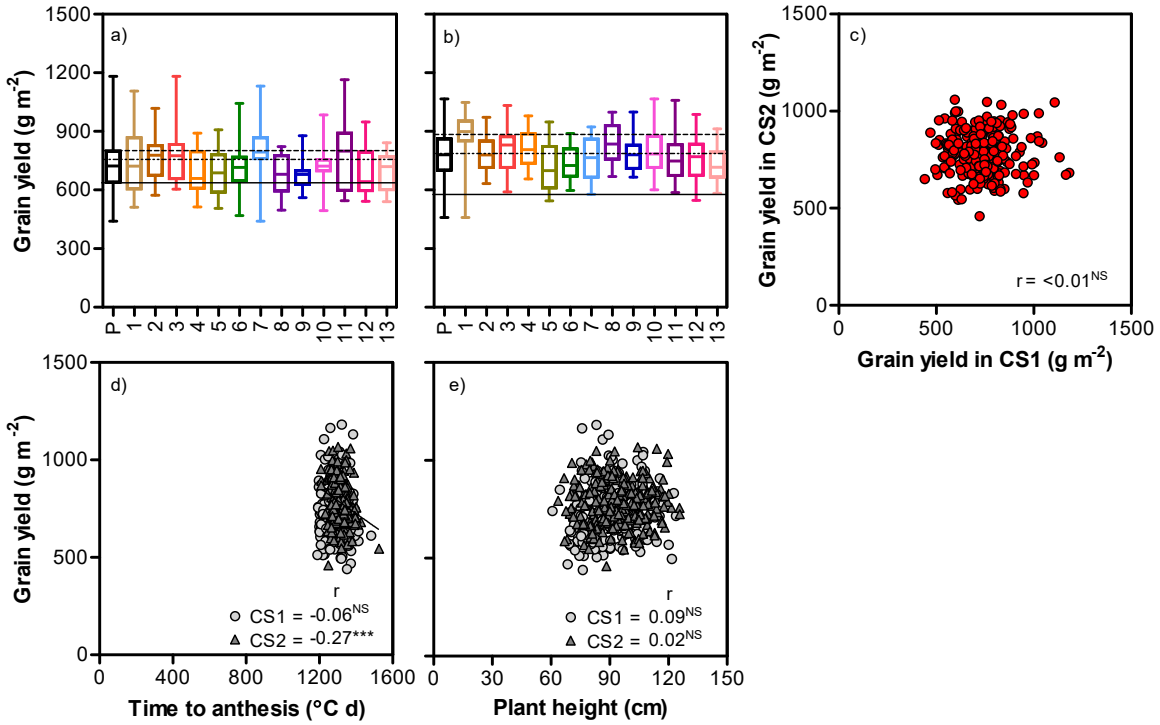


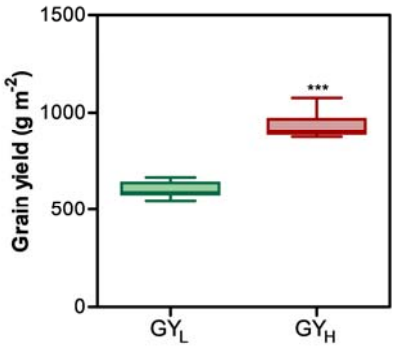
Figure 4. Upper panels: Boxplots showing large variability for grain yield within the whole population (P) and within each family (13 bi-parental crosses) along with three checks Paragon (solid line), Paledor (dotted line) and Garcia (dashed line) in the first (CS1, a) and second cropping season (CS2, b) and inconsistency for grain yield over two seasons (c). Bottom panels: Relationships between grain yield and either total time to anthesis (d) or plant height (e). Note: The crosses corresponding to serial number 1-13 is given in materials and methods; graph c, d and e do not include checks. Coefficients of correlations and their significance level (*** $p<0.001$; NS=non-significant) are shown for the relationships.

Taking into account the large G×E interaction (reflected by the lack of consistency in GY between CS1 and CS2) and our aim to identify trait relationships that can suggest traits relevant for further raising yield through breeding, we identified sub-groups within the sub-set of 231 lines that consistently expressed the extremes of GY over the two seasons: we chose all lines that were part of the bottom and top quartiles of GY in both seasons, low- and high-GY (GY_L and GY_H, respectively). Applying this criterion, there were 13 GY_L and 11 GY_H lines.

411 Naturally these two sub-groups of lines differed in GY highly significantly, with no overlap between
412 them (i.e. the lowest-yielding line of GY_H clearly out yielded the highest-yielding line of GY_L; Fig. 5).
413 However, there were also clear genotypic differences in GY within each of these two sub-groups (Fig.
414 5). By virtue of the data selection made, major genetic variation in GY between the two sub-groups
415 were highly consistent across seasons. We focused on the analysis of traits determining GY in the
416 selected lines within and across these GY_L and GY_H. For the benefit of readers who may be interested
417 in the relationships across the whole sub-set of 231 lines we did also report the relationships for them,
418 naturally for each season separately due to the large G×E interaction in GY, in supplementary figures.
419 The mainstream relationships will be shown across all sources of variation (the 24 genotypes at each
420 of the two growing seasons) as well as using only the genotypes as drivers for the phenotypic
421 correlation (averaged for each genotype across both seasons) as in Elía et al. (2016). In both cases,
422 relationships were established for the two sub-groups separately (highlighting whether the considered
423 trait was relevant or not for the genetic variation within GY_L and GY_H lines) and for all of them
424 together (highlighting the contribution of the trait to produce the consistently highest-yielding lines
425 of the population) and described the variation levels in these traits between these two sub-groups. For
426 these relationships' overall genotypes, we also calculated the genetic correlations (genetic relationship
427 between traits: Alvarado et al., 2020).

428

429



430

431 **Figure 5.** Boxplot depicting variation in grain yield between selected sub-groups of lines being
432 consistently low- and high-yielding in both cropping seasons (GY_L and GY_H, respectively). Significance
433 level: *** $p < 0.001$.

434

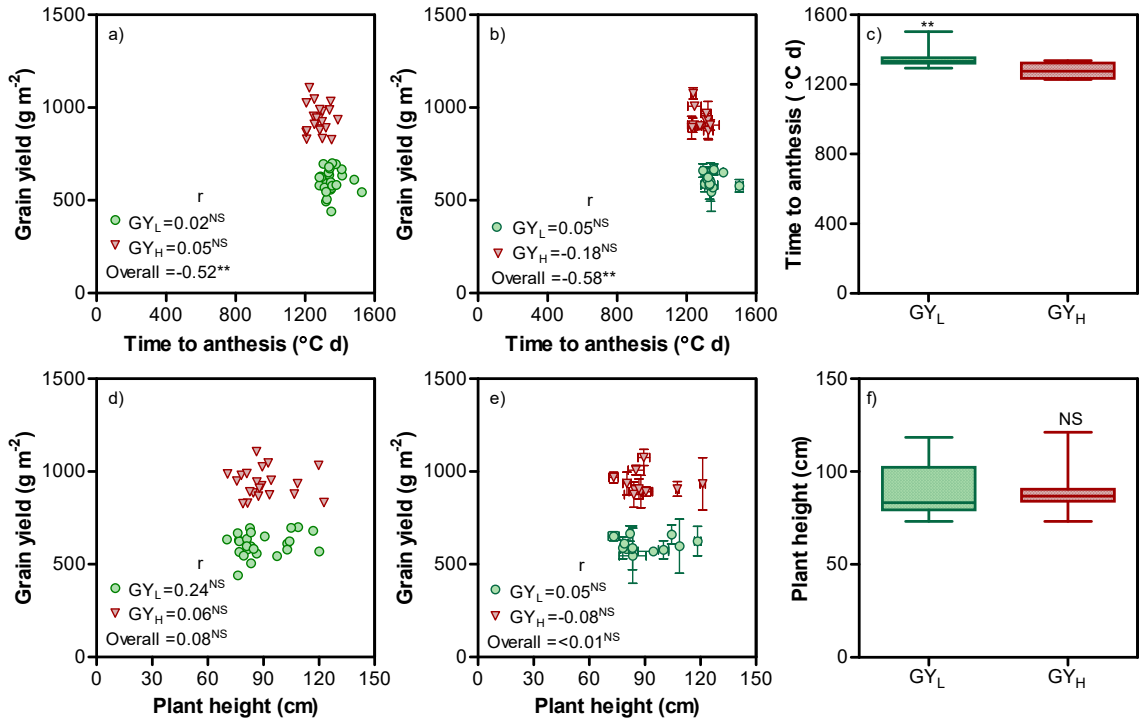
435 **3.3. Determinants of grain yield differences in selected lines**

