Genome-scale Analysis of Escherichia coli FNR Reveals Complex Features of Transcription Factor Binding

Kevin S. Myers1,2, Huihuang Yan3a, Irene M. Ong3, Dongjun Chung4b, Kun Liang4,5c, Frances Tran6d, Sündüz Keleş4,5, Robert Landick3,6,7*, Patricia J. Kiley2,3*

1 Microbiology Doctoral Training Program, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, 2 Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, 3 Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, 4 Department of Statistics, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, 5 Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, 6 Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, 7 Department of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin, United States of America

Abstract

FNR is a well-studied global regulator of anaerobiosis, which is widely conserved across bacteria. Despite the importance of FNR and anaerobiosis in microbial lifestyles, the factors that influence its function on a genome-wide scale are poorly understood. Here, we report a functional genomic analysis of FNR action. We find that FNR occupancy at many target sites is strongly influenced by nucleoid-associated proteins (NAPs) that restrict access to many FNR binding sites. At a genome-wide level, only a subset of predicted FNR binding sites were bound under anaerobic fermentative conditions and many appeared to be masked by the NAPs H-NS, IHF and Fis. Similar assays in cells lacking H-NS and its paralog Stn showed increased FNR occupancy at sites bound by H-NS in WT strains, indicating that large regions of the genome are not readily accessible for FNR binding. Genome accessibility may also explain our finding that genome-wide FNR occupancy did not correlate with the match to consensus at binding sites, suggesting that significant variation in ChIP signal was attributable to cross-linking or immunoprecipitation efficiency rather than differences in binding affinities for FNR sites. Correlation of FNR ChIP-seq peaks with transcriptomic data showed that less than half of the FNR-regulated operons could be attributed to direct FNR binding. Conversely, FNR bound some promoters without regulating expression presumably requiring changes in activity of condition-specific transcription factors. Such combinatorial regulation may allow Escherichia coli to respond rapidly to environmental changes and confer an ecological advantage in the anaerobic but nutrient-fluctuating environment of the mammalian gut.

Citation: Myers KS, Yan H, Ong IM, Chung D, Liang K, et al. (2013) Genome-scale Analysis of Escherichia coli FNR Reveals Complex Features of Transcription Factor Binding. PLoS Genet 9(6): e1003565. doi:10.1371/journal.pgen.1003565

Editor: William F. Burkholder, Institute of Molecular and Cell Biology (IMCB), A*STAR, Singapore

Received February 20, 2013; Accepted April 29, 2013; Published June 20, 2013

Copyright: © 2013 Myers et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by grants from the NSF (MCB0640642) to RL, NIH to PJK (GM045844), RL (GM38660), and SK (HG003747 and HG006716). KSM was supported by the UW-Madison NIH Biotechnology Training Grant (5T32GM08349), and the DOE BACTER Program, (DE-FG02-04ER25627). This work was also funded in part by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FG02-07ER64494). HT sequencing was provided by the DOE Joint Genome Institute (supported by the Office of Science of the U.S. DOE under contract no. DE-AC02-05CH11231) and the UW-Madison Biotechnology Center. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: landick@bact.wisc.edu (RL); pjkiley@wisc.edu (PJK)

a Current address: Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota, United States of America.

b Current address: Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut, United States of America.

c Current address: Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada.

d Current address: Molecular and Computational Biology Program, University of Southern California at Los Angeles, Los Angeles, California, United States of America.

Introduction

Regulation of transcription initiation by transcription factors (TFs) is a key step in controlling gene expression in all domains of life. Genome-wide studies are revealing important features of the complexity of transcription regulation in cells not always apparent from in vitro studies. In eukaryotes, both the inhibition of TF binding by chromatin structure and the combinatorial action of multiple TFs contribute to the genome-wide pattern of TF binding and function [1–5]. In contrast, our knowledge of transcriptional regulation by bacterial TFs stems largely from elegant in vitro experiments that have provided atomic resolution views of TF function [6]. Much less is known about how chromosome structure and combinatorial action affect bacterial TF binding and transcriptional regulation on a genome-wide scale [7]. Previous studies have suggested that, in contrast to the chromatin-restricted TF binding in eukaryotes, the Escherichia coli genome is permissive to TF binding because the occupancy pattern for some TFs correlates well with match to consensus sequence and consequent binding affinity [8–10]. Other studies suggest that nucleoid-associated proteins (NAPs; for example H-NS, Hu, Fis, and IHF) organize the chromosome into discrete domains and structures that may affect transcriptional regulation [7,11–13], but possible global effects of NAPs on TF binding have not been systematically
Regulation of gene expression by transcription factors (TFs) is key to adaptation to environmental changes. Our comprehensive, genome-scale analysis of a prototypical global TF, the anaerobic regulator FNR from *Escherichia coli*, leads to several novel and unanticipated insights into the influences on FNR binding genome-wide and the complex structure of bacterial regulons. We found that binding of NAPs restricts FNR binding at a subset of sites, suggesting that the bacterial genome is not freely accessible for FNR binding. Our finding that less than half of the predicted FNR binding sites were occupied in vivo further challenges the utility of using bioinformatic searches alone to predict regulon structure, reinforcing the need for experimental determination of TF binding. By correlating the occupancy data with transcriptomic data, we confirm that FNR serves as a global signal of anaerobiosis but expression of some operons in the FNR regulon require other regulators sensitive to alternative environmental stimuli. Thus, FNR binding and regulation appear to depend on both the nucleoprotein structure of the chromosome and on combinatorial binding of FNR with other regulators. Both of these phenomena are typical of TF binding in eukaryotes; our results establish that they are also features of bacterial TF binding.

**Author Summary**

Regulation of gene expression by transcription factors (TFs) is key to adaptation to environmental changes. Our comprehensive, genome-scale analysis of a prototypical global TF, the anaerobic regulator FNR from *Escherichia coli*, leads to several novel and unanticipated insights into the influences on FNR binding genome-wide and the complex structure of bacterial regulons. We found that binding of NAPs restricts FNR binding at a subset of sites, suggesting that the bacterial genome is not freely accessible for FNR binding. Our finding that less than half of the predicted FNR binding sites were occupied in vivo further challenges the utility of using bioinformatic searches alone to predict regulon structure, reinforcing the need for experimental determination of TF binding. By correlating the occupancy data with transcriptomic data, we confirm that FNR serves as a global signal of anaerobiosis but expression of some operons in the FNR regulon require other regulators sensitive to alternative environmental stimuli. Thus, FNR binding and regulation appear to depend on both the nucleoprotein structure of the chromosome and on combinatorial binding of FNR with other regulators. Both of these phenomena are typical of TF binding in eukaryotes; our results establish that they are also features of bacterial TF binding.

Many FNR regulated promoters are controlled by multiple TFs (for example CRP, NarL, NarP, and NAPs [7,20,21]), which can have either positive or negative effects on FNR function depending on the promoter architecture. For example, the *narG* promoter is activated by FNR, IHF, and the nitrate-responsive regulator, NarL [23,24]; in contrast, the *dnasA* promoter is activated by FNR, but repressed by NarL. [25,26]. At the *nir* promoter, NarL displaces IHF to overcome a repressive effect of IHF and Fis, and thereby enhances FNR-dependent transcription [27]. Thus, in the presence of the anaerobic electron acceptor nitrate, FNR function can be either enhanced or repressed by NarL depending on the organization of TF-binding sites within the promoter region. In this way, the requirement of additional TFs for combinatorial regulation of promoters bound by FNR resembles transcriptional regulation in eukaryotes [28]. Such complex regulatory patterns cannot currently be inferred simply by identifying the locations of TF binding sites or by the strength of the FNR binding site. Direct measure of occupancy at these sites by each TF and correlation with the resulting transcripts in different growth conditions is needed to understand how complex bacterial regulatory networks coordinate gene expression. As an important first step, Grainger *et al.* used chromatin immunoprecipitation followed by microarray hybridization (ChIP-chip) to examine FNR occupancy using a FLAG-tagged FNR protein in *E. coli* cultures grown anaerobically in a rich medium [29]. Although many new FNR binding sites were identified, these data were not obtained from cells grown in the growth media used for reported transcriptomic experiments [17–19] and thus the datasets cannot readily be compared.

To systematically investigate FNR binding genome-wide, we performed chromatin immunoprecipitation followed by microarray hybridization (ChIP-chip) and high-throughput sequencing (ChIP-seq) for WT FNR from *E. coli* grown anaerobically in a glucose minimal medium (GMM). Computational and bioinformatic analyses were used to refine a FNR position weight matrix (PWM). The PWM was used to determine the relationship between ChIP-seq/ChIP-chip enrichment and match to the PWM, and to identify predicted FNR binding sites not detected by ChIP-seq. To examine the subset of high-quality predicted FNR binding sites that lacked a FNR ChIP-seq peak, we obtained and analyzed aerobic and/or anaerobic ChIP-chip data for NAPs H-NS and IHF along with analysis of previously published aerobic ChIP-seq data for the NAP Fis [30] to determine if NAP occupancy might prevent FNR binding. Further, the effect of H-NS on FNR occupancy was examined directly using ChIP-chip analysis of FNR as well as on O₂ dependent changes in expression in the absence of H-NS and its paralog StfA. After identifying FNR binding sites genome-wide, we performed whole genome transcription profiling experiments using expression microarrays and high-throughput RNA sequencing (RNA-seq) to compare a WT and *Δfis* strain grown in the same medium used for the DNA binding studies. The transcriptional impact of FNR binding genome-wide was investigated by correlating the occupancy data with the transcriptomic data to determine which binding events led to changes in transcription, to identify the direct and indirect regulons of FNR, and to define categories of FNR regulatory mechanisms. Finally, the aerobic and anaerobic ChIP-chip and ChIP-seq distributions of the σ⁷₀ and β subunits of RNAP throughout the genome were analyzed to determine the role of O₂ and FNR regulation on RNAP occupancy and transcription.
Results

TF binding sites were mapped genome-wide in *E. coli* K-12 MG1655 using ChIP-chip and/or ChIP-seq for FNR, σ^70 and β subunits of RNAP, H-NS, and IHF under aerobic or anaerobic growth conditions, as indicated (Figure 1). In addition, we analyzed a publically available Fis data set collected under aerobic conditions [30]. The ChIP-chip distribution of the β subunit of RNAP suggested widespread transcription under both aerobic and anaerobic conditions; as expected, whereas the O2-dependent changes in β occupancy indicated those genes that are differentially regulated by O2. Further, binding, and thus transcription, by the σ^70 housekeeping form of *E. coli* RNAP was observed throughout the chromosome; peak finding algorithms identified a large number of anaerobic σ^70 ChIP-seq peaks (2,106) and aerobic σ^70 ChIP-seq peaks (2,446) (Table S1). About 700 of the σ^70 peaks showed statistically significant O2-dependent changes in occupancy (Table S2). The O2-dependent differences in RNAP occupancy suggest extensive transcriptional reprogramming in response to changes in O2, providing an excellent model system for examining genome-scale changes in transcription.

Comparison of the profiles of other DNA binding proteins indicated that the number of binding sites for NAPs genome-wide was much greater than for the TF FNR. ChIP-seq and ChIP-chip analyses identified 207 FNR peaks, 722 anaerobic H-NS enriched regions, 782 aerobic H-NS enriched regions, 1,020 anaerobic IHF enriched regions (Tables S3, S4, and S5) and published analysis of Fis identified 1,464 aerobic enriched regions [30]. The unbiased distribution of H-NS and IHF throughout the chromosome supports previous genome-wide studies of these NAPs [30–33]. H-NS is known to form filaments that cover multiple kb of DNA [7,12,13,30,34,35] and we observed that half of the identified aerobic (390) and anaerobic (356) H-NS enriched regions were over 1 kb in length, referred to as extended H-NS binding regions (Table S3). Comparison of the aerobic and anaerobic H-NS binding distributions suggests H-NS occupancy is not greatly affected by O2 (Figure 1). For FNR, the number of high-confidence ChIP peaks (207) identified (Table S5) was just a few fold lower than the number of genes found to show FNR-dependent changes in expression (between 300–700) [17–19]. These binding site data were used to determine features of FNR binding genome-wide.

Figure 1. ChIP-seq and ChIP-chip data used in this study. The tracks are (from top): FNR ChIP-seq −O2 (blue) with peaks upstream of a subset of genes labeled, FNR ChIP-chip in Δhns/ΔstxA −O2 (black), σ^70 subunit of RNAP ChIP-seq −O2 (green), σ^70 subunit of RNAP ChIP-seq +O2 (red), H-NS ChIP-chip +O2 (light purple), H-NS ChIP-chip −O2 (orange), IHF ChIP-chip −O2 (purple), Fis ChIP-seq +O2 [30] (aqua), β subunit of RNAP −O2 (yellow), β subunit of RNAP +O2 (dark purple), and genomic coordinates. Locations of FNR binding sites are also shown: predicted FNR binding sites (red lines), FNR ChIP-seq peaks (black lines), FNR peaks upstream of operons showing a FNR-dependent change in expression (blue lines), FNR peaks co-activated by NarL/NarP (green lines), FNR peaks co-activated by CRP (purple lines), and FNR peaks repressed by Fur (yellow lines). doi:10.1371/journal.pgen.1003565.g001
These plots showed a typical binding saturation curve for both fnr IPTG levels in a strain with FNR dimer concentrations (Figure 2). The ChIP peak height did not correlate with similarity to the FNR consensus sequence.

