Illuminating the function of a master regulator of *Caulobacter crescentus* gene expression

by

Andy H. Yuan

B.A. Biology
Harvard University, 2008

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of

Master of Science in Biology

at the

Massachusetts Institute of Technology

June 2012

Signature of Author: Signature redacted
Department of Biology
May 18th, 2012

Certified by: Signature redacted
Michael T. Laub
Associate Professor of Biology
Thesis Supervisor

Accepted by: Signature redacted
Stephen P. Bell
Professor of Biology
Chair, Biology Graduate Committee

Illuminating the function of a master regulator of *Caulobacter crescentus* gene expression

by

Andy H. Yuan

Submitted to the Department of Biology on May 18th, 2012
in partial fulfillment of the requirements for the degree of
Master of Science in Biology

Abstract

Bacteria must allocate cellular resources required for transcription in a manner that accommodates the temporally regulated expression of functionally specialized gene products and the constitutive synthesis of stable RNAs required for ribosome biogenesis and translation. The *Caulobacter crescentus* cell cycle has been well-characterized and represents a powerful model system for studying the regulation of global gene expression patterns and "just-in-time" transcription. Here, we provide evidence that the *Caulobacter* GcrA protein directly regulates transcription at active promoters of highly expressed genes, particularly those encoding stable RNAs. Furthermore, we demonstrate that GcrA chromosome occupancy is dynamic and subject to epigenetic regulation. Our findings establish GcrA as an integral regulator of bacterial physiology and development.

Thesis advisor: Michael T. Laub
Title: Associate Professor of Biology

Introduction

Cell cycle progression requires the fine-tuned temporal regulation of gene expression and the spatial organization of gene products. In the α-proteobacterium *Caulobacter crescentus*, integration of these complex processes establishes cellular asymmetry in part via the dynamic polar localization of histidine kinases and their regulators [1] and culminates in a round of asymmetric cell division. Specifically, *Caulobacter* cell division yields two progeny cells – a motile, replication-incompetent "swarmer" cell and a sessile, replication-competent "stalked" cell – that differ both in their morphological features and developmental fates (Fig. 1).

The molecular phenomena responsible for orchestrating *Caulobacter* development have been the subject of extensive research [2], the cumulative findings of which have identified four essential regulators of the *Caulobacter* cell cycle that directly or indirectly control gene expression levels and the timing of DNA replication: (i) DnaA, the dual-function bacterial replication initiation factor and regulator of transcription, (ii) CtrA, a two-component response regulator and replication inhibitor (iii) CcrM, an N6-deoxyadenosine methyltransferase, and (iv) GcrA, a highly conserved protein, the biological function of which remains elusive.

GcrA was first identified as an essential gene product, the intracellular accumulation of which oscillates out of phase with CtrA [3] and has since then been proposed to play a critical role in *Caulobacter* cell cycle progression. Specifically, GcrA has been postulated to regulate the expression of *ctrA* itself along with approximately 50

genes encoding a suite of proteins involved in replication elongation and chromosome segregation [3, 4, 5].

Here, we present evidence that GcrA positively regulates transcription at active promoters of highly expressed genes, particularly those encoding stable RNAs (ribosomal RNA [rRNA], transfer RNA [tRNA], and ribosomal protein-encoding RNA [RP-RNA]). We also demonstrate that GcrA promoter occupancy is dynamic and influenced by chromosome methylation state. Based on our findings, we propose that the essential function of GcrA be assigned to the regulation of global gene expression in a manner that couples transcription to ribosome biogenesis and other fundamental processes underlying bacterial physiology and development.

Results

Phenotypic consequences of GcrA depletion

Unlike its counterparts DnaA, CtrA, and CcrM, all of which have been studied intensively in *Caulobacter* and several other model bacteria [6, 7, 8], GcrA has been characterized neither phenotypically nor mechanistically. In order to investigate the phenotypic consequences of GcrA depletion, we constructed strain AHY3, which harbors a single chromosomal copy of *gcrA* under the control of a vanillate-inducible promoter [9]. Cells grown in medium supplemented with vanillate were synchronized by density gradient centrifugation, and purified swarmer cells were either placed on agarose-pads and monitored by time-lapse microscopy or released into liquid medium lacking vanillate and subjected to flow cytometry analysis. Cells depleted of GcrA in this manner exhibited several morphological anomalies associated with defects in cell division (Fig. 2A) and accumulated DNA content over time (Fig. 2B).

Reassessing the essential function of GcrA

Previous studies have postulated that GcrA activates one of two promoters driving *ctrA* transcription and thereby propels the *Caulobacter* cell cycle forward [3]. To test this hypothesis, we monitored CtrA levels by western blot over the course of a single cell cycle beginning with synchronized swarmer cell populations isolated from either wild-type cells or strain AHY3. The cell cycle-regulated synthesis and proteolysis of CtrA has been well-established. In brief, CtrA is present in swarmer cells, actively degraded by the ClpXP protease during the swarmer-to-stalked cell transition, and resynthesized in

"pre-divisional" cells [10]. While we observed a significant delay in CtrA reaccumulation in cells lacking GcrA, CtrA levels were nevertheless restored over time (Fig. 3). Furthermore, we found that GcrA remained essential in cells constitutively expressing *ctrA* from a xylose-inducible promoter (data not shown). These observations suggest that the essentiality of GcrA cannot be solely attributed to defects in CtrA accumulation or *ctrA* transcription.

It has also been suggested that GcrA is required for the proper expression of genes encoding regulators of replication elongation and chromosome segregation [3, 4, 5]. To test this second hypothesis, we performed comparative genomic hybridization (CGH) microarray analyses of DNA isolated from mixed populations of either wild-type cells or cells depleted of GcrA. We detected only marginal changes in the copy number of chromosomal loci upon GcrA depletion, suggesting that the essential function of GcrA cannot be assigned solely to defects in replication elongation (Fig. 4). Taken together, these results indicate that existing models for GcrA function remain ill-defined.

A screen for loss-of-function mutations that bypass the essentiality of GcrA

To deepen our understanding of GcrA's essential function, we performed an unbiased genetic selection for suppressors of GcrA via transposon-mediated mutagenesis. Briefly, the *EZ-Tn5* transposon [11] was introduced into strain AHY16, which requires vanillate for *gcrA* expression, and mutagenized cells were directly plated on medium lacking vanillate. This experimental approach led to the identification of several loss-of-function mutations that render *Caulobacter* viable in the absence of GcrA. Interestingly,

the vast majority of these mutations map to genes - specifically, *pstS*, *pstC*, *pstA*, and *pstB* - encoding components of the *Caulobacter* phosphate uptake machinery (Fig. 5A). Previous studies concerning phosphate metabolism in *Escherichia coli* [12] and *Caulobacter* [13] predict that these genetic perturbations activate the PhoR histidine kinase and thereby hyperactivate PhoR's cognate response regulator, PhoB (Fig. 5B). The results of epistasis analyses indicate that PhoR and PhoB are indeed required for suppression, as deletion of either *phoR* or *phoB* in strains bearing a transposon insertion in *pstS* restores the essentiality of GcrA (data not shown).

Remarkably, the same loss-of-function mutations in *pstS*, *pstC*, *pstA*, and *pstB* that render *Caulobacter* viable in the absence of GcrA were previously reported to bypass the essentiality of CcrM [14], an N6-deoxyadenosine methyltransferase recently shown to recognize and modify "GANTC" sequences in a distributive manner [15]. We confirmed these results by performing an independent genetic selection for suppressors of CcrM via transposon-mediated mutagenesis of strain LS2144, in which *ccrM* expression is under the control of a xylose-inducible, glucose-repressible promoter [16]. As the same set of genetic perturbations can bypass the essentiality of both GcrA and CcrM, we propose that a previously uncharacterized relationship likely exists between the function of GcrA and chromosome methylation state.

Changes in global gene expression upon GcrA depletion

We proceeded to perform a series of microarray experiments to assess the effect of GcrA depletion on global gene expression. In partial agreement with a previously

reported microarray study [3], cells depleted of GcrA in an otherwise wild-type genetic background exhibited modest defects in the transcription of several genes, many of which exhibit peak expression in swarmer cells (Fig. 6). While these results are consistent with GcrA-depleted cells exhibiting a defect in swarmer cell birth, they are potentially confounded by the fact that cells rapidly lose viability once *gcrA* expression ceases [3].

We reasoned that depletion of GcrA in a permissive genetic background might unveil otherwise unobservable changes in global transcription patterns. Accordingly, we examined the gene expression profile of GcrA-depleted cells in strain AHY25, which harbors both a single copy of *gcrA* under the control of a vanillate-inducible promoter and a transposon insertion in *pstS*. Under these conditions, we observed significant defects in the expression of a unique set of genes in addition to previously documented changes in swarmer cell-specific transcription. Specifically, we found that GcrA was required for RP-RNA and tRNA expression (Fig. 7). While the design of our microarrays precluded the direct detection of changes in rRNA synthesis, it has been well-established that intracellular levels of RP-RNA correlate with rRNA abundance [17]. Accordingly, out results indicate that GcrA positively regulates transcription of stable RNAs.

