A Comprehensive Study of Spike Fruiting Efficiency in Wheat

Nicole Pretini, Ignacio I. Terrile, Luciana N. Gazaba, Guillermo M. Donaire, Sebastián Arisnabarreta, Leonardo S. Vanzetti, Fernanda G. González*

N. Pretini, CITNOBA (CONICET-UNNOBA), Monteagudo 2772, Pergamino (2700), Buenos Aires, Argentina. I.I. Terrile, EEA Pergamino, INTA, Ruta 32 km 4.5, Pergamino (2700), Buenos Aires, Argentina. L.N. Gazaba and S. Arisnabarreta, Bayer, Estación Experimental Fontezuela, Ruta 8 km 214, Pergamino (2718), Buenos Aires, Argentina. G.M. Donaire, EEA Marcos Juárez, INTA, Ruta 12 Km 3, Marcos Juárez (2580), Córdoba, Argentina. L.S.Vanzetti, EEA Marcos Juárez, INTA, Ruta 12 Km 3, Marcos Juárez (2580), Córdoba, and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290 (C1425FQB), CABA, Argentina. F.G. González, EEA Pergamino, INTA, Ruta 32 km 4.5, and CITNOBA (CONICET-UNNOBA), Monteagudo 2772, Pergamino (2700), Buenos Aires, Argentina.

Received:_____. *Corresponding author (gonzalez.f@inta.gob.ar)

Abbreviations: CHAFF, no-grain spike dry weight at maturity (g spike^{-1}); DH, double haploid; E1 to E5, testing environments, see Table 1; FEm, fruiting efficiency at maturity (grains g_{CHAFF}^{-1}); FE, fruiting efficiency (grains g_{SDWa}^{-1}); FF, fertile florets per spike (n° spike^{-1}); FFE, fertile floret efficiency (florets g_{SDWa}^{-1}); GN, grain number per spike (n° spike^{-1}); GST, grain set (n° grains floret^{-1}); P1, Baguette 19 x BioINTA 2002 (B19xB2002); P2, Baguette Premium 11 x BioINTA 2002 (B11xB2002); SDWa, spike dry weight at anthesis (g spike^{-1}).

ABSTRACT

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/csc2.20143.

The spike fruiting efficiency (FE, grains per unit spike dry weight at anthesis) is a promising trait to improve grain number in wheat. It is often estimated at maturity (FEm) as the grains per unit of no-grain spike dry weight (CHAFF). The fertile floret efficiency (FFE, fertile florets per unit spike dry weight at anthesis) and grain set (n° grains floret^{-1}, GST) were studied to better understand FE determination for the first time. Two double haploid populations (P1 and P2) designed by crossing modern contrasting cultivars for FE (Baguette 19 and Baguette Premium 11 - high FE x BioINTA2002 - low FE) were sown in five environments. The FE and FEm showed an unstable correlation (low or high) among genotypes within environments (due to variable SDWa – CHAFF association), resulting in a worse correlation between GN and FEm than between GN and FE. Then, use FEm as a surrogate to FE for improving GN may yield lower gains than those expected if FE were used. The H^2 of FFE was high, but the variability in fertile florets per spike (FF) among genotypes within environments was correlated with FFE only in the environments with high SDWa. Despite the closely association between FE and FFE, the former was not totally set at anthesis as GST greatly affected FE and GN. Selecting for higher FFE and GST, where GxE determines heavy spikes at anthesis, is an alternative to breed for improved grain number.

INTRODUCTION

The genetic improvement of yield potential is a sustainable alternative to increase wheat production and food security. The main driver of yield potential are the grains produced per unit area, as wheat is still a crop mostly sink-limited during grain filling (Borrás et al., 2004; Acreche and Slafer, 2006; González et al., 2014). Through breeding eras, increased yield potential resulted in more grains per unit area, tightly linked to grains produced per spike, with no consistent trend in spikes per unit land area (Waddington et al., 1986; Perry and D'Antouno, 1989; Siddique et al., 1989; Slafer and Andrade, 1989, 1993; Acreche et al., 2008; Del Pozo et al., 2014; Lovalvo et al., 2018).

Grains per spike are determined in a period close to anthesis (Fischer 1975, 1985). The growth of the spike during the ca. 20 days before anthesis determines the number of florets that are fertile (Fischer & Stockman, 1980; Stockman et al., 1983; Kirby, 1988; González et al., 2003; González et al., 2011a); while the number of fertile florets that set grains (grain set, GST) depends on growth conditions during the ca. 10 days post-anthesis (Fischer 1975, 1985). Then grain number can be described as the product of the spike dry weight at anthesis (SDWa) and the efficiency to set grains or fruiting efficiency (FE, grains per unit of spike dry weight at anthesis) (Fischer, 1983) (Box 1).

INSERT BOX 1

Given the high correlation between grain number and fruiting efficiency in cultivars released during the 90's (Abbate et al., 1998) and after the 00's (González et al., 2011b), FE was proposed as a trait to improve grain number. As the SDWa is complex to measure when the number of plots is large, or in early generations, the no-grain spike dry weight at maturity (or CHAFF) is used as a surrogate to it (e.g. Stapper and Fischer, 1990; González et al., 2011b; Martino et al., 2015; Mirabella et al., 2016; Alonso et al., 2018). Then, the fruiting efficiency is often calculated as the grain number produced per unit of CHAFF, named in the present work as fruiting efficiency at maturity (FEm), following Fischer and Rebetzke (2018). The few reports about the relationship between FE and FEm, working with a reduced number of cultivars, showed contrasting results. Some authors reported a high association (R^2 = 78%, Abbate et al., 2013), while others observed a low relation (R^2 = 32%) with a clear underestimation of the trait and with a change in the ranking of cultivars (Elía et al., 2016). Previous works showed a high effect of genotype and relatively low genotype x environment interaction on FE (Terrile et al., 2017, Rivera-Amado et al., 2019) and FEm (Mirabella et al., 2016; González et al., 2011b, Guo et al., 2016), except for heat stress conditions around anthesis (Terrile et al., 2017). Narrow-sense heritability for FEm has been informed to be

high (Alonso et al., 2018) whereas narrow-sense heritability for FE has not been reported yet. Thus, using a large number of cultivars grown in diverse environments, the present work tried to (i) determine how accurate is it to measure FEm to describe FE, and (ii) estimate narrow-sense heritability of FE.

To understand the physiology of fruiting efficiency and its genotypic variation, we also studied FE as the product of the number of fertile florets produced per unit of spike dry weight at anthesis (i.e. the efficiency to set fertile florets, naming it Fertile Floret Efficiency, FFE) and the grain set (GST) (Box 1) (Fischer, 2011). To the best of our knowledge there are no previous reports trying to understand the differences in FE among modern cultivars from this point of view. There are few reports about GST in modern germplasm, but previous evidences suggest it is high (i.e. >80% of fertile florets set grains, Siddique et al., 1989; González et al., 2003; Elía et al., 2016), except for Guo et al. (2016) who showed GST as low as 60%. The genotypic variation of FFE has never been reported before nor the interaction with environment or its narrow-sense heritability. We hypothesized that (i) differences among wheat cultivars in FE are stablished at anthesis (FFE) and, (ii) narrow-sense heritability of FFE is higher than that of FE because GST, which may be more affected by environment, is excluded in the former but taken into account in the latter.

Furthermore, there are no reports about the relationship between FFE and SDWa. When FE was studied, some authors reported a negative relation to SDWa (Dreccer et al., 2009; Ferrante et al., 2012; Lázaro and Abbate, 2012; Terrile et al., 2017; Lo Valvo et al., 2018), while others showed no relationship at all (García et al., 2014; Elía et al., 2016, Rivera-Amado et al., 2019). The improvement of grain number when selecting for higher FE may be reduced if there is a trade-off between FFE/FE and SDWa.

The aim of the present work was to carry out a comprehensive study of spike fruiting efficiency in wheat. The relationship between FE and FEm (and between SDWa and CHAFF)

was examined, and the genotypic variation and narrow-sense heritability of FE determined. The physiological determinants of FE, i.e. the efficiency to set fertile florets, or fertile floret efficiency (FFE), and grain set (GST), were studied in a large set of genotypes, and the genotypic variation and narrow-sense heritability of FFE evaluated for the first time. Finally, the possible trade-off between FFE/FE and SDWa were also assessed.

