

Dissimilatory Metabolism of Nitrogen Oxides in Bacteria: Comparative Reconstruction of Transcriptional Networks

Dmitry A. Rodionov^{1,4,6}, Inna L. Dubchak², Adam P. Arkin³, Eric J. Alm³, and Mikhail S. Gelfand^{1,4,5}

¹*Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoi Karetny per. 19, Moscow, 127994, Russia;* ²*Genomics and* ³*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, US;* ⁴*State Scientific Center GosNII Genetika, 1st Dorozhny pr. 1, Moscow 117545, Russia;* ⁵*Department of Bioengineering and Bioinformatics, Moscow State University, Vorobievsky Gory 1-73, Moscow 119992, Russia.*

⁶Corresponding author E-mail: rodionov@iitp.ru;

Keywords: comparative genomics, nitrosative stress, denitrification, transcriptional network, bacteria

Running title: Regulation Of Dissimilatory Nitrogen Oxides Metabolism

ABSTRACT

Bacterial response to nitric oxide (NO) is of major importance since NO is an obligatory intermediate of the nitrogen cycle. Transcriptional regulation of the dissimilatory nitric oxides metabolism in bacteria is diverse and involves FNR-like transcription factors HcpR, DNR and NnrR, two-component systems NarXL and NarQP, NO-responsive activator NorR, and nitrite-sensitive repressor NsrR.

Using comparative genomics approaches we predict DNA-binding signals for these transcriptional factors and describe corresponding regulons in available bacterial genomes. Within the FNR family of regulators, we observed a correlation of two specificity-determining amino acids and contacting bases in corresponding DNA signal. Highly conserved regulon HcpR for the hybrid cluster protein and some other redox enzymes is present in diverse anaerobic bacteria including Clostridia, Thermotogales and delta-proteobacteria. NnrR and DNR control denitrification in alpha- and beta-proteobacteria, respectively. Sigma-54-dependent NorR regulon found in some gamma- and beta-proteobacteria contains various enzymes involved in the NO detoxification. Repressor NsrR, which was previously known to control only nitrite reductase operon in *Nitrosomonas* spp., appears to be the master regulator of the nitric oxides metabolism not only in most gamma- and beta-proteobacteria (including well-studied species like *Escherichia coli*), but also in Gram-positive *Bacillus* and *Streptomyces* species. Positional analysis and comparison of regulatory regions of NO detoxification genes allows us to propose the candidate NsrR-binding signal. The most conserved member of the predicted NsrR regulon is the NO-detoxifying flavohemoglobin Hmp. In enterobacteria, the regulon includes also two nitrite-responsive loci, *nipAB* (*hcp-hcr*) and *nipC* (*dnrN*), thus confirming the identity of the effector, i.e. nitrite. The proposed NsrR regulons in *Neisseria* and some other species are extended to include denitrification genes. As the result, we demonstrate considerable interconnection between various nitrogen-oxides-responsive regulatory systems for the denitrification and NO detoxification genes and evolutionary plasticity of this transcriptional network.

SYNOPSIS

The global nitrogen cycle includes the denitrification and nitric oxide (NO)-detoxifying enzymes that mediate inter-conversion of nitrogen oxides between a number of redox states. In bacteria, expression of genes involved in these processes is tightly controlled using a cascade of oxygen and N-oxides responsive transcription factors. Despite the importance of bacterial nitrogen oxides metabolism for ecology and medicine our understanding of its regulation is very limited. In this study, the researchers have applied comparative genomics to the nitrosative stress and denitrification genes in seventy bacterial genomes and identified a large number of new regulatory signals. The regulatory network involves five transcription factors in different combinations, regulating expression of dozens NO-detoxifying genes. The reconstructed network demonstrates considerable interconnection between various nitrogen-oxides-responsive regulatory systems and their evolutionary plasticity. Not only genes are shuffled between regulons in different genomes, but there exists considerable interaction between regulators. Overall, the system seems to be quite conserved, however many regulatory interactions in the identified core regulatory network are taxon-specific, whereas some genomes feature expansion of specific regulons. This study demonstrates the power of comparative genomics in the analysis of complex regulatory networks and their evolution.

INTRODUCTION

Inter-conversion of nitrogen species between a number of redox states forms the biogeochemical nitrogen cycle which has multiple environmental impacts and industrial applications. Bacteria can utilize soluble nitrogen oxides, nitrate and nitrite, as terminal electron acceptors in oxygen-limiting conditions. Two dissimilar pathways of nitrate respiration, ammonification and denitrification, involve formation of a common intermediate, nitrite, but end in different products, ammonia and gaseous nitrogen oxides or dinitrogen, respectively (Fig. 1). At the first step, nitrite is formed by one of three different types of nitrate reductases: soluble assimilatory Nas, membrane-associated respiratory Nar, and periplasmic dissimilatory Nap. The next step of ammonification is conversion of nitrite into ammonia by either respiratory cytochrome *c* nitrite reductase NrfA or detoxifying siroheme-containing enzyme NirBD [1]. In contrast, during denitrification, nitrite is reduced to nitric oxide, nitrous oxide and, finally, dinitrogen, using nitrogen oxide reductases NirK (or NirS), NorB and NosZ, respectively [2].

Nitric oxide (NO) is a signaling and defense molecule in animals, but bacteria are sensitive to high NO concentrations due to its reactivity and membrane permeability [3]. NO and hydroxylamine, two toxic intermediates in 6-electron reduction of nitrite, could be formed during nitrite ammonification [4,5]. In addition to a classical NO reductase (NorB) present in denitrifying species, two other bacterial NO detoxification enzymes have been characterized: an NO reductase (flavorubredoxin NorVW in *Escherichia coli*) [6] and an NO dioxygenase (flavohemoglobin Hmp or Fhp in *E. coli*, *Bacillus subtilis*, *Ralstonia eutropha*, and *Pseudomonas* species) [7-9].

An unusual redox enzyme, called the hybrid cluster protein (Hcp) or 'prismane protein', has been extensively studied in strictly anaerobic (*Desulfovibrio* species) and facultative anaerobic (*E. coli*, *Salmonella typhimurium*, *Acidithiobacillus ferrooxidans*, *Shewanella oneidensis*) bacteria, where it is mostly induced during conditions of nitrite or nitrate stress, suggesting a role in nitrogen metabolism [10-14]. In the latter bacteria, the *hcp* gene always forms a possible operon with NADH oxidoreductase *hcr*, whose product catalyzes reduction of Hcp in the presence of NADH [11]. Until recently, the *in vivo* electron-accepting substrate of Hcp was unknown, and based on the crystal structure, NO was assumed to be a good candidate for this role [12, 15]. *In vitro* studies demonstrated oxygen-sensitive hydroxylamine reductase activity of the *E. coli* Hcp protein, suggesting its possible role in detoxification of reactive by-products of nitrite reduction [16]. More recently, the requirement of the *hcp* gene for *in vivo* hydroxylamine reduction was observed in *Rhodobacter capsulatus* E1F1 [17].

Expression of the denitrification genes is known to be activated by nitrogen oxides and low oxygen tension [18]. Both in denitrifying and ammonifying γ -proteobacteria, the nitrate/nitrite signal is processed by the two component sensor-regulator NarX-NarL and its paralog NarQ-NarP in *E. coli* that control the respiratory nitrate reductase operon *nar* and the nitrite ammonifying loci *nir* and *nrf* [19]. Various transcription factors of the FNR family have been described as NO-sensing regulators of denitrification: DNR in *Pseudomonas* species, NNR in *Paracoccus denitrificans*, and NnrR in *Rhodobacter sphaeroides* and *Bradyrhizobium japonicum* [18]. The DNR/NNR and NnrR proteins cluster phylogenetically in separate subgroups, separately from other family members including FNR, a global regulator of anaerobiosis in facultative anaerobic bacteria [20].

