Thesis for the degree
Master of Science

Submitted to the Scientific Council of the
Weizmann Institute of Science
Rehovot, Israel

עבודת גמר (תזה) לתואר
מוסמך למדעים

מוגשת למועצה המדעית של
מכון ויצמן למדע
רחובות, ישראל

By
Nitai Steinberg

מאת
ניתאי שטיינברג

גידור הימורים בתגובה המותנית-אבולוציונית של *Escherichia coli* לסוכרים

Bet-hedging in the evolutionarily-conditioned response of *Escherichia coli* towards sugars

Advisor:
Prof. Yitzhak Pilpel

מנחה:
פרופ' יצחק פלפל

January 2012

טבת תשע"ב

1. Abstract

Recently, *Mitchell et al.* [1] demonstrated a microbial evolutionary analogy to the classical conditioning of Pavlov. In their work they showed how temporally connected changes in microorganisms' environment have been captured in gene expression programs during the course of evolution. For instance, they demonstrated how *Escherichia coli*, when exposed to lactose, partially activates its maltose operons, as in the gastrointestinal tract lactose could be used as a predictive signal to maltose. However, a useful temporal connection might not be always dependable, and from time to time the signal could appear without the following change. If such instances are frequent, microorganism could use the predictive signal in a non-deterministic manner, allowing a portion of the population to remain inactive, and thus lowering the cost in case that the anticipated change will not come, as a form of bet-hedging strategy. In my study, I sought to elucidate whether such strategy could be manifested in the evolutionarily-conditioned response of *E. coli* towards sugar. In order to do so, I examined the maltose operons activation in response to lactose and maltose at the single cell level using flow cytometry and fluorescent protein reporters for promoter activity. I saw that the distribution of activation of maltose promoters is much wider in the response to lactose compared to maltose. In addition, in certain conditions, the distribution becomes highly skewed to the right in response to lactose, as a portion of the population activates the maltose promoter to high levels. These results might imply a strategy combining evolutionary conditioning and bet-hedging, such that *E. coli* utilizes lactose as a predictive signal to maltose in a non-deterministic manner, in accord to the unpredictability in maltose arrival.

2. Introduction

2.1 Evolutionary conditioning

Most organisms are exposed to frequent changes in their environment. The changes are, in many cases, stochastic and unpredictable. However, some changes might be temporally connected to others. In those instances, it is possible to imagine that organisms will use the temporal connections between changes in order to prepare themselves for the future. To better illustrate this, suppose S1 and S2 are two changes in the environment, while S1 always precedes S2, and preparation for S2 in advance is beneficial. S2 could be, for instance, a stressful change that could be better handled with proper preparation, or, alternatively, the appearance of a valuable nutrient that requires special mechanism in order to be metabolized. In this scenario, the ability to use S1 as a signal for an early preparation for S2 may be advantageous. Thus, if this trait appears in the population it might be positively selected during evolution.

During the last few years, examples for such predictive responses were shown in several microorganisms. In *Saccharomyces cerevisiae*, an asymmetrical cross-protection was shown in which pre-exposure to heat shock or ethanol gives higher survival rates at oxidative stress, but not when the order of stresses is reversed [1]. This asymmetrical cross-protection was suggested to capture the order of the environmental changes which *S. cerevisiae* has been repeatedly exposed to during its evolution, where heat shock and ethanol stress precede an oxidative stress during wine production and the switch from fermentation and respiration [2]. In *Escherichia coli*, the transcriptional responses to elevation in temperature and oxygen limitation were found to be highly correlated [3]. It was suggested that this correlation reflects the coupling between those two conditions when *E. coli* enters its host. In *Vibrio cholerae*, genes that are induced during the late stages of infection were found to be important for the survival of the pathogen in the external environment [4]. Thus, while still in the host, *V. cholerae* activates genes that will help it to survive after the exit. In *Candida albicans*, exposure to glucose induces the oxidative and cationic stress responses [5]. It may imply preparation for the host immune response when the pathogen enters the bloodstream.

As suggested by *Mitchell et al.* [1,6], the positive selection of the ability to prepare in advance to an environmental change based on early predictive signal, can be seen as an evolutionary analogy to the behavioral phenomenon of classical conditioning, first

demonstrated by Ivan Pavlov [7]. It will be referred here shortly as 'Evolutionary Conditioning'. *Mitchell et al.* farther applied this analogy for what Pavlov named 'extinction', a process in which repeated exposures to S1 without the succeeding S2 weaken the in-advance-preparation to S2 during exposure to S1. They suggested that such extinction could happen also for evolutionary conditioning. Providing the cost of preparation for S2 is high, such that wrong prediction could be detrimental, if the prediction usually results in false alarm this trait will be selected against.

In addition to the *S. cerevisiae* example mentioned above, *Mitchell et al.* also demonstrated another example of evolutionary conditioning in *E. coli*.

During its passage thorough the gastro-intestinal tract, *E. coli* is exposed to various nutrients in specific locations (Figure 1a). In the proximal part of the intestine the concentration of lactose is high and the concentration of maltose is low, while in the distal part maltose is high and lactose is low [8,9]. Using strains of the *E. coli* promoter-fused GFP library [10,11], *Mitchell et al.* showed that exposure to lactose activates the maltose metabolism operons, but exposure to maltose does not activate the lactose operon (Figure 1b). They also showed that pre-exposure to lactose gives fitness advantage on maltose. As in the *S. cerevisiae* example, the hierarchy of the sugar metabolism operon activation captures the order of the sugars to which *E. coli* was exposed during its evolution. In addition, they demonstrated, as an example for extinction, that constant exposure to lactose, without the succeeding maltose, weakens the maltose response and, correspondingly, the fitness advantage on maltose of bacteria that were pre-exposed to lactose. *Tagkopoulos et al.* [3] showed a similar example of extinction in the *E. coli* transcriptional responses to temperature elevation and oxygen limitation mentioned above: when *E. coli* was evolved in an environment in which temperature elevation was decoupled from oxygen limitation, the correlation between the transcriptional responses disappeared.

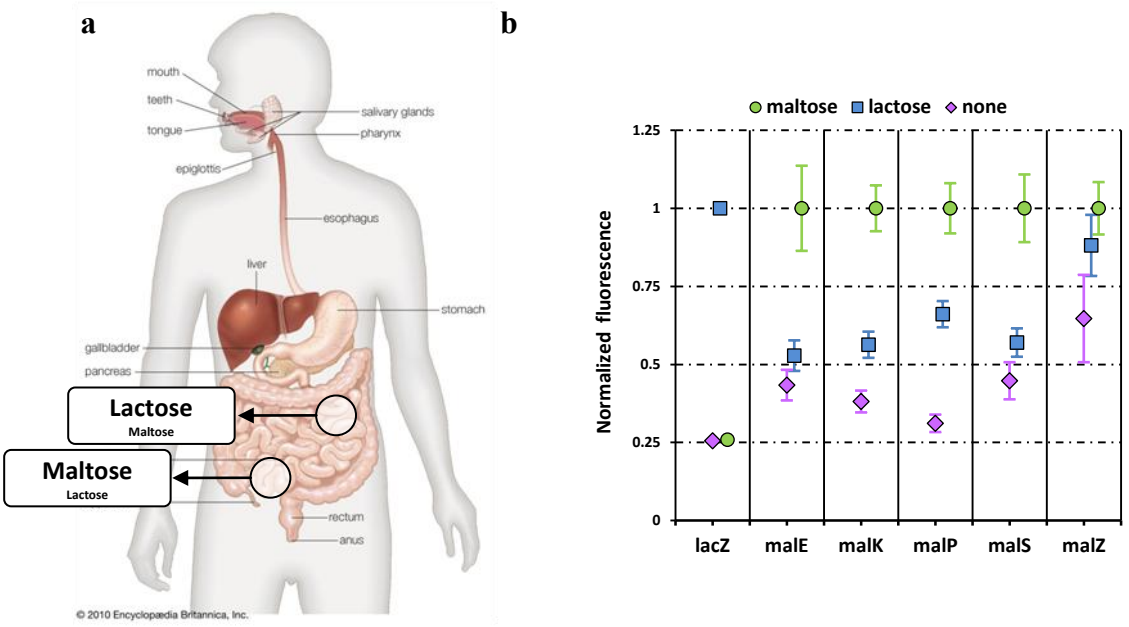


Figure 1 – Evolutionary conditioning in *E. coli* response towards lactose
(**a**) Adapted from [8]. The human gastrointestinal tract, the proximal and the distal part of the small intestine are marked with bright circles. The presumed relative concentrations of lactose and maltose in each of the location are shown.
(**b**) Adapted from [1]. Normalized GFP fluorescence level of strains from the promoter-fused GFP library in exposure to lactose and maltose. The normalization procedure was as follows [1]: (1) Subtraction of the media fluorescence and cell auto fluorescence; (2) GFP fluorescence was divided by the measured OD; (3) Normalized fluorescence was calculated as the average fluorescence measured in a time window of 1 hour of exponential growth. The normalized fluorescence is shown as fraction of the maximal fluorescence observed in each strain.

2.2 Bet-hedging strategies in microorganism

One of the aspects that were not checked by *Mitchell et al*. is how evolutionary conditioning is viewed at the single cell level.

It is now known that heterogeneity can be found in microorganisms even in a population of genetically identical cells grown under the same conditions [12]. This type of phenotypic heterogeneity, that will be referred here generally as 'noise', could be a negative trait, as it weakens the bond between genotype and phenotype, and consequently the adaptation to the environment for which this genotype was selected. However, under certain conditions, noise could be thought to be advantageous at the population level, as eventually it could allow survival of more individuals under unpredictably changing environment [13]. This strategy in which one genotype stochastically leads to various phenotypes that cope differently with different environmental conditions is called in evolutionary theory 'bet-hedging', a term borrowed from gambling, in which one divides one's bets in order to minimize loss

while lowering the maximal potential gain. With the advancement of single cell analyses, bet hedging strategies of microorganisms begin to unravel.

Below I list a few examples of phenotypic heterogeneity in isogenic population of microorganisms that were suggested to represent bet-hedging strategies.

Persistence in E. coli In 1944 it was discovered that a portion of a bacterial population remains alive after antibiotic treatment, but it is not due to genetic resistance as when they were re-cultured the new culture had the same survival percentage as the original one [14]. This phenomenon was termed 'persistence'. It was suggested that the persistent bacteria were cells that were in dormant stage during the antibiotic treatment, and that cells switch stochastically back and forth between the dormant and replicating states. Sixty years later, *Balaban et al.* [15] showed, using a microfluidic device, that this was indeed the case. Persistence can be seen as a bet hedging strategy, as the population splits between dividing cells and non-dividing, persistent, cells. While the dormant cells lower the mean fitness of the population in non stressful conditions as they do not divide, they constitute a subpopulation that can survive a stressful condition, such as exposure to antibiotic.

Competence and sporulation in Bacillus subtilis During starvation an isogenic population of the gram-positive bacterium *B. subtilis* shows a complex behavior, as it splits to different subpopulations that respond differently to the stress. The different responses include: extracellular matrix production, motility, cannibalism, cell lysis that leads to nutrient release, growth arrest, and competence, the ability to take DNA from the environment [16]. Eventually, as the nutritional stress progresses, many (but not all) cells will undergo sporulation, the process of creation of a dormant highly resistant cell, and lysis of the mother cell. It was shown that competence is governed by the noisy expression of a master regulator that in turn, causes the cell to enter and exit the competent state, and that the switching rate between the two states fits the noise level in the master regulator expression [17,18]. In addition, sporulation was shown to be governed by bistable expression of a master regulator [19]. The high and low expression states of the master regulator were found to be heritable, and gave rise, respectively, to sporulating and non-sporulating lineages [20]. In *S. cerevisiae*, onset of meiosis, the first step in yeast sporulation, was also found to be governed by master regulator that variation in its production rates causes variability in the onset of meiosis [21].

Metabolic systems In 1957 it was discovered that the lactose operon of *E. coli*, when activated by intermediate levels of a synthetic non-metabolizable inducer, showed bistability among the population, some cells induced it to full extent, while other kept it off [22]. Since then, the bistability of the lactose operon was extensively analyzed both experimentally and theoretically; more details can be found in the discussion section 5.1. A similar phenomenon was discovered in the arabinose utilization system of *E. coli* [23] and in the galactose signaling network of *S. cerevisiae* [24]. These examples, when regarded as bet-hedging strategies, are a more subtle form of it than the examples above, as no growth arrest is involved. However, activating the metabolic system of one substrate might come at the expense of being able to metabolite another substrate that might be more valuable (e.g. glucose), and thus in unpredictable environments it might be better to have different portions of the population prepared for different sugars.

Laboratory controlled examples In *S. cerevisiae*, researchers placed a uracil biosynthesis gene under a bistable promoter in two engineered strains that have different rates of stochastic switching [25]. They grew the two strains in a fluctuating environment in which in one state the expressing cells had higher fitness, and in the other they had lower fitness, and showed that the overall fitness of each of the strains is high when the rate of switching fits the environmental fluctuation rate. In *Pseudomonas fluorescens*, lab evolution in two alternating environments with bottleneck selection rounds, led to the development of cells that stochastically switch capsule production on and off and shown to have increased fitness in the selection regime of this study [26].

2.3 Bet hedging in evolutionary conditioning

These examples show that the stochastic nature of different cellular processes can create a phenotypic variability in an isogenic population of microorganism. Such heterogeneity may, in certain cases, help the population to cope better with the unpredictably changing environment.

Returning to evolutionary conditioning, it is tempting to imagine a scenario in which evolutionary conditioning is combined with a bet hedging strategy. In natural environments it is likely that even if S1 and S2 are temporally connected, the connection might not be absolutely dependable. In such environments, from time to time, S1 could appear without the following S2. If the cost of futile preparation to S2

is sufficiently high, the uncertainty might be reflected in the distribution of the S2 gene activation in response to S1. Namely, in response to S1 exposure, a portion of the population will activate its S2 genes, while another portion will remain inactive.

