

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/110622>

Please be advised that this information was generated on 2021-09-17 and may be subject to change.

RESEARCH ARTICLE

Open Access

Comparative genome analysis of central nitrogen metabolism and its control by GlnR in the class Bacilli

Tom Groot Kormelink^{1,2,3,5}, Eric Koenders⁵, Yanick Hagemeyer⁵, Lex Overmars^{2,4,5}, Roland J Siezen^{1,2,4,5}, Willem M de Vos^{2,3} and Christof Francke^{1,2,4,5*}

Abstract

Background: The assimilation of nitrogen in bacteria is achieved through only a few metabolic conversions between alpha-ketoglutarate, glutamate and glutamine. The enzymes that catalyze these conversions are glutamine synthetase, glutaminase, glutamate dehydrogenase and glutamine alpha-ketoglutarate aminotransferase. In low-GC Gram-positive bacteria the transcriptional control over the levels of the related enzymes is mediated by four regulators: GlnR, TnrA, GltC and CodY. We have analyzed the genomes of all species belonging to the taxonomic families Bacillaceae, Listeriaceae, Staphylococcaceae, Lactobacillaceae, Leuconostocaceae and Streptococcaceae to determine the diversity in central nitrogen metabolism and reconstructed the regulation by GlnR.

Results: Although we observed a substantial difference in the extent of central nitrogen metabolism in the various species, the basic GlnR regulon was remarkably constant and appeared not affected by the presence or absence of the other three main regulators. We found a conserved regulatory association of GlnR with glutamine synthetase (*glnRA* operon), and the transport of ammonium (*amtB-glnK*) and glutamine/glutamate (i.e. via *glnQHMP*, *glnPHQ*, *gltT*, *alsT*). In addition less-conserved associations were found with, for instance, glutamate dehydrogenase in Streptococcaceae, purine catabolism and the reduction of nitrite in Bacillaceae, and aspartate/asparagine deamination in Lactobacillaceae.

Conclusions: Our analyses imply GlnR-mediated regulation in constraining the import of ammonia/amino-containing compounds and the production of intracellular ammonia under conditions of high nitrogen availability. Such a role fits with the intrinsic need for tight control of ammonia levels to limit futile cycling.

Background

The assimilation and re-distribution of nitrogen within a cell is essentially controlled within the central metabolic conversions between alpha-ketoglutarate, glutamate and glutamine (Figure 1A). The enzymes that catalyze these conversions are glutamine synthetase (GS), glutaminase (G), glutamate dehydrogenase (GDH) and glutamine alpha-ketoglutarate aminotransferase (GOGAT). On a short time-scale, the enzyme activity is controlled via activating and inhibitory molecular interactions. For instance, the activity of GS is suppressed via feedback inhibition (FBI-GS) by the

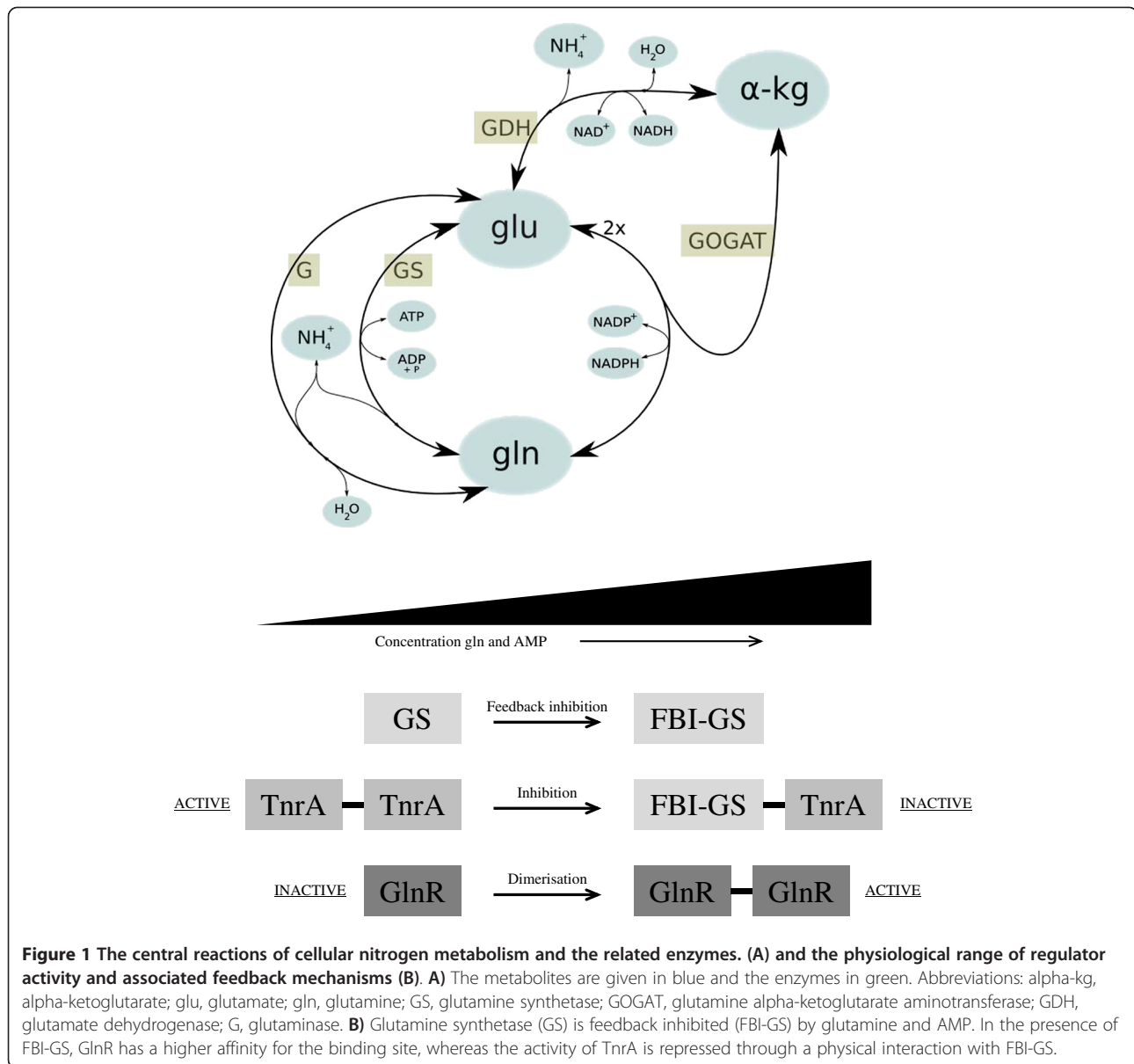
product glutamine and by AMP [1]. Under conditions of nitrogen limitation a high GS activity is maintained to ensure a sufficient level of glutamine [2,3]. On a longer time-scale, the enzyme levels are controlled via the activity of a limited number of transcription regulators.

Marked differences exist in the transcription control of the genes encoding the enzymes involved in central nitrogen metabolism across the bacterial kingdom. In the Gram-positive model organism *Bacillus subtilis*, the expression of these genes is mediated by four major transcription factors: CodY, GlnR, TnrA [4] and GltC [5,6]. Of these, GlnR, TnrA and GltC are specific for nitrogen metabolism whereas the global regulator CodY is linked to both carbon and nitrogen metabolism [7]. GltC is specifically associated with the control of the genes encoding glutamine alpha-ketoglutarate aminotransferase. The transcription factor

* Correspondence: c.francke@cmbi.ru.nl

¹Kluyver Centre for Genomics of Industrial Fermentation, P.O. Box 5057, 2600 GA Delft, The Netherlands

²Ti Food and Nutrition, P.O. Box 557, 6700 AN Wageningen, The Netherlands
Full list of author information is available at the end of the article



GlnR is active during growth with excess nitrogen, whereas TnrA is active during nitrogen-limiting growth [4]. The change in activity of these transcription factors is affected directly by GS and feedback inhibition (FBI) of the enzyme (Figure 1B). In *B. subtilis* GlnR is activated in the presence of FBI-GS [8] and TnrA is inhibited through a physical interaction with FBI-GS [2,9]. It was also shown that TnrA binds to the PII-like regulatory protein GlnK, which is sensitive to ATP, Mg²⁺ and alpha-ketoglutarate [9,10]. The two proteins become tightly associated with the ammonium permease AmtB at a low level of ATP. In *Streptococcus mutans* cross-linking and pull-down assays demonstrated that GlnR also interacts with GlnK and that the interaction enhances the binding of GlnR to its cognate site upstream of the *glnRA* operon [11].

In *B. subtilis* and many other low-GC Gram-positives the genes encoding GlnR (*glnR*) and GS (*glnA*) constitute the operon *glnRA* [12]. In *B. subtilis* GlnR was reported to repress the transcription of the *glnRA* operon (negative autoregulation), and of *tnrA* [13] and the urease gene cluster (*ureABC*) [14]. On the other hand, TnrA was reported to affect the transcription of a larger set of genes/operons [15], for instance activating *glnQHMP* (encoding a glutamine ABC transport system [16]), *amtB-glnK* (i.e. *nrgBA*; encoding an ammonium permease and the regulatory protein GlnK [17]), *nasA* and *nasBC/DEF* (encoding proteins related to nitrite reduction [18]), *gabP* (encoding a gamma amino butyrate transporter [19]) and *pucR* (encoding the purine catabolism regulator [20]), while repressing *alsT* (encoding an H⁺/Na⁺ amino

acid symporter [15]), *gltAD* (encoding glutamate synthase [21,22]) and *ilvBHC-leuABCD* (encoding branched-chain amino acid biosynthesis proteins [23]). Similarly, in the oral Streptococci *S. pneumoniae* and *S. mutans* GlnR was reported to repress the transcription of the *glnRA* operon and of the *glnPQ* operon (encoding another glutamine ABC transport system) in both organisms and of *gdh* (encoding glutamate dehydrogenase) in the former, and the *amtB-glnK* and *citBZ-idh* operons (encoding aconitate hydratase, citrate synthase and isocitrate dehydrogenase [24]) in the latter organism [25,26].

Comparative genome analyses have shown that GlnR, TnrA and CodY are characteristic for the low-GC Gram-positive species although their distribution is not uniform. For instance, whereas GlnR is found in almost all *Bacillus* species, TnrA has been identified only in a few. It is an intriguing question whether in the absence of one of these main regulators the others take over its role. We therefore decided to extend (i.e. from 16 to 173 genomes) a previous comparative analysis reported by [27] to identify the presence of the regulators and the genes they regulate in the low GC Gram-positive species of the class Bacilli. This class includes the well-studied families Bacillaceae, Listeriaceae, Staphylococcaceae, Lactobacillaceae, Leuconostocaceae and Streptococcaceae.