Another NO-responsive transcriptional factor, σ^{54} -dependent NorR, activates expression of the NO reductases *norVW* in *E. coli* and *norAB* in *R. eutropha* [21, 22]. Three

tandem upstream activator sites with the core consensus **GT-(N₇)-AC** were identified as NorR-binding sites observed in both promoter regions. Analysis of the adjacent regions of additional *norR* orthologs in bacterial genomes revealed similar tandem NorR-binding sites upstream of the *norA* and *norB* genes in *Ralstonia* species, *norVW* in *Salmonella* species, *hmp* in *Vibrio cholerae* and *Pseudomonas aeruginosa*, and *hcp* in *Vibrio vulnificus* [23].

A nitrite-sensitive transcriptional repressor, named NsrR, has been identified in lithoautotrophic β -proteobacterium *Nitrosomonas europaea*, where it regulates expression of the copper-nitrite reductase *nirK* [24]. Co-localization of the *nsrR* ortholog and the *hcp* gene in *R. capsulatus* E1F1 suggested that NsrR and nitrite could be involved in the regulation of hydroxylamine assimilation in this α -proteobacterium [17]. NsrR is a member of the Rrf2 family of transcriptional regulators, which includes a putative Rrf2 regulator for a redox operon in *D. vulgaris* [25], the iron-responsive repressor RirA, which controls iron uptake in rhizobia [26], and the IscR repressor for the Fe-S cluster assembly operon in *E. coli* [27].

Despite this diversity of regulatory systems, our understanding of the regulation of the nitrogen oxides metabolism in bacteria is very limited. For example, NO- and/or nitrite-dependent activation of expression of *hmp* in *E. coli* and *B. subtilis*, *hcp-hcr* (*nipAB*) and *dnrN* (*nipC*) in *S. typhimurium*, *norB* and *aniA* (*nirK*) in *Neisseria gonorrhoeae* has been described [28-30], but specific transcriptional factors involved in this control are not yet known. In this study, we analyzed regulation of the nitrosative stress and denitrification genes in available bacterial genomes using comparative genomics approaches [31, 32] and predicted a large number of new regulatory elements for these genes. In addition to complete description of the previously known NorR, DNR and NnrR regulons, we report identification of a novel FNR-like regulator, named HcpR, for the *hcp* and other redox-related genes in anaerobic bacteria. Starting from very limited data, we were able to identify the NsrR-binding signal and describe the NsrR regulons in sequenced γ - and β -proteobacteria, as well as in the *Bacillus* and *Streptomyces* species. Combining published experimental and newly obtained comparative data, we have reconstructed the NO- and nitrite-dependent transcriptional regulatory network for dissimilatory metabolism of nitrogen oxides in bacteria.

RESULTS

HcpR, its signals and core regulon

A member of the CRP/FNR family of transcription factors, HcpR has been initially identified as the regulator of the *hcp* gene encoding the hybrid cluster protein and the *frdX* encoding a ferredoxin-like protein in the *Desulfovibrio* species, anaerobic metal-reducing δ -proteobacteria [33]. The consensus of the candidate HcpR binding sites, **wTGTGAnnnnnnTCACA**w, is similar to the CRP consensus of *E. coli*.

Close *hcpR* orthologs were detected in other δ -proteobacteria, namely two *Geobacter* species, *Desulfotalea psychrophila* and *Desulfuromonas*. However, the same CRP-like signals were not present in these genomes. As the analysis of the regulator multiple alignment revealed a substitution in the Helix-Turn-Helix motif involved in DNA recognition, that could cause this change (see below the subsection “Co-evolution of regulators and their recognition signals” for details), and since the considered species have multiple *hcp* and *frdX* paralogs, we applied the signal detection procedure to a set of corresponding upstream regions and obtained a new FNR-like palindromic signal with consensus sequence **wyTTGACnnnnnGTCA**A_rw, which has a notable distinction from the CRP-like signal in the third position (not G, but T). This signal was observed upstream of most

hcp and *frdX* paralogs in the studied δ -proteobacteria, as well as upstream of some additional genes in *Desulfuromonas* and *Geobacter* species (Fig. 2 and Table S1).

Close orthologs of *hcpR* from δ -proteobacteria are present in two cyanobacteria, *Anabaena variabilis* and *Synechocystis* sp. (*Avar17201* and *slr0449* in Fig. 3), where they are divergently transcribed with the *hcp* and *norB* genes, respectively. Both these genes are preceded by candidate HcpR sites with consensus sequence **TTGACnnnnGTCAA**, and no other similar sites were found in the genomes of these two cyanobacteria (Fig. 2).

To analyze possible regulation of the *hcp* genes in other taxonomic groups of bacteria, we considered their gene neighborhoods and found that genes for FNR-like regulators are often co-localized with *hcp* in most *Clostridium* species, *Bacteroides*, Thermotogales and *Treponema denticola* (Fig. 2). On the phylogenetic tree of the FNR/CRP protein family (Fig. 3), all such regulators form a separate branch, named HcpR2, and additional representatives of this branch always co-occur with *hcp* genes in bacterial genomes. By applying the signal recognition procedure to a set of *hcp* upstream regions from HcpR2-containing genomes, we identified a conserved signal with consensus **GTAACnnnnGTTAC**.

Other types of signals were observed upstream of the *hcp* genes in *Clostridium thermocellum*, *C. difficile* and *Porphyromonas gingivalis*, and upstream of the *hcp* gene in *Thermoanaerobacter tengcongensis*. In the latter species the *hcp* gene has a CRP-like regulatory site and is preceded by the *crp2* gene which is an ortholog of the *B. subtilis* *fnr* gene, making it likely that *crp2* regulates *hcp* (Fig. 2). The predicted HcpR2 regulons in most *Bacteroides* species, *P. gingivalis*, *Fusobacterium nucleatum*, *T. denticola* and Thermotogales contain only *hcp* genes (Fig. 2).

DNR and NnrR core regulons

In two denitrifying *Pseudomonas* species, *P. stutzeri* and *P. aeruginosa*, expression of the *nir*, *nor*, and *nos* genes is regulated by the NO-responsive FNR-like transcriptional activator DNR that binds to a signal similar to the consensus FNR box, **TTGATnnnnATCAA** [34, 35]. By a combination of similarity search and phylogenetic analysis of the CRP/FNR protein family (Fig. 3), we identified DNR orthologs in the genomes of various denitrifying β -proteobacteria, including three *Ralstonia* and two *Burkholderia* species, *C. violaceum* and *Thiobacillus denitrificans* (Fig. 4). To identify the DNR binding signal in denitrifying species, we selected the upstream regions of denitrification genes encoding nitrite, NO and nitrous oxide reductases from genomes containing DNR orthologs and applied the signal detection procedure. The resulting FNR box-like signal with consensus **CTTGATnnnnATCAAG** was identified upstream of most denitrification genes (*nirS*, *nirK*, *norB*, *nosZ*) (Table S2).

No orthologs of the HcpR, DNR, NsrR (below), and NorR (below) regulators were identified in α -proteobacterial genomes. The only exception seems to be *R. capsulatus* E1F1 whose genome contains an *nsrR* ortholog close to the *hcp* gene within the nitrate assimilation *nas* gene cluster [17]. However, in denitrifying species, including *R. sphaeroides* and *B. japonicum*, the FNR-like transcriptional factor NnrR activates expression of nitrite and NO reductases and of the *nnrS* gene [36-38]. Orthologs of *nnrR* were identified in six α -proteobacteria, all of which possess also the *nir* and *nor* genes involved in denitrification (Fig. 5). The NnrR orthologs form a separate branch on the phylogenetic tree of the CRP/FNR family (Fig. 3). To analyze the NnrR regulon, the signal detection procedure was applied to a training set of the *nir*, *nor*, *nos* and *nnrS* upstream regions from α -proteobacteria. The conserved NnrR signal with consensus **ctTTGcgnnnnncgCAAag** was identified upstream of

most denitrification genes (Table S3). The same candidate NnrR sites have been previously identified in *R. sphaeroides* by a comparison of the *nir*, *nor* and *nnrS* upstream regions and then confirmed for the latter gene by site-directed mutagenesis [36].

