Functional characterization of bacterial sRNAs using a network biology approach

Sheetal R. Modi[a,1], Diogo M. Camacho[a,1], Michael A. Kohanski[a,b], Graham C. Walker[c], and James J. Collins[a,b,d,2]

[a]The Howard Hughes Medical Institute, Department of Biomedical Engineering, and Center for BioDynamics, Boston University, Boston, MA 02215; [b]Boston University School of Medicine, Boston, MA 02118; [c]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and [d]The Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115

Edited* by Susan Gottesman, National Cancer Institute, Bethesda, MD, and approved August 3, 2011 (received for review March 19, 2011)

Small RNAs (sRNAs) are important components of posttranscriptional regulation. These molecules are prevalent in bacterial and eukaryotic organisms, and involved in a variety of responses to environmental stresses. The functional characterization of sRNAs is challenging and requires highly focused and extensive experimental procedures. Here, using a network biology approach and a compendium of gene expression profiles, we predict functional roles and regulatory interactions for sRNAs in *Escherichia coli*. We experimentally validate predictions for three sRNAs in our inferred network: IsrA, GlmZ, and GcvB. Specifically, we validate a predicted role for IsrA and GlmZ in the SOS response, and we expand on current knowledge of the GcvB sRNA, demonstrating its broad role in the regulation of amino acid metabolism and transport. We also show, using the inferred network coupled with experiments, that GcvB and Lrp, a transcription factor, repress each other in a mutually inhibitory network. This work shows that a network-based approach can be used to identify the cellular function of sRNAs and characterize the relationship between sRNAs and transcription factors.

microbiology | network inference | sRNA regulation

Small noncoding RNAs (sRNAs) are ubiquitous in all kingdoms of life. These molecules range in length from a few nucleotides to a few hundred nucleotides, and are involved in the regulation of a wide range of physiological processes (1–3). Bacterial sRNAs, some of which have been studied extensively (4), have been implicated in the regulation of bacterial stress responses, iron uptake, quorum sensing, virulence, and biofilm formation (5–8).

The most widely studied class of bacterial sRNAs act as posttranscriptional regulators by base-pairing to target mRNAs. This interaction is facilitated by the Hfq protein, a bacterial Sm-like protein with chaperone function, which acts as a general cofactor in RNA interactions (9). Binding of an sRNA to its mRNA target can result in changes in translational efficiency as well as transcript instability, a process dependent on the RNase-E–including degradosome complex (10). To date, ∼80 sRNAs have been identified in *Escherichia coli*, 30 of which are Hfq-dependent. A number of these sRNAs have been well characterized, such as RyhB, which is known to regulate iron homeostasis (8).

Small RNAs and their targets are being discovered and identified with greater efficiency (11–15); however, the functions of many sRNAs remain unknown. In the present study, we were interested in exploring the possibility of developing and using a network biology approach to elucidate sRNA functional roles. We applied a network inference algorithm to a compendium of *E. coli* microarray expression profiles to reconstruct an sRNA regulatory network. Functional enrichment of the resulting sRNA subnetworks confirmed known functions for some sRNAs and identified putative functions for others. We experimentally validated predicted functional roles for three sRNAs, and an in-depth analysis of the inferred network led to the discovery of a unique sRNA transcription-factor mutual inhibitory network.

Results and Discussion

Small RNA Regulatory Network Inference. We developed a computational biology approach to characterize functional roles for sRNAs in bacteria (see Fig. S1 and *SI Materials and Methods* for a more complete overview of this method). As a first step, we used the Context Likelihood of Relatedness (CLR) algorithm (16) to infer the sRNA regulatory network in *E. coli*. The CLR algorithm is an inference approach based on mutual information and allows for the identification of regulatory relationships between biomolecular entities. This algorithm previously has been used to infer transcriptional regulatory networks (17) by examining the functional relationships between transcription factors and target genes. We applied the CLR algorithm to an existing compendium of *E. coli* microarrays collected under different experimental conditions (Table S1*A*) to reverse engineer and analyze the regulatory subnetworks for Hfq-dependent sRNAs. This process allowed us to infer potential targets of each of these sRNAs with a highly significant false-discovery rate (FDR)-corrected P value ($q < 0.005$) (18). The inferred network (Fig. 1 and Table S2*A*) consists of 459 putative direct and indirect targets for the Hfq-dependent sRNAs, including sRNA–sRNA interactions as well as a number of genes predicted to be coregulated by two sRNAs.

A cellular regulatory scheme in which each transcription factor regulates at least one sRNA (4) has been hypothesized. It is therefore interesting to note that 10 of the sRNAs in the network are predicted to interact with at least one transcription factor, although directionality of regulation is not implied. Transcriptional regulators in the network include LexA (SOS response), FlhD (chemotaxis), and GadE, GadW, and GadX (acid stress response). These network results indicate that regulation of the associated cellular processes may involve a complex interplay between sRNAs, transcription factors, and their respective targets.

We subsequently performed pathway enrichment for each of the inferred sRNA subnetworks, either by Gene Ontology (GO) term enrichment analysis or by using gene function information obtained from EcoCyc (19). These analyses allowed us to classify subnetworks according to function, and thereby implicate the sRNAs as regulators of specific cellular processes (Fig. 1). We were able to identify functional enrichment for seven of the inferred sRNA subnetworks: iron homeostasis (under *ryhB* regulation), amino acid metabolism (*gcvB*), motility and chemotaxis (*micF*), pH adaptation (*gadY*), DNA repair (*glmZ* and *isrA*), protection and adaptation to stress (*cyaR*), and extracellular transport (*dicF*). The involvement of RyhB in iron homeostasis and GadY in the regulation of acid response has been reported previously (7, 20). Our network analyses correctly characterized

Author contributions: S.R.M., D.M.C., M.A.K., G.C.W., and J.J.C. designed research; S.R.M., D.M.C., and M.A.K. performed research; S.R.M. and D.M.C. analyzed data; and S.R.M., D.M.C., M.A.K., G.C.W., and J.J.C. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

Data deposition: The microarray data reported in this paper have been deposited at www.bu.edu/abl.

