Exit from lag phase in populations of *Pseudomonas fluorescens* is determined by strong interactions among cells

Maxime Ardré [1], Guilhem Doulcier [1], Naama Brenner [2], Paul B. Rainey [1,3]

1 Laboratoire Biophysique et Évolution, CBI, ESPCI Paris, Université PSL, CNRS 75005 Paris, France
2 Network Biology Research Laboratories, and Department of Chemical Engineering, Technion–Israel Institute of Technology, Haifa, Israel
3 Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Plön, Germany

* maxime.ardre@espci.psl.eu

Abstract

The relationship between the number of cells colonizing a new environment and time for resumption of growth is a subject of long-standing interest. In microbiology this is known as the "inoculum effect". Its mechanistic basis is unclear with possible explanations ranging from the independent actions of individual cells, to collective actions of populations of cells. Progress requires precise measurement of lag-time distributions while at the same time, experimentally controlling inoculum size. Here we use a millifluidic droplet device in which the growth dynamics of hundreds of populations founded by different numbers of *Pseudomonas fluorescens* cells, ranging from a single cell, to one thousand cells, were followed in real time. Our data show that lag phase decreases with inoculum size. The average decrease, variance across droplets, and distribution shapes, follow predictions of extreme value theory, where the inoculum lag-time is determined by the minimum value sampled from the single-cell distribution. Our experimental results show that exit from lag phase depends on strong interactions among cells, consistent with a "leader-cell" triggering end of lag phase for the entire population.

Introduction

When bacteria encounter new environmental conditions, growth typically follows four phases: a lag phase, during which bacteria acclimate, but do not divide; an exponential phase, during which cells multiply; a stationary phase, where nutrient exhaustion causes cessation of growth; and finally a death phase, during which cells may lyse.

In a fluctuating environment, each phase can play an important role in population persistence. The lag phase has particular significance because of both benefits (enhanced growth) and costs (sensitivity to external stressors) associated with the resumption of growth [14, 25]. The time to resumption of growth — and controlling factors — has significant implications for the entire field of microbiology [13], but especially for infection caused by pathogens and for food safety [3, 17, 22].

Despite its discovery more than 100 years ago [15], cellular and molecular details defining the lag phase, factors triggering resumption of growth, and contributions to

fitness, are not well understood. This is largely a consequence of the difficulties associated with experimental quantification of the dynamics of populations founded by small numbers of cells. Nonetheless, advances over the last decade have shown that bacteria in lag phase are transcriptionally and metabolically active [19], that lag phase is a dynamic state, that single cells are heterogeneous in time to resume division [2, 9, 14, 25], and that numerous factors affect its duration [3].

Arguably the most intriguing aspect of lag phase biology is the apparent inverse relationship between the number of cells in the founding population and duration of lag phase — often referred to as the "inoculum effect". First reported in 1906 [18], the relationship has been shown to hold for a number of different bacteria [10–12, 16], although there exist few recent quantitative investigations. In certain instances, the inoculum effect is observed only under specific culture conditions [1, 10].

Factors controlling the inoculum effect are of special interest [1, 3, 5, 8, 11, 22]. Given that bacterial cells are typically variable in many of their properties, the simplest explanation posits that population lag time is determined by the set of cells with the shortest time to first division. Accordingly, the larger the founding population, the more likely it is that the inoculum contains cells on the verge of division, with these cells contributing disproportionately to the resumption of population growth.

An alternate explanation is that resumption of growth depends on interaction among founding cells, for example, via production of an endogenous growth factor: once a critical threshold concentration is achieved growth resumes and thus the larger the inoculum, the sooner the threshold is achieved. Evidence in support of such control derives from analysis of *Bacillus* [11], *Francisella* (formerly *Pasturella*) *tularensis* [8], *Micrococcus luteus* [23] and *Aerobacter aerogenes* [5].

In instances where exit from lag-phase is determined by interactions among founding cells, models have assumed that all cells are equal contributors to the production of growth activating factors [10, 11]. However, an alternative possibility exists, namely, that population lag time is set by the activity of a single "leader cell" that triggers resumption of growth for the entire population of cells. Distinguishing among competing hypotheses requires precision measurements of population growth, high levels of replication, ability to control inoculum size, and crucially, knowledge of the distribution of lag times for single cells.

Here we use a millifluidic droplet device in which the growth dynamics of hundreds of populations founded by different numbers of *Pseudomonas fluorescens* cells were followed in real time. Data confirm that length of lag phase decreases with inoculum size. We demonstrate that statistical properties of droplets with varying inoculum sizes follow extreme value theory, where population lag time is determined by the minimum value sampled from the single cell distribution. Our experimental results show that exit from lag phase depends on strong interactions among cells consistent with a leader-cell triggering end of lag phase for the entire population.

Materials and Methods

The strain. The ancestral strain of *Pseudomonas fluorescens* SBW25 was isolated from the leaf of a sugar beet plant at the University of Oxford farm (Wytham, Oxford, United Kingdom [20]). The strain was modified to incorporate, via chromosomally integrated Tn*7*, the gene GFP-mut3B controlled by an inducible Ptac promoter.

Preparation of the cells *P. fluorescens* SWB25 was grown in casamino acid medium (CAA). CAA for 1l: 5g of Bacto Casamino Acids Technical (BD ref 223120), 0.25g $MgSO_4 \cdot 7 \cdot H_2O$ (Sigma CAS 10034-99-8), 0.9g KH_2PO_4 (Melford CAS 7758-11-4). Prior to generation of droplets SBW25 was grown from a glycerol stock for 19h in 5ml of CAA incubated at 28°C and shaken at 180 rpm. At 19h this culture was

centrifuged at 3743 RCF for 4min and the supernatant removed from the pellet. The pellet was then resupended in 5ml of sterile CAA. It was then centrifugate and resuspended one more time. This prevents any interference of a growth-activator that may come from the supernatant of the overnight culture. The washed culture was then adjusted to OD 0.8 with CAA and mixed 1:1 in volume with autoclaved 30%v/v glycerol. 100µl aliquots were pipetted in 1ml eppendorf and frozen at -80°C. After freezing, one aliquot was taken to measure viable cells by plating on agar. We found $1.62 \cdot 10^8$ cell·ml^{-1}.

Generation of droplets with a range of inoculum sizes. Each experiment with a range of inoculum sizes was prepared as follows. One frozen aliquot was thawed and diluted in 4ml of sterile CAA (with appropriate intermediate dilutions) by $7.04 \cdot 10^4\times$, $1.76 \cdot 10^4\times$, $1.76 \cdot 10^4\times$, $4.4 \cdot 10^3\times$, $1.1 \cdot 10^3\times$, $2.75 \cdot 10^2\times$, $6.875 \cdot 10^1\times$. We completed the dilutions from frozen stock by adding 29.1µl, 29µl, 28.6µl, 27.3µl, 21.8µl and 0µl, respectively, of sterile 60% v/v glycerol. This step is very import to balance the glycerol coming from the frozen stock and ensure that all the tubes have the same composition of medium (see Supplementary Fig. 5). We then added 50µl of sterile IPTG (100mM) to each sample. Each dilution was then pipetted in wells (250µl/well) of a 96 well microtitre plate to proceed to the generation of the droplets in the Millidrop Azur©. Droplets have a volume of 0.4µl which yields, with our dilutions, a range of inoculum sizes as follows: 1, 4, 16, 64, 256, 1024 cells/drop. We generated 40 replicate droplets for each population of a given founding inoculum size, except for populations founded by 1024 cells, which for technical reasons were restricted to 30 replicates.

The inoculum of droplets follows a Poisson distribution. Importantly, the inoculum size in each droplet is controlled by the Poisson process occurring during the formation of a drop from the 96 well-plates. Therefore the inoculum that we report here is the average inoculum. For a given inoculum the actual number of cells inserted in the droplets follows a Poisson distribution. Hence for a given inoculum the variance of the number of cells between the droplets is equal to the average inoculum.

Generation of droplets with inoculum 1. To generate droplets with an inoculum of 1 cell/drop we diluted in sterile CAA a frozen alicot by $7.04 \cdot 10^4\times$, added 29.1µl of sterile 60% v/v glycerol and 50µl of sterile IPTG (100mM). We generated 230 droplets in the Millidrop Azur that yielded 156 droplets that grew due to the Poisson process inherent to the sampling process.

Calculation of the uncertainty of measurement for the inoculum 1 $\triangle N_{th}$ is the uncertainty on the threshold $N_{th} = 1.6 \cdot 10^8$ cell·ml^{-1}. The uncertainty of the calibration Fig. 10 gives $\triangle N_{th} = 0.7 \cdot 10^8$ cell·ml^{-1} for this value as depicted by the grey area. $\triangle t_{th}$ is the uncertainty of the time when the population reaches beyond the threshold N_{th}. We take its value as equal to the sampling frequency of the machine $\triangle t_{th} = 18$min. λ is the average growth rate of populations in droplets and $\triangle\lambda$ is the uncertainty. These quantities are estimated with the distribution of the growth rate shown Fig. 8: $\lambda = 0.84$ h^{-1} and $\triangle\lambda = 0.02$ h^{-1}. $\triangle N_0$ is the uncertainty on the inoculum size. In the experiment with inoculum 1 shown Fig. 2B, $230 - 156 = 74$ droplets were empty despite being generated from the same mother culture. This is due to the randomness of the pipetting process that fills droplets of bacteria according to a Poisson process. The randomness of the process gives intrinsically an uncertainty on N_0. In the following we explain how this was estimated. Knowing the number of empty droplets in the experiments allows calculation of the precise average inoculum of the experiment: $n = 1.134$ cells/drop. This average takes into account the empty droplets with zero bacteria but we only measure the non-empty droplets. To estimate the average inoculum of the non-empty droplets we simply draw numerically a large series of random numbers with a Poisson probability of parameter $n = 1.134$ and calculate the average and the SD of the non zero values. This yielded an average of 1.7 and an SD of

0.9. Therefore we consider that for our experiment in an ideal case with an infinite number of droplets the uncertainty on the inoculum of the droplets will be intrinsically $\triangle N_0 = 0.9$ cell/drop and that the averaged inoculum (of the filled droplets) is $N_0 = 1.7$ cell/drop. All together these values allow calculation of the uncertainty of the lag time estimated by equation 1. The expression of uncertainty is given by equation 2 and numerical application gives $\triangle\theta = 0.88$ h.

Results

High-throughput quantification of bacterial population dynamics with millifluidic technology.

To investigate the relationship between inoculum size and duration of lag phase, we used a millifluidic device to quantify the dynamics of bacterial population growth across time. The device allows the monitoring of 230 bacterial populations compartmentalized in droplets contained within a tube. Fig. 1A shows a portion of the tube with two droplets filled with cells. Statistical power of the experiment comes from precise control of large numbers of droplets, in terms of both inoculum size and homogeneity of culture conditions: Fig. 1B shows the growth dynamics of 40 replicate populations. Exponential growth and stationary phase are clearly seen, while lag time is concealed behind the detection threshold (grey area). Fig. 1C shows the growth dynamics from a population contained within a single droplet, and the inference of lag time from the best-fit exponential growth curve.

The four phases of *P. fluorescens* growth are described by parameters estimated from time series that report the population density. Density is estimated by the fluorescence intensity of the population composed of GFP-labelled bacteria. The exponential growth rate, λ, is the maximum slope of the time series on a y-semi-logscale (we use a Gaussian processes method that makes no *a priori* assumption about the shape of the growth curve [21]). The final population size is estimated directly from measurements; death phase is not significant in our experiment and is ignored. The lag phase τ, is the time cells spend in a non-dividing phase prior to onset of exponential growth. Hardware limitations mean that fluorescence data are unobtainable for cell densities below 1,600 cells per droplet ($4 \cdot 10^6$ cell/ml) and thus τ must be estimated indirectly. This is done by firstly taking an arbitrary point (t_{th}, N_{th}) in exponential phase where cell density is $N_{th} = 1.6 \cdot 10^8$ cell·ml^{-1}. By rearranging the equation for exponential growth: $N_{th} = N_0 e^{\lambda(t_{th}-\tau)}$, and making τ the subject

$$\tau = t_{th} - \log(N_{th}/N_0)/\lambda \quad (1)$$

A geometrical counterpart of equation 1 gives the population lag time as the time point at which the exponential growth (line in semi-log scale) intersects the horizontal line that depicts the inoculum density (see Fig. 1C).