Co-evolution of regulators of the CRP/FNR family and their recognition signals

The HcpR signals identified in several bacteria demonstrated some diversity which could be correlated with changes in the regulator DNA-binding Helix-Turn-Helix domain. In particular, the CRP-like signal **wTG TGAnnnnnnTCACA w** of *Desulfovibrio* species differs from the FNR-like signal **wyTTGACnnnnnGTCAArw** in other δ -proteobacteria in the third position (not T, but G). Examination of multiple alignment of the CRP/FNR-like proteins revealed one specific amino acid (R₁₈₀) in the HTH motif involved in DNA recognition, which has changed from arginine (like in *E. coli* CRP and *Desulfovibrio* HcpR) to Ser or Pro in other δ -proteobacteria (see the HcpR1 branch of the phylogenetic tree, Fig. 3). Similarly, the difference between the HcpR2 signal **GTAACnnnnnGTTAC** and the signals of δ -proteobacteria is consistent with substitution of Glu-181 in the DNA-recognizing HTH motif to Pro in the HcpR2 proteins (Fig. 3).

The structure of CRP in complex with its DNA operator has been determined [39]. Three positions (1ber, chain A residues 180, 181, and 185) interact with the DNA target site, and mutagenesis studies have shown that point mutations at these positions alter the specificity of the protein [40]. We systematically analyzed HcpR, HcpR2, DNR, NnrR sites identified here as well as several known consensus sequences for other CRP/FNR-family regulators and observed a correlation of two specificity-determining positions, R₁₈₀ and E₁₈₁, and contacting bases in a DNA signal, G₃ and A₆, respectively (Fig. 3).

A different substitution is observed in three bacterial species, where position 181 in the HcpR2 protein is filled by either Ser (*Clostridium thermocellum* and *C. difficile*), or Gln (*Porphyromonas gingivalis*). In agreement with these replacements, the candidate HcpR2 signals in these species differ from the common signal detected for most HcpR2-containing genomes (Fig. 3). Finally, the Crp2 regulator in *T. tengcongensis* has CRP-like regulatory signal and is orthologous to the FNR regulator of *B. subtilis*.

The phylogenetic tree of the CRP/FNR regulators (Fig. 3) represents four main groups of proteins analyzed in this study, DNR, NnrR, HcpR1 and HcpR2. It also includes several well-studied family members with established DNA-binding consensuses. All respective branches on the tree are deeply rooted and thus their phylogenetic relationships to each other are not well resolved and differ from results of a more extensive phylogenetic analysis of the CRP/FNR-like transcriptional regulators [41]. Nevertheless, in both trees the HcpR1 and DNR branches cluster together, whereas HcpR1 and HcpR2 form two quite distinct branches on the phylogenetic tree (Fig. 3). All these proteins lack the canonical FNR-type cysteine motif involved in binding of oxygen-labile Fe-S cluster, thus excluding their role in oxygen sensing [41,42].

NsrR, its signals and core regulon

The above analysis suggests that HcpR controls the *hcp* genes in strictly anaerobic bacteria. However, a large number of facultative anaerobic bacteria possessing *hcp* gene lack *hcpR* orthologs. In *E. coli*, *S. typhimurium* and *S. oneidensis*, *hcp* is expressed only under anaerobic conditions in the presence of nitrite or nitrate [10-12]. In an attempt to explain a possible molecular mechanism of this induction, we first aligned the upstream regions of the *hcp* genes from eight enterobacteria and identified two highly conserved regions (Fig. S1).

The upstream potential signal resembles the consensus sequence of the FNR signal and thus most likely is involved in the anaerobic induction of the *hcp-hcr* operon by FNR [12]. The second potential signal, an imperfect inverted repeat with consensus **gATGyAT-(N₅)-ATrCATc** located downstream of the FNR site, is likely the binding site for a regulatory protein which responds to nitrogen oxides.

Construction of a recognition rule and search in complete *E. coli* genome identified similar sites in upstream regions of the hypothetical gene *dnrN* (*ytfE*) and the *hmp* gene encoding NO-detoxifying flavohemoglobin. Importantly, both these candidate sites are highly conserved in multiple alignments of *dnrN* and *hmp* upstream regions from related enterobacteria (Fig. S2 and S3). Search in the *S. oneidensis* genome identified the same signal upstream of *hcp-hcr*, *SO4302* (*dnrN*), and *SO2805* (*nnrS*), the latter encoding a hypothetical haem-copper-containing membrane protein [36]. The *hmp* gene is absent in this genome (Fig. 6).

Identification of the conserved palindromic signal suggests that some common transcription factor co-regulates the *hcp-hcr*, *dnrN*, *hmp* and *nnrS* genes in enterobacteria and *Shewanella* species. Since bacterial transcription factors often directly regulate adjacent genes [43], we analyzed gene neighbourhoods of genes preceded by the predicted sites (Fig. 4 and 6). In many proteobacteria, including Vibrionales, *Acinetobacter* sp., *Chromobacterium violaceum*, *Ralstonia* and *Bordetella* species, as well as in Gram-positive bacilli and actinobacteria, the flavohemoglobin gene *hmp* is positionally clustered with a hypothetical transcriptional factor from the Rrf2 protein family (Pfam PF02082). The characterized members of the PF02082 family are the Rrf2 repressor for the electron transport operon *hmc* in *Desulfovibrio vulgaris* [25], the iron-sulfur cluster repressor IscR in *E. coli* [27], the iron-responsive regulator RirA in rhizobia [26], and the nitrite-sensitive repressor NsrR for the nitrite reductase operon *nirK* in *Nitrosomonas europaea* [24]. Phylogenetic analysis (D.R., in preparation) demonstrated that all representatives of the Rrf2 protein family associated with *hmp* genes appear to be orthologs of the *N. europaea* NsrR protein and thus we tentatively assign this name to the entire subfamily. Orthologs of the nitrite-sensitive repressor NsrR were identified in all β - and most γ -proteobacteria, being absent only in Pasteurellaceae, Pseudomonadales, and *V. cholerae*. We predict that this transcriptional factor actually binds the identified DNA signal upstream of nitrite/NO-induced genes in enterobacteria and *Shewanella*.

To analyze further the NsrR regulon, we constructed a recognition rule for the NsrR signal and used it to scan the genomes of γ - and β -proteobacteria (Table S4; Fig. 4 and 6). The flavohemoglobin gene *hmp* has an upstream NsrR site in most of these genomes, excluding Pseudomonadales, *V. cholerae* and *Polaromonas* sp. where it is a member of the NO-responsive regulon NorR (see below). The *nnrS* gene, another well-conserved member of the NsrR regulon, was found in some genomes within a possible operon with *nsrR* or *hmp* (Fig. 4 and 6). The *norB* gene encoding an NO reductase in denitrifying bacteria is preceded by NsrR sites in the *Neisseria* species, *C. violaceum*, *Polaromonas* sp., *Ralstonia solanacearum*, and two *Burkholderia* species, *B. mallei* and *B. pseudomallei*. Another key enzyme of the denitrification, the copper-containing nitrite reductase NirK, is predicted to be a member of the NsrR regulon in the *Neisseria* species, *C. violaceum*, and *N. europaea* (Fig. 4), and in the latter bacterium it was recently shown to be a target of this nitrite-sensitive repressor [24]. In addition to γ -proteobacteria, the *hcp* gene was found under NsrR regulation in a β -proteobacterium (*B. cepacia*), and an α -proteobacterium (*R. capsulatus* E1F1).

Orthologs of *nsrR* have been also found in the complete genomes of most *Bacillus* and *Streptomyces* species, where they are clustered with the flavohemoglobin gene *hmp*. The only exception is *B. subtilis*, which has a stand-alone *nsrR* ortholog, *yhdE*. The predicted

NsrR-binding signal appears to be well-conserved in these Gram-positive bacteria, and candidate sites were observed only upstream the *hmp* genes. Multiple experimental studies in *B. subtilis* showed nitrite or NO-dependent induction of expression of *hmp*, however the mechanism of this control was not known [9, 29]. The experimentally mapped *hmp* promoter in *B. subtilis* significantly overlaps with the predicted tandem NsrR sites [44].