[1]S.R.M. and D.M.C. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: jcollins@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1104318108/-/DCSupplemental.

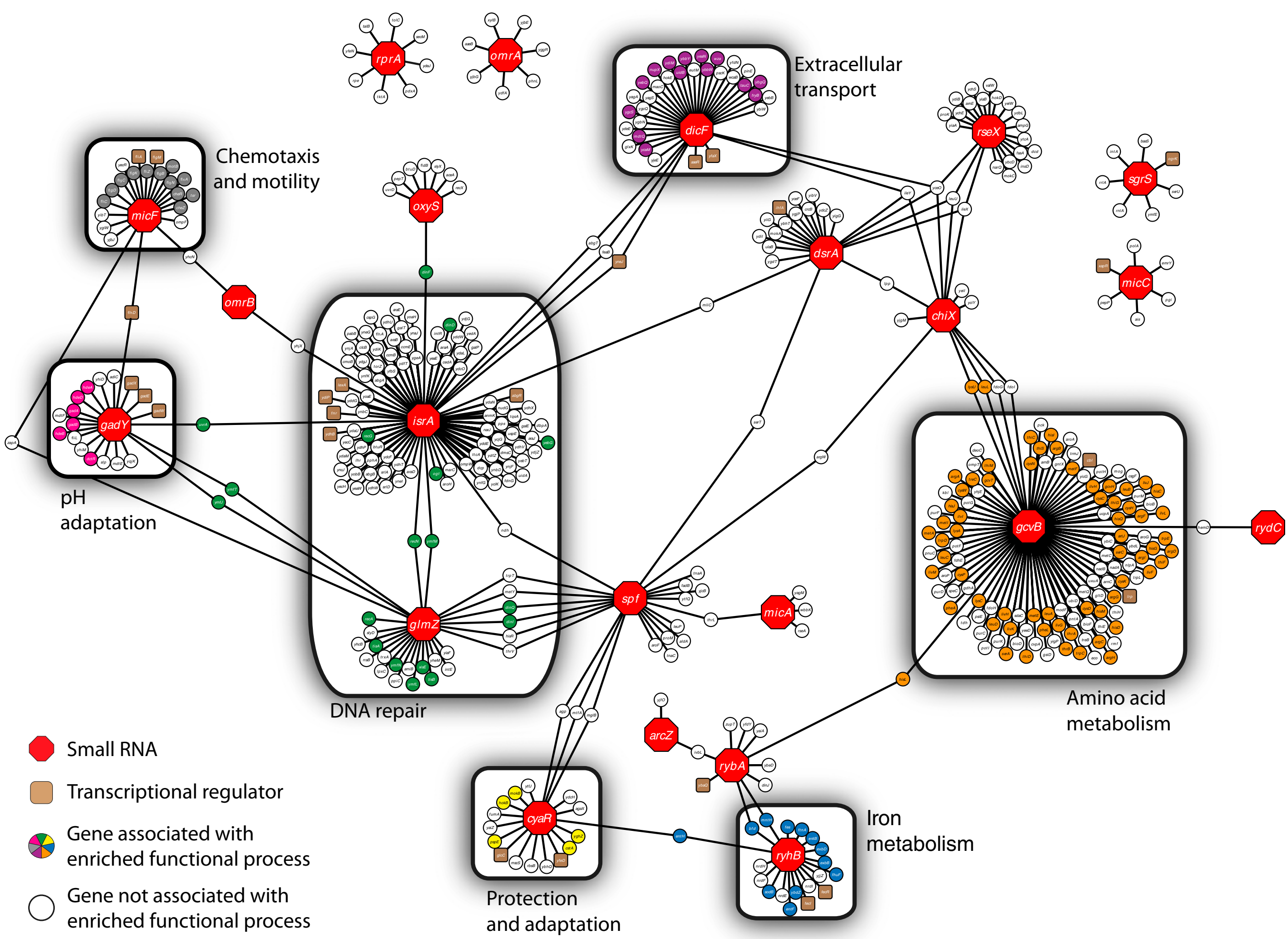

Fig. 1. Small RNA regulatory network in *E. coli*. We inferred highly significant regulatory interactions for Hfq-dependent sRNAs (shown in red) by applying the CLR algorithm to a compendium of *E. coli* microarrays. Genes in the seven sRNA subnetworks are colored if they are involved in the functional process assigned to the subnetwork by enrichment (i.e., green for DNA repair genes, yellow for protection and adaptation genes, blue for iron metabolism genes, orange for amino acid metabolism genes, purple for extracellular transport genes, gray for chemotaxis and motility genes, and pink for pH adaptation genes). Genes in white have functions that are not associated with an enriched process. Genes in brown encode transcriptional regulators.

these functions and additionally, suggested important roles for sRNAs in other cellular responses. Network topology also shows connectivity between functional processes. Although direct connections between functional processes may be tenuous, this predicted architecture shows that expression of intermediary genes varies significantly with multiple sRNAs, alluding to themes of overlapping sRNA regulation to coordinate global behavior. The functional annotations in our network, made possible by the identification of a large number of putative targets for Hfq-dependent sRNAs in *E. coli*, provide a basis for further exploration of the functional roles of sRNAs.

