Neuroscience Area – PhD course in

Molecular Biology

Cell-cell interactions and LuxR solos in plant associated bacteria

Candidate:

Cristina Bez

Supervisor:

Vittorio Venturi

Academic Year 2020-21

Abstract

In the last decade, plant microbiome studies have evidenced that bacteria live as part of complex multispecies communities. Plant health heavily depends on its microbiome and cell-cell signaling among beneficial bacteria as well as pathogens and harmless bacteria are likely to be very important for the establishment and maintenance of microbial communities. This research is now a major challenge in microbiology as cell-cell signaling has thus far been mainly studied in the laboratory in pure cultures. In this thesis, three experimental chapters are presented that are focused on the mechanisms of interspecies signaling in plant associated bacteria and how these contribute in creating a stable multispecies community. The first chapter uses the rice foot rot disease caused by *Dickeya zeae* as a model to decipher the possible interactions between the pathogen and the commensal members of the microbiome. 16S rRNA gene amplicon-based community profiling showed that the pathogen significantly alters the resident bacterial community and its presence is positively correlated with several bacterial species which are likely to team-up with the pathogen and be involved in the disease process. The second chapter focuses on the role of a sub-family of quorum sensing regulators called LuxR solos in bacterial cell-cell interactions. The distribution, frequency and functional role of the LuxR solos regulators is investigated in the ubiquitous plant associated group of fluorescent *Pseudomonas* spp.; nine different sub-groups have been identified and the majority of them are likely to respond to novel exogenous (or possibly endogenous) signals, suggesting that these regulators could play a role in inter-species/inter-kingdom signaling. The last experimental chapter investigates a novel cell-cell communication system, in which a LuxR solo responds to and regulates the biosynthesis of a pigment molecule. In summary, this thesis highlights cell-cell signaling in the microbiome and

the emerging role played by the LuxR solo regulators in providing different ways of bacterial signaling.

Table of contents

List of Figures

List of Tables

Chapter I

General Introduction

1.1 Microbiome – general overview

The total amount of microbes (archaea, bacteria, fungi, viruses, protists and algae), living in a specific environmental niche, is called microbiota (Turnbaugh et al., 2007; Qin et al., 2010). On the other hand, the term microbiome is the combination of the microbiota and their "theatre of activity" such as the microbial structural elements, metabolites/signal molecules, and the surrounding environmental conditions (Berg et al., 2020). It is estimated that the microbial cells which colonize the human body are at least as abundant as our somatic cells and contain far more genes than the human genome (Hughes and Sperandio, 2008); at least 500-1,000 species of bacteria exist in the human body at any one time making up a substantial portion of our biomass. Different people harbor radically different collections of microbes with abundances that vary substantially even among conserved taxa; this degree of personalization is so high that the microbiome is considered a personal signature, having important implications. Additionally, many localized differences in the microbiota of each person depend on where in the body the microbiota is collected from and when over time the microbiota is analyzed (Bäckhed et al., 2012; Hacquard et al., 2015) (**Figure 1.1**).

To date, very little is known on the factors and dynamics that lead and regulates this variation/changes among individuals and whether these variations influence/correlate with people wellness, the progression of diseases or preservation of health status (Bäckhed et al., 2012).Recently many studies evidenced that the microbiome can provide novel biomarkers for several human diseases, hinting opportunities for new preventive and therapeutic modalities (Jiang et al., 2015; Zheng et al., 2016).

Similarly, plants support numerous and diverse microbial communities (archaea, bacteria, fungi, viruses and protists) that are intimately connected to their health and function (Lindow

and Brandl, 2003; Berendsen et al., 2012; Vorholt, 2012; Bulgarelli et al., 2013; Philippot et al., 2013; Lakshmanan et al., 2014).The large collective genome of microbes associated with plants is often termed the “plant-second genome” (Berendsen et al., 2012) due to the importance of microbes in plant growth, stress tolerance, health and productivity (Berendsen et al., 2012; Müller et al., 2016).

The development of meta-omics and computational tools has radically changed the field of microbial community studies (Venter et al., 2004; Gilbert and Dupont, 2011; Vandenkoornhuyse et al., 2015). Especially cultivation-independent methods based on profiling of marker genes (16S, ITS and 23S rRNA genes) or shotgun metagenome sequencing allow considerable progress in the understanding of microbial ecology in the human and plant environment as well as the possible protective or harmful functions of specific microbes for their hosts (Gilbert and Dupont, 2011; Soni et al., 2017).

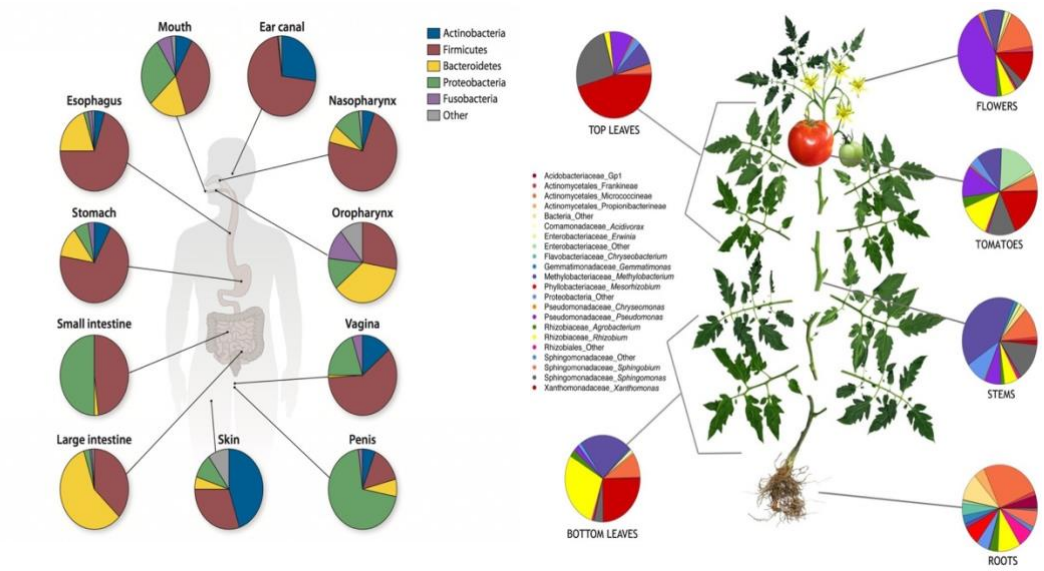


Figure 1. 1 The human microbiome and the plant microbiome are highly personalized and compartimentalized.

However, very little information is currently available on the mechanisms that determine the formation of the microbiome as well as the type of interactions among bacteria. There is therefore a major knowledge gap on the bacteria-bacteria interactions, such as the inter-species communication, which are important for the establishment, maintenance and stability of the microbiome. This aspect is now a major challenge in microbiology as microbial signaling has thus far been mainly studied in the laboratory using simple settings. Deciphering the molecular basis of these interbacterial relationships in the microbiome will therefore have impact on both basic sciences, microbiome-based medicine and translational agriculture.

1.2 Plant microbiome - diversity and dynamics

As mentioned above, plants, like humans and animals, are colonized by a large number of microorganisms resulting in a superorganism that partially relies on this complex community for specific functions and traits (Müller et al., 2016). The plant microbiome confers an additional reservoir of genes that the plant can have access to when needed. It confers fitness advantages to the plant host, including biotic and abiotic stress tolerance, nutrient acquisition, growth promotion, systemic resistance induction and resistance to pathogens (Turner et al., 2013; Edwards et al., 2015; Müller et al., 2016; Hartman et al., 2017)(**Figure 1.2**). Thus, plant health strongly depends on its microbiome and specific assembly rules and molecular mechanisms, both microbe-microbe or microbe-host based are crucial for the establishment of this stable and highly biodiverse microbial community (Reinhold-Hurek and Hurek, 2011; Bulgarelli et al., 2013).

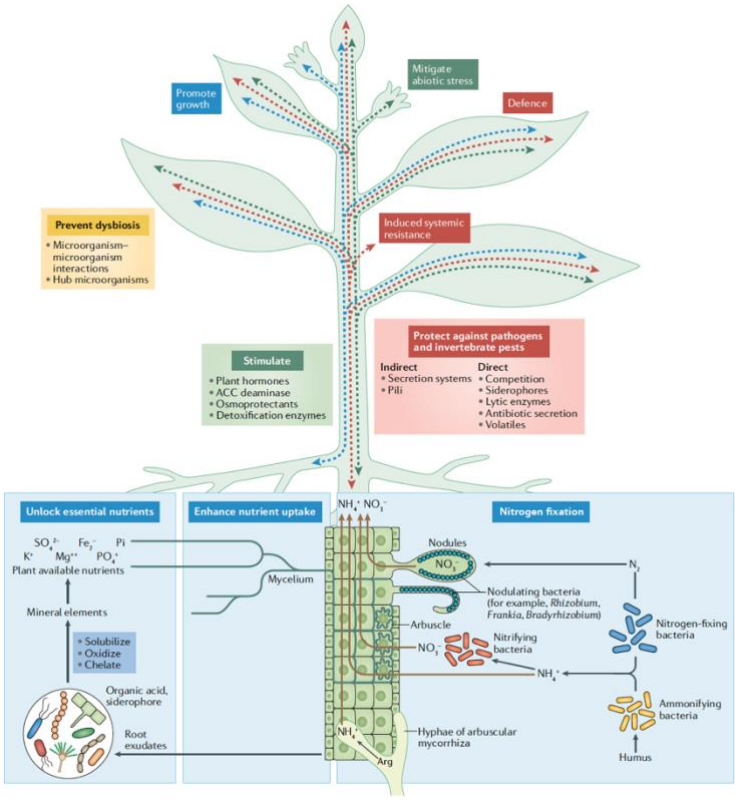


Figure 1. 2 Beneficial effects of the plant – associated microbiome (Trivedi et al., 2020)

Several factors influence microbiome composition such as the soil type (Bulgarelli et al., 2012; Lundberg et al., 2012), plant compartment (Bai et al., 2015; Edwards et al., 2015; Leff et al., 2015; Zarraonaindia et al., 2015), host genotype (Redford et al., 2010; Peiffer et al., 2013; Ofek-Lalzar et al., 2014) plant immune system (Horton et al., 2014), plant development stage and time/season (Redford et al., 2010; Rastogi et al., 2012; Bodenhausen et al., 2014; Chaparro et al., 2014; Maignien et al., 2014). Despite the large number of bacterial phyla existing in nature and the large number of factors influencing the bacterial community formation, plant microbiomes are taxonomically structured and dominated by bacteria which mainly belong to four bacterial phyla: the Proteobacteria, Bacteroidetes, Actinobacteria and Firmicutes (Bulgarelli et al., 2012; Edwards et al., 2015).This phylogenetic conservation infers an organized assembly of microbiomes which is directed by microbe-microbe interaction

strategies that allow them to tightly associate and persist within the plant (Hassani et al., 2018).

Additionally, the interaction between plant and bacteria is also fundamental for the formation of this microbial community. However, unlike mammals, plants are sessile, thus they have to release an array of chemical signals to interact with other organisms (Gopal and Gupta, 2016). Through photosynthesis, plants serve as a rich source of carbon for diverse bacterial communities, feeding them and at the same time using these chemical molecules to mediate several interactions; the latter include mutualistic associations with beneficial microbes, such as rhizobia, mycorrhizae, endophytes or plant-growth-promoting rhizobacteria (PGPR) and parasitic interactions with phytopathogenic microbes and invertebrate herbivores (Frey-Klett et al., 2007; Raaijmakers et al., 2009; Berendsen et al., 2012; Vorholt, 2012). The plant-microbe interaction is called interkingdom signaling and it plays an important role for the microbial community recruitment as well as influencing plant genes expression and important living functions. Plants are able to shape their microbiome, since, upon pathogen attack, they can recruit protective microorganisms to suppress pathogenic bacteria by inducing systemic resistant (ISR) (Lugtenberg and Kamilova, 2009; Doornbos et al., 2012). However, the molecular mechanisms that govern plant-microbe interactions as well as the genes that contribute to bacterial adaptation to plants and root colonization, are not yet very well understood.

1.2.1 Different plant compartments harbor selected set of microbes

The plant microbiota inhabits different plant compartments which are colonized by diverse and specific communities of microorganisms that live either inside the tissues (endophytes)

or on the surface (epiphytes) of the plant (Compant et al., 2005; Compant et al., 2010; Gopal and Gupta, 2016; Müller et al., 2016).

The rhizosphere, the endosphere, the phyllosphere and the seeds constitute the major compartments in which the microbial communities reside. Community composition varies significantly between the rhizosphere, endosphere and phyllosphere, indicating that the plant compartment is the major selective force that shape the composition of the plant-associated microbiome. This could be explained in part by plant-microbiome co-evolution; however, niche adaptation may also have a role in the selective filtering of different microorganisms (Müller et al., 2016) (**Figure 1.3**).

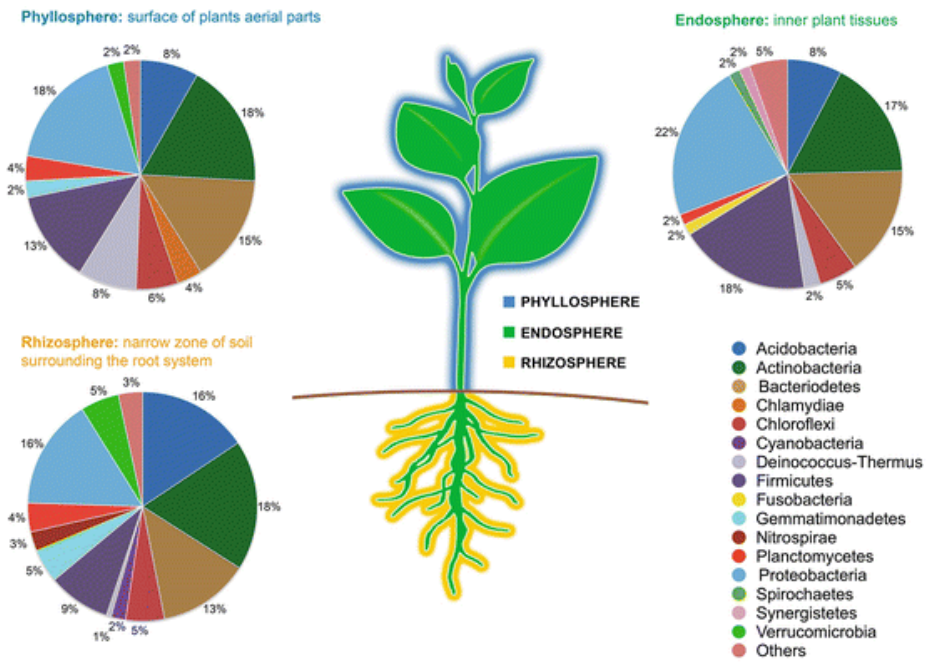


Figure 1. 3 Plant microbiome: composition and functions in plant compartments (Rossmann et al., 2017)

The rhizosphere, which is the soil area immediately surrounding plant roots, hosts a rich and important microbial plant community which in return provides a series of beneficial

outcomes related to plant growth (Bulgarelli et al., 2012). The root surface is the rhizoplane and often it is difficult to distinguish from the rhizosphere. Roots microbiota can be horizontally transferred or originate from the soil environment which contains a highly diverse microbiome (Mendes et al., 2013). As the biodiversity of the bulk soil decreases towards the root surface, cell microbial numbers however increase indicating the more favorable growth conditions of the selected microbes. In the rhizosphere there is a significant increase in bacteria belonging to phylum Proteobacteria which constitute the 50% of the community composition (Fitzpatrick et al., 2018; Hamonts et al., 2018). Rhizosphere soils have a higher water-holding capacity, increased nutrient availability, and greater microbial biomass compared to bulk soils. This is because the root exudates enrich the area in carbohydrates, lipids, and amino acids, as well as in secondary metabolites (terpenoids, flavonoids, and isoflavonoids) that serve a variety of functions including antimicrobials, and pathogen/herbivore defenses (Hardoim et al., 2008; Boller and He, 2009; Deakin and Broughton, 2009). The secretion of these plant exudates in the rhizosphere also drives the recruitment and selection of a special set of beneficial microbes which is called rhizosphere microbiome and which is very important for plant's adaptation to environmental changes and stresses (Long, 1989; Bolan, 1991). Environmental factors like elevated levels of CO2, drought, and nutrient deprivation (especially nitrogen and phosphorus) can influence the root exudate composition and the rhizosphere microbiome. Therefore, similar to the function of the gut, the composition of the rhizosphere microbiome is informative about the healthy or diseased state of the plant (Burdon and Thrall, 2009).

The phyllosphere, on the other hand, constitutes the aboveground surfaces of the plant, mostly composed of the leaves but also including blossoms, fruit, and stems (Vorholt, 2012).

The phyllosphere community mainly comprises bacteria belonging to phylum Proteobacteria, Bacteroidetes, Firmicutes and Actinomycetes. Differences in nutrient availability and environmental conditions between the rhizosphere and the phyllosphere contribute to their distinct microbial diversity and community distribution. The phyllosphere is a transient environment compared to the rhizosphere; seasonal changes in foliage, differences in annual versus perennial lifestyle, fluctuations in nutrients, water availability, and temperature, as well as constant exposure to damaging ultraviolet radiation, can lead to severe changes of the phyllosphere microbiota (Lindow and Brandl, 2003; Vorholt, 2012; Müller et al., 2016). The microbial phyllosphere communities is important for plant's life, protecting the plant to overcome herbivore and related biotic stresses (Saleem et al., 2017).

The endosphere consists of all the inner tissues of the plant at below-ground level as well as above-ground. Microbial network complexity decreases from the soil to endosphere compartments, since it requires the ability of the microorganisms to enter and colonize internal inter-cellular spaces (Hardoim et al., 2008; Berg et al., 2014b). Plant endophytic communities are frequently enriched in members of the Proteobacteria and Firmicutes (twofold in relative abundance compared to the rhizosphere), while they are depleted in members of Bacteroidetes and Acidobacteria (enriched in bulk soils) (Zarraonaindia et al., 2015; Hamonts et al., 2018).

Seeds also represent an important source of microorganisms, which proliferate in the roots of developing plants. In this way bacteria may be vertically transmitted via seeds, resulting in a community that is linked to the plant lifestyle; these bacteria are defined as seed-borne bacteria (Liu et al., 2012). For example, specific microbial groups that confer drought tolerance or disease resistance to the plant can be passed from the mother plants to their

offspring, suggesting a link between the host genome and the microbiome (Naylor and Coleman-Derr, 2018).

The assembly of the total plant-associated microbiome is a multistep process. Early colonizers are transmitted vertically, through seed transmission pathways and they preferentially become associated with above-ground plant tissues. Soil-derived microorganisms are mainly associated with the rhizosphere and they are recruited over the life cycle of their plant host. Microbiome composition is highly variable and dynamic during the early vegetative phase of the plant, while it stabilizes during the vegetative or reproductive phase (Edwards et al., 2015).

1.2.2 Composition and members of plant-associated bacterial communities

Plants host three functionally distinct microbial sub-communities: a ubiquitous and stable core microbiome, a variable microbiome and a highly variable microbial community that is responsive to biotic and abiotic processes/factors (Hernandez-Agreda et al., 2016). The core microbiome is a subset of microbial lineages, which is associated with a given host across a wide range of environmental conditions and provides critical functions (Niu et al., 2017; Fitzpatrick et al., 2018). Genome-wide association studies have shown that the host genome influence and tailor the microbiome composition acting on the heritable taxa which belong to the core microbiome, suggesting that these key microbial taxa could possibly carry genes with functions that are important for the host fitness (Kudjordjie et al., 2019). Many members of the core microbiome of different plant species are common at genus level (for example *Pseudomonas*, *Agrobacterium, Rhizobium, Methylobacterium, Sphingomonas and Erwinia*); suggesting the existence of a universal core plant microbiome (Astudillo-García et al., 2017;

Lemanceau et al., 2017; Yeoh et al., 2017). Importantly, the identification of co-occurring core microorganisms may be useful for manipulating and studying the mechanisms that drive community assembly and the interactions among different members (Xue et al., 2015; Niu et al., 2017). The use of synthetic communities (SymComs) provides an opportunity to validate the interaction networks and metabolic model predictions in natural settings and have emerged as an important tool to demonstrate their applicability in smart agricultural systems (Trivedi et al., 2020).

Bacterial members of microbiomes likely undergo several direct interactions such as predation, parasitism, mutualism and competition which influence microbiome composition. Similarly, symbiotic relationships drive the dynamics between plants and their associated microbial communities (Faust and Raes, 2012).

Recently, the importance of the relationships between hosts and the associated microorganisms gave rise to an important shift in the understanding of the microbial-host coevolution from the “separation” approach to the “holistic approach”. According to the “separation” approach, the microorganisms comprising the plant microbiome can be divided into pathogens, neutral, and symbionts/beneficial depending on their interactions with the host (Lederberg and McCray, 2001) (**Figure 1.4**). Plant growth-promoting bacteria (PGPB) are beneficial microorganisms since they can promote plant growth by direct or indirect mechanisms. The most common traits of PGPB bacteria are (i) the production of phytohormones like auxin, cytokinin, and gibberellin which affect plant growth through modulating endogenous hormone levels; (ii) the secretion of enzymes, such as the 1-aminocyclopropane-1-carboxylate (ACC) deaminase, which reduces the level of stress hormone ethylene in the plant; (iii) the production of a range of enzymes that can detoxify

reactive oxygen species, minimizing plant-induced stress; (iv) the biocontrol activities against plant-pathogen invasion and disease, through the niche exclusion by competition for nutrients, the production of antimicrobial compounds, lytic enzymes, siderophores and volatiles; (v) the ability to act as a mobile component of the plant defense, modulating plant hormones level and inducing plant systemic resistance; (vi) the presence of various microbial structures (such as secretion systems, flagella, pili and effector proteins) that contribute to plant defense by triggering the induced systemic resistance response (Arshad and Frankenberger Jr, 1997; Reinhold-Hurek and Hurek, 2011; Glick, 2014; Hopkins et al., 2017).

Strains of *Pseudomonas* spp., *Arthrobacter* spp. and *Bacillus* spp. and others can increase plant growth through the production of ACC deaminase (Mohamed and Gomaa, 2012). Similarly, bacteria from the rhizosphere microbiome including *Rhizobia* and *Mycorrizha* or *Pseudomonas* spp., have plant growth promotion properties like phosphorus solubilization, nitrogen fixation, indole acetic acid production, which are involved in improved nutrient uptake and plant growth. Besides, other bacteria like *Pseudomonas*, *Streptomyces*, *Bacillus*, *Pantoea*, *Burkholderia*, *Paenibacillus*, *Enterobacter* have been reported for their role in pathogen suppression (Berendsen et al., 2012).

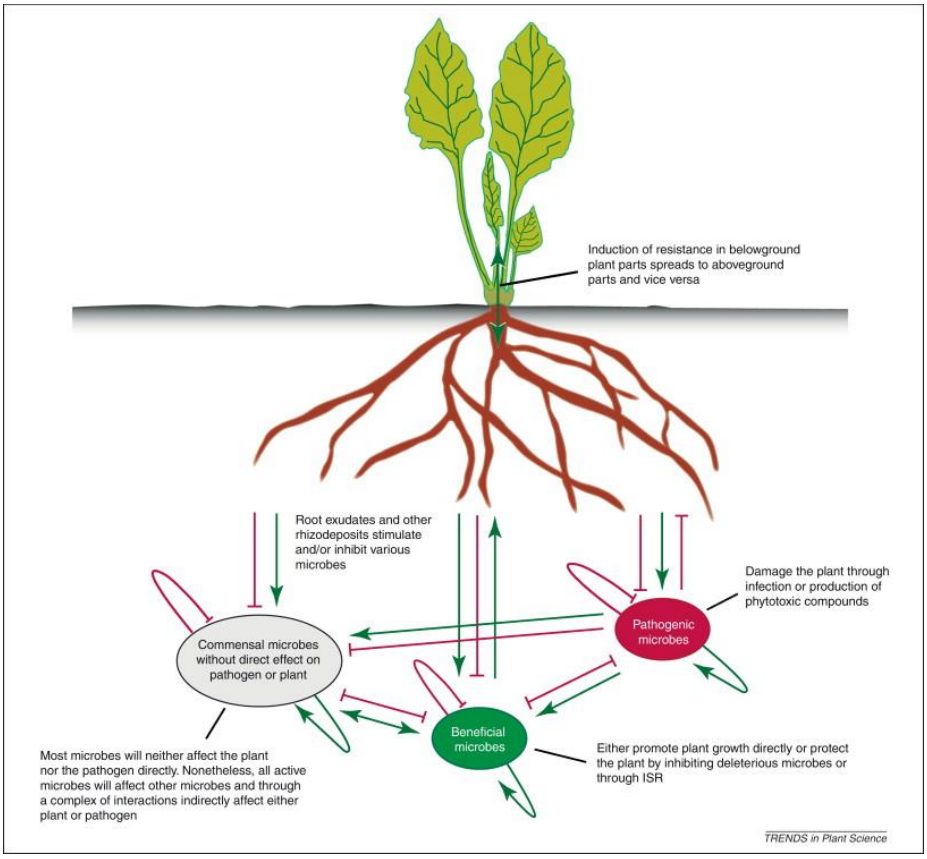


Figure 1. 4 Interactions in the rhizosphere (Berendsen et al., 2012)

Members of the plant microbiome can also cause disease symptoms. For example, *Pseudomonas syringae* is a model plant pathogen having a very broad host range including tomato, tobacco, olive, and green bean. Another example is *Erwinia amylovora* that causes fire blight disease of fruit trees and ornamentals plants (Mansfield et al., 2012). *Xanthomonas* spp*., Pseudomonas* spp*.* and *Dickeya* spp*.* are also associated with many important diseases of cereals. In order to protect their host from these pathogens, plant-associated bacteria can enhance host resistance against pathogen infections either through commensal-pathogen interactions or through modulating plant defense (Mansfield et al., 2012).

The plant microbiome also consists of a complex number of neutral collaborators/commensals not having a direct impact on plant life. Furthermore, commensals

or even beneficial microbes may co-operate with a primary pathogen resulting in a complex disease. Recent studies on opportunistic pathogens showed that host-microbe interactions depend not only on the host, but also on the entire microbiome shifting towards an holistic approach. The interplay between the host and its microbiome is therefore responsible for the healthy or disease state of the holobiont (host and its microbiome). Thus, the healthy state is most often accompanied with eubiosis, high diversity and similarity of the microbiome, while the disease state is linked to dysbiosis, low diversity and high variability and recently it has been called "pathobiome" state (Berg et al., 2015; Larsen and Claassen, 2018).

1.2.3 The pathobiome concept: from single pathogen to pathobiome

Although the great majority of bacteria found in nature live in multispecies communities, microbiological studies have thus mainly focused on single species. The same is true for studies on bacterial plant diseases that have thus far focused on the pathogen, with little attention/importance given on the many other microorganisms present in the infection sites (Vayssier-Taussat et al., 2014; Sweet and Bulling, 2017). The microbiota living in the infection site is now beginning to receive more attention for its possible involvement in the disease process, establishing cooperative, commensal or antagonistic interactions with microbial pathogens (Vayssier-Taussat et al., 2014).

The term 'pathobiome' has recently been coined defining the dynamics of the microbiome in response to stress and to a disease (Ryan, 2013; Vayssier-Taussat et al., 2014). The pathobiome is the totality of microbes interacting with a pathogen and their influence on pathogenesis and disease severity (Rovenich et al., 2014). Several studies are beginning to

highlight the limits to the historical concept of "one microbe-one disease", as described in Koch's postulates. This has been reported in the medical field, for example in human gut diseases (Gevers et al., 2012; Lloyd-Price et al., 2016) or polymicrobial oral diseases like dental carries and periodontitis (Ramsey et al., 2011). Multispecies infections have also been well documented in chronic infections, such as the ones occuring in the lungs of cystic fibrosis patients (Burmølle et al., 2010). Very likely, several plants-pathogens cooperate and team up with other accessories pathogens or plant commensal/resident bacteria, resulting in a multispecies complex interaction that drives the microbial disease (Buonaurio et al., 2015; Jakuschkin et al., 2016). An example is between the olive knot pathogen *Pseudomonas savastanoi* pv. *savastanoi* (*Psv*) and harmless endophytic *Erwinia toletana*. *Psv* causes the olive knot disease and the tumors (knots) contain a multispecies bacterial community, which affect the disease development (Buonaurio et al., 2015). *Psv* undergoes interspecies interactions with the harmless endophyte *E. toletana*; they colocalize and form a stable community, resulting in a more aggressive disease (**Figure 1.5**). These two bacterial species produce the same type of the *N*-acylhomoserine lactone (AHL) quorum sensing (QS) signal, and they share AHLs *in planta* (Hosni et al., 2011; Buonaurio et al., 2015; Caballo-Ponce et al., 2018). Another example of a pathobiome study is on the causal agent of oak powdery mildew, *Erysiphe alphitoides* (Jakuschkin et al., 2016). Here a bacterial and fungal species interact with the pathogen further highlighting that cooperation is taking place with members of the plant microbiome.

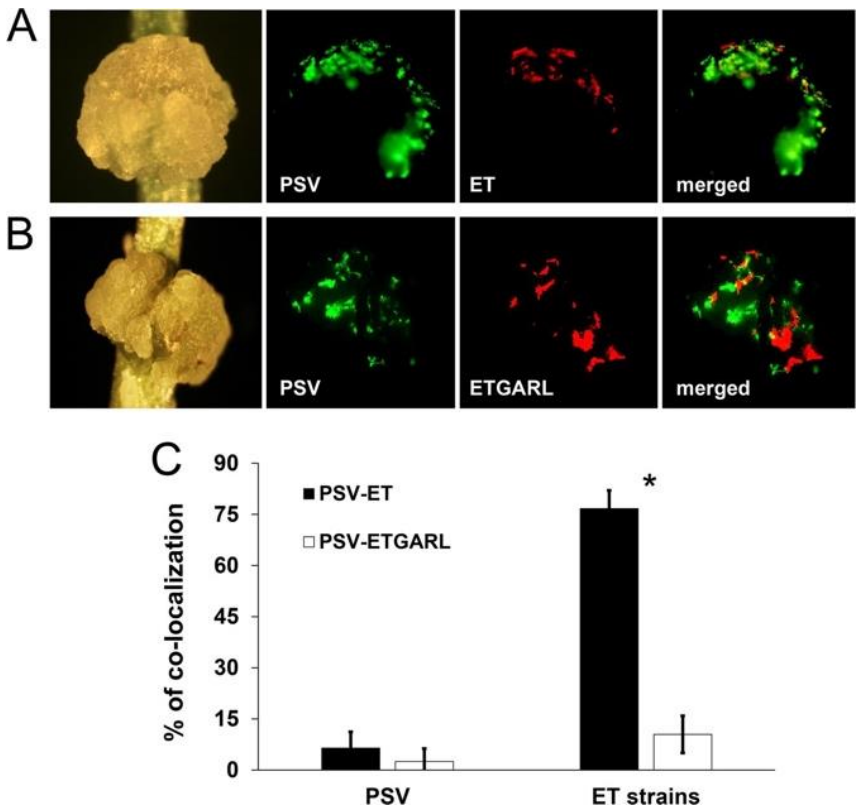


Figure 1. 5 Knots developed at 30 dpi in micropropagated olive plants after coinoculation of GFP-labeled P. savastanoi pv. savastanoi (PSV) with RFP-labeled E. toletana (ET) or ETGARL. (A) Coinoculation using GFP-labeled P. savastanoi pv. savastanoi and RFP-labeled E. toletana (Caballo-Ponce et al., 2018).

It is not clear whether interactions between commensal-resident microbes and incoming pathogens can either prevent or facilitate the establishment of a new infecting disease; very likely both the type of interactions take place. Thus, elucidation of the members and interactions of/in the pathobiome is necessary for understanding the evolution and the pathogenesis of microbial plant diseases. Microbial association networks provide a useful tool to predict positive links (indicating species co-occurrence/mutualism interaction) and negative link (indicating co-exclusion/competition or antagonism) (Jakuschkin et al., 2016).

A future challenge is to decipher the molecular mechanisms of the positive and negative interactions/ecological networks between plant pathogens and the resident microbial community to better understand the mechanisms of disease development and plant disease resistance; this could help in devising ways to control plant diseases.

1.2.4 Microbial networks and inter-microbial interactions play a crucial role for structuring plant-associated microbial communities

As mentioned above, the combination of host-microbe and microbe-microbe interactions is critical for the establishment of complex and diverse plant-associated communities.

Bioinformatic networks and co-occurrence analyses provides clues on the complexity of the microbial interactions but they are not sufficient to describe the nature and the molecular aspects of these interactions (Barberán et al., 2012; Faust and Raes, 2012). Microbial co-occurance networks allow to identify "hub species" which are represented by nodes, that have highest degree of connections with other species. These "hub microorganisms" can influence the community structure through strong biotic interactions with other microbial species and also with the host (van der Heijden and Hartmann, 2016). Acting directly on these hubs, host plants selectively influence the structure of their associated microbiome, which then transmits the information to the broader microbial network. The variation of one or two hub microorganisms has a significant effect on the assembly and organization of the microbiome and this variation is independent from external factors, such as the plant age, location or season (Agler et al., 2016). Moreover, hub species in networks can act as keystone species, which play a key role within the microbiome and have a greater impact on the dynamics of an ecosystem than other species (Banerjee et al., 2018) (**Figure 1.6**). Hub species within a co-occurrence network do not necessary play a role as keystone species (Berry and Widder, 2014). Keystone taxa are often rare species characterized by low abundance which orchestrate functionally and taxonomically diverse microbial groups. Keystone taxa may function alone, or a group of taxa with similar functioning may form a keystone guild and have a strong influence and impact on broad processes (Banerjee et al., 2018). In contrast,

microbial taxa whose abundance and presence does not correlate with other microbes, are called "peripheral species" and they are unaffected by other microbes in the network and have lower rates of microbe-microbe interactions (Barberán et al., 2012). Despite important developments in this field, it is still considered challenging to infer community organization based on co-occurrence networks (van der Heijden and Hartmann, 2016).

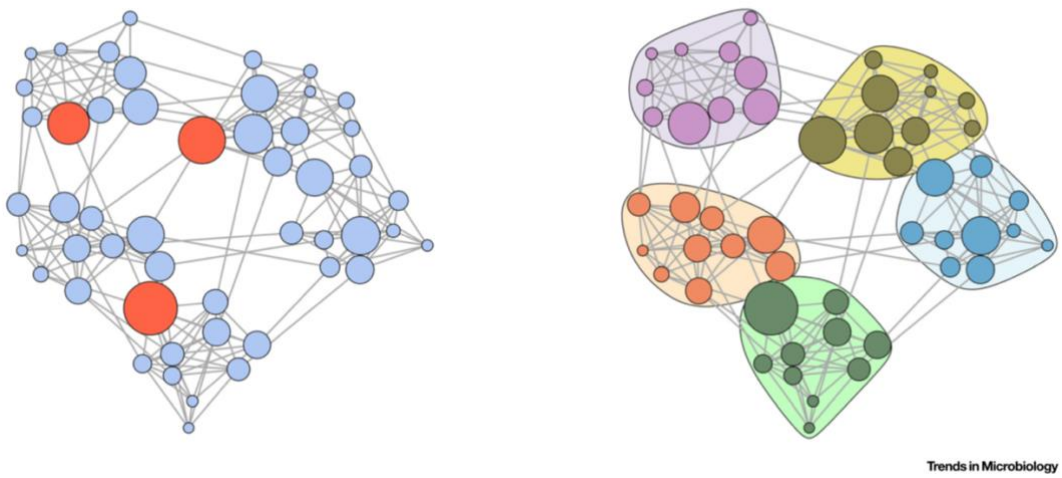


Figure 1. 6 A microbiome network with hub (keystone) species highlighted in red and the same microbiome network with microbes clustered into five distinct groups (Layeghifard et al., 2017).

Several cooperative and competitive interaction mechanisms are employed by the microbiota members to firstly colonize the plant environment and then persist within the plant microbiome. Most mechanistic knowledge of bacteria-bacteria and plant-bacteria interactions so far has been obtained using reductionist approaches such as mono/binary microbial set-ups and does not consider the synergistic interactions taking place in the microbiome.

The genomes of many plant-associated bacteria encode several secondary metabolites (such as non-ribosomal peptides, siderophores, lipopeptides, bacteriocins, toxins, polyketides and

a variety of antimicrobial, volatile compounds and signal molecules) and structures to undergo cell-cell contact with neighboring cells (such as vesicles, filaments, secretion systems and nanotubes) (Berg et al., 2014a; Tyc et al., 2017; D'Souza et al., 2018) which are involved in microorganism-microorganism interactions. In addition, the synergistic sharing of metabolic activities between distantly related bacteria or the nutrient sequestration among members of the microbiome most probably constitute another important biotic interaction in the establishment and maintenance of the microbiome (Little et al., 2008). Similarly, there are many occurrences of syntrophic interactions, in which primary metabolites or cofactors are exchanged and used by other members of the community as the substrate for creating hybrid secondary molecules (Wang and Seyedsayamdost, 2017).

Cell-cell signaling and quorum sensing (QS) is a well-established bacterial mechanism that dynamically regulates a variety of metabolic and physiological activities in response to the host, environment and microbial neighbors. Different bacteria taxa can synthetize the same type of signaling molecule (e.g. homoserine lactone; AHLs), which enable either cooperation or interference (quorum quenching) with other taxa (Whiteley et al., 2017). Metagenomic studies has evidenced that plant-associated microbiomes are enriched of AHLs compared to the bulk soil (Trivedi et al., 2020), suggesting an important role of the quorum sensing in the assembly, dynamics and stability of the microbiome.

1.3 Chemical cell-to-cell signaling in the microbiome

1.3.1 Bacterial cell-to-cell intraspecies signaling

In the complex plant microbiome environment, cell-cell signaling is likely taking place and regulating biological processes in response to environmental cues or to microbial cell density.

Quorum sensing (QS) is a cell-to-cell communication process that enables bacteria to collectively modify behavior in response to changes in the surrounding microbial community (Fuqua et al., 1994). This process involves the production, release, and detection of extracellular signaling molecules, which are called autoinducers. The concentration of these signals in a given environment is proportional to the number of bacteria present. Autoinducers accumulate in the environment as bacterial population density increases, translating extracellular information into internal changes in gene expression. These signals belong to a wide range of chemical classes, and multiple QS systems using different types of signals often occur within a single organism (Fuqua et al., 2001).

In Gram-negative bacteria, especially in the proteobacteria, the most common class of signals are the N-acyl-homoserine lactones (AHLs). Most of the ~25 identified AHLs contain unbranched aliphatic acyl groups that differ only in their length (4–18 carbons) and certain substituents (a 3-oxo or 3-hydroxyl group and/or a cis-alkene). A few have been found to contain aryl tails (e.g.,phenylacetanoyl and cinnamoyl) (Fuqua et al., 2001; Liao et al., 2018). Microorganisms interact using these signals in a variety of ways, including self-talk, cross-talk and eavesdropping (Chandler et al., 2012; Wellington and Greenberg, 2019).

A typical AHL QS system is composed of a LuxI-type protein responsible for synthesizing the AHL signals, which are produced at low but constant levels and which then interacts at quorum concentrations with the cognate LuxR-type transcription factors. The AHL: LuxR-type receptor complexes typically homodimerize and activate the transcription of target genes by binding specific DNA sequences (*lux-boxes*) at QS-regulated promoters. These target genes include the *luxI*-homologue, creating a positive feedback loop that is the hallmark of all known QS systems (Ng and Bassler, 2009). The Lux nomenclature originates from the first LuxI/R system identified in *Vibrio fischeri*, which produces luciferase at high density (via the lux operon) (Nealson et al., 1970; Schuster et al., 2013). The *luxI* and *luxR* homologs genes are almost always coupled and genetically linked in the chromosome (Fuqua et al., 2001)(**Figure 1.7**).

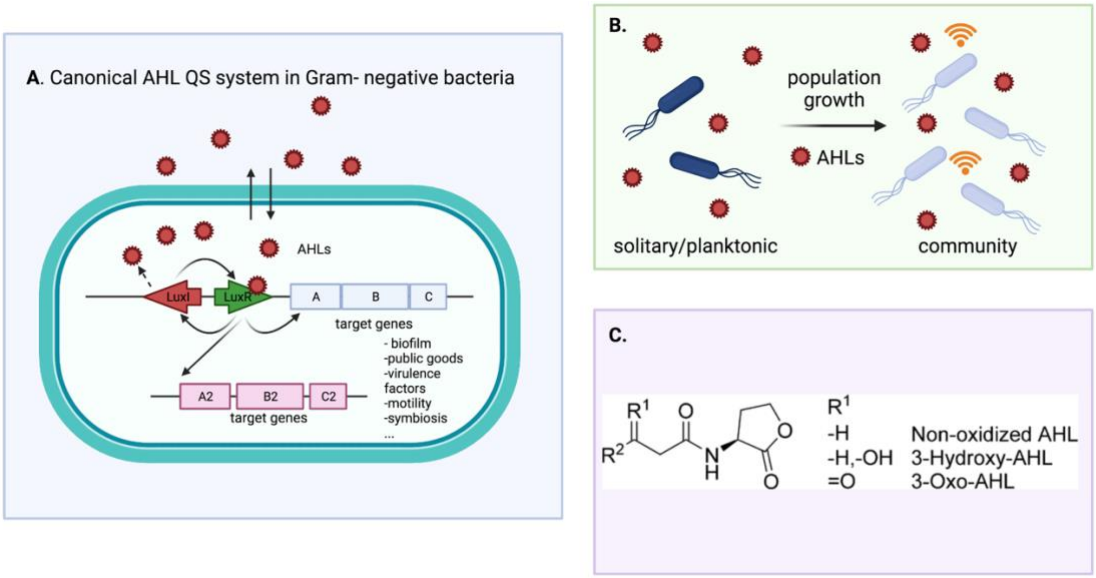


Figure 1. 7 A-B) Canonical AHL-QS system in gram negative bacteria, C) AHLs chemical structure and substituents. Figure created using Biorender.com

LuxI enzymes produce AHLs by deriving the lactone molecule from S-adenosylmethionine (SAM), and, in most cases, the specific acyl chain is obtained from intermediates of fatty acid

biosynthesis; the length of the acyl chain can affect the stability of the molecule and the dynamics of signaling processes (Case et al., 2008).

AHL-QS signaling is used by plant-associated bacteria to regulate and collective coordinate a wide variety of functions that are related to symbiosis, including roots colonization, exopolysaccharide production, biofilm formation, nodulation, or antibiotics production (Rosemeyer et al., 1998; Atkinson and Williams, 2009). Plant associated fluorescent pseudomonads have evolved QS-regulated synthesis of secondary metabolites implicated in antagonistic activities against plant pathogens, such as phenazines and pyoverdines (Pierson 3rd et al., 1994; Stintzi et al., 1998). Similarly, QS via AHLs is also used by many phytopathogens to regulate the expression of virulence-associated factors, motility, biofilm formation and colonization of host surfaces (Von Bodman et al., 2003).

1.3.2 Bacterial cell-to-cell interspecies signaling

To date most investigations on QS have involved mono-culture and reductionistic laboratory set ups; however, bacteria in nature mostly live as poly-microbial consortia, which most likely involve interspecies signaling through the action of diffusible chemical molecules (Duan et al., 2003; Ryan and Dow, 2008).

Several bacterial small molecules can mediate the cross-talk between microbes of different species. For example, some fatty acids and alcohols produced by bacteria have roles as interspecies signals; the diffusible signal factor (DSF) from *Xanthomonas campestris* is an interspecies signal, that can inhibit the hyphae formation of *Candida albicans* (Shank and Kolter, 2009). Microbial volatile organic compounds (VOCs), which are typically small molecules (100–500 Da, usually alkenes, alcohols, benzenoids, aldehydes, ketones,

terpenes), play an important role in long-distance interactions in microbial communities; they can regulate gene expression and influence microbial behaviors such as biofilm formation, virulence and stress tolerance (Audrain et al., 2015).

The role of AHLs in interspecies signaling is not very well studied (Hosni et al., 2011). Hosni et al. used the olive knot community between *Pseudomonas savatanoi* pv. *savastanoi* (*Psv*) and *Erwinia toletana* as a model system for study the role of the interspecies signaling and the sharing of AHL QS molecules between different bacterial species. Results have shown that the two bacteria produce and share the same two AHLs molecule (C6-3-oxo-HSL and C8-3oxo-HSL). Further evidence of AHL cross-talk occurs in *P. aeruginosa* which is triggered by resident microflora (Duan et al., 2003; Sibley et al., 2008).

However, there are many questions on the dynamics of interactions among bacteria in nature; it is still not known the types and extent by which cell-cell signals play a role in the formation of a stable microbial community. A major challenge will be to study this in planta and/or to develop simple model systems for the investigations of interspecies interactions.

1.3.3 Plant-bacteria interkingdom signaling

With the exception of unique relationships such as the rhizobial-legume symbiotic association (Peters et al., 1986), the pathogenesis between agrobacteria and their host plants (Hiei et al., 1994) and type III secretion-mediated pathogenesis (Hueck, 1998), the understanding of the mechanisms regulating plant-microbe interactions (interkingdom signaling) remains rather limited.

Plant compounds potentially involved in chemical signaling with bacteria includes sugars, amino acids, phenolics and polyamines (such as putrescine and arginine) (**Figure 1.8**). For

example, two plant phenolic compounds such as arbutin and D-fructose induce phytotoxin synthesis in the plant pathogen *P. syringae* pv. *syringae* (Wang et al., 2006). This is an example of the evolutionary ability of this plant-pathogen to sense and respond to the plant environment (Wang et al., 2006). Similarly, plant phenolics such as o-coumaric acid (OCA) and t-cinnamic acid (TCA) affect the expression of the type III secretion system (T3SS) in plant pathogenic *Dickeya dadantii* (Yang et al., 2008).

Other signals implicated in interkingdom signaling are the diketopiperazines (DKPs) which are cyclodipeptides, produced by several bacterial species. Interestingly, several plant growth-promoting *Pseudomonas* spp. release different DKPs, which stimulate root biomass and lateral root development (Holden et al., 1999). In addition, phytohormones such as the stringolactones (SLs) can function as ex-planta rhizosphere signaling molecules (Koltai, 2011; Wu et al., 2017). SLs are short-living compounds produced in roots and released to the rhizosphere. SLs are considered the most common metabolite in plant-fungi relationships and they are also involved in the regulation of nodulation of legumes by *Rhizobium* (De Cuyper and Goormachtig, 2017; Bouwmeester et al., 2019). Another potential class of interkingdom signals are the antibiotics which have been proposed to function at low non-inhibitory concentrations (Andersson and Hughes, 2014).

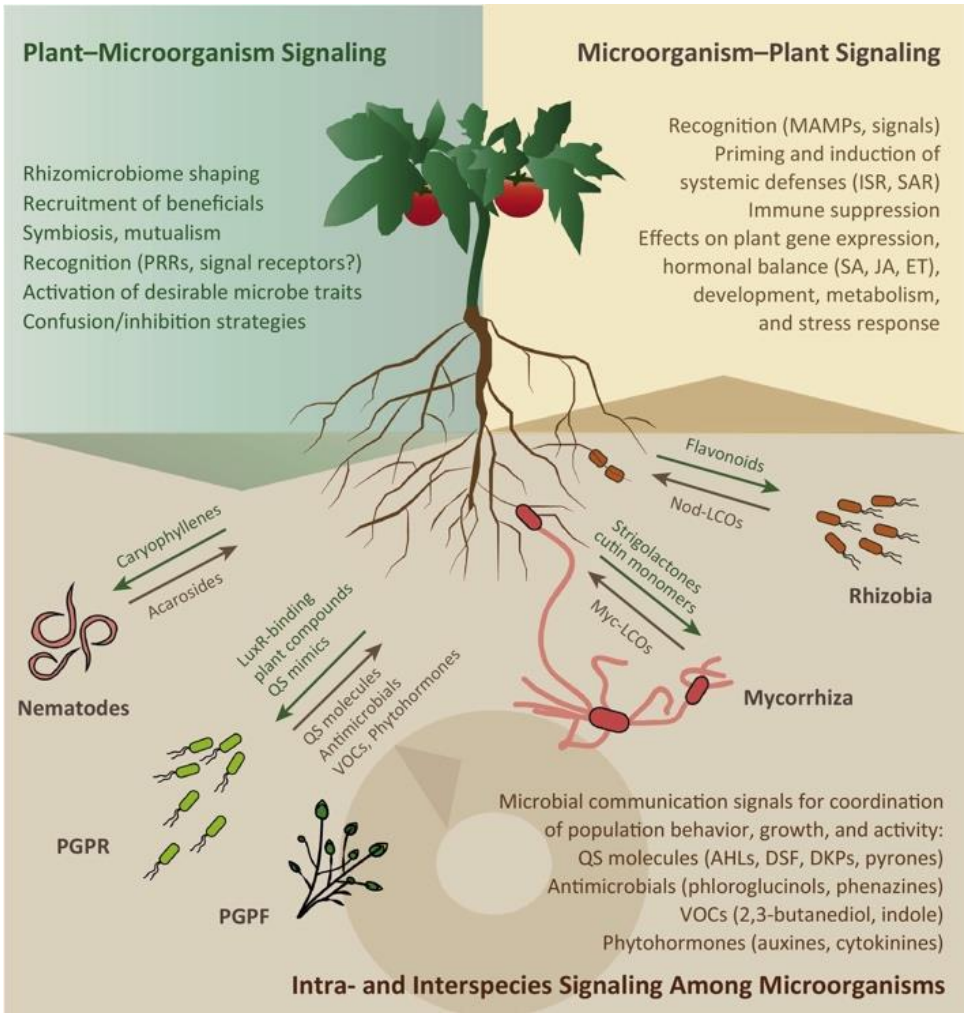


Figure 1. 8 Signaling in the rhizosphere (Venturi and Keel, 2016).

AHLs have also been implicated in plant-bacteria interactions since they exhibit structural similarities to eukaryotic hormones and they display biological roles on eukaryotic cells (Venturi and Keel, 2016). AHLs can affect plant gene expression, altering the levels of many proteins, including those involved in hormonal and defense response as observed in *Arabidopsis thaliana* (Schuhegger et al., 2006; You et al., 2006; von Rad et al., 2008). As AHLs are involved in the regulation of virulence factors in pathogenic bacteria, plants may have evolved strategies to interfere with the AHL signaling system in order to prevent them from initiating a pathogenic attack. Such interference could include the production of signal mimics, signal blockers or signal-degrading enzymes (McClean et al., 1997). Plant agonist

AHL-mimics can lead to pathogen confusion, reducing the pathogenicity since they stimulate a premature expression of virulence genes. Examples of AHL-mimics include the halogenated furanones produced by the Australian marine alga, *Delisea pulchra* which inhibit bacterial cell swarming and attachment responses, thus preventing the formation of bacterial biofilms on the algal surfaces (Givskov et al., 1996). Another example of plant-bacterial symbiosis involves the opportunistic alphaproteobacterium *Rhodopseudomonas palustris* and land plant roots. Under certain conditions, such as plant senescence, *p*-coumaric acid, which is a component of the lignin, accumulates in the rhizosphere. In *R. palustris*, *p*-coumaric acid activates 4-coumaroyl-homoserine lactone synthase (RpaI), which then uses the plant-derived *p*-coumaric acid to generate the hybrid signal *p*- coumaroyl-AHL and consequently regulate many genes involved in symbiosis (Pan et al., 2008; Schaefer et al., 2008).

In conclusion, QS systems are involved in diverse interactions, from competition to mutualism and symbiosis, that allow microbial species to interact to each other and exhibit social behaviors in natural communities. Bacteria are constantly adapting to the chemical language of other species in their proximity and adjust which gene sets are activated or deactivate. For this reason, AHL-QS gene circuits may be a flexible and adaptable genetic tool used by bacteria to adapt to new community members, new languages and ever-changing environment. It is very likely that bacteria are evolving (i) detect a wide variety of chemical signals, (ii) acquire QS circuits from neighbors into their own genomes via horizontally gene transfer, and (iii) alter which set of genes are activated by QS modifying or propagating the *lux-box* promoter regions (Prescott and Decho, 2020). Moreover, LuxR proteins may play a key role in adapting to new chemical languages establishing new communication networks,

thanks to their flexibility to recognize different AHLs and new signals, as it is discussed in below.

1.4 Quorum sensing LuxR-type proteins

QS LuxR-type family proteins are cytoplasmatic receptors/transcriptional regulators that bind and respond to AHLs. They are approximately 230-270 amino acids long and consist of two modular conserved domains LBD (N-terminal ligand-binding domain) and DBD (C-terminal DNA-binding helix-turn-helix domain) separated by a short linker region and acting independently (Slock et al., 1990; Choi and Greenberg, 1991). LuxR-type proteins bind to DNA at a conserved site called a *lux box*, which often consists of an inverted repeat recognition sequence of a 18-20bp palindrome that is usually located at -42.5 from the transcriptional starting site. AHL binding LuxRs, result in a conformational change upon binding to AHLs. Despite the very similar structures of natural AHLs, receptors are in most cases very selective for their cognate signal, ensuring bacteria cooperate and share resources with closely related cells. However, there are several examples of QS receptors that respond promiscuously to multiple signals (Wellington and Greenberg, 2019). The ability of these receptors to crosstalk with distantly related cells could be beneficial in both inter-species competition and cooperation, expanding the function of QS systems in bacterial communities. An example of crosstalk mediated by LuxR receptors occurs in the plant pathogen *Pectobacterium wasabiae;* this bacterium has two AHL-LuxR receptors and one has a broad signal specificity and responds at higher densities to other bacterial species and induce an earlier expression of virulence genes (Valente et al., 2017).

AHL LuxR homologs display a rather low primary sequence similarity (18% to 25%). However, the primary structures of the LBDs (with ~70% homology) and the DBDs (at least for TraR, QscR, and SdiA) are conserved. More specifically, nine amino acid residues that are critical for ligand-binding properties are highly conserved in AHL-binding LuxRs. Six of these are hydrophobic or aromatic and form the cavity of the AHL-binding domain (W57, Y61, D70, P71, W85, G113). The remaining three are in the HTH binding domain. These key residues are highly conserved in AHL-binding LuxR proteins, whereas a lack of conservation raises the possibility of binding to other/novel signaling molecules (Patankar and González, 2009). The dimerization interfaces and orientations of the two domains vary significantly in each structure, despite the binding of closely related (or even identical) AHLs.

Structural data for three length LuxR homologs complexed to native AHL activators (TraR from *Agrobacterium tumefaciens*, QscR from *P. aeruginosa*, and SdiA from *E. coli*) (**Figure 1.9**) (Zhang et al., 2002; Churchill and Chen, 2011; Lintz et al., 2011; Nguyen et al., 2015) and one complexed to a synthetic AHL antagonist (CviR from *Chromobacterium violaceum*)(Chen et al., 2011) have been reported.

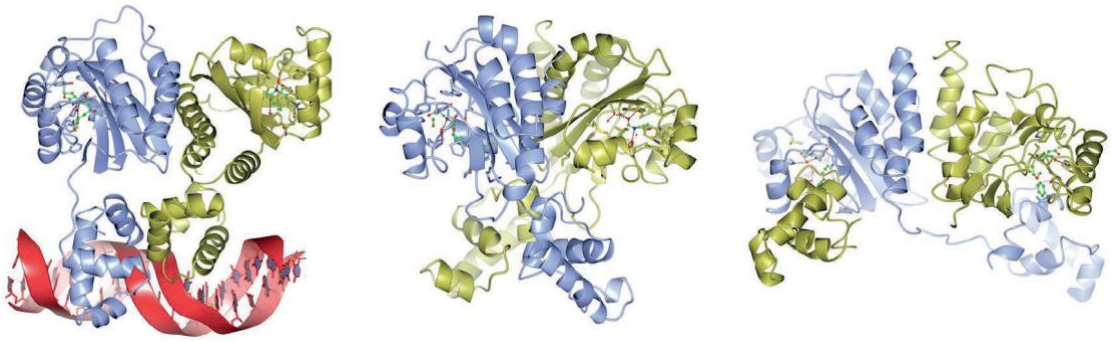


Figure 1. 9 Structure of LuxR type receptors. a) TraR from *Agrobacterium tumefaciens.* b) QscR from *Pseudomonas aeruginosa.* c) CviR from *Chromobacterium violaceum*

1.4.1 LuxR "solos": interspecies and interkingdom signaling

A sub-group of QS LuxRs lack a cognate LuxI-type homolog and they are called LuxR solos (Subramoni and Venturi, 2009b). LuxR solos have the same domain organization as canonical LuxR proteins and they have been found in many proteobacteria. They can expand the regulatory targets of the canonical QS systems by responding to endogenous or exogenous AHL or to non-AHL ligands, suggesting a role in inter-bacterial and host-bacterial interactions (Soares and Ahmer, 2011; Venturi and Fuqua, 2013). (**Figure 1.10**)

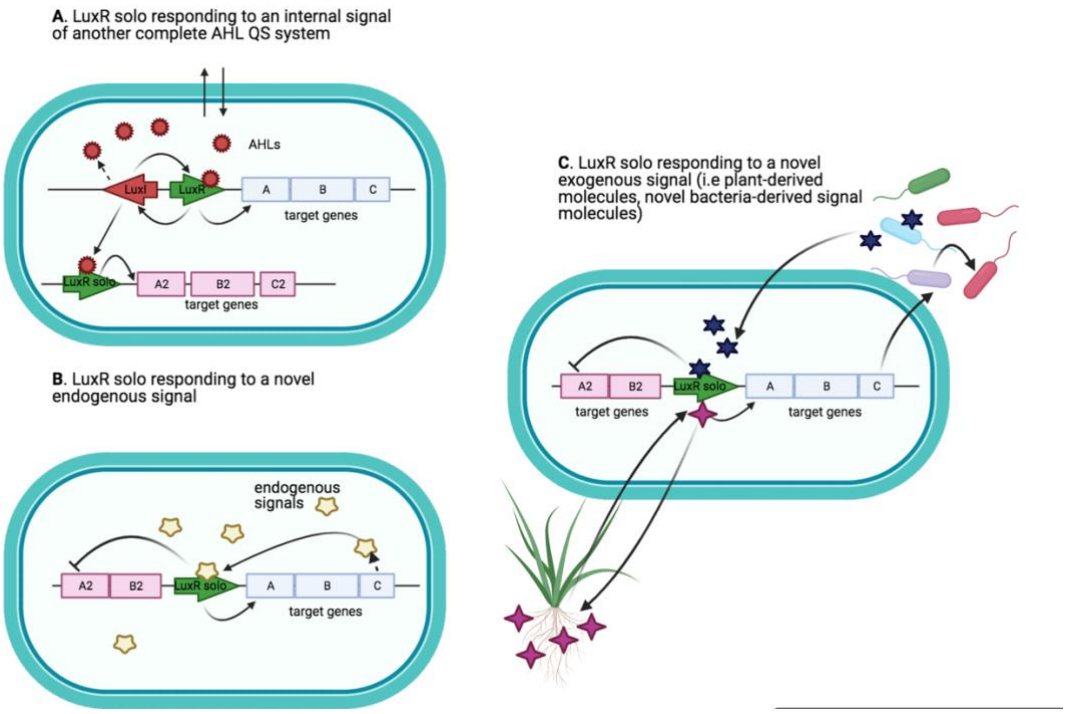


Figure 1. 10 A) LuxR solo responding to signals of a non-adjacent LuxI synthase in the same genome or B)responding to endogenous signals or C) external signals (plant/host or microbiome derived). Figure created using Biorender.com

Few LuxR solos have been functionally characterized so far; some respond to endogenously produced AHLs such as ExpR of *Sinorhizobium meliloti*, BisR of *Rhizobium leguminosarum bv.*

viciae, VjbR of *Brucella melitensis* and QscR of *Pseudomonas aeruginosa* (Ahmer et al., 1998; Chugani et al., 2001; Pellock et al., 2002; Lee et al., 2006).

QscR from *P. aeruginosa* is probably the best-characterized LuxR-solo receptor. QscR has a relaxed ligand-binding specificity and it is activated by the presence of nanomolar concentrations of C8-AHL, C10-AHL, 3-oxo-C10-AHL, C12-AHL, 3-oxo-C12-AHL and C14-AHL (Chugani et al., 2001; Lee et al., 2006).

LuxR solo SdiA from *Salmonella enterica* and *E. coli,* which do not produce AHLs, responds to AHLs synthesized by neighboring bacteria, thus suggesting a role interspecies signaling (Ahmer et al., 1998; Ahmer, 2004). Similarly, the *P. putida* PpoR LuxR solo binds to 3-oxo-C6-AHL and it is very well conserved among *P. putida* species (Subramoni and Venturi, 2009a).

A subfamily of LuxR solos present only in plant associated bacteria (PAB), has evolved to respond to plant signals (González et al., 2013; González and Venturi, 2013; Patel et al., 2013; Venturi and Fuqua, 2013). Compared to canonical QS LuxRs, these LuxR solos lack some conservation in the AHL-binding domain (Ferluga et al., 2007; Ferluga and Venturi, 2009). More precisely, W57 and Y61 are substituted by M and W, respectively (**Table 1.1)**. The evolution of these changes likely corresponds with the ability to bind low-molecular-weight compounds produced by plants.