The obtained data suggest that the nitrite-responsive NsrR regulon has a wide phylogenetic distribution. Its most conserved member is the NO-detoxifying flavohemoglobin Hmp, which is present both in Gram-negative and Gram-positive bacteria. Most other regulon members are involved in the nitrosative stress and denitrification. The identified NsrR signal, a palindrome with consensus **gATGyAT-(N₅)-ATrCATc**, is well conserved in most analyzed bacteria (Fig. 7). The only exception is the NsrR signal in *Neisseria* species, where symmetrical positions G₄ and C₁₆ are replaced by T and A, respectively.

NorR regulon

In Vibrionales, the *hcp-hcr* operon is preceded by a gene that encodes a homolog of the NO-responsive regulator NorR, named NorR2. NorR is a σ^{54} -dependent transcriptional activator that regulates expression of the NO reductase operons, *norVW* in *E. coli* and *norAB* in *R. eutropha* [21, 22]. NorR binds to three tandem operator regions, inverted repeats with degenerate consensus **GT-(N₇)-AC**, which are localized upstream of the σ^{54} promoter site [23]. By applying the signal detection procedure to the *hcp* promoter regions from four Vibrionales genomes, we identified two tandem palindromic sites with consensus **GATGT-(N₇)-ACATC** (Fig. S4). These likely binding sites for the NorR2 protein are localized immediately upstream of candidate σ^{54} promoters well conforming to the consensus, and thus could be involved in the NO-dependent activation of the *hcp-hcr* operon (Table S5). In addition, the *norR2-hcp-hcr* gene loci in Vibrionales contain a single NorR site without an associated σ^{54} promoter located upstream of the *norR2* gene. This site could be involved in the negative autoregulation of the NorR2 expression (Fig. 6).

To analyze analogous NO-responsive regulons in other species, we performed exhaustive similarity search and identified *norR*-like genes only in a limited number of β - and γ -proteobacteria (Fig. 4 and 6). An *E. coli*-like arrangement of the divergently transcribed *norR* and *norVW* genes with conserved tandem NorR-binding sites and a σ^{54} promoter was found in *S. typhimurium*, two *Erwinia* species and two *Vibrio* species. Other NO-detoxification genes possibly regulated by candidate NorR sites are the NO-reductase *norB* in *Ralstonia* spp. and *Shewanella putrefaciens*, and the NO dioxygenase *hmp* gene in *V. cholerae*, *Pseudomonas* spp., *Polaromonas* sp., and *B. fungorum*. In all these cases except *P. stutzeri*, the *norR* gene is clustered with the target genes on the chromosome. In the unfinished genome of *P. stutzeri*, the candidate tandem NorR sites followed by candidate σ^{54} promoters were found upstream of the *hmp* and *dnrN* genes, but the *norR* gene was not found in the sequenced portion of the genome. In addition, we found that *V. cholerae* has a second target for NorR in the genome, the hypothetical gene *nnrS*, which was identified as a member of various NO/nitrite-responsive regulons in other proteobacteria (NsrR, DNR, NnrR, see below).

The consensus sequences of NorR and NorR2 signals identified in various taxonomic groups have only a limited number of universally conserved positions (Fig. 8). Positions G₅ and T₆ and complementary positions A₁₄ and C₁₅ are the most conserved ones throughout the NorR family, being only partially replaced in *Polaromonas* sp. (A₅). Noteworthy, in some *Vibrio* species, two *norR* paralogs are present, *norR1* and *norR2*, which are associated with

the *norVW* and *hcp-hcr* operons, respectively. The NorR1 and NorR2 consensus sequences differ significantly in four positions (C₇, G₁₃ for NorR1 and G₂, C₁₈ for NorR2), allowing for discrimination of the target sites by the NorR paralogs in these species.

Complex regulation of hybrid cluster protein genes

Differences in the predicted mode of regulation of the hybrid cluster proteins (Table 1), which are present in diverse bacterial and archaeal species, are well traced on the phylogenetic tree of this protein family (Fig. 9). Indeed, the *hcp* gene is regulated by HcpR (highlighted in yellow) in many anaerobic bacteria, by NsrR in facultative anaerobic enterobacteria and some β - and α - proteobacteria (in magenta), by NorR in most Vibrionales (in green), and by DNR in *A. ferrooxidans* and *Thermochromatium tepidum* (in blue). It often forms operons with either the NADH oxidoreductase *hcr* (in γ -proteobacteria) or the ferredoxin-like gene *frdX* (in δ -proteobacteria and *Clostridium* spp.), suggesting functional linkage between the Hcp and Hcr/FrdX proteins. In addition to predicted NsrR signals, the *hcp-hcr* operons in enterobacteria are also preceded by candidate binding sites of the anaerobic activator FNR, suggesting their induction during anaerobiosis (Fig. S1). We also investigated the regulatory regions of *hcp* in two Pasteurellaceae lacking all above mentioned nitrogen oxides regulators (*Actinobacillus pleuropneumoniae* and *Mannheimia succiniciproducens*) and found a strong candidate binding site of NarP, a response regulator from the nitrate/nitrite-specific two-component regulatory system NarQ-NarP. The NarP regulon in *E. coli* contains mainly genes from the nitrate/nitrite respiration pathway [18], whereas in the above two Pasteurellaceae, the NarP regulon is extended to include the detoxification genes *hcp-hcr*, *dnrN* and *norB*. Additional analysis using the NarP recognition rule revealed candidate NarP sites upstream of some NsrR-regulated genes in γ -proteobacteria and *Neisseria* species (Fig. 6). In agreement with these findings, the nitrate- and nitrite-induced transcription from the *hcp* promoter in *E. coli* was found to depend on the response regulator proteins NarL and NarP [45]. These observations show that the nitrite induction of the NO-detoxification genes in different genomes can be achieved by multiple transcriptional factors.

Additional members of the regulons

The main regulatory interactions analyzed in this study are shown in Fig. 10 and Table 1. The core regulon members, that is, genes regulated by the nitrogen-oxide responsive factors NsrR, HcpR, DNR, NnrR and NorR in many genomes, are the hybrid cluster protein gene *hcp*, the NO-detoxifying flavohemoglobin *hmp*, two hypothetical genes *dnrN* and *nnrS*, NO reductase operon *norVW*, and multiple denitrification genes, *nir*, *nor*, and *nos*, encoding nitrite, NO and nitrous oxide reductases, respectively. The core regulon members can be regulated by different regulators in various genomes. Further, some genes may be regulated by several regulators simultaneously. All considered regulons contain also a large number of additional members that are summarized in Table 1.

Though the main target of the HcpR regulators are genes encoding the hybrid cluster protein Hcp and the hypothetical ferredoxin-containing protein FrdX, the HcpR regulons are significantly extended in δ -proteobacteria and clostridia (Table S1, Fig. 2 and Table 1). In *Desulfuromonas* and *Geobacter* species, they include the nitrite reductase *nrjHA*, the NADH dehydrogenase *ndh*, and the nitrate reductase *nar* operon. In *Desulfovibrio* species, the HcpR regulon is extended to include the *apsBA* and *sat* loci involved in the sulfate reduction pathway. Among hypothetical genes, the predicted HcpR regulons often contain ferredoxin-,

hemerythrin-, or cytochrome *c*-like genes. For instance, the *CTC0897-CTC0898* operon of *C. tetani* encoding a permease and a ferredoxin-like protein, respectively, is likely regulated by the divergently transcribed paralog of HcpR2 (CTC0896).

Additional members of the NsrR regulon were identified in all enterobacteria (Fig. 6). Two hypothetical transporter genes, that are homologous to the tellurite resistance gene *tehB* and the nitrite extrusion gene *narK*, could be involved in the protection from the nitrosative stress by excretion of nitrite from cytoplasm. In support of these observations, the NsrR regulons were found to include a *narK*-like transporter gene (*nasA*) in β -proteobacterium *C. violaceum*. and a *tehB* homolog in *Photobacterium profundum*. The *V. fischeri* NsrR regulon includes a homolog of the eukaryotic alternative oxidase Aox. In *Legionella pneumophila*, a non-denitrifying γ -proteobacterium without *hmp*, *hcp*, *dnrN* and *nnrS* genes, the *glbN-lpg2536* operon encoding a heme-containing cyanoglobin and a hypothetical ferredoxin reductase was found to be a sole NsrR target. We suggest that both these genes could be involved in the detoxification process by mediating NO oxidation (similar to the flavohemoglobin Hmp).

