Gene Expression
Outline

• Introduction

• Gene Expression - how much and where?

• What is regulated and how?

• Four levels of Gene Regulation
E. coli K-12
MG1655 4,639,221 bp

Annotations:
- **CDS +**
- **CDS -**
- **rRNA**
- **tRNA**

Position Preference
- **dev**
  
Percent AT
- **dev**
  
Resolution: 1856

CTE

"Transcriptomics and adaptive genomics of the asymptomatic bacterium *Escherichia coli* strain 83972"

Molecular Genomes and Genetics, **279**:523-534 (2008).
1. What is Regulated?

2. How is it Regulated?
Regulation of transcription

Global, nonspecific regulation

Gene location relative to Ori

1. Not applicable

Histone-like protein/DNA interaction

2. ~100,000 copies

Availability of sigma factor

3. ~1,000 copies

Sigma factor binding to promoter

4.

mRNA stability

5. Variable

Trans-acting transcription regulators

6. ~5-100 copies

Promoters, cis-acting elements

7. ~1-10 copies*

Local, specific regulation

What we can predict from genome sequences

- Predict which sigma factor binds to promoter
- Predict binding sites
- Sigma factor genes
- Polycistronic spacer sequences, RNA structures stabilizing mRNA
- Regulator genes can be predicted
- Conservation, melting properties

What we cannot predict from genome sequences

- Amount of histone-like protein binding can be cell-cycle dependent
- How much sigma factor is present under which conditions cannot be predicted
- Binding sites may be organism-specific
- Regulated target gene can be hard to predict
- Relative strength may be organism specific

* 1-10 copies of a particular cis element can act per RNA, but >1000 cis-elements can be present per genome
1. Global - chromatin
Curved DNA

torroidal supercoiling

topolgical boundry

plectonemic supercoils

RNA polymerase

Figure 8: Logo plots for IHF and FIS binding sites in E. coli. Shown underneath is the consensus site of Finkel and Johnson [20]. Abbreviations for the nucleotides are: N = A, T, G, or C; R = A or G; Y = C or T; W = A or T. Lower case letters represent less well conserved bases. Note that in the consensus site (as well as in the logo plot above), only the C and G are highly conserved.
The diagram represents the chromatin atlas of E. coli K-12 isolate MG1655, with various annotations and data visualizations. The annotations include CDS +, CDS -, rRNA, tRNA, IHF sites, IHF-REP sites, and FIS binding sites. The position preference, GC skew, and percent AT are also visualized. The data is from the Biochimie, 83:201-212, (2001) publication.
2. Sigma factors
Part 1: Regulation of Transcription

depending on the viewpoint) of these networks are signaling proteins that 'feel' extracellular and intracellular conditions, so that the cell is optimally equipped to respond to any changes.

The promoter upstream of a particular gene or operon will determine which Sigma factor binds, and the choice of Sigma factor we call level 4. Sigma factor binding sites can be predicted with varying accuracy. As the binding site of Sigma 54 is strongly conserved, its prediction is relatively easy. A sequence logo of the Sigma 54 promoter consensus of *E. coli* is shown in Fig. 10.6. The consensus binding site for Sigma 70 is also shown, and as can be seen, this is less conserved, even within a species. This is because there are typically over a thousand Sigma 70 promoter sites in a genome, compared to a few hundred or less for other Sigma binding sites.

For other Sigma factors the recognition sequence may be conserved between closely related species, but a 'general' promoter sequence that will work in every bacterial cell unfortunately doesn’t exist. The nature and number of Sigma factors present in a bacterial species can vary considerably, though within genera there is more conservation. For this reason, genes that have to be artificially expressed in a foreign bacterial species (a frequent application in biotechnology) are usually

![Fig. 10.5](image)

**Fig. 10.5** Initiation of transcription in bacteria. In the first step, Sigma factor binds to the DNA on two locations (in the case of Sigma 70 the -35 and -10 sites). RNA polymerase (a complex of two $\alpha$, one $\beta$ and one $\beta'$ subunit) binds next, after which the DNA wraps around the protein. Sigma induces local strand separation so that RNA polymerase starts producing RNA (*in green*). The Sigma factor is then released and RNA polymerase proceeds along the DNA, moving with a local bubble of melted DNA (*indicated by the arrows*)

The diagram illustrates the process of gene transcription with a focus on RNA regions such as 16S rRNA, tRNA, 23S rRNA, and 5S rRNA. It highlights the presence of 3-5 fis elements and UP elements at positions -35 and -10 upstream of the transcription start site.

