Detection of transcriptional triggers in the dynamics of microbial growth: application to the respiratorily versatile bacterium *Shewanella oneidensis*

Qasim K. Beg[1], Mattia Zampieri[2,3], Niels Klitgord[2], Sara B. Collins[2], Claudio Altafini[3], Margrethe H. Serres[4] and Daniel Segrè[1,2,5,*]

[1]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA, [2]Bioinformatics Program, Boston University, Boston, MA 02215, USA, [3]SISSA (International School of Advanced Studies), 34136 Trieste, Italy, [4]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, USA and [5]Department of Biology, Boston University, Boston, MA 02215, USA

ABSTRACT

The capacity of microorganisms to respond to variable external conditions requires a coordination of environment-sensing mechanisms and decision-making regulatory circuits. Here, we seek to understand the interplay between these two processes by combining high-throughput measurement of time-dependent mRNA profiles with a novel computational approach that searches for key genetic triggers of transcriptional changes. Our approach helped us understand the regulatory strategies of a respiratorily versatile bacterium with promising bioenergy and bioremediation applications, *Shewanella oneidensis*, in minimal and rich media. By comparing expression profiles across these two conditions, we unveiled components of the transcriptional program that depend mainly on the growth phase. Conversely, by integrating our time-dependent data with a previously available large compendium of static perturbation responses, we identified transcriptional changes that cannot be explained solely by internal network dynamics, but are rather triggered by specific genes acting as key mediators of an environment-dependent response. These transcriptional triggers include known and novel regulators that respond to carbon, nitrogen and oxygen limitation. Our analysis suggests a sequence of physiological responses, including a coupling between nitrogen depletion and glycogen storage, partially recapitulated through dynamic flux balance analysis, and experimentally confirmed by metabolite measurements. Our approach is broadly applicable to other systems.

INTRODUCTION

The transcriptional dynamics of microbial growth depends both on the physico-chemical properties of the external environment and on the internal state of the cell (e.g. its growth rate) (1–3). The time-dependent array of environmental stimuli and the wiring of the underlying regulatory network jointly determine the dynamical changes in gene expression (4, 5). Understanding the interplay between these two elements is an important open challenge, especially relevant for metabolically versatile microbes that may rapidly cycle through different environments. Given that a microbial cell needs to dynamically adapt to changing environmental conditions, transient changes in protein expression must be triggered by condition-specific external factors, which interact with the intracellular genetic network. However, other transcriptional programs, such as those involved in cellular division, must be robust to influences external to the cell. Thus, it may be possible to discriminate between

environment-independent circuits (e.g. dedicated to controlling core cellular functions, including growth) and portions of the genetic networks that have specifically evolved to mediate environmental sensing (6–8). The interplay between genetic circuitry and external variables is particularly important during growth phase transitions. Microorganisms are able to sense the depletion of environmental resources and respond with a variety of condition-specific strategies, such as sporulation (9,10), increased motility (11) and the formation or dispersal of biofilms (12,13). DNA microarrays have been widely used to monitor the time-dependent response of various microorganisms to environmental change, showing that their fast physiological adaptation is typically accompanied by a massive transcriptional rearrangement, where selected genes dynamically adjust their level of expression as a function of the external stimuli (14–17).

Here, we combine experimental and computational approaches to study the transcriptional response upon growth phase transitions for the versatile gram-negative bacterium *Shewanella oneidensis* MR-1. *Shewanella oneidensis* is found mainly in nutrient-rich environments, such as marine sediments, often in association with fermentative communities (18–20). *Shewanella* species have the characteristic ability to reduce a broad spectrum of metals and organic compounds (in addition to oxygen), making them very relevant for environmental and bioenergy-related applications, such as bioremediation of metal-contaminated waters, biofuel production and microbial fuel cells (18,19,21–24). Several prior studies on *S. oneidensis* cultures have characterized transcriptional and proteomic responses of this microbe under various media and stress conditions (25–38). These comparative studies have mostly focused on 'binary' changes, i.e. on comparing transcriptional states before and after a given perturbation or environmental shift. However, to our knowledge, there has been no detailed measurement and systems level analysis of transcript levels at multiple time points as *S. oneidensis* transitions between exponential and stationary phase of growth.

Our results are organized as follows: first, we introduce our measurements of time-dependent mRNA expression levels during batch growth of *S. oneidensis* MR-1 under two radically different growth media compositions (minimal lactate and rich LB, i.e. Luria and Bertani's Lysogenic Broth) and identify global transcriptional trends. We then implement a new growth derivative mapping (GDM) approach to compare transcriptional profiles across different time-course experiments and to discriminate genes (and processes) most likely controlled by growth-associated functions. Next, we dive into a detailed system-level analysis of the genetic response to growth phase transitions under the lactate-minimal medium. Specifically, we use a new mathematical approach (dynamic detection of transcriptional triggers or D2T2) to identify the genes through which environmental stimuli are expected to affect the internal dynamics. Our analysis highlights the importance of some specific pathways, whose metabolic relevance is confirmed by dynamic flux balance analysis (dFBA) calculations. In particular, we characterize some aspects of the

transcriptional response to oxygen limitation, detecting the activation of genes previously shown to be relevant for anaerobic respiration. Moreover, we find that nitrogen limitation is coupled to storage of glycogen. Both observations are corroborated by measurement of relevant intracellular and extracellular metabolites, as well as by complementary analyses of literature information and competitive fitness assay data.

MATERIALS AND METHODS

Chemicals and reagents

DL-Lactate (60% solution) and ammonia analysis kit were procured from Sigma-Aldrich (St Louis, MO, USA). All other components of the M4 minimal medium (Supplementary Table S1 in Supplementary text) were of highest purity grade and were also procured either from Sigma-Aldrich or Thermo-Fisher Scientific (Pittsburgh, PA, USA). Qiagen Inc. supplied the RNA protect reagent, RNAse easy kit for isolation of RNA, cDNA purification kit and RNAse-free DNAse enzyme. Other reagents and chemicals used during isolation and purification of RNA and during various steps of *Shewanella* chips hybridization (Affymetrix Inc.) were purchased from several different vendors: Superscript II reverse transcriptase, DTT, random hexamers and BSA from Invitrogen Inc.; Gene chip labeling reagent, One-phor-all buffer and B2 oligo from Affymetrix; DNAse from Pierce Biochemicals; MES stock, lysozyme, Goat IgG and 200 proof ethanol from Sigma-Aldrich; Terminal transferase, Herring sperm DNA and dNTPs from Promega; 0.5 M EDTA solution from Gibco; Biotinylated Anti-Streptavidin antibody from Vector laboratories; SSPE, Streptavidin, SAPE, 10% Tween-20, NaOH and HCl from Thermo-Fisher Scientific; and TE Buffer (pH 8.0), Superase 1 n, 5 M NaCl and nuclease-free water were from Ambion.

Strain, cultivation and sample collections

Shewanella oneidensis MR-1 ATCC 700550 was used in this work. The strain was revived from –80°C glycerol stocks by overnight growth in LB medium. One hundred microliters of an overnight *S. oneidensis* MR-1 pre-culture was inoculated in M4 minimal medium containing lactate (LAC) (Supplementary Table S1) and rich LB medium separately for the inoculum, which was used for inoculation in 1.3l working volume Bioreactor vessel (Bioflo110, New Brunswick Scientific Company) for M4-lactate and LB media runs, respectively. Various growth parameters, viz., temperature (30°C), pH (7.2), aeration (1 l/m) and agitation were controlled using microprocessor probes. The pH was maintained at 7.2 using automatic additions of 2 N NaOH and 10% H_3PO_4 using peristaltic pumps attached to the bioreactor. The dissolved oxygen (dO_2) probe was calibrated before the inoculation and dO_2 in the vessel was maintained at 20% air saturation level using automatic control of O_2 cascade throughout the experiment. We started collecting biomass samples for RNA isolation after the optical density of culture was at least above 0.150. In lactate-minimal medium (LAC), the

time interval for most of the biomass samples between collections was between 1 and 6 h, except after 36 h only two samples were collected at 48 and 50 h. For LB medium, the biomass samples were collected every 30 min between 1.5 and 6 h of growth (exponential phase); however, after late exponential and into stationary phase, the biomass samples were collected at time intervals of 1–4 h until 36 h of growth. Three more samples at 48, 50 and 55 h near the end of bioreactor run were also collected. We collected two biomass samples during each collection, and both the samples were processed for RNA extraction to provide statistical significance in subsequent microarray results.

Extracellular lactate, acetate, pyruvate, ammonium and intracellular glycogen quantification

High pressure liquid chromatography (HPLC) was used for quantifying residual lactate and other organic acids (acetate, pyruvate) produced during growth of *S. oneidensis* MR-1 on lactate. The BioRad column (HPX 87H 300 mm × 7.8 mm) was used on a HPLC (HP-Agilent 1090 Series II). The quantification was carried out by injecting 20 μl of sample using 0.008 N H_2SO_4 in mobile phase at 35°C. The flow rate during quantification was kept at 0.6 ml/h and detection was done at 210 nm. The regression equations from individual standard curve of all three compounds were used to calculate the amount of individual compound from the test sample. The retention times for lactate, acetate and pyruvate were 12.5, 14.9 and 8.9 min, respectively. Residual lactate was also measured using enzymatic analysis kit (R-Biopharm). Residual ammonium in the cell-free supernatant and intracellular glycogen were measured using commercially available kits for ammonium (Sigma-Aldrich, St Louis, MO, USA) and glycogen (BioAssay Systems, Hayward, CA, USA), respectively.

RNA isolation and *S. oneidensis* microarrays

At each time point of interest, 2–30 ml (depending on phase of growth and cell concentration) of live bacterial culture were sampled, immediately mixed with two volumes of Qiagen RNA protect solution, vortexed and centrifuged, and RNA was isolated from biomass pellet using Qiagen's RNA easy isolation kit according to manufacturer's recommendation. During isolation procedure, the RNA sample was treated with Qiagen's DNAase to get rid of any genomic DNA contamination. RNA yields were quantified on Nanodrop and UV 260/280 ratios were calculated to check purity of each RNA sample. The previously described protocol (23,31) was used for microarrays on *S. oneidensis* chips from Affymetrix. Briefly, ~10 μg of each RNA sample were used for cDNA synthesis, followed by reverse transcription, cDNA purification and cDNA fragmentation. This was followed by labeling of cDNA and 16 h of hybridization at 45°C on *S. oneidensis* arrays. The labeled and hybridized arrays were subjected to several cycles of washing and staining using Affymetrix Wash buffers A and B, Goat IgG, Streptavidin, Anti-streptavidin and SAPE according to Affymetrix protocol for prokaryotic arrays. This was followed by scanning of stained arrays on Affymetrix Gene Chip Scanner Model 3000. The *S. oneidensis* MR-1 platform and microarray data has been submitted to Gene Expression Omnibus (accession numbers: GPL8434 and GSE25821 for platform and data, respectively).

Growth derivative mapping

We developed the GDM method to compare gene expression profiles during growth under two nutrient conditions. The details of GDM are available in the Supplementary Text. GDM maps gene expression profiles from a time-dependent to a growth rate-dependent domain, through the following steps: first, we fit the observed OD curves to a logistic growth model of bacterial growth. Then, by matching time points corresponding to similar instantaneous growth rates in the two conditions, we generate a non-linear mapping between the rich and the minimal medium profiles, which we apply systematically to each individual gene. Finally, profiles across conditions can be compared, with the goal of finding genes whose time-dependent expression levels display significant correlation or anti-correlation across the two conditions.

D2T2 algorithm

The D2T2 approach is aimed at integrating gene expression time courses with steady-state profiles to identify genes that are primarily involved in the response to external perturbations. Complete method details can be found in the Supplementary Text. Here, we provide a brief summary.

The main input of the D2T2 algorithm, which ultimately provides information about the trigger genes, is a set of time course data of gene expression. However, the algorithm requires a preliminary step, which yields an influence gene network based on steady-state profiles. In this first step, the influence network model is trained on an independent large compendium of gene expression measurements taken under the assumption of steady-state conditions (39), in analogy to previous reverse engineering approaches (40–43). This initial procedure, which we recently developed and applied to the study of prion infection (44) is further shown here to provide accurate predictions on an *Escherichia coli* multiple perturbation dataset (see Supplementary text). The fundamental simplifying assumption of this model is that the rate of change in abundance of gene i (i.e. $\dot{x}_i$) can be expressed as a linear combination of all other gene expression levels. Under steady state, this assumption reduces our model to a system of algebraic equations $Ax = 0$, where A is a sparse $n \times n$ matrix (n is the number of genes) and each non-zero element (i.e. a_{ij}) in A represents the influence of gene j on gene i. Solution to this first preliminary step is based on a relevance network approach, where pairs of genes showing a high linear proportionality across multiple conditions are first selected in order to make the influence matrix (i.e. A) sufficiently sparse. These relationships are further weighted through a multiple linear regression scheme, which evaluates the influence of a gene

node in driving expression changes of the putative interacting ones. A gene network model is then selected on the basis of a Bayesian criterion, which identifies the best compromise between the model complexity and its predictive ability (See Supplementary text). It is worth noting that the solution of the regression problems is a weighted asymmetric matrix, with all diagonal elements identical to 1 resulting from rescaling each row of A by an undetermined scaling factor (i.e. the diagonal term a_{ii}). We named this rescaled matrix A_c.