Table 1. 1 Conservation of the amino-acids residues in the LBD of LuxR-type and LuxR solo responding to AHLs or plant-derivative compounds

	AHL – binding domain						HTH domain		
LuxR (*V.fisheri*)	W57	Y61	D70	P71	W85	G113	E178	L182	G188
PsoR (*P.fluorescens*)	W	W	D	P	W	G	E	L	G
OryR *(X.oryzae)*	M	W	D	P	W	G	E	L	G
XccR (*X. campestris*)	M	W	D	P	W	G	E	L	G
Sdia (*E. Coli*)	W	Y	D	P	W	G	E	L	G

Members of this sub-family include XccR of *Xanthomonas campestris*, OryR of *Xanthomonas oryzae*, PsoR of *Pseudomonas fluorescens*, XagR of *Xanthomonas axonopodis*, NesR in *Sinorhizobium meliloti*, and PipR of *Pseudomonas* sp.strain GM79 (**Figure 1.11**) (Ferluga et al., 2007; Zhang et al., 2007; Ferluga and Venturi, 2009; Patankar and González, 2009; Schaefer et al., 2016). A common feature that characterizes all these receptors is the adjacency to the virulence-associated proline iminopeptidase (*pip*) genes which are under the regulation of the LuxR solo and they are activated in response to plant signals (Schaefer et al., 2016). The *pip* genes have been implicated in virulence factors, but their mode of action remains unknown. PipR, from *Populus deltoides* root endophyte *Pseudomonas* sp.GM79, activates the downstream *pip* gene in response to an ethanolamine derivative (HEHEAA). This compound forms spontaneously from ethanolamine and serves as an intermediate in plant cell membrane biogenesis and plant-hormones (Coutinho et al., 2018; Luo et al., 2021).

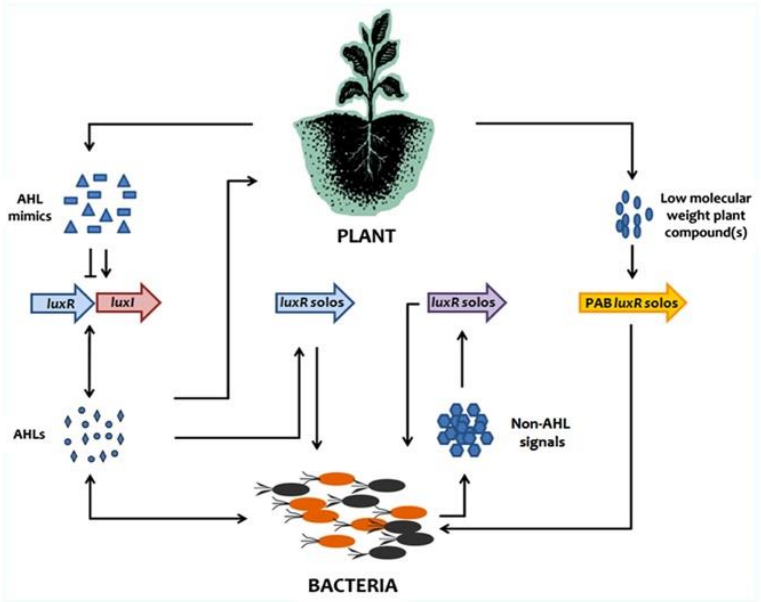


Figure 1. 11 Summary of current mode of action of AHL QS and of LuxR solos in signaling between plants and bacteria (Patel et al., 2013)**.**

Bacteria could have evolved QS-AHL-LuxR-type proteins that bind plant-compounds which allow them to sense their arrival inside the plant and regulate diverse cellular processes necessary for the bacterial life inside the host and its associated-microbial community.

Two LuxR solos PluR and PauR from *Photorhabdus luminescens* and *Photorhabdus asymbiotica* do not bind AHLs. These two species of *Photorhabdus* are human enteric and insect pathogens. These LuxR solos are part of a cell-to-cell signaling system which respond to a different endogenous signaling molecule. PluR senses alpha -pyrones, named photopyrones (PPY) and PauR detects dialkylresorcinols (DARs) and cyclohexanediones (CHDs). Both these molecules activate the expression of the corresponding nearby operon, leading to cell clumping and contributing to the virulence of *Photorhabdus* species (Brameyer et al., 2014, 2015). Interestingly, these two signals can also be bifunctional; PPYs can act as insect toxins at high concentrations, and DARs can act as antibiotics (Brameyer et al., 2015). However, it is disputable whether PluR can be designated as LuxR solo since its cognate signal

synthase has been identified as the adjacent gene *ppyS* (Brachmann et al., 2013). In addition, PluR and PauR both harbor four substitutions at similar positions in the conserved WYDPWG-motif of AHL-sensors, displaying a TYDQCS-motif and a TYDQYI-motif, respectively, which are important for signal-sensing, but not alone sufficient for signal-specificity (Brameyer et al., 2015).

1.4.2 Bacterial communication beyond acyl-homoserine lactones (AHLs)

As described above, LuxR solos are very widespread among proteobacteria; a genomics study revealed that out of 3550 total LuxR-type proteins identified, 2698 encoded for LuxR solos (Hudaiberdiev et al., 2015). However, the signal/ligand that is sensed by the majority of LuxR solos is not yet known. LuxR solos are likely to be major players in cell-cell signaling and most probably evolved bacteria to respond and communicate to many different cell-cell signals (Brameyer et al., 2014).

Other cell-cell signals related to AHLs are produced by bacteria (**Figure 1.12**). For example, members of the genus *Bradyrhizobium* and *Rhodopseudomonas* utilize aryl-AHLs as a chemical language instead of straight-chain fatty acyl-AHLs. These bacteria synthesize and use as signal the aryl-AHLs p-coumaroyl-HSL and cinnamoyl-HSL. Interestingly, p-coumaroyl-AHL is produced using plant-derived p-coumaric acid, so that cell-cell communication can occur only when these bacteria colonize plants (Schaefer et al., 2008; Ahlgren et al., 2011). *Ralstonia solanacearum* and *Xanthomonas campestris*, produce atypical autoinducers called 3-hydroxypalmitic-acid-methyl-ester (3-OH PAME) and (R)-methyl-3-hydroxymyristate, that control virulence traits and the formation of biofilms (Flavier et al., 1997; Kai et al., 2015).

It is also common for a single bacterium to use multiple signal molecules in order to regulate different cellular processes in diverse environmental conditions, host factors and community composition. An example is *P. aeruginosa*, which possesses four QS circuits, which besides using to AHL QS systems, it also uses 2-heptyl-3-hydroxy-4- quinolone (PQS) and the newly identified 2-(2-hydroxyphenyl)- thiazole-4-carbaldehyde (IQS) as languages. Another example is *Vibrio harveyi*, which integrates three different QS signaling molecules, allowing specific inter- and intraspecies communication (Rutherford and Bassler, 2012).

Indole and its derivatives have also recently been shown to be an intercellular, inter-species and inter-kingdom signaling molecule and it plays important roles in bacterial pathogenesis and plant immunity. In particular, indole controls plant defense systems and growth modulating the oxidative stress levels. However, the molecular mechanism of regulation of this possible novel signal and its function is still not completely deciphered (Lee et al., 2015).

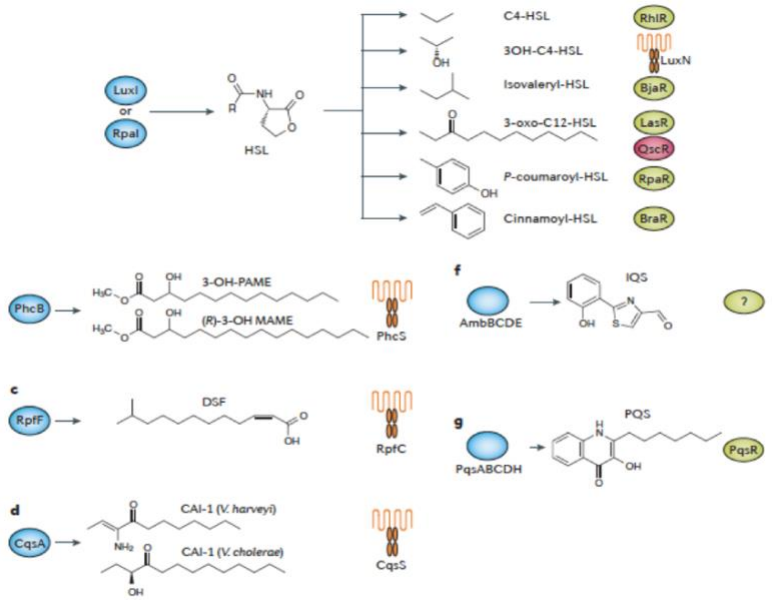


Figure 1. 12 Structures of various autoinducers together with their corresponding synthase (blue) and receptors (green) (Papenfort and Bassler, 2016)

1.5 Scope of this thesis

In natural environments microorganisms live in tight associations forming complex ecological interaction webs and are surrounded by dynamic and intricate exchange of small molecules.

The development of meta-omics and computational tools has allowed us to gain extraordinary insights into the taxonomic and functional composition of bacteria associated with plants. Similarly, GWAS and metagenome-wide association studies have identified key drivers that influence the formation of plant-associated microbiome and have linked individual microbial taxa to traits related to plant fitness and health. From an evolutionary point of view, this microbial biodiversity is fundamental for plant life, acting as a significant ally for the plant in controlling colonization/infection by plant pathogens and plant health. However, the evolution of the cell-cell bacterial interactions that favours the co-existence of highly diverse consortia remains very poorly understood. In the recent years, microbiologists have made considerable progress in unravelling the many different biotic mechanisms of interactions among bacteria, but always using reductionistic approaches. Thus, it remains at large unknown what is their possible role in the assembly, dynamics and stability of plant microbiomes. Moreover, it is important to focus on the molecular mechanisms that the bacteria use for the establishment of stable multispecies microbial communities, in order to protect the plant and enhance the activities of indigenous microbiome for a more sustainable agriculture.

The main aim of this thesis is to shed some light on how bacteria undergo inter-species signaling in the plant microbiome, as well as the identification of novel bacterial signaling systems/circuits that are involved in interspecies community formation. All the signaling studies were performed mainly using rice plants (Oryza sativa) as host-model. The

experimental work of this thesis is divided into three chapters: the first investigates how the rice microbiome changes/fluctuates in response to the attack of an emerging rice pathogen; the foot rot pathogen *Dickeya zeae*. The 16S rRNA amplicon-based community profiling approach has been followed and this plant bacterial disease has been used as a model in order to unravel potential microbial collaborators which interacts with the pathogen (pathobiome concept).

The second chapter is aimed to unravel the role played by the mechanism of QS in the formation and maintenance of plant microbiomes. The distribution, frequency and functional role of the LuxR solos receptors is investigated in the establishment of the multispecies community in the group of fluorescent *Pseudomonas* spp. and to understand which are the different modes of action of these “stand-alone” sensors/receptors. It is believed that this sub-family of QS related bacterial regulators contain members that respond to novel exogenous (or possibly novel endogenous) signals produced by other bacterial species or plants and hence play a role in inter-species/inter-kingdom signaling.

The third and the final data chapter presents studies on a novel cell-to-cell communication system, in which a LuxR solo is involved in the regulation and synthesis of a pigment-like molecule which in turn acts as a possible LuxR solo binding signal; this study expands the diversity of signaling molecules.

In summary, this thesis presents advances in the understanding of the role of interspecies bacterial-bacterial signaling in the plant-microbiome.

1.6 References

Agler, M.T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S.-T., Weigel, D., and Kemen, E.M. (2016) Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology* **14**: e1002352.

Ahlgren, N.A., Harwood, C.S., Schaefer, A.L., Giraud, E., and Greenberg, E.P. (2011) Aryl-homoserine lactone quorum sensing in stem-nodulating photosynthetic bradyrhizobia. *Proceedings of the National Academy of Sciences* **108**: 7183-7188.

Ahmer, B.M. (2004) Cell-to-cell signalling in *Escherichia coli* and *Salmonella enterica*. *Molecular microbiology* **52**: 933-945.

Ahmer, B.M., Van Reeuwijk, J., Timmers, C.D., Valentine, P.J., and Heffron, F. (1998) *Salmonella typhimurium* encodes an SdiA homolog, a putative quorum sensor of the LuxR family, that regulates genes on the virulence plasmid. *Journal of bacteriology* **180**: 1185-1193.

Andersson, D.I., and Hughes, D. (2014) Microbiological effects of sublethal levels of antibiotics. *Nature Reviews Microbiology* **12**: 465-478.

Arshad, M., and Frankenberger Jr, W.T. (1997) Plant growth-regulating substances in the rhizosphere: microbial production and functions. *Advances in agronomy* **62**: 45-151.

Astudillo-García, C., Bell, J.J., Webster, N.S., Glasl, B., Jompa, J., Montoya, J.M., and Taylor, M.W. (2017) Evaluating the core microbiota in complex communities: a systematic investigation. *Environmental Microbiology* **19**: 1450-1462.

Atkinson, S., and Williams, P. (2009) Quorum sensing and social networking in the microbial world. *Journal of the Royal Society Interface* **6**: 959-978.

Audrain, B., Farag, M.A., Ryu, C.-M., and Ghigo, J.-M. (2015) Role of bacterial volatile compounds in bacterial biology. *FEMS Microbiology Reviews* **39**: 222-233.

Bäckhed, F., Fraser, C.M., Ringel, Y., Sanders, M.E., Sartor, R.B., Sherman, P.M. et al. (2012) Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell host & microbe* **12**: 611-622.

Bai, Y., Müller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M. et al. (2015) Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* **528**: 364-369.

Banerjee, S., Schlaeppi, K., and van der Heijden, M.G. (2018) Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology* **16**: 567-576.

Barberán, A., Bates, S.T., Casamayor, E.O., and Fierer, N. (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME journal* **6**: 343-351.

Berendsen, R.L., Pieterse, C.M., and Bakker, P.A. (2012) The rhizosphere microbiome and plant health. *Trends in plant science* **17**: 478-486.

Berg, G., Mahnert, A., and Moissl-Eichinger, C. (2014a) Beneficial effects of plant-associated microbes on indoor microbiomes and human health? *Frontiers in microbiology* **5**: 15.

Berg, G., Krause, R., and Mendes, R. (2015) Cross-kingdom similarities in microbiome ecology and biocontrol of pathogens. *Frontiers in microbiology* **6**: 1311.

Berg, G., Grube, M., Schloter, M., and Smalla, K. (2014b) Unraveling the plant microbiome: looking back and future perspectives. *Frontiers in microbiology* **5**: 148.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T. et al. (2020) Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**: 1-22.

Berry, D., and Widder, S. (2014) Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in microbiology* **5**: 219.

Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., and Vorholt, J.A. (2014) A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet* **10**: e1004283.

Bolan, N. (1991) A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. *Plant and soil* **134**: 189-207.

Boller, T., and He, S.Y. (2009) Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science* **324**: 742-744.

Bouwmeester, H.J., Fonne-Pfister, R., Screpanti, C., and De Mesmaeker, A. (2019) Strigolactones: plant hormones with promising features. *Angewandte Chemie International Edition* **58**: 12778-12786.

Brachmann, A.O., Brameyer, S., Kresovic, D., Hitkova, I., Kopp, Y., Manske, C. et al. (2013) Pyrones as bacterial signaling molecules. *Nature chemical biology* **9**: 573-578.

Brameyer, S., Kresovic, D., Bode, H.B., and Heermann, R. (2014) LuxR solos in *Photorhabdus* species. *Frontiers in cellular and infection microbiology* **4**: 166.

Brameyer, S., Kresovic, D., Bode, H.B., and Heermann, R. (2015) Dialkylresorcinols as bacterial signaling molecules. *Proceedings of the National Academy of Sciences* **112**: 572-577.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., Van Themaat, E.V.L., and Schulze-Lefert, P. (2013) Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology* **64**: 807-838.

Bulgarelli, D., Rott, M., Schlaeppi, K., van Themaat, E.V.L., Ahmadinejad, N., Assenza, F. et al. (2012) Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* **488**: 91-95.

Buonaurio, R., Moretti, C., da Silva, D.P., Cortese, C., Ramos, C., and Venturi, V. (2015) The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. *Frontiers in plant science* **6**: 434.

Burdon, J.J., and Thrall, P.H. (2009) Coevolution of plants and their pathogens in natural habitats. *Science* **324**: 755-756.

Burmølle, M., Thomsen, T.R., Fazli, M., Dige, I., Christensen, L., Homøe, P. et al. (2010) Biofilms in chronic infections–a matter of opportunity–monospecies biofilms in multispecies infections. *FEMS Immunology & Medical Microbiology* **59**: 324-336.

Caballo-Ponce, E., Meng, X., Uzelac, G., Halliday, N., Cámara, M., Licastro, D. et al. (2018) Quorum sensing in *Pseudomonas savastanoi* pv. *savastanoi* and *Erwinia toletana*: role in virulence and interspecies interactions in the olive knot. *Applied and environmental microbiology* **84**: e00950-00918.

Case, R.J., Labbate, M., and Kjelleberg, S. (2008) AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. *The ISME journal* **2**: 345-349.

Chandler, J.R., Heilmann, S., Mittler, J.E., and Greenberg, E.P. (2012) Acyl-homoserine lactone-dependent eavesdropping promotes competition in a laboratory co-culture model. *The ISME journal* **6**: 2219-2228.

Chaparro, J.M., Badri, D.V., and Vivanco, J.M. (2014) Rhizosphere microbiome assemblage is affected by plant development. *The ISME journal* **8**: 790-803.

Chen, G., Swem, L.R., Swem, D.L., Stauff, D.L., O'Loughlin, C.T., Jeffrey, P.D. et al. (2011) A strategy for antagonizing quorum sensing. *Molecular cell* **42**: 199-209.

Choi, S., and Greenberg, E. (1991) The C-terminal region of the *Vibrio fischeri* LuxR protein contains an inducer-independent lux gene activating domain. *Proceedings of the National Academy of Sciences* **88**: 11115-11119.

Chugani, S.A., Whiteley, M., Lee, K.M., D'Argenio, D., Manoil, C., and Greenberg, E. (2001) QscR, a modulator of quorum-sensing signal synthesis and virulence in *Pseudomonas aeruginosa*. *Proceedings of the national academy of sciences* **98**: 2752-2757.

Churchill, M.E., and Chen, L. (2011) Structural basis of acyl-homoserine lactone-dependent signaling. *Chemical reviews* **111**: 68-85.

Compant, S., Clément, C., and Sessitsch, A. (2010) Plant growth-promoting bacteria in the rhizo-and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry* **42**: 669-678.

Compant, S., Duffy, B., Nowak, J., Clément, C., and Barka, E.A. (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Applied and environmental microbiology* **71**: 4951-4959.

Coutinho, B.G., Mevers, E., Schaefer, A.L., Pelletier, D.A., Harwood, C.S., Clardy, J., and Greenberg, E.P. (2018) A plant-responsive bacterial-signaling system senses an ethanolamine derivative. *Proceedings of the National Academy of Sciences* **115**: 9785-9790.

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., and Kost, C. (2018) Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports* **35**: 455-488.

De Cuyper, C., and Goormachtig, S. (2017) Strigolactones in the rhizosphere: friend or foe? *Molecular Plant-Microbe Interactions* **30**: 683-690.

Deakin, W.J., and Broughton, W.J. (2009) Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nature Reviews Microbiology* **7**: 312-320.

Doornbos, R.F., van Loon, L.C., and Bakker, P.A. (2012) Impact of root exudates and plant defense signaling on bacterial communities in the rhizosphere. A review. *Agronomy for Sustainable Development* **32**: 227-243.

Duan, K., Dammel, C., Stein, J., Rabin, H., and Surette, M.G. (2003) Modulation of *Pseudomonas aeruginosa* gene expression by host microflora through interspecies communication. *Molecular microbiology* **50**: 1477-1491.

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S. et al. (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences* **112**: E911-E920.

Faust, K., and Raes, J. (2012) Microbial interactions: from networks to models. *Nature Reviews Microbiology* **10**: 538-550.

Ferluga, S., and Venturi, V. (2009) OryR is a LuxR-family protein involved in interkingdom signaling between pathogenic *Xanthomonas oryzae* pv. *oryzae* and rice. *Journal of bacteriology* **191**: 890-897.

Ferluga, S., Bigirimana, J., Höfte, M., and Venturi, V. (2007) A LuxR homologue of *Xanthomonas oryzae* pv. *oryzae* is required for optimal rice virulence. *Molecular plant pathology* **8**: 529-538.

Fitzpatrick, C.R., Copeland, J., Wang, P.W., Guttman, D.S., Kotanen, P.M., and Johnson, M.T. (2018) Assembly and ecological function of the root microbiome across angiosperm plant species. *Proceedings of the National Academy of Sciences* **115**: E1157-E1165.

Flavier, A.B., Clough, S.J., Schell, M.A., and Denny, T.P. (1997) Identification of 3-hydroxypalmitic acid methyl ester as a novel autoregulator controlling virulence in *Ralstonia solanacearum*. *Molecular microbiology* **26**: 251-259.

Frey-Klett, P., Garbaye, J., and Tarkka, M. (2007) The mycorrhiza helper bacteria revisited. *New phytologist* **176**: 22-36.

Fuqua, C., Parsek, M.R., and Greenberg, E.P. (2001) Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. *Annual review of genetics* **35**: 439-468.

Fuqua, W.C., Winans, S.C., and Greenberg, E.P. (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of bacteriology* **176**: 269.

Gevers, D., Knight, R., Petrosino, J.F., Huang, K., McGuire, A.L., Birren, B.W. et al. (2012) The Human Microbiome Project: a community resource for the healthy human microbiome. *PLoS Biol* **10**: e1001377.

Gilbert, J.A., and Dupont, C.L. (2011) Microbial metagenomics: beyond the genome. *Annual review of marine science* **3**: 347-371.

Givskov, M., de Nys, R., Manefield, M., Gram, L., Maximilien, R., Eberl, L. et al. (1996) Eukaryotic interference with homoserine lactone-mediated prokaryotic signalling. *Journal of bacteriology* **178**: 6618-6622.

Glick, B.R. (2014) Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiological research* **169**: 30-39.

González, J.F., and Venturi, V. (2013) A novel widespread interkingdom signaling circuit. *Trends in plant science* **18**: 167-174.

González, J.F., Myers, M.P., and Venturi, V. (2013) The inter-kingdom solo OryR regulator of *Xanthomonas oryzae* is important for motility. *Molecular plant pathology* **14**: 211-221.

Gopal, M., and Gupta, A. (2016) Microbiome selection could spur next-generation plant breeding strategies. *Frontiers in microbiology* **7**: 1971.

Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S. et al. (2015) Microbiota and host nutrition across plant and animal kingdoms. *Cell host & microbe* **17**: 603-616.

Hamonts, K., Trivedi, P., Garg, A., Janitz, C., Grinyer, J., Holford, P. et al. (2018) Field study reveals core plant microbiota and relative importance of their drivers. *Environmental microbiology* **20**: 124-140.

Hardoim, P.R., van Overbeek, L.S., and van Elsas, J.D. (2008) Properties of bacterial endophytes and their proposed role in plant growth. *Trends in microbiology* **16**: 463-471.

Hartman, K., van der Heijden, M.G., Roussely-Provent, V., Walser, J.-C., and Schlaeppi, K. (2017) Deciphering composition and function of the root microbiome of a legume plant. *Microbiome* **5**: 1-13.

Hassani, M.A., Durán, P., and Hacquard, S. (2018) Microbial interactions within the plant holobiont. *Microbiome* **6**: 1-17.

Hernandez-Agreda, A., Leggat, W., Bongaerts, P., and Ainsworth, T.D. (2016) The microbial signature provides insight into the mechanistic basis of coral success across reef habitats. *MBio* **7**: e00560-00516.

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994) Efficient transformation of rice (*Oryza sativa L.*) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant Journal* **6**: 271-282.

Holden, M.T., Ram Chhabra, S., De Nys, R., Stead, P., Bainton, N.J., Hill, P.J. et al. (1999) Quorum-sensing cross talk: isolation and chemical characterization of cyclic dipeptides from *Pseudomonas aeruginosa* and other gram-negative bacteria. *Molecular microbiology* **33**: 1254-1266.

Hopkins, S.R., Wojdak, J.M., and Belden, L.K. (2017) Defensive symbionts mediate host–parasite interactions at multiple scales. *Trends in parasitology* **33**: 53-64.

Horton, M.W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B.D., Subramanian, S. et al. (2014) Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nature communications* **5**: 1-7.

Hosni, T., Moretti, C., Devescovi, G., Suarez-Moreno, Z.R., Guarnaccia, C., Pongor, S. et al. (2011) Sharing of quorum-sensing signals and role of interspecies communities in a bacterial plant disease. *The ISME journal* **5**: 1857-1870.

Hudaiberdiev, S., Choudhary, K.S., Vera Alvarez, R., Gelencsér, Z., Ligeti, B., Lamba, D., and Pongor, S. (2015) Census of solo LuxR genes in prokaryotic genomes. *Frontiers in cellular and infection microbiology* **5**: 20.

Hueck, C.J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and molecular biology reviews* **62**: 379-433.

Hughes, D.T., and Sperandio, V. (2008) Inter-kingdom signalling: communication between bacteria and their hosts. *Nature Reviews Microbiology* **6**: 111-120.

Jakuschkin, B., Fievet, V., Schwaller, L., Fort, T., Robin, C., and Vacher, C. (2016) Deciphering the pathobiome: intra-and interkingdom interactions involving the pathogen *Erysiphe alphitoides*. *Microbial ecology* **72**: 870-880.

Jiang, H., Ling, Z., Zhang, Y., Mao, H., Ma, Z., Yin, Y. et al. (2015) Altered fecal microbiota composition in patients with major depressive disorder. *Brain, behavior, and immunity* **48**: 186-194.

Kai, K., Ohnishi, H., Shimatani, M., Ishikawa, S., Mori, Y., Kiba, A. et al. (2015) Methyl 3-hydroxymyristate, a diffusible signal mediating phc quorum sensing in *Ralstonia solanacearum*. *ChemBioChem* **16**: 2309-2318.

Koltai, H. (2011) Strigolactones are regulators of root development. *New Phytologist* **190**: 545-549.

Kudjordjie, E.N., Sapkota, R., Steffensen, S.K., Fomsgaard, I.S., and Nicolaisen, M. (2019) Maize synthesized benzoxazinoids affect the host associated microbiome. *Microbiome* **7**: 1-17.

Lakshmanan, V., Selvaraj, G., and Bais, H.P. (2014) Functional soil microbiome: belowground solutions to an aboveground problem. *Plant physiology* **166**: 689-700.

Larsen, O.F., and Claassen, E. (2018) The mechanistic link between health and gut microbiota diversity. *Scientific reports* **8**: 1-5.

Layeghifard, M., Hwang, D.M., and Guttman, D.S. (2017) Disentangling interactions in the microbiome: a network perspective. *Trends in microbiology* **25**: 217-228.

Lederberg, J., and McCray, A.T. (2001) Ome SweetOmics--A genealogical treasury of words. *The Scientist* **15**: 8-8.

Lee, J.-H., Wood, T.K., and Lee, J. (2015) Roles of indole as an interspecies and interkingdom signaling molecule. *Trends in microbiology* **23**: 707-718.

Lee, J.H., Lequette, Y., and Greenberg, E.P. (2006) Activity of purified QscR, a *Pseudomonas aeruginosa* orphan quorum-sensing transcription factor. *Molecular microbiology* **59**: 602-609.

Leff, J.W., Del Tredici, P., Friedman, W.E., and Fierer, N. (2015) Spatial structuring of bacterial communities within individual Ginkgo biloba trees. *Environmental Microbiology* **17**: 2352-2361.

Lemanceau, P., Blouin, M., Muller, D., and Moënne-Loccoz, Y. (2017) Let the core microbiota be functional. *Trends in Plant Science* **22**: 583-595.

Liao, L., Schaefer, A.L., Coutinho, B.G., Brown, P.J., and Greenberg, E.P. (2018) An aryl-homoserine lactone quorum-sensing signal produced by a dimorphic prosthecate bacterium. *Proceedings of the National Academy of Sciences* **115**: 7587-7592.

Lindow, S.E., and Brandl, M.T. (2003) Microbiology of the phyllosphere. *Applied and environmental microbiology* **69**: 1875-1883.

Lintz, M.J., Oinuma, K.-I., Wysoczynski, C.L., Greenberg, E.P., and Churchill, M.E. (2011) Crystal structure of QscR, a *Pseudomonas aeruginosa* quorum sensing signal receptor. *Proceedings of the National Academy of Sciences* **108**: 15763-15768.

Little, A.E., Robinson, C.J., Peterson, S.B., Raffa, K.F., and Handelsman, J. (2008) Rules of engagement: interspecies interactions that regulate microbial communities. *Annu Rev Microbiol* **62**: 375-401.

Liu, Y., Zuo, S., Xu, L., Zou, Y., and Song, W. (2012) Study on diversity of endophytic bacterial communities in seeds of hybrid maize and their parental lines. *Archives of microbiology* **194**: 1001-1012.

Lloyd-Price, J., Abu-Ali, G., and Huttenhower, C. (2016) The healthy human microbiome. *Genome medicine* **8**: 1-11.

Long, S.R. (1989) Rhizobium-legume nodulation: life together in the underground. *Cell* **56**: 203-214.

Lugtenberg, B., and Kamilova, F. (2009) Plant-growth-promoting rhizobacteria. *Annual review of microbiology* **63**: 541-556.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S. et al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**: 86-90.

Luo, S., Coutinho, B.G., Dadhwal, P., Oda, Y., Ren, J., Schaefer, A.L. et al. (2021) Structural basis for a bacterial Pip system plant effector recognition protein. *Proceedings of the National Academy of Sciences* **118**.

Maignien, L., DeForce, E.A., Chafee, M.E., Eren, A.M., and Simmons, S.L. (2014) Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *MBio* **5**: e00682-00613.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P. et al. (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular plant pathology* **13**: 614-629.

McClean, K.H., Winson, M.K., Fish, L., Taylor, A., Chhabra, S.R., Camara, M. et al. (1997) Quorum sensing and *Chromobacterium violaceum*: exploitation of violacein production and inhibition for the detection of N-acylhomoserine lactones. *Microbiology* **143**: 3703-3711.

Mendes, R., Garbeva, P., and Raaijmakers, J.M. (2013) The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS microbiology reviews* **37**: 634-663.

Mohamed, H., and Gomaa, E. (2012) Effect of plant growth promoting *Bacillus subtilis* and *Pseudomonas fluorescens* on growth and pigment composition of radish plants (*Raphanus sativus*) under NaCl stress. *Photosynthetica* **50**: 263-272.

Müller, D.B., Vogel, C., Bai, Y., and Vorholt, J.A. (2016) The plant microbiota: systems-level insights and perspectives. *Annual review of genetics* **50**: 211-234.

Naylor, D., and Coleman-Derr, D. (2018) Drought stress and root-associated bacterial communities. *Frontiers in plant science* **8**: 2223.

Nealson, K.H., Platt, T., and Hastings, J.W. (1970) Cellular control of the synthesis and activity of the bacterial luminescent system. *Journal of bacteriology* **104**: 313-322.

Ng, W.-L., and Bassler, B.L. (2009) Bacterial quorum-sensing network architectures. *Annual review of genetics* **43**: 197-222.

Nguyen, Y., Nguyen, N.X., Rogers, J.L., Liao, J., MacMillan, J.B., Jiang, Y., and Sperandio, V. (2015) Structural and mechanistic roles of novel chemical ligands on the SdiA quorum-sensing transcription regulator. *MBio* **6**: e02429-02414.

Niu, B., Paulson, J.N., Zheng, X., and Kolter, R. (2017) Simplified and representative bacterial community of maize roots. *Proceedings of the National Academy of Sciences* **114**: E2450-E2459.

Ofek-Lalzar, M., Sela, N., Goldman-Voronov, M., Green, S.J., Hadar, Y., and Minz, D. (2014) Niche and host-associated functional signatures of the root surface microbiome. *Nature communications* **5**: 1-9.

Pan, C., Oda, Y., Lankford, P.K., Zhang, B., Samatova, N.F., Pelletier, D.A. et al. (2008) Characterization of anaerobic catabolism of p-coumarate in *Rhodopseudomonas palustris* by integrating transcriptomics and quantitative proteomics. *Molecular & cellular proteomics* **7**: 938-948.

Papenfort, K., and Bassler, B.L. (2016) Quorum sensing signal–response systems in Gram-negative bacteria. *Nature Reviews Microbiology* **14**: 576.

Patankar, A.V., and González, J.E. (2009) Orphan LuxR regulators of quorum sensing. *FEMS microbiology reviews* **33**: 739-756.

Patel, H.K., Suárez-Moreno, Z.R., Degrassi, G., Subramoni, S., González, J.F., and Venturi, V. (2013) Bacterial LuxR solos have evolved to respond to different molecules including signals from plants. *Frontiers in plant science* **4**: 447.

Peiffer, J.A., Spor, A., Koren, O., Jin, Z., Tringe, S.G., Dangl, J.L. et al. (2013) Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences* **110**: 6548-6553.

Pellock, B.J., Teplitski, M., Boinay, R.P., Bauer, W.D., and Walker, G.C. (2002) A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. *Journal of Bacteriology* **184**: 5067-5076.

Peters, N.K., Frost, J.W., and Long, S.R. (1986) A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* **233**: 977-980.

Philippot, L., Raaijmakers, J.M., Lemanceau, P., and Van Der Putten, W.H. (2013) Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology* **11**: 789-799.

Pierson 3rd, L., Keppenne, V.D., and Wood, D.W. (1994) Phenazine antibiotic biosynthesis in *Pseudomonas aureofaciens* 30-84 is regulated by PhzR in response to cell density. *Journal of bacteriology* **176**: 3966-3974.

Prescott, R.D., and Decho, A.W. (2020) Flexibility and adaptability of quorum sensing in nature. *Trends in microbiology* **28**: 436-444.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C. et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *nature* **464**: 59-65.

Raaijmakers, J.M., Paulitz, T.C., Steinberg, C., Alabouvette, C., and Moënne-Loccoz, Y. (2009) The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant and soil* **321**: 341-361.

Ramsey, M.M., Rumbaugh, K.P., and Whiteley, M. (2011) Metabolite cross-feeding enhances virulence in a model polymicrobial infection. *PLoS pathog* **7**: e1002012.

Rastogi, G., Sbodio, A., Tech, J.J., Suslow, T.V., Coaker, G.L., and Leveau, J.H. (2012) Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *The ISME journal* **6**: 1812-1822.

Redford, A.J., Bowers, R.M., Knight, R., Linhart, Y., and Fierer, N. (2010) The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental microbiology* **12**: 2885-2893.

Reinhold-Hurek, B., and Hurek, T. (2011) Living inside plants: bacterial endophytes. *Current opinion in plant biology* **14**: 435-443.

Rosemeyer, V., Michiels, J., Verreth, C., and Vanderleyden, J. (1998) luxI-and luxR-homologous genes of *Rhizobium etli* CNPAF512 contribute to synthesis of autoinducer molecules and nodulation of *Phaseolus vulgaris*. *Journal of Bacteriology* **180**: 815-821.

Rossmann, M., Sarango-Flores, S.W., Chiaramonte, J.B., Kmit, M.C.P., and Mendes, R. (2017) Plant microbiome: composition and functions in plant compartments. In *The Brazilian Microbiome*: Springer, pp. 7-20.

Rovenich, H., Boshoven, J.C., and Thomma, B.P. (2014) Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current opinion in plant biology* **20**: 96-103.

Rutherford, S.T., and Bassler, B.L. (2012) Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harbor perspectives in medicine* **2**: a012427.

Ryan, E.T. (2013) The intestinal pathobiome: its reality and consequences among infants and young children in resource-limited settings. In: Oxford University Press.

Ryan, R.P., and Dow, J.M. (2008) Diffusible signals and interspecies communication in bacteria. *Microbiology* **154**: 1845-1858.

Saleem, M., Meckes, N., Pervaiz, Z.H., and Traw, M.B. (2017) Microbial interactions in the phyllosphere increase plant performance under herbivore biotic stress. *Frontiers in microbiology* **8**: 41.

Schaefer, A.L., Oda, Y., Coutinho, B.G., Pelletier, D.A., Weiburg, J., Venturi, V. et al. (2016) A LuxR homolog in a cottonwood tree endophyte that activates gene expression in response to a plant signal or specific peptides. *MBio* **7**: e01101-01116.

Schaefer, A.L., Greenberg, E., Oliver, C.M., Oda, Y., Huang, J.J., Bittan-Banin, G. et al. (2008) A new class of homoserine lactone quorum-sensing signals. *Nature* **454**: 595-599.

Schuhegger, R., Ihring, A., Gantner, S., Bahnweg, G., Knappe, C., Vogg, G. et al. (2006) Induction of systemic resistance in tomato by N-acyl-L-homoserine lactone-producing rhizosphere bacteria. *Plant, Cell & Environment* **29**: 909-918.

Schuster, M., Joseph Sexton, D., Diggle, S.P., and Peter Greenberg, E. (2013) Acyl-homoserine lactone quorum sensing: from evolution to application. *Annual review of microbiology* **67**: 43-63.

Shank, E.A., and Kolter, R. (2009) New developments in microbial interspecies signaling. *Current opinion in microbiology* **12**: 205-214.

Sibley, C.D., Duan, K., Fischer, C., Parkins, M.D., Storey, D.G., Rabin, H.R., and Surette, M.G. (2008) Discerning the complexity of community interactions using a Drosophila model of polymicrobial infections. *PLoS Pathog* **4**: e1000184.

Slock, J., VanRiet, D., Kolibachuk, D., and Greenberg, E. (1990) Critical regions of the *Vibrio fischeri* LuxR protein defined by mutational analysis. *Journal of Bacteriology* **172**: 3974-3979.

Soares, J.A., and Ahmer, B.M. (2011) Detection of acyl-homoserine lactones by *Escherichia* and *Salmonella*. *Current opinion in microbiology* **14**: 188-193.

Soni, R., Kumar, V., Suyal, D.C., Jain, L., and Goel, R. (2017) Metagenomics of plant rhizosphere microbiome. In *Understanding Host-Microbiome Interactions-An Omics Approach*: Springer, pp. 193-205.

Stintzi, A., Evans, K., Meyer, J.-m., and Poole, K. (1998) Quorum-sensing and siderophore biosynthesis in *Pseudomonas aeruginosa*: lasRllasI mutants exhibit reduced pyoverdine biosynthesis. *FEMS microbiology letters* **166**: 341-345.

Subramoni, S., and Venturi, V. (2009a) PpoR is a conserved unpaired LuxR solo of *Pseudomonas putida* which binds N-acyl homoserine lactones. *BMC microbiology* **9**: 1-15.

Subramoni, S., and Venturi, V. (2009b) LuxR-family 'solos': bachelor sensors/regulators of signalling molecules. *Microbiology* **155**: 1377-1385.

Sweet, M.J., and Bulling, M.T. (2017) On the importance of the microbiome and pathobiome in coral health and disease. *Frontiers in Marine Science* **4**: 9.

Trivedi, P., Leach, J.E., Tringe, S.G., Sa, T., and Singh, B.K. (2020) Plant–microbiome interactions: From community assembly to plant health. *Nature Reviews Microbiology* **18**: 607-621.

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., and Gordon, J.I. (2007) The human microbiome project. *Nature* **449**: 804-810.

Turner, T.R., James, E.K., and Poole, P.S. (2013) The plant microbiome. *Genome biology* **14**: 1-10.

Tyc, O., Song, C., Dickschat, J.S., Vos, M., and Garbeva, P. (2017) The ecological role of volatile and soluble secondary metabolites produced by soil bacteria. *Trends in microbiology* **25**: 280-292.

Valente, R.S., Nadal-Jimenez, P., Carvalho, A.F., Vieira, F.J., and Xavier, K.B. (2017) Signal integration in quorum sensing enables cross-species induction of virulence in *Pectobacterium wasabiae*. *MBio* **8**: e00398-00317.

van der Heijden, M.G., and Hartmann, M. (2016) Networking in the plant microbiome. *PLoS Biology* **14**: e1002378.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., and Dufresne, A. (2015) The importance of the microbiome of the plant holobiont. *New Phytologist* **206**: 1196-1206.

Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.F., Jacques, M.-A., Lebrun, M.-H. et al. (2014) Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Frontiers in cellular and infection microbiology* **4**: 29.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A. et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *science* **304**: 66-74.

Venturi, V., and Fuqua, C. (2013) Chemical signaling between plants and plant-pathogenic bacteria. *Annual review of phytopathology* **51**: 17-37.

Venturi, V., and Keel, C. (2016) Signaling in the rhizosphere. *Trends in plant science* **21**: 187-198.

Von Bodman, S.B., Bauer, W.D., and Coplin, D.L. (2003) Quorum sensing in plant-pathogenic bacteria. *Annual review of phytopathology* **41**: 455-482.

von Rad, U., Klein, I., Dobrev, P.I., Kottova, J., Zazimalova, E., Fekete, A. et al. (2008) Response of *Arabidopsis thaliana* to N-hexanoyl-DL-homoserine-lactone, a bacterial quorum sensing molecule produced in the rhizosphere. *Planta* **229**: 73-85.

Vorholt, J.A. (2012) Microbial life in the phyllosphere. *Nature Reviews Microbiology* **10**: 828-840.

Wang, N., Lu, S.-E., Wang, J., Chen, Z.J., and Gross, D.C. (2006) The expression of genes encoding lipodepsipeptide phytotoxins by *Pseudomonas syringae* pv. *syringae* is coordinated in response to plant signal molecules. *Molecular plant-microbe interactions* **19**: 257-269.

Wang, R., and Seyedsayamdost, M.R. (2017) Opinion: Hijacking exogenous signals to generate new secondary metabolites during symbiotic interactions. *Nature Reviews Chemistry* **1**: 1-8.

Wellington, S., and Greenberg, E.P. (2019) Quorum sensing signal selectivity and the potential for interspecies cross talk. *MBio* **10**: e00146-00119.

Whiteley, M., Diggle, S.P., and Greenberg, E.P. (2017) Progress in and promise of bacterial quorum sensing research. *Nature* **551**: 313-320.

Wu, Y., Dor, E., and Hershenhorn, J. (2017) Strigolactones affect tomato hormone profile and somatic embryogenesis. *Planta* **245**: 583-594.

Xue, C., Penton, C.R., Shen, Z., Zhang, R., Huang, Q., Li, R. et al. (2015) Manipulating the banana rhizosphere microbiome for biological control of Panama disease. *Scientific reports* **5**: 1-11.

Yang, S., Peng, Q., San Francisco, M., Wang, Y., Zeng, Q., and Yang, C.-H. (2008) Type III secretion system genes of *Dickeya dadantii* 3937 are induced by plant phenolic acids. *PloS one* **3**: e2973.

Yeoh, Y.K., Dennis, P.G., Paungfoo-Lonhienne, C., Weber, L., Brackin, R., Ragan, M.A. et al. (2017) Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. *Nature communications* **8**: 1-9.

You, Y.-S., Marella, H., Zentella, R., Zhou, Y., Ulmasov, T., Ho, T.-H.D., and Quatrano, R.S. (2006) Use of bacterial quorum-sensing components to regulate gene expression in plants. *Plant physiology* **140**: 1205-1212.

Zarraonaindia, I., Owens, S.M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax, S. et al. (2015) The soil microbiome influences grapevine-associated microbiota. *MBio* **6**: e02527-02514.

Zhang, L., Jia, Y., Wang, L., and Fang, R. (2007) A proline iminopeptidase gene upregulated in planta by a LuxR homologue is essential for pathogenicity of *Xanthomonas campestris* pv. *campestris*. *Molecular microbiology* **65**: 121-136.

Zhang, R.-g., Pappas, K.M., Brace, J.L., Miller, P.C., Oulmassov, T., Molyneaux, J.M. et al. (2002) Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature* **417**: 971-974.

Zheng, P., Zeng, B., Zhou, C., Liu, M., Fang, Z., Xu, X. et al. (2016) Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. *Molecular psychiatry* **21**: 786-796.

Chapter II

The rice foot rot pathogen *Dickeya zeae* alters the in-field plant microbiome

Keywords: Plant pathobiome, Microbial communities, Biotic interactions, Metabarcoding, Pathogens, *Dickeya zeae*, Rice foot rot disease

2.1 Introduction

It has been long accepted that disease severity is a multifaceted mechanism, being the outcome of interactions between the pathogen, host, and the environment (Brader et al., 2017). Plants are colonized and live in association with a large number of different microorganisms which play important roles in plant health and resistance to biotic and abiotic stresses (Schlaeppi and Bulgarelli, 2015; Vannier et al., 2019; Liu et al., 2020). Members of the plant microbiome undergo direct interactions such as predation, parasitism, mutualism, and competition (Faust and Raes, 2012). What happens then to the plant microbiome when an incoming pathogen establishes itself and causes disease? Does the microbial environment change, tolerate, and/or assist the colonization of the pathogen? There are several reports, especially in humans and animals, demonstrating that a successful pathogen invasion disrupts the resident-microbiota, resulting in a shift from a healthy to a dysbiotic unstable state (Anna Karenina principle) (Clemente et al., 2012; Gilbert, 2016; Zaneveld et al., 2017; Proctor, 2019). It is fundamental to take into consideration that interactions between the pathogen and other microorganisms (harmless, neutral, or even beneficial) of the microbiome can occur and affect positively or negatively the virulence thus adding a fourth dimension to the disease triangle (Brader et al., 2017). Besides, certain plant diseases are complex, involving the interaction/cooperation of different pathogens (Lamichhane and Venturi, 2015). The recent perception that the microbiome contributes to disease formation and severity, along with the discovery of complex diseases has led to the introduction of the term “pathobiome” (Vayssier-Taussat et al., 2014), which is the pathogen(s) integrated with the host-associated microorganisms (encompassing prokaryotes, eukaryotes and viruses). Elucidation of the members of the pathobiome could identify the key biomarker species that

can team-up with the pathogen (Agler et al., 2016). This could become important for understanding pathogenesis, persistence, transmission, and evolution of several plant pathogens and for developing microbiome-based plant protection strategies (Schlaeppi and Bulgarelli, 2015).

Dickeya spp.is one of the top ten important bacterial phytopathogens in the world (Mansfield et al., 2012).There are currently eight species in this genus, including *D. chrysanthemi, D. dadantii, D. dianthicola, D. dieffenbachiae, D. paradisiaca, D. zeae, D. aquatic, D. solani* (Samson et al., 2005; Brady et al., 2012) and among them *D. dadantii, D. solani and D. zeae* cause devasting disease, resulting in a considerable loss in crop yield. *Dickeya* spp. causes soft rot disease in a wide variety of economically important crops and ornamental plants such as *Zea mays*, *Oryza sativa*, *Solanum tuberosum*, and *Musa* spp. in different parts of the world (Sabet, 1954; Jafra et al., 2009; Sławiak et al., 2009; Laurila et al., 2010). Infections by *Dickeya* spp. usually results in maceration and rotting of parenchymatous tissue of the affected organ. In particular, infected rice plants by *D. zeae* present a dark brown decay of the tillers at the site of infection, which later lead to the collapse of the entire plant (Goto, 1979; Barras et al., 1994; Collmer and Bauer, 1994; Nassar et al., 1994; Hussain et al., 2008; Pu et al., 2012). *D. zeae* are often present in latent infections on many host crops and they can persist overwinter in contaminated plant residues. Under certain conditions, such as in decreasing O_2 concentration, high temperature, high humidity and water film on the surface of the plant organs, latency is broken and the bacteria start to grow and cause decay. In these situations, the inoculum produced in one growing season persists to the next, and increases in load over a period of years. A change in spatial distribution of infected plants occurs, it appears as “spots” with the highest density in the center of infected plants. In most cases, penetration

by bacteria occurs via breaks in the periderm caused by bruising in harvesting, insects, nematodes, or fungal infections (Perombelon and Kelman, 1980). Unlike the other members of the genus, *D. zeae* can infect both monocotyledons and dicotyledons; it produces large quantities of pectic enzymes, phytotoxins and bacteriocins that enable it to macerate the plant parenchymatous tissue and result in disease (Samson et al., 2005; Zhou et al., 2011; Cheng et al., 2013). A quorum sensing regulated *zms* gene cluster in *D. zeae* rice isolates, which encodes the biosynthesis of zeamine phytotoxins, is responsible of inhibiting rice seeds germination and growth; no other virulence associated mechanisms is currently known (Zhou et al., 2011; Zhou et al., 2016).

This study performs pathobiome analysis of the rice foot rot disease, caused by *D. zeae*, as a model in order to investigate the effects of this bacterial pathogen to the total resident microbiome and to highlight possible interactions between the pathogen and the members of the microbiota. The bacterial communities of field-grown rice plants with or without symptoms of foot rot over two rice growing seasons and belonging to two different rice cultivars have been characterized via 16S rRNA amplicon sequencing. Our main hypothesis was that *D. zeae* co-operates with other members, either harmless or pathogens, of the pathobiome resulting in the disease development and/or aggravation. Network of interactions and LDA analysis allows the identification of likely positive interactions between the pathogen and members of the pathobiome that are consistent in the two growing-seasons. Culture-dependent methods and *in planta* studies have been performed to support the *in silico* analysis. In summary, this study highlights that the plant microbiome shows marked changes in its composition and structure during the foot rot disease and that this

disease can be viewed in part as the result of the interactions between a primary plant pathogen with members of the pathobiome.

2.2 Material and methods

2.2.1 Rice samples collection and treatment for the microbiome/pathobiome analysis

Rice plants during the early booting phase (16 weeks-old) were collected in two different rice growing seasons (2017 and 2018) belonging to two different rice cultivars grown in two rice fields owned by the Italian rice growers organization called SAPISE located in Vercelli. The plants were sampled at the same growth stage, during the two years and from the same rice fields, in order to decrease as much as possible the variability in the microbial community due to the environmental factors. Two rice cultivars were sampled for this study: Barone CL, characterized by long type seeds, very high field yield potential and high resistance to diseases from the rice field (45°26’33.2”N 8°21’50.0”E) and Sole CL, characterized by round type seeds, high adaptability to different environments and medium resistance to diseases from the rice field (45°30’56.7”N 8°22’22.0”E). From each field, symptomatic and asymptomatic rice plants were collected. The symptomatic plants were selected according to the visible presence of the typical foot rot symptoms (2-3 disease severity index), while the asymptomatic plants were selected randomly and as much as possible close to the symptomatic ones.

The plant material was washed and for each infected sample, close to the crown roots, 3-5 cm region of the brown rot stem has been selected and weighed (3 grams of plant material per sample). The same region of the healthy plant stem has been cut and treated in the same way. For each sample, 2 grams of plant material were macerated in liquid nitrogen using sterilized pestle and mortar and used later for bacterial DNA extraction and 16S rRNA gene

library preparation. The remaining part (1 gram) was resuspended in 4 mL PBS (Phosphate Buffered Saline) solution and stored with 18% glycerol at -80ºC for the culturable analyses.

2.2.2 DNA extraction from plant material, 16S rRNA gene amplicon library preparation and Illumina MiSeq sequencing

DNA was extracted from all plant samples using the PowerMax Soil DNA isolation kit (MO BIO Laboratories, Carlsbad, CA, USA) according to the manufacturer's instructions and using as starting material 0.25 gram from each sample. Total DNA concentration was analyzed by using a Nanodrop™ spectrophotometer (Thermofisher Scientific Inc., Waltham, MA, USA) and normalized to a concentration of 7 ng/uL for preparing the 16S rRNA gene amplicon library. The DNA extracted was used to amplify the V3 and V4 hypervariable region of the 16S rRNA gene using barcoded primers and PCR conditions following Illumina Inc.'s protocol (Illumina Inc., San Diego, CA, USA). Briefly, individual barcoded libraries were directly prepared by PCR using long primers (Klindworth et al. 2013) incorporating the Illumina adapter sequences (16S_Amplicon_PCR_Fw: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG; 16S_Amplicon_PCR_Rv:GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC). Following the first amplification, a cleaning step was performed using the AMPure XP bead clean-up (A63880l; Beckman Coulter Inc., Brea, CA, USA). A second PCR reaction was then performed to attach dual index and Illumina sequencing adapters using the Nextera XT Index Kit; followed by a final AMPure XP bead clean-up. Amplicons size, integrity, and purity were checked using the Bioanalyzer equipment (Agilent Inc., Santa Clara, CA, USA) and the library concentration was measured by fluorimetric quantification using Qubit 2 (Invitrogen

Inc., Carlsbad, CA, USA). Finally, library sequencing was performed using the Illumina Miseq technology using a 250 bp paired-end strategy.

2.2.3 Sequence data processing and statistical analysis

FASTQ files were imported into Qiime2 v2020.11 (Bolyen et al., 2019), quality filtering and OTU picking was done using DADA2 v1.18.0 (Callahan et al., 2016). OTUs consisted of groups of identical sequences. Each OTU was represented by a single sequence, named representative sequence. Representative sequences were aligned using mafft, and a phylogenetic tree was built using fasttree (Price et al., 2009; Katoh and Standley, 2014). Taxonomic assignment was based on the Silva database (release 138) (Quast et al., 2012) using an *ad hoc* classifier trained on the region amplified by the primer pairs used in the present study. After the removal of the OTUs annotated as chloroplasts and mitochondria, a rarefaction analysis using 600 reads per sample has been performed, to have a homogeneous sampling depth; on this dataset, the alpha and beta diversity were calculated. This value was selected after observing the alpha rarefaction plots and witnessing that, at this sampling depth, both the Shannon diversity values and the number of OTUs approached the plateau for all samples. After rarefaction, 59 out of 73 samples were retained (the 80.82%). Despite the loss of the 20% of the samples, the experimental design was still balanced in terms of variety (23 samples retained from Barone and 32 from Sole), symptomatology (16 healthy samples and 39 sick ones), and year (11 samples from year 2017 and 44 from 2018). Significant differences in alpha and beta diversity among categories of samples were assessed using the Kruskal-Wallis (performed on each category) and PERMANOVA tests, respectively. Subsequent statistical analyses and data visualizations were performed using the phyloseq package (McMurdie and Holmes, 2013) in R version 4.0.2 (R core team, 2014).

A Venn diagram was drawn to visualize unique and shared genera in sick and healthy plants, a Venn diagram was generated using the web tool http://bioinformatics.psb.ugent.be/webtools/Venn/.

To calculate the species co-occurrence network, two separate OTU tables were exported, one for each symptomatology category. Correlation values among the OTUs were calculated using fastspar (Watts et al., 2019), an implementation of the SparCC (Sparse Correlations for Compositional data) algorithm (Friedman and Alm, 2012). All the correlation absolute values below 0.3 were filtered out and the table was imported in cytoscape for visualizations and analyses (Shannon et al., 2003).

To identify differential abundances of bacterial taxa between the two conditions tested, a LefSe analysis has been performed, as implemented in the Galaxy server (Segata et al., 2011) (http://huttenhower.org/galaxy). In this analysis differences in the relative abundance of taxa between healthy and diseased conditions were calculated with the Kruskal-Wallis test and the trend identified by the Kruskal-Wallis test was then checked by the Wilcoxon test. For the comparison, symptomatology condition was used as a class of interest, with the Wilcoxon test alpha value set at 0.05, the alpha value of the Kruskal-Wallis test set at 0.05 and the threshold for the LDA analysis score was set at 2.0.

2.2.4 Culturable microbiome analysis

The macerated plant tissue of each independent plant sample (5 samples from the first sampling and 10 from the second sampling) displaying foot-rot symptoms and stored at -80ºC was mixed and diluted in 20 mL of PBS as described above. Serial dilutions were performed and 100 µl of each dilution was spread and grown on four different solid media (TSA, PDA,

M9 and 869-medium) (**Table 2.1**) and the plates were incubated at room temperature (RT) for 2 and 5 days. After 2 days the totality of the colonies grown were harvested with 2 mL of PBS and the same has been done after 5 days of incubation. From the bacterial suspension obtained, the genomic DNA was extracted using the Bacterial genomic DNA isolation kit (Norgen biotech corp., Thorold, Canada) following the manufacturer's instructions. The DNA was used then to amplify the V3 and V4 hypervariable region of the 16S rRNA gene using barcoded primers and PCR conditions following Illumina Inc.'s protocol (Illumina Inc., San Diego, CA, USA) as described above. Amplicon sequencing of the 16S rRNA gene was performed by using the Illumina MiSeq platform with v3-v4 chemistry and 250 paired-end reads on the MiSeq instrument (Illumina) according to the manufacturer's instructions. Sequencing reads were processed and analyzed as described above.

Table 2. 1 Composition of the bacterial growth media used for the culturable pathobiome fraction selection.

Medium	Medium ingredients for 1L	Reference or source
Tryptic soy (ts) broth/agar	15g Pancreatic Digest of Casein; 5g Peptic Digest of Soybean Meal; 5g Sodium Chloride	[1]
M9	Na_2HPO_4 60.5 gr, KH_2PO_4 12 gr, NaCl 2 gr, NH_4Cl 4 gr, 1M $MgSO_4$, 1M $CaCl_2$, Gluocose 0,1%	[4]
PDA	Potato infusion 200 gr, Dextrose 20 gr, agar 20 gr	Difco, BD Laboratories, MD 21152 USA
1/10 869	$CaCl_2$ 0.035 gr, NaCl 0.500 gr, Tryptone 1 gr, Yeast extract 0,500 gr, Glucose D+ 0,100 gr, agar 15 gr	[5]

2.2.5 Isolation of culturable bacteria from rice plants

The same macerated plant tissue from symptomatic samples stored with glycerol at -80ºC and used for the culturable microbiome analysis (see above) was used for the isolation of pure bacterial colonies by plating different dilutions on 1/10 Tryptic Soy Agar (TSA; Difco, BD Laboratories, MD 21152 USA) solid medium. Plates were incubated at RT for 2 and 5 days and pure independent colonies showing distinct colony morphology were picked and streaked again on 1/10 TSA plates to ensure the purity of the culture and then stored at -80ºC in 1/10 TSA and 18% glycerol.

Amplification of the 16S rRNA gene was then performed on the pure colonies isolated by using fD1Funi 16S (5'-AGAGTTTGATCCTGGCTCAG-3') and rP2Runi 16S (5'-ACGGCTACCTTGTTAGGACTT-3') primers to amplify the complete 16S rRNA gene and gain a more precise taxonomic information of the bacterial isolates. Colony PCR was performed after boiling (10' at 98°C) a colony suspension in 50 µl of sterile H_2O. PCR products were purified by using Gel extraction and PCR Clean-Up System purification kit (Euroclone S.p.A). The sequencing performed with primers 907R (5'-CCGTCAATTCMTTTRAGTTT-3') and 785F (5'- GGATTAGATACCCTGGTA-3') was realized by GATC (Eurofins Genomics Company, Germany) and identification of the isolates was obtained by BLAST analysis at NCBI (http://www.ncbi.nlm.nih.gov).

2.2.6 In planta test of a selected bacterial isolate to study its possible cooperation with D. zeae

In order to investigate the effect of a possible partner of *D. zeae* isolated from the foot-rot pathobiome in the development of the disease, co-infection tests *in planta* were performed.

Ampicillin 100, Gentamicin 50 and Nitrofurantoin 100 were used for *Burkholderia* isolation and growth. A spontaneous *D. zeae* rifampicin-resistant was isolated by growing the strain in 1/6 TSB (Tryptic Soy Broth) medium supplemented with gradually increasing amounts of rifampicin (Rif) ranging from 15 to 100 µg ml^{-1}. Finally, the culture was plated on TSA (Tryptic Soy Agar) and single colonies were re-inoculated in TSB containing Rif 100 µgml^{-1}, stored at -80°C.

The effect, on the virulence degree, of the pathogen and its partner presence on 20-day rice seedlings (8 replicate) was tested as follows: seeds were sterilized in 50% commercial bleaching agent (3,62% w/v NaOCl) for 1 hr and, after abundant washes with sterile water, were incubated on sterile water-soaked Whatman paper in the dark at 30 °C for 7 days. Plantlets were transferred independently to a tube containing Hoagland´s semi-solid solution (Steindler et al., 2009) and grown for the other 15 days at 27 °C, 80% humidity, 16 h/8 h light-dark cycle. Inoculation was performed by injuring the stem with needles contaminated with *Burkholderia* spp., grown in 1/10 TSB at OD_{600} of 0.1. The following day, an inoculation in the same position was performed with *D. zeae* Rif^R, grown in 1/10 TSB at OD_{600} of 0.1. For each treatment, four biological replicates were performed and as control, plants were inoculated with single strains (only *D. zeae* or *Burkholderia*) at the same OD_{600} and PBS. Lesion appearance and development was followed for 1 week. Bacteria were re-isolated from the dark rotten lesion in the stem of the plant by macerating it in PBS solution and serial dilutions plated in TSA containing the appropriate antibiotics followed by incubation at 30 °C for 24 hours and counted for CFU/g calculation.

2.3 Results

2.3.1 Rice sampling for pathobiome and microbiome 16S rRNA gene community sequencing

In order to unveil how the plant microbiome can be affected by pathogen establishment, pathobiome analysis in symptomatic rice plants and microbiome analysis in asymptomatic plants was performed and compared. Rice plants (16 weeks-old) were sampled in two growing seasons (2017 and 2018) from growing fields of two rice cultivars (Barone CL and Sole CL). From each of the two fields, cv. Barone asymptomatic and symptomatic plants and similarly cv. Sole asymptomatic and symptomatic plants were collected. The symptomatic plants were selected according to the visible presence of the typical foot-rot symptoms, characterized by yellow and dry leaves, black rot, and foul-smelling base and roots. The disease severity between all the rice plants exhibiting foot-rot symptoms was comparable (2-3 disease severity index) in each sampling year. On the other hand, the asymptomatic plants were collected according to the absence of any visible infection signs and as close as possible to the symptomatic plants, in order to decrease the microbial variability due to the soil and other environmental factors (**Figure 2.1**).

Figure 2. 1 Dickeya zeae rice foot rot symptoms (A) and comparison with asymptomatic samples (B). All the symptomatic samples collected showed a comparable disease severity (2-3 severity index).

