Subpopulations of sensorless bacteria drive fitness in fluctuating environments

Thomas Julou[1,2*], Diana Blank[1,2], Athos Fiori[1,2], Erik van Nimwegen[1,2*]

1 Biozentrum, University of Basel, Basel, Switzerland
2 Swiss Institute of Bioinformatics, Basel, Switzerland

* thomas.julou@unibas.ch * erik.vannimwegen@unibas.ch

Abstract

Populations of bacteria often undergo a lag in growth when switching conditions. Because growth lags can be large compared to typical doubling times, variations in growth lag are an important but often overlooked component of bacterial fitness in fluctuating environments. We here explore how growth lag variation is determined for the archetypical switch from glucose to lactose as a carbon source in *E. coli*. First, we show that single-cell lags are bimodally distributed and controlled by a single-molecule trigger. That is, gene expression noise causes the population before the switch to divide into subpopulations with zero and nonzero *lac* operon expression. While 'sensorless' cells with zero pre-existing *lac* expression at the switch have long lags because they are unable to sense the lactose signal, any nonzero *lac* operon expression suffices to ensure a short lag. Second, we show that the growth lag at the population level depends crucially on the fraction of sensorless cells, and that this fraction in turn depends sensitively on the growth condition before the switch. Consequently, even small changes of basal expression affecting the fraction of sensorless cells can significantly affect population lags and fitness under switching conditions, and may thus be subject to significant natural selection. Indeed, we show that condition-dependent population lags vary across wild *E. coli* isolates. Since many sensory genes are naturally low expressed in conditions where their inducer is not present, bimodal responses due to subpopulations of sensorless cells may be a general mechanism inducing phenotypic heterogeneity and controlling population lags in switching environments. This mechanism also illustrates how gene expression noise can turn even simple sensory gene circuits into a bet-hedging module, and underlines the profound role of gene expression noise in regulatory responses.

Introduction

Most unicellular organisms live in variable environments that require them to adapt to, among other things, changes in available nutrients. Already in the early 1940s Monod observed that during diauxic growth, when bacteria switch from consuming one sugar to another, there is typically a lag period in growth [1]. Reasoning that this lag is somehow associated with the need for cells to 'adapt' to the alternative sugar, subsequent investigations into the origin of this lag eventually led to the discovery of gene regulation [2]. Monod also observed that the length of the lag can vary strongly, suggesting that the amount of time cells need to adapt depends on the type of the switch, but at the time it was not clear why certain switches would require more adaptation time than others.

It is currently widely believed that lag times depend on which of two types of general strategies cells employ to adapt to a particular environmental change. If, like in the *lac* operon studied by Monod, cells have dedicated machinery for sensing the nutrient and up-regulating the appropriate target genes in response, lags are generally expected to be short. Alternatively, in the absence of sensory machinery, cells can employ a bet hedging strategy where, through stochastic gene expression, small subsets of the population of cells are already pre-adapted to other environments [3]. In this case, long lags can result from the time required for the rare subset of pre-adapted cells to expand.

However, as it has become well appreciated that essentially all genes are subject to substantial stochastic fluctuations in expression, there may well be no clean dichotomy between sensing and stochastic bet hedging strategies [4]. That is, even when cells possess dedicated sensing and regulatory machinery for responding to a switch in nutrients, stochastic gene expression will cause phenotypes to vary across cells, and this may well lead to a significant variations in lags across single cells.

The archetypical example of a nutrient switch, which led to the discovery of gene regulation, involves exposing *Escherichia coli* bacteria that were growing on glucose to lactose. In response, the three genes of the *lac* operon will be expressed, including the lactose transporter LacY and the galactosidase LacZ, which both hydrolyses lactose and isomerizes it into allolactose (Fig. 1A). The expression of the *lac* operon is controlled by a dedicated transcription factor, LacI, which is inhibited by allolactose. In the absence of lactose the repressor LacI keeps the operon at low expression, whereas at high external lactose concentrations a positive feedback loop between import of lactose by lacY, production of allolactose by LacZ, and inhibition of LacI by allolactose, causes the operon to switch to a high induced state.

We recently reported a preliminary investigation into the distribution of single-cell lags in the induction dynamics of the *lac* operon and observed that lags are bimodally distributed [5]. Here we characterise the molecular mechanisms underlying this bimodality, the determinants of the fractions of the population in each of the modes, and the resulting impact on population lag and fitness when bacteria are switching between nutrients.

Using a combination of experiments that perturb the state of cells before the first induction and detailed analysis of transcriptional memory in lineages of single cells after induction, we show that cells with short and long lags correspond to cells with either nonzero or zero pre-existing *lac* expression at the time of the switch. That is, we find that the *lac* operon constitutes a single-molecule trigger where any nonzero expression of LacY/Z proteins suffices to ensure a quick response. In contrast, long lags occur because cells with zero *lac* expression cannot sense lactose and have to wait for a spontaneous stochastic burst of *lac* operon expression before they induce. We show that the population lag is determined by the fraction of '*sensorless*' cells with zero *lac* expression and that this fraction can sensitively depend on the precise growth conditions before the switch. Using simulations to infer population growth lags from single-cell lags we show that subtle changes in the fraction of sensorless cells can have substantial consequences for population lag and fitness. Diauxie experiments with both laboratory strains and a diverse set of wild *E. coli* isolate validate our predictions on population lags and show that wild *E. coli* strains also exhibit highly context-dependent lags in *lac* operon induction.

In summary, for the *lac* operon, a sub-population of sensorless cells causes a bimodal response to lactose and the population growth-lag is largely determined by the fraction of sensorless cells at the time of the switch, which in turn sensitively depends on the growth conditions before the switch. This raises the question as to what extent the occurrence of sensorless cells is a general mechanism for generating bimodal responses and determining population lag to an inducing signal. Using existing proteomic data we

show that a large fraction of sensory systems in *E. coli* are expressed at such low levels in many conditions that substantial fractions of sensorless cells are expected to generically occur for many sensory systems. Thus, population lags may well be controlled by fractions of sensorless cells for many environmental switches.

Results

In order to study the stochastic induction dynamics of the *lac* operon, we used a recently developed integrated setup that combines microfluidics, quantitative time-lapse microscopy, and automated image analysis [5]. We monitored growth and gene expression of single *E. coli* cells expressing a LacZ-GFPmut2 fusion from their native *lac* operon locus (Fig. 1A) while growing in Dual Input Mother Machine chips, switching the growth medium between glucose and lactose every 4 hours, starting from a glucose phase (Fig. 1B). As we already reported in our preliminary study based on a much smaller sample [5], GFP levels are below autofluorescence levels in all cells during the initial glucose phase and upon the switch to lactose all cells immediately enter growth arrest for a stochastic period of time, after which growth recommences approximately at the same time as LacZ-GFP expression is first detected (Fig. 1B). Notably, although we do not observe bistability in the sense that almost all cells eventually induce their *lac* operon, the distribution of lag times until induction is clearly bimodal (Fig. 1C; Fig. S2): 30% of cells induce their *lac* operon in less than 50 min while the other 70% take between 1 and 3 h. The bimodality of single-cell lags is not dependent on the abrupt nature of the switch, because an almost identical bimodal distribution was observed for experiments in which cells are exposed to a gradual transition from glucose to lactose over 40 min (Fig. S3). Another striking feature of this regulatory response is that, for virtually all cells, the resumption of growth occurs almost exactly when *lac* expression becomes detectable (Fig. 1B). In order to rigorously quantify the correlation between *lac* induction and growth lags we devised a Bayesian inference procedure to estimate the growth lag of single cells from their (noisy) elongation curves (see Methods). We observe that growth lags and induction lags are very tightly correlated (Pearson correlation coefficient $r^2 = 0.87$, Fig. 1D; Fig. S2), with the exception of rare cells ($\approx 6\%$) which did not restart growing despite having induced their *lac* operon. Consequently, the distribution of induction lags shapes the population growth response when cells are exposed to lactose.

The fact that the growth lags closely track the *lac* induction lags of individual cells, together with the fact that we did not observe any correlation between lags and basic physiological parameters such as the growth rate of the cells or their cell cycle stage at the switch (Fig. S4), strongly suggests that the *lac* induction dynamics is the main determinant of growth after the switch. These observations led us to hypothesize that the lag time of a cell may also be controlled by its pre-existing *lac* expression before the switch. To test whether the fractions of long and short lags can be altered using a perturbation that only affects *lac* expression before the switch, we added a minute amount of the artificial inducer IPTG (5 μM) to the glucose media before the switch. Although this amount of IPTG is far too small to cause induction of the *lac* operon (Fig. S5), it should inhibit some of the LacI proteins and thereby increase pre-existing *lac* expression. Indeed, we observe that this amount of IPTG is sufficient to increase the fraction of short lags by 2.4-fold (Fig. 1E). In contrast, when the LacI repressor is over-expressed approximately 5-fold from a low copy number plasmid, leading to a stronger repression of the *lac* operon, virtually all naive cells have long lags (Fig. 1E). Moreover, LacI overexpression can be compensated by IPTG supplementation, indicating that these perturbations are indeed counteracting each other (Fig. 1E). Since LacI only targets the *lac* operon it seems highly unlikely that these perturbations in

