Contrasted phenotypic plasticities of life-history traits in interacting plants: case of pure versus mixed stands of wheat

Meije Gawinowski[1,2,3], Jérôme Enjalbert[1], Paul-Henry Cournède[2], Timothée Flutre[1]

[1] Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, 91190 Gif-sur-Yvette, France

[2] Université Paris-Saclay, CentraleSupélec, MICS, 91190 Gif-sur-Yvette, France

[3] Université Paris Cité, 75006 Paris, France

Correspondence:
Meije Gawinowski, Timothée Flutre
meije.gawinowski@inrae.fr, timothee.flutre@inrae.fr

Abstract

- **Background and Aims**
  Cultivar mixture is an agronomic practice of diversification increasingly used in the framework of the agroecological transition. However, even though the yield of such mixtures is higher on average than the mean yield of the monocultures, the variance of mixture yield is large. This variability is likely due to the co-occurence of multiple ecophysiological processes shaping plant-plant interactions, yet it remains poorly understood, notably in crops. With winter wheat (Triticum aestivum L.) as a case study, we designed a field experiment to explore phenotypic plasticity at both genotype and plant levels along a gradient of neighborhood heterogeneity.

- **Methods**
  Eight wheat commercial cultivars were grown in pure and mixed stands in field plots for two seasons. Two quaternary mixtures were assembled with cultivars contrasted either for height or earliness. Thanks to a precision sowing in mixtures, genotypes were tracked at plant scale from sowing to harvest, and individual plants were phenotyped for above-ground traits throughout growth. Phenotypic plasticity between pure and mixed stands was then analyzed at both within- and between-genotype scales, according to a new conceptual framework distinguishing mean and variability differences.

- **Key results**
  Some genotypes dominated others in mixed stands, i.e., they produced a significantly higher mean yield, whereas all these genotypes yielded similarly in pure stands. These between-genotype dominance relationships remained stable over the two seasons despite strong contrasts in temperature and light sums. We showed that these dominance relationships in mixed stands were caused by contrasted phenotypic plasticity of yield components and biomass allocation in pure versus mixed stands. Tillering dy-

namics, determined by light competition between individual plants, was a main causal factor explaining between-genotype plasticity in both pure and mixed stands.

- **Conclusions**
  Our innovative experimental design enabled us to measure phenotypic plasticity at both within- and between- genotype levels. Plasticity in tillering dynamics and yield components allowed to decipher the genotype strategies in mixtures.

Keywords: plant-plant interactions, phenotypic plasticity, canalization, cultivar mixtures, bread wheat

Introduction

Agroecological transition, intra-plot diversification and over-yielding in varietal mixtures

Our current farming system mainly relies on industrial agriculture that provides high levels of productivity. It is notably dependent on cultivars selected for their high-yield potential under optimal conditions obtained thanks to a massive usage of chemical inputs (Therond et al., 2017). Besides the large increase in crop production (Evenson and Gollin, 2003), this system is associated with negative externalities (Jenkins, 2003; Bourguet and Guillemaud, 2016; Thompson et al., 2019). In contrast, agroecology promotes low-input practices (Altieri, 1989), with diversification at various spatiotemporal scales as a main lever (Gliessman, 2014). For annual crops at the scale of a plot, it consists in sowing intraspecific or interspecific mixtures (Barot et al., 2017; Justes et al., 2021) with the goal of taking advantage of niche complementarity and facilitation (Brooker et al., 2015). Both practices see a renewed interest from farmers and academics (e.g., the ReMIX project https://www.remix-intercrops.eu/). Yet, a quantitative and dynamic understanding of such mixed stands is still lacking (Gaudio et al., 2019).

Intraspecific mixtures, also called varietal mixtures, even if relying on a lower functional diversity, have already demonstrated various advantages. For our case study of winter wheat, on average varietal mixtures provide a higher yield than the mean of their yield in pure stands (relative yield total RYT > 1 meaning over-yielding), $+2\%$ on average (Borg et al., 2018). This is especially observed under high disease pressure ($+5\%$ on average), a well-known "mixture" effect due to their limitation of airborne fungal epidemics (Wolfe, 1985; Mundt, 2002; de Vallavieille-Pope, 2004). Varietal mixtures also display a better interannual stability than their components in pure stands (Reiss and Drinkwater, 2018). However, even under optimal conditions (non-limiting nutrients thanks to fertilization, and neither disease nor weed pressure with effective pesticide

treatments), where the only remaining limiting factor is light, there still is a large variability in RYT between varietal mixtures ($\pm 40\%$ in the meta-analysis of Borg et al. (2018), and even in the same environment ($\pm 20\%$ on data from (Forst, 2018)).

Differences in relative yields between mixtures and importance of plasticity

Such a variability in RYT between mixtures can be explained by mobilizing functional ecology and evolutionary biology. Neighboring plants of different genotypes can differ, however slightly, in their resource requirements and usages. Such a functional diversity can favor complementarity in resource foraging and facilitation, leading to over-yielding in the mixture (RYT > 1). For disputed resources, competition comes into play and gives rise to the so-called "selection" effect whereby a genotype with particular traits will dominate the others, usually the one with the highest biomass production in monoculture (Loreau, 2001). Such a case can also lead to over-yielding, but not necessarily. Indeed, dominating genotypes that over-invest in their competitive ability can trigger a tragedy of the commons (Anten and Vermeulen, 2016), leading to under-yielding with RYT below 1 (Foucteau et al., 2000; Weiner et al., 2017).

Behind all the processes involved in plant-plant interaction lies the capacity of plants to adjust their strategy, i.e., their phenotypes, depending on their abiotic conditions and biotic neighborhood (Schmitt, 1993). This refers to their phenotypic plasticity, understood here as the tendency for phenotypic variation to be influenced by the environment (Hallgrimsson et al., 2019). The extent of phenotypic plasticity in pure versus mixed stands and the magnitude of plasticity differences between genotypes remain largely unexplored. Studying them also raises specific questions in terms of experimental design.

Scales and metrics of analysis for plant-plant interactions

As plant-plant interactions occur in both pure and mixed stands, as formalized in model 3 of (Forst et al., 2019), neighborhood heterogeneity can be seen as a main environmental variable, lower in pure stands and higher in mixed stands. The reaction norm of interest here for a given genotype will hence be its trait value along this gradient. The case of pure vs mixed stands cultivated under optimal conditions are usually presented as corresponding to passive plasticity, in the sense of "stemming from direct environmental influences on chemical, physiological and developmental processes, and generally not considered anticipatory", best viewed as a "consequence of the environment, such as stunted growth owing to low resource levels" (Sultan, 2003; Forsman, 2015). However, active plasticity (in the sense of "anticipatory, and often highly integrated, phenotypic changes in response to some environmental cue or signal, reflecting modifications of developmental pathways and regulatory genes" Forsman (2015)), can occur. Among the possible mechanisms, neighbor-modulated disease susceptibility was recently demonstrated (Pélissier et al., 2021), and kin selection may also be involved even though no evidence was found so far (Fréville et al., 2019).

Some articles on phenotypic plasticity and genotype-environment interactions focus on the changes in mean trait value between genotypes from one environment to another (Guntrip and Sibly, 1998; Sultan, 2003). Beyond the mean, others rightly highlighted the importance of studying differences in phenotypic variability between environments (Fordyce, 2006; Flatt, 2005; Hallgrimsson et al., 2019). Moreover, studying clonal species (or strongly-selfing species such as wheat) allows to estimate the variability per genotype in a given environment, hence allowing us to disentangle it from the variability between genotypes in the same environment as well as from the reaction norm of each genotype. Of general interest beyond our case study, we hence emphasize here the need to consider two scales of study, within-genotype and between-genotype differences, as well as two metrics of trait values, mean trend and variability (Figures 1 and 2).

At the within-genotype scale (Figure 1), sources of variability include micro-environmental variation as well as developmental plasticity. The cases of decrease (respectively, increase) in phenotypic variability between environments correspond to environmental canalization (resp., decanalization).

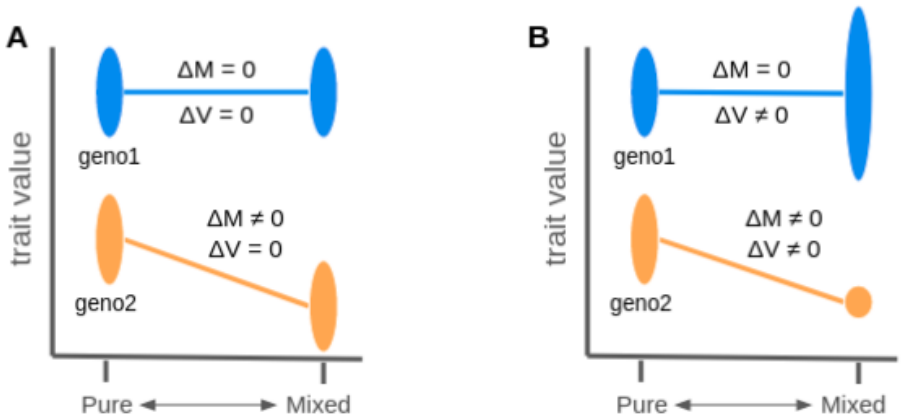


Figure 1 – Within-genotype phenotypic plasticity of a given trait for various genotypes depending on neighborhood heterogeneity, here exemplified by pure versus mixed stands. ΔM (respectively, ΔV) is the mean (resp., the variability) in the mixed stand minus the one in the pure stand. (A) geno1: difference neither in the mean nor in the variance, i.e., absence of phenotypic plasticity; geno2: presence of phenotypic plasticity for the mean only. (B) geno1: presence of phenotypic plasticity for the variability only; geno2: presence of phenotypic plasticity for both the mean and the variability.

At the between-genotype scale (Figure 2), the cases of decrease (respectively, increase) in phenotypic variance between both stands correspond to phenotypic convergence (e.g., Dahlin et al. (2020); resp., divergence), indicative of genetic canalization in mixed stands (resp., in pure stands).

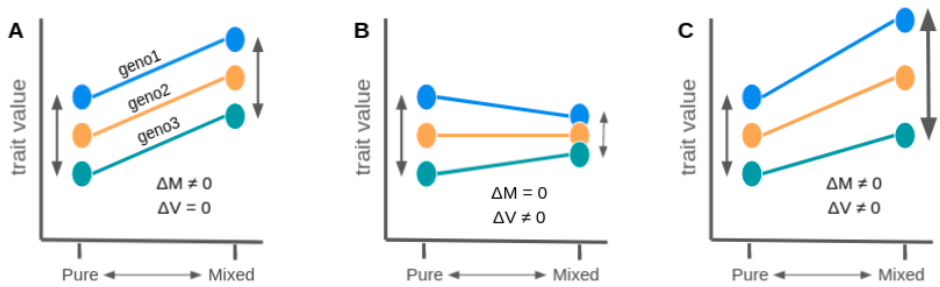


Figure 2 – Between-genotype phenotypic plasticity of a given trait for several genotypes depending on neighborhood heterogeneity, here exemplified by pure versus mixed stands. ΔM (respectively, ΔV) is the mean (resp., the variability) in the mixed stand minus the one in the pure stand. (A) All three genotypes differ from each other in mean only. (B) They differ in variability only. (C) They differ in both mean and variability.

Phenotypic integration

It is also imperative to go beyond the analysis of reaction norms for single traits separately given the importance of phenotypic integration that can lead to compensation between traits (Pigliucci and Preston, 2003). Crops such as cereals are especially well-studied in this regard. Grain yield is decomposed into several components that display contrasted phenotypic plasticities (Sadras and Slafer, 2012). For instance, tillering which conditions the final number of spikes is highly plastic with respect to plant density (Darwinkel, 1978), notably as a result of light quality (Xie et al., 2016). However, irrespective of differences in means between pure and mixed stands, little is known not only about differences in yield variability, but also about the contribution of each yield component to such differences.

