Identification of regulatory proteins involved in the expression of levansucrase in *Pseudomonas syringae*

by

Khaled Abdallah

A Thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
in Biochemistry

Approved Dissertation Committee

Prof. Dr. Matthias Ullrich
Jacobs University Bremen

Prof. Dr. Alexander Lerchl
Jacobs University Bremen

Dr. Astrid Gärdes
Leibniz-Zentrum für Marine Tropenökologie

Date of Defense: August 19, 2016

Department of Life Sciences and Chemistry

Contents

Acknowledgments ... 3

Summary ... 5

List of Abbreviations ... 6

1. Introduction ... 9

1.1 The Soybean ... 9

1.2 Plant pathogens ... 11

1.2.1 *Pseudomonas syringae* ... 12

1.2.2 *Pseudomonas syringae* pv. glycinea PG4180 ... 13

1.3 Levan and levansucrase ... 14

1.4 Bacteriophages of *Pseudomonas* ... 16

1.5 Glucose metabolism in *Pseudomonas* ... 18

2 Aims of the study ... 21

3 Results ... 23

3.1 The bacteriophage-derived transcriptional regulator, LscR, activates the expression of levansucrase genes in *Pseudomonas syringae* ... 25

3.2 Analysis of the prophage-derived genomic region flanked by *trpE* and *trpG* in *P. syringae* pv. glycinea PG4180 ... 59

3.3 Expression of extra-cellular levansucrase in *Pseudomonas syringae* is controlled by the in planta fitness-promoting metabolic repressor HexR ... 81

4 Discussion ... 115

5 References ... 123

Statutory Declaration ... 132

Acknowledgments

I would like to offer my sincere gratitude to my main supervisor, Prof. Dr. Matthias Ullrich, for the regular advice and well-directed criticism which improved my scientific thinking. His generous, insightful, and enthusiastic support throughout this research project enabled me to endeavor my Ph.D. studies and make my stay at Jacobs University one of the most rewarding in my life.

Genuine gratitude is also expressed to my thesis committee members Prof. Dr. Alexander Lerchl and Dr. Astrid Gärdes for their time, valuable comments and suggestions.

I would also like to express my warm and sincere thanks to current and former members of Prof. Ullrich research group for their friendly and family working atmosphere. Particularly I would like to mention Dr. Antje Stahl, Dr. Ahmed Rezk, Dr. Daniel Pletzer, Dr. Gabriela Alfaro-Espinoza, Dr. Shaunak Khandekar, Neha Kumari, and Benjamin Gillard. Furthermore, I would also like to acknowledge the much appreciated help and work of our lab technicians Maike Last and Nina Böttcher. I also thank my lab rotation student Katharina Hartman for her contribution in one of my projects.

I am thankful to the Deutscher Akademischer Austauschdienst (DAAD) for a doctoral fellowship to pursue my Ph.D. studies at Jacobs University, Germany.

My deepest appreciation goes to my best friends Omar Al Hafez, Zayd Al Hafez, Rami Al Oweini, Mohammad Habbal, Adnan Sinan, Amer Bralic, Hashem Al Ghaili, Khalid Shawwa and Mahmoud Ghalayini for their unconditional patience, love and full support. Finally, always indebted to a loving family whose nurture, support and love drove me limits afar and raised me mountains high, I dedicate to them this achievement. Thank you Dad (Ahmad), Mom (Samira), Mohammad, Hanadi, Riham, Manal and my niece Maria, your presence radiates my life and makes me a better person.

Summary

Pseudomonas syringae is a phytopathogenic γ-proteobacterium that induces a wide variety of diseases on various agronomically significant crops, as well as on an unknown number of wild plant species. Virulence of the bacterial blight pathogen of soybean, *P. syringae* pv. *glycinea* PG4180, is favored by cold and humid conditions. This bacterium can synthesize the exopolysaccharide levan when it encounters moderate to high concentrations of sucrose. Interestingly, this process is also temperature-dependent and is mediated by the extracellular enzyme levansucrase. Although the presence of multiple alleles for levansucrase gene in *P. syringae* PG4180 has been the focus of many previous studies, no regulators have been described for expression of levansucrase. In the first study, the genome of *P. syringae* PG4180 was screened for transcriptional activators and results revealed that the prophage-borne transcriptional regulator, LscR, mediates expression of levansucrase. A *lscR*-deficient mutant was generated and exhibited a levan-negative phenotype when grown on a sucrose-rich medium. Zymographic analysis and Western blots confirmed the observed mutant phenotype. Transcriptional analysis and electrophoretic mobility shift assays confirmed that LscR regulate expression levels of levansucrase and glycosyl hydrolase genes by binding to their upstream promoter region. Furthermore, genomic analysis of the region surrounding the *lscR* gene revealed that LscR resides on a ~25 kb fragment of a bacteriophage origin. To determine its nature and function, the prophage region was subjected to nucleotide sequencing followed by BLAST analysis. Results showed that this region does not encode for an active phage as it only possesses genes involved in phage tail morphogenesis. In addition, genome comparative alignment with other *Pseudomonas* strains revealed that the possible origin of this region is phage MU or SfV. In the last section of this study, we investigated a potential repressor for the transcription of *lsc*. Potential binding sites for the hexose metabolism repressor, HexR, were found in the PAPE region upstream of both *lsc* genes. A *hexR* mutant of *P. syringae* PG4180 was generated and transcriptional analysis revealed a slight increase in the expression levels of *lsc*. In addition, *P. syringae* PG4180 were significantly growth-impaired when incubated with sucrose or glucose as sole carbon source and the mutant's capability to multiply *in planta* was reduced. Our data proposed that genes engaged in extra-cellular sugar acquisition are co-regulated with those involved in intra-cellular energy-providing metabolic pathways in *P. syringae*.