GcrA is a conditionally essential gene product

Several studies have established that intracellular stable RNA levels correlate with growth rate and are subject to regulation by changes in nutritional environment [18]. In light of our microarray results, we hypothesized that *gcrA* expression may be dispensable under particular growth conditions. Indeed, we found that cells lacking GcrA are viable

when grown in either minimal medium at 30°C or rich medium at 18°C (data not shown). These results establish GcrA as a conditionally essential gene product and support a model according to which GcrA's essential function involves the growth rate-dependent regulation of ribosome biogenesis.

GcrA is enriched at promoters of highly expressed genes

The results of our gene expression studies – along with experimental evidence that purified GcrA protein is capable of binding DNA *in vitro* (data not shown) – suggest that GcrA functions as a transcription factor directly or indirectly responsible for regulating global gene expression, particularly the expression of stable RNAs. To gain mechanistic insight into GcrA's function, we set forth to identify all sites of GcrA chromosome occupancy by chromatin immunoprecipitation coupled to deep sequencing (ChIP-Seq).

Qualitative analyses of our ChIP-Seq data revealed that GcrA primarily occupies intergenic, promoter-proximal loci throughout the *Caulobacter* genome. Notably, GcrA is not present in preparations of tandem-affinity purified *Caulobacter* RNA polymerase (RNAP) [19], suggesting that GcrA is not a novel RNAP-associated protein (data not shown).

To assess GcrA occupancy in a quantitative manner, we calculated the number of sequences aligned to every *Caulobacter* promoter, loosely defined as the 150 bp stretch of DNA upstream of each annotated *Caulobacter* open reading frame start codon. By employing this strategy, we found that GcrA is most enriched at promoters of genes

encoding stable RNAs (Fig. 8A). Furthermore, by comparing our ChIP-Seq data to previously reported measurements of cell cycle-regulated gene expression levels [20], we observed that GcrA is enriched at promoters of the most highly expressed *Caulobacter* protein-coding genes (p=1.26E-8 as determined by the Wilcoxon rank-sum test; p=2.13E-4 [two-tailed] as determined by Student's unpaired t-test).

We proceeded to utilize the MEME-ChIP algorithm [21] to analyze sequences of the 1,500 promoters most highly-enriched for GcrA occupancy. This computational approach led to the identification of two A/T-rich motifs that either reflect consensus binding sites for GcrA or core promoter elements conserved among GcrA-regulated genes (Fig. 8B).

GcrA chromosome occupancy is both dynamic and subject to epigenetic regulation

As (i) *Caulobacter* cells are readily synchronized by density gradient centrifugation, (ii) changes in global gene expression during *Caulobacter* development have been well-documented [20], and (iii) GcrA levels are cell cycle-regulated [3], we set out to monitor genome-wide GcrA occupancy as a function of the cell cycle with a temporal resolution of 20 minutes. Furthermore, as the results of our independent genetic selections for suppressors of GcrA and CcrM suggest that a relationship exists between GcrA function and chromosome methylation state, we also examined global GcrA occupancy in mixed populations of cells bearing either constitutively-methylated or unmethylated chromosomes.

Intriguingly, we identified several genes, the promoters of which are dynamically occupied by GcrA (Fig. 9A). Moreover, we found that the time at which GcrA occupies a particular gene promoter often correlates with the time at which that gene is maximally expressed upon synchronization and passage through the cell cycle (Fig. 9B). Furthermore, we observed instances where chromosome methylation state directly or indirectly influenced GcrA occupancy (Fig. 10). Taken together, these ChIP-Seq results represent proofs-of-principle underscoring the feasibility of monitoring the genome-wide epigenetic regulation and dynamic behavior of a single DNA-binding protein as a function of the bacterial cell cycle.

Discussion

Our cumulative findings provide evidence that GcrA directly regulates transcription at active promoters of highly expressed genes, particularly those encoding stable RNAs. Moreover, our work establishes an unprecedented link between an epigenetic DNA modification and GcrA promoter occupancy. As (i) stable RNAs typically represent greater than 95% of the total RNA in a bacterial cell [18] and (ii) changes in chromosome methylation state have been suggested to control *Caulobacter* cell cycle progression [5], we propose that GcrA represents an integral regulator of bacterial physiology and development.

The mechanism of GcrA-dependent transcription activation

The bacterial RNAP holoenzyme consists of a catalytic multisubunit core enzyme ($\alpha_2\beta\beta'\omega$) associated with a σ factor, which mediates promoter-specific protein-DNA interactions [22]. Bacterial promoters are typically characterized by two conserved hexamers known as −10 and −35 elements, which interact with σ regions 2 and 4, respectively [23]. Furthermore, A/T-rich sequences known as UP elements can dramatically increase the strength of promoters, particularly those controlling rRNA transcription, by interacting with the C-terminal domains of the two RNAP α subunits (α-CTDs) [24].

While the A/T-rich sequence motifs identified by MEME-ChIP may represent consensus binding sites for GcrA, it is equally if not more likely that they represent previously uncharacterized *Caulobacter* UP elements. Several DNA-binding

transcription factors are known to activate gene expression by making direct protein-protein interactions with UP element-engaged α-CTDs [25, 26]. One such transcription factor is the *E. coli* Fis protein. Originally identified as a gene product required for site-specific recombination, Fis has since then been shown to bind upstream of *E. coli* UP elements, interact with one or both α-CTDs, and thereby activate transcription from rRNA promoters [26].

GcrA bears striking similarities to Fis in that both are small DNA-binding proteins exhibiting high isoelectric points (pI_{GcrA}=11.2 and pI_{Fis}=9.7) required for the maximal expression of stable RNAs. As *Caulobacter* lacks a Fis orthologue, we propose that the mechanism underlying GcrA-dependent transcription activation is likely to be analogous to that of Fis in *E. coli*.

Implications for Caulobacter physiology

The downregulation of stable RNA synthesis in response to amino acid starvation, a phenomenon known as the stringent response, has been the subject of extensive research in *E. coli* [27]. While the starvation signals required for stringent response induction differ between *E. coli* and *Caulobacter* [28], both bacteria adapt to nutrient deprivation by synthesizing the alarmone ppGpp and the RNAP-associated protein DksA. Together, ppGpp and DksA function to inhibit transcription from intrinsically unstable rRNA promoters [29], which form short-lived open complexes and exhibit an atypically high K_M for NTP incorporation [18]. As GcrA and ppGpp-DksA appear to have antagonistic effects on stable RNA transcription, it will be of great interest to study

GcrA's function in the context of the stringent response and its implications for *Caulobacter* survival in oligotrophic environments.

Experimental procedures

Bacterial strains and media

Caulobacter strains were constructed by utilizing a combination of homologous recombination strategies and general transduction as previously described [9, 30]. PYE (rich medium) and M2G (minimal medium) were supplemented with 0.5 mM vanillate, 0.3% xylose, or 0.2% glucose when indicated. Synchronizations were performed on mid-log phase cells using Percoll (GE Healthcare) density gradient centrifugation.

Time-lapse microscopy

Synchronized swarmer cells were immobilized on 1.5% PYE-agarose pads. Imaging was performed at 30°C using an Axiovert 200 microscopy (Zeiss) equipped with a 63x phase objective fitted with an objective heater (Bioptechs) and a culture dish heater (Warner Instruments). Images were acquired at 10 min intervals over the course of 6 hr and analyzed using ImageJ software.

Flow cytometry

Samples were prepared for flow cytometry as previously described [19]. Data collected was subsequently analyzed using FlowJo software.

Immunoblotting

Cell pellets, normalized to the OD_{600} of cultures, were flash-frozen in liquid nitrogen and resuspended in SDS sample buffer. Cells were lysed by heating to 95°C for 10 min. Equal amounts of total protein were subjected to SDS-PAGE on pre-cast 12%

Tris-HCl gels (Bio-Rad) and transferred to PVDF membranes. Proteins were detected using a 1:10,000 (α-GcrA [Covance], α-RpoA [Neoclone]) or 1:5,000 (α-CtrA [Covance]) dilution of primary antibody and a 1:10,000 dilution of an HRP-conjugated secondary antibody (Sigma-Aldrich). Blots were exposed to Kodak Biomax autoradiography films (Sigma-Aldrich), which were subsequently developed and scanned for further analysis.

CGH DNA microarray analysis

Genomic DNA was purified with a Genomic DNA Purification Kit (Qiagen) according to manufacturer instructions and fragmented by sonication to an average length of 500 bp with a Branson Sonifier (15% output, 3x 5 sec bursts). Equal quantities of DNA were labeled with either Cy3-dCTP or Cy5-dCTP using random hexamers (Invitrogen) and Klenow Fragment (NEB). Labeled DNA samples were mixed and purified with a QIAquick PCR Purification Kit (Qiagen). Samples were hybridized to custom Agilent arrays overnight, and arrays were scanned with an Agilent scanner (MIT BioMicroCenter).

