Annual dynamics of *Zymoseptoria tritici* populations in wheat cultivar mixtures: a compromise between the efficiency and durability of a recently broken-down resistance gene?

Carolina Orellana-Torrejon [1,2], Tiphaine Vidal [1], Anne-Lise Boixel [1], Sandrine Gélisse [1], Sébastien Saint-Jean [2], Frédéric Suffert [1 *]

[1] Université Paris-Saclay, INRAE, AgroParisTech, UMR BIOGER, 78850 Thiverval-Grignon, France

[2] Université Paris-Saclay, INRAE, AgroParisTech, UMR ECOSYS, 78850 Thiverval-Grignon, France

*corresponding author: frederic.suffert@inrae.fr

Abstract

Cultivar mixtures slow polycyclic epidemics but may also modify the evolution of pathogen populations by diversifying the selection pressures exerted by their plant hosts at field scale. We compared the dynamics of natural populations of the fungal pathogen *Zymoseptoria tritici* in pure stands and in three binary mixtures of wheat cultivars (one susceptible cultivar and one cultivar carrying the recently broken-down *Stb16q* gene) over two annual field epidemics. We combined analyses of population 'size' based on disease severity, and of population 'composition' based on assessments of changes in the frequency of virulence against *Stb16q* in seedling assays with more than 3000 strains. In the field, disease levels were lower in mixtures, with each cultivar providing the other with reciprocal protection. The three cultivar proportions in the mixtures (0.25, 0.5 and 0.75) modulated the decrease in (i) the size of the pathogen population relative to the two pure stands, (ii) the size of the virulent subpopulation, and (iii) the frequency of virulence relative to the pure stand of the cultivar carrying *Stb16q*. Our findings suggest that optimal proportions may differ slightly between the three indicators considered. We identified potential trade-offs that should be taken into account when deploying a resistance gene in cultivar mixtures: between the dual objectives 'efficacy' and

‘durability’, and between the ‘size’ and ‘frequency’ of the virulent subpopulation. Based on current knowledge, it remains unclear whether virulent subpopulation size or frequency has the largest influence on interepidemic virulence transmission.

Keywords

Cultivar mixture, resistance breakdown, *Stb16q*, virulence, wheat disease epidemics, *Zymoseptoria tritici.*

Introduction

The industrialisation of agrosystems has greatly simplified the implementation of crop protection. In conventional cereal farming, fungal disease control is based on two major pillars: phytosanitary products and disease-tolerant or resistant cultivars. Most fields are planted with a single cultivar, in which effective resistance usually depends on the combination of a few qualitative resistance genes rather than diversified quantitative genetic components. The strong resistance conferred by a single gene is simpler for breeders to handle, but may be easily overcome by pathogens (McDonald & Linde, 2002).

The number of wheat cultivars grown in North-Western Europe is small (Perronne *et al.*, 2017). This leads to the exertion of unidirectional selection pressures on plant pathogen populations, fostering the breakdown of qualitative resistances, and resulting in an increase in the frequency of virulent strains just after their emergence (Brown & Tellier, 2011). Host homogeneity thus limits the 'durability' of qualitative resistance, defined as its efficacy over a prolonged period of widespread use under conditions conducive to the disease (Johnson, 1984). Pathogen evolution, which occurs at a variable rate, usually results in the replacement of the 'oldest' cultivars grown with others carrying 'new' sources of resistance. These dynamics have been widely described in publications on wheat diseases as a response to the common 'boom-and-bust' cycles observed in clonal pathogen populations (e.g. McIntosh & Brown, 1997) and to the more gradual breakdown of resistance in diversified sexual populations (e.g. Cowger *et al.*, 2000). In this context, one of the most striking issues currently facing us is the need to reconcile the efficiency and durability of the resistances available to farmers, by identifying optimal strategies for their deployment over different space-time scales (Fabre *et al.*, 2015).

Increasing crop diversity at field scale has been widely reported to result in more efficient control over plant disease epidemics. This is the case for cultivar mixtures (Wolfe, 1985), in which cultivars of the same species with complementary traits, including those relating to plant immunity, are grown together. Cultivar mixtures have long been considered particularly effective against airborne fungal pathogens (Cowger & Mundt, 2002a; Borg *et al.*, 2018) and were widely used at the end of the last century in some European countries. The use of wheat

mixtures has increased in Europe in recent years, reflecting a growing interest in the beneficial effects of mixtures. In France, for instance, the percentage of the national area occupied by mixtures of bread wheat cultivars has doubled over the last four years (4.8% in 2017, 8.6% in 2018, 11.9% in 2019, and 12.2% in 2020; FranceAgriMer). The factors affecting the efficacy of cultivar mixtures against splash-dispersed diseases, such as Septoria tritici blotch (STB) on bread wheat, have also been investigated, with a combination of experimental and modelling approaches (Gigot *et al.*, 2013; Vidal *et al.*, 2017; Kristoffersen *et al.*, 2019). The proportions of susceptible and resistant cultivars in the mixture are known to have a determinant effect on disease reduction (Finckh *et al.*, 2000; Gigot *et al.*, 2014; Borg *et al.*, 2018). Nevertheless, most of the experimental approaches focused on the impact of mixtures on the overall canopy, with the aim of using the most resistant plants to 'protect' the susceptible plants, and little is known about the consequences for the most resistant cultivar. The value of cultivar mixtures for controlling diseases lies in the complementarity of the cultivar phenotypes, providing a guarantee of efficacy. However, cultivar mixtures may modify the evolution of pathogen populations, by diversifying the selection pressures exerted by their plant hosts at field scale, potentially with different effects on the different cultivars in the mixture. Little is known about pathogen evolution (changes in virulence gene frequencies) in cultivar mixtures, particularly during the early stages of resistance gene breakdown (i.e., when the frequency of virulent strains remains low).

The heterothallic ascomycete fungus *Zymoseptoria tritici*, which causes STB, is responsible for yield losses averaging 20% on susceptible bread wheat cultivars in North-Western Europe (Fones & Gurr, 2015). The development of epidemics is dependent on both splash-dispersed asexual pycnidiospores during the wheat growing period and wind-dispersed ascospores from sexual reproduction on crop residues (Figure 1a; Suffert *et al.*, 2011). The potential intra-annual selective dynamics and strong genetic recombination within *Z. tritici* populations make STB a good candidate model for investigating the issues described above. Cultivars carrying qualitative and quantitative resistances have been developed and used to control STB, leading to an overall increase in the mean resistance of commercial wheat cultivars registered in France in the last decade (from 5.2 in 2015 to 6.0 in 2019, on a scale of 1 to 9, with 9 corresponding to the most resistant cultivar; ARVALIS-Institut du Végétal/CTPS). In total, 21 *Stb* genes and 89 QTLs have been identified and mapped (Brown *et al.*, 2015). *Stb16q*, one of

the most recent major resistance genes to have been incorporated into several cultivars by breeders, initially provided high levels of resistance against STB in both seedlings and adult plants (Ghaffary *et al.*, 2012). *Stb16q* is present in 13 wheat cultivars registered in France between 2012 and 2018, including cv. Cellule (Florimond Desprez, France). Cellule was widely used in France, becoming the third most frequent cultivar grown (8% of the area under wheat in France in 2016), but its use then declined following a decrease in its resistance to STB. Indeed, STB symptoms were first observed on Cellule in 2015 in France (Romain Valade, ARVALIS-Institut du Végétal, pers. com.) and in 2019 in Ireland (Kildea *et al.*, 2020), highlighting the beginning of a breakdown of *Stb16q* resistance.

In this epidemiological study, our goal was to investigate the capacity of mixtures to reduce disease, as a short-term epidemiological outcome, and to interfere with the resistance breakdown process, as an evolutionary outcome. In field experiments, we analysed the effect of the proportion of wheat plants carrying the qualitative resistance gene *Stb16q* within a binary wheat mixture on the size (disease severity) and composition (frequency of virulent strains) of *Z. tritici* pathogen populations during an annual epidemic.

Materials and methods

Overall strategy

We compared the two-year dynamics of natural *Z. tritici* populations in pure stands and in binary mixtures of wheat cultivars, from the early to late stages of a field epidemic (Figure 1a). We combined analyses of population 'size' based on disease severity in the field as a proxy, and of population 'composition' based on estimations of the proportion of virulence against *Stb16q* obtained by screening several thousands of field-collected strains in seedling assays. To this end we developed and applied an *in planta* 'population-phenotyping' strategy in a greenhouse. We then reinforced this strategy by 'confirmatory individual phenotyping' in a growth chamber, to increase the certainty of characterisation for some of the strains.

Cultivars studied

This study focused on two bread wheat (*Triticum aestivum*) cultivars: Apache ('A'; Limagrain, France) was chosen for its moderate susceptibility to STB (rated 4.5, stable over time, on a scale of 1 to 9, where 9 corresponds to the most resistant cultivar, ARVALIS-Institut du Végétal/CTPS), and Cellule ('C'; Florimond Desprez, France), which carries the *Stb16q* resistance gene (rated 7 in 2016, vs. 5 in 2020, ARVALIS-Institut du Végétal/CTPS). Apache was the predominant cultivar in France from 2000 to 2010, accounting for 24% of the area under wheat in 2003, subsequently falling to less than 5% after 2015 (Figure 2a). Cellule, registered in 2012, became the third most frequently used French cultivar, but the area under this cultivar rapidly decreased after 2016 (Figure 2b). Both cultivars were well represented in the Parisian basin and, thus, probably played a significant role in the evolutionary trajectory of the local pathogen populations collected in 2018 and 2019.

Design of the field experiment

We carried out a field experiment at the INRAE experimental station at Thiverval-Grignon (France) in the 2017-2018 and 2018-2019 seasons. A field plot (Figure S1 and S2) was divided into three blocks (a, b and c), each consisting in five microplots. These five microplots corresponded to five randomly distributed treatments (individual cultivars or cultivar mixtures): a pure stand of wheat cv. Cellule (C_1-A_0), a pure stand of wheat cv. Apache (C_0-A_1), and three binary mixtures, with sowing proportions of Cellule (p_C) of 0.25, 0.5 and 0.75 ($C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$). The wheat was sown on October 17, 2017 and November 15, 2018, at a density of 220 seeds·m^{-2}. The actual proportions of wheat axis in each microplot, estimated late in the growing season, close to heading, was similar to the proportions of seed sown (Table S1) except for a few microplots. The seeds were treated with the insecticide tefluthrin at a dose of 0.2 g·kg^{-1} (Attack®, Syngenta France SAS) before sowing, and a dye was included in the seed coating to facilitate differentiation between the two cultivars (red for Cellule, green for Apache; Figure S2b) at the seedling stage. Cultivar identification was made later in the season, according to the bearded (Cellule) or unbearded (Apache) nature of the inflorescence (Figure S2e). Nitrogen fertiliser applications were adjusted by the balance-sheet method with a target yield of 7 t·ha^{-1}. A single herbicide treatment was applied in the early spring. No fungicide treatment was applied.

Characterisation of the pathogen populations size

We used disease severity (i.e., diseased leaf area) as a proxy for the size of the pathogen population (i.e., the number of asexual offspring), assuming the two traits to be correlated (Suffert *et al.*, 2013). Disease severity was estimated as the percentage sporulating area (i.e., the percentage of the leaf area covered by *Z. tritici* pycnidia, scored 0, 1%, 3%, 5%, 10% and then increments of 5% up to 100%). It was assessed visually early in the epidemic (time t1; Zadoks stage 15), before the expression of most of the secondary infections, and then late in the epidemic (time t2; Zadoks stage 70), just after the disease had reached the upper leaves (Figure S2). The t1 assessments of disease severity were performed on various dates (from March 5 to March 12 in 2018 and February 27 to March 18 in 2019), due to experimental constraints. We collected 10 seedlings of each cultivar from each microplot, on which disease severity was scored for the five lowest (first-formed) leaves. At t2, similar disease severity assessments were performed for the five upper (most recently formed) leaves (L1 to L5 from the top to the bottom of the plant) of 10 plants of each cultivar per microplot, on May 11, 2018 and June 18, 2019. The optimal scoring window varies with epidemic dynamic. Scoring was therefore performed earlier in 2018 (intense early epidemic) than in 2019 (moderately intense late epidemic). The flag leaves (L1) were assessed a second time, in all microplots, on June 15, 2018 and June 26, 2019.

A disease severity index (SEV) was calculated at the plant scale. At t1, SEV was calculated as the mean sporulating area over all non-senescent leaves (generally from two to three leaves). At t2, SEV was calculated as the mean sporulating area of representative leaf layers (L4 and L3 in 2018, and L3 and L2 in 2019), excluding those with a mean severity below 1% or completely senescent on more than two microplots.

A mean disease severity index (SEV_v) was then calculated for each cultivar *v* for each treatment (cultivar proportion). An overall mean disease severity index ($\overline{\mathrm{SEV}}$) was also calculated at the whole-canopy scale for each treatment, using SEV_v weighted by the proportions of the cultivars (p_C for Cellule and $1 - p_C$ for Apache), as follows:

$$\overline{\mathrm{SEV}} = p_C \times \mathrm{SEV}_C + (1 - p_C) \times \mathrm{SEV}_A \qquad \text{eq. 1}$$

We assessed the effect of treatment on SEV at t1 and t2, and analysed changes in this effect over the course of each epidemic. To this end, SEV was transformed by ordered quantile normalisation (*bestNormalize* R package). Its variance was partitioned into sources attributable to the following factors and their interactions, in pairs: cultivars (A; C), treatments (C_0-A_1; $C_{0.25}$-$A_{0.75}$; $C_{0.5}$-$A_{0.5}$; $C_{0.75}$-$A_{0.25}$; C_1-A_0), blocks (a; b; c), epidemic stages (t1; t2) and years (2018; 2019), according to the generalised linear model. The model was adjusted by removing non-significant interactions one-by-one. For SEV_v and $\overline{\mathrm{SEV}}$, standard errors were corrected by taking into account the error propagation of formulas.

The effect of treatment on SEV_A and SEV_C at each stage of each year was first assessed with a Tukey HSD post-hoc test for multiple pairwise comparisons on estimated marginal means (*emmeans* and *multcomp* packages of R). The variance of the overall disease severity index $\overline{\mathrm{SEV}}$ was then estimated by calculating the standard error. The effect of cultivar mixtures on $\overline{\mathrm{SEV}}$ was evaluated by plotting the 'theoretical' $\overline{\mathrm{SEV}}$ (assuming an absence of mixture effect) as a linear function of the severity in each pure stand weighted by cultivar proportions (p_C). All data analyses were performed with R software (v4.0.2 R Core Team 2012).

Characterisation of the pathogen populations composition

Sampling of strains in the field — Strains of *Z. tritici* were collected at t1 and t2 in 2018 and 2019, to constitute groups of 30 strains per cultivar and per microplot (total of 3060 strains; Table 1). The sampling effort was doubled in the pure stands of the susceptible cultivar at t1 in 2018 and 2019 (2 x 180 strains). The population present on cv. Apache early in the epidemic (t1), which may be considered a representative sample of the primary inoculum population, was used to estimate the initial (t0) frequency of virulent strains ($f_{t0,vir}$; Figure 1).

We ensured that 30 strains were obtained per cultivar from each microplot, by sampling the 40 lowest (first) leaves of seedlings at t1 and the flag leaves of adult plants at t2. These samples were placed in a humid chamber to promote the exudation of cirrhi. One cirrhus per leaf was picked with a pin and streaked onto a Petri dish containing PDA (potato dextrose agar, 39 $g \cdot L^{-1}$), which was then incubated at 18°C for more than five days in the dark. A few blastospores from a single colony were then picked and spread on a new plate. This operation was repeated twice to obtain pure single-spore strains. Each strain was finally scraped off the plate, to obtain

a substantial mass of blastopores (150 mg), which was deposited at the bottom of cryotubes and stored at -80°C. At t1 in 2018, we sometimes had to collect more than one strain per leaf to counterbalance the low level of severity on certain microplots or the degradation/senescence of the diseased leaves. In both situations, we picked cirrhi from two to three distinct, distant lesions present on the same leaf.

Phenotyping strategy — We characterised the *Z. tritici* strains for 'virulence' (defined by Niks *et al.* (2011) as "the capacity of a pathogen to infect a plant with one or more major genes for resistance"; against *Stb16q* and their 'aggressiveness' (defined as the quantitative term of pathogenicity on a susceptible host; Lannou, 2012). We developed a 'population-phenotyping' method for use on cv. Cellule and cv. Apache seedlings in the greenhouse (Figure S3), to make it possible to reach our objective in a reasonable amount of time. This made it possible to characterise a large number of strains while accepting some imprecisions or individual errors quantified *a posteriori*. This approach was completed by 'confirmatory individual phenotyping' in growth chambers, in which we characterised several dozens of strains that were collected from sporulating lesions on Cellule in the field but appeared avirulent during population-phenotyping. This made it possible to determine the rate of type II error (false avirulence) for population-phenotyping.