For instance, it is possible that in the *E. coli* evolutionary conditioning example examined by *Mitchell et al.*, the activation of maltose metabolism operons during lactose exposure will be noisy, as it is only likely, but not absolutely sure, that the cell will reach the maltose-rich environment. On the other hand, when exposed to maltose, cells might express those genes in a less noisy fashion, as there is no uncertainty in its occurrence. As can be seen in Figure 1b, the activation of the maltose operons in response to lactose is only partial. This might imply that in response to lactose all the cells in the population activate the maltose operons to an intermediate level, or, alternatively, that only part of the population activates the maltose operons, manifesting a sort of a bet-hedging strategy.

In this study, I examined the evolutionarily conditioned response of *E. coli* towards sugars, in which maltose operons are activated in response to lactose exposure, at the single cell level. This allowed me to analyze the distribution of the maltose operon activation levels when the cells are exposed to lactose or to maltose, and to ask whether the distribution's shape might imply a sort of bet hedging strategy in this system. A possible manifestation of bet-hedging in this context could be a broader distribution of activation levels of the maltose operons in response to lactose, compared to maltose.

3. Methods

3.1 Bacterial strains and plasmids

E. coli MG1655 was used in all experiments. *E. coli* DH5α was used for cloning. Plasmids used for experiments and cloning are listed below (maps are in Figure 2):

pUA66/pUA139 (Figure 2a) - The scaffolds used in [10,11] for construction of the promoter-fused GFP library are low copy SC101 origin of replication plasmids, with a kanamycin resistance gene, and GFP gene (*gfpmut2*) with a cloning site for the promoters upstream to it (XhoI-BamHI). pUA139 is the same as pUA66, except that the restriction sites are switched (originally used in [10,11] for cloning of opposite strand promoters using the same primers). Promoters were defined as the intergenic region upstream to a gene, plus 50-150 bp into the adjacent coding regions [10]. The eight library strains kindly provided by Prof. Uri Alon are listed in table 1.

Table 1 – Strains from the promoter-fused GFP library [10,11]

E. coli strain	Plasmid	Promoter	Genomic coordinates of the cloned promoter (according to NCBI Reference Sequence: NC_000913.2)	Operon
MG1655	pUA66	*malK*	4244356 - 4244849	*malK-lamB-malM*
		malT	3550413 - 3551207	*malT*
		malZ	421509 - 421824	*malZ*
		malS	3735105 - 3735593	*malS*
		lacZ	365687 - 365438	*lacZ-lacY-lacA*
		σ^{70}	A synthetic promoter containing a consensus σ70 binding site. Sequence can be found in the supplemental data of [11].	-
	pUA139	*malP*	3551209 - 3550413	*malP-malQ*
		malE	4244846 - 4244356	*malE-malF-malG*

pZS2-123 (Figure 2b) - a three-color reporter scaffold [27], kindly provided by Prof. Michael Elowitz, containing YFP, RFP and CFP genes optimized for bacterial

expression, with promoter cloning sites upstream to each coding region (XhoI-BamHI for YFP, XmaI-AvrII for RFP, and SalI-XmnI for CFP).

This plasmid, as the pUA66/pUA139 plasmids, is a low copy SC101 origin of replication, with a kanamycin resistance gene.

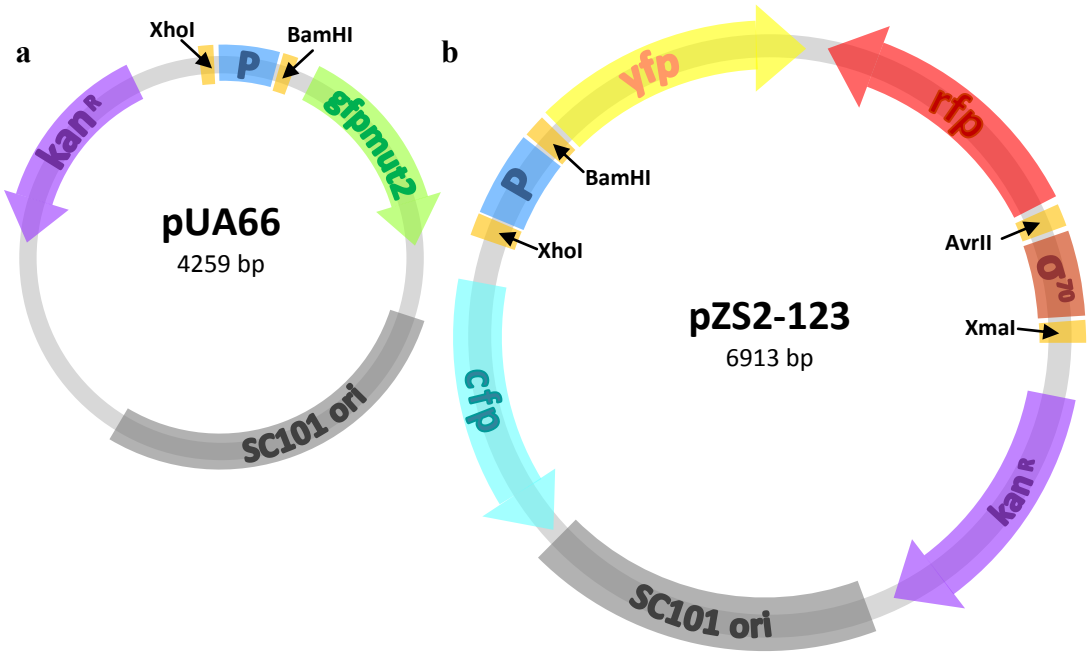


Figure 2 - Schematic maps of pUA66 and pZS2-123 plasmids

(**a**) pUA66, the scaffold of the promoter-fused GFP library [7,8]. **P** (blue region) stands for the varying *gfpmut2* promoter region. pUA139 is the same as pUA66, except that the XhoI and BamHI restrictions sites are in opposite orientation.

(**b**) pZS2-123, the three-color reported plasmid [26]. The original promoter of RFP was replaced by a synthetic promoter with a σ^{70} binding site from [8]. **P** (blue region) stands for the varying YFP promoter region that was replaced each time by one of the promoters from [7,8] mentioned in Table 1.

Length of plasmids in base pairs does not include the varying promoter region.

The promoters of the kanamycin resistance and CFP genes are not presented for simplicity.

3.2 Media

For all flow cytometry and sorting experiments M9 supplemented medium was used (M9 Salts (*MP Biomedicals*), 1mM MgSO4, 0.1mM CaCl2, 0.05% casamino acids, 5 ng/ml thiamine, and 0.1% carbon source). Carbon source was either glycerol or glucose. The glucose containing medium was supplemented with 20mM cyclic AMP (cAMP) to avoid glucose repression.

Sugars, when added, are in the following final concentrations: maltose 5mM, lactose 10mM (these concentrations were chosen by *Mitchell et al.* as the minimal concentration that allows maximal growth).

Cultures for cloning were grown in LB medium (5gr/l Yeast extract (*Becton Dickinson*), 10gr/l triptone (*Becton Dickinson*), and 10gr/l NaCl).

Kanamycin was added to all media (final concentration: 25µg/ml).

3.3 Plasmid purification protocol

As all plasmids were low-copy SC101 origin of replication plasmids, the yield of a standard miniprep protocol (QIAprep Spin Miniprep Kit, *Qiagen*) was too low for later cloning procedures. Thus, a modified protocol was used: 40ml of bacterial stationary culture (grown on LB + kanamycin medium) were divided to 4x10ml; from each 10ml DNA was then extracted using the miniprep kit protocol, with double volumes of P1, P2 and N3 buffers. After the neutralization (N3) step, supernatant from two tubes was loaded on one column (two columns in total) for purification. DNA was eluted using 150µl double-distilled water, and then the two columns were united and the water was evaporated using SpeedVac concentrator (SVC100H, *Savant*). Then, the pellet was re-suspended in 50µl TE buffer.

3.4 Growth conditions and data acquisition in flow cytometry experiments

Incubation temperature was either 37°C or 30°C throughout the experiment. Overnight culture was diluted 1:1000 into 5ml of fresh medium, grown for 1h and treated with sugar (maltose, lactose, or none) in the final concentration mentioned in section 3.2. Cells were farther incubated for 4 hours. During incubation samples were taken each hour, and measured using an LSR-II cytometer (*Becton Dickinson*). Acquisition was performed using the FACSDiva software (*Becton Dickinson*). Excitation lasers, emission filters and voltage settings for the photomultiplier tubes, which were used for the measurement of forward scatter (FSC),

side scatter (SSC) and each of the fluorescent proteins are listed in table 2. Threshold was based on a FSC-SSC 'AND' gate with varying values (200-1000). CFP could not be detected with either of the filters mentioned in table 2, using the original construct under constitutive promoter or under the regulation of the σ^{70} promoter.

Table 2 – Laser, filters and voltage settings for the flow cytometry experiments

Parameter	Laser	Filters		Voltage settings
		LP	BP	
FSC	488nm (Blue)	-	-	500 (Log)
SSC	488nm (Blue)	-	488/10	300 (Log)
GFP/YFP	488nm (Blue)	505	525/50	416 (Log)
RFP	561nm (Yellow-Green)	600	610/20	600 (Log)
CFP	405nm (Violet)	505	525/50	-
	405nm (Violet)	-	470/15	-

3.5 Flow cytometry data analyses

All flow cytometry data processing was performed in Matlab (*Mathworks*), using a graphical user interface I developed (see results section 4.3 for more details on the GUI and the gating process). For the analysis in Matlab, the output from the LSR-II cytometer (binary .fcs files) was converted to text files using the freeware FCSExtract Utility (Dr. Earl F. Glynn, Stowers Institute for Medical Research). Other statistical analyses were performed using various Matlab built-in functions.

3.6 Cell sorting experiment

Growth conditions in the sorting experiment were similar to those used in the flow cytometry experiments (section 3.4 above), M9 carbon source was glycerol. Overnight culture grown at 37°C (containing ~$5x10^8$ cells/ml), was diluted 1:1000 into 60ml of fresh medium, grown for 1h at 37°C, then treated with 10mM lactose, and incubated for additional ~3 hours at 37°C (final concentration is roughly $4x10^6$ cells/ml). Then, the culture was centrifuged, and the pellet was re-suspended in 6ml of fresh lactose containing medium (in order to concentrate the culture to roughly $4x10^7$ cells/ml). Sorting was performed using the FACSAriaII cell sorter (*Becton*

Dickinson), and the FACSDiva software. The sheath fluid for the sorting was the M9+Glycerol+lactose medium itself. Roughly 50,000 cells per second were sorted (in the 'Yield' mode of the instrument). Gating procedure for sorting is described in results section 4.2.2. The sorted cells ($3x10^7$ cells for each of the two sorted populations, in a volume of about ~120ml) were centrifuged and re-suspended in 10ml M9+Glycerol+5mM maltose (concentration of $3x10^6$ cells/ml). Then, a 96 well flat-bottom plate was prepared in a "checkerboard pattern" of the two sorted population (100µl of culture in each well), and 50µl of oil was added on top. The bacteria were grown in an Infinite F500 plate reader (*Tecan Group*), at 37°C, for 6 hours, during which OD and YFP levels were measured in 18 minutes intervals.

4. Results

4.1 Methodology development and assessment

4.1.1 Experimental design

In order to examine the distribution of activation of maltose operons in response to lactose and to maltose, I repeated *Mitchell et al.* experiments, but this time measuring the fluorescence level using flow cytometry, rather than a plate reader as was used in their study, in order to have a single cell resolution.

The maltose regulon consists of ten genes organized in five operons (*malPQ*, *malEFG*, *malK-lamB-malM*, *malZ* and, *malS*, will be noted here by the first gene of each operon), all regulated by one transcription activator, MalT (more details can be found in Discussion section 5.1). The promoter activity of the five maltose regulon's operons, *malT* promoter, and the lactose operon, were examined using the corresponding strains from the promoter-fused GFP library [10,11], each harboring a plasmid with GFP under the regulation of the specific operon's promoter.

Two types of supplemented minimal media were used, which differ in the added carbon source: glycerol or glucose (See Methods section 3.2 for more information). *Mitchell et al.* used the two media as they have different advantages and gave similar yet distinct results. As general transcription in the bacterial cell is affected by growth rate, it was important to minimize the growth rate differences between cultures that were treated with the various sugars. Therefore, *Mitchell et al.* used glucose as a background carbon source, as the addition of another sugar, such as maltose or lactose does not affect growth rate. However, glucose leads to repression of many sugar operons in *E. coli*, among which lactose and maltose, and thus cAMP was added in order to cancel the repression [11]. On the other hand, both glucose and cAMP affects the maltose regulon directly (see discussion section 5.1), and in order to control for this effect, *Mitchell et al.* used also glycerol as an alternative background carbon source, which does not require cAMP, but adding maltose or lactose to it would affect growth rate. Here, I used the two types of media as well.

In contrast to a plate reader, which provides proper conditions for growth, such as constant temperature and shaking, during the measurements, flow cytometers are not applicable for time course experiments directly. Thus, a sample was taken from a culture grown under the optimal conditions each hour, for a total time of four hours (five time points including zero). Each of the strains was measured in three cultures,

one treated with maltose, the other in lactose, and the third served as a control (see Methods section 3.4).

4.1.2 Modification of a three-color reporter plasmid for the current research

After preliminary results were obtained, we realized that a general cellular state reporter is needed, as in the moment of measurement cells in the population might be in different states, affecting generally the transcription and translation in the cell. Such influences might be, for instant, different size, number of ribosomes and RNA polymerases, and cell cycle stage. Thus some of the variability in GFP expression stems from these general differences between cells, and thus normalization to a general cellular state reporter will leave those differences out of the fluorescence variance.

The promoter chosen for general cellular state reporting was a synthetic promoter, constructed by *Kaplan et al.* [11], which contains a binding site for σ^{70}, the housekeeping sigma factor in *E. coli*. σ^{70} is active throughout the exponential phase and its activity decreases when the bacteria enter stationary phase [28], thus the synthetic promoter that contains σ^{70} binding site can be used as a general cellular state reporter during the exponential phase, which is the growth phase of the bacteria in all experiments.