The fis element consensus and UP element 1 consensus are shown for P1, with respective -35 and -10 positions. Similarly, for P2, the UP element 2 consensus and -35 and -10 positions are depicted.
rRNA promoter P1

fis

E. coli/ Shigella

Salmonella

Yersinia

rRNA promoter P2

Up

-35

-10

Up

-35

-10

DTU course 27105 Comparative Genomics
# ECF Sigma Factors in 250 Bacterial Genomes

![Bar chart showing the number of ECF sigma factors in 250 bacterial genomes across various bacterial phyla.](chart-image)

Evolutionary Tinkering with Conserved Components of a Transcriptional Regulatory Network

Hugo Lavoie¹,²✉, Hervé Hogues¹, Jaideep Mallick¹, Adnane Sellam¹,³, André Nantel¹,³, Malcolm Whiteway¹,²✉

Author Summary

Conserved metabolic machineries direct energy production and investment in most life forms. However, variation in the transcriptional regulation of the genes that encode this machinery has been observed and shown to contribute to phenotypic differences between species. Here, we show that the regulatory circuits governing the expression of central metabolic components (in this case the ribosomes) in different yeast species have an unexpected level of evolutionary plasticity. Most transcription factors involved in the regulation of expression of ribosomal genes have in fact been reused in new ways during the evolutionary time separating S. cerevisiae and C. albicans to generate global changes in transcriptional network structures and new ribosomal regulatory complexes.
S. cerevisiae

Rap1 binds RP gene promoters (GO:0022626; \( p = 3.03 \times 10^{-10} \)), glycolytic enzyme promoters (GO:0006096; \( p = 8.98 \times 10^{-6} \)), the silent mating type locus, and the telomeres as reported (Figure 3A) [49,50], while in C. albicans it binds none of the glycolytic genes and a single (RPS5) RP gene. Instead, C. albicans Rap1 binds to telomeres (Figure S2) and to a few (36) intra-chromosomal locations enriched upstream RNA polII transcriptional regulators (GO:0006357; \( p = 1.34 \times 10^{-6} \)). Altogether, apart from the connections of Cbf1 with sulfur starvation and respiration and of Rap1 with telomeric repeats, all edges in the functional network of these generalist TFs appear to have been reorganized between S. cerevisiae and C. albicans.

**Changes in Hierarchical Layers of the TRN**

In addition to regulating coherent groups of functionally related structural genes, TFs can also act in hierarchical layers by controlling the expression of other TFs as well as key regulatory proteins like kinases or kinase regulators. Changing these hierarchies can have important functional consequences on cellular regulation, and therefore we examined the changes in generalist TF regulatory relationships within the networks obtained from our data. First, we found that TF auto-regulation (by feedback or feed-forward), commonly observed in regulatory network motifs [51–53] and detected here by the binding of a TF to its own promoter, could be gained or lost between species. While Cbf1 and Hmo1 bind their promoter region in both species, the probable auto-regulation of Rap1 and Tbf1 defined by protein binding is seen only in S. cerevisiae (Figure 3B). Second, the hierarchical layers of the TRN have been reorganized between species. The regulatory relationships between TFs appear to be plastic and the hierarchical status of TFs can change dramatically: for example, S. cerevisiae Tbf1 binds 11 TFs (\( p = 2.25 \times 10^{-3} \)) while it binds none in C. albicans (Figure 3B). Similarly, C. albicans Rap1 seems to have moved in the regulatory network hierarchy; six TFs rank in the 10 most Rap1-enriched intergenic regions in C. albicans while its S. cerevisiae homolog binds only 13 TFs amongst its 595 target genes (Figure 3B). Most interestingly, two of C. albicans Rap1-regulated TFs are Sfp1 and Dot6 (Figure S6) [42,43,54,55].