In the first sampling (July 2017), 10 asymptomatic (5 Barone and 5 Sole) and 10 (5 Barone and 5 Sole) rice plants showing foot-rot symptoms were collected. The second round of sampling (July 2018) has been performed in the same fields of the first sampling, increasing the number of samples to 28 (14 Barone and 14 Sole) asymptomatic plants and 28 (14 Barone and 14 Sole) foot-rot symptomatic plants. During the second sampling, the number of plant samples was increased since the incidence of the disease was higher.

DNA was purified from infected stems 3-5 cm around the symptoms lesion site, close to the crown roots and from the same plant zone of healthy plants, in order to reduced bacterial community variability due to the plant compartment. Bacterial communities were studied via 16S rRNA amplicon sequencing.

2.3.2 Bacterial community compositional shifts between asymptomatic and symptomatic rice plants

The number of reads passing the quality filter was on average 4082,06 (ranging 70-18557). The average number of OTUs was 66,69 (ranging 15-168). The reads annotated as *Chloroplast* were on average 47,1% ± 29,5 (**Supplementary Table S2.1**) and were removed from the

dataset. The number of OTUs was higher in samples of 2018 ($p < 0,01$) compared to the samples of 2017 and in symptomatic samples compared to the asymptomatic ones ($p < 0,001$) (**Table 2.2**).

Table 2. 2 Average number of OTUs subset by Cultivar, Symptom and Year.

Average number of different OTUs subset by parameters (Cultivar, Symptom and Year) and split in Classes. OTU numbers differed significantly with the symptomatology (Healthy vs Sick) and the Year (2017 vs 2018). Significance between classes was calculated by Wilcoxon-Mann-Whitney test.

Parameter	Classes	Avg. N° OTUs	St.dev.	p-val*
Cultivar	Sole	65,15	44,15	N.S.
	Barone	68,82	34,79	
Symptom	Healthy	35,87	23,07	<0.001
	Sick	79,33	39,01	
Year	2017	33,81	15,65	<0.01
	2018	74,9	40,32	

*after Wilcoxon-Mann-Whitney test

The Shannon diversity values were higher in symptomatic samples compared to the asymptomatic ones in 2018 ($p < 0,001$), while no significant differences were detected in 2017. (**Figure 2.2 A and B**). The PCoA based on the weighted UniFrac distance measure showed that symptomatic and asymptomatic samples formed two distinct clusters (**Figure 2.2 C**). Asymptomatic and symptomatic samples from the year 2018 were significantly ($p = 0.001$) different and clustering separately, while the differences in 2017 were not evident ($p = 0.1$). Asymptomatic and symptomatic samples from the year 2017 were partially overlapping, suggesting that their microbiomes were more similar than the ones from the year 2018.

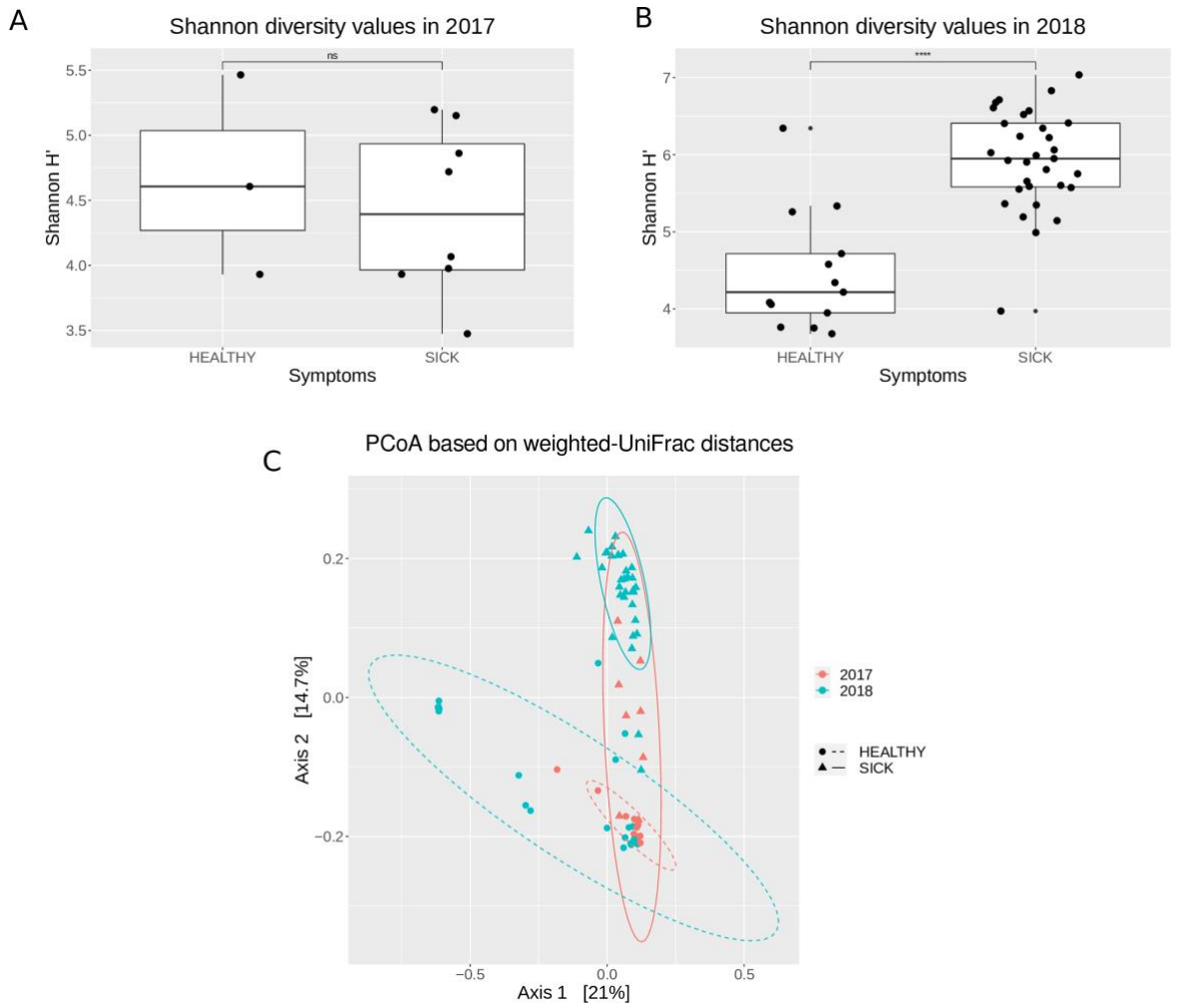
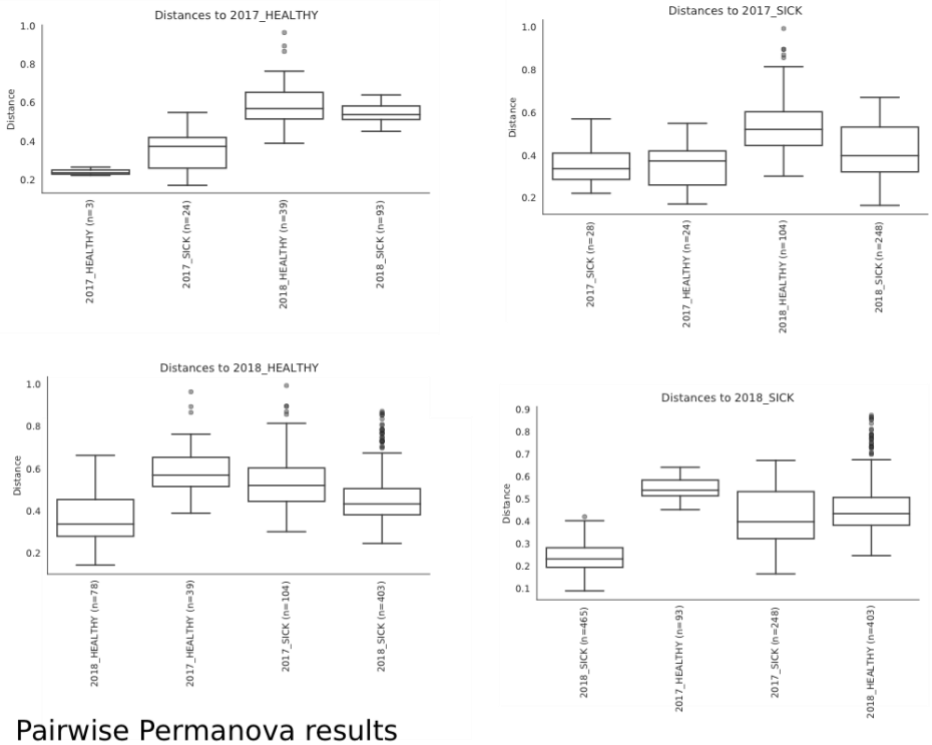


Figure 2. 2 Complexity and composition of the microbial communities of asymptomatic and symptomatic samples. A) Alpha diversity estimations of the bacterial communities associated with asymptomatic (healthy) and symptomatic (sick) samples, using Shannon's H index for year 2017 and **B)** for year 2018; box plot depict medians (central horizontal lines), the inter-quartile ranges (boxes), 95% confidence intervals (whiskers) and outliers (black dots). Asterisk indicate significant differences between two groups of samples (*P<0.05). Statistical analyses were calculated based on wilcoxon-test. **C)** Principal Coordinate Analysis (PCoA) based on weighted UniFrac distances. Ellipses show confidence Intervals (CI) of 95% for each sample type. Statistical significance has been inferred using PERMANOVA (see **Figure 2.3** and **2.4**).

To infer significant differences among the asymptomatic and symptomatic samples considering also the effect of the two years, a PERMANOVA test on the UniFrac distances within each group compared with the ones among different groups was performed (with 999 permutations in all tests). The samples were divided in i) healthy samples in the year 2017; ii) sick samples in the year 2017; iii) healthy samples in the year 2018; and iv) sick samples in the year 2018. Significant differences were detected for the following combinations: healthy samples in 2017 Vs sick samples in 2018 (pseudo-F = 22,24, p = 0,001); healthy samples in 2018 Vs sick samples in 2017 (pseudo-F = 9,27, p =0,001) and as mentioned above, healthy samples in 2018 Vs sick samples in 2018 (pseudo-F = 24,67, p = 0.001,**Figure 2.3**).

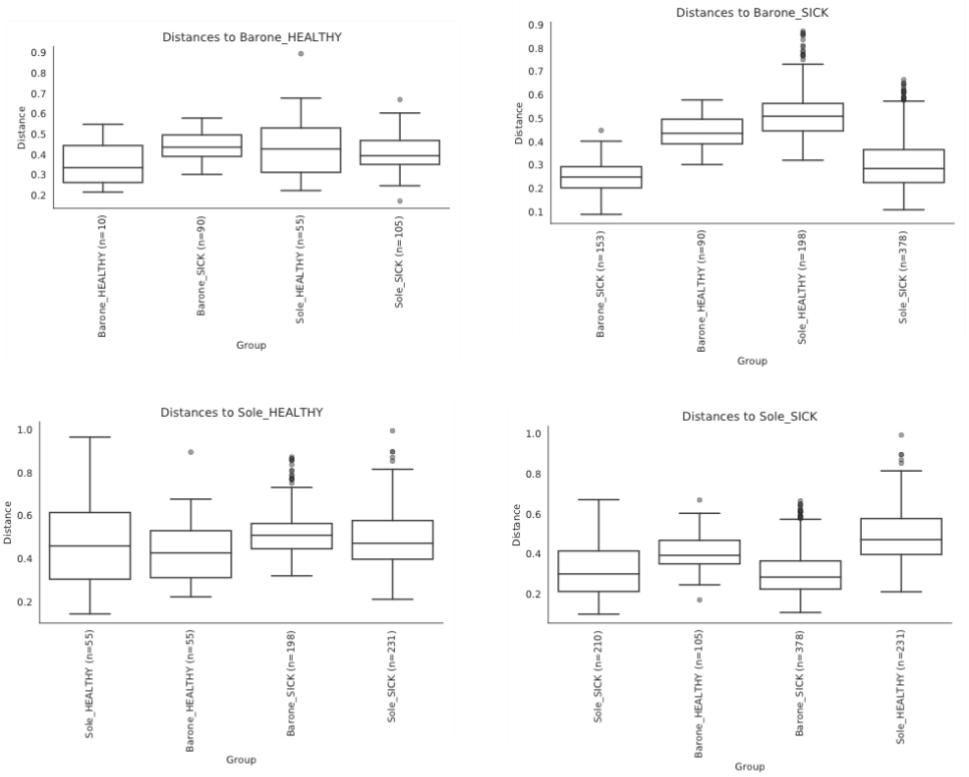


Pairwise Permanova results

Group 1	Group 2	Sample size	Permutations	pseudo-F	p-value	q-value
2017_HEALTHY	2017_SICK	11	999	2.104374	0.115	0.1150
	2018_HEALTHY	16	999	9.273814	0.002	0.0024
	2018_SICK	34	999	22.644662	0.001	0.0015
2017_SICK	2018_HEALTHY	21	999	11.991601	0.001	0.0015
	2018_SICK	39	999	17.668034	0.001	0.0015
2018_HEALTHY	2018_SICK	44	999	24.671807	0.001	0.0015

Figure 2. 3 PERMANOVA test on the UniFrac distances within each group compared with the ones among different groups (with 999 permutations in all tests). The distance is calculated especially considering the variable of the year.

The rice cultivar had no influence on the microbiome composition (pseudo-F = 3,06, p = 0,009). Specifically, the PERMANOVA test performed on the beta diversity between the two cultivars and symptomatology revealed that the microbiome changed more between the conditions of health and disease rather than between the two cultivars (**Figure 2.4**).



Group 1	Group 2	Sample size	Permutations	pseudo-F	p-value	q-value
Barone_HEALTHY	Barone_SICK	23	999	10.827807	0.001	0.002
	Sole_HEALTHY	16	999	1.128690	0.325	0.325
	Sole_SICK	26	999	3.374515	0.020	0.024
Barone_SICK	Sole_HEALTHY	29	999	14.055871	0.001	0.002
	Sole_SICK	39	999	3.428520	0.010	0.015
Sole_HEALTHY	Sole_SICK	32	999	7.104305	0.001	0.002

Figure 2. 4 PERMANOVA test on the UniFrac distances within each group compared with the ones among different groups (with 999 permutations in all tests). The distance is calculated especially considering the variable of the rice cultivar.

The OTUs were annotated using the Silva database (Quast et al., 2012); in total, 20 phyla, 36 classes, 95 orders, 158 families, and 278 genera were identified. The microbiomes were overall dominated by *Proteobacteria* of the alpha- and gamma-class, and the *Bacteroidetes* and *Negativicutes* (**Figure 2.5 A**, **Table 2.3**).

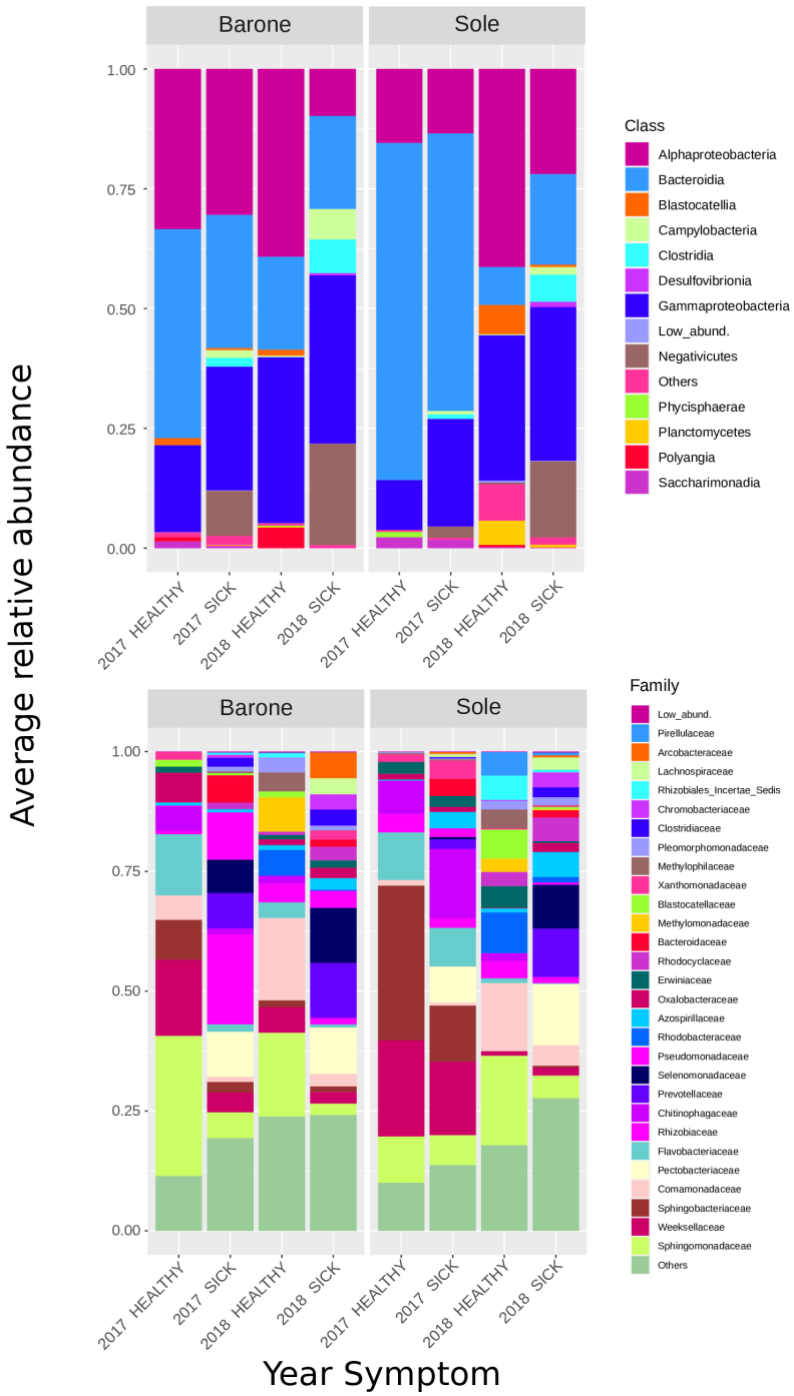


Figure 2. 5 Average relative abundance (% of sequencing reads) of predominant OTUs. Stacked bar-charts showing the relative abundance at the level of A) Class; and B) Family, divided according to the symptoms and year. Taxa whose abundance was <1% have been lumped into "Low abundance" category; taxa that did not classify at that specific taxonomic level were lumped into the category "Others".

The phylum *Firmicutes* and *Campylobacteria* were found at higher average abundance in symptomatic samples (20,82% versus 0,67% in healthy, and 3,33% versus 0,17% in healthy, respectively). On the other hand, the phylum *Acidobacteria* was detected at significant higher average abundance in asymptomatic samples (5,25% versus 0,47% in infected plants) (**Table 2.3**). At the family level the most prevalent taxa were *Pectobacteriaceae, Sphingomonadaceae*, *Comamonadaceae, Azospirillaceae* and *Arcobacteraceae* among the *Proteobacteria*. *Firmicutes* were dominated by *Clostridiaceae* and *Lachnospiraceae*. *Bacteroidetes* were mainly represented by the family *Prevotellaceae*, and *Acidobacteria* were mainly composed of OTUs assigned to the *Blastocatellaceae* family. All families presented distinct differences among asymptomatic and symptomatic microbiomes (**Fig. 2.5 B**, **Table 2.3**).

Table 2. 3 **Predominant OTUs abundance subset by Phyla, Class and Genera**

Average relative abundance of predominant OTUs subset by Phyla, Class and Genera. Significative differences between healthy and sick condition are calculated by t-test and summarized in the table.

	Taxon	Healthy	Sick	p-value
Phyla	Proteobacteria	60.76	47.49	0.005
	Firmicutes	0.67	20.82	<0.001
	Acidobacteriota	5.25	0.47	<0.001
	Campylobacterota	0.17	3.33	0.003
	Bacteroidota	26.85	25.54	n.s.
	Actinobacteriota	2.53	0.26	<0.001
	Planctomycetota	0.03	0.14	n.s.
	Cyanobacteria	0.38	0.15	n.s.
Class	Alphaproteobacteria	34.65	16.58	<0.001
	Clostridia	0.02	5.20	<0.001
	Campylobacteria	0.17	3.33	0.003
	Blastocatellia	3.04	0.23	<0.001
	Bacteroidia	26.85	25.51	n.s.
	Gammaproteobacteria	26.10	31.35	n.s.
	Negativicutes	0.13	15.43	<0.001
	Acidobacteriae	2.18	0.19	0.001
	Desulfovibrionia	0.005	0.62	<0.001

	Bacilli	0.52	0.17	n.s.
	Planctomycetes	2.27	0.25	0.001
	Paludibaculum	0.00	8.65	<0.001
	Candidatus_Koribacter	5.02	1.67	0.016
	Bryobacter	0.13	5.08	<0.001
	Aridibacter	5.29	0.94	0.007
	Holophaga	2.90	2.00	n.s.
	Vicinamibacter	3.88	1.18	n.s.
	Kineococcus	4.79	0.49	<0.001
	Kineosporia	0.00	2.89	<0.001
	Curtobacterium	0.00	2.27	0.02
Genera	Nocardioides	0.51	1.85	0.001
	Bacteroides	1.52	0.79	n.s.
	Dysgonomonas	2.36	0.12	0.002
	Microbacter	2.40	0.05	<0.001
	Paludibacter	0.00	1.83	<0.001
	Prevotella	1.93	0.31	n.s.
	Aurantisolomonas	1.81	0.04	<0.001
	Chitinophaga	0.00	1.39	<0.001
	Dinghuibacter	1.66	0.12	<0.001
	Edaphobaculum	0.02	1.30	<0.001
	Ferruginibacter	1.46	0.14	0.01
	Filimonas	1.36	0.20	n.s.

The number of shared and unique genera between asymptomatic and symptomatic samples is shown in the Venn diagram (**Figure 2.6**). 158 genera were found only in the bacterial communities of symptomatic samples, whereas 40 genera were found exclusively in asymptomatic samples only, while 140 genera were shared (i.e. found in both conditions).

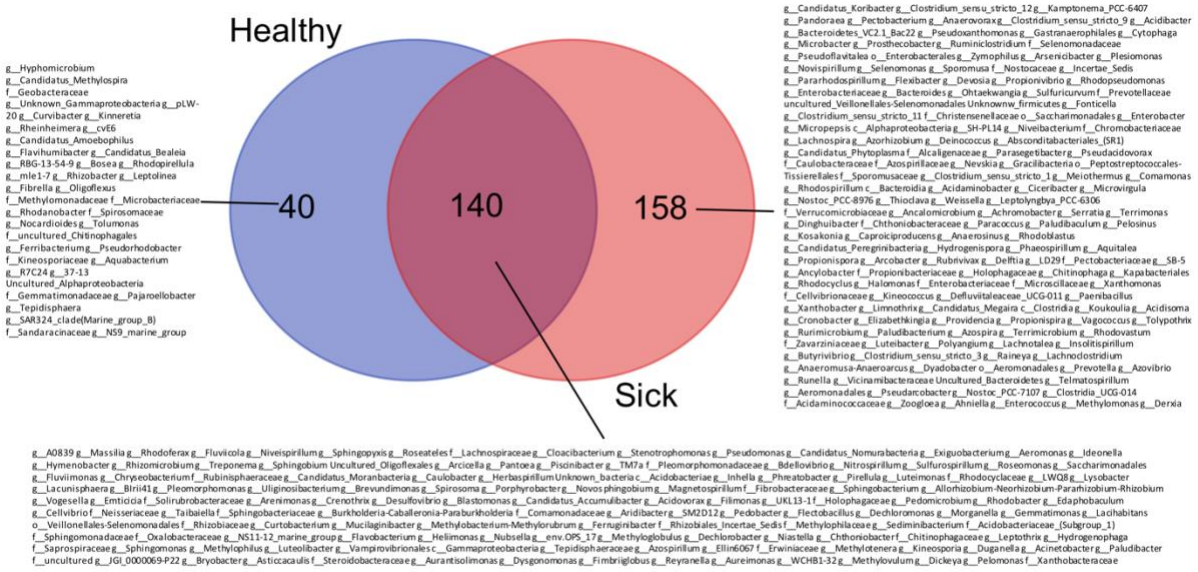
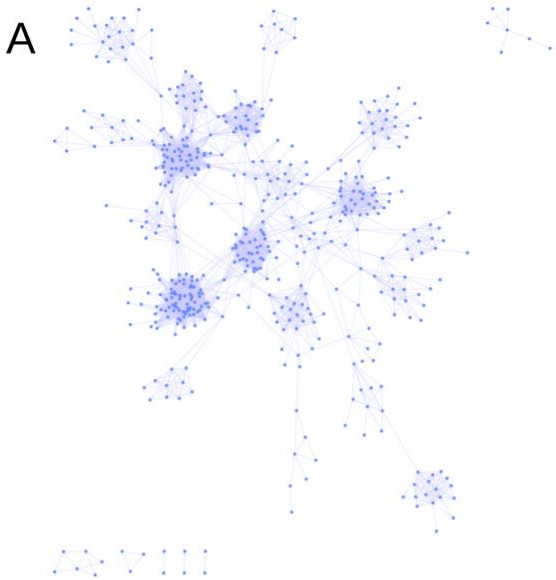
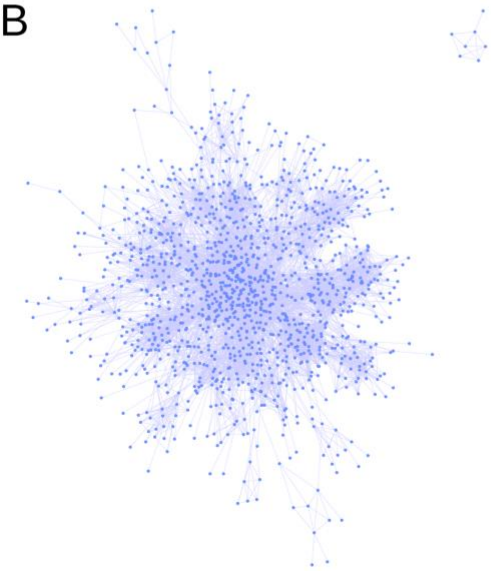
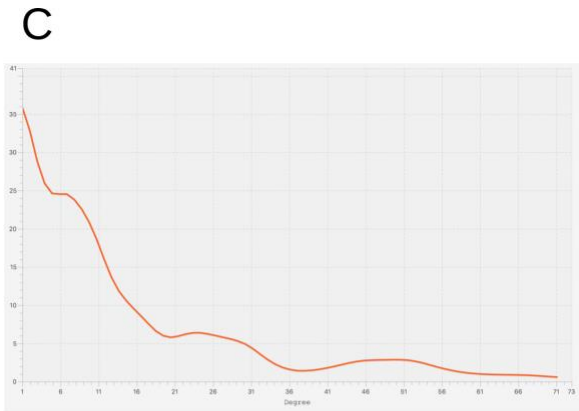


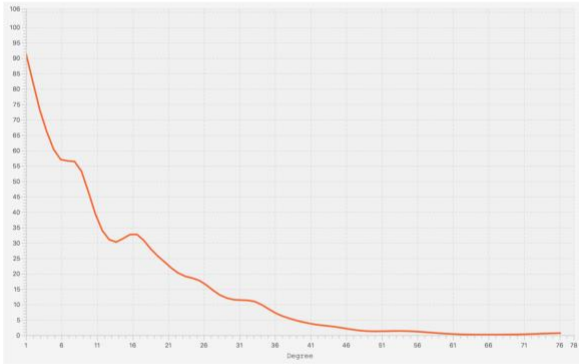
Figure 2. 6 Venn diagram displaying the number of unique and shared OTUs between asymptomatic (healthy) samples and symptomatic (sick) samples along with the name of the taxa unique or shared between the two conditions studied.

The network analysis, based on the SparCC correlation values, revealed that the networks in asymptomatic and symptomatic samples were different. The network deriving from asymptomatic samples microbiome featured a more distinct clustering pattern, a lower number of nodes and edges, and a lower average degree compared to the network derived from symptomatic samples (**Figure 2.7**). The two networks had a comparable number of average neighbors per node, but the network derived from the asymptomatic samples had higher clustering coefficient and density, whereas the network of the symptomatic samples was sparser.









E

	Healthy	Sick
N. Nodes	517	1271
N. Edges	4916	8909
Avg. n. Neighbors	16232	14019
Diameter	10	13
Radius	1	1
Path length	2621	3945
Clustering Coeff.	0.336	0.244
Density	0.016	0.006

Figure 2. 7 Network analysis of microbiomes and pathobiomes. A) OTUs interactions network based on the SparCC values (nodes are the OTUs and the links were established when the SparCC absolute correlation value was above 0.3) among the OTUs in asymptomatic microbiomes; and **B)** in symptomatic microbiomes. **C)** and **D)** Nodes degree distributions for asymptomatic and symptomatic microbiomes, respectively. **E)** Main topological statistics on the two networks.

2.3.3 Dickeya zeae is present only in foot rot symptomatic rice plants

The average relative abundance of *Dickeya* genera reads for symptomatic plants was 10,53% (ranging 0,0 – 31,42% **Figure 2.9 A**). Reads annotated as *Dickeya* were detected only in one asymptomatic plant at the abundance of 0,39%. To further confirm the presence of *D. zeae*, the virulence *zsmA* gene, which is unique to *D. zeae* (Zhou et al., 2011), was amplified using as template purified DNA from plant material. The *zsmA* gene was only successfully amplified from symptomatic samples, demonstrating again the presence of *D. zeae* and its involvement in foot-rot disease (**Figure 2.8**).

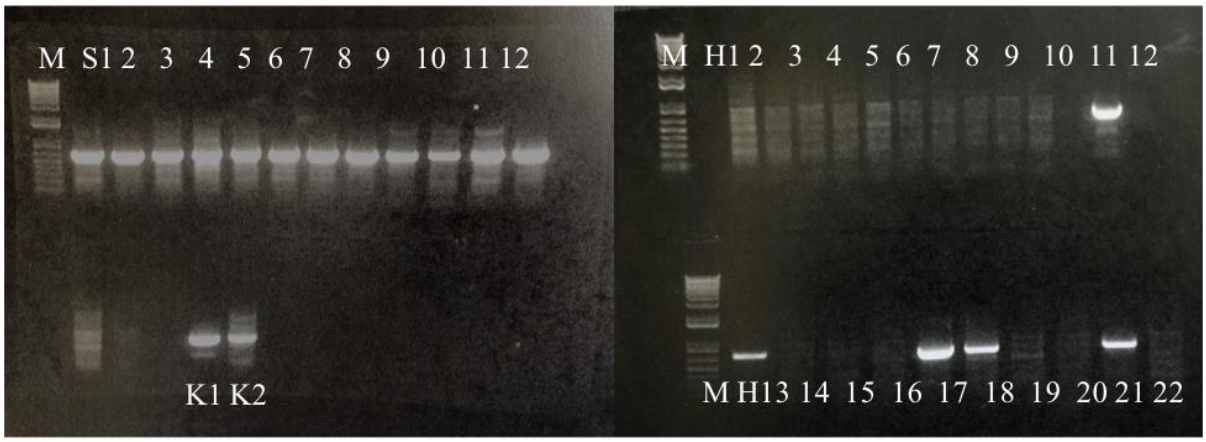


Figure 2. 8 Conventional PCR results of *zsmA* gene amplification; M: 1Kb marker; S1-S12: symptomatic samples from the 2018 sampling; K1: positive control Dz2Q genome DNA (Bertani et al., 2013); K2 positive control SS8 sample from 2017 sampling; H1-H22: asymptomatic samples from 2018 sampling. Samples H13-17-18-21 were removed from the analysis. The primers used for the *zsmA* gene amplification are : ZmsA3 (5'-CCGGCCTCAAGGATGGAGG-3') and ZmsA4 (5'-AGGCGTGGTGGTAGCAGTGAC-3').

In order to explore the possible interactions between the pathogen and the other members of the pathobiome, a co-occurrence network analysis was performed; this analysis revealed 27 significant correlations ($r \geq 0,5$ for co-occurrence and $r \leq -0,2$ for co-exclusion); between *Dickeya* spp. and different OTUs classified at a various taxonomic rank (**Figure 2.9 B).** The strongest positive correlations identified were with OTUs belonging to the family

Selenomonadaceae, genus *Prevotella, Propionispira, Clostridium_sensu_stricto_1* and genus *Azospirillum*. On the other hand, the strongest negative correlations observed were with OTUs belonging to the genus *Crenothrix*, *Rhizobacter,* and *Saccharimonadales* **Figure 2.9 B)**.

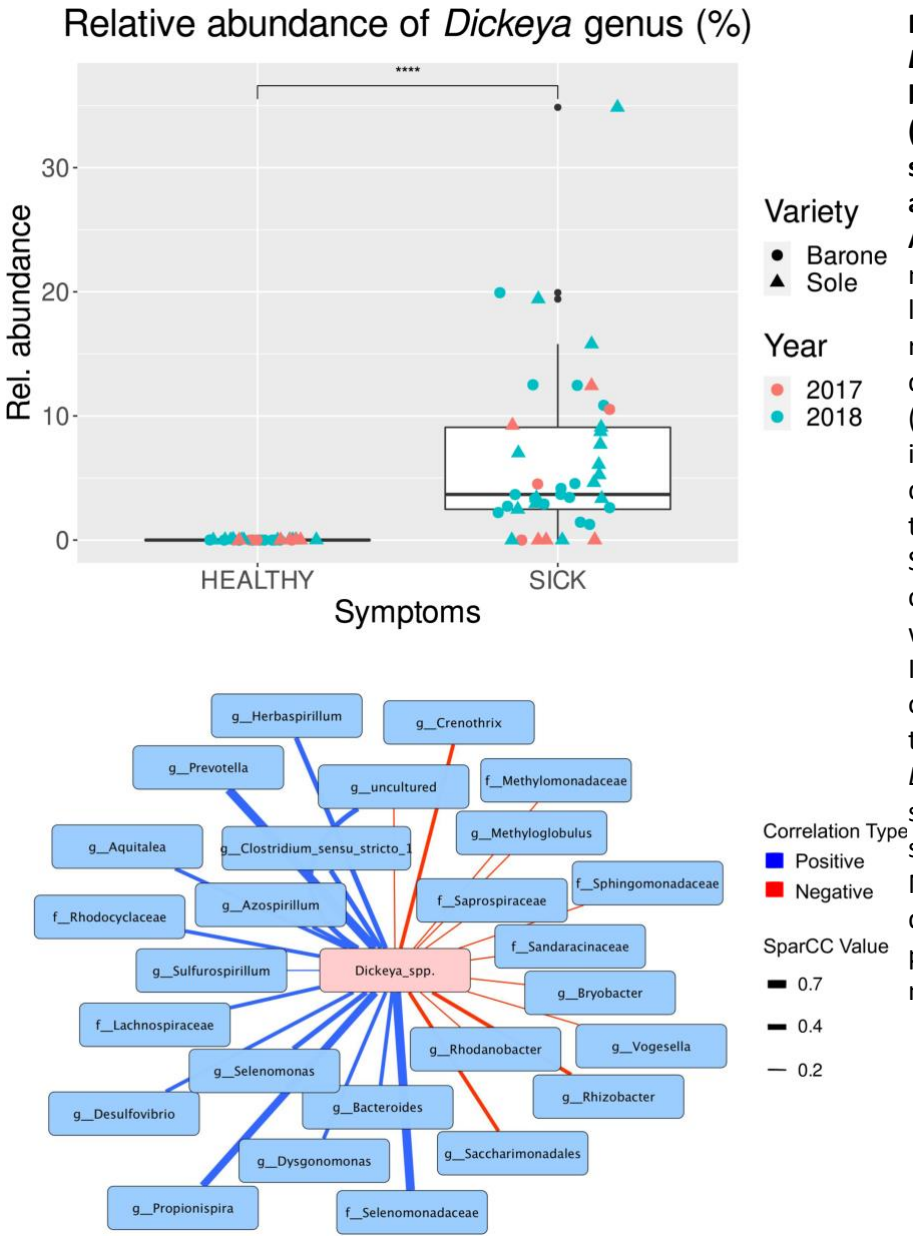


Figure 2. 9 Comparison of *Dickeya* genus abundance between asymptomatic (HEALTHY) and symptomatic (SICK) plants, and interactions network. A) Box plot depicting medians (central horizontal lines), the inter-quartile ranges (boxes), 95% confidence intervals (whiskers). Asterisks indicate significant differences between the two groups of samples. Statistical analyses were calculated based on wilcoxon-test. **B)** Interactions network based on the SparCC values among the OTUs classified as *Dickeya* genus and the most significant taxa showing stronger correlation values. Negative correlations are depicted as red edges, while positive correlations are marked as blue edges.

2.3.4 Several bacterial taxa significantly correlate with D. zeae presence in symptomatic plants

A linear discriminant analysis (LDA) between the bacterial community composition of symptomatic samples (pathobiome) and asymptomatic samples (microbiome) was performed, in order to identify potential biomarker taxa between the two conditions. A graphical representation of the differentially represented-taxa (*p* value < 0,05) from the two growing seasons is shown in **Figure 2.10.** Among the significantly differentially represented-taxa, 76 were more abundant in symptomatic samples and 16 in the asymptomatic ones. The genera *Prevotella*, *Propionispira*, *Clostridium_sensu_stricto_1* and *Azospirillum*, identified in the network analysis as positively interacting with *Dickeya* spp. (**Figure 2.9 B**), were also found to be statistically correlated with the disease condition (**Figure 2.10**). On the other hand, *Crenothrix* and *Rhizobacter*, negatively interacting with *Dickeya* spp. (**Figure 2.9 B**), were statistically correlated with the asymptomatic condition (**Figure 2.10**). Notably, several taxa such as *Pseudomonas*, *Azospirillum*, *Burkholderia*, *Sulfurospirillum, Pleomorphomonas* and *Magnetospirillum* were present in both asymptomatic and symptomatic samples but were significantly enriched in the samples affected by the disease. Some taxa were only present in the symptomatic samples, in particular *Prevotella*, *Propionispira*, *Selenomonas*, *Clostridium_sensu_stricto_1, and Clostridium_sensu_stricto_12,* suggesting that their establishment was significantly stimulated/influenced by the presence of *D. zeae*.

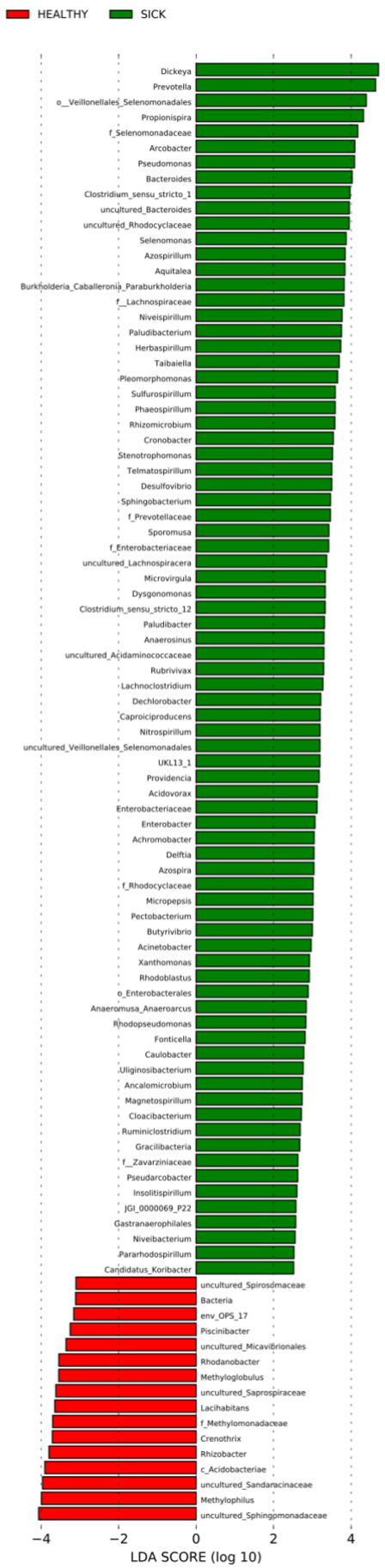


Figure 2. 10 Linear discriminant analysis (LDA) scores of the main bacterial genera which significantly differ in their relative abundance between asymptomatic (HEALTHY) and symptomatic (SICK) samples.

Differential abundance of bacterial genera between two groups of samples tested, was assessed by performing a linear discriminant analysis (LDA). The two symptomatology conditions were used as class to test. Only genera with a *p* value < 0.05 for the Kruskal-Wallis test and an LDA score >2 are displayed.

2.3.5 The culturable fraction of bacteria from rice foot-rot symptomatic plants

To begin to investigate whether some members of the pathobiome are cooperating with the pathogen in the disease process, it was important to isolate some bacterial strains that belong to taxa which were significantly enriched by the presence of *D. zeae* (see above). Five representative samples of symptomatic plants from the 2017 and ten representative samples from the 2018, were used for performing a culturable fraction study via 16S rRNA amplicon sequencing. This allowed us to determine which bacterial genera/species were cultivable under laboratory conditions. Four different growth solid media (including both rich and minimal media) (**see Material and methods Table 2.1**) and two different time points were used for growing bacteria from the macerated plant samples; all the bacteria colonies grown were then collected *en masse*, resuspended and diluted and the total DNA purified as described in the Materials and Methods section. 16S rRNA gene community sequencing was then performed and results are summarized in **Table 2.4.** Among the 298 different taxa detected in the pathobiome (see above), the fraction of bacteria culturable under the conditions tested here was 13%.

Table 2. 4 List of genera detected via 16S rRNA amplicon sequencing deriving from the cultivable fraction of the pathobiome. Four different media and two-time points have been used for the isolation. The first column of the table shown the list of genera detected in the culturable pathobiome study via 16S rRNA gene amplicon sequencing. The ability of each culturable bacteria genus to grown in the media used in this experiment is defined by presence (✓) or absence (✕).

Culturable taxa	869- medium	M9-medium	PDA-medium	TSA-medium
Achromobacter	✓	✓	✕	✓
Acinetobacter	✓	✓	✓	✓
Acidovorax	✓	✓	✕	✓
Aeromonas	✓	✓	✓	✓
Azospirillum	✓	✓	✕	✓
Aquitalea	✓	✓	✕	✓
Burkholderia	✓	✓	✓	✓
Brevundimonas	✓	✓	✕	✓
Chryseobacterium	✓	✓	✓	✓
Citrobacter	✓	✓	✓	✓
Comamonas	✓	✓	✕	✓
Caulobacter	✓	✓	✓	✓
Delftia	✓	✓	✕	✓
Dickeya	✓	✓	✓	✓
Elizabethkingia	✓	✓	✓	✓
Enterobacter	✓	✓	✓	✓
Enterococcus	✓	✓	✓	✓
Exiguobacterium	✓	✓	✕	✓
Flavobacterium	✓	✓	✕	✓
Herbaspirillum	✓	✓	✕	✓
Klebsiella	✓	✓	✓	✓
Kosakonia	✓	✓	✓	✓
Microbacterium	✓	✓	✕	✓

Microvirgula	✓	✓	✕	✓
Morganella	✓	✓	✕	✓
Novosphingobium	✓	✓	✕	✓
Paenibacillus	✓	✓	✕	✓
Pandoraea	✓	✓	✕	✓
Pantoea	✓	✓	✓	✓
Providencia	✓	✓	✕	✓
Pseudomonas	✓	✓	✓	✓
Rhizobium	✓	✓	✕	✓
Salmonella	✓	✓	✓	✓
Serratia	✓	✓	✓	✓
Sphingobacterium	✓	✓	✕	✓
Sphingomonas	✓	✓	✕	✓
Stenotrophomonas	✓	✓	✓	✓
Vagococcus	✓	✓	✕	✓
Xanthomonas	✓	✓	✕	✓

From the same symptomatic samples selected for the culturable pathobiome fraction analysis, a bacterial collection of 100 pure strains was generated (**Supplementary Table S2.2**) and molecular characterized via 16S rRNA complete gene amplification. *In planta* virulence tests were then performed with one isolate, namely *Burkholderia* sp. A56; the reason was because the *Burkholderia* genus significantly correlated with the *D.zeae* presence (see above linear discriminant analysis, **Figure 2.10**) and it was considerably enriched in the pathobiome culturable fraction. The virulence test was performed using single inoculations as well as co-inoculations along with *D. zeae*. No significant differences in the size of the lesion caused by

D.zeae were observed between plants treated only by the pathogen and those co-inoculated by the pathogen and with the *Burkholderia* sp. isolate (**Figure 2.11**).

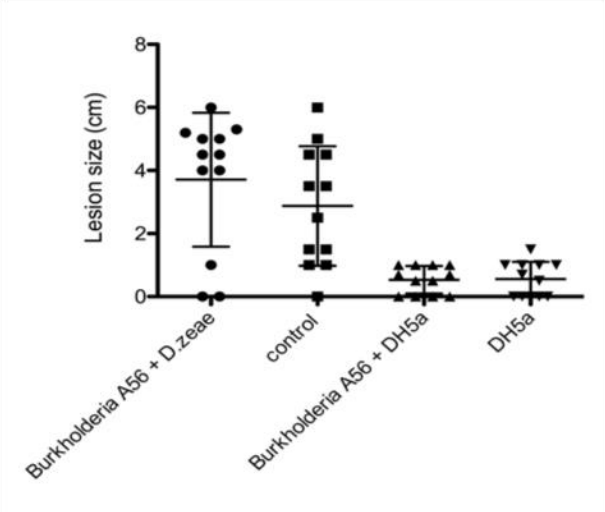


Figure 2. 11 Disease symptoms lesions size (cm) with single inoculation of *Dickeya zeae* (control) and co-inoculation of *Dickeya zeae* and *Burkholderia* sp. A56 .Each dot is an independent biological replicate; 12 plants were tested in each condition.

Bacterial CFU of *D. zeae* recovered from the infection site was on average of 1.2 x 10^7 and 3.2 x 10^7 CFU/g in both plants inoculated alone or co-inoculated, indicating that this strain does not affect the growth of *D. zeae in planta* experiment (**Figure 2.12**). A significant difference (*p*value < 0,001) however was observed in bacterial CFU of *Burkholderia* sp. A56 comparing the single inoculation (on average 6.3 x10^5 CFU/g) versus the co-inoculation with *D. zeae* (1.1 x 10^7 CFU/g) (**Figure 2.12**). This indicated that *Burkholderia* sp. A56 colonized the rice plant significantly more resulting in a higher bacterial load when *D. zeae* was present.

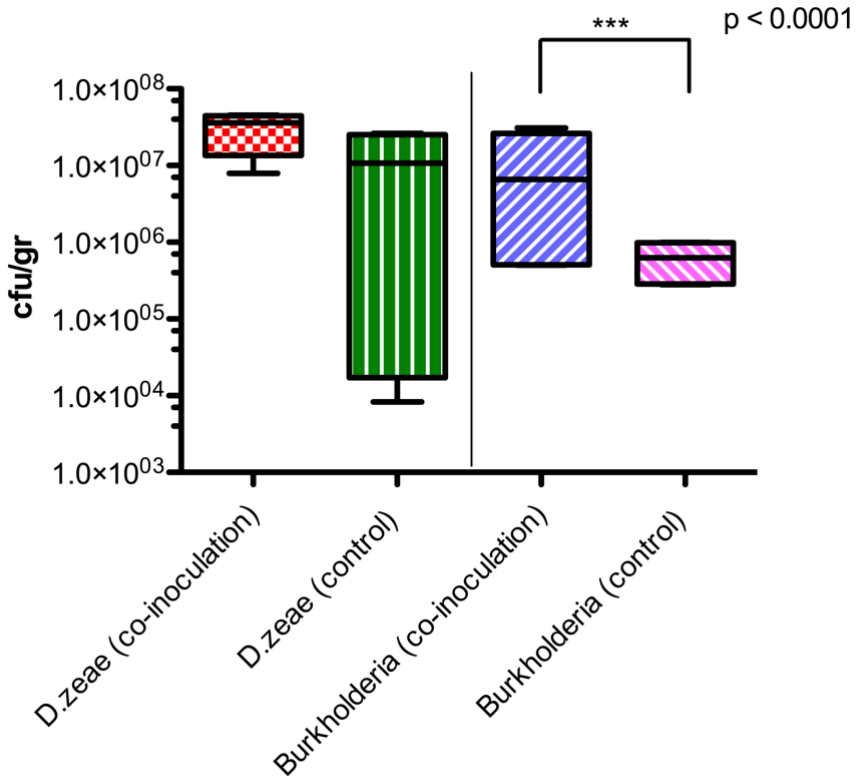


Figure 2. 12 Co-infection test on rice plants by *Dickeya zeae* and *Burkholderia* sp. A56 as possible co-operator. The putative co-operator was inoculated independently and in co-presence of the pathogen. CFU/g of *D.zeae* and *Burkholderia* sp. A56 recovered after the co-inoculation was evaluated in rice plant stem's lesion site/section at 15 dpi by plating serial dilutions from plant tissue and by antibiotic selection (*D.zeae* Rif and *Burkholderia* Amp_{100}, Gm_{40}, Nf_{100}). For each single and co-inoculation treatment, 8 plants were used. The control treatment was realized by single inoculation of pathogen or co-operators independently and PBS. Significance between groups were calculated using t-test; no significant differences were found between the CFU/g of *Dickeya* recovered.

2.4 Discussion

Studies on bacterial plant diseases have thus far mainly focused on single bacterial species causing the disease, with little attention given to the many other microorganisms present in the microbiome. In this study, the foot rot disease caused by *D. zeae* was used as model to investigate how the plant microbiome changes in response to a pathogen attack and to identify the most likely genera that could undergo interactions with the pathogen and possibly play a role in the pathosystem. This study shifts from a pathogen-focused view towards a more ecological perception of the foot rot disease foreseeing the entire microbial community dynamics.

Comparisons of the rice microbiome of asymptomatic and symptomatic plants in two different growing seasons and belonging to two rice cultivars has evidenced the dominant presence of *D. zeae* only in the plants showing the symptoms, suggesting that this approach is suitable for detecting the causal/primary agent of a disease. The number of *Dickeya* genera reads however showed variations in abundance among the single symptomatic samples, suggesting possible different times/stages of the infection. In addition, an increase in the abundance of *Dickeya* genera reads during the year 2018 compared to the 2017 was observed. This difference was not evident in the symptom's severity, since the infected plants either from the 2017 or 2018 year showed a similar degree of symptoms. The disease incidence (the number of plants affected) was lower during the 2017 compared to the 2018 year, this was probably due to the field growing conditions (temperature, soil and water status) and to the earlier stage of the disease.

Significant changes were detected in the bacterial composition and abundance between symptomatic and asymptomatic plants. Bacteria in the microbiome varied primarily according to the symptomatology, regardless of plant growing season or plant cultivar. There was a significant increase of the biodiversity between the asymptomatic and symptomatic plants from the 2018 year, while the difference between the two conditions in the 2017 was less evident; this is likely due to the early phase and low incidence of the disease. Bacterial communities from both asymptomatic or symptomatic samples clustered together independently by the year and the cultivar, suggesting that the seasonal changes affect only partially the microbiome composition of the asymptomatic or symptomatic plants while the influence of the rice cultivar is negligible. It is widely accepted that microbial community structures are affected by abiotic and host-dependent factors (Dastogeer et al., 2020); however, our results demonstrated that these two factors did not affect significantly the microbiome composition in our samples. Several published studies have often demonstrated that some diseases are associated with the changes occurring in the microbiome (Ma et al., 2019); however, it is still uncertain whether diversity change is a cause of microbiome-associated disease or a consequence.

The analyses and comparison of the community structure networks between healthy and disease states revealed that the microbiome of asymptomatic plants had a higher clustering coefficient compared to symptomatic plants. This supports the hypothesis that the bacterial community in a healthy situation is featured by biologically significant associations among OTUs which could be either cooperative or functionally related. In healthy plants, the total bacterial community is formed by a high number of bacterial genera which are present in similar amounts. Moreover, the community composition in asymptomatic samples appears

similar and stable. The fitness of the plant is likely to depend on the organized biodiversity of its microbiome that is effective in rapid adaptation/plasticity to environmental changes (Vandenkoornhuyse et al., 2015). The same analysis derived from symptomatic plants revealed a sparser network, with a lower number of edges and clustering coefficient (**Fig. 3**). Reduced connectivity of the microbial networks and an increase in heterogeneity in pathobiomes is consistent with previous studies and to the Anna Karenina principle (Zaneveld et al., 2017; Li and Convertino, 2019; Sweet et al., 2019).

Upon establishment of *D. zeae,* it was observed (i) conversion of a *Proteobacteria*-rich community to a *Firmicutes*-rich community; and (ii) an increase of the overall relative bacterial diversity. The latter is an exception to the common finding that pathobiomes are associated with decreased biodiversity. It is generally accepted that high diversity defines healthy microbiomes, whereas reduction in diversity may be associated with dysbiosis, as evident in human microbiome studies (Li et al., 2012; Lozupone et al., 2012). However, the inverse relationship between diversity and disease has been recently questioned (Ma et al., 2019) as other cases in which the dysbiotic state is associated with higher alpha diversity have been reported (Ceccarani et al., 2019; Lamelas et al., 2020). For example, bacterial communities associated with tropical coral-host revealed that diseased samples were associated with two to three times more bacterial diversity; this condition likely reflects imbalance in the community structure (Closek et al., 2014). The increase in taxa associated with the disease can also be due to opportunistic bacteria which may compete for available resources and/or benefit of the impaired host defense system.

Interestingly, *Firmicutes*, such as *Lachnospiraceae* and *Clostridiaceae*, were significantly more abundant in symptomatic samples. Moreover, the majority of the taxa that positively

correlate with the disease condition such as *Prevotella*, *Propionispira*, *Pleomorphomonas, Clostridium_sensu_stricto, Sulfurospirillum, Rhodopseudomonas* and many others are obligate or facultative anaerobes, suggesting that under disease conditions oxygen levels changes may favor growth of anaerobic bacteria. Accordingly, the drastic decrease of *Chloroplast* sequences in the symptomatic samples is also attributable to the rapid depletion of O_2 and to the anerobic conditions developed. Oxygen deficiency with increased availability of nutrients, as well as the reduction of the plant immune resistance, is likely to promote rapid growth of rotting bacteria favoring the establishment of decay (Ma et al., 2007). Under these conditions, only a small amount of the pathogen ($<10^2$) is likely to be required to initiate a lesion (Perombelon and Kelman, 1980) and *D. zeae* possibly undergoes several interactions with anaerobic bacteria that are present in the pathobiome as a result of plant rotting ecology.

Co-occurrence network analysis identified several positive and negative interactions with *D. zeae*; these may be involved in the establishment of a stable consortium that could influence the development of the disease. Very strong positive correlations were detected between *D. zeae* and *Prevotella* and *Propionispira*, independent of the year or the cultivar and consistent across samples, supporting the hypothesis that *D. zeae* undergoes specific positive interactions with some members of the pathobiome. The highest correlation occurred with *Prevotella*, which is able to decompose the hemicellulose and pectin, the major constituents of plant cell walls (Ueki et al., 2007). It is therefore likely that strains belonging to this genus play an important role in some steps of the disease development and aggravation.

To begin to address possible interactions between the pathogen and other microbes in the pathobiome-network, several bacterial strains from the pathobiome were isolated. Many

isolates displayed low abundance in the pathobiome community studies, the reason for this is intrinsic to the properties of the microbial communities, which are usually dominated by very few species; therefore, high-throughput cultivation would likely recover many species that have low abundance (Lynch and Neufeld, 2015). Among the taxa positively correlated with the disease condition, only a few strains were isolated as a single and pure colonies. Unfortunately, under the conditions used here (aerobic grown condition, limited number of different media, limited time-points), all the attempts failed to isolate strains belonging to *Prevotella* or *Proprionispora* genera; the reason for this is unknown, it could be that they require specific growth conditions or these strains cannot grow as pure isolates. *In planta* experiments were performed with a *Burkholderia* sp. strain, which significantly correlated with *D. zeae*, according to the LDA analysis and its 16S rRNA sequence match the sequence in the amplicon community analysis. Intriguingly, its plant colonization ability significantly increased when *D. zeae* is present, indicating that it benefits from the presence of the pathogen. The improved growth of *Burkholderia* sp. might be explained in multiple ways: it might be due to (i) the cell-cell signaling and cooperation between this member of the microbiome and the pathogen, (ii) the nutrients that become available as a result of the plant rotting and (iii) the plant derived defense molecules and metabolites, such as phytohormones, aromatic amino acids and phenolic compounds. The plant environment changes in response to a pathogen infection due to host cell wall lignification, synthesis of phytoalexins or increase in phenolics components and hormone-like substances (Glazebrook, 2005). These molecules may serve as nutrient source, signaling molecules, or have a antimicrobial activities, and may thus influence all the members of the community (Eichmann et al., 2021). These co-inoculations did not result in a more severe disease in our plant-disease model. It cannot be excluded that in the wild, *D. zeae* could benefit from the presence of

Burkholderia sp., resulting in more severe symptoms or more rapid development of the foot-rot disease.

The shift of the biodiversity in the pathobiome could be in part due to (i) the production by the pathogen of anti-microbial compounds, (ii) higher carbon and nitrogen source availability, (iii) the local and systemic plant immune response and (iv) plant-microbe signaling and microbe-microbe interactions. A notable example of cooperative behavior is the plant pathogen *Pseudomonas savastanoi* pv*. savastanoi* causing the olive-knot disease which undergoes interspecies cell-cell signaling via quorum sensing signaling molecules with endophytic harmless *Erwinia toletana* resulting in a more aggressive disease where both bacterial species benefit (Hosni et al., 2011; Buonaurio et al., 2015). A possible mechanism of mutual benefit is via complementarity in metabolic pathways resulting in an ameliorated metabolic capacity of the consortium compared to the single species (Buonaurio et al., 2015). Future research is needed to understand how members of the pathobiome modify the activity of the primary pathogen; this will require functional studies of the pathosystem as well community profiling.

In order to determine possible commonalities among different pathosystems, a similar set of experiments of another rice pathogen was performed. The plant pathobiome and microbiome of an important bacterial disease of rice called leaf blight (BLB), caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) (Mew et al., 1993; Mansfield et al., 2012), has been determined. Importantly, microbial pathobiome/microbiome experiments performed here from rice plant samples from Vietnam on the BLB vascular rice disease caused by *Xoo,* did not result in any perturbation/shift in the microbiome showing that the pathobiome has very low biodiversity and that this disease is very different from rice foot-rot. (**Figure 2.13**). *Xoo* is a

vascular pathogen colonizing in solitary the vascular system thus the microbiome does not seem to play a major role in influencing this disease. Furthermore, unlike what has been reported for the rice *D. zeae* pathosystem, no evidence to support the Anna Karenina principle, which predicts higher heterogeneity in destabilized microbial communities, were found. Generalizations on the role of the microbiome and of possible positive biotic interactions in plant disease therefore cannot be made.

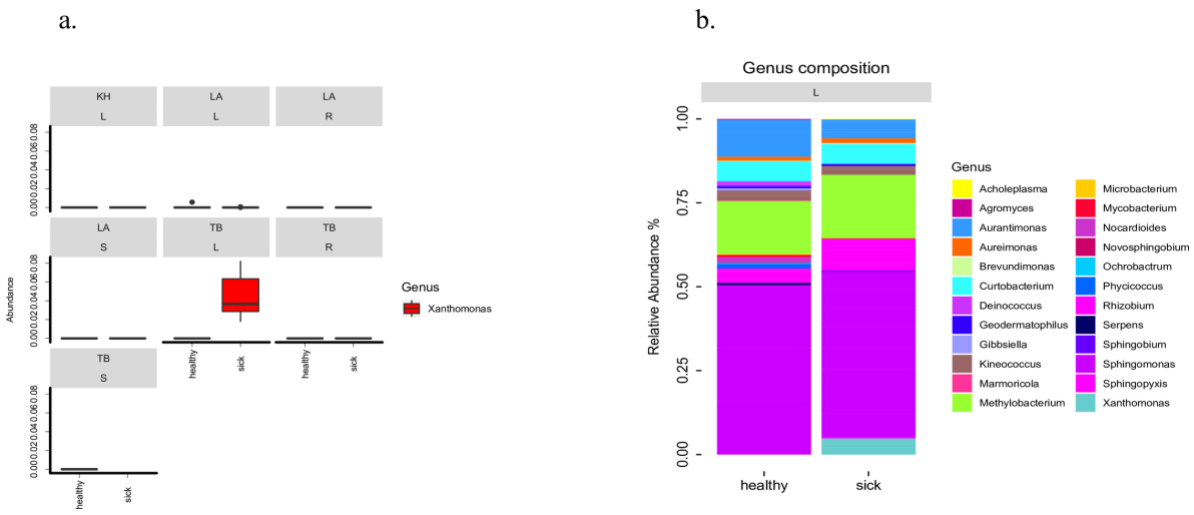


Figure 2. 13 Pathobiome study of rice bacterial leaf blight (BLB). A) Comparison of *Xanthomonas* genus abundance between symptomatic and asymptomatic samples from three different rice cultivars in Vietnam. The presence of *Xanthomonas* has been checked in sample leaves (L), roots (R) and steam (S) from three different rice cultivars (LA= Long An, KH= Khanh Hoa and TB = Thai Binh). Box plot depict medians (central horizontal lines), the inter-quartile ranges (boxes), 95% confidance intervals (whiskers) and outliers (black dots). Statistical analyses were calculated based on wilcoxon-test. **B)** Genus distribution of the OTUs among asymptomatic and symptomatic leaf -samples from Thai Binh cultivar.