MATERIALS AND METHODS

Plant material

Two DH populations, specially designed to study fruiting efficiency in the breeding target environment, were used. The populations were derived from the cross between Baguette 19 x BioINTA 2002 (P1, n= 102) and Baguette Premium 11 x BioINTA 2002 (P2, n= 81). The parental lines were high yielding commercial cultivars released to the Argentine market in 2004 (B11) and 2006 (B19 and B2002), and are well adapted to the North of Buenos Aires and South of Cordoba States (i.e. the central region of the wheat producing area of Argentina). These cultivars were identified as contrasting for FEm in González et al. (2011b) and for FE in Terrile et al. (2017) (B19 and B11 with high FE and B2002 with low FE). The three cultivars have awn-spikes.

Experiments and general conditions

The populations were sown in five (P1) or four (P2) different environments (Table 1). Environments E1 to E4 were field experiments in the central area of wheat production in Argentina, with sowing dates within the optimum for the region (Table 1). The experiments were performed at the EEA Pergamino (33° 51'S, 60° 56'W) and EEA Marcos Juarez (32° 43'S, 62° 06'W) Research Stations of INTA (National Institute of Agricultural Technology and Husbandry, Argentina), under a silty clay loam soil (Typic Argiudol) with 2.5% MO (0-20 cm). Mean soil N at sowing was 16 (E1), 22 (E2) and 25 (E3) kg ha^{-1} for the topmost 60-cm soil layer while soil phosphorus levels were always ≥ 20 ppm in the topmost 20-cm soil

layer. Urea was applied to level up N to 200 kg ha^{-1}, split in two applications (when >50% lines reached beginning of tillering - Z 2.2-, and beginning of stem elongation -Z3.1- Zadoks et al., 1974). In E4, 100 kg N ha^{-1} was applied as UAN 32 during pre-sowing. Regarding P, in E2 and E3, 50 kg ha^{-1} of mono-ammonium phosphate was applied at sowing, while 80 kg ha^{-1} of the same fertilizer was applied in E4. In the four environments (E1 to E4) the experiments were sown after soybean. The E5 was a summer season under greenhouse, where plants were transplanted during February after artificial vernalization in cool chamber (20 days at 5°C, 8 h). This non-agronomic environment was planned to extremely increase the range of environments explored. In E1 to E4, temperature, rainfall and photosynthetic active radiation (PAR) were registered in a weather station (Campbell Scientific) close (<1000 m distance) to the experiments. In E5, a 4-chanell weather station (cavadevises.com) was installed inside the greenhouse.

Complementary sprinkler irrigation and pesticides were applied to prevent abiotic and biotic stresses. In the E1, E2 and E3, metsulfuron methyl (60%) at 7 g ha^{-1}+ dicamba (98%) at 100 ml ha^{-1} were applied prior to Z3.1 for weeds control, and pyraclostrobin (13.3%) + epoxiconazole (5%) were applied at 1 L ha^{-1} twice during growth cycle to prevent diseases (when >50% lines reached Z3.3 and Z3.9). For E4 during pre-sowing a combination of glyphosate (48%) at 2.5 L ha^{-1} + chlorsulfuron (62.5%) + metsulfuron methyl (12.5%) at 15 g ha^{-1} were applied to prevent weeds, and metominostrobin (15%) + tebuconazole (30%) at 0.5 L ha^{-1} was applied twice during growth cycle to prevent diseases (when >50% lines reached Z3.3 and Z3.9). In E1 to E3, chlorpyrifos (48%) at 0.30 L ha^{-1} was applied at heading, while lambdacialotrin (25%) at 0.05 L ha^{-1} was applied in E4. In E5 thiamethoxam (14.1%) + lamdacialotrin (10.6%) was applied twice during growth cycle at 0.075 L ha^{-1}.

In E1 to E4, lines (or genotypes) within each population were arranged in a randomized complete block design with two replications. Plots consisted of (i) two rows 1 m long and

0.21 m apart in E1 (190 plants m^{-2}), and (ii) five rows 2 m long and 0.21 m apart in E2 to E4 (E2: 330 plants m^{-2}, and E3-E4: 280 plants m^{-2}). In E5, five plants per pot (5 l capacity, filled with soil from the area + 1 g pot^{-1} Basacote plus 6M (NPKS + FeCuMnZnBMo) were transplanted and six pots per line were tested in a complete randomized design. Pots were re-arranged each 10 days within the greenhouse to avoid border effects.

INSERT Table 1

Measurements and analyzes

The number of fertile florets were estimated when each plot reached the anthesis stage (Z6.1, Zadoks et al., 1974). In E2 and E3, 0.5 m of a central row was sampled and spikes were separated from the rest of biomass. The spikes were arranged by length and the three median spikes were selected. In E1 five spikes from the most representative layer were cut, while in E5 the main stem spikes of three plants/pot and three pots were selected for fertile florets count. In E4 no measurements at anthesis were performed. For each selected spike, the number of fertile florets (when yellow anthers were visible or the floret score was >9.5 in Waddington and Cartwright, 1983) in one-half side of the spike ("a") and in the terminal spikelet ("t") were counted using a binocular microscope. The number of fertile florets per spike (FF) was estimated multiplying the value "a" by 2 and adding the value "t". The spikes were dried in an oven (70°C) to estimate the spike dry weight at anthesis (SDWa). The FFE was calculated as the ratio between FF and SDWa.

The spikes from another 0.5 m of a central row were cut from the rest of the plant in E2, E3 and E4 at maturity, and three spikes were selected following the procedure described for anthesis. In E1 and E5 individual spikes were selected also following the procedure previously described for anthesis. All the selected spikes (E1 to E5) were oven-dried (70°C) and weighed (total spike weight) before threshing by hand. The grain number of each spike was counted with an automatic counter to obtain the grain number per spike (GN). The FE

was calculated as the ratio between GN and SDWa. The CHAFF was calculated as the subtraction of the grain weight from the total spike weight before threshing. The FEm was calculated as the ratio between GN and CHAFF. The GST was estimated as the relation GN/FF.

Descriptive statistics was performed for each data set (trait x population x environment) before proceeding with the statically analyzes. The possible measurement error (within each environment) was inspected for each genotype through analyzes of coefficient of variation among replications (CVr). The Shapiro-Wilks modified by Mahibbur and Govindarajulu (1997) and Q-Q plots were used to test for normal distribution of populations within each environment. The relationship between these distributions and the value of the parental lines was also studied.

The narrow-sense heritability was calculated as

$h^2 = \sigma^2_G / (\sigma^2_G + \sigma^2_{GE}/E + \sigma^2_{RES}/ER)$

Where σ^2_G is the genetic (additive) variance, σ^2_{GXE} is the GxE interaction variance, E is the number of environments, R is the number of replications, and σ^2_{RES} is the error variance (Hallauer and Miranda, 1981). Before performing Pearson correlations to determine the relationship between traits, the coefficient of variation was inspected to assess the possible magnitude of spurious correlation when a variable is shared (e.g. GN vs FE, FF vs FFE, etc) (Brett, 2004). The Infostat/p software (Di Rienzo et al., 2016) was used for all the statistical analyzes.

RESULTS

Photosynthetically active radiation (PAR) and temperature increased during the crop cycle in E1 to E4 and decreased in E5 (Table 2). Temperature in E1 and E2 was cooler than in E3 and E4 during June, July and Aug., whereas it was similar during Sept., Oct. and Nov. (except for

E3 during Oct. that was 2-3ºC cooler, Table 2). The accumulated rainfall during growth cycle was high (> 700 mm) in E1 and E3, while it was close to the historical values of the region (350-400mm) in E2 and E4 (Table 2). The mean anthesis date took place during Oct. and was very concentrated within each population except for the summer environment E5, where mean anthesis date was on May 2nd and the standard deviation of the population was 23 days (Table 1). The high rainfall during Oct. 2012 (Table 2, E1) resulted in fusarium head-blight infection despite the fungicide applications. Then, only the data of anthesis for the E1 was considered for the present work.

INSERT Table 2

The CVr were estimated for testing measurements errors. In environments E1 to E4, the percentile 50 (P50) of the CVrs ranged from 2 to 15 %, while the P90 varied from 6 to 30 %, considering all traits and both populations. As expected, for the summer environment (E5) the P50 and P90 were greater (P50 ~ 5 - 30 %, P90 ~ 13 - 41%, for all traits). The CVs among genotype means within each environment were within the same magnitude for all the traits, and in no case the relation between the CV of the shared and non-shared variable was > 1.5, reducing the possible spurious correlation (Brett 2004) (see CVs in Tables S1, S2 and S3). All measured variables showed a normal distribution within each environment and transgressive segregation from the parental lines (Tables S1, S2 and S3).