A small number of FNR peaks showed a large degree of variation in peak height across the genome. Previous studies of the repressor LexA reported that ChIP-chip peak height correlated with the match to the consensus sequence [10], suggesting that differences in site occupancy may reflect relative binding affinities to individual sites. Because FNR is a global regulator with a more degenerate binding site than LexA, we tested whether we could use this parameter to gain additional information about FNR binding-site preferences. A PWM (Figure 2 Inset) was constructed from an alignment of sequences from the ChIP-seq peaks and the scores representing the match to the PWM were determined with the algorithm PatSer (Table S5) [36]. In contrast to the studies of LexA [10], we found a poor correlation between the height of the FNR ChIP-seq peak and the match to the FNR PWM for the site predicted within each peak (Figure 3A). The same lack of correlation was also observed with FNR ChIP-chip data, indicating that this was not specific to the detection method. Additionally, there was a lack of correlation between FNR peak height and the number of known FNR binding sites. Furthermore, the majority of the FNR ChIP-seq or ChIP-chip peaks had similar height and the number of known FNR binding sites. Furthermore, there was a lack of correlation between FNR peak height and the number of known FNR binding sites. Furthermore, the majority of the FNR ChIP-seq or ChIP-chip peaks had similar height and the number of known FNR binding sites. Furthermore, the majority of the FNR ChIP-seq or ChIP-chip peaks had similar height, regardless of the score of the FNR motif present (Figure 3A).

One explanation for this latter result is that most FNR binding sites were saturated for binding in vivo. To examine this possibility directly, we performed ChIP-chip experiments over a range of cellular FNR dimer concentrations below the normal anaerobic cellular level of ~2.5 μM [37], controlled by varying IPTG levels in a strain with fnr fused to an IPTG-inducible promoter. Peak areas for 35 selected FNR sites, representing a distribution of peak heights, were quantified for several cellular FNR dimer concentrations (~0.45, ~0.7, ~1.9, and ~2.5 μM). These plots showed a typical binding saturation curve for both novel and previously identified FNR binding sites, and revealed that all sites examined were saturated for binding at the normal cellular FNR dimer level of ~2.5 μM (Figure 3B, Figure S1). However, because the broad distribution of peak heights between different sites was still observed, despite the fact that the sites were maximally occupied, we concluded that variation in peak height was not related to strength of FNR binding (Figure 3, Figure S1). As a control, we tested four FNR peaks that were determined to be non-specific due to enrichment in a Δfnr control ChIP-chip experiment and these peaks showed no change in peak height when FNR levels were varied (Figure S1). Thus, we conclude that differences in peak height in the ChIP-seq and ChIP-chip experiments for FNR were most likely due to differences in cross-linking efficiency or immunoprecipitation at particular genomic locations and not to differences in FNR binding affinity.
Cross-linking of FNR to a subset of genomic locations may be inhibited by other proteins

A well-known challenge in genomic studies is the use of computational tools to accurately predict DNA binding sites, particularly for global regulators like FNR that have degenerate binding sites. To investigate the usefulness of the PWM generated from our set of ChIP binding sites for predicting FNR sites genome-wide, we initially used a PatSer [36] threshold low enough that a FNR motif was identified in each FNR ChIP-seq peak. However, this threshold resulted in >10,000 possible genomic FNR binding sites. In contrast, if we used a precision-recall (PR) curve [38] to determine the optimal threshold to predict FNR binding sites (ln(p-value) of −10.75), then we obtained a more reasonable number (187) of predicted FNR binding sites (Figure 2, Table S6). Surprisingly, fewer than half of these sites (63 of 187) corresponded with a FNR ChIP-seq peak (Table S6), despite the fact that some predicted sites without a corresponding ChIP-seq peak had higher quality PatSer scores than those with a ChIP-seq peak. Although it is possible that some of the predicted sites without a ChIP-seq peak contain flanking sequence elements that disfavor FNR binding, we considered the possibility that many are functional FNR sites but either FNR binding was masked by other DNA binding proteins or FNR cross-linking failed for other reasons.

NAPs are known to affect the binding of some TFs in E. coli [7,12]. To ask if the NAPs H-NS, IHF, or Fis might occlude the 124 predicted FNR binding sites lacking a FNR ChIP-seq peak, we analyzed ChIP-chip data for H-NS and IHF, obtained from the same growth conditions, and publicly available ChIP-seq data for Fis [30]. Nearly all of these FNR sites (111 of 124 sites; ~90%; silent FNR sites) were enriched in IHF, H-NS, or Fis, consistent with the idea that these NAPs occupy the silent FNR sites and thereby block FNR binding (Table S6, Figure S2). Similar occupancy was observed when the 124 predicted FNR sites were compared with H-NS and IHF enrichment from published ChIP-chip and ChIP-seq data performed under different growth conditions [30,31,33]. In comparison, only ~20% (14 of 63 sites) of the FNR sites that coincided with a FNR ChIP-seq peak were enriched in a NAP ChIP signal, significantly less than NAP occupancy at FNR sites lacking a peak (p-value < 0.05). In contrast, we found ~50% of the previously identified LexA binding sites [10] were co-occupied with H-NS. We conclude that the NAPs H-NS, IHF, or Fis likely prevent FNR binding at sites by occlusion.

We also examined whether the silent FNR sites are preferentially ruled by the extended H-NS binding regions. The extended binding regions of H-NS (>1 kb) likely represent H-NS filaments that are known to cover multiple kb of DNA and silence transcription [7,12,13,30,34]. Consistent with this notion, our results showed that the extended H-NS binding regions were negatively correlated with RNApol (β) ChIP-chip occupancy and this silencing occurred in both the presence and absence of O2 (p-value < 0.005) (Figure S3). In contrast, shorter H-NS enriched regions (<1 kb) were both positively and negatively correlated with RNApol ChIP-chip occupancy under aerobic and anaerobic growth conditions. The 46 silent FNR sites bound by H-NS were more likely to be occupied by extended H-NS binding regions (42 sites) than by short H-NS binding regions (4 sites) (p-value < 0.05; example in Figure S3C), suggesting that extended H-NS binding regions may inhibit FNR binding at silent FNR sites.

To investigate the impact of H-NS binding on FNR occupancy, we characterized FNR ChIP-chip peaks in a strain deleted for both hns and stpA; stpA encodes a H-NS paralog that partially compensates for H-NS in a Δhns mutant [39,40]. Many new FNR peaks (196) appeared in the Δhns/ΔstpA strain (Figure 1, Figure 4A–C, Table S7), and a large fraction (81%; 158 FNR peaks) of these new peaks corresponded to H-NS binding regions in the WT strains, indicating that FNR binding was unmasked in the absence of H-NS and StpA. The distribution of the FNR PWM scores of the FNR sites found within the FNR ChIP-chip peaks unmasked by the absence of H-NS and StpA was similar to that found in the WT strain (Figure 4C, Tables S6 and S7). The majority (78% of 99) of silent FNR sites lacking FNR peaks in the Δhns/ΔstpA strain were enriched for IHF and/or Fis, suggesting that these NAPs still occluded FNR binding in the absence of H-NS and StpA (Table S6). Taken together, these results establish that removal of H-NS and StpA allowed FNR to bind to sites covered by H-NS in WT strains.

Nearly all FNR peaks found in the WT strain were retained in the Δhns/ΔfstA mutant (163 of 169 peaks; Figure 4A, Table S7), but a small proportion (~15%) showed a significant increase in peak average (average log2/IP/INPUT value of the binding region) in the Δhns/ΔfstA strain (Figure 4D). The majority of these FNR peaks with increased peak averages were also bound by H-NS in the WT strain, suggesting that removing H-NS allowed for increased cross-linking or immunoprecipitation of FNR at these loci likely due to changes in chromosomal structure in the absence of H-NS and StpA [33]. In contrast, removing H-NS did not affect FNR occupancy or cross-linking at locations lacking H-NS ChIP signal in WT strains. We conclude that H-NS reduces or blocks FNR binding at many locations in vivo.

Operons in the FNR regulon were organized into seven regulatory categories

To determine which FNR binding events from the WT strain caused a change in gene expression, the FNR occupancy data were correlated with the 122 operons differentially expressed (DE) by FNR (Table S8). Surprisingly, less than a half of the 122 operons were correlated with a FNR ChIP-seq peak while less than a fourth of the 207 FNR ChIP-seq peaks were correlated with a FNR-dependent change in expression (Figure S4). To address this unexpected result, we systematically analyzed the regulation of all of these operons by incorporating published data and classified the operons into seven regulatory categories (Figure 5). Category 1 (Table 1) contained operons that were directly activated by FNR because they showed a FNR-dependent increase in anaerobic transcript levels and a FNR ChIP-seq peak within 500 nt of the translation start site of the first gene of an operon. Category 2 (Table 1) contained operons that were directly repressed by FNR (showed a FNR-dependent decrease in expression and had a FNR ChIP-seq peak). Categories 3–5 contained a surprisingly large number of operons (156) with a FNR ChIP-seq peak within 500 nt of the translation start site of the first gene of an operon but no FNR-dependent change in expression. Previously published studies (23 operons) and our additional collation of other relevant TF-binding sites (52 operons) suggest that at least half (75) of these sites may be directly regulated by FNR under alternative growth conditions (Table S9). For example, Category 3 (Tables 2 and 3) contained operons known or proposed to be co-regulated by FNR and another TF under growth conditions not used in our study. Category 4 (Table 4) contained operons known to be repressed by another TF under our growth conditions. Category 5 (Table S9) contained operons with other potential regulatory mechanisms. Category 6 (Table 5, Table S10) contained operons that were indirectly regulated by FNR because no FNR ChIP-seq peak was found within 500 nt of the translation start site despite showing a FNR-dependent change in expression. Finally, Category 7 (Table S11) contained operons with a FNR peak identified only in the
Δhns/ΔstpA strain, which also showed potential FNR regulation in the absence of H-NS and StpA.

**Category 1 - Direct activation by FNR**

The 32 operons directly activated by FNR (Table 1) contain some of the best-studied FNR regulated operons. In addition to operons associated with anaerobic respiration (dmsABC, frdABCD, narGHIJ) [41–43], this category included glycolytic (pykA) and fermentative enzymes (pflB and ackA), which would be expected to promote mixed acid fermentation of glucose to ethanol, acetate, formate and succinate in the absence of an added electron acceptor (Figure 6), the conditions used in this study. As expected, we also found that these promoters showed an increase in σ70 occupancy, as illustrated by representative FNR and σ70 data for FNR activation of dmsABC (Figure 7), providing a proof-of-principle for our approach. While expression of many operons in this category was known to be FNR regulated, only about half had been shown to directly bind FNR (Table 1).

FNR also directly activated operons with functions that illustrate the broader role of FNR in anaerobic metabolism: pepE, a peptidase, suggesting peptide degradation in *E. coli* similar to that observed in *Salmonella* [44]; ymjE, an enzyme involved in biosynthesis of molybdopterin, a cofactor used by anaerobic respiratory enzymes [45]; pyrD, a dihydroorotate dehydrogenase in pyrimidine biosynthesis [46]; and ynfK, a predicted dethiobiotin synthetase and paralog of BioD of the biotin synthesis pathway. The activation of the biofilm TF bssR by FNR suggests a link between biofilm formation and anaerobiosis (Table 1). FNR directly activated the carnitine-sensing TF CaiF, confirming a link between FNR and carnitine metabolism [29,47]. In addition, the FNR-enriched region found upstream of fnrS supports FNR direct transcription activation of this small regulatory RNA [48,49], although the fnrS sRNA was not represented in our gene expression arrays and was too small to be detected by our RNA-Seq protocol (Table 1).

To determine the position of FNR binding sites relative to the TSS, we used the FNR PWM (Figure 2 Inset) to search the FNR binding site.

---

**Figure 4. Identification of FNR occupancy in a Δhns/ΔstpA strain compared to WT.**

A) Venn diagram showing the overlap of FNR peaks identified only in the WT strain (purple), in both the WT and the Δhns/ΔstpA strains (blue) or only in the Δhns/ΔstpA strain (green).

B) Example of a high-quality predicted FNR binding site (blue line) within fimE that showed no FNR binding in the WT strain (blue trace), but did show enrichment of H-NS in the WT strain (purple trace). A FNR ChIP-chip peak was identified in the Δhns/ΔstpA strain (green trace) at the location of the predicted FNR binding site.

C) The 193 FNR peaks found only in the Δhns/ΔstpA strain with a statistical increase in FNR occupancy in the Δhns/ΔstpA strain compared to the WT strain (p-value<0.05). Correlation of ChIP-chip peak average (log2(IP/INPUT) average) and the corresponding FNR PWM score (determined by PatSer [36]).