**Figure 3. Functional involvement of generalist TFs of the RP transcriptional regulatory network has drastically changed between S. cerevisiae and C. albicans.** (A) GO categories significantly enriched (\( p < 1 \times 10^{-2} \)) in the target gene sets of each pleiotropic TF were displayed in a graph representing TF-GO interactions. A simplified representation of the regulatory interactions with major functional categories and chromosomal structures is presented between each species TF-GO interaction network. (B) Evolution of TF hierarchies between S. cerevisiae and C. albicans. A map of all significant regulatory relationships between each pleiotropic TF and the TFs present within each species cellular network was drawn. A dashed circle surrounds transcription factors uniquely found in S. cerevisiae.

doi:10.1371/journal.pbio.1000329.g003
The image shows a simple tri-helical structure with labels N and C.
Two-Component Signal Transduction Systems, Environmental Signals, and Virulence

E. Calva and R. Oropeza

Instituto de Biotecnología, UNAM, Cuernavaca, Morelos 62210, Mexico

Received: 29 April 2005 / Accepted: 19 September 2005 / Online publication: 31 January 2006

Abstract

The relevance toward virulence of a variety of two-component signal transduction systems is reviewed for 16 pathogenic bacteria, together with the wide array of environmental signals or conditions that have been implicated in their regulation. A series of issues is raised, concerning the need to understand the environmental cues that determine their regulation in the infected host and in the environment outside the laboratory, which shall contribute toward the bridging of bacterial pathogenesis and microbial ecology.

Introduction

The two-component signal transduction systems (TCS) in bacteria are constituted by a membrane-bound sensor histidine kinase that perceives environmental stimuli and a response regulator that affects gene expression (Fig. 1). Upon sensing an environmental signal, the kinase becomes autophosphorylated and then transfers the phosphate to the response regulator which, in turn, binds to DNA regulatory sequences affecting gene expression. Paradigms of such TCS include the NtrB/NtrC system involved in nitrogen assimilation; the chemotactic system CheA/CheY, although the latter interacts with the flagellar switch and not with the DNA; the porin regulon EnvZ/OmpR; and the system for sporulation control KinA,KinB/Spo0A. A multiple-step phosphorelay pathway can also be found among TCS, as in the virulence BvgS/BvgA system [38].

TCS have been implicated in virulence in a number of bacteria. Moreover, various environmental signals or conditions have been invoked to influence such TCS (Table 1). Major questions in this research area involve
E. Calva and R. Oropeza: Two-Component Systems and Environmental Signals

**Figure 1.** Signal transduction two-component systems in bacteria. (A) The simple prototype phosphotransfer pathway from one histidine kinase sensor (HK) to the response regulator (RR), as for EnvZ/OmpR and PhoQ/PhoP. (B) Multiple-step phosphorelay pathway involving a phosphotransfer scheme of His-Asp-His-Asp, as for BvgS/BvgA. TM: transmembrane domain; Hpt: histidine-phosphotransfer domain.

Microbial Ecology, 51:166-176, 2006
Table 1. Two-component signal transduction systems (TCS) in pathogenic bacteria