Grain number, fruiting efficiency and spike dry weight at anthesis

The GN[1] was affected by the environment and the genotype in both populations, but the GxE interaction tended to be significant only in P2 (Table 3). Considering the agronomic environments (E1 to E4), the highest GN was produced in E3 (56.4±7.7 and 49.4±10.1 grains spike^{-1}, in P1 and P2, respectively), while the lowest was observed in E2, in both populations

[1] The population mean ± standard deviation of all traits are reported. For detailed data see Tables S1, S2 and S3

(36.4±4.7 and 31.3±4.9 grains spike^{-1}, in P1 and P2, respectively) (Table S1). As expected, the summer environment (E5) yielded even a lower GN value than the E2 (18.4±5.6 grains spike^{-1} in P1) (Table S1).

Similarly to GN, the FE was modified by the environment and the genotype in both populations, but the GxE interaction was significant only in P2 (Table 3). It is noteworthy that GxE interaction was not detected in P1, even when the summer environment with extremely contrasting growing conditions was considered. The environment explained ca. 15% of FE variation (in both populations) while the genotype showed a higher effect ranging from 36 to 26% in P1 and P2, respectively (Table 3). The GxE interaction was important in P2, explaining ca. 27% variation in FE (Table 3). Therefore, the narrow-sense heritability of FE was high (0.61) in P1, but it was not relevant in P2. The FE was larger in E3 than E2 in both populations (145±32 and 138±28 grains g_{SDWA}^{-1} in E3 vs 124±25 and 114±20 grains g_{SDWA}^{-1} in E2, for P1 and P2, respectively), while it was the lowest in the summer environment (105±39 grains g_{SDWA}^{-1} in P1) (Table S1).

In contrast to FE, the SDWa variation was mostly explained by the environment (ca. 80%), despite genotype and GxE interaction were statically significant (explaining 6-8% variation in SDWa). Similarly to GN and FE, the SDWa was greater in E3 than E2 (0.40±0.07 and 0.36±0.06 vs 0.30±0.05 and 0.28±0.04 g spike^{-1} in P1 and P2, respectively), while the E5 showed the lowest value (0.20±0.06 in P1). The E1, which was not tested for GN and FE (due to fusarium, see item 3), showed the highest values of SDWa (0.60±0.07 and 0.58±0.06 in P1 and P2, respectively) (Table S1).

The GN in both populations was highly correlated with FE when genotype variation within environment was considered (Fig.1a). Despite correlation was also positive across environments, great variation in GN was observed for the same FE due to the differences in SDWa mainly associated with the environments (Fig. 1b). The GN was highly associated

with SDWa across environments in both populations, but their relationship for genotype variation within environments was not significant (except for P2 in E3, Fig. 1b).

INSERT FIGURE 1

How accurate is it to measure fruiting efficiency at maturity to estimate fruiting efficiency?

The environment, the genotype and the GxE interaction modified the FEm in both populations (Table 3). The variation of FEm explained by the environment was large in P1 (ca. 60%), while the effects of genotype and GxE interaction were moderate (ca. 15% each, Table 3). In contrast, the three sources of variation (E, G and GxE) explained a similar proportion of FEm variability in P2 (ca. 25% each Table 3). Nevertheless, the narrow-sense heritability of FEm was high in both populations (0.63-0.56 in P1 and P2, respectively). The FEm in E2 was the greatest, contrasting with FE that showed the larger values in E3. The FEm ranged from 136±20 and 114±16 grains g_{CHAFF}^{-1} in E2 to 86±13 and 84±12 grains g_{CHAFF}^{-1} in E3, in P1 and P2, respectively (Table S2). The FEm was the lowest in the summer environment E5, 70±24 grains g_{CHAFF}^{-1} (Table S2). The correlation between GN and FEm was lower than that between GN and FE. In P1 the correlation between them ranged from 0.31 to 0.66 within environments, being only 0.22 when all data was considered together ($p<0.05$, data not shown). It was even worse in P2, not only for genotypes within environments (r's ~ 0.45-0.20, $p<0.07$) but also across environments where GN was negatively correlated with FEm ($r = - 0.25$, $p<0.0001$).

The low relation described previously may be consequence of the not so good association between FEm and FE. The correlation between them varied with environments in P1 but not in P2, ranging from low (r= 0.23) to high (r = 0.75) (Fig. 2a). When all data was considered together, i.e. across environments, covariance of both FE and FEm was moderate in P1 while, unexpectedly, it was negative in P2 (Fig. 2a). This variable covariance resulted in an unstable

estimation of FE by measuring FEm. Although an under-estimation of FE was mainly observed, some over-estimation was also detected depending on the combination genotype x environment (Fig. 2a).

The accuracy of measuring FEm to estimate FE relays on the relation between SDWa and CHAFF at maturity. Similarly to SDWa, the environment explained most CHAFF variation in both populations (60-73%), while the effect of genotype and GxE was reduced (<15%) (Table 3). The CHAFF was heavier in E3 than in E2 in both populations (0.66±0.10 and 0.60±0.13 vs 0.27±0.04 and 0.28±0.04 g spike^{-1} in P1 and P2, respectively) while it was similar to E2 in the summer environment E5 (0.28±0.08 g spike^{-1}). The covariance of SDWa and CHAFF was high in E5 (P1), moderate in E3 (both populations) and not significant in E2 (both populations) (Fig. 2b). When CHAFF weight was heavier than SDWa (most cases), it resulted in the under-estimation of FE and the contrary occurred when it was lighter (Fig. 2b).

INSERT FIGURE 2

Fertile florets, fertile floret efficiency and grain set

The FF was modified by environment, genotype and GxE interaction in both populations (Table 3). It varied from 26.7±4.4 to 55.6±8 florets spike^{-1} (E5 to E1) in P1 and from 51.3±8.6 to 55.6±8.6 florets spike^{-1} (E3 to E1) in P2 (Table S3). As expected, the FF covariation with SDWa was high across genotypes within environments as well as across environments in both populations (Fig. 3a). Differences in FF for the same SWDa are consequence of FFE.

The environment and the genotype modified the FFE in both populations, but the GxE interaction was statically significant only in P1 (Table 3). Comparing with the FE, it showed lower influence of GxE (8-10% vs 23-27% variation explained by interaction), but narrow-sense heritability was similarly high (0.52-0.58). Within the agronomic environments, the FFE varied from 94±12 to 137±20 (E1 to E3) and from 97±13 to 143±16 florets g_{SDWa}^{-1} (E1

to E3) in P1 and P2, respectively (Table S3). The summer environment showed the highest value of FFE (149±35 florets g_{SDWa}^{-1}) contrasting with the FE (which showed the lowest value under this environment). Despite variation observed for FFE in P1 in all the environments, covariation between FF and FFE was statistically significant only in E1 and E3, the two environments with larger SDWa (Fig. 3b). In P2, the correlation between FF and FFE was significant for genotypes within environments in all the cases, but once again, it improved as the spike was heavier (e.g. E1) (Fig. 3b). It is noteworthy that across environments the FF tended to reduce as FFE increased, suggesting a possible trade-off between FFE and SDWa.

INSER FIGURE 3

The FE would be the result of FFE and GST (see Box 1). The correlation between FE and FFE for genotypes within environments ranged from moderate to high (0.59 to 0.66 in P1, and 0.35 to 0.52 in P2, Fig. 4a). When variation across environments was considered, the covariance was reduced in both populations (Fig. 4a). As expected, in both populations GN improved as FF increased (Fig.4b), showing moderate correlation for genotype variation within environments (except for E5 where it was low) and high association across environments (Fig.4b).

Despite the significant relationship between GN and FF, the GST (i.e. the number of grains set per fertile floret) played an important role, more than expected, in FE determination. The environment, the genotype and the GxE interaction modified GST in both populations (explaining from 21 to 29% variation, Table 3), yielding low (6.5% in P1) or null (in P2) narrow-sense heritability. The GST varied from 0.99±0.14 and 0.91± 0.16 grains $floret^{-1}$ in E3 to 0.85±0.13 and 0.75±0.12 grains $floret^{-1}$ in E2, in P1 and P2, respectively. As expected the lowest GST was observed in the summer environment in P1, ca. 0.69±0.21 grains $floret^{-1}$ (Table S3). Anthesis date could modify the growing conditions explored by genotypes

affecting GST, but as it was very concentrated within each population, no relation was detected between anthesis date and GST ($p > 0.12$). In contrast, the GST was positively correlated with the post-anthesis growth of the spike (CHAFF-SDWa), showing high correlation for variation among genotypes within environments ($r > 0.52$ $p < 0.0001$ in E2 and E3 in both populations and $r = 0.22$ $p = 0.0300$ in E5 in P1). As GST improved the FE increased, being the correlation high for genotype variation within and across environments (Fig. 4c).