Denitrifying bacteria like the *Pseudomonas* and *Burkholderia* species contain additional members of the DNR and NnrR regulons (Fig. 4 and 5). For instance, the *hemN* gene encoding an O₂-independent coproporphyrinogen III oxidase involved in the protohaem biosynthesis and relevant for denitrification [46] is preceded by a strong DNR site in the *Burkholderia* species but has candidate NnrR-binding sites in some α -proteobacteria. In *Brucella melitensis*, the NnrR regulon includes the *nosA* gene, encoding an outer membrane copper receptor, shown to be required for the assembly of the copper-containing nitrous oxide reductase in *P. stutzeri* [47]. Finally, a gene of unknown function (*COG4309*) is probably co-transcribed with the NnrR-regulated *nor* genes in three rhizobial species, where as in three β -proteobacteria this gene is a predicted member of the NsrR regulon and it is often co-transcribed with *norB* (Fig. 4). These observations indicate that the COG4309 family could be relevant for function of the NO reductase.

Two close *dnr* paralogs most likely resulting from a recent duplication were found in *A. ferrooxidans*. One is divergently transcribed with the *hcp-COG0543* operon, which is preceded by a strong candidate DNR site. The second paralog clusters with the *cgb-COG0543-COG0446* operon, which also is preceded by a candidate DNR site. The product of the first gene in this operon is similar to a single-domain hemoglobin in *Campylobacter jejuni*, which is indispensable for defense against NO and nitrosating agents [48]. Thus we predict that the recently duplicated DNR paralogs in *A. ferrooxidans* regulate two different NO-detoxifying systems.

DISCUSSION

The results of this study demonstrate considerable variability of the metabolic and regulatory systems for nitrogen oxides (Fig. 10). Many genes change the regulators in different genomes (Table 1). However, overall, the system seems to be quite conserved. Genes involved in denitrification, such as *nir*, *nor* and *nos*, are mainly regulated by two NO-responsive transcriptional activators from the FNR/CRP family, NnrR in α -proteobacteria and DNR in β -proteobacteria and *Pseudomonas* spp. Three different nitrogen oxides-responsive transcription factors appear to regulate genes required for defense against the nitrosative stress: the σ^{54} -dependent activator NorR from the NtrC family in some γ - and β -proteobacteria (present in at least 18 species), the Rrf2 family NsrR repressor widely-distributed in proteobacteria and firmicutes (39 species), and the FNR-like transcription factor HcpR in diverse anaerobic bacteria (22 species). The primary targets of the newly

identified regulator HcpR are the hybrid-cluster protein Hcp, which has a protective role in nitrite stress conditions, and the associated ferredoxin-like proteins. NorR usually regulates cytoplasmic NO reductase *norVW* and sometimes another membrane-bound NO reductase (*norB*) and the NO dioxygenase *hmp*. The NsrR regulon almost always includes the *hmp* and *hcp* genes, as well as one or both genes of unknown function, *dnrN* and *nnrS*.

On the whole, the NsrR, NorR and DNR regulons are differentially distributed in γ - and β -proteobacteria, and the former prevails over the other two. All three regulons are present only in three *Ralstonia* species, where NsrR controls the NO-detoxifying flavohemoglobin, whereas NorR and DNR regulate the denitrification genes (*nir*, *nor* and *nos*). In addition, the NorR and NsrR factors co-occur in ten other non-denitrifying species, complementing each other in the control of the nitrosative stress genes. Finally, NorR and DNR regulators were found only in *P. aeruginosa*, and NsrR and DNR co-occur in four denitrifying β -proteobacteria.

Some regulatory interactions in the identified core regulatory network seem to be taxon-specific (thin lines in Fig. 10B, also see above “Complex regulation of hybrid cluster protein genes”). They include NsrR-regulated *norVW* and *nos* in *P. profundum*, *norB* and *nirK* in various β -proteobacteria; NorR-regulated *dnrN* in *P. stutzeri* and *nnrS* in *P. aeruginosa*; HcpR-regulated *dnrN* in *B. fragilis* and *Desulfuromonas*, *norB* in *Synechocystis* sp.; and DNR-regulated *nnrS* and *hmp* in *P. stutzeri*, *hcp* in *A. ferrooxidans* and *T. tepidum*. The extensions of the NsrR regulon include the denitrification genes *nirK* and *norB* in *Neisseria* species, *Burkholderia* spp. and *C. violaceum*. The former is of particular interest as, in contrast to the latter two lineages, the *Neisseria* species lack the DNR regulator, assuming a lineage-specific substitution of both the transcription factor and its binding sites. Indeed, in the *Neisseria* species, the complete denitrification pathway including nitrite, NO and nitrous oxide reductases, as well as *dnrN* and the two-component regulatory system *narQP*, seems to be regulated by NsrR. The hypothesis that NsrR mediates regulation of denitrification genes in *Neisseria* is further supported by the observation that in *N. gonorrhoeae*, the *norB* gene is strongly induced by NO independently of FNR and NarP [30].

Not only genes are shuffled between regulons in different genomes, but there may exist considerable interaction between regulators. Firstly, in some species the DNR regulon overlaps with other nitrogen oxide-responsive regulons. The upstream regions of *norB* and *nirK* in *C. violaceum*, *COG4309-norB* in *R. solanacearum*, and *nnrS2* in *Burkholderia* species contain two candidate regulatory sites, a downstream NsrR site and an upstream DNR site (Fig. 4), yielding both positive regulation by the NO-responsive activator DNR and nitrite-induced derepression by the NsrR repressor. Secondly, the NO-detoxifying gene *hmp* in *P. stutzeri* is preceded by three candidate NorR sites at positions -192, -173 and -148 (relative to the translational start site), a strong DNR site at position -116, and a putative σ^{54} promoter at position -91 (Fig. 6). This arrangement of regulatory elements indicates dual positive control of the *hmp* expression by different NO-responsive activators, σ^{54} -dependent NorR and σ^{70} -dependent DNR. Finally, in many cases genes are regulated by two additional regulators, the oxygen-responsive factor FNR (*hcp-hcr*, *hmp* and *narK* in enterobacteria) and the nitrite/nitrate-sensitive two-component system NarQ/NarP (*hcp-hcr*, *dnrN*, and *hmp* in enterobacteria, *nnrS* in Vibrionales and *Shewanella* spp.). More complex regulatory interactions are observed in *Neisseria* spp., where NsrR regulates the NarQ/NarP system, whereas the common upstream region of the divergently transcribed genes *norB* and *nirK* contains two candidate NsrR sites, a candidate NarP site in the middle, and an FNR-binding site immediately upstream of *nirK*, the latter being involved in the anaerobic induction of this gene [49].

Various aspects of the described regulatory network for the nitrogen oxides metabolism are verified by a large number of independent experimental observations. Upregulation of the *hcp* gene in response to growth on nitrate or nitrite was reported in *S. oneidensis*, *E. coli*, *S. typhimurium* and *D. vulgaris* [10-12, 14], the same pattern of regulation was observed for *dnrN* in *S. typhimurium* and *P. stutzeri* [12, 34]. Our prediction of positive regulation of *hcp-frdX* and negative regulation of the sulfate reduction genes *apsBA* and *sat* [33] was validated in a macroarray hybridization study, where *hcp* was upregulated 255-fold with 5 mM nitrite, whereas *aprAB* and *sat* were downregulated nearly tenfold in the same conditions [14]. In addition, nitrite induced transcription of thiosulfate reductase *phsA* (a candidate member of the HcpR regulon) and inhibited the membrane-bound electron transport complex *qmoABC*, located just downstream of the *apsBA* genes and thus also possibly repressed by HcpR.