List of Abbreviations

5-ALA	5-aminolevulinic acid
Ap^r	Ampicillin resistance
BCA	Bicinchoninic acid
bp	Base pairs
BLAST	Basic Local Alignment Search Tool
CFU	Colony Forming Unit
Cm^r	Chloramphenicol resistance
dsDNA	Double stranded DNA
dsRNA	Double stranded RNA
DTT	Dithiothreitol
Eda	2-keto-3-deoxy gluconate aldolase
Edd	6-phosphogluconate dehydratase
EDP	Entner-Doudoroff pathway
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
EPS	Exopolysaccharide
Gap	Glyceraldehyde 3-phosphate dehydrogenase
Gm^r	Gentamicin resistance
Glk	Glucose kinase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HexR	Hexose metabolism repressor
H-NS	Histone nucleoid structuring like protein
Hrp	Hypersensitive response and pathogenicity
HS	Hoitink-Sinden
HSA	Hoitink-Sinden supplemented with L-arabinose
kb	Kilo base pairs
KB	King's B
Km^r	Kanamycin resistance
LB	Luria-Bertani

Lsc	Levansucrase
LscR	Levansucrase regulator
MBP	Maltose binding protein
MG	Mannitol-glutamate
MGE	Mobile genetic elements
NCBI	National Center of Biotechnology Information
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAPE	Prophage assosciated promoter element
PCR	Polymerase chain reaction
Pgl	6-phosphoglucose lactonase
PFP	Polyhedral filamentous pleomorphic
qRT-PCR	Quantitative Reverse Transcriptase Polymerase Chain Reaction
RT-PCR	Reverse-Transcriptase Polymerase Chain Reaction
SDS	Sodium dodecyl sulphate
Sp^r	Spectinomycin resisitance
TAE	Tris base, acetic acid and EDTA
ssDNA	Single stranded DNA
ssRNA	Single stranded RNA
Tc^r	Tetracycline resistance
TCA	Tricarboxylic acid cycle
TSS	Transcriptional start site
TTSS	Type three secretion system
Zwf	Glucose-6-phosphate dehydrogenase

1. Introduction

1.1 The Soybean

Soybean is a leguminous plant that is native to the warm and tropical regions in Asia. It has originated as a food crop in Northeastern China around 1100 B.C and since then it has risen to one of the top-traded commodities (Hartman *et al.*, 2011). Soybean is known for its unique nutritional value as their seed content contains around 38% proteins and 18% oil, hence it is one of the most significant resource of plant protein in the world (Hartman *et al.*, 2011).

Soybean have several applications from livestock and aquaculture feeds to usage in industrial products such as plastics, cosmetics and many other products (Hartman *et al.*, 2011). In addition, soybean may offer medicinal value as reducing cholesterol (Rosell *et al.*, 2004), increase iron in the blood (Murray-Kolb *et al.*, 2003), nutritional supplements for diabetics (Azadbakht *et al.*, 2003, Villegas *et al.*, 2008) and others. All these factors increased the demand for soybean over the past 40 years to an extent that 6% of the world's arable land in 2008 are used for soybean production (Hartman *et al.*, 2011). As a matter of fact, global soy production is expected to almost double by 2050 compared to 2012 because of the continuous increase in the world population (**Fig. 1**).

A

Figure 1A is found in the link below:
http://wwf.panda.org/what_we_do/footprint/agriculture/soy/facts/

B

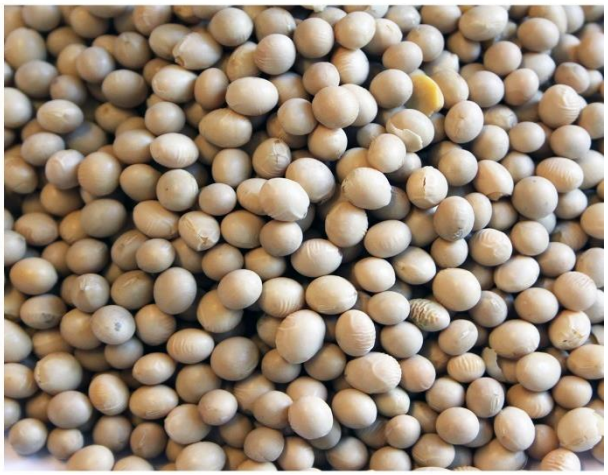


Figure 1: Economic impact of soybean. A) Increase in global soy production in millions of tonnes (http://wwf.panda.org/what_we_do/footprint/agriculture/soy/facts/). B) Soybean seeds.

However, there are several factors that can threaten and challenge this billion-dollar industry. Of these factors are abiotic restraints caused by environmental conditions such as weather, flooding, drought, soil salinity and many others (Lal, 2009). Unfortunately, only few of these problems can be tackled through farming practices. Another factor affecting soybean yield is the biotic restraints such as pests, weeds and pathogens which can infect all parts of a soybean plant (**Fig. 2**) (Strange & Scott, 2005). Although the severity of economic plant damage depends on many factors, it is of great importance to understand the pathogenesis mechanisms of these plant pathogens.