<table>
<thead>
<tr>
<th>TCS sensor/regulator</th>
<th>Bacterium</th>
<th>Criteria for virulence</th>
<th>Environmental signals/conditions&lt;sup&gt;a&lt;/sup&gt;</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BvgS/BvgA</td>
<td><em>Bordetella pertussis</em></td>
<td>Attenuated mutants</td>
<td>Nicotinic acid, Mg&lt;sup&gt;2+&lt;/sup&gt;, temperature</td>
<td>[52, 81]</td>
</tr>
<tr>
<td>EtaS/EtaA</td>
<td><em>Enterococcus faecalis</em></td>
<td>Attenuated mutant</td>
<td>50°C, bile salts</td>
<td>[47, 78]</td>
</tr>
<tr>
<td>PehS/PehR</td>
<td><em>Erwinia carotovora</em></td>
<td>Attenuated mutants</td>
<td>[Ca&lt;sup&gt;2+&lt;/sup&gt;]</td>
<td>[21]</td>
</tr>
<tr>
<td>PmrB/PmrA</td>
<td><em>Erwinia carotovora</em></td>
<td>Attenuated mutants</td>
<td>[Fe&lt;sup&gt;3+&lt;/sup&gt;], pH</td>
<td>[40]</td>
</tr>
<tr>
<td>PhoQ/PhoP</td>
<td><em>Erwinia chrysanthemi</em></td>
<td>Induced in planta; acid-sensitive</td>
<td>[Mg&lt;sup&gt;2+&lt;/sup&gt;], in planta</td>
<td>[50]</td>
</tr>
<tr>
<td>ArcB/ArcA</td>
<td><em>Haemophilus influenzae</em></td>
<td>Attenuated mutant</td>
<td>Anaerobiosis</td>
<td>[12]</td>
</tr>
<tr>
<td>Sensor ORFs HP244,</td>
<td><em>Helicobacter pylori</em></td>
<td>Attenuated mutants</td>
<td>?</td>
<td>[64, 71]</td>
</tr>
<tr>
<td>HP165, HP1364,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulator ORF HP1365</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SenX3/RegX3</td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>Attenuated mutants</td>
<td>Anaerobiosis</td>
<td>[65, 70]</td>
</tr>
<tr>
<td>PhoQ/PhoP</td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>Attenuated mutant</td>
<td>[Mg&lt;sup&gt;2+&lt;/sup&gt;]</td>
<td>[27, 66]</td>
</tr>
<tr>
<td>MprB/MprA</td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>Attenuated mutants</td>
<td>?</td>
<td>[86, 87]</td>
</tr>
<tr>
<td>DevS/DevR</td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>Altered virulence of mutants</td>
<td>Hypoxia</td>
<td>[6, 11, 51, 65]</td>
</tr>
<tr>
<td>TcrY/TcrX, TcrS/TcrR, KdpD/KdpE</td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>Enhanced virulence of mutants</td>
<td>Trc: early exponential phase Kdp: K&lt;sup&gt;+&lt;/sup&gt; limitation</td>
<td>[65]</td>
</tr>
<tr>
<td>(ToxS)ToxR</td>
<td><em>Vibrio cholerae</em></td>
<td>Expression of toxin and colonization genes; diminished virulence of mutant strain</td>
<td>Amino acids, carbon dioxide, bile salts, osmolarity</td>
<td>[8, 13, 32, 37, 39, 55, 72, 75, 76]</td>
</tr>
<tr>
<td>LuxO</td>
<td><em>Vibrio cholerae</em></td>
<td>Attenuated mutant strain</td>
<td>?</td>
<td>[53, 80, 88]</td>
</tr>
<tr>
<td>OmpR</td>
<td><em>Yersinia enterocolitica</em></td>
<td>Attenuated mutant strain</td>
<td>Stress signals</td>
<td>[7, 15]</td>
</tr>
<tr>
<td>PhoP</td>
<td><em>Yersinia pestis,</em> Y. pseudotuberculosis</td>
<td>Attenuated mutant strain</td>
<td>[Mg&lt;sup&gt;2+&lt;/sup&gt;] Stress signals</td>
<td>[28, 63]</td>
</tr>
<tr>
<td>YsrR/YsrS</td>
<td><em>Yersinia enterocolitica</em></td>
<td>Control of a type III secretion system</td>
<td>[NaCl]</td>
<td>[83]</td>
</tr>
</tbody>
</table>

<sup>a</sup>Environmental signals that were shown to function under laboratory conditions or were associated to function.
Can we predict these for bacterial genomes?
Table 2. Number of histidine kinases (HisK) and response regulators (RRs) in this month’s genomes and for eight *Pseudomonas* genomes

<table>
<thead>
<tr>
<th>Organism</th>
<th>HisKA</th>
<th>RR</th>
<th>ECF</th>
<th>Sig70</th>
<th>Sig54</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Colwellia psychrerythraea</em> 34H</td>
<td>38</td>
<td>71</td>
<td>12</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Mycoplasma hyopneumoniae</em> J</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><em>Mycoplasma synoviae</em> 53</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>‘Candidatus Pelagibacter ubique’ HTCC1062</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>‘Psychrobacter arcticum’ 273-4</td>
<td>10</td>
<td>16</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><em>Staphylococcus haemolyticus</em> JCSC1435</td>
<td>11</td>
<td>15</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em> PAO1</td>
<td>53</td>
<td>83</td>
<td>19</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas fluorescens</em> Pf-5</td>
<td>68</td>
<td>113</td>
<td>28</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas fluorescens</em> PFO-1</td>
<td>63</td>
<td>101</td>
<td>21</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas fluorescens</em> SBW25</td>
<td>64</td>
<td>102</td>
<td>25</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas putida</em> KT2440</td>
<td>58</td>
<td>89</td>
<td>19</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas syringae</em> DC3000</td>
<td>60</td>
<td>85</td>
<td>10</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas syringae</em> pv. syringae B728a</td>
<td>61</td>
<td>89</td>
<td>10</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas syringae</em> pv. phaseolicola 1448A</td>
<td>58</td>
<td>85</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>
**Microbiology, 151:3447-34452, 2005**