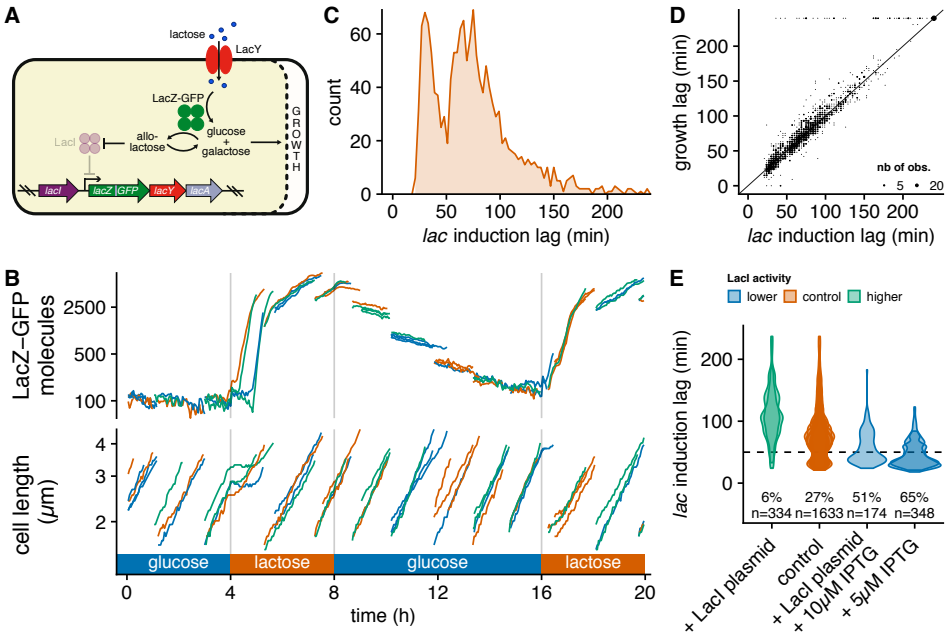

Fig 1. *lac* operon induction lags are bimodally distributed and controlled by LacI activity before the switch.
A. Summary of the *lac* gene circuit. We use a LacZ-GFPmut2 translational fusion integrated at the native genetic locus to monitor *lac* expression in single cells. **B.** Time courses of cell length and number of LacZ-GFP molecules (on log scales) for a subset of cells from a typical experiment where cells are grown in a Dual Input Mother Machine and exposed to two consecutive 4 hour lactose episodes interspersed by a glucose period (which is 8 h in this experiment); for clarity, only the 4 cells near the closed end in one representative growth channel are shown using random colours to distinguish cells, and LacZ-GFP levels are offset by 100 molecules. **C.** Histogram of single-cell induction lags for the *lac* operon at the first lactose exposure ('short' lags under 50 min, 'long' lags above 50 min) *lac* induction lags were defined as the delay after the switch until cells increase their LacZ-GFP by 200 molecules and were estimated from time series of LacZ-GFP expression (shown in B) for 1633 cells in 9 independent replicates (Fig. S1, Fig. S2). **D.** A scatter of growth lags versus *lac* induction lags in individual cells shows that these two lags are tightly correlated (the solid line is a guide for the eye with $y = x$). **E.** Violin plots of the distribution of single-cell induction lags for populations of cells whose LacI activity during growth in glucose was perturbed either by overexpression from a plasmid or by titration with sub-induction concentrations of IPTG. The fraction of cells with short lags is reported for each treatment; each contour corresponds to an independent replicate (with ≈ 150 cells) and the dotted line corresponds to 50 min separating long and short lags.

LacI activity alter anything in the state of the cells besides the strength of *lac* operon repression. These results thus strongly support that single-cell lag times are controlled by pre-existing *lac* operon expression. The fact that we observe a bimodal distribution of lag times thus suggests that there exists a threshold of pre-existing *lac* expression that controls whether a cell has a short or long induction lag. In this interpretation, this threshold expression level would correspond to the critical level of pre-existing *lac* expression required for the positive feedback of the *lac* system to quickly drive cells to the induced state, whereas cells with pre-existing *lac* expression below this threshold would have to wait until a stochastic burst of *lac* expression would drive the system over this threshold.

We next set out to quantify this critical *lac* expression threshold value. Unfortunately, in our system the fluorescence of pre-existing LacZ-GFP molecules is small compared to autofluorescence levels which fluctuates significantly from cell to cell ($\pm$ 20 GFP molecules; Fig. 2A). In addition, small uncontrolled day-t-day variations in illumination contribute additional uncertainty on LacZ-GFP measurements. Altogether this masks any correlation between pre-existing LacZ-GFP levels and induction lags of individual cells (Fig. S4). However, by partitioning cells at the switch into those that exhibited either long lags or short lags, we find that cells with short lags have on average $\approx$ 5 LacZ-GFP molecules more than cells with long lags (p-value = 1e-5; Fig. 2A), indicating that the presence of only a few molecules may be enough to ensure short lags.

In addition, we were able to perform an independent and more precise quantification of the pre-existing *lac* expression threshold using systematic dilution experiments. During the period of growth in glucose that intercedes the two lactose periods, the *lac* operon is fully repressed and, because LacZ-GFP is sufficiently stable that protein decay can be neglected [6], LacZ-GFP levels divide in half at each cell division (Fig. 1B). For any cell at the second switch to lactose, we can thus estimate its LacZ-GFP level from the lacZ-GFP level of its ancestor at the end of the first lactose phase, and the number of cell divisions in the intervening period. By modulating the length of the interceding glucose phase, we can thus systematically vary the distribution of remaining *lac* expression at the second switch to lactose, and measure the fractions of long and short lags as a function of the level of remaining *lac* expression. We observe that even for surprisingly long lags of up to 12 h, virtually all lags are short, and that it takes ca. 20 h (corresponding to circa 16 divisions) to relax back to the distribution of induction lags observed in naive cells (Fig. 2B). Moreover, there is a very strong correlation between the observed lag times and the estimated number of remaining LacZ-GFP molecules (Fig. S6). Virtually all cells that are estimated to have at least 10 remaining LacZ-GFP molecules exhibit short lags, and only cells with one or fewer remaining LacZ-GFP molecules exhibit the same fraction of long lags as naive cells (Fig. 2C). Together, these dilution experiments and direct fluorescence measurements (Fig. 2A) indicate that the critical threshold of pre-existing *lac* expression is on the order of one molecule.

The following picture of a single-molecule trigger emerges: when exposed to lactose, those cells that have at least one molecule of both LacY and LacZ present are able to respond fast, whereas all other cells cannot sense the lactose and have to wait for the next stochastic burst of expression (Fig. 3A). A straightforward and testable prediction follows: the fraction of cells with long induction lags ($\approx$70%) should match the fraction of cells with zero Lac protein in absence of inducer. This prediction is confirmed by single-molecule experiments which observed that $\approx$65% of cells have no LacZ [7], and $\approx$50% no LacY [8].

Another implication of the single-molecule trigger mechanism is that the fractions of short and long lags depend sensitively on the strength of the *lac* operon repression before the switch, i.e. on what fraction of cells have precisely zero expression. This suggests that even subtle changes in the strength of *lac* operon repression due to

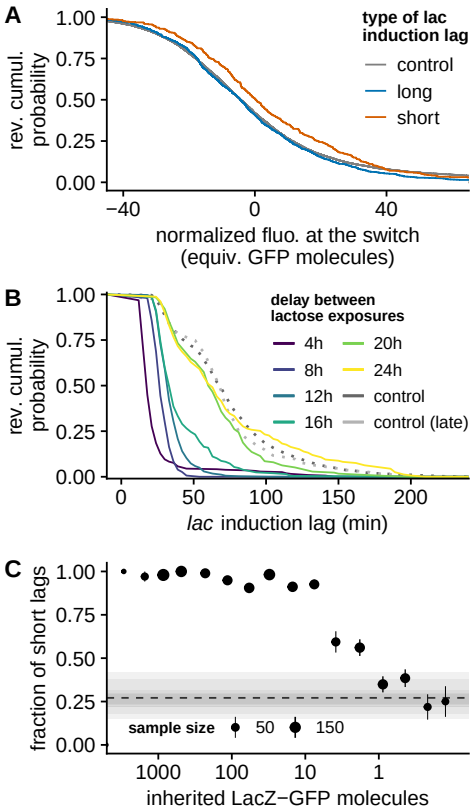

Fig 2. The critical *lac* expression threshold for inducing the *lac* operon is on the order of a single molecule.
A. Distributions of initial fluorescence in cells with short (orange) and long(blue) *lac* induction lags; the grey line is a control showing the distribution of autofluorescence in a wild-type strain without GFP (fluorescence is centered per experiment, see supplementary methods). Note that the shift in fluorescence between cells with short and long lags corresponds to ~ 5 LacZ-GFP molecules. **B.** Reverse cumulative distributions of *lac* induction lags for cells that first grew in lactose and fully induced their *lac* operon, and then grew in glucose for different lengths of time (shown in colours; 1 or 2 replicates per condition); the distribution for naive cells is shown as a reference (dotted dark gray, corresponding to the histogram of Fig. 1C) along with a control where cells are first exposed to lactose after 26 h in the microfluidic chip (light gray). **C.** Fraction of short lags (<50 min) for cells pre-exposed to lactose as a function of the estimated number of inherited LacZ-GFP molecules. The dashed horizontal line shows the overall fraction of short lags in naive cells as a reference and semi-transparent gray rectangles show the mean $\pm$ s.e. for each replicate.

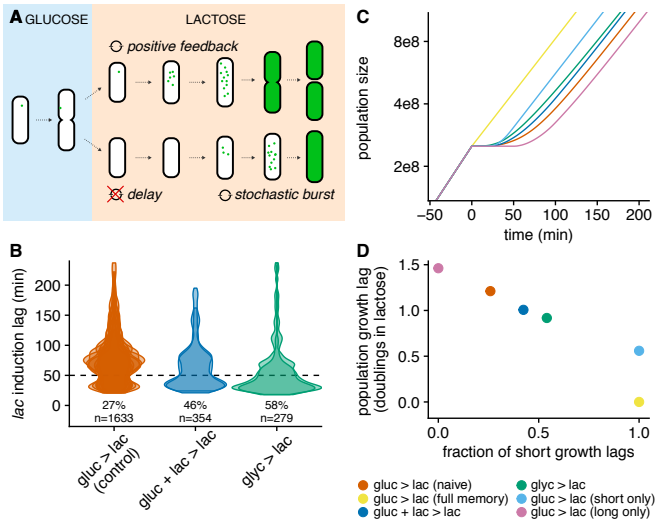

Fig 3. The response of single cells to lactose is context-dependent and shapes the population fitness after the switch.
A. Cartoon depicting how the *lac* induction of a given cell is determined by the stochastic expression in the repressed state, and hence depends on the conditions before the switch. Green dots depict Lac proteins. **B**. Violin plots of the single-cell lag distributions for switches to lactose from growth conditions in which the *lac* operon is slightly less repressed than in pure glucose, i.e. 0.4% glycerol (blue) and 0.2% glucose + 0.2% lactose (green). Each contour corresponds to a separate replicate with ≈150 cells); the fraction of short lags (<50 min, indicated by the dotted line) is indicated below the plots. **C**. Simulated population growth curves for different distributions of single-cell lags, as observed for different switches (different colors). The simulations assume deterministic exponential growth of each cell before and after the switch, and complete growth arrest during the lag. The pink and light blue curves show the population growth that would be obtained with only short and long lags, respectively. **D**. Population growth lags as inferred from the simulated population growth (vertical axis) as a function of the fraction of short lags (horizontal axis). The lag is measured as a delay compared to a population with no arrest (yellow), and reported in number of doublings in lactose (i.e. the number of divisions lost due to growth arrest).