Population-phenotyping — Population-phenotyping was performed from April 2018 to August 2020, by splitting the 3060 strains into 32 batches of 102 (t1) or 93 (t2) strains, which were simultaneously tested in the greenhouse (Figure 3). Each batch was balanced to contain similar numbers of strains from each cultivar, treatment, and block. Eight virulent strains were added to each batch as control strains. Wheat seeds of cv. Apache and Cellule were sown in 0.4-litre pots, to obtain three seedlings per pot. The pots were kept in the greenhouse under natural daylight supplemented with 400 W sodium vapour lamps to homogenise the lighting conditions. The temperature was kept below 20°C during the 15-h light period and above 15°C during the 9-h dark period. Relative humidity was kept at a mean value of 68 $\pm$ 9%. The thermal time at which disease severity was assessed was calculated for each batch (Figure S5).

We simplified the classical method based on accurate adjustment of a blastospore suspension, by considering that the differences in disease severity induced by the differences in concentration within the range 10^5 to 5.0 10^5 spores·mL^{-1} (see Figure 8 in Suffert *et al.*, 2018) were acceptable for our objective. We thus used non-adjusted suspensions in this range of

concentrations, by diluting the blastospores recovered directly with a 10 μL inoculating loop (about 10 mg in the half-filled central ring; Sarstedt, Nümbrecht, Germany) from the freshly thawed contents of cryotubes preserved at -80°C, in 30 mL of water. Two series of independent tests showed that this process provided a good compromise between robustness and repeatability (Figure S4). Two drops of surfactant (0.1% v/v of Tween 20; Sigma, France) were added to each suspension, to increase its adhesion to leaves. All suspensions were prepared the day before inoculation and stored at 5°C. A cotton bud was used to inoculate a 5 cm mid-length leaf section (by two back-and-forth passages over the leaf) of the second and third leaves from the bottom, on three 18-day-old seedlings (2 x 3 replicates). All pots were then enclosed in sealed transparent polyethylene bags moistened with distilled water for 72 h, to promote infection.

Disease severity was assessed visually 21 days post inoculation (mean thermal time 434.8°C·days for the 32 batches; Figure S5), as the percentage of the inoculated leaf surface covered by pycnidia (sporulating area assessed, from 0 to 100%, in increments of 5%). A strain was considered virulent ('vir') against *Stb16q* when pycnidia were observed (at least 5% of sporulating area, the minimum for visual quantification in our experimental conditions), on at least one of the six leaves of Cellule. The strains were otherwise considered avirulent ('avr'). The frequency of virulent strains $f_{t,v,vir}$ ($t \in [t1; t2]$, $v \in [A; C]$) was calculated for each cultivar in each treatment. The effect of the treatment on the number of virulent individuals was assessed with a χ^2 test, or with Fisher's exact test when the expected numbers were small, with Bonferroni correction for pairwise comparison. We analysed the changes in $f_{t,v,vir}$ from t1 to t2, assessing the year effect with a test of equal or given proportions (*stats* package of R). The initial frequency of virulent strains in the local pathogen population ($f_{t0,vir}$) in 2018 and 2019 was estimated by characterising the 180 strains from the pure stand of cv. Apache. We postulated that $f_{t1,A,vir} = f_{t0,vir}$ in C_0-A_1 (filter effect exerted on a strain according to its virulence status with respect to *Stb16q* considered negligible; Figure 1).

We calculated the aggressiveness (AG) and conditional aggressiveness (AGc), defined as the amount of disease "conditional on the individuals being infected" (McRoberts *et al.*, 2003). AG was estimated as the mean sporulating area on all leaves of the susceptible cultivar, whereas AGc was calculated on symptomatic leaves only. For a given strain, AGc was calculated under the condition that at least three of the six inoculated leaves presented

pycnidia; this metric was otherwise considered not to be relevant. We assessed the possible fitness cost of virulence against *Stb16q,* by evaluating the effect of virulence on aggressiveness in a generalised linear model adapted to a Poisson distribution, followed by a Tukey HSD post-hoc comparison (*car R package*). The variance of AG (and AGc) was partitioned into sources attributable to the following factors and their interactions, in pairs: the virulence status of the strain (vir; avir), treatment (C_0-A_1; $C_{0.25}$-$A_{0.75}$; $C_{0.5}$-$A_{0.5}$; $C_{0.75}$-$A_{0.25}$; C_1-A_0), the cultivar of origin (A; C), year (2018;2019), epidemic stage (t1; t2), block (a; b; c), and batch (1-32).

Confirmatory individual phenotyping — A second round of phenotyping was applied to 47 strains considered avirulent against *Stb16q* after the first round of population-phenotyping despite having been isolated from sporulating lesions on Cellule in 2018 (see the Results section). For this purpose, we used Cellule (C) and two isogenic lines of the wheat cultivar Chinese Spring without *Stb6*, one of which carried *Stb16q* (CS-*Stb16q*) whereas the other did not (CS). The strains were split into three batches and tested in a growth chamber (Figure S3). A virulent and an avirulent strain were added to each batch as controls. Inoculum suspensions were adjusted to a density of 5.0 10^5 spores·mL^{-1} with a Malassez cell. Leaves were inoculated by applying the suspension with a paintbrush, on the second leaf from the bottom of nine seedlings of each cultivar. The temperature in the growth chamber was kept below 20°C during the 15-h light period and above 15°C during the 9-h dark period. Relative humidity was maintained at a mean of 88.5 $\pm$ 1.2%. Disease severity was assessed visually 14, 20 and 26 dpi (day post inoculation), as the percentage sporulating area. The virulent, avirulent or non-pathogenic nature of each strain was determined, with a threshold at 10% disease severity for mean sporulating area. A strain was considered (i) virulent against *Stb16q* if the mean sporulating area was equal or higher on CS-*Stb16q* and Cellule, (ii) avirulent if it was equal or higher on CS only, or (iii) non-pathogenic if it was below the threshold of 10% of the three cultivars (Thierry Marcel, INRAE BIOGER, pers. com.; Figure S10). The mean aggressiveness of strains was compared between CS-*Stb16q* and Cellule with a Wilcoxon test. The effect of treatment on the number of avirulent strains was assessed with Fisher's exact test.

Overall impact of cultivar mixtures on the dynamics of virulent subpopulations

We calculated the percentage of the diseased leaf area occupied by virulent strains at t2 ($n_{t2,v,vir}$; v $\in$ [A; C]), as a proxy for the size of the virulent subpopulations responsible for the

last asexual disease cycle. This dimension-less variable was obtained by multiplying the disease severity index ($\mathrm{SEV_v}$) on the flag leaves of each cultivar in each microplot by the frequency of virulent strains estimated at t2 ($f_{\mathrm{t2,v,vir}}$). The mean size of the virulent subpopulations ($\overline{n}_{\mathrm{t2,vir}}$) and the mean frequency of virulent strains ($\overline{f}_{\mathrm{t2,vir}}$) were calculated at the whole-canopy scale for each treatment (cultivar proportion), as for overall disease severity index ($\overline{\mathrm{SEV}}$, eq. 1). The effect of treatment on $n_{\mathrm{t2,v,vir}}$ and $\overline{n}_{\mathrm{t2,vir}}$ was assessed in a Kruskal-Wallis test.

Results

Impact of cultivar mixtures on the dynamics of the epidemic and the overall size of pathogen populations

Disease was generally less severe in the mixtures than in the two pure stands, when the whole canopy was considered, and the protection of each cultivar by the other appeared to be reciprocal.

At t1 in 2018, no significant difference in mean disease severity index ($\mathrm{SEV_v}$) was observed between pure stands and mixtures, regardless of the cultivar considered, except for Apache in $C_{0.5}$-$A_{0.5}$ (p = 0.987, p = 0.042 and p = 0.155 for Apache and p = 0.172, p = 0.388 and p = 0.065 for Cellule in $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, $C_{0.75}$-$A_{0.25}$ respectively; Figure 4). Apache was more attacked than Cellule, consistent with their resistance ratings: $\mathrm{SEV_A}$ was scored at 39% for the pure stands, 37% in $C_{0.25}$-$A_{0.75}$, 29% in $C_{0.5}$-$A_{0.5}$ and 32% in $C_{0.75}$-$A_{0.25}$; $\mathrm{SEV_C}$ was lower, at 11% for the pure stands, 7% for $C_{0.25}$-$A_{0.75}$, and 10% for $C_{0.5}$-$A_{0.5}$, 8% for $C_{0.75}$-$A_{0.25}$. The modelling results showed that the variance of disease severity index (SEV) was not attributable to the block factor (p = 0.090). Overall disease severity index $\overline{\mathrm{SEV}}$ increased with the proportion of Apache. Mean $\overline{\mathrm{SEV}}$ averages of 30%, 19% and 14% were obtained for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively. At t1 in 2019, $\mathrm{SEV_A}$ was scored at 18% for $C_{0.25}$-$A_{0.75}$, 13% for $C_{0.5}$-$A_{0.5}$ and 13% for $C_{0.75}$-$A_{0.25}$. $\mathrm{SEV_C}$ was lower than in 2018, at 6% for $C_{0.25}$-$A_{0.75}$, 6% for $C_{0.5}$-$A_{0.5}$ and 4% for $C_{0.75}$-$A_{0.25}$. $\overline{\mathrm{SEV}}$ increased with the proportion of Apache (6%, 10% and 15% for $C_{0.75}$-$A_{0.25}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.25}$-$A_{0.75}$, respectively).

At t2 in 2018, $\mathrm{SEV_v}$ differed significantly between pure stands and mixtures for Cellule, but not for Apache (p = 0.397, p = 0.148 and p = 0.357 for Apache and p < 0.001, p < 0.001 and p = 0.089 for Cellule in $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, $C_{0.75}$-$A_{0.25}$, respectively; Figure 4). In pure stands, Apache ($\mathrm{SEV_A}$ scored at 23%) was more attacked than Cellule ($\mathrm{SEV_C}$ scored at 16%). $\mathrm{SEV_A}$ and, to a greater extent, $\mathrm{SEV_C}$, were lower in mixtures than in pure stands (20%, 17% and 18% on Apache and 5%, 7% and 13% on Cellule for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively). At the whole-canopy scale, $\overline{\mathrm{SEV}}$ was lower in the three mixtures than in pure stands (17%, 12% and 14% for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively). At t2 in 2019, $\mathrm{SEV_v}$ differed significantly between pure stands and mixtures for both cultivars, this difference being more significant than in 2018 (p < 0.0001 for Apache and p < 0.0001 for Cellule in all $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, $C_{0.75}$-$A_{0.25}$). In pure stands, Apache ($\mathrm{SEV_A}$ scored at 25%) was more attacked than Cellule ($\mathrm{SEV_C}$ scored at 10%). The decreases in $\mathrm{SEV_A}$, $\mathrm{SEV_C}$, and $\overline{\mathrm{SEV}}$ in mixtures relative to pure stands was more pronounced in 2019 than in 2018.

The overall effect of treatment was illustrated by plotting a theoretical $\overline{\mathrm{SEV}}$ as a linear function of the severity in each pure stand weighted by p_C (Figure 4), which accounts only for "dilution" in mixtures. For the two years considered, the actual and theoretical $\overline{\mathrm{SEV}}$ curves were close to each other at t1, but not at t2. This highlights the overall absence of a mixture effect on disease dynamics at t1 and the strong mixture effect at t2. The efficacy of the various treatments relative to pure stands was 19%, 40%, 22% in 2018 and 67%, 71%, 69% in 2019, for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively.

Impact of mixtures on the composition of pathogen populations

In total, 3035 of the 3060 *Z. tritici* strains sampled during the two years were characterised for virulence and aggressiveness, by inoculating 36420 leaves of wheat seedlings in the greenhouse. The cultivar mixtures modulated the frequencies of virulent and avirulent strains from the early (t1) to late (t2) stages in the two epidemic years.

The frequency of virulence at t1 ($f_{\mathrm{t1,A,vir}}$) was estimated at 13% in 2018 (23 virulent strains among the 179 collected on Apache in pure stands) and 14% in 2019 (26/180) (Figure 5). These frequencies are not significantly different (χ^2 test, p = 0.791), and the two years may therefore be considered as 'replicates'. The frequency of virulent strains on Apache did not differ

significantly between treatments at t1 (χ^2 test, p = 0.128 in 2018 and p = 0.290 in 2019), confirming the absence of a mixture effect early in the epidemic. The frequency of virulent strains on Apache in pure stands at t2 ($f_{t2,A,vir}$) was 8% in 2018 (7/90) and 18% in 2019 (16/90), and was not significantly lower than that at t1 (test of equal or given proportions, p = 0.244). Importantly, in mixtures, $f_{t2,A,vir}$ increased with p_C (23%, 54% and 78% in 2018 and 30%, 43% and 53% in 2019, for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively; χ^2 test, $p < 0.001$).

Very few strains were classified as non-pathogenic (0.3% i.e., 5/1525 in 2018 and 0.1% i.e., 2/1510 in 2019). Most of the strains (>95%) induced pycnidia on Apache (Figure S6 and S7), with no significant difference between treatments: for strains collected on Cellule p = 0.286 for t1 and p = 0.573 for t2 in 2018, p = 0.812 for t1 and p = 0.091 for t2 in 2019 (Fisher test); for strains collected on Apache p = 0.093 for t1 and p = 0.078 for t2 in 2018, p = 0.459 for t1 and p = 0.984 for t2 in 2019 (Fisher test or χ^2 test). The proportion of strains that did not induce pycnidia on Apache was low: 4% (31/719) and 3% (23/710) for those collected on Cellule in 2018 and 2019, respectively, and 6% (50/806) and 5% (38/800) for those collected on Apache in 2018 and in 2019, respectively. Similarly, 7% of the strains (52/719) collected on Cellule in 2018 and 10% (74/710) of those collected on this cultivar in 2019 (Figure 6) did not induce pycnidia on Cellule. The proportion of these strains considered avirulent after population-phenotyping did not differ significantly between treatments. Only 2% of the strains (27/1525 in 2018 and 20/1510 in 2019, distributed across all treatments) induced pycnidia on Cellule but not on Apache. The genetic backgrounds of the two cultivars differ, but these differences had a limited impact on the analysis of strain virulence status.

Confirmatory individual phenotyping showed that 51% (24/47) of the strains tested were truly avirulent, 34% (16/47) were virulent, and the results were dubious for 15% (7/47). The percentage of 'false avirulent' strains in the overall pathogen population collected on Cellule in 2018 was thus estimated at a maximum of 3% (23/719). All the avirulent strains among the 47 tested by confirmatory individual phenotyping were collected at t1 (Table S2). Virulent strains were significantly more aggressive on CS-*Stb16q* than on Cellule (67.3% and 58.2%, respectively; Wilcoxon test, $p < 0.0001$; Figure S9). Five of the strains yielding dubious results that induced pycnidia on CS-*Stb16q* were unable to infect Cellule, probably due to another source of resistance not present in CS. Similarly, two strains that induced pycnidia on Cellule

were unable to infect CS-*Stb16q*. All 47 strains induced pycnidia on CS, confirming their pathogenicity.

No significant difference in aggressiveness (neither in AG nor AGc; p = 0.054 and p = 0.597, respectively; Figure S8) was detected between the avirulent and virulent strains that were pathogenic on Apache (2886/3035).

Impact of mixtures on the dynamics of the virulent subpopulation

At whole-canopy scale, cultivar proportions modulated the decreases in the size of the virulent subpopulation and in the frequency of virulent strains relative to the pure stand of the cultivar carrying *Stb16q* (Cellule).

In 2018, the mean size of the virulent subpopulation ($\overline{n}_{t2,vir}$) was significantly smaller for two mixtures, $C_{0.25}$-$A_{0.75}$ and $C_{0.5}$-$A_{0.5}$ (1.1% and 1.5%, respectively; 2.0% in $C_{0.75}$-$A_{0.25}$), than for pure stands of Cellule (6.5% in C_1-A_0) (p = 0.165, Kruskal-Wallis test; Figure 7a), whereas no difference was observed between the mixtures and the pure stand of Apache (1.5% in C_0-A_1). In 2019, the three mixtures significantly decreased $\overline{n}_{t2,vir}$ relative to the pure stand of Cellule C_1-A_0 (p = 0.033, Kruskal-Wallis test; Figure 7b), whereas the decrease relative to the pure stand of Apache C_1-A_0 was significant only for the $C_{0.25}$-$A_{0.75}$ mixture. In 2019, $\overline{n}_{t2,vir}$ in mixtures (1.0%, 1.7%, 1.8% in $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively) was lower than that in pure stand of Cellule by a factor of about 10 (14.8% in C_1-A_0) and lower than that in the pure stand of Apache by a factor of four (6.4% in C_0-A_1). This difference can be explained by the very low frequency of virulent strains on Apache ($f_{t2,A,vir}$) being compensated in 2018, and, to a greater extent in 2019, by the size of the pathogen population on Apache (SEV_A), resulting in a larger number of virulent strains ($\overline{n}_{t2,vir}$). A dissociated analysis of each varietal component of the mixture can provide insight on this point. On Apache, the size of the virulent subpopulation ($n_{t2,A,vir}$) was not significantly different between pure stands and mixtures in 2018 (p = 0.200, Kruskal-Wallis test; 1.5%, 2.9%, 4.5% and 5.3% in C_0-A_1, $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively; Figure S11a) or 2019 (p = 0.578, Kruskal-Wallis test; 6.4%, 2.3%, 3.1% and 2.0% in C_0-A_1, $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively; Figure S11c). On Cellule, the size of the virulent subpopulation ($n_{t2,C,vir}$) increased significantly with p_C in 2018 (p = 0.033; 0.4%, 1.3%, 3.5% and 6.5%,) and 2019 (p = 0.024; 1.3%, 3.8%, 4.3% and 14.8% in $C_{0.25}$-

$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, $C_{0.75}$-$A_{0.25}$ and C_1-A_0, respectively). Thus only $n_{t2,C,vir}$ contributes to the differences at whole-canopy scale.