We have redefined the binding motifs of GlnR and TnrA on basis of the available experimental and sequence data and used them to identify their respective regulons anew. For that purpose we have applied a footprinting approach formulated earlier by us [28] and a similar motif search procedure [29]. The difference in composition of the GlnR regulon was compared for the various taxonomic families within the class Bacilli and for species having only GlnR or also additional regulators. For most families we found a rather stable composition of the GlnR regulon and some species-specific connections, independent of the presence or absence of the other two regulators. The data imply that GlnR-mediated regulation serves predominantly to limit the import of ammonia/amino-containing compounds and, at the same time, to limit the production of intracellular ammonia.

Results and discussion

Presence/absence analysis of the genes encoding the central enzymes and regulators

We identified the orthologs of the genes encoding the enzymes of central nitrogen metabolism (G, GS, GDH and GOGAT), the related transport systems and the regulators CodY, GlnR and TnrA, in the sequenced genomes of species related to the class Bacilli on basis of BLAST searches with the sequences of experimentally verified proteins (see methods for details). In Tables 1 and 2 the analysis results for representative species of the orders Bacillales and Lactobacillales, respectively, are presented;

the results for the complete set of analyzed species are given in Additional file 1. We observed a clear distinction in gene content between the two orders and between the different taxonomic families within the orders.

Remarkably, within the family Lactobacillaceae, *Lactobacillus acidophilus* and its close relatives lack all three regulators. There are only three other species, *Bacillus halodurans*, *Bacillus clausii* and *Bacillus selenitrireducens* that lack a gene encoding GlnR. The global regulator CodY is present in most species except for those of the families Lactobacillaceae and Leuconostocaceae. TnrA is only present in species of the order Bacillales within the families Bacillaceae, Paenibacillaceae and the genus *Exiguobacterium* with the exception of the species and strains of the *Bacillus cereus* group, *Alicyclobacillus acidocaldarius*, *Brevibacillus brevis* and *Lysinibacillus sphaericus*.

Similarly, we observed a large variation in the presence of the enzymes of central nitrogen metabolism, but much less so in the related transport systems. The set of enzymes is complete within the family of the Bacillaceae and mostly reduced in the other families; in many of the Lactobacillaceae, Leuconostocaceae and Streptococcaceae only glutamine synthetase and one of the other enzymes is present. In the case of transport, at least one ammonium transporter AmtB (Amt-family; 1.A.11 in TCDB classification [42]), also referred to as NrgA [40], is present in most species, although the transporter is absent in more than half of the analyzed Streptococcaceae, in three *Bacillus anthracis* strains, in the gut-related Lactobacilli (e. g. *L. johnsonii* and *L. gasseri*) and in some meat-related species (e.g. *Lactobacillus sakei* and *Staphylococcus carnosus*, *Macrococcus caseolyticus*). It was recently put forward that transport of ammonia (NH_4^+) should be active and tightly regulated to limit futile cycling [43]. This control was suggested to be exerted by the small PII-like regulator GlnK, earlier referred to as NrgB (and as GlnB in e.g. *L. lactis*); the corresponding genes are indeed found genetically associated to *amtB* in many of the analyzed species. However, at the same time, it is absent in many others, including all analyzed Lactobacillaceae. Moreover, GlnK was shown to interact with TnrA in *B. subtilis* [9,10] and GlnR in *S. mutans* [11].

Most of the analyzed species carry one operon encoding a glutamine ABC transporter (PAAT-family; 3.A.1.3). Two variants of this system were identified: i) a *glnQHMP* operon found in *Escherichia coli* and *B. subtilis* and ii) a *glnP^HQ* variant found in the Lactobacilli and Streptococci. This latter operon consists of 2 genes, in which *glnP^H* represents a fusion-gene of the permease *glnP* and the substrate-binding protein encoding *glnH*. In fact, in the case of the Streptococci the substrate-binding domain was duplicated to generate a *glnP^HH*Q variant, whereas in some Lactobacilli the *glnP^H* gene has been duplicated. The *E. coli*

Table 1 Distribution and GlnR-mediated regulation of genes related to central nitrogen metabolism within the order Bacillales

NCBI PROJECT NAME \ GENES	<i>codY</i>	<i>tnrA</i>	<i>glnR</i>	<i>glnA</i>	<i>gdh</i>	<i>bdh</i>	glutamin- ase (HA)	glutamin- ase (LA)	<i>gltAB</i>	<i>gltA^{BC}</i>	<i>gltS</i>	<i>amtB</i>	<i>glnK</i>	<i>glnQHMP</i>	<i>glnP^{HQ}</i>	<i>glnT, alsT, yrbD, yfiA</i>	<i>gltP, dctA, gltT, yhcl, nqt</i>	
A) order Bacillales																		
<i>Anoxybacillus flavithermus</i> WK1	1	<u>1</u>	<u>1</u>	<u>2(2)</u>	2	1			1	1		<u>1^d</u>			1	2(1)	1	
<i>Geobacillus kaustophilus</i> HTA426	1	<u>1</u>	<u>1</u>	<u>2(1)</u>	2	2	1		1			<u>1^d</u>		1	1	1	2	
<i>G. thermodenitrificans</i> NG80-2	1	<u>1</u>	<u>1</u>	<u>2(2)</u>	2	2	1		1			<u>1^d</u>			1	1	5(1)	
<i>G. sp.</i> Y412MC61	1	<u>1</u>	<u>1</u>	<u>2(1)</u>	2	2	1		1			<u>2^d</u>	<u>1</u>	1	1	1	2	
<i>G. sp.</i> C56-T3	1	<u>1</u>	<u>1</u>	<u>2(1)</u>	2	1	1		1			<u>1^d</u>		1	1	1	2	
<i>G. sp.</i> WCH70	1	<u>1</u>	<u>1</u>	<u>1</u>	2	1	1		1	1				1		1	3	
<i>Bacillus cereus</i> ATCC 14579	1		<u>1</u>	<u>1^b</u>	1	1	1	1,0		1	1	<u>1</u>		1,2		6(1)	5(1)	
<i>B. anthracis</i> str. Ames	1		<u>1</u>	<u>1</u>	1	1	1	1		1	1	<u>0,1</u>		1		6(1)	5(1)	
<i>B. thuringiensis</i> str. Al Hakam	1		<u>1</u>	<u>1</u>	1	1	1	1		1	1	<u>1</u>		1		7(1)	6(1)	
<i>B. weihenstephanensis</i> KBAB4	1		<u>1</u>	<u>1</u>	2	1	1			1	1	<u>1</u>		1		6(1)	5(1)	
<i>B. atrophaeus</i> 1942	1	<u>1</u>	<u>1</u>	<u>1</u>	3	1	1	1	1		1	<u>1</u>	<u>1</u>	<u>1</u>		3(1)	4	
<i>B. amyloliquefaciens</i> DSM7	1	<u>1</u>	<u>1</u>	<u>1</u>	2	1	1	1 ^c	1		1	<u>1</u>	<u>1</u>	1		2(1)	3	
<i>B. pumilus</i> SAFR-032	1	<u>1</u>	<u>1</u>	<u>1</u>	2	1	1	1	1		1	<u>1</u>	<u>1</u>	1		3	5(1)	
<i>B. subtilis</i> str. 168	1	<u>1</u>	<u>1</u>	<u>1</u>	2	1	1	1	1		1	<u>1</u>	<u>1</u>	1,0		4(1)	4	
<i>B. licheniformis</i> ATCC 14580	1	<u>1</u>	<u>1</u>	<u>1</u>	2(1)	1	1	1	1			<u>2(1)</u>	<u>2(1)</u>			6(1)	4	
<i>B. megaterium</i> DSM319	1	<u>1</u>	<u>1</u>	<u>1</u>	3	1		1	1		1	<u>2(2)</u>	<u>2(2)</u>	1		3(1)	11(1)	
<i>B. pseudofirmus</i> OF4	1	<u>1</u>	2(1) ^e	<u>2(2)</u>	4	1	1	1	1			<u>1</u>			1	3(1)	3	
<i>B. clausii</i> KSM-K16	1	1		2	2	1	1		1			1			1	3	3	
<i>B. halodurans</i> C-125	1	1		2	5	2	1	1	1			1				5	1	
<i>B. selenitireducens</i> MLS10	1	1		1	4	1		1	1 ^b	1		1			1	2		
<i>Lysinibacillus sphaericus</i> C3-41	1		<u>1</u>	<u>2(2)^f</u>	2	1						<u>1</u>	<u>1</u>		1	3	5	
<i>Oceanobacillus iheyensis</i> HTE831	1	1 ^a	<u>1</u>	<u>1</u>	1	2		1	1		1	<u>1</u>	<u>1</u>		1	3	3	
<i>Exiguobacterium</i> sp. AT1b	1	<u>1</u>	<u>1</u>	2(1)	1	1	1		1		1	<u>1</u>				6(1)	1	
<i>Paenibacillus polymyxa</i> E681		<u>1</u>	<u>1</u>	2(1)	1		1		1			<u>1^d</u>		2(1)	1	2	2	
<i>Pb. sp.</i> JDR-2	1	<u>1</u>	<u>1</u>	<u>2(1)</u>	2	1	1		1		1	<u>3(1^d)</u>	<u>1</u>	<u>1</u>			1	
<i>Brevibacillus brevis</i> NBRC 100599	1		<u>1</u>	<u>2(2)</u>	2	1	1	1	1 ^b		1	<u>1^d</u>		1	1	1	3	
<i>Ab. acidocaldarius</i> DSM 446	1		<u>1</u>	<u>3(1)</u>	2	1				1		<u>2(1)</u>	<u>1</u>					
<i>Listeria innocua</i> Clip11262	1		<u>1</u>	<u>1</u>	1				1			<u>1</u>	<u>1</u>		1			
<i>L. monocytogenes</i> EGD-e	1		<u>1</u>	<u>1</u>	1				1			<u>1</u>	<u>1</u>		1			
<i>L. seeligeri</i> 1/2b str. SLCC3954	1		<u>1</u>	<u>1</u>	1				1			<u>1</u>	<u>1</u>		1			
<i>Staphylococcus aureus</i> Mu50	1		<u>1</u>	<u>1</u>	1				1		1	<u>1</u>			<u>1</u>	2	2	

Table 1 Distribution and GlnR-mediated regulation of genes related to central nitrogen metabolism within the order Bacillales (Continued)

<i>S. carnosus</i> TM300	1	<u>1</u>	2(1)	1	1	1	2	3
<i>S. epidermidis</i> ATCC 12228	1	<u>1</u>	<u>1</u>	1	1	<u>1</u>	1	2
<i>S. haemolyticus</i> JCSC1435	1	<u>1</u>	<u>1</u>	1	1	<u>1</u>	1	2
<i>S. lugdunensis</i> HKU09-01	1	<u>1</u>	<u>1</u>	1	1	<u>1</u>	1	2
<i>S. saprophyticus</i> ATCC 15305	1	<u>1</u>	<u>1</u>	2	1	<u>1</u>	2(1)	4
<i>Mc. caseolyticus</i> JCSC5402	1	<u>1</u>	<u>1</u>	1	1		2	3

The table summarizes the data retrieved for all species and strains which can be found in Additional file 1. The number of orthologs/homologs in every genome is indicated. For instance *Anoxybacillus flavithermus* WK1 contains 2 genes encoding glutamine synthetase. Every species is represented by one strain only. Species abbreviations: *Ab.*, *Alicyclobacillus*; *B.*, *Bacillus*; *G.*, *Geobacillus*; *L.*, *Listeria*; *Mc.*, *Macrococcus*; *Pb.*, *Paenibacillus*; *S.*, *Staphylococcus*.