changes in growth conditions before the switch, may substantially affect the distribution of single-cell lags. To investigate the context-dependence of the distribution of single-cell lags we used different nutrients before the switch that we reasoned would alter the strength of repression. In particular, besides being repressed by LacI, the *lac* operon is also positively regulated by CRP [9], and the activity of CRP can be increased by, for example, replacing glucose with glycerol as a carbon source [10]. Since the strength of the *lac* repression should thus be weakened when growing in glycerol, the fraction of short lags should increase, and this is indeed what we observe, i.e. the fraction of naive cells inducing fast increases 2.1-fold to 58% (Fig. 3B). In addition, we also measured single-cell lags when growing cells on a mixture of glucose and lactose before the switch, i.e. as in diauxie experiments. Although *lac* expression remains low on such a mixture [11], whenever the *lac* expression is nonzero, the presence of lactose will cause some of it to be transported into the cells and metabolized into allolactose, inhibiting some fraction of the LacI repressors. Thus, compared to growth on glucose, we expect the repression of the *lac* repression to be slightly weakened on a mixture of glucose and lactose. Note that the direct inhibition of LacY permeases due to the ongoing glucose import ("inducer exclusion") is not compplete and should thus not prevent this effect from occuring [9]. Indeed, we find that under these conditions the fraction of short lags is increased 1.8-fold to 48% (Fig. 3B). These results confirm that, because induction depends on a single-molecule trigger, the fraction of cells with short lags is highly sensitive to the precise growth conditions before the switch.

We wondered to what extent this sensitive dependence of the single-cell lags would translate into substantial differences in population lags, and thus, fitness. In order to infer population growth lags from distributions of single-cell growth lags, we simulated population growth curves during a switch to lactose (Fig. 3C). By comparing these to the growth curve of a population where all cells grow immediately at their maximum growth rate on lactose (yellow line in Fig. 3C) we can calculate a population lag for each of the conditions described above (Fig. 3D). In addition, we calculate population lags for hypothetical populations where cells would have only short lags or only long lags. Remarkably, while a population where all cells switch fast will have a lag on the order of 0.5 division time (due to the time necessary to initiate the positive feedback), a population of naive cells will be retarded by ~ 1.2 divisions compared to a population without lag, and the population lag can be as long as 1.5 division times if no cells have Lac protein at the switch. Thus, we find that, by affecting the fraction of sensorless cells with zero *lac* expression, even subtle changes in growth conditions before the switch can have a strong impact on population lag.

Our measurements of the single-cell lags suggest that population lags, and thus fitness in response to a switch of the available nutrients to lactose, can sensitively depend on the growth conditions before the switch. To test this in real populations, and to investigate to what extent this behavior extends beyond lab strains, we performed classic diauxie experiments with both lab strains and a diverse set of natural *E. coli* isolates [12, 13]. As is well known, when *E. coli* is grown on a combination of glucose and lactose, cells first consume the glucose and naturally switch to consuming lactose when the glucose runs out. To test the context-dependence of population lags, we performed diauxie experiments with two different lactose concentrations.

Based on our previous observations with experiments in the DIMM (Fig. 3B), we reasoned that higher external concentrations of lactose would lead to more lactose being imported into the cells during growth on glucose, thereby weakening the *lac* repression more. Thus, we expect shorter population lags in the higher lactose concentration. To calculate growth lags, we compared the observed diauxic growth curves with population growth in the same conditions supplemented with large amounts of IPTG (200 μM), such that cells express their *lac* operon during the whole experiment and can readily

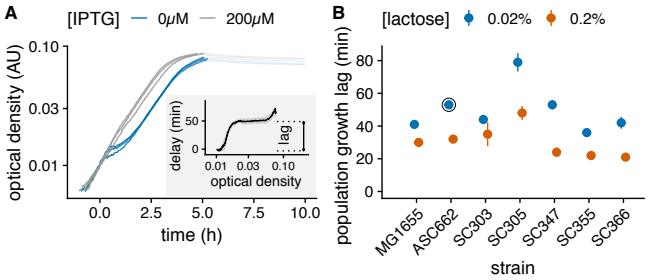

Fig 4. Single-molecule triggering is observed across a diverse set of *E. coli* strains.
A. Estimation of the population growth lag from diauxie experiments. Population growth curves of strain ASC662 during diauxie experiments with 0.02% lactose and 0.005% glucose (blue) and controls where the *lac* operon is expressed throughout with supplemented ITPG (gray). Each line corresponds to one replicate. The delay between the blue growth curve and gray control features a long plateau which corresponds to the population lag (inset). **B**. Population lag in diauxie experiments with medium (0.02%, blue) and high (0.2%, orange) lactose concentrations for both lab strains and a diverse set of natural *E. coli* isolates. Points and error bars correspond to mean and standard error over at least 3 replicates with 0.005% glucose; the circled dot correspond to the growth curves shown in A. Note that the lag is reduced in all strains when more lactose is provided.

metabolise lactose when they exhaust glucose (Fig. 4A). Across all strains, we find that population lags are indeed shorter for higher lactose concentrations, even though the absolute value of the lags and the size of the difference between the two lactose concentrations varies substantially across strains (Fig. 4B, Fig. S7).

Although, based on our single-cell experiments, we interpret these changes in population lags to result from changes in the fractions of cells with nonzero *lac* expression at the time that the glucose is exhausted, a higher lactose concentration may also reduce lags by increasing the rate of lactose import in cells with nonzero *lac* operon expression, i.e. shortening the time of the induction process itself. To disentangle these effects, we performed experiments in which cells where grown on a mixture of glucose and lactose (at a given concentration), and where then transferred to media with only lactose at varying concentrations. Despite the less precise measurement of the population lag with this protocol than in diauxie experiments, we observe that the growth lag after the switch is impacted both by the lactose concentrations before and after the switch, and that these two parameters have effects of similar magnitude on the population lag (Fig. S8).

Our experiments have shown that population lag times, and thus fitness under an environmental switch to lactose, are to a large extent determined by the frequency of sensorless cells with zero *lac* operon expression. Moreover, such lags are observed across a diverse set of natural *E. coli* isolates. Given that even a small relaxation of the strength of *lac* operon repression can significantly reduce population lags, it is remarkable that natural selection has apparently not acted to avoid the occurrence of sensorless cells in any of the strains we tested here. This suggests that another selection pressure acts to ensure that a substantial fraction of cells have zero *lac* expression in the absence of lactose. One could speculate that perhaps LacY may otherwise transport toxic substances into the cell, or waste energy on transporting useless substances [14, 15]. In any case, it appears *E. coli* must trade-off short lags under a switch to lactose, against unwanted side effects of even very low *lac* operon expression when lactose is absent. The strength of *lac* operon repression observed in different strains may reflect

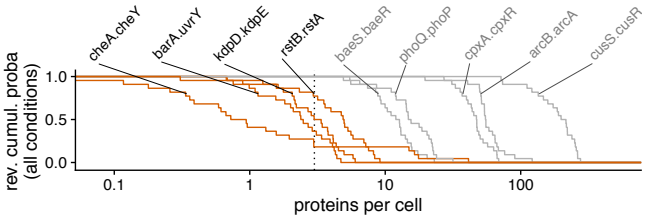

Fig 5. Sensorless cells are expected to occur for many two-component systems in *E. coli*.
Distributions of expression levels of sensor kinases across 28 conditions for 9 of the 28 annotated two-components systems in *E. coli*. The vertical dashed lines correspond to an average of 3 proteins per cell which is the threshold below which more than 5% of the cells is expected to be sensorless. Note that 4 out of 9 two-components systems (highlighted in orange) feature sensorless cells in 6 conditions or more.

this trade-off.

These observations raise the question as to whether other sensory systems in *E. coli* are also repressed in the absence of their signal to the point that substantial fractions of sensorless cells occur that will be unable to respond when the signal appears. We are aware of only one other system that has been shown to exhibit very similar behavior and that is the regulatory response protecting *E. coli* against the toxic and mutagenic effects of DNA alkylation damage. In this system a single-molecule trigger leads to fast or slow expression of Ada proteins, which transiently creates two subpopulations with different phenotypes, in this case mutation rates [16]. In order to address how ubiquitous single-molecule triggers are in bacteria, we focused on two-component systems. Two-components systems all feature a kinase sensor located at the cell envelope whose activation by environmental signals lead to phosphorylation of a cognate response regulator, typically a transcription factor. In many cases the activation of the response regulator upregulates the expression of the kinase, thereby implementing a positive feedback. Thus, two-components systems with very low basal expression of the kinase could exhibit single-molecule trigger dynamics. But even without single-molecule trigger behavior, sufficiently low basal expression of the kinase will cause sensorless cells to appear and they will be unable to respond to the corresponding signal.

Using quantitative proteomics data obtained for *E. coli* in 28 conditions [17], we could query the expression levels for 9 kinase-TF pairs and for another 12 TFs (Fig. S9), out of 27 two-components systems annotated on `ecocyc.org`. Remarkably for 4 of the 9 pairs, cells carry less than 3 copies of the kinase on average in many conditions (Fig. 5). If we conservatively assume Poisson expression fluctuations across cells, then proteins with less than 3 copies on average imply a fraction of at least $e^{-3} \approx 0.05$ sensorless cells. It is also remarkable that for all 9 pairs, the kinase is expressed at a lower level than the TF (typically 200-1000 fold less); since 6 out of the 12 additional TFs are expressed at 10 copies or less per cell (Fig. S9), we expect several of these two-components systems to have sensorless cells as well. This limited exploration of sensory gene circuits in *E. coli* illustrates that the occurrence of sensorless cells and single-molecule triggers may in fact be common in bacteria.

