

Review

Classifying the Topology of AHL-Driven Quorum Sensing Circuits in Proteobacterial Genomes

Zsolt Gelencsér¹, Kumari Sonal Choudhary², Bruna Goncalves Coutinho^{2,3}, Sanjarbek Hudaiberdiev², Borisz Galbáts^{1,4}, Vittorio Venturi^{2,*} and Sándor Pongor^{2,*}

¹ Faculty of Information Technology, PázmányPéter Catholic University, Práter u. 50/a, Budapest 1083, Hungary; E-Mails: gelzsolt@gmail.com (Z.G.); galbo@digitus.itk.ppke.hu (B.G.)

² International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano 99, Trieste 32149, Italy; E-Mails: sonal.kumari@icgeb.org (K.S.C.); coutinho@icgeb.org (B.G.C.); sanjarbek.hudaiberdiev@icgeb.org (S.H.)

³ The Capes Foundation, Ministry of Education of Brazil, Cx postal 250, Brasilia, DF 70.040-020, Brazil

⁴ Biological Research Center, Temesvári krt 62, Szeged 6726, Hungary

* Authors to whom correspondence should be addressed; E-Mails: venturi@icgeb.org (V.V.); pongor@icgeb.org (S.P.); Tel.: +39-040-3751; Fax: +39-040-226-555.

Received: 1 February 2012; in revised form: 24 March 2012 / Accepted: 26 April 2012 /

Published: 27 April 2012

Abstract: Virulence and adaptability of many Gram-negative bacterial species are associated with an *N*-acylhomoserine lactone (AHL) gene regulation mechanism called quorum sensing (QS). The arrangement of quorum sensing genes is variable throughout bacterial genomes, although there are unifying themes that are common among the various topological arrangements. A bioinformatics survey of 1,403 complete bacterial genomes revealed characteristic gene topologies in 152 genomes that could be classified into 16 topological groups. We developed a concise notation for the patterns and show that the sequences of LuxR regulators and LuxI autoinducer synthase proteins cluster according to the topological patterns. The annotated topologies are deposited online with links to sequences and genome annotations at <http://bacteria.itk.ppke.hu/QStopologies/>.

Keywords: quorum sensing; N-AHL; topology; proteobacteria; bacteria

1. Introduction

Bacteria might be simple, single celled organisms but their social behavior means they can form complex communities and engage in coordinated behaviors. Many bacterial species use chemical signals to monitor their environment and regulate population density. A pioneering study in bacterial chemical signaling was the unraveling of the regulation of luminescence in the marine bacterium *Vibrio fischeri* through a self-produced signal termed autoinducer [1]. *V. fischeri* regulates light production in the Hawaiian bobtail squid (*Euprymna scolopes*); when the population of *V. fischeri* reaches a certain density in the light organ of the squid, the concentration of the signaling molecule, exceeds a threshold extra-cellularly, triggering a cascade of cellular events that finally manifest in the production of light. More than 10 years later the signal was identified as an acyl homoserine lactone (AHL; [2]) and the genetic circuit was then also identified and shown to be composed of a signal generator called the *luxI*/LuxI gene/protein and a response-regulatory protein designated as *luxR*/LuxR gene/protein [3]. The name quorum sensing (QS) was then coined in the 1990s [4] to denote this cell-cell communication mechanism which is now recognized as a key trait governing bacterial community behavior [5–7]. In this mechanism, chemical signals trigger cellular actions such as cell division, the production of virulence factors as well as other traits that aid survival. Cross-talk between bacterial species may also determine survival in complex microbiota. Even inter-kingdom communication routes have been identified, including bacterial-fungal and plant-bacterium interactions [8,9]. The importance of quorum sensing has been clear for a number of years as it appears to be a promising target for controlling bacterial colonization and pathogenesis in human disease [10]. This could be of particular importance, as multidrug resistant strains of human pathogens continue to emerge. Potentially beneficial effects in relation to plant health can be attributed to QS in certain species including regulation of the production of anti-microbial compounds and induction of systemic resistance in plants [11].

One of the most well studied cellular and genetic mechanisms of QS is based on the *N*-acylhomoserine lactones (N-AHLs) [4]. Bacterial cells secrete and respond to autoinducers continuously to sense the surrounding environment, and respond to events as they occur. When an autoinducer reaches a critical level, the population of bacteria responds through a coordinated expression of specific target genes, which finally manifest in a particular behavior/phenotype. The N-AHL QS system is based on two proteins belonging to the LuxI and LuxR families [5,12]. LuxI-family proteins are cytoplasmic enzymes and are responsible for N-AHL synthesis [13]. After synthesis, the signal molecule moves freely across cell membranes and accumulates both intra- and extra-cellularly in proportion to cell density. When populations are low, N-AHL dissipates; when populations are high, N-AHL concentration increases. Above a critical concentration or cell density, N-AHLs interact with the LuxR-family [14] which in most cases result in complexes (homodimers) that bind specific promoter DNA sequences (termed *lux*-boxes) located in the promoter region of target QS-regulated genes ([15]. This subsequently affects their expression resulting in particular phenotypes of the organism. Most often one of the targets of AHL QS is the *luxI*-family gene resulting in a positive feedback loop.