In a second step (the core of the D2T2 procedure), we go beyond the static influence matrix A_c computed earlier and explicitly incorporate the measured dynamic changes in the model. This is achieved by modeling the rate of change in abundance of each gene as a linear combination of the 'influencing-gene' expression levels plus an additional external input (i.e. u). Here, we approximate the external function $u(t)$ as a unique step function $[u(t) = 1]$, leading to the following formulation of the problem: $\dot{x} = Ax + B$. Here, the matrix B represents the influence that the perturbation u has in driving changes in the transcriptional rate of each gene. Hence, the condition-specific evolution in time of the system is described by a large-scale system of linear differential equations (ODEs) for the time-dependent gene expression profiles. Typically, the complexity of this type of model is limited by the sampling frequency and number of time points. In order to overcome these limitations, we take advantage of the initial steady-state model to reduce the fitting of the gene expression profiles to only two parameters (i.e. for gene x_i we fit c_{ii} and b_i, see Supplementary text for full details), and hence we write $\Delta x = CA_c\, x + B$, where Δx is the approximation of the continuous derivative in time using the Euler scheme and C is a diagonal matrix ($n \times n$) consisting of normalizing factors c_{ii}.

The identification of these two parameters for each single transcript aims at describing the portion of dynamical changes caused by external factors (i.e. B) from rearrangements induced by internal transcriptional regulatory mechanisms (A_c). In particular, the genes showing an incoherent behavior relative to the initial model (e.g. transient evolution) are chosen as the best candidates for direct targets of environmental changes. The underlying network and the significantly perturbed genes are identified through a rigorous statistical analysis. Once the distribution of empirical P-values has been generated, the Q-value is used to correct the index of significance for multiple testing. Categories associated to a Q-value in the 1% confidence interval are considered to be significant (similar results are obtained for the 5% confidence threshold).

Predicting growth dynamics using a flux balance model

dFBA was implemented as described previously (45), using a static optimization method. The detailed dFBA optimization method is described in Supplementary Text. Here, we provide a brief summary. dFBA performs iterations of flux balance analysis (46,47) with flux bounds computed based on external metabolite concentrations updated at each time step. This approach assumes that biomass and environmental metabolite changes happen at a slower scale relative to the time necessary for the achievement of intracellular steady state of metabolite levels. The *S. oneidensis* MR-1 genome-scale metabolic model iSO783 (48) was used to predict growth of the bacteria in the LAC experimental condition. This model contains 774 reactions and 634 unique metabolites. Initial medium composition was modeled by assuming an initial concentration of nutrients that reflects the known M4 medium components (Supplementary Table S1) with supplemented D- and L-lactate. The simulated level of external O_2 was maintained throughout the dFBA simulation. To test the effect of O_2 on overall growth dynamics, we run the simulation under different constant concentration values, between 2 and $8\,mM/h$. In order to simulate the production of glycogen, we implemented a secondary objective function which maximizes a glycogen sink flux (to what could be thought of as a glycogen storage compartment) when biomass cannot be produced. This is done by imposing an objective function which is a linear combination of biomass production and glycogen production, in proportions of 99 to 1. When this objective is maximized, biomass is always favored, if the appropriate resources are available. Otherwise, excess glycogen can be observed to accumulate over time, simulating effectively the internal sequestration that happens *in vivo*.

RESULTS

Global transcriptional changes during growth on minimal and rich media

Shewanella oneidensis MR-1 was cultured in a batch bioreactor using nutritionally rich medium (LB) and minimal M4 medium containing DL-lactate (LAC) (see 'Materials and Methods' section for details). Time-dependent mRNA expression profiles were obtained using *S. oneidensis* Affymetrix microarrays (BU_*Shewanella*_100k_v1.0, Gene Expression Omnibus platform reference number GPL8434). We collected biomass samples and measured expression at 21 time points in LB and 19 time points in LAC, covering the growth curves from early log phase to a few hours after reaching stationary phase (Figure 1A). These data sets provide snapshots of the regulatory programs employed by the organism throughout the transitions from abundantly available to depleted resources in two metabolically very different conditions (Supplementary Data set 1). An enrichment analysis for main functional categories in each cluster across the two growth conditions is presented in Supplementary Data set 2. Some initial insights into the specific transcriptional changes occurring throughout the two experiments can be obtained using standard K-means clustering (Supplementary Figure S1 in Supplementary text). In particular, in the LAC growth experiment, this analysis reveals how major events in the adaptive response at the level of the gene transcription are mainly occurring during the late-exponential phase.

To illustrate the global trends in gene expression dynamics throughout our experiments, we visualized the
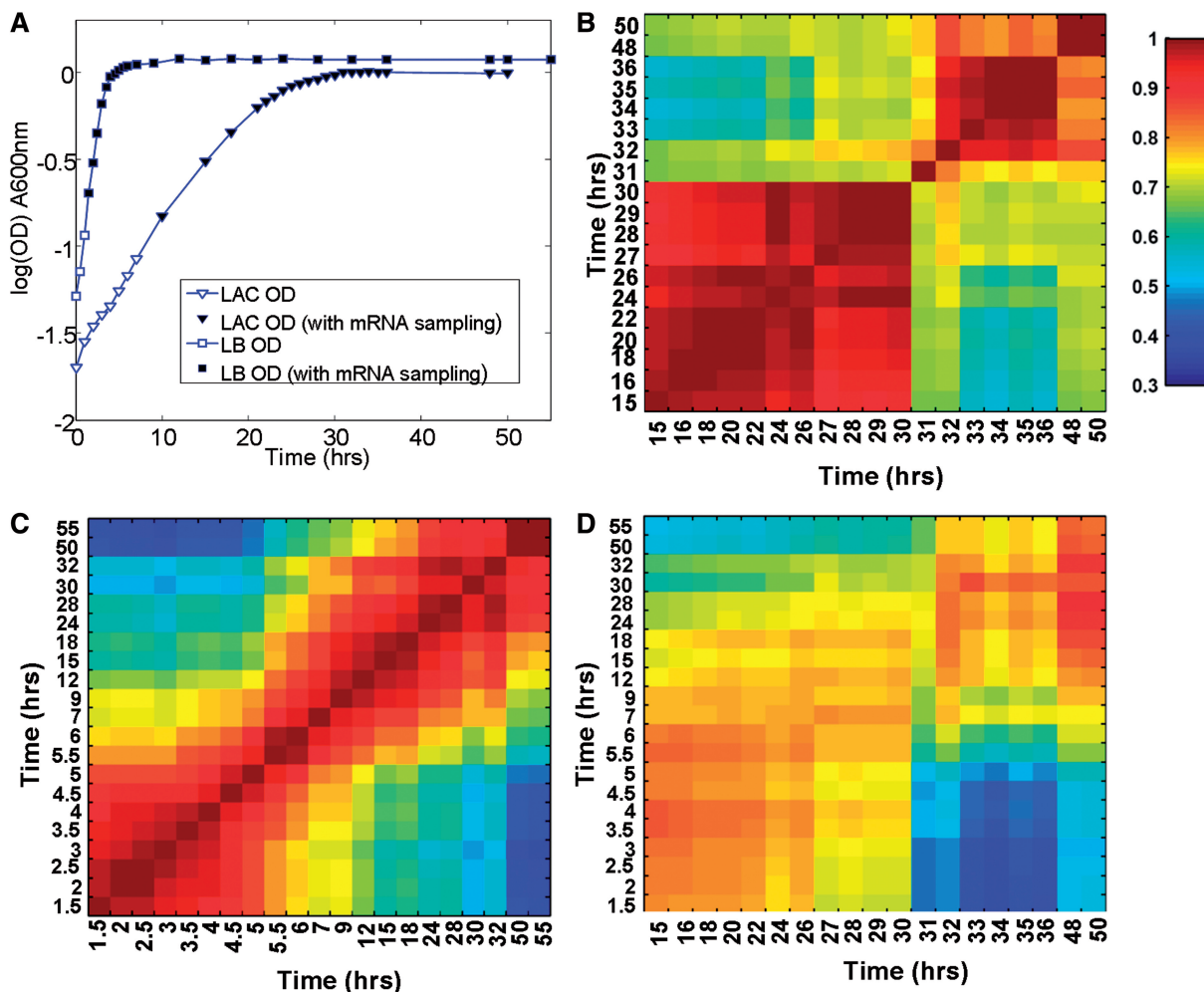
Figure 1. (A) Growth curves in the LB-rich and LAC-minimal medium. Filled symbols (inverted triangle and square) correspond to the time points at which mRNA was extracted for microarray hybridization, for LAC and LB medium, respectively. The heatmaps in (**B–D**) indicate the level of pair-wise similarity in the transcriptome profiles at different phases of growth (B: LAC versus LAC; C: LB versus LB and D: LB versus LAC). These heatmaps indicate the correlation between genome-wide transcriptional profiles at any two time points.

correlation between genome-wide transcriptional profiles at any two time points (Figure 1B–D). During growth on LAC medium, while the first consecutive time points are highly similar, a sudden drastic change marks the transition from late exponential phase into stationary phase, between 26 and 30 h (Figure 1B). This sharp transition suggests a fast, highly coordinated system-level response, likely triggered by environmental changes as explored in detail later. In contrast, during growth on the LB medium, the transcriptional dynamics follows a more gradual and smoother rearrangement (Figure 1C). This is likely a consequence of the much higher number of degrees of freedom that the complex regulatory network can span while switching between the multiple carbon and nitrogen sources as they are gradually used up (49,50). Despite the clear difference between the LAC and LB transcriptional dynamics profiles, it is possible to identify weak correlations across the two experiments. As shown in Figure 1D, early log phase gene states in the LAC and LB growth conditions slightly correlate with each other, and so do transcription profiles during

entry in stationary phase. This result can be interpreted as a suggestion that some components of the growth transcriptional program are environment dependent, while others are environment independent, and likely driven by the growth process itself. It is important to note that absolute times in the two experiments are different, suggesting that the LAC versus LB correlation observed holds with respect to some rescaled timeline as discussed in the following section.

Comparing the transcriptional response across conditions

In order to perform a systematic and quantitative comparative analysis of the LAC and LB time-dependent courses despite their different timescales, we devised an approach (GDM) to rescale the transcriptional profiles based on an analytical estimation of growth-rate (see 'Materials and Methods' section and Supplementary Text for more details) (51), inspired by previous assessments of the dependence of gene expression on growth rates (52,53). As opposed to a previous method that compared gene expression across conditions based on

equal optical densities, GDM is based on an analytical estimation of growth-rate changes. Using the GDM approach, we compared rescaled transcriptional profiles for each gene across the two nutrient conditions, and identified three classes of genes (Supplementary Data set 3): a first class (Figure 2A) comprising those genes that have significant positive correlation between transcriptional profiles across the two conditions (891 genes with a $Q \leq 0.01$); a second class (Figure 2B) including 175 genes that are significantly anti-correlated; and a third class, (the remaining 3164 genes) displaying no significant correlation or anti-correlation across the two growth conditions.