The compendium of microarray expression profiles used to reconstruct our regulatory network encompasses broad perturbations, such as different growth conditions and stress inductions (Table S1*A*). Including chips with sRNA-related genetic perturbations (e.g., *hfq* mutants) and additional environmental perturbations that increase the expression landscape of the cell would improve the algorithm's performance (Table S1*B* and Fig. S2). Furthermore, because our approach relies on RNA expression data, our approach is limited to predicting regulation that affects transcript levels. This finding could explain the absence of functional predictions for the *oxyS* and *spf* subnetworks (Fig. 1), as these sRNAs are known to regulate translation of their targets (21–23). Incorporating data at the translational level, such as from 2D gels and mass spectrometry profiling, would improve the predictive power of our approach in the discovery of regulatory roles for sRNAs.

IsrA and GlmZ Are Involved in the DNA Damage Response. To assess the validity of our network approach, we chose to explore the predicted involvement of the GlmZ and IsrA sRNAs in the cellular response to DNA damage. GlmZ is known to activate GlmS, a protein involved in the biosynthesis of amino sugars (constituents of the cell wall) to regulate expression based on the availability of external sugars (24). In contrast, no information has been published on IsrA (IS061) since its discovery in a bioinformatics-based screen (25). Our network results show that ~15% of the putative targets for these two sRNAs are involved in the DNA damage response, with 53% of these genes being under the regulation of the LexA repressor protein (Fig. 2*A* and Table S2 *B* and *C*).

To investigate the predicted role of these sRNAs, we treated *E. coli* cultures with DNA-damaging agents, specifically, the gyrase inhibitor norfloxacin, mitomycin C (MMC), and UV radiation, and observed their morphology. It is known that the SOS response induces filamentous growth, which is considered to be indicative of the state of DNA damage (26). Although the single-gene deletion strains, Δ*isrA* and Δ*glmZ*, did not exhibit a mor-

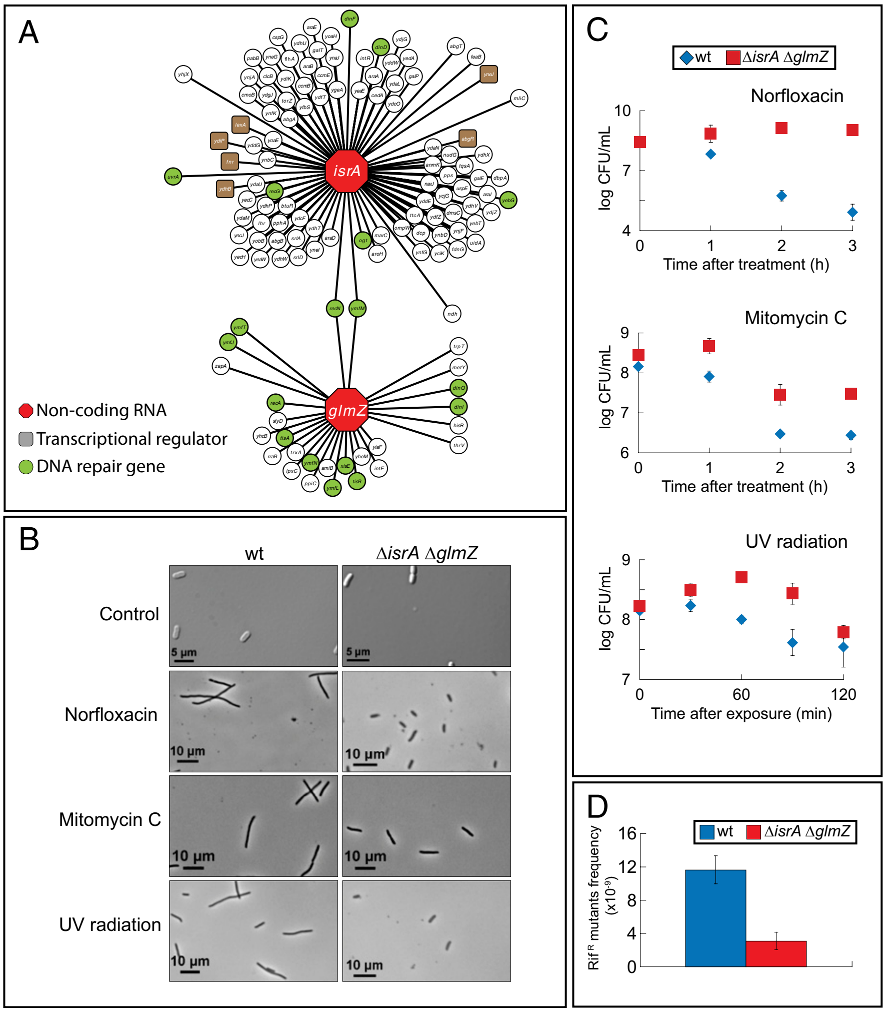

Fig. 2. A functional role for IsrA and GlmZ in the DNA damage response. (*A*) Inferred network connections for IsrA and GlmZ. Of the identified interactions, 18 are involved in DNA damage pathways. Approximately 50% of these DNA repair genes are members of the LexA regulon. (*B*) Representative micrographs of MG1655 (*Left*) and Δ*isrA*Δ*glmZ* MG1655 (*Right*) before (100× objective) and during DNA damage treatment (40× objective). Images show cells during norfloxacin treatment (125 ng/mL, T = 3 h), MMC treatment (2 μg/mL, T = 2 h), and repeated UV exposure (100 J/m^2, T = 1.5 h). See *Materials and Methods* for treatment details and Fig. S3 for full micrograph images. (*C*) Log change in colony-forming units per milliliter (CFU/mL) during DNA damage exposure. Survival of MG1655 (blue diamonds) and Δ*isrA*Δ*glmZ* MG1655 (red squares) following exposure to norfloxacin (125 ng/mL), MMC (2 μg/mL), and repeated UV exposure (100 J/m^2). In this and all other figures, error bars represent ± SE. (*D*) Basal mutation rate (mutations per cell per generation) for MG1655 (blue) and Δ*isrA*Δ*glmZ* MG1655 (red) using a rifampicin-based selection method. Wild-type mutation rate is similar to that previously reported (28).