Furthermore, a plant strategy involves resource allocation between vegetative and reproductive structures (Reekie and Bazzaz, 2011). Differences in reproductive allocation per genotype in different environments can be due to the allometry between total and reproductive biomass, or to the plasticity of this allometry (Weiner, 2004). Either way, given the physiological impacts of the competition for light, such trade-offs are likely to differ between pure and mixed

stands, but in contrasted ways depending on the genotype. As a result, differences in reproductive allocation for a genotype between pure and mixed stands would have an impact on its performance in the mixed stand, both in absolute terms compared to its performance in pure stands and in relative terms compared to the other genotypes in the mixed stand.

Hypotheses, rationale and design

In the framework of a hypothesis-driven research, a naive null hypothesis could be the absence of difference in both mean trend and variability for any trait between pure and mixed stands. However, competition between plants is pervasive (Grace and Tilman, 1990), and can change the relative proportions of genotypes between sowing and harvest, whether in terms of grain weight (yield, i.e., "agronomic fitness") or grain number ("evolutionary fitness"). In this context, it seems relevant to put forward distinct hypotheses for dominating versus dominated genotypes.

For clonal or strongly-selfing species, the plants of a genotype dominating the others in a mixed stand will have (much) less difficulties in acquiring resources, e.g., light. At the within-genotype scale, we expect for a dominating genotype (i) an increase in mean for life-history traits compared to what happens in a pure stand, and also (ii) a decrease in variability given that most (all) of its plants have reached their potential. On the opposite, for the plants of a dominated genotype in a mixed stand, we expect a decrease in mean for life-history traits at the within-genotype scale, reflecting the negative impact of competition. In terms of variability, expectations are less straightforward, and one can imagine scenarios where variability increases or decreases depending on the strength of competition incurred by the dominated plants.

At the between-genotype scale, we expect a higher variability in mixed stands compared to pure ones. Phenotypic plasticity would trigger a more important di-

vergence between genotypes in mixed than pure stands due to a more heterogeneous competition and possibly leading to niche specialization. Overall, a lower productivity of the mixed stand compared to the mean of the pure stands would be interpreted as a tragedy of the commons. But other scenarios also co-exist, such as an over-yielding of the mixed stand driven by an increase in frequency of the most productive genotype in pure stands.

The rationale of our investigation hence was to document the extent of phenotypic plasticity in pure versus mixed stands, assess its impact on the differential productivity between stands, and evaluate its genetic and environmental origins. Because we studied a strongly-selfing species, winter wheat, we could decompose plasticity into two scales, within- and between-genotype, and assess their respective magnitudes both in terms of mean trend and variability using robust statistical procedures based on a non-parametric approach. This also allowed us to test if the within-genotype variability in pure stands was associated with a difference in mean performance between pure and mixed stands (i.e., a reaction norm with a strong slope). Furthermore, stands are made of many individual plants, each adapting its strategy to its neighborhood. To reach this level of details while maintaining agronomic relevance and exploring a diversity of genotypes and assemblies, we proposed a new experimental design and applied it in a field trial with eight genotypes assembled into two four-way mixtures, which main innovation consisted in tracking the genotype of each individual plant in the mixed stands over their whole life cycle. The main phenological stages of each genotype were monitored. Moreover, for all individual plants, we not only assessed, at maturity, their life-history traits (yield components), but we also monitored, all along their growth, their height and tiller number, both traits known to be involved in light competition, to get a sense of the temporal dynamics along which plants interact. Moreover, to extend the environmental context in which such plant-plant interactions occur, we repeated our experimental study during two contrasted growing seasons, and experimented with

two plant densities.

Materials and Methods

Plant material

Among a panel of 210 European cultivars (Touzy et al., 2019), eight were chosen with contrasted heading date and height at maturity (Table S1): Accroc, Aubusson, Bagou, Belepi, Bergamo, Boregar, Expert and Kalahari. Two four-way mixtures were assembled, knowing that French farmers mix on average between three to seven cultivars. The first, "Mixture 1", was composed of Accroc, Aubusson, Bergamo and Expert, the first two having earlier heading dates than the two others. The second, "Mixture 2", was composed of Bagou, Belepi, Boregar and Kalahari, with Belepi and Kalahari being taller at maturity.

Field trial

A field trial was conducted at Le Moulon, Gif-sur-Yvette, France, from sowing on October 30, 2019 to harvest on July 20, 2020. It was fully replicated once, with sowing on November 5, 2020, and harvest on August 2, 2021. Each season, the whole trial was a rectangle of $8m$ wide and $12m$ long, and comprised 12 "nano-plots" ($2m^2$ for pure stands and $1.5m^2$ for mixed stands): 8 pure stands sown with 160 seeds.m-2 and 4 mixed stands also at 160 seeds.m^{-2} (two replicates per mixture). The first season, both mixtures were also observed at density 250 seeds.m^{-2} (no replicate). All nano-plots were sown as a regular grid made of 12 ranks and 30 rows (pure stands) or 22 rows (mixed stands), with an equal row and rank spacing of 8 cm. The spatial distribution of each mixture (Figure S1) was obtained by regularly randomizing each genotype (Lieng et al., 2012), and seeds were sown at the center of PVC rings of $0.2cm$ thickness, $5.5cm$ radius and $5cm$ depth (Figure S2) to facilitate the phenotyping and harvest of individual plants without disturbing their growth. In terms of agronomic management, seeds were

treated prior to sowing, two fungicides as well as a herbicide were applied during the crop cycle, and nitrogen fertilization was applied twice (at the BBCH30 and BBCH55 stages). Corridors between nano-plots were regularly weeded by hand. Average daily temperature, rainfall and global radiation were measured by a local weather station and monitoring data was retrieved from the INRAE CLIMATIK platform (https://agroclim.inrae.fr/climatik/, in French) managed by the AgroClim laboratory of Avignon, France. The average daily temperature between October and August was 12.1$°C$ in 2019-2020 and 10.8$°C$ in 2020-2021, cumulative rainfall was 617 mm in 2019-2020 and 665 mm in 2020-2021, and cumulative global radiation was 4261.22 MJ.m^{-2} in 2019-2020 and 3447.19 MJ.m^{-2} in 2020-2021.

Phenotyping

For each season, from January until June, multiple traits were phenotyped once a month. In pure stands, 20-30 plants were sampled at each time point to record height (PH, in cm), tiller number (TN) and above-ground dry biomass (PW, in g) after a 48h-drying at $60°C$. In mixtures sown at 160 seeds.m^{-2}, height and tiller number were recorded non-destructively at each time point for each individual plant. Only in May, the plants of the last two rows were sampled to record their above-ground dry biomass as in the pure stands. Phenological stages BBCH30 (ear at 1cm) and BBCH55 (half of ear emerged above the flag leaf ligule; heading) were determined for pure stands at plot scale. At harvest all remaining plants were sampled to record, at the individual-plant scale, shoot dry weight (SW, in g), spike dry weight (SPW, in g), spike number (SN), grain number (GN) and grain weight (GW). From these were then computed the total above-ground dry biomass (PW = SW + SPW, in g), the harvest index (HI = GW / PW), the grain number per spike (GNpS = GN / SN) and the thousand kernel weight (TKW = GW / GN * 1000, in g). The number of plants at sowing and at harvest per genotype for each stand for both years is available in Table S2.

Statistical analysis

Global ANOVA

A linear model was fitted on the whole data set at density 160, with grain weight per plant (GW) as the response and with several explanatory factors, "genotype" (eight levels), "stand" (two levels; pure and mixed) and "season" (two levels; 2019-2020 and 2020-2021), and their interactions. Generalized least squares with REML were used to estimate an error variance per season (heteroscedasticity). The significance of each of the three main factors was assessed based on their F test statistic from ANOVAs with sequential tests ("type II"). For each season, as there was no significant difference between the two replicates of any given mixed stand, data from these replicates were subsequently analyzed jointly.

Relative yield totals

The relative yield total of a given mixed stand (Weigelt and Jolliffe, 2003) was computed as $RYT = Y_{mix}/(p_1 \cdot Y_{pure,1} + p2 \cdot Y_{pure,2} + p3 \cdot Y_{pure,3} + p4 \cdot Y_{pure,4})$ where Y_{mix} was the grain weight per plant averaged over all plants in the mixed stand and, for each genotype i, p_i was the proportion of harvested plants (details of number of harvested plants per genotype available in Table S3 in the mixed stand being of genotype i and $Y_{pure,i}$ was the grain weight per plant averaged over all plants in the pure stand.

Dominance rankings between genotypes

Dominance rankings between genotypes of a given mixture were defined based on Tukey's range test on plant grain weight using the R package multcomp (Hothorn et al., 2008). This test was applied per mixture and season. Pairwise differences between genotypes were declared significant if their p-values were below 0.05, followed by a letter-based representation (Piepho, 2004). The same procedure was also applied on the pure-stand data for all genotypes of a given mixture.

Tests of within-genotype phenotypic plasticity

For each genotype per season, the following model was used: $E[y_i] = M_i$, where y_i is the trait value of interest for plant i, M_i is the stand (pure or mixed) and E is the expectation. The null hypothesis is that the mean (respectively, variability) of the trait in the pure stand is equal to its mean (resp., variability) in the mixed stand. The coefficient of variation (CV) was used as the variability metric instead of the variance because of the possible correlation between mean and variance. The hypothesis was tested separately for traits related to yield at harvest as well as for height and tiller number. For both metrics (mean and CV), a bootstrap algorithm was applied to avoid assuming equal variances between pure and mixed stands as well as normality of trait values per stand, notably to test for CV equality (Amiri and Zwanzig, 2010). For each metric, 1000 bootstrap samples were generated to approximate the distribution of the test statistic under the null hypothesis, and a p-value was computed based on this distribution. No correction was applied to correct for multiple testing as justified in an exploratory context (Rothman, 1990; Heller, 2011).

Link between within-genotype variability and slope of reaction norms

For each genotype per season, a linear regression of grain weight as a function of stand (pure or mixed) was fitted to estimate the slope of the reaction norm. These estimates were then regressed onto the coefficients of variation in pure stands: $y_i = \beta_0 + \beta_1 \cdot CVp_i + \epsilon_i$, with ϵ_i $N(0, \sigma^2)$; where y_i is the absolute value of the slope of the reaction norm for grain weight for the ith genotype in a given season, CVp_i is the coefficient of variation of grain weight for this genotype in the pure stand the same season, and ϵ_i is the error term of variance σ^2. The inverse of the standard errors of estimates of the reaction norm slopes were used as weights of the corresponding error terms.

Decomposition of variability in grain weight

Following (Piepho, 1995), the coefficient of variation of grain weight per plant (GW) in a given stand was decomposed as the product of the contribution of spike number (SN), grain number per spike ($GNpS$) and thousand kernel weight (TKW): $CV_{GW} \approx \sigma_{log(GW)} = (c_{SN} + c_{GNpS} + c_{TKW})^{1/2}$. Each c_i corresponds to the contribution of component i (SN, GNpS, TKW), computed as:

$$c_{SN} = Var[log(SN)] + Cov[log(SN), log(GNpS)] + Cov[log(SN), log(TKW)]$$

$$c_{GNpS} = Var[log(SN)] + Cov[log(SN), log(GNpS)] + Cov[log(GNpS), log(TKW)]$$

$$c_{TKW} = Var[log(SN)] + Cov[log(SN), log(TKW)] + Cov[log(GNpS), log(TKW)]$$

Tests of between-genotype phenotypic plasticity

To assess if different genotypes had more similar trait values in mixed or pure stands, we tested (i) mean differences among different cultivars in pure versus mixed stands with a bootstrap approach, as well as (ii) CV differences among different cultivars with the method from Amiri and Zwanzig (2010) previously mentioned.