The flavohemoglobin gene *hmp* is induced by NO and nitrite in *E. coli* and *B. subtilis* [28, 29], and the mechanism of the *hmp* regulation by the nitrite repressor NsrR predicted in this study is conserved in these diverse bacteria. Another NO-mediated mechanism of *hmp* regulation in *E. coli* by the O₂-responsive repressor FNR was proposed [50]. At that, NO was found to inactivate FNR anaerobically, restoring the *hmp* expression. However, our data indicate that, in contrast to the candidate NsrR site, the FNR binding site is conserved only in closely related bacteria *E. coli* and *S. typhimurium* (Fig. S3). Finally, NO induces the *norB* expression in *N. gonorrhoeae*, but it was found to be independent of known nitrogen oxide-responsive regulators [30]. Here we describe possible co-regulation of all denitrification genes in the *Neisseria* species by the nitrite-sensitive repressor NsrR. Recently, transcriptional regulation of the flavohemoglobin gene *fhp* (*hmp* ortholog) by the NO-responsive regulator FhpR (NorR ortholog) has been demonstrated in *P. aeruginosa* [51].

A recent paper by Elvers et al. [52] describes a new nitrosative stress-responsive regulon in ϵ -proteobacterium *C. jejuni* regulated by a member of the CRP/FNR family. This regulator (named NssR) is homologous (26% identity) to the HcpR2 factor from *T. maritima* described in this study. It has an FNR-like binding signal (TTAACnnnnnGTTAA) and specificity-determining positions A₁₈₀ and Q₁₈₁, conforming to the correlation between these two positions and contacting bases in the DNA signal observed in this study. NssR positively regulates expression of the *Campylobacter* globin gene *cgb*, encoding a single-domain hemoglobin that mediates resistance to NO and nitrosative stress [48]. Thus the regulation of the *cgb* gene by HcpR in *A. ferrooxidans* predicted in this study is in agreement with verified NssR-mediated activation of the *cgb* ortholog in *C. jejuni*.

The mechanisms of regulation of nitrogen oxides metabolism in bacteria are of great importance both for ecology and medicine. Nitrifying and denitrifying microorganisms are significant sources of both nitric and nitrous oxides production in the atmosphere, and thus have a great impact in the greenhouse effect [53]. Nitrate has become a pollutant of groundwater and surface water. Nitric oxide and reactive nitrogen intermediates are also part of the arsenal of antimicrobial agents produced by macrophages [54]. Therefore, nitrosative stress tolerance genes, which are inducible after invasion, provide a strong advantage for pathogenic bacteria that need to resist the host defense system. Here we tentatively characterized the transcriptional regulatory network for the genes involved in these significant metabolic processes. Overall, although each particular prediction made in this study may require experimental verification, the emerging overall picture seems to be rather consistent and robust.

MATERIALS and METHODS

Complete and partial bacterial genomes were downloaded from GenBank [55]. Preliminary sequence data were also obtained from the WWW sites of the Institute for Genomic Research (<http://www.tigr.org>), the Wellcome Trust Sanger Institute (<http://www.sanger.ac.uk/>), and the DOE Joint Genome Institute (<http://jgi.doe.gov>). The gene identifiers from Genbank are used throughout. Genome abbreviations are listed in Table 2. Protein similarity search was done using the Smith-Waterman algorithm implemented in the GenomeExplorer program [56]. Orthologous proteins were defined by the best bidirectional hits criterion and named by either a common name of characterized protein or by an identifier in the Clusters of Orthologous Groups (COG) database for uncharacterized proteins [57]. Distant homologs were identified using PSI-BLAST [58]. The SEED tool, which combines protein similarity search, positional gene clustering, and phylogenetic profiling of genes, was applied for comparative analysis and annotation of multiple microbial genomes (<http://theseed.uchicago.edu/FIG/index.cgi>, see the “Nitrosative stress and Denitrification” subsystems). The phylogenetic trees were constructed by the maximum likelihood method implemented in PHYLIP [59] using multiple sequence alignments of protein sequences produced by CLUSTALX [60]. In addition, the InterPro [61], and PFAM [62] databases were used to verify protein functional and structural annotations.

For identification of candidate regulatory motifs we started from sets of potentially co-regulated genes (using previous experimental and general functional considerations). A simple iterative procedure implemented in the program SignalX (as described previously in [63]) was used for construction of common transcription factor-binding signals in sets of upstream gene fragments. Each genome encoding the studied transcription factor was scanned with the constructed profile using the GenomeExplorer software (see the detailed description at <http://bioinform.genetika.ru/projects/reconstruction/index.htm>), and genes with candidate regulatory sites in the upstream regions were selected. The threshold for the site search was defined as the lowest score observed in the training set. Dependent on the signal and the number of sites in the training set, such threshold could be too strict or too permissive. It seems that the threshold choice was adequate in our cases, as little clear false positives were encountered, and, on the other hand, most functionally relevant genes were found to belong to at least one of the studied regulons (Table 3 and Results and Discussion). The upstream regions of genes that are orthologous to genes containing regulatory sites of any of studied nitrogen-related factors were examined for candidate sites even if these were not detected automatically with a given threshold. Among new candidate members of a regulon, only genes having candidate sites conserved in at least two other genomes were retained for further analysis. We also included new candidate regulon members that are functionally related to the nitrogen oxides metabolism. This procedure allowed us to reject a small number of false positive sites identified after scanning of microbial genomes (Table 3). Sequence logos for derived regulatory signals were drawn using WebLogo package v.2.6 [64] (<http://weblogo.berkeley.edu/>).

ACKNOWLEDGEMENTS

This study was partially supported by grants from the Howard Hughes Medical Institute (55000309 to M.G.), the Russian Fund of Basic Research (04-04-49361 to D.R.), the Russian Science Support Fund (M.G.), and the Russian Academy of Sciences (Programs "Molecular and Cellular Biology" and "Origin and Evolution of the Biosphere"). I.L.D., E.J.A. and A.P.A. were in part supported by a US Department of Energy Genomics:GTL

grant (DE-AC03-76SF00098, to A.P.A.). We thank Anna Gerasimova and Dmitry Ravcheev for the FNR and NarP recognition profiles, and Andrey Mironov and Sergey Stolyar for useful discussions.

REFERENCES

1. Simon J (2002) Enzymology and bioenergetics of respiratory nitrite ammonification. *FEMS Microbiol Rev* 26: 285-309.
2. Shapleigh JP (2000) The Denitrifying Prokaryotes. In: Dworkin M et al., editors. *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, 3rd edition, release 3.1, January 20, 2000, Springer-Verlag, New York.
3. Poole RK (2005) Nitric oxide and nitrosative stress tolerance in bacteria. *Biochem Soc Trans* 33: 176-180.
4. Rudolf M, Einsle O, Neese F, Kroneck PM (2002) Pentahaem cytochrome c nitrite reductase: reaction with hydroxylamine, a potential reaction intermediate and substrate. *Biochem Soc Trans* 30: 649-653.
5. Kuznetsova S, Knaff DB, Hirasawa M, Setif P, Mattioli TA (2004) Reactions of spinach nitrite reductase with its substrate, nitrite, and a putative intermediate, hydroxylamine. *Biochemistry* 43: 10765-10774.
6. Gomes CM, Giuffre A, Forte E, Vicente JB, Saraiva LM, et al. (2002) A novel type of nitric-oxide reductase. *Escherichia coli* flavorubredoxin. *J Biol Chem* 277: 25273-25276.
7. Gardner PR, Gardner AM, Martin LA, Salzman AL (1998) Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. *Proc Natl Acad Sci U S A* 95: 10378-10383.
8. Cramm R, Siddiqui RA, Friedrich B (1994) Primary sequence and evidence for a physiological function of the flavohemoprotein of *Alcaligenes eutrophus*. *J Biol Chem* 269: 7349-7354.
9. LaCelle M, Kumano M, Kurita K, Yamane K, Zuber P, et al. (1996) Oxygen-controlled regulation of the flavohemoglobin gene in *Bacillus subtilis*. *J Bacteriol* 178: 3803-3808.
10. Beliaev AS, Thompson DK, Khare T, Lim H, Brandt CC, et al. (2002) Gene and protein expression profiles of *Shewanella oneidensis* during anaerobic growth with different electron acceptors. *OMICS* 6: 39-60.
11. van den Berg WA, Hagen WR, van Dongen WM (2000) The hybrid-cluster protein ('prismane protein') from *Escherichia coli*. Characterization of the hybrid-cluster protein, redox properties of the [2Fe-2S] and [4Fe-2S-2O] clusters and identification of an associated NADH oxidoreductase containing FAD and [2Fe-2S]. *Eur J Biochem* 267: 666-676.
12. Kim CC, Monack D, Falkow S (2003) Modulation of virulence by two acidified nitrite-responsive loci of *Salmonella enterica* serovar Typhimurium. *Infect Immun* 71: 3196-3205.
13. Dominy CN, Deane SM, Rawlings DE (1997) A geographically widespread plasmid from *Thiobacillus ferrooxidans* has genes for ferredoxin-, FNR-, prismane- and NADH-oxidoreductase-like proteins which are also located on the chromosome. *Microbiology* 143: 3123-3136.
14. Haveman SA, Greene EA, Stilwell CP, Voordouw JK, Voordouw G (2004) Physiological and gene expression analysis of inhibition of *Desulfovibrio vulgaris* hildenborough by nitrite. *J Bacteriol* 186: 7944-7950.
15. Cooper SJ, Garner CD, Hagen WR, Lindley PF, Bailey S (2000) Hybrid-cluster protein (HCP) from *Desulfovibrio vulgaris* (Hildenborough) at 1.6 Å resolution. *Biochemistry* 39: 15044-15054.