436 GY variation within the GY_L and GY_H lines was totally unrelated to time to anthesis (Fig. 6a).
437 Regression across all lines, i.e. considering both groups, did show a negative relationship when
438 considering as source of variation both the cultivars and growing seasons (Fig. 6a) as well as when
439 considering the genotypic differences (Fig. 6b). This overall relationship was due to the fact that lines
440 from sub-group GY_L were slightly later than those of GY_H (Fig. 6b), whilst the yield differences in
441 yield within these sub-groups were totally unrelated to time to anthesis (Fig. 6a, b). However, the
442 influence of this difference on GY between the two sub-groups would have been negligible for several
443 reasons. Firstly, the overall GY variation explained by time to anthesis variation was relatively small
444 (c. 35%). Secondly, the groups show extensive overlap to the extent that many GY_H lines have the
445 same time to anthesis that many GY_L lines (Fig. 6b) still having substantially higher yields (Fig. 5). In
446 fact, only few lines in each sub-group accounted for the significance of the difference in time to
447 anthesis between the two yielding categories. Finally, and in relation to that distribution, the average
448 time to anthesis between GY_L and GY_H lines was only slightly different (70 ºC d, equivalent to c. 3 d;
449 Fig. 6b) and that would hardly explain the large difference in average yield (>300 g m^{-2}; Fig. 5).

450 Regarding plant height, although relatively more variable than time to anthesis, the lack of influence
451 of this trait on GY was even more clear, as the relationships were not significant both within and
452 across yielding sub-groups (Fig. 6d, e), neither when considering all sources of variation (Fig. 6d) or
453 only the genotypic differences (Fig. 6e). Indeed, the genetic correlation was even lower ($r_{genetic}$=0.02;
454 P=0.94). The range of variation in plant height between the two sub-groups contrasting in GY was
455 totally overlapped, implying that across them the difference in height was not significant (Fig. 6f).

456 The fact that neither of these two traits had a relevant role is important as we were interested in
457 identifying likely traits responsible for differences in GY of elite material other than time to anthesis
458 and plant height. And this lack of relevance was also evident if the analysis were made with the whole
459 sub-set of 231 lines (see above and Fig. 4d, e).

460



461

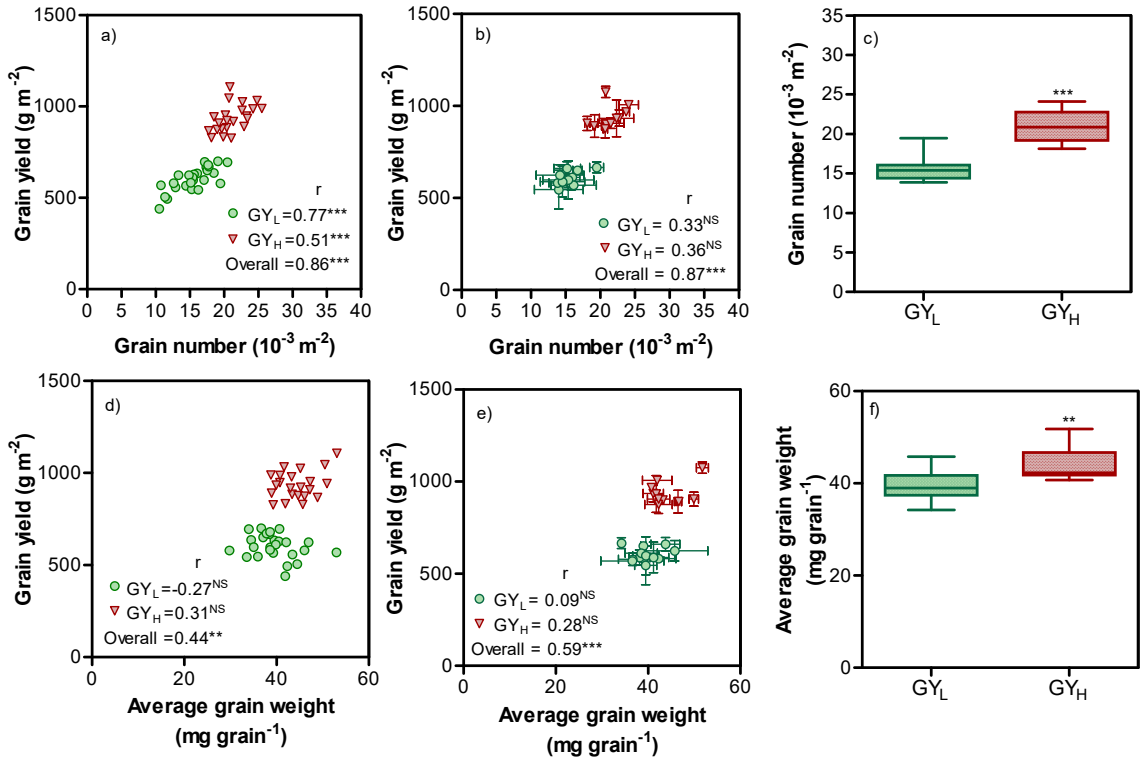
462 **Figure 6.** Relationships between yield and either time to anthesis (a, b) or plant height (d, e) for the
463 selected sub-groups of low and high yielding lines (GY_L and G, respectively), for the combinations of
464 the 24 genotypes and two growing seasons (a, d) or for genotypic means across growing conditions
465 (b, e); and boxplots describing the variation in these traits for the two sub-groups (c, f). Coefficients
466 of phenotypic correlations are shown for each sub-group individually and for the overall data across
467 both sub-groups. Significance level: ** $p<0.01$; NS=non-significant.

