bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Reducing phenotypic and genotypic instabilities of microbial population during continuous cultivation based on stochastic switching dynamics

**Thai Minh Nguyen[1], Samuel Telek[1], Andrew Zicler[1], Juan Andres Martinez[1], Boris Zacchetti[1], Julian Kopp[2], Christoph Slouka[2], Christoph Herwig[2,3], Alexander Grünberger[4], Frank Delvigne[1*]**

[1] Terra research and teaching centre, Microbial Processes and Interactions (MiPI), Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium
[2] Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses, Institute of Chemical, Environmental and Biological Engineering, Vienna University of Technology, 1060 Vienna, Austria
[3] Research Division Biochemical Engineering, Institute of Chemical Environmental and Bioscience Engineering, Vienna University of Technology, Vienna, Austria
[4] Multiscale Bioengineering, Technical Faculty, Bielefeld University, Bielefeld Germany & CeBiTec, Bielefeld University, Bielefeld, Germany

***Corresponding author:** F.Delvigne@uliege.be. Terra research and teaching center, Microbial Processes and Interactions (MiPI), Gembloux Agro-Bio Tech, University of Liège, Gembloux, Avenue de la Faculté 2B, 5030, Belgium. Tel: 0032 81 62 23 09

## Abstract

Predicting the fate of a microbial population (i.e., growth, gene expression…) remains a challenge, especially when this population is exposed to very dynamic environmental conditions, such as those encountered during continuous cultivation processes. Indeed, the dynamic nature of continuous cultivation process implies the potential deviation of the microbial population involving genotypic and phenotypic diversification. This work has been focused on the induction of the arabinose operon in *Escherichia coli* as a model system. As a preliminary step, the GFP level triggered by an arabinose-inducible $P_{BAD}$ promoter has been tracked by flow cytometry in chemostat with glucose-arabinose co-feeding. Ampicillin was used as an "unstable" selective marker, allowing the simultaneous investigation of the effect of phenotypic diversification and genetic instability in continuous cultures. Under classical chemostat operation, the system was very unstable, with only a small fraction of cells (less than 10%) being able to accumulate GFP to a large extent, this fraction rapidly collapsing with time and going below 10% of the total population. On the long run, this phenotypic diversification was followed by an extensive loss of plasmid. In a second set of experiments, continuous cultivation was performed by adding either glucose or arabinose, based on the ability of individual cells for switching from low GFP to high GFP states, according to a technology called segregostat. In segregostat mode of cultivation, on-line flow cytometry analysis was used for adjusting the arabinose/glucose transitions based on the stochastic switching capabilities of the microbial population. This strategy allowed finding an appropriate arabinose pulsing frequency, leading to a prolonged maintenance of the induction level with limited impact of phenotypic diversification and genetic instability for more than 68 generations

**Keywords** : Segregostat, Single cell analysis, Flow cytometry, Arabinose operon, Biological noise, Plasmid stability, GFP reporter

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Introduction

Due to the inherent intracellular biological noise and under specific diversification pressures, microbial cells within the same population tend to split into subpopulations exhibiting different metabolic features [1][2][3]. This cell-to-cell variability in metabolic activities has long been recognized as an adverse effect for bioprocessing [4]. It is believed, specifically in the field of biochemical and bioprocess engineering, that achieving highly homogeneous microbial populations, more tractable, robust and predictable bioprocesses can be attained [5][6]. However, in nature microbial cells have evolved to constantly adapt to different diversification pressures and phenotypic plasticity is known to improve cellular decision-making processes, resulting in most of the cases in a fitness gain for the whole population [7]. Adaptation to carbon limitation [8] and switching to an alternative carbon source [9][10] have been found to be key drivers of microbial phenotypic diversification. Optimal exploitation of the functionalities offered by this diversification process by microbial populations has typically lead to improved fitness in front of environmental perturbations [11], notably through mechanisms such as bet-hedging [12][13]. Bet-hedging relies on the simultaneous occurrence of subpopulations of cells pre-adapted to diverse environmental conditions, leading to an anticipative adaptation to unexpected environmental transition [14]. It has been mathematically demonstrated that this strategy is particularly beneficial in fluctuating environments with given fluctuation frequencies and amplitudes [15]. More specifically, high phenotypic diversity has been found to lead to a gain in fitness if the rate of phenotypic switching is equal or lower than the rate of environmental transitions. Phenotypic switching dynamics are complex and involve several layers of regulation, each at specific time-scales, i.e. transcription or translation, noise propagation through the gene regulatory network (GRN) [16] and possible feedback effects exerted by the cell division process, resulting in the dilution over the intracellular molecular species [17]. These switching mechanisms are then dependent on the specifics of the biological system and its culturing processing conditions, therefore they must be explored and analyzed based on single cell data.

In this work, population stability in a continuous cultivation was controlled via population heterogeneity measurements coupled with tailored carbon source feeding strategies. In this context, the induction of the arabinose operon in *Escherichia coli* following glucose scarcity and switch to arabinose utilization was used as a tunable model system. The strategy relying on the use of environmental perturbations at given frequencies and amplitudes for stabilizing and directing microbial diversification processes was explored through an experimental device previously set-up and published, i.e. the segregostat [18]. The segregostat is a continuous cultivation device connected to on-line flow cytometry for tracking phenotypic diversification processes. More importantly, the device is able to automatically induce environmental transitions based on the phenotypic state of the population. In this work, this device was used to assess whether stochastic switching could be applied for stabilizing microbial populations under continuous cultivation. Surprisingly, when setting this switching rate for maximizing phenotypic diversification dynamics, the segregostat led to a more predictable induction profile of the arabinose-inducible araBAD promoter ($P_{BAD}$), which remained stable over more than 68 cell generations. In comparison to chemostat cultivation, where population bistability was observed, the segregostat led to a monostable population of cells smoothly transitioning between the uninduced and induced states. Besides phenotypic diversification, occurring on short timescales, continuous bioprocessing operations are also prone on longer timescales to genotypic diversification, i.e. mutation driven processes [19] including plasmid loss [20]. Furthermore, classical chemostat led to massive plasmid loss, whereas the segregostat led to a mitigation of the diversification processes and to a more genotypically stable population marked by a higher plasmid stability.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Method

### Strains and media

*E. coli K12 W3110* wildtype was transformed with pBbB8a (4538 bp) plasmid with the $P_{BAD}$::GFP , which is an arabinose inducible promoter system. This plasmid belongs to a library of expression vectors compatible with the BglBrick standard [21]. Pre-cultures and cultures were performed on defined mineral salt medium containing (in $g.L^{-1}$): $K_2HPO_4$ 14.6, $NaH_2PO_4.2H_2O$ 3.6; $Na_2SO_4$ 2; $(NH_4)_2SO_4$ 2.47, $NH_4Cl$ 0.5, $(NH_4)_2$-H-citrate 1, glucose 5, thiamine 0.01. Thiamine was sterilized by filtration (0.2 µm). The medium was supplemented with 3 ml $l^{-1}$ of trace element solution, 3 ml $l^{-1}$ of a $FeCl_3.6H_2O$ solution (16.7 $g.L^{-1}$), 3 ml $l^{-1}$ of an EDTA solution (20.1 g $l^{-1}$) and 2 $ml.L^{-1}$ of a $MgSO_4$ solution (120 $g.L^{-1}$). The trace element solution contained (in $g.L^{-1}$): $CaCl_2.H_2O$ 0.74, $ZnSO_4.7H_2O$ 0.18, $MnSO_4.H_2O$ 0.1, $CuSO4.5H_2O$ 0.1, $CoSO_4.7H_2O$ 0.21 [22]. Ampicillin (100 µg/mL) was added to the medium.