Among the genes that are most significantly positively correlated between the two growth conditions, we found many genes involved in growth-dependent activities, such as ATP and aminoacyl-tRNA biosynthesis, amino acid metabolic processes, ribosomal genes, DNA replication and aerobic respiration. A notable example of correlated expression profiles in LAC and LB was found for the transcriptional regulator (TR) *rpoD*, the primary σ^D factor for exponential growth (Figure 2C). Conversely, *rpoS* (the σ^S factor and master regulator of the general stress response) displays non-correlated patterns across the two growth conditions (Supplementary Data set 3). It is known that the relative concentrations of σ^S and σ^D play a crucial role in dictating the growth strategy of a bacterium, balancing optimal growth and survival to stress events (54,55). Our data support the possibility that *rpoD* transcription is associated with growth-rate processes irrespective of media composition, while other regulators (e.g. *rpoS*) or post-transcriptional control may modulate environment-specific changes (55,56). As a further hint to the complex regulation of these two
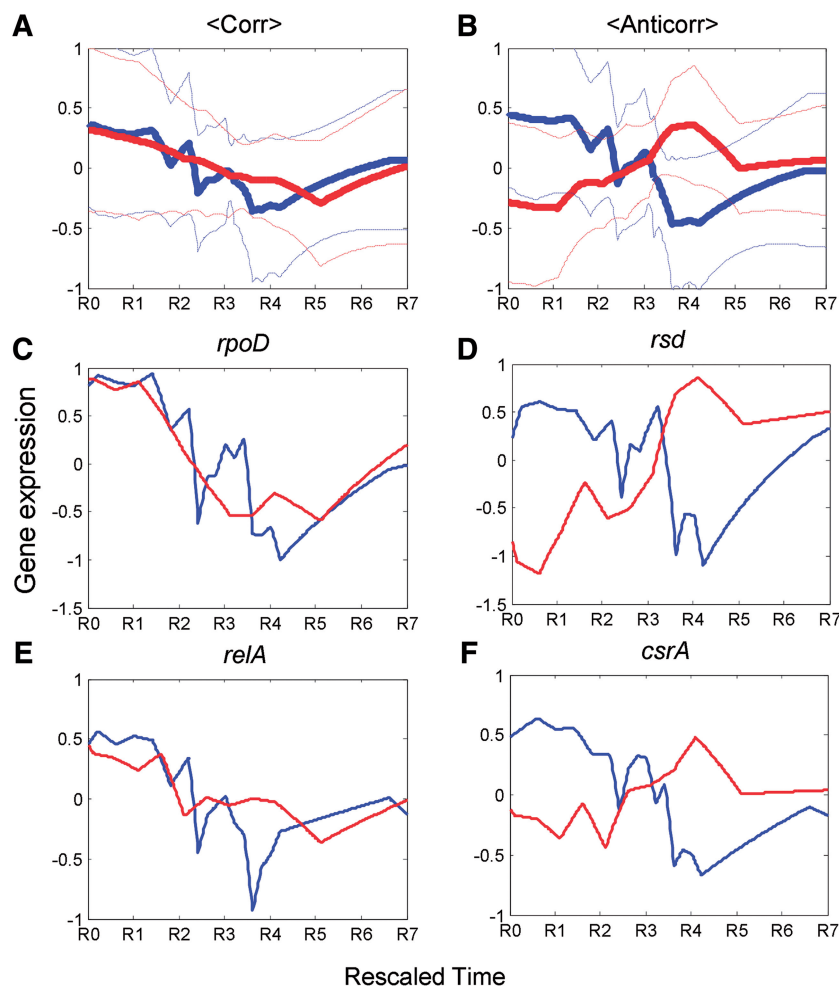

Figure 2. The approach of GDM is used to compare individual gene time-courses across the two conditions. In particular, one can identify genes whose expression (on *X*-axis of all panels) time-courses are correlated (**A**) or anti-correlated (**B**) between LAC and LB growth conditions, in a statistically significant way. The solid red (LAC) and blue (LB) curves in (A and B) are the mean expression value of all genes from each data set. Dashed lines represent expression values that are one standard deviation above or below the average. The time points R0, R1, ..., R7 on the *X*-axis correspond to labels for the rescaled and interpolated time obtained through the GDM approach across two growth conditions (see 'Materials and Methods' in Supplementary Text). For the LAC experiment, R0 = 15 hr, R2 = 20 hr, ..., R7 = 50 hr; for the LB experiment. R0 = 1.5 hr, R2 = 7.6 hr, ..., R7 = 55 hr. The *Q* values were obtained as described in the 'Materials and Methods' section. The other panels show selected examples of important correlated genes, *rpoD* (**C**) and *relA* (**E**), with *Q* values of 0 and 0.004, respectively; and anti-correlated genes, *rsd* (**D**) and *csrA* (**F**) with *Q* values of 0.0025 and 0.0007, respectively.

factors, *relA* (57), a key enzyme in the biosynthesis of the global stringent response regulatory molecule (p)ppGpp (responding to nutrient and energy starvation), shows positively correlated profiles in the LAC and LB conditions (Figure 2E). Also, *rsd*, an anti-sigma factor that sequesters RpoD from the RNA polymerase (58,59) displays anti-correlated expression patterns across the two conditions (Figure 2D), revealing an overall dichotomous behavior of regulators. An additional anti-correlated gene worth highlighting is *csrA* (Figure 2F), which is upregulated in the LAC data set and encodes a known regulator of carbon and glycogen metabolism (20,60). As evidenced later, this finding agrees with the glycogen-related transcriptional changes we observe during LAC growth (see section on Metabolite measurements and dFBA observations).

Thus, using GDM, we showed that it is possible to identify genes whose expression changes are likely associated with the growth process, and are only weakly affected by the drastically different nutritional parameters in the LAC and LB conditions. Conversely, in the next section, we ask what genes are influenced the most by environmental changes, and could be considered the main entry points for downstream transcriptional regulation.

D2T2: identifying genes that mediate the response to environmental changes

Is it possible to use the measured expression profiles to understand environment-dependent genetic interactions (61,62), and in particular to identify the gene nodes (triggers) that mediate the response to changing environmental conditions? A lot of work has been devoted to clustering microarray data and reconstructing regulatory maps (5,42,43,63–65). Here, we focus on a rather different and complementary aspect, specifically on using time-course data to infer key genes that serve as interface between the internal network and the external environment. To overcome issues typically encountered in time-course analyses (such as small number of time samples compared to a large number of transcripts), we propose a new method, which capitalizes on the availability of a large compendium of *S. oneidensis* transcriptional response data to a variety of environmental perturbations (M3D) (39) in addition to our time-dependent profiles. Specifically, our approach, named D2T2 (see Supplementary text for complete details), performs reverse engineering of gene expression time courses through integration with steady-state profiles.

Briefly, the D2T2 algorithm operates in two main steps (Figure 3). A gene network model is first trained on a large compendium of gene expression measurements taken under the assumption of steady state, using a relevance network approach (see 'Materials and Methods' section and Supplementary text for full details). These relationships are further weighted through a multiple linear regression scheme, and refined to reach the best compromise between the model complexity and its predictive ability. The time-dependent gene expression profile, describing the condition-specific dynamic behavior of the system, is then integrated with the previously determined
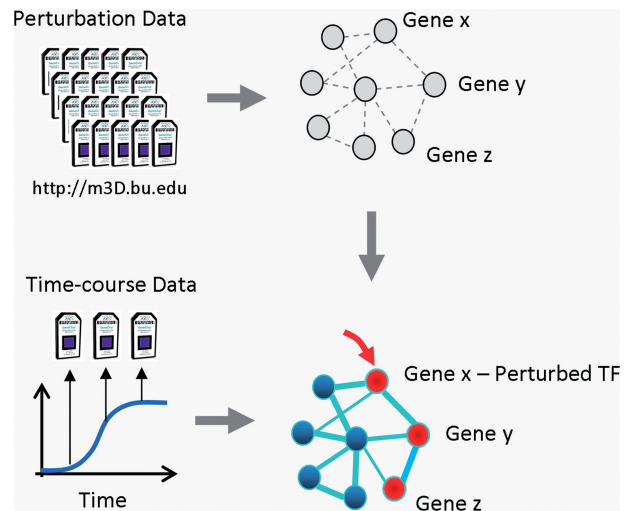

Figure 3. Our new method for reverse engineering time-series gene expression data, D2T2, acts through the integration of 245 gene expression profiles of *S. oneidensis* MR-1 from the M3D database (http://m3d.bu.edu) (39). A connectivity map is extrapolated from this training set and edges are weighted performing two linear regressions. The time-dependent data from our experiments are integrated with the steady-state network, to provide an estimate of what genes have expression changes that cannot be well explained by the internal network itself, but rather require the presence of external triggers.

static network, leading to the predictions of the key responsive genes. More specifically, the genes showing an incoherent behavior relative to the initial model (e.g. transient response) are chosen as the best candidates for being the mediators (or triggers) of the overall transcriptional rearrangement. Both the underlying network and the trigger genes are determined based on statistical inference, with confidence quantifiable through Q-values (Supplementary Data set 4).

D2T2 validation: the time-dependent response of *E. coli* to antibiotics

Before applying the D2T2 algorithm to the new *S. oneidensis* data, we tested its performance on the time-dependent transcriptional profiles in *E. coli* under the effect of a quinolone antibiotic (Norfloxacin) (43), and we asked whether D2T2 could recover the molecular targets and the mechanisms of action of this antibiotic, which had been well characterized before (66–68). In Table 1, we compare the performance of D2T2 with the approach of Time Series Network Identification (TSNI) (43). Full details of the analysis of the results from applying D2T2 to *E. coli* can be found in Supplementary text.

Our ranking of confirmed Norfloxacin targets was consistently and significantly improved relative to the TSNI results, giving e.g. *recA* and *gyrA* among the top hits (see Table 1 for details). Our method also identified two genes, *tisA* and *tisB*, so far not suspected to be implicated in quinolone action, but which could explain the interplay between quinolone antibiotics and the observed membrane depolarization (69). Our finding is in

Table 1. *Escherichia coli* test of D2T2

| | TOP 50 | |
	D2T2	TSNI
gyrA	33	–
recA	2	14
tisA	1	–
tisB	22	–

We report here the comparison of D2T2 with an existing computational method (TSNI) for identifying key mediators of Norfloxacin antibiotic in *E. coli*, based on data and analysis reported in (43). Known key transcriptional mediator (*recA*) and molecular target (*gyrA*) are among the most significant hits recovered by D2T2, outperforming TSNI. It is worth noting that D2T2 takes advantage of independent data set, while other methods, such as TSNI, rely exclusively on the time samples. This is a crucial step in the entire procedure because it reduces the underdetermined nature of reverse engineering time course gene expression data set, and enables the method to work on a genome scale level, while for example TSNI can be applied to a pre-selected subset of genes of the order of hundreds.

agreement with a recent report linking TisB to proton motive force and ATP levels and highlighting the role of this gene in persistency and drug tolerance (70).

Applying D2T2 to *S. oneidensis* during growth in LAC medium

In applying the D2T2 approach to our *S. oneidensis* transcriptional data, we had no discrete perturbation events (as opposed to the *E. coli* antibiotics case described above). However, we envisaged that nutrient availability might be similarly linked to transcriptional triggers whose signals propagate through the internal genetic networks. The D2T2 approach allowed us to focus on what we refer to as perturbed, or 'trigger' genes (290 in the LAC and 430 in LB condition, at $Q \leq 0.01$, Supplementary Data set 4). Among these trigger genes, 17 are known or putative TRs in the LAC and 24 in the LB medium (Table 2). These regulators can be thought of as the responsive triggers that dictate the overall response to the environmental changes during both growth conditions.

In Table 2 and Figure 4A, we showed 17 significantly perturbed TRs that are identified by D2T2 as the major environmental triggers during growth in LAC minimal medium. Based on available literature, we confirmed that most of these identified TRs are known to have an important role as environmental sensors in different organisms (Table 3). Many of these TRs are related to carbon and nitrogen metabolism or depletion, most notably *lldR* (SO3460), a gene recently shown to regulate the transcription of the L-LDH operon (74), and *rpoN*, a well-characterized sigma factor associated with nitrogen limitation (92). Their perturbed signals seem to propagate to the downstream nodes: *lldR* is likely responsible for the subsequent repression of SO0827 (L-lactate permease of the LctP family) and *lldE* (subunit of L-lactate dehydrogenase) (74,88). It is worth noting that both *lldE* and SO0827 show a significant reduction in expression near the end of the exponential growth phase, while *lldR* tends to increase, suggesting that *lldR* acts as a repressor

of *lldE* and SO0827 (Figure 5A). Alternatively, these two genes might be coregulated by other TRs or other signalling circuits, as corroborated by the fact that both score very high in the D2T2 algorithm (Supplementary Data set 4). The potential complexity of the regulatory circuit determining the fate of lactate utilization is supported by the fact that in *E. coli*, *lldR* acts both as a repressor and an activator of the *lldPRD* operon (73).

Other TRs that appear to act as interfaces for environmental sensing include EtrA/Fnr (electron transport regulator) and ArcA (TR for aerobic respiration). In our data, both TRs are up-regulated during mid-log phase and down-regulated upon entry into late-log phase (Figure 5B). EtrA, a homolog of FNR in *E. coli* (18,81), has recently been shown to be involved in fine tuning the expression of anaerobic metabolism genes in *S. oneidensis* (82). Another report (83) showed that the expression of either ArcA or EtrA/Fnr was not influenced by growth conditions. Our D2T2 analysis suggests that both regulators play an important role during aerobic growth, responding to high oxygen demand during fast growth (as supported by data in the section on metabolite profiles and dFBA discussed later), and possibly to the initial sensing of carbon and nitrogen starvation.