Gene expression DNA microarray analysis

Microarray experiments were performed as previously described [19]. In brief, mid-log phase cells were harvested by centrifugation at 20,000 g for 30 sec and immediately flash-frozen in liquid nitrogen. Total RNA was isolated from cell pellets with an RNeasy Kit (Qiagen) according to manufacturer instructions. Equal quantities of

RNA were labeled with either Cy3-dCTP or Cy5-dCTP using random hexamers (Invitrogen) and a Superscript-II Kit (Invitrogen). Template RNA was degraded by the addition of NaOH, and labeled cDNA samples were mixed and purified with a QIAquick PCR Purification Kit (Qiagen). Samples were hybridized to custom Agilent arrays overnight, and arrays were scanned with an Agilent scanner (MIT BioMicroCenter).

ChIP-Seq

20 ml cultures were cross-linked in the presence of 10 mM Na_2HPO_4 (pH 7.6) and 1% formaldehyde for 10 min at room-temperature (RT). Reactions were quenched by the addition of 840 µl of 2.5 M glycine and incubated first for 5 min at RT and subsequently for 15 min on ice. Cross-linked cells were pelleted at 8,500 rpm for 5 min at 4°C and washed three times in 20 ml of cold PBS (pH 7.4).

Cell pellets were resuspended in 500 µl of TES buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA, 100 mM NaCl) supplemented with 35,000 U of Ready-Lyse Lysozyme Solution (Epicentre) and incubated for 15 min at RT. 500 µl of ChIP Buffer (1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl [pH 8.1], 167 mM NaCl) supplemented with EDTA-free Protease Inhibitor Cocktail Tablets (Roche, 1 tablet per 2 ml of ChIP Buffer) were added to lysates, which were subsequently incubated for 10 min at 37°C. DNA was fragmented by sonication to an average length of 500 bp with a Branson Sonifier (15% output, 6x 10 sec bursts), and samples were cleared of debris by centrifugation at 14,000 rpm for 5 min at 4°C.

Supernatants were diluted in ChIP Buffer supplemented with 0.01% SDS to a final protein concentration of 500 μg/ml as determined by the D_C Protein Assay (Biorad). 1 ml samples were pre-cleared for 1 hr at 4°C with 100 μl of Protein-A agarose slurry (Roche) that had been pre-blocked for 2 hr at 4°C in ChIP Blocking Buffer (ChIP Buffer supplemented with 100 μg of BSA and 300 μg of herring sperm DNA). Protein-A agarose was pelleted at 3,000 rpm for 5 min and supernatants were recovered.

2 ul of α-GcrA polyclonal antibody were added to pre-cleared supernatants, and samples were incubated overnight at 4°C. Samples were subsequently incubated with 100 μl of pre-blocked Protein-A Dynabeads (Invitrogen) for 6 hr at 4°C. Immune complexes were captured with a DynaMag magnet (Invitrogen) and washed consecutively for 15 min at 4°C in each of the following buffers: (i) Low Salt Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 150 mM NaCl), (ii) High Salt Wash Buffer (Low Salt Wash Buffer containing 500 mM NaCl), (iii) LiCl Wash Buffer (250 mM LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl [pH 8.1]), and (iv) TE Buffer (10 mM Tris-HCl [pH 8.1], 1 mM EDTA).

Washed immune complexes were eluted from Protein-A Dynabeads twice by vortexing samples in 250 μl of freshly prepared Elution Buffer (1% SDS, 100 mM $NaHCO_3$). 37.5 μl of 4 M NaCl and 2 μl of RNase A (0.5 mg/ml) were added to the collective eluates, and samples were incubated at 65°C overnight to reverse cross-linking. Immunoprecipitated DNA was purified with a QIAquick PCR Purification Kit

(Qiagen), eluted in 100 ul of TE Buffer, and prepared for deep sequencing on an Illumina Genome Analyzer II sequencer (MIT BioMicroCenter).

Acknowledgements

We thank J. Modell for ChIP-Seq computational support, K. Jonas for providing DnaA depletion strain KJ218, and J. Collier for providing CcrM depletion strain LS2144. This work was supported by an NSF predoctorol fellowship to A.H.Y.

Figure Legends

Figure 1. *Schematic of the Caulobacter cell cycle.* See main text for details.

Figure 2. *Phenotypic consequences of GcrA depletion.* (*A*) Strain AHY3 (Δ*gcrA*; P_{vanA}-*gcrA*) was grown in the presence of vanillate, and cells were synchronized by density gradient centrifugation. Swarmer cells were immobilized on 1.5% PYE-agarose pads containing or lacking vanillate and imaged at 10 min intervals over the course of 6 hr. (*B*) Wild-type cells and strain AHY3 were synchronized by density gradient centrifugation, and purified swarmer cells were released into liquid PYE lacking vanillate. Samples were collected at 60 min intervals over the course of 3 hr, and DNA content per cell was measured by flow cytometry.

Figure 3. *The essentiality of GcrA cannot be solely attributed to defects in CtrA accumulation.* Wild-type cells and strain AHY3 were synchronized by density gradient centrifugation, and purified swarmer cells were released into PYE lacking vanillate. CtrA and GcrA levels were monitored at 20 min intervals over the course of a single cell cycle. RpoA (the α-subunit of RNA polymerase) levels were monitored as loading controls.

Figure 4. *The essentiality of GcrA cannot be solely attributed to defects in replication elongation.* (*A*) Strain AHY3 was grown in PYE supplemented with vanillate, washed twice in PYE, and subsequently grown in PYE lacking vanillate. GcrA levels were monitored at 30 min intervals over the course of 3 hr. GcrA protein was no longer

detectable by western blot after 2 hr of growth in PYE lacking vanillate. (*B*) Results of CGH DNA microarray analyses of gDNA isolated from either wild-type cells or strain AHY3 3 hr-post GcrA depletion. Relative DNA content represents the ratio of fluorescence intensities of Cy5-labeled experimental sample DNA and Cy3-labeled reference sample DNA hybridized to custom Agilent arrays. Reference gDNA was isolated from strain KJ218 (Δ*dnaA*; P_{vanA}-*dnaA*) 3 hr-post DnaA depletion.

Figure 5. *Loss-of-function mutations in components of the Caulobacter phosphate-uptake machinery bypass the essentiality of GcrA.* (*A*) Mapped sites of *EZ-Tn5* transposon insertion [11] that render cells viable in the absence of GcrA. (*B*) Schematic of the conserved bacterial phosphate-uptake system [12, 13]. (*Left panel*) Under phosphate-replete conditions, periplasmic inorganic phosphate (P_i) enters the cell via the PstSCAB complex, which in turn inhibits the activity of the PhoR histidine kinase and thereby renders PhoB inactive. (*Right panel*) Growth under phosphate-limiting conditions or the genetic disruption of the PstSCAB complex (loss of PstS function is depicted) relieves inhibition of PhoR activity and results in the hyperactivation of PhoB. IM, inner membrane; OM, outer membrane.

Figure 6. *Changes in global gene expression upon GcrA depletion.* Results of gene expression DNA microarray analysis comparing total RNA isolated from strain AHY3 2 hr-post GcrA depletion to total RNA isolated from wild-type cells. Genes exhibiting a greater than or equal to 3-fold increase (yellow) or decrease (blue) in expression upon

GcrA depletion are shown. GcrA-regulated genes previously identified in [3] (grey) and genes primarily expressed in swarmer cells (black) are highlighted.

Figure 7. *Changes in global gene expression upon GcrA depletion in a permissive genetic background.* Results of gene expression DNA microarray analyses comparing total RNA isolated from strain AHY3 2 hr-post GcrA depletion, strain AHY6 (*pstS*::*Tn5*), and strain AHY25 (Δ*gcrA*; P_{vanA}-*gcrA*; *pstS*::*Tn5*) 2 hr-post GcrA depletion to total RNA isolated from wild-type cells. Genes exhibiting a greater than or equal to 3-fold increase (yellow) or decrease (blue) in expression upon GcrA depletion in strain AHY25 are shown. Genes encoding RP-RNA/tRNA (grey) are highlighted.

Figure 8. *GcrA is highly enriched at promoters of genes encoding stable RNAs.* (*A*) A rank-ordered list of the 50 genes most highly enriched for GcrA promoter occupancy is shown. Genes encoding RP-RNA/tRNA (grey) are highlighted. (*B*) Sequence motifs identified by the MEME-ChIP algorithm [21]. Computational analysis was performed with the following specifications: (i) distribution of motif occurrences = any number of repetitions (ii) number of different motifs = 5, (iii) minimum motif width = 4, (iv) maximum motif width = 10.