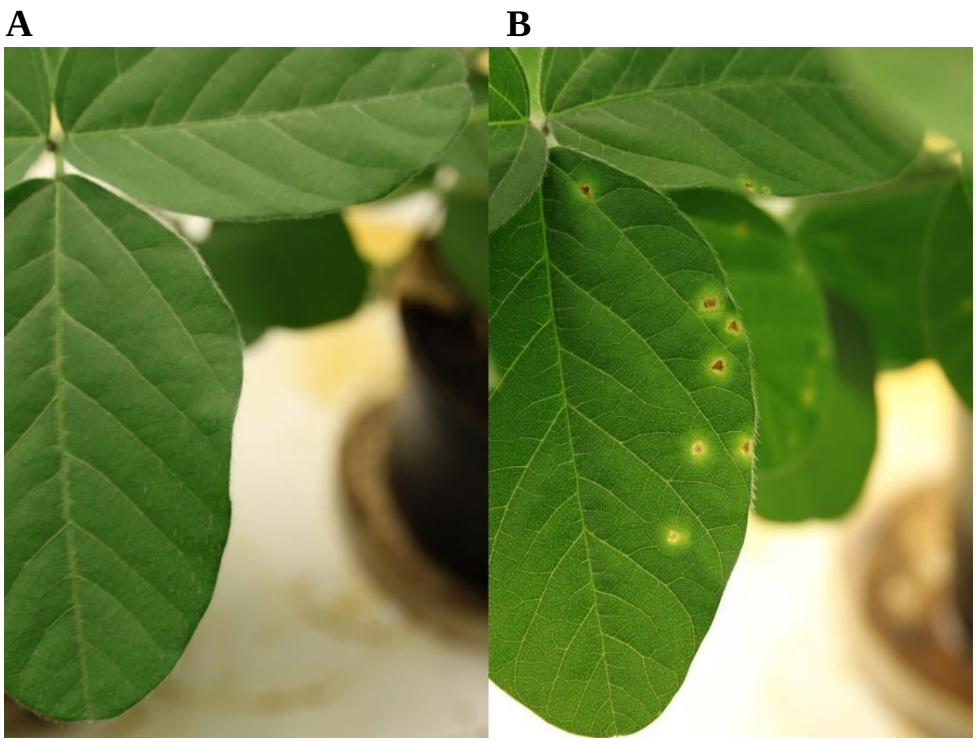


Figure 2: Soybean leaf. A) Healthy soybean leaf. B) Soybean leaf infected by the plant pathogen *P. syringae* pv. glycinea PG4180.

1.2 Plant pathogens

Plant pathogenic bacteria are characterized by their ability to induce local or systemic irregular physiological functioning on vulnerable plants. Such diseases can influence the integrity or development of the plant and can lead to its localized tissue injury or death (**Fig. 3**) (Tampakaki *et al.*, 2009). In addition to the significant difference in shape, size and method of multiplication, plant pathogens also differ in host specificity. Some infect one or a few strains within one plant species, while others can infect most or all plants belonging to a single species, to a genus, or to a number of plant families (Agrios, 2009).

