# Spatially-explicit modeling improves empirical characterization of dispersal: theory and a case study

Petteri Karisto[1]†, Frédéric Suffert[2], and Alexey Mikaberidze[1]‡

[1] Plant Pathology Group, Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland.
[2] INRA, Agro Paris Tech, UMR 1290 BIOGER-CPP, Thiverval-Grignon, France.
† present address: Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.
‡ present address: School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading, UK.

Corresponding author: P. Karisto
Email address: petteri.karisto@ieu.uzh.ch

Keywords: Dispersal kernel, Dispersal gradient, Source geometry, *Zymoseptoria tritici*, Rain-splash dispersal, Bootstrap

Content:
Abstract
Main text
Appendices A, B, C, and D.

6700 words in the main text.

## Abstract

Dispersal is a key ecological process. An individual dispersal event has a source and a destination, both are well localized in space and can be seen as points. A probability to move from a source point to a destination point can be described by a dispersal kernel. However, when we measure dispersal, the source of dispersing individuals is usually an area, which distorts the shape of the dispersal gradient compared to the dispersal kernel. Here, we show theoretically how different source geometries affect the gradient shape depending on the type of the kernel. We present an approach for estimating dispersal kernels from measurements of dispersal gradients independently of the source geometry. Further, we use the approach to achieve the first field measurement of dispersal kernel of an important fungal pathogen of wheat, *Zymoseptoria tritici*. Rain-splash dispersed asexual spores of the pathogen spread on a scale of one meter. Our results demonstrate how analysis of dispersal data can be improved to achieve more rigorous measures of dispersal. Our findings enable a direct comparison between outcomes of different experiments, which will allow to acquire more knowledge from a large number of previous empirical studies of dispersal.

# Introduction

Individuals comprising biological populations often need to move from one location to a different location in order to survive and reproduce. Hence, dispersal is an important component of many life histories. Empirical characterization of dispersal has been a major theme in ecological research for a long time (for example Heald, 1913; Bullock et al., 2017). However, Bullock et al. (2017) found much fewer datasets describing plant dispersal than plant demography, likely indicating that "dispersal is notoriously difficult and resource-consuming to measure".

To measure dispersal, one needs a source of dispersing units and a method to record their displacement. Sources can be natural (e.g. a spawning site) or artificial (a planted patch). To record the displacement, studies on animal movement often use on mark-recapture experiments (Van Houtan et al., 2007; Carrasco et al., 2010), while plant studies commonly use seed traps or genotyping of seedlings around potential parents (Nathan et al., 2000; Goto et al., 2006). Spread of a plant pathogen can be recorded based on visual symptoms and genetic data (Solheim and Hietala, 2017). The appropriate methodology varies depending on the study system.

In the presence of a localized source, a dispersal gradient is expected: many individuals will stay close to the source while fewer individuals will travel further, leading to a decreasing dependency with distance. This pattern can be described mathematically by fitting a decreasing one-dimensional function to gradient measurements (e.g. review of Fitt et al., 1987; Ferrandino, 1996; Werth et al., 2006; Madden et al., 2007). However, the geometry of the source affects the shape of such gradients (Zadoks and Schein, 1979; Ferrandino, 1996; Cousens and Rawlinson, 2001). "Flattening" of gradients due to extended sources is noted qualitatively in previous studies (Zadoks and Schein, 1979; Ferrandino, 1996), but how exactly and how much does the source geometry affect the dispersal gradient?

A more rigorous mathematical description of dispersal is achieved with a dispersal kernel that represents a probability distribution of dispersal to a certain location relative to the source ("dispersal location kernel", Nathan et al., 2012). It is convenient to have a point source for an empirical characterization of dispersal kernels, because a dispersal gradient from a point source will have the same shape as the kernel. Zadoks and Schein (1979) proposed a rule of thumb, stating that a point source should have "a diameter smaller than 1% of the gradient length; but in many experiments, it is up to 5 or 10%". However, to determine whether the source is small enough so that the dispersal gradient captures the shape of the dispersal kernel, the size of the source should be compared with the characteristic distance of dispersal (i.e., the distance over which the dispersal kernel changes substantially), rather than the gradient length. This represents a challenge for the design of dispersal experiments that aim to achieve a point-like source, because whether or not the chosen source size is sufficiently small can be established with certainty only when the measurements are already conducted. As a result, "point" sources of various sizes are found in literature: an adult tree (Werth et al. (2006); cf. Cousens and Rawlinson (2001) presenting effect of tree canopy morphology on the shape of the gradient), circles of 80 cm (Skarpaas and Shea, 2007) and 25 cm diameter (Loebach and Anderson, 2018), 4 $m^2$ square (Emsweller et al., 2018), route of a single sampling dive (D'Aloia et al., 2015).

This challenge can be resolved using a modeling approach that incorporates the spread from any source geometry considering each point within the source as an independent point source (Clark et al., 1999). This would lead to a better, more mechanistic understanding of the dispersal as recommended by Bullock et al. (2006). While such approach has been suggested (e.g. by Greene and Calogeropoulos, 2002) it is rarely adopted, as demonstrated by the previous examples of various "point" sources and, for example, Bullock et al. (2017) who excluded line and area sources from their analysis, because those could not be compared to gradients from point sources.

We investigate the effect of source geometry on the shape of dispersal gradients considering three qualitatively different dispersal kernels: exponential, Gaussian, and power-law. We present possible simplifications, i.e. cases when a non-point source can be considered a point. We provide straightforward mathematical methods to take into account the source geometry in a more general case, when the simplifications are not possible. Finally, we present results of a case study, where we measured rain-splash driven asexual dispersal of a major fungal pathogen of wheat, *Zymoseptoria tritici*, characterizing its dispersal kernel for the first time in natural, field conditions.

# Theory

Dispersal location kernel describes the probability of dispersal from a source point ($p_s = (x_s, y_s)$) to a destination point ($p_d = (x_d, y_d)$) depending on the distance between the points ($r(p_s, p_d) = \sqrt{(x_d - x_s)^2 + (y_d - y_s)^2}$). Note the important difference between dispersal location kernel and dispersal distance kernel (Appendix A); we consider dispersal location kernel hereafter. In an ideal situation, the dispersal kernel could be measured in an experiment with a similar structure: a single point as a source of dispersing individuals and a single point for measuring dispersed individuals at each location. In such an experiment, the dispersal gradient, i.e. the spatial distribution of the dispersed individuals will correspond to the dispersal kernel. In reality, the source or the destination or both are usually areas, i.e. the source has a certain measurable area and the measurements at the destination are performed over a certain area. To describe such situations mathematically, we have to take sum over the individual points comprising the source to calculate their combined contribution to the dispersing population. Similarly, the sum over all points of the destination area gives the total population in the area. Population in a destination area $D$ after dispersal is then calculated as

$$N_1(S, D) = \iint\limits_D \iint\limits_S N_0(p_s)\kappa(r)\, dA_S\, dA_D \tag{1}$$

where $N_0(p_s)$ is the total dispersing population from $p_s$ (more precisely, the density function of the dispersing individuals within S), $S = \{p_s\}$ is the source area, $\kappa(r)$ is the dispersal kernel, and area integrals sum up the contributions of source points and destination points to the total observed population at the destination $D = \{p_d\}$. When the populations before dispersal ($N_0$) and after dispersal ($N_1$) are measured, the Eq. (1) becomes the function of only the kernel parameters, which can be estimated by fitting this function to the data.

Fitting a model with the above structure to empirical data can be challenging. Multiple integrations increase the computational demand making the process slower. Also, analytical solutions are more difficult to achieve with complex formulae. Therefore, simplifications could be useful to improve the analytical understanding and data analysis.

A common simplification is to fit a one-dimensional model to dispersal gradient data to estimate dispersal without accounting for the source geometry: $N_1 = C\kappa$. For example, a function of the form

$$N_1 = Ce^{-x/\alpha} \tag{2}$$

can be used to estimate the dispersal parameter $\alpha$ and the scale parameter $C$, in the case of an exponential kernel. The parameter $C$ here does not have a clear biological meaning. If both the source and the destination are points, the above approach provides a correct estimate for the dispersal parameter $\alpha$, because the function in Eq. (2) is the same as the exponential dispersal kernel [Eq. (3) in Box 1] up to a constant factor. The estimate of the parameter $C$ then contains the normalization factor of the kernel and other biological parameters, such as the population size at the source and the dispersal probability, which

**Exponential kernel** is defined as

$$\kappa_e(r,\alpha) = C_k e^{-r/\alpha} \tag{3}$$

where $k \in \{1,2\}$ is the number of dimensions, $r = r(p_s, p_d) >$ is the Euclidean distance from the source point $p_s = (x_s, y_s)$ to the destination point $p_d = (x_d, y_d)$ (in one dimension $y_s = y_d = 0$), and $C_k$ is normalization factor: $C_1 = 1/(2\alpha)$ and $C_2 = 1/(2\pi\alpha^2)$.

**Gaussian kernel** is defined as

$$\kappa_g(r,\alpha) = C_k e^{-r^2/2\alpha^2}, \tag{4}$$

where $C_1 = 1/\sqrt{2\pi\alpha^2}$ and $C_2 = 1/(\pi\alpha^2)$.

**Power-law kernel** is defined here as

$$\kappa_p(r,\alpha,\lambda) = C_k(\lambda + r)^{-\alpha}, \tag{5}$$

where $C_1 = (\alpha-1)\lambda^{\alpha-1}$, $C_2 = (\alpha-2)(\alpha-1)\lambda^{\alpha-2}/(2\pi)$. $\lambda$ is a scale parameter defining the finite starting point of the distribution in relation to $r^{-\alpha}$ distribution, which is not defined at $r = 0$.

Box 1

cannot be disentangled without additional information. This approach generally works for any kernel function (e.g., Gaussian or power-law kernels), when both the source and the destination are points.

However, when the geometry of the source and/or destination is more complex, the above approach may lead to wrong estimates. The parameter $\alpha$ estimated with this approach, may depend on the particular experimental design and have no relation to the actual kernel shape. However, the shape of the dispersal gradient does match to the shape of the dispersal kernel even when the source and the destination are areas in certain special cases. We discuss these special cases in relation to exponential, Gaussian, and power-law kernels (defined in Box 1).

If the source is extended in the direction of the measured gradient, and the underlying kernel is exponential, the Eq. (2) will still give a correct estimate of $\alpha$. This holds, because exponential kernels are memoryless (Box 2). This property allows to sum up all point sources within the source area along the $x$-axis to an equivalent virtual point

**Memoryless kernel.** Exponential kernels have a special feature: they are memoryless. To be memoryless means that setting any point along the gradient as a starting point, the tail of the distribution will have the same shape as entire distribution. This property explains why exponential kernel can be described unambiguously with the half-distance $\alpha \ln(2)$. From any point on an exponential gradient, moving $\alpha \ln(2)$ further along the gradient will decrease the density by half.

**Separable kernel.** Separable functions are those that can be expressed as a product of functions which depend on only one independent variable each, e.g. $f(x, y) = f_x(x)f_y(y)$. The shape of the dispersal gradient in the $x$-direction does not depend on the $y$-coordinate if the kernel is separable.

Most dispersal kernels found in the literature are neither memoryless nor separable (Nathan et al., 2012).

Box 2

source at $x = 0$ and in this way simplify the fitting process (see Fig. 1B). Thus, the extension of the source in the direction of the gradient will only add more power to the source but not change the shape of the gradient outside of the source, leading to a correct estimate of $\alpha$. This is not true for Gaussian and power-law kernels (Fig. 1 C, D).

If the extension of the source is in the other direction, perpendicular to the source, the simplified approach works with Gaussian kernel (Fig. 1C). Gaussian kernel is separable, which means that the shape of the kernel along $x$-dimension does not change when $y_s$ varies (Box 2). Hence, when measuring the dispersal along the $x$-axis, the extension of source along the $y$-axis only adds to the power of the source but does not modify shape of the gradient. Thus changing the source from a point to a thin line source perpendicular to the gradient leads to a different estimate of $C$ but the same estimate of $\alpha$. This holds for any separable kernel, but not for non-separable exponential or power-law kernels (Fig. 1B, D).

The situation is analogous when we consider extended destinations. Extended destination here implies that multiple measurements are conducted across the destination area in a uniformly random manner, and subsequently an average is taken over these measurements. When the kernel is exponential, both the source and the destination can be elongated in the direction of the dispersal gradient. An exponential function in Eq. (2) fitted to dispersal gradient data acquired in this way will have same dispersal parameter $\alpha$ as the dispersal kernel. In the case of separable kernels, both the source and the destination can be elongated perpendicular to the gradient and the gradient will have the same shape as the original kernel. If the geometry of the source or the destination is more complex, for example rectangles, the presented simplifications fail with each of the three kernels.

# Case study

## Materials and methods of the experiment

### Experimental design and disease measurements

We performed a field experiment to measure dispersal kernels of *Zymoseptoria tritici* in natural, field conditions within a wheat canopy. By analyzing the experimental data we demonstrate both benefits and drawbacks of simplifications related to the theory presented above. Winter wheat cultivar Runal was grown in 1.125 m × 4 m plots in Eschikon, Switzerland (coordinates: 47.449N, 8.682E). Inoculation was performed in inoculation areas in the middle of each plot (Fig. 2 A) on 17–18 May 2017 with 300 ml of spore suspension containing $10^6$ spores/ml of *Z. tritici* strain 1A5 (treatment A), strain 3D7 (D) or both strains (B, $5\times10^5$ sp/ml each) (Zhan et al., 2002). The pathogen strains were chosen because of their capacity to infect cultivar Runal and due to their contrasting production of pycnidia (asexual fruiting bodies) (Stewart et al., 2018). Control plots (C) were not inoculated. Five replicates of each treatment were assigned in a fully randomized design to 20 plots (Fig. B1). Further details of the experimental materials and methods are given in the Appendix B.

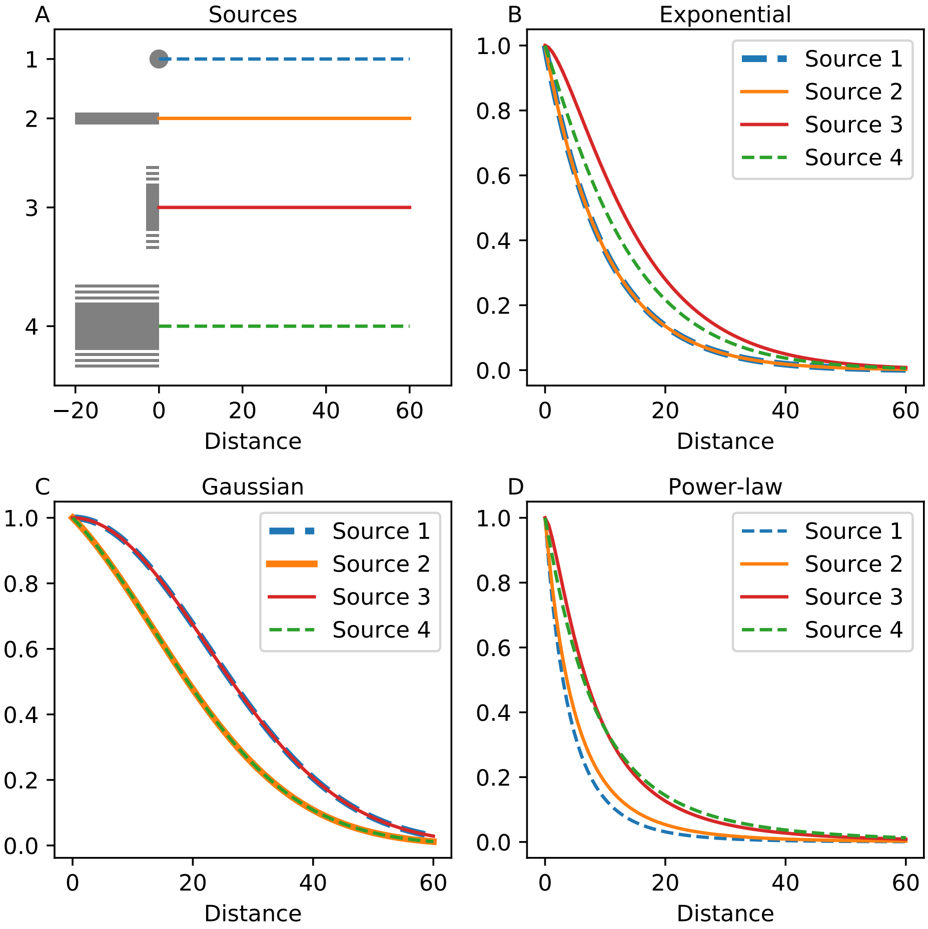


Figure 1: How source geometry affects the shape of dispersal gradients with different kernels (Box 1). (A) Four different sources: 1) point source; 2) Line source $x \in [-20, 0]$; 3) Line source, $y \in [-100, 100]$; 4) Rectangular area source, $(x, y) \in [-20, 0] \times [-100, 100]$. (B, C, D) Simulated gradients along the coloured lines in (A). Kernel parameters are chosen such that the mean dispersal distance is 20 units in all cases. Gradients are normalized to begin at 1. (B) With the exponential kernel ($\alpha = 10$), sources 1 and 2 result in identical gradients. (C) With the Gaussian kernel ($\alpha = 22.5$) the gradients are identical between sources 1 and 3 and between sources 2 and 4. (D) With the power-law kernel ($\alpha = 5, \lambda = 20$) all gradients are different.

Disease levels were measured within 10 cm-wide measurement lines across each plot (Fig. 2A), representing the destination area. Within each measurement line, multiple measurements were conducted in a uniformly random manner. In each measurement, *Z. tritici* incidence was assessed at the leaf scale by visual counting. After that, diseased leaves were collected and analyzed using the automated image analysis (Karisto et al., 2018, and Fig. 2 B-D) to obtain pycnidia counts as a measure of conditional severity. Success of inoculation was confirmed within the inoculated areas (source areas) on 14 June and primary disease gradients were measured in all measurement lines three weeks later, on 4 July.