Reproducibility

Analyses were conducted in R (R Core Team 2020) and figures were produced with the ggplot2 package (Wickham, 2009). Supplementary information, data and code that support the findings of this study are openly available at the INRAE space of the *Recherche Data Gouv* repository https://doi.org/10.57745/LZS8SU.

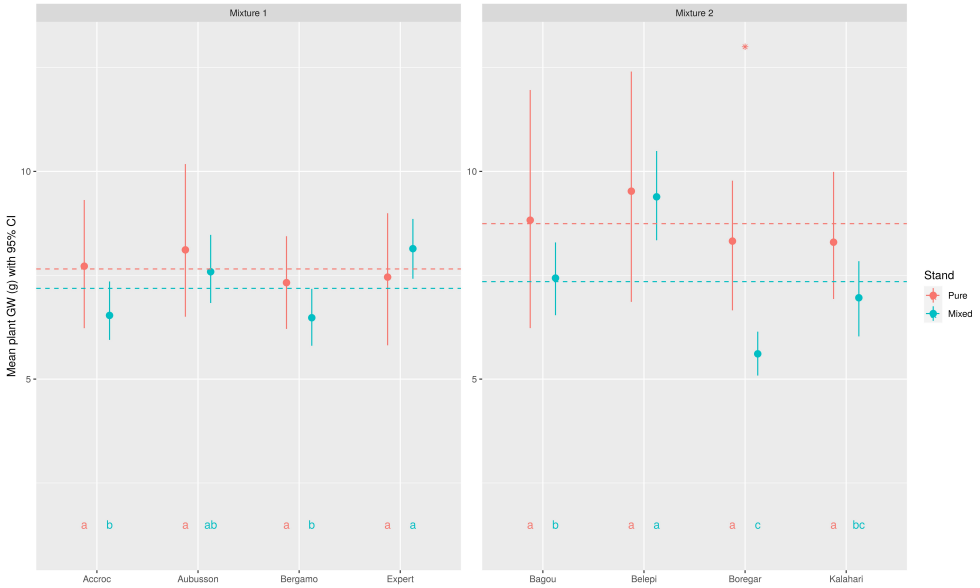
Results

Stand-year interaction effects on productivity, and assessment of dominance rankings

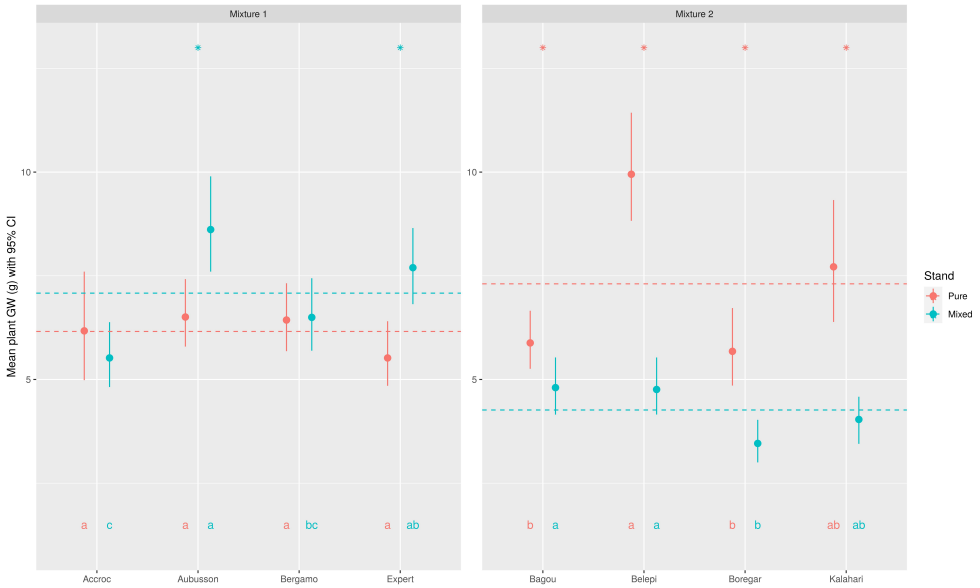
The two seasons of the field experiment were characterized by contrasting trends in terms of climatic variables (Figure S4). In 2019-2020, the global radiation over the whole crop cycle was higher than the average of the previous two decades; similarly for the mean temperature from BBCH30 until grain maturity, and for rainfall (even though there was a deficit during grain filling). In contrast, 2020-2021 had a slight deficit in global radiation from BBCH30 until grain maturity, exacerbated during grain filling, mean temperature was in the average, and an excess of rainfall was observed from BBCH55 up to grain maturity. Overall, the conditions of the first season were more favorable to biomass production and less stressful than the second season, as confirmed by the plant-level above-ground dry biomass averaged over all plants at harvest, around $17.73 \pm 0.35g$ in 2019-2020 and $13.41 \pm 0.25g$ in 2020-2021.

To assess the main factors contributing to the variability of individual plant productivity in the field experiments, an analysis of variance was performed with grain weight per plant (GW) as the response and with "genotype", "stand" and "season" as explanatory factors (Table S3). Results showed that these three factors were all significantly associated as well as all their interactions except "stand" with "season".

At the genotype scale, the comparisons of mean grain weight per plant and per year between pure and mixed stand displayed a season-by-mixed stand interaction (Figure 3). In 2019-2020, both mixtures performed worse than the mean of their components in pure stands, with a relative yield total of 0.92 for Mixture 1 and 0.83 for Mixture 2. However, in 2020-2021 Mixture 1 had an over-yielding of 1.11, while under-yielding reached 0.60 for Mixture 2.



A) 2019-2020



B) 2020-2021

Figure 3 – Mean grain weight per plant and season for each genotype in pure and mixed stands. Dashed lines represent overall means per stand (pure versus mixed). Bars represent 95% confidence intervals obtained by bootstrap. A star at the top of the plot indicates a significant difference (from the bootstrap) in mean grain weight in pure versus mixed stand. A letter-based representation of all pairwise comparisons between genotypes with a Tukey rank test is indicated per stand at the bottom of each plot.

In each mixture per season, the dominating (respectively, dominated) genotypes were defined as those with the highest (resp., lowest) mean grain weight per plant based on the results of a Tukey test (Figure 3, Table S4). In contrast to relative yield totals, the dominating genotypes did not change between seasons. For Mixture 1 (respectively, Mixture 2), cultivars Aubusson and Expert (resp. Bagou and Belepi) were classified as dominating over Accroc and Bergamo (resp. Boregar and Kalahari), most clearly in 2020-2021 (resp. 2019-2020). The same ranking was obtained using the mean grain number per plant (Figure S3), except for Mixture 2 in 2019-2020 where Boregar stood out with a very low grain number. Compared to the main experiments performed at density 160 seeds.m^{-2}, both plant grain weight and spike number at density 250 plants.m^{-2} in 2019-2020 were consistently significantly reduced for all genotypes (Figure S5), and dominance rankings were also different. For Mixture 1, all cultivars had the same rank. For Mixture 2, ranks were not as contrasted as at 160 plants.m^{-2} and Belepi was not dominating. Moreover, for genotypes dominating at density 160 seeds.m^{-2}, such as Expert and Belepi, spike number was even more reduced at density 250 seeds.m^{-2}. Despite these decreases at plant level, spike number and grain weight per m^2 were higher at density 250 seeds.m^{-2} for all cultivars, except Bagou and Belepi (Figure S6), resulting in an overall yield at 250 seeds.m^{-2} higher than at 160 seeds.m^{-2}.

Phenotypic plasticity in pure vs mixed stands at the within-genotype scale and the environmental canalization of reaction norms

As the experimented cultivars were inbred lines, with fixed genotypes, phenotypic plasticity was firstly assessed at the within-genotype scale (Figure 1). Overall (Table S5, Figure S7), distributions were similar between the two seasons, particularly for Mixture 1, yet there were clear differences between mixtures. For Mixture 1, most variables in most genotypes had similar means and CVs in pure

and mixed stands whereas most results for Mixture 2 had different means (still with a similar CVs) indicating more phenotypic plasticity in terms of means for cultivars in Mixture 2. Still, in this mixture, plasticity for trait variability was also observed in 2020-2021 for half of the traits.

Mean of grain weight and its components

Neighborhood genetic heterogeneity, experimented here as pure versus mixed stands, impacted the dominance rankings between genotypes all along the crop cycle, hence the establishment of yield components and finally yield itself. Starting this analysis with an example, this is best illustrated by a comparison between a dominating cultivar, Bagou, and a dominated one, Kalahari, from Mixture 2 in 2019-2020 (Figure 4). We chose them in particular because, even though Kalahari was 19 cm taller than Bagou at maturity, this height differential was not reflected in their dominance ranking, highlighting the need to also consider other traits. For both genotypes, productivity, as assessed by grain weight per plant, was lower on average in mixed stands compared to pure stands but not significantly. Yet, their allocation strategy as assessed by the harvest index (HI) showed opposite trends. The higher HI for Bagou in the mixed stand was due to a stronger decrease in plant weight in the mixed stand than the decrease in grain weight. This drop in plant weight was itself due to a lower spike number in the mixed stand, determined by tillering dynamics before flowering. Interestingly, Kalahari displayed a different behavior. Its lower HI in the mixed stand was the result of a decrease in grain weight and a slight increase in plant weight due to a slight increase in spike number. For this genotype, not only the grain number per spike was slightly lower in the mixed stand, but the thousand kernel weight (TKW) displayed a strong drop, indicative of a post-flowering competition-related stress during the grain filling period.

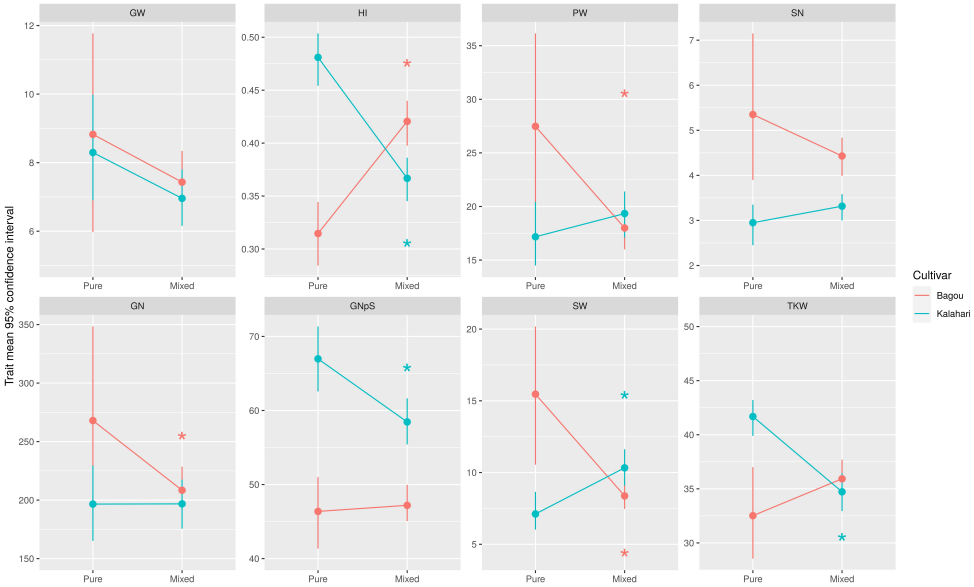


Figure 4 – Reaction norms for plant-level mean values of two contrasted genotypes of Mixture 2, dominating Bagou and dominated Kalahari, in 2019-2020 for different traits at harvest. Bars represent 95% confidence intervals obtained by bootstrap. Stars indicate significant differences between pure and mixed stands.