Recently, a three-color reporter plasmid was constructed [27] (see Figure 2b for a map). This construct contains genes for three fluorescent proteins: YFP, RFP and CFP, under the control of three separate promoters, each with bordering restriction sites, which facilitate modifications and replacement using restriction enzymes cloning techniques. Thus, one of the fluorescent proteins could be used as a general cellular state reporter, and the two others could be used for simultaneously examine two promoters of interest in the same cell, allowing to elucidate potential coordination between the two. However, CFP was not detected in the flow cytometry, not in the original plasmid, where it is under a constitutive promoter, and not when this promoter was replaced with the σ^{70} synthetic promoter. Therefore it could not be used in this experimental system, and following two promoters simultaneously in the same cell was not possible.

The RFP was used as a general cellular state reporter. The promoter of the original plasmid was replaced with the σ^{70} synthetic promoter mentioned above. For that, the synthetic promoter region from the promoter-fused GFP library plasmid was

amplified using PCR, with primers bordering this region but that contained the restriction sites needed for integration into the three-color reporter plasmid. The PCR product was then cut with the proper restriction enzymes and ligated to the cut three-color reporter plasmid.

Then, the YFP promoter was replaced with the different maltose and lactose operon promoters to create seven different plasmids each containing YFP under one of the seven promoters (Table 1, methods section 3.1). For most promoters, the bordering restriction sites in the two construct are the same, so cloning did not involve a PCR amplification step, but a restriction reaction and purification of the proper band from the gel, followed by ligation into the cut three-color plasmid. Only *malP* and *malE* promoters, whose restriction sites are in the opposite orientation, were amplified by PCR with primers containing the reverse restriction sites.

For each plasmid, several clones were first examined for correct integration in colony PCR using primers from both sides of integration site on the three-color reporter plasmid. Positive colonies were validated by sequencing.

In order to examine the potential influence of each of the two different plasmids on the expression of the fluorescent protein, strains containing the same promoter from the promoter-fused GFP library and from the cloned three-color plasmid were examined in the same experiment. Figure 3 presents the Pearson correlation coefficient between the mean fluorescence values of the two type of plasmids for each promoter across the various treatments and time points. Pearson correlation was chosen as it is affected by the relative change in fluorescence rather than the absolute value, and an intrinsic difference between GFP and YFP absolute fluorescence is expected. As can be seen, for most promoters the correlation coefficients are close to 1, thus the different plasmids do not affect the relative mean fluorescence dynamics. However, *malZ* and *malS* each have a non-significant low correlation coefficient on one of the temperatures. The low correlation in only one of the experiments may imply on mixed population, as only one colony was taken for each experiment. In addition, in the verification sequencing results - the one *malZ* colony sent for sequencing gave *malS* sequence and the one *malZ* colony gave *malS* sequence, which may support the notion of mixed populations.

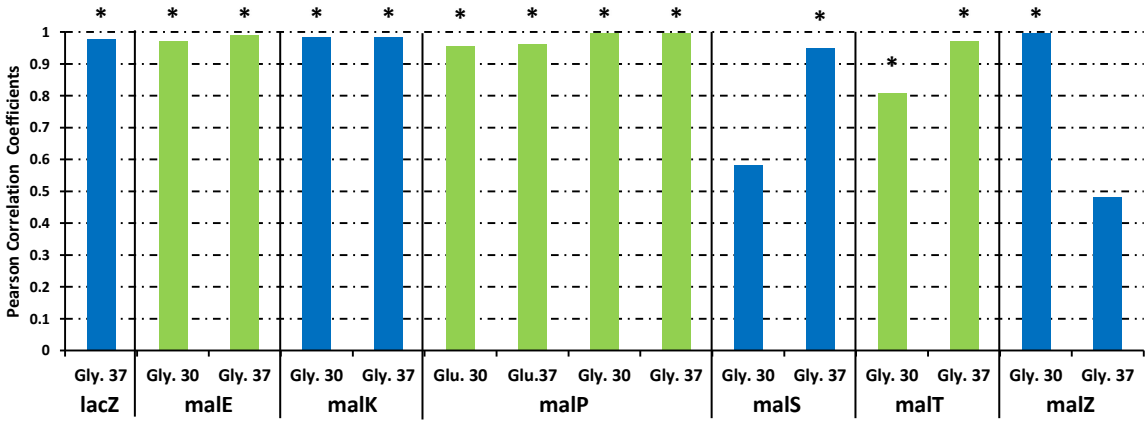


Figure 3 – Pearson correlation coefficients between mean fluorescence of promoter-fused GFP library and the cloned three-color plasmid strains

The Pearson correlation coefficients were calculated using the 5 time points and the three treatments as one 15 point vector. Gating was based only on FSC-SSC (see section 4.1.4). Pearson correlation coefficients with p-value lower than 0.01 are marked with asterisks.

4.1.3 A Matlab graphical user interface for display and analysis of flow cytometry data

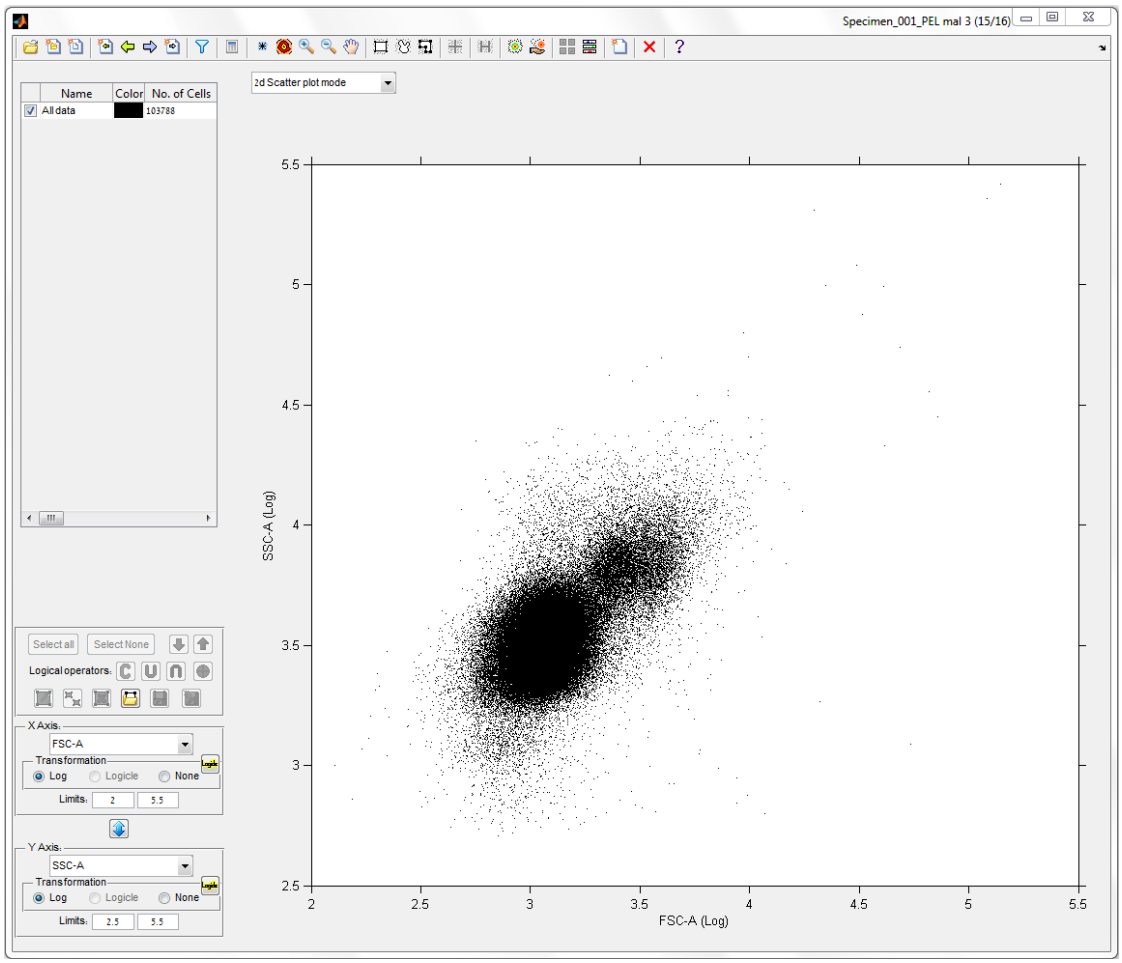
Many flow cytometry users use one of the commercially available softwares for their data analysis, such as *FACSDiva* (*Becton Dickinson*), *FlowJo* (*Tree Star*) and *FCS Express* (*De Novo Software*). These software are indeed adequate for the purposes of many flow cytometry experiments. However for the current aim, a few features that are required for proper analysis are missing in any of those softwares: (1) Gating can be done only by defined shapes (ellipses, polygons, lines) drawn on the plot, without the ability to give a numeric value, or a custom shape; (2) Limited control over visualization of one-dimensional histograms and two-dimensional dot plots and contour plots; (3) Statistical measures of the population regarding only one parameter at a time (e.g. no way to know what is the X,Y values of most dense point on a two-dimensional contour plot); (4) Inability to perform math operations on parameters, such as calculate the ratio between two parameters for each cell (e.g. YFP-to-RFP ratio).

Matlab language is suitable for all of the analysis processes mentioned above, as well as for handling large matrices such as flow cytometry data. While Matlab has some programs designate to biological data analysis (microarrays, bioinformatics etc.), there is no existing program for comprehensive flow cytometry data analysis.

As none of the available commercial software had the essential features mentioned above, I decided to build a new flow cytometry program in Matlab, based on the commercial softwares, with the addition of the missing features. To make it more easy-to-use without the need to delve into the code, the program was designed as a graphical user interface. It will be uploaded to Matlab File Exchange site (Matlab Central, http://www.mathworks.com/matlabcentral/fileexchange/) for the use and development of other Matlab flow cytometry users.

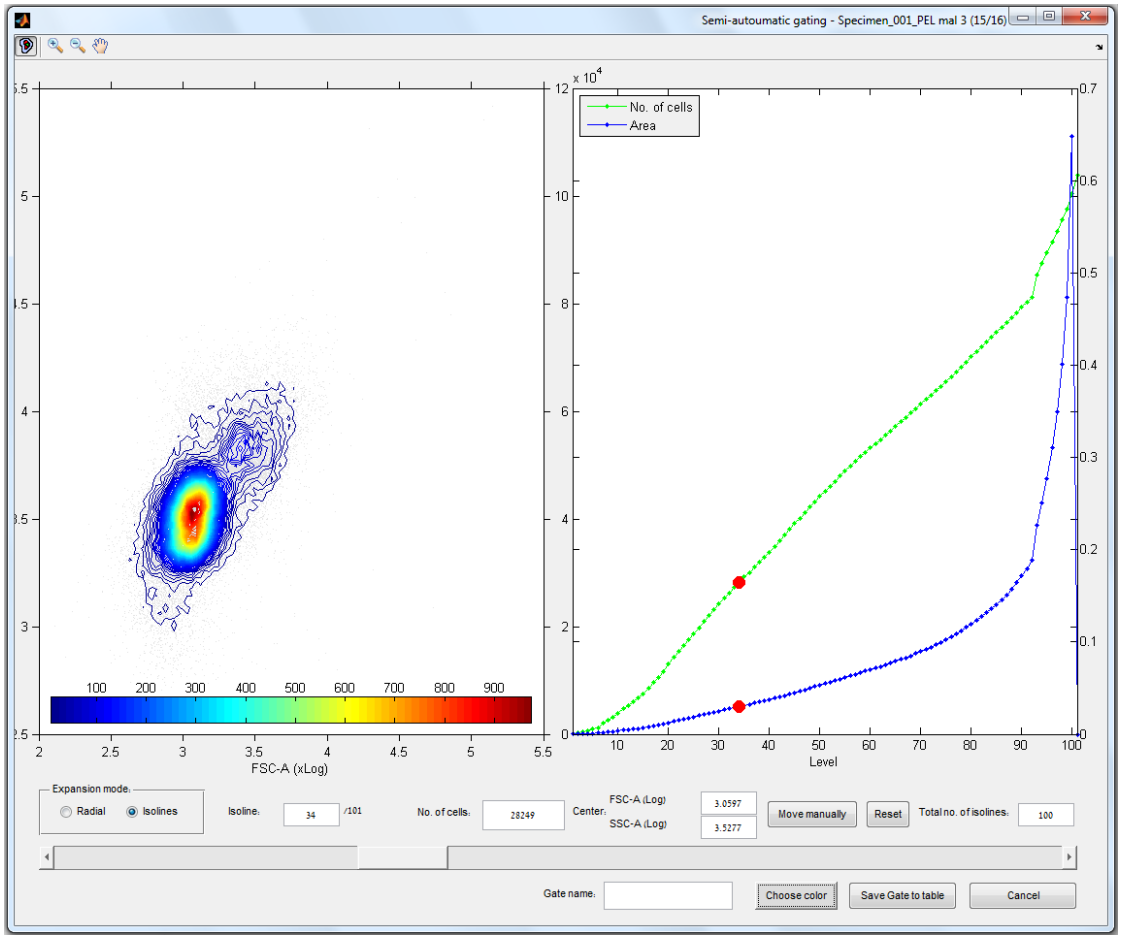
Screenshots of the graphical user interface are presented in Figure 4.