468

469 Considering the two major GY components, it was clear that variations in GY were almost solely
470 explained by GN (Fig. 7). Considering the variation in GY within sub-groups, it was better explained
471 by those in GN (Fig. 7a) than in AGW (Fig. 7d). However, when analysing the relationships
472 minimising the environmental effect the coefficient of correlation was non-significant for both GY_L
473 and GY_H (Fig. 7b), though still higher than the coefficient of correlation for GY and AGW within any
474 of the two sub-groups (Fig. 7e); In addition, it is also true that the highest yielding line in the GY_H
475 sub-group had intermediate GN but the highest AGW within that sub-group (Fig. 7b, e). But most
476 importantly when trying to determine the reasons why the GY_H sub-group out-yielded the GY_L sub-
477 group, GN was a far more robust determinant of GY than AGW considering all lines across both
478 sub-groups (cf. Fig. 7a and 7d, where it can be seen that more than 70% of the overall variation in GY
479 was related to that in GN, while this percentage was less than 40% when considering AGW). This was

480 further reinforced by the fact that the genetic correlation between GY and GN was extremely high
481 ($r_{genetic}$=0.95; P<0.001) and clearly higher than that between GY and AGW ($r_{genetic}$=0.78; P<0.001).

482



483

484 **Figure 7.** Relationships between grain yield and its two major components: grain number (a, b) and
485 average grain weight (d, e) for the selected sub-groups of low and high yielding lines (GY_L and GY_H,
486 respectively), for the combinations of the 24 genotypes and two growing seasons (a, d) or for genotypic
487 means across growing conditions (b, e); and boxplots describing the variation in these traits for the
488 two sub-groups (c, f). Coefficients of correlations are shown for each sub-group individually and for
489 the overall data across both sub-groups. Significance level: ** p<0.01; *** p<0.001; NS=non-
490 significant.

491

492 Therefore, whilst there was virtually no overlap between the ranges of GN of GY_L and GY_H lines,
493 which differed as groups significantly (in average GY_H lines had almost 40% more grains m^{-2} than
494 GY_L lines; Fig. 7c), AGW of GY_L and GY_H lines displayed noticeable overlapping (in average GY_H
495 lines had c. 10% heavier grains than GY_L lines; Fig. 7f).

496 The relationship between these GY components were clearly negative within each of the two sub-
497 groups (Supplementary Fig. S5). However, this did not represent complete compensation as in both

498 sub-groups increasing GN improved GY (Fig. 7a). Furthermore, this relationship not only lost
499 significance but also became positive when all lines were considered together as the GY_L lines did
500 have fewer grains that were slightly lighter than the GY_H lines (Supplementary Fig. S5). Indeed, the
501 genetic correlation between AGW and GN when considering both sub-groups together was positive
502 and significant ($r_{genetic}$=0.54; P<0.01).

503 The proposed focus on GN was reinforced by our analysis of the variation within the sub-set of 231
504 lines in each of the two seasons (Supplementary Fig. S6). GN was significantly related to GY in both
505 seasons and GY was also significantly related to AGW although only in CS2 and the magnitude of the
506 association was marginal in absolute terms and negligible compared with that of GN (cf.
507 Supplementary Fig. S6a and b). In both seasons the two major GY components were negatively related
508 across all lines but again this negative relationship did not result in a clear compensation
509 (Supplementary Fig. S6c).

510 Thus, to understand the traits responsible for the yield advantage of GY_H over the GY_L lines, it is GN
511 which requires further dissection.

512

513 **3.4. Physiological components of grain number**

514 Physiological components of GN, SDWa and FE, explained part of GN variation (Fig. 8). However,
515 their relative relevance seemed to vary with the type of comparison. When comparing the differences
516 across GY_L and GY_H lines it seemed that SDWa was most important as the overall relationship was
517 significant when considering all sources of variation (Fig. 8a) or only the average across the two
518 growing seasons (Fig. 8b), as GY_L lines exhibited significantly lower values than those of GY_H (Fig.
519 8c). However, GN was unrelated to SDWa within each of the two sub-groups (Fig. 8a, b). The
520 contribution of FE to differences in GN across both GY_L and GY_H lines was rather relevant when
521 considering all sources of variation (Fig. 8d), but the main driver for that overall relationship was the
522 difference between seasons: when the relationship was fitted with the genotypic values averaged across
523 both seasons it became not significant (Fig. 8e) and naturally lower than that of SDWa. Although GN
524 in GY_L lines was better explained by FE than SDWa, this was not true for the differences in GN within
525 the GY_H lines (cf. Fig. 8b and 8e). Furthermore, FE was not significantly different between the GY_H
526 and GY_L lines (Fig. 8f). Thus, considering both sub-groups together the increase in SDWa did not
527 imply a compensation due to reductions in FE (cf. Figs. 8c and 8f). Indeed, although there was a clear

negative relationship between the two physiological determinants of GN within each of the two groups, the overall phenotypic correlation was lower (p=0.048; Supplementary Fig. S7). This was because in this relationship GY_H lines were displaced to the right (Supplementary Fig. S7) as a result of their overall higher SDWa (i.e. the increase in SDWa of the GY_H lines compared to the GY_L lines) that did not bring about any compensation in FE (Figs. 8c, f). This lack of compensation was further emphasised by the fact that there was not a genetic correlation between FE and SDWa ($r_{genetic}$=-0.02; P=0.92).

Again, should we have focused the analysis in the comparison in sub-set of 231 lines, the outcome would have been similar. Regardless of the trade-off exhibited by FE and SDWa (Supplementary Fig. S8c), both similarly influenced GN across all lines in each of the two cropping seasons (Supplementary Fig. S8a, b).

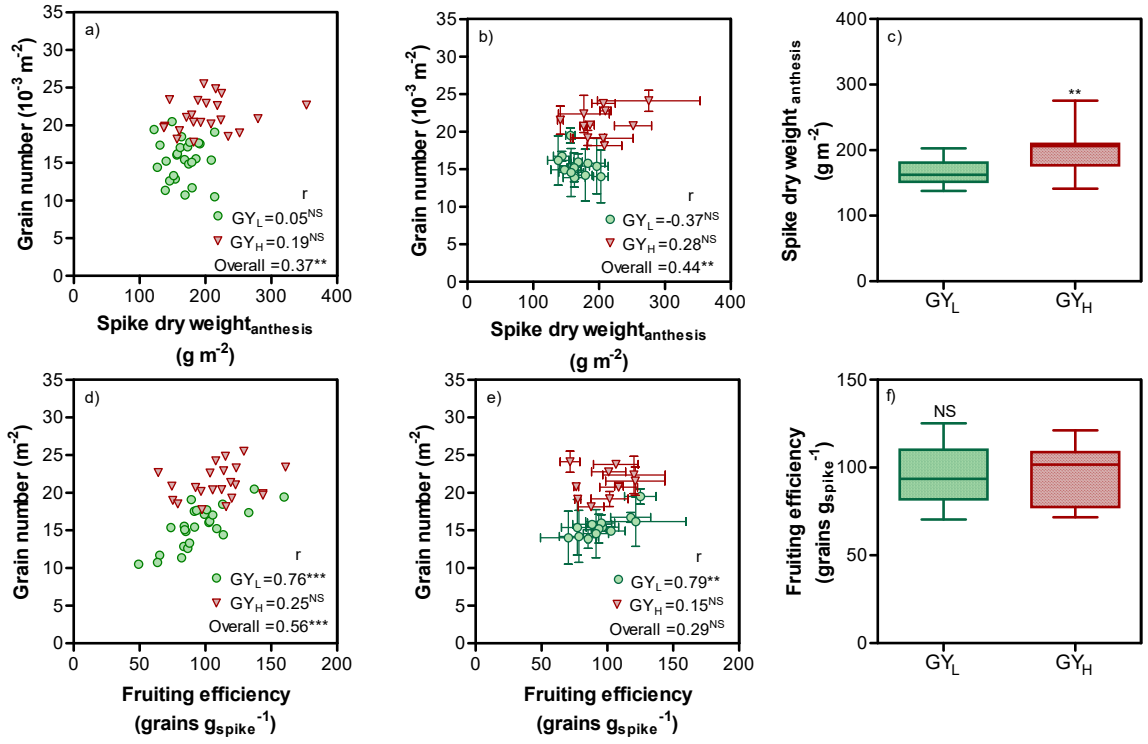


Figure 8. Relationships between grain number and two of its physiological determinants: spike dry weight at anthesis (a, b) and fruiting efficiency (d, e) for the selected sub-groups of low and high yielding lines (GY_L and GY_H, respectively), for the combinations of the 24 genotypes and two growing seasons (a, d) or for genotypic means across growing conditions (b, e); and boxplots describing the variation in these traits for the two sub-groups (c, f). Box plots describing variability for these traits in two sub-groups (b, d). Coefficients of correlations are shown for each sub-group individually and for the overall data across both sub-groups. Significance level: ** p<0.01; *** p<0.001; NS=non-

547 significant.