Figure 9. *GcrA promoter occupancy is dynamic.* (*A*) GcrA chromosome occupancy was monitored by ChIP-Seq (i) as a function of the cell cycle with a temporal resolution of 20 min (*rows 1-6*), (ii) in mixed populations of cells bearing either constitutively-methylated chromosomes (strain LS2144 grown in PYE supplemented with xylose) or unmethylated

chromosomes (strain LS2144 grown in PYE supplemented with glucose for 5 hr) (*rows 7-8*), and (iii) in cells depleted of GcrA (strain AHY3 grown in PYE lacking vanillate for 3 hr) (*row 9*). Data was visualized in the UCSC Genome Browser. Evidence for dynamic GcrA occupancy at P_{CC0761} and P_{CC0762} is shown. (*B*) GcrA occupancy at P_{CC0761} and P_{CC0762} correlates with the time at which *CC0761* and *CC0762* are maximally expressed upon synchronization and passage through the cell cycle [20].

Figure 10. *GcrA promoter occupancy is subject to epigenetic regulation.* GcrA chromosome occupancy was monitored by ChIP-Seq as described (Fig. 9A). (*A*) Evidence that GcrA occupies P_{CC0925} only when chromosomes are unmethylated is shown. (*B*) Evidence that GcrA occupies P_{CC2139} only when chromosomes are either constitutively-methylated or unmethylated is shown.

References

1. Tsokos, C.G, Perchuk, B.S. & Laub, M.T. (2011) A dynamic complex of signaling proteins uses polar localization to regulate cell-fate asymmetry in *Caulobacter crescentus. Dev. Cell.* 20(3): 329-341.

2. Laub, M.T., Shapiro, L. & McAdams, H.H. (2007) Systems biology of *Caulobacter. Ann. Rev. Genet.* 41: 429-441.

3. Holtzendorff, J., Hung, D., Brende, P., Reisenauer, A. Viollier, P.H., McAdams, H.H. & Shapiro, L. (2004) Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science.* 304(5673): 983-987.

4. Collier, J., Murray, S.R. & Shapiro, L. (2006) DnaA couples DNA replication and the expression of two cell cycle master regulators. *EMBO J.* 25(2): 346-356.

5. Collier, J., McAdams, H.H. & Shapiro, L. (2007) A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proc. Natl. Acad. Sci. USA.* 104(43): 17111-17116.

6. Jonas, K., Chen, Y.E. & Laub, M.T. (2011) Modularity of the bacterial cell cycle enables independent spatial and temporal control of DNA replication. *Curr. Biol.* 21(13): 1092-1101.

7. Barnett, M.J., Hung, D.Y., Reisenauer, A., Shapiro, L. & Long, S.R. (2001) A homologue of the CtrA cell cycle regulator is present and essential in *Sinorhizobium meliloti. J. Bacteriol.* 183(10): 3204-3210.

8. Wright, R., Stephens, C. & Shapiro, L. (1997) The CcrM DNA methyltransferase is widespread in the α subdivision of proteobacteria, and its essential functions are conserved in *Sinorhizobium meliloti* and *Caulobacter crescentus. J Bacteriol.* 179(18): 5867-5877.

9. Thanbichler, M., Iniesta, A.A. & Shapiro, L. (2007) A comprehensive set of plasmids for vanillate- and xylose-inducible gene expression in *Caulobacter crescentus. Nucleic Acids Res.* 35(20): e137.

10. Jenal, U. & Fuchs, T. (1998) An essential protease involved in bacterial cell cycle control. *EMBO J.* 17(19): 5658-5669

11. Kirby, J.R. (2007) *In vivo* mutagenesis using EZ-Tn5. *Methods Enzymol.* 421: 17-21.

12. Wanner, B.L. (1993) Gene regulation by phosphate in enteric bacteria. *J. Cell Biochem.* 51(1): 47-54.

13. Gonin, M., Quardokus, E.M., O'Donnol, D., Maddock, J. & Brun, Y.V. (2000) Regulation of stalk elongation by phosphate in *Caulobacter crescentus*. *J Bacteriol.* 182(2): 337-347.

14. McCollum, S.E. (2001) Effects of altering the temporal expression pattern of the essential DNA adenine methyltransferase, CcrM, in *Caulobacter crescentus*. *Dissertation* (Stanford University).

15. Albu, R.F., Jurkowski, T.P. & Jeltsch, A. (2011) The *Caulobacter crescentus* DNA-(adenine-N6)-methyltransferase CcrM methylates DNA in a distributive manner. *Nucleic Acids Res.* 40(4): 1708-1716.

16. Stephens, C., Reisenauer, A., Wright, R. & Shapiro, L. (1996) A cell cycle-regulated bacterial methyltransferase is essential for viability. *Proc. Natl. Acad. Sci. USA.* 93(3): 1210-1214.

17. Zengel, J.M. & Lindahl, L. (1994) Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog. Nucleic Acid Res. Mol. Biol.* 47: 331-370.

18. Paul, B.J., Ross, W., Gaal, T. & Gourse, R.L. (2004) rRNA transcription in *Escherichia coli*. *Ann. Rev. Genet.* 38: 749-770.

19. Gora, K.G., Tsokos, C.G., Chen, Y.E., Srinivasan, B.S., Perchuk, B.S. & Laub, M.T. (2010) A cell type-specific protein-protein interaction modulates transcriptional activity of a master regulator in *Caulobacter crescentus*. *Mol. Cell.* 39(3): 455-467.

20. Laub, M.T., McAdams, H.H., Feldblyum, T., Fraser C.M. & Shapiro, L. (2000) Global analysis of the genetic network controlling a bacterial cell cycle. *Science.* 290(5499): 2144-2148.

21. Machanick, P. & Bailey, T.L. (2011) MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics.* 27(12): 1696-1697.

22. Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J. & Young, B. (1998) The functional and regulatory roles of σ factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.* 63:141-155.

23. Murakami, K.S., Masuda, S. & Darst, S.A. (2002) Bacterial RNA polymerases: the wholo story. *Curr. Opin. Struct. Biol.* 13: 31-39.

24. Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. & Gourse, R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science.* 262(5138): 1407-1413.

25. Chen, H., Tang, H. & Ebright, R.H. (2003) Functional interaction between RNA polymerase α subunit C-terminal domain and σ^{70} in UP element and activator-dependent transcription. *Mol. Cell.* 11(6): 1621-1633.

26. Bokal, A.J., Ross, W., Gaal, T., Johnson, R.C. & Gourse, R.L. (1997) Molecular anatomy of a transcription activation patch: Fis-RNA polymerase interactions at the *Escherichia coli rrnB* P1 promoter. *EMBO J.* 16(1): 154-162.

27. Gourse, R.L., Gaal, T., Bartlett M.S., Appleman, J.A. & Ross, W. (1996) rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Ann. Rev. Genet.* 50: 645-677.

28. Boutte, C.C & Crosson, S. (2011) The complex logic of stringent response regulation in *Caulobacter crescentus*: starvation signaling in an oligotrophic environment. *Mol. Microbiol.* 80(3): 695-714.

29. Perederina, A., Svetlov, V., Vassylyeva, M.N., Artsimovich, I., Yokoyama, S. & Vassylyev, D. (2004) Regulation through the secondary channel – structural framework for ppGpp-DksA synergism during transcription. *Cell.* 118(3): 297-309.

30. Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G. & Laub, M.T. (2005) Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol.* 3(10): e334.