In summary, *D. zeae* alters the resident bacterial community in species composition, abundance and richness leading to the formation of a microbial consortia/pathobiome linked to the disease state. Our results have shown that the driving force of microbial community variation is the presence or absence of the foot-rot symptoms, while the effect of the growing season and the cultivar is negligible. Specific anaerobic bacterial taxa, which are significantly co-present with the pathogen over the two years, were detected, suggesting a possible involvement in the disease development through the facilitation of its colonization, expression of virulence factors, and reduction of host resistance. It is likely that the bacterial foot rot disease is the result of a multi-species interaction and not solely due to one single pathogen. It is important to begin shifting from the paradigm of pathogens to pathobiome in order to understand the importance of microbial communities. Deciphering the molecular basis of interbacterial relationships in the plant pathobiome will be a major future challenge in order to better understand the pathogenicity and epidemiology of microbial plant diseases and target polymicrobial infections.

2.5 References

Agler, M.T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S.-T., Weigel, D., and Kemen, E.M. (2016) Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology* **14**.

Barras, F., van Gijsegem, F., and Chatterjee, A.K. (1994) Extracellular enzymes and pathogenesis of soft-rot *Erwinia*. *Annual review of phytopathology* **32**: 201-234.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A. et al. (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology* **37**: 852-857.

Brader, G., Compant, S., Vescio, K., Mitter, B., Trognitz, F., Ma, L.-J., and Sessitsch, A. (2017) Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annual review of phytopathology* **55**: 61-83.

Brady, C.L., Cleenwerck, I., Denman, S., Venter, S.N., Rodríguez-Palenzuela, P., Coutinho, T.A., and De Vos, P. (2012) Proposal to reclassify Brenneria quercina (Hildebrand and Schroth 1967) Hauben et al. 1999 into a new genus, Lonsdalea gen. nov., as Lonsdalea quercina comb. nov., descriptions of Lonsdalea quercina subsp. quercina comb. nov., Lonsdalea quercina subsp. iberica subsp. nov. and Lonsdalea quercina subsp. britannica subsp. nov., emendation of the description of the genus Brenneria, reclassification of *Dickeya dieffenbachiae* as *Dickeya dadantii* subsp. dieffenbachiae comb. nov., and emendation of the description of *Dickeya dadantii*. *International journal of systematic and evolutionary microbiology* **62**: 1592-1602.

Buonaurio, R., Moretti, C., da Silva, D.P., Cortese, C., Ramos, C., and Venturi, V. (2015) The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. *Frontiers in plant science* **6**: 434.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods* **13**: 581.

Ceccarani, C., Foschi, C., Parolin, C., D'Antuono, A., Gaspari, V., Consolandi, C. et al. (2019) Diversity of vaginal microbiome and metabolome during genital infections. *Scientific reports* **9**: 1-12.

Cheng, Y., Liu, X., An, S., Chang, C., Zou, Y., Huang, L. et al. (2013) A nonribosomal peptide synthase containing a stand-alone condensation domain is essential for phytotoxin zeamine biosynthesis. *Molecular plant-microbe interactions* **26**: 1294-1301.

Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. (2012) The impact of the gut microbiota on human health: an integrative view. *Cell* **148**: 1258-1270.

Closek, C.J., Sunagawa, S., DeSalvo, M.K., Piceno, Y.M., DeSantis, T.Z., Brodie, E.L. et al. (2014) Coral transcriptome and bacterial community profiles reveal distinct Yellow Band Disease states in Orbicella faveolata. *The ISME journal* **8**: 2411-2422.

Collmer, A., and Bauer, D. (1994) *Erwinia chrysanthemi* and *Pseudomonas syringae*: plant pathogens trafficking in extracellular virulence proteins. In *Bacterial Pathogenesis of Plants and Animals*: Springer, pp. 43-78.

Dastogeer, K.M., Tumpa, F.H., Sultana, A., Akter, M.A., and Chakraborty, A. (2020) Plant microbiome–an account of the factors that shape community composition and diversity. *Current Plant Biology*: 100161.

Eichmann, R., Richards, L., and Schäfer, P. (2021) Hormones as go-betweens in plant microbiome assembly. *The Plant Journal* **105**: 518-541.

Faust, K., and Raes, J. (2012) Microbial interactions: from networks to models. *Nature Reviews Microbiology* **10**: 538.

Friedman, J., and Alm, E.J. (2012) Inferring correlation networks from genomic survey data. *PLoS Comput Biol* **8**: e1002687.

Gilbert, P. (2016) *Human nature and suffering*: Routledge.

Glazebrook, J. (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu Rev Phytopathol* **43**: 205-227.

Goto, M. (1979) Bacterial foot rot of rice caused by a strain of *Erwinia chrysanthemi*. *Phytopathology* **69**: 213-216.

Hosni, T., Moretti, C., Devescovi, G., Suarez-Moreno, Z.R., Fatmi, M.B., Guarnaccia, C. et al. (2011) Sharing of quorum-sensing signals and role of interspecies communities in a bacterial plant disease. *The ISME journal* **5**: 1857.

Hussain, M.B., Zhang, H.-B., Xu, J.-L., Liu, Q., Jiang, Z., and Zhang, L.-H. (2008) The acyl-homoserine lactone-type quorum-sensing system modulates cell motility and virulence of *Erwinia chrysanthemi pv. zeae*. *Journal of bacteriology* **190**: 1045-1053.

Jafra, S., Przysowa, J., Gwizdek-Wiśniewska, A., and Van Der Wolf, J. (2009) Potential of bulb-associated bacteria for biocontrol of hyacinth soft rot caused by *Dickeya zeae*. *Journal of applied microbiology* **106**: 268-277.

Katoh, K., and Standley, D.M. (2014) MAFFT: iterative refinement and additional methods. In *Multiple sequence alignment methods*: Springer, pp. 131-146.

Lamelas, A., Desgarennes, D., López-Lima, D., Villain, L., Alonso-Sánchez, A., Artacho, A. et al. (2020) The bacterial microbiome of meloidogyne-based disease complex in coffee and tomato. *Frontiers in plant science* **11**: 136.

Lamichhane, J.R., and Venturi, V. (2015) Synergisms between microbial pathogens in plant disease complexes: a growing trend. *Frontiers in plant science* **6**: 385.

Laurila, J., Hannukkala, A., Nykyri, J., Pasanen, M., Hélias, V., Garlant, L., and Pirhonen, M. (2010) Symptoms and yield reduction caused by *Dickeya spp*. strains isolated from potato and river water in Finland. *European Journal of Plant Pathology* **126**: 249-262.

Li, J., and Convertino, M. (2019) Optimal microbiome networks: macroecology and criticality. *Entropy* **21**: 506.

Li, K., Bihan, M., Yooseph, S., and Methe, B.A. (2012) Analyses of the microbial diversity across the human microbiome. *PloS one* **7**: e32118.

Liu, H., Brettell, L.E., Qiu, Z., and Singh, B.K. (2020) Microbiome-mediated stress resistance in plants. *Trends in plant science*.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* **489**: 220-230.

Lynch, M.D., and Neufeld, J.D. (2015) Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology* **13**: 217-229.

Ma, B., Hibbing, M.E., Kim, H.-S., Reedy, R.M., Yedidia, I., Breuer, J. et al. (2007) Host range and molecular phylogenies of the soft rot enterobacterial genera *Pectobacterium* and *Dickeya*. *Phytopathology* **97**: 1150-1163.

Ma, Z.S., Li, L., and Gotelli, N.J. (2019) Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. *The ISME journal* **13**: 1911-1919.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P. et al. (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular plant pathology* **13**: 614-629.

McMurdie, P.J., and Holmes, S. (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one* **8**: e61217.

Mew, T.W., Alvarez, A.M., Leach, J., and Swings, J. (1993) Focus on bacterial blight of rice. *Plant disease* **77**: 5-12.

Nassar, A., Bertheau, Y., Dervin, C., Narcy, J.-P., and Lemattre, M. (1994) Ribotyping of *Erwinia chrysanthemi* strains in relation to their pathogenic and geographic distribution. *Appl Environ Microbiol* **60**: 3781-3789.

Perombelon, M.C., and Kelman, A. (1980) Ecology of the soft rot erwinias. *Annual review of phytopathology* **18**: 361-387.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular biology and evolution* **26**: 1641-1650.

Proctor, L. (2019) Priorities for the next 10 years of human microbiome research. In: Nature Publishing Group.

Pu, X., Zhou, J., Lin, B., and Shen, H. (2012) First report of bacterial foot rot of rice caused by a *Dickeya zeae* in China. *Plant disease* **96**: 1818-1818.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. et al. (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**: D590-D596.

Sabet, K. (1954) A new bacterial disease of maize in Egypt. *Empire Journal of Experimental Agriculture* **22**: 65-67.

Samson, R., Legendre, J.B., Christen, R., Saux, M.F.-L., Achouak, W., and Gardan, L. (2005) Transfer of *Pectobacterium chrysanthemi* (Burkholder et al. 1953) Brenner et al. 1973 and Brenneria paradisiaca to the genus *Dickeya gen*. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii sp*. nov., *Dickeya dianthicola sp*. nov., *Dickeya dieffenbachiae sp*. nov. and *Dickeya zeae sp*. nov. *International Journal of Systematic and Evolutionary Microbiology* **55**: 1415-1427.

Schlaeppi, K., and Bulgarelli, D. (2015) The plant microbiome at work. *Molecular Plant-Microbe Interactions* **28**: 212-217.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011) Metagenomic biomarker discovery and explanation. *Genome biology* **12**: R60.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**: 2498-2504.

Sławiak, M., Łojkowska, E., and Van der Wolf, J. (2009) First report of bacterial soft rot on potato caused by *Dickeya sp*.(syn. *Erwinia chrysanthemi*) in Poland. *Plant Pathology* **58**.

Steindler, L., Bertani, I., De Sordi, L., Schwager, S., Eberl, L., and Venturi, V. (2009) LasI/R and RhlI/R quorum sensing in a strain of *Pseudomonas aeruginosa* beneficial to plants. *Appl Environ Microbiol* **75**: 5131-5140.

Sweet, M., Burian, A., Fifer, J., Bulling, M., Elliott, D., and Raymundo, L. (2019) Compositional homogeneity in the pathobiome of a new, slow-spreading coral disease. *Microbiome* **7**: 1-14.

Ueki, A., Akasaka, H., Satoh, A., Suzuki, D., and Ueki, K. (2007) *Prevotella paludivivens* sp. nov., a novel strictly anaerobic, Gram-negative, hemicellulose-decomposing bacterium isolated from plant residue and rice roots in irrigated rice-field soil. *International journal of systematic and evolutionary microbiology* **57**: 1803-1809.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., and Dufresne, A. (2015) The importance of the microbiome of the plant holobiont. *New Phytologist* **206**: 1196-1206.

Vannier, N., Agler, M., and Hacquard, S. (2019) Microbiota-mediated disease resistance in plants. *PLoS pathogens* **15**: e1007740.

Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.F., Jacques, M.-A., Lebrun, M.-H. et al. (2014) Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Frontiers in cellular and infection microbiology* **4**: 29.

Watts, S.C., Ritchie, S.C., Inouye, M., and Holt, K.E. (2019) FastSpar: rapid and scalable correlation estimation for compositional data. *Bioinformatics* **35**: 1064-1066.

Zaneveld, J.R., McMinds, R., and Thurber, R.V. (2017) Stress and stability: applying the Anna Karenina principle to animal microbiomes. *Nature microbiology* **2**: 1-8.

Zhou, J., Zhang, H., Wu, J., Liu, Q., Xi, P., Lee, J. et al. (2011) A novel multidomain polyketide synthase is essential for zeamine production and the virulence of *Dickeya zeae*. *Molecular plant-microbe interactions* **24**: 1156-1164.

Zhou, J.N., Zhang, H.B., Lv, M.F., Chen, Y.F., Liao, L.S., Cheng, Y.Y. et al. (2016) SlyA regulates phytotoxin production and virulence in *Dickeya zeae* EC1. *Molecular plant pathology* **17**: 1398-1408.

Chapter III

LuxR solos from environmental fluorescent pseudomonads

Keywords: Quorum sensing, LuxR solos, Fluorescent Pseudomonads, inter-species signaling, plant-associated bacteria

3.1 Introduction

The recent development of omics methodologies and the extensive studies in microbial diversity have evidenced that in nature microbes live as part of complex mixed communities. For this reason, microbes very likely communicate and socially interact with numerous different species in order to cooperate, synchronize and synergize their behavior in response to environmental changes. Quorum Sensing (QS) is one type of social interaction among bacteria, which regulates gene expression in response to cell density, playing a major role in the formation and stability of microbial populations (Waters and Bassler, 2005; Ng and Bassler, 2009). QS cell-cell signaling in bacteria has so far been mostly addressed in simple settings, mainly focusing on single species and thus limiting our understanding of its possible roles in complex mixed communities.

To date, the most common and best studied QS system in Gram-negative proteobacteria is mediated by *N*-acyl homoserine lactone (AHL) signals. The archetypical AHL-QS system is composed by two most commonly genetically adjacent genes; the *luxI* family gene encoding an AHL synthase and its cognate *luxR* family gene, which encodes a transcriptional factor that detects and responds to the cognate AHL (Fuqua et al., 1994; Zhu and Winans, 2001; Fuqua and Greenberg, 2002). LuxR-type family proteins are approximately 250 amino acids long and consist of two domains: an inducer binding domain (IBD) at the N terminus (Shadel et al., 1990; Slock et al., 1990) and a DNA-binding helix-turn-helix (HTH) domain at the C terminus (Choi and Greenberg, 1991). The IBD of canonical LuxRs recognizes AHLs resulting in a conformational change that affects its ability to bind target DNA in gene promoter regions at conserved sites, called *lux boxes* (Devine et al., 1989; Stevens and Greenberg, 1997). LuxRs share 9 highly conserved amino acid residues (Whitehead et al., 2001; Zhang et al., 2002). Six

are hydrophobic or aromatic and form the cavity of the IBD and the remaining three are in the HTH domain (Zhang et al., 2002). LuxR family proteins are a source of adaptability and flexibility in QS circuits, allowing for alterations in response to AHL types or different signal molecules. In particular, signal specificity can be altered by specific changes in some residues of LuxR receptors (Lintz et al., 2011). LuxRs can also be promiscuous, by binding not only to their cognate AHL, but to multiple AHL types and thus responding to non-self signals (Wellington and Greenberg, 2019). This eavesdropping through promiscuous receptors may play a role in interspecies interactions and can affect both interspecies competition and cooperation, expanding the function of QS systems in complex bacterial communities (McClean et al., 1997; Hawver et al., 2016).

Analysis of different genomes of proteobacteria has uncovered the widespread presence of LuxR regulators that occur without the cognate LuxI homolog; these are referred to as LuxR solos or orphans LuxRs (Fuqua, 2006; Case et al., 2008; Patankar and González, 2009; Subramoni et al., 2015). LuxR solos are closely related to QS LuxRs, displaying significant primary structure homology and having the same two-domain organization and modular structure as canonical LuxR proteins. LuxR solos can expand the regulatory targets, by responding to endogenous or exogenous AHLs, also resulting in interspecies signaling. For example, QscR from *P. aeruginosa* responds to endogenously produced AHLs (Chugani et al., 2001; Lequette et al., 2006), SdiA of *Salmonella enterica* and *Escherichia coli*, which do not produce AHLs, responds to AHLs synthetized by neighboring bacteria (Ahmer et al., 1998; Michael et al., 2001; Ahmer, 2004), whereas PpoR from *P. putida* binds to AHLs either self or foreign (Subramoni and Venturi, 2009a, b).

LuxR solos have also been implicated in interkingdom signaling, having evolved to respond to signals produced by eukaryotes (Soares and Ahmer, 2011; Venturi and Fuqua, 2013). A sub-group of LuxR solos which is only found in plant-associated bacteria (PAB) responds to plant low-molecular-weight molecules (González et al., 2013; González and Venturi, 2013; Venturi and Fuqua, 2013). Compared to canonical QS LuxRs, members of this subfamily have some substitutions among the highly conserved amino acids in the IBD, which very likely correspond with their ability to bind low-molecular-weight compounds produced by plants (Ferluga and Venturi, 2009). Members of this sub-family are found in both plant-pathogenic bacteria, such as XccR of *Xanthomonas campestris*, OryR of *Xanthomonas oryzae*, XagR of *Xanthomonas axonopodis*, or beneficial bacteria, like NesR in *Sinorhizobium meliloti*, PsoR of rhizospheric *Pseudomonas fluorescens*, PipR of plant endophytic *Pseudomonas* sp. strain GM79 and PsrR of *Kosakonia* sp.strain KO348 (Ferluga et al., 2007; Zhang et al., 2007; Ferluga and Venturi, 2009; Patankar and González, 2009; Subramoni and Venturi, 2009a; Coutinho et al., 2018; Mosquito et al., 2020). Finally the LuxR solos PluR and PauR from the insect pathogen *Photorhabdus luminescens* and human and insect pathogen *Photorhabdus asymbiotica,* respectively, respond to novel endogenous molecules, namely photopyrones and dialkylresorcinols (Brachmann et al., 2013; Brameyer et al., 2014, 2015). In summary, LuxR solos extend beyond next of kin AHL-driven QS, being used in different ways by bacteria and thus becoming major players in cell-cell communication in bacteria (Brachmann et al., 2013; Prescott and Decho, 2020).

In this study it is intended to begin to map LuxR solos in the model proteobacterial *Pseudomonas* genus by genomics and genetics/molecular biology approaches. The distribution, conservation and possible responses of a set of LuxR solos within the *P.*

fluorescent group was investigated. This group of bacteria is one of the most diverse groups within the *Pseudomonas* genus comprising more than fifty species and many unclassified isolates many of which are found in plant-associated environments (Garrido-Sanz et al., 2016; Garrido-Sanz et al., 2017). Many members of the fluorescent *Pseudomonas* spp. are excellent rhizosphere colonizers and are studied for their plant beneficial properties (O'sullivan and O'Gara, 1992). An analysis for QS LuxRs domains of over 600 genomes has evidenced the predominance of LuxR solos in fluorescent pseudomonads, which were divided into nine different sub-groups based on their neighboring genes and their primary structure. The cartography of their ligand binding site allowed to classify each LuxR solo into potential AHL-binding or non-AHL-binding. LuxR solos genomic knock-out mutants in several *Pseudomonas* sp. strains of different sub-groups have been generated and studied in order to get insights into possible gene targets and mechanisms of action. Overall, our analysis revealed that LuxR solos homologs occur considerably more frequently than complete LuxI/LuxR QS systems within the *P. fluorescent* group and that LuxR solos from closely related genomes or from genomes carrying multiple LuxR solos cluster in different sub-groups. These results highlight the existence of novel and diverse LuxR solos sub-groups, that could be involved in intercellular cell-cell or intracellular signaling regulatory functions. Some could have evolved away from canonical QS LuxRs and possibly bind to new signals/molecules.

3.2 Material and methods

3.2.1 Bacterial species, culture conditions and genomes sequencing

The bacterial strains used in this work were as follows: *P. putida P.16A* and *P. oleovorans* AG1008 (isolated from rhizosphere and endosphere rice plants collection during this project), *P. fluorescens F113* (Shanahan et al., 1992) and *P. jessenii* DSM17150 (Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Germany). All strains were grown in liquid Luria-Bertani (LB), King's Broth (KB) or M9 medium at 30 °C under moderate shaking (120 rpm). When required, antibiotics for *Pseudomonas* strains growth were added at the following concentrations: nitrofurantoin (Nf) 100 µg ml $^{-1}$, ampicillin (Amp) 100 µg ml^{-1}. The mutants of each strain (carrying a knock-out mutation of the *luxR* solo gene) have been grown using 100 µg ml^{-1} kanamycin (Km) as antibiotic. *E. coli* DH5α, S17 and BL21 (DE3) were routinely grown at 37 °C in LB broth and antibiotics were added when required at the following concentrations: Amp 100 µg ml^{-1}, tetracyclin 15 µg ml^{-1}. AHLs were obtained from Sigma-Aldrich (St. Louis, MO, USA).

The complete genomes of twenty fluorescent *Pseudomonas* were sequenced with the Illumina MiSeq platform using 150 bp paired-end reads and following the tagmentation Illumina Nextera XT protocol (Illumina Inc., San Diego, CA, USA). The sequencing was performed by the Exeter University (UK). Sequenced genomic DNA was assembled using Spades 3.9.03 (Bankevich et al., 2012) and the assembled sequence annotated, using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). Genomes were also annotated using RAST (Rapid Annotation using Subsystem Technology) Server (Aziz et al., 2008), uploaded to the Integrated Microbial Genomes and Metagenomes (IMG/M) database and automatically

annotated, using annotation pipeline IMG Annotation Pipeline v.4.16.6 (Markowitz et al., 2009).

Each Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank and is accessible through the BioProject ID PRJNA701950. The Whole Genome Shotgun project of *Pseudomonas sp.*18_A has been deposited under the accession JAFGYG000000000 and the version described in this paper is version JAFGYG010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*29_B has been deposited under the accession JAFGYH000000000 and the version described in this paper is version JAFGYH010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*32_A has been deposited under the accession JAFGYI000000000 and the version described in this paper is version JAFGYI010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*43(2021) has been deposited under the accession JAFGYJ000000000 and the version described in this paper is version JAFGYJ010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*21_B has been deposited under the accession JAFGYK000000000 and the version described in this paper is version JAFGYK010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*67(2021) has been deposited under the accession JAFGYL000000000 and the version described in this paper is version JAFGYL010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*69_B has been deposited under the accession JAFGYM000000000 and the version described in this paper is version JAFGYM010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*71_D has been deposited under the accession JAFGYN000000000 and the version described in this paper is version JAFGYN010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*74_A has been deposited under the accession JAFGYO000000000 and the version described in this paper is version

JAFGY010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*78_B has been deposited under the accession JAFGYP000000000 and the version described in this paper is version JAFGYP010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*79_C has been deposited under the accession JAFGYQ000000000 and the version described in this paper is version JAFGYQ010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*81_B has been deposited under the accession JAFGYR000000000 and the version described in this paper is version JAFGYR010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*86_A has been deposited under the accession JAFGYS000000000 and the version described in this paper is version JAFGYS010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*95_A has been deposited under the accession JAFGYT000000000 and the version described in this paper is version JAFGYT010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*100_A has been deposited under the accession JAFGYU000000000 and the version described in this paper is version JAFGYU010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*16_A has been deposited under the accession JAFGYV000000000 and the version described in this paper is version JAFGYV010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*51_B has been deposited under the accession JAFGYW000000000 and the version described in this paper is version JAFGYW010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*50_B has been deposited under the accession JAFGYX000000000 and the version described in this paper is version JAFGYX010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*30_B has been deposited under the accession JAFGYY000000000 and the version described in this paper is version JAFGYY010000000. The Whole Genome Shotgun project of *Pseudomonas sp.*58(2021) has been deposited under the accession JAFGYZ000000000 and the version described in this paper is version JAFGYZ010000000.

3.2.2 Plasmid and recombinant DNA techniques

The plasmids, constructs and set of specific primers (Sigma-Aldrich) used in this study are listed in **Table 3.1**. pGEM-T Easy vector (Promega Corp., Madison, WI, USA) was used for cloning. When necessary, 5-bromo-4-chloro-3-indolyl-β-d-galactoside (X-Gal) was added at a final concentration of 80 µg ml^{-1}._Routine DNA manipulation steps, such as digestion with restriction enzymes, agarose gel electrophoresis, purification of DNA fragments, ligation with T4 DNA ligase and transformation of_*E. coli*, were performed as described previously (Sambrook et al., 1989). Plasmids were purified by using EuroGold columns (EuroClone, Pero, Milan, Italy); total DNA was isolated by sarkosyl-pronase lysis, as described previously (Better et al., 1983). Digestion with restriction enzymes was conducted according to the supplier's instructions (New England BioLabs, USA). DNA was ligated with T4 DNA ligase (New England BioLabs, USA) according to the manufacturer's recommendations.

Table 3. 1 Plasmids and primers used in this study.

Plasmids/primers	Relevant features	References or sources
pGEM-T	Cloning vector; Amp^R	Promega
pMP220	Promoter probe vector; IncP;Tc^R	(Spaink et al., 1987)
pBBR1MCS-5	Broad-host-range vector; Gm^R	(Kovach et al., 1995)
pLAFR3	Broad-host-range vector; IncP;Tc^R	(Staskawicz et al., 1987)
pEX19Gm	Suicide vector for making deletion mutants, Gm^R	(Dreier and Ruggerone, 2015)
pETM-11	His_6-tagged protein expression vector	Addgene, Watertown, MA
pUC4K	pUC7 derivative, Amp^R and Km^R	Addgene, Watertown, MA
pEX19-PpuR16R	PpuR16R sequence depleted of 20 bp cloned in pEX19Gm	This study
pEX19-PpuR16R_2	PpuR16R_2 sequence depleted of 20 bp cloned in pEX19Gm	This study
pEX19-PfluR_113	PfluR_113 sequence depleted of 20 bp cloned in pEX19Gm	This study
pEX19-PjeR	PjeR sequence depleted of 20 bp cloned in pEX19Gm	This study
pEX19-PolR	PolR sequence depleted of 20 bp cloned in pEX19Gm	This study
pPppu16R220	Ppu16R promoter cloned in pMP220	This study

pPferr220	Ferrodoxin-NADP-reductase promoter cloned in pMP220	This study
pP23S220	23S rRNA methyltransferase promoter cloned in pMP220	This study
pPppu16R2_220	Ppu16R_2 promoter cloned in pMP220	This study
pPfluR220	PfluR_113 promoter cloned in pMP220	This study
pPmoaF220	MoaF promoter cloned in pMP220	This study
pPjeR220	PjeR promoter cloned in pMP220	This study
pPsperm220	Spermidine permease promoter cloned in pMP220	This study
pPolR220	PolR promoter clone\d in pMP220	This study
pPputr220	Putrescine importer promoter cloned in pMP220	This study
pBBR-PfluR	PfluR_113 cloned in ΔPfluR_113	This study
pETM-Ppu16R	Ppu16R sequence cloned in pETM-11	This study
Primers	**Sequence**	**Source**
KmR1	CAACTCTGGCGCATCGGGCT	This study
KmR2	GCGTAATGCTCTGCCACACA	This study
P16A_SOLO_EXT	GAGATTTCCTACACTTCGTTC	This study
P16A_SOLO2_EXT	AGATCGTCAACGACGGC	This study
PF113_SOLO_EXT	TGGTCAGCGAGAGTTTCGTC	This study
PJES_SOLO_EXT	GTGCTCGCTAAAGGATTCAG	This study
POLEOV_SOLO_EXT	ACTCTAGGCCAGGGTGGG	This study
FW_F113_SOLO_compl_Xba	TCTAGACTGTGGGAAGTGGTCA	This study
RV_F113_SOLO_compl_Kpn	GGTACCTGGTTGATCAGAGGAA	This study

3.2.3 Genomic mutant construction and their complementation

In frame deletions of the *luxR solo* genes were generated using the pEX19Gm plasmid as described previously (Hoang et al., 1998). Briefly, each *luxR solo* gene sequence, synthetized by Twist bioscience company (South San Francisco), is listed in the **Supplementary Table S3.1**. The design of the constructs was performed as follows: internal fragments of 20 bp from each gene of interest were deleted and replaced with a restriction site (BamHI or SmaI) in order to clone inside the Km gene cassette previously extracted from pUC4K. Sequentially the fragments were excised with Kpn and XbaI restriction enzymes and cloned in the corresponding site in pEX19Gm. The resulting pEX19Gm-derivative plasmids, listed in **Table**

3.1, were introduced by biparental conjugation in the corresponding *Pseudomonas* genomes. Clones with a chromosomal insertion of the pEX19Gm plasmids were selected on LB agar plates supplemented with 40 μg ml^{-1} gentamycin (Gm) and 100 μg ml^{-1} Nf. Plasmid excision from the chromosome was subsequently selected on LB agar plates supplemented with 10% (w/v) sucrose. All the mutants were verified by PCR using primers (**Table 3.1**) specific to the Km cassette and to the genomic DNA sequences upstream and downstream from the targeted genes. For complementation analysis, the encoding regions of each *luxR solos* full-length genes were amplified by the primers listed in **Table 3.1**. The PCR products were digested with restriction enzymes and then cloned in the expression vector pBBR1MCS-5 vector (Kovach et al., 1995) digested with the same enzymes. The complementation constructs were introduced into corresponding mutants by bi-parental mating selected for Km^R and Gm^R and confirmed by PCR analysis.

3.2.4 β-galactosidase activity assay

In order to identify possible target genes, the promoter regions of several genes adjacent to each *luxR* solo studied were synthetized by Twistbioscience company (South San Francisco) and cloned into promoter probe vector pMP220, which harbors a promoter-less *lacZ* gene, as described in **Table 3.1** and **Supplementary Table S3.1**. pMP-derivative constructs were then introduced independently into the wild type strain and each corresponding *luxR* solo mutants by conjugation. β-galactosidase assays were performed as previously described by Miller (Miller and Lee, 1984), with the modifications of Stachel *et al*. (Stachel et al., 1985). Average Miller unit values and standard deviations were calculated from three independent experiments. When necessary, AHLs (C_4 homoserine lactone -HSL-, C_6-HSL, OHC_6-HSL, OC_6-HSL, C_8-HSL, OHC_8-HSL, OC_8-HSL, C_{10}-HSL, OHC_{10}-HSL, OC_{10}-HSL, C_{12}-HSL, OHC_{12}-HSL, OC_{12}-

HSL) were added at the final concentration of 1μM, as well as a cocktail of polyamines (putrescine, spermine, spermidine) (Sigma-Aldrich, St. Louis, MO, USA) at a final concentration of 0.1 mM.

3.2.4 Statistical Analysis

For analysis of statistical significance, the data were analyzed using GraphPad Prism's t-test or ANOVA and $P < 0.05$ was considered significant for all experiments.

3.2.5 Fluorescent pseudomonads strains isolation

A set of twenty fluorescent pseudomonads strains were purified from a laboratory collection of rhizosphere and endosphere of rice plants (Bertani et al., 2016), stored in glycerol at -80°C. The samples were plated on KB agar medium supplemented with an iron-chelator such as ethylendiamine-N,N'-diacetic acid (EDDA) (Sigma-Aldrich, St.Louis, MO, USA). Fluorescent pseudomonads strains producing fluorescent siderophores in iron-limited conditions were detected exposing the plates under UV-rays. The fluorescent colonies were isolated and stored at -80ºC in a 18% glycerol suspension.

3.2.6 Protein and Sequence Data download

Protein FASTA sequences of 601 genomes from 17 *Pseudomonas* species were downloaded from PATRIC database (Wattam et al., 2014).

3.2.7 Detection of LuxR/luxR and LuxI/luxI protein/genes

Hidden Markov Model (HMM) recognizers were collected from PFAM for Autoinducer binding domain and GerE domain typical of luxR and autoinducer synthase domain from

Interpro for identification of luxI proteins. These HMM recognizers were used to identify luxR and luxI proteins among all Pseudomonas strains using hmmsearch tool (Johnson et al., 2010). Hits stronger than E- value of 10^{-10} were taken as potential homologues of QS genes.

3.2.8 Phylogenetic analysis

Phylogenetic tree for all the *Pseudomonas* strains were built using the MEGAX program package (Kumar et al., 2018) installed from http://www.megasoftware.net using Neighbour-Joining method and then visualized using ggtree package in R (Yu et al., 2017).

3.2.9 Homology modeling and structural alignments

Five web-based servers were exploited to build the 3D homology models of the IBD of each LuxR solo studied. The top-score models generated by the different approaches were then ranked and validated by the protein model quality predictor ProQ (Wallner et al., 2003) including PSIPRED (Buchan and Jones, 2019) for secondary structure prediction. The top-scored model of the prototype of sub-group A, Ppu16R, (having the predicted LGscores and MaxSub values of 4.155 and 0.336, respectively) was obtained by M4T (Fernandez-Fuentes et al., 2007), based on two templates: SdiA from *E. coli* (PDB_ID 4Y13) (Nguyen et al., 2015) and CviR from *Chromobacterium violaceum* (PDB_ID 3QP6) (Chen et al., 2011).

The top-scored model of the prototype of sub-group B (having the predicted LGscores and MaxSub values of 4.078 and 0.725, respectively) was obtained by M4T (Fernandez-Fuentes et al., 2007), based on two templates: QscR from *P. aeruginosa* (PDB_ID 3SZT) (Lintz et al., 2011) and SdiA from *E. coli* (PDB_ID 4Y13) (Nguyen et al., 2015).

The top-scored model of the prototype of sub-group D, PfluR_113, (having the predicted LGscores and MaxSub values of 4.078 and 0.725, respectively) was obtained by RaptorX (Wang et al., 2016), based on CviR from *C. violaceum* (PDB_ID 3QP5) (Chen et al., 2011) as a template.

Regarding the sub-group E, two members, PolR and PjeR, have been modelled. In detail, the top-scored model of PolR (having the predicted LGscores and MaxSub values of 4.063 and 0.870, respectively) was obtained by Phyre2 (Kelley et al., 2015), based on CviR from *C. violaceum* (PDB_ID 3QP5) (Chen et al., 2011), as a template. The top-scored model of PjeR (having the predicted LGscores and MaxSub values of 4.205 and 0.580, respectively) was obtained by M4T (Fernandez-Fuentes et al., 2007), based on five templates: TraR from *Sinorhizobium fredii* (PDB_ID 2Q0O) (Chen et al., 2007) and from *Agrobacterium tumefaciens* (PDB_ID 1H0M) (Vannini et al., 2002), SdiA from *E. coli* (PDB_ID 4LFU) (Kim et al., 2014), CviR from C. violaceum (PDB_ID 3QP5) (Chen et al., 2011), QscR from *P. aeruginosa* (PDB_ID 3SZT) (Lintz et al., 2011).

The top-scored model of the prototype of sub-group F (having the predicted LGscores and MaxSub values of 4.052 and 0.931, respectively) was obtained by RaptorX (Wang et al., 2016), based on CviR from *C. violaceum* (PDB_ID 3QP5) (Chen et al., 2011), as a template.

The top-scored model of the prototype of sub-group G (having the predicted LGscores and MaxSub values of 4.062 and 0.811, respectively) was obtained by M4T (Fernandez-Fuentes et al., 2007), based on CviR from *C. violaceum* (PDB_ID 3QP5 and 3QP6) (Chen et al., 2011), as templates.

The top-scored model of the prototype of sub-group H, Ppu16R_2, (having the predicted LGscores and MaxSub values of 4.556 and 0.560, respectively) was obtained by RaptorX (Wang et al., 2016), based on TraR from *S. fredii* (PDB_ID 2Q0O) (Chen et al., 2007), as a template.

The top-scored model of the prototype of sub-group I (having the predicted LGscores and MaxSub values of 4.037 and 0.737, respectively) was obtained by Phyre2 (Kelley et al., 2015), based on SdiA from *E. coli* (PDB_ID 4LFU) (Kim et al., 2014), as a template.

Sequence alignment was performed by Expresso (Armougom et al., 2006), exploiting structural aligners algorithms like SAP (Taylor, 2000) or TMalign (Zhang and Skolnick, 2005). Each sub-group prototype was also aligned with all the canonical QS LuxR proteins, whose X-ray strutures are available: TraR from *A. tumefaciens* (PDB_ID 1H0M (Vannini et al., 2002)) and from *S. fredii* NGR234 (PDB_ID 2Q0O (Chen et al., 2007)), LasR (PDB_ID 3IX3 (Zou and Nair, 2009)) and QscR (PDB_ID 3SZT (Lintz et al., 2011)) from *P. aeruginosa*, CviR from ***C. violaceum*** (PDB_ID 3QP1 (Chen et al., 2011)) and SdiA from ***E. coli*** (PDB_ID 4Y13) (Nguyen et al., 2015)) . The structure-based homology model of OryR from *X. oryzae* (Covaceuszach et al., 2013), the prototype of the sub-group C, was also included in the structural-based multiple alignment.

3.3 Results

3.3.1 QS LuxRs and LuxR solos in the genomes of environmental fluorescent pseudomonads

In order to investigate the presence, distribution and conservation of QS LuxRs among the *P. fluorescens* complex, a systematic bioinformatic analysis has been performed. A collection of 601 sequenced genomes belonging to 17 different fluorescent pseudomonads species were sourced from the PATRIC database (Wattam et al., 2014) and analyzed to identify putative LuxR solos, according to the criteria described in the Materials and Methods section. All potential QS LuxRs and LuxR solos identified contained the typical two signature Pfam domains; PF03472 autoind_bind domain at N-terminal and PF00196 DNA-binding HTH domain at C-terminal. However, the nine signature conserved residues (six key amino acids in the inducer-binding domain and three key amino acids in DNA-binding domain) found in archetypical QS LuxRs were not all present in many of the LuxR solos detected.

In total, 651 QS LuxR protein sequence hits were identified consisting of 528 LuxR solos and 123 LuxR proteins that are part of 122 complete QS systems (one system had a LuxR-LuxI-LuxR configuration). Out of 601 fluorescent *Pseudomonas* genomes analyzed, only 87 genomes (14.5%) contained complete QS LuxI/R systems (a few genomes had multiple complete QS systems). On the other hand, more than half (approximately 50.5%; 303 genomes) harbor at least one LuxR solo, while only 8.9% of the genomes (55 genomes of 601 total) contains both LuxR solos and a complete QS system(s) (**Supplementary Table S3.2**). In approximately 35% of genomes, it was not detected either a complete QS LuxI/R system nor a LuxR solo.

The vast majority of fluorescent *Pseudomonas* genomes carried more than one copy of a QS *luxR solo* gene. In this regard, the most varied distribution was found in strains belonging to *P. fluorescens* and *P. putida,* which could carry up to four *luxR* solo genes (**Supplementary Table S3.2**). Overall, these observations show that it is much more common for fluorescent pseudomonads to harbor LuxR solos proteins rather than complete QS system(s).

It was also of interest to further analyze the conservation and distribution of the QS LuxR solos among the fluorescent pseudomonads isolated from plant roots. For this purpose, several fluorescent pseudomonads strains have been isolated from the rhizosphere of rice plants, as described in the Materials and Methods section. The complete genome of twenty strains has been determined and mined for QS LuxR solos. None of the genomes carried a complete QS *luxI/R* system(s), whereas all harbored one or multiple *luxR* solos (**Supplementary Table S3.3**). This observation suggested a clear trend for the occurrence of LuxR solos, which could play a role in adaptation in the plant root habitat. In summary, this analysis of 621 fluorescent pseudomonads (601 genomes downloaded from PATRIC and 20 genomes sequenced in this study) highlights that QS LuxI/R systems are not abundantly present. In contrast, LuxR solos are prevalent, indicating a probable evolution away from complete AHL QS LuxI/R systems.

3.3.2 Phylogenetic analysis and functional grouping of LuxR solos of the environmental fluorescent pseudomonads

In order to determine the relatedness between the LuxR solos identified in fluorescent pseudomonads, a phylogenetic analysis was carried out, as detailed in the Materials and

Methods section. Several clades clearly emerged based on their primary structure, as evidenced by the phylogenetic tree (**Figure 3.1)**. Interestingly, these LuxR solos clades do not cluster according to the species taxonomy since several branches of the tree are formed by LuxRs solos belonging to different fluorescent pseudomonads species. In addition, multiple LuxR solos carried by the same genome do not cluster together, indicating low relatedness.

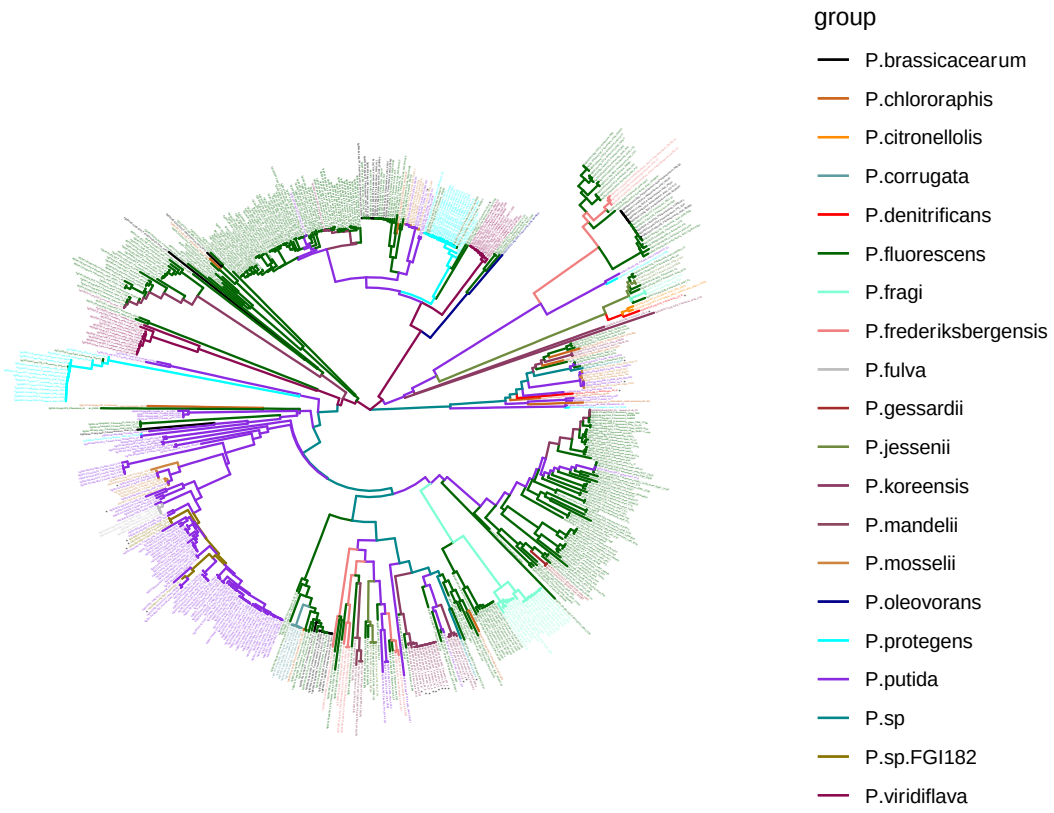


Figure 3. 1 Phylogenetic analyses of multiple LuxR solos as described in the Results section. The tree was inferred by using the Maximum Likelihood method. Colors indicate bacteria species.

It was also of interest to classify closely related LuxR solos into putative functional groups in order to further understand their relatedness and gain insights on their biological role. For this reason, the analysis of the genomic context flanking each LuxR solo was performed, since in bacteria, it is common that adjacent loci are targets for the transcriptional regulators. The primary structure and adjacent loci allowed LuxR solos to be divided into nine different sub-groups **(Figure 3.2)**. Co-members of the sub-groups are likely orthologs and functionally related, and almost all identified putative LuxR solos could be placed into these nine sub-groups.

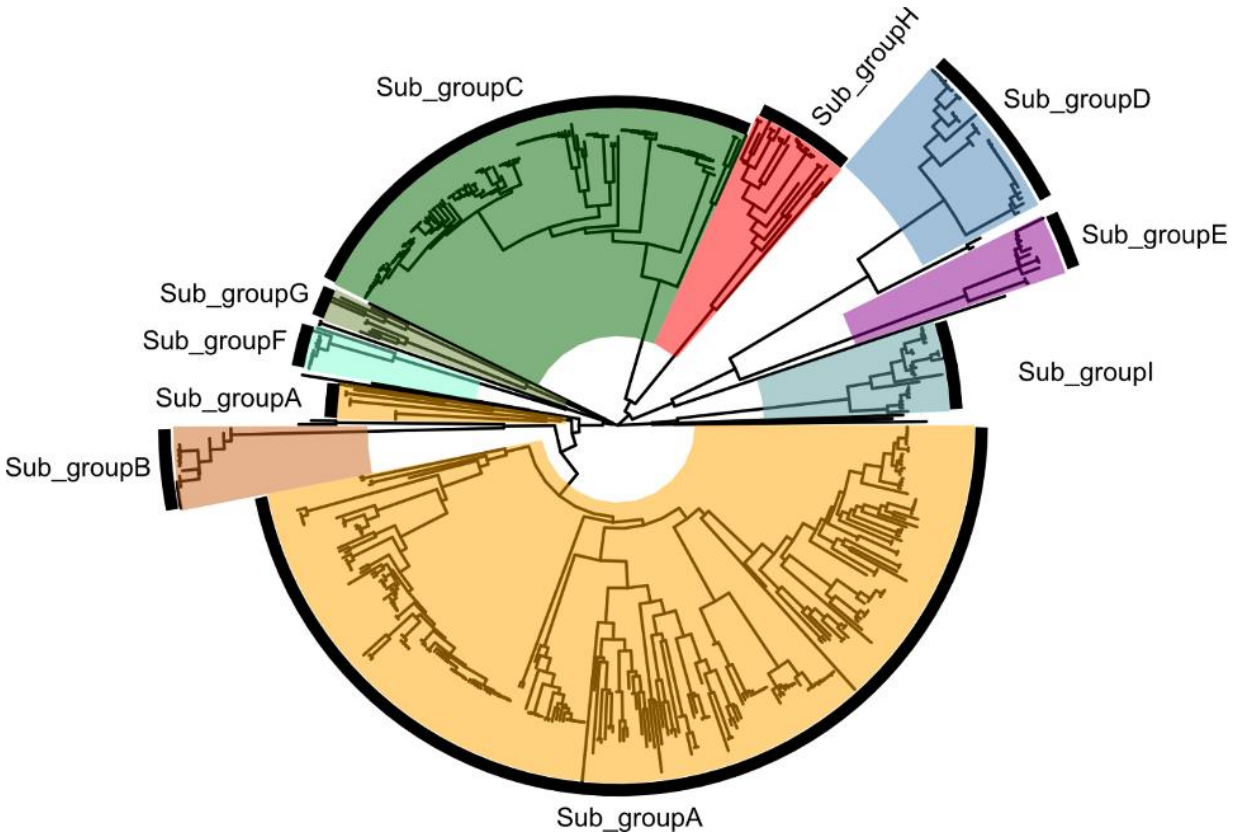


Figure 3. 2 Phylogenetic analysis and functional grouping of 528 LuxR solos carried by fluorescent *Pseudomonas.* Sub-groups are highlighted with a different colored background. LuxR solos which did not fit in any of the sub-groups are not labeled.

Only a few remained ungrouped, showing unique primary structure and flanking gene context **(Figure 3.3)**

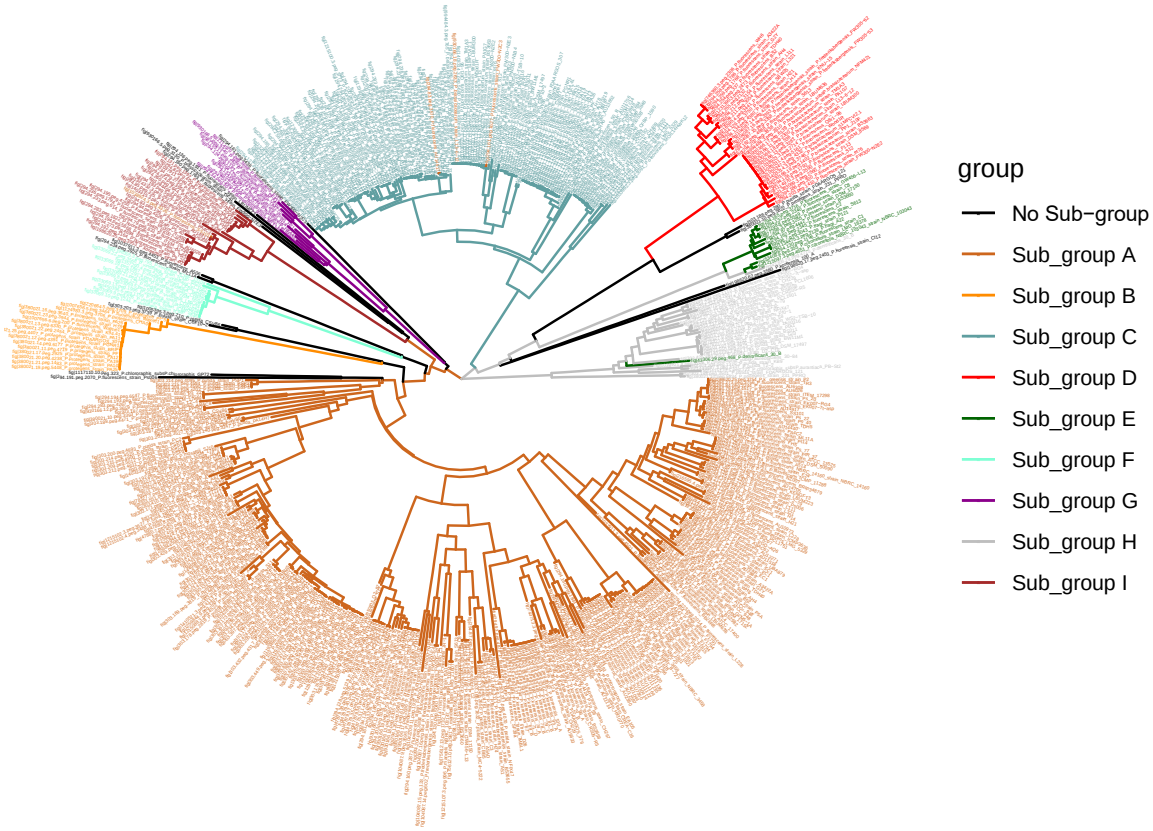


Figure 3. 3 The same phylogenetic analysis presented in the Figure 3.1. In this phylogenetic tree, each LuxR solo hit is named with the full ID and highlighted with a different color according to the sub-group they belong to, as described in the Results section. LuxR solos which did not fit in any of the sub-groups are labeled in black.

LuxR solos were (i) highly conserved within the sub-groups B and F (sequence homology between 99.5%-100%), (ii) medium conserved within the sub-groups D, E, H, and I (75%-90%), and (iii) low conserved within sub-groups A and G (31%-52%). LuxR solos belonging to different sub-groups showed a sequence relatedness around 10-25% (**Table 3.2**).

Table 3. 2 Overview of the primary structure homologies among the different sub-groups of LuxR solos using the pairwise comparison.

		1	2	3	4	5	6	7	8	9
Sub_groupA	1	100.00	25.88	22.52	21.21	23.31	17.83	13.79	13.69	15.05
Sub_groupG	2	25.88	100.00	16.85	17.31	21.92	18.21	11.30	13.79	14.64
Sub_groupI	3	22.52	16.85	100.00	18.63	18.52	15.57	14.24	10.22	14.49
Sub_groupH	4	21.21	17.31	18.63	100.00	16.60	17.38	10.10	11.41	12.95
Sub_groupF	5	23.31	21.92	18.52	16.60	100.00	13.98	12.67	14.51	15.19
Sub_groupC	6	17.83	18.21	15.57	17.38	13.98	100.00	17.43	15.49	12.15
Sub_groupD	7	13.79	11.30	14.24	10.10	12.67	17.43	100.00	15.11	14.74
Sub_groupE	8	13.69	13.79	10.22	11.41	14.51	15.49	15.11	100.00	9.47
Sub_groupB	9	15.05	14.64	14.49	12.95	15.19	12.15	14.74	9.47	100.00

As previously mentioned, LuxR solos belonging to the same sub-group were found in different taxonomic clades of the phylogenetic tree; the nine LuxR solos sub-groups are discussed below.

Sub-group A

LuxR solos of this sub-group occur in almost all the fluorescent *Pseudomonas* species analyzed here. Two very conserved genes always flank these LuxR solos, a) encoding for a ferrodoxin-NADP$^+$ reductase and b) encoding for a 23S rRNA methyltransferase; for this reason, it is likely that these adjacent loci are functionally associated with the flanking *luxR solo* (**Figure 3.4**). Either the ferrodoxin-NADP$^+$ reductase or the 23S rRNA methyltransferase

are involved in primary metabolism participating in a wide variety of redox metabolic pathways, suggesting a possible role for the LuxR solo in regulating a broad range of key metabolic functions. This LuxR solo and the adjacent loci are also highly conserved in all the twenty rice rhizosphere genomes isolated and sequenced in this study (**Figure 3.4** and **Figure 3.3**)

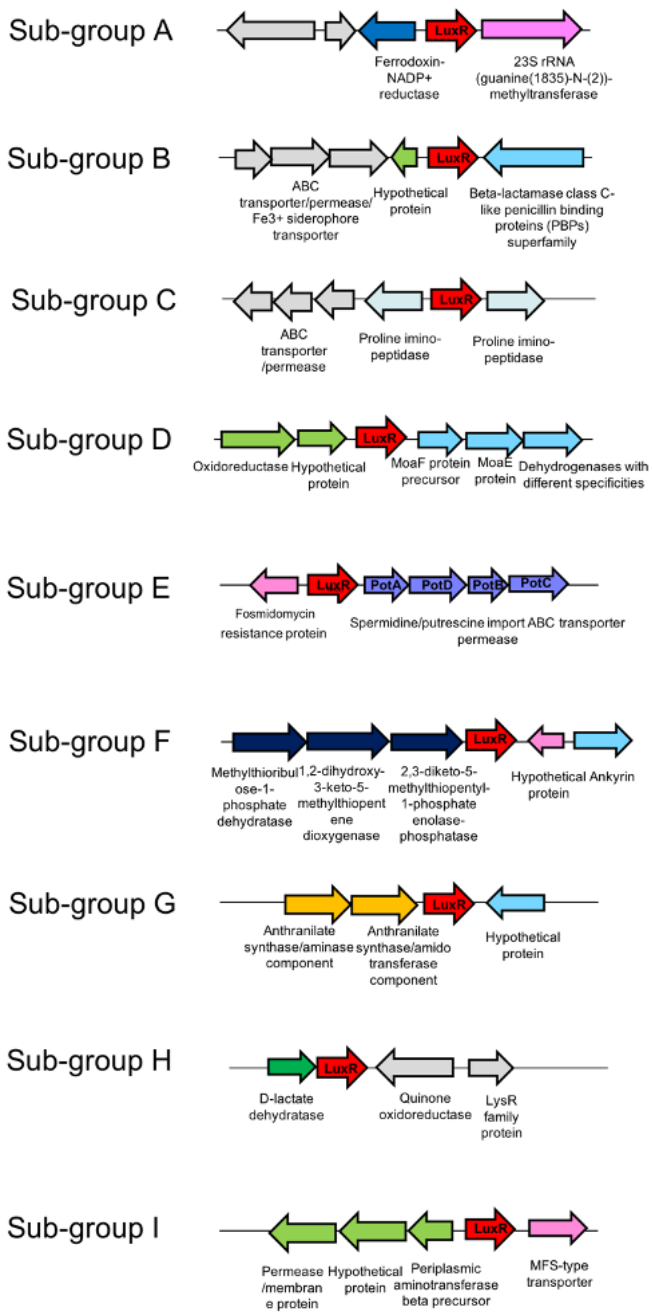


Figure 3. 4 Functional grouping of LuxR solos and genomic context (5kb).

Sub-group B

The sub-group B is only found in *P. protegens* species. The neighboring genes are beta-lactamase class C-like protein on one side and a hypothetical protein of unknown function on the other side (**Figure 3.3; Figure 3.4**).

Sub-group C

This sub-group of LuxR solos is well-studied and is often referred to as PAB LuxR solos that respond to plant low-molecular-weight compounds. They are found in many different species of plant-associated bacteria (Mosquito et al., 2020); examples are OryR and XccR, which are found in *Xanthomonas* plant pathogens, and PipR and PpoR, which are harbored in plant beneficial *Pseudomonas* sp. (Ferluga et al., 2007; Ferluga and Venturi, 2009; Subramoni and Venturi, 2009a; Coutinho et al., 2018). These LuxR solos show some substitutions among the highly conserved amino acids residues in the IBD binding pocket and regulate the adjacently located proline iminopeptidase (**pip**) gene. By responding to plant compound(s), these LuxR solos constitute an interkingdom signaling circuit, involved in plant-bacteria interaction (González and Venturi, 2013).

Sub-group D

This is a small sub-group, which is not frequent among *Pseudomonas* species. These LuxR solos are flanked by two operons with hypothetical functions, most probably involved in primary metabolism. One operon consisting of an oxidoreductase and a hypothetical protein and the other encoding for Moa-like proteins which are likely to be involved in the biosynthesis of the molybdopterin cofactor (MoCo) that is fundamental for the activity of many important enzymes processes (Schwarz and Mendel, 2006) **(Figure 3.4).**

Sub-group E

These LuxR solos are harbored in several different fluorescent pseudomonads species **(Figure 3.3**) and are flanked by genes involved in polyamines membrane transport. Polyamines are aliphatic polycationic molecules (i.e. spermidine, spermine and putrescine), which are widely distributed in bacteria, plants and animals and have been implicated as signaling molecules not only between microorganisms, but also in the interkingdom cell-cell communication (Shi et al., 2019);(Walden et al., 1997; Igarashi and Kashiwagi, 2000). This group of LuxR solos might therefore be involved in the response to polyamine molecules and possibly in plant-bacteria communication.

Sub-group F

This sub-group was limited to the *P. viridiflava* species, possibly suggesting a very specific function for this LuxR solo in regulating currently unknown mechanisms for its lifestyle. The adjacent loci consist of an upstream operon of three genes involved in the L-methionine biosynthesis pathway and a downstream gene coding for an ankyrin-type protein **(Figure 3.3 and 3.4).**

Sub-group G

This sub-group is characterized by the presence of an adjacent operon of two genes encoding for the anthranilate synthase enzymes, which are involved in the phenylalanine/tyrosine metabolism **(Figure 3.4)**. These enzymes catalyze the conversion of chorismate into anthranilate, the biosynthetic precursor of both tryptophan and numerous other secondary metabolites. Thus, it is a possibility that the operon flanking this LuxR solo might be involved in the synthesis of signal molecules.

Sub-group H

This sub-family consists of the *luxR* solo as part of an operon with a D-lactate dehydrogenase gene that encodes for an enzyme, which belongs to the oxidoreductase family and participates in pyruvate metabolism. This sub-group has been found in a small number of *Pseudomonas* species **(Figure 3.3 and 3.4)**.

Sub-group I

This sub-group is formed by LuxR solos that are located adjacent to two different loci, upstream and downstream, that both encode for transporter or permease proteins **(Figure 3.4)**. It is therefore possible that these LuxR solos regulate loci that affect the movement of compounds/molecules through the bacterial membrane.

In summary, these observations revealed that LuxR solos are predominant in fluorescent pseudomonads and allowed classification into several sub-groups, based on the conservation in their primary structure and neighboring loci.