Gene content. The included genes encode the established regulators CodY, TnrA and GlnR [4]; the central enzymes Gln synthetase (*glnA*; [30]), Glu dehydrogenase (*gdh*; [31,32]), 3-hydroxybutyrate dehydrogenase (*bdh* [33,34]), high affinity (HA) and low affinity (LA) glutaminase, glutamate synthase (NADPH-dependent, large and small subunit *gltB* and *gltA* [21,22]; *B. cereus* type *gltA*; ferredoxin dependent *gltS* [35]), aspartate-ammonia ligase/asparagine synthetase (*asnA* [36] or *asnB* [37]), asparaginase (*ansA* or *ansZ* [38,39]); and the transport systems for ammonium (*amtB* with related regulator *glnK*; [40]), gln/glu using ATP (ABC-family, *glnP^HQ* and *glnQHMP*) and glutamine-like proton symport (AGCS-family, *glnT*, *alsT*, *yrbD*, *yflA*) and glutamate-like proton symport (DAACS-family, *gltP*, *gltT*, *yhcl*, *nqt*). In the latter two cases a variation in numbers was observed for different strains. In the other cases the numbers were far less variable (the ones that were variable are indicated as comma separated numbers).

GlnR-mediated regulation. Genes/operons that have a clear upstream GlnR binding site are marked by dark grey boxes (similarity score >87%), whereas genes/operons that are preceded by a less clear site are marked light grey (similarity score 80–87%). In case more genes are present encoding the same function then the number of genes with a clear binding site is indicated between brackets and if in addition the other gene(s) are preceded by a less clear site then the cell is marked light grey.

- A GlnR-binding site is present but downstream of *tnrA* because the gene has opposite direction when compared to all other species.
- For one sequence no ORF was called. However the gene could be identified using tBLASTN.
- One of the sequences in two fragments.
- Seems part of an operon that includes a gene with DUF294 domain and the gene *dnaQ*.
- The *glnR* and *glnA* duplicates in *B. pseudofirmus* OF4 are not part of a single operon and located at different positions on the genome.

Table 2 Distribution and GlnR-mediated regulation of genes related to central nitrogen metabolism within the order Lactobacillales

NCBI PROJECT NAME \ GENES	<i>codY</i>	<i>tnrA</i>	<i>glnR</i>	<i>glnA</i>	<i>gdh</i>	<i>bdh</i>	glutamin- ase (HA)	glutamin- ase (LA)	<i>gltAB</i>	<i>gltA^{BC}</i>	<i>gltS</i>	<i>amtB</i>	<i>glnK</i>	<i>glnQHMP</i>	<i>glnP^{HQ}</i>	<i>glnT, alsT, yrbD, yfIA</i>	<i>gltP, dctA, gltT, yhcl, nqt</i>	
B) order Lactobacillales																		
<i>Enterococcus faecalis</i> V583	1		<u>1</u>	<u>1</u>	1							1		<u>1</u>	<u>1</u>		3	
<i>Lactobacillus acidophilus</i> NCFM				<u>1</u>								1		2				
<i>L. helveticus</i> DPC 4571				<u>1</u>								1		2			1	
<i>L. crispatus</i> ST1				<u>1</u>				1						2				
<i>L. gasserii</i> ATCC 33323				<u>1</u>				1						2			1	
<i>L. johnsonii</i> F19785				<u>1</u>				1						2			1	
<i>L. bulgaricus</i> ATCC 11842				<u>1</u>								1		3				
<i>L. fermentum</i> IFO 3956				<u>1</u>								1			1		1	
<i>L. reuteri</i> JCM 1112				<u>1</u>			1	1				1	1	1	1		1	
<i>L. brevis</i> ATCC 367			<u>1</u>	<u>1</u>								<u>1</u>		<u>1</u>			1	
<i>L. plantarum</i> WCFS1			<u>1</u>	<u>1</u>	1							<u>1</u>	1		<u>1</u>		1	
<i>L. salivarius</i> UCC118			<u>1</u>	<u>1</u>	1							<u>1</u>	<u>1</u>		2(1)^e		1	
<i>L. casei</i> BL23			<u>1</u>	2(1)	1,0				1			1	1		<u>1^b</u>			
<i>L. rhamnosus</i> GG			<u>1</u>	2(1)	1				1			2(1)	1		<u>1^b</u>			
<i>L. sakei</i> 23 K			<u>1</u>	<u>1</u>											1		2	
<i>Pc. pentosaceus</i> ATCC 25745			<u>1</u>	<u>1</u>											<u>1</u>		1	
<i>Leuconostoc citreum</i> KM20			<u>1</u>	<u>1</u>								<u>1</u>	<u>1</u>		<u>1^b</u>			
<i>Ln. mesenteroides</i> ATCC 8293			<u>1</u>	<u>1</u>					1			<u>1</u>	<u>2</u>		<u>1^b</u>			
<i>Oenococcus oeni</i> PSU-1			<u>1</u>	<u>1</u>								<u>1</u>	1		<u>1^d</u>			
<i>Lactococcus lactis</i> Il1403	1 ^a		<u>1</u>	<u>1</u>					1			<u>1</u>	<u>1</u>		<u>1^c</u>			
<i>Streptococcus mutans</i> UA159	1		<u>1</u>	<u>1</u>	<u>1</u>				1			<u>1</u>	<u>1</u>	<u>1</u>	<u>1^c</u>	1		
<i>S. equi</i>	1		<u>1</u>	<u>1</u>	1										<u>1^c</u>	1		
<i>S. uberis</i> 0140 J	1		<u>1</u>	<u>1</u>	<u>1</u>								<u>1</u>		<u>1^c</u>	1		
<i>S. gordonii</i> str. Challis substr. CH1	1		<u>1</u>	<u>1</u>	<u>1</u>										<u>1^c</u>		1	
<i>S. thermophilus</i> CNRZ1066	1		<u>1</u>	<u>1</u>	<u>1</u>							<u>1</u>	1	<u>1</u>	<u>1^c</u>	1		
<i>S. sanguinis</i> SK36	1		<u>1</u>	<u>1</u>	<u>1</u>									<u>1</u>	<u>1^c</u>	1		
<i>S. pneumoniae</i> R6	1		<u>1</u>	<u>1</u>	<u>1</u>									<u>1</u>	<u>1^c</u>	1		

Table 2 Distribution and GlnR-mediated regulation of genes related to central nitrogen metabolism within the order Lactobacillales (Continued)

<i>S. mitis</i> B6	1	<u>1</u>	<u>1</u>	1	<u>1</u>	1	<u>1</u>	<u>1^c</u>	1
<i>S. agalactiae</i> NEM316	1	<u>1</u>	<u>1</u>	1	1		<u>1</u>	<u>1^c</u>	1
<i>S. pyogenes</i> M1 GAS	1	<u>1</u>	<u>1</u>					<u>1^c</u>	1
<i>S. suis</i> BM407	1	<u>1</u>	<u>1</u>	1			<u>1</u>	<u>1^c</u>	1

The table summarizes the data retrieved for all species and strains which can be found in Additional file 1. The number of orthologs/homologs in every genome is indicated. For Gene content and how GlnR-mediated regulation is indicated see the legend of Table 1. Every species is represented by one strain only. Species abbreviations: *L.*, *Lactobacillus*; *Ln.*, *Leuconostoc*; *Pc.*, *Pediococcus*; *S.*, *Streptococcus*.

a) *L. lactis* has a *codY* paralog named *codZ* [41].

b) There is a second *glnP^H* gene located next to the *glnP^H,Q* operon.

c) The *glnP^H* gene encodes two H domains.

d) In *O. oeni* the gene *glnP^H* is in one operon with a gene encoding an asparaginase, whereas a *glnQ* homolog is missing.

e) In *L. salivarius* a second *glnP^H* gene is not part of an operon and is located elsewhere on the genome with respect to the *glnP^H,Q* operon.

system has been related to high-affinity glutamine transport [16], whereas the *L. lactis* system was shown to transport both glutamine and glutamate [44]. Recently, the *E. coli*-type system present in *Streptococcus mutans* was proven also to be involved in the transport of glutamate [45]. Remarkably, most of the species of the order Lactobacillales carry a copy of both types (Table 2). These Lactobacilli lack the genes encoding a glutamate dehydrogenase (*gdh*) or glutamate synthetase (*gltAB*). Therefore these species are unable to synthesize glutamate, which makes it essential to have a glutamate transport system.

Most of the analyzed species encode one or more transporters of the DAACS-family (2.A.13) and AGCS-family (2.A.25), with the exception of the species within the families Listeriaceae and Leuconostocaceae, some Lactobacillaceae and *L. lactis*. These transporter-protein families have been related to the cation symport of dicarboxylates and amino acids. The former family is represented by GltP (glutamate/aspartate [46]), GltT (glutamate [47]), DctA (C4-dicarboxylates including aspartate [48]), YhcL (or TcyP; cystine [49]) and Nqt (putative glutamate in *B. subtilis*), whereas the latter family is represented by GlnT (glutamine [50]), AlsT (amino acid [15]), YrbD (putative amino acid) and YflA (putative amino acid).