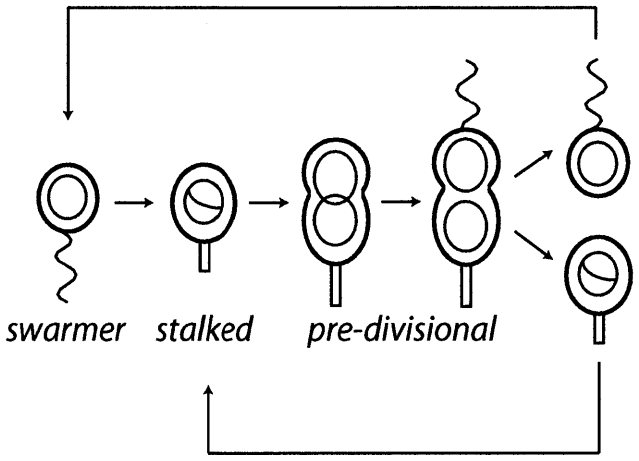

Figure 1

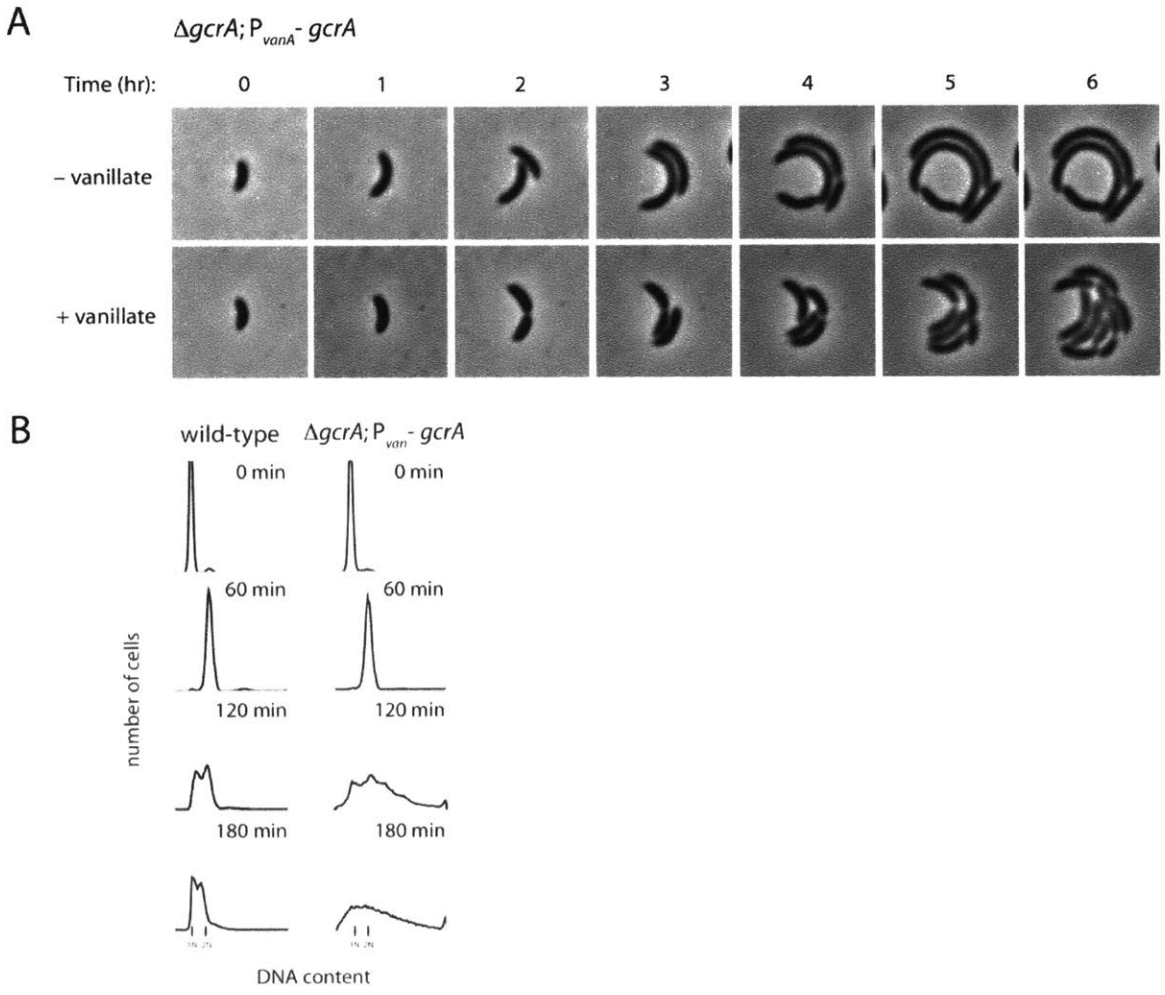
A
ΔgcrA; P vanA - gcrA
Time (hr):
0
1
2
3
4
5
6
– vanillate
+ vanillate
B
wild-type
ΔgcrA; P van - gcrA
0 min
0 min
60 min
60 min
120 min
120 min
180 min
180 min
number of cells
DNA content

Figure 2

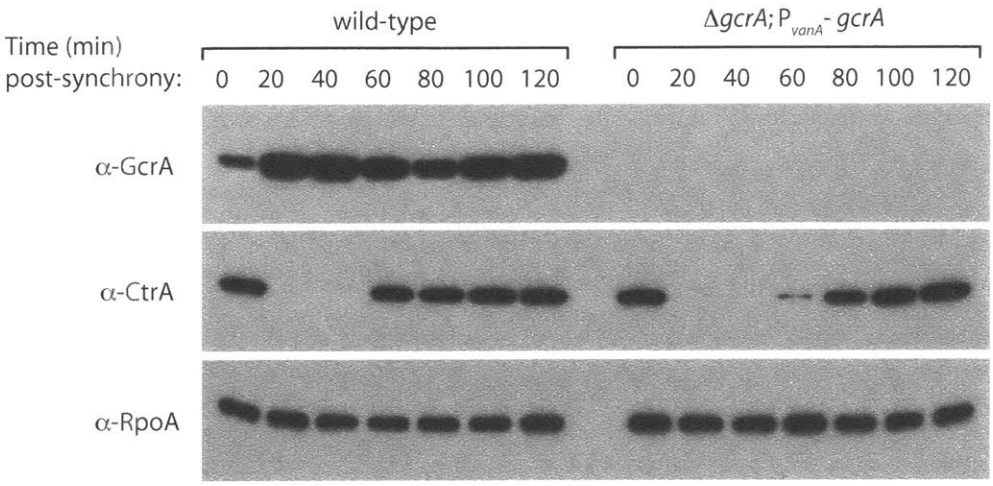
wild-type
ΔgcrA; P vanA - gcrA
Time (min)
post-synchrony:
0 20 40 60 80 100 120
0 20 40 60 80 100 120
α-GcrA
α-CtrA
α-RpoA

Figure 3

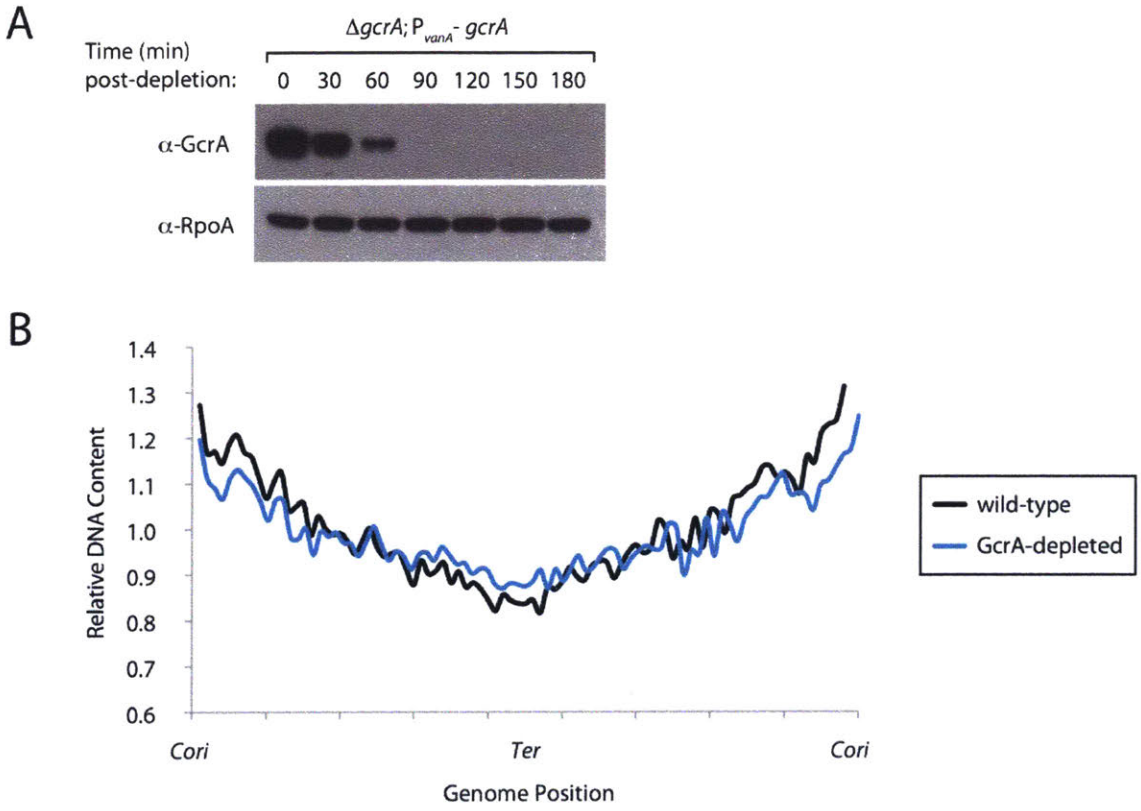
A
ΔgcrA; P vanA - gcrA
Time (min)
post-depletion:
0 30 60 90 120 150 180
α-GcrA
α-RpoA
B
Relative DNA Content
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Cori
Ter
Cori
Genome Position
wild-type
GcrA-depleted