### Chemostat cultivations

Chemostat cultures were performed in a 2MAG© block systems (2mag AG, Munich, Germany). The bioreactor system was equipped with positions for eight fermentation vessels with 16 mL total volume and working volume of 10 mL. The bioreactor block was equipped with magnetic inductive drives with two independent heat exchangers integrated into the bioreactor block, one for temperature control for the reaction broth and the second to control the headspace temperature and prevent evaporation. The system was also equipped with fluorometric sensor spots for dissolved oxygen (DO) positioned at the bottom of each reactor (minireaktor HTBD, Presens, Regensburg, Germany). Each reactor was equipped with a magnetic S-impeller with two permanent magnets (Sm2Co17, IBS magnet, Berlin Germany). In each chemostat, fresh medium was pumped into the culturing chamber at a constant rate, while culture effluent exits at an equal rate though the inlet and outlet tubes with an inner diameter of 0.8 mm (Marprene tubing; Watson-Marlow, Rommerskirchen, Germany). Overnight pre-cultures w performed in 1L baffled flasks containing 100 ml of culture medium and stirred with 200 rpm at 37°C. Fermenter inoculation was adjusted to 0.5 $OD_{600}$ for the initial batch phase. The temperature was maintained at 37°C under continuous stirring rate of 2600 rpm. The DO signal was used as an indicator for switching to the continuous operation modes (a rapid increase of the DO signals after depletion of glucose in the initial batch medium, observed typically after 3–5 h). The medium was continuously fed with the complete minimal medium containing glucose and arabinose at different ratios at a dilution rate of 0.5 $h^{-1}$. Off-line samples were taken from the reactors each 5-10 h and the analysis of the GFP expression levels was performed with a BD Accuri C6 (BD Biosciences, CA, USA) based on FL1-A channel (excitation 488 nm, emission 533 nm).

### Segregostat cultivations

Cultures in segregostat mode were performed in lab-scale stirred bioreactor (Biostat B-Twin, Sartorius) with 1 L working volume. The batch phase was started with 0.5 $OD_{600}$. The pH was maintained at 7.0 by automatic addition of ammonia solution 25%(w/v) or phosphoric acid 25% (w/v). The temperature was maintained at 37°C under a continuous stirring rate of 800 rpm and an aeration rate of 1 vvm. The DO signal was used as an indicator for switching to the continuous operation modes (a rapid increase of the DO signals after depletion of glucose in the initial batch

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

medium, observed typically after 3–5 h). The feeding medium, containing 5 g/L glucose, was continuously fed with a dilution rate of 0.5 $h^{-1}$. According to the sequences controlled by the online FC software, pulse of arabinose was injected in order to quickly increase (approx. 30 seconds) the global arabinose concentration to 0.5 g/L [18]. On-line sampling was performed every 12 minutes and processed according to the dilution sequence set through the following series of steps. First, the sample is automatically transferred to C6 FC (BD Accuri C6, BD Biosciences) and analyzed at a medium flow rate (33 µl $min^{-1}$). All the data related to the different parameters (mean, median, CV) are displayed in real-time during the cultivation. Then a tailor-made feedback control loop MATLAB script based on the FC data regulated the actuator profile for the arabinose addition. Within this script, FC events were gated based on forward scatter (FSC) to distinguish to single cells and clumps. Then GFP positive events were gated based on a 1000 threshold on GFP fluorescence channel area value (FL1-A; excitation 488 nm, emission 533 nm) for the first segregostat (as loose control policy) and with a 10000 threshold for the second segregostat (as tight control policy). Control policy was set for actuator activation when the fraction of GFP positive cells was measured to be below 50%. Actuator activation then exerted arabinose pulsing by the use of a digital control system comprising a peristaltic pump (Watson Marlow, 101 UR). Each segregostat experiment were carried out in duplicate. Data were exported as .fcs files and processed by a custom Python script (see next sections for a description of the data processing steps).

**Online flow cytometry (FC) analyses**

The online flow cytometer platform comprises three modules and can be operated in segregostat mode: (i) a conventional culture device, (ii) a physical interface for sampling and dilution containing peristaltic pumps and mixing chamber, and (iii) a detection device, in this work an Accuri C6 flow cytometer (BD Accuri, San Jose CA, USA) as used. Briefly, sample processing comprises the following steps: (i) sample acquisition, (ii) online FC analysis, (iii) dilution threshold and (iv) feedback control loop. The sample is entered and removed from the mixing chamber based on silicone tubing (internal diameter: 0.5 mm; external diameter: 1.6 mm, VWR, Belgium) and peristaltic pumps (400FD/A1 OEM-pump ~13 rpm and 290 rpm, Watson Marlow). Before and after each experiment, all the connection parts (tubing, pumps and mixed chamber) were cleaned with ethanol and rinsed with filtered PBS as set-up in a previous protocol [23].

**On-line FC data processing and accessibility**

Segregostat experiment involves the generation of a huge amount of FC data. A typical run of 100 hours led to 500 FC analyses, with a total of approximately $10^7$ cells processed. Before processing, these independent .fcs files were compacted in a dataframe (.pkl file extension) based on the Pandas package (https://pandas.pydata.org/) from Python. The codes for generating the .pkl files and the figures are available on a GitLab repository https://gitlab.uliege.be/F.Delvigne/paper_segregostat_arabinose. The code with name "From_fcs_to_pkl_and_statistics.py" can be used for generating most of the figures presented in this paper (i.e., figure 2 and 3). The code with name "Plots_from_pkl.py" can be used for generating the FC dotplots from a .pkl file and for each time points. These dotplots were further assembled into a single .avi movie file based on the ImageJ software [24] shown as Supplemental material (Movies S1 and S2). Raw .fcs data were deposited on FlowRepository and can be accessed by following the links below:

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

- Segregostat with low fluorescence threshold (1000 FU, loose control policy), replicate 1: http://flowrepository.org/id/RvFrEWS971a7Qf5j898Q69CxIFLXBEdUOHTfFHKnPjUH1WLbeV1cIIFzehipclaB
- Segregostat with low fluorescence threshold (1000 FU, loose control policy), replicate 2: http://flowrepository.org/id/RvFrxXrXzOyPLkV3VC511Hjd1oIQSKrmNA3YqtRA4SQfoAyCHWu0Uc1JAPclRU4l.
- Segregostat with high fluorescence threshold (10000 FU, tight control policy), replicate 1: http://flowrepository.org/id/RvFr1PDw2V2EI80kbxdXQRPlGcNWqEL4zGR3eLatTpLxifr8ckeRpleQ1CG8X8ps
- Segregostat with high fluorescence threshold (10000 FU, tight control policy), replicate 2: http://flowrepository.org/id/RvFrNtF1Biv2rENBiIn7SKnPjmWpYDL6KazQbnQbhzp1gdv9T5U7vt3L8gK9lfr3

**Metabolite analysis**

Supernatant glucose and arabinose concentrations were analyzed by high-performance liquid chromatography (Waters Acquity UPLC® H-Class System) using an ion-exchange Aminex HPX-87H column (7.8 × 300 mm, Bio-Rad Laboratories N.V.). The analysis was carried out with an isocratic flow rate of 0.6 mL $min^{-1}$ for 25 min at 50°C. The mobile phase was composed of an aqueous solution of 5 mM $H_2SO_4$. Elution profiles were monitored through a Waters Acquity® Refractive Index Detector (RID) (Waters, Zellik, Belgium). Glucose and arabinose standard solutions (Sigma-Aldrich, Overijse, Belgium) were used to determine the retention times and construct calibration curves.

**Transcripts isolation and quantification**

Samples were taken at different times during fermentation. Samples were immediately frozen and kept at − 80°C. The total RNA was extracted using the NucleoSpin RNA Mini kit ( Macherey-Nagel, Germany) following the manufacturer's protocols. qPCRs measurements were conducted in an Applied Biosystems Step One Real-Time PCR system (Life Technologies, Grand Island, NY, USA). The thermal cycling protocol was performed as follows: initial denaturation for 15 min at 95°C, 40 cycles of 95°C for 10 sec, at 65°C for 15 sec, and at 72°C for 20 sec, and a final dissociation cycle at 95°C for 30 sec, 65°C. qPCR data was normalized to the housing keeping gene *ihfB,* and the sample taken from the end of batch phase was used as a reference [25]. The 2(-Delta Delta C(T)) method was used to analyze the relative changes in gene expression [26].

**Results**

**Chemostat with glucose/arabinose co-feeding leads to massive plasmid loss and population instability**

In a first set of experiments, classical chemostat cultivation has been run at a dilution rate of 0.5 $h^{-1}$. Combinations of glucose and arabinose were considered as co-feeding strategy for activating the arabinose operon during the continuous cultivation phase, and the population heterogeneity was determined based on the signal provided by the GFP reporter system expressed by the means of the $P_{BAD}$ promoter (Figure 1A). Flow cytometry analysis revealed that the co-feeding strategy with arabinose and glucose led to an unstable GFP expression profile (Figure 1B). Indeed, bistability, traduced by the simultaneous presence of GFP negative and GFP positive subpopulations was observed right after the induction upon activation of the co-feeding regime. Bistability, also called “all-or-none” response has previously been observed under similar operating conditions with

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

arabinose/glucose [27] and lactose/glucose co-feeding [28]. On the long-term, the cell population was found to be progressively transitioning to the low expression state, leaving only a few GFP positive (less than 10%) cells after 25 residence times. This behavior was observed for different arabinose/glucose ratios (Figure 1C and 1D). On the short-term, it seems that under continuous co-feeding regime, phenotypic cell switching may lead to the observed bistable behavior with the simultaneous occurrence of GFP positive and GFP negative cells. On the long term, GFP positive cells were decreasing due to the loss of the plasmid carrying the $P_{BAD}$::GFP reporter system (Figure 1E displays the fraction of cells exhibiting plasmid lost at the end of the chemostat cultivation), probably due to a selective pressure associated with the cost of GFP synthesis.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