A third example of TRs whose transcriptional changes are best explained by external factors using our D2T2 algorithm is the pair of nitrogen-related global regulators *lrp* (leucine responsive protein) and *rpoN* (σ^{54}). These regulators affect a number of genes relevant to nitrogen utilization (*ntrB, ntrC, amtB_1, amtB_2, glnA, glnK_1, glnK_2*). In addition, the network of interdependencies obtained by D2T2 shows that these downstream genes are also highly connected to several glycogen-related genes (*glgA, glgB, glgC, glgP, glgX*, see Figure 5C and Supplementary Figure S8). Both RpoN and Lrp, as well as the two-component glutamine regulatory system (NtrBC) are known to be involved in a complex regulatory network that plays a central role during growth on limited nitrogen in *E. coli* (92, 93). One interesting pattern that emerges from our data is a tight relationship between genes associated with the nitrogen starvation response and genes involved in glycogen metabolism, which are up-regulated between 28 and 34 h, with highest peak at ~31 h (Figure 5C). This pattern can be visually observed by analyzing the synchronized expression profiles of the two categories. In addition, the network of gene interdependencies obtained from the D2T2 algorithm places the nitrogen and glycogen related genes into a single highly interconnected cluster (Supplementary Figure S8). Detection of the gluconeogenesis genes (*pckA*, phosphoenolpyruvate carboxykinase and *ppc*, phosphoenolpyruvate carboxylase) by D2T2 further supports the notion that the upstream flow of carbon from lactate to glycogen via gluconeogenesis is regulated by the cell through environmental signals (Figure 5D). These observations are reported for the first time in *Shewanella* species and are consistent with prior knowledge of glycogen production as a storage compound in other bacterial species during exponential and/or stationary phase growth (94–97). Glycogen biosynthesis has been reported to be enhanced during acid stress in *S. oneidensis* (34). A correlation

Table 2. List of TRs identified by the D2T2 approach

SO number	TR	Ontology	Perturbation
Lactate-minimal medium			
SO2263	IcsR	Transcriptional repressor of iron–sulfur cluster assembly genes, IscR	1.5605
SO1820	PolB	DNA polymerase II, PolB	0.2516
SO3460	LldR	TR, LysR family	0.1542
SO1415	N/A	TR, TetR family	0.1535
SO2519	N/A	TR, AraC family	0.1532
SO4140	N/A	TR, LysR family	0.0191
SO0701	N/A	TR, LysR family	−0.1128
SO3494	MexR	TR, TetR family	−0.3144
SO3961	RpoN	RNA polymerase sigma-54 factor, RpoN	−0.4859
SO2356	EtrA/Fnr	Oxygen-sensitive electron transport regulator A, EtrA	−0.6463
SO2852	HypR	TR, GntR family	−0.6534
SO3988	ArcA	Two-component TR for aerobic respiration, ArcA	−0.6690
SO0009	DnaN	DNA polymerase III, beta subunit, DnaN	−0.7673
SO2602	ProQ	Activator of ProP osmoprotectant transporter, ProQ	−0.7942
SO1284	RpoD	RNA polymerase sigma-70 factor, RpoD	−0.8041
SO0874	DksA	RNA polymerase-binding protein, DksA	−1.5216
SO2305	Lrp	Leucine-responsive regulatory protein, Lrp	−1.8174
LB-rich medium			
SO1898	LiuR	Branched chain amino acid metabolism regulator, LiuR	15.6069
SO3988	ArcA	Two-component TR for aerobic respiration, ArcA	5.5129
SO1342	RpoE	RNA polymerase sigma-24 factor, RpoE	1.9123
SO3429	RecX	Regulatory protein, RecX	1.6261
SO2567	N/A	Rra-like regulator of RNAse E	1.0986
SO3649	CgtA	Regulator of ppGpp phosphohydrolase, CgtA	1.0921
SO2640	N/A	TR, MarR family	0.3266
SO3059	N/A	Sigma-54-specific TR, Fis family	0.3066
SO1703	N/A	Pseudogene: TR, TetR family, degenerate	0.1514
SO0386	N/A	Excisionase/response regulator inhibitor-like protein	0.1493
SO0059	KtrE	Two-component response regulator for KtrAB potassium uptake, KtrE	0.0745
SO4020	YidZ	TR, YidZ	0.0487
SO2119	N/A	Protein phosphatase with response regulator receiver modulation	0.0274
SO3138	DctD	Two-component Sigma-54-specific TR of C4-dicarboxylate transport, DctD	−0.0855
SO0989	N/A	TR, LysR family	−0.2576
SO4477	CpxR	Two-component TR for periplasmic stress, CpxR	−0.3319
SO1652	N/A	PHP (Polymerase and Histidinol Phosphatase) domain protein	−0.4092
SO3964	YhbJ	Regulator for small RNA, YhbJ	−0.4539
SO0009	DnaN	DNA polymerase III, beta subunit, DnaN	−0.6046
SO2649	CysB	Transcriptional activator of cys regulon, CysB	−0.6271
SO0360	RpoZ	DNA-directed RNA polymerase, omega subunit, RpoZ	−0.9802
SO2602	ProQ	Activator of ProP osmoprotectant transporter, ProQ	−1.077
SO0096	HutC	TR for histidine utilization, HutC	−1.5814

TRs for which a significant 'perturbing stimulus' (e.g. $Q \leq 0.01$) have been identified by the D2T2 procedure are listed here, with the corresponding putative annotation and estimated perturbation intensities (i.e. b values). The full list of significant genes identified by D2T2 in the LAC and LB conditions can be found in the Supplementary Data set 4 (N/A: name not available).

between glycogen accumulation and low nitrogen in the actinomycete, *Micromonospora echinospora*, was explained to be a result of reduced carbon flux (98). The use of lipid as a storage compound under nitrogen-limiting conditions has also been observed (99). Our study suggests a possible link between glycogen biosynthesis and nitrogen starvation in *S. oneidensis*, though the timing of these effects relative to the carbon starvation response in our experiment remains to be understood.

While we have been focusing here on genes that mediate the response to metabolically relevant processes, we should emphasize that D2T2 also identified other genes associated with the entry into stationary phase, such as TRs and target genes associated with flagellar biosynthesis, pili, chemotactic functions and biofilm formation (see Supplementary Figures S9 and S10; Supplementary Data set 4).

Another subtle aspect emerging from the GDM and D2T2 analyses is that the cells may respond to transcriptional triggers that are internal (i.e. associated to the status of the cell) rather than external (i.e. associated with environmental stimuli). Our initial expectation was that the D2T2 approach would only identify extracellular (i.e. environmental) entry points into the regulatory network. Our results suggest, however, that some intracellular growth-phase sensing may feed back into the system as well, effectively 'communicating' to the cell its growth status, and triggering a specific environment-independent transcriptional response. This phenomenon is specifically suggested by the fact that some genes (e.g. *rpoD*, which is regulated by cAMP/CRP in *E. coli*) appear both as correlated in the GDM analysis (i.e. it is regulated similarly across the two conditions, and hence is weakly dependent on nutrient availability), and as major potential
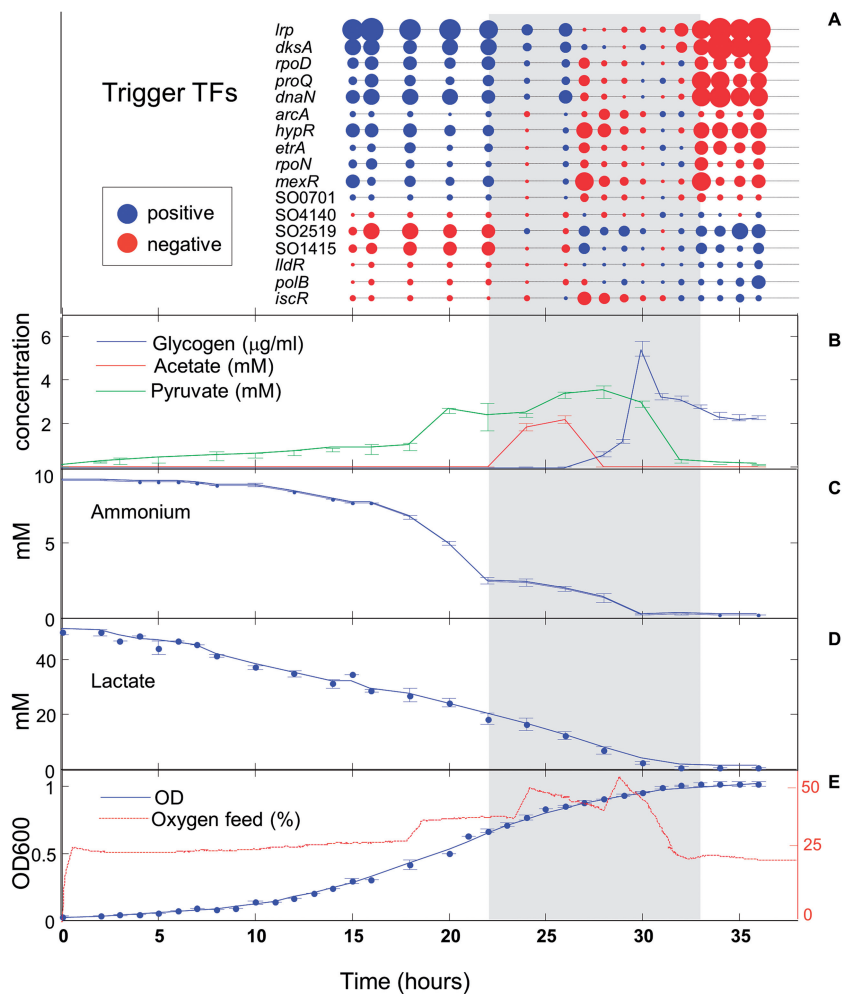
Figure 4. Expression changes (**A**) over time of 17 TRs identified by D2T2 approach (red-low expression, blue-high expression). The size of circle represents level of expression. The shaded region between 22 and 33 hr corresponds to the phase of growth in LAC-minimal medium involving major changes in expression of 17 TRs identified by D2T2 and shown in (**A**). HPLC was used for quantifying acetate and pyruvate (**B**), whereas ammonium (**C**) was quantified using commercially available kit (see 'Materials and Methods' section). Lactate (**D**) was quantified using HPLC and a commercially available kit (see 'Materials and Methods' section). Optical density (**E**) of ~1 ml culture samples was measured during growth of *S. oneidensis* at 600 nm using a spectrophotometer, and the external O_2 feed data was collected using the BioCommand software for Bioflo110 bioreactor from New Brunswick Scientific Company, Edison, NJ.

trigger from the D2T2 analysis (i.e. its transcriptional changes cannot be explained simply by the internal regulatory network architecture). These genes (a total of 107) are listed in Supplementary Data set S6. Our results suggest that post-transcriptional regulation may supplement core regulatory processes providing a feed-forward control of environment on growth strategies (6). An intimate and delicate interplay between growth and environmental control might become a crucial requisite to balance the internal state of the cell and achieve optimal conditions for growth or long-term survival.