Figure 4 – Screenshots of the Matlab graphical user interface

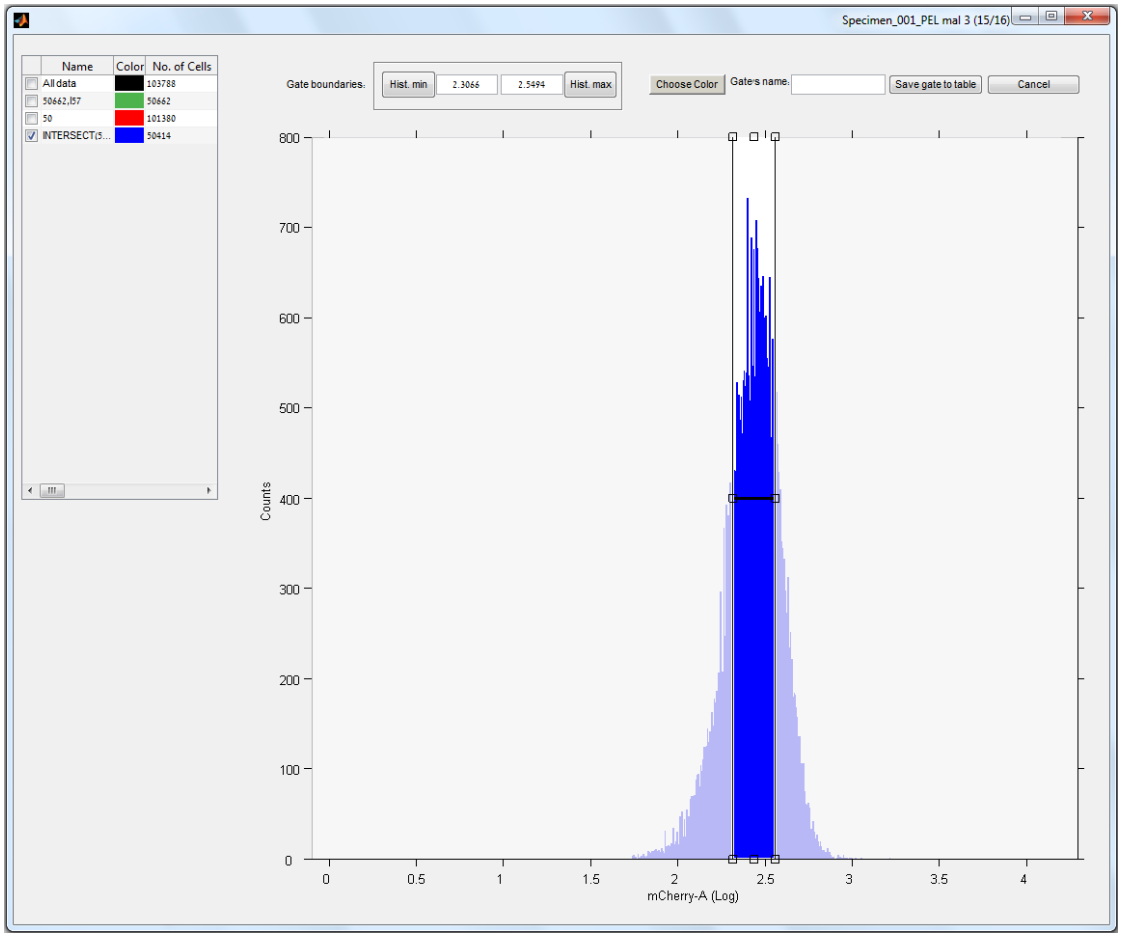


a – Main window. The main window of the graphical user interface enables to present two-dimensional dot plots and one-dimensional histograms of each of the parameters with custom-defined limits, transformations (logarithmic, 'logicle' [29]), and number of bins; gating according to different shapes, limits or distance from the center; logical operations on gates (intersect, union, inverse and exclusive-or); filtering of cells having extreme values and according to time of the read; mathematical operations on the parameters; and moving from file to file while keeping the same settings for easy comparison.

Figure 4



b – Gating according to cell density. A tool for gating according to specified distance from the center of two-dimensional distribution. The center is defined as the point of highest cell density, but can be moved manually. Expansion of the gate can be either done by setting a radius around the center, defined in terms of the units on the two axes of the plot, or according to contour lines of cell density defined with Matlab's 'contourc' function. On the left the contour plot is presented, colors present number of cells per area, on the right a plot of number of cells and area as a function of distance from the center is presented. The user can choose the specific distance according to which the gate will be defined.



c - Gating on one-dimensional histogram

On a one-dimensional histogram view of the main window it is possible to define a gate according to specified limits. The figure shows the main window (see **a**) in gating mode with most of the buttons being hidden or inactivated in order to prevent misuse. The mean and other statistical measures are presented in a table in the main window regular mode (not shown). The default was defined as mean fluorescence value ± 5% of this value, but those limits can be changed.

4.1.4 Gating procedure and influence on between-repetition coefficient of variation

Gating in flow cytometry, refers to selection of subpopulation of cells for further analysis based on one or more of the measured parameters. Gating is performed in order to minimize the variation in the physiological state of cells. Here, the gating was designed both on FSC and SSC, which correlates with different physiological states (cell size and granularity, respectively), and on RFP fluorescence, whose expression under the σ^{70} synthetic promoter is used as a general cellular state reporter (see section 4.1.2).

Using the Matlab graphical user interface described in the previous section, gating was performed in the following steps: (1) On FSC-SSC contour plot, an area around the center of distribution of cells was defined based on the density (see Figure 4b as an example). (2) On the RFP histogram of the FSC-SSC gated population, an area of the mean fluorescence value ±5% of this value was defined (see Figure 4c).

In order to inspect the influence of gating, I examined the variation between experiment repetitions in statistical measures such as mean, noise (variance divided by mean square) and skewness (a measure of asymmetry of a distribution) of the fluorescence distribution, which will be analyzed thoroughly in section 4.2.1. Figure 5 presents the reproducibility in estimating those statistical measures, measured as the between-repetition mean-normalized standard deviation (Coefficient of variation) of each of the statistical measures. As can be seen, the reproducibility of each of the measures is the highest when the FSC-SSC+RFP gate is applied.

The analyses include only *malK* and *malP* as only for them the experiment was repeated, for reasons mentioned in next section.

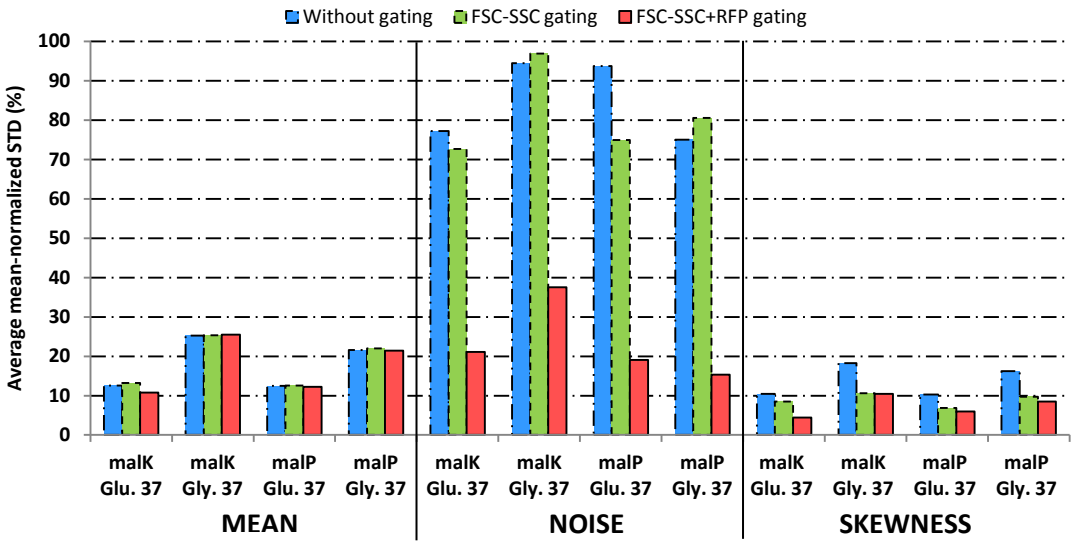


Figure 5 – The influence of gating on the between-repetition mean-normalized standard deviation

The average of between-repetition mean-normalized standard deviation was calculated in the following way: (1) the different statistical measures (mean, noise, skewness) were extracted from the fluorescence distributions (YFP-to-RFP ratio) of each time point of each treatment in the four repetitions; (2) the mean and standard deviation of these statistical measures in the four repetitions for a given time point and a given treatment were calculated; (3) the ratio between the mean and standard deviation was calculated to get the between-repetition mean-normalized standard deviation for a given time point and a given treatment (for the skewness measure that can have a negative value, the most negative value of it in all the data was added to each of the values for the ratio calculations); (4) The average of these ratios was calculated to get the average between-repetition mean-normalized standard deviation. Noise was calculated as variance divided by mean square.

4.1.5 Agreement with Mitchell et al. results

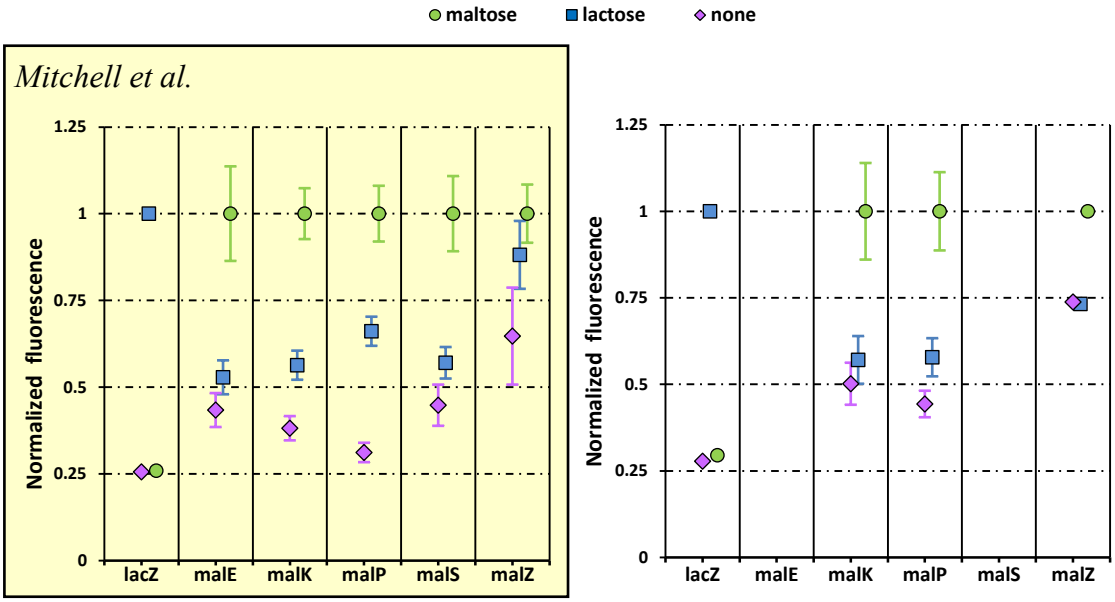
In order to compare the results of the current study to the results of *Mitchell et al.*, I regarded only the mean of the fluorescence distributions in each of the time points of each of the treatments. Figure 6 presents the results of this study compared to *Mitchell et al.* results. This analysis was performed using the FSC-SSC+RFP gated population, and the mean of the YFP to RFP ratio. Yet, since in the plate reader experiments of *Mitchell et al.* there was no gating, nor an additional fluorescent protein for normalization, a more legitimate comparison will be to use the ungated population and the YFP fluorescence alone. In practice though, using the ungated population and unnormalized YFP gave similar results regarding the normalized mean, with larger standard deviations (not shown).

As can be seen, *malP* and *malK*, the two promoters that show the highest relative activation in response to lactose in *Mitchell et al.* study, exhibit the same phenomenon here, namely, lactose exposure leads to partial activation of the promoter. However the gaps between the lactose treated and the control sample are much smaller here

than the gaps in *Mitchell et al.* results, and in the other promoters there is no apparent gap. Therefore I decided to focus on *malP* and *malK* promoters in further experiments. The *malT* promoter (not shown), in consistence with *Mitchell et al.*, showed no activation in response to lactose or to maltose.

The experiments in the current study were performed at 37°C, while the experiments of *Mitchell et al.* were performed at 30°C. Using 30°C in the flow cytometry experiments gave no activation of any of the promoters in the glucose media, thus I decided to perform all the experiments at 37°C. Using 37°C also has the advantage of being the physiologically relevant temperature.

a – M9 + Glucose + cAMP as the growing medium



b – M9 + Glycerol as the growing medium

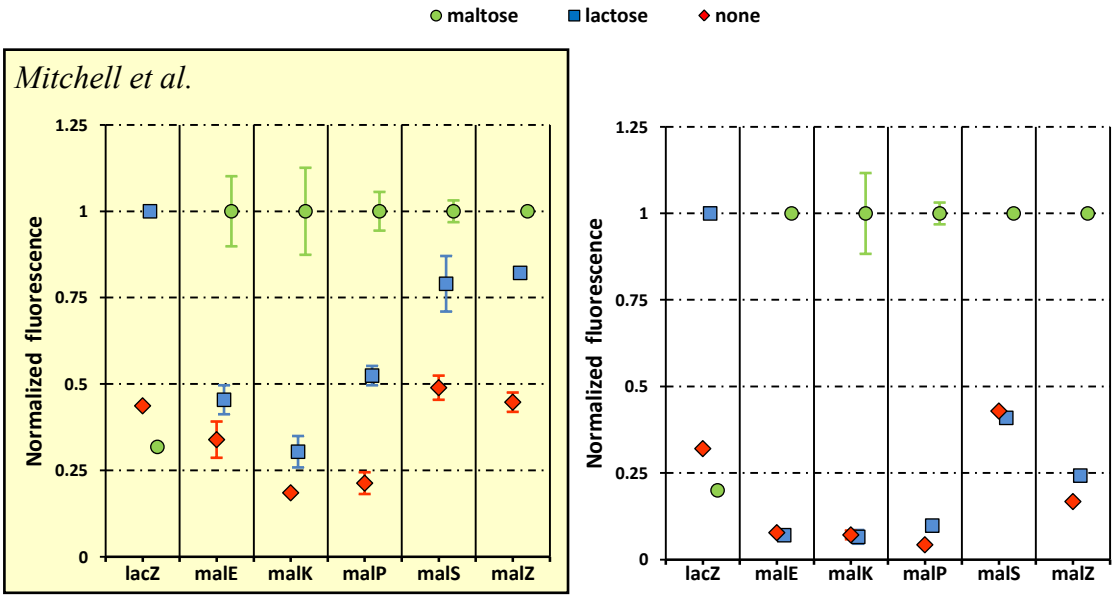


Figure 6 – Comparison to *Mitchell et al.* results

Mitchell et al. results are adapted from [1]. In each plot, mean fluorescence is normalized to the maximal mean fluorescence measured in the various treatments.