INSERT FIGURE 4

Possible trade-off between fertile floret efficiency / fruiting efficiency and spike dry weight at anthesis

The genotypes with heavier SDWa tended to have lower FFE and FE than those with lighter spikes, as the FFE and FE were negatively correlated with SDWa for genotype variation within environments in both populations (Fig. 5a, b). When all data was analyzed together within each population, the negative correlation was still present between FFE and SDWa but it was not significant for FE (Fig. 5a, b). This different response was consequence of the GST, which explained a great proportion of FE variation. Despite the negative correlation, variation in FFE and FE for similar SDWa within each environment could be observed (Fig. 5).

INSERT FIGURE 5

DISCUSSION

Grain number determination and accuracy of measuring fruiting efficiency at maturity to estimate fruiting efficiency

The FE explained most of the genotype's GN variation within environments with a reduced association with SDWa (Fig. 1). These results agree with previous reports showing also a better correlation between grain number and fruiting efficiency than between the former and

spike dry weight at anthesis when variation among genotypes within environments was considered (Abbate et al., 1998; Bustos et al., 2013; García et al., 2014; Elía et al., 2016; Terrile et al., 2017; Lovalvo et al., 2018; Rivera-Amado et al., 2019). Then, it seems that in modern elite genotypes the SDWa is no longer explaining differences in GN, as it used to be when analyzing the effects of Green Revolution in wheat (e.g. Siddique et al., 1989; Slafer and Andrade, 1989, 1993). Nevertheless, both traits FE and SDWa explained GN variation across environments (Fig. 1), which contrast with previous evidences where only differences in spike dry weight at anthesis were important when environmental differences were analyzed (García et al., 2014; Elía et al., 2016; Terrile et al., 2017). Agreeing with García et al. (2014), FE showed transgressive segregation, and for the first time we showed that narrow-sense heritability can be high or null, depending on genetic population.

The FEm showed transgressive segregation and high narrow-sense heritability as recently reported by Alonso et al. (2018). GN was also correlated with FEm, as it was reported previously (González et al., 2011b), but this correlation was worse than that of GN vs FE for genotype differences within and across environments. In contrast to Abbate et al. (2013), the correlation between FE and FEm depended on environment and population, ranging from low to high (or even negative) (Fig. 2). The use of CHAFF dry weight at maturity as a proxy for SDWa was not supported because covariance between them was extremely variable depending on environment. In most cases, the CHAFF was heavier than SDWa (Fischer and Stockman 1980; Stockman et al., 1983; Slafer et al., 2015; Elía et al., 2016) by factors still unknown (Fischer 2011; Slafer et al., 2015), yielding an underestimation of FE by FEm. This would contrast with a recent work where high correlation was observed between FE and FEm, but the latter was 35% higher than the former (Rivera-Amado et al., 2019). Then, the likely improvement in GN would be reduced when selecting for FEm instead of FE.

Fertile florets, fertile floret efficiency and grain set as physiological determinants of fruiting efficiency

As expected, the FF was highly correlated with SDWa for genotype differences within and across environments (Fig. 3) (Fischer and Stockman, 1980; Stockman et al., 1983; González et al., 2003). No previous reports studied the FFE, which showed to be positively associated with genotypic differences in FF only in the environments where the SDWa was heavy, suggesting that the latter is the primary limit determining the number of fertile florets. Then, increasing FFE to improve FF would be a breeding strategy when GxE interaction results in high SDWa.

There was a moderate to high association between FE and FFE for genotype variation within environments, but it was reduced when analyzes were performed across environments (Fig. 4). The FFE showed high narrow-sense heritability, but the FE was not totally determined at anthesis, as we hypothesized, because GST played a more relevant role than expected in determining FE. As GST has been suggested to be larger than ca. 80% in relatively modern cultivars (Siddique et al., 1989; González et al., 2003; Elía et al., 2016) it is generally not measured. In the present work, despite a close relation was observed between GN and FF, the correlation between FE and GST was also important (for genotype variation within and across environments). The GST depended on the environment, the genotype and their interaction in both populations, showing low or null heritability depending on population. The GST may depend on the ovary size of the distal florets (Guo et al., 2016). Then those genotypes with higher FE and FFE may have increased the FF without reducing floret's ovary size at expense of improved dry matter partitioning to the florets instead of the rest of the spike. The environmental conditions explored during the 10 days post-anthesis (Fischer 1975, 1985) would also modify the GST, but as the anthesis dates were very similar within lines in each population, we found no association between anthesis date and GST.

Nevertheless, we could stablished that the post-anthesis growth of the spike explained most of the variations in GST within and across environments. Bustos et al. (2013) reported that differences in FE among four cultivars of high potential grain number were in line with faster post-anthesis growth of the spike. Those results may indirectly support our findings, as higher GST, which was not reported by Bustos et al. (2013), may be behind the correlations they observed.

Is there a trade-off between fertile floret efficiency and fruiting efficiency with spike dry weight at anthesis?

The correlations between FE/FFE and SDWa were strongly negative for genotype variation within environments, while it was still negative across environments for FFE, but not significant for FE (Fig. 5). The lack of negative correlation between FE and SDWa across environments agrees with some previous reports (García et al., 2014; Elía et al., 2016, Rivera-Amado et al., 2019) but disagrees with many others (Dreccer et al., 2009; Ferrante et al., 2012; Lázaro and Abbate, 2012; Terrile et al., 2017; Lo Valvo et al., 2018). This is the first paper reporting the trade-off between FFE and SDWa, suggesting that selecting for higher FFE would result in lighter SDWa, which may reduce the gains in FF and GN. Nevertheless, this relation should be taken with care because one variable includes mathematically the other. To test for the magnitude of spurious correlation we used the equation proposed by Brett (2004), which resulted in moderate to high values (spurious correlation ranging from 0.45 to 0.78 in P1 and from 0.67 to 0.80 in P2). Considering that (i) great variation for genotypes within environments was observed for FFE for the same SDWa, (ii) the FFE showed high heritability and (iii) the SDWa had low genotype impact, it seems reasonable to think that genotypes with both high FFE and SDWa might be designed, if attention is paid to this relation in the breeding process.

CONCLUSIONS

The GN was highly correlated with FE when genotype variation within environments was considered, but FE's narrow-sense heritability was high only in one of the tested populations (B19xB2002). The relationship between CHAFF and SDWa depended on environment yielding an unstable correlation between FE and FEm. For the same reason, correlation between GN and FEm was worse than that between GN and FE. Then, the use FEm as a surrogate to FE for improving GN may return lower gains than those expected if FE were used. The FFE showed high narrow-sense heritability in both populations, but the FF variability among genotypes within environments was correlated with FFE only in the environments with heavy SDWa. Then, increasing FFE to improve FF would be a breeding strategy when GxE interaction results in heavy spikes at anthesis. Despite the high association between FE and FFE, the former was not totally set at anthesis as GST had a great impact on FE and GN. Selecting for larger FFE and GST, where GxE determines heavy spikes at anthesis, seems to be an alternative to breed for improved grain number. The detection and validation of QTL and molecular markers linked with these traits would allow their use in breeding programs.

ACKNOWLEDGEMENTS

We would like to thank Pablo Perez, Nadia Osuna and Luis Blanco for the field technical assistance, and Dr. Andreas Börner for helping to obtain the double haploid populations from the F1 seeds we provided. The present work was funded by the National Agency of Scientific and Technical Promotion of Argentina (PICT 2012-1198), the National Institute of Agricultural and Husbandry Technology (INTA, PNCYO 1127042) Argentina, the Monsanto Beachell-Bourlag Scholarship, the Northwest University of Buenos Aires Province (UNNOBA, SIB 2015, SIB 2017) and the EU FP7 Funding (ADAPATWHEAT 289842). NP

held a National Scientific and Technical Research Council (CONICET) scholarship, Argentina.