Understanding resource limitation using dFBA

The high-throughput mRNA profile analyses using GDM and D2T2 have provided us with insight into genes whose expression is apparently associated with growth phase changes, as well as genes whose regulation is likely influenced by environmental factors, such as changes in nutrient availability. The microarray analyses presented earlier indicate that at different times along the growth curve, changes in the availability of resources (carbon, nitrogen, oxygen) trigger major transcriptional events. In order to better understand the interplay between growth-dependent processes and the effects of resource limitations on the physiology of *S. oneidensis*, we used a time-dependent version of the widespread genome-scale stoichiometric approach of FBA. Specifically, we set to simulate the experimentally measured growth curve and metabolite depletion using a dFBA approach (45) (see 'Materials and Methods' section for details), based on a recently published stoichiometric reconstruction of the *S. oneidensis* metabolic network (48). We initialized dFBA by using the known initial concentrations of available nutrients, and applied consecutive rounds of flux balance updates at equally spaced time intervals. As shown in

Table 3. List of regulatory interactions (input signals and targets) of some selected identified TRs during growth in LAC medium

TR	Gene product	Input signal (Signal class)	Target genes in *S. oneidensis* MR-1[a]	Known target functions (Organism)	References
MexR	Antibiotic resistance regulator	Oxidative state (stress)	*mexF, mexE, mexR*	Multi-drug efflux transporter (*P. Aeruginosa*)	*Pseudomonas aeruginosa* (71), *S. oneidensis* (none)
HypR	TR of proline utilization	Proline (nutrient source)	SO0639, *putA*	Proline degradation, collagenase (*Sinorhizobium meliloti*)	*Sinorhizobium meliloti* (72), *S. oneidensis* (none)
LldR	TR of lactate utilization	Lactate (carbon source, sugar acids)	*lldG, lldF, lldE*	Lactate transport and utilization (*E. coli, S. oneidensis*)	*Escherichia coli* (73), *S. oneidensis* (74)
IscR	TR of FeS cluster assembly	FeS cluster level (cofactor)	*erpA, iscR, iscS, iscU, iscA, hscB, hscA, fdx, dnrN, nfuA*	FeS cluster assembly, anaerobic respiration (*E. coli*)	*Escherichia coli* (75,76), *S. oneidensis* (none)
ArcA	Two-component regulator for aerobic respiration	Small proteins: HptA, ArcS (oxygen tension)	*ybgT, cydB, cydA*	Cytochrome oxidase, DMSO reductase (*S. oneidensis*)	*Escherichia coli* (77), *S. oneidensis* (78,79)
EtrA/ Fnr	Global oxygen-sensitive TR for anaerobiosis response	Oxygen (oxygen tension)	*ybgT, cydB, cydA, ccmF, ccmG, otr, coxB, coxC, ccmH,* SO0269, *frdC, frdC, frdA, frdB, lpdA,* SO0579, *coaX,* SO0581, *tpmT, fbpC, fbpB, fbpA, napB, napH, napG, napA, napD, bfr2, bfr1,* SO1248, SO1249, SO1250, *cyaC, dldD, lldP, ompW, mtrF, mtrE, mtrD, feoA, feoB, tig,* SO2005, *hypF, hyaE, hyaD, hyaC, hyaB, hyaA, ydaO, uspE, fnr,* SO2357, *ccoS, ccoI, ccoH, ccoP, ccoQ, ccoO, ccoN, nrdG, nrdD,* SO2891, SO3054, SO3890, *nrfA, narQ, narP,* SO4142, SO4143, *hupE, moaE, moaD, moaC, moaA, etfQ,* SO4512, *fdhA, fdhB, fdhC, ctaG*	Cytochrome maturation, fumarate reductase, nitrate/nitrite reductase, Fe transport storage, adenylate cyclase, lactate transport/utilization, cytochrome oxidase, formate/NiFe hydrogenase (*S. oneidensis*)	*Escherichia coli* (80), *S. oneidensis* (81–85)
SO1415 TR		Oxygen (oxygen tension)	SO1410, SO1411, SO1412, SO1413, SO1414, SO1415	Anaerobic respiration (*S. oneidensis*)	*Shewanella oneidensis* (86,87)

The identified input signals, documented by specific references for different organisms, include oxygen concentration, osmolarity, oxidative stress, ppGpp (signal for nutrient and energy limitation), carbon (amino acid, lactate) and nitrogen metabolism (via DksA). Pathway genome database group feature was used to extract direct targets of regulators in *S. oneidensis* MR-1 (88–91).
[a]*ybgT, cydB, cydA* are target genes for both ArcA and EtrA/Fnr in *S. oneidensis* MR-1.

Figure 6, the dFBA simulation predicts a number of important features of the growth curve consistent with the transcriptional signals described earlier. In particular, the dFBA predicts that: (i) the major nitrogen source (ammonium) should be depleted from the minimal medium approximately 2 h before lactate depletion. This would explain the large number of nitrogen utilization genes identified in the D2T2 analysis, including the TRs Lrp, RpoN and ProQ (Figure 5C and Supplementary Figure S8); (ii) abundant oxygen availability (8 mM) may not be enough for satisfying the demand of log-phase growing cells. Hence, it is conceivable that during log phase, the cells sense a reduction in oxygen availability. This would be compatible with the transcriptional signals identified with D2T2, in particular TRs EtrA/FNR and ArcA (Figure 5B). Notably, the dFBA model predicts (Figure 6) that under oxygen limitation, *S. oneidensis* will produce acetate at approximately $t = 12\,h$ ($O_2 = 2\,mM$) or $t = 18\,h$ ($O_2 = 6\,mM$); and (iii) if indeed ammonium runs out before lactate, there is potentially a time window during which cells would not be able to produce biomass, but would potentially be able to store the available carbon for survival during starvation periods. This hypothesis may be related to the spike of
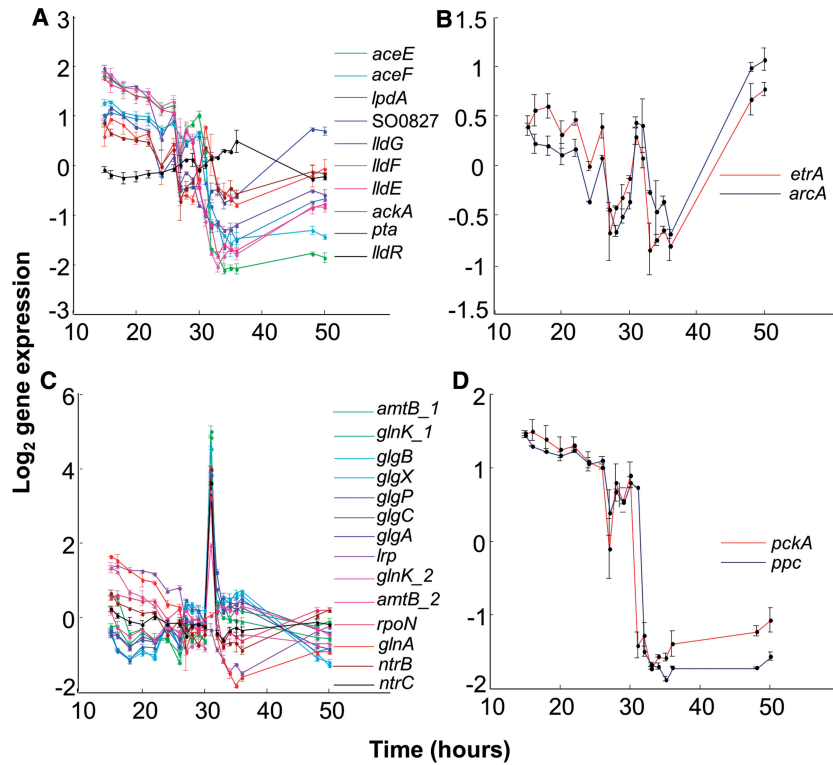
Figure 5. Expression profiles for the genes of selected pathways during batch growth in LAC-minimal medium. Genes responsible for uptake and conversion of lactate into acetate via pyruvate (**A**); know oxygen sensors, EtrA/Fnr and ArcA (**B**); nitrogen and glycogen-related genes (**C**); and two gluconeogenesis genes, *ppc* and *pckA* (**D**).
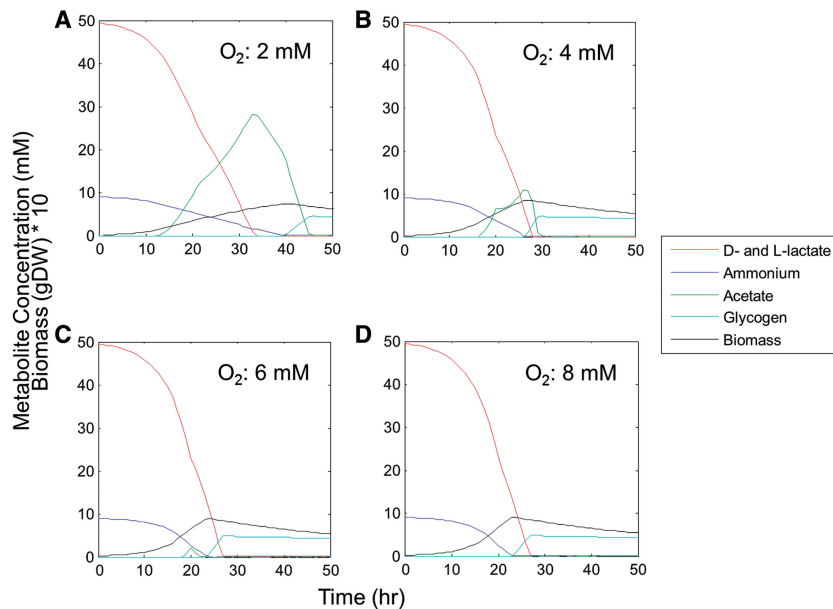

Figure 6. dFBA was used to simulate bacterial growth in minimal medium with lactate (LAC) using a genome scale model of *S. oneidensis* MR-1 for various constant concentrations of oxygen: (**A**) 2 mM O_2, (**B**) 4 mM O_2, (**C**) 6 mM O_2 and (**D**) 8 mM O_2. In each instance, a 2% death rate was implemented. Also, the lactate uptake rate and ATP maintenance cost were constrained to known experimental values. A dual objective function was also implemented to maximize biomass and glycogen production. Acetate production is inversely related to oxygen abundance. In low oxygen environments (e.g. 2 mM in A), abundant acetate is excreted and later utilized, and in high oxygen environments (8 mM in D), acetate production is severely diminished until it is no longer detected.

glycogen-associated genes (Figure 5C) that we observed in conjunction with major changes in nitrogen utilization processes. To verify whether the global stoichiometry of the system would be compatible with the storage of glycogen, we modified the dFBA model to include an outflow ('storage') of glycogen as a potential flux in the objective function. Upon this modification, the dFBA model predicts that glycogen production is compatible with other constraints in the time window between nitrogen and carbon depletion (Figure 6).

Metabolite measurements corroborate transcriptional and dFBA observations

Both the gene expression analyses (Figures 4A and 5) and the dFBA calculations (Figure 6) indicate that carbon, nitrogen and oxygen limitations seem to underlie specific physiological transitions, to which the cells respond with transcriptional control of relevant processes. In order to verify the dFBA-based interpretation of the observed transcriptional changes, we measured at different time points the supernatant concentrations of lactate, acetate and ammonium, and the intracellular concentration of glycogen (Figure 4B–D and 'Materials and Methods' section). Since pyruvate secretion had been reported in previous growth experiments (100,101), we included pyruvate measurement in our experiment as well. In addition, to assess potential oxygen limitation, we also analyzed the time-dependent rate of external oxygen feeding monitored throughout the growth experiment (Figure 4E).

Consistent with the array data and with the onset of stationary phase (Figure 1B), lactate becomes depleted from the culture vessel between 30 and 32 h (Figure 4D). It turns out that the concentration of ammonium drops to zero at ~30 h (Figure 4C), placing the cells into a small (up to 2 h) window in which carbon is still available, but cells are experiencing depletion in nitrogen levels. The nitrogen depletion is consistent with the observed increase in transcript levels of genes involved in nitrogen starvation response (Figure 5C). As suggested also by the dFBA model, this window of excess carbon coupled with nitrogen limitation could explain the observed activation of glycogen-related genes (Figure 5C), similar to previous reports in yeast (102,103). The cells, unable to produce biomass because of the lack of nitrogen, instead turn their metabolic activity to storing the available carbon. Indeed, upon measuring intracellular glycogen over time, we observed an increase from zero up to ~3 µg/ml (corresponding to ~5.5 µg/g dried biomass), peaking at ~30 h, exactly when cells enter the window of ammonium depletion in presence of carbon (Figure 4B–D).

The transcriptional profiles and flux balance predictions also suggested that oxygen limitation occurs during fast exponential growth. In addition, the dFBA model predicted that upon oxygen limitation, acetate would be secreted as a byproduct. While the level of dO$_2$ in our LAC growth experiment was kept constant (~20%), the rate of external oxygen feed used to maintain such level of dO$_2$ in the bioreactor vessel underwent significant changes (Figure 4E). In particular, one can see a steep increase in

O$_2$ demand at different time intervals between mid-exponential to stationary phase of growth. At ~24 h, one can observe the occurrence of a sharp increase in the oxygen uptake signal. This corresponds to the time window (24–27 h) where acetate is excreted into the medium and subsequently consumed (Figure 4B) and is consistent with the up- and down-regulation of genes that convert acetaldehyde to acetate (*aldA*) and acetate to acetyl-CoA, respectively (Figure 5A and Supplementary Data set 4 of perturbed genes in LAC growth condition). Another sudden peak in the demand of oxygen (~27–28 h) corresponds to the onset of nitrogen limitation in the culture, and corresponds to a major peak in the transcription of the oxygen sensors ArcA and EtrA/FNR (Figures 4A and 5B). Note that an alternative explanation for the secretion of acetate into the external environment between 23 and 26 h of growth (Figure 4B) is the sudden change in metabolic strategy occurring when nitrogen starts becoming limiting, at ~22–23 h (Figure 4C). The rate of nitrogen utilization decreases sharply at this time, potentially causing a sudden overflow of carbon flux, that is relieved into the medium as acetate before the glycogen biosynthesis mechanism is ready to dissipate it.