In *Mitchell et al.* the values are calculated based on four repetitions, and error bars represent the standard deviation.

In the current study, the mean of the fluorescence (YFP-to-RFP ratio) distribution of the FSC-SSC+RFP gated population was determined and normalized as mentioned above, using the 3 hours time point of each treatment. For *malK* and *malP* only, the mean fluorescence value was calculated as the mean of mean fluorescence values of four repetitions, the error bars represent the standard deviation of these values.

4.2 Single cell measurements and analysis of E. coli evolutionarily conditioned response towards sugars

4.2.1 Distribution of maltose promoter activity in response to lactose and maltose

Having established a measurement system for the promoter activities in flow cytometry, I could proceed to my original goal - examination of the distribution of promoter activity levels in the population, rather than the mere mean, in order to explore the feasibility of a sort of bet-hedging strategy in the response of the maltose operons to lactose.

Figure 7 presents an example of the fluorescence (YFP to RFP ratio) distribution in each time point of the various treatments in one out of four experimental repetitions of *malP* promoter in the glucose (Figure 7a) and glycerol (Figure 7b) media, and *malK* on the glucose medium (Figure 7c). As can be seen, the partial activation of the maltose operons, in terms of mean fluorescence, in response to lactose compared to the response to maltose is present in those three examples. Interestingly, the shape of the distribution varies in the different treatments as well. In general, the distribution is much broader in the lactose treated samples, than in the maltose or the control samples, and has higher level of noise accordingly (noise is defined as the variance divided by the mean square in all analyses). In addition, in the case of *malP* in the glucose medium, the distribution becomes highly asymmetrical, and has a long right tail, and thus high positive skewness, indicating that a relatively large portion of the population that displays high expression of *malP* in treatment with lactose.

Figure 8 summarizes the three statistical measures for distributions such as those that are presented in Figure 7, on the four repetitions, of the two promoters and two media. Here, again, the mean measure shows the partial activation of the maltose operons in response to lactose in all cases, but *malK* on glycerol, which shows no partial activation. In *malK* on glucose, and *malP*, which show partial activation, the noise level is much higher in the lactose treated samples, compared both to maltose and the non-treated samples. And finally, for *malP* on glucose, the skewness, is positive and much higher in the lactose treated samples, and implies on right tailed distribution, which corresponds to cells that show high expression level. *malP* on glycerol and *malK* on glucose do not show a right tail in the logarithmic scaled histograms in figure 7, but do show higher skewness when skewness was calculated based on the original, non-transformed, values.

Finally, Figure 9 displays the noise as a function of the mean fluorescence of *malK* on glucose, and *malP* on both media. As it is known that noise scales with mean: until a certain value genes with higher mean generally have lower noise [30-32], it was interesting to see whether such a scaling appears in the noise and mean values of the maltose operons in the various treatments and time points. One major difference between the common measurements of noise and this study is that usually noise is measured in steady state, while here I show the change in noise throughout activation. In addition, the noise measured here is the pathway specific noise and the intrinsic noise, as the RFP normalization is supposed to buffer the extrinsic noise. Moreover, the scaling was shown in translational fluorescent protein fusion, while in this study noise and mean of promoter activity were measured. And lastly, having only 15 points (5 time points x 3 treatments) compared to the hundreds of genes measured in the studies that showed the scaling phenomenon, it is not likely to see such scaling, and indeed there is no such scaling, as can be seen in Figure 9.

Nonetheless, Figure 9 exhibits interesting features. As can be seen, there is an elevation in the mean together with elevation in noise in all of the lactose-treated samples. In addition, when comparing points in which the various treatments that produce similar mean values, the lactose-treated samples have higher noise.

The normalization by RFP did not change the general trend in *malP*, and did change it on *malK*, as the unnormalized YFP fluorescence mean becomes smaller from a certain point in time.

a – *malP* on M9+Glucose+cAMP

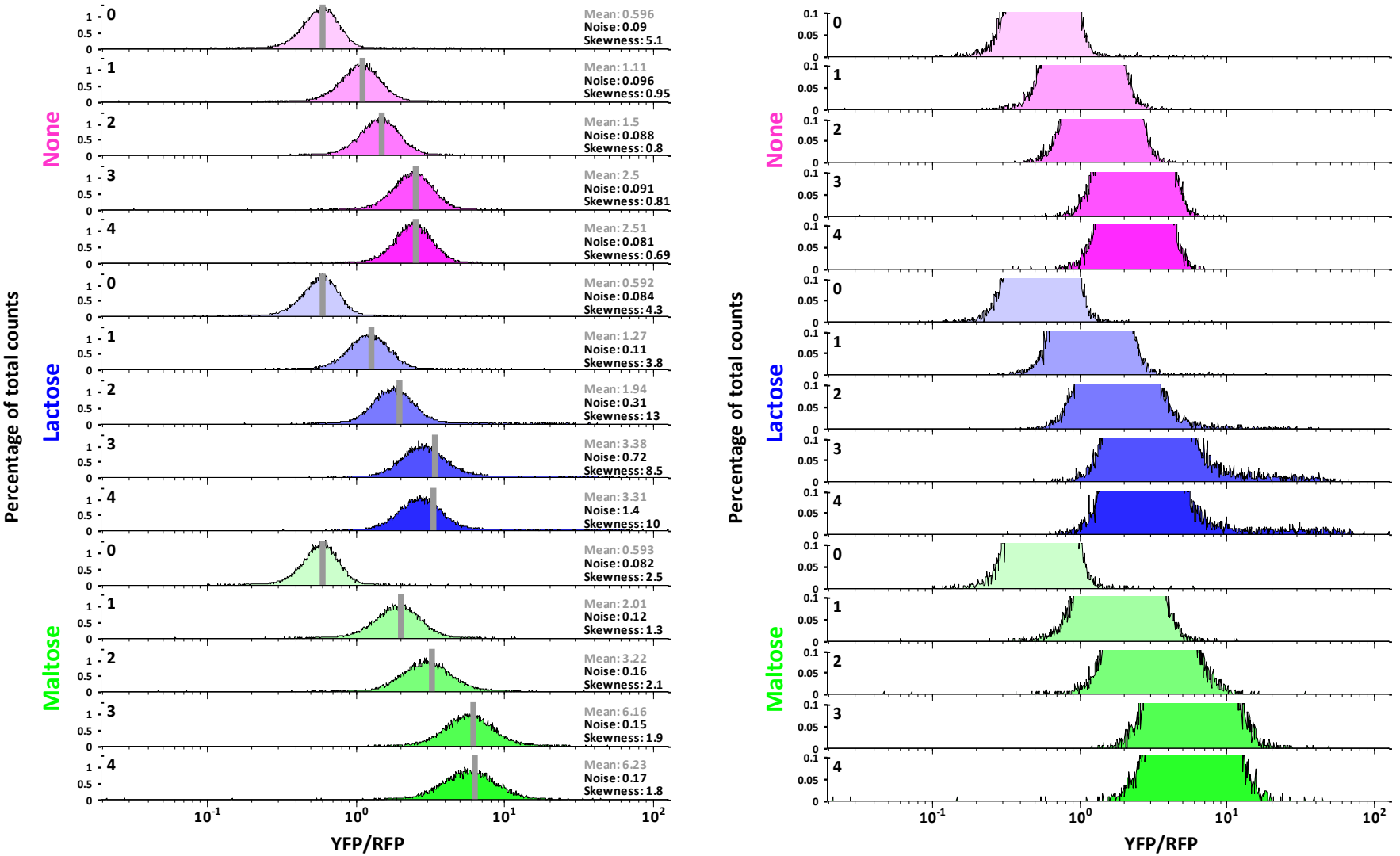
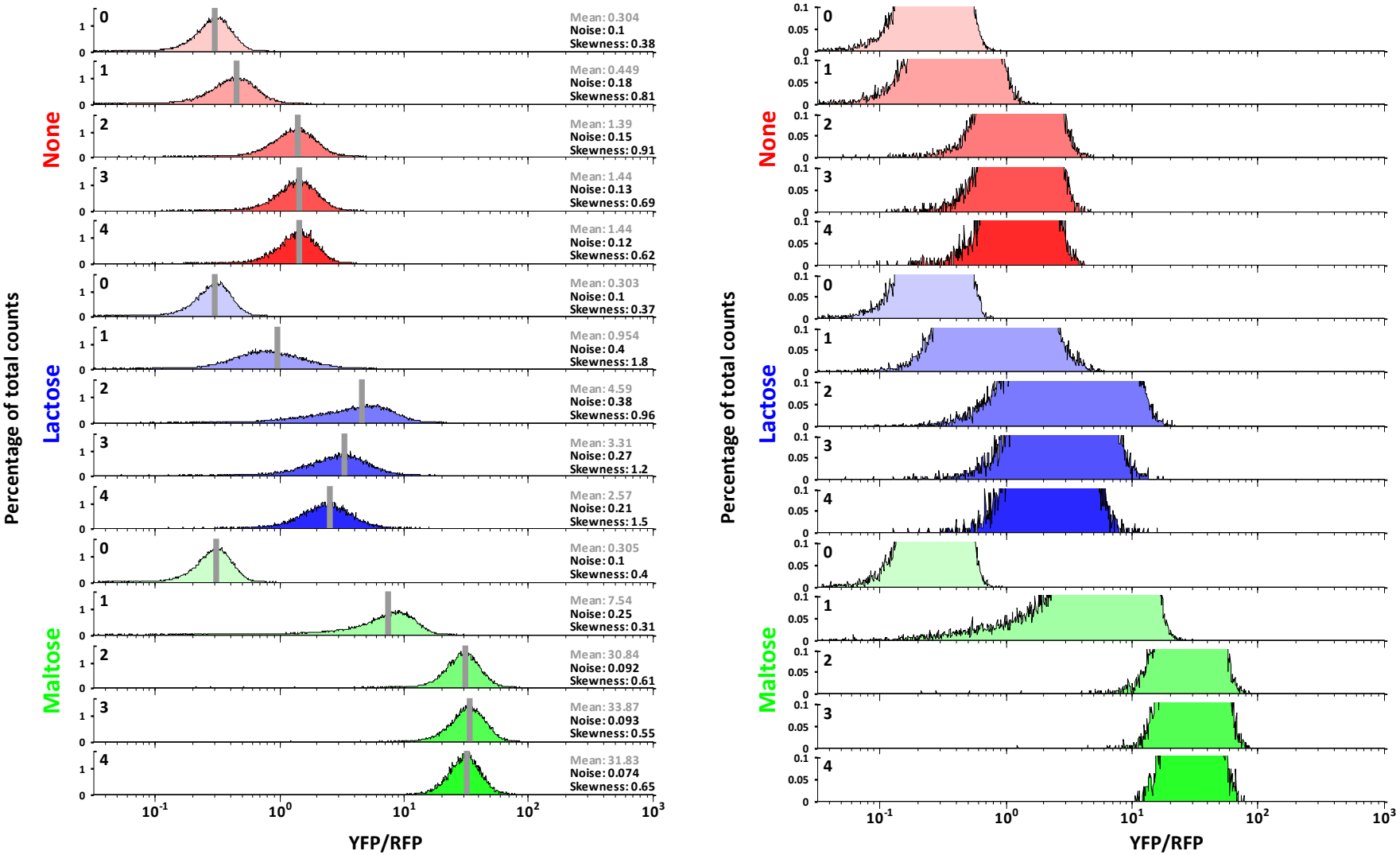