Figure 4

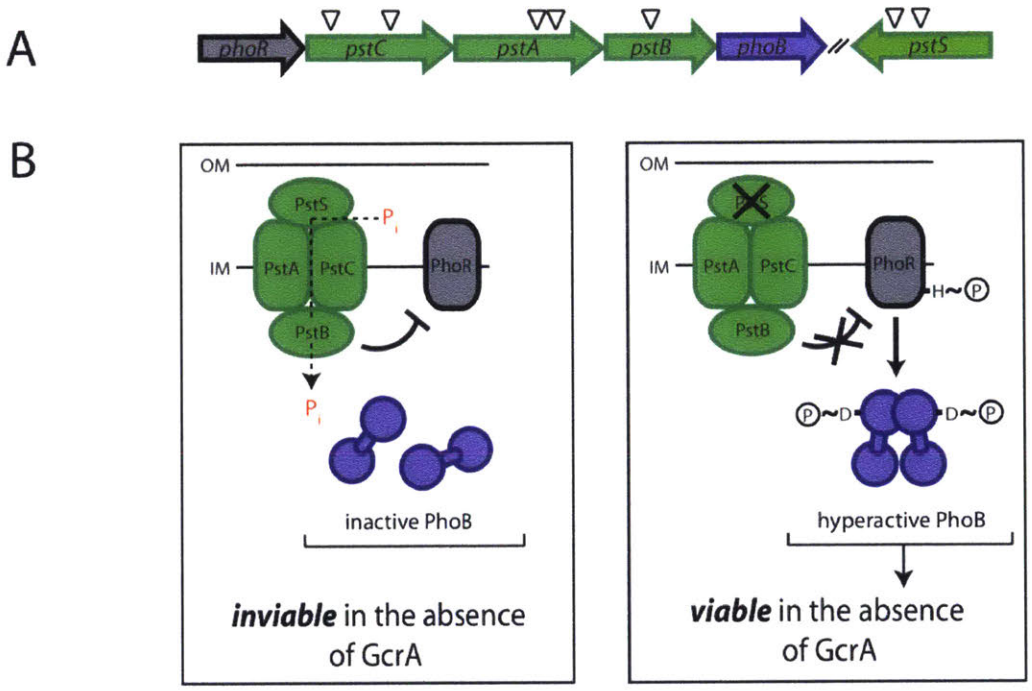
A
phoR
pstC
pstA
pstB
phoB
pstS
B
OM
PstS
Pi
IM
PstA
PstC
PhoR
PstB
Pi
inactive PhoB
inviable in the absence of GcrA
OM
IM
PstA
PstC
PhoR
H~P
PstB
P~D
D~P
hyperactive PhoB
viable in the absence of GcrA

Figure 5

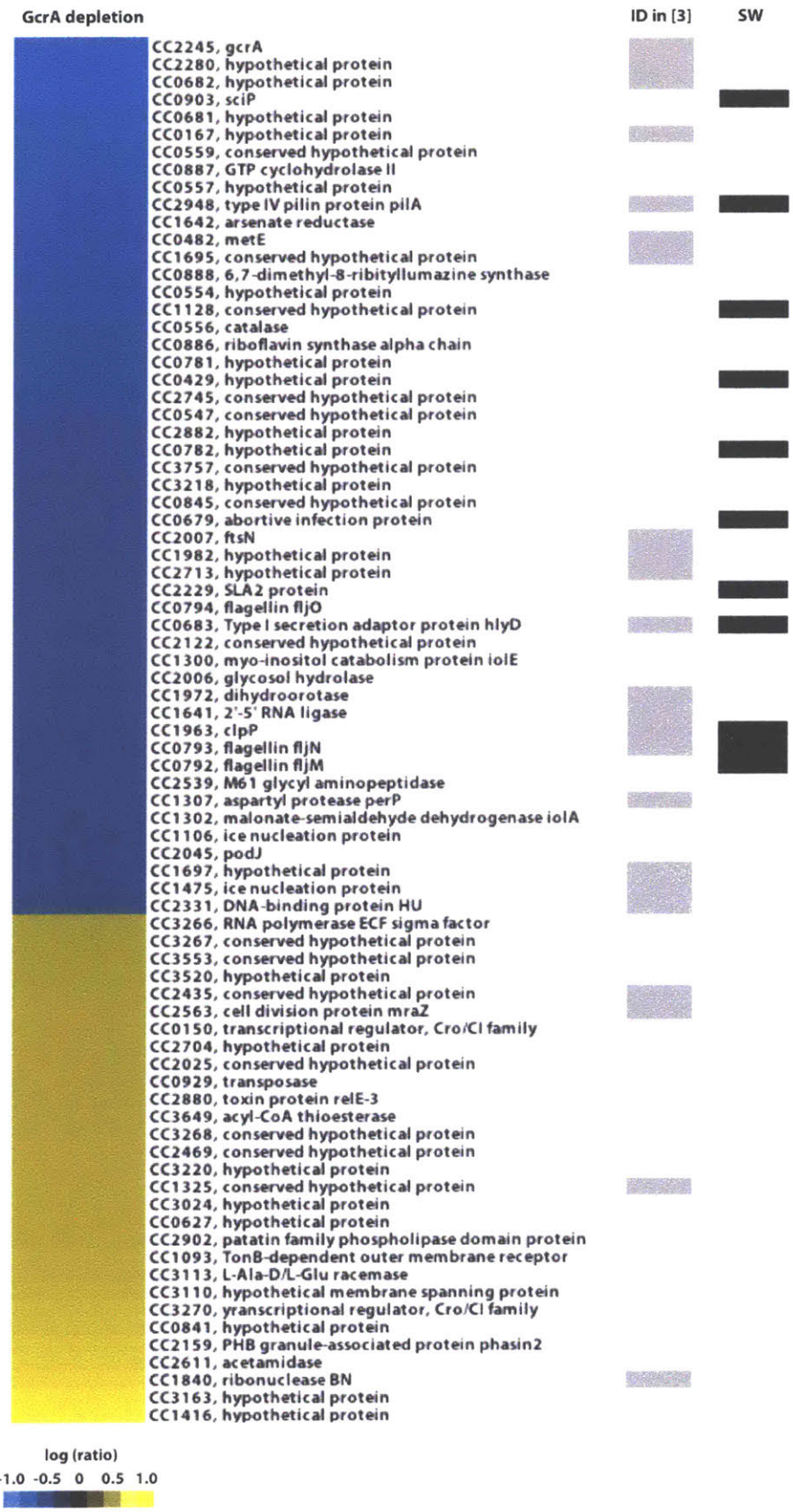
GcrA depletion
ID in [3]
SW
CC2245, gcrA
CC2280, hypothetical protein
CC0682, hypothetical protein
CC0903, sciP
CC0681, hypothetical protein
CC0167, hypothetical protein
CC0559, conserved hypothetical protein
CC0887, GTP cyclohydrolase II
CC0557, hypothetical protein
CC2948, type IV pilin protein pilA
CC1642, arsenate reductase
CC0482, metE
CC1695, conserved hypothetical protein
CC0888, 6,7-dimethyl-8-ribityllumazine synthase
CC0554, hypothetical protein
CC1128, conserved hypothetical protein
CC0556, catalase
CC0886, riboflavin synthase alpha chain
CC0781, hypothetical protein
CC0429, hypothetical protein
CC2745, conserved hypothetical protein
CC0547, conserved hypothetical protein
CC2882, hypothetical protein
CC0782, hypothetical protein
CC3757, conserved hypothetical protein
CC3218, hypothetical protein
CC0845, conserved hypothetical protein
CC0679, abortive infection protein
CC2007, ftsN
CC1982, hypothetical protein
CC2713, hypothetical protein
CC2229, SLA2 protein
CC0794, flagellin fljO
CC0683, Type I secretion adaptor protein hlyD
CC2122, conserved hypothetical protein
CC1300, myo-inositol catabolism protein iolE
CC2006, glycosol hydrolase
CC1972, dihydroorotase
CC1641, 2'-5' RNA ligase
CC1963, clpP
CC0793, flagellin fljN
CC0792, flagellin fljM
CC2539, M61 glycyl aminopeptidase
CC1307, aspartyl protease perP
CC1302, malonate-semialdehyde dehydrogenase iolA
CC1106, ice nucleation protein
CC2045, podJ
CC1697, hypothetical protein
CC1475, ice nucleation protein
CC2331, DNA-binding protein HU
CC3266, RNA polymerase ECF sigma factor
CC3267, conserved hypothetical protein
CC3553, conserved hypothetical protein
CC3520, hypothetical protein
CC2435, conserved hypothetical protein
CC2563, cell division protein mraZ
CC0150, transcriptional regulator, Cro/CI family
CC2704, hypothetical protein
CC2025, conserved hypothetical protein
CC0929, transposase
CC2880, toxin protein relE-3
CC3649, acyl-CoA thioesterase
CC3268, conserved hypothetical protein
CC2469, conserved hypothetical protein
CC3220, hypothetical protein
CC1325, conserved hypothetical protein
CC3024, hypothetical protein
CC0627, hypothetical protein
CC2902, patatin family phospholipase domain protein
CC1093, TonB-dependent outer membrane receptor
CC3113, L-Ala-D/L-Glu racemase
CC3110, hypothetical membrane spanning protein
CC3270, yranscriptional regulator, Cro/CI family
CC0841, hypothetical protein
CC2159, PHB granule-associated protein phasin2
CC2611, acetamidase
CC1840, ribonuclease BN
CC3163, hypothetical protein
CC1416, hypothetical protein
log (ratio)
-1.0 -0.5 0 0.5 1.0