Discussion

In this study of the *lac* operon induction by lactose in single cells, we uncovered that it is governed by a single-molecule trigger: while cells that have at least one of each Lac protein expressed at the moment of the switch will respond readily, cells lacking Lac

proteins are unable to sense lactose and have to wait for the next stochastic burst of *lac* expression, which can delay the response for as long as several divisions. This mechanism, where a positive feedback involved in the response is combined with a very low basal expression level of its sensor such that a fraction of cells is effectively sensorless, gives rise to transiently coexisting induced and uninduced subpopulations. Importantly, despite involving a positive feedback, this molecular mechanism underlying phenotypic variability does not require bistability of the gene circuit, as highlighted by the fact that all cells eventually switch on.

The *lac* operon is one of the most widely studied regulatory systems, which raises the question why this single-molecule trigger has not been observed previously. First, in most previous quantitative studies of the *lac* operon, artificial inducers such as IPTG and TMG were used in order to measure *lac* expression when cells are in a steady state of growth that is not perturbed by the induction [9,18]. In these conditions, true bistability has been reported, with a critical threshold in the order of hundreds of molecules [8]. However, in our experiments lactose is both an inducer and a nutrient, which can only be metabolized when the *lac* operon is sufficiently expressed. Therefore, gene expression and growth are directly coupled, and all cells induce the *lac* operon from a state of growth arrest, which is very different from the previous experiments with artificial inducers. We note two previous studies of the *lac* induction dynamics that are most closely related to our current study. Lambert and Kussell studied *lac* induction in small populations of *E. coli* cells in a microfluidic setup where bacteria are directly exposed to switching growth media [6]. However, in contrast to our study where growth and gene expression is followed at the single-cell level, growth is only tracked for the entire population. This study established the long-lived memory of *lac* operon induction, i.e. that a population of cells is able to respond rapidly as long as Lac proteins have not been diluted below a threshold. However, the setup was not calibrated to measure *lac* expression in number of proteins so that this threshold could not be quantified, and the bimodality of induction lags could not be uncovered because of the lack of single-cell measurements. In a second study *lac* operon induction was studied in single cells by growing cells in microfluidic devices in which media were dispensed by diffusion through a gel, and growth lag upon a switch to lactose was measured in single cells [19]. Because no growth arrest was observed for the large majority of cells in this study, lags were defined as the delay until growth-rate fully recovered, and no bimodality of this recovery time was observed. We hypothesize that, due to the different method of nutrient switching, cells may have been exposed to remaining low levels of glucose for sufficiently long to avoid growth-arrest in this study.

Thus, the fact that our DIMM setup allows not only precise temporal control of the growth media that cells experience, but also accurate quantitative tracking of growth and gene expression in lineages of single cells, were essential for uncovering that a single-molecule trigger underlies the bimodal behavior of *lac* operon induction, and that it also constitutes the molecular basis of cellular memory after exposure to lactose. In addition, we showed that the population lag is determined by the fractions of slow and fast responding cells and these fractions are highly sensitive to growth conditions before the switch, with even very subtle relaxation of the *lac* repression increasing the fraction of fast responding cells.

The occurrence of a subpopulation of sensorless cells can have major consequences for fitness under switching environment, as they may cause substantial population lags, delaying growth by more than an entire division. Moreover, substantial population growth lags as a result of a subpopulation of sensorless cells may be a fairly common feature of bacterial sensory systems as a survey of two-component systems shows that, for a large fraction of them, sensors are so low expressed that a fraction of cells with zero expression is to be expected. This observation, in combination with the observation

that the fraction of sensorless cells is highly sensitive to the strength of repression of the system, raises the question as to how natural selection has acted on the expression levels of sensory systems. Indeed, as our experiments with supplementing a tiny amount of IPTG demonstrate, even very modest derepression of the system would dramatically lower the fraction of sensorless cells, and it seems that bacteria could easily evolve to do so.

However we observed that substantial population lags occurs in all strains of a diverse set of natural *E. coli* isolates. This suggest that, quite generally, there must be a cost to the expression of a sensory system in the absence of its cognate signal. These costs may derive from toxicity (e.g. due to biochemical or transport activity of the sensor) [14, 16] or from energetic costs associated with the sensing [15]. It is possible that the tuning of the sensor's expression to ensure the co-occurrence of a small subpopulation in which the sensor is present with a majority of sensorless cells, provides a good bet hedging strategy for limiting population lags under a sudden appearance of the signal, while at the same time relieving the majority of the population from the costs associated with expressing the sensor. Finally, since the fraction of sensorless cells is highly sensitive to the precise growth conditions, one would expect that natural selection should be able to tune the sensor's expression in a condition-dependent manner, so as to effectively anticipate the appearance of the signal in some conditions more than in others. For example, the *mal* operons have higher basal expression on lactose than on other sugars, leading to shorter population lag after a switch to lactose [20]. More generally, for the *lac* system, because CRP activity increases whenever carbon flux is limiting [21], cells anticipate a switch to lactose more readily in such conditions than in conditions where carbon flux is not limiting and CRP activity is low.

Materials and Methods

1 Bacterial strains and media

The *E. coli* strain used in the microfluidics experiments is ASC662 (MG1655 lacZ-GFPmut2) [22]. We confirmed by sequencing that this strain has no mutation in the *lac* operon apart from the lacZ-GFP translational fusion, and identified the GFP sequence as a GFPmut2 lacking the monomeric substitution A206K; fortunately this doesn't affect its spectrum nor its maturation time [23].

Diauxie experiments were done with this strain ASC662 and with its parent strain MG1655 (CGSC #7740) as well as with 6 natural isolates of *E. coli* [12]: SC303, SC305, SC330, SC347, SC355, SC366.

All experiments were done using M9 minimal media (Sigma-Aldrich) supplemented with 2 mM MgSO4, 0.1 mM CaCl2, and sugars as indicated (typically 0.2% for glucose or lactose and 0.4% for glycerol). TMG and IPTG were diluted from frozen stocks in water (at 0.1 M and 1 M respectively).

All experiments were carried out at 37°C. Note that the melAB operon is not expressed at this temperature [24] so that it should not interfere with the regulation of the *lac* operon.

In order to increase LacI activity in the repressed state (Fig. 1E), plasmid p_lacI_SC101 was transformed into ASC662. It was obtained by restriction-ligation cloning of MG1655's native lacI promoter and gene (amplified by PCR between positions 366,428 and 367,601) into a pSC101 vector (pUA66; using XbaI and XhoI, hence removing the plasmid's promoter and GFP loci). In order to make cloning easier, we used a high copy number derivative of pUA66 where the GFP locus is replaced with pUC19 origin (produced by restriction-ligation cloning with BamHI and XbaI).

2 Microfluidics experiments

2.1 Microfluidic device fabrication

The Dual Input Mother Machine (DIMM) microfluidic design used in this study has been described elsewhere [5] and is freely available online [25]: in brief it combines comb-like structures enabling to grow bacteria in steady state conditions over long times with a special type of junction allowing fast and accurate mixing of two input media. Several microfluidics masters were produced using soft lithography by micro-resist Gmbh; one master with regular growth channels of suitable size (0.7 μm width $\times$ 0.9 μm height) was used for all experiments.

For each experiment, a new chip was produced by pouring PDMS (Sylgard 184 with 1:9w/w ratio of curing agent) on the master and baking it for 4 h or more at 80ºC. After cutting the chip and punching inlets, the chip was bonded to a #1.5 glass coverslip as follows: the coverslip was sonicated 3 to 5 min in acetone, rinsed in isopropanol then water; the chip cleaned from dust using MagicTape, rinsed in isopropanol then water; surfaces were activated with air plasma (40 sec at 1500 μm of Hg) before being put in contact; the assembled chip was cooked 1 h or more at 80ºC.

Before running the experiment, the chip was primed and incubated 1 h at 37ºC using passivation buffer (2.5 mg/mL salmon sperm DNA, 7.5 mg/mL bovine serum albumin) for the mother machine channels and water for the overflow channels.

2.2 Culture conditions and flow control

Bacteria were streaked onto LB agar plates from frozen glycerol stocks thawed from -80ºC. Overnight precultures were grown from single colonies in M9 minimal medium supplemented with the same sugar that the cells were to experience in the initial condition of the experiment. The next day, cells were diluted 100-fold into fresh medium with the same sugar and harvested after 4–6 h.

The experimental apparatus was initialized, pre-warmed and equilibrated. Flow control was achieved using a syringe pump for early experiments and using a pressure controller later on. A total flow of $\approx$3 μL/min was used in all cases (corresponding to a pressure of $\approx$2000 mbar on both inlets). Polystyrene beads (Polybead 1 μm) were added to one of the two media to allow observing the flow ratio at the mixing junction.

The primed microfluidic chip was mounted, connected to media supply and flushed with running media for 30 min or more to rinse passivation buffer. The grown cell culture was centrifuged at 4000$\times$g for 5 min, and the pellet resuspended in a few μL supernatant and injected into the device using a syringe. Since flow was running, the pressure had to be continuously adjusted to make sure the cells stopped flowing in the main channel and could enter the growth channels (typically 10 to 40 min).

After loading, bacteria were incubated during 2 h before starting image acquisition. Every 3 min phase contrast and fluorescence images were acquired for several positions in parallel (typically 6), including the dial-a-wave junction with short phase contrast exposure (10 ms) so as to monitor the flow ratio between the two inputs.

All conditions used in this project are summarized in Table 4. The default protocol consists of 6 h in M9 + 0.2% glucose followed by 4 h in M9 + 0.2% lactose.

2.3 Microscopy and image analysis

An inverted Nikon Ti-E microscope, equipped with a motorized xy-stage and enclosed in a temperature incubator (TheCube, Life Imaging Systems) was used to perform all experiments. The sample was fixed on the stage using metal clamps and focus was maintained using hardware autofocus (Perfect Focus System, Nikon). Images were recorded using a CFI Plan Apochromat Lambda DM $\times$100 objective (NA 1.45,

WD 0.13 mm) and a CMOS camera (Hamamatsu Orca-Flash 4.0). The setup was controlled using μManager [26] and timelapse movies were recorded with its Multi-Dimensional Acquisition engine (customized using runnables). Phase contrast images were acquired using 100 ms exposure (CoolLED pE-100, full power). Images of GFP fluorescence (ex 475/35 nm; em 525/50 nm; bs 495 nm) were acquired using 2 s exposure (Lumencor SpectraX, Cyan LED at 17% with ND4).