D) Correlation of ChIP-chip peak averages (log2(IP/INPUT) average) for FNR ChIP-chip peaks found in both WT and Δhns/ΔstpA strains. Shown are peaks with no statistical difference in occupancy (red points) and those peaks that showed a statistical increase in FNR occupancy (blue points) in the Δhns/ΔstpA strain compared to the WT strain (p-value<0.05).

doi:10.1371/journal.pgen.1003565.g004
enriched regions using a PatSer score threshold low enough to identify FNR sites from every ChIP peak [36]. A majority (89%) of the FNR ChIP-seq peaks in the FNR direct regulon contained one FNR binding site (Table 1). Of the 23 promoters directly activated by FNR for expression, with the lighter blue circles representing FnrS other TFs (CaiF, BssR, PdhR, GadE) that potentially control the indirect regulon, shown by yellow squares. Red circles are operons known or predicted to be co-regulated by FNR and other TFs, while green circles have other potential regulatory mechanisms with FNR. Each box represents different categories of FNR regulation identified in this study. Categories 1 and 2 (upper left and middle boxes) show direct activation and repression of operons by FNR (blue oval). Category 3 (upper right box) show co-activation by other TFs (red star; e.g. CRP, NarL, NarP) and Category 4 (lower left box) shows TF repression that prevents FNR regulation (green rectangle; e.g. Fur). Category 5 (lower middle box) represents operons with other possible regulatory mechanisms with FNR and Category 6 (lower right box) shows the subset of the indirect regulon affected by other TFs and, for example, by the small, regulatory RNA FnrS (red line).

doi:10.1371/journal.pgen.1003565.g005

Figure 5. The FNR transcriptional network and categories of FNR regulation. A) Graphical representation of the FNR transcriptional network. FNR is shown in the blue octagon, while other TFs (CRP, Fur, NarL) are shown as purple diamonds. Circles represent operons with an upstream FNR ChIP-seq peak, while squares represent operons indirectly regulated by FNR. Dark blue circles are operons directly dependent on FNR for expression, with the lighter blue circles representing FnrS other TFs (CaiF, BssR, PdhR, GadE) that potentially control the indirect regulon, shown by yellow squares. Red circles are operons known or predicted to be co-regulated by FNR and other TFs, while green circles have other potential regulatory mechanisms with FNR. Each box represents different categories of FNR regulation identified in this study. Categories 1 and 2 (upper left and middle boxes) show direct activation and repression of operons by FNR (blue oval). Category 3 (upper right box) show co-activation by other TFs (red star; e.g. CRP, NarL, NarP) and Category 4 (lower left box) shows TF repression that prevents FNR regulation (green rectangle; e.g. Fur). Category 5 (lower middle box) represents operons with other possible regulatory mechanisms with FNR and Category 6 (lower right box) shows the subset of the indirect regulon affected by other TFs and, for example, by the small, regulatory RNA FnrS (red line).

doi:10.1371/journal.pgen.1003565.g005
Table 1. Operons with an upstream FNR ChIP-seq peak and a FNR-dependent change in expression under GMM.

| Peak Center (nt) | Operon | B-number of first gene | Function of Operon Product | Number of FNR binding sites | Location top scoring FNR binding site | % occupancy +O2 relative to -O2 | WT -O2 expression relative to WT +O2 expression | Previous Experimental Evidence of FNR Binding | Previous Evidence of FNR Regulated Expression |
|-----------------|--------|------------------------|-----------------------------|-----------------------------|-----------------------------------|------------------------------------|-----------------------------------------------|---------------------------------------------|-----------------------------------------------|
| 1,003,976       | pyID   | b0945                  | Dihydroorotate Dehydrogenase| 1                          | -38.5                             | +                                  | 0                             | [29]                                        | [17]                                      |
| 1,656,036       | yIDFGH- dmsD | b1587                 | Putative Selenate Reductase (yIDFGH); DMS Reductase Maturation Protein (dmsD) | 1                          | -40.5                             | +                                  | +                             | None                                       | [18,19]                                   |
| 1,935,550       | pyKA   | b1854                  | Pyruvate Kinase III         | 1                          | -40.5                             | +                                  | +                             | [29]                                        | [19]                                      |
| 3,611,605       | nkaBCDE | b3476                  | Nickel Transporter         | 1                          | -40.5                             | +                                  | +                             | None                                       | [126]                                     |
| 953,741         | locA-pflB | b0904                 | Formate Transporter (locA); Pyruvate Formate-Lyase (pflB) | 2                          | -40.5                             | +                                  | +                             | [127]                                       | [128]                                     |
| 2,714,605       | yID    | b2579                  | Stress-Induced Alternative Pyruvate Formate-Lyase | 1                          | -40.5                             | +                                  | +                             | [129]                                       | [129]                                     |
| 940,035         | dmABC   | b0984                  | Dimethyl Sulfoxide Reductase| 1                          | -41.5                             | +                                  | +                             | [41]                                        | [41]                                      |
| 1,279,003       | narGHU  | b1224                  | Nitrate Reductase          | 1                          | -41.5                             | +                                  | +                             | [41]                                        | [130]                                     |
| 1,627,208       | ydfZ    | b1541                  | Unknown Function           | 2                          | -41.5                             | +                                  | +                             | [29]                                        | [18,19]                                   |
| 1,837,412       | yjE     | b1757                  | Molybdopterin Synthase Sulfurtransferase | 1                          | -41.5                             | +                                  | +                             | None                                        | [18]                                      |
| 3,491,947       | nirBDC- cysG | b3365                | Nitrite Reductase (nirBDC); Uroporphyrin III C-Methyltransferase (cysG) | 1                          | -41.5                             | +                                  | +                             | [21]                                        | [130]                                     |
| 4,285,670       | nirABCDEFG | b4070                 | Periplasmic Nitrite Reductase| 1                          | -41.5                             | +                                  | +                             | [42]                                        | [42]                                      |
| 34,059          | caiF    | b0034                  | Carnitine Transcriptional Activator | 1                          | -41.5                             | +                                  | 0                             | [29]                                        | [47]                                      |
| 1,277,082       | narK    | b1223                  | Nitrate/Nitrite Antipporter| 1                          | -41.5                             | +                                  | +                             | [41]                                        | [131]                                     |
| 877,441         | lasR    | b0836                  | Regulator of Biofilm Formation | 1                          | -41.5                             | +                                  | +                             | None                                        | None                                       |
| 1,752,688       | ydh/YWKLUF | b1674               | Predicted Oxidoreductase System | 1                          | -42.5                             | +                                  | +                             | [132]                                       | [18,19]                                   |
| 1,407,150       | fnS     | b4699                  | Small Regulatory RNA       | 1                          | -42.5                             | +                                  | N/A                           | [48,49]                                     | [48,49]                                   |
| 4,380,446       | frdABCD  | b4154                  | Fumarate Reductase         | 1                          | -45.5                             | +                                  | +                             | [133]                                       | [43]                                      |
| 3,635,591       | pta     | b3493                  | Phosphate Transporter      | 1                          | -56.5                             | +                                  | 0                             | [29]                                        | [18,19]                                   |
| 1,831,403       | ydiXYZ- ynjABC | b1750            | Predicted Proteins (ydiXYZ-ynjAB); Predicted Transporter (ynjCD) | 1                          | -60.5                             | +                                  | +                             | [18]                                        | [18,19]                                   |
| 913,151         | hcp-hcr | b0873                  | Hybrid-Cluster Protein (hcp); NADH Oxidoreductase (hcr) | 1                           | -72.5                             | No % occupancy +O2 relative to -O2 | +                             | [134]                                       | [134]                                     |
| 2,411,410       | ackA-pta | b2296                  | Acetate Kinase/jackA; Phosphate Acetyltransferase (pta) | 1                          | -74.5                             | +                                  | +                             | [29]                                        | [135]                                     |
| 4,347,259       | dckB-fumB | b4123                 | Dicarboxylate Transporter (dckB); Fumarase B (fumB) | 1                          | -132.5                            | +                                  | +                             | [29]                                        | [136]                                     |
| Peak Center (nt)  | Operon | B-number of first gene | Function of Operon Product | Number of FNR binding sites | Location top scoring FNR binding site | σ70 occupancy -O2 relative to +O2 | WT -O2 expression relative to WT +O2 expression | Previous Experimental Evidence of FNR Binding | Previous Evidence of FNR Regulated Expression |
|------------------|--------|------------------------|-----------------------------|-----------------------------|---------------------------------------|----------------------------------|-------------------------------------------|-------------------------------------------|-------------------------------------------|
| 1,665,279        | ynk    | b1593                  | Predicted Dethiobiotin Synthetase | 1                          | TSS not known                         | +                                | +                                        | [29]                                      | [18,19]                                  |
| 2,415,052        | yfc    | b2298                  | Putative S-transferase         | 1                          | TSS not known                         | +                                | 1                                        | [29]                                      | [19]                                     |
| 3,299,421        | yhbbIV | b3158                  | Predicted Peptidase            | 1                          | TSS not known                         | +                                | +                                        | [29]                                      | [18,19]                                  |
| 3,352,132        | yhcC   | b3211                  | Predicted Fe-S Oxidoreductase  | 1                          | TSS not known                         | +                                | None                                     | [18]                                      |                                          |
| 3,654,199        | yhiD   | b3508                  | Predicted Mg2⁺ ATP-dependent Transporter | 1                          | TSS not in peak region                | +                                | None                                     | [18]                                      |                                          |
| 3,463,910        | yjIL   | b4335                  | Hypothetical Protein           | 1                          | TSS not in peak region                | +                                | None                                     | [17,18]                                  |                                          |
| 4,615,181        | yjIL   | b4380                  | Conserved Protein              | 1                          | TSS not known                         | +                                | None                                     | [18,19]                                  |                                          |
| 655,298          | dcoC   | b6021                  | Dicarboxylate Transporter      | 1                          | TSS not in peak region                | +                                | None                                     | [137]                                    |                                          |
| 4,228,181        | pepE   | b4021                  | Peptidase E                    | 1                          | TSS not known                         | +                                | None                                     | [18]                                      |                                          |

Operons directly repressed by FNR with an overall decrease in expression under anaerobic growth conditions (Category 2)

| Peak Center (nt)  | Operon | B-number of first gene | Function of Operon Product | Number of FNR binding sites | Location top scoring FNR binding site | σ70 occupancy -O2 relative to +O2 | WT -O2 expression relative to WT +O2 expression | Previous Experimental Evidence of FNR Binding | Previous Evidence of FNR Regulated Expression |
|------------------|--------|------------------------|-----------------------------|-----------------------------|---------------------------------------|----------------------------------|-------------------------------------------|-------------------------------------------|-------------------------------------------|
| 1,397,604        | fnr    | b1334                  | Anaerobic Transcriptional Regulator | 1                          | −0.5                                 | −                                | o                                        | [138]                                    |                                           |
| 142,604          | can    | b0126                  | Carbonic Anhydrase 2         | 2                          | −11.5                                | −                                | None                                     | [17]                                     |                                           |
| 2,088,088        | hisGDC | b2018                  | Histidine Biosynthesis       | 1                          | −11.5                                | +                                | o                                        | [18]                                     |                                           |
| 1,165,151        | ndh    | b1109                  | NADH:Ubiquinone Oxidoreductase II | 1                          | −50.5                                | −                                | o                                        | [139]                                    | [139]                                   |
| 2,176,686        | fabB   | b2097                  | Fructose Bisphosphate Aldolase Class I | 1                          | −59.5                                | No σ70 Peaks (σ3 Site) | o                                        | None                                     | [18]                                     |
| 1,030,741        | yccA   | b0970                  | Putative Transport Protein   | 1                          | −71.5                                 | −                                | o                                        | None                                     | None                                     |
| 3,217,259        | yjIG   | b3073                  | Putrescine Aminotransferase  | 1                          | −96.5                                 | −                                | None                                     | None                                     | None                                     |
| 1,014,801        | rmf    | b0953                  | Ribosome Modulation Factor   | 1                          | −98.5                                 | +                                | o                                        | None                                     | None                                     |
| 4,231,491        | lysC   | b4024                  | Aspartate Kinase III         | 2                          | −115.5                                | +                                | o                                        | None                                     | None                                     |
| 1,986,025        | yecR   | b1904                  | Predicted Protein            | 1                          | +30                                   | −                                | o                                        | None                                     | [18,19]                                  |
| 400,328          | irp    | b0382                  | Anti-Adaptor Protein for σ7 Stabilization | 2                          | TSS not in peak region                | −                                | −                                        | None                                     | [18,19]                                  |
| 1,860,717        | msrB   | b1778                  | Methionine Sulfoxide Reductase B | 1                          | TSS not in peak region                | +                                | −                                        | None                                     | [18]                                     |
| 2,342,585        | ndkB   | b2234                  | Ribonucleoside Diphosphate Reductase I | 1                          | TSS not in peak region                | −                                | −                                        | None                                     | [18,19]                                  |

Operons directly repressed by FNR with an overall increase in expression under anaerobic growth conditions (Category 2)

| Peak Center (nt)  | Operon | B-number of first gene | Function of Operon Product | Number of FNR binding sites | Location top scoring FNR binding site | σ70 occupancy -O2 relative to +O2 | WT -O2 expression relative to WT +O2 expression | Previous Experimental Evidence of FNR Binding | Previous Evidence of FNR Regulated Expression |
|------------------|--------|------------------------|-----------------------------|-----------------------------|---------------------------------------|----------------------------------|-------------------------------------------|-------------------------------------------|-------------------------------------------|
| 3,656,030        | gadE   | b3512                  | Transcriptional Activator   | 2                          | −22.5                                 | +                                | +                                        | None                                     | [18]                                     |
| 3,654,924        | hdeD   | b3511                  | Acid-Resistance Membrane Protein | 1                          | −43.5                                 | +                                | None                                     | None                                     | [18]                                     |
| Peak Center (nt) | Operon | B-number of first gene | Function of Operon | Number of FNR binding sites | Location top scoring FNR binding site | σ^{70} occupancy relative to +O_2^9 | WT -O_2 expression relative to WT +O_2 expression | Previous Experimental Evidence of FNR Binding | Previous Evidence of FNR Regulated Expression |
|-----------------|--------|------------------------|--------------------|----------------------------|---------------------------------------|-----------------------------------|---------------------------------|---------------------------------------------|---------------------------------------------|
| 121,987         | pdhR-aceEF-lpdA | b2113 | Pyruvate Dehydrogenase | 1 | −50.5 | − | + | [29] | [140] |
| 770,404         | cydAB   | b0733 | Cytochrome bd-1 Terminal Oxidase | 1 | −53.5 | + | + | [50] | [141] |
| 3,654,924       | hdeAB-yhdD | b3510 | Acid-Resistance Proteins | 1 | −126.5 | + | + | None | [18] |
| 1,131,935       | ampW    | b1256 | Outer Membrane Protein | 1 | −126.5 | + | + | [29] | [19] |
| 2,310,730       | ompC    | b2215 | Outer Membrane Porin C | 1 | TSS not in peak region | + | + | None | [18] |
| 2,848,688       | hycABCDEFGHI | b2725 | Hydrogenase 3 | 1 | TSS not in peak region | No σ^{70} Peaks (σ^{70} Site) | + | None | [18] |

*Genomic location within each FNR ChIP-seq peak with the highest read count (the summit of the peak).