Extending this analysis to all cultivars in 2019-2020 (Figure S8.A), the reaction norms of yield components displayed a decreasing trend from pure to mixed stand and none was increasing. However, most had an increase in HI in the mixed stand, likely due to a decrease in plant weight, itself explained by a decrease in spike number, pointing to the importance of tillering dynamics in the pre-flowering period. In addition, differences between pure and mixed stands also occurred in the post-flowering period, as exemplified by the significant increase in TKW for Belepi (dominating) and a decreasing trend for Accroc (dominated).

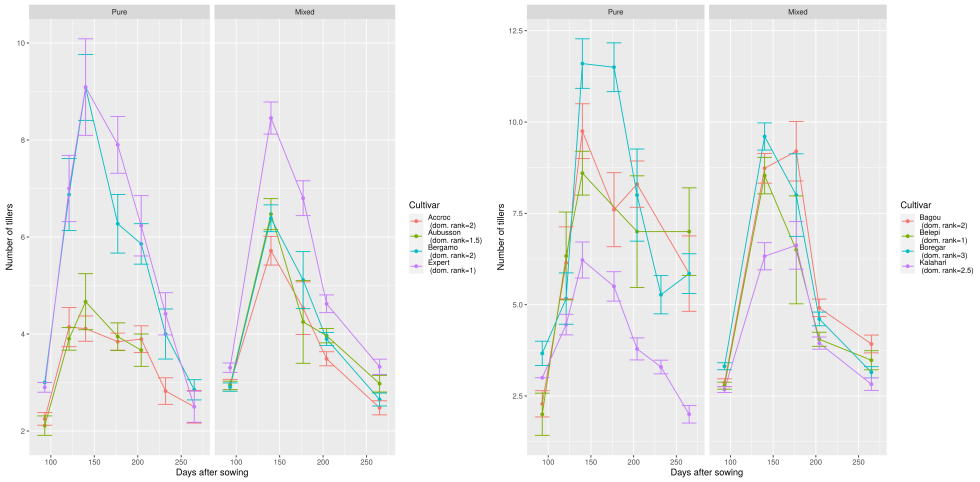
For the second season, 2020-2021 (Figures S8.B and S9), Aubusson and Expert, both dominating in Mixture 1, displayed higher mean trait values in mixed stands, both driven by pre- as well as post-flowering interactions (higher tillering and grain number per spike). For cultivars of Mixture 2, and both dominated

cultivars of Mixture 1, nearly all mean trait values were lower in mixed stands, noticeably HI.

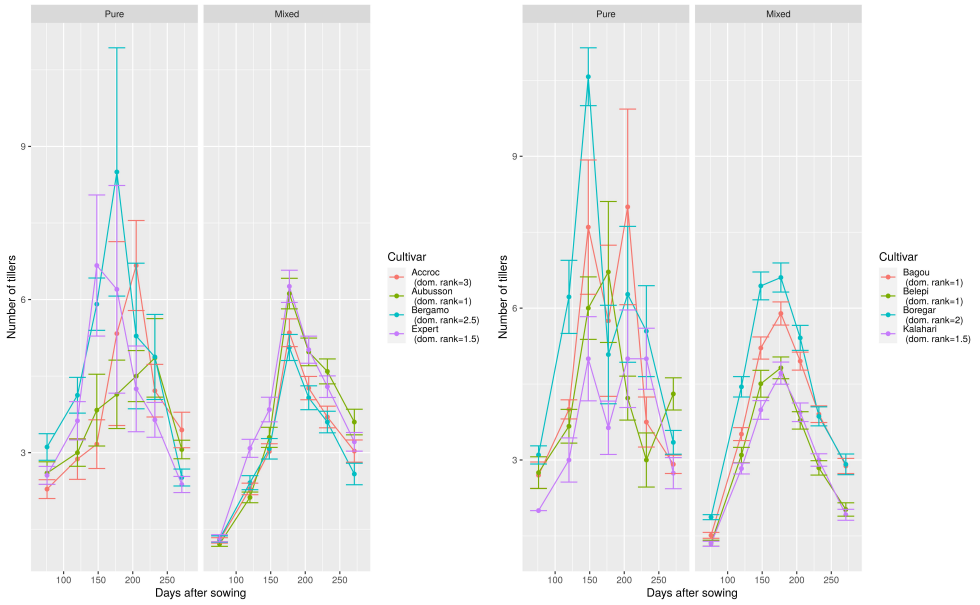
Dynamics of phenotypic plasticity in height and tillering

To understand the dynamics governing the establishment of dominance rankings, key traits involved in competition for light, height and tiller number, were monitored along the growth cycle at plant-level. First of all, for height (Figure S10), no major difference was observed between stands for any given genotype. Mean height between pure and mixed stands were significantly different only at a few dates for some genotypes (Figure S11) and the magnitude of these differences remained weak. One exception was Accroc being taller (+ 10 cm) in mixture in 2020-2021, yet it remained dominated as in the previous season. Overall, the genotype ranking based on height after flowering was mostly conserved between seasons and the tallest genotypes at maturity were not necessarily dominating, as was particularly the case of Kalahari in Mixture 2, especially in 2019-2020 (+ 10 cm taller than the second tallest genotype Belepi). Combining both height and earliness was not associated either with dominance rankings, as exemplified by Mixture 1 in 2019-2020 where Accroc (dominated) and Aubusson (dominating) were earlier than the others, yet shorter than them after flowering. Furthermore, in the majority of cases, height after flowering was not associated with a better grain filling in mixture. This was particularly the case of Kalahari dominated in 2019-2020 that was the tallest in Mixture 2, yet its TKW was lower in the mixture compared to the pure stand. Also, in 2019-2020 for the dominating and tall Expert, the dominated and short Accroc and the dominated and tall Bergamo, all three had their TKW equal in pure and mixed stands. The fourth, Aubusson had a TKW higher in the mixture, but its post-flowering height was average.

Tillering dynamics however were different between pure and mixed stands (Figures 5 and S12). When summarized into two key metrics, tiller number at tillering cessation and tiller regression, they could explain a large part of dominance ranking in mixture as spike number is a major driver of grain weight. For



A) 2019-2020



B) 2020-2021

Figure 5 – Mean number of tillers per plant throughout growth for each genotype in pure and mixed stands. Error bars correspond to standard errors.

instance, in 2019-2020 for Mixture 1, Expert had a similar peak in both stands, but a weaker regression in mixture making him the cultivar with the highest spike number in this stand, in agreement with it being dominating. In contrast, both earlier cultivars Accroc and Aubusson had a peak significantly higher in mixed stands, yet counterbalanced by tiller regression. In 2019-2020 for Mixture 2, even though Boregar had the highest peak in pure stand, its peak in mixture was significantly lower and, as it displayed a strong tiller regression in both stands, it ended up with the third lowest spike number in mixture, hence explaining, at least partly, it being dominated. At the same time, Kalahari had fewer tillers in both stands throughout growth than any other genotype. Even though its tiller regression was not as strong in mixture than in pure stand, it still ended up being dominated with the lowest spike number. In 2020-2021 for Mixture 1, the most dominating genotype, Aubusson, had both a higher peak and a weaker cessation in the mixed stand. For all three others, their lower peak in mixture was counterbalanced by a weaker regression than in pure stand. At maturity, both dominated genotypes hence ended up with the same number of spikes as in pure stand, and the other dominating genotype, Expert, even had more spikes in the mixture. This likely explained the 11% overyielding. In contrast, Mixture 2 in 2020-2021 had a massive 40% under-yielding, as illustrated by the fact that all genotypes had a lower peak and equal or stronger tiller regression in mixture, even the dominating genotypes.

Variability of grain weight and its components

Beyond mean differences, significant differences between pure and mixed stands were also observed for trait variability within each cultivar, as measured by the coefficient of variation. As above, cultivars Bagou and Kalahari from Mixture 2 in 2019-2020 are used to illustrate the analysis (Figure 6). The dominating genotype, Bagou, displayed a decrease in variability in the mixed stand for most trait values whereas the dominated Kalahari showed the opposite trend.

Overall, most CV remained equal or decreased from pure to mixed stands,

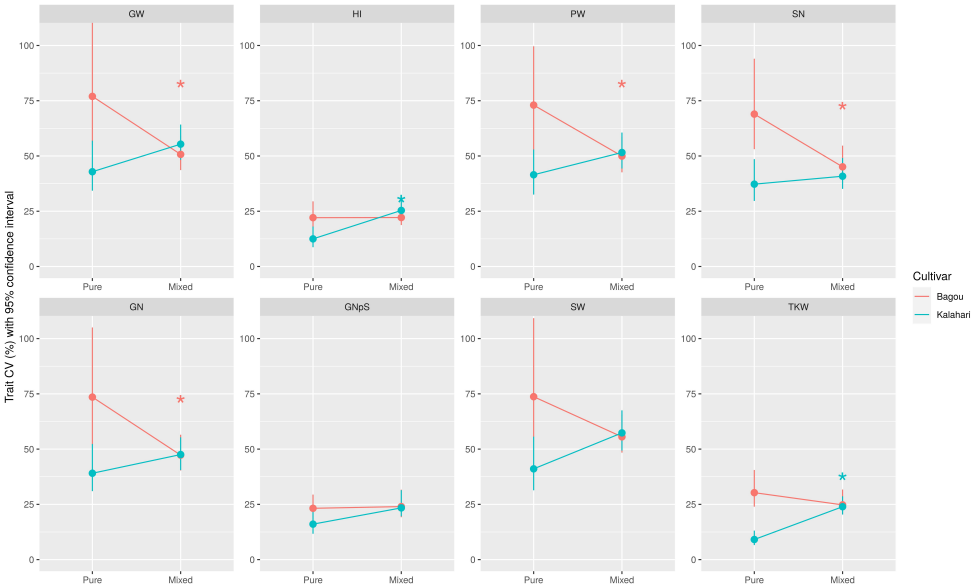


Figure 6 – Reaction norms for plant-level CV values of two contrasted genotypes of Mixture 2 (dominating Bagou in red and dominated Kalahari in cyan) in 2019-2020 for different traits at harvest. Bars represent 95% confidence intervals obtained by bootstrap. Stars indicate significant differences between pure and mixed stands.

with the exception of mixture 2 in 2020-2021 displaying several cases of increased CV in mixtures (Figure S13.A). In 2019-2020, decreases in CV in mixed stands were also occasionally observed for Expert and Bagou, both dominating in their respective mixture, but there was no other case of increase besides Kalahari. In 2020-2021 (Figures S13.B and S14), decreases in CV in mixed stands were not associated with dominance rankings as they were observed not only for Expert (dominating) but also for Bergamo (dominated).

Decomposition of grain weight variability in pure vs mixed stands

The decomposition of grain weight (GW) per plant into a product of spike number per plant (SN), grain number per spike (GNpS) and thousand kernel weight (TKW) allowed us to assess the relative contribution of each of these components to the overall GW variability, and investigate the differences in pure versus mixed stands (Figure S15). In pure stands, for all genotypes in both seasons, spike number contributed the most, usually by far. It was also the major contributing component in mixtures, with the exception of Mixture 2 in 2020-2021 where the contribution of grain number per spike was slightly higher.