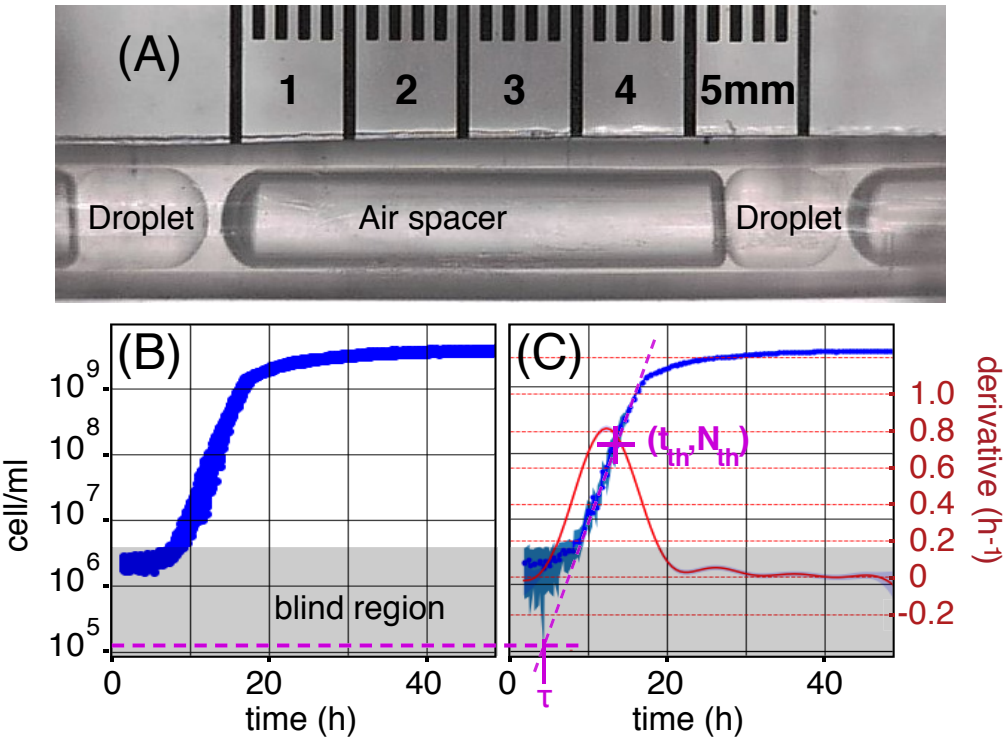

Figure 1. Bacterial population growth in droplets. (A) Two droplets of 0.4 µl are separated by an air spacer (to prevent droplet coalescence) inside the tube of a millifluidic machine. Droplets are prepared by "sipping" samples from a 96 well plate. Typically 230 droplets are produced from six mother cultures that differ solely in the number of founding cells (the *inoculum*). Each mother culture delivers 40 replicate droplets, but for technical reasons that last delivers 30 replicates. Droplets move back-and-forth, via changes in pressure, passing in front of a fluorescence detector every ∼18 minutes. *P. fluorescens* SBW25 cells express GFP from a chromosomally integrated reporter, allowing changes in biomass to be determined based on intensity of the fluorescent signal (excitation at 497 nm emission at 527 nm). Signal intensity is calibrated to cell density by plate counting (Supplementary Fig. 10). The range of detection extends from $4 \cdot 10^6$ to $5 \cdot 10^9$ cell·ml^{-1} ($1.6 \cdot 10^3$ to $2 \cdot 10^6$ cells per droplet). The grey area on subfigures (B) & (C) denote the bacterial density that are below the threshold of detection. **(B)** Fluorescent signal across time from 40 replicate populations (in semi-logscale) in droplets prepared from the same mother culture. The average inoculum in each droplet is $1.6 \cdot 10^5$ cell·ml^{-1}, thus 64 cells per droplet. In this example the signal met the detection threshold at ∼7 h, by which populations are in exponential growth phase. At ∼20 h, stationary phase is reached, marked by cessation of growth. **(C)** A single time series showing population growth within a single droplet coming from the set of replicates shown in (B). The left y-axis is shared between these two plots. The blue line depicts cell density derived using DropSignal [6] and the shaded area represents the standard deviation (SD). Population lag time is inferred by extrapolating the exponential growth (purple dotted line) to its intersection with the inoculum density (purple horizontal dotted line) . The specific point (t_{th}, N_{th}) of the growth curve is fixed to a cell density of $N_{th} = 1.6 \cdot 10^8$ cell·ml^{-1} (64,000 cells per droplet). The right y-axis in red gives the derivative of the time series (continuous bell-shaped line in red) with its shaded area denoting SD (which is small).

Duration of lag phase depends on the number of founding cells.

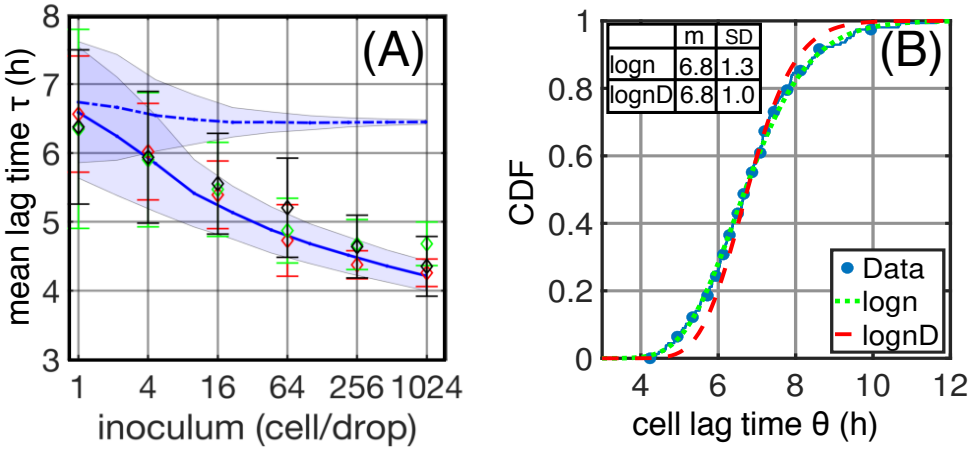

Figure 2. Simulations are consistent with exit from lag phase being dependent upon a single leader cell. (A) Shows the relationship between inoculum size and population lag time (τ). Diamond symbols are the mean lag times of populations for given inoculum sizes with error bars denoting the standard deviation (SD). The three colours correspond to independent experiments. **(B)** shows the distribution of cell lag times (θ) in an experiment that monitored population growth in 156 droplets inoculated with on average a single bacterium (blue points). CDF is the cumulative distribution function, where the y-value give the probability that the cell-lag times assume a value less than or equal to the x-values. The measured distribution is fitted to a log-normal distribution (green dotted line) with a mean of 6.8h and a SD of 1.3h. A Gaussian "de-blurring" applied to these data generates the corrected distribution of cell lag times (red dotted line). The dashed blue line in (A) is the result of simulations in which populations are founded by bacteria with lag times drawn at random from the corrected distribution of cell lag times (see (B)) (the shaded area corresponds to plus and minus SD). The continuous blue line in (A) are results from simulations performed as previously, however, after virtual populations are founded, the lag time of all cells is set to the minimum cell lag time, *ie* synchronization to leader cell (the shaded area corresponds to plus and minus SD).

In Fig. 2A the average lag time from three independent experiments is shown as a function of inoculum size. Lag time decreases monotonically from 6.4 ± 1.1h for droplets inoculated with a single cell, to 4.4 ± 0.3h for an inoculum size of 1,024 cells. Note that the standard deviation (SD), represented by the error-bars in the figure, decreases monotonically and rather slowly with increasing inoculum size.

We also examined the dependence of other growth parameters on inoculum size. Initial experiments showed an effect on final cell density, however this was found to be a consequence of subtle differences in glycerol concentrations arising from dilutions of frozen glycerol-saline stock cultures used to prepare founding inocula. When corrected, no effect of inoculum size on final cell density was observed. This technical, but important experimental observation is explained in Supplementary Fig. 5. Additionally, no effect of inoculum size on mean growth rate was detected, although the variance across droplets decreased. Details are provided in Supplementary Fig. 6.

What might be the basis of the decrease in mean lag-time with inoculum size? There are three mutually exclusive explanations with all three recognising that populations of cells are heterogeneous with regard to individual cell lag times as a consequence of

innate phenotypic variability. Explanation I posits no interaction among cells, with population lag time being set by an event equivalent to a selective sweep that is initiated by the cell (or cells) with the shortest cell lag time.

Explanations II and III involve interactions among cells and can be thought of in terms of two extremes of a continuum. Explanation II posits that all cells contribute equally (mean field) to the production of some growth-stimulatory factor. Explanation III recognises that population lag time might be set by the cell with the shortest lag time and whose activity triggers division of other cells. We demonstrate below that distinguishing between these alternate explanations is possible via quantitative data obtained from the millifluidic droplet device that includes knowledge of the distribution of lag times from populations founded by single cells.

Precise estimation of the distribution of cell-lag times from inocula containing a single bacterium.

To quantify the lag time of individual bacteria, 230 droplets were inoculated by – on average – a single bacterium, resulting in growth in 156 droplets (the inoculation of droplets follows a Poisson process). For each droplet, the lag time was estimated as in Fig. 1A. The resulting distribution of lag times is shown in Fig. 2B (blue dots). In this case, the lag time of each population is equal to the lag time of the founding cell and is denoted θ. The heterogeneity of cell lag times is broad, ranging from $\sim$4 h to $\sim$12 h. The mean value of the data is $m_{exp} = 6.8$ h with SD $\sigma_{exp} = 1.3$ h. A Shapiro-Wilk test applied on the logarithm of the data reveals the underlying distribution to be log-normal (see also the quantile-to-quantile plot shown in Supplementary Fig 9). Fitting a log-normal function (green dashed line in Fig. 2) yields log-normal parameters $\mu = 1.9$ and $s = 0.2$.

Although the fit is good, there is necessarily uncertainty in the estimation of lag times due to the fact that they are calculated by extrapolation (Eq. 1). This uncertainty stands to blur the "true" distribution of lag time. Eq. 1 expresses the dependence of individual lag times, labeled θ, on parameters used in the estimation of this quantity: t_{th}, N_0, N_{th} and λ. The error propagation is estimated by expansion of θ according to a Taylor series, assuming independent variables. This allows the uncertainty $\triangle\theta$ to be calculated as:

$$\triangle\,\theta = \sqrt{(\triangle t_{th})^2 + \left(\frac{\triangle N_{th}}{\lambda N_{th}}\right)^2 + \left(\frac{\triangle N_0}{\lambda N_0}\right)^2 + \left(\frac{\triangle \lambda}{\lambda^2}\right)^2} \qquad (2)$$

where $\triangle t_{th}$, $\triangle N_{th}$, $\triangle N_0$ and $\triangle\lambda$ correspond to the uncertainty of t_{th}, N_{th}, N_0 and λ, respectively. Given the values of these uncertainties, we estimate $\triangle\theta = 0.88$h (see Materials and Methods for details of calculations).

The uncertainty associated with direct measurements thus blurs the "true" distribution of single cell lag times, which is less dispersed, *i.e.*, it has a smaller SD (σ). To estimate the value of σ we consider a Gaussian noise of zero mean and a SD equal to the uncertainty of measurement $\sigma_{noise} = \triangle\theta$. Deconvolution of the Gaussian noise from the measured distribution [4] amounts to subtracting the noise and variance of the noise from that of the measurements to get the mean and variance of the "true" underlying distribution:

$$< \theta >= m_{exp} - m_{noise} \approx 6.8h, \qquad (3)$$

$$\sigma^2 = \sigma^2_{exp} - \sigma^2_{noise} \approx 1.0h^2. \qquad (4)$$

The true distribution remains lognormal. With the mean and variance of the true log-normal it is possible to obtain its parameters μ and s via:

$\mu = log\left(\langle\theta\rangle^2/\sqrt{\sigma^2+\langle\theta\rangle^2}\right)$ and $s^2 = log\left(\sigma^2/\langle\theta\rangle^2+1\right)$. Expression of the true distribution of lag time is thus:

$$f(\theta) = \frac{1}{\theta s\sqrt{2\pi}} e^{-\frac{(\ln(\theta)-\mu)^2}{2s^2}} \quad (5)$$

The red dotted line in Fig. 2B denotes the cumulative distribution function (CDF), corrected for measurement noise and is narrower than that obtained by direct measurement. This distribution can now be used to examine the previously proposed explanations for the dependence of population lag time on inoculum size.