Identification of a GlnR and TnrA specific binding motif

The protein sequences of GlnR and TnrA are highly similar and their reported DNA binding sites show little difference [27]. The palindromic consensus sequence has been defined as TGTNA-N7-TNACA [13,15,51-55]. Gel mobility shift assays indicated that TnrA and GlnR indeed bind to the same sites upstream of the *tnrA* gene and the *glnRA* operon, albeit with different specificity [13]. To achieve a separation of the predicted sites we have employed a genomic footprinting strategy that we formulated previously [28,56] to identify the GlnR-specific binding motif anew. The strategy involved the definition of Groups Of Orthologous Functional Equivalents (GOOFEs) on basis of conserved genomic context. Within these GOOFEs we assumed conservation of binding motif. In all analyzed species that contain *glnR*, the genetic association with *glnA* has been conserved. Moreover, for several species GlnR was shown experimentally to be autoregulatory and therefore the upstream region of the *glnRA* operon within all genomes was scanned for a conserved GlnR binding site. In line with earlier published observations we found a clear and strongly conserved binding site 3–6 nucleotides upstream of a putative-35 region (i.e. TTGAC) of the promoter in all analyzed species and a second binding site overlapping the promoter in many of the *Bacillus* species (Figure 2).

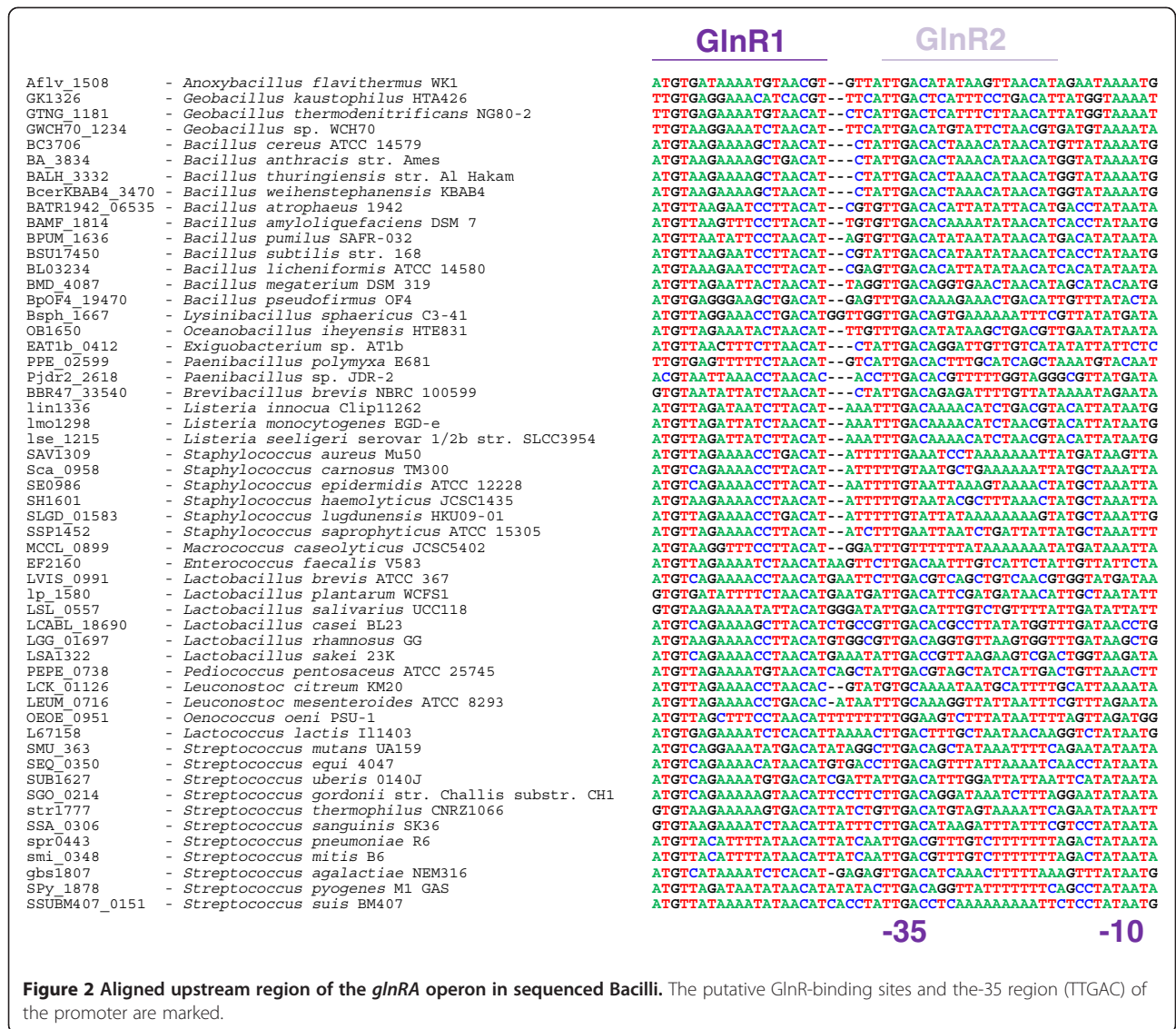
It was shown in a cross-regulation study that the binding site upstream of the promoter of the *glnRA* operon in *B. subtilis* is only involved in GlnR-mediated regulation [13].

Therefore, to tract potential differences between the GlnR and TnrA binding motifs, we used the conserved GlnR-binding sites upstream of the promoter to generate a family specific position frequency matrix (see methods). It appeared that the frequency representations of the motif varied slightly between the Streptococci and the other Bacilli (Figure 3A and B). Both motifs that were generated for GlnR adhered to the consensus motif [13,15,51-55] and were similar to the motif that was previously defined by [27]. Then, a TnrA-specific motif was created on basis of the TnrA sites upstream of *amtB*, *ansZ*, *gapP*, *glnQ*, *nasA*, *nasB*, *nasD*, *oppA*, *pucJ*, *pucR*, *yzkB* and *ywrD*. These binding sites were reported to relate to transcription activation in *B. subtilis* ([15,19,38,52,57-59] and raw data file 1) and are supposed to be TnrA-specific as GlnR has not been reported to activate transcription. The frequency representation of the TnrA-specific motif is given in Figure 3C. A comparison of the GlnR and TnrA specific motifs shows that there is limited difference. Yet the TnrA motif clearly lacks the conserved A and T at the 3' and 5' end, as was noted before. In fact, mutation of the conserved T at the 5' end to a C or a G (but not A) was reported to abolish GlnR-mediated repression of the *glnRA* operon in *B. subtilis* [55], although [13] did not observe such an effect. Our new motifs also suggest that there is a slight preference for a G at position 7 and a C at position 13 which is less pronounced in the GlnR motif in Streptococci.

The predicted GlnR and TnrA regulon of *B. subtilis*

The specific GlnR and TnrA motif were used to search the *B. subtilis* genome for similar sites using the Similar Motif Search (SMS) procedure described in the methods. The results of this search can be found in Additional file 2. Although the differences in GlnR and TnrA motif did not appear strong in first instance, the results of the motif search in *B. subtilis* suggest they are large enough to bring about some separation between GlnR and TnrA binding sites, in line with the observed variable affinities of these transcription factors for the same sites [13].

In principle the highest scoring sites are likely to be genuine binding sites and by using a relatively high cut-off score of 0.89 the majority (>70%) of experimentally validated sites was indeed captured for GlnR as well as for TnrA. Moreover, most other true binding sites scored just below the cut-off. Only 4 out of 22 reported TnrA binding sites were not recovered in this way. Some of the sites were actually found at a relatively large distance from the translation start (e.g. in the case of *ilvB* [60]) and many sites were found located in the shared regulatory region of neighboring genes located on opposite strands (so-called divergents). Although many sites were found in both searches, the similarity score was mostly clearly better for one than for the other. Genes/operons predicted to be controlled by both regulators included the known genes/operons *glnRA*



(glutamine synthesis) and *tnrA*. Additional shared sites were found upstream of *alsT*, *pucH*, *pucJKLM* and the *amtB-glnK* operon. Although these sites have not been attributed to GlnR earlier and were described to be activated by TnrA [15,57], the relatively high similarity score and the evolutionary conservation, also among organisms that lack TnrA, suggest they are true binding sites. In the case of the *amtB-glnK* operon (import of ammonia) it was formerly concluded that it is not repressed by GlnR on the basis of a singular observation. It was found that the *amtB-glnK* operon remained repressed in a GlnR deletion mutant (i.e. *glnR57* [12]) in the presence of glutamine, similar to the wild-type [61]. However this observation does not exclude repression by GlnR in case additional regulators are at play. In fact, in *L. lactis* it was shown that expression of the *amtB-glnK* operon is controlled by GlnR but also by CodY [62]. In *S. mutans* it was shown by electrophoretic

mobility shift assay that GlnR binds to the promoter region of both the *glnRA* and *amtB-glnK* operon [11]. The same study identified GlnK as an activator of GlnR DNA-binding. Besides, the data in Tables 1 and 2 indicate that a putative GlnR-binding site upstream of *amtB* is present across almost all species of the class *Bacilli*. The conservation of these putative binding sites, including the conservation of the flanking A and T nucleotides (see Additional file 3), suggests GlnR represses the *amtB-glnK* operon in all analyzed species, and thus also in *B. subtilis*.

It was proposed that GlnR lacks the capability to recruit RNA polymerase and therefore acts solely as a repressor [13]. Given this lack of activating/recruiting capacity, it is to be expected that GlnR will only act on the expression of one gene in various divergences, like for instance on *tnrA* but not *ykzB* [58] and on *pucH* but not *pucR* [59]. Moreover, in various cases where TnrA was shown to activate

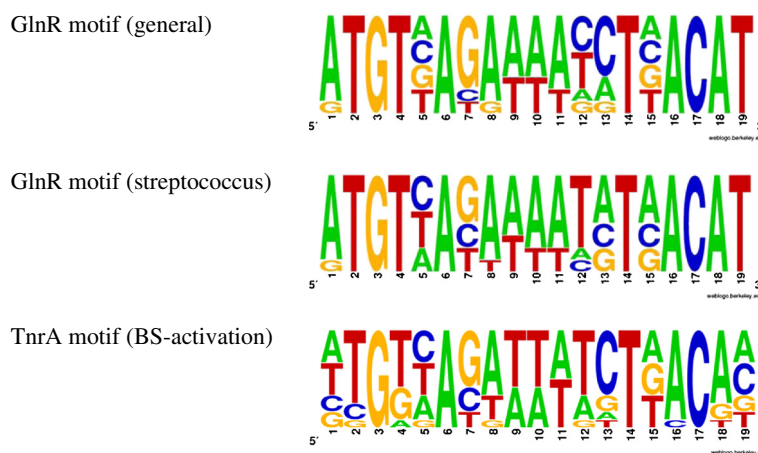


Figure 3 Frequency representation of the putative GlnR and TnrA binding motifs. **A)** GlnR-binding motif in species of the class Bacilli. **B)** GlnR-binding motif in species of the family Streptococci. **C)** TnrA-binding motif in *Bacillus subtilis*.

transcription our analysis suggests the binding site is TnrA-specific, like for *gabP*, *oppABCDF* and *glnQHMP*. Other genes/operons that seem to be TnrA-specific in *B. subtilis* include: *yclG*, encoding an uronase; *ydaB*, encoding an acyl-CoA ligase; *pel*, encoding a pectate lyase; *braB*, encoding a branched-chain amino acid-Na⁺ symporter; and *pucABCDE*, encoding a xanthine dehydrogenase operon.