3.3.3 Comparative cartographic analysis of the identified sub-groups of LuxR solos in fluorescent pseudomonads

In order to gain insights into the architecture of the LuxR solos inducer binding pockets and their signal specificity, a cartographic analysis of the selected solos was applied, based on structure-based homology modeling and structural superimposition, combined with multiple structure-based sequence alignments. Previous studies have found that signal specificity could be altered by specific substitutions of conserved amino acids within the inducer binding

domain (IBD). In particular, the focus was on the pocket residues directly interacting with the ligand that are conserved and belong to Cluster 1 and Cluster 2 (colored in green and in cyan, respectively in **Figure 3.5**) as previously described (Covaceuszach et al., 2013). Residues of the third Cluster, belonging to variable patch and thus being poorly conserved even within members of QS LuxRs, have not been taken into account.

```
                                                 v   v   sv  sv            cccs
                                                         57  61            70  73
1H0M_TraR              DHFGFTGYAYLHI--------QHRHITAVTNYHRQWQSTYF----DKKFEALDPVVK--RARS     78
2Q0O_TraR              HSCGYDRFAYLQK--------DGTQVRTFHSYPGPWESIYL----GSDYFNIDPVLA--EAKR     80
3QP1_CviR              SQAPSERLLLALGRLNNQNQIQRLERVLNVSYPSDWLDQYM----KENYAQHDPILR---IHL    100
2IX3_LasR              SDLGFSKILFGLL---PKDSQDYENAFIVGNYPAAWREHYD----RAGYARVDPTVS--HCTQ     81
3SZT_QscR              GNYGFEFFSFGAR---APFPLTAPKYHFLSNYPGEWKSRYI----SEDYTSIDPIVR--HGLL     83
4LGW_SdiA              QQLEYDYYSLCVR---HPVPFTRPKVAFYTNYPEAWVSYYQ----AKNFLAIDPVLN--PENF     88
Sub-group A_Ppu16R     QALNMEFVGLALH---LHVAARGPQVILYNNYPGAWNERYQ----AEDLIKIDPTVS--KCHH     83
Sub-group B            RETNHTEFSFILR---SDLKSHSKVVYNIDSYSYAWRRDYE----AEKFWIVDPILKI-HSPT     86
Sub-group C_OryR       QSLVYDYAPVPLS---MEGALITPTVFMQRNAPGDMQHVWC----EHGYYQHDPVQQ--RATR     88
Sub-group D            QLVPVSRAAFYCV-----DRQLQVHDFSLHKMSGEMHRDYL-----DNYRQFDPLHPRHCVSS     71
Sub-group E_PjeR       --CPVDHFVLFVY-----EGNHR-PLALFDTFSADKRAVYVDDY-QCGPYLLDPFYL--ACTR     88
Sub-group E_PolR       KALIYDYAPVPLS---HEGVLITPSVFKVRNAPADMQAFWC----ERGYYQIDPVQQ--CASR     88
Sub-group F            RAMFYQ------------FKGVQLLTFVTSNVDDVWGEAYR----RQQMISHDPVVR--HYRN     90
Sub-group G            RQLGFSSYLISLGP--ATPRDTDPLADLITTFPKRWLDQYR----GDGLIDIDPILR--HCRR     75
Sub-group H_Ppu16R_2   FVLFSASAARDEV---------VERIHWVEGDWFGDGQVVDA----RTYVRHCPVTR--HLLA     85
Sub-group I            AELSRRGLGHFEV---YQCGPTLDEAEIITNYPDALDRGDA----ADDFYQRGHLLS--LSKR     77
                                                     r   r            dddd
                                c                   s   s           s  c                c
                                85                 101 105         110 113             129
1H0M_TraR              RKHIFTWSGEHERPT-----LSKDERAFYDHASDFG------IRSGITIPIKTANGFMSMF-T    129
2Q0O_TraR              RRDVFFWTADAWPAR-----GSSPLRRFRDEAISHG------IRCGVTIPVEGSYGSAMML-T    131
3QP1_CviR              GQGPVMWEERFNRA------KGAEEKRFIAEATQNG------MGSGITFSAASERNNIGSILS    150
2IX3_LasR              SVLPIFWEPSIY--------QTRKQHEFFEEASAAG------LVYGLTMPLHGARGELGAL-S    129
3SZT_QscR              EYTPLIWNGEDF----------QENRFFWEEALHHG------IRHGWSIPVRGKYGLISML-S    129
4LGW_SdiA              SQGHLMWNDDLF----------SEAQPLWEAARAHG------LRRGVTQYLMLPERALGFL-S    134
Sub-group A_Ppu16R     TTLPLVWNDDLY----------CEVPQLREAATALG------MTHGWSQSVHDQQHNESQL-S    129
Sub-group B            PYFPTIWSDDFI-----------ESKLFLQAGKDYG------FLSGMSIGLPCSFGNYGVL-S    131
Sub-group C_OryR       RNTPFVWSYRTDGDCAGVEYVGGQHRQVTRYLCDSG------MGTGVTVPLHLPGGAFATF-S    144
Sub-group D            EVAVVPLSLAMA------RQPRRDNRRYHDFLQRYG------VVDVVEVFAHRDGQPQAAI-S    121
Sub-group E_PjeR       GQAPGLWRLRQFAPD-----HFYLGEYFLTYYQQTG------LAEEVAFFVDLGGGAMAVL-S    138
Sub-group E_PolR       RAAPFVWSFRSETRSALREYLGDSHRPVTCYVRDSG------LATGVTVPLHLPGGAFATL-T    144
Sub-group F            NLGFLEWDDAIR--------HYLPPAWYEETLSACN--LAPAISYGYTSHCRGVDGVTSIF-T    142
Sub-group G            ELVPFFWDRERR-------RARGRSRHFWQQRELYG------LRSGLSIPLRYELL-RGTL-S    123
Sub-group H_Ppu16R_2   AREAFFWSKQPG--------DEGERYRVVRLPGGPG------I-HGLQVPVFGPAGLEGAM-S    132
Sub-group I            RIKPFFWSDENI--------LAPAAEAVPVPGPCAP------VGEGASFIVHGNQGFYSIL-N    125
                                f                   r   r           p  f                f
```

Figure 3. 5 Structure-based multiple sequence alignment of the inducer binding domains of the prototypes of the nine identified LuxR solo sub-groups with QS LuxRs. The residues belonging to Cluster 1 and Cluster 2 are highlighted in green and cyan, respectively. The 3D architecture of the boundaries of the ligand-binding site is schematized by r (roof), f (floor), p (proximal wall) and d (distal wall) and its tripartite topology by c (conserved core), s (specificity patch) and v (variable patch).

Ten LuxR solos that represent each of the nine sub-groups discussed above were selected and the molecular determinants of each inducer-binding site were analyzed. This analysis revealed key differences between the binding sites among the representatives of each sub-group (**Figure 3.5 and Table 3.3**): all the comparisons were paralleled to TraR from *A.*

tumefaciens, as the prototype of QS LuxR proteins. According to the molecular cartography and structure-based alignment, only two subgroups (A and B) are very closely related to the archetypical QS LuxRs. They maintained the two conserved hydrogen bonds stabilizing AHL binding (**Figure 3.5 and Table 3.3**), namely one between the ε nitrogen of W57 (according to TraR numbering) and the carbonyl oxygen of the lactone moiety and the second between the ε oxygen of D70 and the nitrogen preceding the acyl moiety. In addition, all the apolar residues belonging to the conserved and specificity patches, which further stabilize the AHL binding by hydrophobic interactions, are maintained with respect to the AHL-binding template, except for one substitution of a residue with similar steric hindrance (L→M). Overall, the cartographic observations suggest that these two sub-groups of LuxR solos very likely bind and respond to AHLs.

Interestingly, also the binding pockets of the members of sub-group F and G are characterized by an overall conservation in all the residues of the conserved and specificity patches, except for the amino acid corresponding to A105 of TraR. The substitutions of this small side chain with residues characterized by much higher steric hindrance (A→L and A→R in sub-groups F and G, respectively) deeply impacts on the shape of the binding sites, partially occluding the hydrophobic pocket in which the lactone ring accommodates. This effect due to a single substitution is anyway very likely to alter sub-groups F and G ligand specificity respect to canonical AHL binding LuxRs.

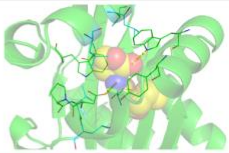
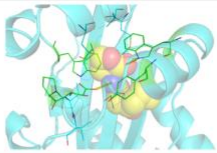
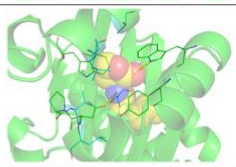
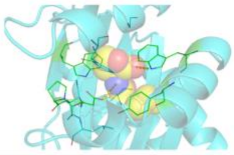
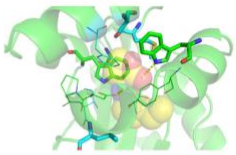
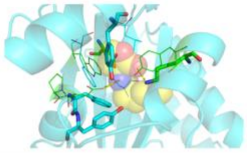
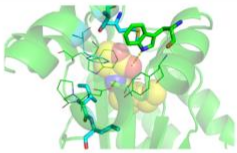
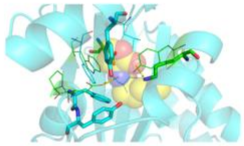
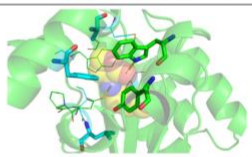
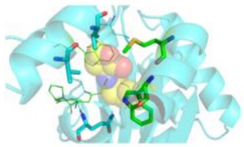
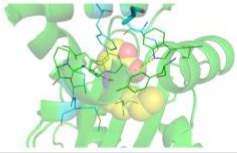
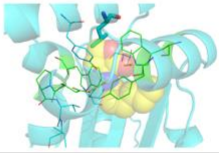
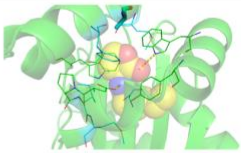
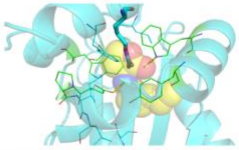
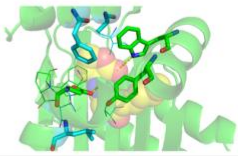
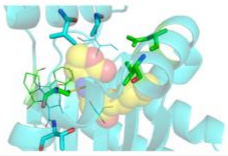
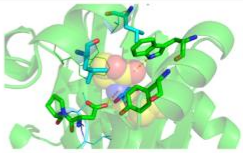
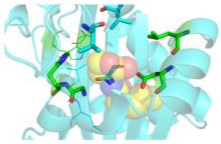
Functional LuxR solo group	template	model
A		
B		
C		
D		
E		
F		
G		
H		
I		

Table 3. 3 Comparison of the inducer-binding sites of the prototypes of the nine identified LuxR solo sub-groups (right column) with the corresponding QS LuxRs templates used for their in silico modelling (left column). Semitransparent cartoon representation, with the side chains of residues belonging to Cluster 1 and Cluster 2 highlighted in green and cyan, respectively: conserved residues are represented by lines while non conserved amino acids are highlighted by sticks. The bound AHL is represented by spheres and its carbon, nitrogen and oxygen atoms are colored in yellow, blue and red respectively. The hydrogen bonds stabilizing the lactone ring binding are highlighted by yellow dotted lines. Figures produced by Pymol (Schrödinger, 2010)

The members of sub-group C, which has been already identified as a member of the PAB LuxR solos group that respond to plant compounds, as previously reported, showed replacement

of amino acids at positions corresponding to the following residues of TraR: W57 (→M, leading to the loss of one of the stabilizing hydrogen bonds); V73 (→Q, impacting on the hydrophobic environment of the cleft); Y61 (→W), F101 (→V) and A105 (→L), which generate different steric hindrance likely altering the shape of the pocket. Overall, these key differences suggest a different specificity towards what is believed to be plant compound(s) (González et al., 2013; Coutinho et al., 2018).

Surprisingly, the remaining groups showed significant modifications in the binding pocket, due to several changes not only in the specificity patch, but also in the invariant amino acids of the conserved patch, thus suggesting that these proteins likely bind other non-AHL compounds (**Table 3.3**)**.** In particular, all the candidates from these latter groups have lost at least one of the two hydrogen bonds stabilizing AHL binding, due to substitutions not only in the residue corresponding to W57 of TraR, which is part of the specificity patch and is not conserved in all the 4 sub-groups, but also in the very conserved amino acid corresponding to D70 of TraR, namely in sub-groups H and I. Additional invariant positions that are not conserved in these sub-groups are the ones corresponding to TraR V72 (→F in PjerR of sub-group E, leading to increased steric hindrance), W85 (→L in sub-group D, leading to decreased steric hindrance), G113 (→E in PjerR of sub-group E and →V in sub-group D, leading to increased steric hindrance that is also combined to a huge variation in the electrostatics of the pocket in the case of PjeR). Regarding the specificity patch, the residues at almost all the positions are substituted with amino acids with side chains that have entirely different steric hindrance and moreover are charged or polar, profoundly impacting not only on the overall shape but also on the hydrophobicity of the pocket that is a prerequisite for AHL binding (**Figure 3.5**). Therefore, these sub-groups of LuxR solos appear to be more distantly related

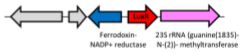
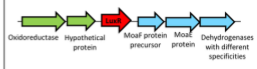
to the canonical QS LuxRs and are possibly able to respond to yet-unknown exogenous or endogenous compounds.

In summary, the cartographic analysis showed variable degrees of conservation in the amino acids forming the binding pocket among the LuxR solos of the fluorescent pseudomonads. Thus, the main hypothesis was that some eavesdrop by binding AHLs, whereas others could have evolved to bind different compounds/signals produced by neighboring species or possibly to currently unknown endogenous signals.

3.3.4 Potential target gene promoter expression analysis of a set of LuxR solos

In order to begin to acquire insights into the mode of action of the LuxR solos, *luxR* solo autoregulation and gene expression studies of the flanking loci were performed. LuxR solos candidates from sub-groups A, D, E and H were selected as they showed some interesting features according to our analyses (Table 3.**4).** The *luxR* solo genes were mutated in fluorescent pseudomonad strains as described in the Material and Methods section and the **Supplementary Table 1**. The transcription of the various loci was studied via gene promoters transcriptionally fused to a lacZ reporter gene in a plasmid construct and assays were performed in the wild type and luxR solo mutant strains. Below the results of these studies on five LuxR solos that belong to the four different sub-groups are presented.

Table 3. 4 LuxR solos candidates for target gene promoter expression analyses.

LuxR solo ID	Organism	NCBI accession number	Functional LuxR solo group	In silico modeling	Genetic locus
Ppu16R	*P. putida 16_A*	This study	A	Probable respond to AHLs	
Ppu16R_2	*P. putida 16_A*	This study	H	Unlikely respond to AHLs	
PfluR_113	*P. fluorescens F113*	CP003150	D	Unlikely respond to AHLs	
PjeR	*P. jessenii DSM 17150*	NIWT01000008	E	Unlikely respond to AHLs	
PolR	*P. oleovorans AG1003*	This study	E	Probable respond to plant-compounds	

Ppu16R of sub-group A

Ppu16R of P. putida 16A is highly identical in its IBD to QS-LuxRs and cartographic analysis predicted that it very likely binds and responds to AHLs. Therefore, it was of interest to study its auto-expression and that of the adjacent genes in the presence/absence of AHLs. Moreover, mining the genome of P. putida 16A revealed that it does not possess any canonical AHL-QS LuxI/R systems, suggesting Ppu16R could be responding to exogenous AHLs. No Ppu16R-dependent promoter activities were detected either in the presence or absence of AHLs under the conditions tested here (Figure 3.6). One possible explanation is that the Ppu16R does not autoregulate its expression and that flanking genetic loci are not its targets or the conditions used in this study do not allow for activating/repressing the expression of these genes**.**

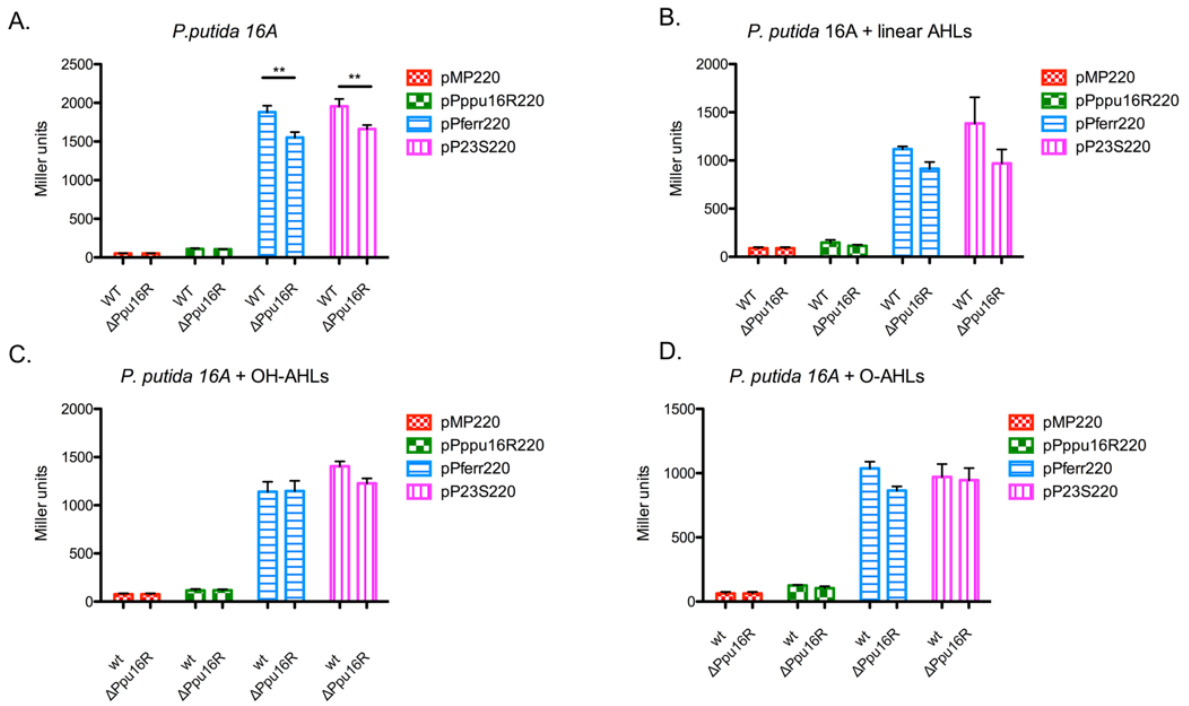


Figure 3. 6 Gene promoter activity in the presence or absence of AHLs in P. putida 16A WT and Δ*Ppu16R.* β-Galactosidase activities (Miller units) of 3 gene promoter transcriptional fusion (*ppu16R, ferrodoxin NADP+ reductase, 23S methyltransferase*) were determined to compare the expression levels between WT and Δ*Ppu16R.* The WT and Δ*Ppu16R* strain with empty plasmid pMP220 was used as control. The AHLs used are the follow: linear AHLs (C_4-AHL;C_6-AHL;C_8-AHL;C_{10}-AHL;C_{12}-AHL), OH- AHLs (OHC_6-AHL;OHC_8-AHL;OHC_{12}-AHL) and O-AHLs (OC_6-AHL;OC_8-AHL;OC_{12}-AHL). All experiments were performed in triplicate. Statistical analysis was calculated using one-way ANOVA followed by Tukey's multiple comparisons by Prism 7 (GraphPad Software, Inc.). The error bars indicate standard deviations.

To further investigate whether Ppu16R could bind AHLs, his-tagged Ppu16R was recombinantly expressed in E. coli in the presence of different AHLs as most commonly, AHL-binding QS LuxRs are stabilized and solubilized when bound to AHLs (Zhu and Winans, 2001). His-tagged Ppu16R resulted in being soluble in the absence of AHLs and the presence of AHLs did not increase solubility (data not shown), not allowing a direct readout of AHL binding (**Figure 3.7**). This LuxR protein solubility independent of AHLs was also observed for the SdiA LuxR from *E. coli* (Nguyen et al., 2015). In summary, these studies have not provided direct evidence for gene targets and AHL binding for this LuxR solo.

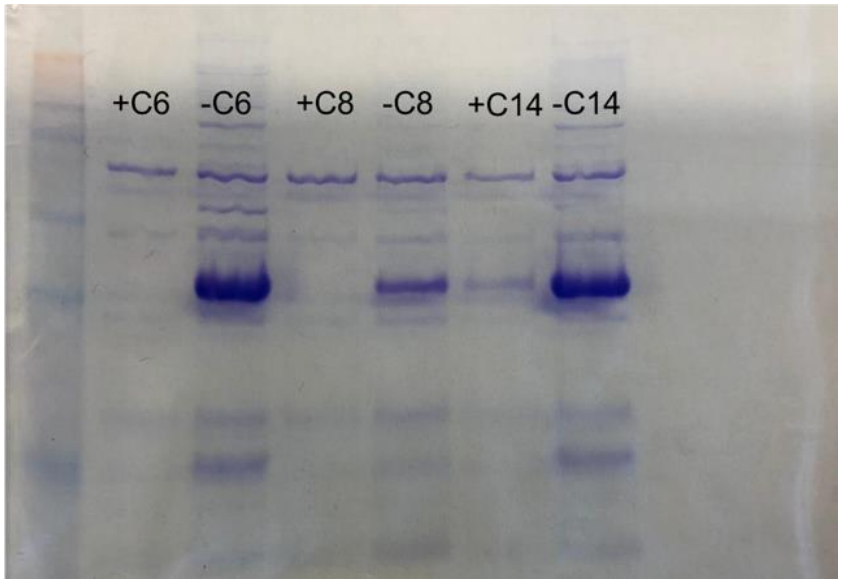


Figure 3. 7 His-tagged *Ppu16R* expression in pETM-11 adding 100mM of each AHL (C6-AHL; C8-AHL; C14-AHL). Soluble fractions purified using MagneHis Protein Purification System (Promega, Corp., Madison, WI, USA) is shown in the protein gel; *Ppu16R* was soluble when unbound to AHLs.

PfluR113 of sub-group D

The PfluR113 solo of *P. fluorescens* F113 belonged to sub-group D, and according to cartographic analysis, it most probably does not bind AHL signals. In addition, this strain does not possess any canonical AHL-QS LuxI/R systems. To understand whether there was autoregulation and whether adjacent operons were regulated by the nearby solo gene, the transcriptional activity of *pfluR113* and the adjacent operons was determined. This established that PfluR113 negatively regulated the transcription of one of the genetically linked operons. A significant increase of the expression of the operon in the Δ*pfluR_113* mutant was determined when the bacterial culture was in an early log-phase, while no significant differences were detected in the stationary phase (**Figure 3.8**). Complementation of the Δ*pfluR_113,* via the wild-type gene harbored in a plasmid, restored the expression levels observed in the wild type strain in the early log-phase. This suggested that PfluR113 plays a role in the growth-phase dependent regulation of the adjacent operon and that this solo may respond to some yet-uncharacterized endogenous signals/molecules.

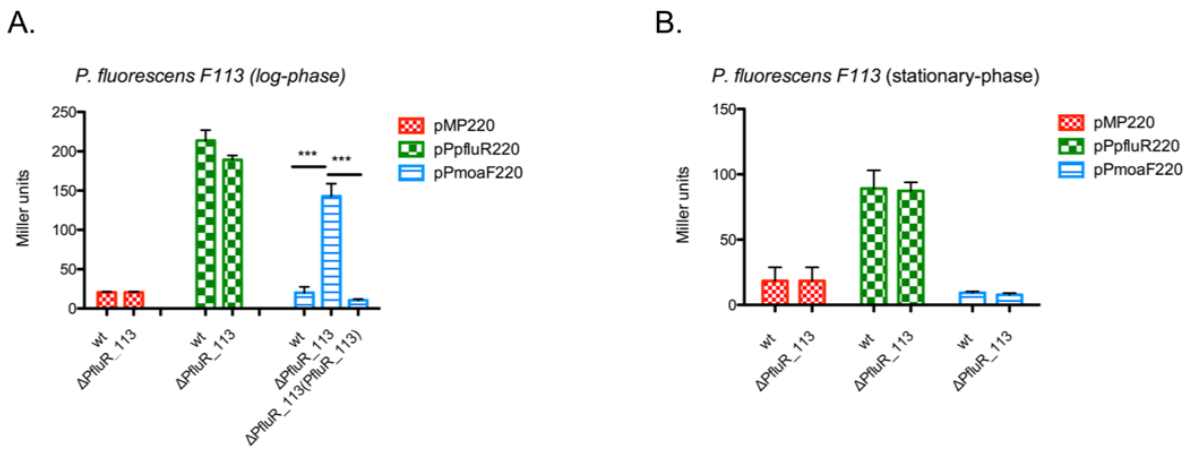


Figure 3. 8 Gene promoter activity in *P. fluorescens* F113 and Δ*pfluR*. β-Galactosidase activities (Miller units) of 2 gene promoter transcriptional fusion (p*fluR,moaF*) were determined to compare the expression levels between WT and Δ*pfluR.* The WT and Δ*pfluR* strain with empty plasmid pMP220 was used as control. The promoter activity was calculated after 4 hrs **(A)** (log-phase) and after overnight **(B)** (stationary-phase). All experiments were performed in triplicate. Statistical analysis was calculated using one-way ANOVA followed by Tukey's multiple comparisons by Prism 7 (GraphPad Software, Inc.). The error bars indicate standard deviations.

PjeR and PolR of sub-group E

It was of interest to investigate whether polyamines could bind to the LuxR solos belonging to the sub-group E, since they were flanked by genes most likely involved in transporting polyamines through the bacterial membrane. Several recent studies have shown that polyamines (i.e. putrescine, spermidine, spermine) play a role in cell-to-cell signaling regulating phenotypes such as surface motility, biofilm formation and cell differentiation (Karatan et al., 2005; Zhou et al., 2007). Moreover, according to the modeling of their ligand-binding pocket, this LuxR solo sub-group most likely responds to non-AHL molecules. As described in Table 3, the expression of *pjeR* from *P. jessenii* DSM 17150 and of the adjacent putative spermidine transporter gene was tested. Similarly, the expression of *polR* from *P. oleovorans* AG1003 and the flanking putative putrescine importer gene was analyzed. *P. jessenii* DSM 17150 or *P. oleovorans* AG1003 do not posses any canonical AHL-QS LuxI/R

systems**.** All the promoter activities were examined in the presence or absence of (i) putrescine, (ii) spermidine and (iii) spermine. The results showed that none of these gene promoters were activated/induced under any conditions tested (**Figure 3.9A and 3.9B**).

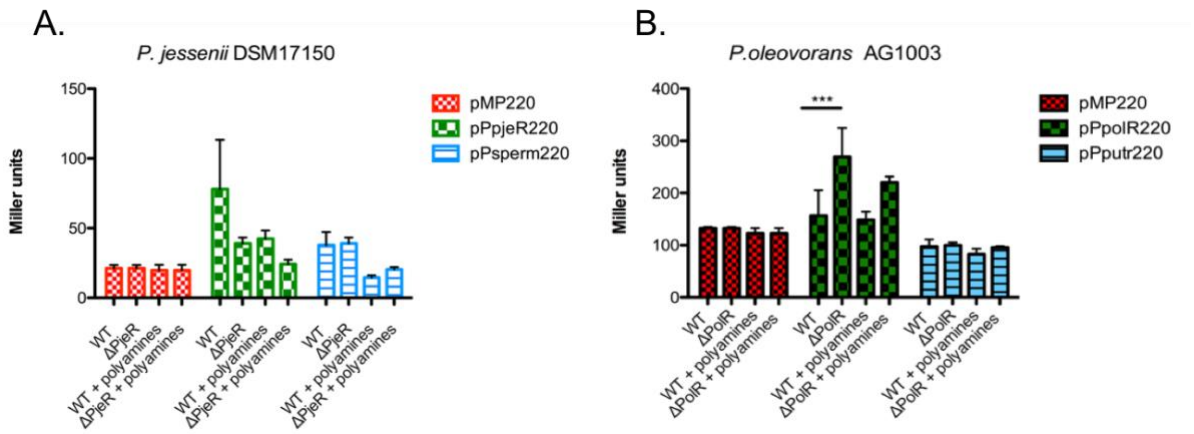


Figure 3. 9 Gene promoter activity in the presence or absence of a cocktail of polyamines (putrescine, spermine, spermidine) at a final concentration of 0.1 mM in **(A)** *P. jessenii* DSM17150 and Δ*pjeR.* β-Galactosidase activities (Miller units) of 2 gene promoter transcriptional fusion (*pjeR, spermidine transporter*) were determined to compare the expression levels between WT and Δ*pjeR.* The WT and Δ*pjeR* strain with empty plasmid pMP220 was used as control. **(B)** *P. oleovorans* AG1003 and Δ*polR.* β-Galactosidase activities (Miller units) of 2 gene promoter transcriptional fusion (*polR,putrescine importer*) were determined to compare the expression levels between WT and Δ*polR.* All experiments were performed in triplicate. Statistical analysis was calculated using one-way ANOVA followed by Tukey's multiple comparisons by Prism 7 (GraphPad Software, Inc.). The error bars indicate standard deviations.

Ppu16R2 of sub-group H

Ppu16R2 is a second LuxR solo harbored by *P. putida* 16A that constitutes an operon with the D-lactate dehydrogenase gene. In this sub-group, the operonic structure is always conserved, suggesting a potential role for this LuxR in pyruvate metabolism via the glyoxylase pathway.

The results obtained (**Figure 3.10A**) showed no *ppu16R2* autoregulation of the operon, neither in the early log-phase or stationary phase.

In summary, these studies revealed that most commonly, the luxR solos are not autoregulated and do not regulate adjacent genes under the conditions that were tested here (**Figure 3.10B**).

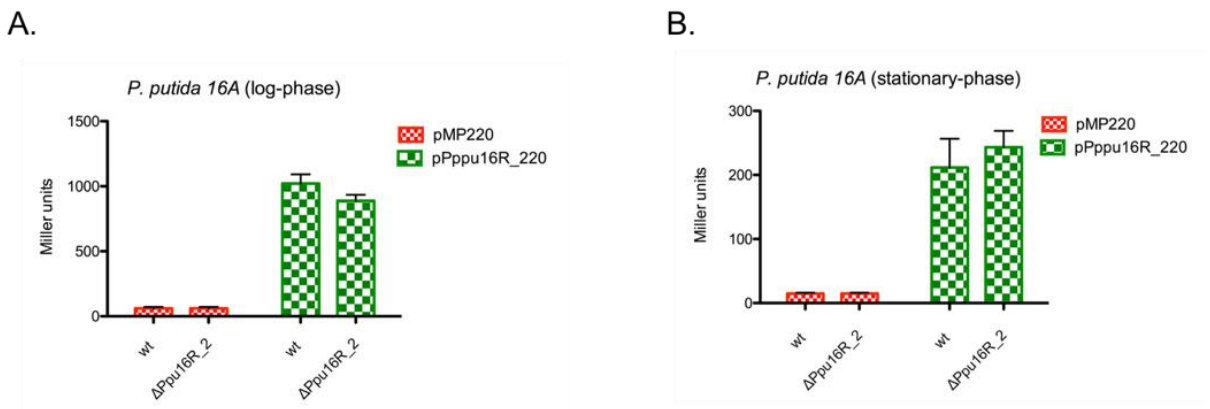


Figure 3. 10 Gene promoter activity in *P. putida* 16A and Δ*ppuR_2.* β-Galactosidase activities (Miller units) of the *ppuR_2* promoter transcriptional fusion were determined to compare the expression levels between WT and Δ*ppuR_2.* The WT and Δ*ppuR_2* strain with empty plasmid pMP220 was used as control. The promoter activity was calculated **(A)** after 4 hrs (log-phase) and **(B)** after overnight (stationary-phase). All experiments were performed in triplicate. Statistical analysis was calculated using one-way ANOVA followed by Tukey's multiple comparisons by Prism 7 (GraphPad Software, Inc.). The error bars indicate standard deviations.

3.4 Discussion

QS LuxR solos are a subfamily of QS LuxR proteins that are very widespread in protebacteria and maintain the N-terminal IBD and C-terminal HTH domain and occur without a cognate LuxI-AHL synthase. To date, only a few LuxR solos have been studied which has shown that they can be involved in intra-species, inter-species, and inter-kingdom signaling.

The main aim of this study was to investigate the distribution and conservation of LuxR solos among members of the fluorescent *Pseudomonas* group, many of which are plant commensals being studied for their biocontrol and plant growth promotion properties (Haas and Défago, 2005). Our analysis of over 600 genomes revealed that the majority of fluorescent *Pseudomonas* carry one or more LuxR solos. Based on their adjacent gene context and primary structure, they have been clustered into nine sub-groups. The modeling analysis revealed that the majority show substitutions at the invariant amino acids of the ligand-binding pocket, raising the possibility of binding to non-AHL ligands or function independently of any ligand.

Only 14.5% of the fluorescent *Pseudomonas* analyzed harbored a complete AHL QS system in its genome, whereas more than half (50.5%) harbor only *luxR solos*. This result is in line with a previous study (Subramoni et al., 2015) that demonstrated that many Gammaproteobacteria carried multiple LuxR solos, particularly plant-associated and environmental isolates. In addition, the isolation and analysis of a set of twenty rice rhizospheric *P. fluorescens* isolates further confirmed the trend for the high occurrence of LuxR solos, since among these genomes only *luxR solos* genes and no complete AHL QS systems have been identified. This result suggests a specific role for single or multiple LuxR solos in bacterial species that colonize plant-associated niches. Rhizosphere *Pseudomonas*

spp. rarely harbor a complete AHL QS system, and its lack of conservation and the unpredictable role played indicates that it is not part of the core genome (Steindler and Venturi, 2007). The absence of complete canonical LuxI/LuxR systems and the highly variable LuxR solos organization can be due to the adaptation of these bacteria to live in mixed communities and the ability to colonize several different environments. Unlike some bacterial species that harbor LuxI/R systems, which colonize specific niches upon reaching high cell densities, fluorescent pseudomonads may have increased their genetic plasticity to be part of mixed complex communities.

Based on sequence similarity, invariant amino acids of the IBD, and conservation of the flanking genes, LuxR solos were placed into putative ortholog sub-groups. The identification of a few LuxR solos which do not cluster into these sub-groups having uncommon flanking genes and primary structure suggests that other LuxR solo sub-groups exist. Nine different sub-groups of LuxR solos have been mapped here, which included the well-studied sub-group of PAB LuxR solos and the other eight uncharacterized sub-groups. Several previous studies have shown that PAB LuxR solos regulate the adjacently located *pip* gene in response to a plant compound. Members of this subfamily are characterized by few substitutions of two important amino acids in the autoinducer binding site (Ferluga et al., 2007; Zhang et al., 2007; Subramoni et al., 2011; Chatnaparat et al., 2012). Our analysis revealed that PAB LuxR solos are very widespread among *P. fluorescens* sequenced genomes, especially among *P. putida*, probably due to its role in adaptating to life next to the plants. Similarly, few members of the sub-group A, characterized by *luxR solos* flanked by two very conserved genes, encoding for a ferrodoxin-$NADP^+$ reductase and a 23S rRNA methyltransferase, have been previously described, such as PpoR from *P. putida* (Fernández-Piñar et al., 2011). These studies revealed

that PpoR plays an important role in iron acquisition however, the molecular mechanism of the response of this sub-group of LuxR solos remains unknown. The sub-group A is the most widespread among *P. fluorescens* species and could be involved in both inter- and intraspecific processes relevant to the fitness of the *P. fluorescens* bacterial group such as the control of some oxidation reactions associated to the rhizosphere, where the levels of toxic bioproducts of the aerobic metabolism of the plant are very high (Laloi et al., 2004; Das and Roychoudhury, 2014). For all the other sub-groups of LuxR solos, there are no reports on their function and response/regulation. Interestingly, it was observed a flexible rearrangement of the genomic context flanking different *luxR solos* and also a variable distribution and abundance of different sub-groups among the species. It is possible that LuxR solos with different functions were acquired by these bacteria from different sources by horizontal gene transfer or genomic rearrangement events, as it is known to be highly prevalent in many *Pseudomonas spp.* (Qiu et al., 2006; Subramoni et al., 2015; Hesse et al., 2018). LuxR solos present in the same genome showed different levels of relatedness to each other, suggesting possible different origins and also possible different ligand binding properties.

To date, there are very few functionally characterized LuxR solos with known ligands (Subramoni and Venturi, 2009a; Brachmann et al., 2013; Brameyer et al., 2014, 2015; Coutinho et al., 2018). Our modeling analysis revealed that only two sub-groups of LuxR solo are likely to bind and respond to AHL signals. One of these is sub-group A, however, our molecular and biochemical studies did not provide evidence for AHL binding. Alternatively**,** they may act independently of AHLs or may bind to different or modified AHL-like molecules produced by neighboring bacteria living in the same mixed community. As this sub-group is widespread among fluorescent Pseudomonas isolated from the rhizosphere, there could also

be a possibility of sensing AHL-like molecules produced by the plant host. Prior studies have shown that AHL availability is higher in the rhizosphere compared to the bulk soil (Bais et al., 2006); it is most likely that varying concentrations or conditions of AHLs are needed for a response by this sub-group. Moreover, it cannot be excluded that some LuxR solos can act independently without the need of an inducing ligand, as previously reported (von Bodman et al., 2003). Alternatively, it is also a possibility that a ligand molecule is endogenously produced upon a stimulus being an intracellular messenger. Additional studies are therefore required to understand the molecular mechanisms of these LuxR solo subfamilies. Non-AHL binding LuxR solo sub-groups could have evolved to respond to different signals, playing different roles in cell-cell communication or having other more classic gene regulatory mechanisms. In particular, differences in the binding pocket conformation possibly suggest different inducer specificity and could result from the adaptation and evolutionary process to colonize, compete and persist in different environments.

Our *in silico* analysis showed that several LuxR solos occur in a transcriptional unit with the neighboring genes; nevertheless, our expression analysis of promoter regions of flanking genes evidenced that most often their regulation is not under the nearby LuxR solos control. This suggests that LuxR solos could have evolved for having different target functions or the expression studies performed here could be influenced by the absence of the LuxR solo ligands/signal molecules or the environmental growth conditions were not appropriate for LuxR solo function.

In summary, this study provides a large picture of LuxR solo distribution, classification, and abundance among the fluorescent pseudomonads group. The results highlight the existence of novel LuxR solos belonging to different sub-groups that are likely to be involved in

establishing possible novel communication networks or to have others regulatory responses. LuxR solos could have evolved away from QS systems (Sabag-Daigle and Ahmer, 2012) to respond to other endogenous or exogenous signals, expanding the regulatory networks for inter-species and inter-kingdom communication. Future work needs to establish their role and the signals they respond to in the plant-associated microbiome.

3.5 References

Ahmer, B.M. (2004) Cell-to-cell signalling in *Escherichia coli* and *Salmonella enterica*. *Molecular microbiology* **52**: 933-945.

Ahmer, B.M., Van Reeuwijk, J., Timmers, C.D., Valentine, P.J., and Heffron, F. (1998) *Salmonella typhimurium* encodes an SdiA homolog, a putative quorum sensor of the LuxR family, that regulates genes on the virulence plasmid. *Journal of bacteriology* **180**: 1185-1193.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B. et al. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic acids research* **34**: W604-W608.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A. et al. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**: 1-15.

Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S., and Vivanco, J.M. (2006) The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57**: 233-266.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S. et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology* **19**: 455-477.

Bertani, I., Abbruscato, P., Piffanelli, P., Subramoni, S., and Venturi, V. (2016) Rice bacterial endophytes: isolation of a collection, identification of beneficial strains and microbiome analysis. *Environmental microbiology reports* **8**: 388-398.

Better, M., Lewis, B., Corbin, D., Ditta, G., and Helinski, D.R. (1983) Structural relationships among Rhizobium meliloti symbiotic promoters. *Cell* **35**: 479-485.

Brachmann, A.O., Brameyer, S., Kresovic, D., Hitkova, I., Kopp, Y., Manske, C. et al. (2013) Pyrones as bacterial signaling molecules. *Nature chemical biology* **9**: 573-578.

Brameyer, S., Kresovic, D., Bode, H.B., and Heermann, R. (2014) LuxR solos in *Photorhabdus* species. *Frontiers in cellular and infection microbiology* **4**: 166.

Brameyer, S., Kresovic, D., Bode, H.B., and Heermann, R. (2015) Dialkylresorcinols as bacterial signaling molecules. *Proceedings of the National Academy of Sciences* **112**: 572-577.

Buchan, D.W., and Jones, D.T. (2019) The PSIPRED protein analysis workbench: 20 years on. *Nucleic acids research* **47**: W402-W407.

Case, R.J., Labbate, M., and Kjelleberg, S. (2008) AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. *The ISME journal* **2**: 345-349.

Chatnaparat, T., Prathuangwong, S., Ionescu, M., and Lindow, S.E. (2012) XagR, a LuxR homolog, contributes to the virulence of *Xanthomonas axonopodis pv. glycines* to soybean. *Molecular plant-microbe interactions* **25**: 1104-1117.

Chen, G., Jeffrey, P.D., Fuqua, C., Shi, Y., and Chen, L. (2007) Structural basis for antiactivation in bacterial quorum sensing. *Proceedings of the National Academy of Sciences* **104**: 16474-16479.

Chen, G., Swem, L.R., Swem, D.L., Stauff, D.L., O'Loughlin, C.T., Jeffrey, P.D. et al. (2011) A strategy for antagonizing quorum sensing. *Molecular cell* **42**: 199-209.

Choi, S., and Greenberg, E. (1991) The C-terminal region of the *Vibrio fischeri* LuxR protein contains an inducer-independent lux gene activating domain. *Proceedings of the National Academy of Sciences* **88**: 11115-11119.

Chugani, S.A., Whiteley, M., Lee, K.M., D'Argenio, D., Manoil, C., and Greenberg, E. (2001) QscR, a modulator of quorum-sensing signal synthesis and virulence in *Pseudomonas aeruginosa*. *Proceedings of the national academy of sciences* **98**: 2752-2757.

Coutinho, B.G., Mevers, E., Schaefer, A.L., Pelletier, D.A., Harwood, C.S., Clardy, J., and Greenberg, E.P. (2018) A plant-responsive bacterial-signaling system senses an ethanolamine derivative. *Proceedings of the National Academy of Sciences* **115**: 9785-9790.

Covaceuszach, S., Degrassi, G., Venturi, V., and Lamba, D. (2013) Structural insights into a novel interkingdom signaling circuit by cartography of the ligand-binding sites of the homologous quorum sensing LuxR-family. *International journal of molecular sciences* **14**: 20578-20596.

Das, K., and Roychoudhury, A. (2014) Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Frontiers in environmental science* **2**: 53.

Devine, J.H., Shadel, G.S., and Baldwin, T.O. (1989) Identification of the operator of the lux regulon from the *Vibrio fischeri* strain ATCC7744. *Proceedings of the National Academy of Sciences* **86**: 5688-5692.

Dreier, J., and Ruggerone, P. (2015) Interaction of antibacterial compounds with RND efflux pumps in *Pseudomonas aeruginosa*. *Frontiers in microbiology* **6**: 660.

Ferluga, S., and Venturi, V. (2009) OryR is a LuxR-family protein involved in interkingdom signaling between pathogenic *Xanthomonas oryzae pv. oryzae* and rice. *Journal of bacteriology* **191**: 890-897.

Ferluga, S., Bigirimana, J., Höfte, M., and Venturi, V. (2007) A LuxR homologue of *Xanthomonas oryzae pv. oryzae* is required for optimal rice virulence. *Molecular plant pathology* **8**: 529-538.

Fernandez-Fuentes, N., Madrid-Aliste, C.J., Rai, B.K., Fajardo, J.E., and Fiser, A. (2007) M4T: a comparative protein structure modeling server. *Nucleic acids research* **35**: W363-W368.

Fernández-Piñar, R., Cámara, M., Soriano, M.I., Dubern, J.F., Heeb, S., Ramos, J.L., and Espinosa-Urgel, M. (2011) PpoR, an orphan LuxR-family protein of *Pseudomonas putida* KT2440, modulates competitive fitness and surface motility independently of N-acylhomoserine lactones. *Environmental microbiology reports* **3**: 79-85.

Fuqua, C. (2006) The QscR quorum-sensing regulon of *Pseudomonas aeruginosa*: an orphan claims its identity. *Journal of bacteriology* **188**: 3169-3171.

Fuqua, C., and Greenberg, E.P. (2002) Listening in on bacteria: acyl-homoserine lactone signalling. *Nature reviews Molecular cell biology* **3**: 685-695.

Fuqua, W.C., Winans, S.C., and Greenberg, E.P. (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of bacteriology* **176**: 269.

Garrido-Sanz, D., Meier-Kolthoff, J.P., Göker, M., Martin, M., Rivilla, R., and Redondo-Nieto, M. (2016) Genomic and genetic diversity within the *Pseudomonas fluorescens* complex. *PloS one* **11**: e0150183.

Garrido-Sanz, D., Arrebola, E., Martínez-Granero, F., García-Méndez, S., Muriel, C., Blanco-Romero, E. et al. (2017) Classification of isolates from the *Pseudomonas fluorescens* complex into phylogenomic groups based in group-specific markers. *Frontiers in microbiology* **8**: 413.

González, J.F., and Venturi, V. (2013) A novel widespread interkingdom signaling circuit. *Trends in plant science* **18**: 167-174.

González, J.F., Myers, M.P., and Venturi, V. (2013) The inter-kingdom solo OryR regulator of *X anthomonas oryzae* is important for motility. *Molecular plant pathology* **14**: 211-221.

Haas, D., and Défago, G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature reviews microbiology* **3**: 307-319.

Hawver, L.A., Jung, S.A., and Ng, W.-L. (2016) Specificity and complexity in bacterial quorum-sensing systems. *FEMS microbiology reviews* **40**: 738-752.

Hesse, C., Schulz, F., Bull, C.T., Shaffer, B.T., Yan, Q., Shapiro, N. et al. (2018) Genome-based evolutionary history of *Pseudomonas spp*. *Environmental microbiology* **20**: 2142-2159.

Hoang, T.T., Karkhoff-Schweizer, R.R., Kutchma, A.J., and Schweizer, H.P. (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants. *Gene* **212**: 77-86.

Igarashi, K., and Kashiwagi, K. (2000) Polyamines: mysterious modulators of cellular functions. *Biochemical and biophysical research communications* **271**: 559-564.

Johnson, L.S., Eddy, S.R., and Portugaly, E. (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC bioinformatics* **11**: 431.

Karatan, E., Duncan, T.R., and Watnick, P.I. (2005) NspS, a predicted polyamine sensor, mediates activation of *Vibrio cholerae* biofilm formation by norspermidine. *Journal of bacteriology* **187**: 7434-7443.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* **10**: 845-858.

Kim, T., Duong, T., Wu, C.-a., Choi, J., Lan, N., Kang, S.W. et al. (2014) Structural insights into the molecular mechanism of Escherichia coli SdiA, a quorum-sensing receptor. *Acta Crystallographica Section D: Biological Crystallography* **70**: 694-707.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop II, R.M., and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* **35**: 1547-1549.

Laloi, C., Apel, K., and Danon, A. (2004) Reactive oxygen signalling: the latest news. *Current opinion in plant biology* **7**: 323-328.

Lequette, Y., Lee, J.-H., Ledgham, F., Lazdunski, A., and Greenberg, E.P. (2006) A distinct QscR regulon in the *Pseudomonas aeruginosa* quorum-sensing circuit. *Journal of bacteriology* **188**: 3365-3370.

Lintz, M.J., Oinuma, K.-I., Wysoczynski, C.L., Greenberg, E.P., and Churchill, M.E. (2011) Crystal structure of QscR, a *Pseudomonas aeruginosa* quorum sensing signal receptor. *Proceedings of the National Academy of Sciences* **108**: 15763-15768.

Markowitz, V.M., Mavromatis, K., Ivanova, N.N., Chen, I.-M.A., Chu, K., and Kyrpides, N.C. (2009) IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* **25**: 2271-2278.

McClean, K.H., Winson, M.K., Fish, L., Taylor, A., Chhabra, S.R., Camara, M. et al. (1997) Quorum sensing and *Chromobacterium violaceum*: exploitation of violacein production and inhibition for the detection of N-acylhomoserine lactones. *Microbiology* **143**: 3703-3711.

Michael, B., Smith, J.N., Swift, S., Heffron, F., and Ahmer, B.M. (2001) SdiA of *Salmonella enterica* is a LuxR homolog that detects mixed microbial communities. *Journal of bacteriology* **183**: 5733-5742.

Miller, J., and Lee, K. (1984) Experiments in molecular genetics. In: Yi Hsien Pub. Co.

Mosquito, S., Meng, X., Devescovi, G., Bertani, I., Geller, A.M., Levy, A. et al. (2020) LuxR solos in the plant endophyte *Kosakonia sp. strain* KO348. *Applied and environmental microbiology* **86**.

Ng, W.-L., and Bassler, B.L. (2009) Bacterial quorum-sensing network architectures. *Annual review of genetics* **43**: 197-222.

Nguyen, Y., Nguyen, N.X., Rogers, J.L., Liao, J., MacMillan, J.B., Jiang, Y., and Sperandio, V. (2015) Structural and mechanistic roles of novel chemical ligands on the SdiA quorum-sensing transcription regulator. *MBio* **6**.

O'sullivan, D.J., and O'Gara, F. (1992) Traits of fluorescent *Pseudomonas spp*. involved in suppression of plant root pathogens. *Microbiology and Molecular Biology Reviews* **56**: 662-676.

Patankar, A.V., and González, J.E. (2009) An orphan LuxR homolog of *Sinorhizobium meliloti* affects stress adaptation and competition for nodulation. *Applied and environmental microbiology* **75**: 946-955.

Prescott, R.D., and Decho, A.W. (2020) Flexibility and Adaptability of Quorum Sensing in Nature. *Trends in Microbiology* **28**: 436-444.

Qiu, X., Gurkar, A.U., and Lory, S. (2006) Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* **103**: 19830-19835.

Sabag-Daigle, A., and Ahmer, B.M. (2012) ExpI and PhzI are descendants of the long lost cognate signal synthase for SdiA. *PLoS One* **7**: e47720.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*: Cold spring harbor laboratory press.

Schrödinger, L. (2010) The PyMOL molecular graphics system, version 1.3 r1. In: August.

Schwarz, G., and Mendel, R.R. (2006) Molybdenum cofactor biosynthesis and molybdenum enzymes. *Annu Rev Plant Biol* **57**: 623-647.

Shadel, G., Young, R., and Baldwin, T. (1990) Use of regulated cell lysis in a lethal genetic selection in *Escherichia coli*: identification of the autoinducer-binding region of the LuxR protein from *Vibrio fischeri* ATCC 7744. *Journal of bacteriology* **172**: 3980-3987.

Shanahan, P., O'Sullivan, D.J., Simpson, P., Glennon, J.D., and O'Gara, F. (1992) Isolation of 2, 4-diacetylphloroglucinol from a fluorescent pseudomonad and investigation of physiological parameters influencing its production. *Applied and environmental microbiology* **58**: 353-358.

Shi, Z., Wang, Q., Li, Y., Liang, Z., HUI, X.L., Zhou, J. et al. (2019) Putrescine is an intraspecies and interkingdom cell-cell communication signal modulating the virulence of *Dickeya zeae*. *Frontiers in Microbiology* **10**: 1950.

Slock, J., VanRiet, D., Kolibachuk, D., and Greenberg, E. (1990) Critical regions of the *Vibrio fischeri* luxR protein defined by mutational analysis. *Journal of Bacteriology* **172**: 3974-3979.

Soares, J.A., and Ahmer, B.M. (2011) Detection of acyl-homoserine lactones by *Escherichia* and *Salmonella*. *Current opinion in microbiology* **14**: 188-193.

Spaink, H.P., Okker, R.J., Wijffelman, C.A., Pees, E., and Lugtenberg, B.J. (1987) Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. *Plant molecular biology* **9**: 27-39.

Stachel, S.E., Messens, E., Van Montagu, M., and Zambryski, P. (1985) Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature* **318**: 624-629.

Staskawicz, B., Dahlbeck, D., Keen, N., and Napoli, C. (1987) Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae pv. glycinea*. *Journal of Bacteriology* **169**: 5789-5794.

Steindler, L., and Venturi, V. (2007) Detection of quorum-sensing N-acyl homoserine lactone signal molecules by bacterial biosensors. *FEMS Microbiology Letters* **266**: 1-9.

Stevens, A.M., and Greenberg, E. (1997) Quorum sensing in *Vibrio fischeri*: essential elements for activation of the luminescence genes. *Journal of bacteriology* **179**: 557-562.

Subramoni, S., and Venturi, V. (2009a) PpoR is a conserved unpaired LuxR solo of *Pseudomonas putida* which binds N-acyl homoserine lactones. *BMC microbiology* **9**: 125.

Subramoni, S., and Venturi, V. (2009b) LuxR-family 'solos': bachelor sensors/regulators of signalling molecules. *Microbiology* **155**: 1377-1385.

Subramoni, S., Florez Salcedo, D.V., and Suarez-Moreno, Z.R. (2015) A bioinformatic survey of distribution, conservation, and probable functions of LuxR solo regulators in bacteria. *Frontiers in cellular and infection microbiology* **5**: 16.

Subramoni, S., Gonzalez, J.F., Johnson, A., Péchy-Tarr, M., Rochat, L., Paulsen, I. et al. (2011) Bacterial subfamily of LuxR regulators that respond to plant compounds. *Applied and environmental microbiology* **77**: 4579-4588.

Taylor, W.R. (2000) Protein structure comparison using SAP. In *Protein structure prediction*: Springer, pp. 19-32.

Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R. et al. (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *The EMBO journal* **21**: 4393-4401.

Venturi, V., and Fuqua, C. (2013) Chemical signaling between plants and plant-pathogenic bacteria. *Annual review of phytopathology* **51**: 17-37.

von Bodman, S.B., Ball, J.K., Faini, M.A., Herrera, C.M., Minogue, T.D., Urbanowski, M.L., and Stevens, A.M. (2003) The quorum sensing negative regulators EsaR and ExpREcc, homologues within the LuxR family, retain the ability to function as activators of transcription. *Journal of bacteriology* **185**: 7001-7007.

Walden, R., Cordeiro, A., and Tiburcio, A.F. (1997) Polyamines: small molecules triggering pathways in plant growth and development. *Plant physiology* **113**: 1009.

Wallner, B., Fang, H., and Elofsson, A. (2003) Automatic consensus-based fold recognition using Pcons, ProQ, and Pmodeller. *Proteins: Structure, Function, and Bioinformatics* **53**: 534-541.

Wang, S., Li, W., Liu, S., and Xu, J. (2016) RaptorX-Property: a web server for protein structure property prediction. *Nucleic acids research* **44**: W430-W435.

Waters, C.M., and Bassler, B.L. (2005) Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**: 319-346.

Wattam, A.R., Abraham, D., Dalay, O., Disz, T.L., Driscoll, T., Gabbard, J.L. et al. (2014) PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic acids research* **42**: D581-D591.

Wellington, S., and Greenberg, E.P. (2019) Quorum sensing signal selectivity and the potential for interspecies cross talk. *MBio* **10**.

Whitehead, N.A., Barnard, A.M., Slater, H., Simpson, N.J., and Salmond, G.P. (2001) Quorum-sensing in Gram-negative bacteria. *FEMS microbiology reviews* **25**: 365-404.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.Y. (2017) ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**: 28-36.

Zhang, L., Jia, Y., Wang, L., and Fang, R. (2007) A proline iminopeptidase gene upregulated in planta by a LuxR homologue is essential for pathogenicity of *Xanthomonas campestris pv. campestris*. *Molecular microbiology* **65**: 121-136.

Zhang, R.-g., Pappas, K.M., Brace, J.L., Miller, P.C., Oulmassov, T., Molyneaux, J.M. et al. (2002) Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature* **417**: 971-974.

Zhang, Y., and Skolnick, J. (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic acids research* **33**: 2302-2309.

Zhou, L., Wang, J., and Zhang, L.-H. (2007) Modulation of bacterial Type III secretion system by a spermidine transporter dependent signaling pathway. *PLoS One* **2**: e1291.

Zhu, J., and Winans, S.C. (2001) The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proceedings of the National Academy of Sciences* **98**: 1507-1512.

Zou, Y., and Nair, S.K. (2009) Molecular basis for the recognition of structurally distinct autoinducer mimics by the Pseudomonas aeruginosa LasR quorum-sensing signaling receptor. *Chemistry & biology* **16**: 961-970.

Chapter IV

Genetics of a novel LuxR solo based cell-to-cell signaling system

Keywords: Quorum sensing, LuxR solos, Fluorescent Pseudomonads, signal molecules, pigments, genomic islands

4.1 Introduction

Many proteobacteria possess so-called LuxR solos; these are quorum sensing LuxR-family regulators that are not paired with a cognate LuxI-family synthase (Fuqua, 2006; Case et al., 2008; Patankar and González, 2009; Subramoni et al., 2015). LuxR solos can exist in both AHL (*N*-acyl-homoserine lactones) producing/communicating bacteria as well as in non-AHL producers (Ahmer et al., 1998; Chugani et al., 2001; Ahmer, 2004; Lequette et al., 2006; Subramoni and Venturi, 2009). LuxR solos are emerging as important players in cell–cell communication and inter-kingdom signaling as they result in alternative ways of cell-cell communication (Venturi and Ahmer, 2015). Some LuxR solos respond to endogenous AHLs whereas some eavesdrop by responding to AHLs exogenously produced by other bacteria, while others respond to a non-AHL signals endogenously or exogenously produced (Sperandio, 2010; Subramoni et al., 2011; Brachmann et al., 2013; Brameyer et al., 2015b). LuxR-family proteins sensing AHLs contain an N-terminal AHL-binding domain having six conserved amino acids essential for binding of AHLs (Shadel et al., 1990; Slock et al., 1990; Choi and Greenberg, 1991). However, this conserved amino acid motif is altered in several LuxR-solos, possibly enabling the sensing of different signaling molecules of yet unidentified bacterial cell-to-cell communication systems (Brameyer et al., 2015a). The role of LuxR solos is still at large unclear; some have been shown to regulate secreted metabolites, including siderophores, redox-active molecules, pigments, and antibiotics, which are commonly encoded in biosynthetic gene clusters (BGCs). In many cases where BGCs are regulated by LuxR-family proteins; the *luxR* gene is adjacently located, which is either part of a QS LuxI/R complete system or is a LuxR solo (Brotherton et al., 2018). Commonly, closely related LuxR-family proteins from different bacteria are associated with different types of BGCs. This

suggests that LuxR homologs have evolved independently, probably via horizontal gene transfer events and that the loss of the AHL synthase gene is a common evolutionary trajectory (Brotherton et al., 2018). However, many open questions remain about the role and mechanisms of LuxR solos in cell-cell communication.

The bioinformatic study presented in Chapter 3 (Bez et al., 2021) revealed that LuxR solos are predominant among the fluorescent *Pseudomonas* spp. group with over 50% of the 600+ genomes analyzed harboring at least one *luxR* solo gene. Amino-acid sequence homology and mapping of the adjacent genetic loci has allowed the subdivision of the majority of these LuxR solos into 9 sub-groups. Only the 2.3% of the LuxR solo hits identified could not be grouped, since they displayed uniqueness in their primary structure and/or neighboring genomic context. Among them, was a *luxR* solo belonging to *P. fluorescent Ps_77* strain which is located genetically adjacent to a large biosynthetic gene cluster consisting of fourteen genes. This LuxR solo is the subject of this study as it could represent a novel intra- and interspecies signaling system in *Pseudomonas* spp. important for microbiome formation and/or establishment.

Members of the *Pseudomonas* fluorescent group are ubiquitous; they are commonly found in soil, foliage, plant root, freshwater, and seawater (Garrido-Sanz et al., 2016; Garrido-Sanz et al., 2017). Their capacity to produce a wide array of bioactive secondary metabolites, including antibiotics, plant hormones and pigments, is of particular interest as well as their ability to exploit many different nutrition sources and have a high potential for adaptation to changing environmental conditions (Cornelis and Matthijs, 2002; Silby et al., 2009; Höfte and Altier, 2010).

P. fluorescens Ps_77 is a blue-pigmenting strain isolated from mozzarella cheese and able to cause food product discoloration (Andreani et al., 2015b). A previous genomic and transcriptomic study identified the biosynthetic pathway involved in the pigment production; it is a genomic region that includes homologs of *trpABCDF* genes, suggesting that tryptophan is involved in the production of the pigment (Andreani et al., 2015b; Andreani et al., 2019). The black/bluish pigment is released extracellularly in the media during the late logarithmic phase/early stationary phase of growth and it is visible after 48 h of incubation. The biological function and the chemical structure of this pigment is currently unknown; some bacterial pigments play a role in oxidative stress resistance, in iron metabolism by functioning as iron-transport agents. This study presents the genetic and molecular characterization of a unique *luxR* solo located genetically adjacent to the large pigment biosynthetic operon of *Pseudomonas* Ps_77. Comparative genome analysis, structure-based modeling, putative target gene promoter expression analysis, and co-culture studies were performed. Results indicate that this LuxR solo is involved in the transcriptional regulation of the adjacent biosynthetic operon. This locus of the *luxR* solo together with the biosynthetic operon is very rare in bacteria and most likely constitutes a genomic island. It is believed that this is a novel bacterial communication mechanism since our experiments evidence that this system is capable of inter-cellular signaling and the pigment-like molecule is most likely the signal for the LuxR solo.

4.2 Material and methods

4.2.1 Bacterial strains and growth conditions

The *P. fluorescens Ps_77* strain used in this work was previously isolated from mozzarella cheese and its blue-pigment production ability was described in (Andreani et al., 2015a; Andreani et al., 2019). The strain was grown in liquid Luria-Bertani (LB), King's Broth (KB), M9 medium, and PDA (Potato Dextrose agar) at 30 °C under moderate shaking (120 rpm) or without shaking. When required, antibiotics were added at the following concentrations: nitrofurantoin (Nf) 100 µg ml^{-1}, ampicillin (Amp) 100 µg ml^{-1}. The three mutants generated (Δ*fluR*, Δ*trpC-like*, Δ*luxI*) have been grown using 100 µg ml^{-1} kanamycin (Km) as antibiotic. *E. coli* DH5α and S17 were routinely grown at 37 °C in LB broth and antibiotics were added when required at the following concentrations: Amp 100 µg ml^{-1}, tetracycline 15 µg ml^{-1}, gentamycin 15 µg ml^{-1}. The *P. frederiksbergensis* OS210_3 was used in the mixed co-culture experiment for demonstrating the cross-talk. Its draft genome is deposited to the Integrated Microbial Genomes and Metagenomes (IMG/M) database (project ID Ga0314296_11). The same growth conditions describe above were adopted.

4.2.2 Plasmid and recombinant DNA techniques

The plasmids, constructs and set of primers (Sigma-Aldrich) used in this study are listed in **Table 4.1**. pGEM-T Easy vector (Promega Corp., Madison, WI, USA) was used for cloning. When necessary, 5-bromo-4-chloro-3-indolyl-β-d-galactoside (X-Gal) was added at a final concentration of 80 µg ml^{-1}. Routine DNA manipulation steps, such as digestion with restriction enzymes, agarose gel electrophoresis, purification of DNA fragments, ligation with

T4 DNA ligase and transformation of *E. coli*, were performed as described previously (Sambrook et al., 1989). Plasmids were purified by using EuroGold columns (EuroClone, Pero, Milan, Italy); total DNA was isolated by sarkosyl-pronase lysis, as described previously (Better et al., 1983). Digestion with restriction enzymes was conducted according to the supplier's instructions (New England BioLabs, USA). DNA was ligated with T4 DNA ligase (New England BioLabs, USA) according to the manufacturer's recommendations.

Table 4. 1 Plasmids and primers used in this study.