A selective sweep initiated by cells with the shortest lag time is inconsistent with the data.

Having obtained the corrected distribution of single cell lag times, it is now possible to investigate whether the relationship between lag time and inoculum size is explained by sweeps initiated by cell(s) with the shortest lag time (Explanation I). In this case cells are independent and do not interact. To test this hypothesis, growth of populations within droplets established from different numbers of founding cells were simulated and the match with experimental data determined. Thousands of virtual droplets were founded by cells drawn from the true distribution (shown in Fig. 2B) with an exponential growth rate drawn from the distribution of experimental growth rates (see Supplementary Fig. 8). Cells were then allowed to replicate within droplets. To precisely mimic the experimental protocol, the time t_{th} at which populations reach $N_{th} = 64,000$ cells (equal to a density of $1.6 \cdot 10^8$ cell·ml^{-1}) was determined. Eq. 1 was then used to calculate the lag time of each population with known N_0 and with known mean growth rate λ. The blue dotted line in Fig. 2A shows the results of these simulations.

In marked contrast to the experimental results, these simulations of independent (non-interacting) cells show almost no dependence of the mean population lag time on inoculum size. In addition, the decrease in variation across droplets, represented by the SD of lag time (error bars), decreases rapidly, which is contrary to the experimental data, where SD decreases very slowly.

The vanishing SD arising from simulations is understandable as a statistical effect. It is well know that several series of N_0 random numbers normally distributed have a mean value that is also normally distributed, and that the SD of these means decreases as $1/\sqrt{N_0}$. This is precisely what is observed. The lag times of founder cells of a population is a series of independent random lag times drawn from the true distribution of single cell lag times. Populations that are founded with more cells provide a more faithful sample of the single cell lag time distribution (Fig. 2B), and therefore the variation in lag times among populations decreases rapidly with increasing inoculum size. The discrepancy between experiment and model shows that the experimentally determined inoculum effect cannot be explained by independent growth of cells despite the variability of cell lag time.

A simple calculation is further instructive. The true CDF of single cell lag times shown in Fig. 2B has a value of 0.025 for a lag time of 5 h. This means that in a droplet founded by 1024 cells there will be 0.025×1024, or approximately 25 cells with a lag time equal to, or shorter than, 5 h. In the same manner one can estimate from the CDF shown in Fig. 2B that for 1024 founding cells, 737 have a lag time between 5.8 h and 7.8 h ($(0.86-0.14)\times 1024 \approx 737$ cells), close to the mean cell lag time (6.8 h). Given the single cell growth rate $\lambda = 0.84$ h^{-1} (Supplementary Fig. 8), the generation time is $g = 0.83$ h. Thus, the 25 cells that start dividing before 5 h pass through $(6.8-5.0)/g \sim 2$ generations before the 737 cells with a lag time of ~ 6.8 h start

multiplying. Two generations of 25 cells yield $25 \times 2^2 = 100$ offspring. Clearly, cells with a short cell lag time do not exert a dominant sweep effect on the population: the growth rate is too slow for these cells to significantly affect population growth. Besides, there exists only the most marginal correlation between growth rate and cell lag time (Supplementary Fig. 8). This conclusively excludes Explanation I, namely, that a sweep initiated by cells with a short lag time underpins the observed inoculum effect.

A leader cell triggering end of lag time of the population is consistent with the data.

We now turn to Explanations II and III that involve interactions among cells within the founding inoculum. At one extreme case (Explanation III), collective growth is governed by a single event that synchronises population growth to the cell with the shortest lag time. This will happen if the cell that divides first signals this effect to other cells within the founding population, such that the entire population exits lag phase almost simultaneously. We first examine the consequences of this assumption and compare it to the data, and then consider the alternative scenario, namely, Explanation II, in which interactions among cells involve all cells contributing equally (mean field) to the production of a growth-stimulating factor.

In statistical terms, we assume that an inoculum of N_0 cells is a random sample from the single cell lag time distribution $f(\theta)$ (Fig 2B). In the case of a leader cell that triggers growth for all other cells, the measured population lag time will be equal to the shortest cell lag time in the sample, θ_{min}. Several populations starting with N_0 cells gives a collection of θ_{min} (one value coming from each population). Extreme Value Theory (EVT) provides a framework for statistical analysis of the extreme value of a sample, such as the shortest lag time θ_{min} among N_0 cells [7]. In the limit of large samples, EVT predicts the dependence of the mean and variance of a collection of θ_{min} coming from sample of size N_0. It also predicts that the distribution of θ_{min} from populations founded by cells of a given inoculum size will approach a limiting fixed shape after appropriate normalization as the sample size increases; the precise shape is determined by the global properties of the single cell distribution $f(\theta)$.

The unique features of our experiment create an ensemble of droplets with controlled inoculum size, and a measurement of the population lag time for each, labelled τ. These data provide the statistical properties required to test the hypothesis that $\tau = \theta_{min}(N_0)$, namely that the population lag time is equal to the minimum cell lag time among the N_0 single cells of the inoculum. For this we use predictions given by EVT on the distribution of θ_{min} and ask whether they hold for the measures of τ.

The first prediction is that both the average and the SD of θ_{min} from populations founded by cells of a given inoculum size decreases slowly with inoculum size N_0. The precise scaling is derived from the single cell distribution $f(\theta)$ (see Appendix in the Supplementary Information), we find the scaling to be:

$$\langle \theta_{min} \rangle \sim A - B\sqrt{\ln N_0} \tag{6}$$

$$\sigma(\theta_{min}) \sim 1/\sqrt{\ln N_0} \tag{7}$$

where A and B are constants. Strikingly, both predictions agree well with the population lag time τ. Fitting the curve of Eq. 6 to the experimental data on the relationship between mean population lag time ($\langle \tau \rangle$) and inoculum size, reveals a close match Fig. 3A&B. The same holds for the SD fitted to Eq. 7. We note that although testing this prediction involves fitting constants, the dependence on sample size N_0 through $\sqrt{\ln(N_0)}$ is nontrivial and unique to the predictions of EVT.

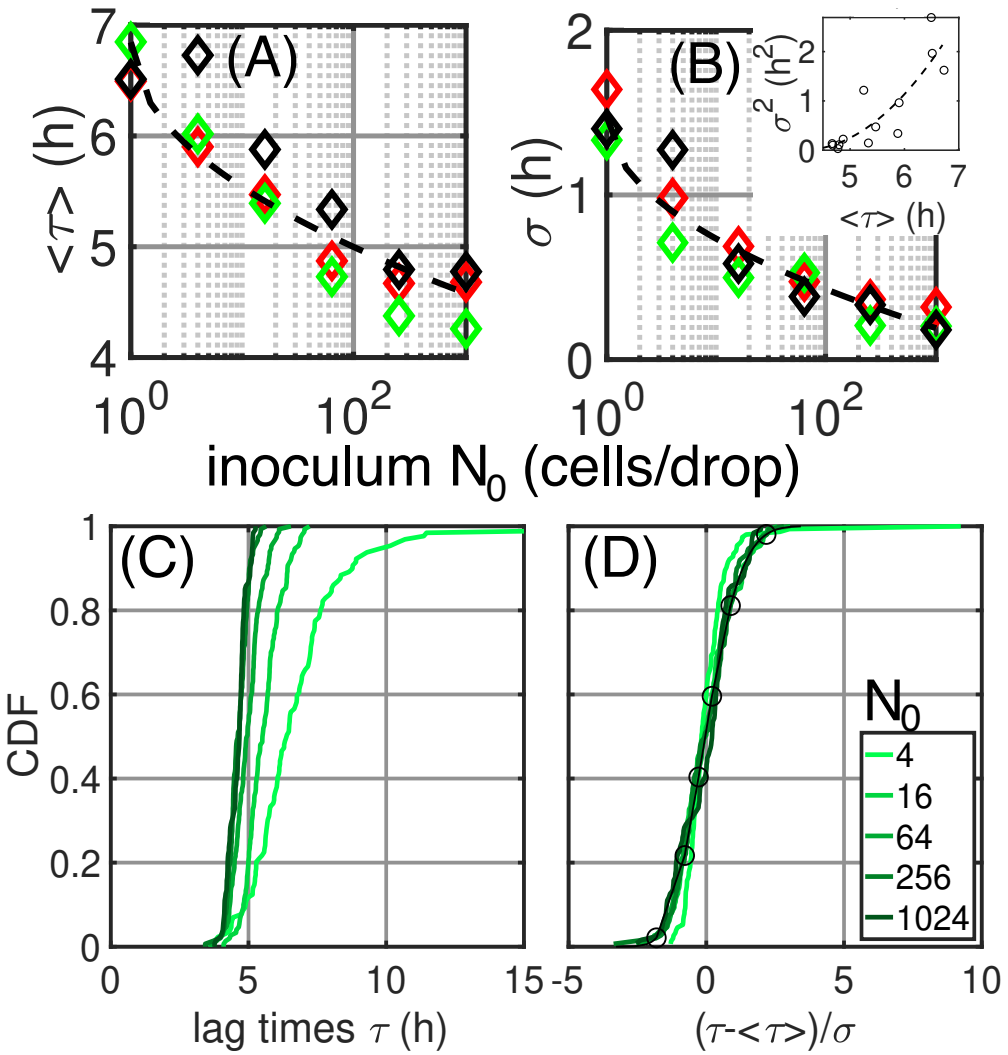

Figure 3. Statistical properties of lag times. For three independent experiments (colors), mean lag times **(A)** over populations and their SD **(B)** are depicted as a function of inoculum size. The scaling relations predicted by EVT are shown in dashed lines: $y = 6.84 - 0.86\sqrt{\ln N_0}$ for the mean and $y = 1/\sqrt{\ln N_0}$ for the SD. The inset in (B) shows the variance (σ^2) as a function of the mean. Each point is calculated over 40 replicate populations (droplets). The dashed line depicts the scaling relation predicted by EVT between the mean and the variance. See Supplementary Appendix for details of the scaling. **(C)** Cumulative Distributions of population lag times for different inoculum sizes (N_0, colors). The curves derive from the pooled data of three independent experiments yielding at least 120 population lag times for each. **(D)** Same data as in (C), scaled by subtraction of empirical mean and division by SD. The black line with circle markers depicts the fit by the universal distribution predicted by the EVT. The y-axis is shared between (C) and (D).

A further prediction is that the distribution of minimal values (θ_{min}), drawn from different sample sizes, tends to a universal shape in the limit of large samples. Although strictly speaking this holds in the limit $N_0 \to \infty$, in practice it may be expected to hold also for finite samples - even relatively small ones. For each sample size N_0, our experiment provides a distribution of population lag times (τ), estimated over all

droplets with the same inoculum size. These CDFs are depicted in Fig. 3 for all inoculum sizes that are above 4 cells. To test whether the prediction on θ_{min} holds for the population lag time τ, we normalize each CDF of τ by subtracting its mean and dividing its SD. Fig. 3B shows the result of this scaling and demonstrates that the distributions of τ collapse on one another is if $\tau = \theta_{min}$. The light blue curve, corresponding to inoculum $N_0 = 4$, deviates from the rest - possibly indicating that this sample size is still too small to acquire the limiting shape.

Moreover, the universal distribution itself is also predicted by EVT. Its CDF has the form:

$$F(\theta_{min}) = e^{-(1+kz)^{-1/k}}, \quad z=(\theta_{min}-\theta_0)/\gamma, \tag{8}$$

with location and scale parameters θ_0, γ and a shape parameter k that reflects properties of the parent distribution $f(\theta)$, specifically the decay at its tails. Fitting the normalized data with this formula reveals an excellent match between the universal distribution and the normalized distributions of τ obtained from populations founded by a given inoculum size (N_0) (black line in Fig. 3B), at least for inoculum sizes above 4 cells per droplet. The analytical formula for the distribution justifies the empirical procedure of normalizing by sample mean and standard deviation used above as a test for the universal shape (see Appendix in Supplementary Information).

As a corollary of the predictions in Eqs. 6,7, the variance and mean of the distributions of extreme values drawn from different sample sizes follow a well-defined relationship (see Appendix). The agreement of this relationship with the population lag time data is shown in the inset of Fig. 3B.