Figure 7

Figure 7

b – *malP* on M9+Glycerol



c – *malK* on M9+Glucose+cAMP

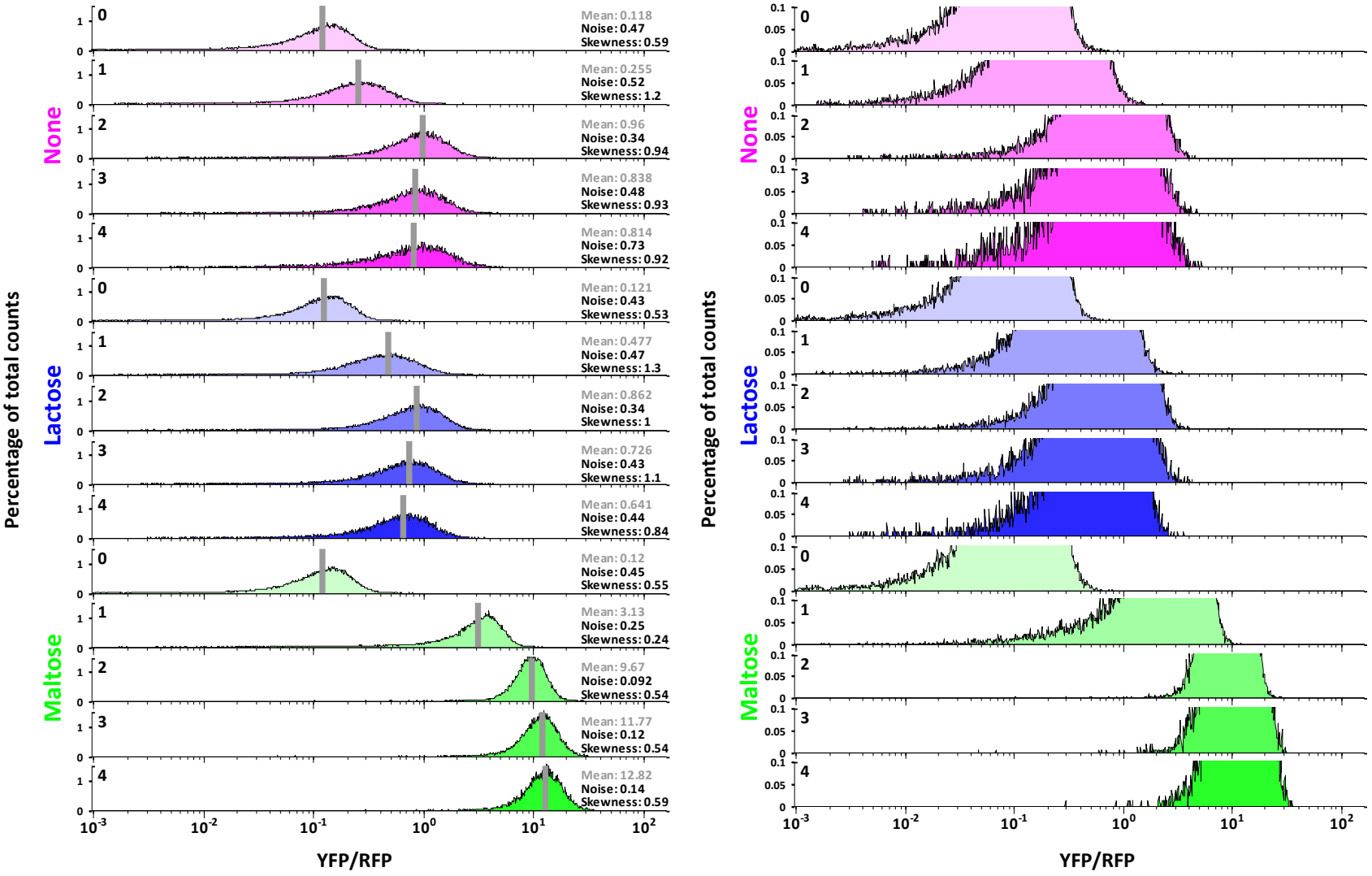
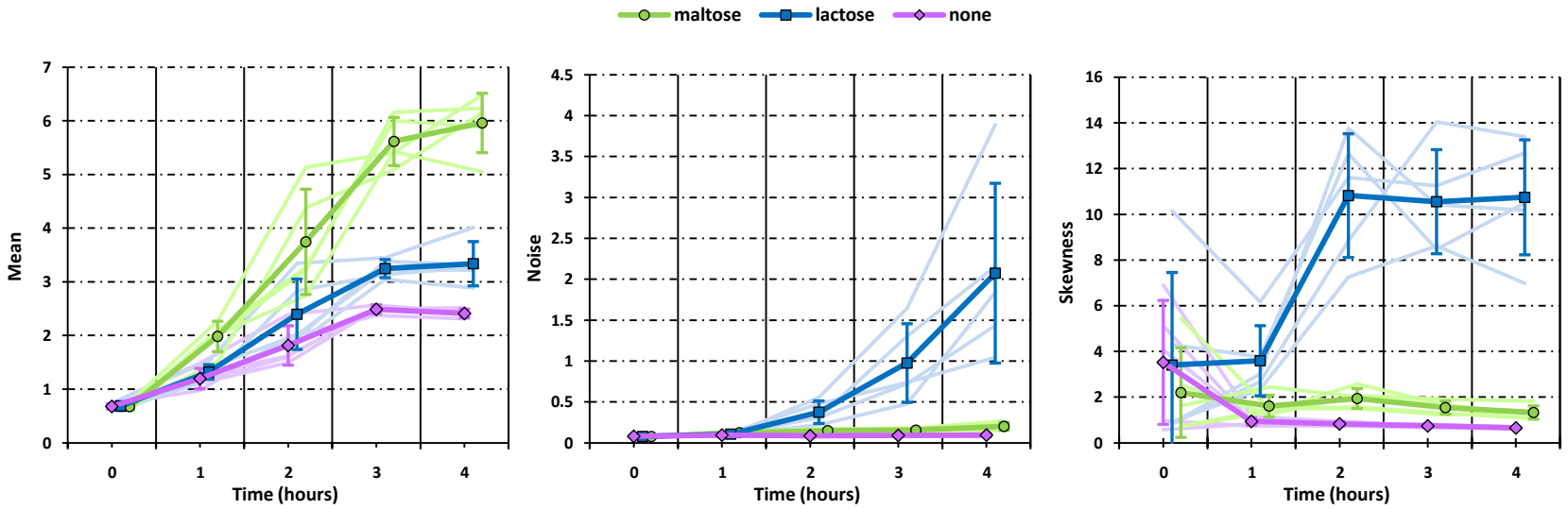


Figure 7

Figure 7 –Histograms of *malP* and *malK* fluorescence

Histograms of the YFP-to-RFP ratio of each of the cells in the FSC-SSC+RFP gated population, in the five time points of the three treatments. The histograms are in logarithmic scale and contain 1,024 bins. The Y axis is the percentage of the counts on a specific bin from the total number of counts. The right plot on each page is a close-up on Y axis (~x10 magnification on the Y axis alone). The mean, noise and skewness parameters appear on each plot. Skewness was calculated using the Matlab 'skewness' function. Noise was defined as the variance divided by the mean square. The mean is marked on each plot with a gray line.

a – *malP* on M9+Glucose+cAMP



b – *malP* on M9+Glycerol

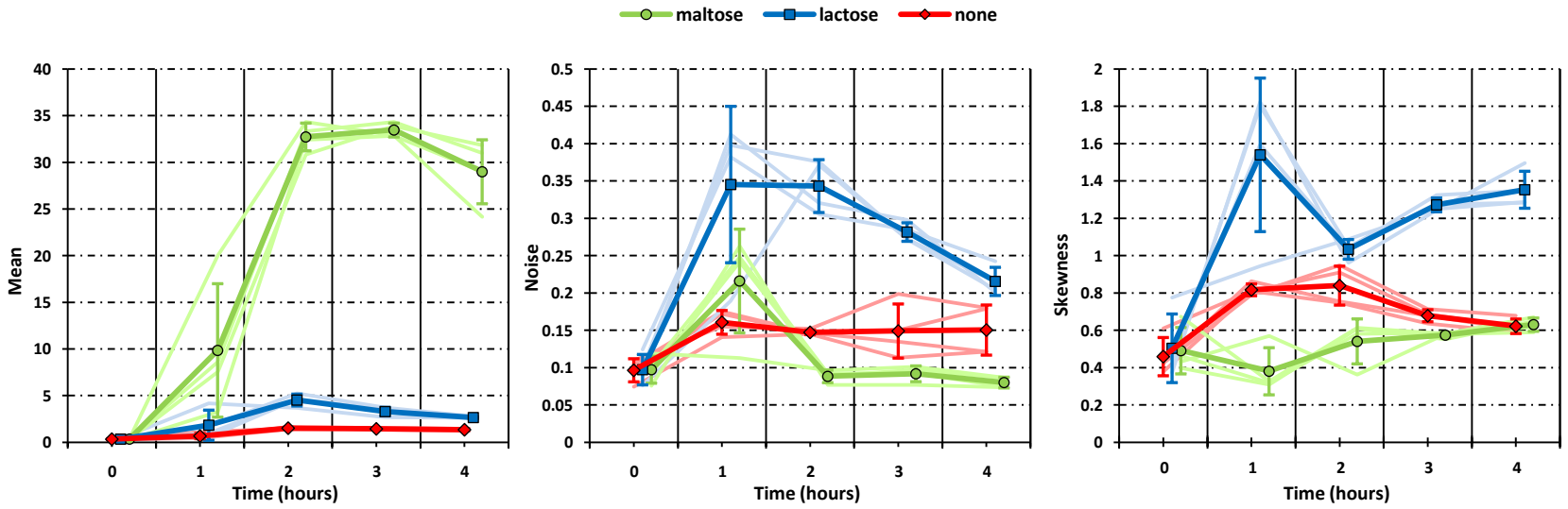
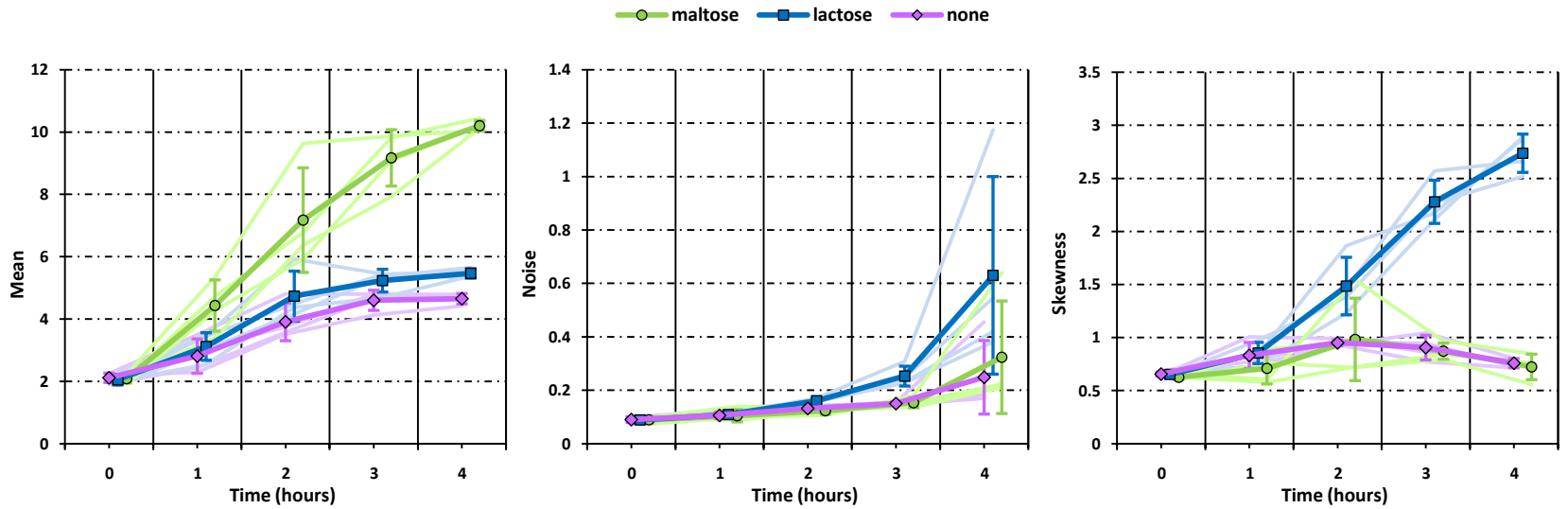


Figure 8

Figure 8

c – *malK* on M9+Glucose+cAMP



d – *malK* on M9+ Glycerol

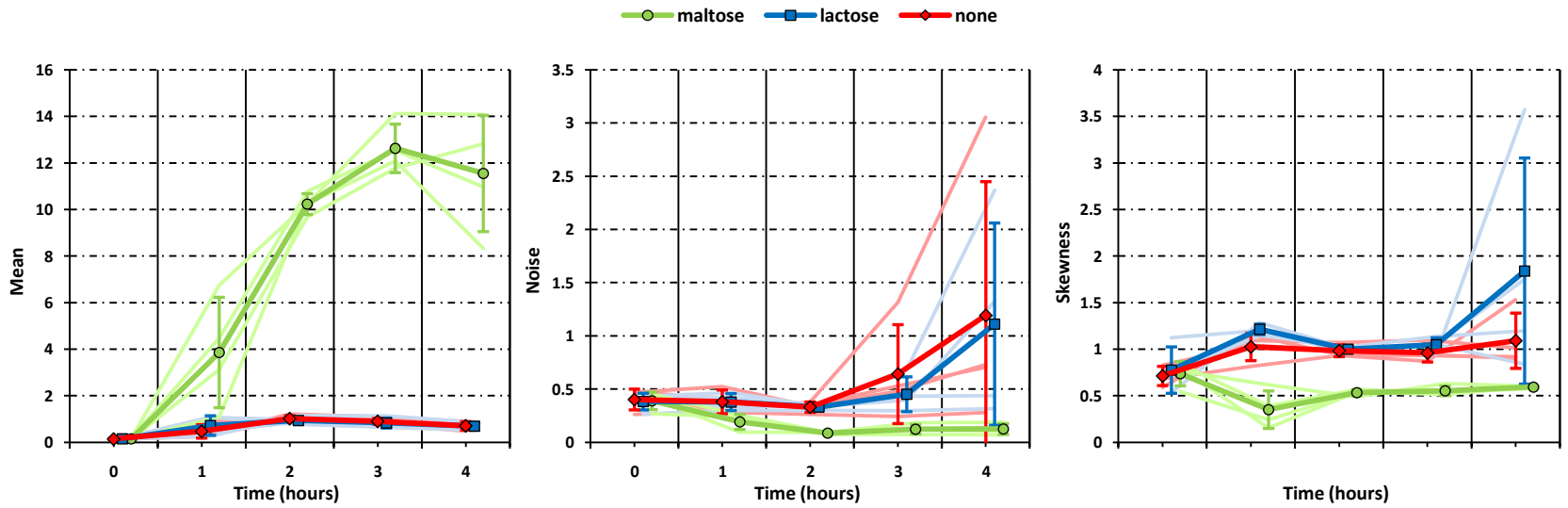


Figure 8 – Summary plots of mean, noise and skewness of the fluorescence distributions of *malK* and *malP*

Light lines represent the mean, noise and skewness values of the fluorescence (YFP-to-RFP ratio) distributions of the four repetitions. Bold lines represent the mean value of each of the measures in the specific time points of the four repetitions, and error bars are the standard deviation of these values. Skewness was calculated using the Matlab 'Skewness' function. Noise was defined as the variance divided by the mean square.