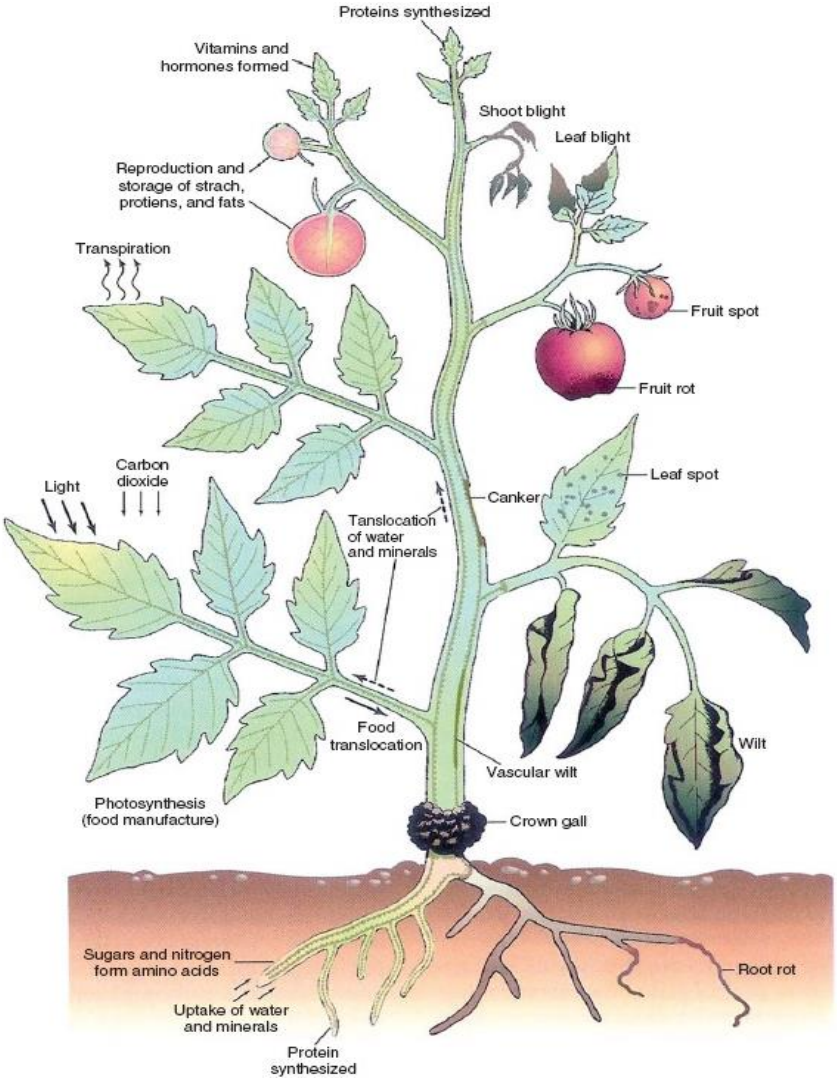


Figure 3: Typical plant of which the left half is showing the basic functions of its main organs, whereas the right half is showing the various symptoms of infection by pathogens and their interference with the basic functions (Agrios, 2009).

1.2.1 *Pseudomonas syringae*

Pseudomonas syringae, a phytopathogenic γ-proteobacterium, is taxonomically subdivided into at least 50 pathovars (Hirano & Upper, 2000) that induces a wide variety of leaf spots, speck, and blight diseases on various agronomically significant crops, as well as on an unknown number of wild plant species (O'Brien *et al.*, 2011). The economic influence of *P. syringae* escalated, with a renaissance of old infections, such as bacterial speck of tomato (pv. *tomato*) (Shenge *et al.*, 2007), and the worldwide rise of important new diseases, for instance bleeding canker of horse-chestnut (pv. *aesculi*) (Green *et al.*, 2010). All these factors ranked *P. syringae* first in the list for top 10 plant pathogenic bacteria as most economically and scientifically important bacterial pathogen (Mansfield *et al.*, 2012).

The ability of *P. syringae* to colonize and proliferate in the apoplast of leaves and other aerial plant tissues depend on many virulence factors such as type III secretion system, effectors, phytohormones and toxins (Rico *et al.*, 2011). Moreover, the two major exopolysaccharides (EPSs) alginate and levan may play a significant role in the virulence and fitness of *P. syringae*. EPSs possess several functions such as offering armor to bacteria from drought, concentrating nutrients and decreasing exposure to lethal substances (Laue *et al.*, 2006). Furthermore, EPSs improve bacterial attachment to surfaces and survival throughout their epiphytic or saprophytic life via the formation of biofilms. Production of biofilms could be significant at some point in early or later stages of tissue colonization, while bacterial dispersion from biofilms may be necessary for movement to the vascular system (Zago & Chugani, 2009). Development of biofilms involves two diverse classes of *Pseudomonas* polysaccharides. Capsular polysaccharides, such as alginate and levan, engulf the bacteria and provide it with a protective layer. Moreover, alginate is involved in water retention while levan played a role in nutrient preservation and protection against starvation (Sutherland, 2001a, Sutherland, 2001b).

Studies showed that phytopathogenic *Pseudomonas* are able to produce alginate similar to what has been previously illustrated for *P. aeruginosa* (Schnider-Keel *et al.*, 2001). Furthermore, different species also secrete levan, which support biofilm persistence and colonization. However, recent results verified that neither levan nor alginate or the alternative sigma factor, AlgT, were important for biofilm development in *P. syringae*. Therefore, results indicated the presence of

additional undiscovered EPS needed for biofilm structure formation and that the main function of levan remains unclear (Laue *et al.*, 2006).

This lead to the discovery of aggregative polysaccharides that do not surround the bacteria but interacts with other matrix components to provide structural integrity to the biofilm (Mann & Wozniak, 2012). Matsukawa and Greenberg found the polysaccharide synthesis locus (*psl*), which was able to generate polysaccharide matrix components (Matsukawa & Greenberg, 2004). Psl is composed of a pentasaccharide that is made up of D-mannose, D-glucose, and L-rhamnose. In addition, another search recognized a fourth locus termed pellicle formation defect (*pel*) whose structure is still undefined (Friedman & Kolter, 2004b, Friedman & Kolter, 2004a).

1.2.2 *Pseudomonas syringae* pv. glycinea PG4180

Virulence of the bacterial blight pathogen of soybean, *Pseudomonas syringae* pv. *glycinea* PG4180, is favored by cold and humid conditions. Symptoms of such infection are water-soaked lesions that develop into necrotic leaf spots surrounded by chlorotic halos (**Fig. 4**) (Li *et al.*, 2006). Thus uncovering the mechanism by which the infection arises and the nature of the causing bacteria is of critical significance for crop protection, enhancement and development of resistant soybean lines and varieties (Ignjatov *et al.*, 2008).

Figure 4: Typical symptoms induced by *Pseudomonas syringae* on the soy bean leaf: necrotic spots are surrounded by chlorotic, yellowish halos.

Bacterial growth on plant leaves is often limited by the presence of carbon compounds; an example would be plant-borne sucrose. The presence of sugars on a leaf surface drives the amount of epiphytic bacterial population growth. A bacterial saccharolytic enzyme such as levansucrase consists of a primary metabolic pathway to use sucrose as the main energy source (Mercier & Lindow, 2000a). During early phases of colonization, *P. syringae* establish a nutritional association with the living cells of its host in two main forms: the acquisition and up-regulation of genes that encode transporters and enzymes that aid absorption of nutrients existing in a specific environment, and the loss or down-regulation of genes that encode transporters and enzymes that are not needed, or which have a harmful influence on growth in the plant environment (Rico *et al.*, 2011).