We investigated two potential trade-offs. We first focused on the relationship between the overall disease severity ($\overline{\mathrm{SEV}}$) — an indicator of the efficacy of the mixtures — and the overall size of the virulent subpopulation ($\overline{n}_{t2,vir}$), an indicator of *Stb16q* breakdown dynamics and, thus, a proxy for the impact of the mixtures on resistance durability (Figure 7c and 7d). In both years, the overall trend was similar for the virulent subpopulation: $\overline{n}_{t2,vir}$ increased with p_C (or remained stable) whereas $\overline{\mathrm{SEV}}$ decreased. In 2018, two mixtures ($C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$) displayed decreases for both $\overline{n}_{t2,vir}$ and $\overline{\mathrm{SEV}}$ relative to the pure stand of Cellule (C_1-A_0) and the pure stand of Apache (C_0-A_1). In the $C_{0.25}$-$A_{0.75}$ mixture, $\overline{\mathrm{SEV}}$ was slightly higher for the same level of $\overline{n}_{t2,vir}$ than in other mixtures. The impact of mixtures was greater in 2019, with all mixtures decreasing both $\overline{n}_{t2,vir}$ and $\overline{\mathrm{SEV}}$, which was, on average, lower (3%, 3% and 2% in $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively) than in pure stands of Apache by a factor of ten (33% in C_0-A_1) and lower than that in pure stands of Cellule by a factor of five (16% in C_1-A_0).

We then focused on the relationship between the mean frequency of virulent strains ($\overline{f}_{t2,vir}$) and the overall size of the virulent subpopulation ($\overline{n}_{t2,vir}$) (Figure 7e and 7f). The quantitative impact of p_C depended on whether we focused on $\overline{n}_{t2,vir}$ or on $\overline{f}_{t2,vir}$: it was generally high for $\overline{f}_{t2,vir}$ but relatively limited for $\overline{n}_{t2,vir}$, especially in 2018 for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$. Finally, regardless of the indicator considered ($\overline{\mathrm{SEV}}$, $\overline{n}_{t2,vir}$ or $\overline{f}_{t2,vir}$) and of relationship trends (see the empirically drawn grey dashed lines in Figure 7), the optimal p_C appeared to be about 0.5. However, our results suggest that moving away from this optimal proportion would have different effects on the three indicators. In other words, there may be a trade-off between the two objectives of the mixture, efficacy, and durability.

Discussion

We found that pathogen population size, estimated on each cultivar but also at the whole-canopy scale, was smaller for mixtures than for pure stands over the course of an STB epidemic. In addition to this evidence of quantitative 'efficacy', we established that mixtures

reduced the size of the virulent subpopulation and modified the frequency of strains displaying virulence against *Stb16q*. Hereafter, we refer to Cellule as the 'resistant' cultivar and to Apache as the 'susceptible' cultivar.

Relevance of the phenotyping method

A method suitable for studies at population scale — Molecular tools are not always available for monitoring the emergence of virulence, and this may complicate the quantification of breakdown dynamics for recently deployed resistance genes. Our *in planta* 'population-phenotyping' method can be used to determine the virulence status of several thousand strains. This method was twice as efficient as the classical method, thus reducing the time required for phenotyping (Figure S3). The various experimental adjustments we made to inoculum preparation, climatic conditions in the greenhouse, visual assessments and the small number of biological replicates (Table S3) introduced some uncertainty, potentially leading to diagnostic errors concerning virulence status at an individual scale ('false avirulence' rate estimated at 3%). However, this uncertainty had only a small impact at population scale.

***Pathogenicity on Cellule vs. virulence against* Stb16q** — Throughout this study, we assumed that a strain pathogenic on cultivar Cellule was virulent against *Stb16q* and that strains non-pathogenic on this cultivar were avirulent. However, as the genetic background of the Cellule cultivar probably includes other sources of resistance, an absence of symptoms on Cellule is not sufficient to demonstrate avirulence against *Stb16q*. We could have resolved this ambiguity by using a cultivar carrying *Stb16q* with no other source of resistance, such as the Chinese Spring isogenic line CS-*Stb16q,* as in the confirmatory individual phenotyping. We did not generalise this approach, for two reasons: (i) seeds of this line were not yet available in sufficient quantities at the time at which the experiments were set up (Cyrille Saintenac, INRAE GDEC, pers. com.); (ii) the host-pathogen interactions tested in this way may diverge from those in the field, as the local pathogen populations would not have encountered the CS genetic background on their evolutionary trajectory. However, the confirmatory individual phenotyping data provide some perspective concerning this approximation: five of the seven strains inducing pycnidia on CS-*Stb16q* were unable to infect Cellule, and two of the strains pathogenic on Cellule were unable to develop sufficient pycnidia on CS-*Stb16q* to exceed the

10% threshold for virulence. The very low frequency of such situations (<1%) suggests that ability to induce pycnidia on Cellule is a reasonably good indicator of virulence against *Stb16q.* Finally, we estimated that the virulence status established based on population-phenotyping of the Cellule response was correct in 97% of cases.

Impact of cultivar mixtures

The cultivars displayed reciprocal protection — Cultivar mixtures reduced the size of the pathogen population in the whole canopy ($\overline{\mathrm{SEV}}$) late in the epidemic (Figure 4), confirming the potential of wheat mixtures to limit the dynamics of STB epidemics (Kristoffersen *et al.*, 2019). In our experiment, mixtures reduced $\overline{\mathrm{SEV}}$ by a mean of 27% in 2018 and 69% in 2019. The size of the total pathogen population was lower on both susceptible ($\mathrm{SEV_A}$) and resistant ($\mathrm{SEV_c}$) cultivars in mixtures than in pure stands. The relative differences in $\mathrm{SEV_v}$ between mixtures and pure stands ranged from 13% to 26% in 2018 and from 64% to 72% in 2019 for the susceptible cultivar (Apache), and from 19% to 69% in 2018 and from 70% to 80% in 2019 for the resistant cultivar. Gigot *et al.* (2013) and Vidal *et al.* (2017) highlighted similar reductions: 35% for the leaf area covered by pycnidia on the three top leaves of susceptible plants and 32% to 34% for the AUDPC, respectively. The resistant cultivar protected the susceptible cultivar by acting as an obstacle to spore dispersal (Cowger & Mundt, 2002a). This protective effect of the resistant cultivar on the susceptible cultivar has been reported before and was particularly strong in this study. Conversely, the presence of susceptible plants decreased the number of virulent strains filtered out by resistant plants relative to a pure stand of the resistant cultivar (Figure S11b and c). The susceptible plants did not concentrate more virulent strains in mixtures than in pure stands (Figure S11a and c), thereby decreasing the number of virulent strains at whole-canopy scale. This may explain the strong protective effect of the susceptible cultivar on the resistant cultivar, which is a crucial result. Most experimental studies have focused on overall disease severity at whole-canopy level (no distinction between cultivars), limiting the understanding of the processes involved. Studies focusing on resistant cultivars have come to divergent conclusions (Zhu *et al.*, 2005; Gigot *et al.*, 2013; Vidal *et al.*, 2017).

Estimation of the local frequency of virulent strains — We estimated the local frequency of virulent strains, using the pathogen population collected on pure stands of the susceptible

cultivar (Apache) early in the epidemic, which should reflect the founding ascospore population (i.e., the primary inoculum pool). The frequency of virulent strains was estimated at 13% in 2018 and 14% in 2019, with no significant difference between years. These results are consistent with the high levels of attacks on Cellule detected in the western central part of France in the last two years (Romain Valade, ARVALIS-Institut du Végétal, pers.com.) and they confirm that the breakdown of *Stb16q* resistance is well underway. This is the first estimation of virulence against *Stb16q*, and more widely, against a STB resistance gene within a local pathogen population.

Some of the strains collected from sporulating lesions on Cellule in the field early in the epidemic (i.e., on seedlings in winter) were avirulent. Two hypotheses, requiring further testing, may explain this: (i) the ontogeny or temperature-modulation of *Stb16q* expression in field conditions, and (ii) the possibility of an avirulent strain co-infecting a host plant by taking advantage of infection with a virulent strain ('stowaway effect'), following systemic induced susceptibility (Seybold *et al.*, 2020).

Changes in the frequency of virulent strains in mixtures — Mixtures modified the composition of the pathogen population, by affecting strain selection during the epidemic. The frequency of virulent strains collected on the susceptible cultivar did not differ between pure stands and mixtures (Figure 5) early in the epidemic, when secondary infections were not yet strongly affecting disease dynamics. The differences expressed later, after several cycles of secondary infection, suggest that the mixture effect results from the modulation of cross-contamination between cultivars driven by pycnidiospores splash dispersal (Figure 1d), consistent with previous findings (Gigot *et al.*, 2013; Vidal *et al.*, 2017). The frequency of virulent strains on the susceptible cultivar increased with the proportion of the resistant cultivar, leading to an overall increase in the frequency of virulent strains in the pathogen population at the whole-canopy scale. The overall effect would probably be different for mixtures of cultivars with different combinations of resistance genes. For determining the consequences of multiple virulence-resistance interactions, a parallel could be drawn with studies focusing on the impact of wheat mixtures on the structure of yellow rust and powdery mildew populations (Villaréal & Lannou, 2000; Lannou *et al.*, 2005) and on the emergence of multivirulent strains. The frequency of such strains was not necessarily significantly higher in mixtures than in pure stands. It has even been shown that the size of the subpopulation of multivirulent strains is

lower in mixtures even when the frequency of these strains is increasing in the population (Chin & Wolfe, 1984; Dileone & Mundt, 1994).

No evidence of a fitness cost associated with AvrStb16q virulence — In theory, and according to some experimental studies, there should be a 'fitness penalty' for pathogen populations with high levels of virulence after adaptation to hosts (Laine & Barrès, 2013). Our findings did not highlight such a cost of virulence against *Stb16q*. This may be because the experimental conditions resulted in small mean sporulating areas (not entirely suitable for such a test) or because this virulence has no cost. This lack of a fitness cost would be consistent with the findings of Cowger & Mundt (2002b), who observed no difference in aggressiveness between pathogen populations phenotyped on susceptible and moderately resistant cultivars carrying a quantitative resistance gene.

Modulation of mixture effects

By the proportion of cultivars — We showed that the proportions of the two cultivars modulated the overall size and composition of the pathogen population, consistent with previous findings (Finckh *et al.*, 2000). All the mixtures efficiently decreased the size of the pathogen population ($\overline{\mathrm{SEV}}$), but the decrease in severity relative to expected levels in the absence of a mixture effect (Figure 4) was greatest for $C_{0.5}$-$A_{0.5}$ in both 2018 (40%) and 2019 (71%). The optimal proportion of the resistant cultivar for reducing $\overline{\mathrm{SEV}}$ seems to lie between 0.5 and 0.75. This is consistent with simulations performed by Gigot *et al.* (2014), which showed the protection of susceptible plants by resistant plants to be optimal for mixtures with more resistant than susceptible plants. Similar conclusions were obtained for leaf rust (Dai *et al.*, 2012). By contrast, our results suggest that the optimal proportion of the resistant cultivar for reducing the size of the virulent subpopulation ($\overline{n}_{t2,vir}$) lies between 0.25 and 0.5 (Figure 7a and b).

By the intensity of the epidemic — The intensity of epidemics, driven by climatic conditions, may account for the observed differences in the magnitude of the mixture effect. The epidemic started earlier in 2018 (based on an empirical comparison of the intensity of early attacks: 39% in 2018 vs. 17% in 2019 for SEV_A in C_0-A_1; Figure 4) and this year was wetter and cooler than 2019. However, STB severity ended up being more intense on flag leaves in 2019

(19% in 2018 vs. 33% in 2019 for $\mathrm{SEV_A}$ in C_0-A_1; data not shown). This apparent contradiction can be explained by late spring rainfall (data not shown) at key moments for the dynamics of the epidemic and by the primary inoculum only rarely being a limiting factor at the start of epidemics (Morais *et al.*, 2016). Mixtures decreased the size of the total pathogen population more effectively in 2019. Mixtures also had a greater impact on the frequency of virulent strains on the susceptible cultivar relative to the pure stand (Figure 5) in 2019 than in 2018, whereas the opposite pattern might have been expected, due to more intense disease dynamics late in the epidemic (Mundt, 2002). Gigot *et al.* (2013) showed that mixtures significantly decreased disease severity in most cases, regardless of disease intensity, whereas Kristoffersen *et al.* (2019) found no correlation with mixture efficacy.

A trade-off between disease control and resistance durability?

The choice of indicator and reference (pure stand of a susceptible or resistance cultivar) is crucial when assessing the impact of mixtures. Our findings suggest that the optimal proportions of cultivars for reducing the size of the virulent subpopulation ($\overline{n}_{\mathrm{t2,vir}}$) may differ slightly from those for reducing the size of the total pathogen population ($\overline{\mathrm{SEV}}$), and those for reducing the overall frequency of virulent strains ($\overline{f}_{\mathrm{t2,vir}}$) (Figure 7). The potential differences in the optima for these three indicators may conceal a trade-off between disease control and resistance durability. Our data confirm that susceptible cultivars can, indeed, be protected by including a resistant cultivar in the host population (one of the historical objectives of mixtures). However, they also suggest that this effect may be counterbalanced by an acceleration of the breakdown dynamics of this resistance, due to an increase in the frequency of virulent strains. The conclusion would clearly be different if a pure stand of the resistant cultivar was used as the reference: the inclusion of susceptible plants decreases the selection pressure exerted by emerging virulent strains, but also, as shown here, protects the resistant plants by decreasing disease severity on these plants. The pathogen population on the susceptible plants of a cultivar mixture may be smaller than that in a pure stand or in a mixture with a higher proportion of the susceptible cultivar, without there necessarily being an impact on the number of virulent strains (Figure S11).

Moreover, there is no reason why the optimal proportion of cultivars minimising the size of the virulent subpopulation should be the same as that minimising the frequency of virulent strains. The 'size' vs. 'frequency' debate is a classical issue in ecology and evolution that should not be neglected here, because demographic processes can act on gene frequencies and population size in different ways, with consequences for genetic diversity and evolution (Brown & Tellier, 2011).

Perspectives

Determination of the effects of mixtures on the durability of the resistances deployed requires a consideration of the complete disease cycle, including sexual reproduction, and an extension of this approach to time scales beyond a single annual epidemic. One of the major issues arising from this study and previous studies (e.g., Finckh *et al.*, 2000; Mundt, 2002; Lannou *et al.*, 2005) is whether a cropping system based on cultivar mixtures would be likely to increase the prevalence of virulent strains after several years. This is both a question of the 'reference' used (pure stand of a susceptible or resistant cultivar) and of the objective assigned to the mixture: decreasing disease severity relative to the pure stand of susceptible plants or increasing the durability of resistance by reducing the pressure exerted on the virulent subpopulation of the pathogen. Investigations of the longer-term effects of cultivar mixtures would need to consider sexual reproduction, as this process may either amplify or attenuate the demogenetic changes occurring during an annual epidemic. Indeed, mixtures may also affect the actual contribution of each subpopulation to sexual reproduction by modulating the number of 'contacts' and, thus, of crosses involving virulent strains, which are essential for the transmission of virulence.

Recent landscape-scale models have investigated the effect of the proportion of fields sown with resistant and susceptible cultivars and their connectivity on the dynamics of resistance breakdown after several generations (e.g. Fabre *et al.*, 2015). In comparisons of different strategies for preventing or delaying the appearance of a 'superpathogen', cultivar mixtures appeared to be effective in the long term for pathogens with high mutation rates and high fitness costs (Rimbaud *et al.*, 2018; Crété *et al.*, 2020). Our experimental findings

confirm that cultivar mixtures may also be used to extend the lifetime of resistance sources, this time at within-field scale, through the introduction of heterogeneity.