Figure 6

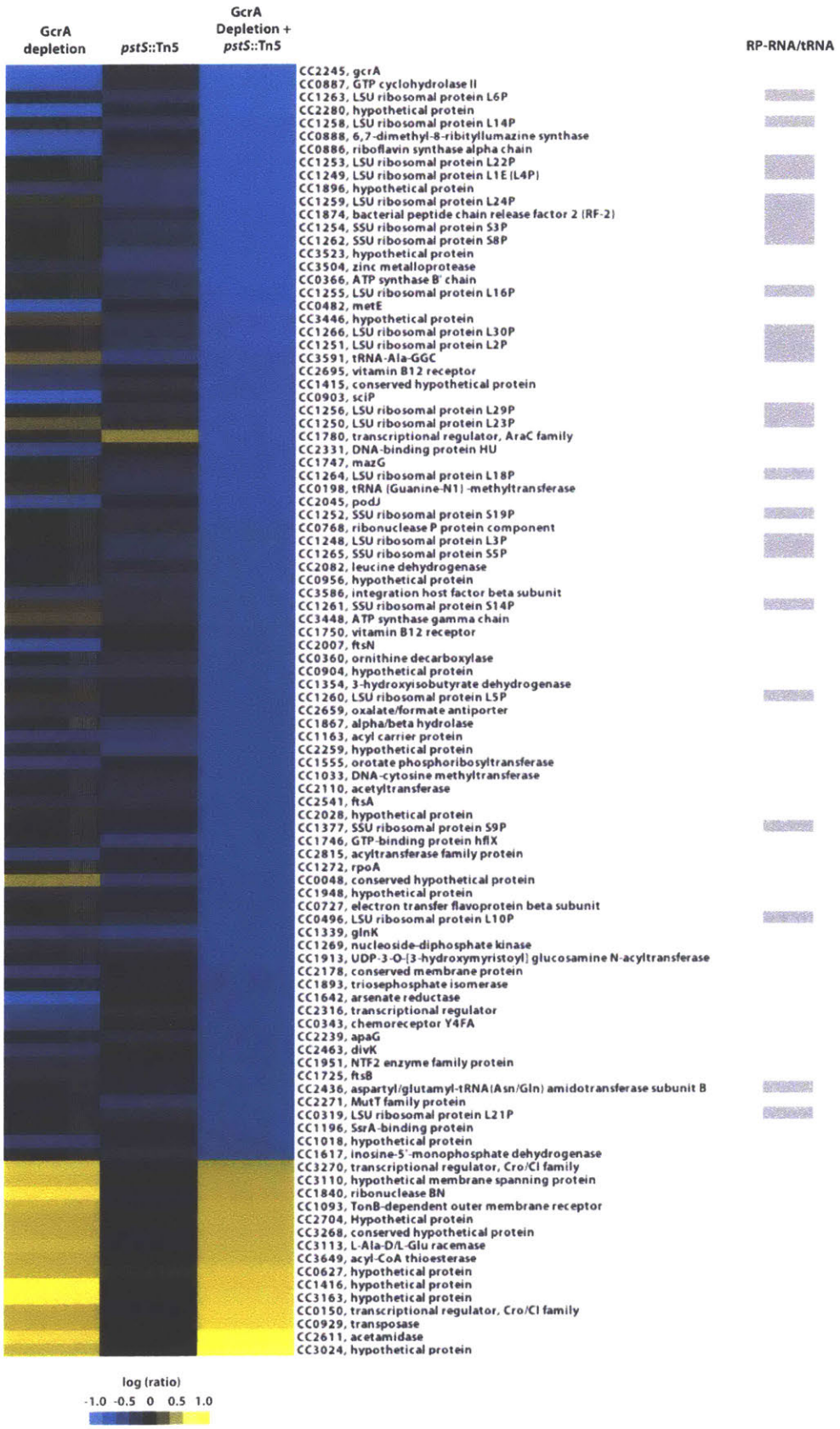
GcrA depletion
pstS::Tn5
GcrA Depletion + pstS::Tn5
RP-RNA/tRNA
CC2245, gcrA
CC0887, GTP cyclohydrolase II
CC1263, LSU ribosomal protein L6P
CC2280, hypothetical protein
CC1258, LSU ribosomal protein L14P
CC0888, 6,7-dimethyl-8-ribityllumazine synthase
CC0886, riboflavin synthase alpha chain
CC1253, LSU ribosomal protein L22P
CC1249, LSU ribosomal protein L1E (L4P)
CC1896, hypothetical protein
CC1259, LSU ribosomal protein L24P
CC1874, bacterial peptide chain release factor 2 (RF-2)
CC1254, SSU ribosomal protein S3P
CC1262, SSU ribosomal protein S8P
CC3523, hypothetical protein
CC3504, zinc metalloprotease
CC0366, ATP synthase B' chain
CC1255, LSU ribosomal protein L16P
CC0482, metE
CC3446, hypothetical protein
CC1266, LSU ribosomal protein L30P
CC1251, LSU ribosomal protein L2P
CC3591, tRNA-Ala-GGC
CC2695, vitamin B12 receptor
CC1415, conserved hypothetical protein
CC0903, sciP
CC1256, LSU ribosomal protein L29P
CC1250, LSU ribosomal protein L23P
CC1780, transcriptional regulator, AraC family
CC2331, DNA-binding protein HU
CC1747, mazG
CC1264, LSU ribosomal protein L18P
CC0198, tRNA (Guanine-N1) -methyltransferase
CC2045, podJ
CC1252, SSU ribosomal protein S19P
CC0768, ribonuclease P protein component
CC1248, LSU ribosomal protein L3P
CC1265, SSU ribosomal protein S5P
CC2082, leucine dehydrogenase
CC0956, hypothetical protein
CC3586, integration host factor beta subunit
CC1261, SSU ribosomal protein S14P
CC3448, ATP synthase gamma chain
CC1750, vitamin B12 receptor
CC2007, ftsN
CC0360, ornithine decarboxylase
CC0904, hypothetical protein
CC1354, 3-hydroxyisobutyrate dehydrogenase
CC1260, LSU ribosomal protein L5P
CC2659, oxalate/formate antiporter
CC1867, alpha/beta hydrolase
CC1163, acyl carrier protein
CC2259, hypothetical protein
CC1555, orotate phosphoribosyltransferase
CC1033, DNA-cytosine methyltransferase
CC2110, acetyltransferase
CC2541, ftsA
CC2028, hypothetical protein
CC1377, SSU ribosomal protein S9P
CC1746, GTP-binding protein hflX
CC2815, acyltransferase family protein
CC1272, rpoA
CC0048, conserved hypothetical protein
CC1948, hypothetical protein
CC0727, electron transfer flavoprotein beta subunit
CC0496, LSU ribosomal protein L10P
CC1339, glnK
CC1269, nucleoside-diphosphate kinase
CC1913, UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase
CC2178, conserved membrane protein
CC1893, triosephosphate isomerase
CC1642, arsenate reductase
CC2316, transcriptional regulator
CC0343, chemoreceptor Y4FA
CC2239, apaG
CC2463, divK
CC1951, NTF2 enzyme family protein
CC1725, ftsB
CC2436, aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B
CC2271, MutT family protein
CC0319, LSU ribosomal protein L21P
CC1196, SsrA-binding protein
CC1018, hypothetical protein
CC1617, inosine-5'-monophosphate dehydrogenase
CC3270, transcriptional regulator, Cro/CI family
CC3110, hypothetical membrane spanning protein
CC1840, ribonuclease BN
CC1093, TonB-dependent outer membrane receptor
CC2704, Hypothetical protein
CC3268, conserved hypothetical protein
CC3113, L-Ala-D/L-Glu racemase
CC3649, acyl-CoA thioesterase
CC0627, hypothetical protein
CC1416, hypothetical protein
CC3163, hypothetical protein
CC0150, transcriptional regulator, Cro/CI family
CC0929, transposase
CC2611, acetamidase
CC3024, hypothetical protein
log (ratio)
-1.0 -0.5 0 0.5 1.0