In addition to the accumulation and consumption of acetate, we also detected accumulation and re-consumption of pyruvate (Figure 4B). A possible explanation is that the rate of lactate uptake and subsequent conversion to pyruvate outpaces the cells ability to process the pyruvate, leading to its secretion. Alternatively, the accumulation of this and other metabolites in the supernatant may be due to the break-down of lactate as a by-product of extracellular detoxification of peroxides in the environment by the bacteria. This aspect of the physiology of *S. oneidensis* is not caught in a straightforward manner by the dFBA model. This discrepancy may be attributed to limitations of the stoichiometric model (e.g. missing or unusually regulated reactions) or of the flux balance framework (e.g. the use of an inadequate objective function), and will require further investigation.

Independent validation of the global role of triggers through phenotypic data analysis

Complementary evidence that the D2T2 trigger genes are indeed involved in sensing external perturbations related to the three main environmental signals relevant to our experimental setup (lactate, nitrogen and oxygen), can be obtained from the analysis of recently published phenotype data (104) (Supplementary Text). This data set consists of fitness assays for 3355 *S. oneidensis* non-essential mutants (including several of the D2T2-identified genes) in 121 diverse conditions. For each mutant, we used a two-sample *t*-test to test whether fitness changes were preferentially observed upon the perturbations of nutritional parameters associated with the aforementioned environmental signals (i.e. anaerobic versus aerobic, lactate versus other carbon source, diverse nitrogen sources). Several of the genes identified by D2T2 under LAC growth show a significant ($P \leq 0.01$) phenotypic signature in at least one of the tested condition (Supplementary Table S2).

In addition, in order to test whether our list of trigger genes was significantly enriched for transcripts showing phenotypic patterns upon oxygen, lactate or nitrogen signals, we performed a permutation test, where we iteratively randomly selected an equal subset of putative trigger genes and counted how many times we could find a significant phenotypic pattern in each of these conditions. The results of this analysis indicate that there is a significant enrichment in our predicted gene list for transcripts showing phenotypes under these three environmental conditions ($P = 0.068$, 0.091 and 0.025, respectively). In particular, there seems to be a prevalence of genes responding to nitrogen availability (last row of Supplementary Table S2), pointing again to the fundamental role of this nutrient depletion in our analysis.

DISCUSSION

We have presented a new set of high-throughput transcriptional data useful for understanding growth phase physiology in *S. oneidensis* MR-1. In order to decode the information embedded in such data, we devised a mapping procedure (GDM) that allows us to compare transcriptional programs across different growth media. Most importantly, we developed a novel method (D2T2) for interrogating time-dependent gene expression about the entry points in the genetic network, and interpreted some of our results in light of flux balance model predictions and metabolite measurements.

Earlier studies on transcriptional regulation in *S. oneidensis* have focused on mRNA changes upon several discrete environmental or genetic perturbations, with no time resolution (25–30,32–35,37,38). Similar time-dependent processes in *Shewanella* have been studied before, but without probing transcriptional profiles (100,101). Our data, which spans two fundamentally different environmental conditions, a minimal lactate and a rich medium (both aerobic), adds a new dimension to the existing knowledge of *S. oneidensis* physiology (39,88,89,105). Our new data are particularly timely, given the increasingly relevant role of *Shewanella* species for studying microbial sediment communities and in bioenergy-related research (18,19), such as bioremediation (106), microbial fuel cells and biohydrogen production (107,108).

A large portion of the computational methods presented in this work is dedicated to the analysis of time-dependent gene expression. In contrast to other analyses of gene expression data, our focus here is intentionally not on obtaining a putative regulatory network. Rather, we sought to take advantage of our unique data sets to propose a novel way of interpreting dynamic gene expression. Our goal was to infer the genes whose dynamical changes cannot be explained by internal gene–gene influences, but likely to depend on other factors such as changes in metabolites in the environment.

By developing a time-rescaling approach (GDM), based on growth-rate estimates derived from measured optical density data, we formulated a direct comparison between time-course profiles for individual genes across the two conditions, revealing transcripts mainly responding to growth-rate variations and weakly dependent on whether simple carbon (lactate) or a complex mixture of nutritional resources are available in the medium. Our current analysis can only tell us which genes, upon rescaling, behave similarly in the two conditions considered. In principle, the same approach could be applied to multiple conditions and can help identify genes and pathways whose regulatory dynamics is truly independent of environmental conditions and changes mainly as a function of the growth phase of the cell.

Although the GDM approach allowed us to tease out transcriptional pattern that are correlated between the two conditions, a different method was necessary in order to highlight the specific environmental cues and transcriptional gates allowing cells to respond to changing nutrient availability. To address this question, we took advantage of the availability of a unique large compendium of *S. oneidensis* microarray data for response to multiple perturbations (39). Specifically, we devised a new algorithm (D2T2) that integrates our new time-series data presented in this manuscript with this compendium's static measurements to decipher how environmental signals feed into the regulation of gene expression and metabolic dynamics.

Our metabolite measurements and literature analyses (Figure 4, Table 3 and Supplementary Table S2) support the D2T2 predictions. However, a direct testing of the 'trigger' role of predicted genes will require complementary experimental approaches. One possible avenue would be the measurement of the response of mutant *S. oneidensis* strains under different conditions. For example the deletion of a TR predicted to mediate the response to a given external metabolite should lead to higher phenotypic changes under conditions in which such metabolite is present than in all other conditions. Along this line, we performed a statistical analysis of recently published data on gene deletion phenotypes under multiple conditions (104), showing overall agreement with our predictions (Supplementary Table S2 and Supplementary Text). In future, more targeted experiments could further help to validate specific D2T2 predictions. For example the activity of a TR could be monitored in a high time-resolution manner using GFP constructs (109), reporting the expression level of its targets upon pulses of the specific putative sensed signal. In addition, one could directly evaluate the sensing role of a TR through protein–metabolite binding assays (110).

Interpreted through the lens of a dFBA simulation, the key transcriptional changes and entry points detected with D2T2 point to specific metabolic events that apparently dictate some of the major decisions taken by *S. oneidensis* throughout the growth curve. Our analysis reveals that both ammonium (nitrogen) and lactate (carbon) limitation constitute key transcriptional triggers. As suggested by transcriptional spikes and dFBA predictions, nitrogen limitation turns out to be a major metabolic determinant in our analysis, and ammonium is predicted to run out slightly before lactate does in dFBA simulations. This is also confirmed by measurements of lactate and ammonium in the culture supernatant. At the same time,

the transcriptional response, as well as a modified dFBA calculation, point to a glycogen-related activity that seemed tightly coupled with the nitrogen starvation response. As verified through direct measurement of intracellular glycogen, this pattern indicates that *S. oneidensis* integrates information about carbon and nitrogen availability in a decision-making process that leads to the accumulation of storage carbon rather than biomass due to lack of nitrogen. This type of behavior, previously observed in other organisms, is to our knowledge reported here for the first time in *S. oneidensis*, suggesting how this lake sediment organism is able to survive under different modes of nutrient availabilities. The dFBA simulations presented in this work allowed us to interpret transcriptional data in light of global resource allocation constraints inherent to metabolism. In this case, the metabolic objective function, a key parameter of constraint-based models of metabolism, was assumed to be the same throughout the growth process. Future research in constraint-based models could take better advantage of data from dynamic processes to try and infer time- or condition-dependent objective functions, such as a shift from maximal growth to minimal energy turnover, in an attempt to mimic the global metabolite management decisions of the cell. Moreover, in integrated models, transcriptional data could be used to directly constrain metabolic fluxes.

Finally, the pipeline of experimental and computational approaches applied and developed for this work could be extended to other microbes and additional conditions. In *Shewanella*, for example, it would be highly interesting to use similar methods to explore what transcriptional changes and environmental triggers accompany the switching between different electron acceptors, or other physical factors e.g. gradual switching between aerobic and anaerobic conditions, as well as between different anaerobic acceptors including metals/metal oxides.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–2, Supplementary Figures 1–10, Supplementary Data sets 1–6 and Supplementary References [111–120].

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Airoldi,E.M., Huttenhower,C., Gresham,D., Lu,C., Caudy,A.A., Dunham,M.J., Broach,J.R., Botstein,D. and Troyanskaya,O.G. (2009) Predicting cellular growth from gene expression signatures. *PLoS Comput. Biol.*, **5**, e1000257.
2. de Lichtenberg,U., Jensen,L.J., Brunak,S. and Bork,P. (2005) Dynamic complex formation during the yeast cell cycle. *Science*, **307**, 724–727.
3. Klumpp,S., Zhang,Z. and Hwa,T. (2009) Growth rate-dependent global effects on gene expression in bacteria. *Cell*, **139**, 1366–1375.
4. Goelzer,A., Bekkal Brikci,F., Martin-Verstraete,I., Noirot,P., Bessieres,P., Aymerich,S. and Fromion,V. (2008) Reconstruction and analysis of the genetic and metabolic regulatory networks of the central metabolism of *Bacillus subtilis*. *BMC Syst. Biol.*, **2**, 20.
5. Lemmens,K., De Bie,T., Dhollander,T., De Keersmaecker,S.C., Thijs,I.M., Schoofs,G., De Weerdt,A., De Moor,B., Vanderleyden,J., Collado-Vides,J. *et al.* (2009) DISTILLER: a data integration framework to reveal condition dependency of complex regulons in *Escherichia coli*. *Genome Biol.*, **10**, R27.
6. Levy,S. and Barkai,N. (2009) Coordination of gene expression with growth rate: a feedback or a feed-forward strategy? *FEBS Lett.*, **583**, 3974–3978.
7. Martinez-Antonio,A., Janga,S.C., Salgado,H. and Collado-Vides,J. (2006) Internal-sensing machinery directs the activity of the regulatory network in *Escherichia coli*. *Trends Microbiol.*, **14**, 22–27.
8. Yi,T.M., Huang,Y., Simon,M.I. and Doyle,J. (2000) Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA*, **97**, 4649–4653.
9. Bischofs,I.B., Hug,J.A., Liu,A.W., Wolf,D.M. and Arkin,A.P. (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay. *Proc. Natl Acad. Sci. USA*, **106**, 6459–6464.
10. Singh,A.H., Wolf,D.M., Wang,P. and Arkin,A.P. (2008) Modularity of stress response evolution. *Proc. Natl Acad. Sci. USA*, **105**, 7500–7505.
11. Amsler,C.D., Cho,M. and Matsumura,P. (1993) Multiple factors underlying the maximum motility of *Escherichia coli* as cultures enter post-exponential growth. *J. Bacteriol.*, **175**, 6238–6244.
12. Nadell,C.D., Xavier,J.B., Levin,S.A. and Foster,K.R. (2008) The evolution of quorum sensing in bacterial biofilms. *PLoS Biol.*, **6**, e14.
13. Wagner,V.E., Bushnell,D., Passador,L., Brooks,A.I. and Iglewski,B.H. (2003) Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *J. Bacteriol.*, **185**, 2080–2095.
14. Balazsi,G., Barabasi,A.L. and Oltvai,Z.N. (2005) Topological units of environmental signal processing in the transcriptional regulatory network of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **102**, 7841–7846.
15. Gasch,A.P., Spellman,P.T., Kao,C.M., Carmel-Harel,O., Eisen,M.B., Storz,G., Botstein,D. and Brown,P.O. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell.*, **11**, 4241–4257.
16. Tao,H., Bausch,C., Richmond,C., Blattner,F.R. and Conway,T. (1999) Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media. *J. Bacteriol.*, **181**, 6425–6440.
17. Chechik,G., Oh,E., Rando,O., Weissman,J., Regev,A. and Koller,D. (2008) Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network. *Nat. Biotechnol.*, **26**, 1251–1259.
18. Fredrickson,J.K., Romine,M.F., Beliaev,A.S., Auchtung,J.M., Driscoll,M.E., Gardner,T.S., Nealson,K.H., Osterman,A.L.,

Pinchuk,G., Reed,J.L. *et al.* (2008) Towards environmental systems biology of *Shewanella*. *Nat. Rev. Microbiol.*, **6**, 592–603.

19. Hau,H.H. and Gralnick,J.A. (2007) Ecology and biotechnology of the genus *Shewanella*. *Annu. Rev. Microbiol.*, **61**, 237–258.

20. Heidelberg,J.F., Paulsen,I.T., Nelson,K.E., Gaidos,E.J., Nelson,W.C., Read,T.D., Eisen,J.A., Seshadri,R., Ward,N., Methe,B. *et al.* (2002) Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. *Nat. Biotechnol.*, **20**, 1118–1123.

21. Bretschger,O., Obraztsova,A., Sturm,C.A., Chang,I.S., Gorby,Y.A., Reed,S.B., Culley,D.E., Reardon,C.L., Barua,S., Romine,M.F. *et al.* (2007) Current production and metal oxide reduction by *Shewanella oneidensis* MR-1 wild type and mutants. *Appl. Environ. Microbiol.*, **73**, 7003–7012.