Taken together, the scaling of the mean and SD of τ according to the inoculum size (Eq. 6 & 7), the resulting relationship between variance and mean of τ, the collapse of normalized distributions of τ at different sample (inoculum) sizes, and the fit of the normalized distribution to the theoretical formula Eq. 8, are consistent with population lag time being equal to the shortest cell lag time θ_{min} in the inoculum. We conclude that, at the time of the shortest lag time in the inoculum, many cells must start growing in parallel, *ie* $\tau = \theta_{min}$.

A single leader cell determines population lag time.

The agreement of statistical properties with predictions from EVT suggest that exit from lag phase is triggered by a single event – possibly a single leader cell – that signals the exit from lag to all other cells. To test this hypothesis we performed a set of simulations where an additional condition was introduced. As for previous simulations at each inoculum size, thousands of virtual populations were generated (see simulation code in the Appendix of Supplementary Information). For each population the cell lag time of each founder cell was drawn at random from the experimental single cell lag time distribution Fig. 2B. However, here, a new condition was added, namely, that all founder cells have a single cell lag time that is equal to the minimal cell lag time of the leader cell. Hence for each virtual population, we first draw the cell lag time values of founding cells, and then set these values to the minimal value of the shortest cell lag time in that population. Note that the minimal cell lag times differ across populations because each population is founded by cells drawn at random from the distribution Fig. 2B. This mimics the situation in which a leader cell with the shortest cell lag time in the population triggers growth of all the other cells when it exits lag phase ($\tau = \theta_{min}$). As in the previous simulation, the numerical population was allowed to grow exponentially with population lag time being estimated as per the experiments.

The results are depicted by the plain line in Fig. 2A and are a close match to the experimental data over three orders of magnitude. Note that this is not a fit: the only input is the true single cell lag distribution measured for populations founded by one

cells (Fig. 2B). Additionally, results from the simulations match the slow decrease of lag time variability among droplets observed in the experiments (shaded area Fig. 2A). Our results thus provide an explanation for the relationship between size of the founding inoculum and population lag time and are fully consistent with resumption of growth of all the cells of the inoculum triggered by an event effected by a *leader cell*. This rules out Explanation II that posits that all cells contribute equally to the exit of lag phase.

Thus far it is not possible to be certain as to whether population growth is triggered by a single cell, or a small pool of cells, each having a short lag time. To investigate if the experimental population lag time can be explained by the presence of a single leader cell, rather than the presence of a small pool of leader cells, we performed a further simulation. In this simulation we consider that cells produce a growth activator. When concentration of the activator reaches a threshold, it triggers growth of all cells in the inoculum that are still in lag phase. The concentration threshold of the growth activator determines the size of the small pool of cells that contribute to its production. If the threshold is low, it is sufficient that a single cell produces the growth activator. If the threshold is set high, then multiple cells can end their inner lag phase before the threshold is reached and contribute to the production of the activator. The duration of lag phase for every cell, is as before, set by drawing a random value from the experimental cell-lag distribution Fig. 2B (see code in Supplementary appendix).

Running simulations for a range of activator thresholds at different inoculum sizes allows investigation of the number of cells that are able to end lag phase before the occurrence of the triggering event. The results are depicted in the Supplementary Information Figs 14&15. Evident is a strong dependence of population lag time on inoculum size only when the threshold concentration of the growth activator is low. Strikingly, we see that for a large range of activator concentrations, only a single cell as the time to exit its inner lag phase before the triggering of the growth of all cells in the population. This leads us to conclude that our experimental observation are fully consistent with Explanation III, where a single leader cell to end lag phase for the entire population.

Discussion

An inverse relationship between the number of bacterial cells founding a new environment and the time to exit lag phase was first noted more than 100 years ago. Despite its significance, rigorous validation has been lacking, and understanding of the causes and controlling factors remain incomplete. Lack of progress has stemmed largely from difficulties associated with experimental analysis of populations founded by few cells. Progress requires ability to follow the growth dynamics of replicate populations established from precisely determined numbers of founding cells. Additionally required is knowledge of the distribution of lag times for hundreds of single cells. Once known, statistical approaches can be used to link the distribution of population lag times, to the behaviour of single cells. The nature of the relationship stands to shed insight on cell-level causes of the inoculum effect.

Here, taking advantage of new opportunities presented by millifluidic technologies we have obtained quantitative evidence from highly replicated populations founded by controlled numbers of cells, that in populations of *P. fluorescens* SBW25, the time to resume growth after transfer to a new environment is strongly influenced by size of the founding inoculum. Moreover, the same technology has allowed determination of the duration of lag phase for a sample of individual cells. The average decrease in time to growth resumption, variance across droplets, and distribution shapes, follow predictions of extreme value theory, consistent with the inoculum lag-time being determined by the minimum value sampled from the single-cell distribution. Our experimental results thus

show that exit from lag phase depends on strong interactions among cells, consistent with a "leader cell" triggering end of lag phase for the entire population.

This finding builds on recent work in which the time to first division of single bacteria maintained in isolated cavities of microfluidic devices has been measured [2, 9, 14, 25]. From such studies it is clear that there is substantial variation in cell-level lag time with evidence that this variance can have profound fitness consequences for population growth. For example, in fluctuating environments, heterogeneity in the time for individual cells to resume growth, can facilitate survival in the face of environmental change [9], especially that wrought by periodic antibiotic stress [2, 14, 25].

Although microfluidic chambers used for analysis of isolated cells allow precision measures of the distribution of lag times for single cells, such experimental devices do not allow for interactions among cells, thus making problematic any attempt to connect the distribution of single cell lag times to population lag times. In fact, extrapolation of population lag times from knowledge of the distribution of single cell lag times would be legitimate exclusively in the case of there being no interaction among cells.

Leaving aside previous evidence that interactions mediated by growth factors are integral to the inoculum effect [10, 11], linking cell and population level behaviours necessarily requires the consideration of interactions among cells and thus it is essential to obtain measures of both the lag time of individual cells and the lag time of populations established from known inoculum sizes in precisely the same environment. Moreover, the environment should be well mixed (spatially homogeneous and devoid of surface effects) so that should emergent population-level behaviours be relevant, mediated via, for example, production of diffusible growth factors, then the experimental protocol allows for their effects to be captured. In this regard the millifluidic device has proven fit for purpose.

In seeking an explanation for the observed inoculum effect we considered three mutually exclusive explanations. Central to Explanation I was absence of interactions among cells with the inoculum effect being explained by disproportionate growth of the set of cells with the shortest time to first cell division. Both simulations and simple calculations led to unequivocal rejection of this hypothesis.

Explanations II and III recognised the possibility of interactions among cells. Because of the power of extreme value theory combined with well understood statistical properties we chose to focus on whether population lag times were determined by the minimum value sampled from the single-cell distribution (Explanation III). EVT makes predictions as to the distribution of minimal cell lag times across droplets, which surprisingly, also held for the distribution of population lag times measured in the experiments, leading to the conclusion that the population lag time is equal to the minimal cell lag time present in the inoculum. Simulations of population growth in droplets based on this evidence delivered an almost perfect match to experimental data.

While conformity to Explanation III means that Explanation II in a strict sense (in which all cells contribute equally (mean-field) to exit from lag phase) cannot hold, the fact that there is a continuum of possibilities led us to perform additional simulations to address whether our data are consistent with resumption of growth being triggered by just a single leader cell. The findings fit surprisingly well and are fully consistent with a single leader cell being sufficient to trigger resumption of growth for the entire population.

An obvious next question concerns identity of the growth activator. While detailed investigation are beyond the scope of this study, we nonetheless, considered the possibility that iron chelation might trigger the effect. Such a possibility has been previously suggested [10]. To this end we repeated the initial experiment in which the time to resumption of growth was determined in replicate populations founded by

different numbers of cells as in Fig 2A. Instead of SBW25 a mutant deficient in production of the iron chelating compound pyoverdin was used *P. fluorescens* SBW25 *pvdS*G229A [24]. No change in the inoculum effect was observed (see Supplementary Fig 13), thus ruling out pyoverdin as the activating molecule.

The relationship between the number of cells founding growth in a new environment and duration of lag phase has profound implications for microbiology. While much remains to be understood, including generality and molecular bases, the rigorous quantification achieved here via a millifluidic device provides unequivocal proof of an inoculum effect in *P. fluorescens* SBW25. Moreover our statistical analysis of the distribution of population lag times is consistent with the activity of a single leader cell triggering simultaneous growth for all cells in the founding population.

Acknowledgments

We thank Millidrop for development of the prototype Azur3 and their willingness to engage in active collaboration. Discussion and input from Arjan De Visser, Jerome Bibette, Andrew Farr, Lukas Geyrhofer, Isabelle Rivals, Clara Moreno-Fenoll, Jean Baudry, Nicolas Bremon, Jean-Baptiste Dupin, Wilfried Sire, Ankur Chaurasia is gratefully acknowledged. The work was supported by an HFSP grant "Interrogating bacterial social interactions in droplets" RGP0010/2015

Author contributions

M.A. & P.B.R. designed the experiment and determined the strains and relevant experimental conditions, M.A. performed the experiments, M. A. & G.D. developed the software to analyse the data, M.A & N. B. made the modeling and simulation, M.A., N.B., P.B.R. wrote the paper.