Image analysis was performed using MoMA [5] as described in its documentation [27]. Raw image datasets were transferred to a centralised storage and preprocessed in batch. For each experiment, 30 growth channels were picked randomly (after discarding channels with structural defects or no cells growing at the first switch of condition) and curated manually in MoMA. MoMA's default post-processing was used in order to refine the measurements of cell length and total fluorescence. Fluorescence arbitrary units were converted to number of GFP molecules using the procedure and conversion factors described previously [5].

2.4 Quality control and data filtering

We established previously that cells slightly decrease their growth rate ($\sim 10-15\%$) in response to our illumination conditions and reach a steady state of growth after 2 hours [5]. Accordingly, the first 2 hours were discarded from further analysis in all experiments.

Given the design of the DIMM microfluidic chip, only one strain and one condition can be used in each experiment. As a result this study is based on more than 30 independent experiments performed over almost 3 years (Table 4). In order to minimize the confounding effect that the initial state of cells might affect their response to environmental change, we discarded experiments showing atypical growth patterns (used as a proxy for physiological state) before the switch of condition occurred. For each full cell cycle observed before the first switch, we fit a linear model to log(length) as a function of time and extract from this fit the growth rate, the Pearson correlation coefficient of the fit, and the predicted initial cell length. The distributions of correlation coefficients did not exhibit systematic changes across dates. However, the growth rate and initial cell length showed some variability across dates and, as expected, there was a clear correlation between cell size and growth rate in each condition (Fig. S1). The figure shows that 3 experiments appear as obvious outliers with slow growing cells and these were discarded (20151218, 20161021, 20180313); in addition, we discarded 2 more experiments with slow growing cells (20160526, 20170108). Finally, we also examined the reproducibility of the *lac* induction lags distributions between replicates and closely-related conditions and discarded 2 additional experiments because of their aberrant distributions of lags (20180123, 20180615). No reason could be found for these last two experiments to be qualitatively different from our expectations (leaving the door opened to an undocumented protocol mistake).

In addition, since our study focuses on the response of *growing* cells to changes of environmental conditions, we checked that the fraction of cells not growing before the switch in the analyzed growth channels is small in all experiments ($< 5\%$, typically $\sim 1\%$).

2.5 Estimation of *lac* induction lags

The time until the *lac* operon is induced in a given cell after a switch to lactose is estimated from the time series of LacZ-GFP level as follows: we compute the pre-induction level computed as the mean value in the 9 min following the switch (for which we checked visually that no induction occurred) and measure the delay after the switch until the cell has increased its level by 200 molecules (which is a conservative

threshold since the fluctuations of background level have a standard deviation of ≈ 20); hence we need to observe a cell for at least 4 time points after the switch to be able to successfully detect an induction. In order to successfully estimate a lag for as many cells as possible, if a cell divided before inducing, the total GFP of its two daughters was summed and their cell traces considered until the induction threshold was passed, or one of the two daughters was lost (because it left the channel). At the first switch to lactose, a finite induction lag was measured for 90% of the 1633 cells, while 7% exited the growth channels before inducing and 3% did not induce during the 4h in lactose. We note that this *lac* induction lag will include the maturation time of the LacZ-GFP fusion proteins (half time = 5.6 ± 0.4 min [23]). Since we established previously that the fact that rare cells exit the growth channels before inducing does not affect significantly the probability distribution of lags [5], we report here the raw distribution of measured lags.

2.6 Estimation of growth lags

When bacteria are exposed to lactose for the first time, they all transiently stop growing. In order to estimate the growth lag, *i.e.* the time until growth restarts after the switch to lactose, from noisy time series of cell length, we devised a Bayesian inference procedure based on a model where each cell is arrested at log-length L_0 during a lag τ and then elongates exponentially at a constant rate λ. The logaritm of cell length Λ follows:

$$\Lambda_\tau(t) = \begin{cases} L_0 & \text{for } t \leq \tau \\ \lambda t + L_0 & \text{for } t > \tau. \end{cases}$$

Assuming deviations from this model are only due to measurement noise, we can easily marginalize over it and over L_0. However marginalizing over λ is less straightforward; we used the Laplace method to approximate the corresponding integral, leading to the following log-likelihood function

$$\log\left[p\left(\vec{L}\,\middle|\,\tau\right)\right] \propto -\log\left\|\vec{\Lambda}^c_\tau\right\| + (2-T)\log\left\|\vec{L}^c - \frac{\vec{L}^c\cdot\vec{\Lambda}^c_\tau}{\left\|\vec{\Lambda}^c_\tau\right\|^2}\vec{\Lambda}^c_\tau\right\|$$

where the c superscript stands for the zero mean vector $\vec{x}^c = \vec{x} - \bar{x}$. We then compute this log-likelihood numerically for all possible values of τ and picked $\tau^\star$ as the one with the highest log-likelihood value. The others maximum likelihood quantities directly follow from previous calculations and read $\lambda^\star = \frac{\vec{L}^c\cdot\vec{\Lambda}^c_{\tau^\star}}{\|\vec{\Lambda}^c_{\tau^\star}\|^2}$, $L_0^\star = \bar{L} - \lambda^\star\bar{\Lambda}_{\tau^\star}$, and $\sigma^{2\star} = \mathrm{Var}\left(\vec{L} - L_0^\star - \lambda^\star\vec{\Lambda}_{\tau^\star}\right)$.

We applied this procedure to the time series of log-length measured after the switch. Similarly to the induction lags, in order to successfully estimate a lag for as many cells as possible, if a cell divided before inducing, the total length of its two daughters was summed and their cell traces considered until the total length had tripled since the switch, or one of the two daughters was lost (because it left the channel). At the first switch to lactose, a finite growth lag was measured for 85% of the 1633 cells, while 9% exited the growth channels before resuming growth and 6% remain arrested during the 4h in lactose.

2.7 Estimating LacZ-GFP levels before the switch

To compare the initial LacZ-GFP levels between cells with short and long lags (Fig. 2A), we used the same average over the 3 time points after the switch. Since autofluorescence is large relative to LacZ-GFP fluorescence in uninduced cells, it is not possible to detect any correlation between LacZ-GFP levels and the *lac* induction lag (Fig. S4) Another

confounding factor is that every experiment has a slightly different offset for fluorescence levels (data not shown); when pooling observations from different experiments, this difference is larger than the difference of LacZ-GFP levels. Consequently, we compared the distribution of pre-existing LacZ-GFP levels in cells with short and long lags after correcting each cell by subtracting the mean in the same experiment (Fig. 2A). As a control, we measured the autofluorescence of wild-type cells (grey line in the lower panel) and noticed that the variability of the fluorescence is the same in uninduced cells (blue line), which confirms that it is dominated by autofluorescence (Fig. 2A). In addition we corrected day-to-day offsets in the fluorescence measurements by subtracting to each cell the average fluorescence in each experiment. Comparing the distributions of corrected LacZ-GFP levels revealed that cells with short lags (orange line) have on average 5 LacZ-GFP molecules more than cells with long lags (blue line; Fig. 2A).

In order to analyze the sensitivity of the *lac* operon, we would like to know the number of β-galactosidase molecules in each bacterium at the onset of the switch to lactose. However, the combination of limited camera sensitivity and spontaneous autofluorescence (with fluctuations on the order of ±20 GFP per cell), prevents us from accurately measuring levels lower than 100 molecules. In addition, repeated illumination induces bleaching so that not all GFP molecules are fluorescent (which we previously quantified to correspond to ≈20% per cell cycle) [5]. However, at the second switch, the number of β-galactosidases inherited from the ancestors' induction during the first exposure can be computed as follows: the lineage is traced backed in time and both L_m, the maximum LacZ-GFP level during the first exposure, as well as n_d, the number of divisions since this time, are recorded; the expected number of inherited molecules equals $L_m/2^{n_d}$. Note that this formula assumes no degradation (which is reasonable since the longest delay is 24 h and β-galactosidase is stable over this time scale) and even partitioning at every division, and that this constitutes a lower bound for the number of β-galactosidase molecules in the cell since neither bleaching nor stochastic expression happening during the delay period are accounted for.

3 Diauxie experiments

3.1 Culture conditions

All diauxie experiments were performed in 96-well plates (Greiner #655090) incubated with 600 rpm shaking. Glycerol stocks were prepared in a single plate for the 8 strains of interest (with 2 replicates for each strain coming from independent clones) and stored at -80°C. A preculture was inoculated with a pin replicator in M9 + 0.2% glucose and grown overnight to saturation. These cultures were diluted (1600 to 10000×, using 2 serial dilutions) to the appropriate growth media, covered with 50 μL mineral oil (to prevent evaporation), and OD was measured approx. every 2 min during 24 to 36 h in a spectrophotometer (Synergy 2, Biotek) with temperature control and shaking (*fast continuous* setting). At least 2 independent replicates were measured per condition.

3.2 Growth curves analysis

Wells were inoculated at very low cell concentration so that at least 2 h were recorded with OD below the detection limit. The background absorbance was estimated in each well as follows: the average absorbance was computed in sliding windows of 10 consecutive measurements and the lowest value of all windows with coefficient of variation below 0.02 was taken as background. For each condition, all growth curves were aligned so that OD = 0.01 at $t = 0$ where OD = absorbance − background. A few wells showing aberrant records (e.g. discontinuous growth curve) or outlier growth

patterns were discarded from further analysis. In addition strain SC330 displayed very variable growth between replicates (in particular variable lags) and could not be analysed further since too many growth curves were truncated.

To estimate the population lag for a given condition, we analyse the delay between the corresponding growth curves and the growth curves in the same condition supplemented with 200 μM IPTG. In particular, using simple linear interpolation between the time points at which OD was measured, we determined for each growth curve a continuous function $t = f(\mathrm{OD})$ of the time t for which a particular OD was first reached. Using these, a mean delay and its standard error was computed as a function of OD using all replicates for a given condition (Fig. 3E). The population lag was then estimated visually as the time at which a plateau occurs in the delay, i.e. the delay value with smallest derivative.