phological phenotype that was different from wild-type (Fig. S3), the double-deletion strain, Δ*isrA*Δ*glmZ,* did show substantially less filamentation (Fig. 2*B* and Fig. S3).

We next sought to determine the effects of IsrA and GlmZ on cell survival under DNA damage. We measured cell viability following treatment with norfloxacin, MMC, and UV, and found that the double-deletion strain was significantly less sensitive than wild-type to each of the treatments (Fig. 2*C*).

Because cells experience low levels of DNA damage under normal physiological conditions (27), we were also interested in determining if the basal mutation rate of the Δ*isrA*Δ*glmZ* strain differed from wild-type in unperturbed conditions. We found that the mutation rate of the sRNA double mutant is approximately threefold less than that of wild-type (Fig. 2*D*).

Together, these results suggest that IsrA and GlmZ function as DNA repair regulators, possibly in a redundant manner. We speculate that these two sRNAs act by differentially affecting the regulation of specific genes within the LexA regulon. The SOS system is an important stress response that has been shown to play a central role in a variety of mechanisms, including antibiotic tolerance (17) and antibiotic-resistant gene transfer (29). Our work indicates that the sRNAs IsrA and GlmZ may play critical regulatory roles in this important cellular stress response.

GcvB Is Involved in the Regulation of Amino Acid Availability. We next chose to examine the inferred subnetwork for the GcvB sRNA. GcvB has been shown to regulate peptide transport (30, 31) and acid stress in *E. coli* (32). Functional enrichment of the genes predicted to interact with GcvB in our inferred network (Fig. 3*A* and Table S2*D*) revealed that ∼50% of them are involved in amino acid transport and metabolic processes, suggesting a broader role for GcvB in nutrient availability.

To assess this predicted role for GcvB, we measured growth rates of strains cultured in minimal media supplemented with different amino acids. Growth experiments in varying nutrient conditions have been used to implicate a number of genes in metabolism and transport. In our study, we compared the doubling times of two mutants, a Δ*gcvB* strain and a strain constitutively expressing *gcvB*, to that of wild-type. We observed that the growth rate of the mutants did not differ from wild-type in unsupplemented conditions (Fig. S4). However, the growth rates in one or both of the *gcvB* strains were significantly different when supplemented with leucine, serine, phenylalanine, or threonine (Fig. 3*B*). These results support our network-derived hypothesis that GcvB plays a broad role in the regulation of amino acid availability and metabolism.

GcvB Represses Lrp. Among the genes predicted to interact with GcvB is *lrp*, which encodes the Lrp transcription factor, an important regulator of amino acid availability (33). We hypothesized that GcvB regulates amino acid pathways through modulation of Lrp.

Sequence information and corresponding secondary structure have been used to uncover sRNA targets (31). Accordingly, we used TargetRNA, a sequence-based algorithm with high-performance capabilities (34), to explore the possibility of a physical interaction between GcvB and Lrp. TargetRNA identifies 21 putative targets of GcvB in *E. coli* using default parameters (Fig. S5).