Acknowledgements

We thank Christophe Montagnier (INRAE Experimental Unit, Thiverval-Grignon, FR), Nathalie Retout, Laurent Gerard, Nicolas Lecutier, Auriane Pinton, Miléna Bonnard, Jérôme Lageyre, Martin Willigsecker and Benjamin Boudier (INRAE BIOGER, Thiverval-Grignon, FR) for technical assistances. We thank Dr. Cyrille Saintenac (INRAE GDEC, Clermont-Ferrand, FR) for providing us with the Chinese Spring isogenic lines of wheat. We thank Dr. Romain Valade (ARVALIS-Institut du Végétal, Thiverval-Grignon, FR) and Dr. Thierry Marcel (INRAE BIOGER, Thiverval-Grignon, FR) for sharing information and providing relevant advices throughout this project. We thank Dr. Julie Sappa for her help correcting our English.

Funding

This research was supported by a grant from the 'Fond de Soutien à l'Obtention Végétale' (FSOV PERSIST project; 2019-2022) and by a PhD fellowship from the French Ministry of Education and Research (MESR) awarded to Carolina Orellana-Torrejon for the 2018-2022 period. The BIOGER laboratory also receives support from Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

References

Borg J, Kiaer LP, Lecarpentier C, Goldrienger I, Saint-Jean S, Enjalbert S, 2018. Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research* **221**, 298–313.

Brown JKM, Chartrain L, Lasserre-Zuber P, Saintenac C, 2015. Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. *Fungal Genetics and Biology* **79**, 33–41.

Brown JKM, Tellier A, 2011. Plant-parasite coevolution: bridging the gap between genetics and ecology. *Annual Review of Phytopathology* **29**, 345–367.

Chin K, Wolfe M, 1984. Selection on *Erysiphe graminis* in pure and mixed stands of barley. *Plant Pathology* **33**, 535–545.

Cowger C, Hoffer ME, Mundt CC, 2000. Specific adaptation by *Mycosphaerella graminicola* to a resistant wheat cultivar. *Plant Pathology* **49**, 445–451.

Cowger C, Mundt CC, 2002a. Effects of wheat cultivar mixtures on epidemic progression of Septoria tritici blotch and pathogenicity of *Mycosphaerella graminicola*. *Phytopathology* **92**, 617–623.

Cowger C, Mundt CC, 2002b. Aggressiveness of *Mycosphaerella graminicola* isolates from susceptible and partially resistant wheat cultivars. *Phytopathology* **92**, 624–630.

Crété R, Pires RN, Barbetti MJ, Renton M, 2020. Rotating and stacking genes can improve crop resistance durability while potentially selecting highly virulent pathogen strains. *Scientific Reports* **10**, 19752.

Dai J, Wiersma JJ, Nolen DL, 2012. Performance of hard red spring wheat cultivar mixtures. *Agronomy Journal* **104**, 17–21.

Dileone J, Mundt C, 1994. Effect of wheat cultivar mixtures on populations of *Puccinia striiformis* races. *Plant Pathology* **43**, 917–930.

Fabre F, Rousseau E, Mailleret L, Moury B, 2015. Epidemiological and evolutionary management of plant resistance: optimizing the deployment of cultivar mixtures in time and space in agricultural landscapes. *Evolutionary Applications* **8**, 919–932.

Finckh M, Gacek E, Goyeau H *et al.*, 2000. Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie* **20**, 813–837.

Fones H, Gurr S, 2015. The impact of Septoria tritici blotch disease on wheat: An EU perspective. *Fungal Genetics and Biology* **79**, 3–7.

Ghaffary SMT, Faris JD, Friesen TL *et al.*, 2012. New broad-spectrum resistance to Septoria tritici blotch derived from synthetic hexaploid wheat. *Theoretical and Applied Genetics* **124**, 125–142.

Gigot C, Saint-Jean S, Huber L *et al.*, 2013. Protective effects of a wheat cultivar mixture against splash-dispersed Septoria tritici blotch epidemics. *Plant Pathology* **62**, 1011–1019.

Gigot C, de Vallavieille-Pope C, Huber L, Saint-Jean S, 2014. Using virtual 3-D plant architecture to assess fungal pathogen splash dispersal in heterogeneous canopies: a

case study with cultivar mixtures and a non-specialized disease causal agent. *Annals of Botany* **114**, 863–876.

Johnson R, 1984. A critical analysis of durable resistance. *Annual Review of Phytopathology* **22**, 309–330.

Kildea S, Byrne JJ, Cucak M, Hutton F, 2020. First report of virulence to the Septoria tritici blotch resistance gene *Stb16q* in the Irish *Zymoseptoria tritici* population. *New Disease Reports* **41**, 13–13.

Kristoffersen R, Jørgensen LN, Eriksen LB, Nielsen GC, Kiær LP, 2019. Control of Septoria tritici blotch by winter wheat cultivar mixtures: Meta-analysis of 19 years of cultivar trials. *bioRxiv*, 658575.

Laine A-L, Barrès B, 2013. Epidemiological and evolutionary consequences of life-history trade-offs in pathogens. *Plant Pathology* **62**, 96–105.

Lannou C, 2012. Variation and selection of quantitative traits in plant pathogens. *Phytopathology* **50**, 319–38.

Lannou C, Hubert P, Gimeno C, 2005. Competition and interactions among stripe rust pathotypes in wheat-cultivar mixtures. *Plant Pathology* **54**, 699–712.

McDonald BA, Linde C, 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**, 349–379.

McIntosh RA, Brown GN, 1997. Anticipatory breeding for resistance to rust diseases in wheat. *Annual Review of Phytopathology* **35**, 311–326.

McRoberts N, Hughes G, Madden LV, 2003. The theoretical basis and practical application of relationships between different disease intensity measurements in plants. *Annals of Applied Biology* **142**, 191–211.

Morais D, Sache I, Suffert F, Laval V, 2016. Is the onset of Septoria tritici blotch epidemics related to the local pool of ascospores? *Plant Pathology* **65**, 250–260.

Mundt CC, 2002. Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytopathology* **40**, 381–410.

Niks RE, Parlevliet JE, Lindhout P, Bai Y, 2011. *Breeding crops with resistance to diseases and pests.* Wageningen: Wageningen Academic Publishers.

Perronne R, Makowski D, Goffaux R, Montalent P, Goldringer I, 2017. Temporal evolution of varietal, spatial and genetic diversity of bread wheat between 1980 and 2006

strongly depends upon agricultural regions in France. *Agriculture, Ecosystems & Environment* **236**, 12–20.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rimbaud L, Papaix J, Barrett LG, Burdon JJ, Thrall PH, 2018. Mosaics, mixtures, rotations or pyramiding: What is the optimal strategy to deploy major gene resistance? *Evolutionary Applications* **11**, 1791–1810.

Seybold H, Demetrowitsch TJ, Hassani MA *et al.*, 2020. A fungal pathogen induces systemic susceptibility and systemic shifts in wheat metabolome and microbiome composition. *Nature Communications* **11**, 1–12.

Suffert F, Delestre G, Gélisse S, 2018. Sexual reproduction in the fungal foliar pathogen *Zymoseptoria tritici* is driven by antagonistic density dependence mechanisms. *Microbial Ecology* **77**, 110–123.

Suffert F, Sache I, Lannou C, 2011. Early stages of Septoria tritici blotch epidemics of winter wheat: build-up, overseasoning, and release of primary inoculum. *Plant Pathology* **60**, 166–177.

Suffert F, Sache I, Lannou C, 2013. Assessment of quantitative traits of aggressiveness in *Mycosphaerella graminicola* on adult wheat plants. *Plant Pathology* **62**, 1330–1341.

Vidal T, Boixel A-L, Durand B, de Vallavieille-Pope C, Huber L, Saint-Jean S, 2017. Reduction of fungal disease spread in cultivar mixtures: Impact of canopy architecture on rain-splash dispersal and on crop microclimate. *Agricultural and Forest Meteorology* **246**, 154–161.

Villaréal LMMA, Lannou C, 2000. Selection for increased spore efficacy by host genetic background in a wheat powdery mildew population. *Phytopathology* **90**, 1300–1306.

Wolfe MS, 1985. The current status and prospects of multiline cultivars and variety mixtures for disease resistance. *Annual Review of Phytopathology* **23**, 251–273.

Tables

Table 1. Number of *Z. tritici* strains sampled on each cultivar ('A' for Apache, 'C' for Cellule) for each treatment (C_0-A_1, $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, $C_{0.75}$-$A_{0.25}$ and C_1-A_0), at the early (t1) and late (t2) stages of the 2018 and 2019 epidemics, and phenotyped to determine virulence status.

Year	Stage	C_0-A_1	$C_{0.25}$-$A_{0.75}$		$C_{0.5}$-$A_{0.5}$		$C_{0.75}$-$A_{0.25}$		C_1-A_0	Total
		A	C	A	C	A	C	A	C	
2018	t1 [1]	179	90	90	90	90	90	87	90	806
	t2 [2]	90	90	90	89	90	90	90	90	719
2019	t1 [3]	180	82	88	88	84	90	89	90	791
	t2 [4]	90	90	90	90	90	90	89	90	719

[1] collected from March 5th to 12th; [2] June 13th; [3] from February 27th to March 13th; [4] June 25th

Figures

Figure 1. (a) Overview of the development of a Septoria tritici blotch epidemic adapted to (b, c) two pure stands and to (d) a mixture of the two wheat cultivars, one carrying the *Stb16q* resistance gene, and the other lacking this gene (cv. Cellule in proportion p_C and cv. Apache in proportion $1 - p_C$). The size of the pathogen population can be expressed as $n_{t,v,vir}$ (size of the virulent subpopulation on the cultivar $v \in$ [A; C] at time $t \in$ [t1; t2]) and $\overline{n}_{t,vir}$ (size of the virulent subpopulation on the whole canopy at time t). The composition of the pathogen population is expressed as $f_{t,v,vir}$ (frequency of virulent strains on cultivar v at t) and $\overline{f}_{t,vir}$ (frequency of virulent strains on the whole canopy at time t), with $f_{t0,vir}$ the initial frequency of virulent strains in the local pathogen population.

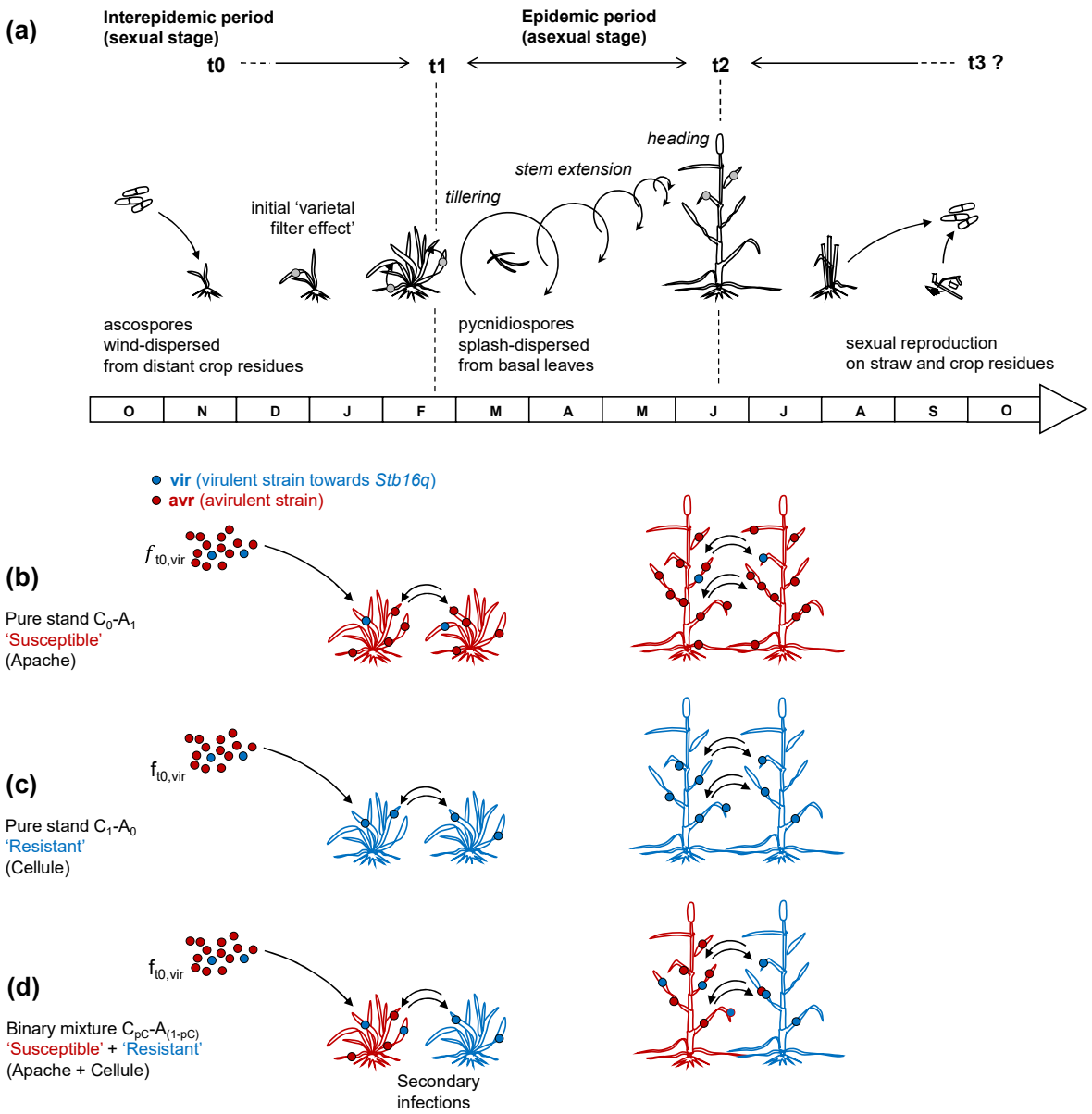


Figure 2. Changes in the percentage of the area under wheat sown with (a) cv. Apache, (b) cv. Cellule, and (c) the 13 cultivars carrying *Stb16q* in France (black dots) and in the Parisian basin (Région Ile-de-France 'IdF'; white dots) from 2005 to 2020 (FranceAgriMer datasets).

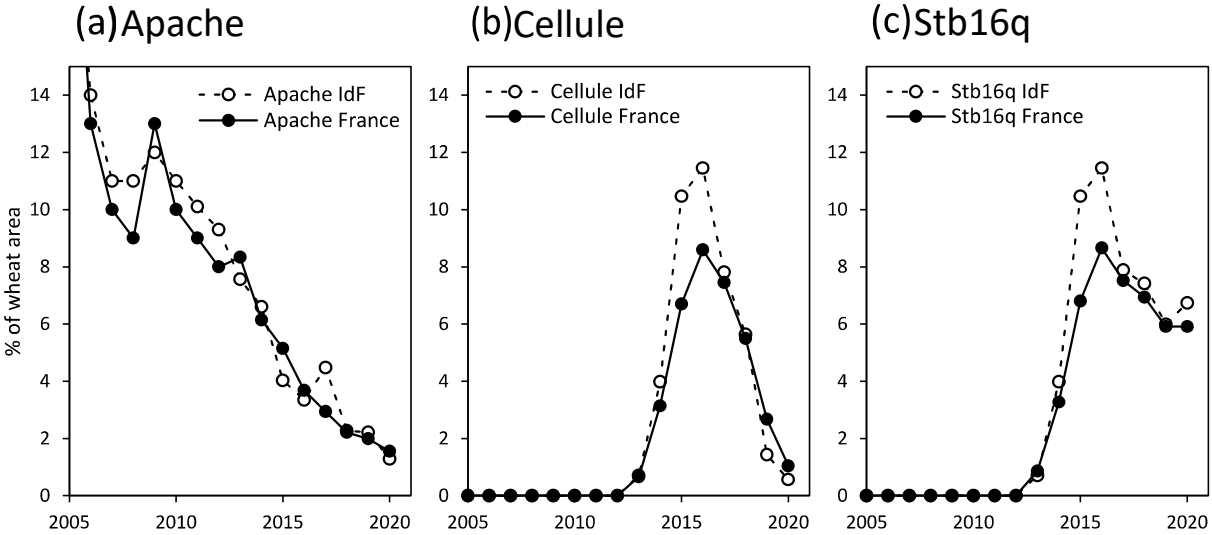


Figure 3. (a) Example of two 102-strain batches, one obtained just after inoculation (right) and the other one week later (left). (b-e) Typical incompatible-compatible interaction between the FS1804 (avirulent against *Stb16q*) and FS2034 (virulent against *Stb16q*) strains on six Cellule and Apache leaves (three second leaves on the left and three third leaves on the right).