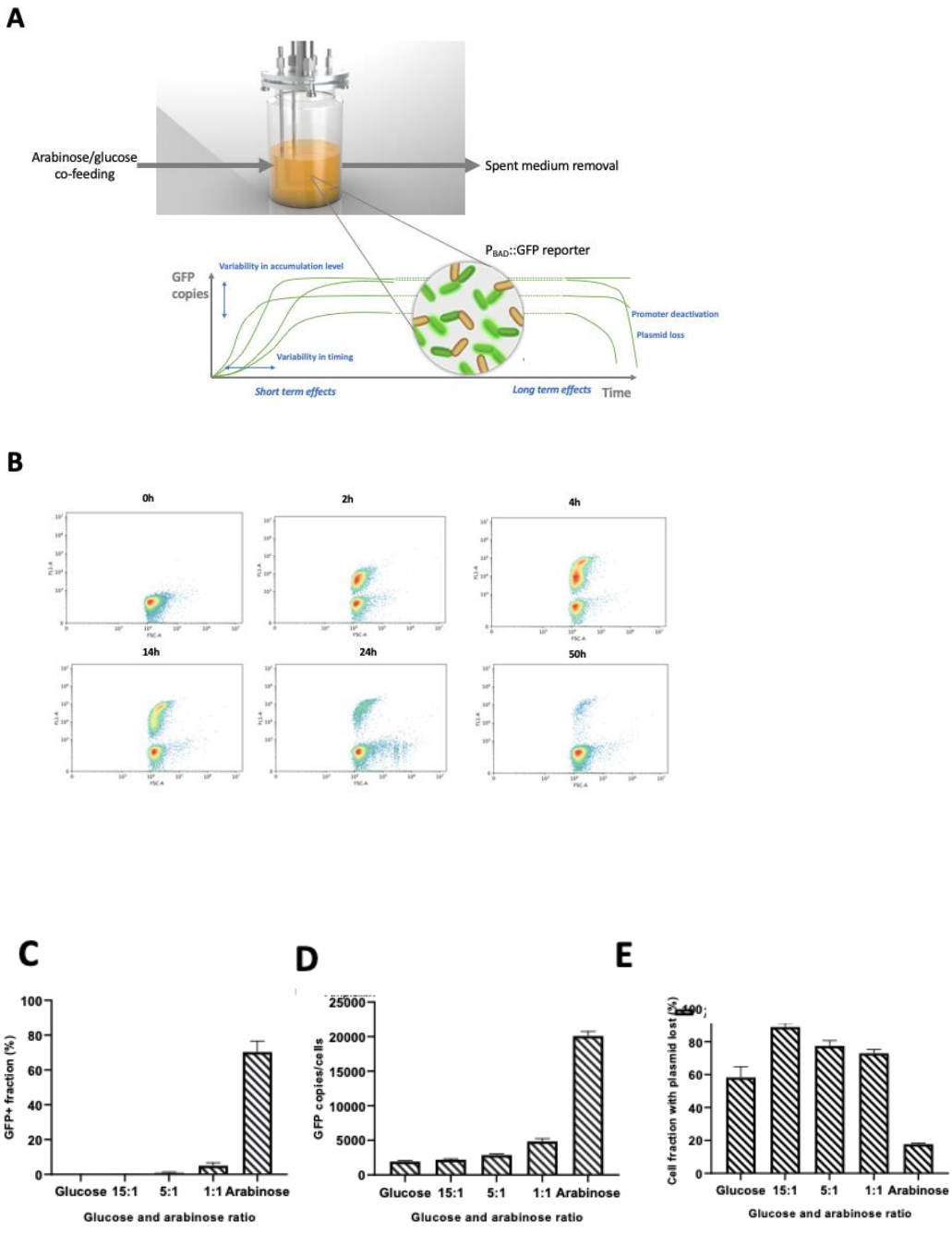



**Figure 1**: **A** Sketch of the chemostat set-up used for the glucose-arabinose co-feeding experiments. **B** Flow cytometry monitoring (x-axis: forward scatter signal; y-axis : green fluorescence signal accounting for the accumulation of GFP inside cells) of a chemostat with glucose-arabinose co-feeding (dilution rate 0.5 h-1; ration glucose-arabinose 5:1). **C** Evolution of the GFP positive cell fraction, the mean GFP copies/cell and the cell fraction with plasmid lost in function of the glucose/arabinose ratio.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Automated addition of arabinose pulses based on phenotypic switching leads to population stability and robust gene expression profiles**

In order to overcome both phenotypic heterogeneity and genetic instability, a cell-machine interface relying on a $P_{BAD}$::*GFP* reporter and on-line FC was designed to automatically adjust the addition of glucose and arabinose during continuous cultivation based on the phenotypic switching capabilities of microbial cells [18]. It has been shown previously that microbial cells are able to switch stochastically from one phenotypic state to another in response to environmental perturbations [15]. In the present work, this specific behavioral feature was exploited for optimizing the GFP expression profiles for the whole population. The cell-machine interface set-up is inspired on the segregostat concept that has been previously used for minimizing phenotypic heterogeneity related to the membrane permeabilization process during *E. coli* continuous cultivation [18][29]. In its current version, the segregostat comprises a continuous cultivation device connected to an on-line flow cytometry device allowing sampling the state of the population at regular time intervals. This device then allows optimizing the glucose-arabinose switches based on the actual phenotypic switching dynamics of the microbial population. For this purpose, it is critical to explore and characterize the phenotypic switching dynamics. To accomplish the latter, different control policies, deriving in different pulsing regimes can be constructed. In the first segregostat experiments, low fluorescence threshold (1000 RFU), corresponding to the upper limit of the low state (Figure 2A) was chosen. When at least 50% of the population has crossed this threshold, then an environmental fluctuation, here represented by either a glucose or arabinose pulse, was triggered. According to this control policy, it was possible to maintain the whole population fluctuating between the high and the low state (Figure 2B). Additionally, no bistability behavior was observed, the whole population following the glucose to arabinose transitions by dynamically adapting gene expression (Supplemental Material, Movie S1). Additionally, gene expression levels were monitored the cultivation by RT-qPCR, and residual arabinose and glucose was determined by HPLC (Supplemental Material, Figure S1). Furthermore, these results pointed out that it is also possible to maintain GFP expressing cells over long periods of time with no plasmid loss, as GFP expressing cells were found within the culture until the end of the experiment (68 generations).

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

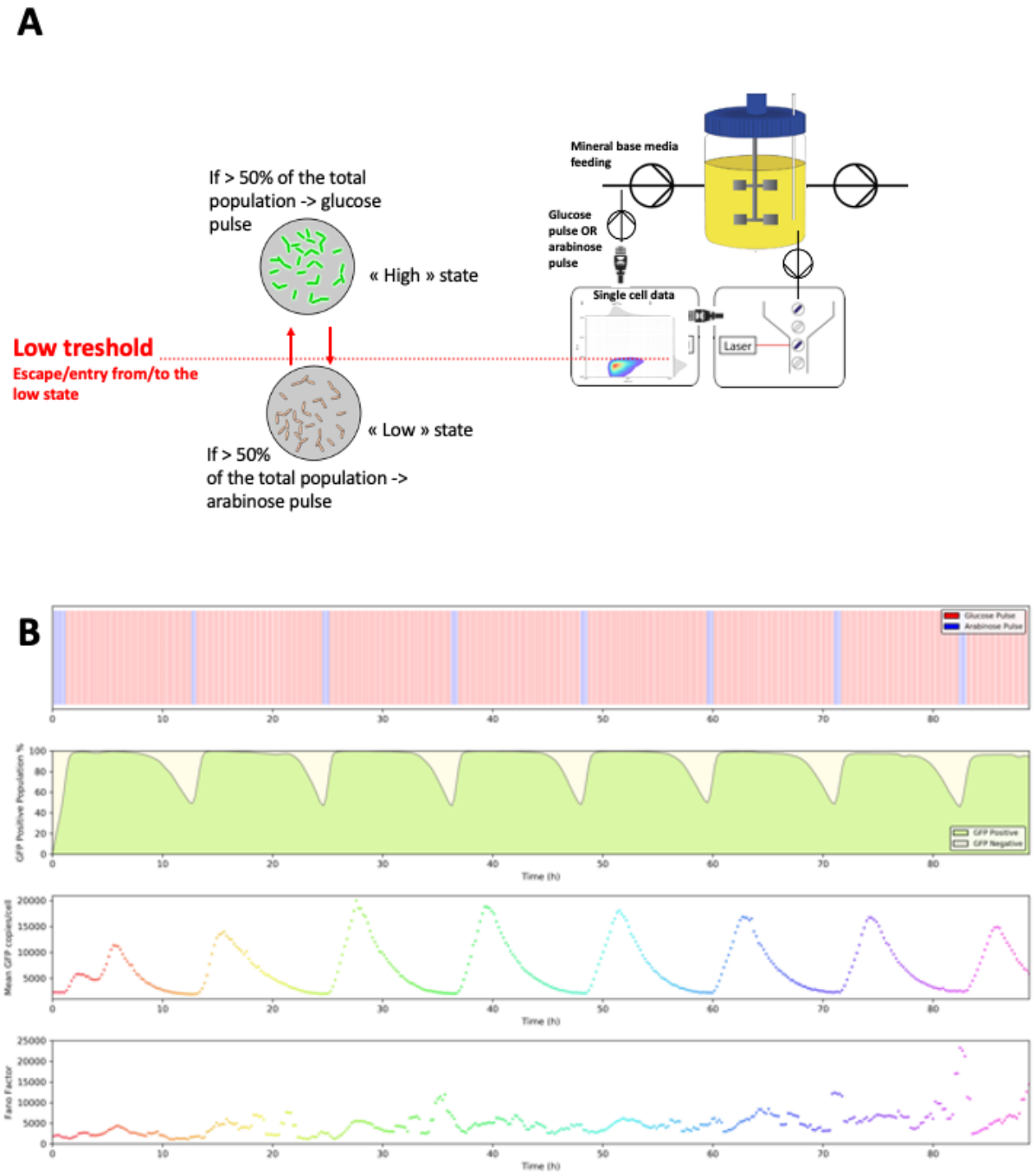



**Figure 2**: **A** Adjustment of glucose and arabinose pulsing based on the phenotypic switching ability of cells. In this case, a low threshold has been selected (loose control procedure). **B** Population dynamics in function of the automated addition of glucose/arabinose pulses. Mean GFP copies per cells and Fano factor (ratio between the variance and the mean of GFP distribution in the GFP positive fraction of cells) have been computed by considering the GFP positive fraction only.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Forcing the entry into the high state leads to long-term population instabilities**