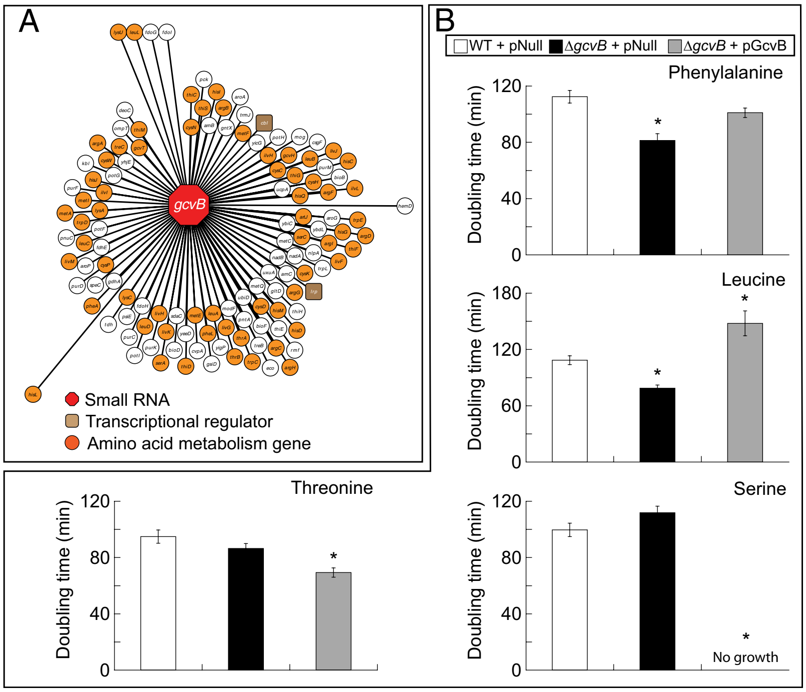

Fig. 3. A functional role for GcvB in amino acid availability. (*A*) Inferred network connections for GcvB. From our CLR results and GO term enrichment, GcvB shows a large number of interactions (51%) with genes involved in amino acid metabolism and transport, including the transcriptional regulator Lrp. Approximately 28% of these amino acid-related genes are members of the Lrp regulon. (*B*) Doubling times for MG1655 + pZA31-null (white bars), Δ*gcvB* MG1655 + pZA31-null (black bars), and Δ*gcvB* MG1655 + pZA31-*gcvB* (gray bars) in M9 minimal media, calculated using OD_{600} values. Leucine and phenylalanine were supplemented at 2 mM, and serine and threonine were supplemented at 1 mM. No growth was observed for Δ*gcvB* MG1655 + pZA31-*gcvB* in serine-supplemented media. Asterisks represent significant ($P < 0.05$) differences in growth rate compared with MG1655 + pZA31-null.

Although the algorithm does not predict *lrp* to be a target of GcvB, it does predict two targets in our network, *trpE* and *livK*. The GcvB-*trpE* and GcvB-*livK* binding regions predicted by TargetRNA overlap and together span positions 65 to 91 on the GcvB transcript. We searched for homologies of this region's complement within 100 bp of the *lrp* translational start and identified a putative binding site for GcvB in the *lrp* 5′UTR (Fig. S6*A*).

To examine this putative direct interaction between GcvB and Lrp, we used an *lrp* gene fusion to GFP to function as a reporter of translational control by GcvB. Our translational fusion consisted of the 5′UTR of Lrp and the first 15 amino acids fused to the N terminal of GFP, and was constructed using the modular plasmid system described by Urban and Vogel (11). This system was designed to confirm sRNA-mediated control of mRNA targets through its ability to uncouple both species from the chromosomal regulatory network and to reliably suppress pleiotropic effects of sRNA expression on target fusion transcription. Expressing GcvB, we observed an approximately twofold decrease in fluorescence of the *lrp*::*gfp* fusion compared with a control plasmid (Fig. 4*A*, *Left*). We used a *dppA*::*gfp* fusion as a positive control for our expression system (Fig. 4*A*, *Left*) and obtained results for this known GcvB target that were consistent with those previously reported (11). As a control to address potential indirect regulation of *lrp*::*gfp*, we experimentally demonstrated that the MicF sRNA, which is not predicted to interact with Lrp, had no effect on the Lrp fusion (Fig. S6*B*).

To obtain additional evidence of the interaction between GcvB and Lrp, we mutated the predicted binding region in the 5′UTR of *lrp*. Four base-pair mutations were made to our target fusion—specifically, A(-9)T, C(-8)G, A(-7)T, and A(-6)T—where base position is with respect to the *lrp* translational start. These mutations eliminated GcvB repression of the Lrp transcript (Fig. 4*A*, *Right*). Taken together, these results demonstrate the direct posttranscriptional repression of Lrp by GcvB and offer an sRNA-transcription-factor regulation scheme for the control of amino acid availability.

***gcvB* Is Regulated by Lrp.** Analysis of our microarray compendium revealed that expression of *gcvB* and *lrp* are anticorrelated, independent of growth phase, suggesting that these genes function in a complex regulatory circuit. Mutually regulating elements endow networks with interesting properties, such as bistability and memory (35, 36). There is precedence for this type of motif at the posttranscriptional level in eukaryotes (37), and many other network architectures have been demonstrated in bacterial sRNA regulation (38, 39). However, mutually inhibitory networks involving sRNAs have not been found in bacteria. We hypothesized that Lrp and GcvB function together in a mutually inhibitory network for controlled pathway regulation of cellular amino acid availability.

Building on our results that establish GcvB regulation of Lrp (Fig. 4*A*), we sought to explore Lrp regulation of *gcvB*. In elucidating the relationship of Lrp to *gcvB*, it was important to do so within the context of known regulation. GcvB is activated by the glycine cleavage system regulator, GcvA, under glycine-rich conditions (40). This interaction is dependent upon Lrp binding and is negatively regulated by GcvR when glycine is limiting (41). Using quantitative PCR, we measured relative expression levels of *gcvB* in wild-type, Δ*gcvA*, and Δ*lrp*, with and without glycine addition (Fig. 4*B*). We found that *gcvB* expression is significantly lower in Δ*gcvA* under glycine-rich conditions, confirming known regulation. Interestingly, we also found that *gcvB* transcript levels are ∼30-fold greater in Δ*lrp* compared with wild-type, independent of glycine addition. Because Lrp is a central regulator of cellular processes, it is possible that Lrp mediates negative regulation of *gcvB* indirectly. To investigate dependence on GcvA, we compared relative *gcvB* expression in Δ*gcvA*Δ*lrp* and Δ*gcvA* (Fig. S6*C*) and found that significantly higher levels of *gcvB* are present in Δ*gcvA*Δ*lrp* strain, demonstrating that Lrp regulation is not mediated exclusively through GcvA.

We next used sequence analysis to look for evidence that may suggest a direct interaction of Lrp on *gcvB*. We used ClustalW2 (42) to search for known Lrp-binding consensus sequences (43, 44) in the 500-bp region upstream of *gcvB*. Homology results indicate that there are two putative binding sites for Lrp in this region (Fig. S6*D*). These data suggest a direct interaction of Lrp on *gcvB*; however, an indirect regulation scheme remains possible and cannot be excluded.