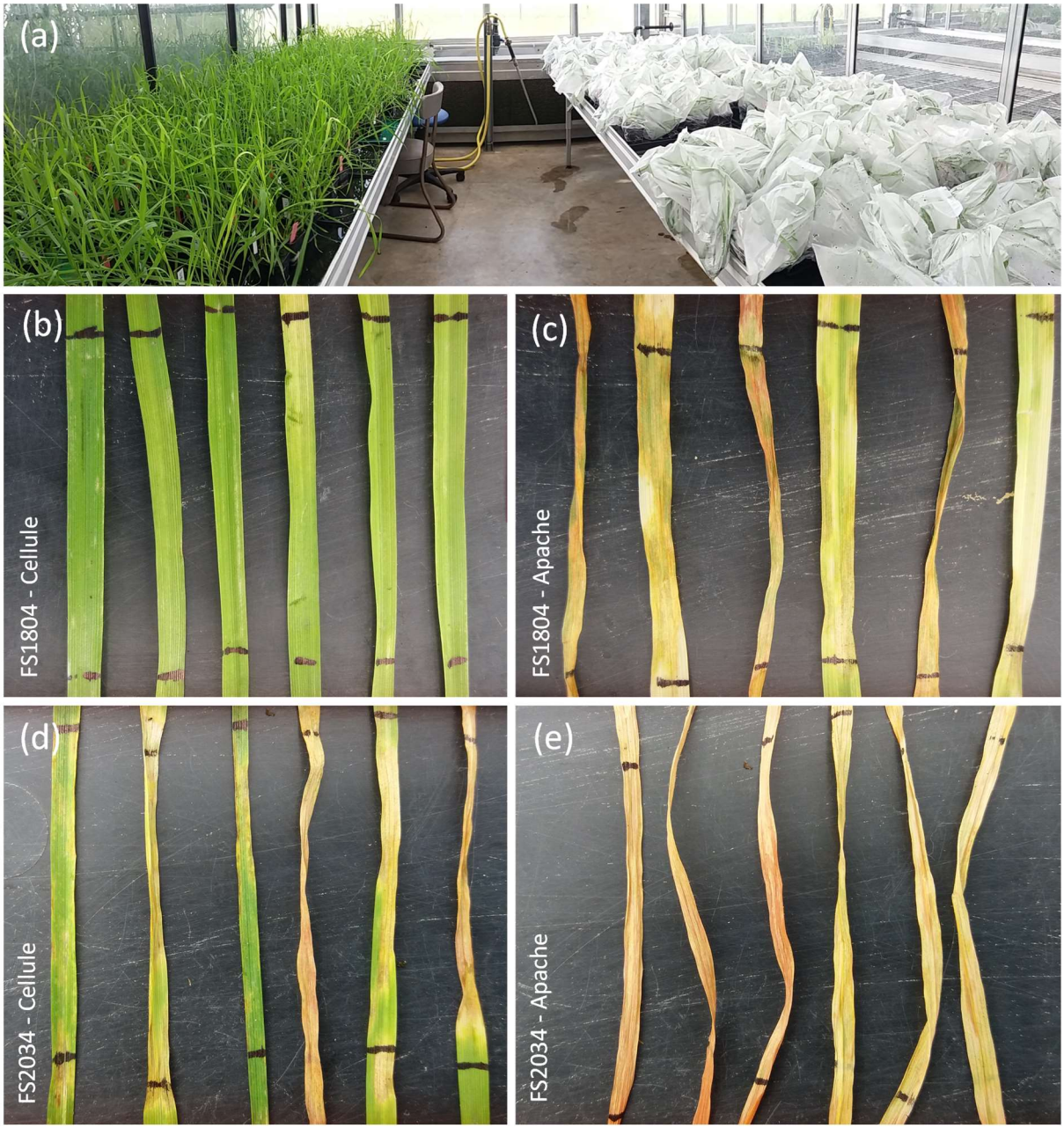


Figure 4. Mean disease severity index SEV_v (mean percentage of the leaf area displaying sporulation on cultivar *v*) according to the proportion of cv. Cellule (p_C) in the canopy. SEV_A (green solid line) and SEV_C (red solid line) were calculated from disease assessments in the field. The overall severity index $\overline{SEV}$ (black solid line) was calculated at the whole-canopy scale as the mean of SEV_A and SEV_C weighted by the proportion of each cultivar. Bars represent the standard error. The 'theoretical' $\overline{SEV}$ (in the absence of a mixture effect) was plotted as a linear function of the severity in each pure stand weighted by p_C (black dotted lines). SEV_A and SEV_C in pure stands at t1 in 2019 are not shown, because disease severity was assessed one week earlier than in mixtures.

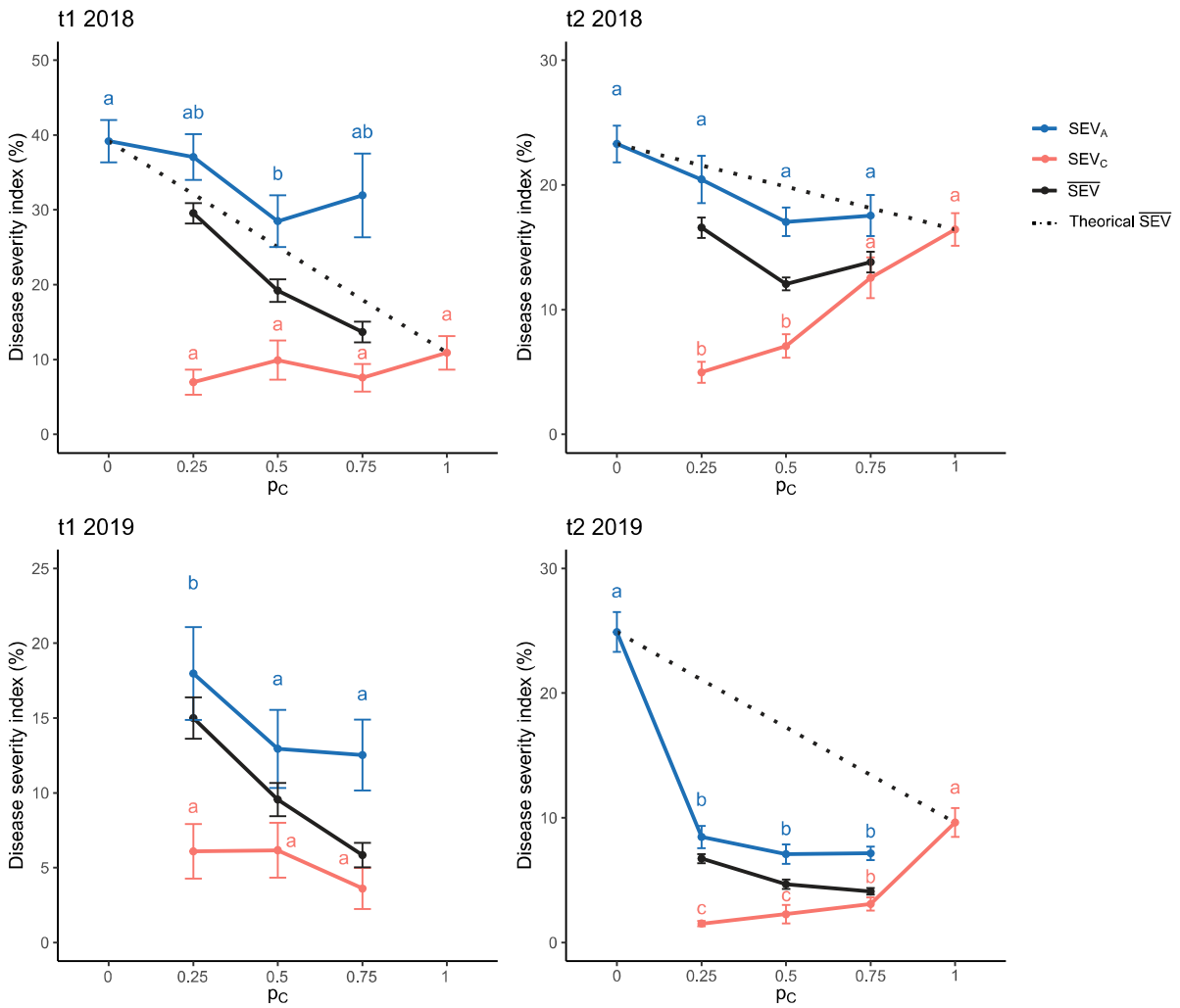


Figure 5. Percentage of *Z. tritici* strains considered virulent (grey) against *Stb16q* (i.e., causing sporulating lesions on cv. Cellule) and avirulent (white) among the populations collected on cv. Apache in pure stands (p_C = 0) and mixtures (p_C = 0.25, 0.5 and 0.75 for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively) at the early (t1) and late (t2) stages of the 2018 and 2019 epidemics. The effect of treatment (pure stands and cultivar mixtures) on the number of virulent strains (in the boxes) was assessed by performing a χ^2 test with Bonferroni correction for pairwise comparison.

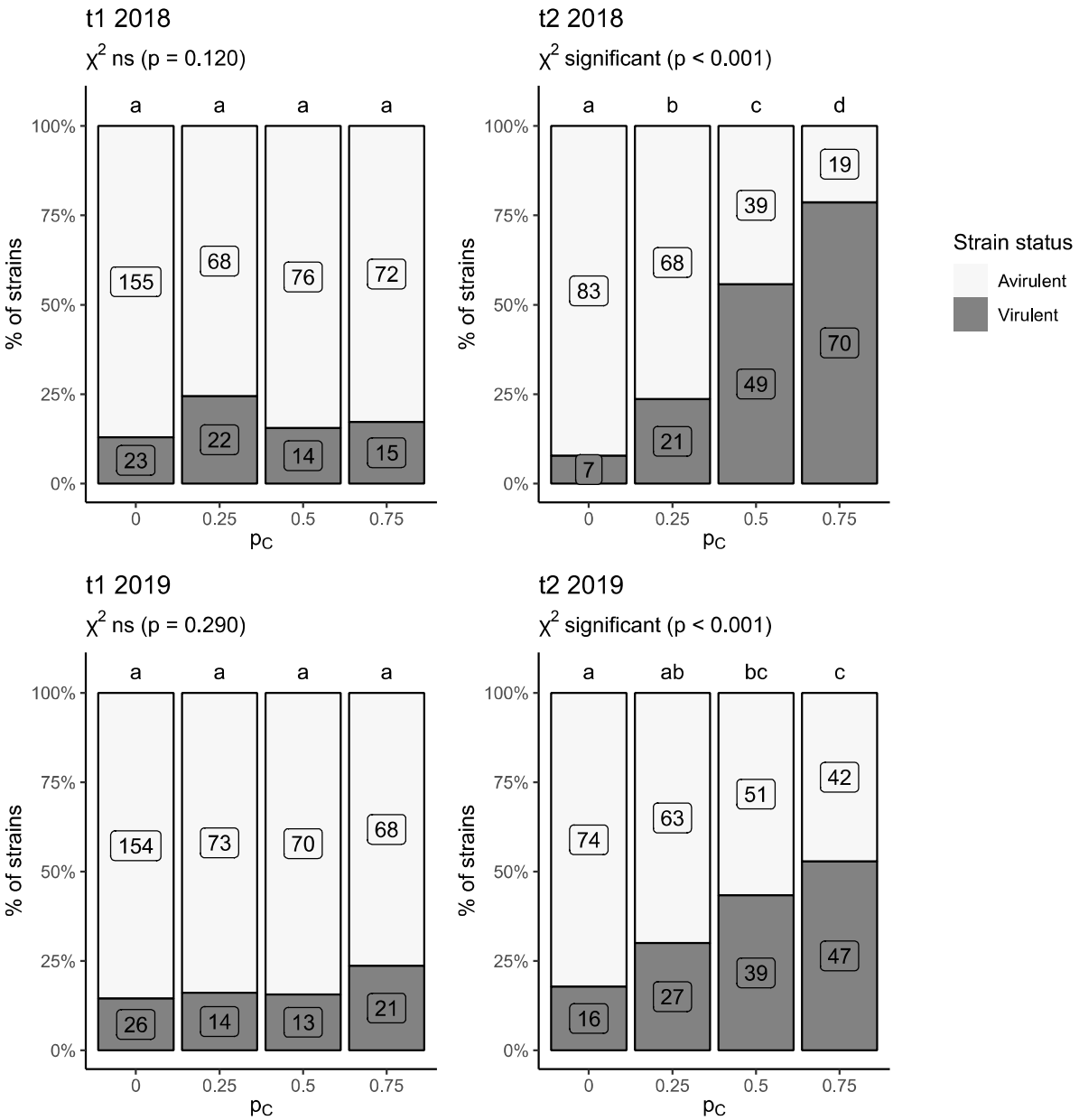


Figure 6. Percentage of *Z. tritici* strains considered virulent (grey) against *Stb16q* (i.e., causing sporulating lesions on cv. Cellule) and avirulent (white) among the populations collected on cv. Cellule in pure stands (p_C = 1) and mixtures (p_C = 0.25, 0.5 and 0.75 for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively) at the early (t1) and late (t2) stages of the 2018 and 2019 epidemics. The effect of treatment (pure stands and cultivar mixtures) on the number of virulent strains (in the boxes) was assessed by performing a χ^2 test, or a Fisher's exact test when the expected numbers were small, with Bonferroni correction for pairwise comparison.

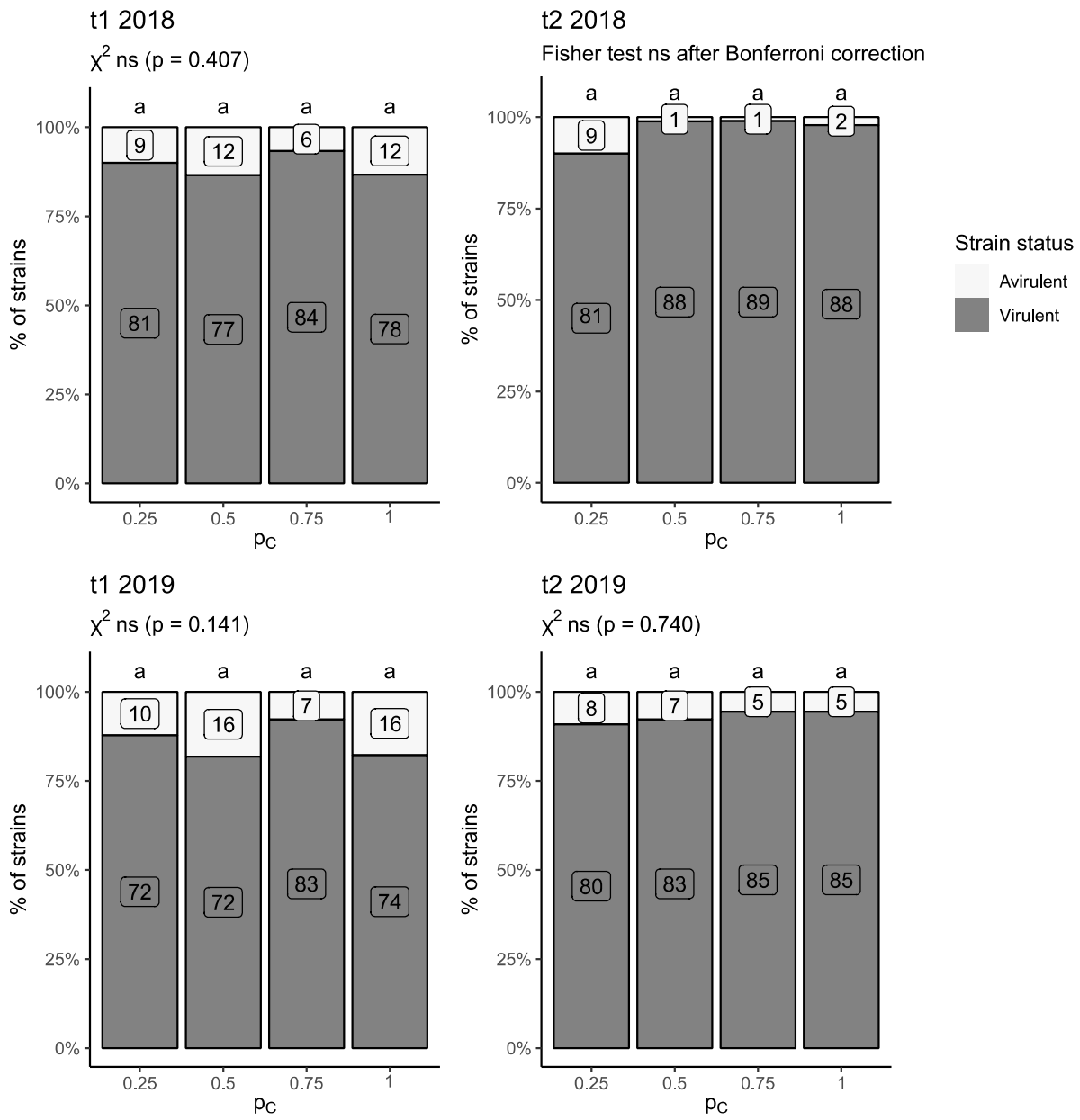
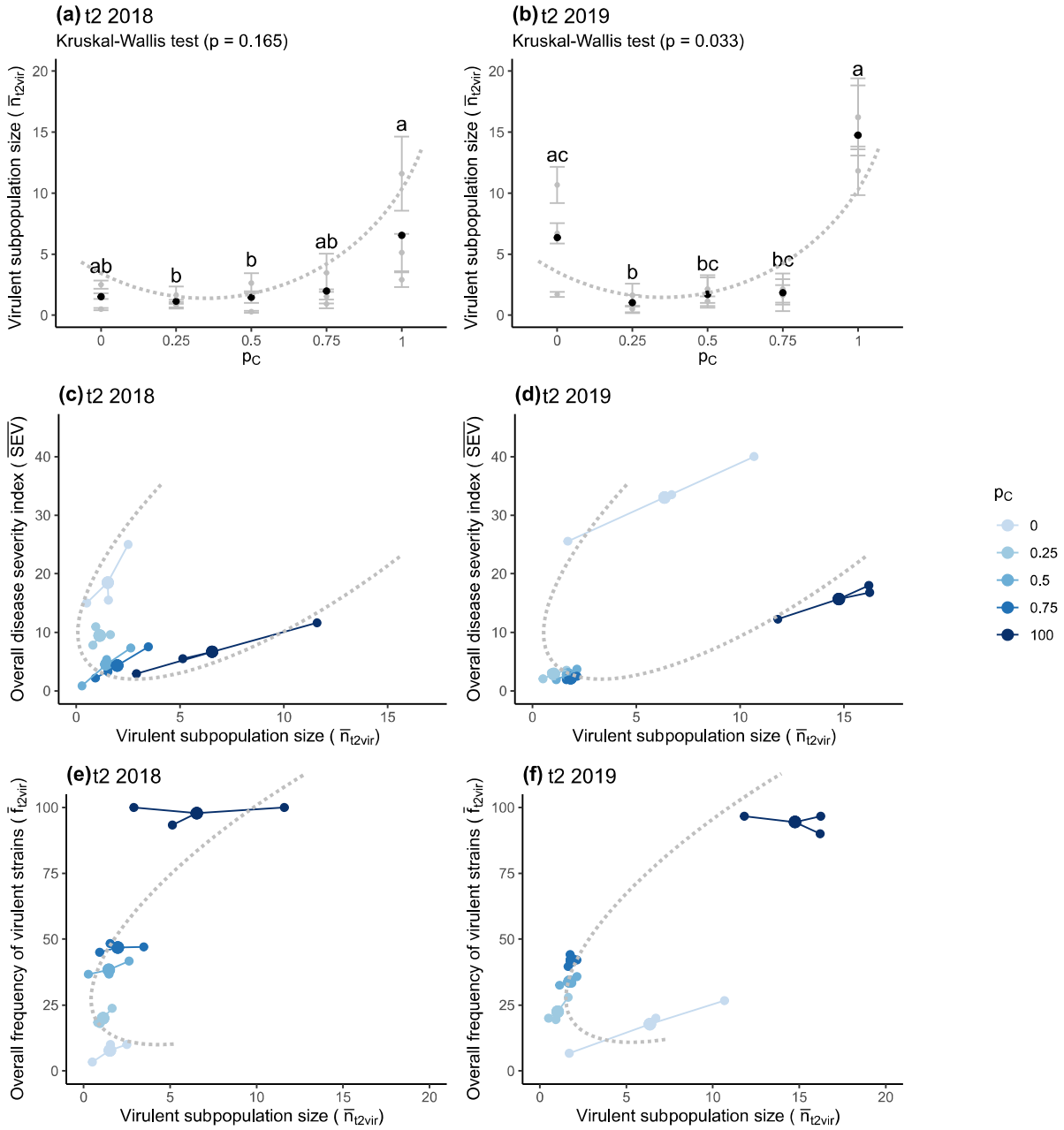