548

549 The negative relationship between GN and AGW (Supplementary Fig. S6c) is mirrored by the negative
550 relationship between AGW and FE (Supplementary Fig. 9). It is the latter which best explains
551 phenotypic differences in AGW (Supplementary Fig. S9a). This was also evident when analysis in sub-
552 set of the 231 lines was considered (Supplementary Fig. S9b). However, there was no genetic basis for
553 this negative phenotypic correlation as the genetic correlation was inexistent ($r_{genetic}$=-0.01; P=0.96).

554

555 **4. Discussion**

556 The lack of association for yield between the two growing seasons was not expected in principle.
557 Although large G×E interaction for complex traits like yield are the most common scenario in the
558 agronomic literature, more consistency is common when crop management is the same. It must be
559 noted, however, that from the total of 1937 lines of the NAM population grown in the first field
560 experiment we discarded the 1706 lines with poorest performance and carried out the detailed
561 experiments with the selected 231 lines (i.e. 12% of the best performing offspring), in which the
562 conclusions are based. This selection has the "disadvantage" that we removed all the genotypes with
563 poor performance (and those are surely a major component of maintaining the consistency between
564 line yields across years in experiments with large populations). But the advantage is that this resembles
565 what breeders do in eliminating poor performers in each cycle of selection. As we aimed at identifying
566 traits responsible for GY differences among lines derived from crosses of elite germplasm, beyond
567 the effects of time to anthesis and plant height, investigating a population that reflects standard
568 breeders' practise is relevant.

569 Although not considering time to anthesis, we were interested in the partitioning of phenological time
570 in the duration of phases occurring before and after TS. In line with previous research (Halloran and
571 Pennell, 1982; Slafer, 2003; Whitechurch et al., 2007) there was a large variation in the two pre-anthesis
572 phases; and there was clear negative relationship between these two phases. This confirms that it
573 would be possible to lengthen the duration of the LRP at the expense of shortening the duration of
574 the phase from sowing to TS (Miralles et al., 2000; Scarth et al., 1985; Slafer et al., 2001). In this
575 context, importance of these observations rests on the hypothesis that a longer LRP would
576 accommodate increased partitioning of biomass to the growing spike (González et al., 2005b; Miralles

et al., 2000; Reynolds et al., 2005; Slafer et al., 2005) with the beneficial knock on effect of increased floret survival and final grain number (Ferrante et al., 2013; Sadras and Slafer, 2012), given all other things are constant. However, the relationships between GY and duration of LRP were positive but rather weak, implying that within this experimental material the duration of this phase was not the most critical trait determining yield differences among lines (as also recently seen in Australia; Zhang et al., 2019).

GN was the main component explaining variations in GY pointing us towards a prioritisation of this trait to explore future genetic gains (Slafer et al., 2014), while not suggesting that maximising AGW is unimportant in the ultimate high yielding varieties produced by breeders (as illustrated by the fact that within the selected lines used for the more detailed characterisation the highest yielding lines had a distinctly higher AGW than the others). To plan for the optimisation of both traits it is important to increase understanding of their negative correlation, which was observed here, as in so many previous studies (Miralles and Slafer, 1995; Siddique et al., 1989a; Slafer et al., 2014). A key question is whether the AGW/GN negative relationship is due to competition for carbohydrates during grain filling. Should there be competition among grains, increasing GN might result in a kind of zero-sum game, with GY not changing significantly. Although, this kind of interpretation of a negative relationship seems quite intuitive, there is good evidence for the less intuitive scenario in which grains do not compete for resources during grain filling. It follows that the source of the trade-off lies elsewhere and is probably not the consequence of a competitive dynamic (Acreche and Slafer, 2006; Miralles and Slafer, 1995). This means that further increases in GN are critical in bringing about major improvements in GY (Fischer, 2011; Reynolds et al., 2012; Sanchez-Garcia et al., 2013; Slafer et al., 2014). These extra grains will have access to sufficient resources for filling as supported by studies from source-sink manipulations during grain filling in which grain growth is unresponsive (Abbate et al., 1997; Borrás et al., 2004b; M. P. Reynolds et al., 2005; Serrago et al., 2013 and references quoted there in; Slafer and Savin, 1994) showing that an excess of assimilates are available at this developmental stage (e.g. Bingham et al., 2007; Borrill et al., 2015); although some exceptions can be found and only for particular seasons under extremely high yielding conditions (e.g. Lynch et al., 2017). In other words, that the crop is rather conservative at the time of establishing GN (Sadras and Slafer, 2012), which would be the reason for the differences in plasticity of GN and AGW (being GN determination strongly responsive to source strength and AGW relatively unresponsive).

The current study did not involve source-sink interventions like defoliations, shading, de-graining, or

608 thinning the plots and so on, nonetheless a quantitative analysis can be used to estimate whether strong
609 source limitation during the grain filling period was at play. For this purpose, we related AGW to the
610 post-anthesis accumulation of crop growth on a per grain basis (i.e. the ratio between total above
611 ground crop dry weight accumulated from anthesis to physiological maturity and GN). This showed
612 that AGW differences between lines were hardly due to severe source limitations in the low AGW
613 genotypes (Fig. 9a). Firstly, there was no clear trend to reduce AGW with reductions in post-anthesis
614 whole crop growth per grain. Secondly, the distributions of the data-points in the figure would be
615 compatible with no source-limitation. Almost all the lines in GY_H were very close to the 1:1 line
616 indicating that the grain growth had more than enough resources: cases in which data points are below
617 the 1:1 line would have never being source-limited to the level that even some of the growth produced
618 after anthesis was allocated to other sinks, while cases in which grain weight exceeded the post-anthesis
619 crop growth per grain would still be sink-limited, as post-anthesis growth is only part of the source
620 available to fill the grains (part of the demand of the growing grains can be satisfied by remobilisation
621 of pre-anthesis reserves). This is also true for the data-points from GY_L that fell at the left side of the
622 1:1 line (Fig. 9a). To have a scale that can be more readily compared with the literature we transformed
623 the independent variable into a percentage of AGW (Fig. 9b). Calculated as percent difference between
624 AGW and post-anthesis crop growth per grain with respect to AGW. A negative value means the
625 percentage of GY that was allocated to other organs (i.e. clearly unrealised yield potential due to post-
626 anthesis sink limitation), and positive values represent the percentage of AGW that has been realised
627 thanks to the remobilisation of pre-anthesis reserves. The dotted line at 35% indicates a practical limit
628 up to which developing grains can access translocated pre-anthesis reserves derived from Savin and
629 Slafer (1991). This is a rather conservative figure as there have been examples in the literature where
630 up to 50% of the final grain weight was contributed by translocation of reserves accumulated before
631 anthesis (e.g. Borrás et al., 2004; Gent, 1994) which produced an elegant demonstration of the fact
632 that only with a large contribution of pre-anthesis reserves to grain growth the observed AGW would
633 have been possible. In that work it was estimated that, depending on the cultivar and season, up to
634 55% of the final AGW could be contributed from pre-anthesis reserves and several examples of such
635 large contribution have been observed (see examples in the review by Blum, 1998).
636