Collectively, our analyses indicate that Lrp and GcvB repress each other (directly or indirectly) in a mutually inhibitory network (Fig. 4*C*). We speculate that *E. coli* can use this dual repression scheme to create a controlled response to changing

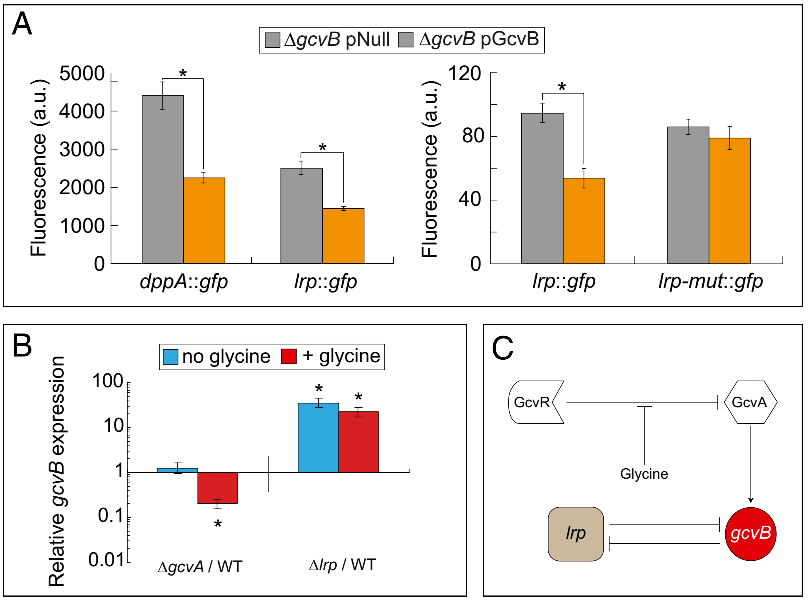

Fig. 4. Mutual inhibitory relationship of GcvB and Lrp. (*A*) GFP translational fusions for *dppA* and *lrp* (*Left*) and *lrp* and *lrp*-mut (*Right*). *dppA*::*gfp* (plasmid pSK-015) was grown in EZ rich media to model conditions in which it was originally tested (11). *lrp*::*gfp* and *lrp*-mut::*gfp* were grown in M9 minimal media to assess the effects of *gcvB* expression in nutrient-limiting conditions. Unregulated target fusion specific fluorescence (expressing control vector) is shown in gray, and regulated target fusion specific fluorescence (expressing *gcvB*) is shown in orange. See *Materials and Methods* and *SI Materials and Methods* for details on fluorescence measurements and calculations. Asterisks represent significant ($P < 0.05$) differences between unregulated and regulated target fusion specific fluorescence. (*B*) Fold-difference in *gcvB* expression in Δ*gcvA* and Δ*lrp* relative to wild-type during growth in M9 minimal media. Blue bars represent relative expression when exogenous glycine was absent, and red bars represent relative expression when glycine (300 μg/mL) was added to the media. Error bars represent propagated error measures. (*C*) GcvB-Lrp regulatory subnetwork resulting from translational fusion, expression data, and known regulatory interactions. Unsequestered GcvA activates *gcvB* expression when glycine is present. GcvB directly represses Lrp and Lrp directly or indirectly represses *gcvB*.

amino acid availability in the environment, allowing for robust adaptation and resource conservation.

Conclusions

In this work, we have shown how a network biology approach can be used to elucidate bacterial sRNA functional and regulatory roles. Although our network map does not discern between direct and indirect sRNA-gene interactions, our methodology can also be used to improve current methods of sRNA target identification. Direct interactions within our network can be uncovered by filtering network predictions with sequence-alignment tools or other target detection methods (14), as we illustrate for GcvB-mediated regulation of Lrp. Examining expression levels of predicted sRNA targets in relevantly perturbed *hfq* mutant strains may provide additional confirmation of direct interactions.

The approach described in this work, which relies on compendia of expression data, can be readily extended to other organisms and used to characterize sRNAs in pathogens as well as microRNAs in eukaryotes. Efforts along these lines could enhance our understanding of the posttranscriptional regulatory events that lead to pathogenicity or disease states, like cancer. Furthermore, as efforts to discover novel sRNAs lead to larger and more extensive RNA-seq expression datasets, our network biology approach could enrich the information found in these expression profiles to infer function of newly discovered sRNAs.

The present study also highlights how large-scale biomolecular networks can be used to guide the discovery and detailed experimental investigation of small-scale networks. In our case, a large-scale, reconstructed sRNA regulatory network enabled us to uncover an intriguing mutually inhibitory network made up of a small RNA and a transcription factor.

Understanding the regulatory roles of noncoding RNAs in prokaryotic and eukaryotic regulation presents an exciting challenge. Reverse genetics methodologies aspire to illuminate the subtle and complex interactions of RNA regulation, and the extent of their influence is slowly being uncovered. Our work shows that network biology approaches can make significant contributions to these efforts and facilitate the efficient reconstruction of functional and regulatory maps.

Materials and Methods

Network Inference. The sRNA network was inferred using the CLR algorithm. The algorithm uses mutual information to score the similarity between expression levels of two genes in a set of microarrays and applies an adaptive background correction step to eliminate false correlations and indirect influences (16). A gene pair is predicted to interact if their mutual information score is larger than an FDR-corrected score (18) at a given significance threshold ($q < 0.005$). The data used as input to the algorithm was an existing compendium of 759 Affymetrix *E. coli* Antisense2 microarray chips normalized as a group with RMA. The compendium includes arrays from the Many Microbe Microarray Database (E_coli_v3_Build_3), as well as 235 arrays run in-house, with experiments involving antibiotic treatment, biofilm growth, different growth media, acid shifts, anaerobic growth, as well as various perturbations of coding genes (Table S1*A*).