The predicted GlnR regulon in oral Streptococci

GlnR binding-site predictions were performed for the oral Streptococci *S. pneumoniae* and *S. mutans* on basis of the Streptococci-specific motif (results in Additional file 4). For *S. pneumoniae* D39 and *S. mutans* UA159, the genes/operons predicted to be controlled by GlnR were compared to the genes/operons whose transcription was most affected in a GlnR mutant [25,26]. We found good agreement between prediction and experiment for both organisms (see Table 3). In the case of *S. pneumoniae* D39, the most significantly up-regulated genes/operons, *glnP^{HH}Q* and *gdh* were represented by the best hits in our analysis. The analysis also revealed the presence of a clear binding site in front of 2 other genes/operons, in line with the predictions of [27]. These included the second glutamine ABC transporter (*glnQHMP*) and an operon containing enzymes of the urea cycle (*arcAB*). The clear regulatory connection between GlnR and the *arcAB* operon (encoding arginine deiminase and ornithine carbamoyltransferase) was found in all sequenced *S. pneumoniae* strains, but was absent in the other species. The absence of a change in *arcAB* and *glnQHMP* expression upon inactivation of *glnR* may be explained by the presence of additional regulatory interactions.

In the case of *S. mutans* UA159 the genes and operons found to be mostly affected in the knockout mutant [26]

were the *nrgA*-SMU_1657c operon, coding for the ammonium transporter AmtB and its nitrogen regulatory protein GlnK; the *citB-citZ-idh* operon coding for aconitate hydratase, citrate synthase and isocitrate dehydrogenase; the *glnQHMP* and *glnP^{HH}Q* operons encoding glutamine ABC transporters; and Smu.807, coding for a putative membrane protein, which is in a divergon with *glnP^{HH}Q*. The best hits resulting from our analysis are also located upstream of the same operons. Moreover, we found a clear binding site preceding the genes *gdh* and *thrC*. The *citB-citZ-idh* operon has been shown to be essential for glutamate biosynthesis in *S. mutans* [24].

Conserved genetic associations of GlnR and the effect of the other regulators

GlnR binding-site predictions were performed for selected genomes that represented all sequenced species of the class Bacilli. We then collected the function annotation of all proteins encoded by genes/operons downstream of a putative GlnR-binding site that fitted the selection criteria (see methods) to generate an overview of the regulatory connections that are conserved between more than three species (accumulated in Additional file 4). The results are summarized in Tables 1, 2 and 4. As expected, we found a conserved regulatory connection between GlnR and the *glnRA* operon in all analyzed species and with *tnrA* in all Bacilli. Only in a few species of the order Bacillales the related GlnR-binding sites deviated from the consensus (e.g. in some *Geobacillus* species). Another connection that was conserved in almost all of the analyzed species was that with *amtB* (often *amtB-glnK*).

Various additional conserved connections were found, although these appeared far more family-specific. For instance, in the order Lactobacillales a genomic association with the genes of the two glutamine ABC transporter

Table 3 Predicted GlnR regulon in *Streptococcus pneumoniae* (A) and *Streptococcus mutans* (B)

gene/operon name	locus tag*	annotation	score (rank)	rank in exp.
A) <i>Streptococcus pneumoniae</i>				
<i>glnR</i> ,A	SP_0501,0502	GlnR regulator, Glutamine synthetase	0.95 (1)	k.o.
<i>glnP^{HI}</i> ,Q	SP_1241,1242	Gln/Glu ABC transport system	0.92 (2)	1,2, SP_1243: 3
<i>gdh</i>	SP_1306	Glutamate dehydrogenase	0.90 (5)	4
<i>arcA,B</i> ^a	SP_2148,2150	Arginine deiminase, Ornithine carbamoyltransferase	0.91 (3) ^b	-
<i>glnQ,H,M,P</i> ^a	SP_0610-0607	Gln/Glu ABC transport system	0.84 ^b	-
B) <i>Streptococcus mutans</i> (UA159)				
<i>glnR</i> ,A	Smu.363,364	GlnR, Glutamine synthetase	0.95 (1)	k.o.
<i>amtB,glnK</i>	Smu.1658,1657c	Ammonium transport system, GlnK regulator	0.95 (1)	1
<i>glnQ,H,M,P</i>	Smu.1519-1522	Gln/Glu ABC transport system	0.95 (2)	2
<i>citB,Z,idh</i>	Smu.670,671,672	Aconitate hydratase, Citrate synthase, Isocitrate dehydrogenase	0.95 (1)	3
<i>Smu.807</i>	Smu.807	Putative membrane protein	0.95 (1)	4
<i>glnP^{HI}</i> ,Q	Smu.806c,805c	Gln/Glu ABC transport system	0.95 (1)	5
<i>gdh</i>	Smu.913	Glutamate dehydrogenase	0.95 (1)	-
<i>Smu.68, thrC</i>	Smu.68,70	Hypothetical, Threonine synthase	0.92 (5)	-

The GlnR-binding site identifications were made as described in the methods for all strains with a published genome (data in Additional file 4). The composition of the regulon appeared identical between strains although the similarity scores of particular binding sites varied slightly. The table lists the numbers obtained with strains TIGR4 and UA159, respectively. Column 4: The ranking is based on the scores obtained with the similar motif search procedure, which provides various sites with identical scores and thus identical ranking. The absence of certain high scoring sites (e.g. ranked 4) was caused by the conservative criteria we applied for a site to qualify as a putative binding site. Column 5 gives the observed ranking on basis of the transcriptional response towards a *glnR* knock out mutation (k.o.) as derived from [25] and the ranking of operons that are downregulated in a *S. mutans glnR* knock out after exposure to acid stress for 30 minutes [26]. The ranking score was calculated by dividing the fold change in the *glnR* mutant by the fold change in the wild-type. Column 1 (a): In some strains the *arcA* and *glnQ* ORFs have not been called. The published sequence of strain R6 suggests *arcA* and *glnQ* are truncated in this strain. Column 4 (b): The composition of the binding site upstream of *arcAB* and *glnQHMP* varies between strains.

* Although [25] studied the effects in strain *D39*, the used microarrays were based on strains *TIGR4* and *R6*. For reasons of comparison we have therefore listed the locus tags in strain *TIGR4*.

encoding variants *glnP^{HI}*Q or *glnQHMP* were identified, whereas this association appeared to be replaced by one with the sodium/proton amino acid symporters encoded by *gltT* (glutamate, [47]) and *alsT* [15] in various species within the family of the Bacillaceae. The AlsT protein is very similar to GlnT, a cation-glutamine symporter, i.e. showing a high degree of sequence conservation and having about the same length and the same number of predicted transmembrane helices. Although the protein is sometimes referred to as an alanine transporter, AlsT could well be a cation-glutamine or asparagine symporter.

Other associations that were conserved included that with the *gdh* gene (encoding glutamate dehydrogenase) and the *citBZ-icd* operon (encoding aconitase, citrate synthase and isocitrate dehydrogenase [24]) in many Streptococci, the *ureABC* operon (encoding urease [64]) and the *nasDEF* operon (encoding nitrite reductase [18]) in several *Bacillus* species and with several genes involved in purine and asparagine/aspartate transport and metabolism (e.g. *pucl* and *puclH* (encoding an allantoin transporter and allantoinase, respectively [20]), *ywoCD* (encoding an amidase and a transport protein), *ansA* and *ansZ* (both encoding asparaginase [38], *aspA* (encoding aspartase) and *asnA* (encoding aspartate-ammonia ligase).

We also found a clear GlnR-binding site upstream of several genes involved in regulation, for instance of *mcp* (chemotaxis, found in several *Geobacillus* species) and of the *ycsFGI-kipIAR-yksK* operon (cellular development, found in several *Bacillus* species). In the initial description of the *ycsFGI-kipIAR-yksK* operon [65], *yksF* was related to the lactam (e.g. 2-pyrrolidinone) utilization gene *lamB* of *Aspergillus nidulans* [66], and *kipA* (orf12) was related to a urea amidolyase of yeast. Later, KipI was identified as a protein inhibitor of auto-phosphorylation of kinase A, the sensor histidine kinase responsible for processing post-exponential phase information and for providing phosphate input to the phosphorelay that activates developmental transcription via phosphorylated Spo0A, and KipA as a protein that counteracts the inhibition [63]. YcsG showed similarity to BraB (branched chain amino acid transport system II) of *Pseudomonas aeruginosa* [67]. The operon was found repressed upon growth on good nitrogen sources like ammonia and glutamine and derepressed on poor nitrogen sources [68], in line with repression mediated by GlnR. The association with the *ycsFGI-kipIAR-yksK* operon connects GlnR-mediated regulation to the regulation of sporulation in some *Bacilli*.