Supplementary Appendix

Codes of simulation

First we give the code of the simulation to generate the Figs. 2A and 14&15:

```
close all
%clear all
Vdrop = 4.4e-4; %ml

Ndrop=1000;
inocula = logspace(0,3,10); %range of inocula to simulate
dt = 1/60;
timeSpan = 0:dt:30;

Nthresh = 1.6e8; %cell/ml. Threshold to calculate the lag time like in the
    experiments.

SDNoise = 0.88; %SD of the noise

MM = 6.8246; %mean value of the Exp distribution of lag for logn
VV = 1.3322^2 - SDNoise^2; %SD of the Exp distribution of lag for logn
%calculate the new param of the logn to remove the noise
mu = log(MM^2 / sqrt(MM^2 + VV));
s = sqrt(log(VV/MM^2 + 1));
```

```
mg = 0.8430; %average grate %1/h from inoculum 1
vg = 0.02; %SD grate from inoculum 1

%% simulation of the simple growth of bacteria in the droplets. Measure of the
    lag for each droplet.

tmes = zeros(Ndrop,length(inocula));

clear timeSeries tmes
for inoc = 1:length(inocula) %loop over the inoculum
   for i = 1: Ndrop %loop over the drop

       r = round(poissrnd(inocula(inoc))); %draw a random inoculum according to
           the poisson distribution.

       timeSeries = ones(r,length(timeSpan))*nan; %timeSerie of the growth for
           each bacteria in this drop
       clear tlag
       tlag = lognrnd(mu,s,r,1); %draw randomly the cell-lag according to the
           corrected cell-lag distribution measured.

       grate = normrnd(mg,vg,r,1); %draw randomly the cell growth rate
           according to the growth rate distribution measured.

       timeSeries(:,:)= exp(grate.*(timeSpan-tlag)); % proceed to the
           exponential growth of every bacteria of this drop.
       timeSeries(timeSeries<1)=1; % the timeseries must start at 1 before the
           division of the bacteria

       %calculation of the cell concentration in this drop along the
       %time by suming the timeSerie of its bacteria
       l = size(timeSeries);
       if l(1)==1
           totDrop = timeSeries/Vdrop;
       else
           totDrop = nansum(timeSeries)/Vdrop;
       end

       %measure the lag of the droplet by finding the time at which the
       %cell concentration gets above Nth like in the experiments
       if r~=0
           tau = timeSpan(find(totDrop>Nthresh,1,'first'));
           if isempty(tau)
               tmes(i,inoc) = nan;
           else
               tmes(i,inoc) = ...
               timeSpan(find(totDrop>Nthresh,1,'first'))...
               -log(Nthresh*Vdrop/r)/nanmean(grate);
           end
       else
           tmes(i,inoc) = nan;
       end
   end
end
```

```
for i = 1:length(inocula)
    tstat(i) = nanmean(tmes(:,i));
    tstsdstat(i) = nanSD(tmes(:,i));
end

%% synchronistation demonstration of the synchronisation effect
clear tmes tsync tstsdsync
tmes = zeros(Ndrop,length(inocula));
for inoc = 1:length(inocula)

    for i = 1: Ndrop

        r = round(poissrnd(inocula(inoc))); %draw a random inoculum according to
            the poisson distribution.

        if r~=0
            timeSeries = ones(r,length(timeSpan))*nan; %timeSerie of the growth
                for each bacteria in this drop

            clear tlag
            %log-normal
            tlag = lognrnd(mu,s,r,1); %draw randomly the cell-lag according to
                the corrected cell-lag distribution measured.

            tlag = ones(r,1) .* (min(tlag)); %set all the cell-lags to the
                minimal cell-lag of the leader

            grate = normrnd(mg,vg,r,1); %draw randomly the cell growth rate
                according to the growth rate distribution measured.


            timeSeries(:,:)= exp(grate.*(timeSpan-tlag)); % proceed to the
                exponential growth of every bacteria of this drop.
            timeSeries(timeSeries<1)=1; % the timeseries must start at 1 before
                the division of the bacteria

            %calculation of the cell concentration in this drop along the
            %time by suming the timeSerie of its bacteria
            l = size(timeSeries);
            if l(1)==1
                totDrop = timeSeries/Vdrop;
            else
                totDrop = nansum(timeSeries)/Vdrop;
            end

            %measure the lag of the droplet by finding the time at which the
            %cell concentration gets above Nth like in the experiments
            tau = timeSpan(find(totDrop>Nthresh,1,'first'));
            if isempty(tau)
                tmes(i,inoc) = nan;
            else
                tmes(i,inoc) = ...
                timeSpan(find(totDrop>Nthresh,1,'first'))...
                -log(Nthresh*Vdrop/r)/nanmean(grate);
            end
        else
            tmes(i,inoc) = nan;
```

```
        end
    end
end

for i = 1:length(inocula)
    tsync(i) = nanmean(tmes(:,i));
    tSDsync(i) = nanSD(tmes(:,i));
end

%% plot the curves of stat effect vs sync effect need to run the cell above
    first.

%experimental points
lag = [ 6.564471 6.018853 5.392048 4.729353 4.375776 4.259628; ...
    6.349308 5.901676 5.470768 4.871843 4.669578 4.680528; ...
    6.377465 5.937269 5.553422 5.204688 4.640315 4.353197]';

slag = [...
    0.843766 0.700523 0.491303 0.519735 0.203902 0.199777;...
    1.445141 0.974685 0.682811 0.469918 0.363288 0.316424; ...
    1.120196 0.956412 0.731424 0.720511 0.455629 0.434152 ]';

N0 = repmat([1 4 16 64 256 1024], size(lag,2), 1)';


%plot the figure

figure('Renderer', 'painters', 'Position', [10 10 900 900]),
hold on

alpha = 0.1;
y = tsync; % your mean vector;
x = log(inocula);
SD_dev = tSDsync;
curve1 = y + SD_dev;
curve2 = y - SD_dev;
x2 = [x, fliplr(x)];
inBetween = [curve1, fliplr(curve2)];
fill(x2, inBetween, 'b','FaceAlpha',alpha);
plot(x, y, '.-b', 'LineWidth', 5,'MarkerSize',20,'DisplayName', 'Leader');

y = tstat; % your mean vector;
x = log(inocula);
SD_dev = tstsdstat;
curve1 = y + SD_dev;
curve2 = y - SD_dev;
x2 = [x, fliplr(x)];
inBetween = [curve1, fliplr(curve2)];
fill(x2, inBetween, 'b','FaceAlpha',alpha);
plot(x, y, '.-.b', 'LineWidth', 5,'MarkerSize',20,'DisplayName', 'Stat');

clr = {'b' 'r' 'g' 'k'};
for i = 1:size(lag,2)
    h = errorbar(log(N0(:,i)), lag(:,i), slag(:,i),'d', 'MarkerSize',20,
        'color', clr{i+1}, 'LineWidth',3, 'CapSize', 40);
end
```

```
set(gca,'FontName','Helvetica')
xlim([-0.2 log(1500)])
xticks( log(N0(:,1)));
xticklabels(N0(:,1));
xlabel('inoculum (cell/drop)');
ylabel('mean lag time (h)');
title(['N0=1 noise SD=' num2str(SDNoise) ])
box('on')
grid('on')
set(gca,'LineWidth',4)
set(gca,'GridAlpha', 0.5)
set(gca,'FontSize',60)
hold off
```

```
%%%The following code corresponds to the simulation producing the figures
    13&14%%%

close all
%clear all
Vdrop = 4.4e-4; %ml

Ndrop=1000;
inocula = logspace(0,3,10); %range of inocula to simulate
dt = 1/60; %heures
timeSpan = 0:dt:30; %heures

Nthresh = 1.6e8; %cell/ml. Threshold to calculate the lag time like in the
    experiments.

stdNoise = 0.88; %SD of the noise 0.87 with lambda and 2 sigma for the calib
stdNoiseTitle = stdNoise;
MM = 6.8246; %mean value of the Exp distribution of lag for logn
VV = 1.3322^2 - stdNoise^2; %std of the corrected distribution of the
    experimental lag that gollows a logn
%calculate the new param of the logn to remove the noise
mu = log(MM^2 / sqrt(MM^2 + VV));
s = sqrt(log(VV/MM^2 + 1));

mg = 0.8430; %average grate %1/h from inoculum 1
vg = 0.02; %SD grate from inoculum 1

spanThActv = logspace(-5,2,11);

rActv = mg/log(2); %the rate of production of growth activator is the inverse
    of the doubling time of cells.

%% synchronistation demonstration of the synchronisation effect
clear lagPop stdLagPop  decsyncPop leadsyncPop

tmes = zeros(Ndrop,length(inocula));
k=0;
for thActv = spanThActv
    k = k +1;
    clear tmes tsync tstsdsync decsync dec lead leadsync
```

```
dec = ones(Ndrop,length(inocula))*nan;
lead = ones(Ndrop,length(inocula))*nan;
tmes = ones(Ndrop,length(inocula))*nan;

for inoc = 1:length(inocula)

    for i = 1: Ndrop

        r = round(poissrnd(inocula(inoc))); %draw a random inoculum
            according to the poisson distribution.
        if r~=0
            timeSeries = ones(r,length(timeSpan))*nan; %timeSerie of the
                growth for each bacteria in this drop
            timeSeriesActv = zeros(r,length(timeSpan)); %timeSerie of the
                growth activator for each bacteria in this drop
            clear tlag
            %lognormal
            tlag = lognrnd(mu,s,r,1); %draw randomly the cell-lag according
                to the corrected cell-lag distribution measured.
            grate = normrnd(mg,vg,r,1); %draw randomly the cell growth rate
                according to the growth rate distribution measured.


            timeSeriesActv = exp(grate.*(timeSpan-tlag))-1; %production of
                molecule is linear with number of cells so it follows the
                exp growth
            timeSeriesActv(timeSeriesActv<0) = 0; % molecule contration
                cannot be negative.
            timeSeriesActv = (timeSpan-tlag).*rActv.*timeSeriesActv;
                %multiplication by time and production rate.

            %calculate the total concentration of molecule produced by all
            %the cell in the drop. Need of condition for drop with one cell
            %(no need to sum over cells)
            if r>1
                actv = sum(timeSeriesActv);
            else
                actv = timeSeriesActv;
            end

            %find the time at which the concentration of molecule gets
            %above a given threshold.
            tActv = timeSpan(find(actv>thActv,1,'first'));

            dec(i,inoc) = tActv - min(tlag); % difference of lag time of the
                leader cell and the lag time due to production of molecule.
            lead(i,inoc) = sum(tlag<=tActv);


            tlagActv = tlag;
            tlagActv(tlagActv>tActv)=tActv; %every cells lag time end when
                the molecule gets above the threshold.


            timeSeries(:,:)= exp(grate.*(timeSpan-tlagActv)); % proceed to
                the exponential growth of every bacteria of this drop.
            timeSeries(timeSeries<1)=1; % the timeseries must start at 1
```

```
                    before the division of the bacteria

                %calculation of the cell concentration in this drop along the
                %time by suming the timeSerie of its bacteria
                l = size(timeSeries);
                if l(1)==1
                    totDrop = timeSeries/Vdrop;
                else
                    totDrop = nansum(timeSeries)/Vdrop;
                end

                %measure the lag of the droplet by finding the time at which the
                %cell concentration gets above Nth like in the experiments
                tau = timeSpan(find(totDrop>Nthresh,1,'first'));
                if isempty(tau)
                    tmes(i,inoc) = nan;
                else
                    tmes(i,inoc) =
                        timeSpan(find(totDrop>Nthresh,1,'first'))-log(Nthresh*Vdrop/m)/nan
                end
            else
                tmes(i,inoc) = nan;
            end
        end
    end

    for i = 1:length(inocula)
        tsync(i) = nanmean(tmes(:,i));
        tstdsync(i) = nanstd(tmes(:,i));
        decsync(i) = nanmean(dec(:,i));
        leadsync(i) = nanmean(lead(:,i));
    end

    lagPop(k,:)=tsync;
    stdLagPop(k,:)=tstdsync;
    decsyncPop(k,:)=decsync;
    leadsyncPop(k,:)=leadsync;
end

%%
close all


Y = repmat(log10(spanThActv)',1,10);
X = repmat(log10(inocula),11,1);
Z = lagPop;

%population lag time vs threshold and inoculum
figure('Renderer', 'painters', 'Position', [10 10 900 900]),
surf(X,Y,Z);

%title('population lag time')
xlabel('inoculum')
xticks(log10(round(inocula)))
xticklabels(round(inocula))
```

```
ylabel('threshold')
yticks(log10(spanThActv(1:2:end)))
yticklabels(num2str(spanThActv(1:2:end)','%1.0e'))
zlabel('population lag time')
c = colorbar;
set(gca,'FontSize',30)
c.Location='northoutside';
view(35.207879105520632,39.388548057259705)

% population lag time minus leader cell lag time
figure('Renderer', 'painters', 'Position', [10 10 900 900]),
Y = repmat(log10(spanThActv)',1,10);
X = repmat(log10(inocula),11,1);
decsyncPop(decsyncPop<=dt)=0;
Z = decsyncPop;

surf(X,Y,Z);

%title('time difference between cell leader lag time and lag time of
    population')
xlabel('inoculum')
xticks(log10(round(inocula)))
xticklabels(round(inocula))

ylabel('threshold')
yticks(log10(spanThActv(1:2:end)))
yticklabels(num2str(spanThActv(1:2:end)','%1.0e'))
zlabel('lag pop - lag leader cell')
c = colorbar;
set(gca,'FontSize',30)
caxis([0, 2]);
c.Location='northoutside';
view(35.207879105520632,39.388548057259705)

%number of cells that multiply before synchro
figure('Renderer', 'painters', 'Position', [10 10 900 900]),
Y = repmat(log10(spanThActv)',1,10);
X = repmat(log10(inocula),11,1);
Z = leadsyncPop;

surf(X,Y,Z);

%title('number of cell leaders')
xlabel('inoculum')
xticks(log10(round(inocula)))
xticklabels(round(inocula))

ylabel('threshold')
yticks(log10(spanThActv(1:2:end)))
yticklabels(num2str(spanThActv(1:2:end)','%1.0e'))
zlabel('number of leader cells')
set(gca,'FontSize',30)
c = colorbar;
caxis([1 5]);
c.Limits = [1 5];
c.Ticks = [1 2 3 4 5 6 7 8];
c.Location='northoutside';
```

```
view(35.207879105520632,39.388548057259705)
```

Statistical properties for extreme values

Extreme Value Theory tells us that the minima of finite samples, drawn from some parent distribution, exhibit predictable statistical properties in the limit of large sample sizes [7]. These properties are testable experimentally in our system, where we can control the sample size - inoculum size N_0, and repeat the sampling many times. The theory predicts how the mean and variance of sample minima depend on N_0, given the parent distribution. It moreover contains the powerful statement that the distribution shape converges to a universal one in the limit $N_0 \to \infty$, which reflects the decay of the original distribution tails. Similar statements hold for maximum as well as other extreme values (e.g. second largest, etc).

To develop some intuition for the scaling of moments with sample size, imagine drawing N_0 random variables from a normal Gaussian distribution. To estimate the minimal drawn number, we divide the real line into N_0 equal-probability bins (see illustration below). On average, there will be one number drawn from each bin; therefore we may estimate that the minimal value lies within the lowest bin, in the range $(-\infty, x_1)$. Clearly, the larger is the sample, the more bins we can use and still have an average of one number in the lower one; as the sample becomes large, we are increasing our chance of sampling low-probability events in the tail of the distribution.