*Operon downstream of the FNR ChIP-seq peak, and operon designation was obtained from EcoCyc [70].

*Identification number (B-number) for the first gene in each operon, obtained from EcoCyc [70].

*Functional description of the products of the operons, obtained from EcoCyc [70].

*Number of predicted FNR motifs identified within the FNR ChIP-seq peak region.

*Location of FNR PWM (motif with best PatSer score used if more than one motif identified) relative to the transcription start site (if known). Start site locations obtained from EcoCyc [70] and Kim et al [142].

*Analysis of σ^{70} occupancy (as determined using ChIP-seq data) under aerobic compared to anaerobic growth conditions. Increases (+) and decreases (−) in σ^{70} occupancy were statistically determined using a one-sided, paired t-test and p-values were corrected using the Bonferroni method. Also listed are those sites with no σ^{70} ChIP-seq peaks (Table S2).

*The expression of each operon in WT −O_2 cultures was compared to expression in WT +O_2 cultures. Each operon was determined to have a significant (>2 fold) increase (+), decrease (−) and no change (o) in WT −O_2 relative to WT +O_2 (Table S12).

*Reference for experimentally determined FNR binding at the location of the FNR ChIP-seq peak, if previously identified. Those operons without experimentally determined FNR binding data are marked “None”.

*Reference for FNR-dependent change in expression of each operon, if previously identified. Those without previous FNR expression data are marked “None”.

DOI:10.1371/journal.pgen.1003565.0001
These operons include *hidD*, *gadE* and *hidAB*-*yhiD*, involved in acid stress response, and *ompC* and *ompW*, encoding outer membrane proteins. The finding that strains lacking *ompC*, *msh*, and *tpo* show decreased viability compared to single or double mutants [51] suggests that these proteins may function in a common stress response, potentially necessary under microaerobic growth conditions. Interestingly, for the 16 promoters directly repressed by FNR with a known TSS, the FNR binding sites were broadly responsive, potentially necessary under microaerobic growth conditions. Finally, comparison of the transcriptomic data to changes in σ70 holo-RNAP ChIP-seq occupancy under aerobic and anaerobic growth conditions revealed that nearly all FNR-regulated operons are expressed using σ70 RNAP. Increases or decreases in σ70 enrichment under anaerobic conditions correlated well, for the most part, with the expression changes for promoters activated or repressed by FNR, respectively, as well as expression changes in anaerobic and aerobic WT cultures (Table 1, Tables S2 and S12). Three operons, which lacked σ70 enrichment, have shown to be dependent on σ5 (hcr-hcr) [52], σ5 (hycABCDEFGHI) [53] and σ3 (fobB) [54], raising the possibility that alternative σ factors transcribe a subset of the FNR direct regulon.

**Category 3 – Co-activation by another TF and FNR**

Comparison of our FNR data with published regulatory data suggested that many FNR regulated operons were co-activated by TFs not active during growth in GMM, specifically NarL, NarP and CRP. For example, FNR-dependent transcription of *napF-DAGHBC*, encoding the periplasmic nitrate reductase, requires co-activation by the NO3−/NO2− sensing response regulator NarP [55]. Transcriptional data [19] showed FNR and NarL or NarP dependent activation in the presence of NO3− and/or NO2− (Table 2) [19] for nine operons that we found associated with FNR ChIP-seq peaks but lacking a FNR-dependent change in expression in our transcriptomic experiments, suggesting co-activation by NarL or NarP when NO3− and/or NO2− is present.

Another possible co-activator of operons in this group is CRP, which is inactive under glucose fermentation conditions presumably because of decreased cAMP [56]. Although previous studies have shown that *ansB* is co-activated by FNR and CRP [57], we did not observe binding of FNR upstream of *ansB* in this study, potentially due to differences in growth conditions. Nevertheless, 12 operons within this group showed an increase in anaerobic expression in transcriptomic data obtained from WT strains grown with carbon sources other than glucose (e.g. glycerol, mannose, arabinose or xylose) compared to growth in glucose (Table 3) [19,58] (Park and Kiley, Personal Communication). A majority (nine) contained distinct CRP and FNR binding sites, suggesting potential competition between FNR and CRP for binding when both TFs are active (Table 3).

**Category 4 – Repression by another TF prevents FNR regulation**

We propose that FNR activation of ten operons is repressed by Fur under the iron replete conditions used here, similar to the known regulation of *feoABC*, encoding a ferrous iron uptake transporter [59]. In addition to *feoABC*, nine additional operons known to be bound by Fur had a FNR ChIP-seq peak but lacked a FNR-dependent change in expression, suggesting that Fur repression masked FNR regulation of these operons (Table 4).

**Category 5 – Other potential regulatory mechanisms with FNR**

Expression of several of the remaining operons associated with FNR ChIP-seq peaks are known to require other TFs but were not known to be co-regulated by FNR, potentially explaining the lack

---

**Table 2. Operons associated with a FNR ChIP-seq peak and lacking a FNR-dependent change in expression in GMM but are activated by FNR in the presence of NO3−, NO2−, NarL or NarP (Category 3) according to Constantinidou et al. [19].**

| Peak Center (nt)a | First Gene of Operon Downstreamb | B-number of First Gene in Operonc | Cellular Functiond | Impact of NO3−, NO2−, NarL or NarP on FNR Regulatione |
|------------------|---------------------------------|----------------------------------|-------------------|-------------------------------------------------|
| 816075           | maoA                            | b0781                             | Molybdenopterin Biosynthesis Protein A | Activated by FNR in NO3−                          |
| 1185000          | pepT                             | b1127                             | Peptidase T       | Activated by FNR in NO3−                          |
| 1545300          | fdnG                             | b1474                             | Formate Dehydrogenase N | Activated by FNR in NO3− and by NarL |
| 2301675          | napF                             | b2208                             | Periplasmic Nitrate Reductase | Activated by FNR in NO3− , NO2− and by NarP |
| 2619000          | upp                              | b2498                             | Uracil Phosphoribosyltransferase | Activated by FNR in NO3− |
| 3538050          | feoA                             | b3408                             | Ferrous Iron Transport Protein | Activated by FNR in NO3− |
| 4131675          | katG                             | b3942                             | Hydroperoxidase I  | Activated by FNR in NO3− and NO2− |
| 4360558          | cadC                             | b4133                             | Metal-Sensitive Transcriptional Activator | Activated by FNR in NO3− |
| 4460925          | mdID                             | b4238                             | Ribonucleoside-Triphosphate Reductase | Activated by FNR in NO3− and NO2− |

aGenomic location within each FNR ChIP-seq peak with the highest read count (the summit of the peak).
bFirst gene of the operon downstream of the FNR ChIP-seq peak, and operon designation was obtained from EcoCyc [70].
cIdentification number (B-number) for the first gene in each operon, obtained from EcoCyc [70].
dFunctional description of the products of the operons, obtained from EcoCyc [70].
eGrowth conditions (presence of NO3−, NO2−, NarL or NarP) in which FNR activated the operon, according to [19].

doi:10.1371/journal.pgen.1003565.t002
of FNR-dependent regulation under our growth conditions. A subset of these FNR-regulated operons may be co-regulated by OxyR (active under oxidative stress), CadC (active at low external pH) or PhoP (active in low Mg$^{2+}$ concentration) (Table S9). In a recent SELEX study [60], three BasR binding sites were identified upstream of operons containing FNR peaks but without a FNR-dependent change in expression, suggesting BasR could possibly act as a subset of these operons may be co-regulated by these regulators. The indirect regulation by FNR could be easily explained for 11 operons targeted by the small RNA FnrS, which is directly activated by FNR (Figure 2) [36]. Only one operon, hmp, contained a predicted FNR-binding site and previous data also supported FNR binding to hmp [61]. Thus, 69 operons are indirectly regulated by FNR. The indirect regulation by FNR could be easily explained for 11 operons targeted by the small RNA FnrS, which is directly activated by FNR [48,49]. These RNAs increased in the FNR strain because of the lower FnrS levels (Table 5) [48,49].

Category 6 – Indirect FNR regulation through hierarchical transcriptional regulator action

Surprisingly, a large number of operons (70) were differentially expressed by FNR but were not associated with a FNR ChIP-seq peak, suggesting they are regulated by FNR indirectly (Category 6, Table S10). To determine whether any of these operons had a FNR site upstream that was missed by ChIP-seq, sequences 500 nt upstream of these operons were searched using the FNR PWM and the algorithm PatSer with the PR curve determined threshold (Figure 2) [36]. Only one operon, hmp, contained a predicted FNR-binding site and previous data also supported FNR binding to hmp [61]. Thus, 69 operons are indirectly regulated by FNR. The indirect regulation by FNR could be easily explained for 11 operons targeted by the small RNA FnrS, which is directly activated by FNR [48,49]. These RNAs increased in the FNR strain because of the lower FnrS levels (Table 5) [48,49].

Category 7 – FNR regulation in the absence of H-NS and StpA

To determine whether FNR binding to sites unmasked by the absence of H-NS and StpA caused a change in expression, we assayed if any of the corresponding genes were differentially expressed by O$_2$ only in the Δhns/Δstpa strain. Of the 158 new FNR peaks unmasked in the Δhns/Δstpa strain, 18 genes showed an anaerobic increase in expression (Table S11), and consistent with this, many of the promoters contained a FNR binding site at a position associated with activation (e.g. near –41.5).

### Table 3. Operons associated with a FNR ChIP-seq peak and lacking a FNR-dependent change in expression in GMM but are potentially co-activated by CRP and FNR (Category 3).

| Peak Center (nt)$^a$ | First Gene of Operon Downstream$^b$ | B-number of First Gene$^c$ | Cellular Function$^d$ | Increase in Expression in Other Carbon Sources$^e$ | Distinct FNR and CRP Sites$^f$ | Reference$^g$ |
|---------------------|------------------------------------|---------------------------|-----------------------|-----------------------------------------------|-------------------------------|----------------|
| 1019175             | ompA                               | b0957                     | Outer Membrane Protein 3A | +                              | +                             | [58,143], (Park and Kiley, Personal Communication) |
| 1157025             | ptsG                               | b1101                     | Glucose PTS Permease     | +                              | +                             | [58,144] |
| 1608750             | uxaB                               | b1521                     | Altronate Oxidoreductase | +                              | −                             | [19,145], (Park and Kiley, Personal Communication) |
| 2531550             | ptsH                               | b2415                     | Phosphonolpyruvate Dependent Phosphotransferase System | +                              | −                             | [58,146] |
| 2632200             | guaB                               | b2508                     | IMP Dehydrogenase        | +                              | −                             | [58,147] |
| 3229415             | fadH                               | b3081                     | 2,4-Dienoyl-CoA Reductase | +                              | +                             | [19,148], (Park and Kiley, Personal Communication) |
| 3242850             | uxaC/txuT                          | b3092/b3093               | Glucuronate and Galacturonate Isomerase (uxaC)/Hexuronate Transporter (exuT) | +/+                           | +/+                           | [19,145], (Park and Kiley, Personal Communication) |
| 3408300             | dusB                               | b3260                     | tRNA Dihydouridine Synthase | +                              | +                             | [58,145] |
| 3490500             | ppiA                               | b3363                     | Peptidyl-Prolyl Cis-Trans Isomerase A | +                              | +                             | [58,149], (Park and Kiley, Personal Communication) |
| 3544500             | gntT                               | b3415                     | Gluconate Transporter    | +                              | +                             | [19,150], (Park and Kiley, Personal Communication) |
| 4366425             | aprA                               | b4139                     | Aspartate Ammonia-Lyase  | +                              | +                             | [19,136], (Park and Kiley, Personal Communication) |

$^a$Genomic location within each FNR ChIP-seq peak with the highest read count (the summit of the peak).

$^b$First gene of the operon downstream of the FNR ChIP-seq peak, and operon designation was obtained from EcoCyc [70]. For peaks located within divergent promoters with both operons activated by CRP, both genes are identified, separated by “/”.

$^c$Identification number (B-number) for the first gene in each operon, obtained from EcoCyc [70]. For peaks located within divergent promoters with both operons activated by CRP, both genes are identified, separated by “/”.

$^d$Functional description of the products of the operons, obtained from EcoCyc [70]. For peaks located within divergent promoters with both operons activated by CRP, the functions of both operons are listed, separated by “/”.

$^e$Increase of expression (+) or no change in expression (0) of gene as determined using microarray analyses when WT E. coli was grown with carbon sources other than glucose (e.g. glycerol, xylose, mannose, arabinose). For peaks located within divergent promoters the expression changes of both operons are listed, separated by “/”.

$^f$Indication if the FNR site different than the CRP site (++) or overlapping the CRP site (+). For peaks located within divergent promoters the FNR and CRP site positions of both operons are listed, separated by “/”.