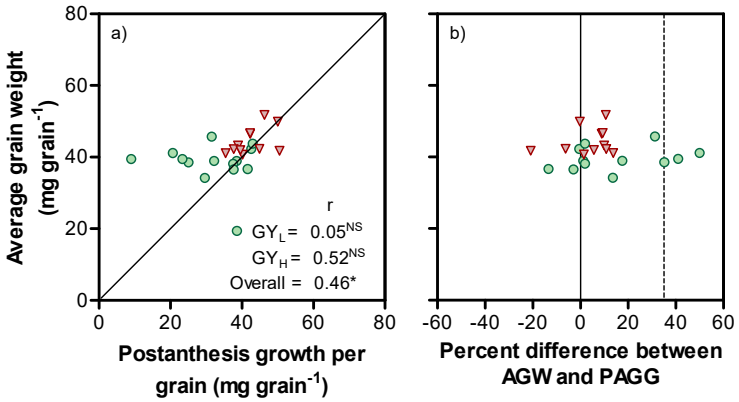


Figure 9. Relationships between average grain weight and either (i) post-anthesis growth per grain (PAGG) in absolute (a), or ii) percent difference between AGW and PAGG with respect to AGW (b) for the selected sub-groups of low and high yielding lines (GY_L and GY_H, respectively). Coefficients of correlations are shown for each sub-group individually and for the overall data across both sub-groups. Significance level: * $p<0.05$; NS=non-significant. Plain lines represent the situation when AGW was equal to post-anthesis growth per grain. The dotted line represents a 35% contribution from pre-anthesis reserves to final grain weight, which is more than a highly likely contribution that can be expected (Austin *et al.*, 1980 in barley; Savin and Slafer, 1991 in wheat).

Furthermore, there was a relationship between GY and total dry weight (at physiological maturity) explaining the genotypic GY differences within and across the GY_L and GY_H lines (Fig. 10a). The most frequent interpretation of this relationship would be that lines with improved growth capacity produced more grains that, when filled, resulted in a proportionally larger GY. However, this seems not to be the most likely explanation in this case. Looking at the differences in total accumulated dry weight from sowing to anthesis (TDWa; Fig. 10b) and from anthesis to maturity i.e., cumulative growth after anthesis (Fig. 10c), it seems that the more common cause-consequence hypothesis can be inverted to reach an interpretation that is at least as likely. Indeed, there was only a marginal difference in TDWa between GY_L and GY_H lines, with a large overlap in this trait between lines of the two sub-groups (Fig. 10b), while the difference became relevant in post-anthesis growth (Fig. 10c). Thus, it may well be that the physiological basis for the higher GY of the GY_H lines is more efficient translation of pre-anthesis growth into GN. These lines increased the sink strength during grain filling lowering the extent of sink limitation. This, in turn, would reduce the down regulation of post-anthesis canopy photosynthesis (that has been shown to exist due to insufficient sink demand in different conditions; e.g. Serrago et al., 2013) driving the improved crop growth during grain filling. This would be in line with previous studies showing that higher GN would increase post-anthesis

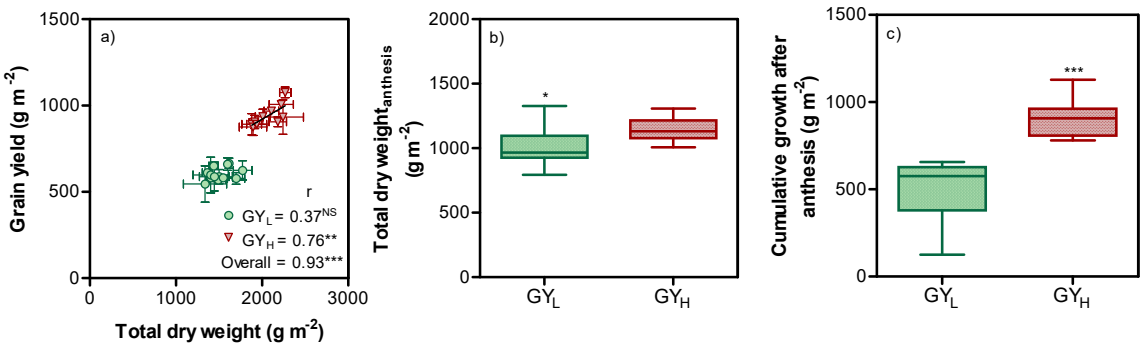
663 growth (Acreche and Slafer, 2009; Reynolds et al., 2005), through its positive effect on canopy
664 photosynthesis.

665

666

667

668



669

670 **Figure 10.** Relation between total dry weight (at maturity) and grain yield (a); box plots showing
671 variations in total dry weight at anthesis (b) and cumulative growth after anthesis (c) for the selected
672 sub-groups of low and high yielding lines (GY_L and GY_H, respectively). Coefficients of correlations
673 are shown for each sub-group individually and for the overall data across both sub-groups.
674 Significance level: ** $p<0.01$; *** $p<0.001$; NS=non-significant.

675

676 Both physiological components of GN, SDWa and FE, seemed to have been relevant to improve GY,
677 depending on whether the analysis was focused within or across sub-groups. This is in line with recent
678 results from Australia in a study combining many commercial cultivars, elite lines and a MAGIC
679 population (Zhang et al., 2019). As lines did not differ much in TDWa their differences in SDWa
680 implies a better dry matter partitioning to the juvenile spike growing immediately before anthesis in
681 high GY_H lines. This is critical because wheat GY is clearly source-limited just prior to anthesis (Borrás
682 et al., 2004; Slafer and Savin, 1994) and SDWa is critical in determining post-anthesis sink strength
683 (Fischer, 2011; Slafer, 2003). This is because the development of labile florets in the juvenile spikes
684 immediately before anthesis is highly sensitive to allocation of resources (Ferrante et al., 2013, 2010;