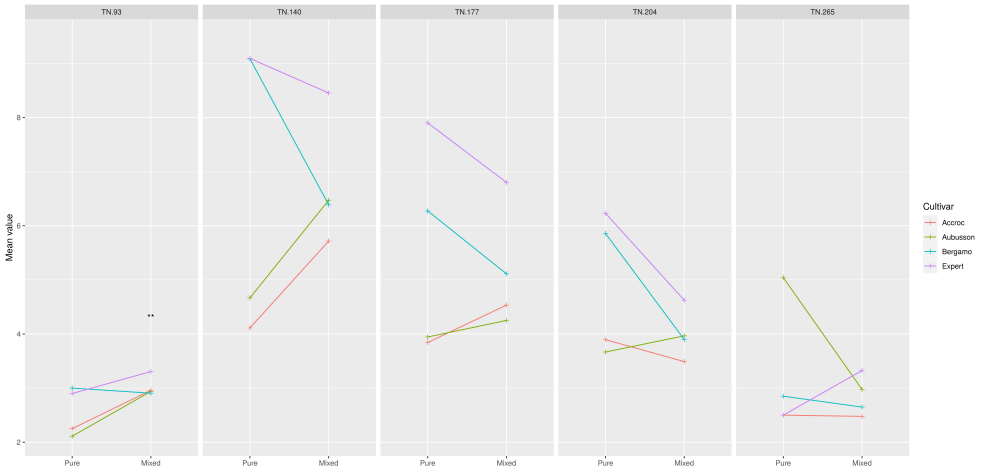
Moreover, for the three cases with a significant decrease in CV in mixed stands for grain weight (Bagou dominating in 2019-2020; Bergamo dominated and Expert dominating in 2020-2021), all three yield components, SN, GNpS and TKW, contributed to the decrease. Yet, the biggest contribution came from the spike number, especially for Bagou. More specifically, this contribution of spike number was mainly due to a variance in pure stands higher than in mixed stands and, to a lesser extent, covariances of spike number with GNpS and TKW also higher in pure stands (Figure S16).

Link between within-genotype CV in pure stand and the slope of reaction norms in pure-vs-mixed stand

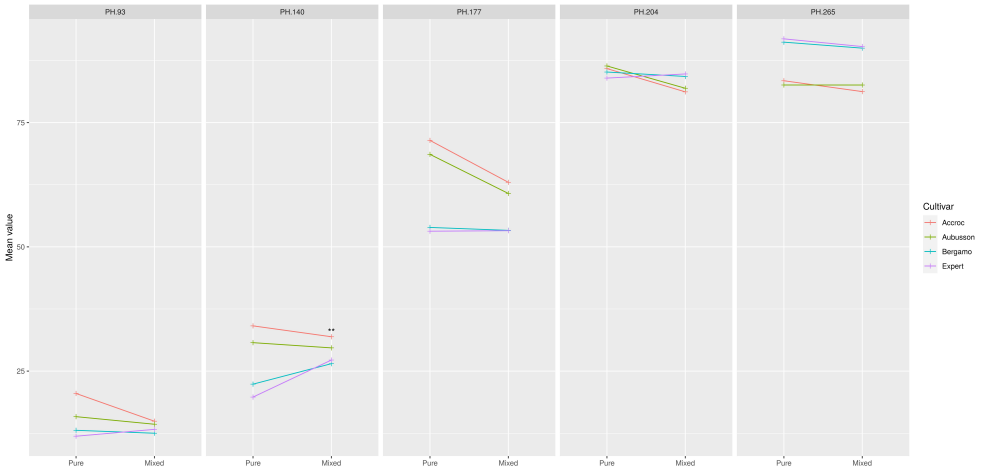
In any panel of wheat cultivars, some genotypes have a yield across environments (sites, years, managements) more stable than other genotypes. We thus tested if the magnitude of variability, as measured by the CV, for plant-level GW in a pure stand was associated with the strength of the reaction norm between pure and mixed stands, as assessed by the absolute value of its slope. However no significant relation was found (Figure S17 and Table S6).

Phenotypic plasticity in pure vs mixed stands at the between-genotype scale and the genetic canalization of reaction norms

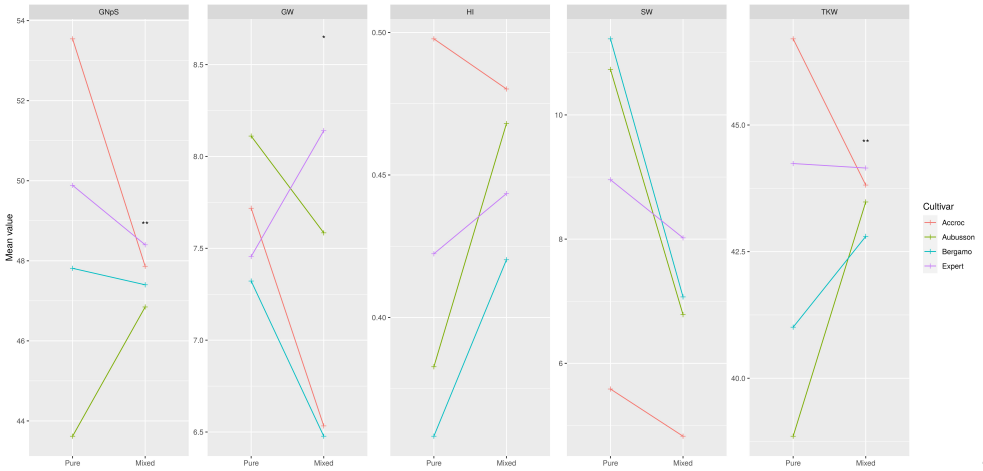
Differences in mean and CV between pure and mixed stands were also tested per season at the between-genotype scale. Overall, for the eight traits at harvest, both mixtures and both seasons exhibited similar distributions (Table S7, Figures S18, S19, S21), with a majority of cases without CV difference, i.e., cases of genetic canalization or decanalization were scarce. Mean and CV differences between stands were also tested for height and tiller number throughout growth (Figures S20 and S22). As differences in means between genotypes were directly related to dominance rankings presented above, we focus now on differences in CV. Even though the statistical power was low (sample size equaling 4 in four-way mixtures), a few cases of phenotypic convergence in mixture were significant, for early tiller number and TKW in 2019-2020 (Figure 7) and for tiller number close to the peak in 2020-2021 (Figure S23). Nevertheless, a visual examination of the reaction norms hinted at several other cases of phenotypic convergence (noticeable narrowings in mixed stands). Similarly, over both mixtures and seasons, the majority of tillering dynamics also appeared less variable compared to in pure stands (Figure 5).



A) Tillering dynamics



B) Plant height throughout growth



C) Productivity and allocation traits at maturity

Figure 7 – Reaction norms in pure versus mixed stands of traits involved in plant-plant interactions with respect to light (tillering and height) and of productivity and allocation traits at maturity for cultivars of Mixture 1 in 2019-2020. Stars indicate significant CV differences between genotypes, * if higher in the mixed stand, ** if lower.

When considering all CV differences and not only the significant ones (50%), a majority (83%) were lower in mixtures than in pure stands (Table S8 and Figures S21 and S22). The only exception to such a phenotypic convergence was Mixture 1 in 2020-2021 for harvest variables (Figure S23). The overyielding mixture hence was also the only one displaying a majority of phenotypic divergence between genotypes across multiple traits at harvest.

Explanations of relative yields by multi-trait within-genotype and between-genotype phenotypic plasticities in pure versus mixed stands

Results detailed above on within-genotype and between-genotype phenotypic plasticities in pure versus mean stands in terms of mean traits can be integrated to explain the relative yield totals as well as the dominance rankings in the various modalities of the field experiments (Figure 8).

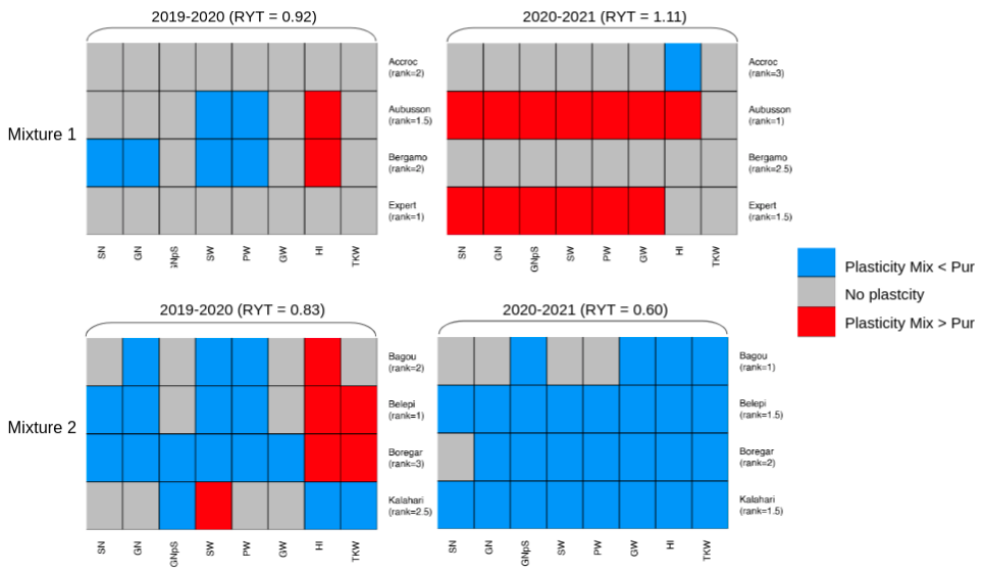


Figure 8 – Heatmap of within-genotype phenotypic plasticity in pure versus mixed stands for traits at harvest, in terms of mean trend, among cultivars of both mixtures for both seasons. Significant differences based on a bootstrap test are blue (respectively, red) if the mean in the mixed stand is lower (resp., higher) than in the pure stand, and gray otherwise. The relative yield total (RYT) is indicated per mixture and season, and dominance ranking is indicated per cultivar.

For Mixture 1 in 2020-2021, both dominating cultivars Aubusson and Expert had higher means in the mixture for yield components SN and GNpS compared to in pure stand, while the dominated Accroc and Bergamo showed no difference between stands for these traits. Both dominating genotypes took advantage of a lower competition for light in the mixed stand, notably Expert with a reduced tiller regression, and Aubusson displayed a higher HI. The dominated genotypes did not suffer too much from this competition compared to in pure stands. Overall, it resulted in an over-yielding.

For Mixture 1 in 2019-2020 however, dominating Expert and dominated Accroc were not plastic between pure and mixed stand. Interestingly, Accroc displayed almost no within-genotype plasticity between pure and mixed stands for both seasons, while also being dominated. Moreover, dominated Bergamo suf-

fered from the competition for light before flowering with a lower tiller peak ending in a lower SN in mixture. Although dominating Aubusson displayed a higher HI in the mixed stand, no difference occurred during grain filling for any genotype, hence leading overall to a moderate under-yielding.

For Mixture 2 in 2019-2020, dominating Belepi had the same number of tillers at the peak but a much stronger regression in the mixed stand, similarly for the second most dominating genotype, Bagou. The most dominated genotype, Boregar, not only had a lower tiller peak in mixture, but also a stronger regression. Such strong regressions in the mixed stand may have been caused by the fourth genotype, Kalahari, which happens to be the tallest since before the heading stage. Yet, this did not help it particularly during grain filling as it had a lower TKW than in the pure stand. Overall, it resulted in a substantial under-yielding. In contrast to the three other modalities, no genotype of Mixture 2 in 2020-2021 took advantage of being in a mixed stand, i.e., all four genotypes had lower means in the mixture for yield components SN, GNpS and TKW. Amplified further by a lower HI, this led to a massive under-yielding.