$^g$Reference for CRP activation of each operon or regulation in alternative carbon sources of each operon.

doi:10.1371/journal.pgen.1003565.t003

---

### Table 3. Operons associated with a FNR ChIP-seq peak and lacking a FNR-dependent change in expression in GMM but are potentially co-activated by CRP and FNR (Category 3).
example, hemolysin E (hlyE), in agreement with previous results [62], and the anaerobic NAP Dan (ttdR) [63] showed increased expression under anaerobic conditions only. This suggests a possible role of Dan in the absence of H-NS and StpA. Only two genes showed a decrease in expression in the absence of H-NS and StpA (yncD and feaR) under only anaerobic growth conditions. However, the expression of the vast majority of genes having FNR bound at unmasked sites resulted in changes under both aerobic and anaerobic growth conditions, indicating that changes in nucleoid structure that occur in the absence of H-NS and StpA could cause misregulation of transcription. For example, H-NS and Rho coordinate to regulate transcriptional termination and the absence of H-NS may cause increased transcriptional readthrough of Rho-dependent terminators [64]. Thus, it seems likely that our analysis provides an underestimate of the impact of H-NS on FNR function, since physiological conditions that alter H-NS activity are likely to have less severe effects on nucleoid structure.

### Table 4. Operons associated with a FNR ChIP-seq peak and lacking a FNR-dependent change in expression in GMM but are repressed by Fur (Category 4).

| Peak Center (nt) | First Gene of Operon Downstream | B-number of First Gene | Cellular Function | Reference |
|------------------|---------------------------------|------------------------|------------------|-----------|
| 611865           | fes                             | b0585                  | Enterochelin Esterase | [151,152] |
| 621417           | fepD/entS                        | b0590/b0591            | Ferric Enterobactin Transporter (fepD)/Enterobactin Efflux Transporter (entS) | [152-154] |
| 623975           | fepB/entC                        | b0592/b0593            | Ferric Enterobactin Transporter (fepB)/Isochorismate Synthase (entC) | [152,155,156] |
| 1298775          | oppA                             | b1243                  | Oligopeptide Transporter | [152] |
| 1308975          | tonB                             | b1252                  | Iron-Siderophore Transport | [152,157] |
| 1634625          | nohA                             | b1548                  | Predicted Packaging Protein (Qin Prophage) | [152,158] |
| 1735575          | purR                             | b1658                  | Hypoxanthine Transcriptional Repressor | [159,160] |
| 3150150          | exbB                             | b3006                  | Iron-Siderophore Transport | [152] |
| 3273150          | garP                             | b3127                  | Galactarate/Glucarate/Glycerate Transporter | [152,160] |
| 3538050          | feoA                             | b3408                  | Ferrous Iron Transport | [59] |

*Genomic location within each FNR ChIP-seq peak with the highest read count (the summit of the peak).

1First gene of the operon downstream of the FNR ChIP-seq peak, and operon designation was obtained from EcoCyc [70]. For peaks located within divergent promoters with both operons repressed by Fur, both genes are identified, separated by “/”.

2Identification number (B-number) for the first gene in each operon, obtained from EcoCyc [70]. For peaks located within divergent promoters with both operons repressed by Fur, both B-numbers are identified, separated by “/”.

3Functional description of the products of the operons, obtained from EcoCyc [70]. For peaks located within divergent promoters with operons repressed by Fur, the functions of both operons are listed, separated by “/”.

4Reference for Fur repression of each operon.

doi:10.1371/journal.pgen.1003565.t004

example, hemolysin E (hlyE), in agreement with previous results [62], and the anaerobic NAP Dan (ttdR) [63] showed increased expression under anaerobic conditions only. This suggests a possible role of Dan in the absence of H-NS and StpA. Only two genes showed a decrease in expression in the absence of H-NS and StpA (yncD and feaR) under only anaerobic growth conditions. However, the expression of the vast majority of genes having FNR bound at unmasked sites resulted in changes under both aerobic and anaerobic growth conditions, indicating that changes in nucleoid structure that occur in the absence of H-NS and StpA could cause misregulation of transcription. For example, H-NS and Rho coordinate to regulate transcriptional termination and the absence of H-NS may cause increased transcriptional readthrough of Rho-dependent terminators [64]. Thus, it seems likely that our analysis provides an underestimate of the impact of H-NS on FNR function, since physiological conditions that alter H-NS activity are likely to have less severe effects on nucleoid structure.

### Table 5. Operons lacking a FNR ChIP-seq peak but with a FNR-dependent change in expression in GMM that are known to be regulated through the action of the small regulatory RNA FnrS (Category 6).

| Operon* | B-number of First Geneb | Cellular Functionc | FNR Regulationd |
|---------|-------------------------|--------------------|-----------------|
| gpmA    | b0755                   | 2,3-Bisphosphoglycerate-Dependent Phosphoglycerate Mutase | Repressed |
| cydDC   | b0887                   | Glutathione/Cysteine Transporter | Repressed |
| chaA    | b1216                   | Sodium/Proton Antipporter | Repressed |
| adhP    | b1478                   | Ethanol Dehydrogenase/Alcohol Dehydrogenase | Repressed |
| sodB    | b1656                   | Iron-Containing Superoxide Dismutase | Repressed |
| yobA-yebZY | b1841             | Conserved Protein | Repressed |
| lid     | b2133                   | Lactate Dehydrogenase | Repressed |
| tolE-tolB | b2153             | GTP Cyclohydrolase I | Repressed |
| eco     | b2209                   | Ecotin Homodimer/Serine Protease Inhibitor | Repressed |
| tolX-tolC | b2303             | Dihydroneopterin Triphosphate 2-Epimerase | Repressed |
| yggG    | b2936                   | Predicted Metallopeptidase | Repressed |

*Operon showing a statistically significant FNR-dependent change in expression compared to WT but lacking a FNR ChIP-seq peak upstream. Operon definitions obtained from EcoCyc [70].

1Identification number (B-number) for the first gene in each operon, obtained from EcoCyc [70].

2Cellular function of the product of the operon, obtained from EcoCyc [70].

3FNR regulation of each operon as identified in the transcriptomic experiments performed in this study.

doi:10.1371/journal.pgen.1003565.t005
Discussion

By combining genome-wide FNR occupancy data from ChIP-seq and ChIP-chip experiments with transcriptomic data, we uncovered new features of bacterial transcriptional regulation and the FNR regulon. Our findings suggest that in vivo FNR occupies only a subset of predicted FNR binding-sites in the genome, and that FNR binding can be blocked by NAPs like H-NS. Furthermore, the lack of correlation between match to consensus of FNR binding sites and ChIP enrichment suggests that variations in ChIP signal result from changes in cross-linking efficiency or epitope access rather than variable occupancy. We found that the FNR regulon is malleable; the set of genes controlled by FNR can be readily tailored to changing growth conditions that may activate or inactivate other TFs, allowing flexible reprogramming of transcription. This strategy would allow the regulon to expand or contract depending on available nutrients, providing a competitive advantage in the ecological niche of E. coli of the mammalian gut [65].
FNR peak height does not correlate with the match to the FNR consensus site

The finding that there was little relationship between peak height and the quality of the FNR motif differs from the results found for LexA, which showed a correlation between peak height and match to consensus [10]. Our data suggest that FNR peak height may be more related to the efficiency of cross-linking or immunoprecipitation since sites that appear to be saturated for binding displayed significantly different peak heights. Thus, at least for FNR, peak height cannot be used to assess relative differences in site occupancy between chromosomal sites. Cross-linking or immunoprecipitation of FNR may be less efficient than for LexA because the larger number of other regulators bound at FNR-regulated promoters may affect accessibility to the cross-linking agent or FNR immunoprecipitation.

FNR sites having either a strong match to consensus (for example, ydfZ (TTGATaaaaAACAA) or a weak match (for example, fudA (TCGATctcgTCAAA) were saturated for binding at FNR dimer concentrations at its cellular level (~2.5 μM) [37]; thus, in vivo most accessible FNR sites are likely to be fully

Figure 7. Representative examples of FNR directly activated and repressed promoters. Representative examples of FNR directly activated (dmsA, panel A) and repressed (ndh, panel B) promoters confirmed in our study. Shown are the ChIP-seq data traces of anaerobic FNR (blue), anaerobic σ^70 (green) and aerobic σ^70 (red). ChIP-seq peak heights are represented on the y-axis (log2 kernel density [115]). The chromosomal location is shown on the x-axis with genes represented by arrows pointing in the direction of transcription. Known promoters are represented as red arrows and previously identified FNR binding sites are shown by blue brackets. The inset shows the linear expression levels for each gene from the RNA-seq experiments comparing FNR^+ and Δfnr -O_2 expression. Linear tag density [125] for each gene is shown on the y-axis.

doi:10.1371/journal.pgen.1003565.g007
occupied. These data also revealed that FNR occupancy was not significantly different for strong and weak sites over the tested range of FNR dimer concentrations, suggesting that in vivo FNR binding is unlikely to be dictated solely by the intrinsic affinity of FNR binding sites.

**Genome-wide data reveal FNR binding throughout the chromosome is influenced by other cellular factors beyond the presence of a FNR motif**

Our finding that not all predicted FNR binding sites are bound by FNR in vivo offers new insight into the accessibility of the genome for binding TFs. Previous studies have predicted anywhere from 12 to 500 FNR binding sites in the *E. coli* genome [66–69], depending on the algorithm used. Of the 187 FNR binding sites predicted here, only 63 contained a corresponding FNR ChIP-seq peak in the WT strain, suggesting many high-quality FNR sites are not bound. Although some of these silent sites may result from false negatives in the ChIP experiments (e.g. failure to immunoprecipitate FNR bound at some sites), only five of the 124 silent FNR sites (uncN, ahdA, hylA, hmp and traD) showed any evidence of FNR regulation in prior studies [70]. Rather, several lines of evidence suggest that binding of NAPs or other TFs masks FNR binding at many of these sites in vivo. First, we observed that binding sites for the NAPs IHF, H-NS, and Fis were statistically overrepresented at the positions of silent FNR binding sites, suggesting these proteins occlude FNR binding. Second, we found that in the absence of H-NS and StpA, additional FNR binding sites became available for FNR binding as detected by ChIP, suggesting that NAPs influence FNR site availability in vivo.

A similar effect has been observed in eukaryotes, where extensive research on TF site availability has shown that chromatin structure in vivo can block binding of TFs (e.g. Pho4, Leu3 and Rap1) to high quality DNA binding sites [1,2,4]. Additionally, known changes to chromosomal structure by IHF, Fis, and H-NS have been shown to inhibit DNA binding of other proteins [7,12,71]. Thus, if the binding profiles of NAPs change under alternative growth conditions, then the occluded FNR binding sites would likely become available for FNR binding.

Nonetheless, the fact that the 207 FNR-enriched regions from this study included 80% of the 63 regions identified by Grainger et al. (Table S5), despite the difference in the growth conditions and experimental design [29], suggests that the overlapping subset of FNR binding events may reflect a core set that is insensitive to growth conditions or binding of other TFs. Furthermore, binding events specific to each growth condition may be reflective of either changes in accessibility of FNR to binding sites due to changes in DNA-binding protein distribution or perhaps increases in activity of a second TF that binds cooperatively. Other regulators, such as CRP, a closely related member of the FNR protein family, also appear to have more binding sites available genome-wide than are occupied in vivo under tested growth conditions. Shimada et al. identified 254 CRP-cAMP binding sites using Genomic SELEX screening, which was 3–4 fold more than the number of CRP sites previously identified by ChIP-chip experiments [72,73]; thus not all chromosomal CRP sites appear to be accessible for binding, although additional experiments would be required to explicitly examine the accessibility of CRP binding sites throughout the genome. Taken together, these results suggest that the restrictive effect of chromosomal structure could influence TF binding beyond FNR.

Environmental stimuli that change NAP distribution would also change TF binding site accessibility and affect transcription. For example, as *E. coli* enters the mammalian GI tract, it experiences a temperature increase from ~25°C to 37°C, and this increase in temperature has been shown to affect transcription of a number of operons, including increased expression of anaerobic-specific operons [74,75]. Because H-NS binding is sensitive to changes in temperature [76,77], an explanation for these temperature-dependent transcriptional changes [74,75] could be genome-wide decreases in H-NS binding and distribution; these changes could increase the accessibility of the binding sites for FNR and other TFs to regulate transcription. Supporting this explanation, several genes with a temperature-dependent increase in expression showed FNR binding and regulation in the absence of H-NS and StpA, including *hlyE, fedR, yueV*, and *torG.* The activity of NAPs can also be affected by the binding of other condition-specific TFs. For example, ChIP-chip and Genomic SELEX analysis of the stationary phase LysR-type TF, LeuO, suggested that binding of LeuO antagonized H-NS activity, but not necessarily H-NS binding, throughout the genome in *Salmonella enterica* and *E. coli* [78,79].

Thus, a picture emerges from our data that binding of FNR is dependent on characteristics of the genome beyond the presence of a FNR binding site; this restrictive effect of chromosome structure by NAPs may affect binding of other TFs in bacteria. NAPs have been shown to occlude and affect binding of TFs and other DNA binding proteins, such as restriction endonucleases and DNA methylation enzymes, suggesting a general role of NAPs in regulating genome accessibility by bending, wrapping and bridging the DNA structure [7,12,13,27,42,76,80,81]. Additionally, NAPs influence DNA supercoiling, which has been shown to affect binding of the TFs Fis and OmpR in *S. enterica* [82,83], providing another mechanism by which NAPs can change the chromosomal structure to influence TF-DNA binding. Taken together, our results support a dynamic model of complex genome structure that affects TF binding to control gene regulation in bacteria.

**Condition-specific expression of the FNR regulon likely requires other transcription factors**

Although expression of a subset of the operons in the FNR regulon appeared to require only FNR for regulation (Categories 1 and 2), our findings point to widespread cooperation between FNR and other TFs for condition-specific regulation (Categories 3 and 4). Changes in activity of these TFs would result in FNR regulation to adapt to changes in environment, such as growth in non-catabolite repressed carbon sources (CRP) [57], anaerobic respiration of nitrate (NarL and NarP) [19], and growth in iron-limiting conditions (Fur) [84]. Although this co-regulation provides insight into growth conditions that should allow FNR-dependent changes in gene expression, the synergistic regulators for many promoter regions bound by FNR are currently unknown (Category 5), but would likely be identified in future genome-scale studies using different growth conditions, particularly microaerobic growth, which has been shown to affect FNR regulation of virulence genes in the pathogen *Shigella flexneri* [85].