## Statistical analysis

**Fitting disease gradients.** The disease intensity (numbers of pycnidia per leaf) at $t_1$ in a given measurement line is a result of dispersal of spores and successful infections from the source area to the area of the measurement line (i.e. the destination area). Assuming spatially uniform success of infections in all plots, the observed disease gradient is the result of the dispersal gradient of spores and it provides the effective dispersal gradient of the pathogen population. Following the equation (1) with the exponential kernel (that fits well when dispersal is driven by water splashes, according to Fitt et al., 1987; Saint-Jean et al., 2004), the dispersal process of the pathogen can be described mathematically using two area integrals: one over the source area and the other over the destination area. The disease intensity at the time $t_1$ in a measurement line at a distance $x^*$ (destination $D = \{(x_d, y_d)\} = [x^*-5, x^*+5] \times [b, w-b]$) from the inoculation area (source $S = \{(x_s, y_s)\} = [0, 40] \times [0, w]$) is given by

$$I_{t_1}(x^*) = \frac{I_0\beta}{10(w-2b)} \iint_D \iint_S \frac{e^{-\sqrt{(x_d-x_s)^2+(y_d-y_s)^2}/\alpha}}{2\pi\alpha^2} dA_S\, dA_D \qquad (6)$$

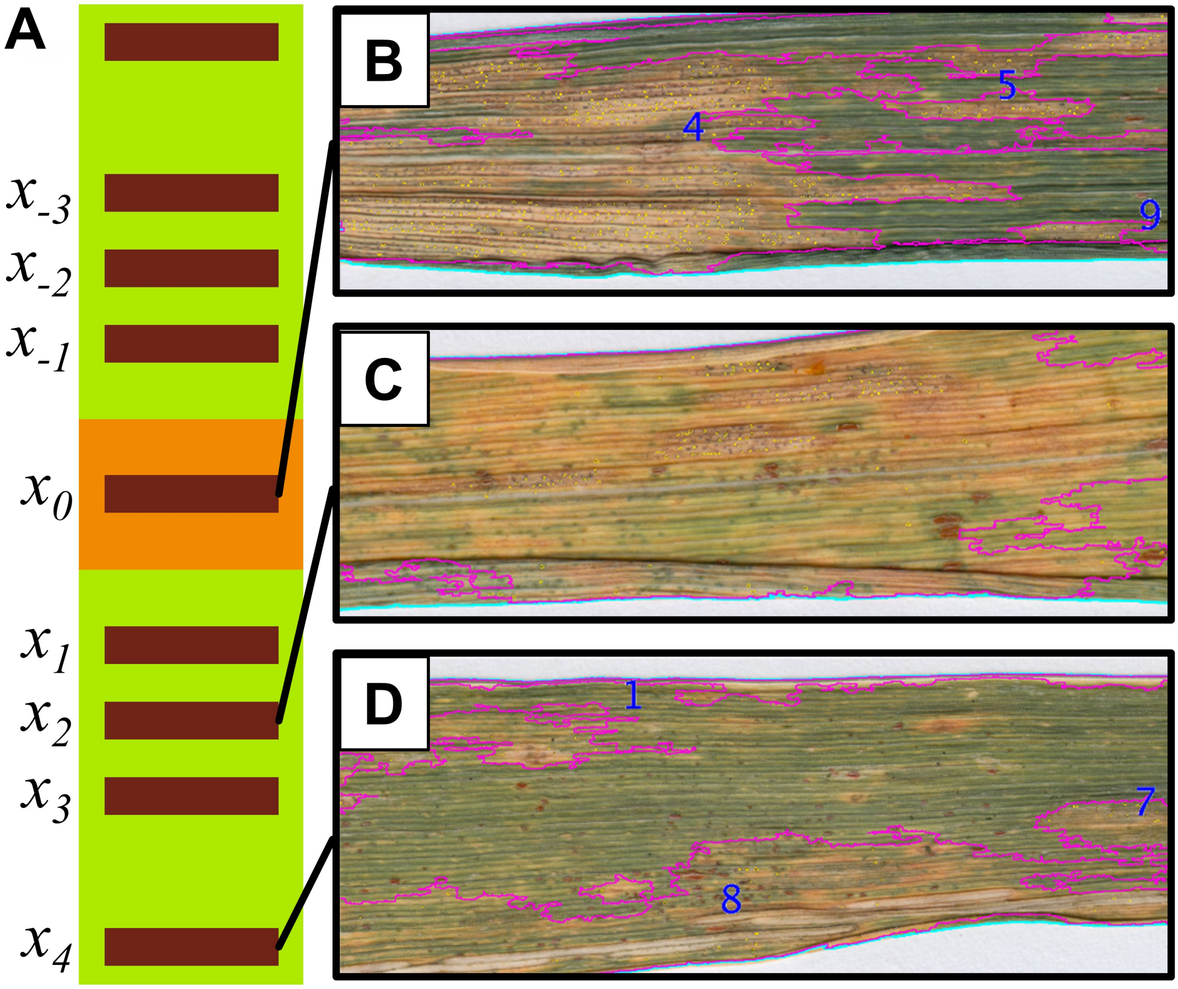


Figure 2: (A) Design of the experimental plot. 40 cm-wide inoculation area in the middle of the plot (orange). Distances from the middle of the inoculated area ($x_0$) to the middle of each measurement line were 0 cm, 40 cm, 60 cm, 80 cm, and 120 cm. (B, C, D) Overlay images illustrate the automated image analysis. Leaves collected from measurement lines $x_0$ (B), $x_2$ (C) and $x_4$ (D) of the treatment D at the sampling date $t_1$. Cyan, purple and yellow lines mark borders of leaves, lesions and pycnidia, respectively.

where $I_0$ is the intensity (pycnidia/leaf) in the inoculation area at $t_0$; $\beta$ is the transmission rate (unitless) describing how many new pycnidia there will be produced in the measured leaf layer per unit of measured intensity in the source leaves; $w = 112.5\,\mathrm{cm}$ is plot width; $b = 12.5\,\mathrm{cm}$ is width of the border excluded from measurement lines; and $\alpha$ is the dispersal parameter describing the dispersal kernel. The integration over the measurement line divided by area of the line, $10(w - 2b)$, gives the average disease intensity in a measurement line, representing uniform sampling of leaves.

Note that 10 cm width of measurement lines was practically the smallest possible width that could be achieved in our field measurements, because the foliage of even a single straw spans more than 10 cm, limiting the spatial resolution of our measurements. For this reason, we simplified the model by neglecting the width of measurement lines and assigning all disease intensity values recorded within each measurement line to the distance from the source that corresponds to the middle of the line. With this simplification, disease gradients were calculated according to

$$I_{t_1}(x^*) = \frac{I_0\beta}{w-2b}\int_b^{w-b}\int_0^{40}\int_0^{w}\frac{e^{-\sqrt{(x^*-x_s)^2+(y_d-y_s)^2}/\alpha}}{2\pi\alpha^2}dx_s\,dy_s\,dy_d. \tag{7}$$

We compare results obtained using Eq. (6) and Eq. (7) to results obtained using several simplifying assumptions. As implied by Madden et al. (2007) and the analysis of Fitt et al. (1987), dispersal is often modeled as a one-dimensional process. However, this simplification leads to correct estimates of the dispersal kernels only under certain circumstances, as discussed above in Theory section. To test what kind of error we make using the one-dimensional approach, we constructed the following function describing the dispersal according to an exponential kernel in one dimension:

$$I_{t_1}(x^*) = I_0\beta\int_0^{l_s}\frac{e^{-r(x_s,x^*)/\alpha}}{2\alpha}dx_s. \tag{8}$$

Here the integral takes sum over the length of the source along the plot, $l_s = 40\,\text{cm}$. When measuring the gradient outside of the inoculation area, ($x \geq 40\,\text{cm}$) the integral can be solved analytically and the equation (8) is greatly simplified:

$$I_{t_1}(x^*) = \frac{I_0\beta}{2}\left(e^{l_s/\alpha} - 1\right) e^{-x^*/\alpha} \tag{9}$$

leading to the same structure as in the equation (2) (with $C = \frac{I_0\beta}{2}\left(e^{l_s/\alpha} - 1\right)$). Equation (9) can be used directly to fit empirical disease gradients when considering the measurement line as a point. When taking into account the real width of the measurement line (10 cm) we calculate the mean intensity within a measurement line taking sum over it and dividing by the width. Continuing from the equation (9) the gradient is now calculated as

$$\begin{aligned} I_{t_1}(x^*) &= \frac{I_0\beta}{2}\left(e^{l_s/\alpha} - 1\right)\int_{x^*-5}^{x^*+5} \frac{e^{-x_r/\alpha}}{10} dx_r \\ &= \frac{I_0\beta\alpha}{20}\left(1 - e^{l_s/\alpha}\right)\left(e^{-5/\alpha} - e^{5/\alpha}\right) e^{-x^*/\alpha} \end{aligned} \tag{10}$$

which still retains the same form as the equation (2) but with a different constant than in the equation (9) ($C = \frac{I_{i,t_0}\beta\alpha}{20}\left(1 - e^{l_s/\alpha}\right)\left(e^{-5/\alpha} - e^{5/\alpha}\right)$). In this example, we retained the same shape of the gradient despite adding the spatial extension of both the source and the destination along the direction of the gradient. This highlights the memoryless property of exponential kernels (Box 2).

We constructed the functions similar to the one in Eq. (8) also using the Gaussian kernel and the power-law kernel (Box 1). These were compared to Eq. (9) based on AIC-score (Akaike information criterion Akaike, 1973). The differences in the AIC-scores were lower than the single parameter penalty, for this reason we conclude that neither of the three kernels can be considered superior.

**Estimation of dispersal and transmission parameters.** The data we used in the analysis were obtained in the following way. First, we collected incidence measurements (counts) and acquired conditional severity measurements using image analysis. Second, the severity data in each measurement line were multiplied with the corresponding incidence to obtain unconditional severity measurements. Third, we calculated mean for each measurement line to obtain five data points for each distance in each treatment. These means over each measurement line were used for fitting. The dispersal gradient functions (Eqs. (6), (7), (9), (10)) were fitted to the data to estimate $\alpha$ and $I_0\beta$.

To compare treatments and dispersal directions we used the bootstrapping approach. We re-sampled our collected samples with replacement to create a large set of variable bootstrap samples. Variation in the bootstrap samples reflects the variation that we expect to observe if the actual experiment was repeated several times (see for example Davison and Hinkley, 1997). Bootstrapping allowed us to model explicitly the variation related to incidence counts and variation related to leaf collection, separately of each other. This approach also allowed us to assess uncertainties in parameter estimates without making any assumptions about the distributions of the data or the parameter values.

We created 100 000 bootstrap samples for each measurement line. First, we re-sampled the infected leaves, i.e. generated 100 000 new samples of original size sampling from original leaf data with replacement. Second, we simulated the incidence counts on the measurement lines to create a distribution of incidence values. Based on the measured plant density ($730\,\mathrm{stems/m^2}$) we had on average 82 leaves within a measurement line. In this way, we simulated incidence counts with a population of 82 plants and all possible incidence values (from 0/82 to 82/82) 100 000 times each and recorded the "real" incidence value each time when the simulation gave the same incidence as in the observed data. Third, the mean of each bootstrap set of leaf severity was multiplied by incidence value drawn from the corresponding incidence distribution to obtain the unconditional mean severity for each measurement line. Fourth, we grouped these unconditional means of measurement lines into sets of five representing the five replicates. As a result, we obtained 100 000 bootstrap replicates of the entire experiment.

The one-dimensional disease gradient in equation (9) was fitted to each of the 100 000 bootstrap replicates. Two-dimensional disease gradient function in equation (7) was fitted to a subset of 10 000 bootstrap replicates. As a result, we obtained a large number of bootstrap point estimates of parameters $\alpha$ and $\beta I_0$ for each treatment and direction. These estimates were used to conduct statistical tests.

**Statistical tests.** Parameter differences were tested using a simple bootstrap hypothesis test (Davison and Hinkley, 1997, p. 162), where the observed difference between parameter values in different conditions is compared to a distribution of differences obtained with bootstrap samples. Significance level ($p$-value) of the test is calculated by dividing the number of cases where the difference in the test statistic $t_i$ is greater than or equal to the observed difference $t_{obs}$ by the number of bootstrap replicates (R) plus the observed case:

$$p = \frac{1 + \#\{t_i \geq t_{obs}\}}{R + 1} \tag{11}$$

If only a few bootstrap samples give a more extreme difference than the observed one, then the observed difference is considered significant. We tested the differences between parameter estimates using Eq. (11).

Additionally, we tested differences between $\alpha$- and $\beta I_0$-estimates simultaneously using a two-dimensional hypothesis test based on the joint distribution of differences in $\alpha$ and $\beta I_0$ (analogous to Johansson et al., 2014). A kernel density estimate of the joint distribution was obtained to define the degree of "extremity" of a point in the two-dimensional parameter space. The point reflecting observed difference in the two parameters was compared to the distribution of differences between bootstrap replicates. The observed difference is considered significantly different from zero, if it is located in a sufficiently sparse area, such that less than 5% of the bootstrap estimates are located in regions with equal or lower density (the “equidensity test”).

We present 95% confidence intervals for the parameters derived from the distribution of bootstrap results, i.e. the limits of 2.5th and 97.5th percentile of the distribution. Differences in disease levels between treatments A, B and D were tested at $t_0$, $x_0$ and $t_1$, $x_{\pm 1}$ with the Kruskal-Wallis test and the pairwise Dunn's posthoc comparison with the Bonferroni correction.

**Statistics implementation.** All data analysis was implemented in Python (versions 3.5.2 and 3.6.0) and the code is provided together with the data. Fitting was performed using lmfit-package (v. 0.9.10, Newville et al., 2014). Numerical integrations were implemented with 'quad' and 'dblquad' functions in scipy-package (v. 1.0.1, Jones et al., 2001–). Fitting of the two-dimensional functions (Eq. 6 and 7) was performed using the high performance computing cluster Euler of the ETH Zurich. Kruskal-Wallis test was conducted with 'kruskal' function in scipy-package and Dunn's test with function 'posthoc_dunn' in package scikit-posthocs (v. 0.3.8, Terpilowski, 2018).

## Results and discussion of the experiment

The inoculations with *Z. tritici* strain 1A5 (treatment A), strain 3D7 (D) and their mixture (B) were successful: at $t_0$ we observed increased disease levels in the inoculation areas of all three treatments (Fig. 3 A). Subsequently, disease gradients reflecting the dispersal gradients were obtained. At $t_1$, there was a gradient of disease severity from higher levels at $x_{\pm 1}$ to lower levels at $x_{\pm 4}$ (Fig. 3 B). Genotyping of re-isolated strains confirmed the successful spread (Appendix C). In total, 4190 plants were inspected for incidence counts; 2527 leaves were collected and analyzed using the digital image analysis. Total analyzed leaf area was $4.56\,\mathrm{m}^2$ and the total number of observed pycnidia was 1 131 608. The entire dataset including raw data, bootstrap replicates, best fitting parameter estimates and weather data is available in DATADRYAD (TBA after acceptance in the journal).

## Pathogen dispersal

Fitting the equation (7) to the observed disease gradients allowed us to estimate parameters $\alpha$ (dispersal parameter) and $\beta I_0$ (transmission rate × initial intensity at the source). In treatment A, estimates of $\alpha$ were very low and estimates of $\beta I_0$ were very high, compared to treatments B and D (Table 1). These unrealistic results (discussed below) were likely due to an insufficient disease intensity within the inoculation area and consequently a shallow gradient outside the inoculation area (Fig. 3 A and B). Less successful pathogen spread in treatment A than in other treatments was confirmed by comparing disease levels between treatments A and D at $t_0$, $x_0$ and at $t_1$, $x_{\pm 1}$. At $t_0$ $x_0$, the disease intensity was significantly lower in treatment A than in treatment D (Kruskal-Wallis test $p = 0.005$, pairwise Dunn's test $p = 0.004$). Further, at $t_1$ $x_{\pm 1}$ the intensity in treatment A was lower than the intensity in both treatments B and D (Kruskal-Wallis $p = 3.4 \times 10^{-26}$; Dunn's test A vs B $p = 2.6 \times 10^{-18}$; Dunn's test A vs D $p = 6.0 \times 10^{-24}$). For this reason, the next steps of analysis were conducted only using data obtained in treatments B and D.

Comparison of the best-fitting parameters (Eq. (7), Table 1) between the positive and the negative directions revealed no significant difference neither in treatment D (equidensity p-value: $p_{2D} = 0.21$, one-dimensional hypothesis test Eq. (11) for parameter $\alpha$: $p_\alpha = 0.17$, parameter $\beta I_0$: $p_{\beta I_0} = 0.13$, Fig. 4 A), nor in treatment B ($p_{2D} = 0.74$, $p_\alpha = 0.60$, $p_{\beta I_0} = 0.95$). This similarity between directions suggests isotropic dispersal.