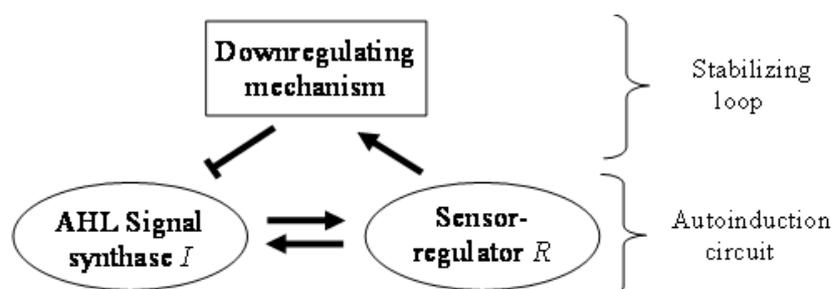
As sequencing technologies have improved and many more bacterial genomes have been sequenced, it has become apparent that homologues of key *V. fischeri* N-AHL QS genes are present in the genomes of many Gram-negative bacteria [16,17]. N-AHLs are one of the most studied QS

signaling molecules discovered so far; as of today over 70 species have been reported to possess N-AHL QS systems (see [17] for an earlier review). Most N-AHL producing bacteria possess canonical QS circuits that broadly resemble the original principles of the design that were described in the nineties. However, a number of interesting variations have since been described in the AHL synthesis/AHL response gene layout and regulation of the QS genes. The goal of this work is to show, based on a new survey of the genomic databases, how these variations correlate with the topological arrangement of the QS regulatory genes.

2. A Generalized Regulatory Framework of N-AHL-Based Quorum Sensing

In general terms, N-AHL-based QS signaling is often referred to as mere autoinduction, requiring only a synthase and a sensor/regulator protein. However, in the absence of regulation of QS, autoinduction would increase signal levels without limit. A down-regulation loop (Figure 1), which turns on at higher signal concentrations is the simplest way to limit and stabilize the signal levels. There are a variety of mechanisms that can play this role in QS systems. For example, the master regulator TraM can form a non-functional heterodimer with the TraR regulator in *Agrobacterium tumefaciens* [18,19]. The DNA-binding negative regulator RsaL acts as a homo-dimer by binding on the bi-directional *rsaL-luxI* promoter [20]. RsaM is another small regulatory protein believed to be acting in a similar way to RsaL in many *Burkholderia* species [21]. The crucial role of these negative regulators is highlighted by the fact that their deletion leads to signal overproduction and a less virulent bacterial phenotype. Bacterial genomes also encode enzymes that degrade AHLs in response to the stress signal ppGpp, which can also efficiently down-regulate QS signaling [22,23]. Down-regulation can also be achieved by RNA-based mechanisms. For example, some QS systems act by activating in their ground state an RNA-binding protein that inhibits transcription of synthases, and then gradually removing the inhibition when signal concentrations are increased [24,25]. Meanwhile, another class of QS regulation system is believed to use sRNA species that considerably decrease the mRNA of the LuxR genes at low population density [26–28]. For the sake of completeness we mention that downregulation can also be achieved by simple resource limitation where a reduction in number of QS cells would decrease of the QS signal concentration itself.

Figure 1. Regulatory outline of N-AHL-based QS signaling. Pointed arrows indicate activation, and the hammerhead arrow indicates inhibition.



Regulatory circuits in which an element can both activate and inhibit another element are termed incoherent feed forward loops (IFFLs) [29,30]. In contrast to simple feed forward arrangements, IFFLs can exhibit a number of complex behavior patterns (for a review see [31]). While simple feed forward

circuits have no inherent limits on their output, IFFL networks have bounded output, which ensures robustness against fluctuations in the input signal levels. The net result is the stabilization of the output signals mentioned above. Another noteworthy property of IFFL networks is adaptation: under appropriate circumstances the circuits will react to the change rather than to the level of the signal [32,33]. Theoretically, this property may also allow bacteria to react to changes rather than the level of the population density.

There are interesting variations also in the regulatory mechanism of the autoinduction loop itself. Although this mechanism is often thought to follow the *luxR/luxI* scenario in which the dimeric LuxR protein forms a regulatory complex with N-AHL, there are a growing number of cases where this happens in a different way. For example, a few members of the LuxR family are able to fold, dimerize, bind DNA, and regulate transcription in the absence of AHLs. Moreover, these proteins are antagonized by their cognate AHLs in the sense that some of the receptors/DNA complexes are disrupted by AHLs *in vitro* [34].

In the canonical AHL QS circuits the signal synthase is a LuxI type enzyme that synthesizes the signal molecule from S-adenosyl-methionine and an intra-cellularly synthesized acyl moiety delivered by an acyl carrier protein (ACP). However, more and more systems have been discovered where the acyl moiety is provided by acyl coenzyme A. Some of these systems use exogenous acyl (e.g., *p*-coumarate) moieties for signal synthesis, which provides a route to respond to environmental stimuli [35–38]. Characteristically, these non-canonical signal synthases tend to attach branched or non-aliphatic acyl side-chains.

3. Distribution and Frequency of QS Genes in Whole Bacterial Genomes

It was noted very early that *luxR* and *luxI* genes of the same QS circuit were most often located vicinal or very near to each other in the genomes. Other instances of *luxR* genes have also been identified in locations that are detached from QS circuits lacking a cognate N-AHL synthase. Such *luxR* genes are termed solos [9] or orphans [39] and are thought to allow bacteria to sense, among others, environmental stimuli and/or AHLs produced by neighboring bacteria. Case *et al.* analyzed the presence and absence of *luxR* and *luxI* genes in 512 genomes [17] in 2007 and found that such genes were exclusively present in proteobacterial genomes. Goryachev meanwhile, provided an excellent overview of the regulatory design principles of QS, distinguishing two chromosomal arrangements—one tandem *luxR/luxI* arrangement called type A, and one convergent arrangement called type B [40,41].