16. Wolfe MT, Heo J, Garavelli JS, Ludden PW (2002) Hydroxylamine reductase activity of the hybrid cluster protein from *Escherichia coli*. J Bacteriol 184: 5898-5902.
17. Cabello P, Pino C, Olmo-Mira MF, Castillo F, Roldan MD, et al. (2004) Hydroxylamine assimilation by *Rhodobacter capsulatus* E1F1. requirement of the *hcp* gene (hybrid cluster protein) located in the nitrate assimilation nas gene region for hydroxylamine reduction. J Biol Chem 279: 45485-45494.
18. Zumft WG (2002) Nitric oxide signaling and NO dependent transcriptional control in bacterial denitrification by members of the FNR-CRP regulator family. J Mol Microbiol Biotechnol 4: 277-286.
19. Stewart V (1993) Nitrate regulation of anaerobic respiratory gene expression in *Escherichia coli*. Mol Microbiol 9: 425-434.
20. Korner H, Sofia HJ, Zumft WG (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. FEMS Microbiol Rev 27: 559-592.
21. Tucker NP, D'Autreaux B, Studholme DJ, Spiro S, Dixon R (2004) DNA binding activity of the *Escherichia coli* nitric oxide sensor NorR suggests a conserved target sequence in diverse proteobacteria. J Bacteriol 186: 6656-6660.
22. Pohlmann A, Cramm R, Schmelz K, Friedrich B (2000) A novel NO-responding regulator controls the reduction of nitric oxide in *Ralstonia eutropha*. Mol Microbiol 38: 626-638.
23. Busch A, Pohlmann A, Friedrich B, Cramm R (2004) A DNA region recognized by the nitric oxide-responsive transcriptional activator NorR is conserved in beta- and gamma-proteobacteria. J Bacteriol 186: 7980-7987.
24. Beaumont HJ, Lens SI, Reijnders WN, Westerhoff HV, van Spanning RJ (2004) Expression of nitrite reductase in *Nitrosomonas europaea* involves NsrR, a novel nitrite-sensitive transcription repressor. Mol Microbiol 54: 148-158.
25. Keon RG, Fu R, Voordouw G (1997) Deletion of two downstream genes alters expression of the *hmc* operon of *Desulfovibrio vulgaris* subsp. *vulgaris* Hildenborough. Arch Microbiol 167: 376-383.
26. Todd JD, Wexler M, Sawers G, Yeoman KH, Poole PS, et al. (2001) RirA, an iron-responsive regulator in the symbiotic bacterium *Rhizobium leguminosarum*. Microbiology 148: 4059-4071.
27. Schwartz CJ, Giel JL, Patschkowski T, Luther C, Ruzicka FJ, et al. (2001) IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. Proc Natl Acad Sci U S A 98: 14895-14900.
28. Poole RK, Anjum MF, Membrillo-Hernandez J, Kim SO, Hughes MN, et al. (1996) Nitric oxide, nitrite, and Fnr regulation of *hmp* (flavo-hemoglobin) gene expression in *Escherichia coli* K-12. J Bacteriol 178: 5487-5492.
29. Moore CM, Nakano MM, Wang T, Ye RW, Helmann JD (2004) Response of *Bacillus subtilis* to nitric oxide and the nitrosating agent sodium nitroprusside. J Bacteriol 186: 4655-4664.
30. Householder TC, Fozo EM, Cardinale JA, Clark VL (2000) Gonococcal nitric oxide reductase is encoded by a single gene, *norB*, which is required for anaerobic growth and is induced by nitric oxide. Infect Immun 68: 5241-5246.
31. Osterman A, Overbeek R (2003) Missing genes in metabolic pathways: a comparative genomics approach. Curr Opin Chem Biol 7: 238-251.
32. Gelfand MS, Novichkov PS, Novichkova ES, Mironov AA (2000) Comparative analysis of regulatory patterns in bacterial genomes. Brief Bioinform 1: 357-371.
33. Rodionov DA, Dubchak I, Arkin A, Alm E, Gelfand MS (2004) Reconstruction of regulatory and metabolic pathways in metal-reducing delta-proteobacteria. Genome Biol 5: R90.

34. Vollack KU, Zumft WG (2001) Nitric oxide signaling and transcriptional control of denitrification genes in *Pseudomonas stutzeri*. *J Bacteriol* 183: 2516-2526.
35. Arai H, Mizutani M, Igarashi Y (2003) Transcriptional regulation of the *nos* genes for nitrous oxide reductase in *Pseudomonas aeruginosa*. *Microbiology* 149: 29-36.
36. Bartnikas TB, Wang Y, Bobo T, Veselov A, Scholes CP, et al. (2002) Characterization of a member of the NnrR regulon in *Rhodobacter sphaeroides* 2.4.3 encoding a haem-copper protein. *Microbiology* 148: 825-833.
37. Kwiatkowski AV, Shapleigh JP (1996) Requirement of nitric oxide for induction of genes whose products are involved in nitric oxide metabolism in *Rhodobacter sphaeroides* 2.4.3. *J Biol Chem* 271: 24382-24388.
38. Mesa S, Bedmar EJ, Chanfon A, Hennecke H, Fischer HM (2003) *Bradyrhizobium japonicum* NnrR, a denitrification regulator, expands the FixLJ-FixK2 regulatory cascade. *J Bacteriol* 185: 3978-3982.
39. Schultz SC, Shields GC, Steitz TA (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 253: 1001-7.
40. Luscombe NM, Thornton JM (2002) Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity. *J Mol Biol* 320: 991-1009.
41. Korner H, Sofia HJ, Zumft WG (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* 27: 559-592.
42. Crack J, Green J, Thomson AJ (2004) Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR). *J Biol Chem* 279: 9278-9286.
43. Korbel JO, Jensen LJ, von Mering C, Bork P (2004) Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nat Biotechnol* 22: 911-917.
44. LaCelle M, Kumano M, Kurita K, Yamane K, Zuber P, et al. (1996) Oxygen-controlled regulation of the flavohemoglobin gene in *Bacillus subtilis*. *J Bacteriol* 178: 3803-3808.
45. Filenko NA, Browning DF, Cole JA (2005) Transcriptional regulation of a hybrid cluster (prismane) protein. *Biochem Soc Trans* 33: 195-197.
46. Zumft WG (1997) Cell biology and molecular basis of denitrification. *Microbiol Mol Biol Rev* 61: 533-616.
47. Lee HS, Abdelal AH, Clark MA, Ingraham JL (1991) Molecular characterization of *nosA*, a *Pseudomonas stutzeri* gene encoding an outer membrane protein required to make copper-containing N₂O reductase. *J Bacteriol* 173: 5406-5413.
48. Elvers KT, Wu G, Gilberthorpe NJ, Poole RK, Park SF (2004) Role of an inducible single-domain hemoglobin in mediating resistance to nitric oxide and nitrosative stress in *Campylobacter jejuni* and *Campylobacter coli*. *J Bacteriol* 186: 5332-5341.
49. Lissenden S, Mohan S, Overton T, Regan T, Crooke H, et al. (2000) Identification of transcription activators that regulate gonococcal adaptation from aerobic to anaerobic or oxygen-limited growth. *Mol Microbiol* 37: 839-855.
50. Cruz-Ramos H, Crack J, Wu G, Hughes MN, Scott C, et al. (2002) NO sensing by FNR: regulation of the *Escherichia coli* NO-detoxifying flavohaemoglobin, Hmp. *EMBO J* 21: 3235-3244.
51. Arai H, Hayashi M, Kuroi A, Ishii M, Igarashi Y (2005) Transcriptional regulation of the flavohemoglobin gene for aerobic nitric oxide detoxification by the second nitric oxide-responsive regulator of *Pseudomonas aeruginosa*. *J Bacteriol* 187: 3960-3968.
52. Elvers KT, Turner SM, Wainwright LM, Marsden G, Hinds J, et al. (2005) NssR, a member of the Crp-Fnr superfamily from *Campylobacter jejuni*, regulates a nitrosative stress-responsive regulon that includes both a single-domain and a truncated haemoglobin. *Mol Microbiol* 57: 735-750.