As there was no significant difference between the two directions, we combined the data from the two directions and estimated the parameters using the combined dataset. We observed a significant difference between treatments B and D using the combined dataset ($p_{2D} = 0.014$, $p_\alpha = 0.020$, $p_{\beta I_0} = 0.018$, Fig. 4 B). Dispersal parameter $\alpha$ was

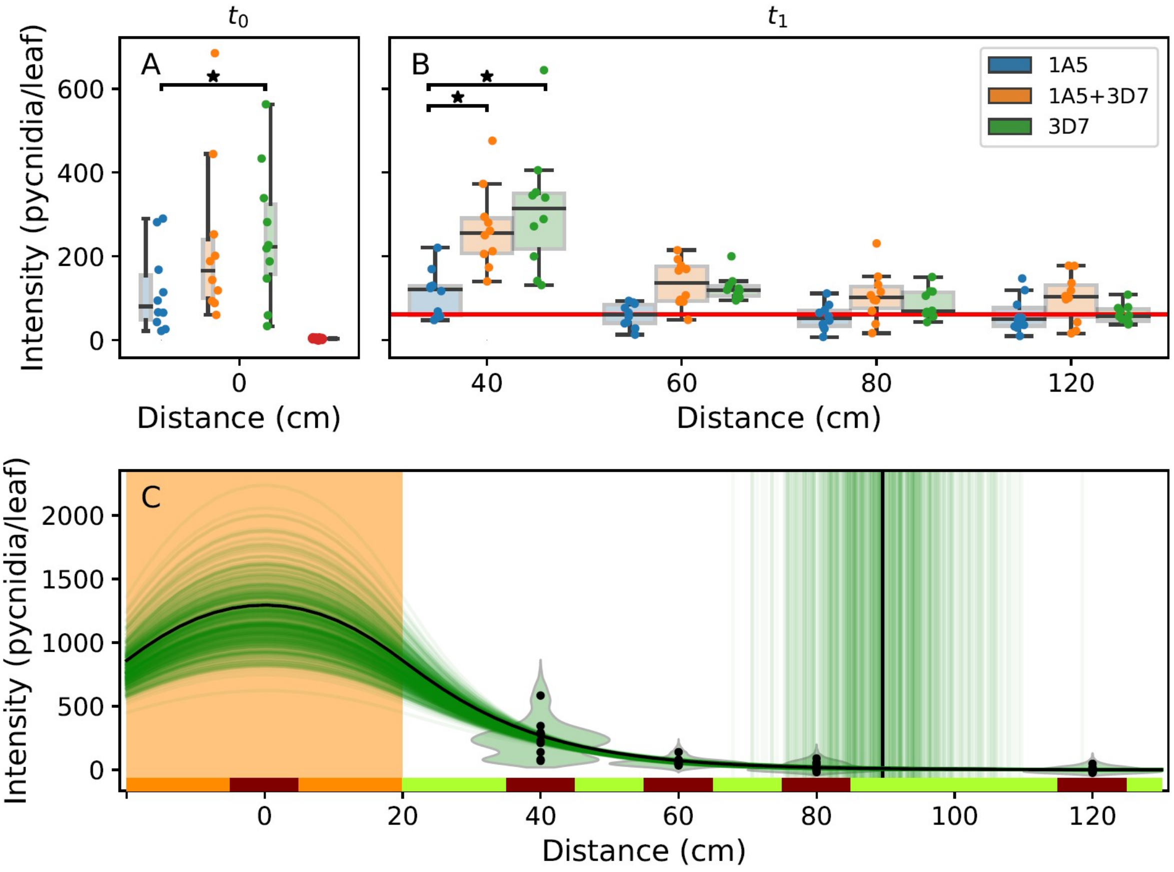


Figure 3: Disease levels and derived dispersal gradients. Data in blue shows treatment A (strain 1A5), in orange treatment B (mixture of 1A5+3D7), in green treatment D (strain 3D7), and in red treatment C (control). A: Disease levels in the inoculation area at $t_0$. Means of replicates shown with dots. Layers F and F-1 shown for treatments A, B, and C; while layer F-2 for treatment C (no disease on F-1 nor F). B: Disease gradients at $t_1$ after the spread event. Blue, orange and green dots and boxes show F leaf data. Red lines: mean over control plots. Black horizontal bars with asterisks show significant pairwise differences between treatments at $t_0$ and at $t_1$ (40 cm). C: Disease gradients of treatment D, using Eq. (9). Black dots show the observed replicate means (same as green dots in (B)) and black curve shows the best fitting gradient function. Vertical black line shows 99th percentile of the dispersal distance (Eq. (A3): $x_{99}$). Distributions of bootstrap data, disease gradients, and $x_{99}$ values of 500 bootstrap replicates shown in green violin plots, curves and horizontal lines, respectively.

higher in treatment B while $\beta I_0$ was higher in treatment D (Table 1). Thus the mixture of the two strains dispersed further, but imposed lower infection pressure on host plants than the strain 3D7 alone. Estimated values $\alpha_D = 13.5\,\mathrm{cm}$ and $\alpha_B = 21.4\,\mathrm{cm}$ fall close to the range estimated in spore dispersal experiments with artificial rain (Fitt et al., 1987). The one-dimensional estimates of $\alpha$ (Table 2) correspond to half-distances 10.5 cm (strain 3D7) and 17.0 cm (strain mixture), which match well to the range of 6–16 cm reported by Fitt et al. (1987). We conclude that experiments in controlled conditions translate well to the field conditions, at least when using simplistic one-dimensional fitting. The dispersal occurred due to two short rain showers (Fig. B2). During a longer rainy period the spores may disperse in multiple splash events leading to longer average dispersal distances and flatter disease gradients (Fitt et al., 1989).

Table 1: Best fitting parameters from fitting Eq. (7).

| Treatment<br>-direction | $\alpha$ (95% CI), cm | $\beta I_0$ (95% CI), pycnidia/leaf |
|---|---|---|
| A | | |
| -positive | 2.5 | 99992 |
| -negative | 4.9 | 3518 |
| -combined | 2.6 | 99975 |
| B | | |
| -positive | 23.1 (15.9 - 37.5) | 1271 (1052 - 1618) |
| -negative | 20.0 (14.8 - 28.4) | 1281 (1049 - 1586) |
| -combined | 21.4 (16.7 - 28.6) | 1271 (1095 - 1468) |
| D | | |
| -positive | 15.3 (12.3 - 19.7) | 1559 (1147 - 2112) |
| -negative | 12.0 (9.3 - 15.5) | 2387 (1613 - 3638) |
| -combined | 13.5 (11.5 - 16.1) | 1915 (1475 - 2430) |

Note: Confidence intervals (CI) derived from bootstrapping, but not for biologically implausible results of treatment A. Values of $\beta I_0$ were limited below 100 000.

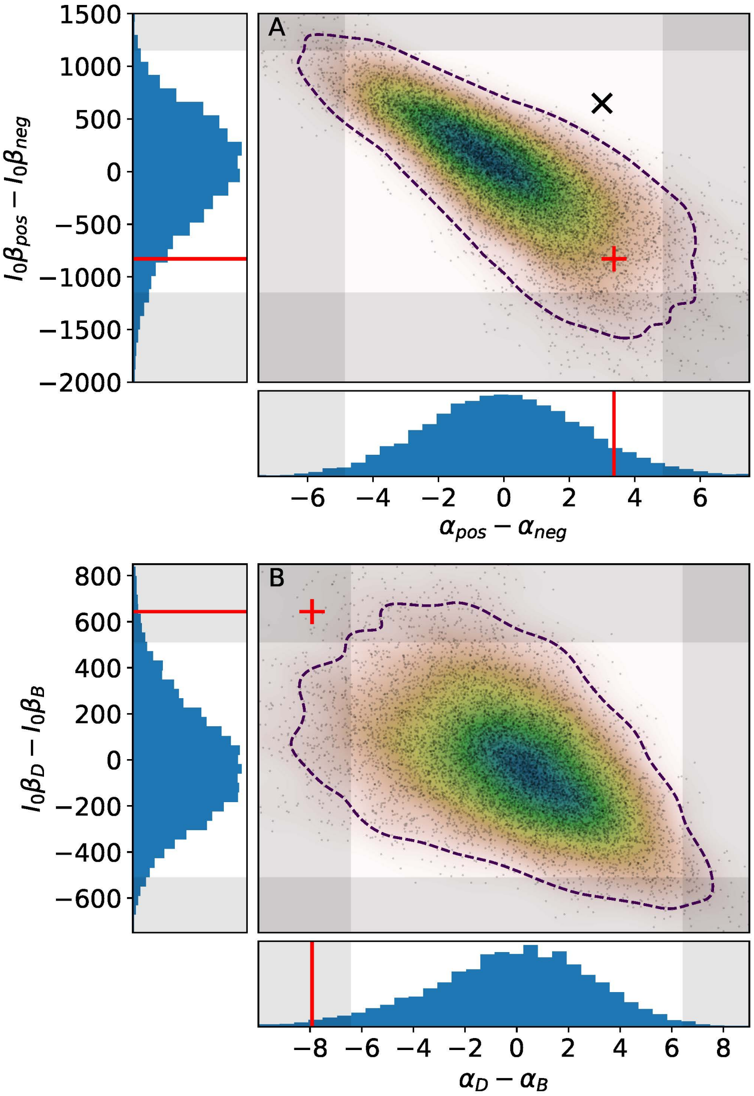


Figure 4: Visualization of one- and two-dimensional bootstrap tests, in histograms and main panels, respectively. (A) Comparison between two directions in treatment D. (B) Comparison between treatments D and B. Histograms show single parameter distributions while heat maps and black dots show joint distributions (10 000 replicates). Observations (red line; red plus) in the 5% extreme of the distribution (shaded area; outside of the dashed line), are considered significant. The differences are significant between treatments (B) but not between directions (A). Black cross in panel (A) shows a hypothetical observation where the difference would be deemed non-significant for each parameter separately (non-shaded area), but the joint test reveals a significant difference (outside of the dashed line).

## Disease transmission

Besides the dispersal parameter, we were also able to estimate transmission rate of the disease. The fitting yielded estimates of $\beta I_0$, from which we extracted $\beta$ by dividing $\beta I_0$ by estimates of $I_0$. We estimated $I_{0,B} = 227$ pycnidia/leaf and $I_{0,D} = 249$ pycnidia/leaf. Based on those we calculated $\beta_D = 7.7$ (unitless) and $\beta_B = 5.6$. Estimation of $\beta$ was only possible possible because we defined the scale parameter in a biologically meaningful manner.

Parameter estimates for the strain 1A5 were not realistic. It is not biologically plausible that spores of the strain 1A5 would disperse by only 2-5 cm, while spores of the strain 3D7 spread some 14 cm, because pycnidiospores of the two different strains are expected to have the same physical properties. Likewise, the transmission rate estimates of 1A5 were unrealistically high (Table 1). However, we inferred the parameter $\beta I_0$ for the strain 1A5 assuming that the physical process of spore transport via rain droplets is the same for the two strains. Under this assumption ($\alpha_A = \alpha_D = 13.5$ cm), we found that the dispersal of the strain 1A5 was isotropic ($p = 0.38$). Furthermore, with the two directions combined, we estimated $\beta I_0 = 349$ pycnidia/leaf and $\beta = 3.0$ for treatment A ($I_{0,A} = 118$ pycnidia/leaf).

The transmission rates of strain 1A5 and the strain mixture were lower than that of strain 3D7 (1A5 vs 3D7: $p = 1.00 \times 10^{-4}$, mixture vs 3D7: $p = 0.0498$). The intermediate transmission of the strain mixture is likely the result of a combination of transmission rates of the two strains. Strain 1A5 is known to produce fewer and smaller pycnidia than strain 3D7 on cultivar Runal in greenhouse (Stewart et al., 2018). Our results in field conditions are consistent with previous findings, as 1A5 produced fewer pycnidia within the source area (Fig. 3 A) and had a lower estimate of the transmission rate.

In this study system, the infectivity depends on weather conditions (Henze et al., 2007). Also, the disease levels within the source source and along the disease gradient were measured only on highest leaves, but dispersal occurred likely to and possibly from the lower leaf layers, which were not included in our analysis. Hence, the reported transmission rates should be considered relative to each other rather than as absolute values.

## Genotyping

In total, 153 individuals of *Z. tritici* were isolated from separate pycnidia on the leaves collected from the experimental plots and genotyped using targeted PCR-primers (Appendix C). The genotyping of the re-isolated strains supported the conclusions drawn from the phenotypic data. (i) The strains spread out from the source area; we detected them on the measurement lines (9 isolates out of 19 were detected as strain 1A5 at $x_{\pm 1}$ on treatment A, 45/55 as 3D7 on treatment D). (ii) There was a decreasing disease gradient; proportion of putative 1A5 and 3D7 isolates was lower further away from the inoculation area ($2 + 37$ (1A5 + 3D7) out of 49 isolates at $x_{\pm 1}$ vs $1 + 8/30$ at $x_{\pm 3}$, on treatment B). (iii) The strain 1A5 was transmitted less successfully than 3D7; proportion of the putative 3D7 strain individuals in treatment D was higher than proportion of 1A5 individuals in treatment A (see (i)) and the same effect was visible in treatment B (see (ii)). Thus, the genotypic and phenotypic data were in agreement.

## How good are the simplifications?

Simplifying the analysis of dispersal data by reducing the source and the destination to one dimension or even to points allows for simpler and faster calculation than in the two-dimensional analysis. However, these simplifications may lead to less accurate estimates of parameters. We compare the simplifications used based on (i) accuracy of the estimates, (ii) effect on statistical tests, and (iii) computational time.

(i) The values of parameter $\alpha$ were higher in one- than in two-dimensional approach (Table 2). If one would use the parameter estimates derived with one-dimensional approach (as found in literature) in a two-dimensional model, the dispersal would be overestimated as the values of $\alpha$ tend to be higher. This overestimation relates to the general "flattening" of disease gradients from extended sources. Compared to a point source, an extended source contributes to the gradient mostly through the tail of the dispersal kernel, which tends to be more flat than the beginning of the kernel. When an extended source is considered as a point, the flattening effect of the source geometry is accounted for in larger estimates of the width of the kernel.

On the other hand, the relationship between the population spread and dispersal parameter $\alpha$ is different between one- and two-dimensional models. This difference becomes clear for example when dispersal is described based on "mean dispersal distance" which is $\alpha$ for one-dimensional exponential kernel but $2\alpha$ for two-dimensional. In treatment D, the corresponding mean dispersal distances are $2\alpha_{2D} = 27\,\mathrm{cm}$ and $\alpha_{1D} = 15\,\mathrm{cm}$ - a considerable difference. Clearly, a one-dimensional model of dispersal should not be used for deriving dispersal distances on a population level. One should also be careful not to confuse the half-distance of an exponential dispersal location kernel with median dispersal distance of the population (see Appendix A).

(ii) All the statistical tests based on the bootstrap replicates gave similarly significant or non-significant results for one-dimensional and two-dimensional parameter estimates. (iii) Regarding the computational time, one-dimensional fitting of the disease gradient based on Eq. 9 to 100 000 bootstrap replicates was easily performed on a PC in a few hours, while fitting the two-dimensional function (Eq. 7) required a few days of computational time for only 10 000 replicates. When using the most complex function with two area integrals [Eq. (6)] it took more than 12 hours on a PC to obtain the estimates for only the observed data with one replicate.

Table 2: Comparison of parameter estimates between different functions.

| | Treatment B | | Treatment D | |
|---|---|---|---|---|
| Equation | $\alpha$, cm | $\beta I_0$, pycn./leaf | $\alpha$, cm | $\beta I_0$, pycn./leaf |
| Eq. 6, 2D extended destination | 21.4 | 1265 | 13.5 | 1889 |
| Eq. 7, 2D line destination | 21.4 | 1271 | 13.5 | 1915 |
| Eq. 10, 1D extended destination | 24.6 | 1123 | 15.1 | 1994 |
| Eq. 9, 1D point destination | 24.6 | 1131 | 15.1 | 2031 |

Note: Estimates from using the combined data of the two directions. "2D" and "1D" stand for two- and one-dimensional models, respectively.

## Discussion

We propose an approach for estimating dispersal kernel parameters, where the source geometry is explicitly incorporated in the model. This provides a solution for correcting inaccurate estimates caused by unjustified simplifying assumptions (See Fig. 1 and Table 2). This approach also provides a quantitative answer to the question "By how much and in which way does the source geometry affect the observed dispersal gradient?", instead of more qualitative statements regarding the "flattening" of the gradient with a larger source (Zadoks and Schein, 1979; Ferrandino, 1996; Cousens and Rawlinson, 2001). Using our method, we are able to relax the requirement of having a point source in a dispersal experiment. This helps designing experiments by increasing the power of the source and consequently the amount of collected data, which may be a limiting factor in many systems. With our approach, one can use results acquired from different experimental designs (e.g. those cited in Fitt et al., 1987) to estimate dispersal kernels in each case. Those can then be compared to each other directly, in contrast to dispersal gradients that reflect differences in experimental designs and cannot be compared if the designs are different. Most importantly, our approach allows to estimate actual kernel parameters in a much wider range of empirical studies than it was recognized previously, that includes all studies with spatially extended sources. In this way, "we can move from descriptions of pattern to a grasp of process" (Bullock et al., 2006).

We show, with simulations (Fig. 1) that different source geometries may lead to similar gradients when the kernel is either memoryless or separable. However, most kernels are neither memoryless nor separable, and thus distortions of the gradient shape are expected with varying source geometry. In any case, such simulations can be used when planning an experiment to guide the experimental design and to test predicted outcomes. Simulated outcomes of an experiment can also help to determine when the source can be considered a point and what kind of errors this simplification may introduce. Clearly, our two-dimensional models of dispersal are simplifications of the three-dimensional process (e.g. Vidal et al., 2018). If the third dimension is of great importance, as perhaps in aquatic environments or with tree canopies (Cousens and Rawlinson, 2001), modeling of source geometry and dispersal processes in three dimensions may be necessary.

We used this approach to analyze the data we acquired on *Z. tritici*, showing how different simplifying assumptions lead to different results. The explicit consideration of source geometry allowed us not only to estimate kernel parameter $\alpha$ but also a biologically relevant transmission rate $\beta$, instead of a meaningless normalization factor. Our experiment was conducted using an artificial experimental design with passively dispersing organisms, but similar approach can be used in observational studies in nature and with actively dispersing organisms whenever the source area can be characterized and the dispersal process can be described with the help of dispersal kernels.

In the common case of anisotropic dispersal (Soubeyrand et al., 2007), the validity of the simplifications based on separability or memorylessness of the specific functional forms of kernels, will generally not hold. However, the more general integration method that we presented can be modified to take into account the anisotropy of the kernel. In the modified model, the probability of dispersal from a source point to a destination point should depend not only on the distance between the points, as in our case, but also on the direction from the source to the destination. In this way, also anisotropic dispersal kernels can be inferred from measurements of dispersal gradients with the explicit consideration of the source geometry.

Our theoretical analysis shows that most pronounced differences between the dispersal gradients originating from different source geometries appear close to the source, while at larger distances from the source these differences disappear (Fig. 1B, C, D). The effect is seen in each of the three very different types of kernels, indicating that it is a universal feature. Therefore, even when the size of the source is much smaller than the gradient length, it could be that the size of the source is still comparable to the characteristic dispersal distance (i.e., the distance over which the dispersal kernel changes substantially). In this case, measurements close to the source will be substantially distorted due to the finite area of the source. Therefore, simple rules of thumb stating that to be considered as a point, the size of the source should be smaller than 1 % of the length of the gradient (Zadoks and Schein, 1979; McCartney et al., 2006), can be quite misleading, and result in inaccurate estimates of dispersal parameters. This emphasizes the importance of explicit modeling of the source geometry as we have done it here, considering that it is often the case that most measurements are conducted close to the source even when the overall gradient is long (Werth et al., 2006; Skarpaas and Shea, 2007; Loebach and Anderson, 2018).

Experiments that measure dispersal are difficult and laborious (Bullock et al., 2017). Geometry of the source, location of sampling areas, amount of sampling at different locations and other components of the experimental design may have a large effect on the precision and generalizability of the results. We support Skarpaas et al. (2005) calling for optimization of dispersal study designs by simulations, to make the most out of the effort.

# Acknowledgements

PK and AM gratefully acknowledge financial support from the Swiss National Science Foundation through the Ambizione grant PZ00P3_161453. Genetic Diversity Centre at ETH Zurich helped with the genetic analysis. PK and AM would like to thank Andreas Hund and Hansueli Zellweger for maintenance of the field experiment. FS would like to thank Christophe Montagnier, Sandrine Gélisse and Nicolas Lecutier for maintenance of the similar field experiment prepared at the facilities of INRA Bioger in Thiverval-Grignon, France.