It has been shown that it is possible to avoid bistability and enhance population stability by modulating the pulsing frequencies of arabinose and glucose during continuous cultivation. However, by considering a low threshold, relaxed oscillations at the level of GFP accumulation profile were observed (mean GFP copies/cell) (Figure 2B). In order to optimize GFP accumulation, a high threshold was considered for running a second segregostat as tight control policy (Figure 3A). This threshold was set on the lower limit of the high state (10000 RFU) and is thus exerts more control pressure on the GFP level compared to the previously loose control experiment (the first segregostat with threshold at 1000 RFU), constraining the system behavior to higher GFP production. Based on this new threshold, GFP copies/cell oscillates around a higher value by comparison to what was obtained based on the loose control policy (Figure 3B). It is important to notice that no bistability was observed during the first 40 hours of the continuous culture (Supplemental Material, Movie S2). However, the GFP oscillation pattern was found to be less regular than the one observed for the lower threshold and, after 40 hours, the GFP positive subpopulation tends to decrease due to plasmid loss (Figure 3B).

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

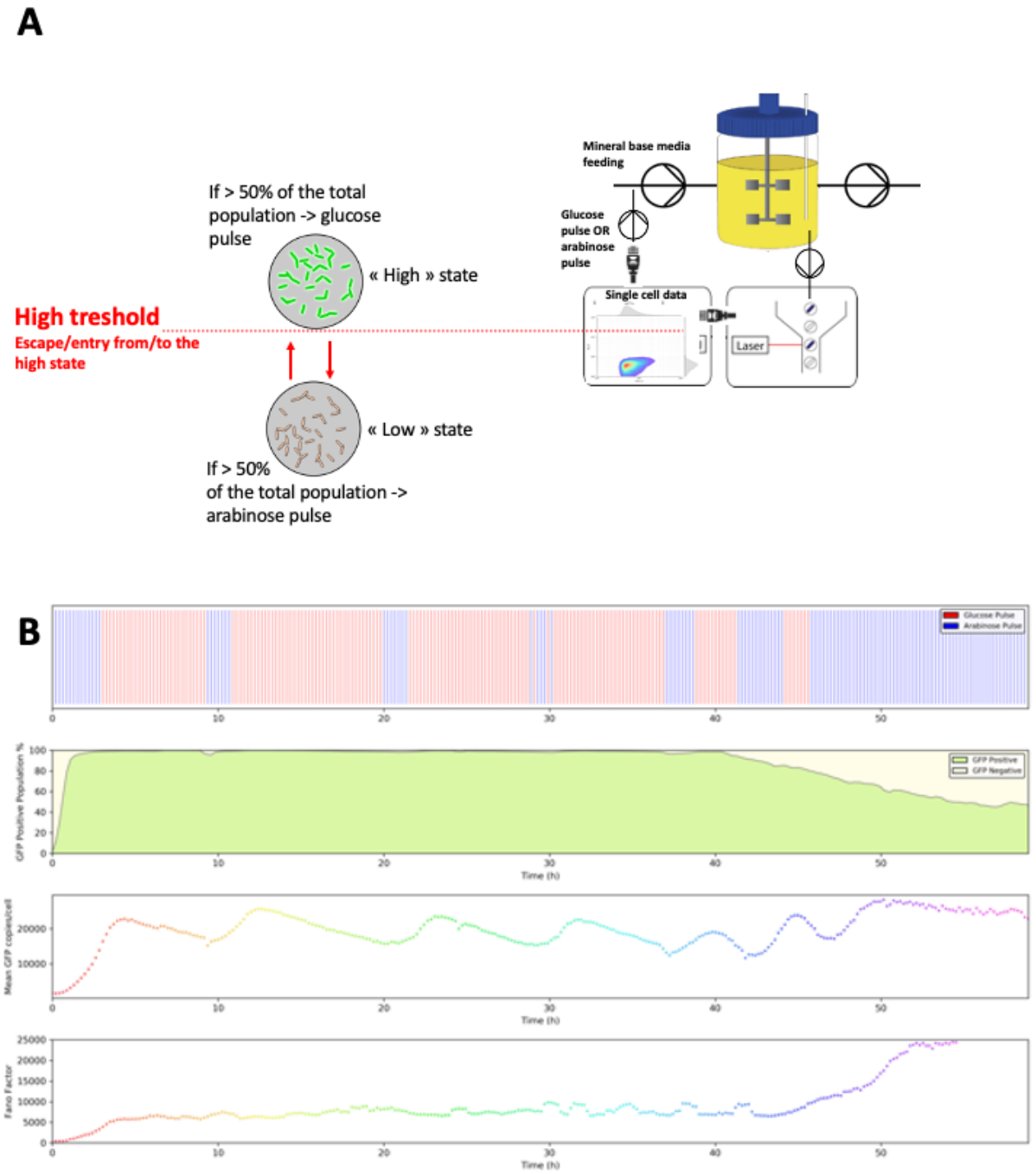



**Figure 3**: **A** Adjustment of glucose and arabinose pulsing based on the phenotypic switching ability of cells. In this case a high threshold has been selected (tight control procedure). **B** Population dynamics in function of the automated addition of glucose/arabinose pulses. Mean GFP copies per cells and Fano factor (ratio between the variance and the mean of GFP distribution in the GFP positive fraction of cells) have been computed by considering the GFP positive fraction only.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Discussion

The use of the segregostat platform to analyze the dynamic behaviour of *E. coli* populations between glucose and arabinose consuming phenotypes related to environmental perturbations, shed new light on the role of phenotypic heterogeneity on the dynamics and stability of microbial populations. In comparison to bistability in GFP distribution observed in chemostat with glucose/arabinose co-feeding, cultivation performed in the segregostat mode led to an unimodal GFP distribution with the microbial population smoothly transiting between the high (i.e., high GFP content driven by $P_{BAD}$ promoter) and low induction states (See movies S1 and S2 supplied in Supplemental Material for an animated evolution of GFP distribution with time). This led to a fully predictable oscillating pattern for GFP expression that was stable during the whole continuous cultivation comprising several generations (68 generations). It is important to acknowledge that further process optimization is needed for the application of these platform and pulsating control regimes to real case scenarios, such as damping the oscillations in GFP content in order to maximize protein production for a given process, cataloguing and characterize different induction promoter systems with different activation profiles, and construction of more responsive or dynamical control policies. However, the second segregostat experiment based on high threshold pointed out that there is a tipping point below which population collapse. In this case the collapse is mainly due to plasmid loss and should be alleviated by using, for example, more stable expression vectors e.g., kanamycin-based vectors.