53. Levine JS, Augustsson TR, Anderson IC, Hoell JM Jr (1984) Tropospheric sources of NO_x: lightning and biology. *Atmos Environ* 18: 1797-1804.
54. Fang FC (1997) Perspectives series: host/pathogen interactions. Mechanisms of nitric oxide-related antimicrobial activity. *J Clin Invest* 99: 2818-2825.
55. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. *Nucleic Acids Res* 33: D34-8.
56. Mironov AA, Vinokurova NP, Gelfand MS (2000) GenomeExplorer: software for analysis of complete bacterial genomes. *Mol Biol* 34: 222-231.
57. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, et al. (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* 29: 22-8.
58. Altschul SF, Koonin EV (1998) Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. *Trends Biochem Sci* 23: 444-447.
59. Felsenstein J (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. *J Mol Evol* 17: 368-376.
60. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882.
61. Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, et al. (2001) The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res* 29: 37-40.
62. Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, et al. (2002) The Pfam protein families database. *Nucleic Acids Res* 30: 276-80.
63. Gelfand MS, Koonin EV, Mironov AA (2000) Prediction of transcription regulatory sites in Archaea by a comparative genomic approach. *Nucleic Acids Res* 28: 695-705.
64. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188-1190.

Figure and Table legends

Figure 1. **The bacterial inorganic nitrogen cycle.** The ammonification, denitrification, detoxification, nitrogen fixation and nitrification pathways are shown by colored solid lines with genes names involved in the pathway. The dashed black line shows possible non-enzymatic interconversions of nitrogen oxides. The dotted line shows additional formation of nitric oxide and hydroxylamine during nitrite ammonification.

Figure 2. **Genomic organization of genes regulated by HcpR.** Yellow circles with numbers denote candidate HcpR sites with different consensus sequences. These numbers correspond to the HcpR profile numbers in Figure 3.

Figure 3. **Maximum likelihood phylogenetic tree of the FNR/CRP family of transcriptional regulators.** The third column contains sequences of helix-turn-helix motifs in the proteins. Two specificity-determining positions correlated with DNA signals are colored (R₁₈₀ and E₁₈₁ in proteins correlate with G₃ and A₆ in DNA sites, respectively). The fourth column includes sequence logos for presumably homogeneous and large site sets and sequence consensi for small sets of DNA sites and for well-established signals of other factors (FNR, CRP, CooA, NtcA, ArcR). The last column indicates the name of a search profile constructed in this study.

Figure 4. **Genomic organization of genes regulated by NsrR, NorR and DNR in β -proteobacteria.** Magenta, green and blue circles denote candidate NsrR, NorR and DNR sites, respectively. Candidate σ^{54} promoters associated with NorR sites are shown by angle arrows. Experimentally known sites of NorR and DNR are marked by 's'. Additional sites of the NarP and FNR factors are indicated by purple squares and black triangles, respectively.

Figure 5. **Genomic organization of genes regulated by NnrR and NsrR in α -proteobacteria.** Orange and magenta circles denote candidate NnrR and NsrR sites, respectively. Experimentally known NnrR sites are marked by 's'.

Figure 6. **Genomic organization of genes regulated by NsrR, NorR and DNR in γ -proteobacteria.** Magenta, green and blue circles denote candidate NsrR, NorR and DNR sites, respectively. Candidate σ^{54} promoters associated with NorR sites are shown by angle arrows. Experimentally known sites of NorR and DNR are marked by 's'. Additional sites of the NarP and FNR factors are indicated by purple squares and black triangles, respectively.

Figure 7. **Sequence logos for the identified NsrR-binding sites in various bacterial taxa.**

Figure 8. **Sequence logos for the identified NorR-binding sites in various species of proteobacteria.**

Figure 9. **Maximum likelihood phylogenetic tree of the hybrid cluster (prismane) proteins.** Genes predicted to be regulated by the nitrogen oxides-related factors are highlighted by respective colors. Additionally, predicted FNR-regulated genes are denoted by black circles. Genes positionally linked to the NADH oxidoreductase *hcr* and the hypothetical ferredoxin *frdX* genes are shown by red and blue lines, respectively. Archaeal genes are shown by pointed lines.

Figure 10. **Regulatory interactions between genes for dissimilatory nitrogen oxides metabolism in bacteria.** **A. Distribution of regulators and regulated genes.** The number of cases when a gene is regulated by a specific transcription factor is indicated by the length of a colored bar in the histogram. The white bar in the histogram shows the cases when the gene is present in a genome possessing at least one of the studied regulons, but is not regulated by any of them. **B. Combined regulatory network.** Arrows denote regulatory interactions, with the thickness reflecting the frequency of the interaction in the analyzed genomes.

Table 1. **Predicted members of regulatory and metabolic networks of the nitrogen oxides dissimilatory metabolism.**

Table 2. **The list of genome abbreviations used in this study.**

Table 3. **Distribution of predicted regulatory sites in bacterial genomes.** As very little experimental information is available, the predictions are classified based on functional and comparative genomics criteria. At that, *candidate sites*, CS, are defined as sites upstream of genes with reasonable function, or belonging to other related regulons, or sites conserved in several genomes. Of these, the number of sites used to make a profile is given in parentheses. *False sites*, FS, are non-conserved sites located upstream of clearly functionally irrelevant genes. *Uncertain sites*, US, are non-conserved sites located upstream of hypothetical genes. Dash, the factor is not present in the genome. N/A, total statistics inapplicable due to genome incompleteness. In most cases NorR has multiple sites for one gene. Thus, 2 and 3 in the NorR column means one gene, and 4-6 means two genes.

Supporting Information.

Table S1. Candidate HcpR-binding sites.

Table S2. Candidate DNR-binding sites.

Table S3. Candidate NnrR-binding sites.

Table S4. Candidate NsrR-binding sites.

Table S5. Candidate NorR-binding sites and corresponding σ^{54} promoters .

Figure S1. Multiple sequence alignment of the upstream regions of the *hcp-hcr* operons from enterobacteria.

Figure S2. Multiple sequence alignment of the upstream regions of the *dnrN* genes from enterobacteria.

Figure S3. Multiple sequence alignment of the upstream regions of the *hmp* genes from enterobacteria.

Figure S4. Multiple sequence alignment of the upstream regions of the *hcp-hcr* operons from *Vibrio* species.