22. Meshulam-Simon,G., Behrens,S., Choo,A.D. and Spormann,A.M. (2007) Hydrogen metabolism in *Shewanella oneidensis* MR-1. *Appl. Environ. Microbiol.*, **73**, 1153–1165.

23. Rosenbaum,M.A., Bar,H.Y., Beg,Q.K., Segre,D., Booth,J., Cotta,M.A. and Angenent,L.T. (2011) *Shewanella oneidensis* in a lactate-fed pure-culture and a glucose-fed co-culture with *Lactococcus lactis* with an electrode as electron acceptor. *Bioresour. Technol.*, **102**, 2623–2628.

24. Sani,R.K., Peyton,B.M. and Dohnalkova,A. (2008) Comparison of uranium(VI) removal by *Shewanella oneidensis* MR-1 in flow and batch reactors. *Water. Res.*, **42**, 2993–3002.

25. Beliaev,A.S., Klingeman,D.M., Klappenbach,J.A., Wu,L., Romine,M.F., Tiedje,J.A., Nealson,K.H., Fredrickson,J.K. and Zhou,J. (2005) Global transcriptome analysis of *Shewanella oneidensis* MR-1 exposed to different terminal electron acceptors. *J. Bacteriol.*, **187**, 7138–7145.

26. Beliaev,A.S., Thompson,D.K., Khare,T., Lim,H., Brandt,C.C., Li,G., Murray,A.E., Heidelberg,J.F., Giometti,C.S., Yates,J. III *et al.* (2002) Gene and protein expression profiles of *Shewanella oneidensis* during anaerobic growth with different electron acceptors. *OMICS*, **6**, 39–60.

27. Bencheikh-Latmani,R., Williams,S.M., Haucke,L., Criddle,C.S., Wu,L.Y., Zhou,J.Z. and Tebo,B.M. (2005) Global transcriptional profiling of *Shewanella oneidensis* MR-1 during Cr(VI) and U(VI) reduction. *Appl. Environ. Microbiol.*, **71**, 7453–7460.

28. Brown,S.D., Martin,M., Deshpande,S., Seal,S., Huang,K., Alm,E., Yang,Y.F., Wu,L.Y., Yan,T.F., Liu,X.D. *et al.* (2006) Cellular response of *Shewanella oneidensis* to strontium stress. *Appl. Environ. Microbiol.*, **72**, 890–900.

29. Brown,S.D., Thompson,M.R., VerBerkmoes,N.C., Chourey,K., Shah,M., Zhou,J.Z., Hettich,R.L. and Thompson,D.K. (2006) Molecular dynamics of the *Shewanella oneidensis* response to chromate stress. *Mol. Cell. Proteomics*, **5**, 1054–1071.

30. Chourey,K., Thompson,M.R., Morrell-Falvey,J., VerBerkmoes,N.C., Brown,S.D., Shah,M., Zhou,J.Z., Doktycz,M., Hettich,R.L. and Thompson,D.K. (2006) Global molecular and morphological effects of 24-hour chromium(VI) exposure on *Shewanella oneidensis* MR-1. *Appl. Environ. Microbiol.*, **72**, 6331–6344.

31. Driscoll,M.E., Romine,M.F., Juhn,F.S., Serres,M.H., McCue,L.A., Beliaev,A.S., Fredrickson,J.K. and Gardner,T.S. (2007) Identification of diverse carbon utilization pathways in *Shewanella oneidensis* MR-1 via expression profiling. *Genome Inform.*, **18**, 287–298.

32. Gao,H., Wang,Y., Liu,X., Yan,T., Wu,L., Alm,E., Arkin,A., Thompson,D.K. and Zhou,J. (2004) Global transcriptome analysis of the heat shock response of *Shewanella oneidensis*. *J. Bacteriol.*, **186**, 7796–7803.

33. Gao,H.C., Yang,Z.M.K., Wu,L.Y., Thompson,D.K. and Zhou,J.Z. (2006) Global transcriptome analysis of the cold shock response of *Shewanella oneidensis* MR-1 and mutational analysis of its classical cold shock proteins. *J. Bacteriol.*, **188**, 4560–4569.

34. Leaphart,A.B., Thompson,D.K., Huang,K., Alm,E., Wan,X.F., Arkin,A., Brown,S.D., Wu,L.Y., Yan,T.F., Liu,X.D. *et al.* (2006) Transcriptome profiling of *Shewanella oneidensis* gene expression following exposure to acidic and alkaline pH. *J. Bacteriol.*, **188**, 1633–1642.

35. Qiu,X.Y., Daly,M.J., Vasilenko,A., Omelchenko,M.V., Gaidamakova,E.K., Wu,L.Y., Zhou,J.Z., Sundin,G.W. and Tiedje,J.M. (2006) Transcriptome analysis applied to survival of *Shewanella oneidensis* MR-1 exposed to ionizing radiation. *J. Bacteriol.*, **188**, 1199–1204.

36. Tang,Y.J.J., Hwang,J.S., Wemmer,D.E. and Keasling,J.D. (2007) *Shewanella oneidensis* MR-1 fluxome under various oxygen conditions. *Appl. Environ. Microbiol.*, **73**, 718–729.

37. Thompson,M.R., VerBerkmoes,N.C., Chourey,K., Shah,M., Thompson,D.K. and Hettich,R.L. (2007) Dosage-dependent proteome response of *Shewanella oneidensis* MR-1 to acute chromate challenge. *J. Proteome Res.*, **6**, 1745–1757.

38. Yang,Y., Harris,D.P., Luo,F., Wu,L., Parsons,A.B., Palumbo,A.V. and Zhou,J. (2008) Characterization of the *Shewanella oneidensis* Fur gene: roles in iron and acid tolerance response. *BMC Genomics*, **9(Suppl 1)**, S11.

39. Faith,J.J., Driscoll,M.E., Fusaro,V.A., Cosgrove,E.J., Hayete,B., Juhn,F.S., Schneider,S.J. and Gardner,T.S. (2008) Many microbe microarrays database: uniformly normalized affymetrix compendia with structured experimental metadata. *Nucleic Acids Res.*, **36**, D866–D870.

40. di Bernardo,D., Thompson,M.J., Gardner,T.S., Chobot,S.E., Eastwood,E.L., Wojtovich,A.P., Elliott,S.J., Schaus,S.E. and Collins,J.J. (2005) Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nat. Biotechnol.*, **23**, 377–383.

41. Faith,J.J., Hayete,B., Thaden,J.T., Mogno,I., Wierzbowski,J., Cottarel,G., Kasif,S., Collins,J.J. and Gardner,T.S. (2007) Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.*, **5**, e8.

42. Gardner,T.S., di Bernardo,D., Lorenz,D. and Collins,J.J. (2003) Inferring genetic networks and identifying compound mode of action via expression profiling. *Science*, **301**, 102–105.

43. Bansal,M., Della Gatta,G. and di Bernardo,D. (2006) Inference of gene regulatory networks and compound mode of action from time course gene expression profiles. *Bioinformatics*, **22**, 815–822.

44. Zampieri,M., Legname,G., Segre,D. and Altafini,C. (2011) A system-level approach for deciphering the transcriptional response to prion infection. *Bioinformatics*, **27**, 3407–3414.

45. Mahadevan,R., Edwards,J.S. and Doyle,F.J. III (2002) Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.*, **83**, 1331–1340.

46. Kauffman,K.J., Prakash,P. and Edwards,J.S. (2003) Advances in flux balance analysis. *Curr. Opin. Biotechnol.*, **14**, 491–496.

47. Price,N.D., Reed,J.L. and Palsson,B.O. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.*, **2**, 886–897.

48. Pinchuk,G.E., Hill,E.A., Geydebrekht,O.V., De Ingeniis,J., Zhang,X., Osterman,A., Scott,J.H., Reed,S.B., Romine,M.F., Konopka,A.E. *et al.* (2010) Constraint-based model of *Shewanella oneidensis* MR-1 metabolism: a tool for data analysis and hypothesis generation. *PLoS Comput. Biol.*, **6**, e1000822.

49. Baev,M.V., Baev,D., Radek,A.J. and Campbell,J.W. (2006) Growth of *Escherichia coli* MG1655 on LB medium: determining metabolic strategy with transcriptional microarrays. *Appl. Microbiol. Biotechnol.*, **71**, 323–328.

50. Beg,Q.K., Vazquez,A., Ernst,J., de Menezes,M.A., Bar-Joseph,Z., Barabasi,A.L. and Oltvai,Z.N. (2007) Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl Acad. Sci. USA*, **104**, 12663–12668.

51. Aach,J. and Church,G.M. (2001) Aligning gene expression time series with time warping algorithms. *Bioinformatics*, **17**, 495–508.

52. Bremer,H., Dennis,P. and Ehrenberg,M. (2003) Free RNA polymerase and modeling global transcription in *Escherichia coli*. *Biochimie*, **85**, 597–609.

53. Liang,S.T., Xu,Y.C., Dennis,P. and Bremer,H. (2000) mRNA composition and control of bacterial gene expression. *J Bacteriol*, **182**, 3037–3044.

54. Dong,T. and Schellhorn,H.E. (2009) Control of RpoS in global gene expression of *Escherichia coli* in minimal media. *Mol. Genet. Genomics*, **281**, 19–33.

55. Ferenci,T. (2003) What is driving the acquisition of mutS and rpoS polymorphisms in *Escherichia coli? Trends Microbiol.*, **11**, 457–461.

56. Battesti,A., Majdalani,N. and Gottesman,S. (2011) The RpoS-mediated general stress response in *Escherichia coli* (*). *Annu. Rev. Microbiol.*, **65**, 189–213.

57. Potrykus,K. and Cashel,M. (2008) (p)ppGpp: still magical? *Annu. Rev. Microbiol.*, **62**, 35–51.

58. Costanzo,A., Nicoloff,H., Barchinger,S.E., Banta,A.B., Gourse,R.L. and Ades,S.E. (2008) ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigmaE in *Escherichia coli* by both direct and indirect mechanisms. *Mol. Microbiol.*, **67**, 619–632.

59. Piper,S.E., Mitchell,J.E., Lee,D.J. and Busby,S.J. (2009) A global view of *Escherichia coli* Rsd protein and its interactions. *Mol. Biosyst.*, **5**, 1943–1947.

60. Baker,C.S., Morozov,I., Suzuki,K., Romeo,T. and Babitzke,P. (2002) CsrA regulates glycogen biosynthesis by preventing translation of *glgC* in *Escherichia coli. Mol. Microbiol.*, **44**, 1599–1610.

61. Ernst,J., Vainas,O., Harbison,C.T., Simon,I. and Bar-Joseph,Z. (2007) Reconstructing dynamic regulatory maps. *Mol. Syst. Biol.*, **3**, 74.

62. Smith,E.N. and Kruglyak,L. (2008) Gene-environment interaction in yeast gene expression. *PLoS Biol.*, **6**, e83.

63. Bar-Joseph,Z. (2004) Analyzing time series gene expression data. *Bioinformatics*, **20**, 2493–2503.

64. Ernst,J., Beg,Q.K., Kay,K.A., Balazsi,G., Oltvai,Z.N. and Bar-Joseph,Z. (2008) A semi-supervised method for predicting transcription factor-gene interactions in *Escherichia coli. PLoS Comput. Biol.*, **4**, e1000044.

65. Tegner,J., Yeung,M.K., Hasty,J. and Collins,J.J. (2003) Reverse engineering gene networks: integrating genetic perturbations with dynamical modeling. *Proc. Natl Acad. Sci. USA*, **100**, 5944–5949.

66. Dwyer,D.J., Kohanski,M.A. and Collins,J.J. (2009) Role of reactive oxygen species in antibiotic action and resistance. *Curr. Opin. Microbiol.*, **12**, 482–489.

67. Dwyer,D.J., Kohanski,M.A., Hayete,B. and Collins,J.J. (2007) Gyrase inhibitors induce an oxidative damage cellular death pathway in *Escherichia coli. Mol. Syst. Biol.*, **3**, 91.

68. Kohanski,M.A., Dwyer,D.J., Hayete,B., Lawrence,C.A. and Collins,J.J. (2007) A common mechanism of cellular death induced by bactericidal antibiotics. *Cell*, **130**, 797–810.

69. Kohanski,M.A., Dwyer,D.J., Wierzbowski,J., Cottarel,G. and Collins,J.J. (2008) Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell*, **135**, 679–690.

70. Dorr,T., Vulic,M. and Lewis,K. (2010) Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli. PLoS Biol.*, **8**, e1000317.

71. Chen,H., Yi,C., Zhang,J., Zhang,W., Ge,Z., Yang,C.G. and He,C. (2010) Structural insight into the oxidation-sensing mechanism of the antibiotic resistance of regulator MexR. *EMBO Rep.*, **11**, 685–690.