685 Fischer, 1985; González et al., 2005a; Siddique et al., 1989b; Slafer et al., 2015), which in turn is the
686 mechanistic basis for the strong and consistent relationship between GN (or number of fertile florets)
687 and SDWa (Fischer, 1985 and a plethora of papers confirming this relationships in different
688 background conditions, in response to various different treatments). In the past, breeding has
689 improved GY through increasing GN exploiting this mechanism. Specifically, modern semi-dwarf
690 cultivars out yielded their older traditional height (tall) predecessors due to an improved dry matter
691 partitioning to the spike before anthesis (e.g. Brooking and Kirby, 1981; Calderini et al., 1995; Fischer
692 and Stockman, 1986; Flintham et al., 1997; Miralles et al., 1998; Shearman et al., 2005; Siddique et al.,
693 1989a; Slafer and Andrade, 1993). But these gains were achieved through plant height reduction. In
694 the present study we showed that there is opportunity to further improve partitioning of dry matter
695 to the spike beyond reductions in plant height (Foulkes et al., 2011) that would be instrumental to
696 further improve GY through reducing the degree of sink-limitation during the post-anthesis phases
697 of development. A recent paper by Rivera-Amado, et al. (2019) clearly illustrates how this further
698 improvement in pre-anthesis partitioning to juvenile spikes would be possible. The other trait that can
699 help in reducing the sink-limitation during grain filling is FE (e.g. Slafer et al., 2015). In this study, FE
700 explained part of the GN differences within the segregants from elite parents. Although we observed
701 trade-off between SDWa and FE that had been also reported before (e.g. Ferrante et al., 2012; Lázaro
702 and Abbate, 2012), this was mainly within sub-groups, as the reason why the GY_H lines had more
703 grain set than high GY_L lines was that differences in SDWa were not compensated by reductions in
704 FE. Indeed, there was no genetic correlation between FE and SDWa. Thus, it seemed feasible to
705 identify genotypes with best combinations of both traits maximising GN, and therefore crossing
706 parents with high SDWa and high FE could increase the likelihood of transgressive segregation for
707 GN (and GY).

708

709 **Acknowledgement**

710 Funding for the experimental work was partly provided by projects AGL2015-69595-R, from the
711 Spanish Research Agency (AEI), and IWYP25FP, from the International Wheat Yield Partnership
712 (IWYP). We thank Semillas Batlle for their technical support in maintenance of field experiments. We
713 are grateful to the team of laboratory of crop physiology for assisting with field cum laboratory work,
714 and Prof. Ignacio Romagosa for helping with the experimental design and statistical analysis. PB held
715 a pre-doctoral research contract from the agency for Management of University and Research

716 (AGAUR) from the *Generalitat de Catalunya*.

717

718 **References**

719 Abbate, P.E., Andrade, F.H., Culot, J. P, Bindraban, P.S., 1997. Grain yield in wheat: Effects of
720 radiation during spike growth period. Field Crops Res. 54, 245–257.
721 https://doi.org/10.1016/S0378-4290(97)00059-2

722 Abbate, P.E., Andrade, F.H., Culot, J.P., 1995. The effects of radiation and nitrogen on number of
723 grains in wheat. J. Agric. Sci. 124, 351–360. https://doi.org/10.1017/S0021859600073317

724 Acreche, M.M., Briceño-Félix, G., Sánchez, J.A.M., Slafer, G.A., 2008. Physiological bases of genetic
725 gains in Mediterranean bread wheat yield in Spain. Eur. J. Agron. 28, 162–170.
726 https://doi.org/10.1016/j.eja.2007.07.001

727 Acreche, M.M., Slafer, G.A., 2009. Variation of grain nitrogen content in relation with grain yield in
728 old and modern Spanish wheats grown under a wide range of agronomic conditions in a
729 Mediterranean region. J. Agric. Sci. 147, 657–667.
730 https://doi.org/10.1017/S0021859609990190

731 Acreche, M.M., Slafer, G.A., 2006. Grain weight response to increases in number of grains in wheat
732 in a Mediterranean area. Field Crops Res. 98, 52–59. https://doi.org/10.1016/j.fcr.2005.12.005

733 Aisawi, K.A.B., Reynolds, M.P., Singh, R.P., Foulkes, M.J., 2015. The physiological basis of the
734 genetic progress in yield potential of CIMMYT spring wheat cultivars from 1966 to 2009. Crop
735 Sci. 55, 1749–1764. https://doi.org/10.2135/cropsci2014.09.0601

736 Alvarado, G., Rodríguez, F.M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., Pérez-Rodríguez,
737 P., Lopez-Cruz, M.A., 2020. META-R: A software to analyze data from multi-environment
738 plant breeding trials. Crop J. https://doi.org/10.1016/j.cj.2020.03.010

739 Araus, J.L., Slafer, G.A., Reynolds, M.P., Royo, C., 2002. Plant breeding and drought in C3 cereals:
740 What should we breed for? Ann. Bot. 89, 925–940. https://doi.org/10.1093/aob/mcf049

741 Bingham, I.J., Blake, J., Foulkes, M.J., Spink, J., 2007. Is barley yield in the UK sink limited?. II.
742 Factors affecting potential grain size. Field Crops Res. 101, 212–220.

https://doi.org/10.1016/j.fcr.2006.11.004

Bingham, J., Wellington., 1981. The Achievements of Conventional Plant Breeding. Philos. Trans. R. Soc. B Biol. Sci. 292, 441–455. https://doi.org/10.1098/rstb.1981.0039

Blum, A., 1998. Improving wheat grain filling under stress by stem reserve mobilisation (Reprinted from Wheat: Prospects for global improvement, 1998). Euphytica 100, 77–83. https://doi.org/10.1023/A:1018303922482

Borrás, L., Slafer, G.A., Otegui, M.E., 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: A quantitative reappraisal. Field Crops Res. 86, 131–146. https://doi.org/10.1016/j.fcr.2003.08.002

Borrill, P., Fahy, B., Smith, A.M., Uauy, C., 2015. Wheat grain filling is limited by grain filling capacity rather than the duration of flag leaf photosynthesis: A case study using NAM RNAi plants. PLoS One 10 (8). https://doi.org/10.1371/journal.pone.0134947

Brooking, I.R., Kirby, E.J.M., 1981. Interrelationships between stem and ear development in winter wheat: The effects of a Norin 10 dwarfing gene, Gai/Rht2. J. Agric. Sci. 97, 373–381. https://doi.org/10.1017/S0021859600040806

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference: Understanding AIC and BIC in model selection. Sociol. Methods Res. 33, 261–304. https://doi.org/10.1177/0049124104268644

Calderini, D.F., Dreccer, M.F., Slafer, G.A., 1995. Genetic improvement in wheat yield and associated traits. A re□ examination of previous results and the latest trends. Plant Breed. 114, 108–112. https://doi.org/10.1111/j.1439-0523.1995.tb00772.x

Calderini, D.F., Slafer, G.A., 1999. Has yield stability changed with genetic improvement of wheat yield? Euphytica 107, 51–59. https://doi.org/10.1023/A:1003579715714

Chairi, F., Vergara-Diaz, O., Vatter, T., Aparicio, N., Nieto-Taladriz, M.T., Kefauver, S.C., Bort, J., Serret, M.D., Araus, J.L., 2018. Post-green revolution genetic advance in durum wheat: The case of Spain. Field Crops Res. 228, 158–169. https://doi.org/10.1016/j.fcr.2018.09.003

de Oliveira Silva, A., Slafer, G.A., Fritz, A.K., Lollato, R.P., 2020. Physiological Basis of Genotypic Response to Management in Dryland Wheat. Front. Plant Sci. 10, 1644. https://doi.org/10.3389/fpls.2019.01644

771 Elía, M., Savin, R., Slafer, G.A., 2016. Fruiting efficiency in wheat: Physiological aspects and genetic
772 variation among modern cultivars. Field Crops Res. 191, 83–90.
773 https://doi.org/10.1016/j.fcr.2016.02.019

774 Ferrante, A., Savin, R., Slafer, G.A., 2013. Is floret primordia death triggered by floret development
775 in durum wheat? J. Exp. Bot. 64, 2859–2869. https://doi.org/10.1093/jxb/ert129