Figure 7. Impact of mixtures on different indicators of the *Z. tritici* virulent subpopulations and their relationships at whole-canopy scale. (a, b) Size of the virulent subpopulations at t2 ($\overline{n}_{t2,vir}$, dimensionless variable; black dots) according to the proportion of cv. Cellule (p_C) in 2018 and 2019. Grey dots indicate the mean value for each block. Bars represent the standard error. (c, d) Disease severity assessed at the whole-canopy scale at t2 ($\overline{SEV}$, determined as the percentage of the leaf area displaying sporulation on flag leaves; blue dots) according to the size of the virulent subpopulation ($\overline{n}_{t2,vir}$). (e, f) Frequency of virulent strains assessed at whole-canopy scale at t2 ($\overline{f}_{t2,vir}$; blue dots) according to the size of the virulent subpopulations in 2018 and in 2019. Small dots indicate the mean value for each block, and large dots are the barycentres (mean for each treatment). The dashed grey lines illustrate a hypothetical mixture efficiency trend.



Supplementary Tables

Table S1. Actual proportion of wheat plants cv. Apache (A) and cv. Cellule (C) in microplots in pure stands (C_0-A_1 and C_1-A_0) and in mixtures ($C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, and $C_{0.75}$-$A_{0.25}$) at the late stage of epidemics in 2018 and 2019. Each proportion was estimated from the proportion of wheat axis of each cultivar at heading assessed on the basis of the presence (C) or absence (A) of bearded heads (counted along three lines of 1 m on July 4th, 2018 and July 1st, 2019). The total number of heads per metre is indicated in italics, in brackets.

Year	Block	C_0-A_1	$C_{0.25}$-$A_{0.75}$		$C_{0.5}$-$A_{0.5}$		$C_{0.75}$-$A_{0.25}$		C_1-A_0
		A	C	A	C	A	C	A	C
2018	a	1	0.12	0.88	0.52	0.48	0.61	0.39	1
		(82)	*(68)*		*(68)*		*(75)*		*(79)*
	b	1	0.25	0.75	0.27	0.73	0.71	0.29	1
		(76)	*(82)*		*(75)*		*(92)*		*(82)*
	c	1	0.20	0.80	0.46	0.54	0.69	0.31	1
		(67)	*(82)*		*(77)*		*(73)*		*(97)*
	mean	1	0.19	0.81	0.42	0.58	0.67	0.33	1
		(75)	*(77)*		*(73)*		*(80)*		*(86)*
2019	a	1	0.14	0.86	0.35	0.65	0.76	0.24	1
		(65)	*(86)*		*(69)*		*(74)*		*(79)*
	b	1	0.39	0.61	0.60	0.40	0.61	0.39	1
		(103)	*(77)*		*(82)*		*(95)*		*(73)*
	c	1	0.31	0.69	0.35	0.65	0.82	0.18	1
		(91)	*(72)*		*(65)*		*(95)*		*(69)*
	mean	1	0.28	0.72	0.43	0.57	0.73	0.27	1
		(86)	*(78)*		*(72)*		*(88)*		*(73)*

Table S2. The origin (treatment, from the following list: $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$, $C_{0.75}$-$A_{0.25}$, C_1-A_0; period, t1 or t2, in 2018) of the 47 *Z. tritici* strains shown to be avirulent against *Stb16q* established in the first population-phenotyping, for which virulence status was validated by confirmatory individual phenotyping. The total number of strains tested is indicated in brackets. There was no significant difference in the number of avirulent strains at t1 between treatments (Fisher's test, $p = 1$).

	$C_{0.25}$-$A_{0.75}$	$C_{0.5}$-$A_{0.5}$	$C_{0.75}$-$A_{0.25}$	C_1-A_0
t1	7 (11)	7 (11)	4 (7)	6 (12)
t2	0 (4)	0 (1)	0 (0)	0 (1)

Table S3. Main sources of uncertainty and benefit-disadvantage analysis of the 'population phenotyping' method compared to the classical 'confirmatory individual phenotyping' method (see also Figure S3). 'False negative' refers to virulent strains that appeared avirulent in the population phenotyping but were actually virulent in the confirmatory individual phenotyping.

Sources of uncertainty	Disadvantages	Solutions for reducing disadvantages and estimating uncertainty	Benefits
Preparation of the inoculum suspensions	The targeted concentration 10^5 spores·mL^{-1} is neither precisely measured nor adjusted, resulting in a loss of accuracy.	The mean concentration of suspensions was estimated ±3.0 10^5 spores·mL^{-1} from a sample of 40 strains, with 85% in the range 1.0 x 10^5-5.0 x 10^5 spores·mL^{-1} (Figure S4). The impact of such variation (see Figure 8 in Suffert *et al.*, 2018) can be considered negligible for characterising the virulence status of strains and acceptable for comparing the orders of magnitude of aggressiveness between virulent and avirulent strains. Uncertainty was further reduced by having the same person prepare all the batches of inoculum suspensions.	Overall preparation time cut to a third (±30 isolates vs. >100 isolates per half day)
Climatic conditions in the greenhouse	Semi-controlled environmental conditions (temperature, moisture and light can only be maintained within a range of variation and not set at given values). The outdoor climatic conditions may influence the climatic conditions in the greenhouse and, thus, the physiology of	The impact of climatic conditions on the development of *Z. tritici in planta* has been clearly established (McCorison & Goodwin, 2020). Hygrometric variations and variations of light exposure may explain some of the 'false negatives' and the poor results obtained with two non-conforming batches. Comparing these less suitable conditions with a classical phenotyping method in a growth chamber, a low threshold was used to identify strains as virulent (at least 5% of the area sporulating on at least one of the six inoculated leaves). Air temperature monitoring (Figure S5) made it possible to check that the thermal time from inoculation to disease assessment	Overall time for *in planta* phenotyping reduced by at least 40% due to space saving in our facilities

	wheat plants between batches due to seasonality.	was significantly longer than the mean latent period of *Z. tritici* for all batches. The two non-conforming batches were identified (using eight virulent strains as control), and removed before retesting.	
Small number of biological replicates: - Only six leaves inoculated - Potential leaf layer effect - No repetition of the batches	The number of 'false negatives' is expected to be inversely related to the number of biological replicates (number of leaves inoculated with a given strain) and to be higher in the absence of repetitions.	Confirmatory individual phenotyping was performed with a larger number of biological replicates in more suitable conditions (growth chamber). Half the strains collected on cv. Cellule in the field that appeared avirulent after the first population-phenotyping were actually virulent. Thus, the maximal percentage of 'false negatives' was estimated ±3%, which was considered acceptable with respect to the goal of this study.	Overall time for *in planta* phenotyping reduced by at least 40% due to space saving in our facilities. Possibility of testing a very large number of strains (several thousand rather than several hundred)
Visual assessment	Visual assessment is theoretically less accurate than image analysis for area estimation (e.g. Stewart *et al.*, 2016).	The difficulties detecting pycnidia made it impossible to use image analysis in practice for high-throughput population phenotyping on seedlings. Uncertainty was reduced further by having the same person perform all visual assessments.	Overall time for *in planta* phenotyping reduced by at least 40% in our facilities

Supplementary Figures

Figure S1. Design of the field experiment performed at the INRAE experimental station in Grignon (Yvelines, France, 48°51'N, 1°58'E) in 2018 and repeated in 2019. The field plot, previously under maize and ploughed, was divided into three blocks (a, b, c), each consisting in five microplots (3.5 wide m x 8 m long, with an inter-row spacing of 0.175 m) corresponding to five randomly distributed treatments: a pure stand of wheat cv. Apache (C_0-A_1; 220 Apache seeds·m^{-2}), a pure stand of wheat cv. Cellule (C_1-A_0; 220 Cellule seeds·m^{-2}), and three binary mixtures with proportions of Cellule of 0.25 ($C_{0.25}$-$A_{0.75}$; 55 Cellule seeds·m^{-2} + 165 Apache seeds·m^{-2}), 0.5 ($C_{0.5}$-$A_{0.5}$; 110 Cellule seeds·m^{-2} + 110 Apache seeds·m^{-2}) and 0.75 ($C_{0.75}$-$A_{0.25}$-; 165 Cellule seeds·m^{-2} + 55 Apache seeds·m^{-2}). Seed lots were prepared with the established proportions by taking into account the specific weight of each cultivar seed (TKW Apache = 45 g; TKW Cellule = 46 g). The microplots were separated by 5.25 m-wide strips of triticale cv. Vuka (220 seeds·m^{-2}) resistant to *Z. tritici* to prevent disease spread between neighbouring plots. Each microplot was split longitudinally into two equal parts. In the first part (western side), lines 2-3-4 were used for strain collection, and lines 6-7-8 were used for the estimation of disease severity early in the epidemic (t1); the second part (eastern side) was used similarly, but at the late stage of the epidemic (t2).

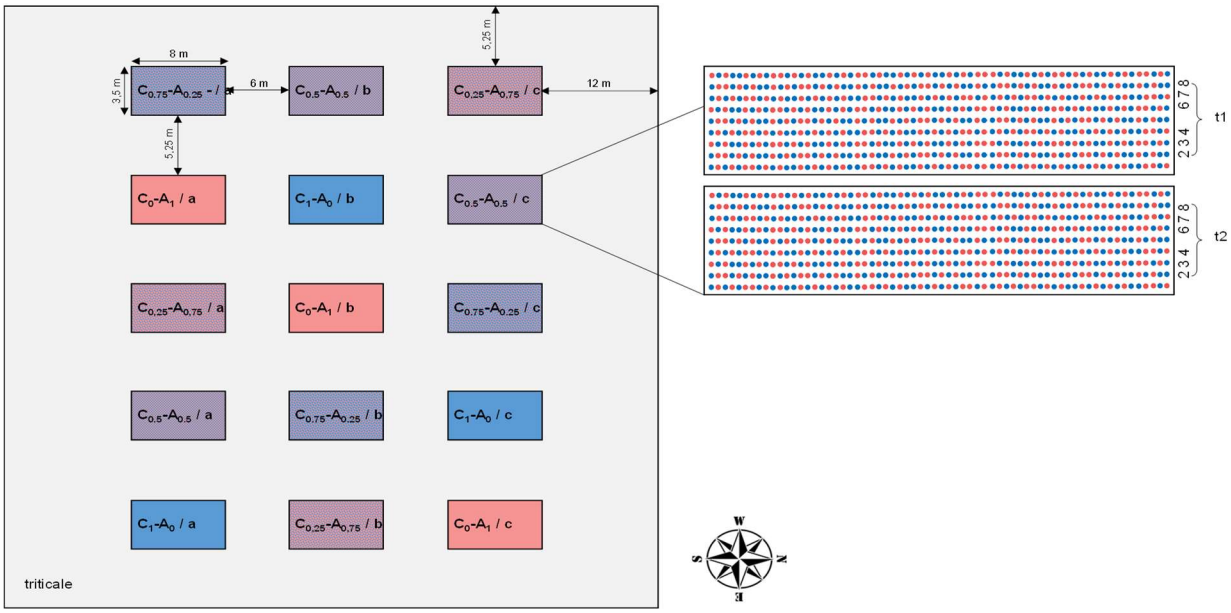


Figure S2. Overview of field conditions during disease assessments and sampling at the early (t1) and late (t2) stages of the epidemic. (a) Microplot at t1 (March 7, 2018). (b) Wheat seedlings of cv. Apache (red seed) and cv. Cellule (green seed). (c) Sporulating lesion caused by *Z. tritici* on an Apache seedling. (d) Microplot at t2 (May 11, 2018). (e) Adult plants of Apache (head not bearded) and Cellule (bearded head). (f) Cluster of lesions on an Apache adult leaf.