72. Maclean,A.M., White,C.E., Fowler,J.E. and Finan,T.M. (2009) Identification of a hydroxyproline transport system in the legume endosymbiont *Sinorhizobium meliloti. Mol. Plant. Microbe. Interact.*, **22**, 1116–1127.

73. Aguilera,L., Campos,E., Gimenez,R., Badia,J., Aguilar,J. and Baldoma,L. (2008) Dual role of LldR in regulation of the *lldPRD* operon, involved in L-lactate metabolism in *Escherichia coli. J. Bacteriol.*, **190**, 2997–3005.

74. Pinchuk,G.E., Rodionov,D.A., Yang,C., Li,X., Osterman,A.L., Dervyn,E., Geydebrekht,O.V., Reed,S.B., Romine,M.F., Collart,F.R. *et al.* (2009) Genomic reconstruction of *Shewanella oneidensis* MR-1 metabolism reveals a previously uncharacterized machinery for lactate utilization. *Proc. Natl Acad. Sci. USA*, **106**, 2874–2879.

75. Schwartz,C.J., Giel,J.L., Patschkowski,T., Luther,C., Ruzicka,F.J., Beinert,H. and Kiley,P.J. (2001) IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. *Proc. Natl Acad. Sci. USA*, **98**, 14895–14900.

76. Yeo,W.S., Lee,J.H., Lee,K.C. and Roe,J.H. (2006) IscR acts as an activator in response to oxidative stress for the suf operon encoding Fe-S assembly proteins. *Mol. Microbiol.*, **61**, 206–218.

77. Compan,I. and Touati,D. (1994) Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA. *Mol. Microbiol.*, **11**, 955–964.

78. Lassak,J., Henche,A.L., Binnenkade,L. and Thormann,K.M. (2010) ArcS, the cognate sensor kinase in an atypical Arc system of *Shewanella oneidensis* MR-1. *Appl. Environ. Microbiol.*, **76**, 3263–3274.

79. Gralnick,J.A., Brown,C.T. and Newman,D.K. (2005) Anaerobic regulation by an atypical Arc system in *Shewanella oneidensis. Mol. Microbiol.*, **56**, 1347–1357.

80. Kiley,P.J. and Beinert,H. (1998) Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol. Rev.*, **22**, 341–352.

81. Saffarini,D.A. and Nealson,K.H. (1993) Sequence and genetic characterization of *etrA*, an *fnr* analog that regulates anaerobic respiration in *Shewanella putrefaciens* MR-1. *J. Bacteriol.*, **175**, 7938–7944.

82. Cruz-Garcia,C., Murray,A.E., Rodrigues,J.L., Gralnick,J.A., McCue,L.A., Romine,M.F., Loffler,F.E. and Tiedje,J.M. (2011) Fnr (EtrA) acts as a fine-tuning regulator of anaerobic metabolism in *Shewanella oneidensis* MR-1. *BMC Microbiol.*, **11**, 64.

83. Gao,H., Wang,X., Yang,Z.K., Chen,J., Liang,Y., Chen,H., Palzkill,T. and Zhou,J. (2010) Physiological roles of ArcA, Crp, and EtrA and their interactive control on aerobic and anaerobic respiration in *Shewanella oneidensis. PLoS One*, **5**, e15295.

84. Beliaev,A.S., Thompson,D.K., Fields,M.W., Wu,L.Y., Lies,D.P., Nealson,K.H. and Zhou,J.Z. (2002) Microarray transcription profiling of a *Shewanella oneidensis etrA* mutant. *J. Bacteriol.*, **184**, 4612–4616.

85. Maier,T.M. and Myers,C.R. (2001) Isolation and characterization of a *Shewanella* putrefaciens MR-1 electron transport regulator etrA mutant: reassessment of the role of EtrA. *J. Bacteriol.*, **183**, 4918–4926.

86. Yang,Y., Harris,D.P., Luo,F., Xiong,W., Joachimiak,M., Wu,L., Dehal,P., Jacobsen,J., Yang,Z., Palumbo,A.V. *et al.* (2009) Snapshot of iron response in *Shewanella oneidensis* by gene network reconstruction. *BMC Genomics*, **10**, 131.

87. Charania,M.A., Brockman,K.L., Zhang,Y., Banerjee,A., Pinchuk,G.E., Fredrickson,J.K., Beliaev,A.S. and Saffarini,D.A. (2009) Involvement of a membrane-bound class III adenylate cyclase in regulation of anaerobic respiration in *Shewanella oneidensis* MR-1. *J. Bacteriol.*, **191**, 4298–4306.

88. Novichkov,P.S., Laikova,O.N., Novichkova,E.S., Gelfand,M.S., Arkin,A.P., Dubchak,I. and Rodionov,D.A. (2010) RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. *Nucleic Acids Res.*, **38**, D111–D118.

89. Karpinets,T.V., Romine,M.F., Schmoyer,D.D., Kora,G.H., Syed,M.H., Leuze,M.R., Serres,M.H., Park,B.H., Samatova,N.F. and Uberbacher,E.C. (2010) *Shewanella* knowledgebase: integration of the experimental data and computational predictions suggests a biological role for transcription of intergenic regions. *Database*, **2010**, baq012.

90. Novichkov,P.S., Rodionov,D.A., Stavrovskaya,E.D., Novichkova,E.S., Kazakov,A.E., Gelfand,M.S., Arkin,A.P., Mironov,A.A. and Dubchak,I. (2010) RegPredict: an integrated system for regulon inference in prokaryotes by comparative genomics approach. *Nucleic Acids Res.*, **38**, W299–W307.

91. Keseler,I.M., Collado-Vides,J., Santos-Zavaleta,A., Peralta-Gil,M., Gama-Castro,S., Muniz-Rascado,L., Bonavides-Martinez,C., Paley,S., Krummenacker,M., Altman,T. *et al.* (2011) EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Res.*, **39**, D583–D590.

92. Tintut,Y., Wang,J.T. and Gralla,J.D. (1995) A novel bacterial transcription cycle involving sigma 54. *Genes Dev.*, **9**, 2305–2313.

93. Reitzer,L. (2003) Nitrogen assimilation and global regulation in *Escherichia coli. Annu. Rev. Microbiol.*, **57**, 155–176.

94. Belanger,A.E. and Hatfull,G.F. (1999) Exponential-phase glycogen recycling is essential for growth of *Mycobacterium smegmatis. J. Bacteriol.*, **181**, 6670–6678.

95. Leckie,M.P., Ng,R.H., Porter,S.E., Compton,D.R. and Dietzler,D.N. (1983) Regulation of bacterial glycogen synthesis. Stimulation of glycogen synthesis by endogenous and exogenous cyclic adenosine 3':5'-monophosphate in *Escherichia coli* and the requirement for a functional CRP gene. *J. Biol. Chem.*, **258**, 3813–3824.

96. Preiss,J., Yung,S.G. and Baecker,P.A. (1983) Regulation of bacterial glycogen synthesis. *Mol. Cell Biochem.*, **57**, 61–80.

97. Seibold,G.M. and Eikmanns,B.J. (2007) The glgX gene product of *Corynebacterium glutamicum* is required for glycogen degradation and for fast adaptation to hyperosmotic stress. *Microbiology*, **153**, 2212–2220.

98. Hoskisson,P.A., England,R., Sharples,G.P. and Hobbs,G. (2004) Modulation of glycogen and trehalose levels in *Micromonospora echinospora* (ATCC 15837). *Antonie Van Leeuwenhoek*, **86**, 225–233.

99. Hernandez,M.A., Mohn,W.W., Martinez,E., Rost,E., Alvarez,A.F. and Alvarez,H.M. (2008) Biosynthesis of storage compounds by *Rhodococcus jostii* RHA1 and global identification of genes involved in their metabolism. *BMC Genomics*, **9**, 600.

100. Elias,D.A., Tollaksen,S.L., Kennedy,D.W., Mottaz,H.M., Giometti,C.S., McLean,J.S., Hill,E.A., Pinchuk,G.E., Lipton,M.S., Fredrickson,J.K. *et al.* (2008) The influence of cultivation methods on *Shewanella oneidensis* physiology and proteome expression. *Arch. Microbiol.*, **189**, 313–324.

101. Tang,Y.J.J., Meadows,A.L. and Keasling,J.D. (2007) A kinetic model describing *Shewanella oneidensis* MR-1 growth, substrate consumption, and product secretion. *Biotechnol. Bioeng.*, **96**, 125–133.

102. Hazelwood,L.A., Walsh,M.C., Luttik,M.A., Daran-Lapujade,P., Pronk,J.T. and Daran,J.M. (2009) Identity of the growth-limiting nutrient strongly affects storage carbohydrate accumulation in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.*, **75**, 6876–6885.

103. Schulze,U., Liden,G., Nielsen,J. and Villadsen,J. (1996) Physiological effects of nitrogen starvation in an anaerobic batch culture of *Saccharomyces cerevisiae*. *Microbiology*, **142(Pt 8)**, 2299–2310.

104. Deutschbauer,A., Price,M.N., Wetmore,K.M., Shao,W., Baumohl,J.K., Xu,Z., Nguyen,M., Tamse,R., Davis,R.W. and Arkin,A.P. (2011) Evidence-based annotation of gene function in *Shewanella oneidensis* MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genet.*, **7**, e1002385.

105. Syed,M.H., Karpinets,T.V., Leuze,M.R., Kora,G.H., Romine,M.R. and Uberbacher,E.C. (2009) ShewregDB: database and visualization environment for experimental and predicted regulatory information in *Shewanella oneidensis* MR-1. *Bioinformation*, **4**, 169–172.

106. Gadd,G.M. (2010) Metals, minerals and microbes: geomicrobiology and bioremediation. *Microbiol.-SGM*, **156**, 609–643.

107. Rosenbaum,M., Cotta,M.A. and Angenent,L.T. (2010) Aerated *Shewanella oneidensis* in continuously fed bioelectrochemical systems for power and hydrogen production. *Biotechnol. Bioeng.*, **105**, 880–888.

108. Rosenbaum,M., He,Z. and Angenent,L.T. (2010) Light energy to bioelectricity: photosynthetic microbial fuel cells. *Curr. Opin. Biotechnol.*, **21**, 259–264.

109. Zaslaver,A., Mayo,A.E., Rosenberg,R., Bashkin,P., Sberro,H., Tsalyuk,M., Surette,M.G. and Alon,U. (2004) Just-in-time transcription program in metabolic pathways. *Nat. Genet.*, **36**, 486–491.

110. Li,G., Ruan,X., Auerbach,R.K., Sandhu,K.S., Zheng,M., Wang,P., Poh,H.M., Goh,Y., Lim,J., Zhang,J. *et al.* (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell*, **148**, 84–98.

111. Everitt,B. and Dunn,G. (2001) *Applied Multivariate Data Analysis*, 2nd edn. Arnold; Oxford University Press, London, New York.

112. Zheng,Q. and Wang,X.J. (2008) GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Res.*, **36**, W358–W363.

113. Storey,J.D. (2002) A direct approach to false discovery rates. *J. R. Stat. Soc.: Series B (Statistical Methodol.)*, **64**, 479–498.

114. Bansal,M. and di Bernardo,D. (2007) Inference of gene networks from temporal gene expression profiles. *IET Syst. Biol.*, **1**, 306–312.

115. Cosgrove,E.J., Zhou,Y., Gardner,T.S. and Kolaczyk,E.D. (2008) Predicting gene targets of perturbations via network-based filtering of mRNA expression compendia. *Bioinformatics*, **24**, 2482–2490.

116. Balaji,S., Babu,M.M., Iyer,L.M., Luscombe,N.M. and Aravind,L. (2006) Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. *J. Mol. Biol.*, **360**, 213–227.

117. Chang,D.E., Smalley,D.J. and Conway,T. (2002) Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. *Mol. Microbiol.*, **45**, 289–306.

118. Thormann,K.M., Duttler,S., Saville,R.M., Hyodo,M., Shukla,S., Hayakawa,Y. and Spormann,A.M. (2006) Control of formation and cellular detachment from *Shewanella oneidensis* MR-1 biofilms by cyclic di-GMP. *J. Bacteriol.*, **188**, 2681–2691.

119. Rakshe,S., Leff,M. and Spormann,A.M. (2011) Indirect modulation of the intracellular c-Di-GMP level in *Shewanella oneidensis* MR-1 by MxdA. *Appl. Environ. Microbiol.*, **77**, 2196–2198.

120. Saville,R.M., Dieckmann,N. and Spormann,A.M. (2010) Spatiotemporal activity of the mshA gene system in *Shewanella oneidensis* MR-1 biofilms. *FEMS Microbiol. Lett.*, **308**, 76–83.