1.3 Levan and levansucrase

Levan is a β-2,6- linked fructose homopolymer with some β-2,1- linked branching points (**Fig. 5**) and is produced by an enzymatic reaction using levansucrase (Kulminskaya *et al.*, 2003). The generation of levan occurs when PG4180 encounter moderate to high concentration of sucrose. Unlike other *P. syringae* pathovars, PG4180 carries a non-sense mutation that lead to a truncation of the gene coding for the alternative sigma factor, AlgT, thus leaving levan as the major EPS of this strain (Schenk *et al.*, 2006).

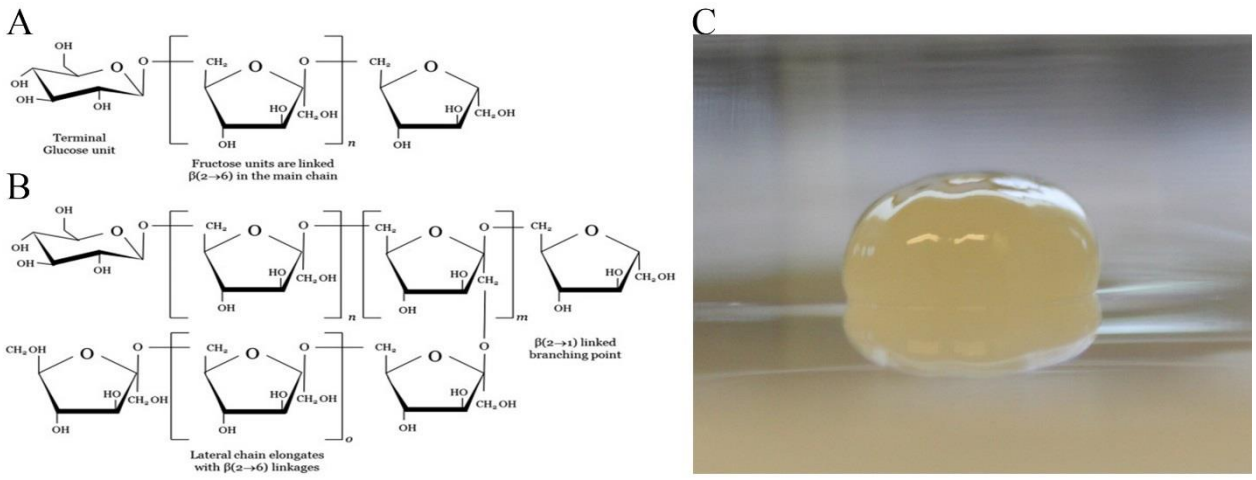


Figure 5: Schematic representation of A) linear and B) branched levan (Oner *et al.*, 2016)**.** C) Levan formation on 5% sucrose agar plates.

Levansucrase (Lsc) (EC 2.4.1.10) is an extracellular enzyme that belongs to the superfamily of ß-fructosidases. It is characterized by its high stability, no need for co-factors, and ease of enzymatic detection due to its glucose releasing activity.

Lsc has a 5-bladed ß-propeller fold and catalyzes the following reactions (Gross & Rudolph, 1987):

1- Synthesis of levan from sucrose by transfructosylation while releasing glucose
2- Hydrolysis of levan to monosaccharides of fructose
3- Exchange of [^{14}C] glucose in the reaction fructose-2,1-glucose + [^{14}C] glucose to fructose-2,1-[^{14}C] glucose + glucose

Although Lsc was found in gram-negative bacteria such as *Erwinia amylovora* as well as in gram-positive bacteria such as *Bacillus subtilis* (SacB), however the enzymatic properties, protein location, peptide sequences and transcriptional regulation differ between both types. Lsc is constitutively secreted in gram-negative bacteria in contrast to its substrate-dependent expression in gram-positive bacteria (Hettwer *et al.*, 1995).

The *lsc* gene encoding levansucrase was first identified using nucleotide sequence analysis of a 1,248-bp open reading frame (ORF) obtained from PG4180. This sequence revealed high similarity to the nucleic acid and deduced amino acid sequences of *lsc* derived from *E. amylovora* and *Zymomonas mobilis* (Hettwer *et al.*, 1998). However, marker exchange mutagenesis of *lsc* didn't exhibit a levan-negative phenotype, indicating the presence of at least one extra allele in PG4180. Southern blot hybridizations confirmed the presence of three *lsc* genes: *lscA* and *lscC* located on the chromosome, while *lscB* was located on a 60-kb plasmid (Li & Ullrich, 2001). PCR screening confirmed the occurrence of multiple Lsc isoenzymes in further *P. syringae* pathovars. Interestingly, only *lscB* and *lscC* were actually expressed in strain PG4180 and showed 98% similarity in their nucleotide sequences, whereas *lscA* was transcriptionally inactive and differed by approximately 14% in its nucleotide sequence compared to *lscB* and *lscC* (Li & Ullrich, 2001).

Furthermore, a study performed by Li *et al.* (2006) was the first to present an evidence for thermo responsive expression and secretion of levansucrase in the plant pathogen *P. syringae* pv. *glycinea*. Genes *lscB* and *lscC* were shown to be functional and expressed in a temperature dependent manner: maximum mRNA production for these genes occurred at 18°C, when the pathogen is

highly virulent, and is repressed at 28°C, the bacterium's most favorable growth temperature (Li *et al.*, 2006).