Table 4 Distribution and function of additional genes-operons that are putatively linked to GlnR-mediated regulation

species	gene/operon	annotation (literature)
<i>A. flavithermus</i> , <i>G. C56-R3</i> , <i>G. kaustophilus</i> , <i>G. Y412MC61</i> , <i>L. seeligeri</i> , <i>P. JDR2</i>	<i>mcp</i>	regulation of chemotaxis
<i>B. megaterium</i> , <i>G. kaustophilus</i> , <i>G. Y412MC61</i> , <i>P. JDR2</i> , <i>P. polymyxa</i>	<i>pucR</i>	regulation of <i>puc</i> operon (purine catabolism) [20]
<i>B. amyloliquefaciens</i> , <i>B. atrophaeus</i> , <i>B. licheniformis</i> , <i>B. megaterium</i> , <i>B. pumilus</i>	<i>ycsFGI-kiplAR-ycsK</i>	includes regulation of sporulation [63]
<i>B. amyloliquefaciens</i> , <i>B. atrophaeus</i> , <i>L. salivarius</i>	<i>gabP</i>	transport of gamma-aminobutyrate [19]
<i>A. flavithermus</i> , <i>B. cereus</i> , <i>L. rhamnosus</i> , <i>O. oeni</i> , <i>P. JDR2</i> , <i>P. polymyxa</i> , <i>S. gordonii</i> , <i>S. sanguinis</i>	<i>opp</i> , <i>dpp</i>	transport of oligopeptide and/or dipeptide
<i>B. amyloliquefaciens</i> , <i>B. atrophaeus</i> , <i>B. megaterium</i> (2x), <i>B. subtilis</i> , <i>G. kaustophilus</i> , <i>G. Y412MC61</i>	<i>pucl</i>	transport of allantoin
<i>B. amyloliquefaciens</i> , <i>B. atrophaeus</i> (& <i>ywoB</i>), <i>B. subtilis</i> (& <i>ywoB</i>), <i>P. JDR2</i> (- <i>ywoD</i>)	<i>ywoCD</i>	transport (<i>ywoD</i>) and hydrolysis (<i>ywoC</i>)
<i>B. anthracis</i> , <i>B. cereus</i> , <i>B. thuringiensis</i> , <i>B. weihenstephanensis</i> , <i>P. JDR2</i> (<i>cpdB</i> only)	K01238<> <i>cpdB</i>	hydrolysis nucleoside 2',3'-cyclic phosphate + H ₂ O <=> nucleoside 3'-phosphate
<i>L. seeligeri</i> , <i>O. oeni</i> , <i>Pb JDR2</i> , <i>P. polymyxa</i>	<i>ansA,ansZ</i>	L-asparagine + H ₂ O <=> L-aspartate + NH ₃ [38]
<i>B. amyloliquefaciens</i> , <i>B. atrophaeus</i> , <i>B. licheniformis</i> , <i>B. pumilus</i> , <i>B. subtilis</i>	<i>argC</i>	N-Acetyl-L-glutamate 5-semialdehyde <=> N-Acetyl-L-glutamate 5-phosphate
<i>L. plantarum</i> , <i>L. rhamnosus</i> , <i>L. citreum</i>	<i>aspA</i>	L-aspartate <=> fumarate + NH ₃
<i>S. mutans</i> , <i>S. gordonii</i> , <i>S. sanguinis</i> , <i>S. suis</i> , <i>S. thermophilus</i>	<i>citZ-citB-idh</i>	conversion of citrate to alpha-ketoglutarate [24]
<i>B. amyloliquefaciens</i> , <i>B. atrophaeus</i> , <i>B. megaterium</i> , <i>B. pseudofirmus</i> , <i>G. C56-T3</i> , <i>G. kaustophilus</i> , <i>G. Y412MC61</i> , <i>P. sp. JDR2</i>	<i>nasDEF</i>	reduction of nitrite to NH ₃ [18]
<i>B. atrophaeus</i> , <i>B. subtilis</i> , <i>G. Y412MC61</i>	<i>pucH</i>	allantoin + H ₂ O <=> allantoinate [20] [<i>pucF</i> : allantoinate <=> ureidoglycolate + 2 NH ₃]
<i>B. anthracis</i> (BC), <i>B. cereus</i> (C), <i>B. thuringiensis</i> (BC), <i>B. weihenstephanensis</i> (C), <i>S. mutans</i> (C)	<i>thrBC</i>	L-homoserine + ATP <=> O-phosphohomoserine + ADP O-phosphohomoserine + H ₂ O <=> threonine + P [<i>ilvA</i> : threonine <=> 2-oxobutanoate + NH ₃] [24]
<i>B. amyloliquefaciens</i> , <i>B. megaterium</i> , <i>B. pseudofirmus</i> , <i>B. subtilis</i> , <i>P. JDR2</i>	<i>ureABC</i>	urea + H ₂ O <=> CO ₂ + 2 NH ₃ [64]

The related data can be found in Additional file 4.

Another important finding was that in *B. subtilis* many operons related to the purine degradation pathway are controlled by GlnR and/or TnrA, like *pucABCDE*, *pucH*, *pucl*, *pucJKLM* and *ureABC*. The relation between purine catabolism and control by TnrA was established before experimentally. It was observed that a *tnrA* mutant strain could not use purines or its metabolic intermediates as a nitrogen source during nitrogen limited conditions [20]. Nevertheless, the extent to which both GlnR and TnrA are connected to the related operons is surprising.

We observed no clear dependency between the composition of the predicted GlnR regulon and the presence or absence of the other nitrogen-related regulators CodY and TnrA. For instance, there are only a few differences between the predicted GlnR regulon of *B. subtilis* and *B. cereus* suggesting that GlnR does not take over regulatory roles of TnrA. Similarly, the presence or absence of CodY does not seem to affect the size of the GlnR regulon in the Lactobacillaceae. In *L. lactis*, a species that

has CodY, it was shown experimentally that at least three genes/operons (*amtB*, *glnRA* and *glnP^{HH}*, *Q*) are repressed by GlnR [62]. We indeed identified clear GlnR-binding sites in the upstream region of these three genes/operons in *L. lactis*. In *L. plantarum* and *L. monocytogenes*, two species that lack CodY, the same genes/operons appear to be preceded by a GlnR-binding site and only a few additional genes were found connected to GlnR indicating that GlnR does not take over the role of CodY in these species. The predicted GlnR regulon was smallest, consisting of only *glnRA*, in the meat isolates *Macrococcus caseolyticus* (a CodY and GlnR containing *Staphylococcus*) and *L. sakei* (a GlnR containing *Lactobacillus*).

Conclusions

We have analyzed all sequenced Bacilli for the presence of genes encoding central nitrogen metabolism and transport of the related metabolites, and identified their connection

to the nitrogen metabolism regulator GlnR. Although there is a considerable variety in the presence of the central enzymes GS, G, GDH and GOGAT, and in the number of available transport systems for the central nitrogen-related metabolites, the composition of the GlnR regulon is relatively invariable between species. Moreover, we hardly found an effect of the absence or presence of the other regulators CodY, TnrA and GltC on the size of the predicted GlnR regulon.

We made an initial conservative regulon prediction by restricting the regulatory association to those connections that are conserved between at least three species. In general, our findings are also in line with previous comparative in silico analysis performed on a limited number of species [27]. Careful redefinition of a specific GlnR-binding and a specific TnrA-binding motif caused a slight but clear separation in the predicted regulons. It is likely that the conserved A/Ts at the 3' and 5' end of the GlnR motif, which are absent in the TnrA motif, contribute significantly to the separation. For *B. subtilis*, *S. pneumoniae*, *S. mutans* and *L. lactis* our predictions complied with the available experimental data. Moreover, within the Bacilli we identified several new potential members of the GlnR regulon, including the *ywoCD* operon and the *ycsFGI-kiplAR-ycsK* operon.

Our analysis confirmed that for most species the size of the GlnR regulon is relatively small. The main regulatory associations in the species of the class Bacilli are with the incorporation of ammonium into central metabolism (or with the production of ammonium at high glutamine concentrations!) via glutamine synthetase (*glnRA* operon), and with ammonium (*amtB-glnK*) and glutamine/glutamate transport (i.e. via *glnQHMP*, *glnPHQ*, *gltT*, *alsT*). At the same time, the lesser conserved associations point to a somewhat broader role. Many of the conserved associations include genes that are either directly (e.g. *ansA*, *arcA*, *aspA*, *gdh*, *nasDEF*, *ureABC*) or more indirectly (by controlling intermediate steps, e.g. *citBZ-idh*, *pucH*, *thrBC*) relate to the intracellular production of ammonia or are related to the import of aminated compounds (e.g. *gabP*, *opp-dpp*, *pucI*, *ywoCD*). Thus, it appears that the main conserved role of GlnR is to prevent the influx and intracellular production of glutamine and ammonium under conditions of high nitrogen availability. The connection of GlnR-mediated repression with the control of intracellular ammonia concentration is interesting. Such a role fits with the intrinsic need for tight control of ammonia levels as put forward by [43], who argue that transport of ammonia (NH_4^+) should be tightly regulated to limit futile cycling by diffusion of ammonia out of the cell.

Methods

Data and tools

Complete genomic sequences and initial annotations were obtained from NCBI ([68]; June 2011). Multiple sequence

alignments were made with ClustalX [69], and BioEdit [70] was used to analyze sequences and alignments. Specific bootstrapped neighbor-joining trees, with 'correction for multiple substitutions', were created using ClustalX and the trees were analyzed using LOFT [71] and Dendroscope [72]. The Microbial Genome Viewer 2.0 (<http://mgv2.cmbi.ru.nl>) was used to examine the function information within the genomic context. Frequency representations of aligned sequences were created with Weblogo [73]. Microarray data from *glnR* gene knockouts in *Streptococcus pneumoniae* and *Streptococcus mutans* used in this research were extracted from the Gene Expression Omnibus from NCBI [74].

The raw data resulting from the various analyses can be found at http://www.cmbi.ru.nl/bamics/supplementary/GrootKormelinketal_2012_GlnRregulon/. Data file 1: GlnR and TnrA motifs used for SMS; data file 2: GlnR motif search in *Bacilli* (w/o *Streptococci*); data file 3: GlnR motif search in *Streptococci*; data file 4: GlnR and TnrA motif search in *Bacillus subtilis*.

Classification and annotation of protein sequences

To obtain all proteins of a certain family, a prominent representative was chosen (listed in the legend of Table 1) and a BLAST search [75] was performed (cut-off $< e^{-5}$) on all publicly available sequenced Bacilli genomes. Then the list of collected sequences (given as Additional file 1) was inspected. For all enzymes, the sequences could be grouped into specific clusters based on BLAST e-value only. In practice we found a group of sequences with comparable (very) low e-values ($< e^{-30}$) separated from the rest of the sequences with considerably higher e-values (separation $< e^{-15}$). In the case of the transcription regulators the separation remained clear, however with higher e-values due to a short length of the regulator protein sequence. In the case of the transporters, for the Amt-family (1.A.11), the DAACS-family (2.A.13) and the AGCS-family (2.A.25) an e-value cut-off also sufficed to collect all family members, whereas for the PAAT-family; 3.A.1.3 (ABC transport) the coding sequences of the putative glutamine/glutamate substrate binding domains were aligned, the alignment was inspected by eye, deviant sequences were removed, and a bootstrap neighbour joining tree was generated (see [76]). The tree was divided into clusters on basis of the branching. For each cluster, single species representatives were considered orthologous.