Plasmids/primers	**Relevant features**	**References or sources**
pGEM-T	Cloning vector; Amp^R	Promega
pMP220	Promoter probe vector; IncP;Tc^R	(Spaink et al., 1987)
pMPGFP	Promoter GFP probe vector; IncP;Tc^R	(Devescovi, 2017)
pBBR1MCS-5	Broad-host-range vector; Gm^R	(Kovach et al., 1995)
pBBR1MCS-3	Broad-host-range vector; Tc^R	(Kovach et al., 1995)
pKNOCK-Km	Conjugative suicide vector; Km^R	(Alexeyev 1999)
pEX19Gm	Suicide vector for making deletion mutants, Gm^R	(Dreier and Ruggerone, 2015)
pUC4K	pUC7 derivative, Amp^R and Km^R	Addgene, Watertown, MA
pEX19-*fluR*	*fluR* sequence depleted of 20 bp cloned in pEX19Gm	This study
pEX19-*trpC-like*	*trpC-like* sequence depleted of 20 bp cloned in pEX19Gm	This study
pKNOCK*luxI*	Central region of the *luxI* sequence cloned in pKNOCK-Km	This study
pMP220pr*fluR*	*fluR* promoter cloned in pMP220	This study

pMP220pr*Operon*	Biosynthetic operon promoter cloned in pMP220	This study
pMPGFPpr*Operon*	Biosynthetic operon promoter cloned in pMPGFP	This study
pBBR5-*fluR*	Complete *fluR* and promoter sequence promoter cloned in pBBR1MCS-5	This study
pBBR3-*fluR*	Complete *fluR* and promoter sequence promoter cloned in pBBR1MCS-3	This study
fluR for pEX	ggtaccCAGTTTTTAAACGTCAAAATATTGGGAGCTTTTTTACTTCCC ATTTGAAATTCATCTAGGTATAAAACTCTCAGGGGAAACCAGTAG ATAAAGGAATAACGATGCGCAACTGGTACAATGACCTGCTTGAAT GGGCTGGGCAGGTCGAGTCTGAAAATGATTTTTTGTATAGGGCAT GCAAGCTTGCACAGTCATTGGAGTTTGAATGGTGCAGCTATCACG TCCAACCGCCCCTCCCCATTTCCAAGCCTGTCATCGCCTTTGCAAG CAATTATCCAAAGGCATGGCAGCGACGCTATCGCGACATGGACTA TGTGCAATTGGACCCGGTCGTCAAAAAGGCGCGGCTGACCCAACT GCCCTTTGTCTGGGAAAGTACACTGCTCGAACAGGAACCCTGCTT CTGGAAGGAAGCGGGTGACGCGGGCCTGAGGGggatccATCAGCA GCAGCGGCACCTTCAGCATGTTGACGCTCGCCCGCAACGAAGAGC CGCTCACCATCAGCGAATTGAATGACAAAGAGCTGAAAATGCGCT GGCTGGCCGATGCCACTCACGTTGCACTGAGCCGCCTGTTCAAAC CGCAGGAACTCGAGGAATCCTACTGGCGGTTGACCGCCCGGGAA ATCGAGATCCTGCGCTGGACAGCCGACGGAAAAACCCAGTGCGA GATCTCACAAATTCTCTCGGTATCTTTCGACACCGTGAAGTTTCAC AGTAAAAATGCTATCGCCAAGCTGGGCACTACCAATAAAACCGCA GCCGTGGTGAGAGCGACCGTTCTGGGCGTGCTGGGCTAATGAGC CTGCTATCCTCTCATCAGGCCGCATGGCGCTCATCCCGAAAGCATC GGGCGtctaga	This study
trpC-like for pEX	ggtaccTAATTTTTGTTGGGTGGCTTGTTCCATGACTGTCATGCGAG ATAATTATATGCTTGAAGAGATCGTTGCGTTTAAGGCTGTTGAAA CGGCCCAAAGGAAAAGTCATCACTCCCTGGGTTCCCTAGAGCGTC GTATAGCTGATGCCAGGGCGCCGCGGGCGTTTGCCCAAGCGATT GCGACCTCAAGCGTTGCGGTTATTGCAGAAGCCAAGTATCGCTCT CCTTCAAAGGGCGTATTACGTGCCGACTATGATCCTCTGGCCTTGG CCCATGCTTATCAGGCGGGCGGTGCCAGTGCGCTTTCGGTGCTGG CCGATAGTCGGTTTTTCGGCAACGCACCCTACGTGGTGGGGCTGT TGGCGAATGCGCCGGGCTTGAACCTGCCAGTGATGTACAAAGACT TTATTGTCGATGAGTTCCAGGTCTACGAGGCCCGCGCGCTGGGGG CTGACGCGATTCTGATCATCGTGCggatccCGACTCTATACGCTGGC CCTGGAATTGGGGCTCGACGTATTGGTGGAAACATTTGATGAGGC TGATATAGACCAGGCCTTGAGTGTCGGCGCAGGGATCGTCGGTAT CAATAACCGTGACCTGGATACATTCAAGGTCAACTTTGATCGCACC GCTGAATTGTTTGAGTTGTTACCTGGTCAGGTCATTGGGGTAGCC GAAAGTGGTATCTCCGGAGTTGCCGATTTTAATCGGATTAATACT ATCGGGTTTCGTGCGGCGCTGATGGGTGAATATTTATTGGGTGCT GAAGATCCAACTCGGCAGTTGCGTTTTTTGACCGCGGGAGGCGAT CCCAATTGACCCTGATGCACTCGATAATAATGGGCTATGGGCACT GCGGAAAGAACTTGCACCACGTATGtctaga	This study

prOperon	CCAGCCCATTCAAGCAGGTCATTGTggatccACCAGTTGCGCATCGTTATTCCTTTATCTACTGGTTTCCCCTGAGAGTTTTATACCTAGATGAATTTCAAATGGGAAGTAAAAAAGCTCCCAATATTTTGACGTTTAAAAACTGACAGTTAGTTGGAGCTTTCGTGCTGCATTAAATGGCTCGTTATTCGTGTGGGATATTAAATGACATGGTGTCAATTGTATCTTGGTTGCAGTTTTTTAAAGGCTCAAGCGCCTACCCAACAAGGTAGGTAGACGGGGTGCTGGCTTTTCCTTTTAATTAATTTGCTGCGAAAAGAGTGGCTAGTTAGTTGTCGAGAATAACAAGGCGGGTTTGATTTCTCTTCTTGGGGCTTTACCCAATAATTTTTGTTGGGTGGCTTGTTCCATGACTGTCATGCGAGATAATTATATGCTTGAAGAGgaattcATCGTTGCG	This study
pr*fluR*	CGCAACGATggatccCTCTTCAAGCATATAATTATCTCGCATGACAGTCATGGAACAAGCCACCCAACAAAAATTATTGGGTAAAGCCCCAAGAAGAGAAATCAAACCCGCCTTGTTATTCTCGACAACTAACTAGCCACTCTTTTCGCAGCAAATTAATTAAAAGGAAAAGCCAGCACCCCGTCTACCTACCTTGTTGGGTAGGCGCTTGAGCCTTTAAAAAACTGCAACCAAGATACAATTGACACCATGTCATTTAATATCCCACACGAATAACGAGCCATTTAATGCAGCACGAAAGCTCCAACTAACTGTCAGTTTTAAACGTCAAAATATTGGGAGCTTTTTTACTTCCCATTTGAAATTCATCTAGGTATAAAACTCTCAGGGGAAACCAGTAGATAAAGGAATAACGATGCGCAACTGGTgaattcACAATGACCTGCTTGAATGGGCTGG	This study
Primers	**Sequence**	**Source**
Km_cassette_Rv1	CAACTCTGGCGCATCGGGCT	This study
Km_cassette_Rv2	GCGTAATGCTCTGCCACACA	This study
FluR_EXT_Fw	ACC AGG CCC GGT TTG AAG TCG A	This study
TrcP_EXT_Fw	TGC CAT CTC AGG CAT GGC TTC	This study
FluRcompl_Fw	GCTCTAGAGC CGCAACGATCTCTTCAAGCAT	This study
FluR_compl_Rv	GGGGTACC CC GCT CAT TAG CCC AGC AC	This study
LuxI_pKNOCK_Fw	TCT AGA CAA CGC GAA TTC GA	This study
LuxI_pKNOCK_Rv	GGT ACC AGA ACC GAG GCG	This study
LuxI_mut_control	CGA CAA TTG CGA TAA GGA CA	This study

4.2.3 Construction of Ps_77 ΔfluR, ΔtrpC-like and ΔluxI and complementation

In frame deletions of the *fluR solo* gene and *trpC-like* gene were generated using the pEX19Gm plasmid as described previously (Hoang et al., 1998). Briefly, each gene sequence, synthetized by Twist bioscience company (South San Francisco), is listed in the **Table 4.1**. The

design of the constructs was performed as follows: internal fragments of 20 bp from each gene of interest were deleted and replaced with a restriction site (BamHI) in order to clone inside the Km gene cassette previously extracted from pUC4K. Sequentially the fragments were excised with Kpn and XbaI restriction enzymes and cloned in the corresponding site in pEX19Gm. The resulting pEX19Gm-derivative plasmids, listed in **Table 4.1**, were introduced by biparental conjugation in the corresponding *P. fluorescens Ps_77* genome. Clones with a chromosomal insertion of the pEX19Gm plasmids were selected on LB agar plates supplemented with 40 μg ml^{-1} gentamycin (Gm) and 100 μg ml^{-1} Nf. Plasmid excision from the chromosome was subsequently selected on LB agar plates supplemented with 10% (w/v) sucrose. All the mutants were verified by PCR using primers (**Table 4.1**) specific to the Km cassette and to the genomic DNA sequences upstream and downstream from the targeted genes.

The *luxI*-homolog mutant was generated using the suicide vectors from the pKNOCK series (Alexeyev, 1999). To generate Ps_77Δ*luxI*, an internal fragment (415 bp) of the *luxI* gene was amplified by PCR using the primers listed in **Table 4.1** and cloned as a *Kpn*I-*XbaI* fragment into the corresponding sites of pKNOCK-Km resulting in pKNOCKluxI. This latter plasmid, was delivered to *P. fluorescent Ps_77* genome by biparental conjugation and transformants were selected after appropriate antibiotic selection (Km).

For complementation analysis, the *fluR* full-length gene (including its gene promoter) was amplified with the primers described in **Table 4.1**; the sequences were verified via DNA sequencing and the resulting fragments were cloned in pBBR1MCS-5 vector and pBBR1MCS-3 (Kovach et al., 1995). The plasmids containing the fragments were individually delivered by biparental conjugation in the mutant strains Ps_77Δ*fluR* and selected for Km^R and Gm^R or Tc^R

respectively and the resulting complemented mutant strains were named Ps_77Δ*fluR* (pBBR*fluR*). Mutants and complemented mutants were verified by colony PCR.

4.2.4 Gene promoter studies via β-galactosidase activity assay

Transcriptional activity studies of two gene promoters (*fluR* gene promoter and the biosynthetic operon promoter) were studied in *P. fluorescens Ps_77 WT, ΔfluR, ΔtrpC-like and ΔluxI.* Gene transcriptional fusion plasmids were constructed in the pMP220 promoter probe vector which harbors a promoterless *lacZ* gene (Spaink et al., 1987). The two promoter sequences were synthesized by Twistbioscience company (South San Francisco), as described in **Table 4.1,** and digested using *BamHI* and *EcoRI* and cloned in the corresponding sites in promoter vector pMP220. β–galactosidase, activity of *P. fluorescens Ps_77* transconjugants harboring the transcriptional plasmid fusion constructs was determined as previously described by (Miller, 1972), with the modifications of (Stachel et al., 1985). Determination of each promoter activity was performed in independent biological triplicates and as control the empty pMP220 promoter probe vector was used.

4.2.5 Statistical Analysis

For analysis of statistical significance, the data were analyzed using GraphPad Prism's t-test or ANOVA and $P < 0.05$ was considered significant for all experiments.

4.2.6 Gene promoter studies by confocal microscopy and cytofluorimetry

The *gfp* reporter gene was also used for studying the promoter activities by co-cultures of *P. fluorescens Ps_77* wild type and derivative mutants in order to further investigate the possible involvement of the genomic island in cell-to-cell signaling. Gene

transcriptional fusion plasmids were constructed in the pMPGFP promoter probe vector which harbors a promoterless *gfp* gene (da Silva et al., 2014; Devescovi et al., 2017). The promoter sequence driving transcription of the biosynthetic operon was synthetized by Twistbioscience company (South San Francisco), as described in **Table 4.1,** and digested using *BamHI* and *EcoRI* and cloned in the corresponding sites in promoter vector pMPGFP. The plasmid containing the promoter sequence was delivered by biparental conjugation in the *Ps_77* WT and in the two mutant strains Δ*fluR* and Δ*trpC-like* and appropriately selected. As control, the empty pMPGFP promoter probe vector was delivered in the *Ps_77* WT and in the two mutant strains Δ*fluR* and Δ*trpC-like.* 6 different set-ups were used in the co-culture as schematically shown in the **Figure 4.1**. Each single pure culture has been grown in liquid LB separately for 18 hrs (OD_{600} 1), diluted to OD_{600} 0.1 and then mixed with the same ratio of another bacterial culture (**Figure 4.1**). After 48 hrs of co-growth and the visible appearance of the dark/bluish pigment production, the presence of GFP was visualized and/or quantified by cytofluorimetry and confocal microscopy. *P. frederiksbergensis* OS210_3 was also used in the mixed culture experiment for demonstrating any possible cross-talk since it harbors a highly similar biosynthetic operon and *luxR* solo. The same experiment has been performed also using the cell-free supernatant of the bacterial cultures grown 48 hrs, and then added to the pellet of 48 hrs grown constructs to test the presence of quorum sensing signal compounds.

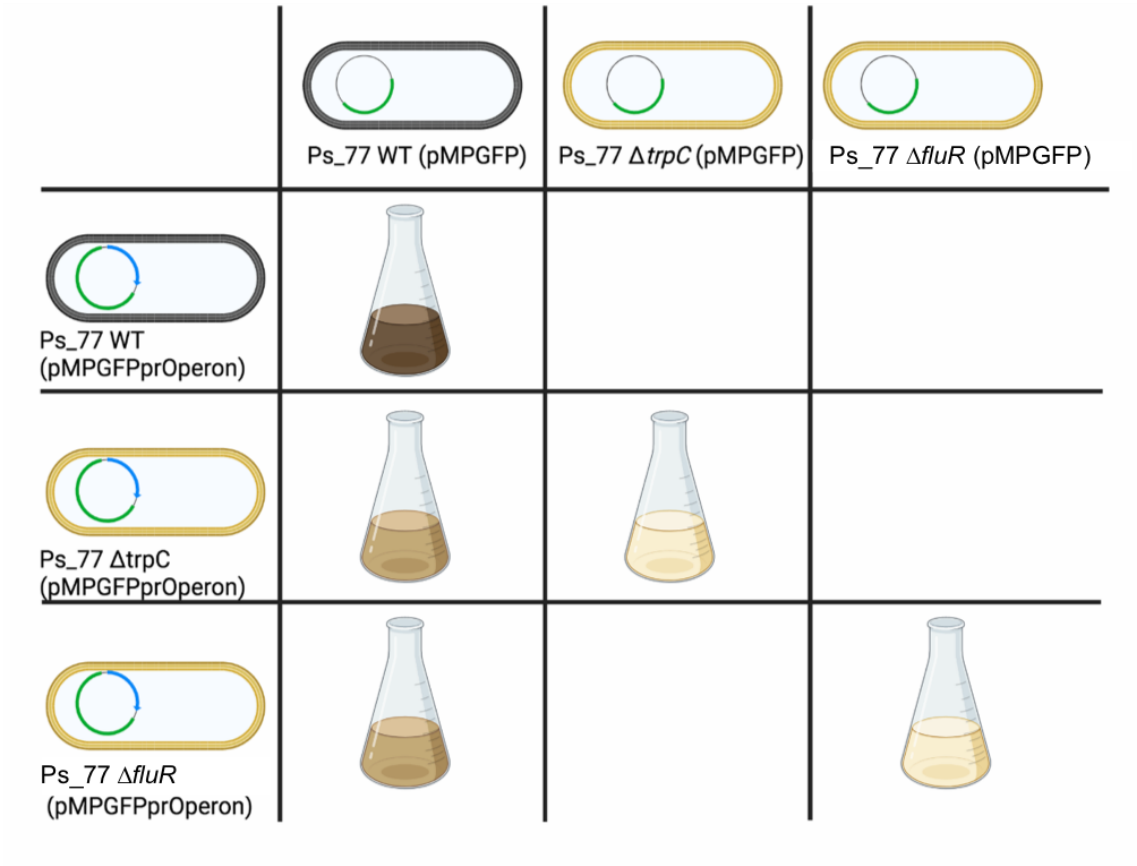


Figure 4. 1 Schematic representation of the mixed cultures set-up for evaluating the involvement of the LuxR solo and the entire genomic island in cell-to-cell signaling. Each experiment has been performed in triplicate. Figure created using Biorender.com

Cytofluorimetry

GFP fluorescence was analyzed using a FACS Calibur cytofluorimeter (Becton Dickinson), equipped with the Argon-ion laser and operating at 488 nm. 5(6)-FAM-SE green fluorescence (FL1) was collected using 530 ± 30 band pass filter. Data was collected using logarithmic amplification either for FSC and SSC or FL1. A FSC threshold was set to gate out debris and it has been gated in R1, while the GFP-positive population-events have been gated in R2. For each sample a total of 100000 events have been acquired. The data were then analyzed by CellQuest software from Becton Dickinson.

Confocal microscopy

For the confocal microscopy detection, bacterial suspensions were washed twice in PBS and diluted 1:1000; around 5 μl of each sample were added to a glass slide, and a cover slip was glued onto the glass slide with nail polish. Samples were imaged using a confocal microscopy setup (Zeiss Airyscan equipped with a 63× objective with an NA of 1.3), and the images were processed using ZEN lite software. Images of green fluorescence were acquired at an excitation wavelength of 488 nm, using the FITC filter. The images were acquired and GFP expression was analyzed using ImageJ software (National Institute of Health, U.S.A.).

4.2.7 Structure homology modeling and structural alignment of FluR

Five web-based servers were used to build the 3D structure-based homology model of the IBD of FluR solo. The top-score models generated by the servers were then ranked and validated by the protein model quality predictor ProQ (Wallner and Elofsson, 2003) and by PSIPRED (Buchan et al., 2010) for the secondary structure prediction. The IntFOLD server (Buenavista et al., 2012) produced the highest quality 3D models for FluR from *Pseudomonas fluorescent Ps_77* according with the ranking obtained by ProQ being the predicted LG-score and MaxSub value of 4.078 and 0.725 respectively for FluR. The template used for FluR modeling were SdiA from *E. coli* (PDB_ID 4LGW_A) (Nguyen et al., 2015) and CviR from *Chromobacterium violaceum* (PDB_ID 3QP6) (Chen et al., 2007).

Sequence alignment was performed by Expresso (Armougom et al., 2006), exploiting structural aligners algorithms like SAP (Taylor, 2000) or TMalign (Zhang and Skolnick, 2005). Each sub-group prototype was also aligned with all the canonical QS LuxR proteins, whose X-ray structures are available: TraR from *A. tumefaciens* (PDB_ID 1H0M (Vannini et al., 2002)) and from *S. fredii* NGR234 (PDB_ID 2Q0O (Chen et al., 2007)), LasR (PDB_ID 3IX3 (Zou and

Nair, 2009)) and QscR (PDB_ID 3SZT (Lintz et al., 2011)) from *P. aeruginosa*, CviR from *C. violaceum* (PDB_ID 3QP1 (Chen et al., 2007)) and SdiA from *E. coli* (PDB_ID 4Y13) (Nguyen et al., 2015)) . The structure-based homology model of OryR from *X. oryzae* (Covaceuszach et al., 2013), the prototype of the sub-group C, was also included in the structural-based multiple alignment.

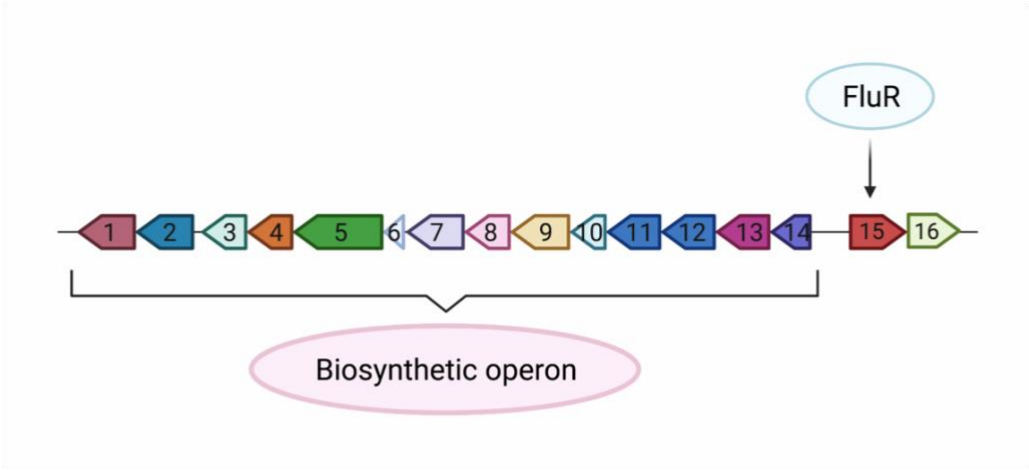
4.2.8 FluR primary sequence and putative genomic island homology analysis

The amino acid sequence of each gene of the complete genomic island were concatenated and queried by a broader blast search, using the sequence aligner Diamond, against all genomes in the US Department of Energy IGM/M database and NCBI (National Center for Biotechnology Information). Only operons with 13 genes, without variations in the order of the genes and stronger than E- value of 10^{-25} were retained in the analysis. The same analysis has also been done retaining all the hits that have at least 7 or more genes in the operon. The concatenation of each protein has been done using the command union from the Emboss package. The alignment of the operons has been performed using MAFFT v.7 software. The homologies were calculated using the pairwise comparison with MEGAX program package (Kumar et al., 2018).

4.3 Results

4.3.1 Identification of a novel LuxR solo possibly harbored in a genomic island

The genome of *P. fluorescens Ps_77* harbors one complete QS *luxI*/*luxR* system and three additional *luxR* solos, one belonging to the Sub-group A, one to the Sub-group C (Chapter 3 and Bez et al., 2021) and one which is ungrouped, showing unique primary structure and flanking gene context. The latter, designated as FluR, is located in operon with a gene encoding for an efflux pump on the (+) strand and genetically adjacent to a large biosynthetic operon including fourteen genes on the (-) strand. The expression of these two operons is under a divergent promoter of 389 bp. A graphical representation of the genomic organization of this locus is reported in **Figure 4.2**.



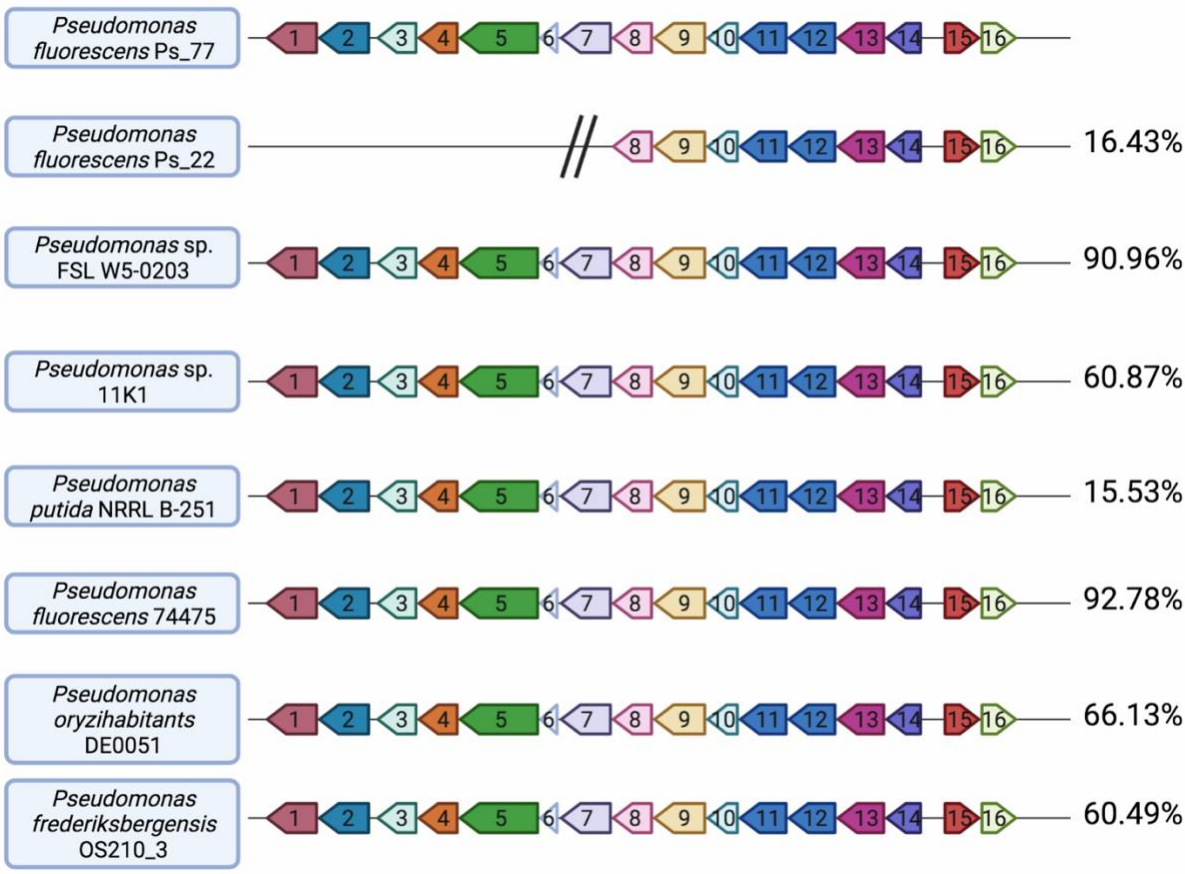
1. UDP-N-acetyl-D-galactosamine dehydrogenase (430 aa)
2. Anthranilate phosphorybosiltransferase –*trpD* (321 aa)
3. Phosphorybosilantranilate isomerase-*trpF* (200 aa)
4. Tryptophan synthase alpha chain-*trpA* (266 aa)
5. Tryptophan synthase beta chain (628 aa)
6. hypothetical protein (138 aa)
7. Glucose--fructose oxidoreductase precursor (305 aa)
8. putative oxidoreductase YjmC (361 aa)
9. putative succinyl-diaminopimelate desuccinylase (424 aa)
10. Inosose dehydratase / Sugar phosphate isomerase/epimerase (289 aa)
11. UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronate aminotransferase (371 aa)
12. UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronate aminotransferase (377 aa)
13. Oxidoreductase family, NAD-binding Rossmann fold (352 aa)
14. Indole-3-glycerol phosphate synthase-*trpC* (262 aa)
15. LuxR (237 aa)
16. Efflux pump/transporter (208 aa)

Figure 4. 2 Genes map and annotation of the genomic system studied. Figure created using Biorender.com

FluR displays its highest identity (30.33%) in its primary structure with SdiA from the *Enterobacteriaceae* family and it has the two PFAM domains conserved among the QS LuxR-type family proteins, one autoinducer binding domain (PF03472) and one bacterial regulatory protein, LuxR type DNA-binding HTH domain (PF00196).

The biosynthetic operon (a total of 16,160 base-pairs) contains genes having homologies with genes involved in L-tryptophan biosynthesis: *trpD* (ID 2742377930; Anthranilate phosphoribosyl transferase), *trpF* (ID 2742377931; N-(5'-phosphoribosyl) anthranilate isomerase), *trpA* (ID 2742377932; Tryptophan synthase alpha chain) and *trpC* (ID 2742377932; Indole-3-glycerol phosphate synthase) (**Supplementary file S4.1**). Genome mining revealed the presence of a *trp* gene cluster also in the core genome, suggesting that this biosynthetic operon harbors accessory tryptophan related genes. The presence of ICEs genes (integrative and conjugative genes) located upstream and downstream of this locus and a significantly different GC content, hints the possibility that this is a genomic island which has been acquired via horizontal gene transfer events (HGT).

Search studies of genome data banks showed that the complete locus is very rare and limited to only 7 *Pseudomonas* spp. isolates; namely *P. fluorescens* strains Ps_22, *Pseudomonas* 7445, *P. putida* NRRLB-251, *P. oryzihabitans* DE0051, *Pseudomonas frederiksbergensis* OS210_3, *Pseudomonas sp*. strains FSL W5-0203 and 11K1 (**Figure 4.3 A-B**). This locus showed a variable sequence relatedness varying around 15 and 92% of identity. The most conserved sequence compared to *P. fluorescens Ps_77* was found in *Pseudomonas* sp.7445 (sequence homology 92.78%), while the lowest conserved sequence was with *P. putida* NRRLB-251 (sequence homology of only 15.53%) (**Figure 4.3 A-B).**



B.

		1	2	3	4	5	6	7	8
11K1	1		64.30	94.37	62.42	15.98	58.96	12.83	60.87
74475	2	64.30		63.90	94.73	15.74	68.89	15.94	92.78
frederick	3	94.37	63.90		62.09	15.78	58.92	12.72	60.49
FSLW5	4	62.42	94.73	62.09		15.55	66.63	15.55	90.96
NRRLB-251	5	15.98	15.74	15.78	15.55		15.01	29.17	15.53
P.oryzihabitans	6	58.96	68.89	58.92	66.63	15.01		13.68	66.13
Ps_22	7	12.83	15.94	12.72	15.55	29.17	13.68		16.43
Ps_77	8	60.87	92.78	60.49	90.96	15.53	66.13	16.43	

Figure 4. 3 A) Comparative genome analysis and homology. The concatenation of the proteins has been performed using the package Emboss and the alignment using MAFFT v.7 software. **B)** Overview of the protein sequence homologies among the different hits obtained using the pairwise comparison.

4.3.2 Comparative cartography analysis of FluR

In order to gain structural insights underlying substrate specificity of FluR, multiple structure-based sequence alignment and structure-based homology modelling were used. The focus was on the specific substitutions of conserved amino acids within the inducer binding domain (IBD) and in particular on the pocket residues directly interacting with the ligand that are conserved and belong to Cluster 1 and Cluster 2 (colored in green and in cyan, respectively in **Figure 4.4**) and to the pocket residues identified as variable, belonging to Cluster 3, colored in pink, as previously described (Covaceuszach et al., 2013).

```
                                 v  v   sv sv      cccs          c
1HOM_TraR   LHIQ-------------HRHITAVTNYHRQWQSTYFDKKFEALDPVVKRARSRKHIFTWSGE----
2Q0O_TraR   LQKD-------------GTQVRTFHSYPGPWESIYLGSDYFNIDPVLAEAKRRRDVFFWTAD----
2IX3_LasR   GLLP---KDSQD-----YENAFIVGNYPAAWREHYDRAGYARVDPTVSHCTQSVLPIFWEPS----
3S2T_QscR   GARA---PFPLT-----APKYHFLSNYPGEWKSRYISEDYTSIDPIVRHGLLEYTPLIWNGE----
4LGW_SdiA   CVRH---PVPFT-----RPKVAFYTNYPEAWVSYYQAKNFLAIDPVLNPENFSQGHLMWNDD----
3QPS_CviR   ALGRLNNQNQIQ-----RLERVLNVSYPSDWLDQYMKENYAQHDPILRIHL-GQGPVMWEERFNRA
OryR        DYAP----VPLSMEGALTPTVFMQRNAPGDMQHVWCEHGYYQHDPVQQRATRRNTPFVWSYR----
FluR        HVQP---PLPIS-----KPVIAFASNYPKAWQRRYRDMDYVQLDPVVKKARLTQLPFVWEST----
                                         r   r      dddd          r
                            s   s   s  c              c
1HOM_TraR   HE-----RPTLSKDERAFYDHASDFGIRSGITIPIKTA-NGFMSMFTMASDKPVI-DLDRE-ID-A
2Q0O_TraR   AW-----PARGSSPLRRFRDEAISHGIRCGVTIPVEGS-YGSAMMLTFASPERKV-DISGV-LDPK
2IX3_LasR   IY--------QTRKQHEFFEEASAAGLVYGLTMPLHGA-RGELGALSLSVEAENRAEANRF-MESV
3S2T_QscR   DF----------QENRFFWEEALHHGIRHGWSIPVRGK-YGLISMLSLVRSSESIA-ATEI-LEKE
4LGW_SdiA   LF----------SEAQPLWEAARAHGLRRGVTQYLMLP-NRALGFLSFSRCSAR-E-IPILSDELQ
3QPS_CviR   KG----------AEEKRFIAEATQNGMGSGITFSAASERNNIGSILSIAGRE-PGR-NA----ALV
OryR        TDGDCAGVEYVGGQHRQVTRYLCDSGMGTGVTVPLHLP-GGAFATFSAAIDAVAA-EALRL-AESQ
FluR        LL----------EQEPCFWKEAGDAGLRVGWTCSSISS-SGTFSMLTLARNEEPLT-ISEL-NDKE
                            r   r   p  f              f
```

Figure 4. 4 Structure-based multiple sequence alignment of the regulatory domains of FluR with members of canonical QS LuxR family and LuxR solos. The residues belonging to Cluster 1, to Cluster 2 and Cluster 3 are highlighted in green, cyan and in pink, respectively. The 3D architecture of the boundaries of the ligand-binding site is schematized by *r* (roof), *f* (floor), *p* (proximal wall) and *d* (distal wall) and its tripartite topology by *c* (conserved core), *s* (specificity patch) and *v* (variable patch).

The following primary sequences were included in the multiple alignment: TraR from *Agrobacterium tumefaciens* -TraR_*At*- (PDB_ID 1H0M_A) (Vannini et al. 2002) and from *Sinorhizobium fredii* NGR234 -TraR_*Sf*- (PDB_ID 2Q0O_A) (Chen et al. 2007) prototypes of the canonical QS LuxR proteins, SdiA from *Escherichia coli* (PDB_ID 4LGW_A) (Kim et al. 2014), QscR (PDB_ID 3SZT_B) (Lintz et al. 2011) and LasR (Zou and Nair 2009) from *P. aeruginosa,* prototypes of AHL-biding LuxR solos and OryR from *Xanthomonas oryzae*, prototypes of the Plant Associated Bacteria (PAB) LuxR solos subfamily (**Figure 4.4**).

According to the molecular cartography and structure-based alignment, FluR is very closely related to the archetypical QS LuxRs. It maintained the two conserved hydrogen bonds stabilizing AHL binding (**Figure 4.5**), namely one between the ε nitrogen of W57 (according to TraR numbering) and the carbonyl oxygen of the lactone moiety and the second between the ε oxygen of D70 and the nitrogen preceding the acyl moiety.

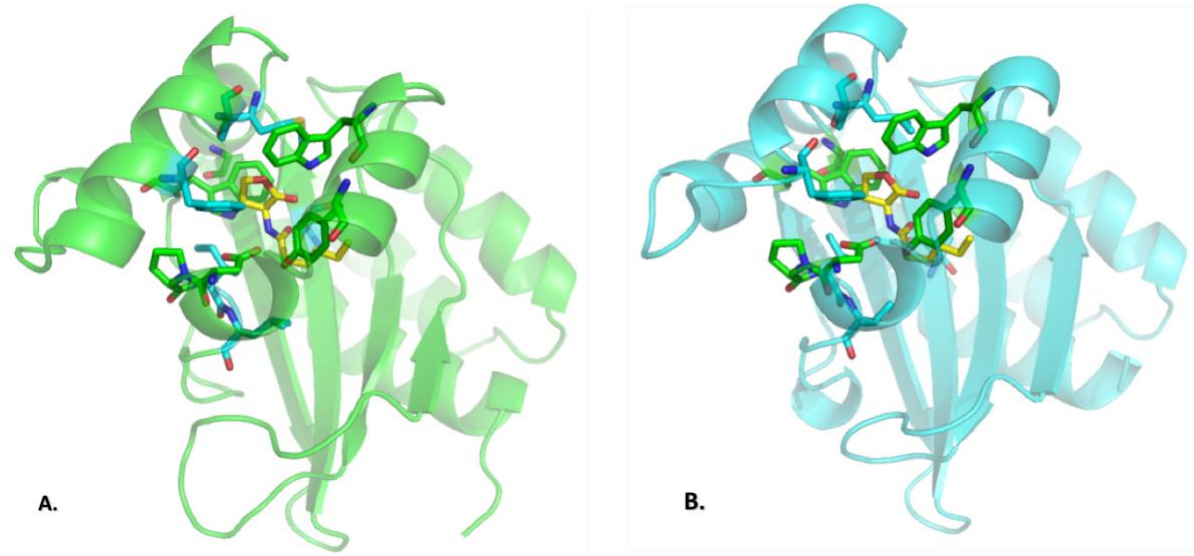


Figure 4. 5 Comparison of the ligand binding site of the QS LuxR solo SdiA (ID_4LGW) **(A)** with FluR **(B).** Semitransparent cartoon representation, with the side chains of residues belonging to Cluster 1 and Cluster 2 highlighted in green and cyan, respectively. conserved residues are represented by lines while non conserved amino acids are highlighted by sticks. Figures produced by Pymol (Schrödinger, 2010).

In addition, all the apolar residues belonging to the conserved and specificity patches, which further stabilize the AHL binding by hydrophobic interactions, are maintained with respect to the AHL-binding template. Interestingly residues belonging to both the conserved and the specificity patches in FluR (**Figure 4.4)** are identical to those of the corresponding regions in the canonical QS LuxR solo SdiA (**Figure 4.5)**, whose crystal structure (ID_4LGW) has been used as a template for homology modeling. Overall, the comparative structural analysis suggests that this LuxR could bind and respond to AHLs (**Figure 4.5**).

4.3.3 FluR detects an endogenous signal and activates the expression of the adjacent biosynthetic operon

It was of interest to acquire insights into the mode of action and genetics of the FluR solo and whether it was involved in the transcriptional regulation of the adjacent operon. The *fluR* solo and the first gene of the biosynthetic operon were mutated as described in the Material and Methods section. Subsequently, the transcriptional activity of *fluR* and the adjacent operon was studied via gene promoters transcriptionally fused to a *lacZ* reporter gene in a plasmid construct and assays were performed in *P. fluorescens Ps_77* wild type, *ΔfluR* and *ΔtrpC-like*.

As already mentioned above, *P. fluorescens Ps_77* wild type produces a black/bluish pigment at the early and late stationary phase under certain conditions. In fact, *P. fluorescens Ps_77* produced the black/bluish pigment when grown in plate media, in liquid media however, it only produced pigment when grown under static non-shaking conditions. If the liquid culture was shaken, no visible pigment was produced even after many days of growth (**Figure 4.6 A**). Interestingly, both the *ΔfluR* and *ΔtrpC-like* mutants lost the ability to produce the pigment, suggesting that the *luxR* solo and the biosynthetic cluster were involved in pigment

production. To further confirm this hypothesis, *ΔfluR* mutant complemented *in trans* with the *fluR* gene restored the pigment production (**Figure 4.6 B**).

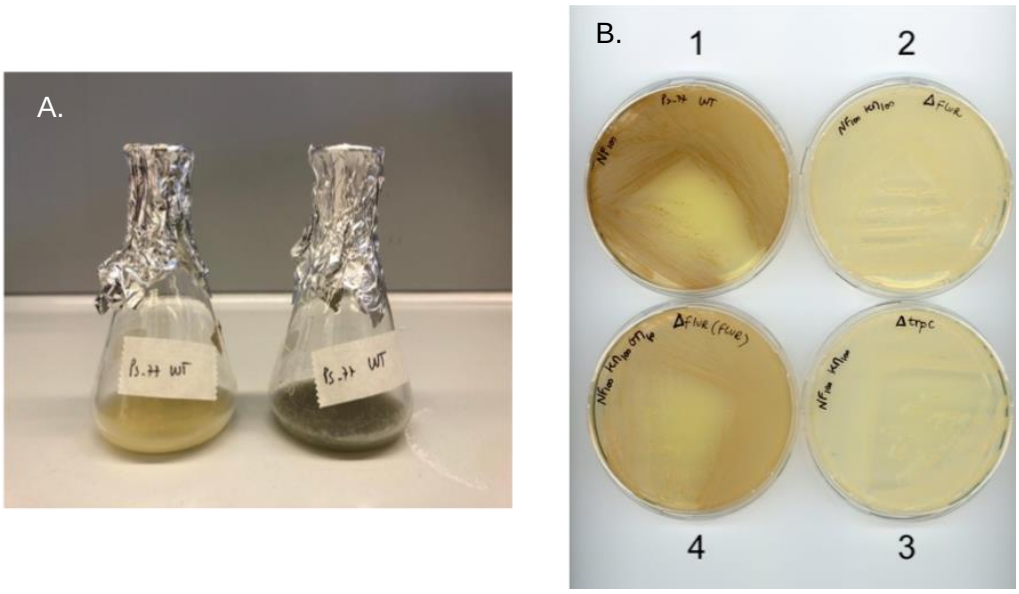


Figure 4. 6 (A) Pigment production assessment by *P. fluorescens* Ps_77 wild type in liquid under shaking and non-shaking growth condition, respectively. **(B)** Pigment production assessment by (1) *P. fluorescens* Ps_77 wild type, (2) *ΔfluR* and (3) *ΔtrpC-like* mutants and (4) *ΔfluR(fluR)* when grown in plate media, after 48 hrs.

Gene promoter studies established that *fluR* is negatively autoregulated since an increase of its promoter expression level was observed in the *ΔfluR* mutant compared to the wild type under shaking (**Figure 4.7 A**) and under static growth conditions (**Figure 4.7 B**).

On the other hand, a significant induction of the expression of the genetically adjacent biosynthetic operon has been detected in the wild type compared to the *ΔfluR* mutant (**Figure 4.7 C-D**), suggesting that this operon is under FluR regulation. Interestingly, the increase of the operon transcriptional levels was observed only under static growth conditions, while no significative differences were detected under shaking growth condition. As mentioned above, this correlates with the observation that *P. fluorescens* Ps_77 does not

produce pigment in liquid under shaking conditions; pigment production by the wild-type in liquid media was only visually observed when bacteria were grown in static conditions.

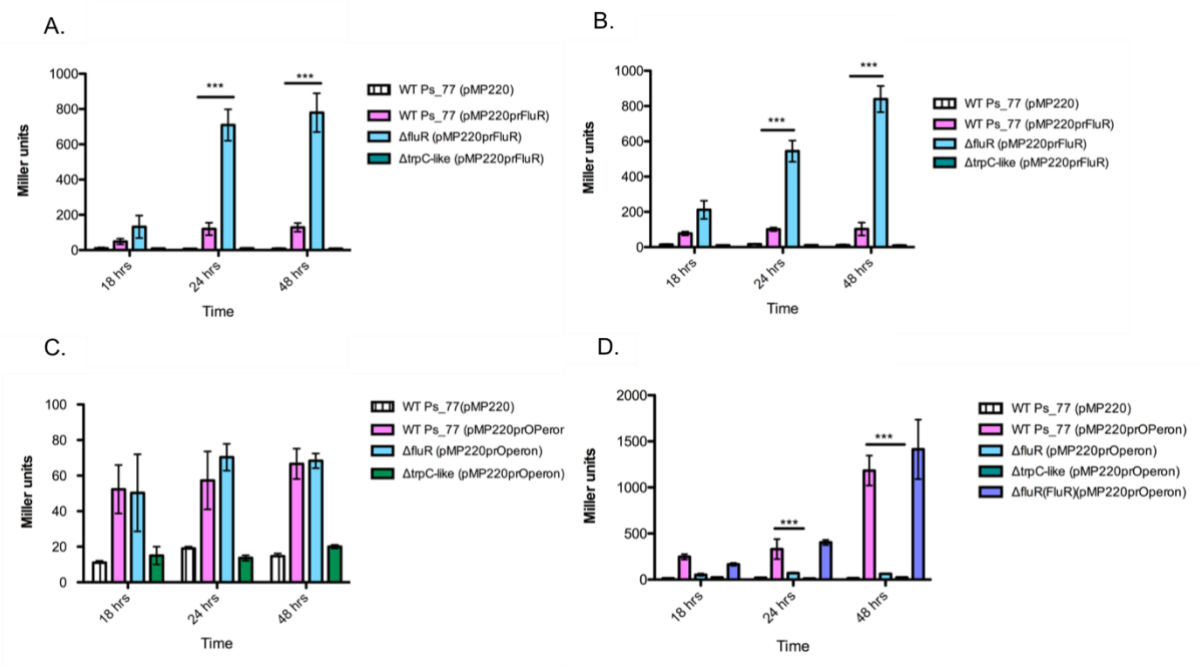


Figure 4. 7 Gene promoter activity in *P. fluorescens* Ps_77 WT, Δ*fluR and* Δ*trpC-like* . β-Galactosidase activities (Miller units) of the FluR gene promoter transcriptional fusion was determined in shaking **(A)** and static conditions **(B)** after 18, 24 and 48 hrs, comparing the expression levels between WT and Δ*fluR and* Δ*trpC-like* . Similarly, β-Galactosidase activities (Miller units) of the biosynthetic operon promoter transcriptional fusion was determined in shaking **(C)** and static conditions **(D)** after 18, 24 and 48 hrs, comparing the expression levels between WT and Δ*fluR and* Δ*trpC-like* . The WT with empty plasmid pMP220 was used as control. All experiments were performed in triplicate. Statistical analysis was calculated using one-way ANOVA followed by Tukey's multiple comparisons by Prism 7 (GraphPad Software, Inc.). The error bars indicate standard deviations.

Gene promoter activities of *fluR* and the biosynthetic operon were restored to the wild-type levels when the *ΔfluR* mutant was complemented *in trans* with the *fluR* gene cloned in a plasmid (**Figure 4.7D**). Moreover, in the *ΔtrpC-like* mutant any transcriptional activity neither of the biosynthetic operon genes or of the *fluR* was detected (**Figure 4.7**). These results suggest that the biosynthetic operon is involved in its transcriptional regulation, thus it was

hypothesized that the pigment synthetized by the large operon might be the signal-inducer for the FluR solo.

4.3.4 P. fluorescens Ps_77 possesses a LuxI/R system which is not involved in pigment production

P. fluorescens Ps_77 possesses an archetypical LuxI/R quorum sensing system hence it was of interest to determine whether it was involved in the regulation of the pigment possibly via the LuxR solo since the cartography studies indicated that it could bind AHLs (see above). The *luxI*-homolog gene has been mutated as described in Material and Methods section and the gene promoter activity of the biosynthetic operon and *fluR* has been determined in this mutant (**Figure 4.8**). No *luxI*-dependent gene promoter activity was detected, as no differences were noticed between the wild type or the *ΔluxI* mutant. In addition, the *luxI* null mutant was not impaired in pigment production.

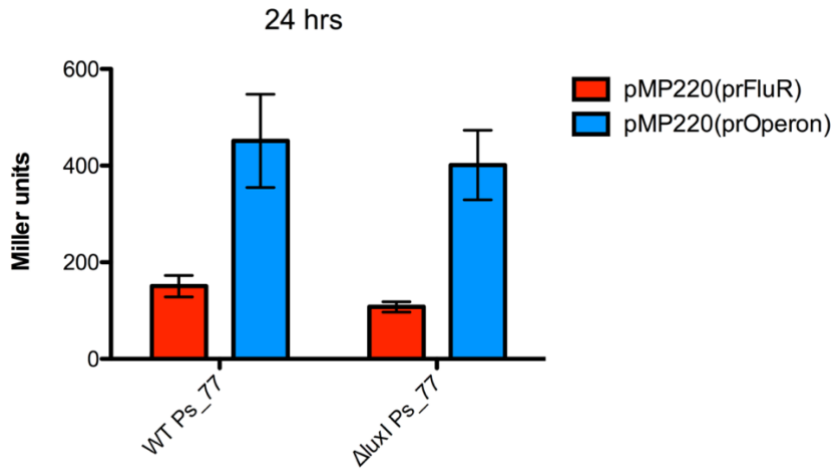


Figure 4. 8 Gene promoter activity in *P. fluorescens* Ps_77 and Δ*luxI.* β-Galactosidase activities (Miller units) of 2 gene promoter transcriptional fusion (*fluR, trpC*) were determined to compare the expression levels between WT and Δ*luxI.* The promoter activity was calculated after 24 hrs (log-phase). All experiments were performed in triplicate. Statistical analysis was calculated using one-way ANOVA followed by Tukey's multiple comparisons by Prism 7 (GraphPad Software, Inc.). The error bars indicate standard deviations.

In summary, these results suggested that FluR is involved in the transcriptional regulation of the adjacent pigment biosynthetic operon which is likely acting as a signal for the LuxR solo creating a positive feedback loop. Moreover, it was excluded any involvement of the complete QS LuxI/R system and AHL signal molecules in the regulation of the pigment by the LuxR solo.

4.3.5 The pigment biosynthetic operon is regulated by cell-cell signaling

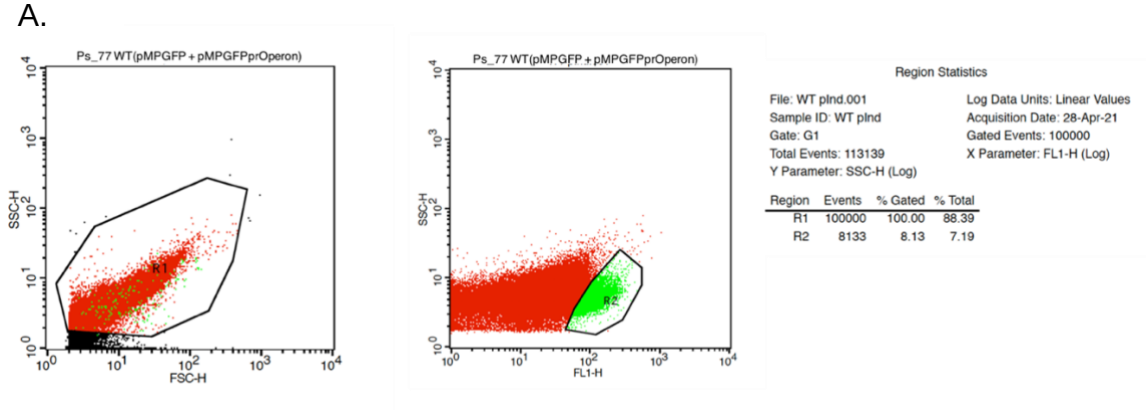
Since it was observed an induction of the biosynthetic operon expression in *P. fluorescens* Ps_77 wild-type and an inhibition in the *P. fluorescens* Ps_77 *ΔtrpC-like* mutant strain, we hypothesized that the pigment-like molecule was involved in the mechanism of regulation of the biosynthetic operon via FluR. In order to investigate this hypothesis and gather evidence that the system is capable of cell-cell signaling, the promoter transcriptional activity of the biosynthetic operon in the mutant *ΔtrpC-like*, which is no longer able to produce the pigment, was tested by a co-culture technique as a way to detect possible cell-to-cell signaling via diffusible signal molecules.

The *gfp* reporter gene was initially used for studying the biosynthetic operon gene promoter activities in *P. fluorescens* Ps_77 WT, *ΔfluR* and *ΔtrpC-like;* the expression of the *gfp* gene was detected by confocal microscopy and cytofluorimetry as described in Material and methods section. As expected, the expression of the biosynthetic operon was detected when grown under static conditions in the wild type strain Ps_77 (pMPGFPprOperon) carrying the promoter with the gfp transcriptional fusion co-cultured with Ps_77 harboring the empty reporter plasmid (pMPGFP).

The amount of the bacterial population expressing the *gfp* gene was around the 8.13%, as shown in **Figure 4.9 A.** The promoter activity was also observed by confocal microscopy, as

shown in **Figure 4.9 B**. This result is in line with gene promoter studies using the β-galactosidase reporter as described above where it was observed induction of expression of the biosynthetic operon in static conditions after 48 hrs of grown. These same growth conditions were applied in all other co-culture experiments.

A.



B.

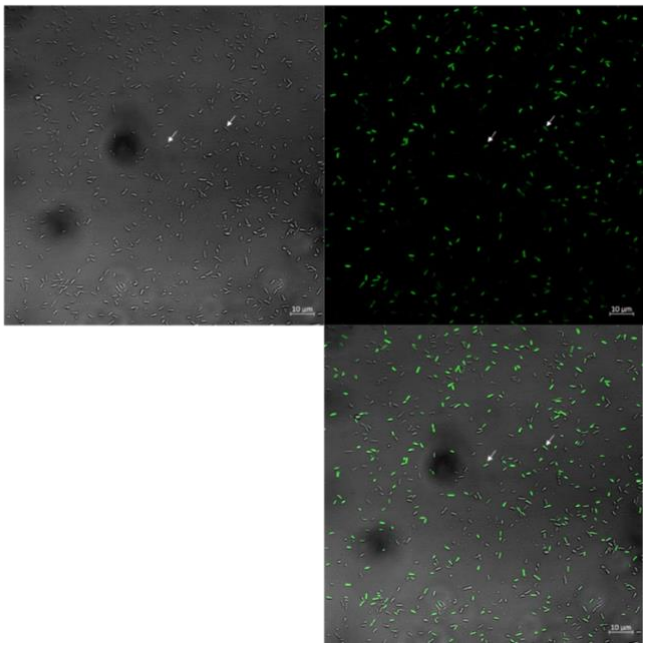


Figure 4. 9 The % of GFP-positive bacteria cells in the Ps_77 wild type strain harboring (pMPGFPprOperon) visualized by cytofluorimetry **(A)** and fluorescence microscopy **(B)** after 48 hrs of grown and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

Similarly, the expression of the biosynthetic operon was visualized and quantified in the mutant *P. fluorescens* Ps_77 *ΔfluR* (pMPGFPprOperon) co-cultured with *ΔfluR* (pMPGFP). The amount of the population expressing the *gfp* gene was around the 0.11%, as shown in **Figure 4.10 A** and observed in **Figure 4.10 B.** This result is a further confirmation that no expression of the biosynthetic operon occurs when the LuxR solo is absent.

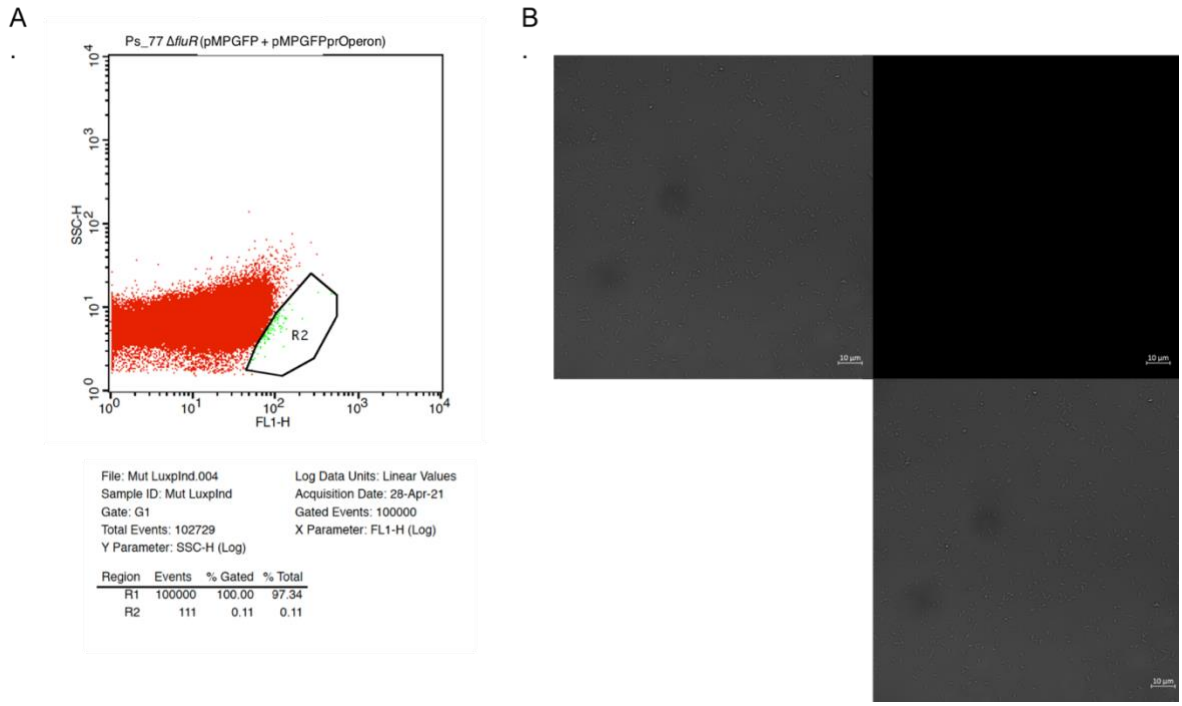


Figure 4. 10 The % of GFP-positive bacteria cells in the Ps_77 *ΔfluR* strain harboring (pMPGFPprOperon) visualized by cytofluorimetry (**A**) and fluorescence microscopy (**B**) after 48 hrs of grown and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

The expression of the biosynthetic operon, was visualized and quantified also in the mutant *P. fluorescens* Ps_77 *ΔtrpC-like* (pMPGFPprOperon) co-cultured with *ΔtrpC-like* (pMPGFP). Only the 0.25% of the population tested expressed the *gfp* gene, as shown in **Figure 4.11 A** and **B**, again indicating that this promoter is not active when the biosynthetic operon is mutated resulting in no pigment production.

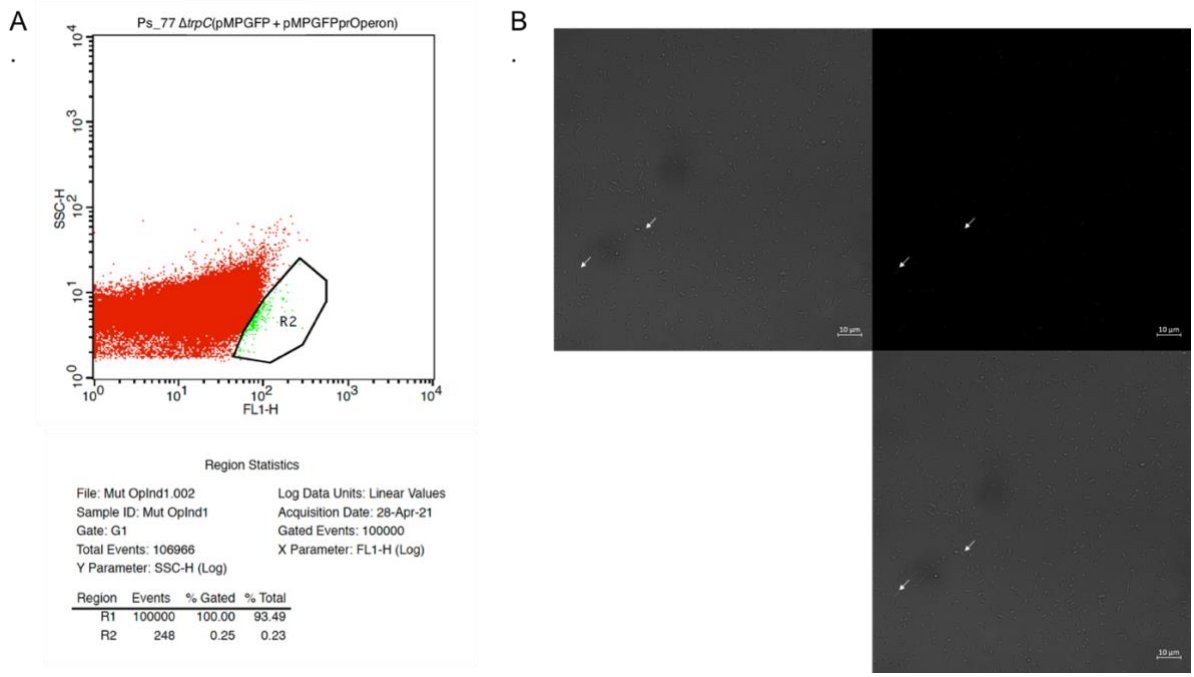


Figure 4. 11 The % of GFP-positive bacteria cells in Ps_77 *ΔtrpC-like* strain harboring (pMPGFPprOperon) visualized by cytofluorimetry **(A)** and fluorescence microscopy **(B)** after 48 hrs of grown and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

The expression of the biosynthetic operon was then visualized and quantified, in the mutant *P. fluorescens* Ps_77 *ΔtrpC-like* (pMPGFPprOperon) co-cultured with wild type Ps_77 (pMPGFP). The amount of the population expressing the *gfp* gene increased considerably around 2.42% (**Figure 4.12 A** and **B**). This result demonstrated that the presence of the wild type induces the expression of the biosynthetic operon promoter in the mutant *ΔtrpC-like* likely via LuxR solo binding.