In addition to all *lscB* or *lscC* variants found in different *P. syringae* pathovars, the corresponding *lscA* variants lack a ~450-bp prophage borne DNA region located upstream of *lscB* and *lscC*. This region, in addition to a 48-bp DNA encoding the common N-terminal sequences of all 1,296-bp Lsc, is termed phage-associated promoter element (PAPE) (Srivastava *et al.*, 2012). The importance of the PAPE was revealed in a study by Khandekar et al 2014 where they demonstrated that the 450-bp upstream region, which carries the translational start sequence of *lsc* variants BC and a prophage-associated *com* gene, is needed for expression of *lsc*. While the 48-bp DNA N-terminal sequence is shown to increase level of *lsc* expression (Khandekar *et al.*, 2014). Moreover, the PAPE is also found upstream of several genes encoding glycosyl hydrolases in PG4180, B728a, and 1448a (Srivastava *et al.*, 2012). Thus it was concluded that *lscA* might have been the ancestral *lsc* gene in *E. amylovora* and *P. syringae*, which explains why it was expressed from its native promoter in the Lsc-deficient *E. amylovora* mutant, Ea7/74-LS6 (Srivastava *et al.*, 2012). Interestingly, the prophage origin of the PAPE region and its function in levansucrase expression hints to an important role for bacteriophages in evolution of bacteria.

1.4 Bacteriophages of *Pseudomonas*

Up to 50% of the daily produced bacteria are killed by bacteriophages, which are viruses that infect bacteria. Encapsulated within a protein or lipoprotein coat, a bacteriophage virion carries either single-stranded or double-stranded DNA or RNA molecule (Ceyssens & Lavigne, 2010). Bacteriophages can be divided into two groups, lytic and temperate. Lytic phages replicate and lyse their host bacterial cells while temperate phages integrate into the DNA of their bacterial host or exist as independent plasmids. Integrated phages are named prophages and inserted at precise integration sites in the host genome, although their location can vary depending on the phage species (Casjens, 2003).

Although prophages reduce the fitness of its lysogenic host by the metabolic load to replicate extra DNA (Canchaya *et al.*, 2003b), they supply their hosts with important agents that are needed to survive in specific ecological niches (Srividhya *et al.*, 2007). In fact, bacterial pathogens encode several virulence factors, such as toxins, pili (fimbriae), adhesins and secretion systems, which are

of phage origin (Wagner & Waldor, 2002). Therefore, prophages can represent a considerable part of the bacterial genome where the most severe case was found in *Escherichia coli* O157:H7 Sakai. This strained harbored 18 prophage genome elements, which referred to 16% of its total genome content (Canchaya *et al.*, 2003b).

The classification of newly identified phages relies on two main criteria set by the International Committee on the Taxonomy of Viruses (ICTV): Virion morphology (tail type, polyhedral, filamentous and pleomorphic) and nucleic acid composition (dsDNA, ssDNA, dsRNA or ssRNA) (Ceyssens & Lavigne, 2010).

More than 97% of the phages found in *Pseudomonas* belong to the order *Caudovirales* which are characterized by their dsDNA and tailed phages. All the others (<3%) belong to the order PFP, which are distinguished by a polyhedral, filamentous or pleomorphic morphology (Ackermann, 2007).

Caudovirales phages

Three main phage families are found under the order of *Caudovirale*:

1- *Myoviridae*: Have a long, helical tail that contracts upon infection.

2- *Siphoviridae*: Represent around 47% of *Pseudomonas* phages and described by a long and flexible tail (Knezevic *et al.*, 2009). An important species under this family is the B3-like *Pseudomonas* phages. They harbor type IV pili-specific phages that show similarity to the myovirus Mu. Their 40-kb genome is integrated randomly into the host genome independent of the prophage's cycle (Morgan *et al.*, 2002).

3- *Podoviridae*: Have a short tail and exhibit slight variation in capsid size and tail structure.

PFP phages

Three main phage families are found under the order of *PFP* phages:

1- *Inoviridae*: Infects Gram-negative bacteria by attaching to the bacterial pilus. Moreover, it is characterized by rod-shaped filamentous phages and a closed ssDNA genome (Mooij *et al.*, 2007).

2- *Cystoviridae*: Includes phages ϕ6- ϕ14 and characterized by a genome made of three segments of dsRNA: large, medium and small. Moreover, a lipid containing membrane surrounds a polyhedral shell which in turn encloses the phage genome (Mindich *et al.*, 1999). *P. syringae* pv. *phaseolicola* is naturally infected by ϕ6 which carries a genome size around 13.3 kb and encodes for 14 gene products. Unlike ϕ8, ϕ12, and ϕ13 which bind to rough LPS, ϕ6 identify and bind to the type IV pili (Qiao *et al.*, 2010).

3- *Leviviridae*: Harbors its ssRNA genome in an icosahedral capsid that is stabilized by calcium ions. The size of the genome is around 3.5 kb and carries four genes that encode a protein with lysis function, coat and maturation proteins a replicase (Olsthoorn *et al.*, 1995).