Overall, our results suggest that the regulation of a subset of FNR-dependent promoters in *E. coli* may depend on combinatorial regulation with other TFs, a mechanism that resembles regulation of eukaryotic promoters [8,20,86]. These experimental data support previous *in silico* regulatory models generated using published data [87–89], suggesting combinatorial regulation may be common in *E. coli*. Further, ChIP-chip and ChIP-seq analyses of other TFs in *E. coli* (e.g. CRP, Fis, and IHF) and *Salmonella typhimurium* (e.g. Sfh, a H-NS homolog), identified many TF binding sites that did not correlate with changes in gene expression in corresponding TF-specific transcriptomic experiments [30,33,73,90,91]. These results raise the possibility of potential
combinatorial regulation for other TFs, although additional analysis is required to support this notion.

The indirect FNR regulon also involves other regulators

We found that FNR directly controls expression of five secondary regulators, most of which are also regulated by specific cofactors, suggesting that the scope of the indirect FNR regulon (Category 6) is also likely to change depending on growth conditions. Of the five regulators, three act in an apparent hierarchical manner. The small RNA FnrS, which is upregulated by FNR and is suggested to stimulate mRNA turnover, decreased the mRNA levels of multiple FnrS target genes in GMM [48,49]. Expression of the TF CaiF was also activated by FNR, but the genes regulated by CaiF were not expressed in GMM because CaiF requires the effector carnitine to be active [92]. FNR activated BssR, a TF involved in biofilm formation. About ~40 operons are thought to be controlled by BssR [93], but none of the five BssR-dependent operons in the FNR indirect regulon that we tested by qRT-PCR showed any change in expression in a BssR- strain (data not shown); thus, under our growth conditions, BssR appeared to be inactive.

FNR also directly repressed the expression of two TFs, including the pyruvate sensing TF PdhR which represses several operons in the absence of pyruvate [94,95]. Although one might expect that PdhR repressed genes would increase anaerobically, many of these genes are redundantly repressed by ArcA (Park and Kiley, Personal Communication); thus the impact of PdhR may be negligible under anaerobic growth in GMM. Similarly, the TF GadE, which is active at low pH [96], was also directly repressed by FNR and accordingly the operons in the GadE regulon were not identified as part of the indirect FNR regulon in GMM.

Finally, we note the caveat that some operons that appear indirectly regulated by FNR may change expression as a result of indirect physiological and metabolic effects in an FNR- strain, which may alter the activity of other TFs, resulting in misregulation of operons. For example, our data show that FNR does not directly regulate arcA transcription, but previous results have suggested that ArcA activity may be affected by the metabolic changes that occur when fur is deleted [97]. Thus, although a subset of ArcA regulatory targets (29 operons) showed potential indirect FNR regulation, such effects were likely caused by changes in the phosphorylation state of ArcA resulting from metabolic changes in an FNR- strain (Table S13) (Park and Kiley, Personal Communication).

In conclusion, our results reveal complex features of TF binding in bacteria and expand our understanding of how E. coli responds to changes in O2 and other environmental stimuli.

A subset of predicted FNR binding sites appear to be inhibited by NAPs and are available in the absence of H-NS and StpA, suggesting that the bacterial genome is not freely accessible for TF binding and that changes in TF binding site accessibility could result in changes in transcription. Finally, correlation of the occupancy data with transcriptomic data suggests that FNR serves as a global signal of anaerobiosis but the expression of a subset of operons in the FNR regulon requires other regulators sensitive to alternative environmental stimuli. This strategy is reminiscent of global regulation by CRP-cAMP [73] in that FNR, like CRP, is bound at many promoters under specific conditions without corresponding changes in mRNA levels, suggesting a common strategy whereby promoters are primed to be activated when the appropriate growth conditions are encountered.

Materials and Methods

Strains and growth conditions

All strains were grown in MOPS minimal medium supplemented with 0.2% glucose (GMM) [98] at 37°C and sparged with a gas mix of 95% N2 and 5% CO2 (anaerobic) or 70% N2, 5% CO2, and 25% O2 (aerobic). Cells were harvested during mid-log growth (OD600 of ~0.3 using a Perkin Elmer Lambda 25 UV/Vis Spectrophotometer). E. coli K-12 MG1655 (F-, Δ-λchp-1) and PK4811 (MG1655 ΔfurΔsfrRΔSmpR) [99] were used for the ChiP-chip, ChiP-seq and transcriptomic experiments unless otherwise specified. All data obtained in this study used GMM as the growth media, and although we know that not all promoters directly regulated by FNR are expressed under these conditions, this has the advantage that both mutant and parental strains exhibit the same growth rate.

For experiments that varied the in vivo concentration of FNR, a strain that contained a single, chromosomal copy of WT fur under the control of the Ptau promoter at the lac attachment site was constructed. Following digestion of pPK823 [99] with XbaI and HindIII, the DNA fragment containing fur was cloned into the XbaI and HindIII sites of pDH660 (ApR) [100] to form pPK6401. Plasmid pPK6401 was transformed into DHB6521 [100] and the Ptau-fur construct was stably integrated into the lac attachment site using the Lambda InCh system as described [100] to produce PK6410. Ptau transduction was used to move the Ptau-fur, ApR allele into strain PK8257, which contains the FNR activated ydfZ promoter-lacZ fusion and deletion of lacI. This strain was transformed with pACYClatF-λ-CAM [101] to generate PK8263. To determine the effect of FNR on the expression of the BssR regulon, a ΔbssR strain was constructed by P1vir transduction of ΔbssR::kanR from the Keio collection [102] into MG1655 to generate PK8923. To determine the role of H-NS on FNR binding, first stpA was recombined with the CmR gene, erm, using λ red recombination and the pSIM plasmid [103]. P1vir transduction introduced the Δhns::kanR allele from the Keio collection [102] into the strain lacking stpA to generate the Δhns/ΔstpA strain.

RNA isolation

Total RNA was isolated as previously described [104]. The concentration of the purified RNA was determined using a NanoDrop 2100, while the integrity of the RNA was analyzed using an Agilent 2100 Bioanalyzer and the RNA Nano LabChip platform (Agilent).

Whole genome transcriptomic microarray analysis

Total RNA (10 μg) from two biological replicates each of MG1655 (Δoxy and Δo2) and PK4811 was reverse transcribed using random hexamers (Sigma) and the SuperScript II Double-Stranded cDNA Synthesis Kit (Invitrogen) following the manufacturer’s protocol. The cDNA (1 μg) was fluorescently labeled with Cy3-labeled 9 mers (Tri-Link Biotechnologies) with Klenow Fragment (NEB) for 2 hours at 37°C and recovered using ethanol precipitation. Labeled dsDNA (2 μg) was hybridized onto the Roche NimbleGen E. coli 4plex Expression Array Platform (4×72,000 probes, Catalog Number A6697-00-01) for ~16 hours at 42°C in a NimbleGen Hybridization System 4 (Roche NimbleGen) following the manufacturer’s protocol. The hybridized microarrays were scanned at 532 nm with a pixel size of 5 μm using a GenePix 4000B Microarray Scanner (Molecular Devices), and the PMT was adjusted until approximately 1% of the total probes were saturated for fluorescence intensity. The data were
normalized using the Robust Multichip Average (RMA) algorithm in the NimbleScan software package, version 2.5 [105]. ArrayStar 3.0 (DNASTAR) was used to identify genes that showed at least a two-fold change in expression between the WT and Δfur strains and were significantly similar among biological replicates, using a moderated t-test (p-value<0.01) [106]. Genes were organized into operons using data from EcoCyc [70]. An operon was called differentially expressed (DE) if only one gene within an operon showed a statistically significant change in expression. NimbleGen microarrays identified 214 statistically significant DE genes that were contained within 130 operons.

The anaerobic MG1655 and FNR- samples from the normalized whole genome expression microarray data from Kang et al. [18] were also analyzed. Genes were determined to be DE if they had a change in expression greater than or equal to two-fold and if the genes were found to be statistically similar between biological replicates using a t-test (p-value<0.01). An operon was called DE if only one gene within an operon showed a statistically significant change in expression. This analysis identified 204 significant DE genes in 130 operons. Sixty operons were found to be DE in both the NimbleGen and Kang et al. data sets (Table S8). Of the 70 operons found DE in only the Kang et al. data set, 41 operons were just below the significance threshold in the NimbleGen data set and 11 operons resulted from activation of the flagellar regulon due to an insertion upstream of fliDC, which was absent in the isolate of MG1655 used in this study.

The Δfur/ΔfisA aerobic and anaerobic expression data were obtained from stand specific, single stranded cDNA hybridized to custom designed, high-density tiled microarrays containing 378,000 probes from alternate strands, spaced every 12 bp through the genome as described previously [107] except Cy3 was used instead of Cy5. Microarray hybridization and scanning were performed as described above except that the PMT was adjusted until approximately 1% of the total probes were saturated for fluorescence intensity of each dye used. The PMT was adjusted until approximately 1% of the total probes were saturated for fluorescence intensity of each dye used. The NimbleGen Hybridization System 4 (Roche NimbleGen) was used to extract the scanned data. ChIP-chip data were normalized within each microarray using quantile normalization [108]. All probe data were normalized using the software package SOAP, version 2.20 [109], allowing no more than two mismatches. Reads aligning to repeated elements in the genome (for example rRNA) were removed from analysis. For reads that had no mapping locations for the first 36 bp, the 3–30 bp subsequences were used in the subsequent mapping to the reference genome. Reads that had unique mapping locations and did not match annotated rRNA genes were used for further analysis. For each gene, the tag density was estimated as the number of aligned sequencing tags divided by gene size in kb and normalized using quantile normalization. The tag density data were analyzed for statistically significant differential expression using baySeq, version 2.6 [109] with a FDR of 0.01, and genes were organized into operons using data from EcoCyc [70]. An operon was called DE if only one gene within an operon showed a statistically significant change in expression. The RNA-seq analysis identified 133 statistically significant DE operons (197 genes). Altogether, microarray and RNA-seq experiments identified 258 operons DE by FNR and slightly fewer than half of these operons (122) were found in at least two of the transcriptomic experiments (Figure S5, Table S9).

**Chromatin immunoprecipitation followed by hybridization to a microarray chip or high-throughput sequencing**

ChiP assays were performed as previously described [110], except that the glycine, the formaldehyde and the sodium phosphate mix were sparged with argon gas for 20 minutes before use to maintain anaerobic conditions when required. Samples were immunoprecipitated using polyclonal antibodies raised against FNR, IHF or H-NS, which had been individually absorbed against mutant strains lacking the appropriate protein. In the case of FNR, affinity purified antibodies were used in some experiments, purified using the method previously described [111]. For RNA Polymerase, a σ70 monoclonal antibody from NeoClone (W0004) or a RNA Polymerase β monoclonal antibody from NeoClone (W0002) were used for immunoprecipitation. For FNR, neither lengthening the cross-linking time nor increasing or decreasing the amount of FNR antibody used in the ChiP protocol showed significant changes in the FNR ChiP-chip peak heights or number of peaks identified. For ChiP-chip, FNR (three samples), FNR- (one sample), β (two samples), H-NS (two samples) and IHF (two samples) were fluorescently-labeled using Cy3 (INPUT) and Cy5 (IP) and hybridized for ~16 hours at 42°C in a NimbleGen Hybridization System 4 (Roche NimbleGen) to custom designed, high-density tiled microarrays containing 378,000 probes from alternate strands, spaced every ~12 bp through the genome. The hybridized microarrays were scanned at 532 nm (Cy3) and 635 nm (Cy5) with a pixel size of 5 μm using a GenePix 4000B Microarray Scanner (Molecular Devices), and the PMT was adjusted until approximately 1% of the total probes were saturated for fluorescence intensity of each dye used. The NimbleScan software package, version 2.5 (Roche NimbleGen) was used to extract the scanned data. ChiP-chip data were normalized within each microarray using quantile normalization (“normalize.quantiles” in the R package VSN, version 3.26.0) [112] to correct for dye-dependent intensity differences as previously described [113]. Biological replicates were normalized between microarrays using quantile normalization as previously described [113], and the normalized log2 ratio values (IP over INPUT) were averaged. There was a strong correlation between enriched regions of ChiP-chip biological replicates (R = 0.7). ChiP-chip peaks for FNR, H-NS and IHF were identified in each data set by the peak finding algorithm CMARRT, version 1.3.
(FDR of 0.01) [114] and proportional Z-tests were used to determine significant differences between proportional data.