Besides its utilization as an effective population control device, segregostat data can also be used for understanding the dynamics of key biological processes involved in phenotypic switching dynamics and microbial population stability. Biological noise is known to play a significant role in the process of adaptation to environmental perturbations [15][16][30], such as the switch from glucose to arabinose considered in this study. As stated in the introduction section, biological noise is quite well characterized for a single gene [31], but the picture becomes more complex in the case of complex GRN . In this case, it is particularly difficult to predict how noise propagate from a gene to another in the regulatory network [32]. The arabinose operon is under the control of a feedforward motif, allowing the cell to tightly regulate the expression of the genes needed for efficient arabinose utilization [33]. This tight regulation mechanism has been selected in order to mitigate the metabolic burden associated with the expression of the *araBADEFGH* genes by optimally inducing this operon when glucose is absent and arabinose becomes available in the environment [34]. It is also known that, under specific environmental conditions, this motif can lead to bistability with cells inducing the operon whereas others in the same population remains at a low state. This strategy could be interpreted as a bet-hedging mechanisms where microbial population effectively exploit noise in order to optimize its fitness in uncertain environmental conditions [14]. Bistability has been observed in our case during chemostat experiment with glucose-arabinose co-feeding. Interestingly, bistability can be avoided when pulsing glucose and arabinose at a particular frequency, which is in accordance with the phenotypic switching ability of cells. In this case indeed, only a unimodal GFP expressing population was observed. However, this population also exhibited noise but with a mean GFP expression level oscillating at a frequency corresponding to the environmental fluctuations. This specific mode of diversification, called dynamic heterogeneity, has been previously predicted based on mathematical modelling by Thattai and van Oudenaarden [15]. The results presented in this work point out that cell physiology in segregostat is significantly different from the one that is observed in chemostat, involving a fundamentally different modes of diversification dynamics. Further work is needed for characterizing this physiology and the underlying mechanisms behind dynamic heterogeneity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Only a few studies are actually dealing with the response of cell population to repetitive stimuli and new insights are needed at this level. The main issue is that the whole spectrum of amplitude-frequency related to environmental perturbation is too broad to be screened in a time efficient way. One of the remarkable features of the segregostat is that the environmental perturbation is triggered by the capacity of the microbial population to split into different phenotypic states, leading to the systematic determination of the optimal rate of environmental perturbations to be tested among the whole spectrum of amplitude/frequency.

In conclusion, the segregostat can be used for automatically adjusting the rate of environmental perturbation to the intrinsic dynamics of the GRN under consideration without any prior knowledge about the architecture and dynamical feature of this GRN. This specific feature could be exploited in the future for ensuring the stability of microbial population in continuous cultivation, or to make the acquisition of fundamental knowledge about the role of biological noise in microbial response strategies to environmental fluctuations. The former perspective is in frame with a developing field of research i.e., cybergenetics. Cybergenetics uses computer interface for controlling in real-time biological processes [35][36]. This approach has been demonstrated to be very useful for controlling a variety of GRNs, including the antithetic motif [37][38] and the toggle switch [39].

**Acknowledgements**

TMN is the recipient of a VIED PhD grant provided by the Vietnamese government. JAM and BZ are the recipient of a postdoctoral grant provided through two Era-cobiotech projects, YBISCUS and ComRaDes respectively (funding through the European Union H2020 horizon program and Service Public de Wallonie (SPW)).

**References**

1. Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of metabolism and growth at the single-cell level. Nature. 2014;514:376–9.

2. Nikolic N, Barner T, Ackermann M. Analysis of fluorescent reporters indicates heterogeneity in glucose uptake and utilization in clonal bacterial populations. BMC Microbiol. 2013;13:258.

3. Nikolic N, Schreiber F, Dal Co A, Kiviet DJ, Bergmiller T, Littmann S, et al. Cell-to-cell variation and specialization in sugar metabolism in clonal bacterial populations. PLoS Genet. 2017;13:e1007122.

4. Delvigne F, Zune Q, Lara AR, Al-Soud W, Sorensen SJ. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol. 2014;32:608–16.

5. Binder D, Drepper T, Jaeger K-E, Delvigne F, Wiechert W, Kohlheyer D, et al. Homogenizing bacterial cell factories: Analysis and engineering of phenotypic heterogeneity. Metab Eng. 2017;42:145–56.

6. Delvigne F, Zacchetti B, Fickers P, Fifani B, Roulling F, Lefebvre C, et al. Improving control in microbial cell factories: from single cell to large-scale bioproduction. FEMS Microbiol Lett. 2018;

7. Perkins TJ, Swain PS. Strategies for cellular decision-making. Mol Syst Biol. 2009;5:326.

8. Schreiber F, Littmann S, Lavik G, Escrig S, Meibom A, Kuypers MMM, et al. Phenotypic heterogeneity driven by nutrient limitation promotes growth in fluctuating environments. Nat

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Microbiol. 2016;1:16055.

9. Kotte O, Volkmer B, Radzikowski JL, Heinemann M. Phenotypic bistability in Escherichia coli's central carbon metabolism. Mol Syst Biol. 2014;10.

10. 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:e61686.

11. Ackermann M. A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol. 2015;13:497–508.

12. Solopova A, Gestel J, Weissing FJ, Bachmann H, Teusink B, Kok J, et al. Bet-hedging during bacterial diauxic shift. Proc Natl Acad Sci U S A. 2014;111:7427–32.

13. van Boxtel C, van Heerden JH, Nordholt N, Schmidt P, Bruggeman FJ. Taking chances and making mistakes: non-genetic phenotypic heterogeneity and its consequences for surviving in dynamic environments. J R Soc Interface. 2017;14.

14. Veening J-W, Smits WK, Kuipers OP. Bistability, epigenetics, and bet-hedging in bacteria. Annu Rev Microbiol. 2008;62:193–210.

15. Thattai M, van Oudenaarden A. Stochastic gene expression in fluctuating environments. Genetics. 2004;167:523–30.

16. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature. 2010;467:167–73.

17. van Heerden JH, Kempe H, Doerr A, Maarleveld T, Nordholt N, Bruggeman FJ. Statistics and simulation of growth of single bacterial cells: illustrations with B. subtilis and E. coli. Sci Rep. 2017;7:16094.

18. Sassi H, Nguyen TM, Telek S, Gosset G, Grunberger A, Delvigne F. Segregostat: a novel concept to control phenotypic diversification dynamics on the example of Gram-negative bacteria. Microb Biotechnol. United States; 2019;12:1064–75.

19. Gresham D, Hong J. The functional basis of adaptive evolution in chemostats. FEMS Microbiol Rev. 2015;39:2–16.

20. Rugbjerg P, Myling-Petersen N, Porse A, Sarup-Lytzen K, Sommer MOA. Diverse genetic error modes constrain large-scale bio-based production. Nat Commun. 2018;9:787.