Figure 7

A

CCNA #	CC #	Gene Annotation	RP-RNA/tRNA
CCNA R0021		tRNA-Pro-CGG	
CCNA R0022		tRNA-Gln-TTG	
CCNA R0090	CC3591	tRNA-Ala-GGC	
CCNA 02783	CC2700	two-component response regulator	
CCNA R0091		tRNA-Leu-CAG	
CCNA R0035		tRNA-Asn-GTT	
CCNA 02436	CC2351	hypothetical protein	
CCNA 02782	CC2699	hypothetical protein	
CCNA R0061		tRNA-Thr-GGT	
CCNA 01297	CC1239	23S rRNA Gm2251 methyltransferase	
CCNA 03718	CC3604	NADH-ubiquinone oxidoreductase subunit	
CCNA 00321	CC0319	LSU ribosomal protein L21P	
CCNA R0046		tRNA-Asp-GTC	
CCNA R0031		tRNA-Tyr-GTA	
CCNA 02033	CC1956	nuoA	
CCNA R0023		possible IS element	
CCNA 03430	CC3321	LSU ribosomal protein L36P	
CCNA R0027		tRNA-Gly-GCC	
CCNA 00322	CC0320	conserved hypothetical protein	
CCNA 01108	CC1055	nucleoside-diphosphate-sugar epimerases	
CCNA 03062	CC2967	cell wall hydrolase family protein	
CCNA 03717	CC3603	ribonuclease D	
CCNA R0034		tRNA-Asn-GTT	
CCNA 02623	CC2540	ftsZ	
CCNA 03638		hypothetical protein	
CCNA 01791	CC1720	CTP synthase	
CCNA 03312		hypothetical protein	
CCNA 03431	CC3322	membrane-bound lytic murein transglycosylase B	
CCNA 03429	CC3320	HAD family hydrolase	
CCNA 02125	CC2045	podJ	
CCNA R0077		tRNA-Trp-CCA	
CCNA R0008		tRNA-Arg-ACG	
CCNA 03609		cell surface antigen sca2	
CCNA 02543	CC2459	LSU ribosomal protein L33P	
CCNA 00007	CC0007	SSU ribosomal protein S20P	
CCNA 00976	CC0927	sppA	
CCNA 03689	CC3574	alanine dehydrogenase	
CCNA 00161	CC0162	hypothetical protein	
CCNA 02126	CC2046	conserved hypothetical protein	
CCNA 03308	CC3203	hypothetical protein	
CCNA 03813	CC3699	creS	
CCNA 03405	CC3296	hypothetical protein	
CCNA R0014		stationary phase expressed sRNA	
CCNA 03091	CC2996	hypothetical protein	
CCNA 03406	CC3297	SSU ribosomal protein S21P	
CCNA 01171	CC1114	shpA	
CCNA R0001		tRNA-Thr-CGT	
CCNA 03736	CC3621	hypothetical protein	
CCNA 00808	CC0769	LSU ribosomal protein L34P	

B

Motif 1:

Motif 2:

Figure 8

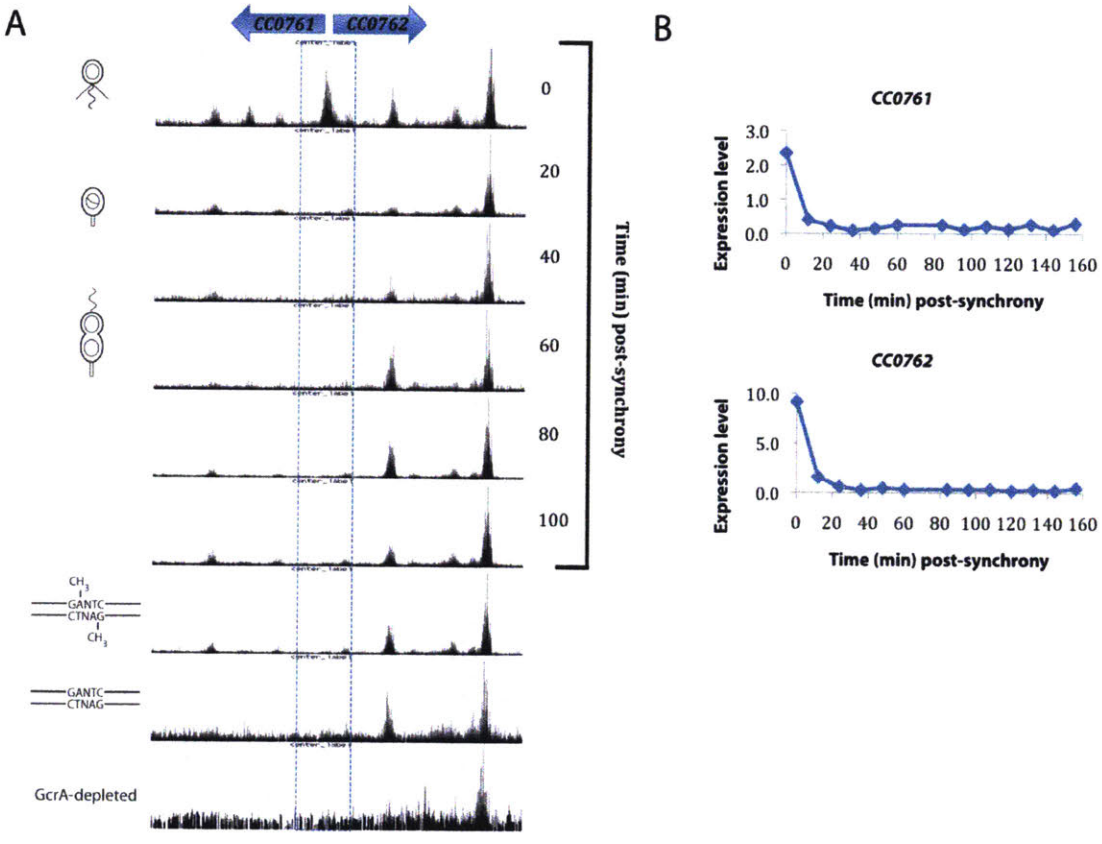
A
CC0761
CC0762
0
20
40
60
80
100
Time (min) post-synchrony
GANTC
CTNAG
GANTC
CTNAG
GcrA-depleted
B
CC0761
Expression level
3.0
2.0
1.0
0.0
0 20 40 60 80 100 120 140 160
Time (min) post-synchrony
CC0762
Expression level
10.0
5.0
0.0
0 20 40 60 80 100 120 140 160
Time (min) post-synchrony

Figure 9

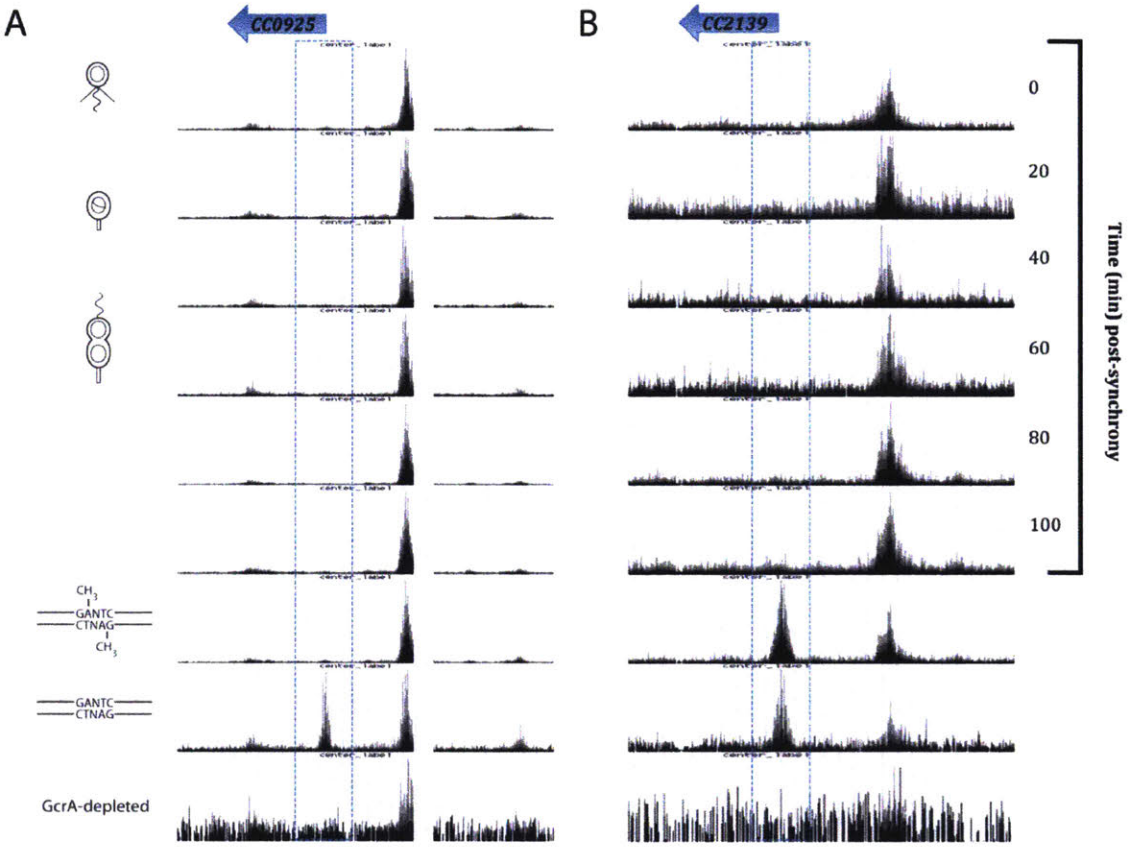
A
CC0925
B
CC2139
0
20
40
60
80
100
Time (min) post-synchrony
CH3
GANTC
CTNAG
CH3
GANTC
CTNAG
GcrA-depleted

Figure 10