Motif definition and Similar Motif Scoring (SMS)

The upstream region of the conserved *glnRA* operon was retrieved for all species and the promoter region was aligned (see Figure 2). Then the conserved sequence upstream of the promoter was collected. For *B. subtilis* it was shown that this site is only involved in GlnR-mediated regulation [13]. The collection was used to generate a position frequency scoring matrix for each taxonomic family

(raw data file 1). It appeared that the frequency matrix was very similar for all species except for the Streptococci, where it was slightly different (illustrated in Figure 3). For the definition of a TnrA-specific binding site the upstream regions of genes whose transcription was shown to be activated by TnrA in *B. subtilis* (raw data file 1) were retrieved and the binding site was identified on basis of the published characteristic GlnR/TnrA motif and the short distance upstream of the promoter. The collection of sites was then used to generate a position frequency scoring matrix.

The position frequency matrices were used to identify potential binding sites in the analyzed genomes using a similarity search method formulated before by us [29]. The method relies on the fact that one of the most common practices observed in literature to reconcile prediction with experiment is to minimize the number of differences between the target and the query (or the 'consensus'). In fact, this criterion can be captured in a straightforward scoring using only the position frequency matrix: Given any number of input sequences of size i , the nucleotide frequency $f_{N(j)}$ (where $N \in A,C,T,G$; and frequency is in terms of fraction) at every position $j = 1$ to i can be used directly to provide all target sequences of size i with a score by just adding up the input-based frequencies that relate to the nucleotide composition of the target. Division of the score by the length of the sequence i results in a 'similarity' score that can range from 0 to 1. Dividing this number by the highest attainable score given the input matrix then yields a relative 'similarity' score. In case the input sequences are representative for high-affinity sites, the ranking of target sequences according to score should approximately correspond to a ranking based on affinity. The method was tested and appeared at least as good to identify putative regulatory elements on basis of known input motifs as the commonly used tool MAST [77], yet providing a similarity score that is far easier to interpret and use. We identified putative GlnR regulon members for all species on the basis of two simple criteria: i) a relative similarity score $>87\%$; and ii) a position between 250 and 0 bases upstream of the predicted translation start. In some cases experimentally verified more distant sites were also included as well as known intergenic sites. The results are given in Additional file 4.

Additional files

Additional file 1: Distribution of genes related to central nitrogen metabolism.

Additional file 2: *Bacillus subtilis* GlnR and TnrA regulon.

Additional file 3: GlnR motif search in *amtB* upstream regions.

Additional file 4: Conserved GlnR-mediated regulation.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This project was carried out within the research program of the Kluuyer Centre for Genomics of Industrial Fermentation which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

Author details

¹Kluuyer Centre for Genomics of Industrial Fermentation, P.O. Box 5057, 2600 GA Delft, The Netherlands. ²Ti Food and Nutrition, P.O. Box 557, 6700 AN Wageningen, The Netherlands. ³Laboratory of Microbiology, Wageningen University & Research Centre, Dreijenplein 10, 6700 HB Wageningen, The Netherlands. ⁴Netherlands Bioinformatics Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands. ⁵Centre for Molecular and Biomolecular Informatics, NCMLS, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.

Authors' contributions

TGK conceived and designed the study, identified the GlnR and TnrA motifs and predicted their regulons, and drafted and revised the manuscript. EK helped designing the study and performed motif searches. YH implemented and validated the motif search algorithm. LO helped designing the study and revising the manuscript. RJS and WMV coordinated the study and helped revising the manuscript. CF coordinated the study and helped designing it, carried out the classification and annotation of protein sequences and helped drafting and revising the manuscript. All authors read and approved the final manuscript.

Received: 23 December 2011 Accepted: 20 March 2012

Published: 18 May 2012

References

1. Deuel TF, Prusiner S: Regulation of glutamine synthetase from *Bacillus subtilis* by divalent cations, feedback inhibitors, and L-glutamine. *J Biol Chem* 1974, **249**(1):257-264.
2. Wray LV Jr, Zalieckas JM, Fisher SH: *Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell* 2001, **107**(4):427-435.
3. Eisenberg D, Gill HS, Pfluegl GM, Rotstein SH: Structure-function relationships of glutamine synthetases. *Biochim Biophys Acta* 2000, **1477**(1-2):122-145.
4. Fisher SH: Regulation of nitrogen metabolism in *Bacillus subtilis*: vive la difference! *Mol Microbiol* 1999, **32**(2):223-232.
5. Bohannon DE, Sonenshein AL: Positive regulation of glutamate biosynthesis in *Bacillus subtilis*. *J Bacteriol* 1989, **171**(9):4718-4727.
6. Belitsky BR, Janssen PJ, Sonenshein AL: Sites required for GltC-dependent regulation of *Bacillus subtilis* glutamate synthase expression. *J Bacteriol* 1995, **177**(19):5686-5695.
7. Sonenshein AL: Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Microbiol* 2007, **5**(12):917-927.
8. Wray LV Jr, Fisher SH: *Bacillus subtilis* GlnR contains an autoinhibitory C-terminal domain required for the interaction with glutamine synthetase. *Mol Microbiol* 2008, **68**(2):277-285.
9. Kayumov A, Heinrich A, Fedorova K, Ilinskaya O, Forchhammer K: Interaction of the general transcription factor TnrA with the PII-like protein GlnK and glutamine synthetase in *Bacillus subtilis*. *FEBS J* 2011, **278**(10):1779-1789.
10. Heinrich A, Woyda K, Brauburger K, Meiss G, Detsch C, Stulke J, Forchhammer K: Interaction of the membrane-bound GlnK-AmtB complex with the master regulator of nitrogen metabolism TnrA in *Bacillus subtilis*. *J Biol Chem* 2006, **281**(46):34909-34917.
11. Castellen P, Rego FG, Portugal ME, Benelli EM: The *Streptococcus mutans* GlnR protein exhibits an increased affinity for the *glnRA* operon promoter when bound to GlnK. *Braz J Med Biol Res* 2011, **44**(12):1202-1208.
12. Schreier HJ, Brown SW, Hirschi KD, Nomellini JF, Sonenshein AL: Regulation of *Bacillus subtilis* glutamine synthetase gene expression by the product of the *glnR* gene. *J Mol Biol* 1989, **210**(1):51-63.
13. Zalieckas JM, Wray LV Jr, Fisher SH: Cross-regulation of the *Bacillus subtilis* *glnRA* and *tnrA* genes provides evidence for DNA binding site discrimination by GlnR and TnrA. *J Bacteriol* 2006, **188**(7):2578-2585.