We recently started to survey the topological arrangements of N-AHL-based QS genes in pseudomonads [20]. In the present review we extend upon this work and survey the main categories of topological arrangements of QS genes in currently available bacterial genome sequences. We analyzed 1,403 complete genomes using standard bioinformatics methods (see Supplementary Materials for detailed methodology). This analysis was based on the reading frames annotated by the depositing authors and showed that 169 complete genomes contained QS genes that were close to other QS genes (*i.e.*, according to a distance threshold of 3,000 bp). All these were proteobacterial genomes. As the proteins encoded by *luxR* and *luxI* genes have spurious sequence similarities to a number of other protein families, we manually checked the length, sequence coverage and residue conservation of each similarity. We adopted this cautious approach in order to find a more reliable set of examples. For the

same reason, draft genomes, solo *luxR* genes and other orphaned occurrences of QS genes were excluded from the detailed analysis. Unannotated genes were included only if they made part of a known topological arrangement. Of the 4.8 million genes we analyzed, 674 were *luxR* genes (unannotated: 33), 294 were *luxI* genes (unannotated: 13), 44 were *rsaL* genes (unannotated: 16) and 37 were *rsaM* (all were unannotated). The results of our search have been deposited to an open online repository from where the arrangements and the individual sequences can be retrieved (<http://bacteria.itk.ppke.hu/QStopologies/>).

It is a reasonable question to ask if the found occurrences reflect the frequency of QS genes in nature. We feel that this is not the case. First, we limited our analysis to cases where *luxR* and *luxI* genes are located near each other. Second, our analysis was based on similarity to known LuxR and LuxI proteins. Consequently, we left out lonely occurrences of *luxR* genes (solos), *luxR* genes that may control a different type of signal synthases, or signal synthases under the control of a different regulator protein. There were a few potential similarities outside the proteobacterial genomes, including the cyanobacterium *Gloeotheca* PCC6909 which was suggested to have a QS system [42]. Nevertheless, we decided to limit our analysis to proteobacterial genomes where most of the well-established examples are found. Third, the analysis was limited to complete genomes, and this is a highly biased dataset. With these limitations we found QS genes in about 12% of complete proteobacterial genomes and it would be tempting to argue that this is in agreement with the expected frequency of AHL positive strains in proteobacteria (6–12%) [17]. At the moment we think however that this apparent agreement has to be confirmed with more rigorous sampling methods.

4. Topological Arrangements of QS Genes in Bacterial Genomes

Gene topology is a broad term that can include the arrangement of genes within chromosomes, with respect to the replication origin or other chromosomal elements. In this work we use the words “topological arrangement” or briefly “topology” to denote the arrangement within a close neighborhood of the QS regulatory genes, to denote whether the genes are convergent, divergent, synthase upstream, receptor upstream, or synthase and receptor separated by other genes, *etc.* The preliminary overview of the searches showed a variety of topological types. To illustrate these, we developed a concise notation based on a PROSITE-like syntax [43]. The *luxR*, *luxI*, *rsaL* and *rsaM* genes were abbreviated as *R*, *I*, *L* and *M*, respectively, and *X* is used for all other genes. An arrow above each gene symbol then shows the direction of transcription. With this notation, for example, $\vec{R}\vec{I}$ denotes adjacent *luxR* and *luxI* genes transcribed in the same direction. $\vec{R}X(>5)\vec{I}$ denotes the same pattern with more than five genes between the *luxR* and *luxI*, without specifying the direction of transcription of the *X* gene. A summary of the topological categories found in our survey is given in Table 1. Somewhat arbitrarily, we divided the patterns into two groups: simple topologies and complex topologies. Simple topologies consist of *luxI/luxR* genes that are either vicinal or are separated by few genes. Such arrangements are characterized by typical patterns of transcriptional orientation that are conserved in many proteobacteria. This category makes up the majority of the observed cases in Table 1 and is the main subject of this review. On the other hand, there are complex topologies in which the *luxR/luxI* genes are separated by a larger and more variable number of intervening genes. These topologies are characteristic of *Agrobacterium*

and *Rhizobium* species. As there are many excellent reviews on these species [44,45], we mention this class only for the sake of completeness.

Table 1. Typical topological patterns found in complete bacterial genomes.

ID	Pattern	Gene topology	Occurrence in complete Proteobacterial genomes				
			Total	alpha	beta	gamma	delta
Short, conserved topological patterns							
R1	$\vec{R}\vec{I}$		96	71	14	11	0
R2	$\vec{R}\overleftarrow{I}$		53	2	2	46	3
R3	$\overleftarrow{R}\vec{I}$		11	1	3	7	0
R4	$\overleftarrow{I}\vec{R}$		2	2	0	0	0
L1	$\vec{R}\overleftarrow{L}\vec{I}$		15	0	7	8	0
M1	$\overleftarrow{R}\vec{M}\vec{I}$		30	0	20	10	0
M2	$\vec{R}\vec{M}\vec{I}$		1	0	1	0	0
X1	$\vec{R}\vec{X}\overleftarrow{I}$		1	0	0	1	0
X2	$\overleftarrow{R}\overleftarrow{X}\vec{I}$		2	2	0	0	0
X3	$\vec{R}\overleftarrow{X}\vec{I}$		4	0	2	2	0
X4	$\overleftarrow{R}\vec{X}\vec{I}$		1	1	0	0	0
X5	$\vec{R}\vec{X}\vec{I}$		2	1	1	0	0
Longer, unusual topological patterns							
M3	$\vec{R}X(2-11)\vec{M}\vec{I}$		6	0	6	0	0
M'	$\vec{M}\vec{I}$		2	0	2	0	0
X6	$\vec{R}\overleftarrow{X}(7)\vec{I}$		1	1	0	0	0
X7	$\overleftarrow{I}X(>7)\vec{R}$		5	5	0	0	0