# Appendix A: Dispersal distance kernel

**Dispersal distance kernel.** Dispersal kernels are not only used to describe distributions of locations of dispersed individuals, but also to summarize dispersal distances, such as mean distance travelled. Dispersal distance kernel is a one-dimensional function describing the probability of individuals to end up at a certain distance from the source. It can be derived from a two-dimensional dispersal location kernel by integrating it around the source, essentially by multiplying it with $2\pi r$ (Nathan et al., 2012). The shape of the dispersal location kernel can differ substantially from the shape of the dispersal distance kernel (e.g. Cousens and Rawlinson, 2001; Nathan et al., 2012). The dispersal distance kernel corresponding to exponential dispersal location kernel is given by

$$\kappa_{e,dist}(r) = 2\pi r \frac{e^{-r/\alpha}}{(2\pi\alpha^2)} = \frac{re^{-r/\alpha}}{\alpha^2} = \frac{r^{k-1}e^{-r/\alpha}}{\alpha^k\Gamma(k)}, \quad k = 2, \tag{A1}$$

where $\Gamma(2) = 1$ is the gamma function. Equation (A1) gives the one-dimensional gamma-distribution with the shape parameter $k = 2$ and the scale parameter $\alpha$.

It is important to keep in mind that means and medians of dispersal location kernels do not generally correspond to means and medians of population dispersal distances (i.e. means and medians of dispersal distance kernels). The example of the exponential kernel is of particular importance, as this kernel is often described with the half-distance. Considering the one-dimensional exponential kernel, the dispersal parameter $\alpha$ gives the mean and $\alpha \ln(2)$ (half-distance) gives the median of the distribution. However, in the case of the two-dimensional exponential location kernel the mean dispersal distance is not $\alpha$ but $2\alpha$ (i.e. mean of the gamma-distribution in Eq. (A1)). Furthermore, median, or any percentile, of the dispersal distance distribution can be determined by solving the equation

$$\int_0^{x_L} \kappa_{e,dist}(r)dr = 0.01L \tag{A2}$$

where $x_L$ is the $L^{\text{th}}$ percentile of dispersal distance ($L \in [0, 100]$). After integration, the Eq. (A2) reads

$$e^{-x_L/\alpha}\left(1+\frac{x_L}{\alpha}\right) = 1 - 0.01L. \tag{A3}$$

We solve Eq. (A3) numerically to obtain the median dispersal distance $x_{50} \approx 1.7\alpha \gg 0.69\alpha \approx \ln(2)\alpha$. Considering the limits of population dispersal, Golan and Pringle (2017) defined 99th percentile of the dispersal distance distribution as a limit for the long-distance dispersal in fungi. At $L = 99$, we find $x_{99} \approx 6.6\alpha$. These numbers have applied relevance, for example in conservation biology or in precision agriculture when a treatment is targeted to a certain fraction of a dispersing population. Mean or median dispersal distances or other characteristic numbers should be determined using the two-dimensional location kernel and the corresponding distance kernel.

The limit of long-distance dispersal ($x_{99}$, Eq. (A3)) corresponding to observed values of $\alpha$ is 90 cm for treatment D and 142 cm for treatment B. In an agricultural field, a visible disease focus (Zadoks and van den Bosch, 1994) and significant host damage (Shaw and Royle, 1993) would occur close to the source due to higher density, while the edge of the population is likely to incur less damage because of lower pathogen density. This hidden pathogen population in the tails of the distribution should be taken into account when attempting spatially targeted treatments, for example in precision agriculture involving focal fungicide spraying.

# Appendix B: Field experiment

## The study system

*Zymoseptoria tritici* (formerly *Mycosphaerella graminicola*) is a major fungal pathogen of wheat in temperate areas (Jørgensen et al., 2014; Dean et al., 2012). It causes the disease Septoria tritici blotch (STB), which is visible as brownish lesions on wheat leaves. The lesions reduce the photosynthetic ability of the host and cause yield losses of 5-10%, even when resistant cultivars and fungicides are used in combination. Annually in Europe some 1.2 billion dollars are spent for fungicides mainly aimed to control STB (Torriani et al., 2015).

Infection by *Z. tritici* begins when spores deposited on wheat leaves germinate and penetrate the leaves through stomata (Kema et al., 1996). The fungus grows in the apoplast for several days without visible symptoms (Duncan and Howard, 2000). In optimal conditions, necrotic lesions appear in the invaded host tissue after about ten days and asexual fruiting bodies called pycnidia begin to form (Kema et al., 1996; Duncan and Howard, 2000). Asexual pycnidiospores ooze from pycnidia within water-soluble cirri and are spread mostly by rain splash. If a spore falls within the wheat canopy and stays on a healthy leaf instead of being washed down, it can infect new host tissue either on the same or on the neighbouring plants. Upon successful infection, the spore again creates a lesion and produces new pycndia within the lesion. The pathogen undergoes several rounds of asexual reproduction per growing season. Zhan et al. (1998, 2000) estimated that ≈66% of infections on flag leaves came from asexual spores, leading to conclusion that asexual reproduction is the most important source of infection on flag leaves.

Initial inoculum by air-borne ascospores is often considered uniform across a wheat field and not a limiting factor for epidemics (Morais et al., 2016). Therefore, much of interest has been on vertical dispersal of the spores from initial infection of seedlings to emerging leaf layers (Shaw, 1987; Lovell et al., 1997; Bannon and Cooke, 1998; Lovell et al., 2004; Vidal et al., 2018). Interaction between the pathogen and the host has been described by Robert et al. (2018) as a race, where the pathogen need to "climb" up to the next leaf layer before current layer becomes senescent and its resources are depleted. The plant, in turn, "tries" to save the newly emerging leaves making them escape the infection by fast stem elongation.

However, horizontal dispersal greatly influences the ability of a particular clonal lineage to grow in numbers and can play a major role in the dynamics of emerging fungicide resistance or ability to overcome host resistance genes. Resistance gene pyramids or homogeneous host cultivar mixtures may select for multiple virulences in a single pathogen strain. Thus, spatial adjustment of control strategies has been suggested as a potentially more sustainable solution and optimal spatial scale of such heterogeneity is determined by the spatial scale of the pathogen's horizontal spread (Mundt and Browning, 1985; Brophy and Mundt, 1991; Newton et al., 2009; Sapoukhina et al., 2010; Newton and Guy, 2011; Djidjou-Demasse et al., 2017). Dispersal of *Z. tritici* or similarly spreading species *Parastagonospora nodorum* (formerly *Septoria nodorum*) has been studied in controlled conditions using either infected straw or spore suspension together with artificial rain and often spore traps (Brennan et al., 1985; Saint-Jean et al., 2004; Vidal et al., 2017). Bannon and Cooke (1998) studied the effect of wheat-clover intercrop on dispersal from plates via artificial rain and merely noted a reduction of dispersed spores at the 15 cm distance. No experiment has so far been conducted in field conditions to estimate parameters of dispersal kernel of the disease spread from infected plants to the surrounding healthy canopy.

Spatial spread directly influences the number of new hosts that a pathogen can potentially invade and it also affects the spatial distribution of the pathogen population. For a polycyclic pathogen, such as *Zymoseptoria tritici*, small differences in monocyclic spread can result in considerable differences in the epidemic outcomes after multiple disease cycles. Thus, understanding the mechanisms and the scale of the spread will improve our ability to predict and control potentially disastrous epidemics of the disease.

## Plant materials and agronomic practices

The experiment was performed at the Field Phenotyping Platform (FIP) site of Eschikon Field Station of the ETH Zurich, Switzerland (Kirchgessner et al., 2017). Experimental plots were sown with winter wheat (*Triticum aestivum*) cultivar Runal on 1 November 2016. Sowing density was 440 seeds/m$^2$ and the observed stem density on 19 June 2017 was 730 stems/m$^2$. Field maintenance included herbicide Herold SC (0.6 l/ha; Bayer) on 2 November 2016, and stem shortener Moddus (0.5 l/ha; Syngenta) on 13 April 2017. Fungicide Input (1.25 l/ha; Spiroxamin 300 g/l, Prothioconazol 160 g/l; Bayer) was applied on 13 March 2017 to suppress the background infection.

Similar experiment was prepared also at the facilities of INRA Bioger in Thiverval-Grignon, France (coordinates: 48.840N, 1.952E). The experimental design was similar with minor modifications. Due to unconducive weather conditions the inoculation failed to produce measurable primary disease gradients. Therefore, the data is not presented.

## Experimental design

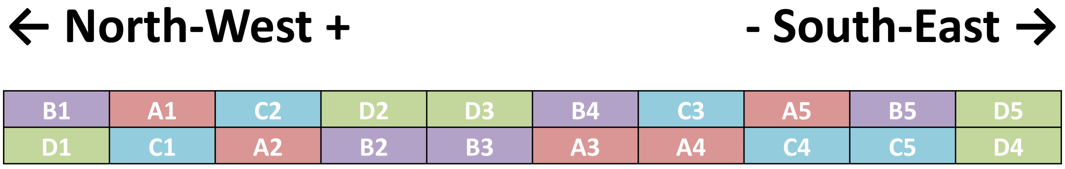
The experimental plots were 1.125 m × 4 m rectangles consisting of nine long rows of wheat with 12.5 cm spacing between the rows. Plots were randomly assigned to four treatments with five replicates of each treatment as shown in Fig. B1. The four treatments were: inoculation with strain ST99CH_1A5 (short identifier 1A5, treatment A), strain ST99CH_3D7 (3D7, treatment D), both strains (B) and no inoculation (C). Strains were collected in Switzerland in 1999 as described by Zhan et al. (2002) (see also www.septoria-tritici-blotch.net/isolate-collections.html).

In each plot, there was a 40 cm-wide inoculated area across the plot in the middle. Disease measurements were conducted in the middle of the inoculated area ($x_0 = 0\,\mathrm{cm}$) and at eight locations outside of the inoculated area, four on each side at distances $x_{\pm 1} = \pm 40\,\mathrm{cm}$, $x_{\pm 2} = \pm 60\,\mathrm{cm}$, $x_{\pm 3} = \pm 80\,\mathrm{cm}$ and $x_{\pm 4} = \pm 120\,\mathrm{cm}$ from the center of the inoculated area (see Fig. 2 A). A measurement line consisted of a line across the experimental plot at the given distance, extended with 5 cm margins along the plot and excluding 12.5 cm borders at the edge of the plot to reduce edge effects. Measurements were conducted uniformly over space in the rectangular area of each measurement line.

## *Z. tritici* inoculation

Inoculum was prepared by growing the fungus for seven days in yeast-sucrose-broth (https://dx.doi.org/10.17504/protocols.io.mctc2wn). The liquid culture was then filtered, spores were pelleted in centrifuge and re-suspended into sterile water to harvest blastospores. The washed spore suspension was diluted to achieve the concentration of $10^6$ spores/ml. For treatment B the final spore concentration was $10^6$ spores/ml so that each strain was present with the concentration of $5 \times 10^5$ spores/ml. Finally, we added 0.1 % (v/v) of Tween20 and kept the inoculum suspension on ice until spraying.

Inoculation was performed by spraying 300 ml of the spore suspension onto the inoculation site of each plot using a hand-pump pressure sprayer. The plots were inoculated during the late afternoon to avoid direct sunlight. All treatments were inoculated with



| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| B1 | A1 | C2 | D2 | D3 | B4 | C3 | A5 | B5 | D5 |
| D1 | C1 | A2 | B2 | B3 | A3 | A4 | C4 | C5 | D4 |

Figure B1: Arrangement of the plots in the field.

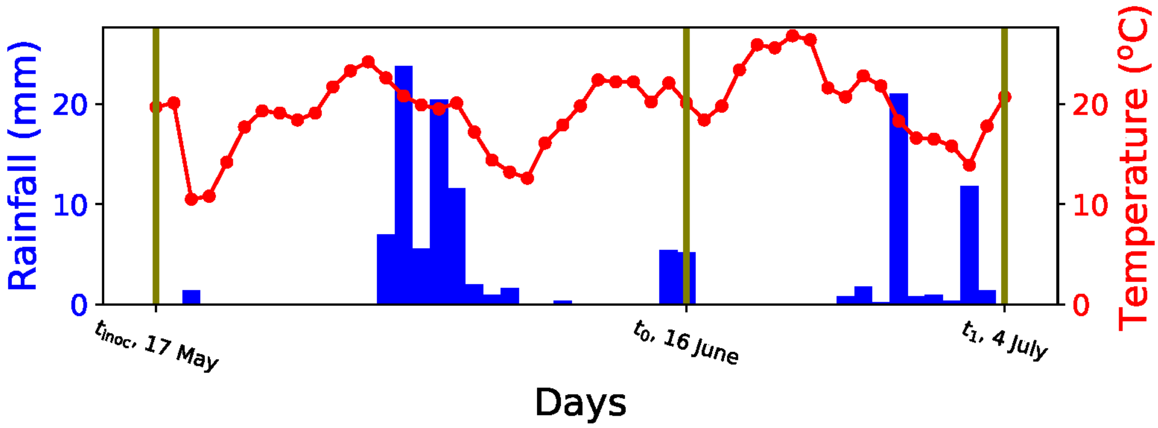


Figure B2: Weather conditions during the experiment. Daily precipitation in the blue bars, daily mean temperature in the red line, dates of inoculation, sampling dates $t_0$ and $t_1$ are shown using vertical lines.

the same sprayer, which was rinsed with water and 70% ethanol to clean all parts before inoculating each treatment. Entire canopy within the inoculation area was inoculated until runoff. During spraying, the inoculation area was bordered with plastic sheets to avoid the spillover of the inoculum to other plots. After spraying, the border sheets were folded over the canopy to enclose the plants in plastic tents maintaining high humidity overnight. The tents were removed early next morning to avoid overheating of plants. The inoculation was repeated next evening in the same manner. Pictures of inoculation are shown in Appendix D.

First attempt to inoculate was made on 5 and 6 April, when F-3 layer (the third leaf layer below flag leaf) was mostly emerged (approximate growth stage, GS 22, Zadoks et al. (1974)), and inoculation success was assessed on 24 April and again on 3 May. Due to cold weather the inoculation success was extremely low: we observed low levels of disease in the F-3 leaf layer and the plants were in the beginning of stem elongation (F-1 emerging, GS 35). Average incidence in F-3 layer in the inoculated area 3 May was 6.1%, 2.9%, 0% and 4.9% for treatments A, B, C, and D, respectively. We considered the inoculation as failed, because the secondary spread from such low initial infection levels would likely cause only negligible gradients and due to stem elongation and senescence the highest leaf layers would likely escape the spread (Robert et al., 2018). We decided to inoculate again the higher leaf layers to achieve stronger, measurable disease gradients. Dates of this main inoculation were 17 and 18 May 2017, when flag leaves had already emerged (GS 39–41).

## Assessment of the disease gradient

The disease assessment combining incidence and severity measurements was performed twice. At $t_0$, on 14 June 2017 (GS 70) only the inoculation areas were assessed to confirm the success of inoculation across the measurement line $x_0$. Flag leaves outside the inoculation area were visually confirmed to be healthy without further assessment. At $t_1$ on 4 July 2017 (GS 85) all measurement lines of treatments A, B and D were sampled. One line on each plot of treatment C was assessed for reference.

At $t_0$, incidence of the disease was measured at the leaf scale in the following manner. Thirty to forty straws were inspected on each measurement line. The highest diseased leaf layer was recorded for each straw. The leaves lower than that were assumed to be diseased as STB is usually more prevalent in the lower leaf layers. Additionally, naturally senescent leaf layers were recorded. In this way, incidence was estimated for all non-senescent leaf layers. After estimating the incidence, eight infected leaves were collected from up to two consecutive leaf layers that had incidence higher than 20%. The collected leaves were then mounted on paper sheets and scanned with 1200 dpi resolution. The resulting images were analyzed using automated image analysis method measuring two aspects of severity of the infection that represent the host damage and pathogen reproduction, as described in Karisto et al. (2018). Host damage was measured as the percentage of leaf area covered by lesions (PLACL) and pathogen reproduction as the pycnidia count per leaf. The sampled leaf layers at $t_0$ were the flag leaf layer (F) and the layer below it (F-1).

At $t_1$, the plants were already mostly chlorotic and hence the incidence measurement was not possible in the field. Instead, we collected about 24 leaves from each measurement line at random. The leaves were taken into lab and each leaf was visually inspected for the presence of pycnidia. Incidence was recorded based on the presence of pycnidia on the collected leaves and only leaves with pycnidia were scanned for severity measurement. Due to vast chlorosis, the measurement of host damage was considered unreliable and only pathogen reproduction was used in the subsequent analysis. Thus, we measured the disease intensity as numbers of pycnidia per leaf.

We estimated number of asexual reproduction and dispersal events between $t_0$ and $t_1$ using the following arguments. First, based on the data from Shaw (1990) regarding latent period lengths of *Z. tritici* at different temperatures (as revisited in Karisto et al. (2018), Fig. A1), latent period after inoculation was approximated to be longer than 20 days (average daily temperature during first 19 days was 19 °C). Thus, there was likely no spread from inoculation area during the rainy period at 13–17 days after inoculation (dai) (Fig. 2). This was confirmed with visual assessments of the inoculation areas on 8 June (22 dai), when we observed few tiny lesions and mostly no pycnidia, concluding that substantial spread had not been possible by then. Second, at $t_0$ (28 dai) there was substantial disease (Fig. 3 A) in the inoculation areas and there were two strong showers in the night after $t_0$. Third, there was no rain for one week before nor after $t_0$. Thus, we conclude that there was most likely only one asexual spread event at $t_0$, which caused the disease gradients outside of the inoculation areas at $t_1$ (38 dai).

In summary, the inoculation was successful and led to increased levels of disease in the inoculation areas after a latent period of 3–4 weeks, at $t_0$. Three weeks later, at $t_1$ there were clearly visible symptoms outside of the inoculation area. The observed symptoms at $t_1$ can be entirely accounted to the raining event and consequent asexual spread of the pathogen at $t_0$.

## Discussion of experimental aspects

**Measurement of pathogen population, not host damage.** Our estimates of dispersal kernel correspond to the effective dispersal of the pathogen population, instead of the basic dispersal kernel of all spores. Difference between these may arise from possibly density-dependent post-dispersal mortality (Nathan et al., 2012; Klein et al., 2013). At high spore densities, that can be found close to the source, leaves can become saturated with the infection leading to a decreased infection efficiency of spores (Karisto et al., 2019). In the tail of the distribution the density is however so low, that saturation may not be a major factor. Dispersal of spores could be measured with spore traps placed within the canopy. However, that would leave open how many of the spores actually attach to healthy plants, how many of them are successful, and how much the established population disperses. Using healthy plants as spore traps leads to the measurement of a more epidemiologically relevant combination of dispersal and infection processes.