**HisKA**

- Actinobacteria (n=20)
- Aquificae (n=1)
- Bacteroidetes
- Chlorobacteria (n=7)
- Chlamydiae
- Verrucomicrobia (n=9)
- Chloroflexi (n=1)
- Cyanobacteria (n=15)
- Deinococcus
- Thermus (n=3)
- Firmicutes (n=60)
- Fusobacterium (n=1)
- Planctomycetes (n=1)
- Proteobacteria (n=123)
- Spirochetes (n=6)
- Thermotogales (n=1)

**Number of HisKA**

- 0
- 20
- 40
- 60
- 80
- 100

DTU course 27105  Comparative Genomics

17 June, 2014
Fig. 2. Box and whisker plot of the distribution of the two-component system proteins, in terms of groups, ‘orphans’ and operons, in 13 different bacterial phyla. Note that this plot represents the total number of proteins (i.e. both RRs and HKs), so the numbers on the x-axis reflect a combination of the two distributions for RRs and HKs. The colour scheme is the same as for Fig. 1. A group is defined as occurring within 15 kbp of another RR or HK. ‘Orphan’ proteins are defined as those remaining RRs (and a few HKs) that are not found within a group. An operon refers to HKs and RRs occurring within 2 kbp of each other and oriented in the same direction.
4.  small ncRNAs

The role of RNAs in the regulation of virulence-gene expression
Pascale Romby¹, François Vandenbesch² and E Gerhart H Wagner³

Bacterial pathogens sense their environment, and in response, virulence genes are induced or repressed through spatial and temporal regulation. They are also subjected to stress conditions, which require appropriate responses. Recent research has revealed that RNAs are key regulators in pathogens. Small RNAs regulate the translation and/or stability of mRNAs that encode virulence proteins, or proteins with roles in adaptive responses, which are triggered by environmental cues and stresses. In most cases, these small RNAs act directly on target RNAs by an antisense mechanism. Other small RNAs act indirectly, by sequestration of regulatory proteins. Direct sensing of environmental signals can occur through induced structural changes in mRNAs.

Addresses
¹ UPR 9002 CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg Cedex, France
² Faculté de Médecine Laennec, National Reference Center for Staphylococci, INSERM E0230, IFRE, 7 rue Guillaume Paradis, 69372 Lyon Cedex 08, France
³ Institute of Cell and Molecular Biology, Biomedical Center, Uppsala University, Box 596, 751 24 Uppsala, Sweden

Corresponding author: Romby, Pascale (p.romby@ibmc.u-strasbg.fr)

Current Opinion in Microbiology 2006, 9:229–236

This review comes from a themed issue on Cell regulation
Edited by Werner Goebel and Stephen Lory

Available online 10th March 2006
1369-5274/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
DOI 10.1016/j.mib.2006.02.005

pathogenic bacteria? The majority of E. coli sRNAs are present in pathogenic strains; this and the known involvement of sRNAs in stress-response regulation suggests they might be important for virulence. Recent genome searches have identified sRNAs in other pathogens (Table 1) such as Staphylococcus aureus [7], Pseudomonas aeruginosa [8] and Vibrio cholerae [9*,10**]. These sRNAs directly or indirectly regulate virulence genes, or affect adaptive stress-responses, which are important for bacterial survival in a host. This review focuses on recently discovered pathogenesis-related RNAs and their mechanisms of action.