Within the class of simple topologies, the majority of the cases are made up of the $\vec{R}\vec{I}$ and the $\vec{R}\overleftarrow{I}$ topologies that Goryachev termed type A and type B, respectively [40,41]. In Table 1 we see combinations that are outside these two known categories. For instance, all four arrangements that are possible for two vicinal genes appear. In a characteristic subgroup of the simple patterns there is

a single intervening gene between the *luxR* and *luxI* genes. These intervening genes show interesting commonalities in terms of function (Table 2). Out of 48 such single intervening genes 11 code RsaL and 29 code RsaM proteins that are both known to negatively regulate quorum sensing. In both cases of these proteins, typical topologies were observed: $\vec{R}\vec{L}\vec{I}$ and $\vec{R}\vec{M}\vec{I}$, respectively. RsaL (*L* in our notation), is a member of the tetra helical superclass of H-T-H proteins [46] which are recognized as widespread QS repressors in bacteria, binding to DNA as dimers. For example, RsaL (which is predominant in pseudomonad genomes [47]) in *P. aeruginosa* prevents expression of the *R* gene by binding to DNA next to the *lux*-box [21]. In our analysis we also found that homologues of RsaL frequently occur outside QS circuits in various bacterial genomes (data not shown). The significance of this is not clear at this stage. In contrast, we found that RsaM (*M* in our notation), a protein of unknown structure that negatively regulates QS in *P. fuscovaginae* [21], seems to only occur in the context of QS circuits (see also [48]). Even though RsaM seems to occur mainly in the $\vec{R}\vec{M}\vec{I}$ arrangement, it also appears in a number of other topologies (M1, M2 and M3; Table 1). Of the rest of the X genes found in RXI topology, *mupX* of *P. fluorescens* NCIMB10586 is an amidase-hydrolase, which is able to digest/degrade the AHL signal of the same species [49]. MupX can therefore also be considered as a negative QS regulator. From the rest of the X genes, two are involved in DNA-mobilization (an integrase and a transposase), and the rest are hypothetical proteins of unknown function.

Table 2. Intervening genes in short, conserved topological patterns

Gene-types	No. found in complete bacterial genomes	Potential role	Example of genomes
<i>rsaL</i>	11	Negative regulator	<i>P. putida</i> , <i>P. fuscovaginae</i> , <i>P. aeruginosa</i> LESB58, <i>P. aeruginosa</i> PAO1
<i>rsaM</i>	29	Negative regulator	<i>B. pseudomallei</i> K96243
<i>mupX</i>	1	Negative regulator	<i>P. fluorescens</i> NCIMB 10586
Integrases/transposases	2	DNA mobilization	<i>B. vietnamiensis</i> G4 <i>Methylobacterium nodulans</i> ORS 2060
LuxR like regulator	1	?	<i>Gluconacetobacter diazotrophicus</i> PAI 5
Unknown function	5	?	<i>B. mallei</i> NCTC 10247, <i>Saccharophagus degradans</i> 2-40

Conserved gene overlaps are another commonality observed in short, conserved topologies. In our analysis, two topologies were found to contain this feature, namely $\vec{R}\vec{L}\vec{I}$ (L1) and the $\vec{R}\vec{I}$ (R2), which occurred in 15 and 53 proteobacterial genomes respectively. The L1 type QS circuits contain an overlap between transiently transcribed *luxR* and *rsaL* gene. The length of the overlap varies in different species. The L1 type QS circuit of *P. aeruginosa* contains an overlap of 10 bp and the same overlap is 20 bp long in *P. fuscovaginae*. In contrast, *P. putida* has an L1 circuit where the R and L are close (4 bp apart) but not overlapping [48]. The R2 topology contains an overlap of 2 to 79 bp between the convergently transcribed *luxR* and *luxI* genes. Previously it has been suggested that the expression of one member of a convergent or overlapping gene pair might antagonize the expression of the second member resulting in activation or repression of different functions or phenotypes [34]. This may happen by either failed recognition by the second RNA polymerase molecule, or hybridization of the two complementary mRNAs, although this is yet to be confirmed. Even though overlapping genes are

not uncommon in tightly co-regulated gene circuits of bacteria [50], e.g., restriction modification systems [36], the overlaps of convergently transcribed genes are less common in the literature. However, a few examples of convergent antisense RNA regulation have been reported also in bacteria, including S-box transcripts that are convergent to the *ubiG-mccBA* operon in *C. acetobutylicum* [51].