Furthermore, within-genotype mean differences in pure versus mixed stand tended to be associated with between-genotype CV differences. For instance, for Mixture 2 in 2019-2020, cultivars Belepi and Boregar ended up with less spikes in mixtures than in pure stands, and their neighbors Bagou and Kalahari with as many spikes in mixed as in pure stands. Overall, it resulted in a homogenization of the mean spike number per plant among these cultivars and, as a consequence, a decrease in between-genotype CV. In contrast, within-genotype CV differences (Figure S24) were not associated with between-genotype differences in trait means. Indeed, the main trend of phenotypic convergence in mixed stands (except at harvest for Mixture 1 in 2020-2021) was observed in parallel to within-genotype CV decreases for traits at maturity in Mixture 1 in 2019-2020, increases in Mixture 2 in 2020-2021, and both in Mixture 2 in 2019-2020.

Discussion

Any individual plant encounters multiple and various heterogeneities all along its life cycle, whether from abiotic or biotic factors, starting with its closest plant neighbors. As a result of these interactions, plants are well-known to adjust their development, whether in an anticipatory or reactionary manner, to the peculiarities of their immediate surroundings, with potentially major consequences on their fitness (Sultan, 2003). In an agricultural setting, such a phenotypic plasticity is invoked when comparing the productivity of varietal mixtures assembled from cultivars initially selected in monocultures (Barot et al., 2017). The proportions of genotypes sown in a mixture may indeed differ between sowing and harvest, as well as the total yield of the mixture when compared to the mean of the monocultures. Yet, the links between the phenotypic plasticity between plants of the same genotype and the fitness increase of a genotype in addition to or in the detriment of the others, remain poorly studied.

An innovative experimental design

Many previous studies documented the discrepancy between genotype performances in pure versus mixed stands and aimed at deciphering the underlying mechanisms, but were limited in one way or another by their experimental design. In field experiments, it is indeed almost impossible to distinguish different cereal cultivars at the vegetative stages. It remains difficult after flowering, except with morphological cues such as height (Jennings and Aquino, 1968), spike aristation, or chaff color differences (Finckh and Mundt, 1992), that greatly restricts the choice of cultivars. Total yields of the micro-plot design, with mixing occurring within sowing ranks, hence are analyzed with empirical "trait-blind" models that, despite recent advances (Forst et al., 2019), cannot study the phenotypic plasticity of mixed genotypes between pure and mixed stands.

Alternatively, nursery designs allow the distinction of genotypes but intro-

duce a systematic bias in plant-plant interaction as mixing only occurs between sowing ranks, and they are furthermore restricted to binary or ternary mixtures (Song et al., 2009; Dahlin et al., 2020; Montazeaud et al., 2020). Besides, studies in pots allow analyses at the scale of individual plants, yet at the cost of being only partially representative of field conditions (Montazeaud et al., 2018; Fréville et al., 2019). Hence, by tracking the genotype of plants in mixed stands throughout growth until harvest, our new experimental design enabled us to measure phenotypic plasticity and canalization at both the within- and between-genotype scales. Moreover, monitoring tillering dynamics as well as assessing biomass allocation and all yield components under realistic field conditions provided us with unprecedented access to the determinants of yield.

Neither genotype productivities in pure stands nor height and earliness contrasts explained dominance ranking in mixtures

First of all, the genotypes we experimented with were all elite cultivars, hence with a yield potential expected to be similar, which was confirmed in the field trials (generally, no significant difference in pure-stand yields between the genotypes belonging to the same mixture). Yet, dominance, i.e., genotypes ranking higher than the others in terms of yield, occurred in mixed stands, whether or not the mixed stands were under- or over-yielding. This result confirms that mixture assembly cannot be based only on genotype productivities in pure stands (Dawson and Goldringer, 2012). Moreover, in the only exception corresponding to Mixture 2 in 2020-2021, not only the higher-yielding genotypes in pure stands were not necessarily dominant in the mixture, but all genotypes also produced less in the mixed stand compared to their pure stand, hence explaining the massive under-yielding.

Second, we chose our cultivars based on contrasts of earliness and height at

maturity, and the mixtures we assembled covered various combinations of these traits, e.g., a genotype taller than all others during the whole life cycle (such as Kalahari in Mixture 2), or a genotype earlier than the others for the first half of the cycle (such as Accroc and Aubusson in Mixture 1). Despite that, our results showed that contrasts of height at maturity or earliness were not associated with the dominance rankings in mixed stands, i.e., taller or early genotypes did not end up with a higher grain yield contrary to what could have been expected. As illustrated by the tallest genotype in Mixture 2 (Kalahari), its harvest index in 2019-2020 lower in the mixture compared to in the pure stand suggested an over-investment in vegetative organs at the expense of reproductive ones. This behavior typical of a tragedy of the commons (Anten and Vermeulen, 2016) hence contributed to the under-yielding of the mixed stand.

Overall, neither genotype productivities in pure stands nor height and earliness contrasts sufficed in explaining the dominance rankings in mixtures.

Impact of yield components in determining dominance rankings in mixture

Our results highlighted the importance of monitoring the yield components (life-history traits) and resource allocation per genotype in both pure and mixed stands, all along the growth cycle, in order to explain the relative yields and dominance rankings at harvest. Indeed, the relative yield total of each mixture was explained thanks to the comparison of the grain yield between genotypes in the mixture (dominance rankings), itself explained by the yield components, and thanks to the knowledge of phenotypic plasticity in terms of mean differences between pure and mixed stands. Even though grain filling during the post-flowering period was non-negligible, our results showed the preponderance of spike number in explaining yield itself but also relative yield totals. As spike number results from tillering dynamics, tracking them allowed us to identify

which phases differed between pure and mixed stands, most notably tiller number at tillering cessation and tiller regression. Differences between genotypes for one or both of these variables were the main factors in determining dominance rankings in pure stands. Coupled with between-genotype phenotypic plasticity in terms of mean differences between pure and mixed stands, they became main factors in determining dominance rankings in mixtures.

Links between phenotypic plasticity, dominance ranking and relative yield total in mixture

In terms of dominance rankings, beyond the explanations based on yield components as discussed above, the results allowed us to test our initial hypotheses about phenotypic plasticity. At the within-genotype scale, we confirmed the initial hypothesis according to which the mean of certain life-history traits (most notably spike number) for the dominating genotypes would increase in mixed stands compared to pure ones, and decrease for dominated genotypes.

Moreover, significant changes of within-genotype variability were also detected between pure and mixed stands, and this variability was indeed reduced in mixtures for most dominating genotypes as initially hypothesized, even though it was not systematic (e.g., Bergamo in 2020-2021). For dominated genotypes, both cases of increase and decrease were observed, as initially suspected. Interestingly, the mixture with the largest underyielding (Mixture 2 in 2020-2021) was associated with an increase in within-genotype variability for most traits of most genotypes. Neither obvious experimental issue nor identified stress (disease, drought) occurred during the experiment for this mixture, hence a self-generated, unknown stress may have caused this for all genotypes in the mixture.

In terms of mean differences at the between-genotype scale, given the same yield potential of the genotypes in pure stand, the “selection effect” hypothesis

according to which an over-yielding of the mixed stand would be driven by the dominance of the most productive genotype in pure stands, could be neither confirmed nor overturned.

In terms of between-genotype variability differences however, our results displayed an interesting contrast, even though the observed trends were not all statistically significant given the small sample size. Contrary to our initial hypothesis, a higher variability between genotypes in mixtures was not systematically observed, rather the opposite. Moreover, the change in variability from pure to mixed stand was associated with the relative yield total: a decrease in CV with under-yielding versus an increase with over-yielding.

This phenotypic convergence in mixed stands, that would correspond to genetic canalization, was previously observed for barley cultivar mixtures and called "adaptive similarity" (Dahlin et al., 2020). It has also been reported for plant height and relative stretching rates, with smaller cultivars getting taller to avoid shading (Fiorucci and Fankhauser, 2017), as well as for root growth (Craine and Dybzinski, 2013). Here we observed this adaptive similarity for different traits including height at maturity for both mixtures (but especially for Mixture 1 with homogeneous heights among cultivars at maturity), but also for tiller number and some yield related traits. Regarding height, the theory according to which taller cultivars could display an altruistic behavior by getting smaller (e.g., Montazeaud et al. (2020)), was not supported by the data here as smaller cultivars in the same mixture were simultaneously getting taller, and overall the mixture was under-yielding. As stated in (Dahlin et al., 2020), adaptive similarity corresponds to a reduction in niche differentiation, hence leading to a decrease in complementarity and facilitation, and an increase in competition. Besides these cases of phenotypic convergence, the only overyielding modality was also associated with phenotypic divergence for harvest variables. This hardly is a definite proof, but it suggests a hypothesis to be tested further, that phenotypic diver-

gence, hence niche differentiation between genotypes, leads to overyielding in cultivar mixtures.

Furthermore, there could have been a link between the variability in pure stands and the reaction norm between pure and mixed stands. However, when looking at grain weight, no association was found between within-genotype variability in pure stands and the absolute value of the slope of the reaction norm between pure and mixed stands. Yet, such possible links remain to be more amply studied.

Perspectives and open questions

Overall, our study highlights the importance of phenotyping yield components between as well as within genotypes in mixtures. Still, many areas of improvement remain for future work. First of all, at the between-genotype level, future studies should aim at assessing genotype proportions of yield to determine which genotypes are dominant or dominated. Given the presence of mixture-season interaction, this should be done with multi-environment trials and envirotyping. Moreover, the importance of tillering dynamics in explaining dominance relationships calls for experimenting at several plant densities (Darwinkel, 1978).

At the within-genotype level, our results highlight the importance of phenotyping tillering dynamics, but it was the main time-consuming effort during the growth cycle. Indeed, although our experimental design has allowed us to acquire valuable data, it required significant human means. Furthermore, competition for light is well-known to have an impact not only on tillering in terms of light quality as mediated by the red/far-red ratio (Xie et al., 2016), but also on biomass production in terms of light quantity as mediated by the interception efficiency (Craine and Dybzinski, 2013). Phenotyping light interception of individual plants in a canopy, even though in a glasshouse, was recently demonstrated on maize, allowing the subsequent quantification of plant-scale radia-

tion use efficiency (Cabrera-Bosquet et al., 2016). Applying such an approach on wheat, along with the development of an automated phenotyping of tillering dynamics, surely represents a worthy endeavor. Outdoor in the field, one could envision combining a yet-to-design genotype-aware precision seeder with high-throughput phenotyping along the growth cycles thanks to unmanned aerial vehicles (Liu et al., 2016; Holman et al., 2016).

Furthermore, we focused our study on the competition for light, hence restricting the phenotyping to aerial plant architectures, ignoring below-ground organs. Similarly, the number of studied genotypes and mixtures were limited, restricting general conclusions about genotype performances in mixtures. Future studies should explore more the gradient of neighborhood heterogeneity, by using more contrasted cultivars in mixtures, e.g., by mixing elite cultivars with landraces as in the Wheatamix project (Dubs et al., 2018) and by using mixtures of different orders (binary, ternary quaternary, etc.). It would also be very informative to compare mixtures managed with and without fungicide in terms of dominance relationships between genotypes whose traits cover a gradient of both light competition and disease resistance (Finckh and Mundt, 1992).

In addition to improving the experimental design and phenotyping methods, exploring larger combinatorics of mixtures with more genotypes using process-based models of plant growth appears as a relevant strategy. This would particularly be the case of FSPMs as they explicitly represent each plant individually (Gaudio et al., 2019; Blanc et al., 2021).

Supplementary Information

Supplemental material is available at https://doi.org/10.57745/LZS8SU and consists of nine tables and twenty-four figures.