For ChIP-seq experiments, 10 ng of immunoprecipitated and purified DNA fragments from the FNR (two biological replicates) and σ70 samples (two biological replicates from both aerobic and anaerobic growth conditions), with 10 ng of input control, were submitted to the University of Wisconsin-Madison DNA Sequencing Facility (FNR samples and one σ70 sample) or the Joint Genome Institute (one σ70 sample) for ChIP-seq library preparation. Samples were sheared to 200–500 nt during the IP process to facilitate library preparation. All libraries were generated using reagents from the Illumina Paired End Sample Preparation Kit (Illumina) and the Illumina protocol “Preparing Samples for ChIP Sequencing of DNA” (Illumina part #11257047 RevA) as per the manufacturer’s instructions, except products of the ligation reaction were purified by gel electrophoresis using 2% SizeSelect agarose gels (Invitrogen) targeting either 275 bp fragments (σ70 libraries) or 400 bp fragments (FNR libraries). After library construction and amplification, quality and quantity were assessed using an Agilent DNA 1000 series chip assay (Agilent) and QuantIT PicoGreen dsDNA Kit (Invitrogen), respectively, and libraries were standardized to 10 μM. Cluster generation was performed using a cBot Single Read Cluster Generation Kit (v4) and placed on the Illumina cBot. A single-end read, 36 bp run was performed, using standard SBS kits (v4) and SCS 2.6 on an Illumina Genome Analyzer Ix. Basecalling was performed using the standard Illumina Pipeline, version 1.6. Sequence reads were aligned to the published *E. coli* K-12 MG1655 genome (U00096.2) using the software packages SOAP, version 2.20, [108] and ELAND (within the Illumina Genome Analyzer Pipeline Software, version 1.6), allowing at most two mismatches. Sequence reads with sequences that did not align to the genome, aligned to multiple locations on the genome, or contained more than two mismatches were discarded from further analysis (<10% of reads). For visualization the raw tag density at each position was calculated using QuEST, version 1.2 [115], and normalized as tag density per million uniquely mapped reads. The read density was determined for each base in the genome for the IP and INPUT samples for FNR and σ70 samples. For FNR, peaks were identified using three peak finding algorithms: CisGenome, version 1.2, NCIS, version 1.0.1, and MOSAICS, version 1.6.0 [116–118] (FDR for all of 0.05), while σ70 peaks were identified using NCIS, version 1.0.1 (FDR of 0.05). Further discussion of these algorithms is in Text S1. Differences between aerobic and anaerobic σ70 ChIP-seq occupancy were determined using a one-sided, paired t-test (p-value < 0.01) comparing 100 bp surrounding the center of each peak. To normalize between the background (un-enriched) signal. The p-values were adjusted using the Bonferroni method to correct for multiple testing. There was a strong correlation between ChIP-seq biological replicates (R = 0.8) as well as between ChIP-chip and ChIP-seq data (Figure S6). All data were visualized in the MochiView browser [119].

Additional ChIP-chip -O2 data sets were performed for WT FNR and a Δfnr [99] control. The 15 FNR peaks identified only in ChIP-chip had low IP/INPUT ratios and were eliminated since ChIP-seq is known to have increased signal to noise relative to ChIP-chip [120]. The Δfnr -O2 ChIP-chip data identified 71 peaks that corresponded to peaks in the FNR -O2 ChIP-seq data, indicating they were not FNR specific, and were removed from the FNR ChIP-seq dataset (Table S5).

**FNR PWM construction and identification of predicted FNR binding sites at FNR ChIP-seq peaks**

To construct the FNR PWM, the sequence of a region of ~100 bp around the nucleotide with the largest tag density within each of the FNR ChIP-seq peaks (the summit of each peak) found by all three peak finding algorithms was analyzed. MEME was used to identify over-represented sequences [121] and the Delila software package was used to construct the PWMs [122]. To search all ChIP-seq peaks for the presence of the FNR PWM, a region of 200 bp around the summit of each FNR ChIP-seq peak was searched with the FNR PWM using PatSer, version 3e [36], and the top four matches to the FNR PWM, as determined by PatSer PWM score, were recorded at each ChIP-seq peak. The standard deviation of the PatSer scores for the four FNR predicted binding sites at each ChIP-seq peak was determined and used as a threshold to determine the number of predicted binding sites at each peak. If the PatSer predicted FNR binding site at a peak with the highest PatSer score was more than one standard deviation greater than the PatSer predicted FNR binding site with the second best PatSer score, that peak was identified as having only one predicted FNR binding site. For FNR peaks (~11%) with the two best PatSer predicted FNR binding site scores less than one standard deviation apart, a Grubbs test for outliers was used to identify outliers within the four PatSer predicted FNR binding sites at a peak (δ of 0.15, critical Z of 1.04). If a PatSer predicted FNR binding site at a FNR peak was identified as an outlier, it was removed from analysis and the standard deviation was re-calculated using the remaining three PatSer binding site scores at that peak. The remaining PatSer predicted FNR binding sites at the FNR peak were then re-examined as described above. After removing outlier PatSer predicted FNR binding sites, a peak was determined to contain two predicted FNR binding sites if the two best predicted FNR binding sites at a peak had PatSer scores less than one standard deviation apart.

The precision-recall curve was constructed using the FNR PWM and searching throughout the genome using PatSer, version 5e [36]. Precision was defined as True Positives (locations with a FNR ChIP-seq peak and a predicted FNR binding site) divided by True Positives plus False Positives (locations with a predicted FNR binding site but no FNR ChIP-seq peak). Recall was defined as True Positives divided by True Positives plus False Negatives (locations with a FNR ChIP-seq peak but no FNR predicted binding site). A high precision value means all true positives have been captured, but there is a high false negative rate. A high recall value means all true positives have been captured, but there is a high false positive rate.

**Controlling expression of fnr with an IPTG-inducible promoter and performing ChIP-chip and analysis**

The strain with *fnr* under the control of *P*<sub>pol</sub> (PK6263) was used to study changes in [FNR] on ChIP-chip peak height. Cultures were grown anaerobically overnight in MOPS+0.2% glucose and were subcultured to a starting OD<sub>600</sub> of ~0.01 in MOPS+0.2% glucose plus Cm20 and various [IPTG] (4 μM IPTG, 8 μM IPTG, and 16 μM IPTG). After this initial step, growth, ChIP-chip experiments (two biological replicates of 4 and 8 μM IPTG) and three biological replicates of 16 μM IPTG were used) and initial analysis were identical to the procedures described above. Estimates of FNR concentration were determined by quantitative Western blot as previously described [37]. A novel method of normalization was developed to compare peak areas between IPTG concentrations for 35 peaks that showed a large distribution in peak heights and 4 peaks that were classified as false positives by
enrichment in the Δfnr Chip-chip sample. The peak finding algorithm CMARRT identified peaks in the WT FNR Chip-chip sample, and this peak region was trimmed to include the center 50% of the peak region. This trimmed region was used for each [IPTG] sample for consistency. For each of the 39 peaks examined, the probe values in a region of ~3000 bp beyond the peak boundary (~1500 bp upstream and downstream of the peak boundary) was selected for analysis from each sample. Within the ~3000 bp region, the probes beyond the peak boundary were considered background for each sample. The median of the background (un-enriched) probes was calculated and the log2 IP/INPUT probe values for the entire peak region (enriched and un-enriched) were shifted by the negative median value of the background probes. The peak average (average of log2 IP/INPUT values) and standard deviation was determined for 39 peak regions to compare between samples at each [IPTG] and WT Chip-chip samples. A one-sided, paired t-test was performed between all conditions (p-value<0.05) to determine statistically significant changes in average peak values.

Comparing FNR enrichment in WT and Δhns/ΔslyA genetic backgrounds using Chip-chip analysis

Growth, Chip-chip experiments, normalization and peak calling was performed as described above. To normalize between WT and Δhns/ΔslyA samples, the enriched regions (peaks) for each sample were shifted by the negative median log2 IP/INPUT value of the background (un-enriched) probes. The peak averages (average of log2 IP/INPUT values) were determined for each condition (WT and Δhns/ΔslyA) at each FNR peak found in either strain background. A one-sided, paired t-test with Bonferroni correction was performed between the two conditions (p-value<0.05) to determine the statistically significant change in peak averages. For peaks found in both WT and Δhns/ΔslyA, peaks were identified as significantly higher in Δhns/ΔslyA using a one-sided, paired t-test with Bonferroni correction performed between the two conditions (p-value<0.05) and if the FNR peak average in the Δhns/ΔslyA strain was greater than the standard deviation found for WT peak average.

Data deposition and visualization

The Chip-chip and Chip-seq data can be visualized on GBrowse at the following address: “http://heptamer.tamu.edu/cgi-bin/gb2/gbrowse/MG1655/”. All genome-wide data from this publication have been deposited in NCBI's Gene Expression Omnibus (GSE41195) (Table S14) [123].

Supporting Information

Figure S1 Comparison of the average IP/INPUT ratio for 35 FNR Chip-chip peaks over a range of FNR concentrations. Comparison of the average Chip-chip peak height for FNR in WT cultures (open symbols) (~2.5 μM FNR) and PK8263 (P tac fnr) cultures (closed symbols) at three [IPTG] concentrations: 4 μM IPTG (~450 nM FNR), 8 μM IPTG (~700 nM FNR), 16 μM IPTG (~1.9 μM FNR), [FNR] determined by quantitative Western samples, the enriched regions (peaks) for each sample were shifted by the negative median log2 IP/INPUT value of the background probes. The peak average (average of log2 IP/INPUT values) and standard deviation was determined for 39 peak regions to compare between samples at each [IPTG] and WT Chip-chip samples. A one-sided, paired t-test was performed between all conditions (p-value<0.05) to determine statistically significant changes in average peak values.

Figure S2 Overlap of NAP enrichment at silent FNR sites. Venn diagram showing the overlap of detected enrichment for Fis (red), N-NS (blue) and IHF (yellow) at silent FNR sites (Table S6).

Figure S3 Computational analysis of H-NS binding in the genome and at silent FNR sites. A) Histograms showing the distribution of log2 IP/INPUT values for every bp of β RNAP (blue) and H-NS (red) Chip-chip data (+O2) within aerobic H-NS binding regions over 1000 bp in length. The H-NS mean value is statistically greater than the β RNAP mean value (p-value<0.05). B) Histograms showing the distribution of log2 IP/INPUT values for every bp of β RNAP (blue) and H-NS (red) Chip-chip data (−O2) within anaerobic H-NS binding regions over 1000 bp in length. The H-NS mean value is statistically greater than the β RNAP mean value (p-value<0.05). C) A representative region of the E. coli genome showing a predicted FNR binding site (blue line) within fimC that is enriched in H-NS Chip-chip data (purple trace) and lacking a FNR Chip-seq peak (blue trace). The x-axis shows the genomic coordinates while the y-axis shows the Chip-chip peak height (log2 IP/INPUT value). Genes are represented by arrows pointing in the direction of transcription. D) Distribution of H-NS binding region lengths at silent FNR sites in anaerobic GMM. E) Distribution of H-NS binding region lengths throughout the genome in anaerobic GMM.

Figure S4 Correlation of FNR occupancy data and FNR transcriptomic data. Venn diagram showing the overlap of operons identified in the genome-wide data sets used in this study. The green overlap represents those operons directly dependent on FNR binding for expression (Categories 1 and 2). The yellow section represents operons with a FNR Chip-seq peak upstream but lacking a FNR-dependent change in expression (Categories 3, 4 and 5). The blue section represents the DE operons found in at least two of the three transcriptomic data sets used in this study but lack a corresponding FNR Chip-seq peak upstream (Category 6).

Figure S5 Overlap between the operons found differentially expressed by FNR. Venn diagram showing the overlap among the DE operons in the three transcriptomic data sets used in this study. Designations are: ‘RNA-seq Expression’ (operons found differentially expressed between WT and Δfnr in the RNA-seq experiment); ‘Microarray Expression - A’ (operons found differentially expressed between WT and Δfnr in the expression microarray experiments performed in this study); ‘Microarray Expression - B’ (operons found differentially expressed between WT and Δfnr from reanalysis of the Kang et al. data [18]).

Figure S6 Correlation of Chip-chip and Chip-seq enrichment levels for FNR. Shown is the correlation between Chip-chip peak height on the y-axis (highest IP/INPUT value of enriched region) and Chip-seq peak height on the x-axis (highest log2 read count value of enriched region) for each peak found in both the Chip-chip and Chip-seq data. The line indicates the linear correlation best-fit (R2 value of 0.8).

Table S1 Peaks identified using NCIS from σ70 Chip-seq data from cultures grown under aerobic and anaerobic growth conditions in GMM.
Table S2  Peaks identified to have statistically significant change in occupancy from 647 ChIP-seq data from cultures grown under anaerobic and aerobic growth conditions.

Table S3  Peaks found in H-NS ChIP-chip data from aerobic and anaerobic growth conditions in GMM.

Table S4  Peaks found in IHF ChIP-chip data from anaerobic growth conditions in GMM.

Table S5  FNR ChIP-chip and ChIP-seq peaks identified in this study.

Table S6  FNR predicted binding sites identified throughout the E. coli genome.

Table S7  FNR ChIP-chip peaks identified in the Δhup/ΔatpA strain.

Table S8  Operons found to be differentially expressed when comparing transcriptomic data from WT and Δfnr strains.

Table S9  FNR ChIP-seq peaks upstream of operons with no corresponding FNR-dependent change in expression (Categories 3, 4 and 5).

Table S10  Operons found to be differentially expressed when comparing transcriptomic data from WT and Δfer strains but lacking a corresponding FNR ChIP-seq peak upstream (Category 6).

Table S11  Genes with an O2 dependent change in expression only in the Δhup/ΔatpA strain.

Table S12  Operons found to be differentially expressed when comparing WT aerobic and anaerobic transcriptomic data.

Table S13  Operons regulated by ArcA that show a FNR-dependent change in expression but lack a FNR ChIP-seq peak.

Table S14  Source of all genome-scale data analyzed in this study.

Text S1  File containing supporting methods and references for the information found in the supporting tables.

Acknowledgments

We thank Gary Stormo for advice about PatSer and PWM searching, Erin Mettert and Erik Jensen for strain construction, Sarah Teter for affinity purified FNR antibody, and Nicole Beauchene and Dan Park for access to unpublished data. We also thank members of the Landick and Donohue lab for advice on data analysis and members of the Kiley lab for comments on the manuscript.