Regulatory RNAs and signaling pathways
The transcription of many pathogenesis-related RNAs is dependent on growth phase. Promoters are tightly regulated, frequently as part of well-understood regulons, responding to specific signals. In several pathogens, the secretion of virulence factors is regulated by cell-density sensing (quorum sensing), a process that involves communication through secreted signaling molecules [11]. Several regulatory RNAs are the main effectors of quorum-sensing systems [9*,10**,12]. In V. cholerae, the sensory signals converge on a response regulatory protein, LuxO (Figure 1). At low cell density, when the autoinducer is absent, phosphorylated LuxO activates the transcription of four redundant Qrr RNAs (quorum regulatory RNAs) that regulate the mRNA of the downstream target gene hapR [10**]. In S. aureus, the effector of quorum sensing is encoded by the agr system, which is composed of two divergent transcription units. The first operon combines a density-sensing cassette (agrD and agrR) and a two-component sensory-regulatory system.
Exploring genomic dark matter: A critical assessment of the performance of homology search methods on noncoding RNA

Eva K. Freyhult,¹ Jonathan P. Bollback,² and Paul P. Gardner³,⁴

¹The Linnaeus Centre for Bioinformatics, Uppsala University, 75124 Uppsala, Sweden; ²Evolution Department, Biological Institute, University of Copenhagen, 2100 Copenhagen, Denmark; ³Molecular Evolution Group, Institute of Molecular Biology and Physiology, University of Copenhagen, 2100 Copenhagen, Denmark

Homology search is one of the most ubiquitous bioinformatic tasks, yet it is unknown how effective the currently available tools are for identifying noncoding RNAs (ncRNAs). In this work, we use reliable ncRNA data sets to assess the effectiveness of methods such as BLAST, FASTA, HMMer, and Infernal. Surprisingly, the most popular homology search methods are often the least accurate. As a result, many studies have used inappropriate tools for their analyses. On the basis of our results, we suggest homology search strategies using the currently available tools and some directions for future development.

[Supplemental material is available online at www.genome.org and http://www.binf.ku.dk/~pgardner/bralibase/bralibase3/1]

Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline

Zasha Weinberg¹,*, Jeffrey E. Barrick²,³, Zizhen Yao⁴, Adam Roth², Jane N. Kim¹, Jeremy Gore¹, Joy Xin Wang¹,², Elaine R. Lee¹, Kirsten F. Block¹, Narasimhan Sudarsan¹, Shane Nep⁵, Martin Tompa⁴,⁵, Walter L. Ruzzo⁴,⁵ and Ronald R. Breaker¹,²,³
sRNAs in regulatory cascades and connections between signals and global regulators. **(a)** The global transcriptional regulators Fur and SgrR sense signals such as intracellular iron concentration and the accumulation of glucose phosphate [2*], respectively, and regulate the transcription of many target genes, in part through controlling transcription of sRNAs. Iron limitation causes inactivation of Fur protein, and consequently derepression of RyhB synthesis. **(b)** Schematic models of *S. aureus* and *V. cholerae* quorum-sensing circuits, adapted from [12] and [9**], respectively. Activation is denoted by arrows and repression by lines. Abbreviation: AIP, autoinducer peptide.
Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea and their metagenomes.
The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium


**Fig. 7.** Conservation of *S. Typhimurium* sRNAs within enteric bacteria. Heat map shows the conservation of *S. Typhimurium* SL1344 sRNAs in 29 genome sequences of bacteria belonging to the family Enterobacteriaceae. Homology was identified with Exonerate software (Materials and Methods). Columns and rows represent sRNAs and bacterial genomes, respectively. In the heat map, red indicates the highest homology as 95–100% identity, and pink shows 85–95% identity. The three blue colors indicate between 85% and 55% identity, and white shows <55% sequence identity. Colored bars at the bottom indicate six phylogenetic groups of *S. Typhimurium* sRNAs: black (conserved in *Typhimurium*), gray (conserved in *Typhimurium*, Paratyphi, Newport, Virchow, Saintpaul and Schwarzengrund), blue (conserved in all *Salmonella enterica* subspecies 1 serovars), yellow (conserved in all *S. enterica* subspecies 1 serovars plus *Salmonella arizonae* and *Salmonella bongori*), orange (conserved in all *Salmonella* and *E. coli* strains), and green (conserved throughout enteric bacteria).
Questions:

● How are most (nearly all) genes regulated?

● Where is the best place along the chromosome to put a gene that you want to be highly expressed?

● What is a helix-turn-helix domain?