1.5 Glucose metabolism in *Pseudomonas*

Glucose is a byproduct released during the synthesis of levan, therefore it was important to investigate the effect of levansucrase on glucose metabolism in *Pseudomonas*. There are at several different routes or pathways for glucose metabolism in *Pseudomonas*. A clear indication of biochemical diversity is resembled in the three different pathways which convert sugar into 6-phosphogluconate are illustrated in the model organism *P. putida* (**Fig. 6**). The glucose kinase pathway is one of the pathways where glucose is carried to the cytoplasm by the glucose transport system. Another pathway is the gluconokinase in which gluconokinase mediates the direct phosphorylation of gluconate or the oxidation of glucose to gluconate via glucose dehydrogenase (Daddaoua *et al.*, 2010). The last pathway is one where 2-ketogluconate is phosphorylated and internalized into 6-phosphogluconate by 2-ketogluconate kinase (Kim *et al.*, 2008, Daddaoua *et al.*, 2010).

The end result of the three pathways is 6-phosphogluconate. The latter enters the Entner-Doudoroff pathway (EDP) by the Edd enszyme (phosphogluconate dehydratase). After that, the Eda enzyme (2-keto-3-deoxy gluconate aldolase) hydrolyzes it to produce glyceraldehydes-3-phosphate and pyruvate (Kim *et al.*, 2008). Furthermore, Gap-1 enzyme metabolizes Glyceraldehyde 3–phosphate whereas pyruvate enters the krebs cycle after it is decarboxylated to Acetyl-CoA (Daddaoua *et al.*, 2009, del Castillo *et al.*, 2008).

The lack of some of the glycolysis metabolic genes, such as that encoding for 6-phosphofructokinase which changes fructose-6-phosphate into fructose 1-6-bisphosphate, renders the EDP as the dominant glucose-utilization pathway in all pseudomonads. It should be noted that EDP is present in many archeal, bacterial and even some eukaryotic species (*Entamoeba* spp, *Aspergiluus* spp. etc.), hence it's not restricted to pseudomonads. Apart from *Pseudomonas*, this pathway is dominant in organisms such as *Xanthomonas* and *Rhizobium*. In enterobacteria, EDP is used for exploitation of specific carbon sources such as gluconate, on the other hand, a modified version of the EDP is exploited in carbohydrate catabolism in *Clostridia* or *Archeae* (Daddaoua *et al.*, 2009, del Castillo *et al.*, 2008). Furthermore, some hypotheses suggest that this particular pathway is believed to be part of specific biosynthetic pathways since it is known to be present in several non-carbohydrate utilizers (Daddaoua *et al.*, 2009, del Castillo *et al.*, 2008). Surprisingly, the enzymes involved in the gluconate pthway are not present in *P. syringae* pv. Phaseolicola 1448A, the closest relative of *P. syringae* – based on the KEGG pathway system (Winsor *et al.*, 2011).

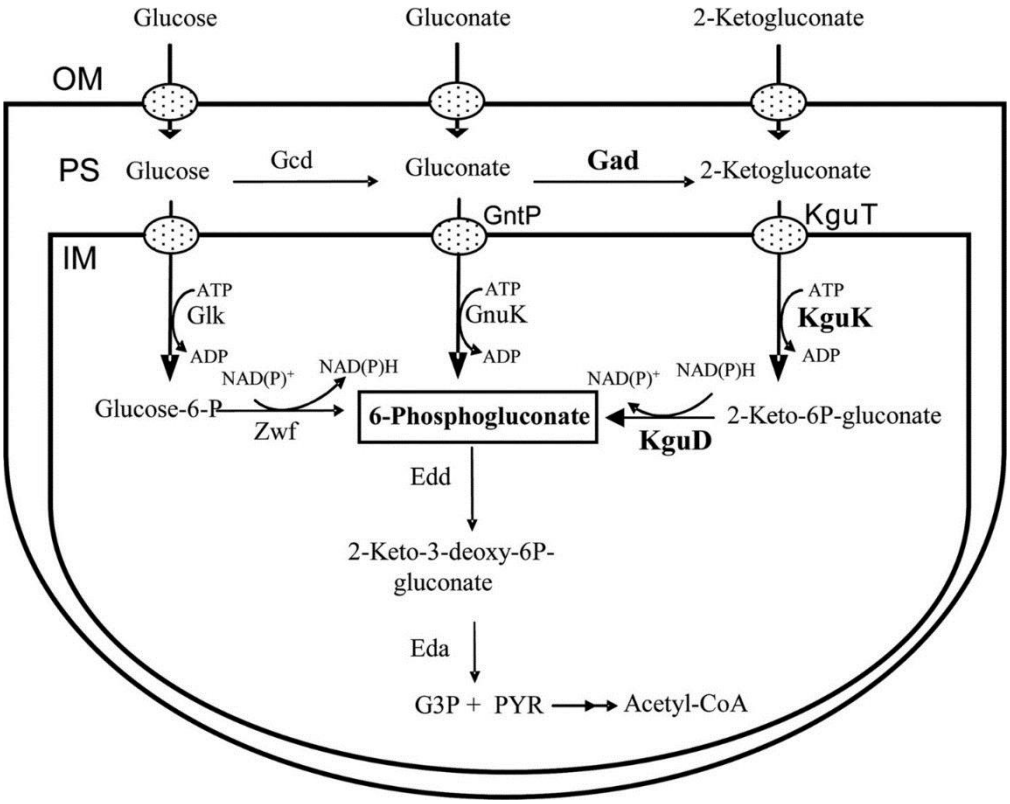


Figure 6: Schematic presentation of glucose metabolism in *P. putida* KT2440. OM, outer membrane; PS, periplasmic space; IM, inner membrane; Gcd, glucose dehydrogenase; Gad, gluconate dehydrogenase; KguD, 2-ketogluconate reductase; Glk, glucokinase; GnuK, gluconokinase; KguK, 2-ketogluconate kinase; Zwf-1, glucose-6-phosphate 1-dehydrogenase; Pgl, 6-phosphoglucose lactonase; Edd, phosphogluconate dehydratase; Eda, 2-keto-3-deoxy gluconate aldolase; GntP, gluconate permease; KguT, 2-ketogluconate transporter; PYR, pyruvate (Daddaoua *et al.*, 2010).