To gain insight into the functional roles of Hfq-dependent sRNAs, we performed pathway enrichment for each of the inferred sRNA subnetworks, either by GO term enrichment analysis (P value < 0.05, minimum GO term depth of 3) or by using gene function information obtained from EcoCyc (19). See *SI Materials and Methods* for more details and references on network analysis.

Media and Growth Conditions. Cultures were grown at 37 °C in Luria-Bertani broth (Fisher Scientific), M9 minimal media supplemented with 0.4% glucose (Fisher Scientific), or EZ Rich Defined media (Teknova). Antibiotics were added to the growth media for selection at the following concentrations: chloramphenicol (30 μg/mL; Acros Organics) and ampicillin (100 μg/mL; Fisher Scientific). Amino acids (Sigma), when supplemented, were used at the following concentrations: leucine and phenylalanine (2 mM), serine and threonine (1 mM), and glycine (300 μg/mL). Optical densities were taken using a SPECTRAFluor Plus plate spectrophotometer (Tecan).

Strains and Plasmids. All experiments were performed with *E. coli* MG1655 (ATCC 700926)-derived strains (Table S3*A*). Gene deletions were derived either from P1 phage transduction from the Keio collection (45) or by the PCR-based phage-λ red recombinase method (46). Expression vectors of *gcvB* were derived from the pZ system (47). Translational fusion-related plasmids, pXG-1, pXG-0, pXG-10, and pSK-015, were used for constructing and examining the target fusions mentioned above, and were kindly provided by Jörg Vogel, Institute for Molecular Infection Biology, University of Würzburg, Würzburg, Germany. Site-directed mutagenesis was performed using the Phusion Site-Directed Mutagenesis kit (New England Biolabs). See *SI Materials and Methods* for details on plasmid construction.

DNA Damage Sensitivity Assays. Cultures of the various strains were grown in 25 mL LB medium in 250 mL flasks to an OD_{600} of 0.3 (time 0), at which time they were exposed to three different DNA damaging agents (norfloxacin, MMC, and UV light). For norfloxacin-treated cultures, norfloxacin was added at a concentration of 125 ng/mL. For MMC-treated cultures, MMC was added at a concentration of 2 μg/mL. UV treatment was delivered using a Stratalinker UV box (Stratagene) such that cultures were exposed to 100 J/m^2 of UV radiation every 30 min over the course of 2 h. Strain viability was assessed by collecting aliquots of the cultures at time 0 (just before exposure to the

DNA damaging agent) and at 1-h intervals (norfloxacin and MMC) or 30-min intervals (UV) following exposure to the DNA damaging agent. Viability of the various strains at each time point was determined by measuring the CFU per milliliter, as described previously (48). Briefly, serially diluted cells were spot plated on LB-agar plates and grown overnight at 37 °C. Colonies were counted at those dilutions with $\sim$10 to 50 cells, and CFU per milliliter was calculated using the following formula: CFU/mL = [(# of colonies) $\times$ (dilution factor)]/0.01 mL. Average CFU per milliliter was determined based on the results of three biological replicates.

Translational Fusion Experiments and Calculations. At inoculation, cultures were induced with 1 mM IPTG (Invitrogen) for *gcvB* expression and grown to an OD_{600} of 0.3 in M9 minimal media (*lrp*::*gfp* and *lrp*-mut::*gfp*) or EZ rich media (pSK-015). Fluorescence measurements were taken, and relative fluorescence values were calculated as previously described (11) using the following formula: fold-change mediated by sRNA = (regulated target fusion specific fluorescence)/(unregulated target fusion specific fluorescence), where regulated target fusion specific fluorescence = [fluorescence(Δ*gcvB* + pZA12-gcvB + target fusion) − fluorescence(Δ*gcvB* + pZA12-*gcvB* + pXG-0)] and unregulated target fusion specific fluorescence = [fluorescence(Δ*gcvB* + pZA12-null + target fusion) − fluorescence(Δ*gcvB* + pZA12-null + pXG-0)]. The effect of *micF* expression on *lrp*::*gfp* was calculated similarly. See *SI Materials and Methods* for more details.

cDNA Synthesis and qPCR. Quantitative PCR was performed using the Roche LightCycler 480 and the LightCycler 480 SYBR Green I Master Kit (Roche Applied Science) according to the manufacturer's instructions. Relative quantification of *gcvB* was determined by the $\Delta\Delta$Cp method using *rrsH* (16S ribosomal RNA) as a reference gene. Fold-changes were calculated by comparing relative expression across the same conditions in the strains mentioned in the text. See *SI Materials and Methods* and Table S3*B* for more details.

Statistical Analysis, All data are representative of mean values of replicates, except in Fig. 2*D*, where the median was used to calculate the mutation rate. Error bars represent $\pm$ SE, which was propagated when necessary as described by others (49). Statistical significance was calculated between data sets using a two-tailed *t* test assuming unequal variance in the population.

ACKNOWLEDGMENTS. We thank Susan Gottesman and Gisela Storz for insightful discussions and comments on our work, as well as detailed information on Hfq-dependent small RNAs in *Escherichia coli*; and Jörg Vogel for the pXG-0, pXG-1, pXG-10, and pSK-015 plasmids. This work was supported by the National Institutes of Health Director's Pioneer Award Program Grant DP1 OD003644 and The Howard Hughes Medical Institute, and National Institutes of Health Grants CA21615 and GM31030 (to. G.C.W.). G.C.W. is an American Cancer Society Professor.