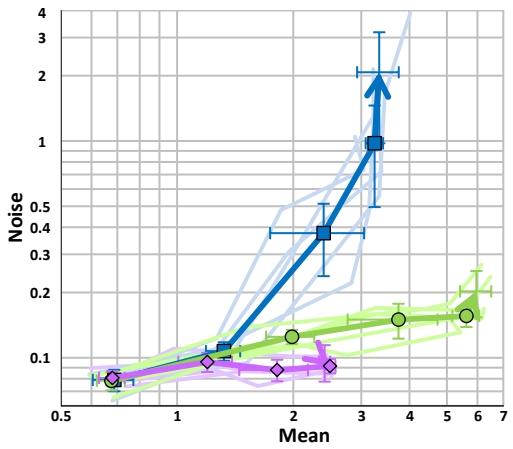
Figure 9

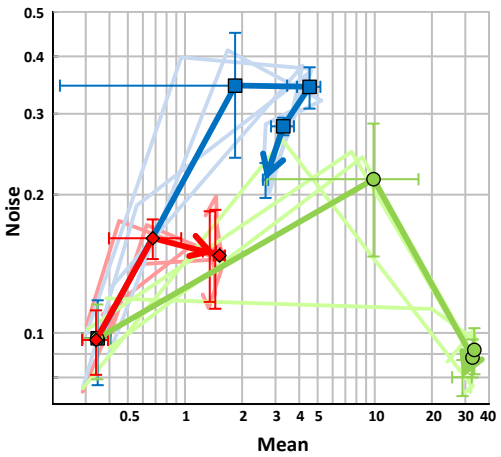
a *malP* on M9 + Glucose + cAMP

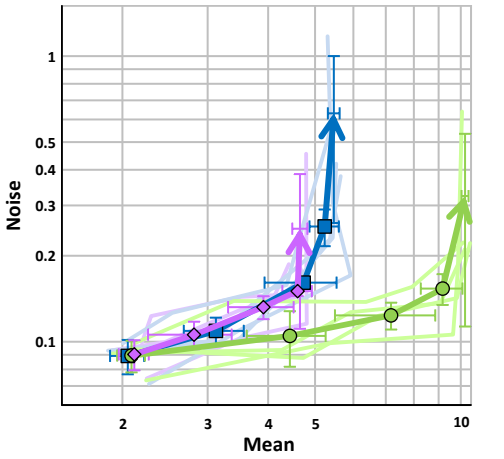
b *malP* on M9 + Glycerol

c *malK* on M9 + Glucose + cAMP

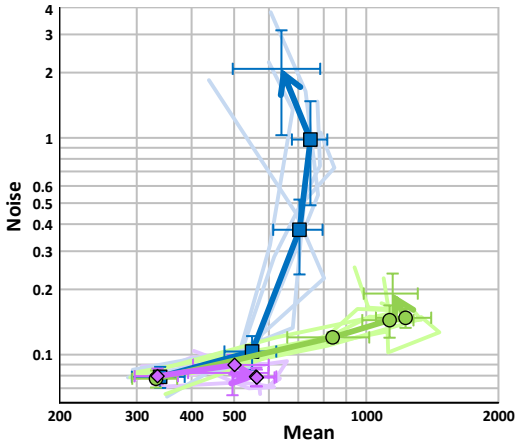
YFP / RFP

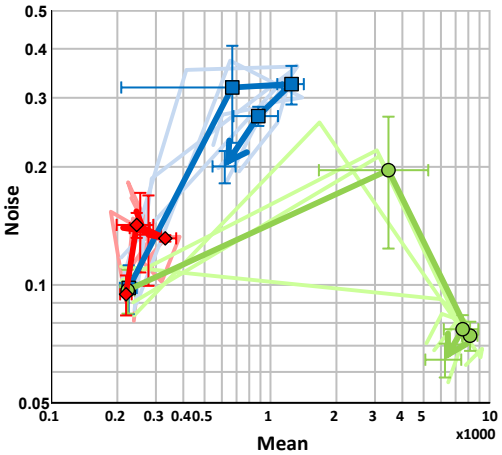






YFP





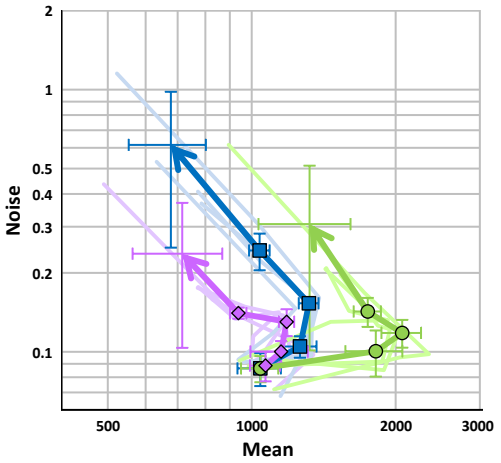


Figure 9 – Noise as a function of mean

Log scale plots of noise against mean values of fluorescence distributions. Light lines represent the values of the four repetitions. Bold lines represent the mean values of the [mean,noise] values of the four repetitions, and error bars are the standard deviation of these values on each of the axes. Arrows represent time flow.

Upper panels present the mean and noise values based on YFP-to-RFP ratio distributions, while the lower panels present the mean and noise values based on YFP distribution. Noise was defined as the variance divided by the mean square.

4.2.2 Sorting experiment (preliminary results)

For a bet hedging strategy, it is necessary to show a direct link to fitness. In this case, it is essential to show that the cells that express the maltose operons highly in response to lactose have higher fitness in maltose compared to cells that expressed their maltose genes lowly. In order to compare the fitness of the two populations on maltose, I used a cell sorter to separate between the high and low expressing cells that were exposed to lactose, and then grew them on maltose and extracted their growth curves (see methods section 3.6 for more information).

Gates for the sorting were defined in the FACSDiva software as it is the only software that is connected to the cell sorter, thus I could not use the Matlab GUI and used the FACSDiva gating tools instead. First, a gate for the central region of the FSC-SSC distribution was defined in a similar way to the definition in the Matlab GUI. Second, a gate for the main area of the RFP distribution was defined (roughly the mean±~10%) on the histogram of RFP values. Lastly, on the YFP histogram of the gated cells, two gates were defined for the sorting: (1) right tail gate (~20% of the gated population from the highest fluorescence values), and (2) left tail gate close to the center of the population (~20% of the gated population, the right border of the gate was at a distance of ~3,000 fluorescence units from the left border of the right tail gate, and the left border was defined accordingly to reach to ~20% of the gated population). The YFP and RFP gates were changed during the experiment to accommodate the apparent shift in fluorescence.

The sorted cells were then grown for 6 hours in a plate reader, and their growth curves were extracted and presented in Figure 10. Figure 10a presents the mean OD of all the wells of each population as a function of time, and a linear fit on each of the curves. As can be seen, there is a very small difference between the two curves, the left tail curve, i.e. the low expressing cells, have slightly higher slope, namely, growth rate was slightly higher in the population of lower promoter activity level, in contrast to

the expectations from a bet-hedging scenario. Figure 10b presents a mean normalized OD, which is the OD of each well divided by the mean OD of the four surrounding wells which are of the other population (according to 'checkerboard' pattern). This procedure is done to reveal small differences that might be masked by variability that is caused by uncontrolled effects of the location of wells on the plate on growth rate. As can be seen, the right tail cells show a decrease in their normalized OD levels, thus the right tail cells have a lower growth rate on maltose, again, opposite than expected. One possible reason for the negative results is the time passed between activation and sorting (~4 hours) in which cells could change their maltose operon activation levels. YFP fluorescence level of the two populations, as measured using the plate reader was almost identical (not shown). In the discussion section 5.2, another explanation is given to these negative results.

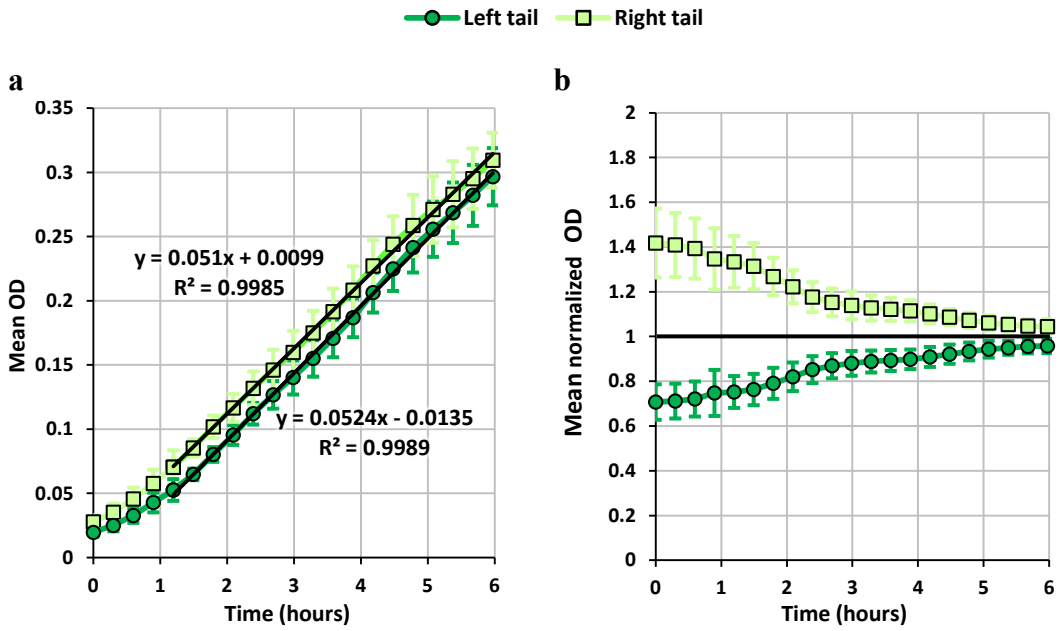


Figure 10 – Growth curves on maltose of the sorted cells
Growth curves of cells sorted according to their YFP expression (*malP* promoter). (**a**) The mean OD of all the wells of each population, error bars presents the standard deviation. A linear fit is marked with a black line. (**b**) Mean normalized OD, which is the OD of a well divided by the mean OD of the four surrounding wells (above, below and to the sides of a given well) from the other population, averaged by these ratios from all the wells of a given populations, error bars presents the standard deviation of these ratios.

5. Discussion

In this study, I continued the study of *Mitchell et al.* and explored evolutionary conditioning of *E. coli* response towards sugars at the single cell level.

One of the possibilities I wanted to investigate is the existence of a sort of a bet-hedging strategy in this system. In *Mitchell et al.* study, bet-hedging was presented as the opposing strategy to evolutionary conditioning, as in pure bet-hedging there are no predictions, cells switch stochastically between states and a portion of the population is always prepared for environmental change, while in evolutionary conditioning cells use external signals in order to activate their response to the sequential environmental change. However, a combination of these two extreme strategies is also possible. In such a combined strategy, cells will use an external signal in order to activate the response to the sequential environmental change, but they will do it in a stochastic manner, some cells will be activated while others will not. This strategy may be advantageous in environments in which the signal and the following change are connected, but it is not guaranteed that the following change will always come. Such environments might be more common than environments in which the connection is absolute.

Bet-hedging strategy would be manifested in this system if the signal causes wider activation distribution than the environmental state to which this response is directed. In the most prominent manifestation, it will lead to a bimodal distribution in which a portion of the population will be activated, and another portion will remain inactive. Using flow cytometry, I examined the activation distribution of cells harboring a fluorescent protein reporter plasmid for various maltose operon promoters in response to lactose and maltose. In accord with *Mitchell et al*. results, lactose leads only to partial activation of the maltose operons regarding the mean fluorescence. However the activation distribution is much wider in the lactose treated cells. In the case of *malP* in the glucose medium, the symmetry of the distribution is also disrupted in response to lactose, and an extended right tail of highly activated cells appears. Nonetheless, true bimodal distribution could not be identified in any of the treatments. The noise as a function of mean plots present elevation in mean coupled with elevation in noise in the lactose-treated samples, and points having similar means show higher noise in the lactose-treated samples, not according to the general scaling between noise and mean.

These findings might imply a combined conditional-bet-hedging strategy in this system. In the next sections I will discuss a possible molecular mechanism for the evolutionary conditioning, aspects of the sorting experiment, and future directions.

5.1 A possible mechanism for the partial maltose operons activation in response to lactose

One of the questions not fully addressed in *Mitchell et al.* study, is how mechanistically the activation of maltose operons in response to lactose is achieved. This question is even more critical in respect to the current study, as another question arises, how does the same promoter give rise to different activity distributions, and noise levels in particular, in response to different stimuli?

In order to try and address these two questions, I will first describe here, briefly, the maltose regulon. A comprehensive review can be found in [33]. The maltose regulon functions in the metabolism of maltodextrins, polymers of glucose molecules joined in an α(1→4) bond, among which is maltose that consists of two such molecules. It has a complex network of regulation which has been the subject of an ongoing research since Jacob and Monod, and some parts of it are still missing. Here, only a few of the nodes and edges of this network will be described (Figure 11).

MalT, the transcription activator of the maltose regulon, as other sugar operons in bacteria, is under catabolite repression, and thus its transcription depends on the presence of the complex of Catabolite Activator Protein (CAP) and cAMP whose concentration is low in glucose rich environment. Conversely, it is also under the regulation of a global repressor, Mlc, which, when glucose is transported, binds to the unphosphorylated glucose transporter and thus relieves the repression. MalT activity as a transcription activator depends on the presence of ATP and the inducer molecule. The inducer molecule in the maltose regulon is not maltose itself, as in the lactose operon, but maltotriose, a maltodextrin consists of three glucose molecules. Maltotriose is present in the cell even in the absence of maltodextrins in the media as a product of glycogen metabolism and gluconeogenesis. The maltose metabolism end products, glucose and glucose-1-phosphate are thought to be used for the endogenous formation of maltotriose. The regulon's cytoplasmic metabolic enzymes (MalP, MalQ and MalZ) also affect maltotriose concentration. In addition to the transcriptional regulation, MalT activity is regulated by interaction of at least three proteins, one of them is MalK, a subunit of the maltose transporter: when it is not bound to the

transporter, it binds to MalT and inhibits its activity as a transcription activator, thus the maltose regulon will be fully activated only when maltodextrins are transported into the cell, and MalK participates in the transporting.