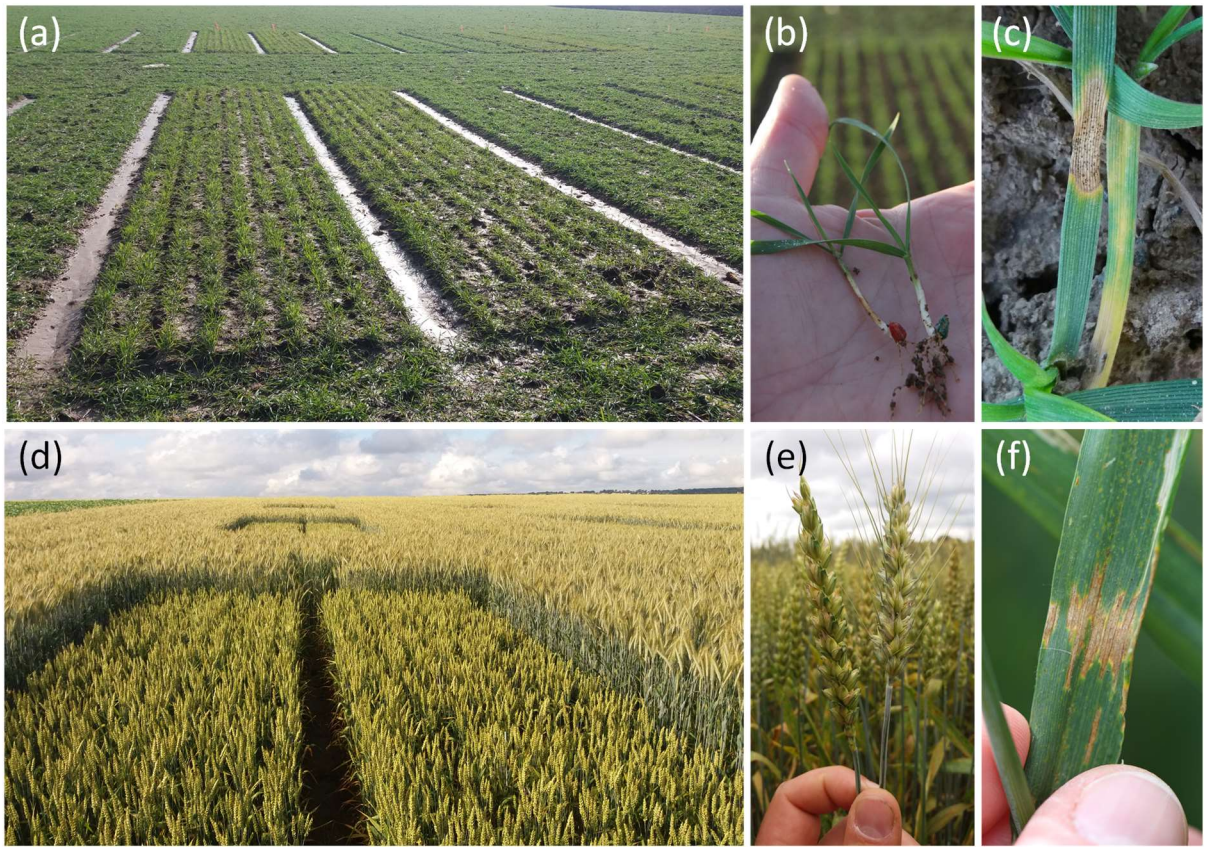


Figure S3. Simplified method of inoculation developed for 'population-phenotyping', estimated to be twice as efficient as the classical method (30 isolates per half day x 36 days vs. 100 isolates per half day x 22 days).

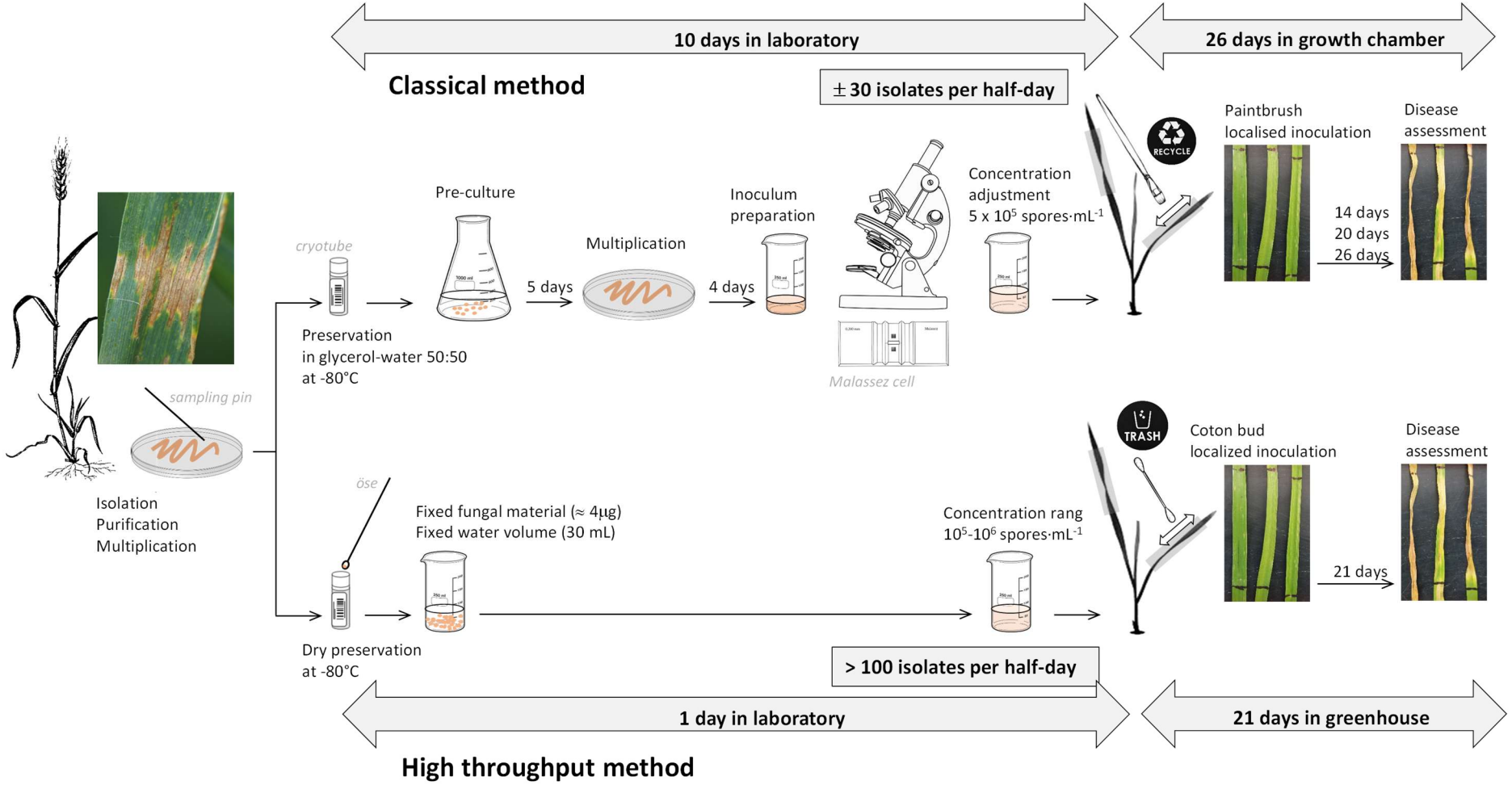


Figure S4. Variability of *Z. tritici* blastospore concentration in the non-adjusted suspensions used for 'population-phenotyping'. The simplified method involved diluting, in a volume of 30 mL of water, the mass of blastospores recovered with a 10-µL inoculating loop (central ring half-filled i.e., about 10 mg) from the freshly thawed contents of cryotubes stored at -80°C. Two independent series of tests (black and white dots) were performed with 2 x 20 strains. The mean concentration (white diamond) was 3.0 x 10^5 spores·mL^{-1}, with 85% of samples within the 10^5-5.0 x 10^5 spores·mL^{-1} range.

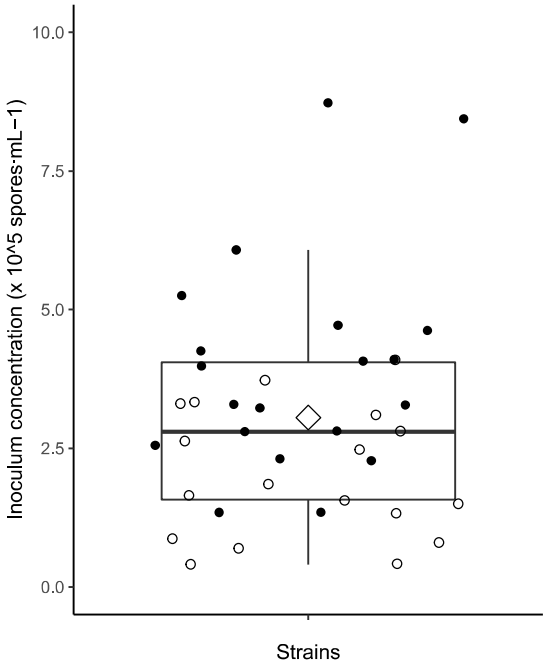


Figure S5 Thermal conditions to which the 32 batches of *Z. tritici* strains were exposed during population-phenotyping in the greenhouse. Thermal time, expressed in degree-days post inoculation (°C·days^{-1}), was calculated by summing daily mean air temperatures (0°C as the base temperature). The mean thermal time per population is plotted (white diamonds). Temperature was recorded every 15 min with a thermal sensor (digital probes with a SPY RF N recorder; JRI, Argenteuil, France). Batches were phenotyped (i) from March 12th to June 6th, 2018 for strains collected at t1 in 2018, (ii) from June 25th to August 21st 2018 and from July 5th to August 17th 2019 for strains collected at t2 in 2018, (iii) from February 14th to March 27th and May 18th to August 11th 2020 for strains collected at t1 in 2019, (iv) from July 12th to October 17th 2019 for strains collected at t2 in 2019.

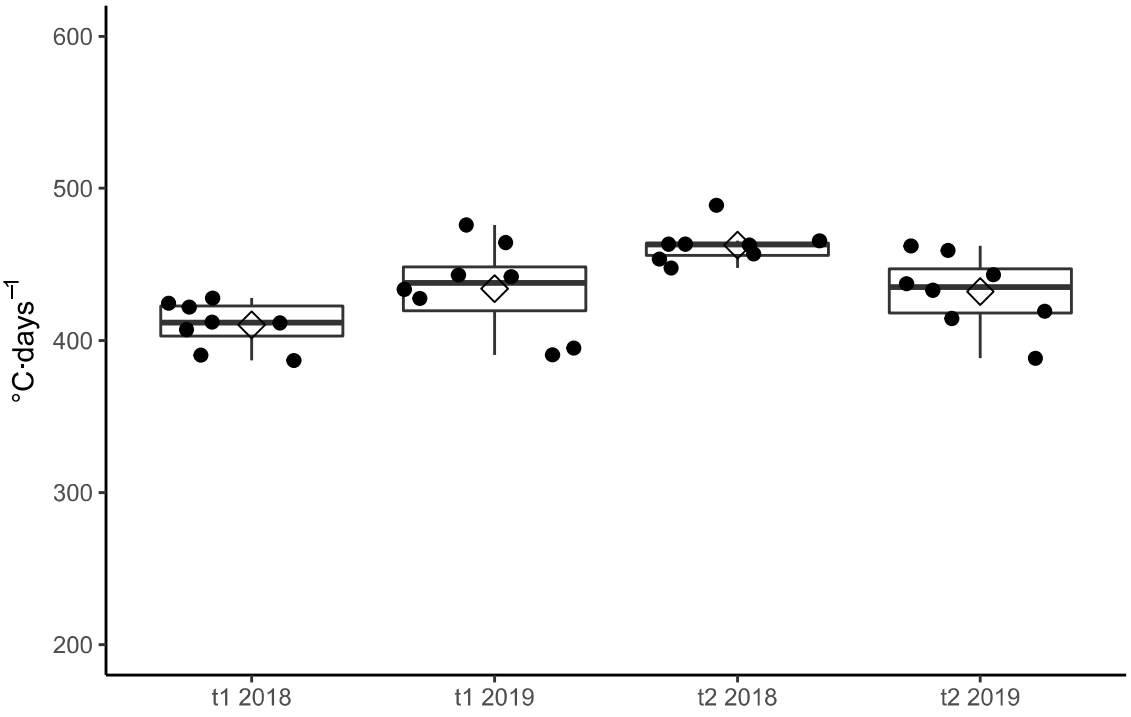


Figure S6. Percentage of *Z. tritici* strains considered 'pathogenic' (i.e., causing sporulating lesions on Apache; grey) and non-pathogenic (white) among the populations collected on Cellule in pure stands (p_C = 1) and mixtures (p_C = 0.25, 0.5 and 0.75 for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively) at the early (t1) and late stages (t2) of the 2018 and 2019 epidemics. The effect of treatment (pure stands and cultivar mixtures) on the number of virulent strains (in the boxes) was assessed by performing Fisher's exact test (the expected numbers were small), with Bonferroni correction for pairwise comparison.

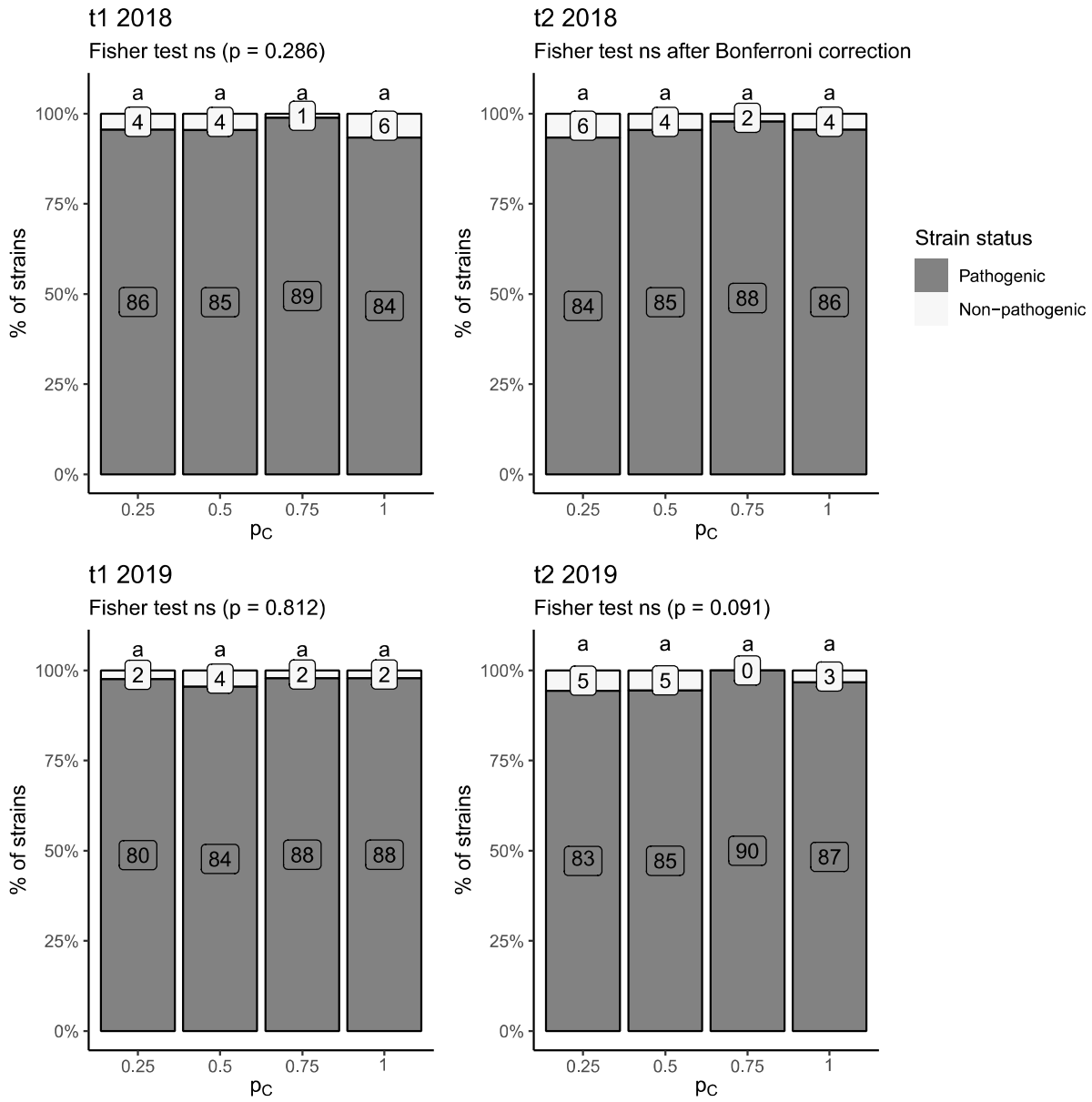


Figure S7. Percentage of *Z. tritici* strains considered pathogenic (i.e., causing sporulating lesions on Apache in the greenhouse; grey) and non-pathogenic (white) among the populations collected on Apache in pure stands (p_C = 0) and mixtures (p_C = 0.25, 0.50 and 0.75 for $C_{0.25}$-$A_{0.75}$, $C_{0.5}$-$A_{0.5}$ and $C_{0.75}$-$A_{0.25}$, respectively) at the early (t1) and late stages (t2) of the 2018 and 2019 epidemics. The effect of treatment (pure stands and cultivar mixtures) on the number of virulent strains (in the boxes) was assessed by performing a χ^2 test, or Fisher's exact test when the expected numbers were small, with Bonferroni correction for pairwise comparison.