Author Contributions

Conceived and designed the experiments: KSM RL PJK. Performed the experiments: KSM FT. Analyzed the data: HY IMO DC KL SK KSM. Contributed reagents/materials/analysis tools: DC KL SK. Wrote the paper: KSM RL PJK.

References

1. Lièb JD, Liu X, Botstein D, Brown PO (2001) Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. Nat Genet 20: 327–334.
2. Liu X, Lee C-K, Granek JA, Clarke ND, Lièb JD (2006) Whole-genome comparison of LexA binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. Genome Res 16: 1517–1526.
3. Farnham PJ (2009) Insights from genomic profiling of transcription factors. Nat Rev Genet 10: 605–616.
4. Zhou X, O’Shea EK (2011) Integrated approaches reveal determinants of genome-wide binding and function of the transcription factor Pho4. Mol Cell 42: 826–836.
5. Gerstein MB, Kundaje A, Hariharan M, Landt SG, Yan K-K, et al. (2012) Architecture of the human regulatory network derived from ENCODE data. Nature 489: 91–100.
6. Lee DJ, Minchin SD, Busby SJW (2012) Activating transcription in bacteria. Annu Rev Microbiol 66: 125–152.
7. Browning DF, Grainger DC, Busby SJ (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. Curr Opin Microbiol 13: 773–780.
8. Struhl K (1999) Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell 98: 1–4.
9. Wade JT, Struhl K, Busby SJW, Grainger DC (2007) Genomic analysis of protein-DNA interactions in bacteria: insights into transcription and chromosome organization. Mol Microbiol 65: 21–26.
10. Wade JT, Reppas NB, Church GM, Struhl K (2005) Genomic analysis of LexA binding reveals the permissive nature of the Escherichia coli genome and identifies unconventional target sites. Genes Dev 19: 2619–2630.
11. Macvanin M, Adhya S (2012) Architectural organization in E. coli nucleoid. Biochim Biophys Acta 1819: 830–835.
12. Rimsky S, Travers A (2011) Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. Curr Opin Microbiol 14: 136–141.
13. Dillon SC, Dorman CJ (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat Rev Microbiol 8: 185–195.
14. Spira S (1994) The FNR family of transcriptional regulators. Antonie Van Leeuwenhoek 66: 23–36.
15. Green J, Crack JG, Thomson AJ, LeBrun NE (2009) Bacterial sensors of oxygen. Curr Opin Microbiol 12: 145–151.
16. Fleischhacker AS, Kiley PJ (2011) Iron-containing transcription factors and their roles as sensors. Curr Opin Chem Biol 15: 335–341.
17. Salmon K, Hung S-P, Mekjian K, Baldi P, Hatfield GW, et al. (2003) Global gene expression profiling in Escherichia coli K12. The effects of oxygen availability and FNR. J Biol Chem 278: 29837–29855.
18. Kang Y, Weber KD, Quy Y, Kiley PJ, Blattner FR (2005) Genome-wide expression analysis indicates that FNR of Escherichia coli K-12 regulates a large number of genes of unknown function. J Bacteriol 187: 1135–1146.
19. Constantinidou C, Hobman JL, Griffiths L, Patel MD, Penn CW, et al. (2006) A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL and NarQP as Escherichia coli K12 adapts from aerobic to anaerobic growth. J Biol Chem 281: 4892–4895.
20. Barnard A, Wolfe A, Busby S (2004) Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes. Curr Opin Microbiol 7: 102–108.
21. Browning DF, Busby SJ (2000) The regulation of bacterial transcription initiation. Nat Rev Micro 2: 57–65.
22. Lunetto MA, Rhodius V, Lamberg K, Kiley P, Busby S, et al. (1998) Identification of a contact site for different transcription activators in region 4 of the Escherichia coli RNA polymerase σ70 subunit. J Mol Biol 284: 1353–1365.
23. Stewart V (1982) Requirement of Fnr and NarL functions for nitrate reductase expression in Escherichia coli K-12. J Bacteriol 151: 1320–1325.
24. Schroder I, Darie S, GunSalus RA (1993) Activation of the Escherichia coli nitrate reductase (narGHIJ) operon by NarL and Fnr requires integration host factor. J Biol Chem 268: 771–774.
25. Lamberg KE, Kiley PJ (2000) FNR-dependent activation of the class II dnaC and narG promoters of Escherichia coli requires FNR-activating regions 1 and 3. Mol Microbiol 38: 817–827.
26. Pearson SM, Alberts JA, GunSalus RP (2002) Oxygen and nitrate-dependent regulation of dnaABC operon expression in Escherichia coli sites for Nar and NarL protein interactions. BMC Microbiol 2: 13.
27. Browning DF, Cole JA, Busby SJW (2000) Suppression of FNR-dependent transcription activation at the Escherichia coli nar promoter by Fin, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. Mol Microbiol 37: 1250–1269.
28. Kornberg RD (1999) Eukaryotic transcriptional control. Trends Cell Biol 9: M46–M49.
50. Cotter PA, Melville SB, Albrecht JA, Gunsalus RP (1997) Aerobic regulation of nuclear proteins with coding and non-coding segments of the Escherichia coli genome. Nucleic Acids Res 34: 4648–4652.

51. Samuel Raj V, Füll C, Yoshida M, Sakata K, Kashiwagi K, et al. (2002) Differential binding profiles of StpA in wild-type and ΔSbtA mutant cells: a comparative analysis of cooperative binding by chromatin immunoprecipitation microarray analysis. J Bacteriol 184: 2380–2391.

52. Melville SB, Gunsalus RP (1996) Isolation of an oxygen-sensitive FNR protein from Escherichia coli. J Bacteriol 178: 4235–4238.

53. Lutz S, Bohme R, Albrecht JA, Gunsalus RP (1997) Aerobic regulation of the Escherichia coli histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res 3: 141–153.

54. Boysen A, Møller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M (2010) Translational regulation of gene expression by an aerobically induced small non-coding RNA in Escherichia coli. J Biol Chem 285: 10980–10972.

55. Darwin AJ, Ziegelhoffer EC, Kiley PJ, Stewart V (1998) Fnr, NarP, and NarL are involved in repression and activation. Mol Microbiol 25: 605–615.

56. Lang B, BreuertT, Ebert M, Geertz M, et al. (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. Nucleic Acids Res 35: 6303–6337.

57. Jennings MP, Beacham IR (1993) Co-dependent positive regulation of the σ54 promoter of Escherichia coli by CRP and the FNR protein: a molecular analysis. Mol Microbiol 9: 153–164.

58. Fabich AJ, Leatham MP, Grissom JE, Wiley G, Lai H, et al. (2011) Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. Nucleic Acids Res 39: 2073–2091.

59. Cotter DC, Hurd D, Goldberg MD, Busby SJW (2007) Integration of three signals at the fnr gene promoter: a role for Fis in catalytic repression. Mol Microbiol 57: 496–510.

60. Boysen A, Møller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M (2010) Transcriptional regulatory sites in the complete genome sequence of Escherichia coli K-12. Bioinformatics 19: 4–40.

61. Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, et al. (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS ONE 4: e5289.

62. Green J, Baldwin ML (1997) The molecular basis for the differential regulation of the hly-encoded haemolysin of Escherichia coli by FNR and HbS lies in the improved activating region 1 contact of HbS. Microbiology 143 (Pt 12): 3715–3725.

63. Teramoto J, Yoshimura SH, Takeyasu K, Ishihama A (2010) A novel nucleic acid promoter of Escherichia coli induced under anaerobic growth conditions. Nucleic Acids Res 38: 3605–3618.

64. Peters JM, Mooney RA, Grass JA, Jessen ED, Tran F, et al. (2012) Rheo- and NutG suppress pervasive anti-sense transcription in Escherichia coli. Genes Dev 26: 2621–2633.

65. Petersen SE, Winter MG, Xavier MN, Thiemannrit P, Poon V, et al. (2013) Host-derived nutrient boosts growth of E. coli in the inflamed gut. Science 339: 708–711.

66. Thiery D, Salgado H, Huerta AM, Collado-Vides J (1998) Prediction of transcriptional regulatory sites in the complete genome sequence of Escherichia coli K-12. Bioinformatics 14: 1–40.

67. Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, et al. (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS ONE 4: e5289.

68. White-Ziegler CA, Malhowski AJ, Young S (2007) Human body temperature variability functions of the napF transcription in vivo Decrease in cell viability in an RMF, and ArcA in repression and activation. Mol Microbiol 25: 605–615.

69. Fabich AJ, Leatham MP, Grissom JE, Wiley G, Lai H, et al. (2011) Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. Nucleic Acids Res 39: 2073–2091.

70. Kammler M, Schon C, Hanke K (1993) Characterization of the ferrous iron uptake system of Escherichia coli. J Bacteriol 175: 6212–6219.

71. Ogasawara H, Shinohara Y, Yamamoto K, Ishihama A (2012) Novel regulation targets of the metal-response BasB-BasC two-component system of Escherichia coli. Microbiology 158: 1482–1492.

72. Cruz-Ramos H, Crack J, Wu G, Hughes MN, Scott C, et al. (2002) NO sensing by CRP: regulation of the Escherichia coli NO-detoxifying flavohaemoglobin, Hup. EMBO J 21: 3235–3244.

73. Green J, Baldwin ML (1997) The molecular basis for the differential regulation of the hly-encoded haemolysin of Escherichia coli by FNR and HbS lies in the improved activating region 1 contact of HbS. Microbiology 143 (Pt 12): 3715–3725.

74. Lutz S, Bohme R, Albrecht JA, Gunsalus RP (1997) Aerobic regulation of the Escherichia coli histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res 3: 141–153.

75. Samuels VA, Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, et al. (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS ONE 4: e5289.
143. Gibert I, Barbé J (1990) Cyclic AMP stimulates transcription of the structural gene of the outer-membrane protein OmpA of *Escherichia coli*. FEMS Microbiol Lett 56: 307–311.

144. Shin D, Choe N, Hsu S, Ryu S (2003) Selective regulation of ptsG expression by Fis. Formation of either activating or repressing nucleoprotein complex in response to glucose. J Biol Chem 278: 14776–14781.

145. Zheng D, Constantinidou C, Hobman JL, Minchin SD (2004) Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. Nucleic Acids Res 32: 5874–5893.

146. Postma PW, Lengeler JW, Jacobson GR (1993) Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol Rev 57: 543–594.

147. Hutchings MI, Drabble WT (2000) Regulation of the divergent gspE1 and rpoD promoters of *Escherichia coli* by the cyclic AMP receptor protein. FEBS Microbiol Lett 187: 115–122.

148. Feng Y, Cronan JE (2010) Overlapping repressor binding sites result in additive regulation of *Escherichia coli* FadH by FadR and ArcA. J Bacteriol 192: 4289–4299.

149. Nørregaard-Madsen M, Mygind B, Pedersen R, Valentin-Hansen P, Søgaard-Andersen L (1994) The gene encoding the periplasmic cyclophilin homologue, PPIase A, in *Escherichia coli*, is expressed from four promoters, three of which are activated by the cAMP-CRP complex and negatively regulated by the CytR repressor. Mol Microbiol 14: 989–997.

150. Peekhaus N, Conway T (1998) Positive and negative transcriptional regulation of the *Escherichia coli* glucorionate regulon gene *gntT* by GntR and the cyclic AMP (cAMP)/cAMP receptor protein complex. J Bacteriol 180: 1777–1785.

151. Zhang Z, Gosset G, Barabote R, Gonzalez CS, Cuevas WA, et al. (2005) Functional interactions between the carbon and iron utilization regulators, Crp and Fur, in *Escherichia coli*. J Bacteriol 187: 908–910.

152. Chen Z, Lewis KA, Shultzberger RK, Lyakhov IG, Zheng M, et al. (2007) Discovery of Fur binding site clusters in *Escherichia coli* by information theory models. Nucleic Acids Res 35: 6762–6777.

153. Lavrrar JL, Christoffersen CA, McIntosh MA (2002) Fur-DNA interactions at the bidirectional *fipD-thcA* promoter region in *Escherichia coli*. J Mol Biol 322: 983–995.

154. Christoffersen CA, Brickman TJ, McIntosh MA (2001) Regulatory architecture of the iron-regulated *fipD-thcA* bidirectional promoter region in *Escherichia coli*. J Bacteriol 183: 2059–2070.

155. Brickman TJ, Ozenberger BA, McIntosh MA (1990) Regulation of divergent transcription from the iron-responsive *fipA-thcA* promoter-operator regions in *Escherichia coli*. J Mol Biol 212: 669–682.

156. Zhang J, Zeuner Y, Kleefeld A, Unden G, Janshoff A (2004) Multiple site-specific binding of Fis protein to *Escherichia coli* nadA promoter DNA and its impact on DNA topology visualised by means of scanning force microscopy. Chembiochem 5: 1206–1209.

157. Young GM, Poole K (1994) Repression of *tanR* transcription during anaerobic growth requires Fur binding at the promoter and a second factor binding upstream. Mol Microbiol 11: 943–954.

158. Vassnova N, Kosyrev D (2000) A method for direct cloning of Fur-regulated genes: identification of seven new Fur-regulated loci in *Escherichia coli*. Microbiology 146: 3171–3182.

159. Stojilkovic I, Baumler AJ, Hanke K (1994) Fur regulon in gram-negative bacteria: identification and characterization of new iron-regulated *Escherichia coli* genes by a fur titration assay. J Mol Biol 236: 331–345.

160. McHugh JP, Rodriguez-Quinones F, Abdul-Tehrani H, Svistunenko DA, Poole RK, et al. (2003) Global iron-dependent gene regulation in *Escherichia coli*. A new mechanism for iron homeostasis. J Biol Chem 278: 29470–29486.