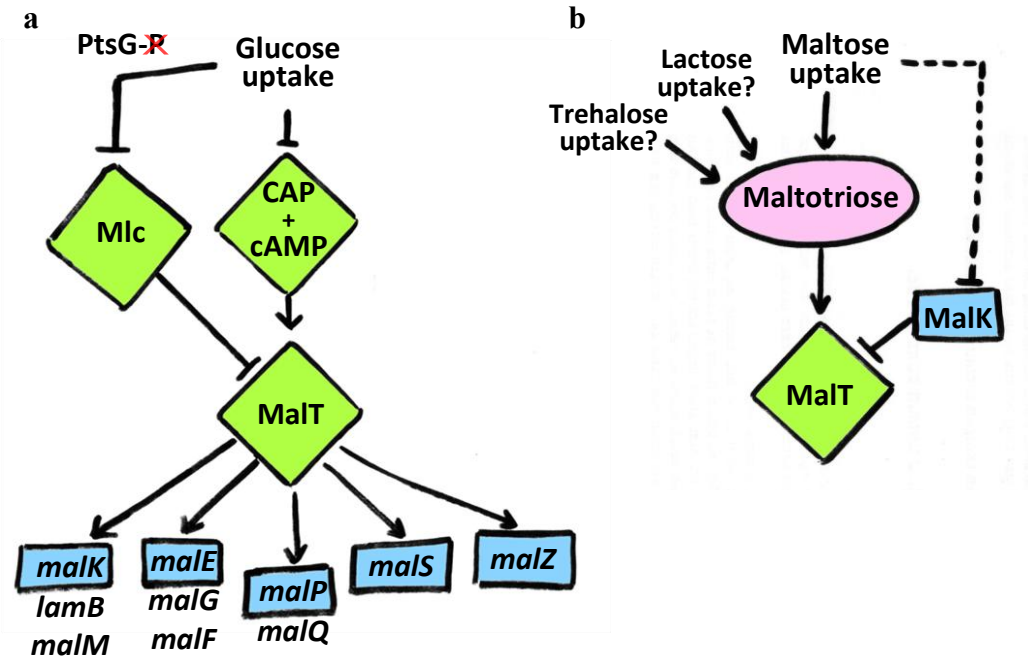


Figure 11 – Schematic representation of part of the regulation network in the maltose regulon (**a**) **The transcriptional regulation network**. MalT, the activator of the maltose regulon, is under the regulation of CAP-cAMP complex, which activated its transcription, and Mlc, which repress its transcription. Glucose uptake leads to the repression of adenylate cyclase (not shown), and thus lowers the intracellular cAMP levels and the activation of *malT* transcription by CAP-cAMP complex. In addition, glucose uptake leads to dephosphorylation of the glucose transporter PtsG, to which Mlc can bind in the unphosphorylated state, thus lowering the repression of *malT* transcription by Mlc. (**b**) **Post-transcriptional regulation**. MalT functions as a transcription activator only in its active form. Its inducer is maltotriose, which is produced in higher rate in response to uptake of maltose, and maybe trehalose and lactose as well. In addition, MalT is inhibited by MalK, a subunit of the maltose transporter, when maltose is not transporter to the cell. When the transporter is active, MalK binds to the transporter and releases MalT.

This network, which is not yet fully characterized, could be the basis of complex responses of the maltose regulon to various stimuli. For instance, besides lactose, also trehalose was shown to partially activate the maltose regulon. Trehalose is a disaccharide that like maltose consists of two glucose molecules, linked in an α(1→1)α bond. It was suggested that the products of trehalose metabolism, glucose and glucose-6-phosphate (which then can be converted to glucose-1-phosphate) are used for intracellular maltotriose production as in the case of maltose and thus lead to the activation of the maltose regulon [33].

Lactose metabolism end products are glucose and glucose-1-phosphate as well, so activation of the maltose regulon by lactose might be reached by its metabolism. Two unpublished findings in *Mitchell et al.* study are in accord with metabolism-dependent activation: (1) lacZ deletion mutants do not show maltose regulon activation in response to lactose; (2) IPTG, a synthetic inducer of the lactose operon does not activate the maltose regulon. One complication is the result of the extinction experiment - lab evolution under constant lactose without maltose that caused disappearance of the maltose activation in response to lactose, while leaving the maltose regulon responsive to maltose. If this is indeed a simple metabolic activation, and given that metabolism has not been changed, a distinct mechanism is required to achieve extinction. However, in such an elaborate network of regulation as found in the maltose regulon, such mechanism could be found.

Towards this end I have re-analyzed the whole genome sequencing results of the extinction strain. One of the mutations I found in this strain is in the *ptsG* gene, which encodes the glucose transporter to which Mlc, a MalT repressor, binds. This mutation could have an influence on the maltose regulon, and may be the basis for the extinction phenotype. This hypothesis could be examined, for instance, by introducing the same mutation into a wild-type strain, where it would be expected to induce extinction of the maltose response to lactose.

Metabolism-dependent activation might not be the only possibility of maltose regulon activation by lactose, as MalT might be activated in a different fashion, or a different, yet undiscovered layer of regulation might exists in the maltose regulon.

Until now only a possible mechanism for the regulon activation by lactose was discussed. However, the question of how the different activation distributions are achieved remained.

In the lactose operon, as was mentioned before, induction by intermediate concentration of IPTG or TMG, another synthetic inducer, causes bimodal distribution of the lactose operon activation distribution. It is tempting to postulate the involvement of bistability in lactose induction in the noisy induction of maltose regulon in response to lactose. For instance, taking the lactose metabolism-dependent activation into account, bistability in the maltose regulon can be obtained if cells that induce the lactose operon will have more of the lactose metabolism products and will activate the maltose regulon to a higher level than cells that do not induce the lactose operon. However, so far the lactose operon showed bistability only to the synthetic

inducers and it is under debate whether lactose itself could induce bistability in the system at all [34]. In the current study, the fluorescence distribution of the *lacZ* promoter showed no bistability in response to lactose, thus the involvement of the lactose operon activation in the noisy maltose regulon activation might not be very plausible.

However even without the direct involvement of the bistable lactose operon activation, a noisy maltose regulon activation by lactose might be achieved. Using the bistable lactose operon as an example: bistability in the lactose operon was obtained by synthetic inducer exposures only in a limited range of intermediate concentrations. The bistability in the arabinose operon of *E. coli* [23], and the galactose signaling network of *S. cerevisiae* [24], appears only in intermediate levels of inducer as well. If lactose metabolism leads to intermediate levels of maltotriose, or, alternatively, maltotriose is produced to the same level as in maltose, but other regulations of MalT are not relieved by lactose (such as the transport dependent MalK de-inhibition), this will lead to intermediate levels of active MalT. Those intermediate levels of active malT, which might be achieved in other ways of regulation as well, might in turn result in higher noise in the maltose regulon activation by lactose than in maltose. An interesting experiment will be to use lower concentrations of maltose than the concentrations used in this study, and compare the activation distribution to the lactose treated samples in order to see whether such a noisy activation could be achieved.

All these possibilities are rather speculative, and their plausibility needs to be examined experimentally, possibly by using mutants of the regulatory network components, such as MalT, PtsG and Mlc; closer examination of the mutation that the extinction experiment revealed in PtsG; and using the three color plasmid in order to simultaneously measure different promoters of the regulon and the regulators in the same cell, and elucidate various connections in the regulatory network.

5.4 Future directions

A prerequisite for a bet hedging strategy is the link between the noise in gene expression, or the cellular response, and fitness. In this system, it is necessary to show that cells that activate the maltose regulon to a higher level during lactose exposure have: (1) lower fitness in the lactose environment as they are expressing unnecessary

genes; and (2) higher fitness in the maltose environment as they were prepared in advance for maltose arrival. For this purpose, the sorting experiment was preformed.

One not trivial assumption of using the fluorescent protein reporters for promoter activity is that they correlate with the concentrations of the original proteins that those promoters regulate, the maltose regulon proteins that are supposed to affect the fitness. This assumption might be valid, as effect of regulation by the same transcription factor, might be stronger than other extrinsic influences. Indeed in their study, *Cox et al.*, who developed the three color plasmid, demonstrated a high correlation between RFP and YFP expressed from different promoters on the plasmid, both regulated by LacI [27]. However, if the correlation between the original maltose operon and YFP is not that high, for instance, in case that there is an influence of the genomic versus plasmidic location of the promoter and the gene, or due to potential high intrinsic noise, this might be yet another explanation for the negative results.

Ideally, in order to measure the original protein level, one could use a fusion protein of the original protein with a fluorescent protein. I tried to use strains of the recently constructed YFP genomic fusion library [32] for this purpose, but using flow cytometry I could not detect any fluorescence above auto-fluorescence in the maltose exposed cells. An alternative approach will be to use a plasmid construct containing the fusion protein, or even a transcriptional fusion, having the fluorescence protein on the same operon, as operons were shown to buffer noise [34].

An alternative to the sorting experiment is the more sensitive time-lapse microscopy, in which single cells can be tracked and measured for fluorescence intensity and growth rate. This experiment could give even more information regarding lineages, and switching rates. We recently started collaboration with Prof. Naama Barkai's lab (Dr. Nurit Avraham Tayar) to perform this experiment, using a live imaging system they developed.

To conclude, in my study I showed evidence for bet-hedging in the *E. coli* evolutionarily conditioned response towards sugars. Showing the connection between expression levels in lactose to fitness in maltose will be necessary in order to establish this as a bet-hedging strategy. The results of the different activation distributions in response to each of the stimuli raised again the question regarding the mechanism behind the evolutionary conditioning and the extinction phenotype. Elucidating this question is essential in order to understand how the two distinct distributions are achieved.

6. Literature

1. Mitchell A, Romano GH, Groisman B, Yona A, Dekel E, Kupiec M, Dahan O, and Pilpel Y (2009) Adaptive prediction of environmental changes by microorganisms. *Nature* 460, 220-224.
2. Bauer FF, and Pretorius LS (2000) Yeast stress response and fermentation efficiency: how to survive the making of wine – a review. *S. Afr. J. Enol. Vitic.* 21, 27-51.
3. Tagkopoulos I, Liu YC, Tavazoie S (2008) Predictive behavior within microbial genetic networks. *Science* 320, 1313-1317.
4. Schild S, Tamayo R, Nelson EJ, Qadri F, Calderwood SB, and Camilli A (2007) Genes induced late in infection increase fitness of *Vibrio cholerae* after release into the environment. *Cell Host Microbe* 2, 264-277.
5. Rodaki A, Bohovych IM, Enjalbert B, Young T, Odds FC, Gow NAR, and Brown AJP (2009) Glucose promotes stress resistance in the fungal pathogen *Candida albicans*. *Mol. Biol. Cell* 20, 4845–4855.
6. Mitchell A, and Pilpel Y (2011) A mathematical model for adaptive prediction of environmental changes by microorganisms. *Proc. Natl. Acad. Sci. USA* 108, 7271-7276.
7. Pavlov IP (1927) Conditioned Reflexes: an investigation of the physiological activity of the cerebral cortex. *Oxford University Press*.
8. Savageau MA (1998) Demand theory of gene regulation II. Quantitative application to the lactose and maltose operons of *Escherichia coli*. *Genetics* 149, 1677-1691.
9. Dahlqvist A, and Borgstrom B (1961) Digestion and Absorption of Disaccharides in Man. *Biochem. J.* 81, 411-418.
10. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, and Alon U (2006) A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* 3, 623-628.
11. Kaplan S, Bren A, Zaslaver A, Dekel E, and Alon U (2008) Diverse two-dimensional input functions control bacterial sugar genes. *Mol. Cell* 29, 786–792.
12. Eldar A, and Elowitz MB (2010) Functional roles for noise in genetic circuits. *Nature* 467, 167-173.

13. de Jong IG, Haccou P, and Kuipers OP (2011) Bet hedging or not? A guide to proper classification of microbial survival strategies. *BioEssays* 33, 215–223.
14. Bigger JW (1944) Treatment of staphylococcal infections with penicillin by intermittent sterilisation. *Lancet* 244, 497-500.
15. Balaban NQ, Merrin J, Chait R, Kowalik L, and Leibler S (2004) Bacterial persistence as a phenotypic switch. *Science* 305, 1622-1625.
16. Balazsi G, van Oudenaarden A, and Collins JJ (2011) Cellular decision making and biological noise: from microbes to mammals. *Cell* 144, 910-925.
17. Süel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, and Elowitz MB (2007) Tunability and noise dependence in differentiation dynamics. *Science* 315, 1716-1719.
18. Maamar H, Raj A, and Dubnau D (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science* 317, 526-529.
19. Chung JD, Stephanopoulos G, Ireton K, and Grossman AD (1994) Gene expression in single cells of *Bacillus subtilis*: Evidence that a threshold mechanism controls the initiation of sporulation. *J. Bacteriol.* 176: 1977–1984.
20. Veening JW, Stewart EJ, Berngruber TW, Taddei F, Kupiers OP, and Hamoen LW (2008) Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc. Natl. Acad. Sci. USA* 105, 4393-4398.
21. Nachman I, Regev A, and Ramanathan S (2007) Dissecting timing variability in yeast meiosis. *Cell* 131, 544-556.
22. Novick A, and Weiner M (1957) Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci. USA* 43,553-566.
23. Siegele DA, and Hu JC (1997) Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci. USA* 94, 8168–8172.
24. Acar M, Becskei A, and van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435, 228-232.
25. Acar M, Mettetal JT, and van Oudenaarden A (2008) Stochastic switching as a survival strategy in fluctuating environments. *Nat. Gen.* 40, 471-475.
26. Beaumont HJE, Gallie J, Kost C, Ferguson GC, and Rainey PB (2009) Experimental evolution of bet hedging. *Nature* 462, 90-93.
27. Cox RS, Dunlop MJ, and Elowitz MB (2010) A synthetic three-color scaffold for monitoring genetic regulation and noise. *J. Biol. Eng.* 6, 306.

28. Sharma UK, and Chatterji D (2010) Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of σ^{70} activity. *FEMS Microbiol. Rev.* 34, 646–657.
29. Parks DR, Roederer M, and Moore WA (2006) A new "Logicle" display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. *Cytometry A* 69A,541–551.
30. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, and Barkai N (2006) Noise in protein expression scales with natural protein abundance. *Nat. Genet.* 38, 636 – 643.
31. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, and Weissman JS (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* 441, 840-846.
32. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, and Xie XS (2010) Quantifying *E. coli* proteome and transcriptome with single molecule sensitivity in single cell. *Science* 329, 533-538.
33. Boos W, and Shuman H (1998) Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism and regulation. *Microbiol. Mol. Biol. Rev.* 62, 204-229.
34. Davidson CJ, and Surette MG (2008) Individuality in Bacteria. *Annu. Rev. Genet.* 42, 253–68.
35. Tabor JJ, Bayer TS, Simpson ZB, Levy M, and Ellington AD (2008) Engineering stochasticity in gene expression. *Mol Biosyst.* 4, 754-761.