Acknowledgments

The authors thank Didier Tropée and Gaël Vieceli for their valuable help and support with the field experiments, as well as T. Randrianarisoa, N. Vazeux-Blumental, H. Belcram, A. Postec, L. Malicet-Chebbah, B. Rouger, E. Forst, G. Van Frank, E. Blanc, A. Merbhène, M. Turbet-Delof, I. Goldringer, A. Hospital, M. Colas, F. Ammar, A. Sidik Meite, K. Ménard and N. Tinomme. The authors also thank T. Moittie (Asur Plant Breeding), M. Balduz (Lemaire Deffontaines), C. Duquet and M. Prevost (Limagrain Europe), C. Michelet (RAGT2n) and S. Caiveau (Syngenta France SAS) for providing the seeds. This work was supported by the doctoral school FIRE of the Learning Planet Institute and the BAP department from INRAE (PerfoMix project).

References

M. A. Altieri. Agroecology: a new research and development paradigm for world agriculture. *Agriculture, Ecosystems & Environment*, 27(1/4):37–46, 1989.

S. Amiri and S. Zwanzig. An Improvement of the Nonparametric Bootstrap Test for the Comparison of the Coefficient of Variations. *Communications in Statistics - Simulation and Computation*, 39(9):1726–1734, Sept. 2010. ISSN 0361-0918, 1532-4141. doi: 10.1080/03610918.2010.512693. URL http://www.tandfonline.com/doi/abs/10.1080/03610918.2010.512693.

N. P. Anten and P. J. Vermeulen. Tragedies and Crops: Understanding Natural Selection To Improve Cropping Systems. *Trends in Ecology & Evolution*, 31(6):429–439, June 2016. ISSN 01695347. doi: 10.1016/j.tree.2016.02.010. URL https://linkinghub.elsevier.com/retrieve/pii/S0169534716000550.

S. Barot, V. Allard, A. Cantarel, J. Enjalbert, A. Gauffreteau, I. Goldringer, J.-C. Lata, X. Le Roux, A. Niboyet, and E. Porcher. Designing mixtures of varieties for multifunctional agriculture with the help of ecology. A review. *Agronomy for Sustainable Development*, 37(2):13, Apr. 2017. ISSN 1774-0746, 1773-0155. doi: 10.1007/s13593-017-0418-x. URL http://link.springer.com/10.1007/s13593-017-0418-x.

E. Blanc, P. Barbillon, C. Fournier, C. Lecarpentier, C. Pradal, and J. Enjalbert. Functional–Structural Plant Modeling Highlights How Diversity in Leaf Dimensions and Tillering Capability Could Promote the Efficiency of Wheat Cultivar Mixtures. *Frontiers in Plant Science*, 12:734056, Sept. 2021. ISSN 1664-462X. doi: 10.3389/fpls.2021.734056. URL https://www.frontiersin.org/articles/10.3389/fpls.2021.734056/full.

J. Borg, L. Kiær, C. Lecarpentier, I. Goldringer, A. Gauffreteau, S. Saint-Jean, S. Barot, and J. Enjalbert. Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research*, 221:298–313, May 2018. ISSN 03784290. doi: 10.1016/j.fcr.2017.09.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0378429017309498.

D. Bourguet and T. Guillemaud. The Hidden and External Costs of Pesticide Use. In E. Lichtfouse, editor, *Sustainable Agriculture Reviews*, volume 19, pages 35–120. Springer International Publishing, Cham, 2016. ISBN 978-3-319-26776-0 978-3-319-26777-7. doi: 10.1007/978-3-319-26777-7_2. URL http://link.springer.com/10.1007/978-3-319-26777-7_2. Series Title: Sustainable Agriculture Reviews.

R. W. Brooker, A. E. Bennett, W. Cong, T. J. Daniell, T. S. George, P. D. Hallett, C. Hawes, P. P. M. Iannetta, H. G. Jones, A. J. Karley, L. Li, B. M. McKenzie, R. J. Pakeman, E. Paterson, C. Schöb, J. Shen, G. Squire, C. A. Watson, C. Zhang, F. Zhang, J. Zhang, and P. J. White. Improving intercropping: a synthesis of research in agronomy, plant physiology and ecology. *New Phytologist*, 206(1):107–117, Apr. 2015. ISSN 0028-646X, 1469-8137. doi: 10.1111/nph.13132. URL https://onlinelibrary.wiley.com/doi/10.1111/nph.13132.

L. Cabrera-Bosquet, C. Fournier, N. Brichet, C. Welcker, B. Suard, and F. Tardieu. High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. *New Phytologist*, 212(1):269–281, Oct. 2016. ISSN 0028-646X, 1469-8137. doi: 10.1111/nph.14027. URL https://onlinelibrary.wiley.com/doi/10.1111/nph.14027.

J. M. Craine and R. Dybzinski. Mechanisms of plant competition for nutrients, water and light. *Functional Ecology*, 27(4):833–840, Aug. 2013. ISSN 02698463.

doi: 10.1111/1365-2435.12081. URL https://onlinelibrary.wiley.com/doi/10.1111/1365-2435.12081.

I. Dahlin, L. P. Kiær, G. Bergkvist, M. Weih, and V. Ninkovic. Plasticity of barley in response to plant neighbors in cultivar mixtures. *Plant and Soil*, 447(1-2):537–551, Feb. 2020. ISSN 0032-079X, 1573-5036. doi: 10.1007/s11104-019-04406-1. URL http://link.springer.com/10.1007/s11104-019-04406-1.

A. Darwinkel. Patterns of tillering and grain production of winter wheat at a wide range of plant densities. *Netherlands Journal of Agricultural Science*, 26(4):383–398, Nov. 1978. ISSN 0028-2928. doi: 10.18174/njas.v26i4.17081. URL https://library.wur.nl/ojs/index.php/njas/article/view/17081.

J. C. Dawson and I. Goldringer. Breeding for Genetically Diverse Populations: Variety Mixtures and Evolutionary Populations. In E. T. Lammerts van Bueren and J. R. Myers, editors, *Organic Crop Breeding*, pages 77–98. Wiley, 1 edition, Jan. 2012. ISBN 978-0-470-95858-2 978-1-119-94593-2. doi: 10.1002/9781119945932.ch5. URL https://onlinelibrary.wiley.com/doi/10.1002/9781119945932.ch5.

C. de Vallavieille-Pope. Management of disease resistance diversity of cultivars of a species in single fields: controlling epidemics. *Comptes Rendus Biologies*, 327(7):611–620, July 2004. ISSN 16310691. doi: 10.1016/j.crvi.2003.11.014. URL https://linkinghub.elsevier.com/retrieve/pii/S1631069104001234.

F. Dubs, X. Le Roux, V. Allard, B. Andrieu, S. Barot, A. Cantarel, C. d. Vallavielle-Pope, A. Gauffreteau, I. Goldringer, C. Montagnier, T. Pommier, E. Porcher, S. Saint-Jean, J. Borg, S. Bourdet-Massein, D. Carmignac, A. Duclouet, E. Forst, N. Galic, L. Gerard, M. Hugoni, A. Hure, A. Larue, J.-C. Lata, C. Lecarpentier, M. Leconte, E. L. Saux, I. L. Viol, P. L'hote, P. Lusley, M. Mouchet, A. Ni-

boyet, R. Perronne, E. Pichot, S. Pin, S. Salmon, D. Tropée, A. Vergnes, T. Vidal, and J. Enjalbert. *An experimental design to test the effect of wheat variety mixtures on biodiversity and ecosystem services*. July 2018. URL https://hal.archives-ouvertes.fr/hal-01843564.

R. E. Evenson and D. Gollin. Assessing the Impact of the Green Revolution, 1960 to 2000. *Science*, 300(5620):758–762, May 2003. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1078710. URL https://www.science.org/doi/10.1126/science.1078710.

M. R. Finckh and C. C. Mundt. Stripe Rust, Yield, and Plant Competition in Wheat Cultivar Mixtures. *Phytopathology*, 82(9):905, 1992. ISSN 0031949X. doi: 10.1094/Phyto-82-905. URL http://www.apsnet.org/publications/phytopathology/backissues/Documents/1992Abstracts/Phyto82_905.htm.

A.-S. Fiorucci and C. Fankhauser. Plant Strategies for Enhancing Access to Sunlight. *Current Biology*, 27(17):R931–R940, Sept. 2017. ISSN 09609822. doi: 10.1016/j.cub.2017.05.085. URL https://linkinghub.elsevier.com/retrieve/pii/S0960982217306930.

T. Flatt. The Evolutionary Genetics of Canalization. *The Quarterly Review of Biology*, 80(3):287–316, Sept. 2005. ISSN 0033-5770, 1539-7718. doi: 10.1086/432265. URL https://www.journals.uchicago.edu/doi/10.1086/432265.

J. A. Fordyce. The evolutionary consequences of ecological interactions mediated through phenotypic plasticity. *Journal of Experimental Biology*, 209(12):2377–2383, June 2006. ISSN 1477-9145, 0022-0949. doi: 10.1242/jeb.02271. URL https://journals.biologists.com/jeb/article/209/12/2377/9395/The-evolutionary-consequences-of-ecological.

A. Forsman. Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity*, 115(4):276–284, Oct. 2015. ISSN 0018-067X, 1365-2540. doi: 10.1038/hdy.2014.92. URL http://www.nature.com/articles/hdy201492.

E. Forst. *Développement de méthodes d'estimation de l'aptitude au mélange pour la prédiction des performances et la sélection de mélanges variétaux chez le blé tendre et co-conception d'idéotypes de mélanges adaptés à l'agriculture biologique*. PhD thesis, Paris-Saclay, Gif-sur-Yvette, Mar. 2018.

E. Forst, J. Enjalbert, V. Allard, C. Ambroise, I. Krissaane, T. Mary-Huard, S. Robin, and I. Goldringer. A generalized statistical framework to assess mixing ability from incomplete mixing designs using binary or higher order variety mixtures and application to wheat. *Field Crops Research*, 242:107571, Oct. 2019. ISSN 03784290. doi: 10.1016/j.fcr.2019.107571. URL https://linkinghub.elsevier.com/retrieve/pii/S0378429018311675.

V. Foucteau, P. Brabant, H. Monod, O. David, and I. Goldringer. Correction models for intergenotypic competition in winter wheat. *Agronomie*, 20(8):943–953, Dec. 2000. ISSN 0249-5627, 1297-9643. doi: 10.1051/agro:2000170. URL http://www.edpsciences.org/10.1051/agro:2000170.

H. Fréville, P. Roumet, N. O. Rode, A. Rocher, M. Latreille, M. Muller, and J. David. Preferential helping to relatives: A potential mechanism responsible for lower yield of crop variety mixtures? *Evolutionary Applications*, 12(9):1837–1849, Oct. 2019. ISSN 1752-4571, 1752-4571. doi: 10.1111/eva.12842. URL https://onlinelibrary.wiley.com/doi/10.1111/eva.12842.

N. Gaudio, A. J. Escobar-Gutiérrez, P. Casadebaig, J. B. Evers, F. Gérard, G. Louarn, N. Colbach, S. Munz, M. Launay, H. Marrou, R. Barillot, P. Hinsinger, J.-E. Bergez, D. Combes, J.-L. Durand, E. Frak, L. Pagès, C. Pradal,

S. Saint-Jean, W. Van Der Werf, and E. Justes. Current knowledge and future research opportunities for modeling annual crop mixtures. A review. *Agronomy for Sustainable Development*, 39(2):20, Apr. 2019. ISSN 1774-0746, 1773-0155. doi: 10.1007/s13593-019-0562-6. URL http://link.springer.com/10.1007/s13593-019-0562-6.

S. Gliessman. *Agroecology: The Ecology of Sustainable Food Systems*. CRC Press, 2014.

J. B. Grace and D. Tilman, editors. *Perspectives on plant competition*. Academic Press, San Diego, 1990. ISBN 978-0-12-294452-9.