21. Lee TS, Krupa RA, Zhang F, Hajimorad M, Holtz WJ, Prasad N, et al. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. J Biol Eng. 2011;5:12.

22. Xu B, Jahic M, Enfors S-O. Modeling of Overflow Metabolism in Batch and Fed-Batch Cultures of Escherichiacoli. Biotechnol Prog. 1999;15:81–90.

23. Brognaux A, Han S, Sorensen SJ, Lebeau F, Thonart P, Delvigne F. A low-cost, multiplexable, automated flow cytometry procedure for the characterization of microbial stress dynamics in bioreactors. Microb Cell Factories. 2013;12.

24. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, et al. ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics. 2017;18:529.

25. Flores N, Flores S, Escalante A, de Anda R, Leal L, Malpica R, et al. Adaptation for fast growth on glucose by differential expression of central carbon metabolism and gal regulon genes in an Escherichia coli strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Metab Eng. 2005;7:70–87.

26. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods. 2001;25:402–8.

27. Sagmeister P, Schimek C, Meitz A, Herwig C, Spadiut O. Tunable recombinant protein expression with E. coli in a mixed-feed environment. Appl Microbiol Biotechnol. Germany; 2014;98:2937–45.

28. Wurm DJ, Hausjell J, Ulonska S, Herwig C, Spadiut O. Mechanistic platform knowledge of concomitant sugar uptake in Escherichia coli BL21(DE3) strains. Sci Rep. 2017;7:45072.

29. Acar M, Mettetal JT, van Oudenaarden A. Stochastic switching as a survival strategy in fluctuating environments. Nat Genet. 2008;40:471–5.

30. Ackermann M. A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol. 2015;13:497–508.

31. Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci U S A. 2002;99:12795–800.

32. Pedraza JM, van Oudenaarden A. Noise propagation in gene networks. Science. United States; 2005;307:1965–9.

33. Dekel E, Mangan S, Alon U. Environmental selection of the feed-forward loop circuit in gene-regulation networks. Phys Biol. England; 2005;2:81–8.

34. Alon U. An introduction to systems biology. 2nd ed. CRC Press; 2020.

35. Banderas A, Le Bec M, Cordier C, Hersen P. Autonomous and Assisted Control for Synthetic Microbiology. Int J Mol Sci. 2020;21.

36. Carrasco-López C, García-Echauri SA, Kichuk T, Avalos JL. Optogenetics and biosensors set the stage for metabolic cybergenetics. Curr Opin Biotechnol. England; 2020;65:296–309.

37. Milias-Argeitis A, Rullan M, Aoki SK, Buchmann P, Khammash M. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. Nat Commun. 2016;7:12546.

38. Briat C, Khammash M. Perfect Adaptation and Optimal Equilibrium Productivity in a Simple Microbial Biofuel Metabolic Pathway Using Dynamic Integral Control. ACS Synth Biol. 2018;7:419–31.

39. Lugagne J-B, Sosa Carrillo S, Kirch M, Kohler A, Batt G, Hersen P. Balancing a genetic toggle switch by real-time feedback control and periodic forcing. Nat Commun. 2017;8:1671.

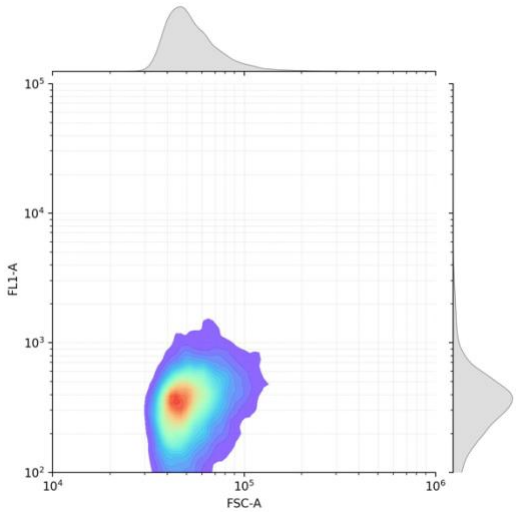
bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Supplemental material**

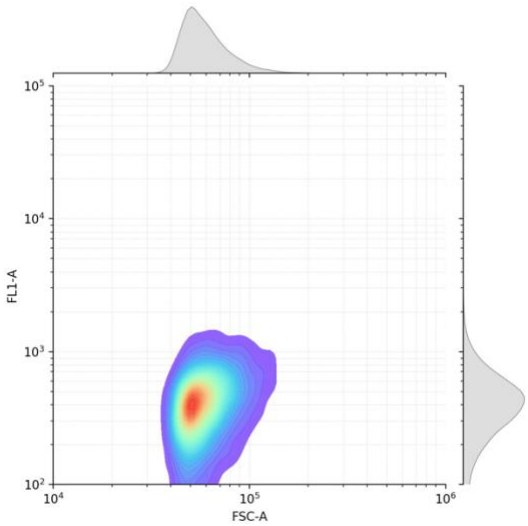
This supplemental material is intended to give more information on the experiments carried out based on the segregostat cultivation device. This system has been designed in another study where it has been used for controlling the fraction of outer-membrane permeable cells in continuous culture for *Escherichia coli* MG1655 and *Pseudomonas putida* KT2440 (Sassi et al., 2019). Briefly, the segregostat comprises a continuous cultivation device (in our case, a stirred bioreactor with a working volume of 1L) where environmental perturbations can be automatically generated based on a controlling device triggered based on an on-line flow cytometer (FC). In this study, the system has been adapted for controlling the switch from glucose to arabinose utilization by *E. coli*. As described before, switching to alternative carbon source is based on a tightly regulated gene regulatory network (GRN) called feedforward motif (Mangan and Alon, 2003) and is prone to stochasticity (Megerle et al., 2008). This stochasticity can induce strong cell-to-cell variability, and in some extreme cases, lead to a phenomenon called bistability. Cell population exhibiting bistability comprises two subpopulations of cells with low and high level of expression genes related to arabinose utilization. This phenomenon has been previously observed in continuous culture with glucose and arabinose co-feeding (Sagmeister et al., 2014).

The segregostat can be used for avoiding the occurrence of bistability and set the population into a totally different switching mode. According to the literature, two switching mode can be observed, i.e. either stochastic switching or responsive switching (Thattai and van Oudenaarden, 2004)(Kussell, 2005). The stochastic switching mode of diversification implies that several subpopulations of cells can be observed simultaneously according to bistability/multistability and is known to increase the resilience of the whole population through a mechanism called bet-hedging (for more explanation about bet-hedging, please have a look at (Veening et al., 2008)). This is typically what happen when arabinose and glucose are simultaneously used as co-feeding in chemostat. However, when using the segregostat (two types of control policies have been used in this work – see Figures 2 and 3 in the main text), it is possible to generate an alternance of glucose and arabinose environment stimulating the responsive switching mode of diversification. For an easy visualization of this mode of diversification, on-line flow cytometry movies have been generated (Movie S1 and Movie S2) and can be used for the visualization of the responsive switching mode. Indeed, following pulses of arabinose, all the cells in the population are switching to the high GFP content state. Of course, this phenomenon occurs with some cell-to-cell variability at the level of the GFP content, but it can be observed that globally the population coordinates itself for following the glucose to arabinose transitions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.