776 Ferrante, A., Savin, R., Slafer, G.A., 2012. Differences in yield physiology between modern, well
777 adapted durum wheat cultivars grown under contrasting conditions. Field Crops Res. 136, 52–
778 64. https://doi.org/10.1016/j.fcr.2012.07.015

779 Ferrante, A., Savin, R., Slafer, G.A., 2010. Floret development of durum wheat in response to
780 nitrogen availability. J. Exp. Bot. 61, 4351–4359. https://doi.org/10.1093/jxb/erq236

781 Fischer, R.A., 2011. Farrer review. Wheat physiology: a review of recent developments. Crop Pasture
782 Sci. 62, 95. https://doi.org/10.1071/CP10344

783 Fischer, R.A., 1985. Number of kernels in wheat crops and the influence of solar radiation and
784 temperature. J. Agric. Sci. 105, 447–461. https://doi.org/10.1017/S0021859600056495

785 Fischer, R.A., Stockman, Y., 1986. Increased Kernel Number in Norin 10-Derived Dwarf Wheat:
786 Evaluation of the Cause. Funct. Plant Biol. 13, 767. https://doi.org/10.1071/pp9860767

787 Flintham, J.E., Borner, A., Worland, A.J., Gale, M.D., 1997. Optimizing wheat grain yield: Effects of
788 Rht (gibberellin-insensitive) dwarfing genes. J. Agric. Sci. 128, 11–25.

789 Flohr, B.M., Hunt, J.R., Kirkegaard, J.A., Evans, J.R., Swan, A., Rheinheimer, B., 2018. Genetic gains
790 in NSW wheat cultivars from 1901 to 2014 as revealed from synchronous flowering during the
791 optimum period. Eur. J. Agron. 98, 1–13. https://doi.org/10.1016/j.eja.2018.03.009

792 Foulkes, M.J., Slafer, G.A., Davies, W.J., Berry, P.M., Sylvester-Bradley, R., Martre, P., Calderini,
793 D.F., Griffiths, S., Reynolds, M.P., 2011. Raising yield potential of wheat. III. Optimizing
794 partitioning to grain while maintaining lodging resistance. J. Exp. Bot. 62, 469–486.
795 https://doi.org/10.1093/jxb/erq300

796 Gent, M.P.N., 1994. Photosynthate reserves during grain filling in winter wheat. Agron. J. 86, 159–
797 167. https://doi.org/10.2134/agronj1994.00021962008600010029x

González, F.G., Miralles, D.J., Slafer, G.A., 2011. Wheat floret survival as related to pre-anthesis spike growth. J. Exp. Bot. 62, 4889–4901. https://doi.org/10.1093/jxb/err182

González, F.G., Slafer, G.A., Miralles, D.J., 2005a. Floret development and survival in wheat plants exposed to contrasting photoperiod and radiation environments during stem elongation. Funct. Plant Biol. 32, 189–197. https://doi.org/10.1071/FP04104

González, F.G., Slafer, G.A., Miralles, D.J., 2005b. Pre-anthesis development and number of fertile florets in wheat as affected by photoperiod sensitivity genes Ppd-D1 and Ppd-B1. Euphytica 146, 253–269. https://doi.org/10.1007/s10681-005-9021-3

Griffiths, S., Simmonds, J., Leverington, M., Wang, Y., Fish, L., Sayers, L., Alibert, L., Orford, S., Wingen, L., Herry, L., Faure, S., Laurie, D., Bilham, L., Snape, J.W., 2009. Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. Theor. Appl. Genet. 119, 383–395. https://doi.org/10.1007/s00122-009-1046-x

Gumedze, F.N., Dunne, T.T., 2011. Parameter estimation and inference in the linear mixed model. Linear Algebra Appl. 435, 1920–1944. https://doi.org/10.1016/j.laa.2011.04.015

Halloran, G.M., Pennell, A.L., 1982. Duration and Rate of Development Phases in Wheat in Two Environments. Ann. Bot. 49, 115–121. https://doi.org/10.1093/oxfordjournals.aob.a086220

Hirotugu Akaike, 1974. A new look at the statistical model identification. IEICE Trans. Autom. Control 19, 718–723. https://doi.org/10.1093/ietfec/e90-a.12.2762

Kirby, E.J.M., 1988. Analysis of leaf, stem and ear growth in wheat from terminal spikelet stage to anthesis. Field Crops Res. 18, 127–140. https://doi.org/10.1016/0378-4290(88)90004-4

Lázaro, L., Abbate, P.E., 2012. Cultivar effects on relationship between grain number and photothermal quotient or spike dry weight in wheat. J. Agric. Sci. 150, 442–459. https://doi.org/10.1017/S0021859611000736

Lo Valvo, P.J., Miralles, D.J., Serrago, R.A., 2018. Genetic progress in Argentine bread wheat varieties released between 1918 and 2011: Changes in physiological and numerical yield components. Field Crops Res. 221, 314–321. https://doi.org/10.1016/j.fcr.2017.08.014

Lynch, J.P., Doyle, D., McAuley, S., McHardy, F., Danneels, Q., Black, L.C., White, E.M., Spink, J., 2017. The impact of variation in grain number and individual grain weight on winter wheat

826 yield in the high yield potential environment of Ireland. Eur. J. Agron. 87, 40–49.
827 https://doi.org/10.1016/j.eja.2017.05.001

828 Maeoka, R.E., Sadras, V.O., Ciampitti, I.A., Diaz, D.R., Fritz, A.K., Lollato, R.P., 2020. Changes in
829 the Phenotype of Winter Wheat Varieties Released Between 1920 and 2016 in Response to In-
830 Furrow Fertilizer: Biomass Allocation, Yield, and Grain Protein Concentration. Front. Plant
831 Sci. 10, 1786. https://doi.org/10.3389/fpls.2019.01786

832 Miralles, D.J., Katz, S.D., Colloca, A., Slafer, G.A., 1998. Floret development in near isogenic wheat
833 lines differing in plant height. Field Crops Res. 59, 21–30. https://doi.org/10.1016/S0378-
834 4290(98)00103-8

835 Miralles, D.J., Richards, R.A., Slafer, G.A., 2000. Duration of the stem elongation period influences
836 the number of fertile florets in wheat and barley. Funct. Plant Biol. 27, 931.
837 https://doi.org/10.1071/PP00021

838 Miralles, D.J., Slafer, G.A., 2007. Sink limitations to yield in wheat: how could it be reduced? J.
839 Agric. Sci. 145, 139. https://doi.org/10.1017/S0021859607006752

840 Miralles, D.J., Slafer, G.A., 1995. Individual grain weight responses to genetic reduction in culm
841 length in wheat as affected by source-sink manipulations. Field Crops Res. 43, 55–66.
842 https://doi.org/10.1016/0378-4290(95)00041-N

843 Prieto, P., Ochagavía, H., Savin, R., Griffiths, S., Slafer, G.A., 2018. Dynamics of floret
844 initiation/death determining spike fertility in wheat as affected by Ppd genes under field
845 conditions. J. Exp. Bot. 69, 2633–2645. https://doi.org/10.1093/jxb/ery105

846 Reynolds, M.P., Foulkes, J., Furbank, R., Griffiths, S., King, J., Murchie, E., Parry, M., Slafer, G.A.,
847 2012. Achieving yield gains in wheat. Plant, Cell Environ. 35, 1799–1823.
848 https://doi.org/10.1111/j.1365-3040.2012.02588.x

849 Reynolds, M.P., Foulkes, M.J., Slafer, G.A., Berry, P., Parry, M.A.J., Snape, J.W., Angus, W.J., 2009.
850 Raising yield potential in wheat. J. Exp. Bot. 60, 1899–1918.
851 https://doi.org/10.1093/jxb/erp016

852 Reynolds, M.P., Pellegrineschi, A., Skovmand, B., 2005. Sink-limitation to yield and biomass: a
853 summary of some investigations in spring wheat. Ann. Appl. Biol. 146, 39–49.