Measurement of pathogen reproduction in terms of numbers of pycnidia per leaf gives us a proxy of the pathogen population size at each measurement point. Traditionally, plant diseases are observed visually based on host damage, but novel methodology allows for a different approach. While host damage is an important agronomic factor, pathogen reproduction is more relevant for pathogen ecology and evolution. Moreover, pathogen reproduction is more powerful than host damage for predicting the host damage at a later time point (Karisto et al., 2018).

**Sampling distances.** The measurement lines were at closest 20 cm (±5 cm) from the edge of the inoculated area. Measuring the gradient closer to the source and even inside the source could make the fitting more accurate, because differences between gradients would be easier to detect closer to the source area where variations are more pronounced.

However, closer to the source, the reliability of data might suffer from saturation and also from dispersal via direct contact (Fitt et al., 1989). Optimal measurement distances have to be determined for each study system based on biological understanding and prior knowledge about the dispersal kernel.

We measured the disease also inside the inoculation area, but those were excluded from fitting to include only secondary infections. The increase in the disease intensity at $x_0$ from $t_0$ to $t_1$ was not only due to secondary infections but also from extremely long latent periods (Karisto et al., 2019). Additionally, possible saturation was strongest at $x_0$. Therefore, measurement of newly spread infection was not possible inside the source area.

# Appendix C: Genotypic test with PCR

**Primer design.** We designed four primer pairs targeted at each of the two strains. The primers were aimed to be first fully specific for the target isolates 1A5 and 3D7 within the set of four commonly used lab strains 1A5, 1E4, 3D1, 3D7 and second as specific to the target strain as possible in the field. Specificity here means that the primers designed for 1A5 should produce an amplicon in PCR only with 1A5 genome and not with other strains. Strain specific primers would allow for a convenient detection of the focal sub-population after the experiment as in a mark-recapture experiment.

To design the primers, we used presence-absence data of predicted genes from Hartmann and Croll (2017). We chose target regions that were present in the target strain (either 1A5 or 3D7) and absent in the other three isolates (1E4, 3D1, and either 3D7 or 1A5). From those potential targets, we selected ten least frequent regions in the 27 Swiss isolates analyzed by Hartmann and Croll (2017). After selecting the target regions, we designed four primer pairs that would be suitable for high throughput qPCR in same conditions: amplicon length 100-150 bp, melting temperature around 60 °C. The primers were designed to amplify regions in different chromosomes of the target strain to minimize the possibility of finding all of them in a single strain in the field. Details of the designed primers are given in Table C1.

**Validation of primer specificity.** First validation of the primers was done with qPCR among the four strains 3D7, 1A5, 3D1 and 1E4 (Tables C2 and C3, Figures C1 and C2). Successful amplification of the target DNA and no amplification on non-target DNA suggested that each of the eight primer pairs was specific to their target strain among the four strains, indicating successful primer design based on the genomes.

Primers' specificity was then validated in a natural population using multiplex-PCR (Table C4, Table C5) combining each specific primer pair with a primer pair that is

Table C1: Primers

| Primer | Sequence (5'-3') | $T_m$ (°C) | Amplicon (bp) |
|---|---|---|---|
| 1A5.5_FWD | AGC AGT CCT CGT AGC ATA ACG | 59.93 | 135 |
| 1A5.5_REV | GAC CTC CTA TGA TGC GGC AA | 59.89 | |
| 1A5.6_FWD | GGG AGG CCC TGG TTG ATT AC | 60.11 | 135 |
| 1A5.6_REV | CTT GTA AGA GCG AGG GGC AA | 60.04 | |
| 1A5.9_FWD | TTC TCT CTA TAG CCC GCC CT | 59.52 | 137 |
| 1A5.9_REV | GAG TAG ACT CTA GAG GAA ACC TAG T | 58.11 | |
| 1A5.10_FWD | CTC GGC CAG GAA GTG ATT GT | 60.04 | 137 |
| 1A5.10_REV | GAG CAG TGG AGC CCA AGA AT | 60.03 | |
| 3D7.2_FWD | CGA CAT CGG TTC AGA GAT GGA A | 60.16 | 146 |
| 3D7.2_REV | GTA CCT TCG ATT CGT GCG GT | 60.46 | |
| 3D7.6_FWD | CTT GGG TGC AAT GAA CGG AC | 59.76 | 139 |
| 3D7.6_REV | TGA GAA ACA GTC GTG TGG CA | 59.82 | |
| 3D7.9_FWD | CAG CTC GAC TTG TGA GTC CT | 59.4 | 136 |
| 3D7.9_REV | CGT GCA AAC GCT GCA TGA T | 60.15 | |
| 3D7.10_FWD | GGT GCC CTC GTC GGA ATA C | 60.23 | 123 |
| 3D7.10_REV | TTG GGG AAG GAG ACC ATT CG | 59.38 | |
| Zt_gen_FWD | ATT GGC GAG AGG GAT GAA GG | 60.5 | 101 |
| Zt_gen_REV | ATT TTC GTG TCC CAG TGC GTG TA | 60.5 | |

Note: Primer name starts with the target strain followed by chromosome number, except for Zt_gen primers. The latter designed by Duvivier et al. (2013).

Table C2: qPCR reaction mix, 20$\mu$l

| Reagent | Concentration | Volume ($\mu$l) | Final concentration |
|---|---|---|---|
| Water | | 6 | |
| EvaGreen Mix | 5X | 4 | 1X |
| Strain.Chr_FWD | 1$\mu$M | 2 | 200nM |
| Strain.Chr_REV | 1$\mu$M | 2 | 200nM |
| Target DNA | 1ng/$\mu$l | 6 | 6ng/20$\mu$l |

Note: We used qPCR mixture HOT FIREPol EvaGreen qPCR Mix Plus (ROX) (Solis BioDyne).

specific to *Z. tritici* generally (Zt_gen primers) (Duvivier et al., 2013). Zt_gen provided a positive control for success of the PCR: if it created an amplicon, the reaction was successful. Primers were tested against 37 natural strains isolated from the control plots of the experiment. Reaction with primers 1A5.9 did not work reliably, indicated by the lack of Zt_gen amplicon. Numbers of false positives for other primer pairs were 4, 8, 20,

Table C3: qPCR reaction cycles

| Step | Temperature (°C) | Time (s) |
|---|---|---|
| 1 | 95 | 900 |
| 2 | 95 | 15 |
| 3 | 60 | 20 |
| 4 | 72 | 20 |

Note: Steps 2–4 repeated 40 times

Table C4: PCR reaction mix, 20$\mu$l

| Reagent | Concentration | Volume ($\mu$l) | Final concentration |
|---|---|---|---|
| Water | | 6.34 | |
| KAPA Buffer | 2X | 10 | 1X |
| Zt_gen_FWD | 10$\mu$M | 0.5 | 250nM |
| Zt_gen_REV | 10$\mu$M | 0.5 | 250nM |
| Strain.Chr_FWD | 5$\mu$M | 1 | 250nM |
| Strain.Chr_REV | 5$\mu$M | 1 | 250nM |
| KAPA3G Polymerase | 2.5U/$\mu$l | 0.16 | 2U/100$\mu$l |
| Spore solution | $10^4$-$10^6$sp/ml | 2 | |

Note KAPA3G Plant PCR Kit (Kapa Biosystems).

Table C5: PCR reaction cycles

| Step | Temperature (°C) | Time (s) |
|---|---|---|
| 1 | 96 | 180 |
| 2 | 95 | 20 |
| 3 | 60 | 15 |
| 4 | 72 | 15 |
| 5 | 72 | 30 |

Note: Steps 2–4 repeated 35 times.

13, 12, 7 and 6 for 1A5.5, 1A5.6, 1A5.10, 3D7.2, 3D7.6, 3D7.9 and 3D7.10 respectively (Figures C3, C4, C5, C6, C7, C8, C9, C10). Importantly, none of the false positives of 1A5.5 and 1A5.6 overlapped with each other, hence using combined data of those two gave no false positives. The six false positives of 3D7.10 were amplified with all the other 3D7-primers and none of the 1A5 primers. Thus, it is possible that they were the actual strain 3D7 either left on the field from previous years of field experiments or it

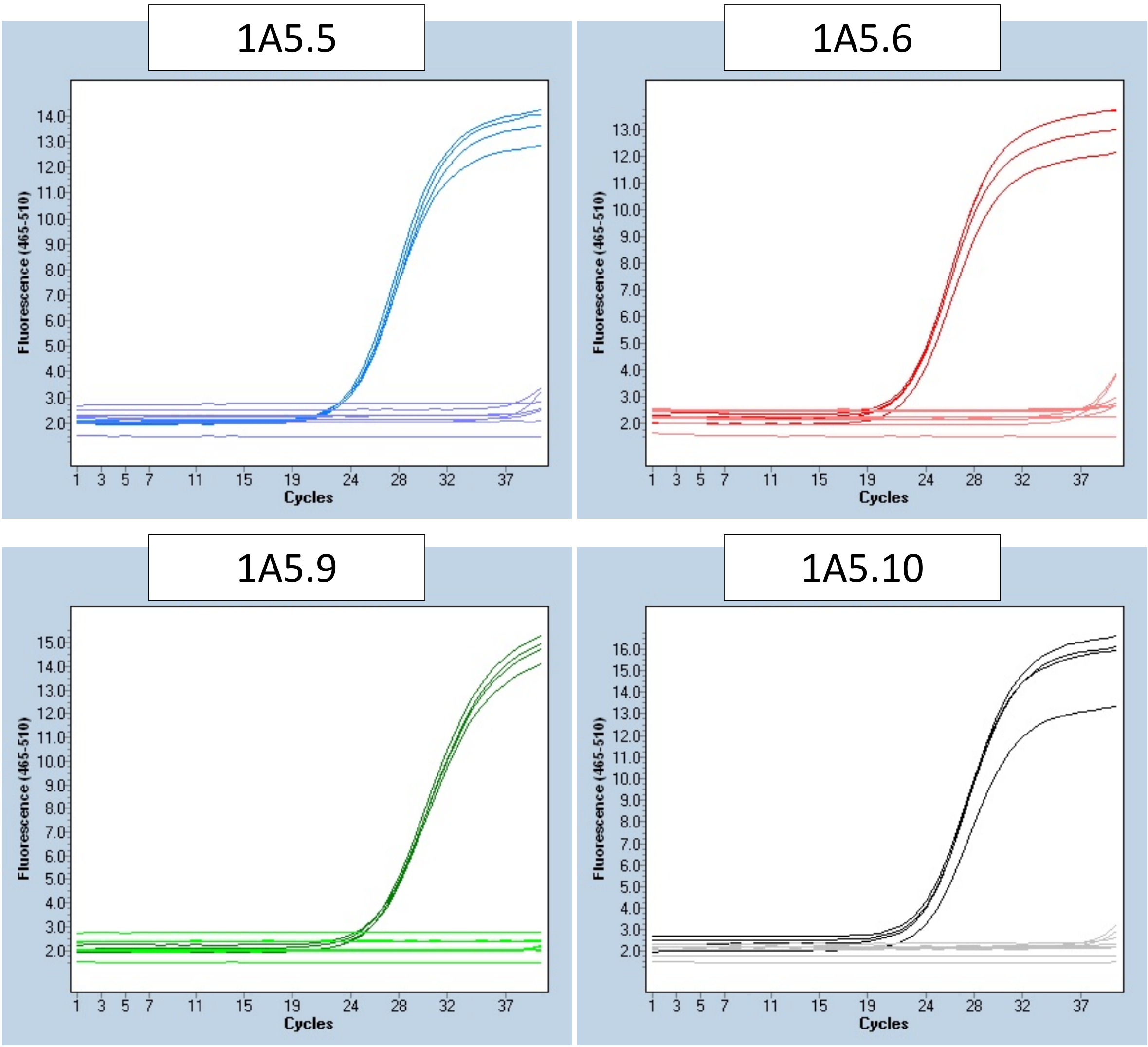


Figure C1: Amplification plots of the 1A5 targeting primers in qPCR. The amplified curves represent four replicates of 1A5, while the lower curves represent two replicates of each of 1E4, 3D1, 3D7 and water.

was a spill-over from the current treatments.

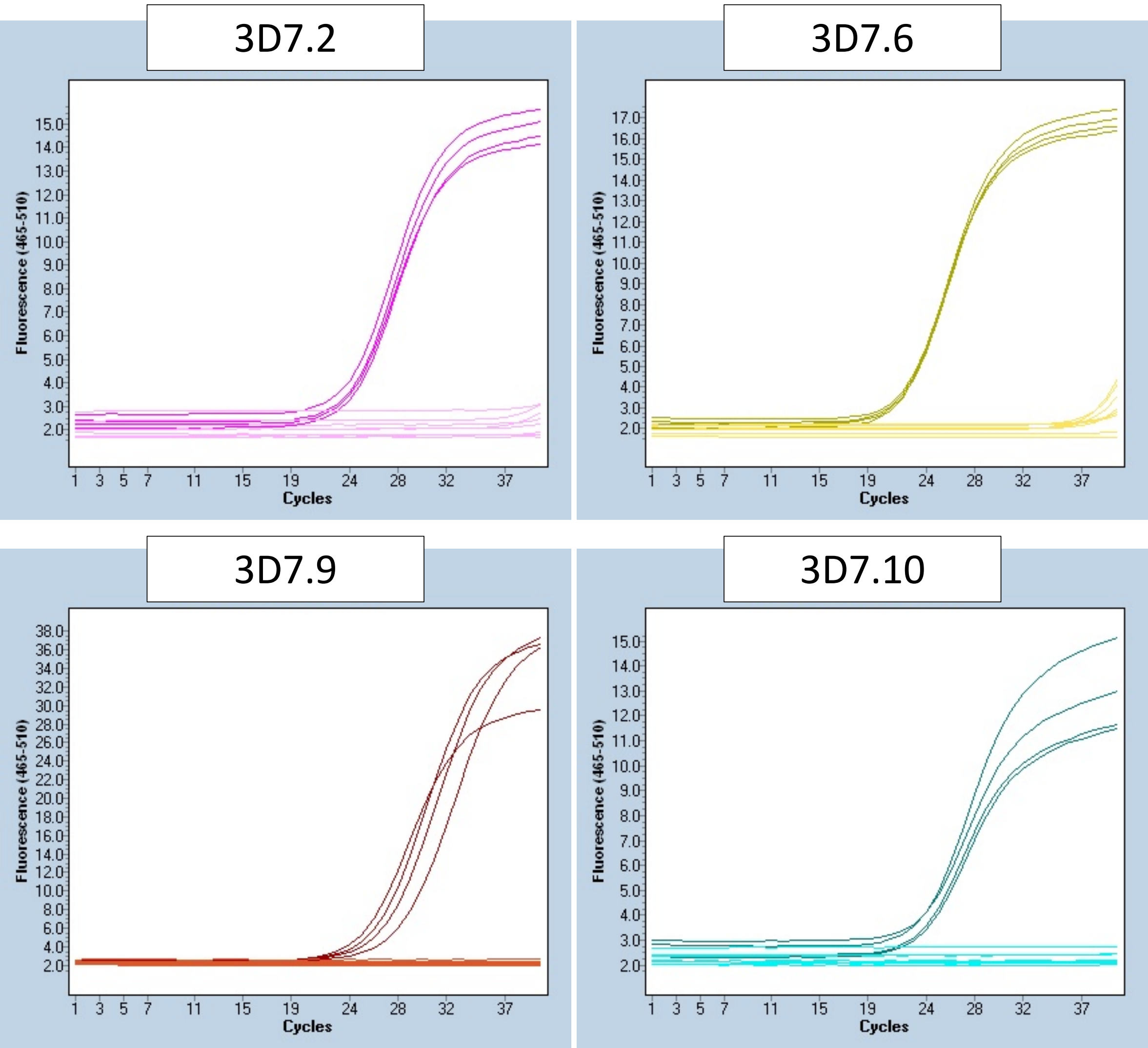


Figure C2: Amplification plots of the 3D7 targeting primers in qPCR. The amplified curves represent four replicates of 3D7, while the lower curves represent two replicates of each of 1A5, 1E4, 3D1 and water.