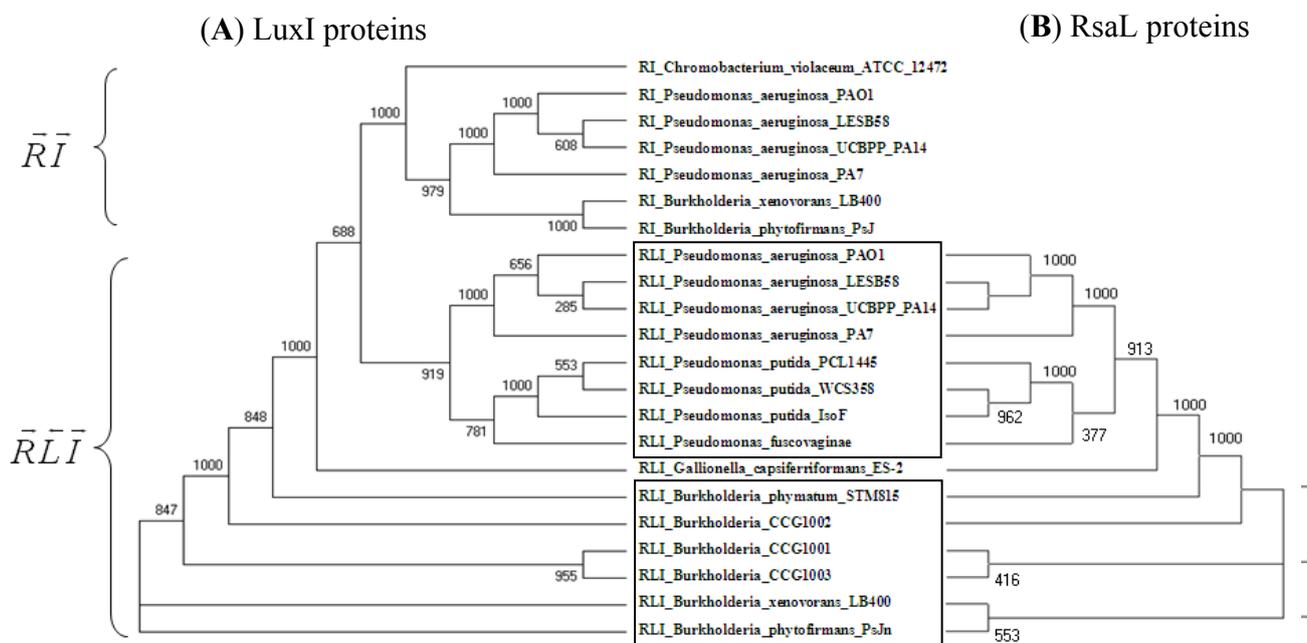
In contrast to simple topologies, the group of longer, complex topological patterns shows greater variety although the number of occurrences we found was considerably less. The groups of *Agrobacterium* and various *Rhizobia* often contain many genes between the *luxR* and *luxI* type genes. An interesting example in this group is the MI pattern of *B. ambifaria*, where two well distinguishable QS genes (*rsaM* and *luxI*) appear in their usual tandem topology, but without an annotated or computationally identifiable *luxR* homolog in their vicinity.

5. Taxonomic Distribution of QS Gene Topology Patterns

The number of complete genomes does not allow us yet to draw definitive conclusions about the preference of certain patterns to appear in various classes or species of bacteria. It is apparent however that the $\vec{R}\vec{I}$ topology is predominant in α -proteobacteria, while the $\vec{R}\vec{I}$ pattern is more frequent in γ -proteobacteria. Furthermore, the $\vec{R}\vec{L}\vec{I}$ and the $\vec{R}\vec{M}\vec{I}$ topologies seem to occur in the β and γ classes but not in α -proteobacteria. The question arises whether the known QS proteins, such as LuxI and LuxR, cluster simply according to the known taxonomy or rather according to the topological pattern. Cladograms of 154 LuxI and LuxR protein sequences (Supplementary Figures 3 and 4) show a clear tendency for clustering according to topological patterns. For instance, the $\vec{R}\vec{L}\vec{I}$, $\vec{R}\vec{M}\vec{I}$, $\vec{R}\vec{I}$ and $\vec{R}\vec{I}$ patterns form clearly distinguishable groups both in the LuxR and in the LuxI trees. Also, if two LuxI (or LuxR) proteins from the same topology group are nearest neighbors within the tree, they are very rarely from the same genome. In more detail, of the LuxI-type proteins that occur in genomes with multiple QS circuits, 38% have a nearest sequence similarity neighbor from the same family, 35% from the same genus, 26% from the same species, and approximately 1% from the same genome. The frequency with which *luxR* genes occur is similar: 37%, 36%, 27%, and 1%, respectively. This indicates that the patterns may have formed before the modern strains diverged from each other.

An example in Figure 2 shows that LuxI proteins present in $\vec{R}\vec{L}\vec{I}$ patterns clearly separate from those present in $\vec{R}\vec{I}$ patterns within the same genome and cluster together with the respective genes of another (β or γ) class (Figure 2(A)). At the same time, the clustering of the RsaL proteins (Figure 2(B)) is identical to the clustering of their accompanying LuxI genes (Figure 2(A)). In other words, QS proteins seem to cluster according to gene topology at various taxonomic levels, which suggests that the *luxR*, *luxI* as well as the intervening genes may have evolved together. Gene neighborhoods are known to evolve via complex rearrangements, with different combinations of genes from a neighborhood fixed in different lineages [53,54]. However, when the genes are part of fine-tuned regulatory networks, such as QS circuits, strict constraints may be imposed on the process of rearrangement.