To get the qualitative nature of the scaling, we seek a relation between the sample size N_0, also the number of bins, and the values of the continuous variables x in the lowest bin. Comparing the probabilities,

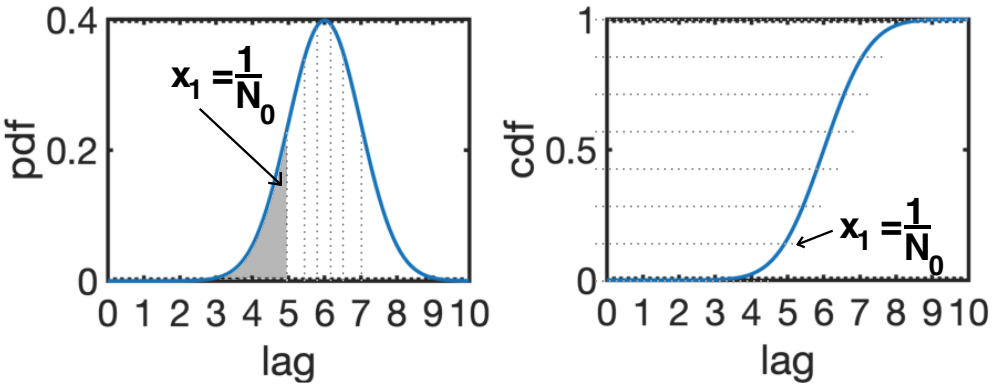

Figure 4. Drawing a finite sample of size N_0 from a continuous Gaussian distribution. To estimate the minimal value drawn out of N_0 values, we divide the real line into N_0 equi-probable bins; in the figure, $N_0 = 7$. The estimate that the minimum resides in the lowest bin can be used to derive the scaling $\sqrt{\ln N_0}$ in the limit of large N_0. Left we represent the probability density function of a Gaussian (mean 6h, SD 1h). The probability to have value in the lowest bin between 0 and $x_1 = \frac{1}{N_0}$ is equal to the area of this bin highlighted in grey. Right we represent the cumulative density function of the same Gaussian. The probability to have a value below or equal to $x_1 = \frac{1}{N_0}$ is equal to the y-value of the lower dotted line pointed out by the arrow.

$$\frac{1}{N_0} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_1} e^{-x^2/2} dx.$$

For large N_0, x_1 is very small and the integral is dominated by the uppermost limit. This means that if a random number is drawn from the lower bin, it is highly likely

equal (or close to) the upper end of the bins; the chance of getting other values inside the bin are exponentially smaller. Therefore we approximate

$$\int_{-\infty}^{x_1} e^{-x^2/2} dx \approx e^{-x_1^2/2}.$$

With this relation, we can solve for $x_1(N_0)$ in Eq. to find

$$x_1 \sim \sqrt{\ln(N_0)}.$$

This value of x_1 gives the scaling behavior of the minimal random number out of a large sample N_0, on average. Computing the pre-factors correctly and estimating the sub-dominant terms requires more careful approximations.

Although this argument is made for Gaussian random variables, the same scaling holds also for the lognormal ones. To account for the mean and variance of the parent distribution, in Fig. 3 we shift and scale by the nonzero mean and non-unit SD.

The most general form of the extreme value distribution was specified in Eq. 8 where θ_0 is the location parameter, γ the scale parameter and k the shape parameter. The mean and variance of this distribution generally depend on its parameters - and this dependence varies between $k = 0$ (corresponding to the Gumbel distribution) and $k \neq 0$ (corresponding to the Frechet or Weibull families). From a practical perspective, we show below that empirically it is not required to know the actual parameters of the distribution in order to test for their scaling property; it is sufficient to empirically subtract the average and divide by the standard deviation.

For all cases where the first two moments exist, they are

$$\langle \theta_{min} \rangle = \begin{cases} \theta_0 + \frac{\gamma}{k}\left[\Gamma(1-k) - 1\right] & k \neq 0\,, \\ \theta_0 + \gamma\epsilon & k = 0 \end{cases}$$

$$\sigma^2(\theta_{min}) = \begin{cases} \frac{\gamma^2}{k^2}\left[\Gamma(1-2k) - \Gamma(1-k)^2\right] & k \neq 0\,, \\ \gamma^2 \frac{\pi^2}{6} & k = 0. \end{cases}$$

Where ϵ is the Euler–Mascheroni constant and Γ the gamma function. Although these are cumbersome expressions, they have the simple form

$$\langle \theta_{min} \rangle = \theta_0 + \gamma f_1(k)$$

$$\sigma^2(\theta_{min}) = \gamma^2 f_2^2(k).$$

Therefore, even without estimating the nonlinear parameters of a distribution, the two-parameter scaling by the two first moments amounts to an affine transformation of the variable:

$$z = \frac{\theta_{min} - \langle \theta_{min} \rangle}{\sigma(\theta_{min})} = \frac{\theta_{min} - [\theta_0 + \gamma f_1(k)]}{\gamma \sqrt{f_2(k)}} = a\theta_{min} + b$$

with a, b constants that can depend on k. It is not difficult to show (see section below) that if a variable is distributed according to a Generalized Extreme Value distribution (GEV) with shape parameter k, $\theta_{min} \sim GEV(\theta_0, \gamma, k)$, then affine- transformed variables $z = a \times \theta_{min} + b$ are also GEV-distributed, with modified scale and shift parameters but with the same shape parameter: $a \times \theta_{min} + b \sim GEV(\tilde{\theta}_0, \tilde{\gamma}, k)$. This property is known as the invariance of the GEV distributions under affine transformations. Therefore, the distribution collapse of sampled data after empirical

scaling by mean and SD provides a test for their common shape, and thus for their consistency with the extreme value properties.

Using the general expressions for mean and variance, we may express $\sigma(\theta_{min})$ as a function of $\langle\theta_{min}\rangle$

$$\sigma^2(\theta_{min}) = \frac{f_2(k)}{f_1(k)^2}\left(\langle\theta_{min}\rangle - \theta_0\right)^2. \tag{9}$$

This relation was also tested in the measured data, as is depicted in the inset of the Fig. 3B.

Invariance of GEV under affine transformation

We here show the invariance of distribution shape under affine transformation, for the entire GEV family of distributions. Suppose x is distributed according to Generalized Extreme Value distribution $x \sim GEV(x_0, \gamma, k)$ where x_0 is the centering parameter, γ the scaling parameter and k the shape parameter. The cumulative form of the GEV is given by:

$$F_X(x) = \exp\left[-\left(1 + \frac{k}{\gamma}(x - x_0)\right)^{-\frac{1}{k}}\right]$$

Now consider the variable $y = ax + b$: the cumulative form of the distribution of y ($F_Y(y)$) can be derived from the distribution of x ($F_X(x)$): $F_Y(y) = \text{Prob}\left(Y \leqslant y\right) = \text{Prob}\left(aX + b \leqslant y\right) = \text{Prob}\left(X \leqslant \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$. Therefore using the expression of F_X we can find F_Y. Using $y_0 = ax_0 + b$,

$$\begin{aligned} F_Y(x(y)) &= \exp\left[-\left(1 + \frac{k}{\gamma}(\frac{y-b}{a} - x_0)\right)^{-\frac{1}{k}}\right] \\ &= \exp\left[-\left(1 + \frac{k}{\tilde{\gamma}}(y - y_0)\right)^{-\frac{1}{k}}\right] \end{aligned}$$

where $\tilde{\gamma} = a\gamma$. This shows that y is also distributed according to a GEV distribution, with modified shift and scale parameters but with the same shape parameter k.

It can be seen directly from the general expression of the mean and variance of the GEV family, that the scaled variable

$$z = \frac{x - \langle x\rangle}{\sigma(x)} = \frac{x}{\gamma f_2(k)} - \frac{x_0 + \gamma f_1(k)}{\gamma f_2(k)} \sim \text{GEV}\left(\frac{f_1(k)}{f_2(k)}, \frac{1}{f_2(k)}, k\right)$$

This GEV distribution of z has a mean $\langle z\rangle = 0$, a standard deviation $\sigma(x) = 1$ and only a single parameter k. This property of extreme values distribution is used in the main text Fig.3D to demonstrate the equality between population lag times and leader cell lag time.

Supplementary figures

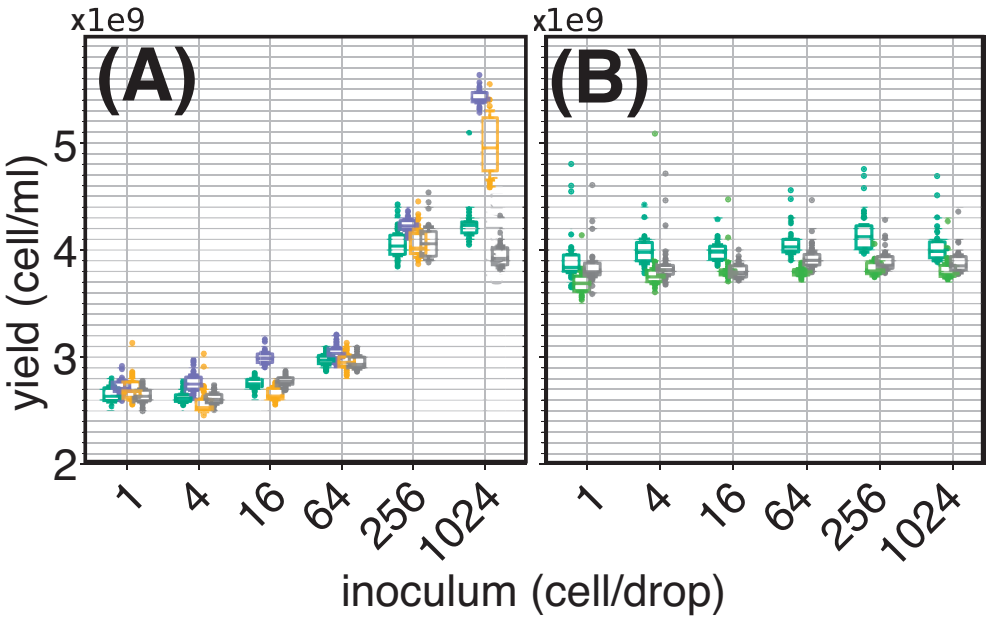

Figure 5. The maximal density of bacteria in the droplets depends on the abiotic inoculum. We measure the maximal density of bacteria reached for 6 different inoculum (1, 4, 16, 64, 256, 1,024 cells) in rich medium. Each box plot is calculated from 40 droplets (30 for the inoculum 1,024). The different colours correspond to independent replicates. **(A)** shows the yield of droplets prepared from a 100µl glycerol stock with a "naive" serial dilution that does not compensate the glycerol concentration across bacterial inocula. We observe that the maximal density of bacteria depends on the inoculum. At high dilution, *ie* small bacterial inocula, the maximal density reaches $2.6 \cdot 10^9$ cell·ml^{-1} whereas at low dilution, *ie* large inoculum, the maximal density goes above $4 \cdot 10^9$ cell·ml^{-1}. (A) and (B) share the same y-axis. **(B)** shows the maximal density of an "aware" serial dilution such as the density of glycerol is kept constant across the bacterial inocula. To do so the glycerol density is balanced by addition of an appropriate volume of 60% glycerol in the inocula. We see that balancing the glycerol in the droplets results in a constant maximal density of bacteria, whatsoever the inoculum of bacteria in the droplets. Thus, traces of glycerol coming from the frozen stocks influence the maximal density of bacteria. Diluting "naively" the glycerol of the frozen stock by 70x (together with the cells) yield an increase of 150% of the maximal bacterial density reached in droplets compare to a dilution of 70,000x. Thence, The millifluidic technology has the sensitivity to measure precisely such abiotic effect. In our work we always took care of balancing the glycerol concentration in the culture to keep it constant across inocula.