1. Barrangou R, et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712.
2. Chen CZ, Li L, Lodish HF, Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303:83–86.
3. Schembri F, et al. (2009) MicroRNAs as modulators of smoking-induced gene expression changes in human airway epithelium. *Proc Natl Acad Sci USA* 106: 2319–2324.
4. Waters LS, Storz G (2009) Regulatory RNAs in bacteria. *Cell* 136:615–628.
5. Gottesman S, et al. (2006) Small RNA regulators and the bacterial response to stress. *Cold Spring Harb Symp Quant Biol* 71:1–11.
6. Lenz DH, et al. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* 118:69–82.
7. Massé E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci USA* 99:4620–4625.
8. Massé E, Vanderpool CK, Gottesman S (2005) Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187:6962–6971.
9. Møller T, et al. (2002) Hfq: A bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* 9:23–30.
10. Gottesman S (2004) The small RNA regulators of *Escherichia coli*: Roles and mechanisms*. *Annu Rev Microbiol* 58:303–328.
11. Urban JH, Vogel J (2007) Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res* 35:1018–1037.
12. Hobbs EC, Astarita JL, Storz G (2010) Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: Analysis of a bar-coded mutant collection. *J Bacteriol* 192:59–67.
13. Mandin P, Gottesman S (2009) A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol Microbiol* 72:551–565.
14. Backofen R, Hess WR (2010) Computational prediction of sRNAs and their targets in bacteria. *RNA Biol* 7:33–42.
15. Huang JC, et al. (2007) Using expression profiling data to identify human microRNA targets. *Nat Methods* 4:1045–1049.
16. Faith JJ, et al. (2007) Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol* 5(1):e8.
17. Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, Collins JJ (2007) A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* 130:797–810.
18. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—A practical and powerful approach to multiple testing. *J R Stat Soc, B* 57:289–300.
19. Karp PD, et al. (2002) The EcoCyc Database. *Nucleic Acids Res* 30:56–58.
20. Opdyke JA, Kang JG, Storz G (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* 186:6698–6705.
21. Altuvia S, Zhang A, Argaman L, Tiwari A, Storz G (1998) The *Escherichia coli* OxyS regulatory RNA represses fhlA translation by blocking ribosome binding. *EMBO J* 17: 6069–6075.
22. Møller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 16: 1696–1706.
23. Zhang A, et al. (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J* 17:6061–6068.
24. Urban JH, Vogel J (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate glmS mRNA translation. *PLoS Biol* 6(3):e64.
25. Chen S, et al. (2002) A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Biosystems* 65:157–177.
26. Brent R, Ptashne M (1980) The lexA gene product represses its own promoter. *Proc Natl Acad Sci USA* 77:1932–1936.
27. Imlay JA (2008) Cellular defenses against superoxide and hydrogen peroxide. *Annu Rev Biochem* 77:755–776.
28. Kohanski MA, DePristo MA, Collins JJ (2010) Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol Cell* 37:311–320.
29. Beaber JW, Hochhut B, Waldor MK (2004) SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* 427:72–74.
30. Pulvermacher SC, Stauffer LT, Stauffer GV (2008) The role of the small regulatory RNA GcvB in GcvB/mRNA posttranscriptional regulation of oppA and dppA in *Escherichia coli*. *FEMS Microbiol Lett* 281:42–50.
31. Pulvermacher SC, Stauffer LT, Stauffer GV (2009) The small RNA GcvB regulates sstT mRNA expression in *Escherichia coli*. *J Bacteriol* 191:238–248.
32. Jin Y, Watt RM, Danchin A, Huang JD (2009) Small noncoding RNA GcvB is a novel regulator of acid resistance in *Escherichia coli*. *BMC Genomics* 10:165.
33. Calvo JM, Matthews RG (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* 58:466–490.
34. Tjaden B (2008) TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res* 36(Web Server issue):W109–W113.
35. Alon U (2007) Network motifs: Theory and experimental approaches. *Nat Rev Genet* 8:450–461.
36. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342.
37. Johnston RJ, Jr., Chang S, Etchberger JF, Ortiz CO, Hobert O (2005) MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc Natl Acad Sci USA* 102:12449–12454.
38. Beisel CL, Storz G (2011) The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol Cell* 41: 286–297.
39. Beisel CL, Storz G (2010) Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* 34:866–882.
40. Urbanowski ML, Stauffer LT, Stauffer GV (2000) The gcvB gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. *Mol Microbiol* 37:856–868.
41. Heil G, Stauffer LT, Stauffer GV (2002) Glycine binds the transcriptional accessory protein GcvR to disrupt a GcvA/GcvR interaction and allow GcvA-mediated activation of the *Escherichia coli* gcvTHP operon. *Microbiology* 148:2203–2214.
42. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
43. Cui Y, Wang Q, Stormo GD, Calvo JM (1995) A consensus sequence for binding of Lrp to DNA. *J Bacteriol* 177:4872–4880.
44. Hu M, Qin ZS (2009) Query large scale microarray compendium datasets using a model-based bayesian approach with variable selection. *PLoS ONE* 4:e4495.
45. Baba T, et al. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* 2:2006.0008.
46. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640–6645.
47. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 25:1203–1210.
48. Dwyer DJ, Kohanski MA, Hayete B, Collins JJ (2007) Gyrase inhibitors induce an oxidative damage cellular death pathway in *Escherichia coli*. *Mol Syst Biol* 3:91.
49. Gardner TS, di Bernardo D, Lorenz D, Collins JJ (2003) Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* 301:102–105.