3.3 Diauxie control

In order to disentangle the effects of the lactose concentration on *lac* expression in the glucose-lactose mixture but also the initial rate of lactose import when glucose gets exhausted, we performed additional experiments where cultures grown on mixture of glucose and lactose were transferred manually to fresh growth medium with lactose only. Transferring cells between media after washing them required to use larger volumes; we hence adapted the diauxie protocol to 15 mL Falcon tubes and used only strain ASC662.

Overnight cultures in 2 mL M9 + 0.2% glucose were inoculated from frozen glycerol stocks and grown to saturation with 200 rpm shaking. After 200×dilution to 2 mL fresh media with 0.2% glucose and variable lactose concentrations (no lactose, 0.02%, 0.2%), bacteria were grown until OD reached 0.08-0.2. 1 mL of each culture was then sampled and washed twice in M9 + 0.02% lactose. Each sample was used to start four 150 μL cultures in a 96-wells plate, two of which were supplemented with 20% lactose to reach a concentration of 0.2%. All wells were covered with 50 μL mineral oil and OD was measured approx. every 2 min during 12 h as described above.

4 Population lag simulations

Bacterial growth was simulated in discrete time (with 1 min steps) assuming deterministic exponential growth in glucose + lactose and in lactose media, with doubling time 49 min and 58 min respectively. Since the effect of nutrient concentration on growth rate was not taken into account, the population size at the start or at the switch to lactose are irrelevant. For a given empirical distribution of single-cell lags, the distribution was discretized into a 100 subpopulations with different lag values. For each subpopulation growth was simulated by keeping a constant population until its lag time was reached, followed by simple exponential growth. The population growth curve after the switch was obtained by summing the growth curves of all 100 subpopulations. Finally the population lag was inferred as the delay compared to the simulation where all cells immediately grow after the switch to lactose, measured once all cells have resumed growth.

Acknowledgments

We are thankful to all members of the van Nimwegen lab for their input, in particular Chris Field and Guillaume Witz. Image and data analysis were performed at sciCORE, the scientific computing center of the University of Basel. We thank Daan Kiviet and Sander Tans for sharing strain ASC662.

References

1. Monod J. The Growth of Bacterial Cultures. Annual Review of Microbiology. 1949;3(1):371–394. doi:10.1146/annurev.mi.03.100149.002103.

2. Jacob F, Monod J. Genetic Regulatory Mechanisms in the Synthesis of Proteins. Journal of Molecular Biology. 1961;3(3):318–356. doi:10.1016/S0022-2836(61)80072-7.

3. Veening JW, Smits WK, Kuipers OP. Bistability, Epigenetics, and Bet-Hedging in Bacteria. Annual Review of Microbiology. 2008;62:193–210. doi:10.1146/annurev.micro.62.081307.163002.

4. Wolf L, Silander OK, van Nimwegen E. Expression Noise Facilitates the Evolution of Gene Regulation. eLife. 2015;4:e05856. doi:10.7554/eLife.05856.

5. Kaiser M, Jug F, Julou T, Deshpande S, Pfohl T, Silander OK, et al. Monitoring Single-Cell Gene Regulation under Dynamically Controllable Conditions with Integrated Microfluidics and Software. Nature Communications. 2018;9(1):212. doi:10.1038/s41467-017-02505-0.

6. Lambert G, Kussell E. Memory and Fitness Optimization of Bacteria under Fluctuating Environments. PLoS Genet. 2014;10(9):e1004556. doi:10.1371/journal.pgen.1004556.

7. Cai L, Friedman N, Xie XS. Stochastic Protein Expression in Individual Cells at the Single Molecule Level. Nature. 2006;440(7082):358–362. doi:10.1038/nature04599.

8. Choi PJ, Cai L, Frieda K, Xie XS. A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell. Science (New York, NY). 2008;322(5900):442–446. doi:10.1126/science.1161427.

9. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A. Multistability in the Lactose Utilization Network of *Escherichia coli*. Nature. 2004;427(6976):737–40.

10. Bren A, Park JO, Towbin BD, Dekel E, Rabinowitz JD, Alon U. Glucose Becomes One of the Worst Carbon Sources for *E. coli* on Poor Nitrogen Sources Due to Suboptimal Levels of cAMP. Scientific Reports. 04 25, 2016;6:24834. doi:10.1038/srep24834.

11. Görke B, Stülke J. Carbon Catabolite Repression in Bacteria: Many Ways to Make the Most out of Nutrients. Nature Reviews Microbiology. 2008;6(8):613–624. doi:10.1038/nrmicro1932.

12. Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ. Presence and Growth of Naturalized Escherichia Coli in Temperate Soils from Lake Superior Watersheds. Applied and Environmental Microbiology. 2006;72(1):612–621. doi:10.1128/AEM.72.1.612-621.2006.

13. Sakoparnig T, Field C, van Nimwegen E. Whole genome phylogenies reflect long-tailed distributions of recombination rates in many bacterial species. bioRxiv. 2019;doi:10.1101/601914.

14. Dykhuizen D, Hartl D. Transport by the Lactose Permease of *Escherichia coli* as the Basis of Lactose Killing. Journal of Bacteriology. 1978;135(3):876–882.

15. Eames M, Kortemme T. Cost-Benefit Tradeoffs in Engineered Lac Operons. Science. 2012;336(6083):911–915. doi:10.1126/science.1219083.

16. Uphoff S, Lord ND, Okumus B, Potvin-Trottier L, Sherratt DJ, Paulsson J. Stochastic Activation of a DNA Damage Response Causes Cell-to-Cell Mutation Rate Variation. Science. 2016;351(6277):1094–1097. doi:10.1126/science.aac9786.

17. Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, et al. The Quantitative and Condition-Dependent *Escherichia coli* Proteome. Nature Biotechnology. 2016;34(1):104–110. doi:10.1038/nbt.3418.

18. Robert L, Paul G, Chen Y, Taddei F, Baigl D, Lindner AB. Pre-Dispositions and Epigenetic Inheritance in the *Escherichia coli* Lactose Operon Bistable Switch. Molecular Systems Biology. 2010;6:357. doi:10.1038/msb.2010.12.

19. Boulineau S, Tostevin F, Kiviet DJ, ten Wolde PR, Nghe P, Tans SJ. Single-Cell Dynamics Reveals Sustained Growth during Diauxic Shifts. PloS One. 2013;8(4):e61686. doi:10.1371/journal.pone.0061686.

20. Mitchell WJ, Saffen DW, Roseman S. Sugar Transport by the Bacterial Phosphotransferase System. In Vivo Regulation of Lactose Transport in Escherichia Coli by IIIGlc, a Protein of the Phosphoenolpyruvate:Glycose Phosphotransferase System. The Journal of Biological Chemistry. 1987;262(33):16254–16260.

21. You C, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, et al. Coordination of Bacterial Proteome with Metabolism by Cyclic AMP Signalling. Nature. 2013;500(7462):301–306. doi:10.1038/nature12446.

22. Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of Metabolism and Growth at the Single-Cell Level. Nature. 2014;514(7522):376–379. doi:10.1038/nature13582.

23. Balleza E, Kim JM, Cluzel P. Systematic Characterization of Maturation Time of Fluorescent Proteins in Living Cells. Nature Methods. 2018;15(1):47–51. doi:10.1038/nmeth.4509.

24. Kennedy EP. Chapter IV: The Lactose Permease System of Escherichia coli. Cold Spring Harbor Monograph Archive. 1970;01(0). doi:10.1101/087969100.1.49.

25. DIMM: The Dual Input Mother Machine;. Available from: `https://metafluidics.org/devices/dual-input-mother-machine/`.

26. Edelstein A, Amodaj N, Hoover K, Vale R, Stuurman N. Computer Control of Microscopes Using μManager. Current Protocols in Molecular Biology. 2010;Unit 14.20:1–17. doi:10.1002/0471142727.mb1420s92.

27. MoMA Wiki;. Available from: `https://github.com/fjug/MoMA/wiki`.

Supporting information

Table S1 List of experiments used in this study with summary statistics. Experiments that have been discarded from further analysis are greyed out. For experiments acquired before November 2015, data are taken from [5].

condition	date	# growth channels	# full cell cycles	# observations	# cells at switch	# estimated lags	# arrested cells at switch
mg1655	20150812	6	364	17283	NA	NA	NA
glucose	20150616	31	2021	92270	NA	NA	NA
glucose	20150617	30	1920	91276	NA	NA	NA
lactose	20150624	30	1649	86776	NA	NA	NA
lactose	20150630	30	1671	85949	NA	NA	NA
memory 4h	20150703	28	1431	76295	462	445	3
memory 4h	20150708	30	1600	88410	536	511	6
memory 6h	20151204	30	1848	107195	548	524	10
memory 8h	*20151218*	*29*	*1821*	*115062*	*NA*	*NA*	*NA*
memory 8h	20180206	31	NA	66894	332	312	1
memory 12h	20180207	30	1517	84568	353	340	3
memory 12h	20180216	30	1633	82626	344	329	4
memory 12h (old)	*20160526*	*31*	*1178*	*83954*	*NA*	*NA*	*NA*
memory 16h	20160912	39	2437	134488	444	419	8
memory 20h	20161014	34	2224	133238	382	355	10
memory 24h	20161007	31	2343	135928	350	334	6
memory 24h	*20180313*	*9*	*57*	*9192*	*NA*	*NA*	*NA*
late	*20161021*	*22*	*865*	*83093*	*NA*	*NA*	*NA*
late	*20170108*	*30*	*1789*	*103600*	*NA*	*NA*	*NA*
late	20180516	30	2370	104792	166	155	0
late	*20180615*	*10*	*794*	*34532*	*NA*	*NA*	*NA*
ramp40min	20171121	30	371	29252	164	154	1
ramp40min	20180319	32	496	33684	195	185	1
preIPTG5uM	20180514	31	385	28585	178	167	0
preIPTG5uM	20180531	29	410	26780	170	161	0
lacIoe	20180116	31	796	55447	177	161	0
lacIoe	*20180123*	*29*	*758*	*49517*	*NA*	*NA*	*NA*
lacIoe	20180214	31	761	46967	157	142	1
lacIoe preIPTG10uM	20180604	33	410	28272	174	160	1
glyc to lac	20170919	26	139	28832	137	123	NA
glyc to lac	20170920	29	NA	32475	142	132	NA
glc+lac to lac	20171114	31	842	47798	180	171	0
glc+lac to lac	20180108	31	804	50630	174	170	4
glc+lac to lac	*20180606*	*9*	*579*	*25851*	*NA*	*NA*	*NA*

Fig. S1 Physiological state of the cells during the initial growth conditions. For each experiment, the medians and 95% posterior intervals are plotted of both growth rate and cell length at birth.