Figure 2. Clustering of LuxI and their negative regulators RsaL shows identical tendencies. (A) shows tree of LuxI proteins present RI and RLI pattern, and (B) shows RsaL proteins in RLI pattern. The sequences were taken from genomes harboring *rsaL* genes located between *luxR* and *luxI* genes. The clustering was carried out by the Phylip program package [52]. The numerical value at each node indicates the bootstrap value supporting every split in the lineage (out of 1,000 bootstrap replicates).



6. Conclusions and Open Problems

We can conclude that the topological patterns in QS genes seem to follow a few basic rules—such as the conservation of genes between *luxR* and *luxI* genes, repeating patterns in certain taxa, *etc.*—but their variety is apparently greater than previously indicated [35,36]. The clustering of QS genes suggests that the topological units might act as regulatory modules that evolve together. In order to further test these hypotheses, we plan to widen the scope of the analyzed databases so as to include in the survey draft genomes and individually sequenced genomic regions. We also plan to pinpoint conserved DNA motifs within the neighborhood of the QS circuits. Maintaining and manually curating such a collection is beyond the possibilities of a single research group. Rather, we consider using bioinformatic tools for updating the current collection in the form of an openly accessible Internet repository so that researchers of the field could add their own annotations in the future.

This work is a preliminary, computational census of topological patterns found in complete proteobacterial genomes. It is expected that other kind of patterns may be found in published papers and in other sources, especially draft genomes, and we plan to develop computational protocols for finding these occurrences. On the other hand, there are a number of problems that are not touched upon in our work but may be answered as more detailed data become available. First, are the topological patterns associated with distinct biological functions? As QS circuits are involved in activating a large variety of genes in various bacteria, we speculate that there may be no simple correspondence between

topology and the regulated functions. While we agree that these questions should be tackled by experiment, some supporting evidence can be gathered by comparing the gene neighborhood of QS circuits. The intervening genes (which are a conspicuous subgroup of neighborhood genes) appear to belong to only a few types, and in some genomes they are also found outside the immediate vicinity of QS genes. Another avenue would be to categorize the QS patterns according to the chemical nature of the signals produced by the *luxI* and/or sensed by the *luxR* gene.

The present classification suggests that the patterns seem to be relatively conserved and their distribution among the various taxonomic groups is not random. This does not mean that one can make broad statements about the functional reasons of this apparent conservation. We tend to speculate that the known principles of gene expression (co-transcription, repression by proteins or RNA) can be combined into a finite set of topological patterns that allow stable positive autoregulation and control of linked genes. It would be tempting to speculate that such modules then can move between genomes but this problem, as well as further evolutionary questions are outside the immediate scope of the present paper.

Acknowledgements

KSC, BGC and SH are graduate students at ICGEB, Trieste. ZG and BG are graduate and undergraduate students, respectively, at the Faculty of Information Technology, Pázmány Péter Catholic University, Budapest (PPKE-ITK). BGC Ph.D. programme was funded by a grant from CAPES, Brazil. Work at the Szeged Biological Research Center was partly funded by grant K. 84335 of the Hungarian Scientific Research Fund, OTKA. Research at PPKE-ITK was partly funded by grant TÁMOP-4.2.1.B-11/2/KMR-2011-0002 of the Hungarian National Development Agency NFU.

References

1. Nealson, K.H.; Platt, T.; Hastings, J.W. Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol.* **1970**, *104*, 313–322.
2. Eberhard, A.; Burlingame, A.L.; Eberhard, C.; Kenyon, G.L.; Nealson, K.H.; Oppenheimer, N.J. Structural identification of autoinducer of *Photobacterium fischeri* luciferase. *Biochemistry* **1981**, *20*, 2444–2449.
3. Engebrecht, J.; Nealson, K.; Silverman, M. Bacterial bioluminescence: Isolation and genetic analysis of functions from *Vibrio fischeri*. *Cell* **1983**, *32*, 773–781.
4. Fuqua, W.C.; Winans, S.C. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J. Bacteriol.* **1994**, *176*, 2796–2806.
5. Fuqua, C.; Parsek, M.R.; Greenberg, E.P. Regulation of gene expression by cell-to-cell communication: Acyl-homoserine lactone quorum sensing. *Annu. Rev. Genet.* **2001**, *35*, 439–468.
6. Fuqua, W.C.; Winans, S.C.; Greenberg, E.P. Quorum sensing in bacteria: The LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* **1994**, *176*, 269–275.
7. Waters, C.M.; Bassler, B.L. Quorum sensing: Cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 319–346.