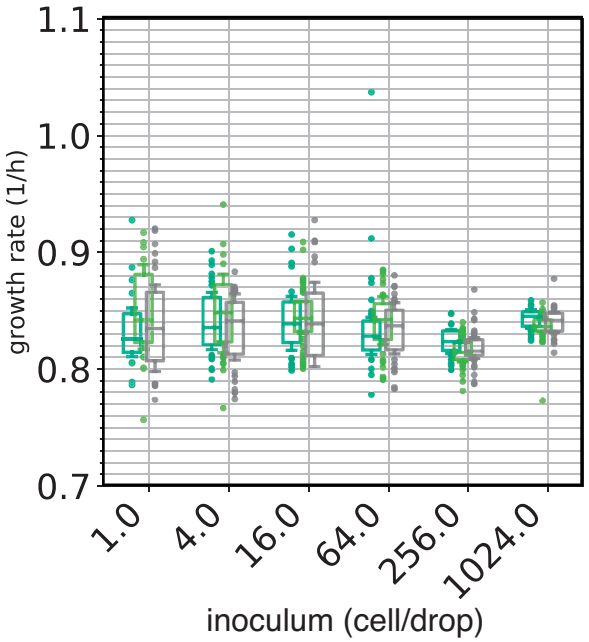

Figure 6. The growth rate of the droplets does not depend on inoculum size. Boxplots of growth rate as a function of inoculum size estimated from time-series of 40 droplets (except 1,024 which is 30). The data are the same as in Fig 2A that shows the lag time of these droplets. The growth rate is approximately constant, with a median at $0.84 \pm 0.02\ h^{-1}$. The colors correspond to three independent experiments.

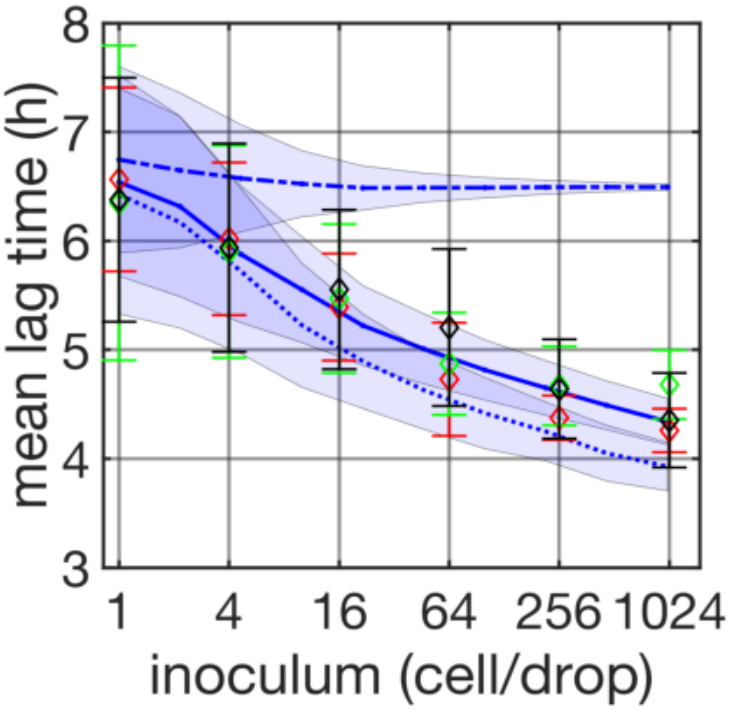

Figure 7. Effect of deconvolving the measurement error on single-cell lag-time distribution. The same data shown in Fig. 2A are shown here. Results of the simulation where no deconvolution is done on the cell lag time distribution Fig. 3B is shown in dotted line. This simulation uses directly the raw cell-lag. It under-estimates significantly the mean lag time as a function of inoculum size.

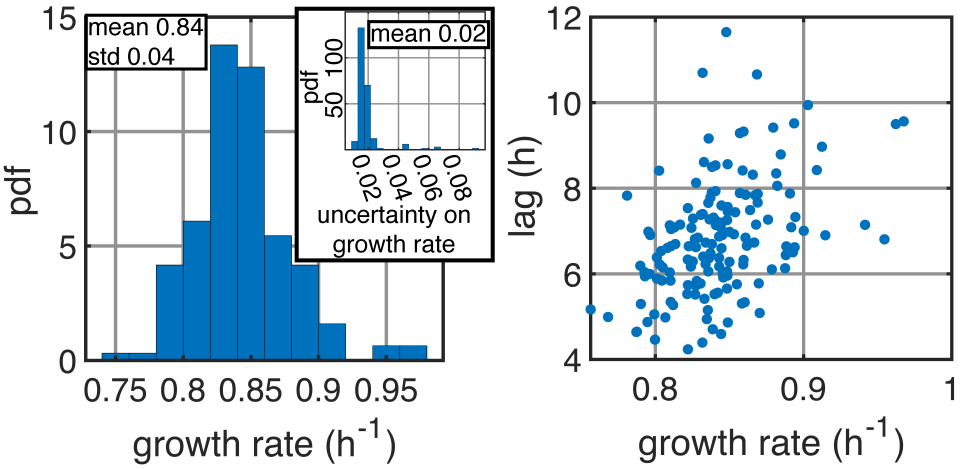

Figure 8. Statistical properties of experiment with inoculum 1. The left panel shows the probability density function of the growth rate in the experiment with inoculum 1 (Fig. 3B). The area of each bar is the number of observations in the bin. The growth rate is defined as the maximum value of the derivative of the growth curves (Fig. 1C). The estimation of the maximal value of the derivative for a droplet is given with its SD (shaded area around the derivative Fig. 1C). The inset reports the histogram of the SDs. The mean value of the SDs is taken as the typical error of the growth rate $\triangle\lambda = 0.02$. The right panel shows that the lag time and the growth rate of each droplets is weakly correlated across (Pearson coefficient of 0.43). Every points correspond to one droplet of the experiment.

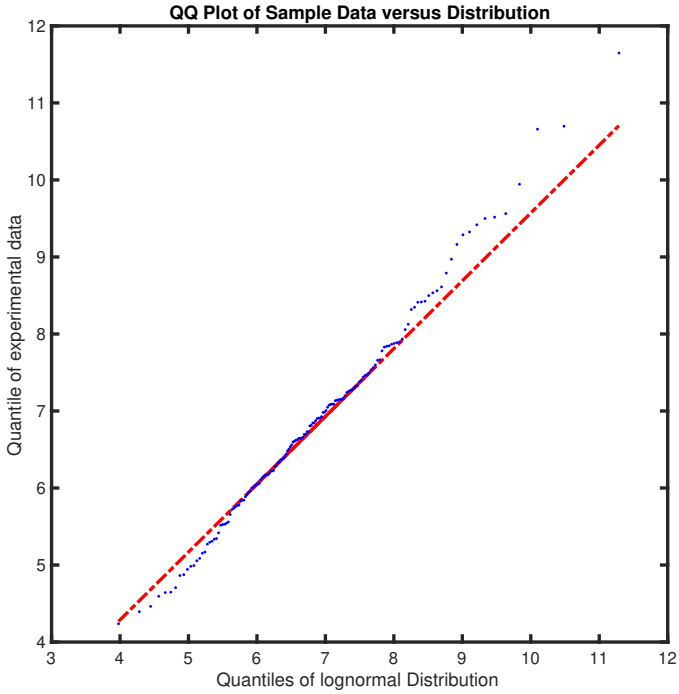

Figure 9. Quantile to quantile plot (qqplot) of cell-lag time To determine the distribution that underpin the cell lag time (θ) Fig 2B we plot the quantile of a log-normam distribution versus the quantile of the distribution of the experimental measurements. The resulting quantile to quantile plot is well fitted by a line of slope 1 indicating that the experimental values follows a log-normal distribution.

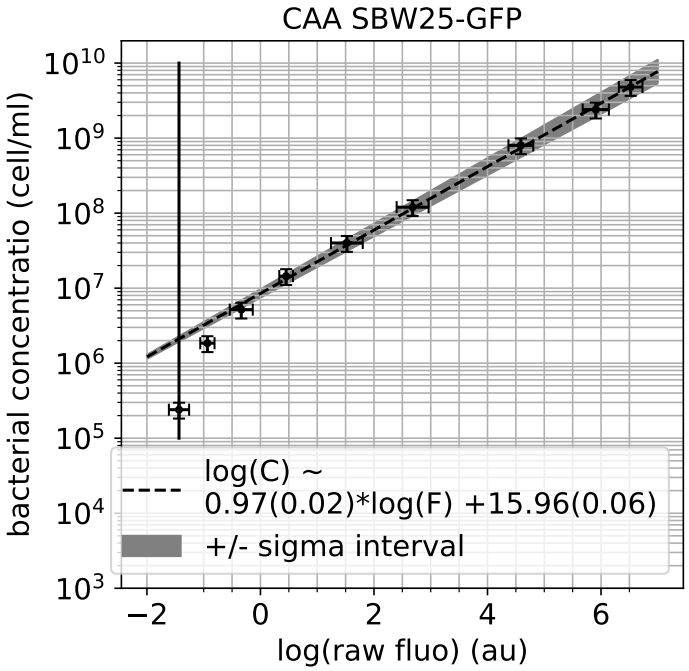

Figure 10. Calibration curve of the voltage given by the fluorescent detector in the channel GFP, called "raw fluo", to a bacterial concentration in cell·ml^{-1}. The dashed line depicts a linear fit on the experimental points. Its equation is given in the inset with uncertainty. The vertical line depict the value of raw fluorescence for pure CAA medium (the blank).

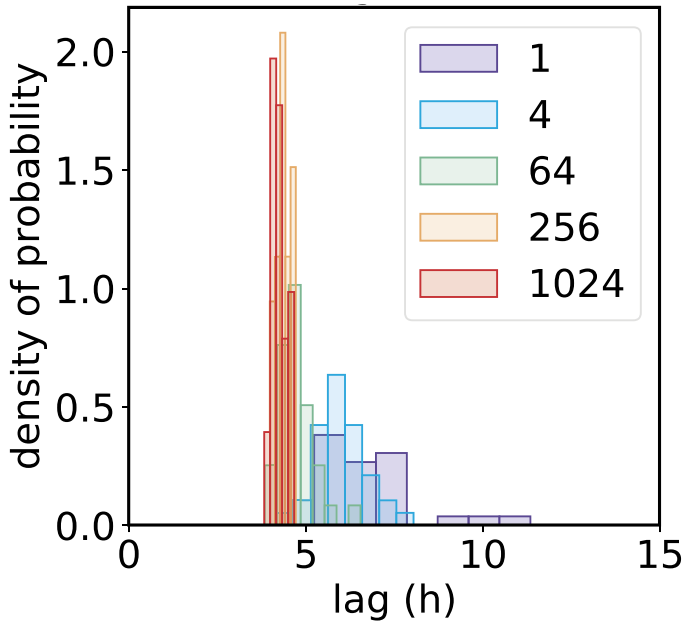

Figure 11. Histograms of lag time for a range of inoculum. It is the same data than Fig. 3A but showing the full histograms instead of points. Measured population lag time for a given inoculum are arranged in histograms. The colors of the histograms displayed in legend indicate their corresponding inoculum (in cells/drop) .

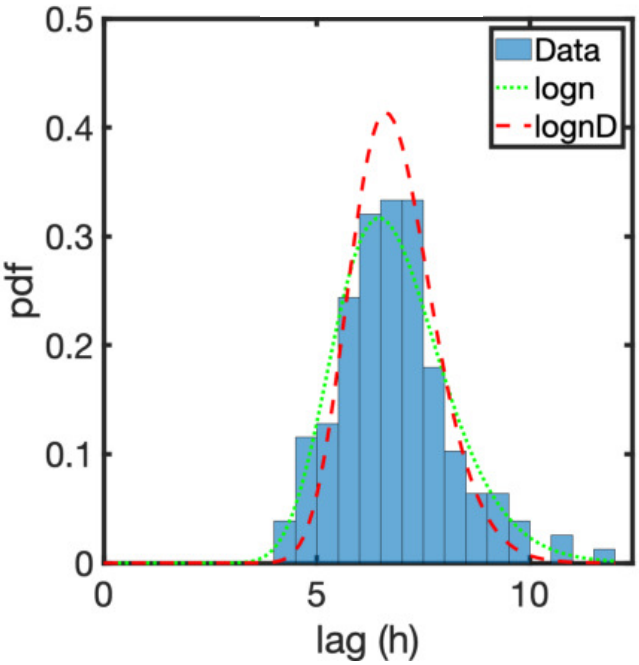

Figure 12. Probability density function estimates of the log-normal distribution of lag time. It the same data than Fig. 3B but in another way of representation often used. The "logn" in green dotted line is the fit on the data (blue bars). The "lognD" is the corrected log-normal distribution of lag after deconvolution of the gaussian noice.