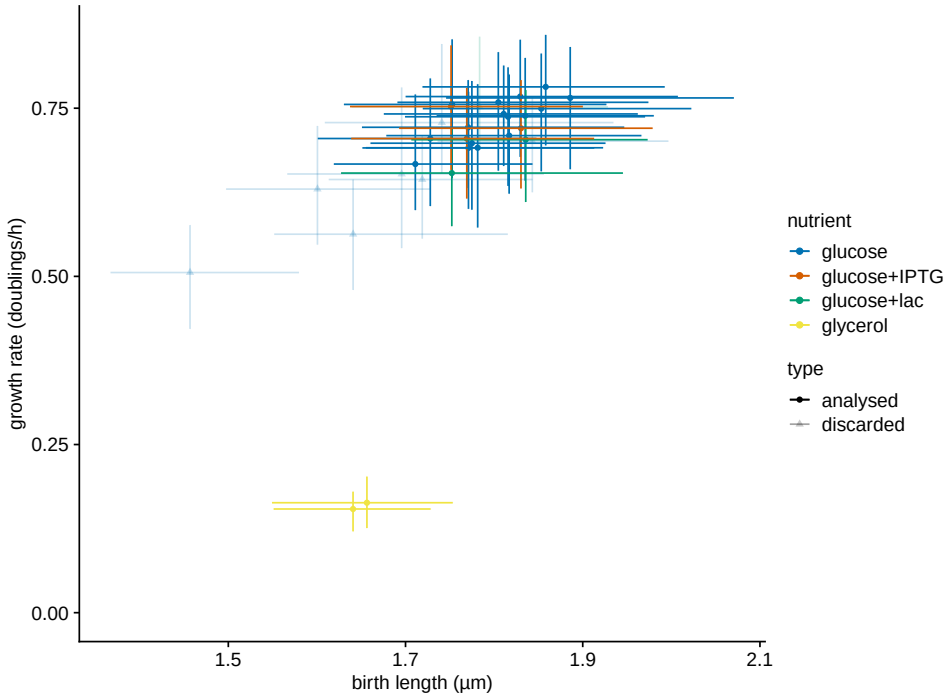

Fig. S2 Controls on the statistics of the *lac* operon single-cell lags. A. Distribution of induction lags for the *lac* operon in naive cells, coloured per day. Due to the limited sample size in each experiment, the bin width was increased to 6 min. Note that the bimodality of *lac* induction lags distributions is a robust feature. **B.** Distribution of induction lags for the *lac* operon in naive cells, stratified per position in the growth channels. Position is indicated as cell rank, counted from the cell closest to the channel open end in the upper panel and closest to the closed end in the lower panel. Note that cells close to the open end tend to exit the channel before inducing their *lac* operon, hence the distributions are noisier due to smaller sample size. **C.** Distribution of growth lags in naive cells. The bin width is the same as the experimental acquisition frequency (3 min). The corresponding distribution of *lac* induction lags (Fig. 1C) is shown for comparison. Note that the distribution of growth lags shows a less marked bimodality, which might result from a combination of less accurate estimation of the growth lag and additional sources of noise being involved in restarting growth once the *lac* operon is expressed.

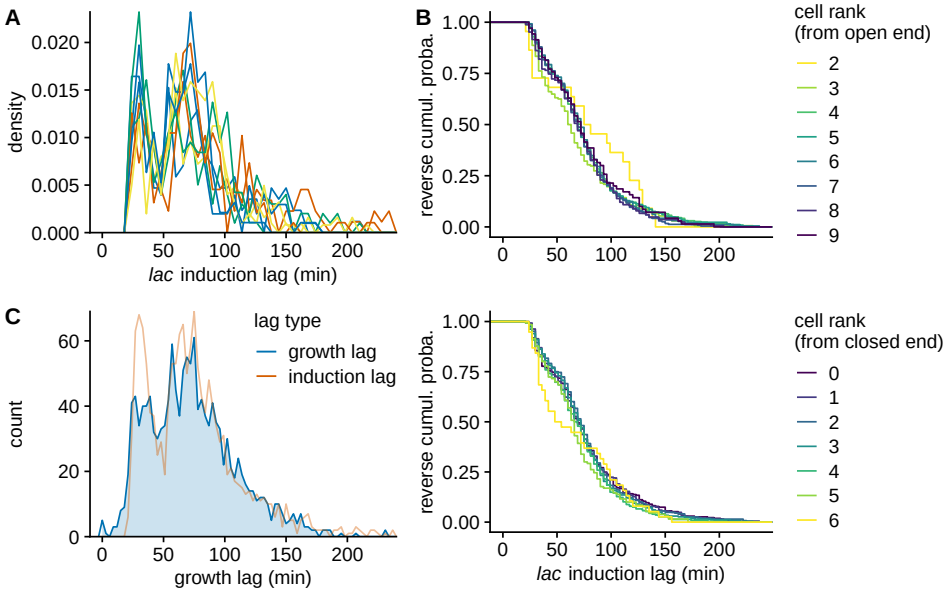

Fig. S3 ***lac*** **induction lags for a gradual transition from 0.2% glucose to 0.2% lactose over 40 min.** Comparison of the distribution of induction lags for the *lac* operon in naive cells exposed to a gradual transition from 0.2% glucose to 0.2% lactose over 40 min (2 independent replicates, blue curve) with the distribution of lags under a sudden switch (orange curve, Fig. 1C). Note that, since we do not know at what point in the 40 minute transition the critical concentrations of glucose/lactose are reached, the lags for each replicate with a gradual transition were offset by a delay that maximized the overlay with the lags under a sudden switch. The fact that the distribution of lags under the gradual transition is almost identical to the distribution under a sudden switch shows that the stochastic single-cell responses remain equally synchronized under the gradual transition, suggesting that there is a common critical concentration of glucose/lactose across all cells. Due to the limited sample size for the gradual transition, the bin width was increased to 6 min.

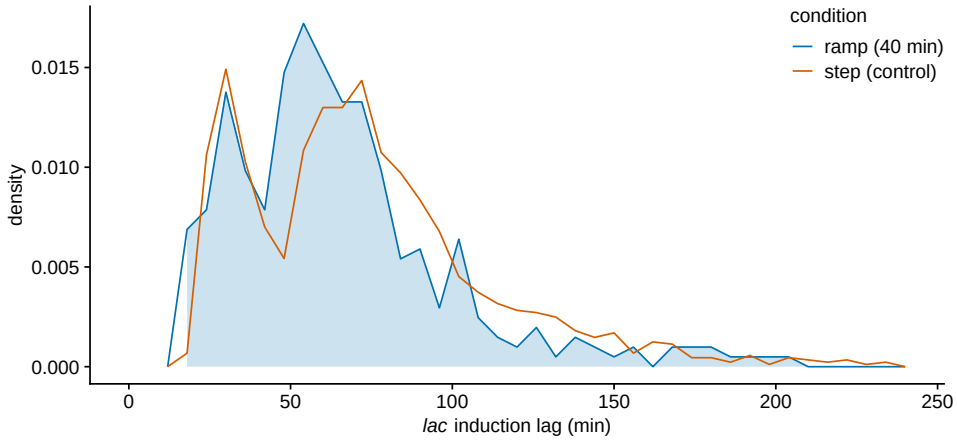

Fig. S4 Induction lag does not correlate with physiological traits at the time of the switch. Neither cell cycle progression (measured either as the time since birth normalised to the average division time in this condition, or as length added since birth as suggested by the "adder" model of cell cycle control), nor fluorescence at the switch (measured in units of LacZ-GFP molecules), nor growth rate correlate with *lac* induction time. Note that fast switching cells are indicated in orange and slow switching cells in blue.

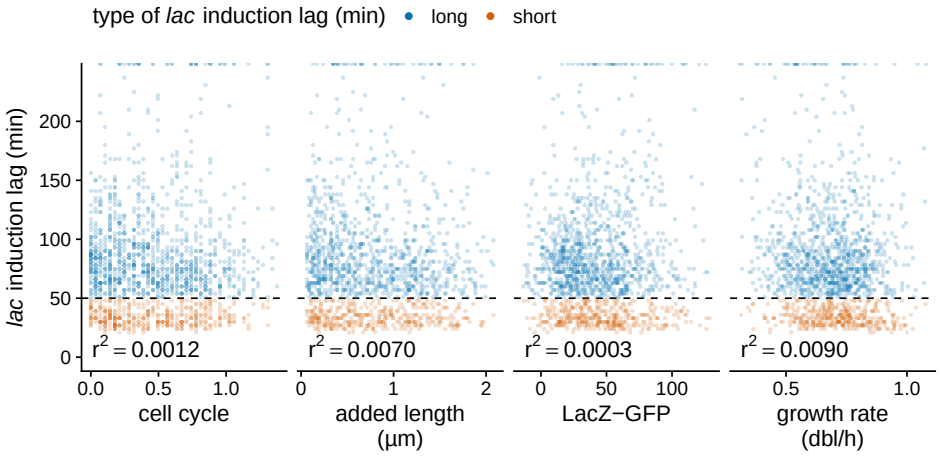

Fig. S5 LacZ-GFP level before exposure to lactose when LacI activity is reduced by low level of IPTG. Note that this measurements are imprecise due to relatively large fluctuations in autofluorescence (between cells) and in illumination intensity (between replicates). Although this treatment increases the fraction of fast-switching cells, no detectable change of LacZ-GFP levels can be measured which supports that the increase of basal expression is less than 50 molecules. In comparison, bacteria carry 3000-6000 LacZ-GFP molecules at full induction (Fig. 1B).