8. Pacheco, A.R.; Sperandio, V. Inter-kingdom signaling: Chemical language between bacteria and host. *Curr. Opin. Microbiol.* **2009**, *12*, 192–198.
9. Subramoni, S.; Venturi, V. LuxR-family ‘solos’: Bachelor sensors/regulators of signalling molecules. *Microbiology* **2009**, *155* (Pt 5), 1377–1385.
10. Bjarnsholt, T.; Givskov, M. Quorum sensing inhibitory drugs as next generation antimicrobials: Worth the effort? *Curr. Infect. Dis. Rep.* **2008**, *10*, 22–28.
11. Venturi, V.; Subramoni, S. Future research trends in the major chemical language of bacteria. *HFSP J.* **2009**, *3*, 105–116.
12. Whitehead, N.A.; Barnard, A.M.; Slater, H.; Simpson, N.J.; Salmond, G.P. Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol. Rev.* **2001**, *25*, 365–404.
13. More, M.I.; Finger, L.D.; Stryker, J.L.; Fuqua, C.; Eberhard, A.; Winans, S.C. Enzymatic synthesis of a quorum-sensing autoinducer through use of defined substrates. *Science* **1996**, *272*, 1655–1658.
14. Hanzelka, B.L.; Greenberg, E.P. Evidence that the N-terminal region of the *Vibrio fischeri* LuxR protein constitutes an autoinducer-binding domain. *J. Bacteriol.* **1995**, *177*, 815–817.
15. Welch, M.; Todd, D.E.; Whitehead, N.A.; McGowan, S.J.; Bycroft, B.W.; Salmond, G.P. N-acyl homoserine lactone binding to the CarR receptor determines quorum-sensing specificity in *Erwinia*. *EMBO J.* **2000**, *19*, 631–641.
16. Fuqua, C.; Winans, S.C.; Greenberg, E.P. Census and consensus in bacterial ecosystems: The LuxR-LuxI family of quorum-sensing transcriptional regulators. *Annu. Rev. Microbiol.* **1996**, *50*, 727–751.
17. Case, R.J.; Labbate, M.; Kjelleberg, S. AHL-driven quorum-sensing circuits: Their frequency and function among the Proteobacteria. *ISME J.* **2008**, *2*, 345–349.
18. Chen, G.; Malenkos, J.W.; Cha, M.R.; Fuqua, C.; Chen, L. Quorum-sensing antiactivator TraM forms a dimer that dissociates to inhibit TraR. *Mol. Microbiol.* **2004**, *52*, 1641–1651.
19. Qin, Y.; Smyth, A.J.; Su, S.; Farrand, S.K. Dimerization properties of TraM, the antiactivator that modulates TraR-mediated quorum-dependent expression of the Ti plasmid *tra* genes. *Mol. Microbiol.* **2004**, *53*, 1471–1485.
20. Venturi, V.; Rampioni, G.; Pongor, S.; Leoni, L. The virtue of temperance: Built-in negative regulators of quorum sensing in *Pseudomonas*. *Mol. Microbiol.* **2011**, *82*, 1060–1070.
21. Mattiuzzo, M.; Bertani, I.; Ferluga, S.; Cabrio, L.; Bigirimana, J.; Guarnaccia, C.; Pongor, S.; Maraitte, H.; Venturi, V. The plant pathogen *Pseudomonas fuscovaginae* contains two conserved quorum sensing systems involved in virulence and negatively regulated by RsaL and the novel regulator RsaM. *Environ. Microbiol.* **2011**, *13*, 145–162.
22. Kobayashi, H.; Kaern, M.; Araki, M.; Chung, K.; Gardner, T.S.; Cantor, C.R.; Collins, J.J. Programmable cells: Interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8414–8419.
23. Wang, C.; Zhang, H.B.; Wang, L.H.; Zhang, L.H. Succinic semialdehyde couples stress response to quorum-sensing signal decay in *Agrobacterium tumefaciens*. *Mol. Microbiol.* **2006**, *62*, 45–56.
24. Cui, Y.; Chatterjee, A.; Hasegawa, H.; Chatterjee, A.K. *Erwinia carotovora* subspecies produce duplicate variants of ExpR, LuxR homologs that activate *rsmA* transcription but differ in their interactions with N-acylhomoserine lactone signals. *J. Bacteriol.* **2006**, *188*, 4715–4726.

25. Sjoblom, S.; Brader, G.; Koch, G.; Palva, E.T. Cooperation of two distinct ExpR regulators controls quorum sensing specificity and virulence in the plant pathogen *Erwinia carotovora*. *Mol. Microbiol.* **2006**, *60*, 1474–1489.
26. Henke, J.M.; Bassler, B.L. Three parallel quorum-sensing systems regulate gene expression in *Vibrio harveyi*. *J. Bacteriol.* **2004**, *186*, 6902–6914.
27. Lenz, D.H.; Miller, M.B.; Zhu, J.; Kulkarni, R.V.; Bassler, B.L. CsrA and three redundant small RNAs regulate quorum sensing in *Vibrio cholerae*. *Mol. Microbiol.* **2005**, *58*, 1186–1202.
28. Lenz, D.H.; Mok, K.C.; Lilley, B.N.; Kulkarni, R.V.; Wingreen, N.S.; Bassler, B.L. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **2004**, *118*, 69–82.
29. Alon, U. Network motifs: Theory and experimental approaches. *Nat. Rev.* **2007**, *8*, 450–461.
30. Milo, R.; Shen-Orr, S.; Itzkovitz, S.; Kashtan, N.; Chklovskii, D.; Alon, U. Network motifs: Simple building blocks of complex networks. *Science* **2002**, *298*, 824–827.
31. Kim, D.; Kwon, Y.K.; Cho, K.H. The biphasic behavior of incoherent feed-forward loops in biomolecular regulatory networks. *Bioessays* **2008**, *30*, 1204–1211.
32. Novak, B.; Tyson, J.J. Design principles of biochemical oscillators. *Nat. Rev.* **2008**, *9*, 981–991.
33. Tyson, J.J.; Novak, B. Functional motifs in biochemical reaction networks. *Annu. Rev. Phys. Chem.* **2010**, *61*, 219–240.
34. Tsai, C.S.; Winans, S.C. LuxR-type quorum-sensing regulators that are detached from common scents. *Mol. Microbiol.* **2010**, *77*, 1072–1082.
35. Ahlgren, N.A.; Harwood, C.S.; Schaefer, A.L.; Giraud, E.; Greenberg, E.P. Aryl-homoserine lactone quorum sensing in stem-nodulating photosynthetic bradyrhizobia. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7183–7188.
36. Hirakawa, H.; Oda, Y.; Phattarasukol, S.; Armour, C.D.; Castle, J.C.; Raymond, C.K.; Lappala, C.R.; Schaefer, A.L.; Harwood, C.S.; Greenberg, E.P. Activity of the *Rhodopseudomonas palustris* p-coumaroyl-homoserine lactone-responsive transcription factor RpaR. *J. Bacteriol.* **2011**, *193*, 2598–2607.
37. Lindemann, A.; Pessi, G.; Schaefer, A.L.; Mattmann, M.E.; Christensen, Q.H.; Kessler, A.; Hennecke, H.; Blackwell, H.E.; Greenberg, E.P.; Harwood, C.S. Isovaleryl-homoserine lactone, an unusual branched-chain quorum-sensing signal from the soybean symbiont *Bradyrhizobium japonicum*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16765–16770.
38. Schaefer, A.L.; Greenberg, E.P.; Oliver, C.M.; Oda, Y.; Huang, J.J.; Bittan-Banin, G.; Peres, C.M.; Schmidt, S.; Juhaszova, K.; Sufrin, J.R.; *et al.* A new class of homoserine lactone quorum-sensing signals. *Nature* **2008**, *454*, 595–599.
39. Lequette, Y.; Lee, J.H.; Ledgham, F.; Lazdunski, A.; Greenberg, E.P. A distinct QscR regulon in the *Pseudomonas aeruginosa* quorum-sensing circuit. *J. Bacteriol.* **2006**, *188*, 3365–3370.
40. Goryachev, A.B. Understanding bacterial cell-cell communication with computational modeling. *Chem. Rev.* **2011**, *111*, 238–250.
41. Goryachev, A.B. Design principles of the bacterial quorum sensing gene networks. *Wiley Interdiscip. Rev.* **2009**, *1*, 45–60.