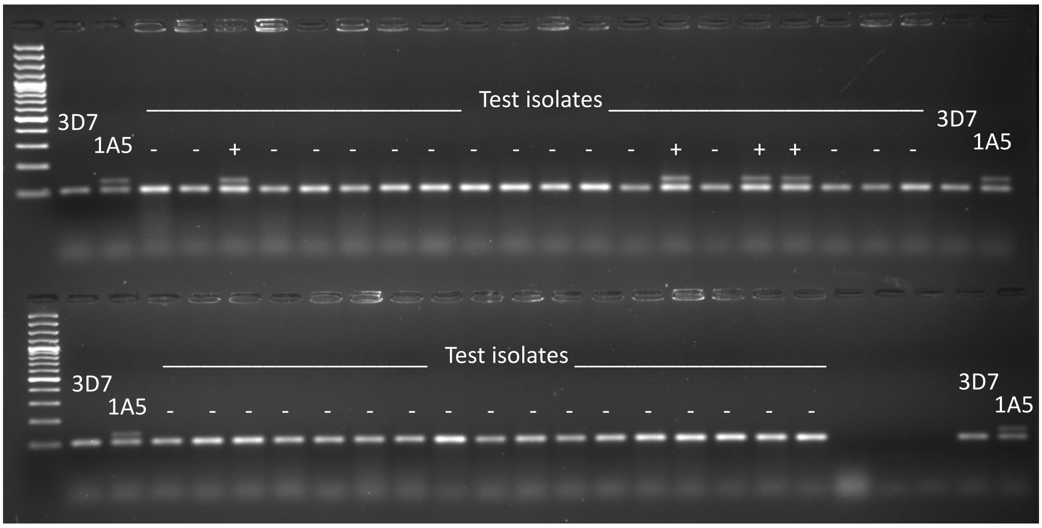


Figure C3: Amplicons from Zt_gen (shorter) and 1A5.5 (longer). Isolate 1A5 as positive control and 3D7 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

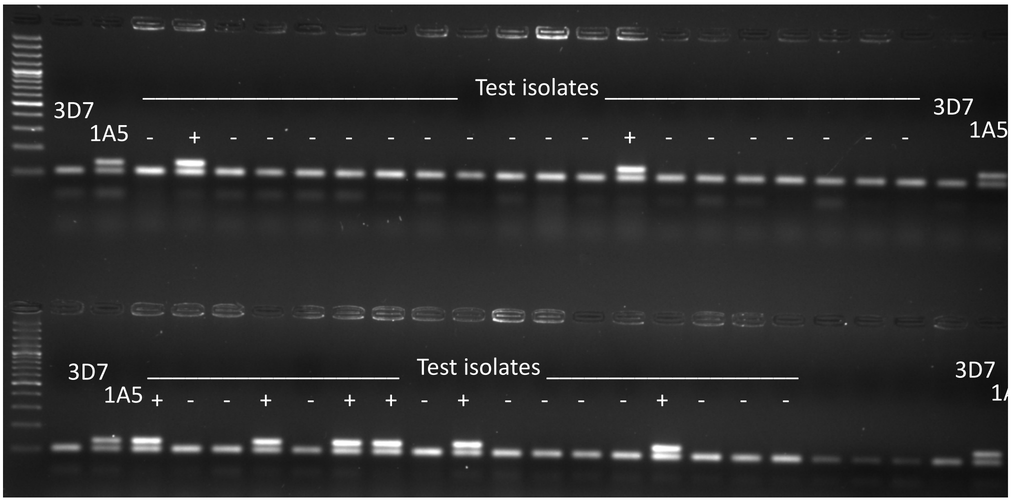


Figure C4: Amplicons from Zt_gen (shorter) and 1A5.6 (longer). Isolate 1A5 as positive control and 3D7 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

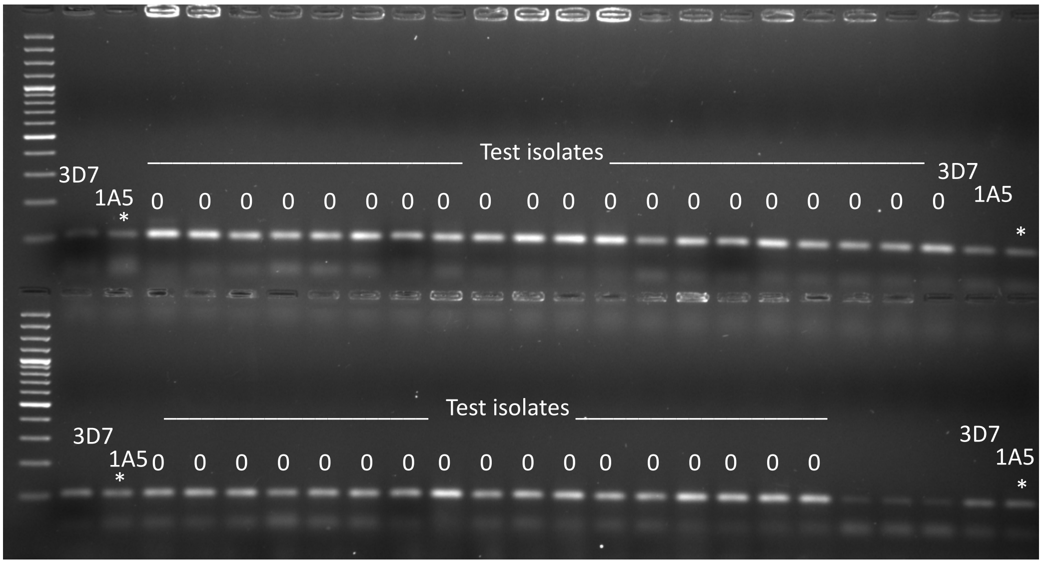


Figure C5: Amplicons from Zt_gen (shorter) and 1A5.9 (longer, not present). Isolate 1A5 as positive control and 3D7 as negative control. Zeros indicate that no conclusions were drawn from the reactions, as the positive controls were not amplified (*).

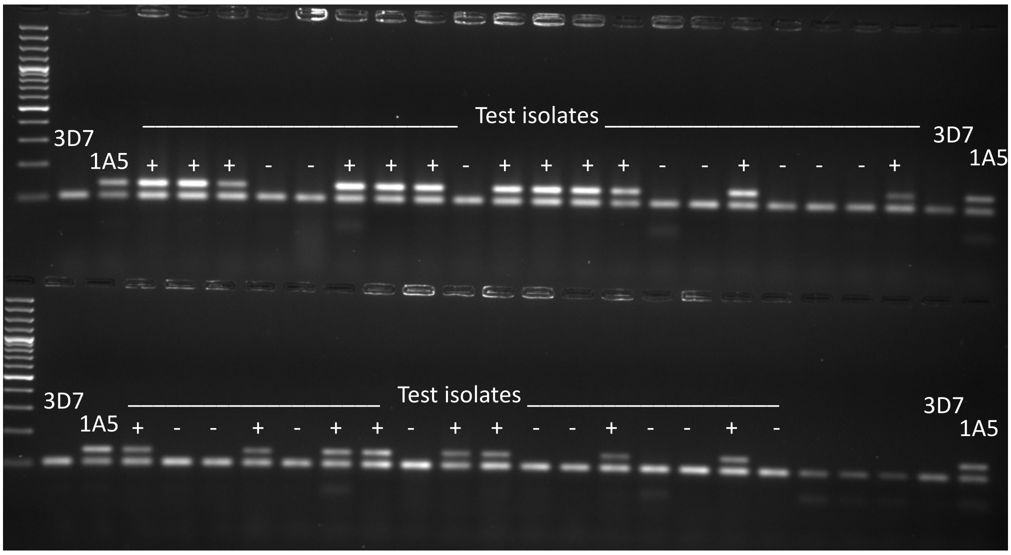


Figure C6: Amplicons from Zt_gen (shorter) and 1A5.10 (longer). Isolate 1A5 as positive control and 3D7 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

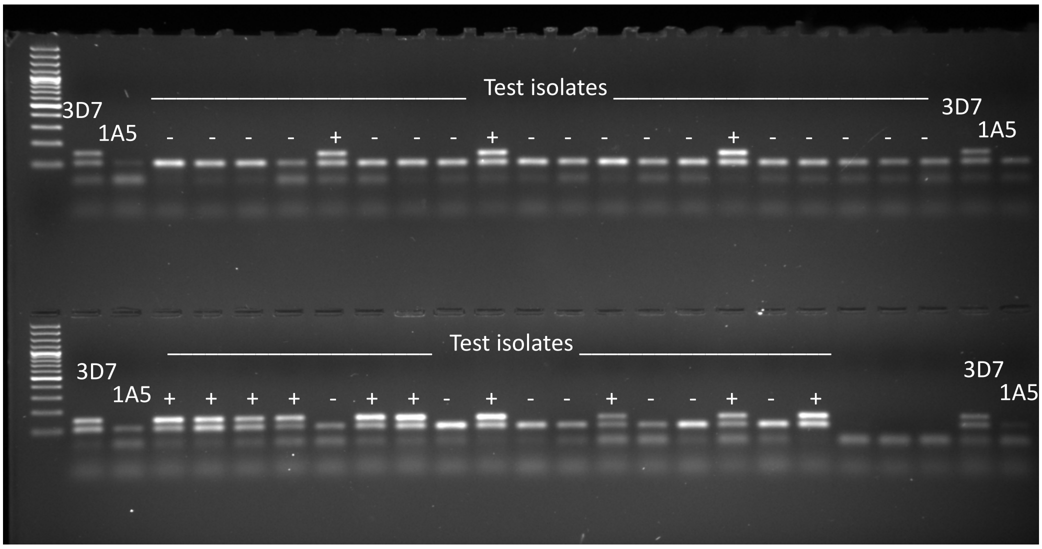


Figure C7: Amplicons from Zt_gen (shorter) and 3D7.2 (longer). Isolate 3D7 as positive control and 1A5 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

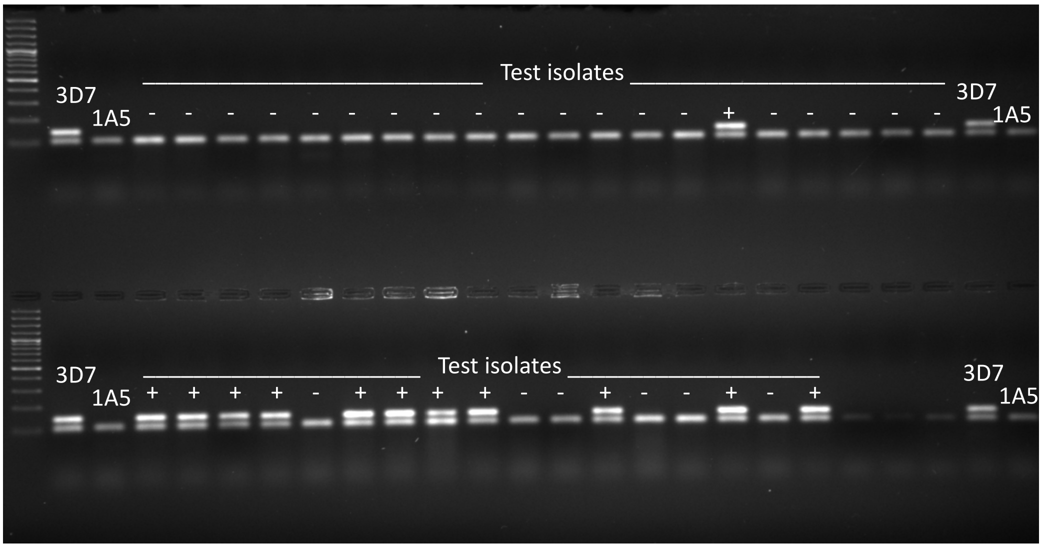


Figure C8: Amplicons from Zt_gen (shorter) and 3D7.6 (longer). Isolate 3D7 as positive control and 1A5 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

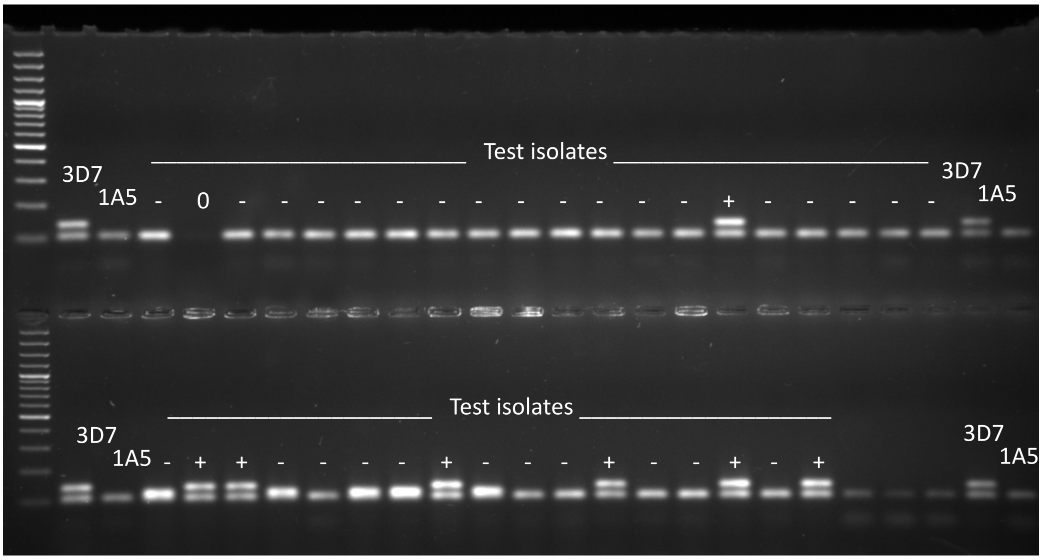


Figure C9: Amplicons from Zt_gen (shorter) and 3D7.9 (longer). Isolate 3D7 as positive control and 1A5 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

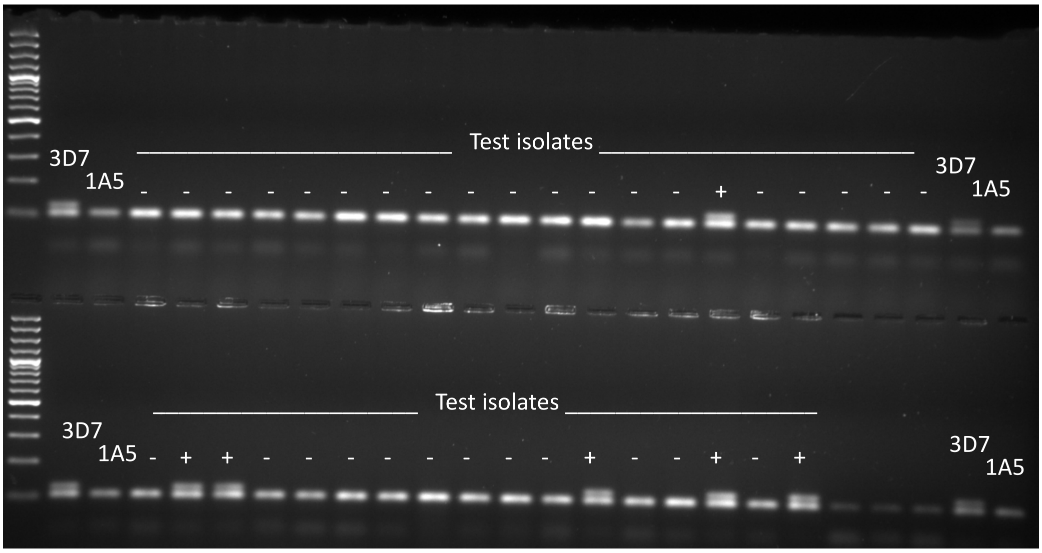


Figure C10: Amplicons from Zt_gen (shorter) and 3D7.10 (longer). Isolate 3D7 as positive control and 1A5 as negative control. Plus and minus indicate successful reaction (Zt_gen amplicon) and presence or absence of target amplicon, respectively.

**Genotyping of the re-isolated strains.** After validation, the primers 1A5.5, 1A5.6, 3D7.9 and 3D7.10 were chosen for genotyping the strains isolated from experimental material. If both of the two primer pairs targeting either 3D7 or 1A5 showed amplification, we called that a detection. On control plots, 6/37 strains tested were detected as 3D7 (16% false positives) while 0/37 strains were detected as 1A5 (0% false positives). On a plot of treatment A (replicate 1), 9/19 strains (47%) at $x_{\pm 1}$ were detected as 1A5 (Fig. C11). In contrast, on a plot of treatment D (replicate 1), 45/55 strains (82%) at $x_{\pm 1}$ were detected as 3D7 (Figs. C12, C13). Thus, frequency of 3D7 was higher than 1A5 outside the inoculation area, as implied by the disease gradients (Fig. 3B). On a plot of treatment B (replicate 1), at $x_{\pm 1}$ 2/49 were 1A5 and 37/49 were 3D7 while $x_{\pm 3}$ 1/30 was 1A5 and 8/30 were 3D7 (Figs. C14, C15 for 1A5, and Figs. C16, C17 for 3D7). As expected, the proportion of the target strains decreased with distance. Lower proportion of 1A5 is likely a result of two-fold effect of weaker transmission: first, the strain produced fewer pycnidia in the inoculation area (treatment B, replicate 1, at $x_0$ $t_0$: 1A5 4/15, 3D7 10/15, Figs. C11, C12, C13) and second, those pycnidia multiplied themselves with lower success.

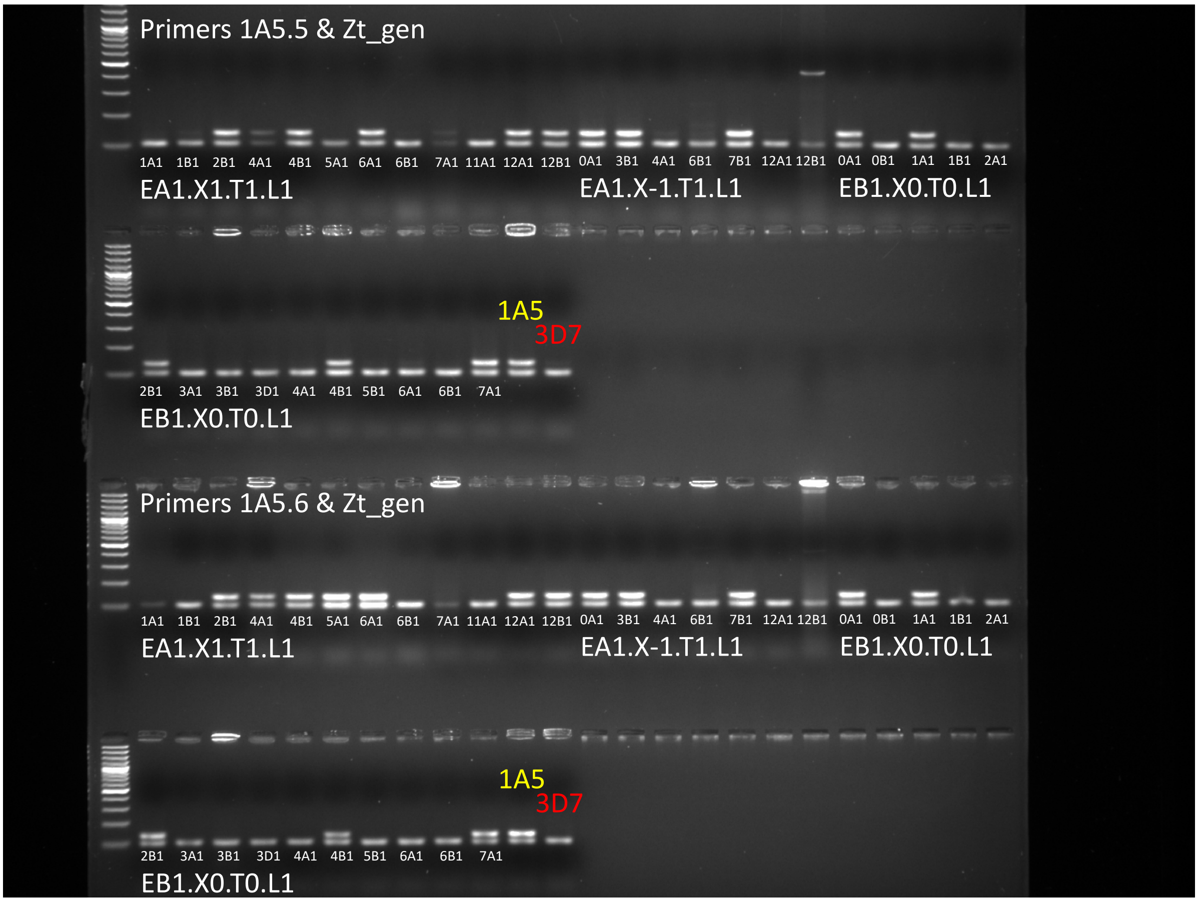


Figure C11: Amplicons from Zt_gen (shorter), 1A5.5 (longer, upper rows) and 1A5.6 (longer, lower rows). Isolate 1A5 as positive control and 3D7 as negative control. First part of isolate labels consist of location (E = Eschikon), treatment (e.g. A), and replicate (e.g. 1); second part contains measurement line (x1 for $x_1$); third, time point (T1 for $t_1$); fourth, leaf layer (L1 = Flag); and finally the isolate itself (e.g. 1A1: leaf 1, area A, isolate 1).