AUTHOR'S CONTRIBUTION

IT and FGG were in charge of experiments during 2012 and 2013; NP, LG and IIT during 2015, with help from LV, SA, GD, and FGG; and NP during 2016. NP and FGG analyzed data, while FGG designed and wrote the ms with revision from NP and LV.

CONFLICT OF INTEREST

Authors state that there is no conflict of interest.

REFERENCES

Abbate PE, Andrade FH, Lázaro L, Bariffi JH, Berardocco HG, Inza VH, Marturano F. 1998. Grain yield increase in recent Argentine wheat cultivars. Crop Sci. 38: 1203-1209.

Abbate PE, Pontaroli A, Lázaro L, Gutheim F. 2013. A method of screening for spike fertility in wheat. J. Agric. Sci. 151: 322-330.

Acreche MM, Slafer GA. 2006. Grain weight response to increases in number of grains in wheat in a Mediterranean area. Field Crops Res. 86: 52-59.

Acreche MM, Briceño-Félix G, Sánchez JAM, Slafer GA. 2008. Physiological bases of genetic gains in Mediterranean bread wheat yield in Spain. Eur. J. Agron. 28: 162-170.

Alonso MP, Panelo J, Mirabella NE, Pontaroli AC. 2018. Selection for high spike fertility index increases genetic progress in grain yield and stability in bread wheat. Euphytica 214: 112.

Amado-Rivera C, Trujillo-Negrellos E, Molero G, Reynolds MP, Sylvester-Bradley R, Foulkes MJ. 2019. Optimizing dry matter partitioning for increased spike growth, grain number and harvest index in spring wheat. Field Crops Res 240: 154-167.

Borrás L, Slafer GA, Otegui ME. 2004. Seed dry weight response to source–sink manipulations in wheat, maize and soybean: a quantitative reappraisal. Field Crops Res. 86: 131–146.

Brett MT. 2004. When is a correlation between non-independent variables "spurious"? OIKOS 105: 647-656.

Bustos-Korts DV, Hasan A, Reynolds M, Calderini D. 2013. Combining high grain number and weight through a DH-population to improve grain yield potential of wheat in high-yielding environments. Field Crops Res. 145:106-115.

Del Pozo A, Mathus I, Serret MD, Araus JL. 2014. Agronomic and physiological traits associated with breeding advances of wheat under high productive Mediterranean conditions. The case of Chile. Environ. Exp. Bot. 130: 180-189.

Di Rienzo JA, Casanoves F, Balzarini MG, Gonzalez L, Tablada M, Robledo CW. InfoStat versión 2016. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina.

Dreccer MF, van Herwaarden AF, Chapman SC. 2009. Grain number and grain weight in wheat lines contrasting for stem water-soluble carbohydrate concentration. Field Crops Res. 112: 43–54.

Elía M, Savin R, Slafer GA. 2016. Fruiting efficiency in wheat: physiological aspects and genetic variation among modern cultivars. Field Crop Res. 191: 83–90.

Ferrante A, Savin R, Slafer GA. 2012. Differences in yield physiology between modern, well adapted durum wheat cultivars grown under contrasting conditions. Field Crop Res. 136: 52–64.

Fischer RA. 1975. Yield potential in a dwarf spring wheat and the effect of shading. Crop Sci. 5: 607-614.

Fischer RA. 1983. Wheat. In: Smith WH, Banta, SJ (Eds.) Symposium on potential productivity of field crops under different environments. International Rice Research Institute 129-153.

Fischer RA. 1985. Number of kernels in wheat crops and the influence of solar radiation and temperature. J. Agric. Sci. 100: 447-461.

Fischer RA. 2011. Wheat physiology: a review of recent developments. Crop and Pasture Sci. 62: 95–114.

Fischer RA, Rebetzke GJ. 2018. Indirect selection for potential yield in early-generation, spaced plantings of wheat and other small-grain cereals: a review. Crop and Pasture Sci. 69: 439-459.

Fischer RA, Stockman YM. 1980. Kernel number per spike in wheat (Triticum aestivum L.) – responses to pre-anthesis shading. Aust. J. Plant Physiol. 7: 169–180.

García GA, Serrago RA, González FG, Slafer GA, Reynolds MP, Miralles DJ. 2014. Wheat grain number: Identification of favorable physiological traits in an elite doubled-haploid population. Field Crops Res. 168: 126-134.

González FG, Slafer GA, Miralles DJ. 2003. Grain and floret number in response to photoperiod during stem elongation in fully and slightly vernalized wheats. Field Crops Res. 81: 17-27.

González FG, Miralles DJ, Slafer GA. 2011a. Wheat floret survival as related to spike growth during pre-anthesis. J. Exp. Bot. 62: 4889-4901.

González FG, Terrile II, Falcón MO. 2011b. Spike fertility and duration of stem elongation as promising traits to improve potential grain umber (and yield): Variation in modern Argentinean wheats. Crop Sci. 51: 1693-1702.

González FG, Aldabe ML, Terrile I, Rondanini DP. 2014. Grain weight response to different postflowering source: sink ratios in modern high-yielding Argentinean wheats differing in spike fruiting efficiency. Crop Sci. 54: 297-309.

Guo Z, Slafer GA, Schnurbusch T. 2016. Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. J Exp Bot 67: 4221-4230.

Hallauer AR, Miranda FJB. 1981. Quantitative genetics in maize breeding. 2nd. Ed. Ames: Iowa State University Press, 468.

Kirby EJM. 1988. Analysis of leaf, stem and ear growth in wheat from terminal spikelet stage to anthesis. Field Crops Res. 18: 127-140.

Lázaro L, Abbate PE. 2012. Cultivar effects on relationship between grain number and photothermal quotient or spike dry weight in wheat. The J. Agric. Sci. 150(4): 442-459.

Lo Valvo PJ, Miralles DJ, Serrago RA. 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.

Mahibbur M, Govindarajulu Z. 1997. A modification of the test of Shapiro and Wilk for normality. J. Appl. Stat. 24(2): 219-236.

Martino DL, Abbate PE, Cendoya ME, Gutheim F, Mirabella NE, Pontaroli AC. 2015. Wheat spike fertility: inheritance and relationship with spike yield components in early generations. Plant Breeding 134(3): 264-270.

Mirabella NE, Abbate PE, Ramirez IA, Pontaroli AC. 2016. Genetic variation for wheat spike fertility in cultivars and early breeding materials. J. Agric. Sci. 154(1): 13-22.

Perry MW, D'Antuono MF. 1989. Yield improvement and associated characteristics of some Australian spring wheat cultivars introduced between 1860 and 1982. Aust. J. Agric. Res. 40: 457-472.

Siddique KHM, Kirby EJM, Perry MW. 1989. Ear stem ratio in old and modern wheat-varieties – relationship with improvement in number of grains per ear and yield. Field Crop Res. 21: 59-78.

Slafer GA, Andrade FH. 1989. Genetic improvement in bread wheat (*Triticum aestivum*) yield in Argentina. Field Crops Res. 21: 289-296.

Slafer GA, Andrade FH. 1993. Physiological attributes related to the generation of grain yield in bread wheat cultivars released at different eras. Field Crops Res. 31: 351-367.

Slafer GA, Elía M, Savin R, García G, Terrile I, Ferrante A, Miralles D J. 2015. Fruiting efficiency: an alternative trait to further raise wheat yield. Food and Energy Sec. 4(2): 92-109.

Stapper M, Fischer RA. 1990. Genotype, sowing date and plant spacing influence on high-yielding irrigated wheat in southern New South Wales. II. Growth, yield and nitrogen use. Aust. J. Agric. Res. 41: 1021-1041.

Stockman YM, Fischer RA, Brittain EG. 1983. Assimilate supply and floret development within the spike of wheat. Aust. J. Plant Physiol. 10: 585-594.

Terrile II, Miralles DJ, González FG. 2017. Fruiting efficiency in wheat (Triticum aestivum L): trait response to different growing conditions and its relation to spike dry weight at anthesis and grain weight at harvest. Field Crops Res. 201: 86–96.