**Movie S1 (click on the figure for animation. For the .pdf version of this document, the movies can be uploaded separately)**: movie involving on-line flow cytometry analyses of cells cultivated in segregostat with loose control policy (low fluorescence threshold set at 1000 FU). x-axis represents forward scatter (FSC) and y-axis represent the FL1 channel where GFP fluorescence was recorded.




**Movie S2 (click on the figure for animation. For the .pdf version of this document, the movies can be uploaded separately)**: movie involving on-line flow cytometry analyses of cells cultivated in segregostat with tight control policy (high fluorescence threshold set at 10,000 FU). x-axis represents forward scatter (FSC) and y-axis represent the FL1 channel where GFP fluorescence was recorded.

In segregostat, cell population is thus exposed to an alternance of glucose and/or arabinose rich environment, this alternance being generated by the phenotypic diversification capacity of the population itself (determined based on automated flow cytometry). Arabinose and glucose concentration have been monitored, as well as the expression level of one of the key gene belonging to the arabinose operon i.e., *araB* (Figure 1C). These data point out a very logical trend, the decrease of araB expression level being correlated to the disappearance of arabinose from the segregostat, under the two control policies used (loose and tight control). The main difference between the two control policies is related to plasmid stability, tight control leading to a more significant loss of $P_{BAD}$::*gfp* plasmid (Figure 2S).

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

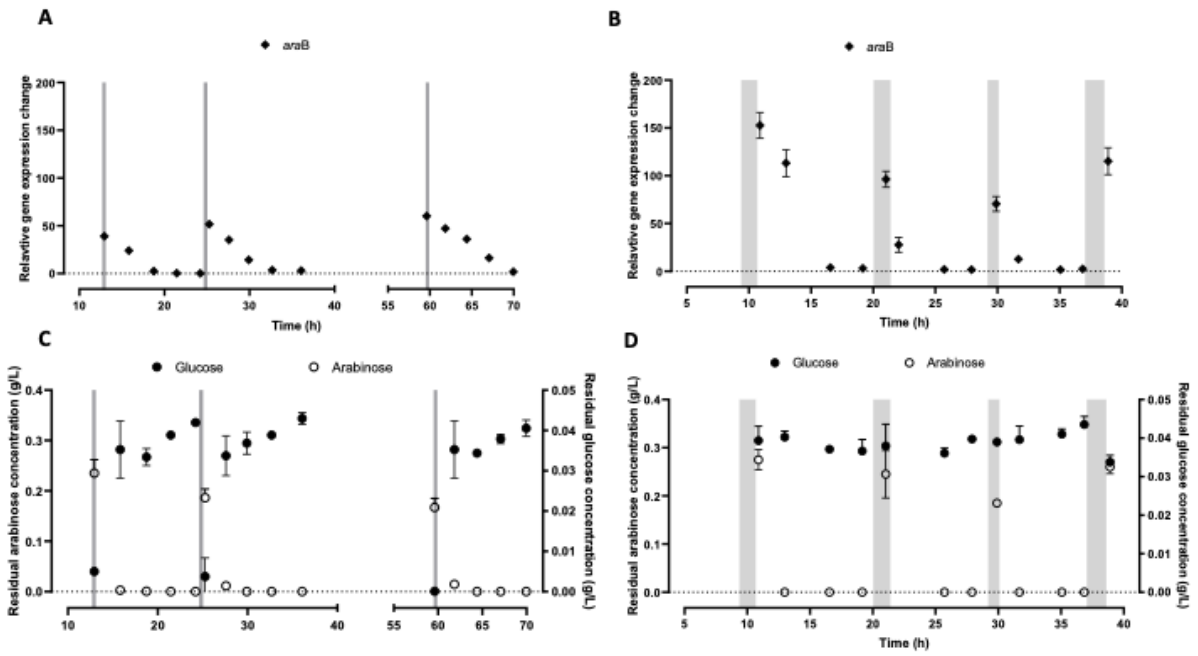



**Figure S1**: Dynamics of *araB* expression level (◆) (monitored by RT-qPCR) and residual glucose (●) and arabinose (○) concentrations for the segregostat with loose (panels A and C) and tight (panels B and D) control respectively. Grey areas correspond to the arabinose pulsing phases. Results were expressed the mean of three technical replicates.

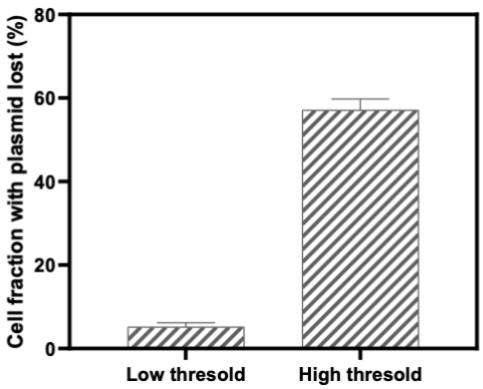



**Figure S2**: Comparison of the cell fraction exhibiting plasmid loss for segregostat cultivation operated based on low (loose control procedure) and high (tight control procedure) fluorescence threshold. Samples were taken at the end of experiments. Upon dilution in sterile PBS, cell suspensions were spread on LB agar plate with and without ampicillin. Plates were incubated overnight at 37°C and the fraction of ampicillin-sensitive cells was determined. Results were the mean of three replicates (standard deviation shown in the figure).

bioRxiv preprint doi: https://doi.org/10.1101/2021.01.13.426484; this version posted January 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## References

Kussell, E. (2005). Phenotypic diversity, population growth and information in fluctuating environments. Science *309*, 2075–2078.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. U. S. A. *100*, 11980–11985.

Megerle, J.A., Fritz, G., Gerland, U., Jung, K., and Rädler, J.O. (2008). Timing and dynamics of single cell gene expression in the arabinose utilization system. Biophys. J. *95*, 2103–2115.

Sagmeister, P., Schimek, C., Meitz, A., Herwig, C., and Spadiut, O. (2014). Tunable recombinant protein expression with E. coli in a mixed-feed environment. Appl. Microbiol. Biotechnol. *98*, 2937–2945.

Sassi, H., Nguyen, T.M., Telek, S., Gosset, G., Grünberger, A., and Delvigne, F. (2019). Segregostat: a novel concept to control phenotypic diversification dynamics on the example of Gram-negative bacteria. Microb. Biotechnol. *12*, 1064–1075.

Thattai, M., and van Oudenaarden, A. (2004). Stochastic gene expression in fluctuating environments. Genetics *167*, 523–530.

Veening, J.-W., Smits, W.K., and Kuipers, O.P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. *62*, 193–210.