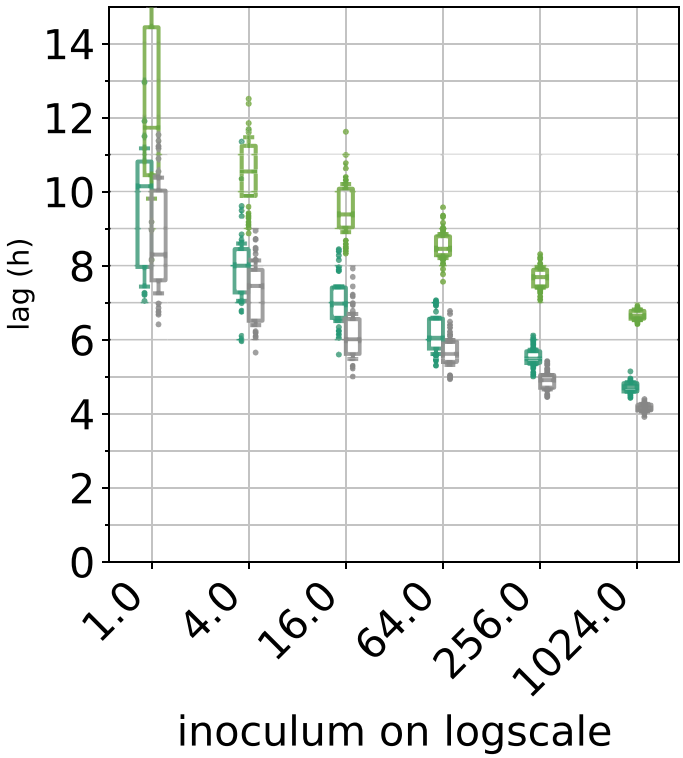

Figure 13. Inoculum effect on the lag for the strain *P. fluo* SBW25 PvdS229. This strain does not produce pyoverdine due to a mutation in *pvdS* a gene that encodes the extracytoplasmic family sigma factor *PvdS* and which directs expression of the pyoverdin biosynthetic genes (Cunliffe et al. 1995). The pyoverdine is an iron chelator the allows the pseudomonads to forage iron in their environment. The presence of the inoculum effect despite the non-production of pyoverdin indicate that this metabolite does not play a role in the coordinated exit of the lag phase. The three colors correspond to three independent experiments in fresh CAA. Data are represented in boxplot for every inoculum (with a little jitter for a better visualisation)

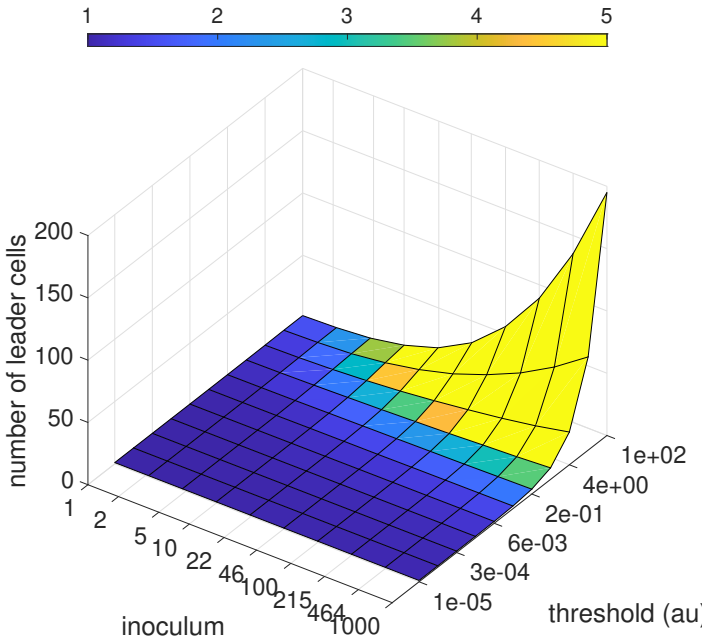

Figure 14. Lag time of the population as a function of the inoculum size and the growth activator threshold. The x-axis shows the dependence of the lag time of the populations as a function of the inoculum size (in number of cells). The y-axis shows the dependence of the lag time of the populations as a function of the threshold of the growth activator (in arbitrary unit). The z-axis and the color-map depict the lag time of the population (in hours). The threshold reported on the y-axis delimit the end of lag phase for the cells that do not multiply yet in the population. The higher it is the higher the growth activator concentration needs to be in order to end the lag phase of the still non-dividing cells. We see that the higher is the threshold, the higher is the population lag time, and the weaker is the dependence of the lag time with the inoculum size. Our experimental observations correspond to the case where the threshold is small which denote a strong dependence on the inoculum size.

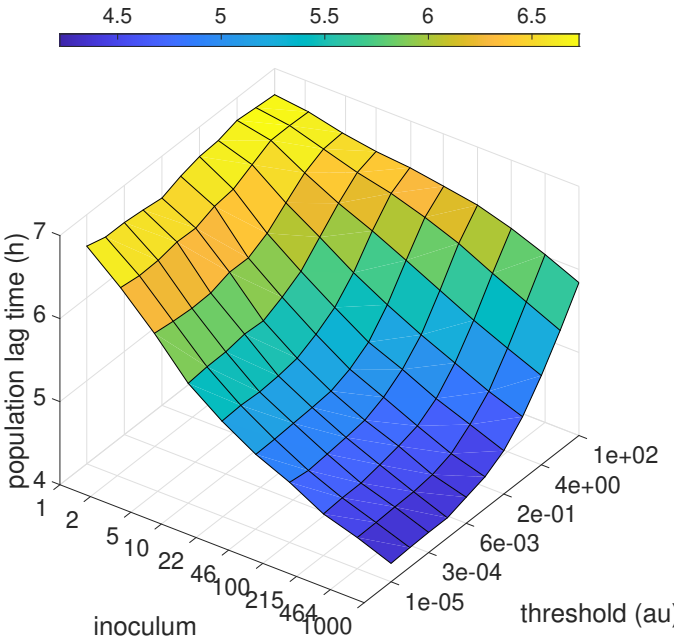

Figure 15. Number of cells that produce growth activator before the triggering event ending the lag phase. The x-axis shows the dependence of the lag time of the populations as a function of the inoculum size (in number of cells). The y-axis shows the dependence of the lag time of the populations as a function of the threshold of the growth activator (in arbitrary unit). The z-axis and the color-map depict the number of cells that exit their lag phase before that the growth activator exceed the given threshold value (in hours). The color-map is saturated for a value of 5 cells (yellow) but eventually the number of leader cells go above 5 in a certain range of the plot. In a large range of inoculum size and growth activator's threshold only one cell has the time to exit its lag phase before the triggering event. This is visible by the large region of the surface that remains dark blue. For values of threshold where the surface is yellow the dependence of the lag time with inoculum size shown Fig 14 does not match our experimental observation (strong dependence of lag time with inoculum size).

References

1. J.-C. Augustin, A. Agne‘, A. Brouillaud-Delattre, L. Rosso, and V. Carlier. Significance of Inoculum Size in the Lag Time of Listeria monocytogenes. Technical Report 4, 2000.

2. N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, sep 2004.

3. R. L. Bertranda. Lag phase is a dynamic, organized, adaptive, and evolvable period that prepares bacteria for cell division. *Journal of Bacteriology*, 201(7), apr 2019.

4. L. G. Blackwood. Normality transformations for environmental data from compound normal-lognormal distributions. *Environmental Monitoring and Assessment: An International Journal Devoted to Progress in the Use of*

Monitoring Data in Assessing Environmental Risks to Man and the Environment, 35(1):55–75, 1995.

5. S. Dagley, E. A. Dawes, and G. A. Morrison. Factors influencing the early phases of growth of aerobacter aerogenes. *Microbiology*, 4(3):437–447, 1950.

6. G. Doulcier. Dropsignal - millifluidic droplet trains analysis, feb 2019.

7. P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media, 2013.

8. M. Halmann and J. Mager. An endogenously produced substance essential for growth initiation of pasteurella tularensis. *Microbiology*, 49(3):461–468, 1967.

9. T. Julou, L. Zweifel, D. Blank, A. Fiori, and E. van Nimwegen. Subpopulations of sensorless bacteria drive fitness in fluctuating environments. *PLoS biology*, 18(12):e3000952, 2020.

10. A. S. Kaprelyants and D. B. Kell. Do bacteria need to communicate with each other for growth? *Trends in Microbiology*, 4(6):237–242, jun 1996.

11. C. E. Lankford, J. R. Walker, J. B. Reeves, N. H. Nabbut, B. R. Byers, R. J. Jones4, and R. J. Jones. Inoculum-Dependent Division Lag of Bacillus Cultures and Its Relation to an Endogenous Factor(s) ("Schizokinen"). Technical report, 1966.

12. B. R. M Lodge and C. N. Hinshelwood. Physicochemical aspects of bacterial growth. Part IX. The lag phase of Bact. lactis aerogenes. *Journal of the Chemical Society*, 51:213–219, 1943.

13. J. Monod. The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394, oct 1949.

14. S. Moreno-Gámez, D. J. Kiviet, C. Vulin, S. Schlegel, K. Schlegel, G. S. van Doorn, and M. Ackermann. Wide lag time distributions break a trade-off between reproduction and survival in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 117(31):18729–18736, aug 2020.

15. M. Müller. Ueber den Einfluss von Fiebertemperaturen auf die Wachsthumsgeschwindigkeit und die Virulenz des Typhus-bacillus. *Zeitschrift für Hygiene und Infektionskrankheiten*, 20:245–280, 1895.

16. W. J. Penfold. On the nature of bacterial lag. *Journal of Hygiene*, 14(2):215–241, 1914.

17. F. Pérez-Rodríguez. *Development and application of predictive microbiology models in foods*, chapter 18, pages 321–362. John Wiley & Sons, Ltd, 2014.

18. O. Rahn. Ueber den einfluss der stoffwechselprodukte auf das wachstum der bakterien. *Centralbl. f. Bakteriol. u. Parasitenk.*, 16:417–429, 1906.

19. M. D. Rolfe, C. J. Rice, S. Lucchini, C. Pin, A. Thompson, A. D. S. Cameron, M. Alston, M. F. Stringer, R. P. Betts, J. Baranyi, M. W. Peck, and J. C. D. Hinton. Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. *Journal of Bacteriology*, 194(3):686–701, 2012.

20. M. W. Silby, A. M. Cerdeño-Tárraga, G. S. Vernikos, S. R. Giddens, R. W. Jackson, G. M. Preston, X. X. Zhang, C. D. Moon, S. M. Gehrig, S. A. Godfrey, C. G. Knight, J. G. Malone, Z. Robinson, A. J. Spiers, S. Harris, G. L. Challis, A. M. Yaxley, D. Harris, K. Seeger, L. Murphy, S. Rutter, R. Squares, M. A. Quail, E. Saunders, K. Mavromatis, T. S. Brettin, S. D. Bentley, J. Hothersall, E. Stephens, C. M. Thomas, J. Parkhill, S. B. Levy, P. B. Rainey, and N. R. Thomson. Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. *Genome Biology*, 10(5):1–16, may 2009.

21. P. S. Swain, K. Stevenson, A. Leary, L. F. Montano-Gutierrez, I. B. Clark, J. Vogel, and T. Pilizota. Inferring time derivatives including cell growth rates using Gaussian processes. *Nature Communications*, 7(May):1–8, 2016.

22. I. Swinnen, K. Bernaerts, E. Dens, A. Geeraerd, and J. Van Impe. Predictive modelling of the microbial lag phase: a review. *International Journal of Food Microbiology*, 94(2):137–159, 2004.

23. T. V. Votyakova, A. S. Kaprelyants, and D. B. Kell. Influence of viable cells on the resuscitation of dormant cells in ¡i¿micrococcus luteus¡/i¿ cultures held in an extended stationary phase: the population effect. *Applied and Environmental Microbiology*, 60(9):3284–3291, 1994.

24. X.-X. Zhang and P. B. Rainey. Exploring the sociobiology of pyoverdin-producing Pseudomonas. *Evolution*, 67(11):1–21, nov 2013.

25. E. Şimşek and M. Kim. Power-law tail in lag time distribution underlies bacterial persistence. *Proceedings of the National Academy of Sciences of the United States of America*, 116(36):17635–17640, sep 2019.