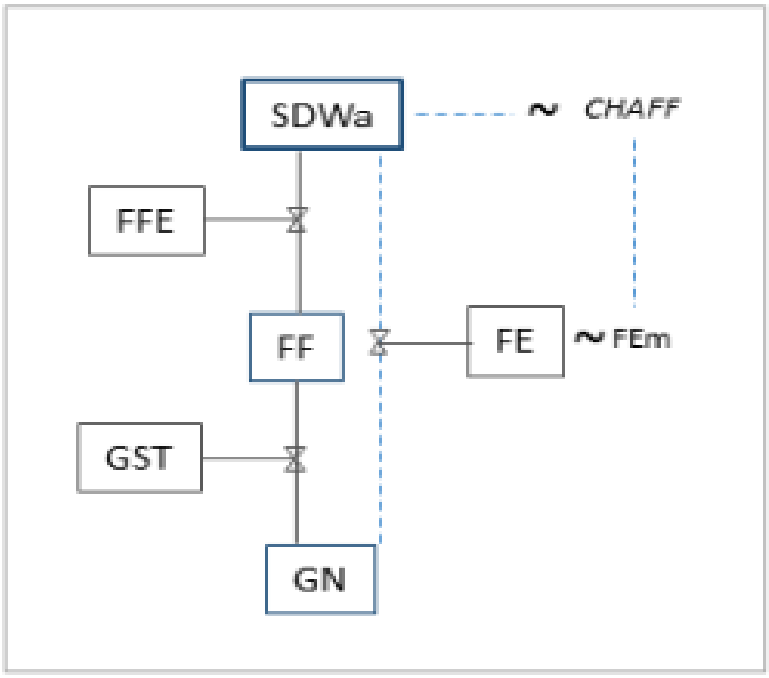
Waddington SR, Cartwright PM. 1983. A quantitative scale of spike initial and pistil development in barley and wheat. Ann. Bot. 51: 119-130.

Waddington SR, Ranson JK, Osmanzai M, Saunders SA. 1986. Improvement in the yield potential of bread wheat adapted to the northwest Mexico. Crop Sci. 26: 698-703.

Zadoks JC, Chang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. Weed Res. 14: 415-421.

BOX Captions

Box 1. Grain number per spike (GN) determination. Spike dry weight at anthesis (SDWa), fruiting efficiency (FE), fertile florets per spike (FF), fertile floret efficiency (FFE) and grain set (GST). The CHAFF (no-grain spike dry weight at maturity) is usually used as surrogate to SDWa and the fruiting efficiency at maturity (FEm) calculated instead of FE.



Box 1

Figure Captions

Figure 1. Relation between grain number spike^{-1} (GN) and (a) fruiting efficiency (FE) and (b) spike dry weight at anthesis (SDWa). Pearson correlations and p-value for genotypes within each environment are indicated. The underlined r shows correlation across environments. Both populations are presented.

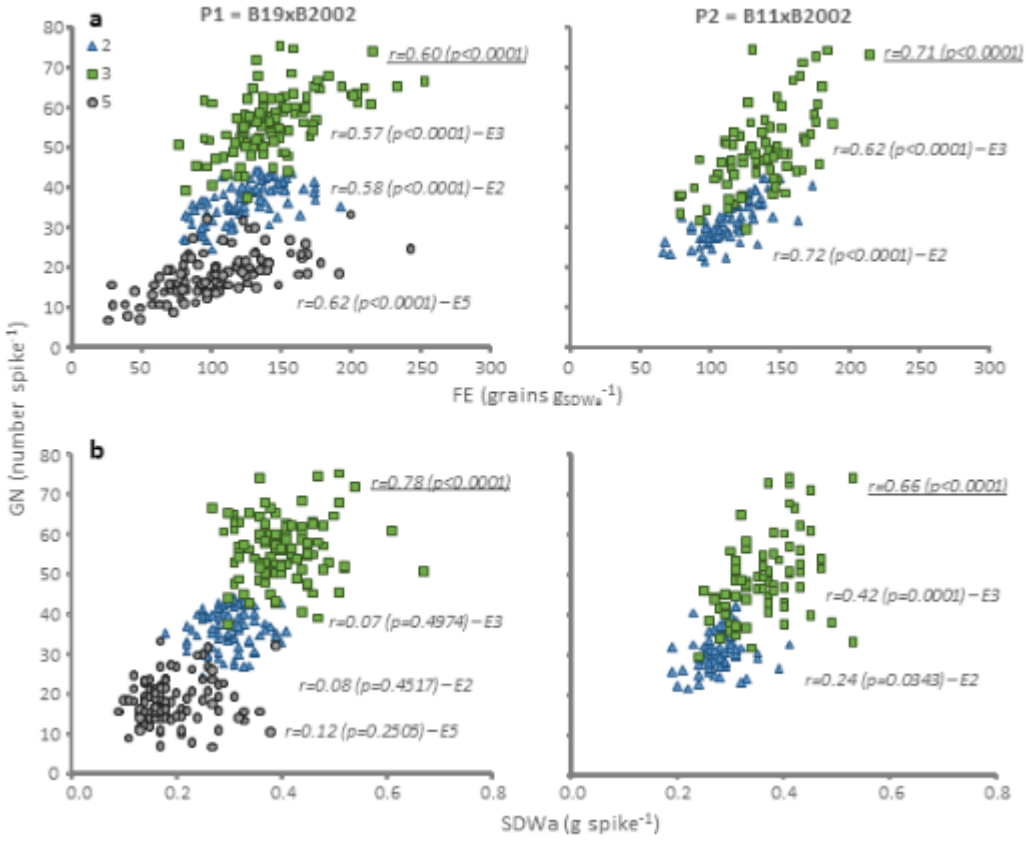


Figure 2. Relation between (a) fruiting efficiency (FE) and its surrogate at maturity, i.e. fruiting efficiency at maturity (FEm) and (b) spike dry weight at anthesis (SDWa) and chaff weight at harvest (CHAFF). Pearson correlations and p-value for genotypes within each environment are indicated. The underlined r shows correlation across environments. Both populations are presented.

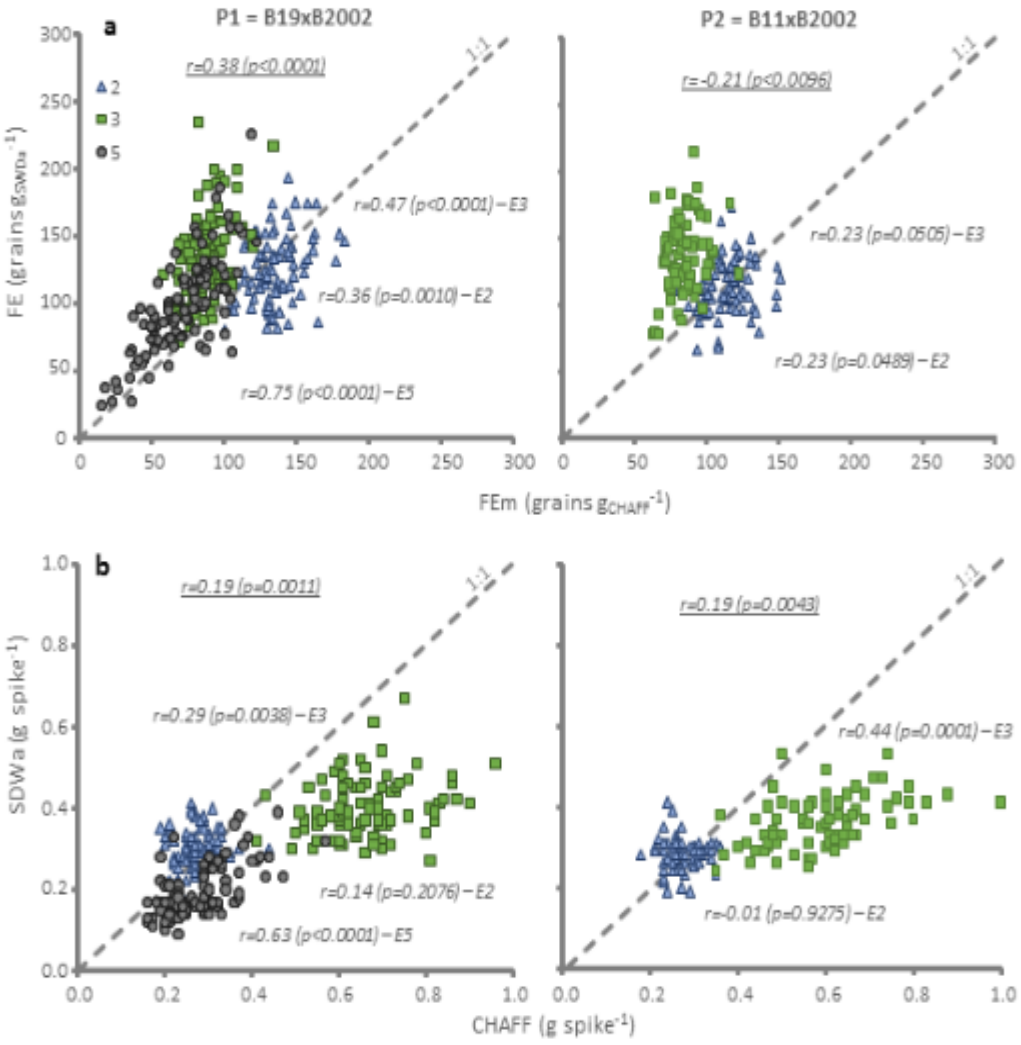


Figure 3. Relation between fertile florets spike^{-1} (FF) and (a) spike dry weight at anthesis (SDWa) and (b) fertile floret efficiency (FFE). Pearson correlations and p-value for genotypes within each environment are indicated. The underlined r shows correlation across environments. Both populations are presented.