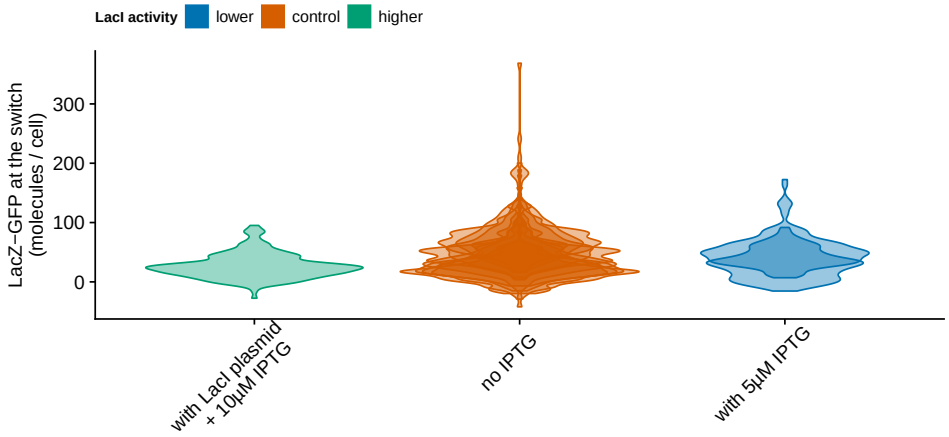

Fig. S6 ***lac*** **induction lag at the second switch as a function of the estimated number of inherited LacZ-GFP.** Each dot corresponds to a cell with its estimated lag shown along the vertical axis, its estimated number of remaining LacZ-GFP molecules along the horizontal axis, and its color corresponding to the amount of time the cell spend in glucose between the two lactose phases. The dotted line corresponds to the 50 min. threshold that separates short from long lags. We stratified the cells into 8 groups depending on their estimated numbers of inherited LacZ-GFP molecules remaining at the second switch and the violin plots show the distributions of lag times of each group, with its horizontal position centered on the average of the group. Note that long lags only reappear for cells with less than 10 inherited molecules of LacZ-GFP.

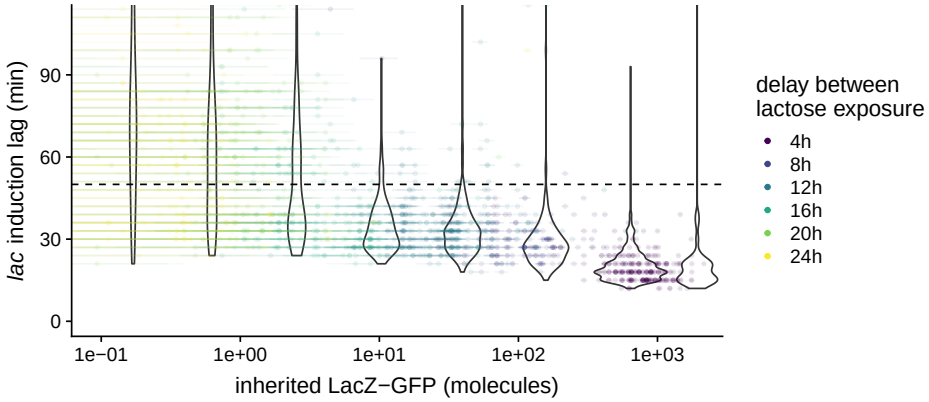

Fig. S7 Population growth curves during diauxie experiments. All growth curves used to compute population lags reported in Fig. 3F are shown. For each mixture of glucose and lactose (orange lines), the corresponding control with constitutive *lac* operon expression was obtained by supplementing IPTG (blue lines). Each line corresponds to a biological replicate; delays were computed for OD below the carrying capacity (solid sections).

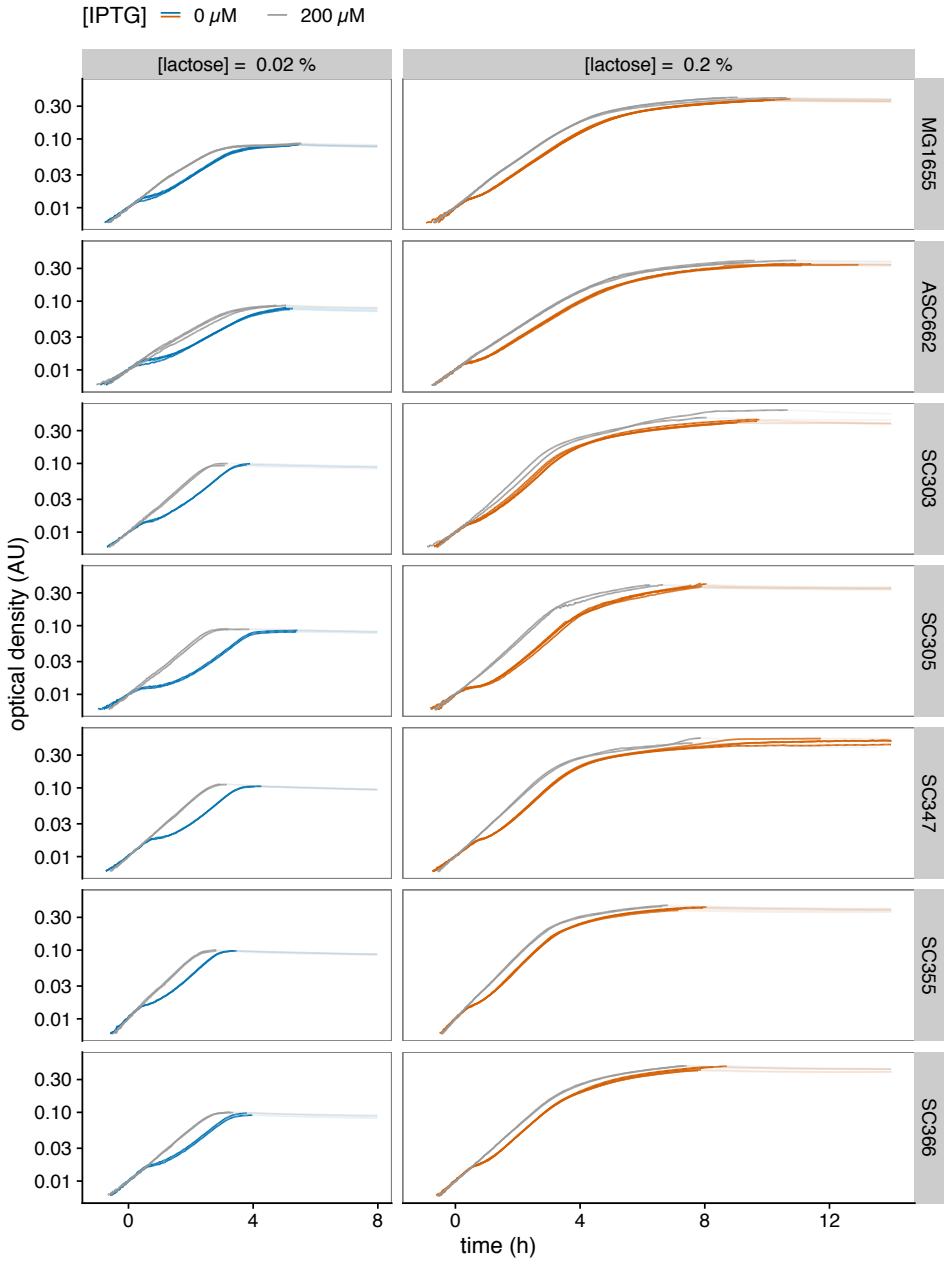

Fig. S8 Lactose concentration during growth in glucose determines the population lag at the diauxic switch. The strain ASC662 was manually transferred from a mixture of glucose (0.005%) and lactose of a given concentration after 4 to 5 h of growth to media with lactose only (OD≈0.1-0.15), at a given concentration. Experiments were performed with three different lactose concentrations before the switch and two after the switch, for a total of 6 combinations. For each combination, the population lag after the transfer was measured as the delay until OD increased by 0.001 (corresponding to an approx. 10% increase of population size).

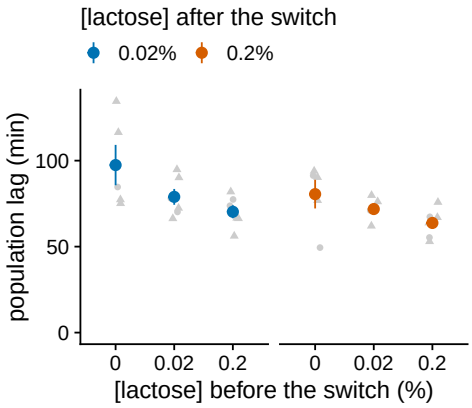
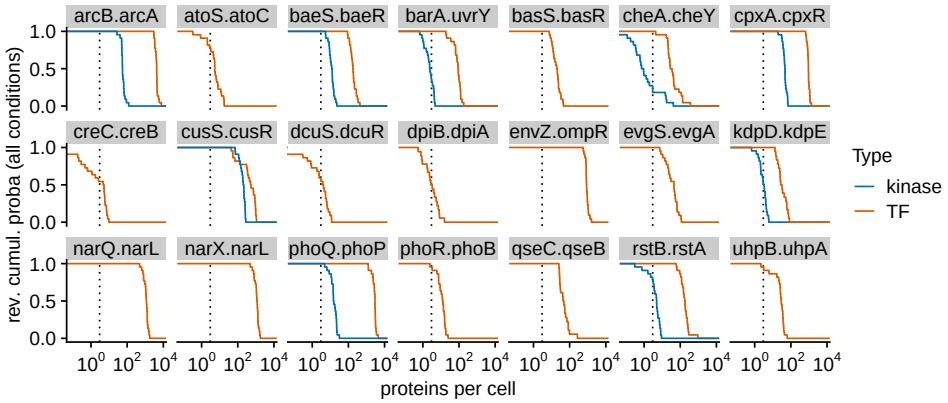
Fig. S9 Many two-components systems in *E. coli* have expression levels so low that no proteins will occur in a substantial number of cell. Distributions of expression levels of sensor kinases (blue) and response regulator transcription factors (orange) across 28 conditions as measured using quantitative proteomics [17] for 21 of the 28 two-components systems annotated on `ecocyc.org`. Each line shows the reverse cumulative distribution of average number of proteins per cell measured across the 28 conditions. The vertical dashed lines correspond to an average of 3 proteins per cell which is the threshold below which more than 5% of the cells is expected not to have any of the corresponding protein (assuming a Poisson distribution for the abundance of low expressed proteins, i.e. $e^{-3} \approx 0.05$).