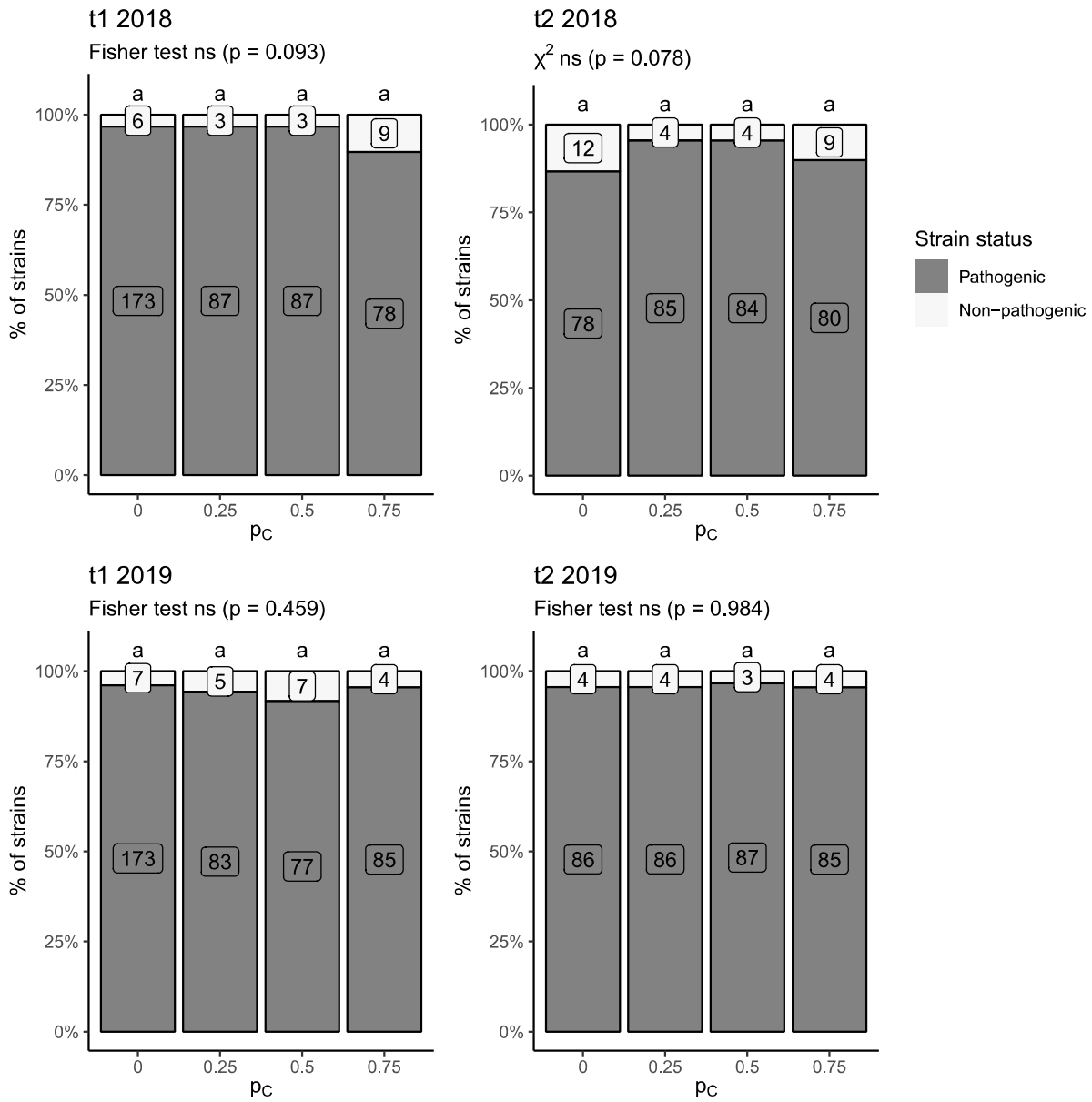


Figure S8. Comparison of the aggressiveness (AG for aggressiveness and AGc for conditional aggressiveness) of 2886 pathogenic strains according to their virulence status with respect to *Stb16q*. Aggressiveness, expressed as the percentage sporulating leaf area, was assessed on Apache. For AG, each dot represents the mean sporulating area induced by a strain on six leaves, and for AGc, the mean sporulating area on symptomatic leaves only. For each subpopulation the number of strains (n) and the mean are indicated. No significant difference between avirulent and virulent strains was observed for either AG (p = 0.054) or AGc (p = 0.054).

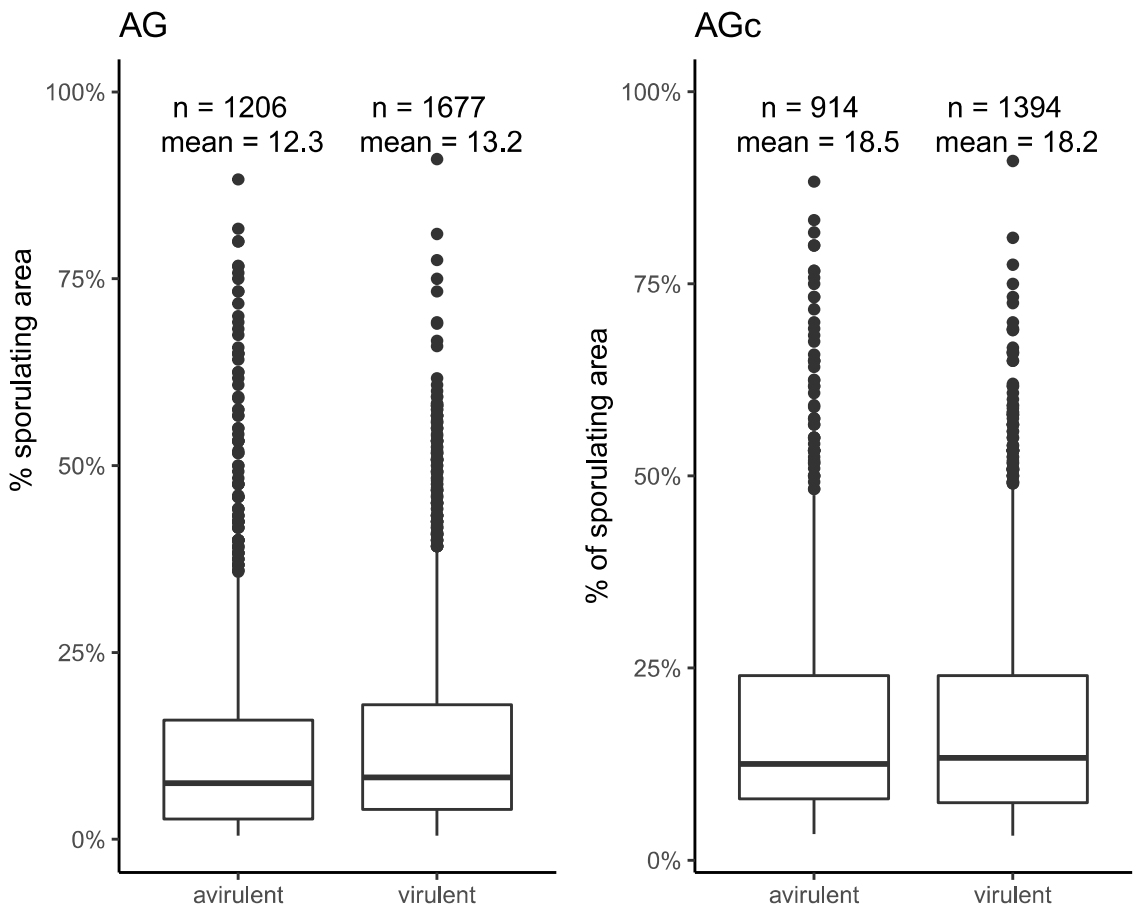


Figure S9. Aggressiveness of the 47 *Z. tritici* strains on Cellule (C) and on the isogenic line Chinese Spring with *Stb16q* (CS-*Stb16q*) assessed during confirmatory individual phenotyping. Each point is the mean percentage of the total area of the inoculated section sporulating on nine leaves per strain × wheat cultivar interaction. The virulence threshold was fixed at 10% (dotted lines). The 16 strains considered virulent on both Cellule and CS-*Stb16q* (mean aggressiveness 58.2% and 65.3%, respectively) are indicated with black dots (yellow square); two strains considered virulent only on Cellule and five strains considered virulent only on CS-*Stb16q* are indicated with white dots (blue rectangles); 24 strains considered avirulent on both Cellule and CS-*Stb16q* are indicated with crosses (red square).

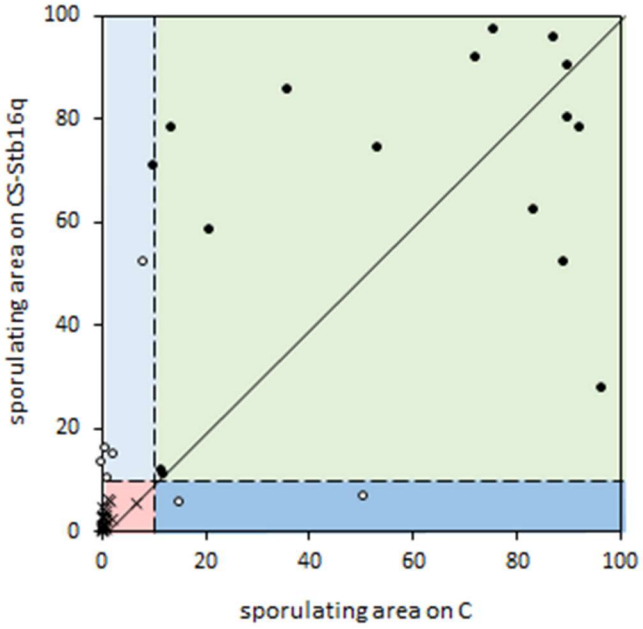


Figure S10. Analysis of disease severity data leading to the fixing at 10% of the relevant threshold (dashed line) for determining the virulence status of *Z. tritici* strains with respect to *Stb16q* 20 days post inoculation (dpi) in confirmatory individual phenotyping. Strains FS1766 (avirulent) and FS5948 (virulent) were tested three times (3 independent batches × 9 inoculated leaves as replicates) on the isolines Chinese Spring (CS), Chinese Spring with *Stb16q* (CS-*Stb16q*) and Cellule (C). Each point is the disease severity assessed on a leaf (i.e., sporulating area, expressed as a percentage of the inoculated leaf section).

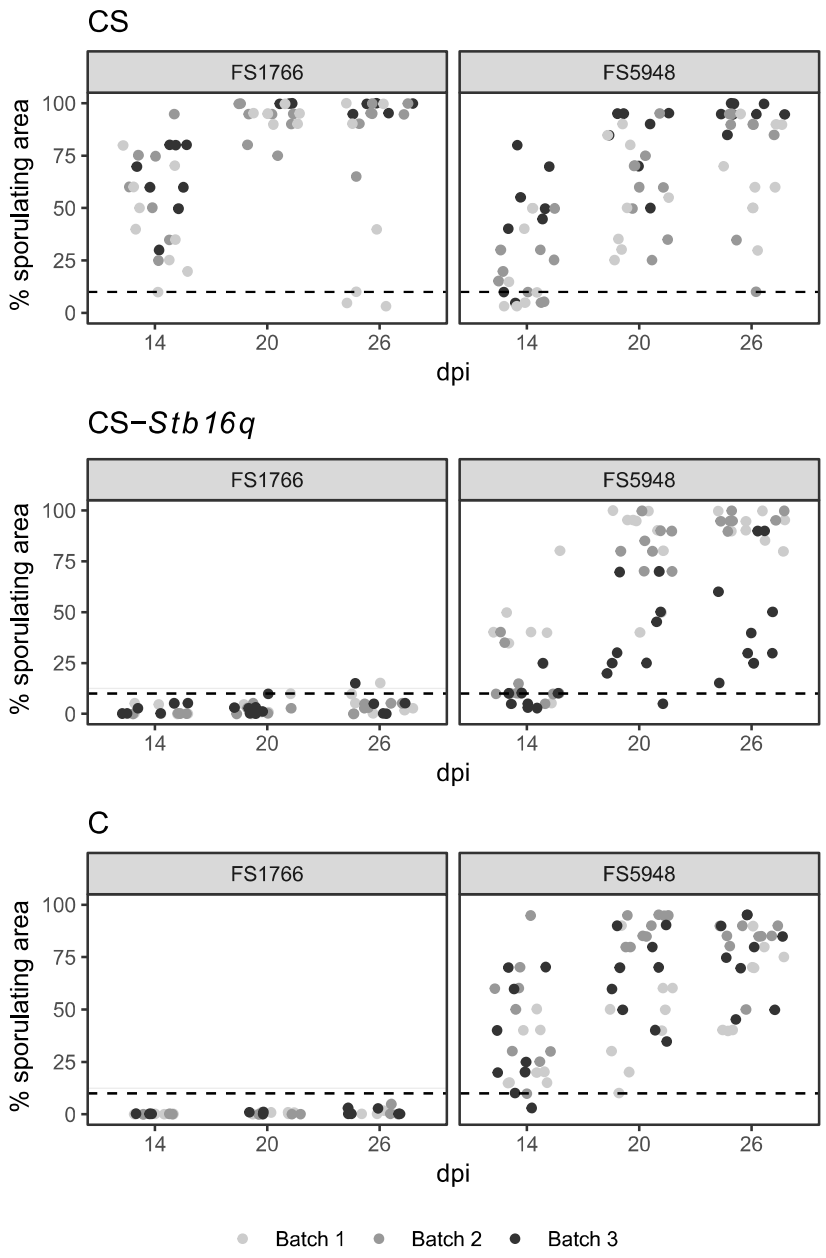
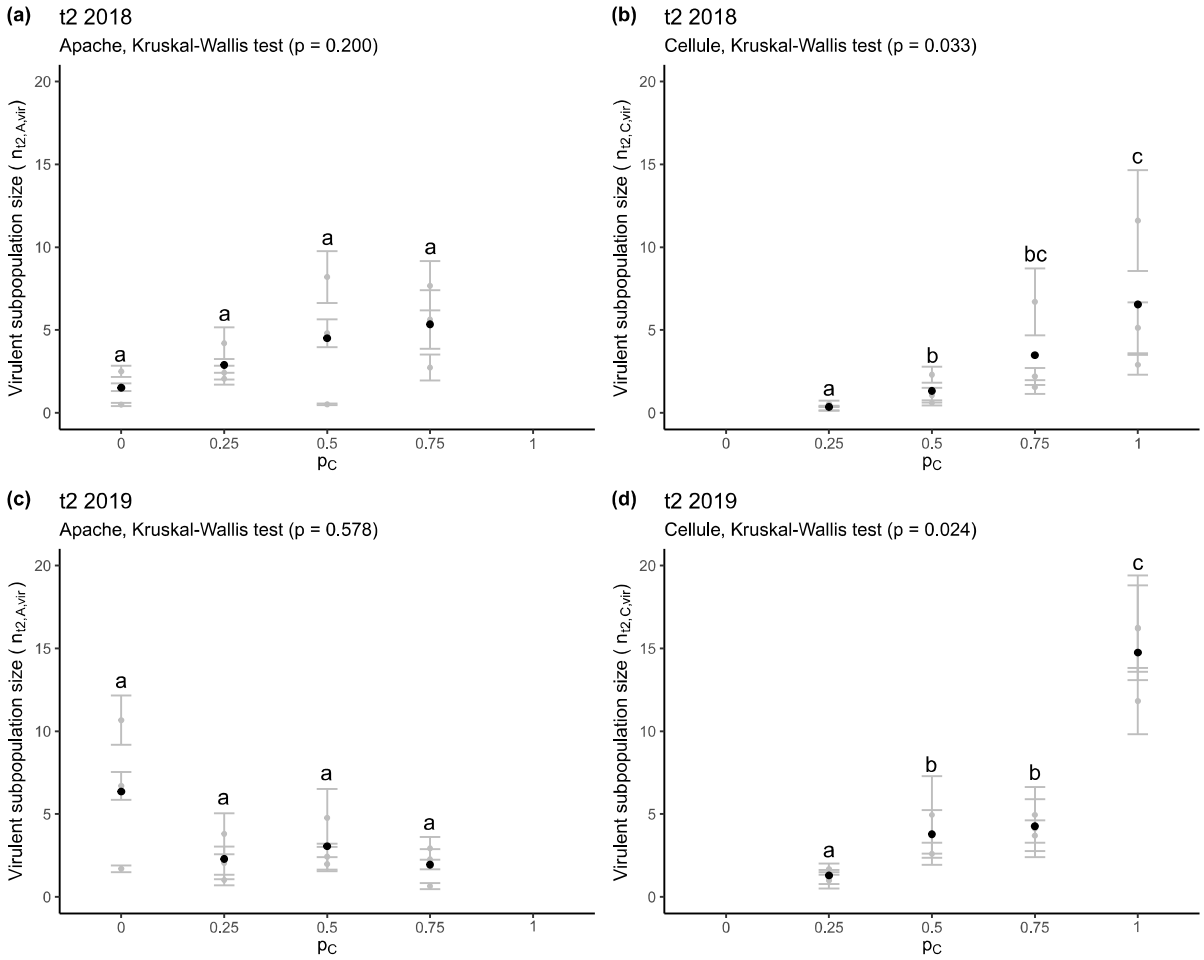



Figure S11. Impact of treatment (pure stands and mixtures) on the *Z. tritici* virulent subpopulation on each cultivar at the late stage of the epidemic (t2). (a, c) Mean size of the virulent subpopulation on Apache ($n_{t2,A,vir}$; dimensionless variable) according to the proportion of Cellule (p_C) in 2018 and 2019. (b, d) Mean size of the virulent subpopulation on Cellule ($n_{t2,C,vir}$) according to the proportion of Cellule (p_C) in 2018 and 2019. Grey dots indicate the mean value for each block. Bars represent the standard error. Error propagation was taken into account.