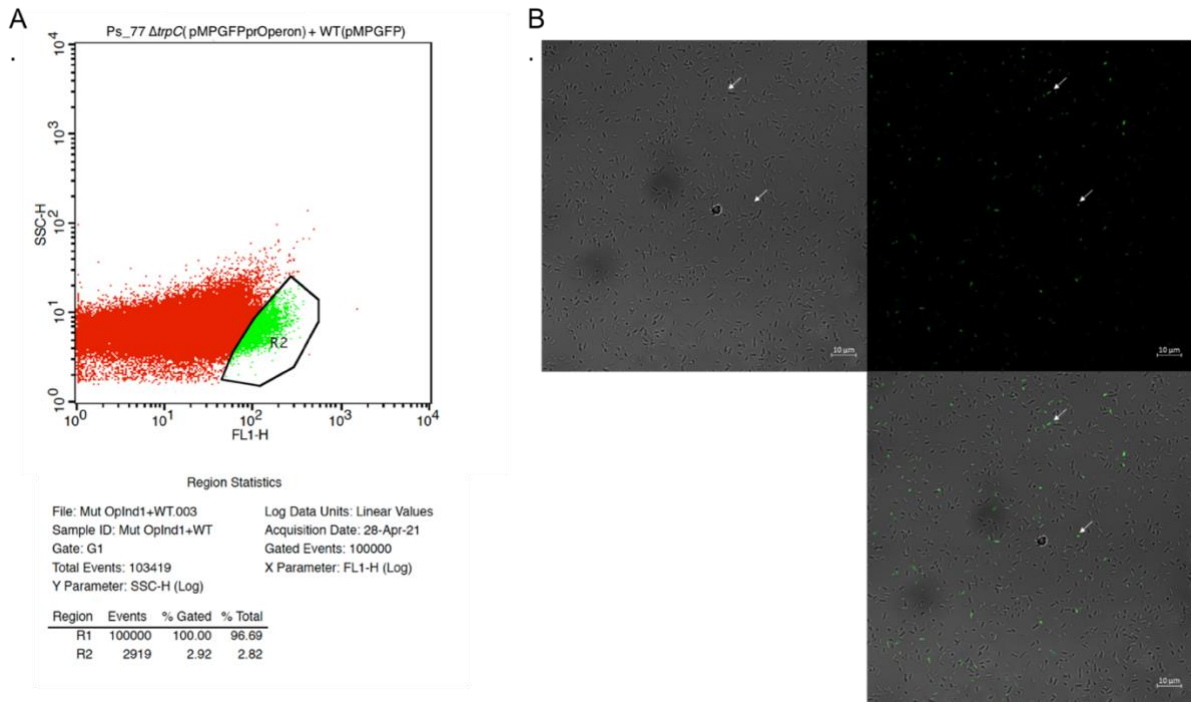


Figure 4. 12 The % of GFP-positive bacteria cells in Ps_77 *ΔtrpC-like* strain harboring (pMPGFPprOperon) and co-cultured with Ps_77 wild type strain visualized by cytofluorimetry **(A)** and fluorescence microscopy **(B)** after 48 hrs of grown and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

In order to further confirm this result, the expression of the biosynthetic operon was tested also in the mutant *P. fluorescens* Ps_77 *ΔfluR* (pMPGFPprOperon) co-cultured with the wild type Ps_77 (pMPGFP). As expected, no increase in the *gfp* reporter gene expression was detected visa the co-culture presence of the wild-type strain when the LuxR solo is not present (**Figure 4.13 A** and **B**).

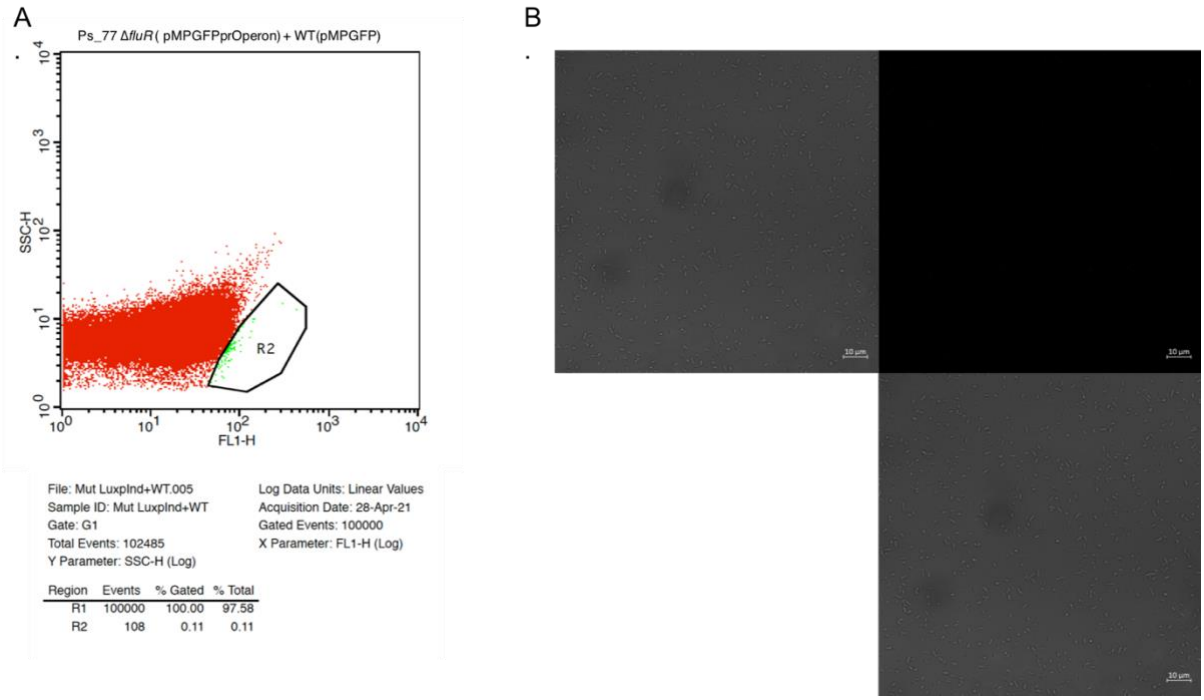
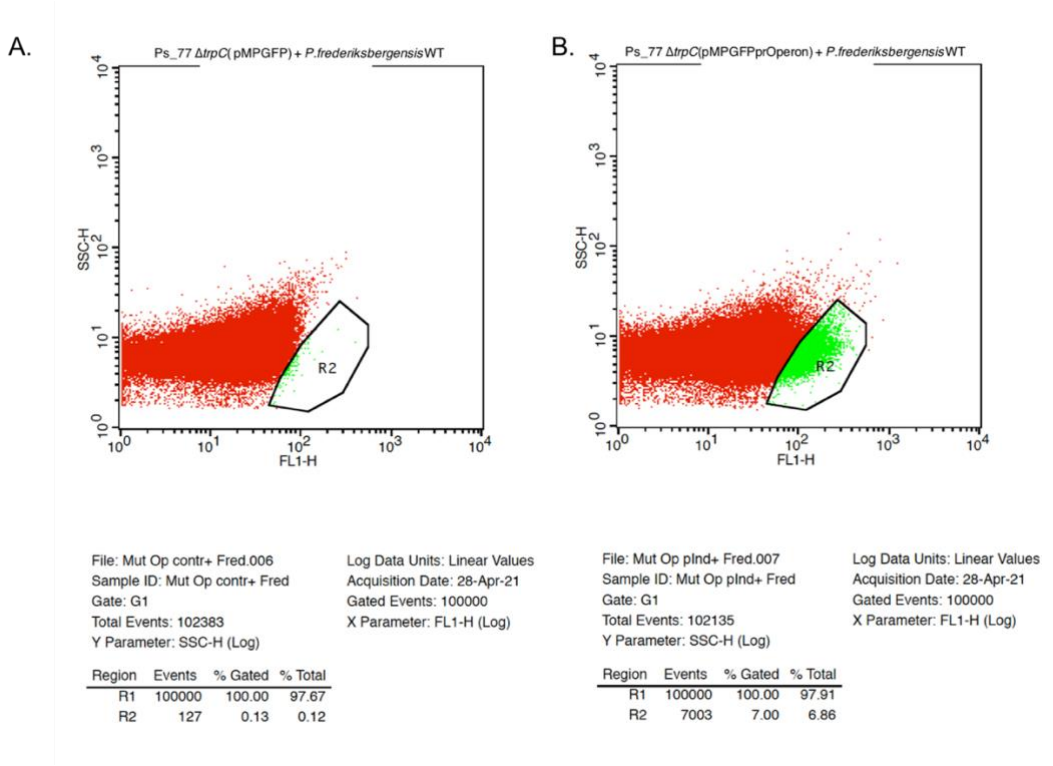


Figure 4. 13 The % of GFP-positive bacteria cells in Ps_77 *ΔfluR* strain harboring (pMPGFPprOperon) and co-cultured with Ps_77 wild type strain were visualized by cytofluorimetry **(A)** and fluorescence microscopy **(B)** after 48 hrs of grown and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

The co-culture experiment was also performed with another *Pseudomonas* isolate, namely *P. frederiksbergensis* OS210_3 which produces a similar pigment and possesses highly homologous locus to the biosynthetic operon and *luxR* solo (see above, **Figure 4.3 A**). The expression of the biosynthetic operon, was then visualized and quantified in the mutant *P. fluorescens* Ps_77 *ΔtrpC-like* (pMPGFPprOperon) co-cultured with *Ps. frederiksbergensis*

OS210_3 (pMPGFP). Compared to the control (**Figure 4.14 A**) the amount of population expressing the *gfp* gene increased around 7% (**Figure 4.14 B** and **C**). This result clearly indicated that the pigment-like molecule produced by *P. frederiksbergensis* OS210_3 significantly induced the LuxR solo of strain Ps_77.



C.

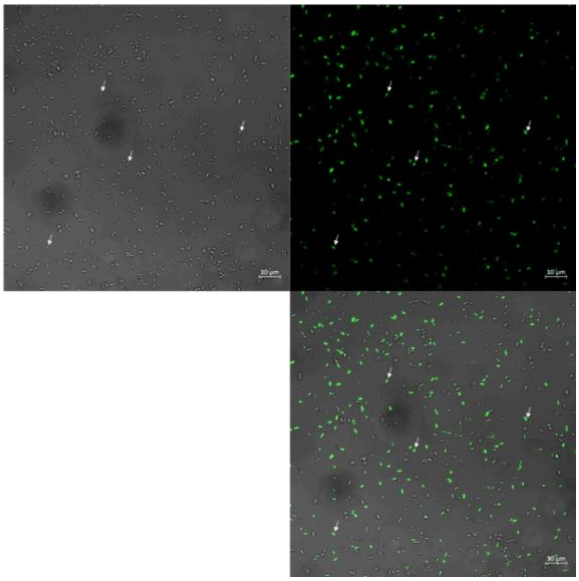


Figure 4. 14 The % of GFP-positive bacteria cells in Ps_77 *ΔtrpC-like* strain (pMPGFP) co-cultured with *P. frederiksbergensis* wild type strain was visualized by cytofluorimetry **(A)** and % of GFP-positive bacteria cells in Ps_77 *ΔtrpC-like* strain (pMPGFPprOperon) co-cultured with *P. frederiksbergensis* wild type strain was visualized by cytofluorimetry (B) and fluorescence microscopy **(C)** after 48 hrs of grown and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

To further investigate the diffusible nature of the pigment-like molecule and to demonstrate its involvement in cell-cell signaling, the same set-up described above has been performed however using the *P. fluorescens* Ps_77 wild type cell-free supernatant. The results obtained are described in the **Figure 4.15** are in line with the one presented above demonstrating induction of gene promoter activity**.**

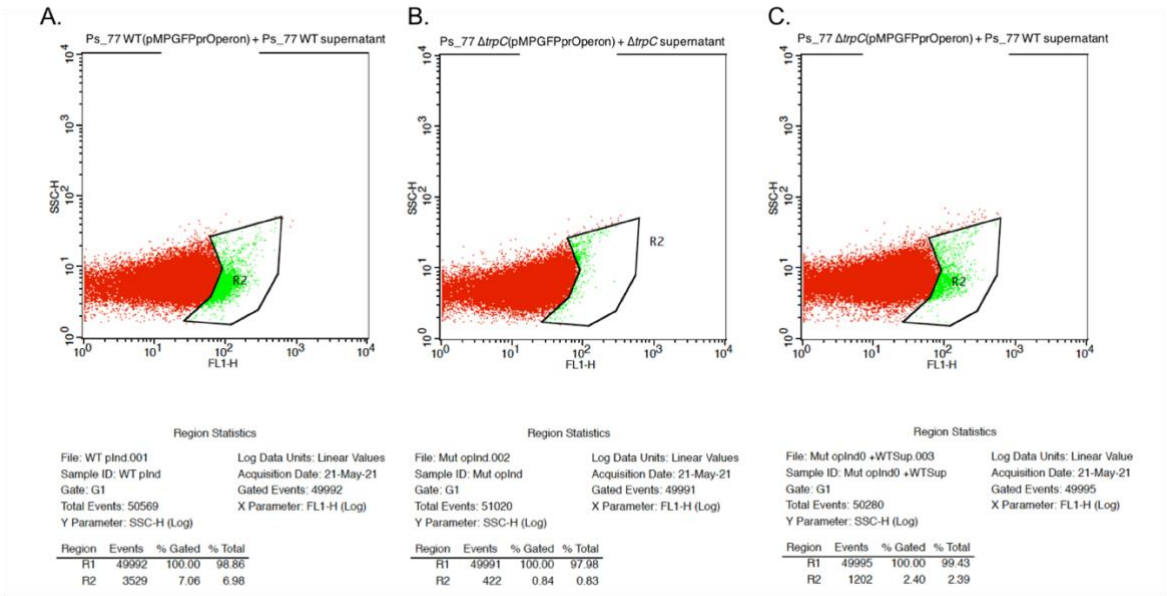


Figure 4. 15 The % of GFP-positive bacteria cells in *P. fluorescens* Ps_77 wild type (pMPGFPprOperon) supplemented with wild type cell-free supernatant (A); % of GFP-positive bacteria cells in Ps_77 *ΔtrpC-like* strain (pMPGFPprOperon) supplemented with Ps_77 *ΔtrpC-like* strain cell-free supernatant (B); % of GFP-positive bacteria cells in Ps_77 *ΔtrpC-like* strain (pMPGFPprOperon) supplemented with Ps_77 wild type strain cell-free supernatant (C). All the bacterial cultures were grown for 48 hrs and under pigment-production conditions. Imaging were performed using FITC (488 nm) filter.

In summary, these results indicate that the FluR solo responds to the pigment-like molecule via cell-cell signaling and is also involved in inter-species signaling being able to respond to a similar pigment-like molecule produced and released by other pigment-producing *Pseudomonas* species. In conclusion, it is believed that this represents a new type of bacterial cell-to-cell communication circuit.

4.4 Discussion

LuxR solos are very widespread among the *P. fluorescens* complex and from a set of approximately 600 genomes, they can be placed at least into nine putative sub-groups according to the sequence similarity, invariant amino acids of the IBD, and conservation of the flanking genes (Chapter 3; Bez et al., 2021). A few LuxR solos however do not cluster into these sub-groups having uncommon flanking genes and primary structure, suggesting that other far less common LuxR solo sub-groups exist.

In the present study, a unique *luxR* solo has been genetically and molecularly characterized. It is harbored by *P. fluorescens* Ps_77 and located adjacent to a large pigment biosynthetic gene cluster which together most likely constitutes a rare bacterial genomic island. In order to get insight into the molecular basis of this rare LuxR solo, designated as FluR, and to begin to understand its functional role and possible involvement in the regulation of the adjacent gene cluster, *in silico* and *in vivo* experiments have been performed. FluR is involved in the synthesis of a pigment-like molecule which in turn act as signal, inducing directly or indirectly for the LuxR solo and creating a positive feedback loop. The entire system could represent a new QS communication circuit formed by a LuxR solo and a pigment signal molecule.

P. fluorescens Ps_77 genome mining revealed the presence of ICEs elements (integrative and conjugative genes) located upstream and downstream of this putative genomic island harboring the *luxR* solo and pigment biosynthetic operon; the integration sites along with the significantly different GC contents, suggest that it might have been acquired by horizontal gene transfer (HGT) or via genomic rearrangement events. ICEs allow bacteria to rapidly adapt to new environmental conditions and to colonize new niches by promoting the

mobilization of genomic islands (Burrus and Waldor, 2004). Genomic islands are common in many *Pseudomonas* spp. (Qiu et al., 2006; Subramoni et al., 2015; Hesse et al., 2018) and reflects the diverse lifestyle of these ubiquitous bacteria (Matilla et al., 2007; Loper et al., 2012). Several BGCs have been previously described that are encoded near or contain a *luxR* homolog genes but it is not clear whether they are regulated by QS (Brotherton et al., 2018).The presence of a secondary metabolic gene cluster, a regulatory and transport genes and phage integrase/transposase sites might suggest that it constitutes an autonomous stand-alone genomic island, which is likely involved in the adaptation and/or survival of this bacterium in specific conditions.

Interestingly, comparative whole-genome analysis showed that this putative genomic island is very rare among bacteria being limited thus far to only 7 *Pseudomonas* spp. isolates; this low presence might be due to the high-energy cost required for the biosynthetic pathway, which in turn it might confer specific advantages in very particular ecological niches. Moreover, all the 7 homologous genomic islands harbored by different *Pseudomonas* spp. isolates, are located between two ICEs sites further supporting the hypothesis that it is a stand-alone genomic island.

In order to further confirm the autonomous existence and functional role of this putative genomic island, several attempts have been made to mobilize it in an heterologous *Pseudomonas* strain which does not possess it; unfortunately because of the large size, it is technically challenging and all attempts so far failed. This experiment can possibly additionally demonstrate that this is a stand-alone system being able to function in other bacteria.

Inactivation of either the regulator FluR or the genes in the biosynthetic operon abolished pigment production. In addition, results have shown that FluR regulates the expression of the

adjacent biosynthetic operon and consequently the production of the pigment. *P. frederiksbergenis* OS210_3 possesses a highly homologous locus and it also produces a pigment further indicating that this putative genomic island is necessary for pigment biosynthesis. Genetic complementation of the *fluR* mutant restored the ability of the Ps_77 strain to synthetize the pigment confirming the involvement of the LuxR solo in the regulation of the pigment.

The chemical structure of the pigment as well as its biological role and biosynthetic pathway are currently unknown. However, the large biosynthetic gene cluster harbors accessory copies of *trp* genes, which encode for enzymes involved in the tryptophan pathway and oxidative stress response. Tryptophan is the most complex and energy-consuming amino acid, characterized by an indole side chain, which is aromatic with a binuclear ring structure (Barik, 2020). Several *trp* derivatives are used by bacteria for specialized functions, such as production of quinolone, indole-3-acetic acid, violacein, serotonin, melatonin, tryptophol and several related indole derivatives (De Troch et al., 1997; Xie et al., 2003; Antônio and Creczynski-Pasa, 2004; Barik, 2020). In particular, indole can be oxidized to hydroxyindole by diverse oxygenases, forming a black/bluish-colored indole-derivative (Zarkan et al., 2020). The presence of genes involved in the tryptophan pathway and indole-formation in the biosynthetic cluster, suggests that the pigment could be likely an indole-based compound characterized by the presence of an indole moiety.

Pigmentation in bacteria is often an adaptative response that protects microorganisms from environmental stress conditions, such as ultraviolet radiation, toxic heavy metals or oxidative stress (Pavan et al., 2020). For example, the production of indole-based pigments is important for plant-associated bacteria during the first stages of root colonization, when

bacteria need to cope with reactive oxygen species (ROS) and phenolic compounds derived from plant defenses (Kim et al., 2000; Pavan et al., 2020).

Since it was determined that *P. fluorescens* Ps_77 produces the likely indole-derived pigment only in static growth conditions, when the oxygen concentration declines, it is a possibility that the pigment could have also an antioxidant function, acting as terminal electron acceptor for anaerobic respiration.

Attempts have been made to extract the pigment in order to get insight into its chemical structure; unfortunately, because of its insolubility in water and in several organic solvents, it was not possible to characterize. The peculiar chemical behavior could be explained by the presence of organic substituents on the indole-derived molecule, which are possibly responsible of its hydrophobicity. However, the diffusion of the blue/black pigment in the plate or liquid medium occurs despite its insolubility in water. Most probably, the diffusion of the pigment in the medium could be complexed with other compounds or due to substitutions or hydroxylation that convert it in a soluble form outside of the bacterial cell. It is also a possibility that the active form of the molecule is un-pigmented, being a physiological precursor, which is then modified or subsequently polymerized, under specific conditions and acquires pigment properties.

Modeling analysis of FluR showed high conservation of amino acid residues forming the binding pocket just like in *E. coli* SdiA, indicating that might bind AHLs. SdiA orthologs are well conserved among members of the *Enterobacteriaceae* family and in particular in *Escherichia*, *Kosakonia*, *Salmonella*, *Enterobacter*, *Citrobacter*, *Cronobacter*, *Klebsiella*, *Pantoea*, and *Erwinia* (Sabag-Daigle and Ahmer, 2012), while is quite rare among *Pseudomonas* genus, supporting further the possibility that this LuxR solo as well as the entire putative genomic island have been acquired by horizontally gene transfer event. Interestingly, SdiA homologs

detects and responds broadly to a wide range of exogenous AHLs signals produced by neighboring bacteria (Michael et al., 2001; Smith and Ahmer, 2003); however our results demonstrated that FluR detects and responds to an endogenously produced self-signal which very likely is not an AHL. This is also evidenced by the fact that the QS LuxI/R system, harbored by *P. fluorescens* Ps_77, is not involved in the regulation of the pigment via the LuxR solo.

The genetic and molecular findings reported here support the hypothesis that the FluR inducer is the indole-based pigment or its derivative/precursor synthesized by the FluR target biosynthetic operon. The current working model is therefore that FluR detects and possibly binds the indole-derivative pigment and autoregulates itself and regulates the directly adjacent operon, creating a positive feedback loop, as schematized in the **Figure 4.16**.

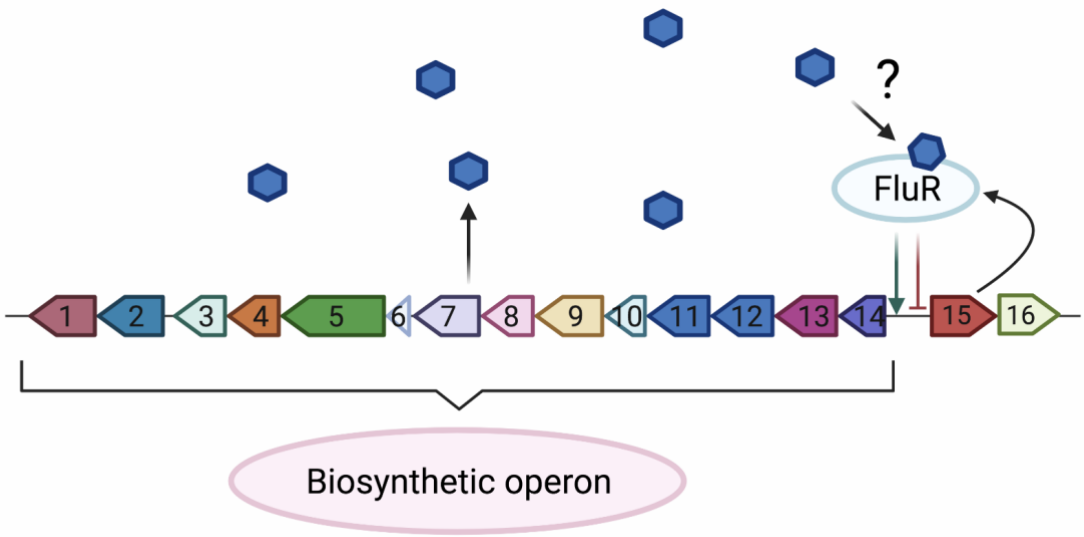


Figure 4. 16 Working model describing the mechanism of regulation of the putative genomic island. Figure created using Biorender.com

Emerging data suggests that indole and its derivatives may act as intercellular, interspecies and interkingdom signaling molecules (Lee et al., 2015; Zarkan et al., 2020), playing important roles in bacterial pathogenesis and physiology such as biofilm and spore formation, virulence and plasmid stability (Lee and Lee, 2010). Several bacteria and some plants are able to

produce indole-like compounds; however the genetic and molecular mechanisms of indole signaling remains unclear and controversial (Lee et al., 2008; Lee et al., 2009). It is currently disputed whether indole is a signaling molecule or it is just an intercellular sub-product of bacteria metabolism (Wang et al., 2001; Winzer et al., 2002; Chant and Summers, 2007). Several studies demonstrated that at high concentration, indole can interfere with AHL-mediated quorum sensing, such as in *P. aeruginosa* (Lee et al., 2009). Intriguingly, the *E. coli* SdiA has been reported to be involved in indole signaling (Lee et al., 2008; Lee et al., 2009; Lee and Lee, 2010; Lee et al., 2015); in fact in a *sdiA* mutant, the effect of indole on biofilm formation was lost, and no significant changes of gene expression with a response to indole were shown (Lee et al., 2008). It is possible that SdiA responds to indole in a promiscuous way or that other abiotic stimuli are needed and indole act as endogenously produced intracellular messenger. However, there is no evidence of the direct binding of indole to SdiA, and it remains unclear if and how SdiA interact in regulating gene expression.

Based on the co-culture experiments, it is believed that this is a novel bacterial cell-cell communication mechanism occurring most probably via an indole-based pigment which acts as a signal inducer for the LuxR solo. Several signaling molecules beside AHLs have been implicated in QS-regulated processes, like aryl-HSLs, photopyrones, dialkylresorcinols, 2-heptyl-3-hydroxy-4-quinolone (PQS) and (2-hydroxyphenyl)-thiazole-4-carbaldehyde (IQS). Similarly the LuxR solos are emerging as important players in cell-cell communication circuits (Schaefer et al., 2008; Brachmann et al., 2013; Lee et al., 2013; Brameyer et al., 2015b). Hence, many novel LuxR solo based cell-cell communication circuits very likely await discovery.

Moreover, we demonstrated that this novel molecule can also function as an inter-species signal since it was observed that it induces gene expression in a different bacterial species. It

is common for bacteria that secondary metabolites with a known function act as 'repurposed' compounds, functioning as signals mediating inter-species crosstalk (Shank and Kolter, 2009). For example, the diffusible signal factor (DSF) from *Xanthomonas campestris*—cis-11-methyl-2-dodecenoic acid— and the BDSF-cis-2-dodecenoic acid from *Burkholderia cenocepacia* are intra-species signals controlling biofilm formation and virulence, however, they can also act as interspecies signals and block the *C. albicans* filamentation by interfering with farnesol signaling (Wang et al., 2004; Torres et al., 2007).

In summary, a putative novel cell-cell communication system that most probably uses an indole-based pigment as a signaling compound was discovered. It is possible that this system has evolved away from canonical AHL-QS system and the LuxR solo became able to promiscuously detect molecules. However, it is disputable whether FluR can be designated as LuxR solo since its cognate signal synthase seems to be the adjacent pigment biosynthesis cluster. Further research is needed to identify the chemical nature of the pigment and its interaction with the LuxR solo.

4.5 References

Ahmer, B.M. (2004) Cell-to-cell signalling in Escherichia coli and Salmonella enterica. *Molecular microbiology* **52**: 933-945.

Ahmer, B.M., Van Reeuwijk, J., Timmers, C.D., Valentine, P.J., and Heffron, F. (1998) Salmonella typhimurium encodes an SdiA homolog, a putative quorum sensor of the LuxR family, that regulates genes on the virulence plasmid. *Journal of bacteriology* **180**: 1185-1193.

Alexeyev, M.F. (1999) The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. *Biotechniques* **26**: 824-828.

Andreani, N., Martino, M., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R. et al. (2015a) Reprint of 'Tracking the blue: a MLST approach to characterise the Pseudomonas fluorescens group'. *Food microbiology* **45**: 148-158.

Andreani, N.A., Carraro, L., Zhang, L., Vos, M., and Cardazzo, B. (2019) Transposon mutagenesis in Pseudomonas fluorescens reveals genes involved in blue pigment production and antioxidant protection. *Food microbiology* **82**: 497-503.

Andreani, N.A., Carraro, L., Martino, M.E., Fondi, M., Fasolato, L., Miotto, G. et al. (2015b) A genomic and transcriptomic approach to investigate the blue pigment phenotype in Pseudomonas fluorescens. *International journal of food microbiology* **213**: 88-98.

Antônio, R.V., and Creczynski-Pasa, T.B. (2004) Genetic analysis of violacein biosynthesis by Chromobacterium violaceum. *Genet Mol Res* **3**: 85-91.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B. et al. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic acids research* **34**: W604-W608.

Barik, S. (2020) The uniqueness of tryptophan in biology: Properties, metabolism, interactions and localization in proteins. *International journal of molecular sciences* **21**: 8776.

Better, M., Lewis, B., Corbin, D., Ditta, G., and Helinski, D.R. (1983) Structural relationships among Rhizobium meliloti symbiotic promoters. *Cell* **35**: 479-485.

Bez, C., Covaceuzach, S., Bertani, I., Choudhary, K.S., and Venturi, V. (2021) LuxR Solos from Environmental Fluorescent Pseudomonads. *Msphere* **6**: e01322-01320.

Brachmann, A.O., Brameyer, S., Kresovic, D., Hitkova, I., Kopp, Y., Manske, C. et al. (2013) Pyrones as bacterial signaling molecules. *Nature chemical biology* **9**: 573-578.

Brameyer, S., Bode, H.B., and Heermann, R. (2015a) Languages and dialects: bacterial communication beyond homoserine lactones. *Trends in microbiology* **23**: 521-523.

Brameyer, S., Kresovic, D., Bode, H.B., and Heermann, R. (2015b) Dialkylresorcinols as bacterial signaling molecules. *Proceedings of the National Academy of Sciences* **112**: 572-577.

Brotherton, C.A., Medema, M.H., and Greenberg, E.P. (2018) luxR homolog-linked biosynthetic gene clusters in Proteobacteria. *MSystems* **3**: e00208-00217.

Buchan, D.W., Ward, S., Lobley, A.E., Nugent, T., Bryson, K., and Jones, D.T. (2010) Protein annotation and modelling servers at University College London. *Nucleic acids research* **38**: W563-W568.

Buenavista, M.T., Roche, D.B., and McGuffin, L.J. (2012) Improvement of 3D protein models using multiple templates guided by single-template model quality assessment. *Bioinformatics* **28**: 1851-1857.

Burrus, V., and Waldor, M.K. (2004) Shaping bacterial genomes with integrative and conjugative elements. *Research in microbiology* **155**: 376-386.

Case, R.J., Labbate, M., and Kjelleberg, S. (2008) AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. *The ISME journal* **2**: 345-349.

Chant, E.L., and Summers, D.K. (2007) Indole signalling contributes to the stable maintenance of Escherichia coli multicopy plasmids. *Molecular microbiology* **63**: 35-43.

Chen, G., Jeffrey, P.D., Fuqua, C., Shi, Y., and Chen, L. (2007) Structural basis for antiactivation in bacterial quorum sensing. *Proceedings of the National Academy of Sciences* **104**: 16474-16479.

Choi, S., and Greenberg, E. (1991) The C-terminal region of the Vibrio fischeri LuxR protein contains an inducer-independent lux gene activating domain. *Proceedings of the National Academy of Sciences* **88**: 11115-11119.

Chugani, S.A., Whiteley, M., Lee, K.M., D'Argenio, D., Manoil, C., and Greenberg, E. (2001) QscR, a modulator of quorum-sensing signal synthesis and virulence in Pseudomonas aeruginosa. *Proceedings of the national academy of sciences* **98**: 2752-2757.

Cornelis, P., and Matthijs, S. (2002) Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. *Environmental Microbiology* **4**: 787-798.

Covaceuszach, S., Degrassi, G., Venturi, V., and Lamba, D. (2013) Structural insights into a novel interkingdom signaling circuit by cartography of the ligand-binding sites of the homologous quorum sensing LuxR-family. *International journal of molecular sciences* **14**: 20578-20596.

da Silva, D.P., Castañeda-Ojeda, M.P., Moretti, C., Buonaurio, R., Ramos, C., and Venturi, V. (2014) Bacterial multispecies studies and microbiome analysis of a plant disease. *Microbiology* **160**: 556-566.

De Troch, P., Dosselaere, F., Keijers, V., de Wilde, P., and Vanderleyden, J. (1997) Isolation and characterization of the Azospirillum brasilense trpE (G) gene, encoding anthranilate synthase. *Current microbiology* **34**: 27-32.

Devescovi, G., Kojic, M., Covaceuszach, S., Cámara, M., Williams, P., Bertani, I. et al. (2017) Negative regulation of violacein biosynthesis in Chromobacterium violaceum. *Frontiers in microbiology* **8**: 349.

Dreier, J., and Ruggerone, P. (2015) Interaction of antibacterial compounds with RND efflux pumps in *Pseudomonas aeruginosa*. *Frontiers in microbiology* **6**: 660.

Fuqua, C. (2006) The QscR quorum-sensing regulon of Pseudomonas aeruginosa: an orphan claims its identity. *Journal of bacteriology* **188**: 3169-3171.

Garrido-Sanz, D., Meier-Kolthoff, J.P., Göker, M., Martín, M., Rivilla, R., and Redondo-Nieto, M. (2016) Genomic and genetic diversity within the Pseudomonas fluorescens complex. *PloS one* **11**: e0150183.

Garrido-Sanz, D., Arrebola, E., Martínez-Granero, F., García-Méndez, S., Muriel, C., Blanco-Romero, E. et al. (2017) Classification of isolates from the Pseudomonas fluorescens complex into phylogenomic groups based in group-specific markers. *Frontiers in microbiology* **8**: 413.

Hesse, C., Schulz, F., Bull, C.T., Shaffer, B.T., Yan, Q., Shapiro, N. et al. (2018) Genome-based evolutionary history of *Pseudomonas spp*. *Environmental microbiology* **20**: 2142-2159.

Hoang, T.T., Karkhoff-Schweizer, R.R., Kutchma, A.J., and Schweizer, H.P. (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants. *Gene* **212**: 77-86.

Höfte, M., and Altier, N. (2010) Fluorescent pseudomonads as biocontrol agents for sustainable agricultural systems. *Research in Microbiology* **161**: 464-471.

Kim, Y.C., Miller, C.D., and Anderson, A.J. (2000) Superoxide dismutase activity in Pseudomonas putida affects utilization of sugars and growth on root surfaces. *Applied and environmental microbiology* **66**: 1460-1467.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop II, R.M., and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* **35**: 1547-1549.

Lee, J., Attila, C., Cirillo, S.L., Cirillo, J.D., and Wood, T.K. (2009) Indole and 7-hydroxyindole diminish Pseudomonas aeruginosa virulence. *Microbial biotechnology* **2**: 75-90.

Lee, J., Zhang, X.-S., Hegde, M., Bentley, W.E., Jayaraman, A., and Wood, T.K. (2008) Indole cell signaling occurs primarily at low temperatures in Escherichia coli. *The ISME journal* **2**: 1007-1023.

Lee, J., Wu, J., Deng, Y., Wang, J., Wang, C., Wang, J. et al. (2013) A cell-cell communication signal integrates quorum sensing and stress response. *Nature chemical biology* **9**: 339-343.

Lee, J.-H., and Lee, J. (2010) Indole as an intercellular signal in microbial communities. *FEMS microbiology reviews* **34**: 426-444.

Lee, J.-H., Wood, T.K., and Lee, J. (2015) Roles of indole as an interspecies and interkingdom signaling molecule. *Trends in microbiology* **23**: 707-718.

Lequette, Y., Lee, J.-H., Ledgham, F., Lazdunski, A., and Greenberg, E.P. (2006) A distinct QscR regulon in the Pseudomonas aeruginosa quorum-sensing circuit. *Journal of bacteriology* **188**: 3365-3370.

Lintz, M.J., Oinuma, K.-I., Wysoczynski, C.L., Greenberg, E.P., and Churchill, M.E. (2011) Crystal structure of QscR, a Pseudomonas aeruginosa quorum sensing signal receptor. *Proceedings of the National Academy of Sciences* **108**: 15763-15768.

Loper, J.E., Hassan, K.A., Mavrodi, D.V., Davis, E.W., Lim, C.K., Shaffer, B.T. et al. (2012) Comparative genomics of plant-associated Pseudomonas spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. *PLoS genetics* **8**: e1002784.

Matilla, M.A., Espinosa-Urgel, M., Rodríguez-Herva, J.J., Ramos, J.L., and Ramos-González, M.I. (2007) Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome biology* **8**: 1-13.

Michael, B., Smith, J.N., Swift, S., Heffron, F., and Ahmer, B.M. (2001) SdiA of Salmonella enterica is a LuxR homolog that detects mixed microbial communities. *Journal of bacteriology* **183**: 5733-5742.

Miller, J.H. (1972) Experiments in molecular genetics.

Nguyen, Y., Nguyen, N.X., Rogers, J.L., Liao, J., MacMillan, J.B., Jiang, Y., and Sperandio, V. (2015) Structural and mechanistic roles of novel chemical ligands on the SdiA quorum-sensing transcription regulator. *MBio* **6**: e02429-02414.

Patankar, A.V., and González, J.E. (2009) An orphan LuxR homolog of Sinorhizobium meliloti affects stress adaptation and competition for nodulation. *Applied and environmental microbiology* **75**: 946-955.

Pavan, M.E., López, N.I., and Pettinari, M.J. (2020) Melanin biosynthesis in bacteria, regulation and production perspectives. *Applied microbiology and biotechnology* **104**: 1357-1370.

Qiu, X., Gurkar, A.U., and Lory, S. (2006) Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* **103**: 19830-19835.

Sabag-Daigle, A., and Ahmer, B.M. (2012) ExpI and PhzI are descendants of the long lost cognate signal synthase for SdiA.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*: Cold spring harbor laboratory press.

Schaefer, A.L., Greenberg, E., Oliver, C.M., Oda, Y., Huang, J.J., Bittan-Banin, G. et al. (2008) A new class of homoserine lactone quorum-sensing signals. *Nature* **454**: 595-599.

Schrödinger, L. (2010) The PyMOL molecular graphics system, version 1.3 r1. In: August.

Shadel, G., Young, R., and Baldwin, T. (1990) Use of regulated cell lysis in a lethal genetic selection in Escherichia coli: identification of the autoinducer-binding region of the LuxR protein from Vibrio fischeri ATCC 7744. *Journal of bacteriology* **172**: 3980-3987.

Shank, E.A., and Kolter, R. (2009) New developments in microbial interspecies signaling. *Current opinion in microbiology* **12**: 205-214.

Silby, M.W., Cerdeño-Tárraga, A.M., Vernikos, G.S., Giddens, S.R., Jackson, R.W., Preston, G.M. et al. (2009) Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. *Genome biology* **10**: 1-16.

Slock, J., VanRiet, D., Kolibachuk, D., and Greenberg, E. (1990) Critical regions of the Vibrio fischeri LuxR protein defined by mutational analysis. *Journal of Bacteriology* **172**: 3974-3979.

Smith, J.N., and Ahmer, B.M. (2003) Detection of other microbial species by Salmonella: expression of the SdiA regulon. *Journal of bacteriology* **185**: 1357-1366.

Spaink, H.P., Okker, R.J., Wijffelman, C.A., Pees, E., and Lugtenberg, B.J. (1987) Promoters in the nodulation region of the Rhizobium leguminosarum Sym plasmid pRL1JI. *Plant molecular biology* **9**: 27-39.

Sperandio, V. (2010) SdiA sensing of acyl-homoserine lactones by enterohemorrhagic E. coli (EHEC) serotype O157: H7 in the bovine rumen. *Gut microbes* **1**: 432-435.

Stachel, S., An, G., Flores, C., and Nester, E. (1985) A Tn3 lacZ transposon for the random generation of beta-galactosidase gene fusions: application to the analysis of gene expression in Agrobacterium. *The EMBO Journal* **4**: 891-898.

Subramoni, S., and Venturi, V. (2009) PpoR is a conserved unpaired LuxR solo of Pseudomonas putida which binds N-acyl homoserine lactones. *BMC microbiology* **9**: 1-15.

Subramoni, S., Florez Salcedo, D.V., and Suarez-Moreno, Z.R. (2015) A bioinformatic survey of distribution, conservation, and probable functions of LuxR solo regulators in bacteria. *Frontiers in cellular and infection microbiology* **5**: 16.

Subramoni, S., Gonzalez, J.F., Johnson, A., Péchy-Tarr, M., Rochat, L., Paulsen, I. et al. (2011) Bacterial subfamily of LuxR regulators that respond to plant compounds. *Applied and environmental microbiology* **77**: 4579-4588.

Taylor, W.R. (2000) Protein structure comparison using SAP. In *Protein structure prediction*: Springer, pp. 19-32.

Torres, P.S., Malamud, F., Rigano, L.A., Russo, D.M., Marano, M.R., Castagnaro, A.P. et al. (2007) Controlled synthesis of the DSF cell–cell signal is required for biofilm formation and virulence in Xanthomonas campestris. *Environmental Microbiology* **9**: 2101-2109.

Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R. et al. (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *The EMBO journal* **21**: 4393-4401.

Venturi, V., and Ahmer, B.M. (2015) LuxR solos are becoming major players in cell–cell communication in bacteria. *Frontiers in cellular and infection microbiology* **5**: 89.

Wallner, B., and Elofsson, A. (2003) Can correct protein models be identified? *Protein science* **12**: 1073-1086.

Wang, D., Ding, X., and Rather, P.N. (2001) Indole can act as an extracellular signal in Escherichia coli. *Journal of Bacteriology* **183**: 4210-4216.

Wang, L.H., He, Y., Gao, Y., Wu, J.E., Dong, Y.H., He, C. et al. (2004) A bacterial cell–cell communication signal with cross-kingdom structural analogues. *Molecular microbiology* **51**: 903-912.

Winzer, K., Hardie, K.R., and Williams, P. (2002) Bacterial cell-to-cell communication: sorry, can't talk now—gone to lunch! *Current opinion in microbiology* **5**: 216-222.

Xie, G., Keyhani, N.O., Bonner, C.A., and Jensen, R.A. (2003) Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiology and Molecular Biology Reviews* **67**: 303-342.

Zarkan, A., Liu, J., Matuszewska, M., Gaimster, H., and Summers, D.K. (2020) Local and universal action: the paradoxes of indole signalling in bacteria. *Trends in Microbiology* **28**: 566-577.

Zhang, Y., and Skolnick, J. (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic acids research* **33**: 2302-2309.

Zou, Y., and Nair, S.K. (2009) Molecular basis for the recognition of structurally distinct autoinducer mimics by the Pseudomonas aeruginosa LasR quorum-sensing signaling receptor. *Chemistry & biology* **16**: 961-970.

Chapter V

Summarizing discussion

In nature bacteria most commonly live in polymicrobial communities rather than in isolation or monoculture. These communities can be complex, with high species richness and unevenness, and their structures are continually influenced by changing environmental factors. In the last decade, next-generation-sequencing and computational biology has unraveled the different bacterial groups populating diverse environmental niches. Plant microbiomes, for example, are always dominated by bacteria which mainly belong to four bacterial phyla. This phylogenetic conservation of the bacterial community composition infers an organized assembly of microbiomes which is directed by mechanisms which are at large unknown. Interactions are the fulcrum of communities; they are required for microbiome development, shaping and maintenance. Possible mechanisms of interaction include metabolic interactions, competition for resources, cell-cell signaling, cell-cell contacts and production of secondary metabolites which can negatively or positively affect the presence of specific members of the bacterial community. Very likely, several of these mechanisms will overlap working together (**Figure 5.1**). A plethora of studies have shown the presence and role of cell-cell signaling demonstrating that bacteria are capable of exchanging different chemical signaling molecules to communicate, regulate and synchronize their behaviors; however the majority of signals and molecular mechanisms are currently unknown.

Quorum Sensing (QS) is one type of social interaction among bacteria, which regulates gene expression in response to cell density, playing a major role in the formation and stability of microbial populations. Bacteria are constantly adapting to the chemical language of neighboring species and program gene expression accordingly. For this reason, AHL-QS gene circuits may be a flexible genetic tool used by bacteria to adapt to new community members,

new languages and ever-changing environments. Despite the prevalence of AHL QS among Proteobacteria, the understanding of its role and other cell-cell signaling mechanisms at the community level is at large unknown. LuxR solos are emerging as important players in cell–cell bacterial and inter–kingdom signaling as they result in alternative ways of cell-cell communication. They are extremely widespread among Proteobacteria, however they are receiving little attention by the scientific community. They are likely to play an important role in adapting to novel chemical languages establishing new communication networks, thanks to their flexibility to recognize different signals.

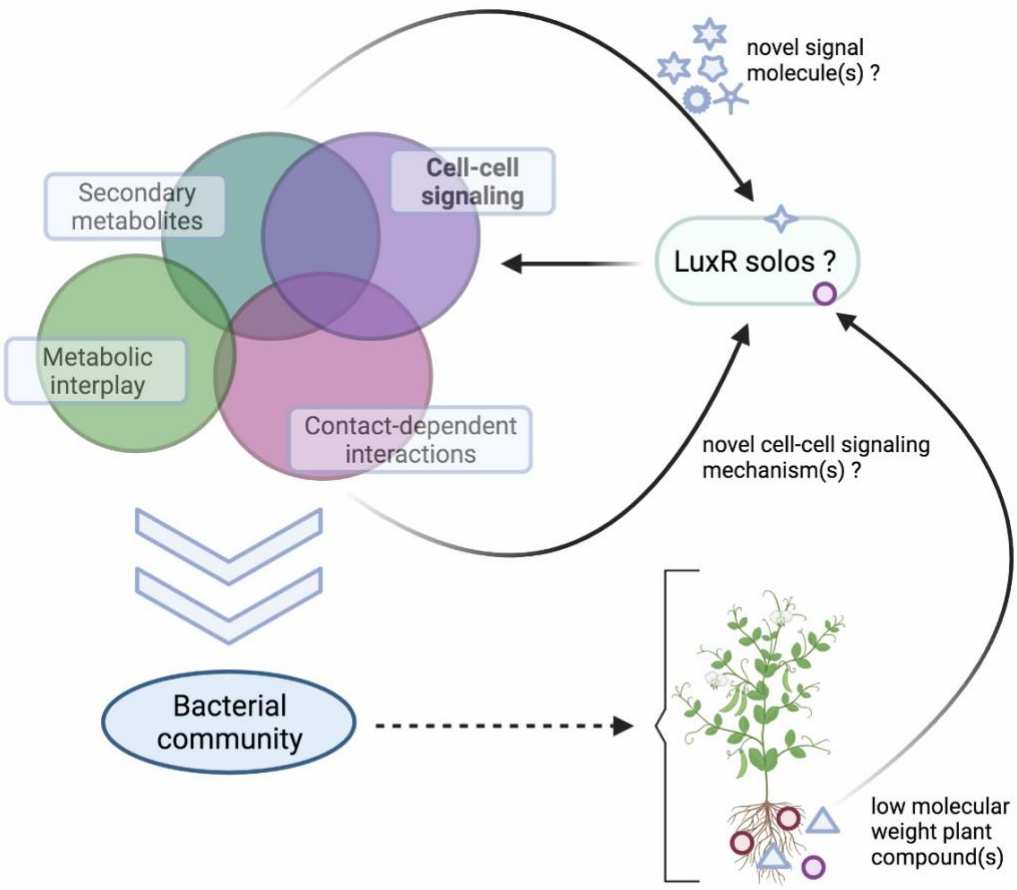


Figure 5. 1 Diagrammatic representation of the key mechanisms of biotic interactions among bacteria in the plant microbiome. Several secondary metabolites can also act as cell-cell signaling molecules as well as being transported by contact-dependent interactions. Similarly, cell-cell signaling molecules can be moved from cell-cell via contact dependent mechanisms. The LuxR solo may be a key player in cell-cell signaling, mediating interactions among bacteria. Figure created using Biorender.com

Unravelling which are the signaling molecules and communication systems/circuits that may influence group behavior and population dynamics will be very informative to understand how natural microbiomes are established; LuxR solos are likely to be important players in this event.

All the results presented in this thesis are advancements in understanding the nature of bacterial cell-cell interactions among members of multispecies communities, via NGS and molecular studies.

5.1 Pathobiome studies of a bacterial plant disease provide insights into possible biotic interactions between the plant microbiome and the incoming pathogen

The Chapter 1 presents pathobiome analysis of the rice foot rot disease, caused by *D. zeae*, as a model in order to investigate the effects of this bacterial pathogen to the total resident microbiome and to highlight possible interactions between the pathogen and the members of the microbiota. *D. zeae* alters the resident bacterial community in species composition, abundance and richness leading to the formation of a microbial consortia linked to the disease state. Network of interactions and LDA analysis allows the identification of likely positive interactions between the pathogen and members of the pathobiome that are consistent in the two growing-seasons. Specific anaerobic bacterial taxa, which were significantly co-present with the pathogen over the two years, were detected, suggesting a possible involvement in the disease development. Culture-dependent methods and *in planta* studies were performed with a *Burkholderia* sp. strain, which significantly correlated with *D.*

zeae, according to the LDA analysis. Intriguingly, its plant colonization ability significantly increased when *D. zeae* was present, indicating that it benefits from the presence of the pathogen. These pathobiome studies showed that they are suitable for identifying key bacterial players and probable bacterial biotic interactions that are taking place in the plant microbiome; however, the combination of traditional bacterial genetics and/or molecular approaches together with up-to-the-minute novel technologies are required to better understand the bacterial interactions taking place in the plant microbiome. Deciphering the molecular basis of interbacterial relationships will be a major future challenge in order to better understand the pathogenicity and epidemiology of microbial plant diseases and target polymicrobial infections.

5.2 QS LuxR solos are extremely predominant among fluorescent pseudomonads. Why are they so widespread and which is their function?

LuxR solo regulators are closely related to QS-LuxRs but are unpaired to a cognate LuxI and initial studies are beginning to show that they are involved in interspecies and inter-kingdom signaling by sensing endogenous and/or exogenous signals. The Chapter 2 provides a picture of LuxR solos distribution, classification, and abundance among the fluorescent pseudomonads group. Results evidenced that LuxR solos occur much more frequently than complete QS-systems within the *Pseudomonas* fluorescens complex and they can be divided into at least nine different functionally sub-groups based on their neighboring genes and their primary structure. The cartography of their ligand binding site suggested that only two sub-

groups are likely to respond to classical AHLs. Differences in the LuxR binding pocket conformation possibly result in the ability to respond/bind to multiple and novel signals. Overall, these results highlight the existence of novel and diverse LuxR solos sub-groups, that very likely evolved away from canonical QS LuxRs and likely bind new signals/molecules and to are involved in novel regulatory mechanisms. Plant microbiomes probably contain a very large number of diverse bacterially produced molecules such as QS signals, volatiles and secondary metabolites which can play cell-cell signaling roles amongst members within the microbiome. LuxR solo regulators will play roles in adaptation, competition and persistence in different environments, due to their high structural plasticity and ability to eavesdrop by responding to different signals produced by neighboring bacteria and the host. To date, there are very few functionally characterized LuxR solos with known ligands and additional studies are therefore required to understand the molecular mechanisms of these LuxR solo subfamilies, their functional role and the signals they respond to in the plant-associated microbiome.

5.3 Genetics of a novel cell-to-cell communication circuit involving a novel LuxR solo and a pigment signal molecule

Bacteria produce and sense a wide variety of small molecules, many of them are used for intercellular and intracellular communication. It is believed that several novel signal molecules and alternative signaling circuits exist, which are currently unknown. In Chapter 4 a unique *luxR* Pseudomonas solo has been genetically and molecularly characterized which did not belong to any sub-group. The LuxR solo gene designated as *fluR* is located adjacent

to a large pigment biosynthetic gene cluster which together likely constitutes a rare bacterial genomic island. Results indicate that FluR detects an endogenous signal and positively regulates the directly adjacent operon which encodes for a black-bluish pigment. The pigment-like molecule in turn acts as a signal, inducing directly or indirectly FluR creating a positive feedback loop. The entire system could constitute a new QS communication circuit formed by a LuxR solo and a signal molecule which is most likely an indole-type pigment. Indole has already been reported as a cell-cell signaling molecule; however its natural role and mechanism of action is still unclear and it is considered an enigmatic signaling molecule. Further research is needed to identify the chemical nature of the pigment and its direct or indirect interaction with the LuxR solo.

5.4 Future directions

Most mechanistic knowledge of bacteria-bacteria interactions so far has been obtained using reductionist approaches such as pure or simple microbial set-ups. Meta-“omics” studies has allowed us to gain extraordinary insights into the taxonomic, functional composition and network of interactions among bacteria associated with plants; they are fundamental for predicting potential chains of direct and indirect positive or negative interactions, but further validations are needed. Very little information is currently available on the role of QS in the context of complex plant microbiomes. Results reported in this thesis are an opening for further research on the role played by the chemical interspecies signaling and in particular by LuxR solos in the process of formation and maintenance of plant microbiomes. LuxR solos are very versatile allowing the co-existence and interaction of highly diverse consortia of bacteria in constantly changing and dynamic environments (**Figure 5.1**).

Possible future directions will be: (i) to map LuxR solos in many more proteobacteria genomes (using genomics and computational biology); (ii) to predict and identify the numerous unknown chemical species that bind to LuxR solos by screening potential sources of signals such as plant exudates, supernatants of microorganisms, soil extracts or endogenously produced via genetically adjacently novel small molecule biosynthetic operons (with machine learning and chemical biology); (iii) to determine the LuxR solo gene targets and define their mechanisms (using biochemistry and analytical chemistry); and (iv) to delineate their roles in bacteria-bacteria interactions and place in the larger microbial community (via genetics and bacterial ecology). In particular, the role of LuxR solos in lifestyles of microbial communities can be studied by microbiome analysis in pertinent ecological environments and by using bacterial synthetic communities (SynComs); this allows to move away from standard models towards more biologically pertinent ones.

6. List of Publications

Bez, C., Esposito, A., Bertani, I., Thuy, H.D., Hong, M.N., Valè, G., Licastro, D., Bertani, I., Piazza, S., and Venturi, V. (2021). The rice foot rot pathogen *Dickeya zeae* alters the in-field plant microbiome. Environmental Microbiology (in press).

Venturi, V., and Bez, C. (2021). A call to arms for cell-cell interactions among bacteria in the plant microbiome. Trends in Plant Science (in press).

Bertani, I., Zampieri, E., Volante, A., Bez, C., Venturi, V., and Monaco, S. (2021). Isolation and characterization of Pseudomonas chlororaphis strain ST9. Plants (in press).

Bez, C., Covaceuzach, S., Bertani, I., Choudhary, K.S., and Venturi, V. (2021) LuxR Solos from Environmental Fluorescent Pseudomonads. mSphere 6: e01322-01320.

Bez, C., Javvadi, S.G., Bertani, I., Devescovi, G., Guarnaccia, C., Studholme, D.J. et al. (2020) AzeR, a transcriptional regulator that responds to azelaic acid in Pseudomonas nitroreducens. Microbiology 166: 73.

Mosquito, S., Meng, X., Devescovi, G., Bertani, I., Geller, A.M., Levy, A., Bez, C., and Venturi, V.. (2020). LuxR solos in the plant endophyte Kosakonia sp. strain KO348. Applied and environmental microbiology 86: e00622-00620.

Musonerimana, S., Bez, C., Licastro, D., Habarugira, G., Bigirimana, J., and Venturi, V. (2020) Pathobiomes revealed that Pseudomonas fuscovaginae and Sarocladium oryzae are independently associated with rice sheath rot. Microbial Ecology 80: 627-642.

Supplementary Materials

Supplementary Table S2.1

List of the culturable bacterial collection of 100 pure/single strains isolated as single colonies by the same macerate of the symptomatic samples used for the culturable pathobiome fraction isolation and analysis. The 16S rRNA gene has been amplified by PCR using the primers fD1Funi 16S (5'-AGAGTTTGATCCTGGCTCAG-3') and rP2Runi 16S (5'- ACGGCTACCTTGTTAGGACTT-3').