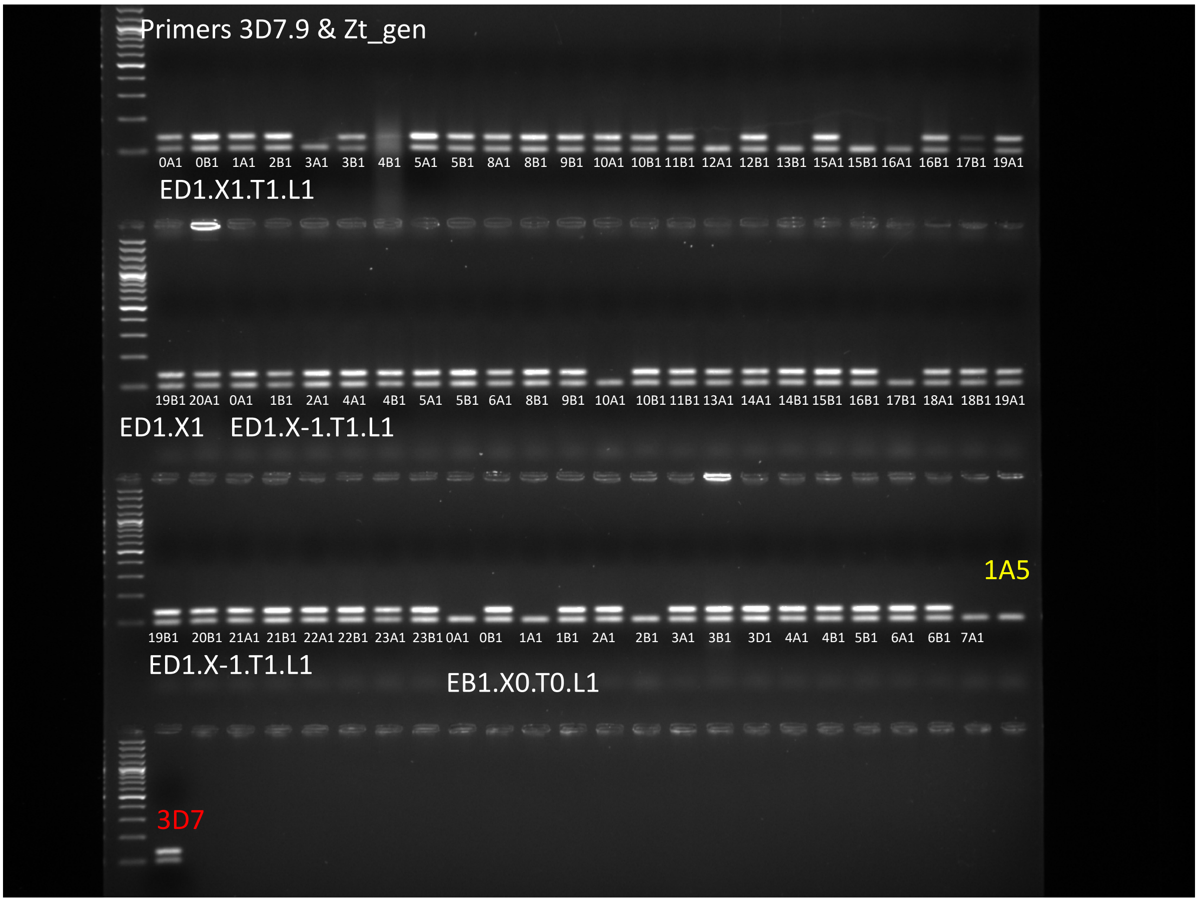


Figure C12: Amplicons from Zt_gen (shorter), 3D7.9 (longer). Isolate 3D7 as positive control and 1A5 as negative control. See Fig. C11 for label decoding.

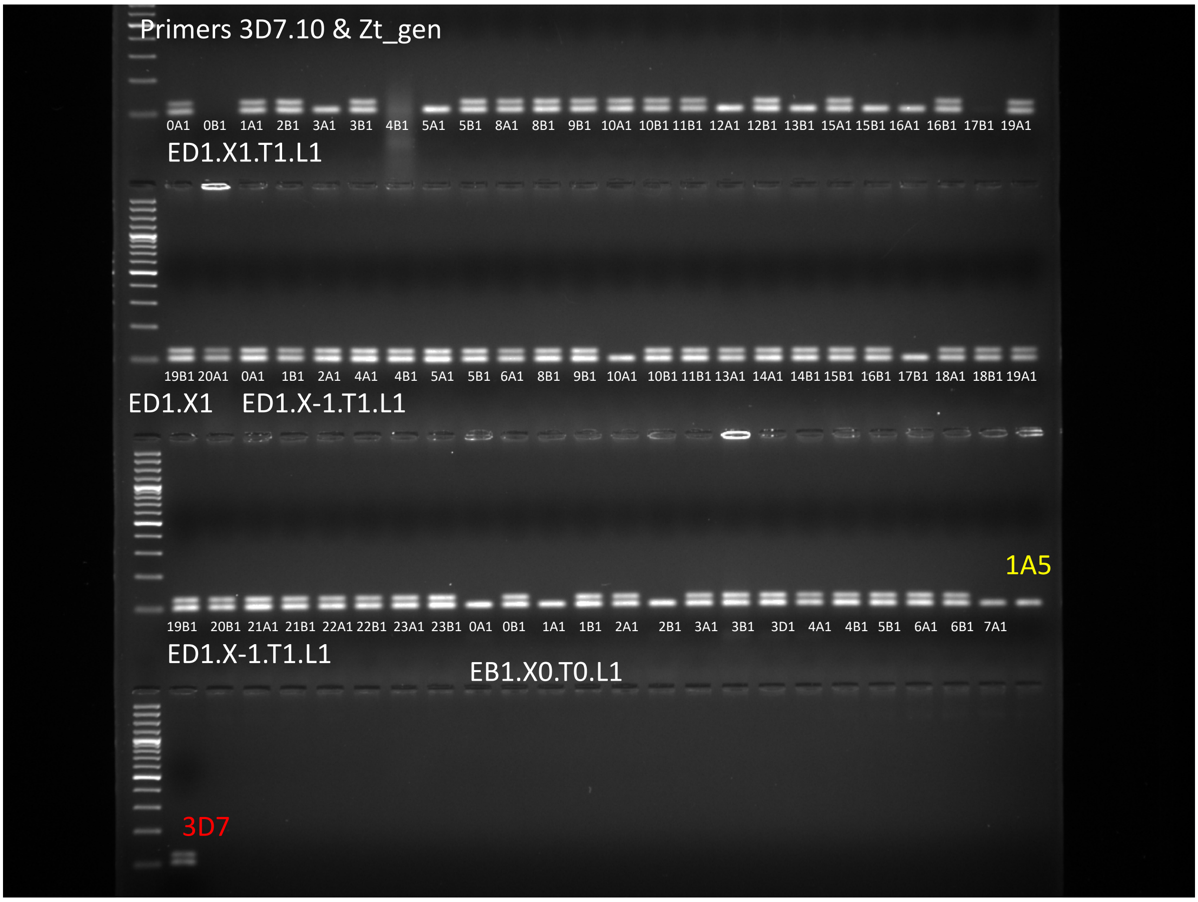


Figure C13: Amplicons from Zt_gen (shorter), 3D7.10 (longer). Isolate 3D7 as positive control and 1A5 as negative control. See Fig. C11 for label decoding.

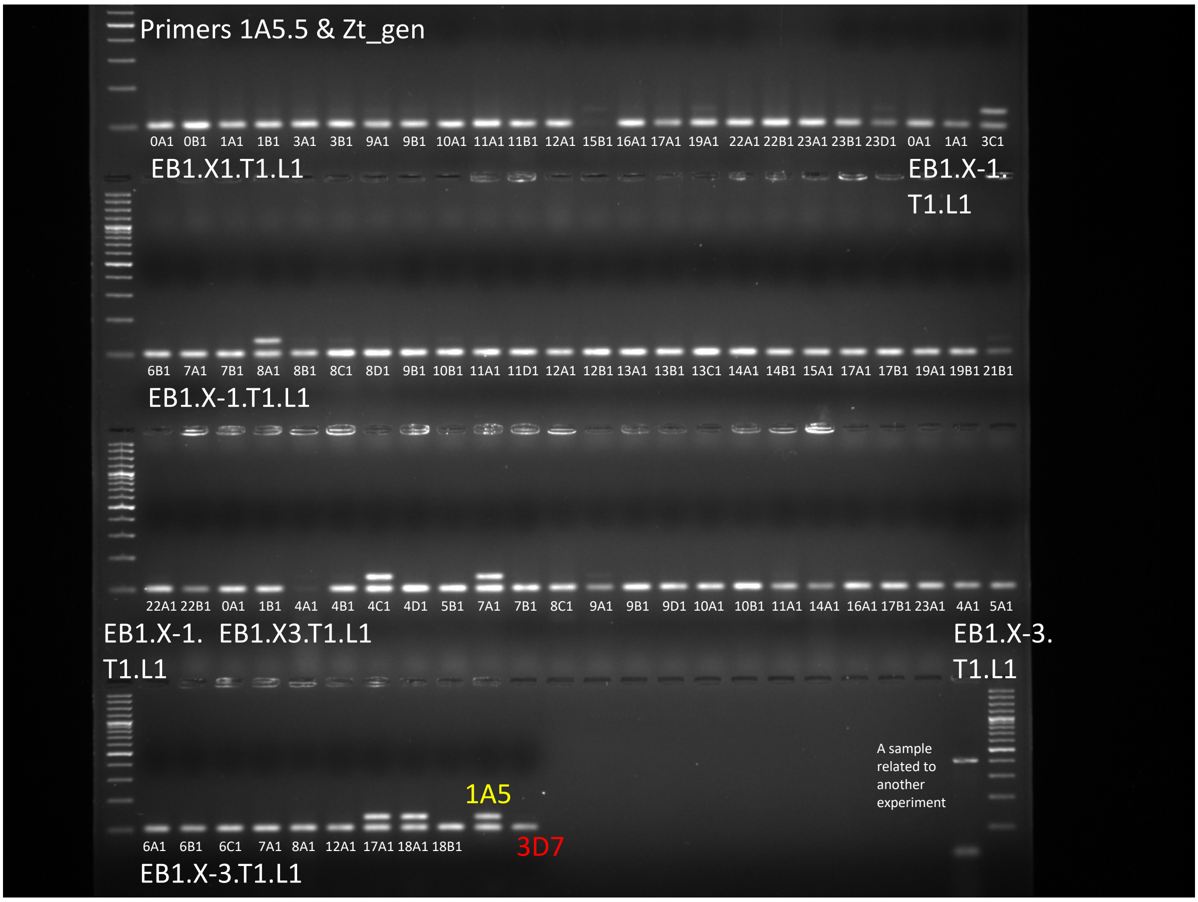


Figure C14: Amplicons from Zt_gen (shorter), 1A5.5 (longer). Isolate 1A5 as positive control and 3D7 as negative control. See Fig. C11 for label decoding.

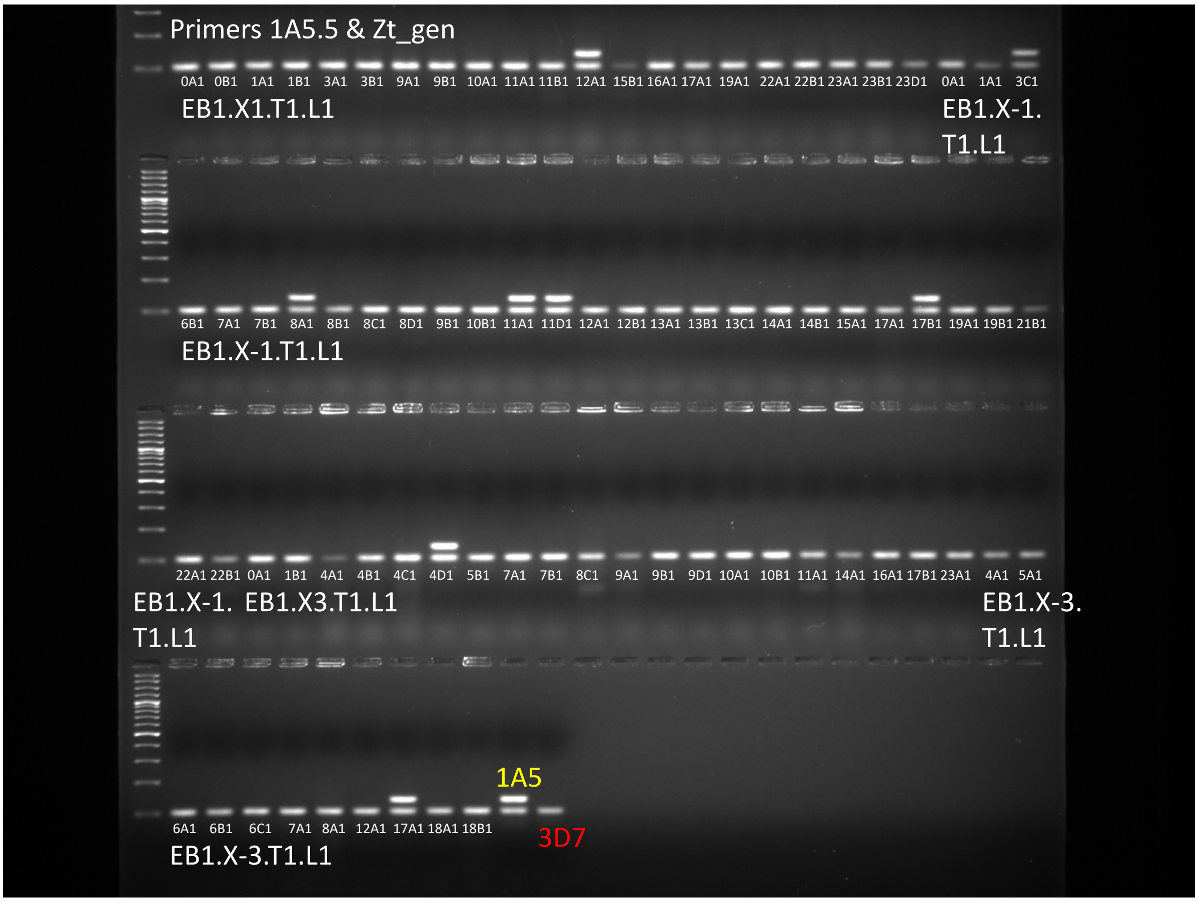


Figure C15: Amplicons from Zt_gen (shorter), 1A5.6 (longer). Isolate 1A5 as positive control and 3D7 as negative control. See Fig. C11 for label decoding.

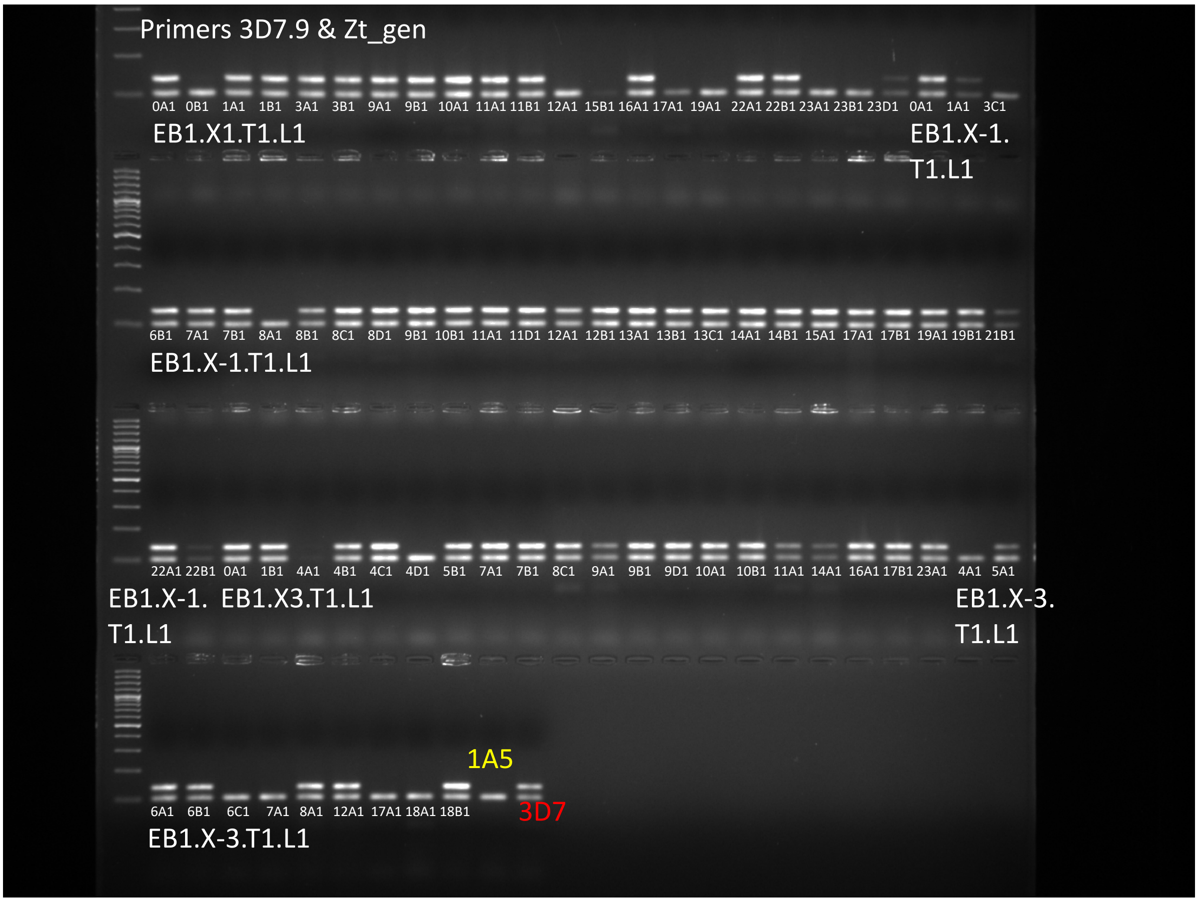


Figure C16: Amplicons from Zt_gen (shorter), 3D7.9 (longer). Isolate 3D7 as positive control and 1A5 as negative control. See Fig. C11 for label decoding.