14. Brandenburg JL, Wray LV Jr, Beier L, Jarmer H, Saxild HH, Fisher SH: Roles of PucR, GlnR, and TnrA in regulating expression of the *Bacillus subtilis* ure P3 promoter. *J Bacteriol* 2002, **184**(21):6060–6064.
15. Yoshida K, Yamaguchi H, Kinehara M, Ohki YH, Nakaura Y, Fujita Y: Identification of additional TnrA-regulated genes of *Bacillus subtilis* associated with a TnrA box. *Mol Microbiol* 2003, **49**(1):157–165.
16. Nohno T, Saito T, Hong JS: Cloning and complete nucleotide sequence of the *Escherichia coli* glutamine permease operon (*glnHPQ*). *Mol Gen Genet* 1986, **205**(2):260–269.
17. Wray LV Jr, Zalieckas JM, Ferson AE, Fisher SH: Mutational analysis of the TnrA-binding sites in the *Bacillus subtilis* *nrgAB* and *gabP* promoter regions. *J Bacteriol* 1998, **180**(11):2943–2949.
18. Nakano MM, Hoffmann T, Zhu Y, Jahn D: Nitrogen and oxygen regulation of *Bacillus subtilis* *nasDEF* encoding NADH-dependent nitrite reductase by TnrA and ResDE. *J Bacteriol* 1998, **180**(20):5344–5350.
19. Ferson AE, Wray LV Jr, Fisher SH: Expression of the *Bacillus subtilis* *gabP* gene is regulated independently in response to nitrogen and amino acid availability. *Mol Microbiol* 1996, **22**(4):693–701.
20. Schultz AC, Nygaard P, Saxild HH: Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. *J Bacteriol* 2001, **183**(11):3293–3302.
21. Belitsky BR, Wray LV Jr, Fisher SH, Bohannon DE, Sonenshein AL: Role of TnrA in nitrogen source-dependent repression of *Bacillus subtilis* glutamate synthase gene expression. *J Bacteriol* 2000, **182**(21):5939–5947.
22. Picossi S, Belitsky BR, Sonenshein AL: Molecular mechanism of the regulation of *Bacillus subtilis* *gltAB* expression by GltC. *J Mol Biol* 2007, **365**(5):1298–1313.
23. Mader U, Hennig S, Hecker M, Homuth G: Transcriptional organization and posttranscriptional regulation of the *Bacillus subtilis* branched-chain amino acid biosynthesis genes. *J Bacteriol* 2004, **186**(8):2240–2252.
24. Cvitkovitch DG, Gutierrez JA, Bleiweis AS: Role of the citrate pathway in glutamate biosynthesis by *Streptococcus mutans*. *J Bacteriol* 1997, **179**(3):650–655.
25. Kloosterman TG, Hendriksen WT, Bijlsma JJ, Bootsma HJ, van Hijum SA, Kok J, Hermans PW, Kuipers OP: Regulation of glutamine and glutamate metabolism by GlnR and GlnA in *Streptococcus pneumoniae*. *J Biol Chem* 2006, **281**(35):25097–25109.
26. Chen PM, Chen YY, Yu SL, Sher S, Lai CH, Chia JS: Role of GlnR in acid-mediated repression of genes encoding proteins involved in glutamine and glutamate metabolism in *Streptococcus mutans*. *Appl Environ Microbiol* 2010, **76**(8):2478–2486.
27. Doroshchuk NA, Gelfand MS, Rodionov DA: Regulation of nitrogen metabolism in gram-positive bacteria. *Mol Biol (Mosk)* 2006, **40**(5):919–926.
28. Francke C, Kerkhoven R, Wels M, Siezen RJ: A generic approach to identify Transcription Factor-specific operator motifs; Inferences for LacI-family mediated regulation in *Lactobacillus plantarum* WCFS1. *BMC Genomics* 2008, **9**:145.
29. Francke C, Groot Kormelink T, Hagemeijer Y, Overmars L, Sluijter V, Moezelaar R, Siezen RJ: Comparative analyses imply that the enigmatic sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics* 2011, **12**:385.
30. Fisher SH, Rosenkrantz MS, Sonenshein AL: Glutamine synthetase gene of *Bacillus subtilis*. *Gene* 1984, **32**(3):427–438.
31. Belitsky BR, Sonenshein AL: Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. *J Bacteriol* 1998, **180**(23):6298–6305.
32. Commichau FM, Gunka K, Landmann JJ, Stulke J: Glutamate metabolism in *Bacillus subtilis*: gene expression and enzyme activities evolved to avoid futile cycles and to allow rapid responses to perturbations of the system. *J Bacteriol* 2008, **190**(10):3557–3564.
33. Hosaka T, Ui S, Ohtsuki T, Mimura A, Ohkuma M, Kudo T: Characterization of the NADH-linked acetylacetoin reductase/2,3-butanediol dehydrogenase gene from *Bacillus cereus* YUF-4. *J Biosci Bioeng* 2001, **91**(6):539–544.
34. Nicholson WL: The *Bacillus subtilis* *ydjL* (*bdhA*) gene encodes acetoin reductase/2,3-butanediol dehydrogenase. *Appl Environ Microbiol* 2008, **74**(22):6832–6838.
35. Vanoni MA, Curti B: Glutamate synthase: a complex iron-sulfur flavoprotein. *Cell Mol Life Sci* 1999, **55**(4):617–638.
36. Cedar H, Schwartz JH: The asparagine synthetase of *Escherichia coli*. I. Biosynthetic role of the enzyme, purification, and characterization of the reaction products. *J Biol Chem* 1969, **244**(15):4112–4121.
37. Scofield MA, Lewis WS, Schuster SM: Nucleotide sequence of *Escherichia coli* *asnB* and deduced amino acid sequence of asparagine synthetase B. *J Biol Chem* 1990, **265**(22):12895–12902.
38. Fisher SH, Wray LV Jr: *Bacillus subtilis* 168 contains two differentially regulated genes encoding L-asparaginase. *J Bacteriol* 2002, **184**(8):2148–2154.
39. Kim SI, Germond JE, Pridmore D, Soll D: *Lactobacillus bulgaricus* asparagine synthetase and asparaginyl-tRNA synthetase: coregulation by transcription antitermination? *J Bacteriol* 1996, **178**(8):2459–2461.
40. Detsch C, Stulke J: Ammonium utilization in *Bacillus subtilis*: transport and regulatory functions of NrgA and NrgB. *Microbiology* 2003, **149**(Pt 11):3289–3297.
41. Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* 2001, **11**(5):731–753.
42. Saier MH Jr, Yen MR, Noto K, Tamang DG, Elkan C: The Transporter Classification Database: recent advances. *Nucleic Acids Res* 2009, **37**(Database issue):D274–D278.
43. Booger FC, Ma H, Bruggeman FJ, van Heeswijk WC, Garcia-Contreras R, Molenaar D, Krab K, Westerhoff HV: AmtB-mediated NH3 transport in prokaryotes must be active and as a consequence regulation of transport by GlnK is mandatory to limit futile cycling of NH4(+)/NH3. *FEBS Lett* 2011, **585**(1):23–28.
44. Schuurman-Wolters GK, Poolman B: Substrate specificity and ionic regulation of GlnPQ from *Lactococcus lactis*. An ATP-binding cassette transporter with four extracytoplasmic substrate-binding domains. *J Biol Chem* 2005, **280**(25):23785–23790.
45. Krastel K, Senadheera DB, Mair R, Downey JS, Goodman SD, Cvitkovitch DG: Characterization of a glutamate transporter operon, *glnQHMP*, in *Streptococcus mutans* and its role in acid tolerance. *J Bacteriol* 2010, **192**(4):984–993.
46. Wallace B, Yang YJ, Hong JS, Lum D: Cloning and sequencing of a gene encoding a glutamate and aspartate carrier of *Escherichia coli* K-12. *J Bacteriol* 1990, **172**(6):3214–3220.
47. Gaillard I, Slotboom DJ, Knol J, Lolkema JS, Konings WN: Purification and reconstitution of the glutamate carrier GltT of the thermophilic bacterium *Bacillus stearothermophilus*. *Biochemistry* 1996, **35**(19):6150–6156.
48. Groeneveld M, Weme RG, Duurkens RH, Slotboom DJ: Biochemical characterization of the C4-dicarboxylate transporter DctA from *Bacillus subtilis*. *J Bacteriol* 2010, **192**(11):2900–2907.
49. Burguiere P, Auger S, Hullo MF, Danchin A, Martin-Verstraete I: Three different systems participate in L-cystine uptake in *Bacillus subtilis*. *J Bacteriol* 2004, **186**(15):4875–4884.
50. Satomura T, Shimura D, Asai K, Sadaie Y, Hirooka K, Fujita Y: Enhancement of glutamine utilization in *Bacillus subtilis* through the GlnK-GlnL two-component regulatory system. *J Bacteriol* 2005, **187**(14):4813–4821.
51. Gutowski JC, Schreier HJ: Interaction of the *Bacillus subtilis* *glnRA* repressor with operator and promoter sequences in vivo. *J Bacteriol* 1992, **174**(3):671–681.
52. Nakano MM, Yang F, Hardin P, Zuber P: Nitrogen regulation of *nasA* and the *nasB* operon, which encode genes required for nitrate assimilation in *Bacillus subtilis*. *J Bacteriol* 1995, **177**(3):573–579.
53. Wray LV Jr, Zalieckas JM, Fisher SH: Purification and in vitro activities of the *Bacillus subtilis* TnrA transcription factor. *J Mol Biol* 2000, **300**(1):29–40.
54. Wray LV Jr, Ferson AE, Fisher SH: Expression of the *Bacillus subtilis* *ureABC* operon is controlled by multiple regulatory factors including CodY, GlnR, TnrA, and Spo0H. *J Bacteriol* 1997, **179**(17):5494–5501.
55. Schreier HJ, Rostkowski CA, Nomellini JF, Hirschi KD: Identification of DNA sequences involved in regulating *Bacillus subtilis* *glnRA* expression by the nitrogen source. *J Mol Biol* 1991, **220**(2):241–253.
56. de Been M, Bart MJ, Abee T, Siezen RJ, Francke C: The identification of response regulator-specific binding sites reveals new roles of two-component systems in *Bacillus cereus* and closely related low-GC Gram-positives. *Environ Microbiol* 2008, **10**(10):2796–2809.
57. Wray LV Jr, Ferson AE, Rohrer K, Fisher SH: TnrA, a transcription factor

- required for global nitrogen regulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 1996, **93**(17):8841–8845.
58. Robichon D, Arnaud M, Gardan R, Pragai Z, O'Reilly M, Rapoport G, Debarbouille M: **Expression of a new operon from *Bacillus subtilis*, *yzkB-ykoL*, under the control of the TnrA and PhoP-phoR global regulators.** *J Bacteriol* 2000, **182**(5):1226–1231.
 59. Beier L, Nygaard P, Jarmer H, Saxild HH: **Transcription analysis of the *Bacillus subtilis* PucR regulon and identification of a cis-acting sequence required for PucR-regulated expression of genes involved in purine catabolism.** *J Bacteriol* 2002, **184**(12):3232–3241.
 60. Tojo S, Satomura T, Morisaki K, Yoshida K, Hirooka K, Fujita Y: **Negative transcriptional regulation of the *ilv-leu* operon for biosynthesis of branched-chain amino acids through the *Bacillus subtilis* global regulator TnrA.** *J Bacteriol* 2004, **186**(23):7971–7979.
 61. Atkinson MR, Fisher SH: **Identification of genes and gene products whose expression is activated during nitrogen-limited growth in *Bacillus subtilis*.** *J Bacteriol* 1991, **173**(1):23–27.
 62. Larsen R, Kloosterman TG, Kok J, Kuipers OP: **GlnR-mediated regulation of nitrogen metabolism in *Lactococcus lactis*.** *J Bacteriol* 2006, **188**(13):4978–4982.
 63. Wang L, Grau R, Perego M, Hoch JA: **A novel histidine kinase inhibitor regulating development in *Bacillus subtilis*.** *Genes Dev* 1997, **11**(19):2569–2579.
 64. Cruz-Ramos H, Glaser P, Wray LV Jr, Fisher SH: **The *Bacillus subtilis* ureABC operon.** *J Bacteriol* 1997, **179**(10):3371–3373.
 65. Akagawa E, Kurita K, Sugawara T, Nakamura K, Kasahara Y, Ogasawara N, Yamane K: **Determination of a 17,484 bp nucleotide sequence around the 39 degrees region of the *Bacillus subtilis* chromosome and similarity analysis of the products of putative ORFs.** *Microbiology* 1995, **141**(Pt 12):3241–3245.
 66. Richardson IB, Katz ME, Hynes MJ: **Molecular characterization of the *lam* locus and sequences involved in regulation by the AmdR protein of *Aspergillus nidulans*.** *Mol Cell Biol* 1992, **12**(1):337–346.
 67. Hoshino T, Kose K, Uratani Y: **Cloning and nucleotide sequence of the gene *braB* coding for the sodium-coupled branched-chain amino acid carrier in *Pseudomonas aeruginosa* PAO.** *Mol Gen Genet* 1990, **220**(3):461–467.
 68. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetverin V, Church DM, DiCuccio M, Federhen S, *et al*: **Database resources of the National Center for Biotechnology Information.** *Nucleic Acids Res* 2011, **39**(Database issue):D38–D51.
 69. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **25**(24):4876–4882.
 70. Tippmann HF: **Analysis for free: comparing programs for sequence analysis.** *Brief Bioinform* 2004, **5**(1):82–87.
 71. van der Heijden RT, Snel B, van Noort V, Huynen MA: **Orthology prediction at scalable resolution by phylogenetic tree analysis.** *BMC Bioinformatics* 2007, **8**:83.
 72. Huson DH, Richter DC, Rausch C, DeZulian T, Franz M, Rupp R: **Dendroscope: An interactive viewer for large phylogenetic trees.** *BMC Bioinformatics* 2007, **8**:460.
 73. Crooks GE, Hon G, Chandonia JM, Brenner SE: **WebLogo: a sequence logo generator.** *Genome Res* 2004, **14**(6):1188–1190.
 74. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R: **NCBI GEO: mining tens of millions of expression profiles—database and tools update.** *Nucleic Acids Res* 2007, **35**(Database issue):D760–D765.
 75. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**(17):3389–3402.
 76. Teusink B, van Enckevort FH, Francke C, Wiersma A, Wegkamp A, Smid EJ, Siezen RJ: **In silico reconstruction of the metabolic pathways of *Lactobacillus plantarum*: comparing predictions of nutrient requirements with those from growth experiments.** *Appl Environ Microbiol* 2005, **71**(11):7253–7262.
 77. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching.** *Nucleic Acids Res* 2009, **37**(Web Server issue):W202–W208.

doi:10.1186/1471-2164-13-191

Cite this article as: Groot Kormelink *et al.*: Comparative genome analysis of central nitrogen metabolism and its control by GlnR in the class Bacilli. *BMC Genomics* 2012 **13**:191.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