Strain lab ID	16S species similarity
A1	*Pantoea sp.*
A2	*Pseudomonas otitis*
A3	*Dickeya zeae*
A4	*Chryseobacterium sp.*
A5	*Chryseobacterium sp.*
A6	*Pseudomonas sp.*
A7	*Pseudomonas otitis*
A8	*Providencia sp.*
A9	*Pseudomonas sp.*
A10	*Stenotrophomonas sp.*
A11	*Serratia marcescens*
A12	*Pseudomonas aeruginosa*
A13	*Providencia sp.*
A14	*Providencia sp.*
A15	*Pseudomonas aeruginosa*
A16	*Pseudomonas aeruginosa*
A17	*Pseudomonas aeruginosa*
A18	*Pseudomonas sp.*
A19	*Panotea sp.*
A20	*Chryseobacterium sp.*
A21	*Pseudomonas otitis*
A22	*Pseudomonas otitis*
A23	*Chryseobacterium sp.*
A24	*Pseudomonas aeruginosa*
A25	*Kosakonia sp.*
A26	*Pseudomonas otitis*
A27	*Burkholderia sp.*
A28	*Pantoea sp.*
A29	*Pseudomonas otitis*
A30	*Stenotrophomonas sp.*

A31	*Pseudomonas sp.*
A32	*Chryseobacterium sp.*
A33	*Pseudomonas aeruginosa*
A34	*Pseudomonas otitis*
A35	*Chryseobacterium sp.*
A36	*Chryseobacterium sp.*
A37	*Pseudomonas otitis*
A38	*Pseudomonas otitis*
A39	*Pseudomonas aeruginosa*
A40	*Pseudomonas aeruginosa*
A41	*Pseudomonas sp.*
A42	*Chryseobacterium sp.*
A43	*Pseudomonas sp.*
A44	*Providencia sp.*
A45	*Pantoea sp.*
A46	*Providencia sp.*
A47	*Aeromonas*
A48	*Aeromonas*
A49	*Chryseobacterium sp.*
A54	*Stenotrophomonas sp.*
A55	*Xanthomonas*
A56	*Burkholderia sp.*
A59	*Pseudomonas otitis*
B26	*Pseudomonas aeruginosa*
B27	*Pseudomonas otitis*
B28	*Stenotrophomonas sp.*
B29	*Stenotrophomonas sp.*
B30	*Pseudomonas otitis*
B31	*Stenotrophomonas sp.*
B32	*Stenotrophomonas sp.*
B33	*Stenotrophomonas sp.*
B34	*Pseudomonas sp.*
B35	*Pseudomonas aeruginosa*
B36	*Dickeya zeae*
B37	*Dickeya zeae*
B38	*Chryseobacterium sp.*
B39	*Burkholderia sp.*
B40	*Morganella sp.*
B41	*Burkholderia sp.*
B42	*Morganella sp.*
B44	*Chryseobacterium sp.*
B45	*Chryseobacterium sp.*
B47	*Providencia sp.*

B48	*Providencia sp.*
B49	*Providencia sp.*
B50	*Providencia sp.*
B51	*Pseudomonas aeruginosa*
B93	*Dickeya zeae*
B94	*Pseudomonas otitis*
B95	*Dickeya zeae*
B105	*Chryseobacterium sp.*
B107	*Burkholderia sp.*
B120	*Pseudomonas otitis*
B121	*Pseudomonas otitis*
B122	*Chryseobacterium sp.*
B123	*Pseudomonas otitis*
B124	*Morganella sp.*
B125	*Stenotrophomonas sp.*
B132	*Panotea sp.*
B133	*Providencia sp.*
B134	*Providencia sp.*
B135	*Burkholderia sp.*
B136	*Dickeya zeae*
B137	*Burkholderia sp.*
B138	*Dickeya zeae*
B139	*Serratia marcescens*
B140	*Pseudomonas otitis*
B141	*Dickeya zeae*
B142	*Chryseobacterium sp.*

Supplementary Table S3.1

>P. putida 16A LUXRSOLO for pEX - 851 bp

GGTACCGTTCCGTGAAATGCCGATATTGGATCCAGCGCCAAGCGACGGGAGCTTCATCGATGCCACATTGGAAAAGCGAG
CAGCTCCAGCAACTGCTGGATGAACAAGAACCGAAGGAGCTGTTCGGTCAAGCGGTGAAACTGGCCCAGGCGCTCAACAT
GGAGTTCGTTGGTTTGGCACTGCACCTGCATGTGGCAGCCCGTGGCCCACAGGTGATCCTCTACAACAACTATCCCGGTG
CCTGGAATGAGCGTTACCAAGCTGAAGACTTAATCAAGATAGACCCAACAGTATCAAAATGCCACCATACAACACTGCCA
CTGGTCTGGAACGACGATCTCTACTGTGAGGTGCCACAACTGCGCGAAGCCGCCACCGCGCTTGGCATGACCCATGGCTG
GAGTCAGTCCCGGGAACGAGAGTCAGTTGAGTGTCTCCAGGCCCCTGGGGGTAGTCACCCGCGAGGAACTCTTCGCGAAG
AGTGCCCAGGTAATGTGGCTGTGCAACACCTTGCACGCCGTGCTCAGCACACATCATCTGCAGAAGTTCAGCCCGGTGCC
GCAACTGAGCGAACGCGAACTCGAAGTGCTCAAGTGGTCCGCAGCGGGCAAGACCGCCGCCGACGTGGCGATGATCCTCT
CGCTGTCGACCAGTACCGTGAATTTCCATATCCGCAGCGTGATCACCAAGACCAACGCCTCCAACAAGGCCGGCGCCATC
GCCATCGCCGCCCTGCGCGGCCTGCTCTGACAGCCCGTCCCCCACCCAGGCGCAAAGCCCTGTAGAATCACCCGCCGCAA
GCCCGCCGCGCAGCGCGCCGGGCAGGATTCGCAGCAGAGCCCCACTCTAGA

>P. putida 16A LUXRSOLO2 FOR pEX- 899 bp

GGTACCACCGCCGACTGCTCACCGGCGACAGCCCGCTGGCCGGCAACGCCTTGGGCAAGCTGGCTGCCGCGGCCTTGCTC

AAGGCGACCAACGGCGGGTGAACGGCATGCCGTTCGACGAGGTGGTACAGGCGGTCCGGCAGATCGAGCAGGCCACCACC

CTGGCGACGATCCAGACGGCCGTGCGCACGTTCGCCCGGCCGCTGGGCTATGACCGCTTCGTGCTGTTCAGCGCCAGCGC

TGCGCGGGACGAGGTGGTCGAGCGCATCCACTGGGTGGAAGGGGACTGGTTCGGCGACGGCCAGGTGGTCGATGCCCGGA

CCTACGTGCGCCATTGTCCGGTGACCCGCCATCTGCTGGCAGCCCGTGAAGCGTTCTTCTGGAGCAAGCAGCCTGGCGAC

GAAGGTGAACGCTATCGGGTGGTTCGGTTGCCCGGCGGGCCGGGCATCCAGGATCCTCGGCCCGGCAGGCCTGGAGGGCG

CGATGAGCCTGGGCGGTGTGCGCATCGACGCTTCGGCGCCGGCGCGCCTGGCCCTGACCCTGCTGGCCAACGCGGCGTTC

CTAGGCGCCCGGCGGCTACTGGAGGCACCGCCCGGCGAGGGTCAGCTGTCGGCGCGTGAGCGCCAGGTGCTGGCCTGGAC

CGCCGCCGGGCAACGCCAGGCCGATATCGCCGCCACCCTGGGCCTGTCGGTGCGCACCGTGGAAAACCACCTGCGCGCCG

CTCGCCGGCGCCTGGGGGTGAGCACCACCGCCCAGGCGATCAGGATCGCCCTGGGCAGCGGCGCGCTCGACTGACTCAGC

CGTTGTGCGCAACCTGCAACACCACCTTGCCGATATGCCGGCCAGACTCCATCAGGGCATGGGCCTGGCTGGCCGCTGCC

AGCGGGAAGGTCGTCTAGA

>P. fluorescens F113 LUXRSOLO FOR pEX – 861 bp

GGTACCGGGAAGTGGTCAGCGAGAGTTTCGTCAAACACTACTTCGCGGTAGCCTGATCGGACTTGGCTGCCGACCCCACC
GGCAGCCCTCTCGGCGAAGGTGCGACCATGAACGTCATCCCCACCCAGGACATCGCCGGCCAGTGCCTGCACGCCTTTAC
CCAGCTGGTTCCCGTCAGCAGGGCGGCGTTCTACTGCGTTGATCGGCAACTGCAAGTCCATGACTTCAGCCTGCATAAGA
TGAGCGGCGAGATGCACCGCGACTACCTGGACAACTATCGCCAATTCGACCCGCTGCACCCGCGCCACTGCGTGTCCAGC
GAGGTGGCAGTGGTGCCGTTGAGCCTGGCGATGGCCCGCCAACCGCGACGTGACAACCGTCGTTATCACGACTTCCTGCA
GCGCTACGGCGTGGTCGACGTAGTGGAAGTCTTCGCCGGATCCAAGCCGCGATTTCGTTGCTGCGCACTGCCGAGCAAGG
TGTCTTCACCAGCCAGCAGTTGAGTCAACTCAGCGCGTTGCAGGCCTTGCTGCAACTGGCCGTGGCCCATTTACCAGCCC
ATGAGGACCCACTGGGCGACTTGACGCCCAAGGAGCGCCAGATTGCCTGGCTGCTGCGCCAAGGCGTCAGCAACAAGCAA
TTGGCCCGGGAGCTGGAGGTGGGCTTGCCGACCATCAAGACTCACCTGATCCACCTGTTTCGCAAGGTTGGCGTGAGCAG
CCGTACCGAGTTGGTCAGCGCACTGTTCCTCTGATCAACCATTTGGTTGATAGGCCGCCGTTGCCCGGCACCCCATGATC
CGAGTGGCTAATTCAACCACCGGAGCCTTGCAATGAGCACATCTTCAGACTGGATTCTAGA

>P. jessenii LUXRSOLO FOR pEX- 923 bp

GGTACCATGGCGTACTTTGTGCATCCAAAGGCCATCCTCTTCTGCCTGTAGCCGCCATGCGCCCCAAGGAACTGAAACTC
TGGACCACCGGTGTCGCCCACCTGCTCGACCTGCCTGCCGGGCATGGGCGGTTGTCGGGGTTGAGCCATTGGCTTCGGCA
AATCTGCCCGGTCGATCATTTCGTGTTGTTTGTCTACGAAGGCAATCACCGGCCACTGGCGTTGTTCGATACGTTCTCGG
CGGACAAGCGGGCGGTGTATGTCGACGACTATCAGTGCGGGCCTTACCTGCTGGATCCCTTCTACCTGGCTTGTACACGC
GGGCAGGCGCCGGGGCTGTGGCGTCTGCGCCAGTTCGCCCCCGACCACTTCTACCTCGGCGAGTACTTCCTGACCTATTA
CCAGCAAACCGGACTGGCCGAGGAAGTGGCGTTCTTCGTCGACCTCGGTGGTGGCGCCATGGCAGTGTCCCGGGGCCAGT
TGTGCATTCAGCCGGGACGAAATGCAACTGGTCGAGTGCGCGCAAGCGGTGGTCGAGCAGGTCATCAATGAGGCCTGGCG
CTTGCGTCAGGCTCAGGAACCGCGACCGGCGCAGGACCTGGACTTCAAGATCCGCGAAGCCTTCGACCACTTCGGTGCGC
ACATCCTCACTGGACGCGAGCAGGAAATCGTTCAACTGTTGCTGCGCGGCCACTCCAGCGCCTCGGTGGCCGAACAACTG

AACATCAGCCCCGGCACGGTAAAGATTCACCGCAAGAACATCTACGCCAAGCTGGGCATTGGCAGCCAGTCGGAGTTGCT
GGGGTTGTTTATTCGGGAGCTGACGGAGGATCCGTTGCAGGCCTCCGGCCTCAAGCTGCAAGCCTGATACCTGTCGGCTA
GCAGCCCCCCCCCTGTGGGAGCGAGCTTGCTCGCGATTCTAGA

>P. oleovorans LUXRSOLO FOR pEX - 814 bp

GGTACCCTTCGCGTTCCCCATCGCGATCCAGGGGAGGAGCTCCAAGATGCATGCCGATCTGCTTACCGTCAGCAGCCGCC
TCGCCGGACTGCCCACGCTCGACGCCTGTCTGGATGCCGTCCAGGCCAGCGCCGACCACCTGGGGTGCAAGGCGCTGATC
TACGACTATGCCCCGGTGCCCCTCAGTCACGAGGGCGTGCTCATCACCCCCTCGGTATTCAAGGTGCGCAACGCCCCGGC
GGACATGCAGGCCTTCTGGTGCGAGCGCGGCTATTACCAGATCGACCCGGTGCAGCAATGCGCCAGTCGGCGCGCGGCGC
CCTTCGTCTGGTCGTTTCGCAGCGAAACCCGCAGCGCCCTGCGCGAGTACCTGGGTGACAGTCACCGCCCGGTCACCTGC
TATGTGCGTGACAGCGGCCTGGCCACGGATCCACCTGCCGGGCGGTGCCTTCGCCACCCTGACCGGCATCCTCGATGCCG
ATGCGCGGGACGCCGAGGCCGCCGGCGAAGCCTGCCTGGGCGCCTTCCTGGTGCTGGCCCATGCCTTCCAGGCCCGGGCC
GACGAACTGCTCGCGCCGGCTGAACGCCGCTGCGGTCAGGTCCGCCTGACGCCGCGCGAGCGCGAATGCCTGCAGTACTC
GGCCAAGGGCCTCACCGCCAAGCGCATCGCCGCGACCCTCAACCGCTCCACCGCCACGGTCAACCTGCACCTCAACTCCG
CCGCCCGCAAACTGGGCGCGCGCAACCGGGTCGAGGCGGTGGTGCGGGGCTTGCACTATCGCCTGGTCGAGGTCTGAGCG
ACTTGGCCTCTAGA

>PppuR16A - 312 bp

CAGGGAAAACAGGTTGGGCGTCAGGGGCTGGACGTCGAGCAGAGTCTGGCGGGTGAATTTAGATCTGTCGGCGCTGGCGG
TCATGACGGGCTCCATGGATCAGATGCGCACAGTTTCGCGCAAAGTGCCCCAGATAAACACCGTCGGTTTGTAAGGTATC
GCTTATGAGTGATCGCAATCAGCTAAAAAAGCGTAACAGCTATAGCAAAAAAAACCAGCGCAGACTTAGGAGATTTCCTA
CACTTCGTTCCGTGAAATGCCGATATTGGATCCAGCGCCAAGCGACGGGAGCTTCATCGATGCCACGAATTC

>P23S - 318 bp

CTGTCGACCAGTACCGTGAATTTCCATATCCGCAGCGTGATCACCAAGACCAACGCCTCCAACAAGGCCGGCGCCATCGC
CATCGCCGCCCTGGATCCGCGCGGCCTGCTCTGACAGCCCGTCCCCCACCCAGGCGCAAAGCCCTGTAGAATCACCCGCC
GCAAGCCCGCCGCGCAGCGCGCCGGGCAGGATTCGCAGCAGAGCCCCACGTCATGCCCCTGTGAATTCTCAACAGCCCCT
TCGCCGACCTCGACCTGATCCGCCAGCCGGCCCAGGCCAATGACCCGCTGCTGGCCTTCGACGCCGCCGACCAGTACC

>Pferrodoxin- 312 bp

CAGGGAAAACAGGTTGGGCGTCAGGGGCTGGACGTCGAGCAGAGTCTGGCGGGTGAATTTGAATTCGTCGGCGCTGGCGG
TCATGACGGGCTCCATGGATCAGATGCGCACAGTTTCGCGCAAAGTGCCCCAGATAAACACCGTCGGTTTGTAAGGTATC
GCTTATGAGTGATCGCAATCAGCTAAAAAAGCGTAACAGCTATAGCAAAAAAAACCAGCGCAGACTTAGGAGATTTCCTA
CACTTCGTTCCGTGAAATGCCGATATTGGATCCAGCGCCAAGCGACGGGAGCTTCATCGATGCCACAGATCT

>PppuR16A_2- 314 bp

GCGCGTTCGGTGCTGCAACGGCATCGGCAACAGTTGCGGTTGGTCGAGGTAGACGACGCAGGGGTATTGCGCGAAGATCT
CGTCGATACGCCCGCCGCGCTCGGCTAGACTGCCCGGAGTTCATTCGCAGCCTTTTCGCGACACAAGGCCGTTCCAGGGA
CCGCGCGGTCTGATGTACCGATGCCACAGGCAGCTGTTGAGGGAAAACCCTCATTGCCGCTGATCCGGTGGCCTCCTAGC
ATGGATCCGGTCAACCCCCAGGCACAGGAGCACAATCATGACCAACAGAATTCGCAGCGACGACAAACGCCCGA

>PpfluR_113 - 314 bp

AGGGCACGCCAGTGCTGTGGGGGATCCAAGTGGTCAGCGAGAGTTTCGTCAAACACTACTTCGCGGTAGCCTGATCGGAC
TTGGCTGCCGACCCCACCGGCAGCCCTCTCGGCGAAGGTGCGACCATGAACGTCATCCCCACCCAGGACATCGCCGGCCA
GTGGAATTCCCTGCACGCCTTTACCCAGCTGGTTCCCGTCAGCAGGGCGGCGTTCTACTGCGTTGATCGGCAACTGCAAG
TCCATGACTTCAGCCTGCATAAGATGAGCGGCGAGATGCACCGCGACTACCTGGACAACTATCGCCAATTCGAC

>PmoaF - 313 bp

GGCCCGGGAGCTGGAGGTGGGCTTGCCGACCATCGGATCCAAGACTCACCTGATCCACCTGTTTCGCAAGGTTGGCGTGA
GCAGCCGTACCGAGTTGGTCAGCGCACTGTTCCTCTGATCAACCATTTGGTTGATAGGCCGCCGTTGCCCGGCACCCCAT
GATCCGAGTGGCTAATTCAACCACCGGAGCCTTGCAATGAGCACATCTTCAGACTGGATCAGAATTCCCGTCGGCGCCCT
GGCCGATGGCTTTGCCCCCGAAGCCTTCATCCTGCCAAACCTGGCCGACCTGGACGGCAAGACCTTCACCCTG

>PpjeR - 166 bp

GGATCCTCAGCCGATCATTTTTGATACCTGCCCTGTAGCACTCGGTAACACCAAGCCCTGTTGCGGTGCGTCTTTGCCACG
ACGCACCGCCAACCGCCGCGCTTCACACGGGCAATACCGCCATGGCGTACTTTGTGCATCCAAAGGCCATCCTCTTCTGCC
TGTAGCCGCCGAATTC

>Psperm - 383 bp

GGATCCTGCAAGCCTGATACCTGTCGGCTAGCAGCCCCCCCCCTGTGGGAGCGAGCTTGCTCGCGATTGCGGCCTTACATTCAACA
TCCATGTGTCTGACACACCGCCATCGCGAGCAAGCTCGCTCCCACAGTTCTATCGGTGATCTGGAGATTGTGGTCAGGGCGGATCAA
GCTGATCCAGCGCCCCTGGGCTCAGTTTGCAGGTTAATACCTGTCGGTTGGCAATAATCCCTTGTGGGAGCGA
GCTCGCTCGCGATTGCGGCCTTACATTCAACATCCATGTGTCTGACACACCGCCATCGCGAGCAAGCTCGCTCCCACAAT
GGTCGGCGGTGTAAACAAAATTTGTATTCACCTCAAATCCCTGTAGGAGCGAGCCCGCTCGCGGAATTC

>PpolR - 451 bp

GGATCCGCCGACATAGGCATAGACCGAACCGGCCACCGGGAAGGCGCGCGACATCTGCTGGTAGCTGACCGCGGTGAACA
GCATGGCGACGAAGCCGATCAGGTAGGTGAGGGGCACCATGCCGTGGCTGGCATCGAACACCCCGCCAAAGATGGAAAAG
GGCGCGATGGGCACCATGAACACCAGGCCGTAGATCAGCAGGTCCTTCAACCCGAGCTGGCGCTTGAGTTCCTGGGTGTA
GCCGAAGCGGGCGAGGTCGGCGGCGTCGCCGGACGGAGCGTGCTGAGTCATGGGGGCTTCCTCGAAGGGGCGGGTTAGGC
TCGACTCTAGGCCAGGGTGGGGCCTGGCCCCATCCGCAAAAGTGCTAGGTTTGCAGCAGTGCTGTCGCGACTGGCAGAAC
CTTTGCTTCGCGTTCCCCATCGCGATCCAGGGGAGGAGCTCCAAGGAATTC

Supplementary Table S3.2

File .xlx attached.

Supplementary Table S3.3

File .xlx attached.

Supplementary Table S4.1

664502..665128
>efflux pump/transporter
MFPLDIWLTYTAACLLLVLSPGPDNLLAIARGLSQGRLAAAISGMASGTGILFHVTTASLGLTLLMQTSV
LAFWIVKVIGASYLLWLGIKVLRSRSLINFQPAARQPLKKIFLTGFLSAALNPKPGLFVLAFIPQFVNPK
LGSVTVQMMVYGAWFAGLTALGFALMGIFATSLSTWLQRKPKVINGLNVGAGLTFVVSGLSIATLPQK

663704..665128
>**FluR**
MRNWYNDLLEWAGQVESENDFLYRACKLAQSLEFEWCSYHVQPPLPISKPVIAFASNYPKAWQRRYRDMD
YVQLDPVVKKARLTQLPFVWESTLLEQEPCFWKEAGDAGLRVGWTCSSISSSGTFSMLTLARNEEPLTIS
ELNDKELKMRWLADATHVALSRLFKPQELEESYWRLTAREIEILRWTADGKTQCEISQILSVSFDTVKFH
SKNAIAKLGTTNKTAAVVRATVLGVLG

662526..663314
>Indole-3-glycerol phosphate synthase
MLEEIVAFKAVETAQRKSHHSLGSLERRIADARAPRAFAQAIATSSVAVIAEAKYRSPSKGVLRADYDPL
ALAHAYQAGGASALSVLADSRFFGNAPYVVGLLANAPGLNLPVMYKDFIVDEFQVYEARALGADAILIIV
RILSPETFRRLYTLALELGLDVLVETFDEADIDQALSVGAGIVGINNRDLDTFKVNFDRTAELFELLPGQ
VIGVAESGISGVADFNRINTIGFRAALMGEYLLGAEDPTRQLRFLTAGGDPN

661471..662529
>Oxidoreductase family, NAD-binding Rossmann fold
MTLMHSIIMGYGHCGKNLHHVCLRKLQVLPTLSELCERVHAVDPQVSSPASAHLASHEQLLPPDSIREGV
GVVHICTSPALHLQHVREALHAGYRYIILEKPMVISQAQATELLALQRTFKAHILVVAVWAHSSLIKLMA
QRLQASGGGLAQLQIVHNKPRFSRTLQRHGEHVFDIEMPHQVSLALLLAGDELKLVDAHSEPLHLDGETR
PSMKSGSLCLEGPQGERVRLTSDLSSPTRERRLSMELEDGAFCQGYLPVSADDSYSQFEAYSPSGVRIAG
QVLADEPLTTCLEAYYRYFLACERGESPLPPLGSSIAFNQRVVELLEQARHLADAAAARAVGSIDRPYEL
RA

660327..661460
>UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronate aminotransferase
MTVPFFAMNQSFSSTLPLIEESLDAFVDQASWVNDKQVRLLEQAIEAYTGAPFAIATGNATDSLIISLLA

LGIGPDDEVIVPCYSFFASLSCVLHVGATPVFVDIEPGSYGLDCAHVEACITPRTKAIMPVHLFRQMVDM

QALKAIAQRHRLLLIEDSAEGIGMRWDGEHAGLIGDIGVLSFFPTKTLGALGDAGMILTRDPLLASRARQ

IMDNGRDAAGLAQRLGYNSRMDDLQALWLRARMLELEPDIARRAVHCALYDKYLEPLTQWVQRPVTLVRN

CPQRTVDYVYLIEVPLRDALAAFLAARHIGTEAYYPLPLHLQPVCEHLGFQAGDLPVAERASTRALGLPL

YPDLTPAAIGRVCEAIGEFYAAQGADL

659215..660330

>UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronate aminotransferase
MIKYYDYRRRHPDARSDILQIMEHQVGRGEFILKDAVSTFERALADKVGARHAICVSSGTSAMTLGLIAA
GIGPGDEVITPAFCYVAAASAIIQAGATPVFADVLPECFTLAPQSVARLITPRTRAVLVAHVFSGLANIA
AIRAVLPARVLILEDSATAFGARYDNKPAGTLGDVGVYSFFPAKPLGGLGDGGAVITEDDEVARRVRMLR
NHGQDGTRRFYHQMLGFNSRMDDLNAAWLSRQLCDNDRQLARKQAIALRYDAAVAPSAGFLTAQRRGTAD
FSPHAYVVRCRSRDRFVRHMNQANVQTKVHFGTLLPEQPAFQQWAGAKAAYAQARALAGECVAIPLCAGM
DEHQVARVVDSVLEFAHESAL

658362..659231
>Inosose dehydratase
MNQPFDLTGVADEGASSIADQIACHRALGWNHLELRSIDAVPVARLADEAWRAVVDQLQQHAMAVPVLAS
QIGNWGSRIAGDFAVDLAELDRLLEIAAPLGTRMIRIMSYPNSGWEEARWRAAVIERLCRLTEVAQRHDV
VLVHENCAGWAGRSAEHTLDMLHAVGSDHLKLLFDVGNGLSYGYQALDFLREVWPHVAHVHLKDGLGSGQ
GVHYQELGEGEAGAVECIDFLLRNGYGGWFSIEPHLSLIPHLQVSDDNPLRKQATYVAYARRVEHILNSR
QGNAHVIHG

657104..658378
>putative succinyl-diaminopimelate desuccinylase
MLFTDRYVQRLVQLLEIDTVTPMESHQASQIAHANAVFADWAASLGMRVVFAGPGEVSVHDDYVTPRTVE
RLCHEHEGFLDWQPHTVLEIGHGPKHRTLMFNFHMDTVSPHLPVLLREGCVHGRGAVDNKGPGLAVLAAI
EEVQHTHPQVLRDIRLLIQVVAGEEGGAMGVYGTRYLCERGLVGALNVFVEPSGDGYFDASTTSMTYEIR
MDGNDSTDDFPERGDNASLILSFMAQEMARSLAEPVKALDVKMTLAGIHTGLHHNRVYGSGHCLFNFSYR
SADAGRQVAEQVDQAYAAALECCRVALSGLHPFSFTIERLLSTCSASWLKRDLPVLNNRDAAMETVLMNA
GINRNTCSAEAFTCDAMWAQADDAYSIVWGPGSLALNGAHTALEHVRLDDLESFTRAVHRLITQFATATA
HTPN

656008..657093
>putative oxidoreductase YjmC
MDFYQVTAPALHAFMVAAFRGARFDVAQARRAADVLHYADLNGHDTHGVANLANIYLAGARSGEIDPRAE
GAWVADQGACATFDAQGGLGLLAGQVAMARALDKARDFGIGCTVVRNSSHFGAAGFYASMGLDQAMIAMA
MTNLGHAPVAHPLGSVAPLLGTNPISFAAPPTKGSVPFVLDMSTTVCASGKIKQAVRLEQNVPQGWLFDA
QGETSGNPRDYLDQRASLPMLGGAFAEQGGHKGLGLGLMVEVLCGALAGAQTAADKGQDGRNNIGHFFLA
INPGFFGSSNAFTGSLDALLGSISGAPVHPAYAPLSYPGQPDSVTRAERLKNGIRLDSHLVSQLDEVADH
CAIPRLARAAS

655094..656011
>Glucose--fructose oxidoreductase precursor
MTIGVGIIGMGVISHYYLKAFERKQACRLVAVCDKAPARLQAYVDSPSVRTYGDYHSLLEDPLVDAVVIN
LPNNLHFQACIDALKAGKHVCCEKPLTLDLEQAEQLRDMARRLNLTLLTAFHRRYNTYLINAVEHDVFAQ
AVHVRARYHERIEDHAGQDTWYLNAAACGGGCIADNGPNVFDTLHVAVGPLRVVNLQVRHGETGIDLGAN
ISLVTPQGLAVSAELSWDYALGEQKDLVVTYADGTTATVDLLQDSAGFKTSLYHEYEGVLAHLASSIQGL
AEDGTMGVEAVRLVRDCYAMTRSPA

654681..655097
>hypothetical protein
MNSELRGTPSKGPIKARLIKLLFHKQLTRGMTLIEFQSRCVRRTEVHELVTTDQLDARPGDRIDRVGFIG
FVEVLEAGVLEAGDAFYIDGRCIGHVLGFDECHFPNHYNILIGTDRSLSGNDIEGCVLGNDVEFKELI

652798..654684
>Tryptophan synthase beta chain
MTEIFEVAVTAAVRDAFPGAGVLAVWHKGGAPASAQVLDWSLSRAIWSEVDKDQLLLHPAVAPFCEYYRQ
VAINPRKSPPSVANFIYRAFCRPDARLPRINAIVDTVNWVAVSTMTSLGAFDARSIVGELCLDVSVEGDW
FEPVGSESREAIPGGRLVLRDREKILSLFSIRDTVHTAIRGASCDLLLLGCLMPGVNPLQVRSALSLLDQ
KLRGDTAPPSAEVPAKGPWYDSFGGSFIAETLSPPVAELNESYERIIASESFQQRYQALLKHYVGRQTPL
TLAENFSRHLGVKAYLKREDLAHTGAHKINNALGQALLAQAMGKRRVVAETGAGQHGVATAAACALLGIE
CVIYMGLRDMQRQALNVQRMRLMGATVVPAEGGSQTLKDAINDALRDWVAHADTTYYLLGSALGPHPYPA

IVRYFQSVIGKEARQQFAALEDGALPDAVVACVGGGSNAIGLFSAFIDEPQVKLCGVEAAGQGEASGLHS
IRFGDSGQSRLGVLQGCQSYVLQDEHGQIMETHSIAPGLDYAMVGPEHAQLRDNGRAQYLQATDEEAIDA
LKLLSRCEGIIPALESAHAVAGAIKLAKRLPAGARIIINISGRGDKDMETISRLVADTVQEGEANESH

652011..652811
>Tryptophan synthase alpha chain
MNPIDRSFAGLALQGRKALMPYLTVGYPARDSLPGLMAAAVEGGADIVEVGIPFSDPVADGPVIQATSQR
AIENGMTLTLALQQIADLPRDDHTPPIVVMTYTNLLMSHGYARFAAQARAAGVSGLIVPDMPVEASDELR
AQLEPAGIHMIFMVTPNLSDARLQRIAAQAKGFLYLVSVLGTTGERSEMADISAFIARVRQHSTLPLAVG
FGVGTPEQARHLWEQVEGVIIGSALARRLYDVDQAPIEAQRYLASFTDRVRNHAGA

651422..652024

>phosphoribosylanthranilate

MPVPRGLIKVCGVRSLEDVQACVEAGADMIGLVMVPGSKRQLLEHQALVL
RERIAGDAQVVGVFMDQPWDEVRRLSEVLALDHVQLHGSESGQQWEQLGR
CTIRRIKPACYHVGCQEPTILPLVDAGAGDGVALDWPPGVSYPDALIAGG
LTVNSVGALIRQLKPLGVDVSSGVEAQPGIKSARLIHEFCASARAAFQQI

650290..651255
>anthranilate phosporibosyltransferase
MQQAFRTMMRGELPDSLVAALLVRLPTKNLTVAELTSATQVVREFLIPVN
AHCAQPPIDLCGTGGDSQGTFNVSTTASFVAAAAGCYVAKHGNRSVSSSC
GSADLLEAVGINLALTPPQIALCLEQVGMGFMFTPLHRPALPGLNKVRRE
LAVRTVFNAMGPLTNPAGAKIQLVGVFSRELVPVMAHTLRALGSERALVV
HGEEGLDEISLSGTTFVAELRGGEIREYTLHPHEVGLDTASLDAIRVRSP
EHAKAMFMAVLNNEAGPPRDIVLLNSAAALYLAGKVETLAEGVPLAAQTI
KTGKALGKYQSLVAYTQSFLR

648969..650261
>UDP-N-acetyl-D-galactosamine dehydrogenase
MNSTFGTVAVVGLGYVGLPLAVEFGKHMHTIGFDISTTTLEHYRDAIDPS
GELTSQQLQAAIHLSYSHVPESLNQADIIIVAVPTPVDIAHQPDLSPLLR
ATECVGRHMRRGALVIFESTVYPGATEEQCVPMLELHSGLTWKRDFNVGY
SPERINPGDRTHTLTRVIKVVSGDCSDSLDCVARLYELIVEPGVHRAPSI
KVAEAAKVIENTQRDLNIALMNELSIIFSGLSIDTTQVLEAANTKWNFLD
FKPGLVGGHCIGVDPYYLTYKAETTGYHPQIILAGRRINDGMGKWIAEKT
IKMLIASGRPIKGATVNMLGVTFKENCSDVRNSKVIDIVHELVSYGVTVH
IHDPLADTAMVQREYGLTLTPWEQLPQADALVVAVPHDEFKTRSMEKMSK
KLAPRGCIVDVKSILDRQQFVDRGYSFWRL

Complete genomic locus Ps_77

TCAGAGTCGCCAGAAACTGTAGCCGCGATCAACAAACTGCTGTCGGTCAAGTATCGACTTGACATCGACT

ATACAGCCACGTGGTGCTAATTTCTTGCTCATTTTTTCCATGGATCGAGTCTTAAACTCGTCATGGGGAA

CCGCTACCACGAGGGCGTCGGCTTGGGGCAACTGCTCCCACGGCGTCAACGTCAGGCCATATTCGCGCTG

GACCATCGCCGTATCCGCCAATGGGTCATGGATATGCACCGTCACACCATAGGAAACCAATTCATGCACG

ATGTCGATCACCTTTGAGTTCCTGACATCACTGCAGTTTTCCTTGAAGGTCACCCCTAGCATATTGACCG

TGGCGCCTTTGATCGGGCGACCGGATGCGATCAGCATCTTGATGGTTTTTTCGGCGATCCATTTACCCAT

CCCGTCATTAATACGCCGACCCGCGAGGATAATTTGCGGATGATAGCCGGTGGTCTCGGCTTTATAGGTC

AGGTAGTAAGGATCTACACCGATGCAGTGGCCACCGACCAGGCCCGGTTTGAAGTCGAGAAAATTCCATT

TTGTATTGGCGGCCTCCAGCACTTGGGTGGTGTCTATCGACAGACCGGAAAAAATGATCGACAGCTCATT

CATAAGGGCAATGTTCAGGTCGCGCTGGGTATTCTCGATGACCTTGGCCGCTTCAGCCACTTTGATCGAG

GGAGCCCGGTGGACACCCGGCTCGACAATCAACTCATACAGGCGCGCGACGCAATCGAGCGAGTCACTGC

AGTCACCGGATACCACCTTGATCACCCGCGTCAGCGTATGAGTTCGGTCTCCCGGATTTATCCGTTCGGG

CGAATAACCCACATTGAAATCGCGCTTCCAGGTCAGTCCGGAATGCAGCTCCAGCATCGGCACGCATTGC

TCTTCCGTCGCCCCTGGGTACACCGTTGACTCAAAGATCACCAAGGCACCGCGCCTCATATGCCGCCCTA

CGCACTCAGTGGCTCTGAGCAAAGGCGAAAGATCCGGCTGATGGGCAATATCGACCGGCGTCGGTACGGC

CACAATAATGATATCCGCCTGGTTCAGGCTCTCGGGAACGTGGCTATAGGACAGGTGAATGGCGGCCTGC

AATTGCTGGGAGGTCAACTCGCCACTGGGGTCGATCGCATCGCGGTAATGCTCAAGCGTGGTGGTGGAAA

TGTCGAAGCCAATGGTGTGCATGTGCTTGCCGAACTCAACCGCCAACGGCAACCCGACATACCCCAGCCC

CACGACAGCAACGGTTCCAAATGTACTGTTCATCGTTCTTTCCTTGAAGTAGTCACTTCGTTCAGCGCAA

AAAGCTTTGTGTGTAAGCCACTAGGGATTGGTATTTGCCCAGGGCTTTACCCGTCTTGATTGTTTGTGCG

GCCAGCGGTACCCCTTCGGCCAGGGTCTCGACCTTGCCCGCCAGGTAAAGCGCAGCCGCGCTATTGAGCA

GGACGATGTCCCGGGGTGGACCGGCTTCGTTATTCAGCACGGCCATAAACATGGCCTTGGCGTGTTCGGG

GCTGCGTACCCGGATAGCGTCTAGTGACGCAGTGTCGAGGCCCACCTCATGGGGATGCAGGGTGTACTCA

CGGATCTCGCCCCCACGCAGCTCTGCGACGAATGTGGTTCCCGAGAGTGAGATTTCGTCCAGTCCCTCTT

CACCGTGAACCACCAGTGCGCGCTCGCTGCCCAACGCTCTCAGGGTGTGAGCCATCACGGGGACCAGTTC

CCGGCTGAATACACCCACCAGTTGGATTTTTGCGCCCGCCGGATTCGTCAATGGCCCCATCGCGTTAAAC

ACCGTGCGAACCGCAAGTTCTCTGCGCACCTTGTTAAGTCCCGGCAAAGCAGGACGGTGCAAAGGGGTAA

ACATAAACCCCATGCCCACTTGCTCAAGACACAACGCTATTTGCGGGGGGGTCAGTGCCAGGTTGATGCC

AACGGCCTCCAATAAGTCGGCGCTGCCGCAACTTGAGCTAACACTGCGGTTGCCGTGTTTGGCCACATAA

CAGCCCGCCGCTGCCGCCACAAAACTGGCGGTGGTAGAGACATTGAAGGTCCCTTGTGAATCGCCGCCGG

TTCCGCATAAATCAATCGGGGGCTGCGCACAATGGGCGTTCACCGGAATCAGAAACTCACGCACGACTTG

AGTCGCAGAGGTCAATTCGGCGACGGTCAGGTTTTTGGTCGGCAACCTCACCAACAGTGCCGCCACTAGT

GAATCAGGCAACTCCCCACGCATCATCGTTCTGAACGCCTGTTGCATTGTTGCCTGGCACACTGGTAACC

CTTCAATAAGCCTGGATACAACTGACTGAACATTATCTGGCACTTTCTAATCCTCACACGCGTTAGCATT

CGCGCAGCCGGGAGCCGCGTCATGTACTTGCAGACACCCGATACACTGTTGTTTTATCTAACTTTCTTGA

ACATTAAATCTGTTGAAAAGCCGCCCTCGCACTGGCGCAAAACTCATGAATCAACCGAGCTGACTTGATG

CCCGGTTGTGCCTCGACGCCGGATGAAACATCAACCCCCAACGGTTTCAACTGTCGTATCAAGGCCCCCA

CGCTATTCACCGTCAGGCCGCCGGCAATCAGCGCATCGGGGTACGACACGCCCGGCGGCCAATCAAGGGC

CACCCCGTCACCGGCTCCTGCATCCACCAGCGGTAGTATCGTTGGCTCCTGGCAACCAACGTGATAGCAC

GCAGGTTTGATCCGCCGGATCGTGCAACGGCCCAATTGCTCCCACTGCTGCCCTGACTCACTGCCGTGGA

GTTGGACGTGATCCAGTGCCAGCACTTCGCTCAAACGCCGAACCTCATCCCAGGGTTGATCCATGAATAC

GCCAACCACCTGAGCGTCACCGGCAATACGCTCACGCAAAACCAATGCCTGGTGTTCCAGCAGTTGACGC

TTTGAACCTGGCACCATCACCAGACCAATCATGTCCGCCCCCGCCTCCACGCATGCCTGTACATCCTCCA

GGGAACGCACACCGCACACTTTGATCAGGCCCCTAGGCACCGGCATGATTACGCACTCGGTCCGTGAAAC

TCGCCAGGTAACGTTGAGCTTCGATGGGAGCCTGGTCGACGTCATACAGGCGACGGGCCAGGGCACTGCC

GATAATCACGCCCTCCACCTGCTCCCACAGGTGCCGCGCCTGTTCCGGCGTGCCCACCCCGAAACCCACA

GCCAGGGGGAGCGTGGAGTGCTGTCTCACCCTGGCGATAAACGCGGAAATATCCGCCATTTCGCTGCGTT

CCCCGGTCGTGCCAAGTACCGAGACCAGGTATAAAAAGCCCTTGGCTTGCGCAGCAATACGCTGCAGCCG

GGCATCGCTCAGGTTCGGGGTGACCATGAAAATCATATGGATGCCCGCCGGTTCCAGTTGCGCCCGCAAT

TCGTCGCTGGCCTCCACCGGCATGTCCGGGACAATCAACCCGCTGACACCCGCCGCCCTAGCCTGCGCCG

CGAAGCGCGCATAACCATGGCTCATCAACAAATTGGTGTACGTCATTACCACGATCGGTGGCGTGTGGTC

ATCACGGGGCAAGTCGGCAATCTGCTGGAGAGCCAGAGTCAAGGTCATGCCATTTTCGATGGCGCGCTGG

CTGGTGGCCTGGATCACAGGGCCATCGGCCACCGGGTCGGAAAACGGAATCCCCACCTCGACAATATCCG

CCCCACCTTCCACTGCGGCCGCCATCAATCCAGGCAGAGAGTCACGCGCCGGGTAACCCACGGTGAGGTA

TGGCATCAGGGCTTTGCGCCCCTGCAGGGCCAGGCCTGCAAAACTTCGGTCAATGGGATTCATTGGCTTC

ACCCTCCTGGACAGTGTCGGCTACCAGCCGACTGATGGTTTCCATGTCTTTGTCACCACGACCACTGATG

TTGATGATGATGCGGGCGCCAGCGGGAAGCCTCTTGGCCAGCTTGATTGCCCCGGCGACCGCGTGGGCAC

TTTCCAGGGCCGGAATAATGCCTTCGCAGCGTGACAACAGTTTCAGCGCATCAATCGCCTCTTCATCCGT

CGCCTGCAGGTATTGCGCACGCCCGTTGTCCCGCAGTTGTGCATGCTCCGGACCGACCATGGCATAGTCC

AGGCCCGGCGCGATGCTATGGGTTTCCATGATTTGCCCGTGTTCATCCTGCAACACGTAGGATTGGCAGC

CCTGCAGGACTCCAAGGCGACTTTGTCCCGAATCGCCAAAACGAATCGAGTGCAAGCCTGAAGCCTCGCC

TTGACCGGCGGCTTCAACACCGCACAGCTTGACCTGAGGCTCATCGATAAAGGCCGAAAACAAGCCGATC

GCATTCGAGCCGCCACCGACACACGCAACGACGGCATCGGGCAACGCACCGTCCTCCAGGGCTGCAAATT

GCTGCCGCGCTTCCTTGCCGATCACACTCTGGAAGTAACGCACGATCGCCGGATAAGGGTGCGGCCCCAA

GGCTGACCCCAGCAGGTAGTAAGTGGTATCGGCATGTGCCACCCAATCCCGCAACGCGTCATTGATCGCA

TCCTTGAGAGTCTGGCTGCCCCCTTCGGCAGGCACCACCGTCGCGCCCATCAGGCGCATACGCTGAACGT

TCAAGGCCTGGCGCTGCATATCCCGCAGGCCCATATAGATCACGCATTCGATGCCCAGCAACGCGCACGC

GGCCGCCGTTGCGACCCCATGCTGGCCCGCGCCGGTTTCGGCCACCACCCGCCGTTTACCCATGGCCTGG

GCGAGCAATGCCTGGCCCAGGGCATTGTTGATTTTGTGGGCGCCGGTGTGGGCCAGGTCCTCGCGCTTAA

GGTAAGCCTTGACGCCAAGGTGCCGGGAGAAATTCTCCGCCAGGGTCAATGGGGTTTGTCGGCCGACGTA

ATGCTTGAGCAAAGCCTGGTAGCGTTGCTGGAACGACTCACTGGCGATGATTCGCTCGTAGCTTTCGTTC

AGTTCAGCCACGGGCGGCGAAAGGGTTTCGGCGATAAAACTGCCGCCAAAGGAGTCATACCACGGCCCCT

TGGCTGGCACCTCGGCGCTTGGCGGAGCGGTGTCGCCGCGTAGTTTCTGGTCCAAGAGGCTCAATGCCGA

ACGAACCTGCAGCGGGTTGACGCCTGGCATCAGGCAGCCCAATAACAGCAGGTCACAGCTGGCCCCACGG

ATCGCGGTATGCACCGTATCGCGAATGGAAAACAGCGACAGAATCTTCTCCCTGTCGCGCAGCACCAGGC

GCCCACCGGGAATCGCCTCCCTGGATTCGCTGCCAACCGGCTCGAACCAGTCGCCTTCCACACTCACGTC

AAGGCACAGCTCGCCGACAATCGACCTTGCGTCGAAGGCCCCCAACGAGGTCATCGTCGACACGGCTACC

CAATTGACGGTGTCAACAATCGCGTTGATCCGTGGCAGCCGTGCATCAGGCCTGCAGAAGGCCCGGTAGA

TAAAATTGGCGACCGACGGTGGGCTCTTCCGAGGGTTGATCGCAACTTGCCGGTAGTACTCACAGAACGG

CGCGACGGCGGGATGCAGCAATAGCTGATCCTTGTCCACCTCCGACCAGATAGCCCGACTCAACGACCAG

TCCAGTACCTGGGCAGACGCCGGCGCCCCCCCTTTGTGCCAGACCGCCAGGACGCCCGCCCCCGGAAACG

CATCACGTACGGCGGCGGTCACCGCCACCTCGAATATCTCTGTCATATCAGCTCCTTGAACTCGACATCG

TTGCCCAGGACACACCCTTCAATATCGTTACCGGACAGCGACCTGTCCGTACCGATGAGGATGTTGTAGT

GATTTGGAAAATGGCACTCGTCGAACCCCAATACGTGGCCGATACAGCGTCCATCGATGTAAAAGGCATC

GCCTGCCTCTAACACACCCGCTTCGAGCACTTCGACAAAACCAATGAAGCCGACCCGATCGATGCGATCA

CCCGGTCGAGCATCGAGTTGATCGGTGGTAACCAGTTCATGCACCTCGGTGCGCCGCACACAGCGGGATT

GAAACTCGATAAGGGTCATGCCGCGCGTCAGTTGCTTGTGGAACAGCAGCTTGATCAGGCGCGCCTTGAT

CGGGCCTTTGGACGGGGTGCCGCGGAGCTCGGAGTTCATGCCGGCGACCGCGTCATGGCATAGCAATCAC

GCACCAGCCGCACTGCCTCCACGCCCATCGTGCCGTCTTCGGCTAAGCCCTGGATCGAGCTCGCCAGGTG

GGCCAGGACACCTTCATATTCGTGGTACAGCGACGTCTTGAAACCGGCCGAGTCCTGCAACAGATCGACC

GTGGCGGTGGTTCCGTCGGCATAGGTGACGACCAAGTCTTTCTGCTCGCCAAGGGCGTAGTCCCAGCTCA

GTTCAGCGCTGACAGCAAGCCCCTGGGGCGTGACCAGGGAAATGTTTGCCCCCAGGTCGATGCCTGTCTC

TCCATGGCGAACCTGCAGGTTCACCACACGCAAGGGGCCCACGGCCACGTGCAGGGTATCGAACACATTC

GGCCCGTTGTCTGCGATGCACCCGCCACCACACGCTGCCGCATTGAGGTACCAGGTGTCTTGCCCAGCAT

GGTCTTCGATGCGTTCGTGATAACGGGCGCGCACATGCACCGCCTGGGCAAACACATCGTGCTCCACCGC

ATTGATAAGGTAGGTGTTATAGCGTCGGTGAAAGGCGGTCAGCAGTGTCAGGTTCAACCGGCGGGCCATG

TCGCGCAATTGCTCGGCCTGCTCCAGGTCAAGGGTCAATGGCTTTTCGCAGCAAACATGTTTACCCGCCT

TGAGGGCGTCGATGCAGGCCTGGAAATGCAGGTTGTTCGGCAGGTTGATGACGACCGCGTCAACCAGTGG

GTCTTCCAGCAACGAATGATAATCGCCGTAGGTGCGAACACTGGGAGAATCGACGTAAGCCTGCAACCGC

GCGGGAGCCTTGTCGCAAACGGCGACGAGCCGGCAGGCCTGCTTGCGCTCGAATGCCTTGAGGTAATAGT

GGGAAATGACGCCCATGCCGATAATGCCCACACCGATTGTCATGATGCAGCCCTCGCCAGCCGTGGAATT

GCGCAGTGGTCCGCCACTTCGTCCAGCTGCGAAACCAAGTGGCTGTCCAACCGAATGCCATTTTTCAGGC

GCTCGGCCCTTGTGACTGAGTCGGGTTGACCGGGGTAGCTCAGGGGAGCGTAGGCGGGGTGAACCGGGGC

CCCGCTGATAGACCCCAGCAAGGCATCGAGACTGCCGGTGAAAGCGTTGCTGCTGCCAAAGAAGCCAGGG

TTTATCGCCAGGAAAAAGTGCCCGATATTATTACGCCCGTCCTGCCCTTTATCGGCAGCCGTTTGCGCAC

CCGCCAGTGCCCCACAGAGGACTTCGACCATCAGCCCAAGCCCCAGCCCCTTGTGTCCGCCCTGCTCGGC

AAATGCACCGCCCAACATGGGCAAGCTGGCTCTTTGGTCAAGGTAGTCCCGGGGGTTACCGGACGTTTCA

CCCTGGGCGTCGAACAACCAGCCCTGCGGGACGTTCTGCTCCAGCCGCACGGCCTGCTTGATCTTGCCAC

TGGCACACACAGTGGTGCTCATATCCAGCACGAAGGGCACGCTCCCCTTGGTGGGTGGTGCTGCGAAGCT

GATGGGGTTGGTGCCCAGCAAGGGCGCGACACTGCCCAGCGGGTGAGCCACCGGTGCGTGCCCCAGGTTG

GTCATGGCCATTGCGATCATTGCCTGGTCCAGGCCCATGCTTGCGTAAAAACCTGCCGCACCAAAATGCG

AGCTGTTGCGCACCACAGTACAACCGATGCCAAAGTCCCGCGCCTTGTCCAACGCACGAGCCATGGCGAC

TTGCCCGGCGAGCAGGCCCAGCCCGCCTTGTGCGTCGAATGTGGCACAGGCGCCCTGGTCCGCGACCCAG

GCACCCTCGGCACGCGGGTCGATTTCACCGCTTCGCGCCCCTGCCAGGTAAATATTGGCGAGGTTGGCCA

CACCGTGGGTGTCATGGCCATTGAGGTCGGCGTAATGCAGTACATCGGCTGCCCGCCGCGCTTGCGCAAC

ATCGAACCTGGCCCCTCGGAAGGCTGCCACCATAAAAGCGTGCAGCGCGGGCGCTGTCACCTGATAGAAA

TCCATAGCGATATCCTTAATTCGGGGTGTGGGCAGTCGCCGTAGCGAACTGTGTGATCAAACGATGAACC

GCGCGGGTGAATGATTCGAGGTCATCCAGGCGTACGTGTTCGAGCGCGGTATGCGCACCGTTGAGTGCCA

GTGATCCCGGCCCCCAGACGATGCTGTAGGCGTCATCTGCCTGGGCCCACATCGCATCGCAGGTGAAGGC

TTCGGCACTGCAAGTGTTGCGGTTGATGCCGGCGTTCATTAGGACGGTCTCCATGGCGGCATCGCGGTTG

TTCAACACCGGCAAGTCACGCTTGAGCCAACTGGCCGAACACGTGGATAACAGGCGTTCGATGGTGAAGC

TGAAGGGATGAAGCCCGGACAACGCCACCCGGCAACACTCCAGCGCAGCAGCGTAGGCCTGGTCGACCTG

CTCGGCCACCTGGCGCCCTGCATCGGCACTGCGATAGCTGAAATTGAACAGGCAATGCCCGCTGCCGTAG

ACGCGGTTATGGTGCAAACCGGTATGGATTCCCGCCAGTGTCATCTTCACGTCCAGGGCCTTGACCGGTT

CAGCCAGCGAGCGCGCCATCTCTTGCGCCATGAACGAGAGGATCAGCGACGCGTTGTCACCCCGCTCGGG

AAAGTCGTCCGTCGAGTCGTTCCCATCCATGCGGATTTCATAGGTCATCGAGGTCGTGGAGGCATCGAAG

TAACCGTCGCCCGACGGCTCGACAAACACGTTCAGCGCGCCAACAAGGCCCCGCTCACACAGATAGCGCG

TGCCATATACCCCCATGGCACCACCCTCCTCGCCCGCCACGACCTGGATCAACAGGCGAATGTCCCGCAG

GACCTGCGGATGGGTATGCTGCACCTCTTCAATCGCGGCCAGCACGGCCAGCCCTGGCCCCTTGTTATCC

ACCGCACCACGGCCATGTACACAACCTTCGCGTAGCAGTACCGGCAGATGGGGGGAAACGGTGTCCATGT

GGAAATTGAACATCAGGGTTCGATGCTTCGGCCCGTGCCCAATCTCCAGCACCGTGTGAGGCTGCCAGTC

GAGGAAACCCTCATGCTCGTGGCACAGGCGCTCGACCGTCCGGGGCGTGACGTAGTCGTCGTGGACACTC

ACTTCTCCCGGCCCGGCAAACACCACCCGCATGCCCAGGCTCGCGGCCCAATCGGCAAAGACGGCATTAG

CGTGAGCAATCTGTGAGGCTTGATGGCTTTCCATCGGCGTTACCGTATCGATTTCCAGCAGCTGTACCAG

GCGCTGCACATACCTATCCGTGAATAACATGGGCGTTCCCTTGCCGTGAGTTCAAAATGTGCTCGACGCG

GCGAGCGTAAGCCACATAGGTAGCCTGCTTGCGCAGCGGGTTATCGTCGGACACCTGCAAATGCGGGATC

AGGCTCAGGTGCGGCTCGATGGAAAACCAACCGCCATACCCGTTTCGCAGCAGGAAATCGATGCATTCGA

CAGCTCCCGCTTCACCTTCGCCCAGCTCCTGGTAGTGCACACCTTGACCACTGCCCAGGCCATCCTTGAG

GTGGACATGCGCGACGTGAGGCCAGACTTCACGCAGGAAGTCGAGCGCCTGATAGCCGTAGCTCAGACCG

TTACCCACATCAAACAGCAGCTTCAAGTGATCCGAGCCAACGGCATGGAGCATGTCCAAGGTGTGTTCGG

CGCTGCGCCCGGCCCAGCCCGCGCAGTTTTCATGAACCAGCACCACATCGTGCCGTTGCGCGACCTCTGT

CAGCCGGCATAACCGCTCGATCACCGCCGCGCGCCAACGAGCTTCCTCCCAGCCACTGTTGGGGTAAGAC

ATGATACGAATCATCCGGGTACCCAGCGGCGCAGCGATCTCCAGCAACCGATCCAGCTCGGCGAGGTCGA

CGGCAAAGTCCCCGGCAATCCGGCTGCCCCAGTTACCGATTTGCGAGGCCAGCACGGGCACCGCCATGGC

GTGTTGCTGCAGTTGGTCGACCACCGCCCTCCAGGCTTCATCCGCCAAGCGGGCAACCGGCACTGCGTCA

ATGCTGCGCAGTTCCAGGTGGTTCCAGCCCAGTGCACGGTGACATGCAATCTGATCGGCGATCGAACTTG

CGCCTTCGTCGGCCACGCCGGTAAGGTCAAAGGGCTGATTCATGGGCGAACTCCAATACGCTGTCGACAA

CCCGTGCTACCTGGTGTTCGTCCATACCCGCGCACAACGGAATCGCCACACACTCGCCCGCCAGGGCTCG

TGCCTGAGCGTACGCTGCCTTCGCGCCAGCCCACTGCTGGAAGGCAGGTTGCTCCGGCAGCAGGGTGCCA

AAGTGGACCTTTGTCTGCACATTCGCCTGGTTCATGTGCCGCACAAACCGGTCGCGGGAACGACAGCGCA

CCACATAGGCATGGGGCGAAAAATCCGCTGTGCCACGACGCTGCGCCGTCAGGAAACCCGCCGACGGTGC

AACAGCCGCATCGTAGCGAAGGGCAATGGCTTGCTTGCGCGCCAATTGCCGGTCGTTGTCGCACAACTGC

CGGCTCAACCAGGCGGCGTTAAGGTCATCCATCCGGCTGTTGAACCCGAGCATCTGGTGATAGAAACGAC

GCGTACCGTCTTGGCCATGGTTGCGCAACATGCGCACCCGGCGAGCCACTTCATCATCCTCGGTAATCAC

TGCACCACCATCGCCCAAACCTCCCAAGGGCTTGGCTGGAAAAAACGAATAAACCCCTACGTCTCCCAGG

GTGCCGGCCGGTTTGTTGTCATACCGCGCGCCGAACGCGGTGGCGCTATCTTCGAGAATCAGCACGCGGG

CCGGCAACACCGCACGTATTGCGGCGATGTTGGCCAAGCCGGAAAAGACATGGGCCACCAGCACGGCCCG

GGTACGCGGTGTAATCAGCCGCGCGACAGATTGCGGGGCCAGCGTGAAGCATTCGGGCAATACATCGGCA

AACACCGGTGTGGCGCCCGCCTGGATGATTGCCGAGGCTGCAGCCACGTAACAGAACGCTGGCGTAATCA

CCTCGTCGCCAGGTCCGATGCCGGCCGCAATCAGCCCCAGCGTCATCGCCGAGGTACCCGAGCTGACACA

GATGGCATGCCGGGCTCCGACCTTGTCGGCCAGGGCCCGTTCGAACGTGCTGACCGCGTCCTTGAGGATG

AATTCTCCCCGCCCAACCTGATGCTCCATAATCTGCAAAATATCAGACCGGGCATCAGGATGACGGCGCC

TGTAGTCGTAGTACTTGATCATAGGTCTGCTCCTTGCGCTGCGTAAAACTCACCGATGGCCTCGCATACC

CGGCCGATGGCGGCCGGTGTCAGGTCCGGATAAAGCGGCAGGCCAAGGGCTCGGGTCGAGGCACGCTCTG

CCACCGGCAAATCACCCGCCTGAAAGCCAAGGTGCTCACACACCGGTTGCAGGTGCAGTGGCAGCGGATA

GTAGGCTTCCGTGCCGATGTGGCGTGCCGCAAGGAAGGCGGCCAGCGCATCGCGCAGCGGCACTTCGATC

AGGTATACGTAGTCGACTGTGCGCTGAGGACAGTTGCGGACCAATGTCACCGGCCGTTGCACCCACTGTG

TGAGGGGCTCAAGGTACTTATCGTATAGCGCACAGTGCACTGCCCGCCGGGCAATGTCGGGCTCCAGCTC

CAGCATTCGTGCGCGTAGCCACAAGGCCTGCAAATCATCCATGCGGCTGTTGTACCCCAAGCGCTGGGCC

AGTCCCGCCGCATCGCGCCCGTTGTCCATGATCTGCCTTGCCCGGCTGGCCAGCAACGGGTCACGGGTCA

GGATCATGCCGGCGTCACCCAACGCTCCCAGGGTCTTGGTCGGGAAAAACGAAAGCACGCCAATGTCCCC

GATCAACCCGGCATGCTCGCCGTCCCAGCGCATGCCAATGCCTTCTGCGCTGTCTTCGATCAGCAGCAAC

CGATGCCGTTGCGCGATGGCTTTCAACGCCTGCATATCAACCATTTGCCGGAAAAGATGGACGGGCATGA

TCGCCTTGGTTCGAGGCGTAATGCAGGCCTCGACGTGCGCACAATCCAGGCCGTAGCTGCCGGGCTCGAT

ATCCACGAACACCGGCGTCGCCCCCACATGCAATACACAGGACAACGACGCGAAGAAGCTGTAGCAGGGC

ACGATCACTTCGTCGTCCGGGCCAATCCCCAGGGCCAGCAGACTGATGATCAGCGAGTCTGTCGCGTTAC

CGGTGGCGATCGCAAATGGCGCGCCGGTGTAGGCCTCGATCGCCTGTTCCAACAAGCGCACCTGTTTGTC

ATTCACCCAACTGGCCTGATCGACGAAAGCGTCCAGGCTCTCTTCAATTAGCGGCAAAGTGCTGGAAAAC

GATTGGTTCATGGCAAAAAATGGCACTGTCATGGGAGGCCTCCTAGGCTCTGAGTTCGTACGGTCGGTCA

ATTGAACCCACAGCCCGGGCAGCGGCGGCGTCGGCCAAATGGCGCGCCTGCTCCAGAAGCTCGACGACTC

GCTGATTGAAGGCAATCGAGCTGCCAAGAGGCGGGAGTGGCGACTCGCCCCGCTCGCACGCCAGAAAATA

CCGGTAGTACGCCTCCAGACAGGTCGTAAGGGGCTCGTCCGCCAAAACCTGCCCGGCAATGCGCACGCCC

GAGGGGGAGTAGGCCTCGAACTGCGAATAACTGTCGTCGGCGCTGACGGGCAAATAGCCCTGGCAAAACG

CACCATCCTCCAATTCCATGGACAAGCGCCGCTCACGTGTGGGCGAGCTCAAGTCCGACGTCAATCGGAC

TCGTTCACCTTGTGGCCCCTCCAGGCACAGAGACCCGGACTTCATCGACGGCCGGGTTTCACCGTCCAGG

TGCAATGGCTCGCTGTGCGCGTCGACGAGCTTGAGTTCGTCTCCCGCCAATAAGAGCGCCAGACTGACCT

GATGAGGCATTTCAATATCAAATACGTGCTCTCCATGGCGCTGCAACGTGCGCGAAAAACGCGGCTTGTT

ATGCACGATCTGCAATTGCGCGAGGCCGCCACCGGACGCCTGGAGCCGCTGCGCCATGAGCTTGATCAGT

GAGCTGTGCGCCCATACCGCAACTACCAGGATGTGTGCCTTGAAAGTACGCTGCAGGGCCAACAGTTCAG

TGGCTTGTGCCTGGCTGATCACCATCGGCTTTTCGAGAATGATGTAGCGATACCCGGCGTGGAGTGCCTC

CCGCACATGCTGCAAATGCAACGCGGGCGACGTGCATATATGCACCACCCCGACACCTTCGCGGATGCTG

TCGGGTGGCAGCAGTTGCTCATGACTGGCCAGGTGCGCAGACGCTGGCGATGAAACCTGCGGGTCCACGG

CATGCACCCGCTCGCACAACTCACTCAGAGTCGGCAACACCTGTAACTTGCGCAGGCATACGTGGTGCAA

GTTCTTTCCGCAGTGCCCATAGCCCATTATTATCGAGTGCATCAGGGTCAATTGGGATCGCCTCCCGCGG

TCAAAAAACGCAACTGCCGAGTTGGATCTTCAGCACCCAATAAATATTCACCCATCAGCGCCGCACGAAA

CCCGATAGTATTAATCCGATTAAAATCGGCAACTCCGGAGATACCACTTTCGGCTACCCCAATGACCTGA

CCAGGTAACAACTCAAACAATTCAGCGGTGCGATCAAAGTTGACCTTGAATGTATCCAGGTCACGGTTAT

TGATACCGACGATCCCTGCGCCGACACTCAAGGCCTGGTCTATATCAGCCTCATCAAATGTTTCCACCAA

TACGTCGAGCCCCAATTCCAGGGCCAGCGTATAGAGTCGCCGGAACGTTTCGGGCGAAAGGATCCGCACG

ATGATCAGAATCGCGTCAGCCCCCAGCGCGCGGGCCTCGTAGACCTGGAACTCATCGACAATAAAGTCTT

TGTACATCACTGGCAGGTTCAAGCCCGGCGCATTCGCCAACAGCCCCACCACGTAGGGTGCGTTGCCGAA

AAACCGACTATCGGCCAGCACCGAAAGCGCACTGGCACCGCCCGCCTGATAAGCATGGGCCAAGGCCAGA

GGATCATAGTCGGCACGTAATACGCCCTTTGAAGGAGAGCGATACTTGGCTTCTGCAATAACCGCAACGC

TTGAGGTCGCAATCGCTTGGGCAAACGCCCGCGGCGCCCTGGCATCAGCTATACGACGCTCTAGGGAACC

CAGGGAGTGATGACTTTTCCTTTGGGCCGTTTCAACAGCCTTAAACGCAACGATCTCTTCAAGCATATAA

TTATCTCGCATGACAGTCATGGAACAAGCCACCCAACAAAAATTATTGGGTAAAGCCCCAAGAAGAGAAA

TCAAACCCGCCTTGTTATTCTCGACAACTAACTAGCCACTCTTTTCGCAGCAAATTAATTAAAAGGAAAA

GCCAGCACCCCGTCTACCTACCTTGTTGGGTAGGCGCTTGAGCCTTTAAAAAACTGCAACCAAGATACAA

TTGACACCATGTCATTTAATATCCCACACGAATAACGAGCCATTTAATGCAGCACGAAAGCTCCAACTAA

CTGTCAGTTTTTAAACGTCAAAATATTGGGAGCTTTTTTACTTCCCATTTGAAATTCATCTAGGTATAAA

ACTCTCAGGGGAAACCAGTAGATAAAGGAATAACGATGCGCAACTGGTACAATGACCTGCTTGAATGGGC

TGGGCAGGTCGAGTCTGAAAATGATTTTTTGTATAGGGCATGCAAGCTTGCACAGTCATTGGAGTTTGAA

TGGTGCAGCTATCACGTCCAACCGCCCCTCCCCATTTCCAAGCCTGTCATCGCCTTTGCAAGCAATTATC

CAAAGGCATGGCAGCGACGCTATCGCGACATGGACTATGTGCAATTGGACCCGGTCGTCAAAAAGGCGCG

GCTGACCCAACTGCCCTTTGTCTGGGAAAGTACACTGCTCGAACAGGAACCCTGCTTCTGGAAGGAAGCG

GGTGACGCGGGCCTGAGGGTGGGCTGGACATGTTCCAGTATCAGCAGCAGCGGCACCTTCAGCATGTTGA

CGCTCGCCCGCAACGAAGAGCCGCTCACCATCAGCGAATTGAATGACAAAGAGCTGAAAATGCGCTGGCT

GGCCGATGCCACTCACGTTGCACTGAGCCGCCTGTTCAAACCGCAGGAACTCGAGGAATCCTACTGGCGG

TTGACCGCCCGGGAAATCGAGATCCTGCGCTGGACAGCCGACGGAAAAACCCAGTGCGAGATCTCACAAA

TTCTCTCGGTATCTTTCGACACCGTGAAGTTTCACAGTAAAAATGCTATCGCCAAGCTGGGCACTACCAA

TAAAACCGCAGCCGTGGTGAGAGCGACCGTTCTGGGCGTGCTGGGCTAATGAGCCTGCTATCCTCTCATC

AGGCCGCATGGCGCTCATCCCGAAAGCATCGGGCGGGTACCCAAACACAAAAGGAGAATAATTGTGTTCC

CATTGGATATTTGGCTGACCTATACCGCCGCCTGTCTGTTACTGGTGCTTTCCCCCGGGCCGGATAACCT

CCTGGCAATCGCCAGGGGGCTCAGCCAGGGCAGACTGGCGGCTGCCATCTCAGGCATGGCTTCGGGCACC

GGCATCCTGTTTCACGTGACAACCGCCTCCCTGGGGTTGACGCTCTTGATGCAGACCTCAGTGCTGGCCT

TCTGGATCGTCAAGGTGATCGGGGCCAGCTACTTGCTTTGGCTCGGTATCAAAGTCTTGCGCTCGCGTAG

CCTGATCAACTTTCAGCCTGCCGCGCGCCAGCCCTTGAAAAAAATATTCCTAACCGGTTTTCTGTCTGCG

GCACTGAATCCCAAGCCGGGCCTCTTTGTGCTTGCCTTCATTCCTCAGTTCGTCAACCCCAAGCTGGGCT

CGGTCACCGTGCAAATGATGGTGTATGGCGCCTGGTTCGCGGGTTTGACCGCACTGGGCTTTGCCCTCAT

GGGGATTTTTGCGACGAGCCTTTCGACGTGGCTGCAACGCAAGCCCAAAGTGATCAACGGGTTAAATGTG

GGGGCAGGACTGACCTTCGTTGTTTCCGGTCTTTCGATAGCAACCCTGCCCCAAAAATAA