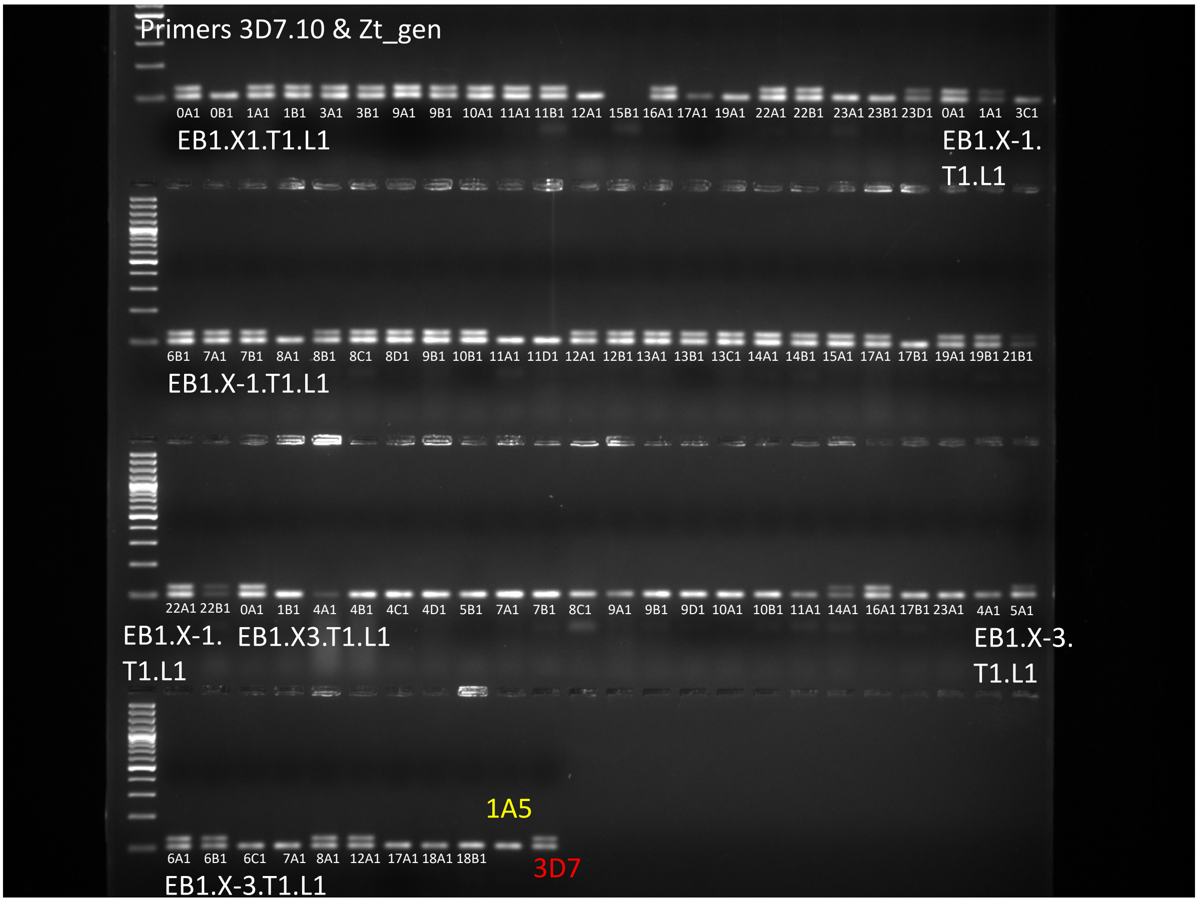


Figure C17: Amplicons from Zt_gen (shorter), 3D7.10 (longer). Isolate 3D7 as positive control and 1A5 as negative control. See Fig. C11 for label decoding.

# Appendix D: Pictures of inoculation

Figure D1: P. Karisto setting up the inoculation tent on one plot.

Figure D3: A. Mikaberidze spraying the spore suspension inside the tent to inoculate the canopy of the source area.

Figure D2: Inoculation tent prepared.

Figure D4: A tent closed after inoculation to maintain high humidity.

# Literature Cited

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* B. Petrov and F. Csaki, editors. Proc. 2nd International Symposium on Information Theory. Budapest.

Bannon, F., and B. Cooke. 1998. Studies on dispersal of Septoria tritici pycnidiospores in wheat–clover intercrops. Plant Pathology **47**:49–56.

Brennan, R., B. D. Fitt, G. Taylor, and J. Colhoun. 1985. Dispersal of Septoria nodorum pycnidiospores by simulated raindrops in still air. Journal of Phytopathology **112**:281–290.

Brophy, L. S., and C. C. Mundt. 1991. Influence of plant spatial patterns on disease dynamics, plant competition and grain yield in genetically diverse wheat populations. Agriculture, ecosystems & environment **35**:1–12.

Bullock, J. M., L. Mallada González, R. Tamme, L. Götzenberger, S. M. White, M. Pärtel, and D. A. Hooftman. 2017. A synthesis of empirical plant dispersal kernels. Journal of Ecology **105**:6–19.

Bullock, J. M., K. Shea, and O. Skarpaas. 2006. Measuring plant dispersal: an introduction to field methods and experimental design. Plant Ecology **186**:217–234.

Carrasco, L., T. Harwood, S. Toepfer, A. MacLeod, N. Levay, J. Kiss, R. Baker, J. Mumford, and J. Knight. 2010. Dispersal kernels of the invasive alien western corn rootworm and the effectiveness of buffer zones in eradication programmes in Europe. Annals of Applied Biology **156**:63–77.

Clark, J. S., M. Silman, R. Kern, E. Macklin, and J. HilleRisLambers. 1999. Seed dispersal near and far: patterns across temperate and tropical forests. Ecology **80**:1475–1494.

Cousens, R., and A. Rawlinson. 2001. When will plant morphology affect the shape of a seed dispersal "kernel"? Journal of Theoretical Biology **211**:229–238.

D'Aloia, C. C., S. M. Bogdanowicz, R. K. Francis, J. E. Majoris, R. G. Harrison, and P. M. Buston. 2015. Patterns, causes, and consequences of marine larval dispersal. Proceedings of the National Academy of Sciences **112**:13940–13945.

Davison, A. C., and D. V. Hinkley. 1997. Bootstrap methods and their application. Cambridge university press.

Dean, R., J. A. van Kan, Z. A. Pretorius, K. E. Hammond-Kosack, A. Di Pietro, P. D. Spanu, J. J. Rudd, M. Dickman, R. Kahmann, J. Ellis, and G. D. Foster. 2012. The Top 10 fungal pathogens in molecular plant pathology. Molecular plant pathology **13**:414–430.

Djidjou-Demasse, R., B. Moury, and F. Fabre. 2017. Mosaics often outperform pyramids: insights from a model comparing strategies for the deployment of plant resistance genes against viruses in agricultural landscapes. New Phytologist **216**:239–253.

Duncan, K. E., and R. J. Howard. 2000. Cytological analysis of wheat infection by the leaf blotch pathogen *Mycosphaerella graminicola*. Mycological research **104**:1074–1082.

Duvivier, M., G. Dedeurwaerder, M. De Proft, J.-M. Moreau, and A. Legrève. 2013. Real-time PCR quantification and spatio-temporal distribution of airborne inoculum of Mycosphaerella graminicola in Belgium. European Journal of Plant Pathology **137**:325–341.

Emsweller, L. N., D. L. Gorchov, Q. Zhang, A. G. Driscoll, and M. R. Hughes. 2018. Seed Rain and Disturbance Impact Recruitment of Invasive Plants in Upland Forest. Invasive Plant Science and Management **11**:69–81.

Ferrandino, F. 1996. Length scale of disease spread: fact or artifact of experimental geometry. Phytopathology **86**:806–811.

Fitt, B. D., H. McCartney, and P. Walklate. 1989. The role of rain in dispersal of pathogen inoculum. Annual review of phytopathology **27**:241–270.

Fitt, B. D. L., P. H. Gregory, A. D. Todd, H. A. McCartney, and O. C. Macdonald. 1987. Spore dispersal and plant disease gradients; a comparison between two empirical models. Journal of Phytopathology **118**:227–242.

Golan, J. J., and A. Pringle. 2017. Long-distance dispersal of fungi. Microbiology spectrum **5**:FUNK–0047–2016.

Goto, S., K. Shimatani, H. Yoshimaru, and Y. Takahashi. 2006. Fat-tailed gene flow in the dioecious canopy tree species Fraxinus mandshurica var. japonica revealed by microsatellites. Molecular Ecology **15**:2985–2996.

Greene, D. F., and C. Calogeropoulos, 2002. Measuring and modelling seed dispersal of terrestrial plants. Pages 3–23 *in* Dispersal ecology: the 42nd symposium of the British Ecological Society.

Hartmann, F. E., and D. Croll. 2017. Distinct trajectories of massive recent gene gains and losses in populations of a microbial eukaryotic pathogen. Molecular biology and evolution **34**:2808–2822.

Heald, F. 1913. The Address of the President for 1912: The Dissemination of Fungi Causing Disease. Transactions of the American Microscopical Society **32**:5–29.

Henze, M., M. Beyer, H. Klink, and J.-A. Verreet. 2007. Characterizing meteorological scenarios favorable for Septoria tritici infections in wheat and estimation of latent periods. Plant Disease **91**:1445–1449.

Johansson, R., P. Strålfors, and G. Cedersund. 2014. Combining test statistics and models in bootstrapped model rejection: it is a balancing act. BMC systems biology **8**:46.

Jones, E., T. Oliphant, P. Peterson, et al., 2001–. SciPy: Open source scientific tools for Python. URL `http://www.scipy.org/`.

Jørgensen, L. N., M. S. Hovmøller, J. G. Hansen, P. Lassen, B. Clark, R. Bayles, B. Rodemann, K. Flath, M. Jahn, T. Goral, J. H. Czembor, P. Cheyron, C. Maumene, C. de Pope, R. Ban, G. C. Nielsen, and G. Berg. 2014. IPM Strategies and Their Dilemmas Including an Introduction to www. eurowheat. org. Journal of Integrative Agriculture **13**:265–281.

Karisto, P., S. Dora, and A. Mikaberidze. 2019. Measurement of infection efficiency of a major wheat pathogen using time-resolved imaging of disease progress. Plant Pathology **68**:163–172.

Karisto, P., A. Hund, K. Yu, J. Anderegg, A. Walter, F. Mascher, B. A. McDonald, and A. Mikaberidze. 2018. Ranking Quantitative Resistance to Septoria tritici Blotch in Elite Wheat Cultivars Using Automated Image Analysis. Phytopathology **108**:568–581.

Kema, G. H., D. Yu, F. H. Rijkenberg, M. W. Shaw, and R. P. Baayen. 1996. Histology of the pathogenesis of Mycosphaerella graminicola in wheat. Phytopathology **86**:777–786.

Kirchgessner, N., F. Liebisch, K. Yu, J. Pfeifer, M. Friedli, A. Hund, and A. Walter. 2017. The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system. Functional Plant Biology **44**:154–168.

Klein, E. K., A. Bontemps, and S. Oddou-Muratorio. 2013. Seed dispersal kernels estimated from genotypes of established seedlings: does density-dependent mortality matter? Methods in Ecology and Evolution **4**:1059–1069.

Loebach, C. A., and R. C. Anderson. 2018. Measuring short distance dispersal of Alliaria petiolata and determining potential long distance dispersal mechanisms. PeerJ **6**:e4477.

Lovell, D., S. Parker, T. Hunter, D. Royle, and R. Coker. 1997. Influence of crop growth and structure on the risk of epidemics by Mycosphaerella graminicola (Septoria tritici) in winter wheat. Plant pathology **46**:126–138.

Lovell, D., S. Parker, T. Hunter, S. Welham, and A. Nichols. 2004. Position of inoculum in the canopy affects the risk of septoria tritici blotch epidemics in winter wheat. Plant Pathology **53**:11–21.

Madden, L. V., G. Hughes, and F. van den Bosch. 2007. The study of plant disease epidemics. American Phytopathological Society (APS Press).

McCartney, H., B. D. Fitt, and J. S. West, 2006. Dispersal of foliar plant pathogens: mechanisms, gradients and spatial patterns. Pages 159–192 *in* B. M. Cooke, D. G. Jones, and B. Kaye, editors. The epidemiology of plant diseases. Springer.

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

Mundt, C., and J. Browning. 1985. Development of crown rust epidemics in genetically diverse oat populations: effect of genotype unit area. Phytopathology **75**:607–610.

Nathan, R., E. Klein, J. J. Robledo-Arnuncio, and E. Revilla, 2012. Dispersal kernels: review. Chapter 15, pages 187–210 *in* J. Clobert, M. Baguette, T. G. Benton, and J. M. Bullock, editors. Dispersal ecology and evolution. Oxford University Press Oxford.

Nathan, R., U. N. Safriel, I. Noy-Meir, and G. Schiller. 2000. Spatiotemporal variation in seed dispersal and recruitment near and far from Pinus halepensis trees. Ecology **81**:2156–2169.

Newton, A., G. Begg, and J. Swanston. 2009. Deployment of diversity for enhanced crop function. Annals of Applied Biology **154**:309–322.

Newton, A., and D. Guy. 2011. Scale and spatial structure effects on the outcome of barley cultivar mixture trials for disease control. Field Crops Research **123**:74–79.

Newville, M., T. Stensitzki, D. B. Allen, and A. Ingargiola. 2014. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. Zenodo .

Robert, C., G. Garin, M. Abichou, V. Houlès, C. Pradal, and C. Fournier. 2018. Plant architecture and foliar senescence impact the race between wheat growth and Zymoseptoria tritici epidemics. Annals of botany **121**:975–989.

Saint-Jean, S., M. Chelle, and L. Huber. 2004. Modelling water transfer by rain-splash in a 3D canopy using Monte Carlo integration. Agricultural and Forest Meteorology **121**:183–196.

Sapoukhina, N., Y. Tyutyunov, I. Sache, and R. Arditi. 2010. Spatially mixed crops to control the stratified dispersal of airborne fungal diseases. Ecological Modelling **221**:2793–2800.

Shaw, M. 1987. Assessment of upward movement of rain splash using a fluorescent tracer method and its application to the epidemiology of cereal pathogens. Plant Pathology **36**:201–213.

Shaw, M. 1990. Effects of temperature, leaf wetness and cultivar on the latent period of Mycosphaerella graminicola on winter wheat. Plant Pathology **39**:255–268.

Shaw, M., and D. Royle. 1993. Factors determining the severity of epidemics of Mycosphaerella graminicola (Septoria tritici) on winter wheat in the UK. Plant Pathology **42**:882–899.

Skarpaas, O., and K. Shea. 2007. Dispersal patterns, dispersal mechanisms, and invasion wave speeds for invasive thistles. The American Naturalist **170**:421–430.

Skarpaas, O., K. Shea, and J. M. Bullock. 2005. Optimizing dispersal study design by Monte Carlo simulation. Journal of Applied Ecology **42**:731–739.

Solheim, H., and A. M. Hietala. 2017. Spread of ash dieback in Norway. Baltic Forestry **23**:144–149.

Soubeyrand, S., J. Enjalbert, A. Sanchez, and I. Sache. 2007. Anisotropy, in density and in distance, of the dispersal of yellow rust of wheat: experiments in large field plots and estimation. Phytopathology **97**:1315–1324.

Stewart, E. L., D. Croll, M. H. Lendenmann, A. Sanchez-Vallet, F. E. Hartmann, J. Palma-Guerrero, X. Ma, and B. A. Mcdonald. 2018. Quantitative trait locus mapping reveals complex genetic architecture of quantitative virulence in the wheat pathogen Zymoseptoria tritici. Molecular plant pathology **19**:201–216.

Terpilowski, M., 2018. Scikit-posthocs. URL `https://scikit-posthocs.readthedocs.io`.

Torriani, S. F., J. P. Melichar, C. Mills, N. Pain, H. Sierotzki, and M. Courbot. 2015. Zymoseptoria tritici: a major threat to wheat production, integrated approaches to control. Fungal Genetics and Biology **79**:8–12.

Van Houtan, K. S., S. L. Pimm, J. M. Halley, R. O. Bierregaard Jr, and T. E. Lovejoy. 2007. Dispersal of Amazonian birds in continuous and fragmented forest. Ecology letters **10**:219–229.

Vidal, T., C. Gigot, C. de Vallavieille-Pope, L. Huber, and S. Saint-Jean. 2018. Contrasting plant height can improve the control of rain-borne diseases in wheat cultivar mixture: modelling splash dispersal in 3-D canopies. Annals of botany **121**:1299–1308.

Vidal, T., P. Lusley, M. Leconte, C. De Vallavieille-Pope, L. Huber, and S. Saint-Jean. 2017. Cultivar architecture modulates spore dispersal by rain splash: A new perspective to reduce disease progression in cultivar mixtures. PloS one **12**:e0187788.

Werth, S., H. H. Wagner, F. Gugerli, R. Holderegger, D. Csencsics, J. M. Kalwij, and C. Scheidegger. 2006. Quantifying dispersal and establishment limitation in a population of an epiphytic lichen. Ecology **87**:2037–2046.

Zadoks, J., and F. van den Bosch. 1994. On the spread of plant disease: a theory on foci. Annual review of phytopathology **32**:503–521.

Zadoks, J. C., T. T. Chang, and C. F. Konzak. 1974. A decimal code for the growth stages of cereals. Weed research **14**:415–421.

Zadoks, J. C., and R. D. Schein. 1979. Epidemiology and plant disease management. Oxford University Press Inc.

Zhan, J., G. H. Kema, C. Waalwijk, and B. A. McDonald. 2002. Distribution of mating type alleles in the wheat pathogen Mycosphaerella graminicola over spatial scales from lesions to continents. Fungal Genetics and Biology **36**:128–136.

Zhan, J., C. Mundt, and B. McDonald. 1998. Measuring immigration and sexual reproduction in field populations of Mycosphaerella graminicola. Phytopathology **88**:1330–1337.

Zhan, J., C. Mundt, and B. McDonald. 2000. Estimation of rates of recombination and migration in populations of plant pathogens—a reply. Phytopathology **90**:324–326.