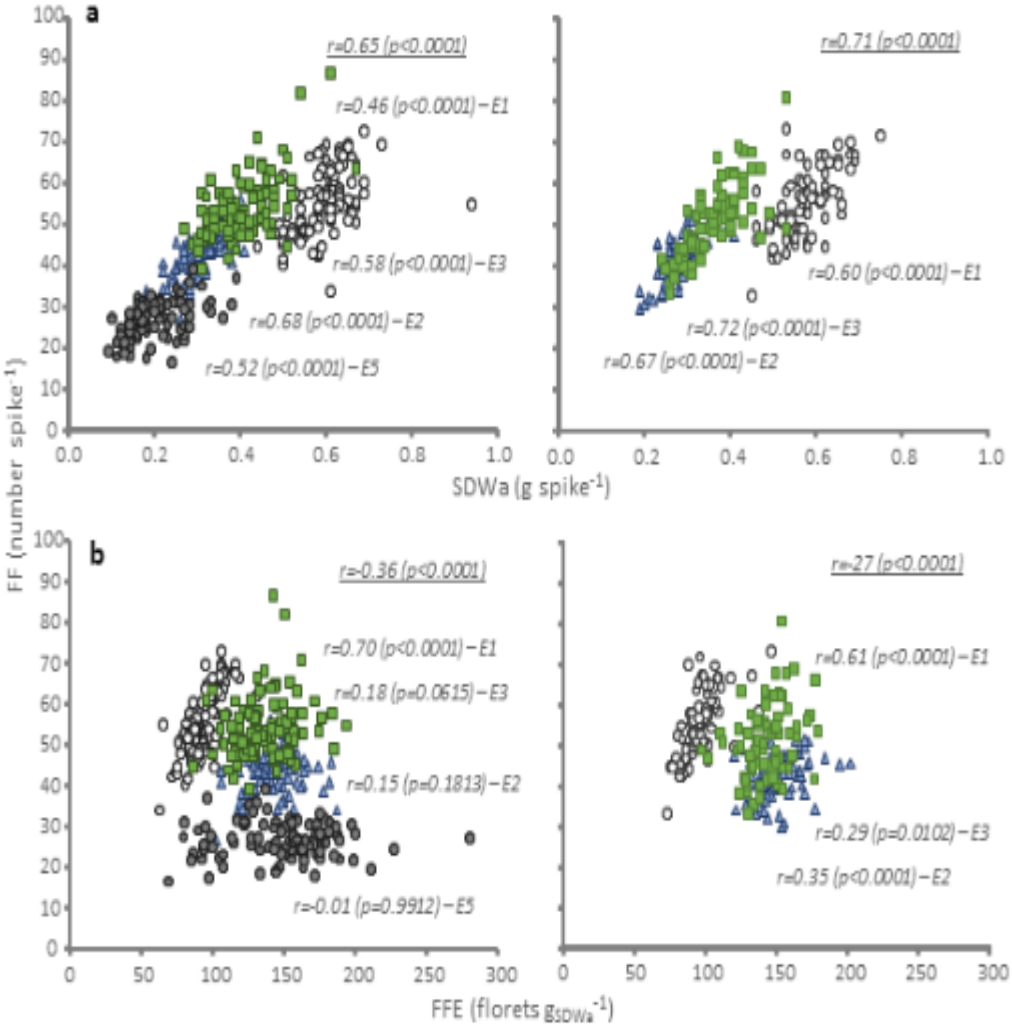


Figure 4. Relation between (a) fruiting efficiency (FE) and fertile floret efficiency (FFE), (b) grains spike^{-1} (GN) and fertile florets spike^{-1} (FF), and (c) fruiting efficiency (FE) and grain set (GST). Pearson correlations and p-value for genotypes within each environment are indicated. The underlined r shows correlation across environments. Both populations are presented.

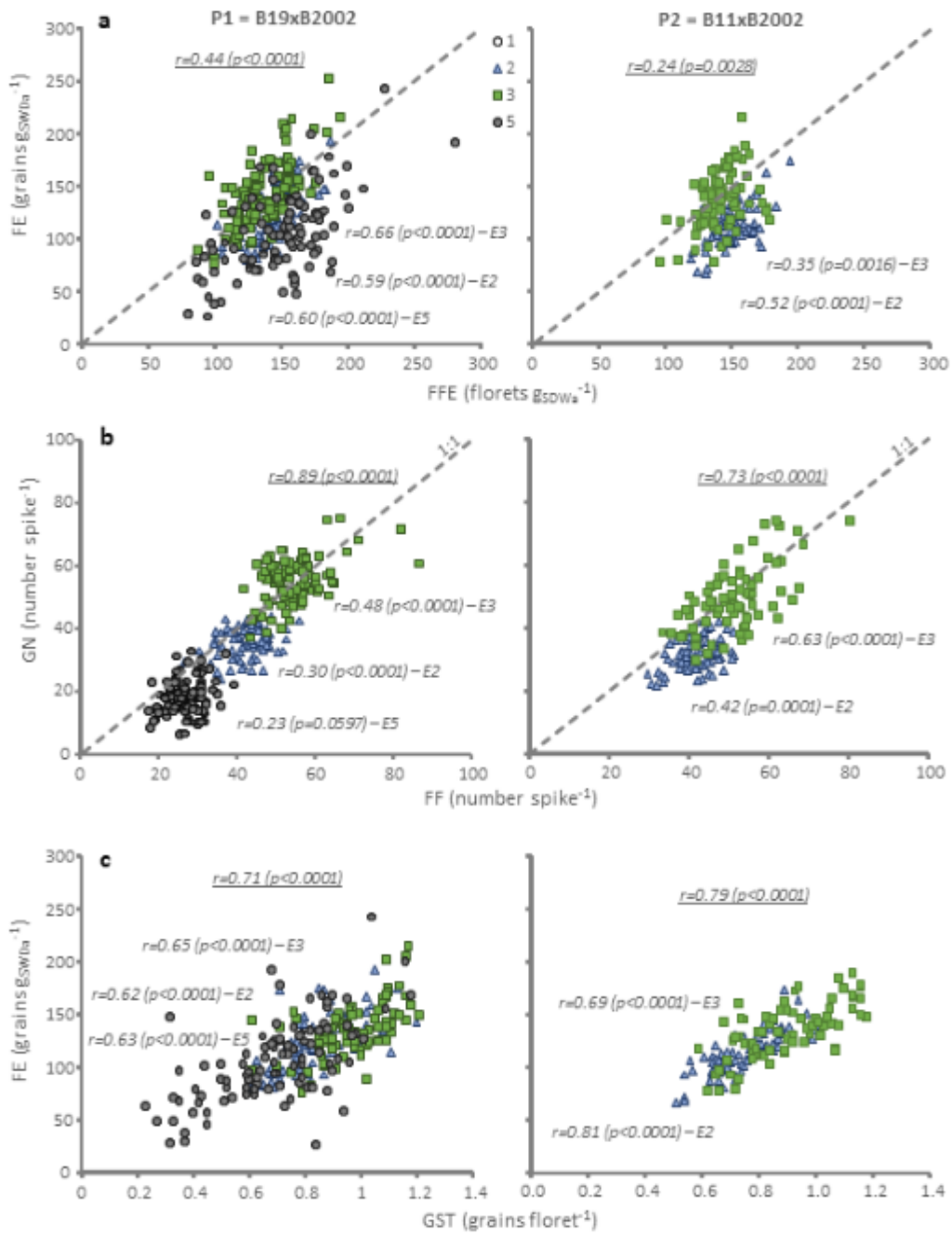


Figure 5. Relation between fertile floret efficiency (FFE) and fruiting efficiency (FE) vs spike dry weight at anthesis (SDWa). Pearson correlations and p-value for genotypes within each environment are indicated. The underlined r shows correlation across environments. Both populations are presented.