854 https://doi.org/10.1111/j.1744-7348.2005.03100.x

855 Rivera-Amado, C., Trujillo-Negrellos, E., Molero, G., Reynolds, M.P., Sylvester-Bradley, R., Foulkes,
856 M.J., 2019. Optimizing dry-matter partitioning for increased spike growth, grain number and
857 harvest index in spring wheat. Field Crops Res. 240, 154–167.
858 https://doi.org/10.1016/j.fcr.2019.04.016

859 Sadras, V.O., Slafer, G.A., 2012. Environmental modulation of yield components in cereals:
860 Heritabilities reveal a hierarchy of phenotypic plasticities. Field Crops Res. 127, 215–224.
861 https://doi.org/10.1016/j.fcr.2011.11.014

862 Sanchez-Garcia, M., Royo, C., Aparicio, N., Martín-Sánchez, J.A., Álvaro, F., 2013. Genetic
863 improvement of bread wheat yield and associated traits in Spain during the 20th century. J.
864 Agric. Sci. 151, 105–118. https://doi.org/10.1017/S0021859612000330

865 Savin, R., Slafer, G.A., 1991. Shading effects on the yield of an Argentinian wheat cultivar. J. Agric.
866 Sci. 116, 1–7. https://doi.org/10.1017/S0021859600076085

867 Scarth, R., Kirby, E.J.M., Law, C.N., 1985. Effects of the Photoperiod Genes Ppd1 and Ppd2 on
868 Growth and Development of the Shoot Apex in Wheat. Ann. Bot. 55, 351–359.
869 https://doi.org/10.1093/oxfordjournals.aob.a086912

870 Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.
871 https://doi.org/10.1214/aos/1176348654

872 Serrago, R.A., Alzueta, I., Savin, R., Slafer, G.A., 2013. Understanding grain yield responses to
873 source-sink ratios during grain filling in wheat and barley under contrasting environments. Field
874 Crops Res. 150, 42–51. https://doi.org/10.1016/j.fcr.2013.05.016

875 Shaw, L.M., Turner, A.S., Laurie, D.A., 2012. The impact of photoperiod insensitive Ppd-1a
876 mutations on the photoperiod pathway across the three genomes of hexaploid wheat (*Triticum*
877 *aestivum*). Plant J. 71, 71–81. https://doi.org/10.1111/j.1365-313X.2012.04971.x

878 Shearman, V.J., Sylvester-Bradley, R.K., Scott, R.K., Foulkes, M.J., 2005. Physiological processes
879 associated with wheat yield progress in the UK. Crop Sci. 45, 175–185.
880 https://doi.org/10.2135/cropsci2005.0175

881 Siddique, K.H.M., Belford, R.K., Perry, M.W., Tennant, D., 1989a. Growth, development and light

882 interception of old and modern wheat cultivars in a Mediterranean-type environment. Aust. J.
883 Agric. Res. 40, 473–487. https://doi.org/10.1071/AR9890473

884 Siddique, K.H.M., Kirby, E.J.M., Perry, M.W., 1989b. Ear: Stem ratio in old and modern wheat
885 varieties; relationship with improvement in number of grains per ear and yield. Field Crops Res.
886 21, 59–78. https://doi.org/10.1016/0378-4290(89)90041-5

887 Slafer, G.A., 2003. Genetic basis of yield as viewed from a crop physiologist's perspective. Ann.
888 Appl. Biol. 142, 117–128. https://doi.org/10.1111/j.1744-7348.2003.tb00237.x

889 Slafer, G.A., Abeledo, L.G., Miralles, D.J., Gonzalez, F.G., Whitechurch, E.M., 2001. Photoperiod
890 sensitivity during stem elongation as an avenue to raise potential yield in wheat. Euphytica 119,
891 191–197. https://doi.org/10.1023/A:1017535632171

892 Slafer, G.A., Andrade, F.H., 1993. Physiological attributes related to the generation of grain yield in
893 bread wheat cultivars released at different eras. Field Crops Res. 31, 351–367.
894 https://doi.org/10.1016/0378-4290(93)90073-V

895 Slafer, G.A., Andrade, F.H., 1989. Genetic improvement in bread wheat (Triticum aestivum) yield in
896 Argentina. Field Crops Res. 21, 289–296. https://doi.org/10.1016/0378-4290(89)90010-5

897 Slafer, G.A., Andrade, F.H., Satorre, E.H., 1990. Genetic-improvement effects on pre-anthesis
898 physiological attributes related to wheat grain-yield. Field Crops Res. 23, 255–263.
899 https://doi.org/10.1016/0378-4290(90)90058-J

900 Slafer, G.A., Araus, J.L., Royo, C., García Del Moral, L.F., 2005. Promising eco-physiological traits
901 for genetic improvement of cereal yields in Mediterranean environments. Ann. Appl. Biol. 146,
902 61–70. https://doi.org/10.1111/j.1744-7348.2005.04048.x

903 Slafer, G.A., Elia, M., Savin, R., García, G.A., Terrile, I.I., Ferrante, A., Miralles, D.J., González,
904 F.G., 2015. Fruiting efficiency: An alternative trait to further rise wheat yield. Food Energy
905 Secur. 4, 92–109. https://doi.org/10.1002/fes3.59

906 Slafer, G.A., Rawson, H.M., 1994. Does temperature affect final numbers of primordia in wheat?
907 Field Crops Res. 39, 111–117. https://doi.org/10.1016/0378-4290(94)90013-2

908 Slafer, G.A., Savin, R., 1994. Source-sink relationships and grain mass at different positions within
909 the spike in wheat. Field Crops Res. 37, 39–49. https://doi.org/10.1016/0378-4290(94)90080-

Slafer, G.A., Savin, R., Sadras, V.O., 2014. Coarse and fine regulation of wheat yield components in response to genotype and environment. Field Crops Res. 157, 71–83. https://doi.org/10.1016/j.fcr.2013.12.004

Stockman, Y., Fischer, R.A., Brittain, E., 1983. Assimilate Supply and Floret Development Within the Spike of Wheat (*Triticum aestivum* L.). Funct. Plant Biol. 10, 585. https://doi.org/10.1071/pp9830585

Whitechurch, E.M., Slafer, G.A., Miralles, D.J., 2007. Variability in the duration of stem elongation in wheat and barley genotypes. J. Agron. Crop Sci. 193, 138–145. https://doi.org/10.1111/j.1439-037X.2007.00260.x

Whitechurch, E.M., Snape, J.W., 2003. Developmental responses to vernalization in wheat deletion lines for chromosomes 5A and 5D. Plant Breed. 122, 35–39. https://doi.org/10.1046/j.1439-0523.2003.00749.x

Zhang, H., Richards, R., Riffkin, P., Berger, J., Christy, B., O'Leary, G., Acuña, T.B., Merry, A., 2019. Wheat grain number and yield: The relative importance of physiological traits and source-sink balance in southern Australia. Eur. J. Agron. 110. https://doi.org/10.1016/j.eja.2019.125935