42. Sharif, D.I.; Gallon, J.; Smith, C.J.; Dudley, E. Quorum sensing in Cyanobacteria: N-octanoyl-homoserine lactone release and response, by the epilithic colonial cyanobacterium *Gloeotheca* PCC6909. *ISME J.* **2008**, *2*, 1171–1182.
43. Falquet, L.; Pagni, M.; Bucher, P.; Hulo, N.; Sigrist, C.J.; Hofmann, K.; Bairoch, A. The PROSITE database, its status in 2002. *Nucleic Acids Res.* **2002**, *30*, 235–238.
44. Sanchez-Contreras, M.; Bauer, W.D.; Gao, M.; Robinson, J.B.; Allan Downie, J. Quorum-sensing regulation in rhizobia and its role in symbiotic interactions with legumes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2007**, *362*, 1149–1163.
45. White, C.E.; Winans, S.C. Cell-cell communication in the plant pathogen *Agrobacterium tumefaciens*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2007**, *362*, 1135–1148.
46. Rampioni, G.; Polticelli, F.; Bertani, I.; Righetti, K.; Venturi, V.; Zennaro, E.; Leoni, L. The *Pseudomonas* quorum-sensing regulator RsaL belongs to the tetrahelical superclass of H-T-H proteins. *J. Bacteriol.* **2007**, *189*, 1922–1930.
47. de Kievit, T.; Seed, P.C.; Nezezon, J.; Passador, L.; Iglewski, B.H. RsaL, a novel repressor of virulence gene expression in *Pseudomonas aeruginosa*. *J. Bacteriol.* **1999**, *181*, 2175–2184.
48. Gelencsér, Z.; Galbáts, B.; Gonzalez, J.F.; Choudhary, K.S.; SHudaiberdiev, S.; Venturi, V.; Pongor, S. Chromosomal arrangement of AHL-driven quorumsensing circuits in *Pseudomonas*. *ISRN Microbiol.* **2012**, 484176.
49. El-Sayed, A.K.; Hothersall, J.; Thomas, C.M. Quorum-sensing-dependent regulation of biosynthesis of the polyketide antibiotic mupirocin in *Pseudomonas fluorescens* NCIMB 10586. *Microbiology* **2001**, *147* (Pt 8), 2127–2139.
50. Krakauer, D.C. Stability and evolution of overlapping genes. *Evolution* **2000**, *54*, 731–739.
51. Thomason, M.K.; Storz, G. Bacterial antisense RNAs: How many are there, and what are they doing? *Annu. Rev. Genet.* **2010**, *44*, 167–188.
52. Ropelewski, A.J.; Nicholas, H.B.; Gonzalez Mendez, R.R. MPI-PHYLIP: Parallelizing computationally intensive phylogenetic analysis routines for the analysis of large protein families. *PLoS One* **2010**, *5*, e13999.
53. Coenje, T.; Vandamme, P. Organisation of the S10, spc and alpha ribosomal protein gene clusters in prokaryotic genomes. *FEMS Microbiol. Lett.* **2005**, *1*, 117–126.
54. Rogozin, I.B.; Makarova, K.S.; Murvai, J.; Czabarka, E.; Wolf, Y.I.; Tatusov, R.L.; Szekely, L.A.; Koonin, E.V. Connected gene neighborhoods in prokaryotic genomes. *Nucleic Acids Res.* **2002**, *30*, 2212–2223.