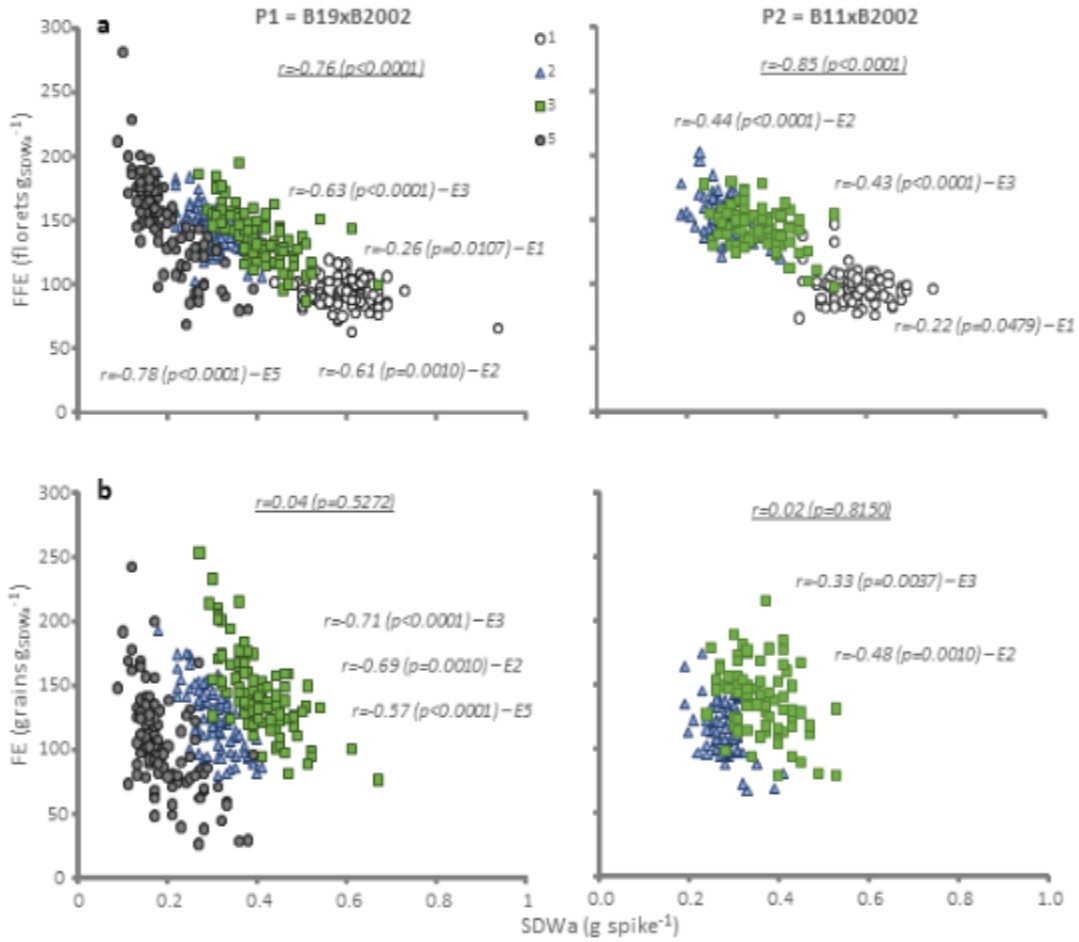


Table 1. Environments used for phenotyping P1= B19xB2002 and P2 = B11xB2002. Sowing and anthesis dates for each environment are presented.

Env.	Research Station	Year	DH population	Sowing date	Anthesis date[2]
E1	EEA Pergamino	2012	P1	08-Jun	23-Oct (±6 days)
			P2	06-Jun	24-Oct (±4 days)
E2	EEA Pergamino	2013	P1	19-Jun	26-Oct (±4 days)
			P2	19-Jun	25-Oct (±4 days)
E3	EEA Pergamino	2015	P1	11-Jun	20-Oct (±6 days)
			P2	11-Jun	22-Oct (±6 days)
E4	EEA Marcos Juárez	2015	P1	10-Jun	11-Oct (±5 days)
			P2	10-Jun	12-Oct (±6 days)
E5	EEA Pergamino	2016	P1	23-Feb[1]	02-May (±23 days)

[1]transplanting date after 20 days of vernalization in cool chamber, [2]mean ± standard deviation

Table 2. Growing conditions for each environment. Mean temperature, rain + irrigation and photosynthetic active radiation (PAR) during growth cycle.

		Month					
	Environment	June	July	August	September	October	November
Temperature	E1	9.8	5.8	10.9	14.3	16.8	20.6
(°C)	E2	9.9	9.8	9.9	13.2	17.6	20.3
	E3	12.1	10.8	13.6	13.6	14.6	19.2
	1967-2015 PE[1]	*10.2*	*9.7*	*11.3*	*13.4*	*16.6*	*19.5*
	E4	13.5	11.9	14.3	14.9	20.8	20.9
	1967-2015 MJ[2]	*10.8*	*10.4*	*12.1*	*14.6*	*18.0*	*20.9*
Rain +	E1	4	9	230	79	301	168

Irrigation	E2[3]	7	28	28	57	62	172
(mm)	E3	54	67	262	62	119	195
	1967-2015 PE	*37*	*36*	*42*	*55*	*106*	*104*
	E4[4]	6	45	94	23	48	193
	1967-2015 MJ	*20*	*23*	*20*	*46*	*95*	*109*
PAR	E1	121	149	140	183	223	278
(Mj m^{-2})	E2	112	118	187	172	255	281
	E3	118	136	156	191	208	267
	E4	117	120	140	200	247	278

		February	March	April	May	June
Temperature (°C)	E5	25.2	22.9	20.5	17.2	10.9
PAR (Mj m^{-2})	E5	313	251	128	114	105

[1] PE: EEA Pergamino historical data, [2]MJ: EEA Marcos Juárez historical data, [3] Irrigation during August was 26 mm and during September 25 mm, [4] No irrigation was performed despite low rain (e.g. September) because table water was <1m depth during growth cycle.

Table 3. Percent of total variation (considering sum of squares) explained by environment (E), genotype (G) and GxE interaction for: grain number (GN), fruiting efficiency (FE), spike dry weight at anthesis (SDWa), fruiting efficiency at maturity (FEm), spike chaff at harvest (CHAFF), fertile florets at anthesis (FF), fertile floret efficiency (FFE) and grain set (GST), in P1 = B19xB2002 and P2 = B11x B2002.

	Env.[1]	Trait	E	G	GXE
P1	1,2,3,5	GN	76.0	7.0	8.4
			<0.0001	<0.0001	0.1274
	2,3,5	FE	14.7	35.8	23.4

			<0.0001	<0.0001	0.2354
	1,2,3,5	SDWa	80.0	5.6	8.0
			<0.0001	<0.0001	<0.0001
	2,3,4,5	FEm	59.3	15.7	13.8
			<0.0001	<0.0001	0.0004
	2,3,4,5	CHAFF	73.6	6.2	10.0
			<0.0001	<0.0001	0.0868
	1,2,3,5	FF	71.5	9.5	10.1
			<0.0001	<0.0001	0.0032
	1,2,3,5	FFE	79.9	5.6	8.0
			<0.0001	<0.0001	<0.0001
	2,3,5	GST	27.7	21.2	29.6
			<0.0001	0.0007	0.0015
P2	2,3,4	GN	41.3	22.5	16.5
			<0.0001	<0.0001	0.0844
	2,3	FE	15.7	25.8	27.3
			<0.0001	0.0765	0.0194
	1,2,3	SDWa	77.0	6.5	7.9
			<0.0001	<0.0001	0.0216
	2,3,4	FEm	27.0	28.3	22.0
			<0.0001	<0.0001	0.044
	2,3,4	CHAFF	61.0	13.7	14.3
			<0.0001	<0.0001	<0.0001
	1,2,3	FF	31.3	28.9	19.8
			<0.0001	<0.0001	0.0066
	1,2,3	FFE	63.1	13.3	10.2

		<0.0001	<0.0001	0.2096
2,3	GST	21.1	26.4	27.9
		<0.0001	0.0444	0.0042

[1]Environments where traits were phenotyped.