J. Guntrip and R. M. Sibly. Phenotypic plasticity, genotype-by-environment interaction and the analysis of generalism and specialization in Callosobruchus maculatus. *Heredity*, 81(2):198–204, Aug. 1998. ISSN 0018-067X, 1365-2540. doi: 10.1046/j.1365-2540.1998.00354.x. URL http://link.springer.com/10.1046/j.1365-2540.1998.00354.x.

B. Hallgrimsson, R. M. Green, D. C. Katz, J. L. Fish, F. P. Bernier, C. C. Roseman, N. M. Young, J. M. Cheverud, and R. S. Marcucio. The developmental-genetics of canalization. *Seminars in Cell & Developmental Biology*, 88:67–79, Apr. 2019. ISSN 10849521. doi: 10.1016/j.semcdb.2018.05.019. URL https://linkinghub.elsevier.com/retrieve/pii/S1084952117302872.

R. Heller. Discussion of "Multiple Testing for Exploratory Research" by J. J. Goeman and A. Solari. *Statistical Science*, 26(4), Nov. 2011. ISSN 0883-4237. doi: 10.1214/11-STS356A.

F. H. Holman, A. B. Riche, A. Michalski, M. Castle, M. J. Wooster, and M. J. Hawkesford. High Throughput Field Phenotyping of Wheat Plant Height and Growth Rate in Field Plot Trials Using UAV Based Remote Sensing. *Remote*

Sensing, 8(12), 2016. doi: 10.3390/rs8121031. URL https://www.mdpi.com/2072-4292/8/12/1031.

T. Hothorn, F. Bretz, and P. Westfall. Simultaneous Inference in General Parametric Models. *Biometrical Journal*, 50(3):346–363, 2008.

M. Jenkins. Prospects for Biodiversity. *Science*, 302(5648):1175–1177, Nov. 2003. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1088666. URL https://www.science.org/doi/10.1126/science.1088666.

P. R. Jennings and R. C. Aquino. STUDIES ON COMPETITION IN RICE. III. THE MECHANISM OF COMPETITION AMONG PHENOTYPES. *Evolution*, 22(3): 529–542, Sept. 1968. ISSN 00143820. doi: 10.1111/j.1558-5646.1968.tb03990.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1558-5646.1968.tb03990.x.

E. Justes, L. Bedoussac, C. Dordas, E. Frak, G. Louarn, S. Boudsocq, E.-P. Journet, A. Lithourgidis, C. Pankou, C. Zhang, G. Carlsson, E. S. Jensen, C. Watson, and L. Li. THE 4 C APPROACH AS A WAY TO UNDERSTAND SPECIES INTERACTIONS DETERMINING INTERCROPPING PRODUCTIVITY. *Frontiers of Agricultural Science and Engineering*, 2021. ISSN 2095-7505. doi: 10.15302/J-FASE-2021414. URL https://journal.hep.com.cn/fase/EN/10.15302/J-FASE-2021414.

H. Lieng, C. Richardt, and N. A. Dodgson. Random Discrete Colour Sampling, 2012. URL http://diglib.eg.org/handle/10.2312/COMPAESTH.COMPAESTH12.081-087.

T. Liu, W. Wu, W. Chen, C. Sun, X. Zhu, and W. Guo. Automated image-processing for counting seedlings in a wheat field. *Precision Agriculture*, 392 (406), 2016. doi: 10.1007/s11119-015-9425-6. URL https://doi.org/10.1007/s11119-015-9425-6.

M. Loreau. Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science*, 294(5543):804–808, Oct. 2001. ISSN 00368075, 10959203. doi: 10.1126/science.1064088. URL https://www.sciencemag.org/lookup/doi/10.1126/science.1064088.

G. Montazeaud, C. Violle, H. Fréville, D. Luquet, N. Ahmadi, B. Courtois, I. Bouhaba, and F. Fort. Crop mixtures: does niche complementarity hold for belowground resources? An experimental test using rice genotypic pairs. *Plant and Soil*, 424(1-2):187–202, Mar. 2018. ISSN 0032-079X, 1573-5036. doi: 10.1007/s11104-017-3496-2. URL http://link.springer.com/10.1007/s11104-017-3496-2.

G. Montazeaud, C. Violle, P. Roumet, A. Rocher, M. Ecarnot, F. Compan, G. Maillet, F. Fort, and H. Fréville. Multifaceted functional diversity for multifaceted crop yield: Towards ecological assembly rules for varietal mixtures. *Journal of Applied Ecology*, 57(11):2285–2295, Nov. 2020. ISSN 0021-8901, 1365-2664. doi: 10.1111/1365-2664.13735. URL https://onlinelibrary.wiley.com/doi/10.1111/1365-2664.13735.

C. Mundt. Performance of wheat cultivars and cultivar mixtures in the presence of Cephalosporium stripe. *Crop Protection*, 21(2):93–99, Mar. 2002. ISSN 02612194. doi: 10.1016/S0261-2194(01)00067-9. URL https://linkinghub.elsevier.com/retrieve/pii/S0261219401000679.

H. P. Piepho. A simple procedure for yield component analysis. *Euphytica*, 84(1):43–48, Feb. 1995. ISSN 0014-2336, 1573-5060. doi: 10.1007/BF01677555. URL https://link.springer.com/10.1007/BF01677555.

H.-P. Piepho. An Algorithm for a Letter-Based Representation of All-Pairwise Comparisons. *Journal of Computational and Graphical Statistics*, 13(2):456–466, June 2004. ISSN 1061-8600, 1537-2715. doi: 10.1198/

1061860043515. URL http://www.tandfonline.com/doi/abs/10.1198/1061860043515.

M. Pigliucci and K. Preston. Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecology Letters*, 6(3):265–272, Mar. 2003. ISSN 1461-023X, 1461-0248. doi: 10.1046/j.1461-0248.2003.00428.x. URL http://doi.wiley.com/10.1046/j.1461-0248.2003.00428.x.

R. Pélissier, C. Violle, and J.-B. Morel. Plant immunity: Good fences make good neighbors? *Current Opinion in Plant Biology*, 62:102045, Aug. 2021. ISSN 13695266. doi: 10.1016/j.pbi.2021.102045. URL https://linkinghub.elsevier.com/retrieve/pii/S1369526621000455.

E. Reekie and F. Bazzaz. *Reproductive Allocation in Plants*. Elsevier Science, 2011. ISBN 978-0-08-045433-7. URL http://www.totalboox.com/book/id-879441384717483270. OCLC: 968085321.

E. R. Reiss and L. E. Drinkwater. Cultivar mixtures: a meta-analysis of the effect of intraspecific diversity on crop yield. *Ecological Applications*, 28(1):62–77, Jan. 2018. ISSN 1051-0761, 1939-5582. doi: 10.1002/eap.1629. URL https://onlinelibrary.wiley.com/doi/10.1002/eap.1629.

K. J. Rothman. No adjustments are needed for multiple comparisons. *Epidemiology (Cambridge, Mass.)*, 1(1):43–46, Jan. 1990. ISSN 1044-3983.

V. O. Sadras and G. A. Slafer. Environmental modulation of yield components in cereals: Heritabilities reveal a hierarchy of phenotypic plasticities. *Field Crops Research*, 127:215–224, Feb. 2012. ISSN 03784290. doi: 10.1016/j.fcr.2011.11.014. URL https://linkinghub.elsevier.com/retrieve/pii/S0378429011003807.

J. Schmitt. REACTION NORMS OF MORPHOLOGICAL AND LIFE-HISTORY TRAITS TO LIGHT AVAILABILITY IN IMPATIENS CAPENSIS. *Evolution*, 47

(6):1654–1668, Dec. 1993. ISSN 0014-3820, 1558-5646. doi: 10.1111/j.1558-5646.1993.tb01258.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1558-5646.1993.tb01258.x.

L. Song, F. Li, X. Fan, Y. Xiong, W. Wang, X. Wu, and N. Turner. Soil water availability and plant competition affect the yield of spring wheat. *European Journal of Agronomy*, 31(1):51–60, July 2009. ISSN 11610301. doi: 10.1016/j.eja.2009.03.003. URL https://linkinghub.elsevier.com/retrieve/pii/S1161030109000264.

S. E. Sultan. Phenotypic plasticity in plants: a case study in ecological development. *Evolution and Development*, 5(1):25–33, Jan. 2003. ISSN 1520-541X, 1525-142X. doi: 10.1046/j.1525-142X.2003.03005.x. URL http://doi.wiley.com/10.1046/j.1525-142X.2003.03005.x.

O. Therond, M. Duru, J. Roger-Estrade, and G. Richard. A new analytical framework of farming system and agriculture model diversities. A review. *Agronomy for Sustainable Development*, 37(3), June 2017. ISSN 1774-0746, 1773-0155. doi: 10.1007/s13593-017-0429-7. URL http://link.springer.com/10.1007/s13593-017-0429-7.

R. L. Thompson, L. Lassaletta, P. K. Patra, C. Wilson, K. C. Wells, A. Gressent, E. N. Koffi, M. P. Chipperfield, W. Winiwarter, E. A. Davidson, H. Tian, and J. G. Canadell. Acceleration of global N2O emissions seen from two decades of atmospheric inversion. *Nature Climate Change*, 9(12):993–998, Dec. 2019. ISSN 1758-678X, 1758-6798. doi: 10.1038/s41558-019-0613-7. URL http://www.nature.com/articles/s41558-019-0613-7.

G. Touzy, R. Rincent, M. Bogard, S. Lafarge, P. Dubreuil, A. Mini, J.-C. Deswarte, K. Beauchêne, J. Le Gouis, and S. Praud. Using environmental clustering to identify specific drought tolerance QTLs in bread wheat (T. aestivum L.). *Theoretical and Applied Genetics*, 132(10):2859–2880, Oct. 2019. ISSN 0040-

5752, 1432-2242. doi: 10.1007/s00122-019-03393-2. URL http://link.springer.com/10.1007/s00122-019-03393-2.

A. Weigelt and P. Jolliffe. Indices of plant competition: *Competition indices*. *Journal of Ecology*, 91(5):707–720, Oct. 2003. ISSN 00220477. doi: 10.1046/j.1365-2745.2003.00805.x. URL http://doi.wiley.com/10.1046/j.1365-2745.2003.00805.x.

J. Weiner. Allocation, plasticity and allometry in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, 6(4):207–215, Jan. 2004. ISSN 14338319. doi: 10.1078/1433-8319-00083. URL https://linkinghub.elsevier.com/retrieve/pii/S1433831904700773.

J. Weiner, Y.-L. Du, C. Zhang, X.-L. Qin, and F.-M. Li. Evolutionary agroecology: individual fitness and population yield in wheat (Triticum aestivum). *Ecology*, 98(9):2261–2266, Sept. 2017. ISSN 00129658. doi: 10.1002/ecy.1934. URL http://doi.wiley.com/10.1002/ecy.1934.

H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. ISBN 978-0-387-98140-6. URL http://ggplot2.org.

M. S. Wolfe. The Current Status and Prospects of Multiline Cultivars and Variety Mixtures for Disease Resistance. *Annual Review of Phytopathology*, 23 (1):251–273, Sept. 1985. ISSN 0066-4286, 1545-2107. doi: 10.1146/annurev.py.23.090185.001343. URL https://www.annualreviews.org/doi/10.1146/annurev.py.23.090185.001343.

Q. Xie, S. Mayes, and D. L. Sparkes. Optimizing tiller production and survival for grain yield improvement in a bread wheat × spelt mapping population. *Annals of Botany*, 117(1):51–66, Jan. 2016. ISSN 0305-7364, 1095-8290. doi: 10.1093/aob/mcv147. URL https://academic.oup.com/aob/article-lookup/doi/10.1093/aob/mcv147.