2 Aims of the study

The major aim of the present study is to identify regulatory protein(s) responsible for the expression of levansucrase in *P. syringae* pv. glycinea PG4180. A screening of a *P. syringae* genomic cosmid library will be conducted in the heterologous host *P. putida* KT2440. The cosmid of interest will be subjected to *in vitro* transposon mutagenesis to find the gene(s) responsible for the *lsc* expression phenotype. In order to investigate its function in *P. syringae*, a gene-deficient mutant will be generated and tested for a levan-negative phenotype. Proteins from total cell lysates and cell-free supernatants from *P. syringae* PG4180 will be assessed for presence of levansucrase using Western blots and zymograms. Finally, the mechanism used to regulate expression of levansucrase will be investigated using qRT-PCR and electrophoretic mobility shift assay.

The second aim is to assess the function of the prophage region spanning between *trpE* and *trpG* and harboring an essential transcriptional regulator for levansucrase expression. For this purpose, the cosmid harboring the prophage region will be subjected to nucleotide sequencing and open reading frames will be annotated using BLAST-N/BLAST-P. Additionally, we will perform GC profile analysis and genome alignment with other *Pseudomonas* strains.

The third aim of this study is to investigate the role of a hexose metabolism repressor HexR, known for its role in cellular glucose metabolism, in regulating levansucrase expression. A *hexR* mutant will be generated in *P. syringae* PG4180 and tested for its growth in minimal medium containing different carbon sources. In addition, *lsc* expression will be studied using qRT-PCR and Western blotting. *In planta* assays were conducted using *P. syringae* PG4180 wild type and its *hexR* mutant to evaluate effect of HexR on bacterial fitness and virulence.

3 Results

Results of the present doctoral thesis are presented in the following three manuscripts:

The bacteriophage-derived transcriptional regulator, LscR, activates the expression of levansucrase genes in *Pseudomonas syringae*

Khaled Abdallah, Katharina Hartman, Daniel Pletzer, Daria Zhurina, and Matthias S. Ullrich

Published in *Molecular Microbiology* (2016). doi: 10.1111/mmi.13536

KA designed experiments, generated mutagenic construct, conducted levan formation, Western blots, zymogram, qRT-PCR, and prepared the manuscript. KA and KH conducted protein purification. KA and DP mutagenized the lscR gene. DZ screened for regulators and conducted transposon mutagenesis. MU designed the study, supervised the work, discussed the results, and contributed to manuscript writing. All authors read and approved the final manuscript.

Analysis of the prophage-derived genomic region flanked by *trpE* and *trpG* in *Pseudomonas syringae* pv. glycinea PG4180

Khaled Abdallah and Matthias Ullrich

KA conducted all the experimental design, practical work and prepared the manuscript. MU designed the study, supervised the work, discussed the results, and contributed to manuscript writing. The manuscript is in preparation.

Expression of extra-cellular levansucrase in *Pseudomonas syringae* is controlled by the in planta fitness-promoting metabolic repressor HexR

Amna Mehmood, Khaled Abdallah, Shaunak Khandekar, Daria Zhurina, Abhishek Srivastava, Nehaya Al-Karablieh, Daniel Pletzer, Gabriela Alfaro-Espinoza, and Matthias S. Ullrich

BMC Microbiology (2015) 15: 48

AM designed experiments, conducted qRT-PCR and RT-PCR, and prepared the manuscript; KA designed experiments, conducted *in planta* experiments, and prepared the manuscript; SK conducted growth curve experiments and Western blots; AS and DP conducted the in silico analysis lsc upstream sequences; DZ and GA mutagenized the hexR gene; NK conducted the hypersensitive response assay; MU designed the study, supervised the work, discussed the results, and contributed to manuscript writing. All authors read and approved the final manuscript.

3.1 The bacteriophage-derived transcriptional regulator, LscR, activates the expression of levansucrase genes in *Pseudomonas syringae*

Summary

Synthesis of the exopolysaccharide levan occurs in the bacterial blight pathogen of soybean, *Pseudomonas syringae* pv. glycinea PG4180, when this bacterium encounters moderate to high concentrations of sucrose inside its host plant. The process is mediated by the temperature-dependent expression and secretion of two levansucrases (Lsc), LscB and LscC. Previous studies showed the importance of a prophage-associated promoter element in driving the expression of levansucrase genes. Herein, heterologous screening for transcriptional activators revealed that the prophage-borne transcriptional regulator, LscR, from *P. syringae* mediates expression of levansucrase. A *lscR*-deficient mutant was generated and exhibited a levan-negative phenotype when grown on a sucrose-rich medium. This phenotype was confirmed by zymographic analysis and Western blots which demonstrated absence of levansucrase in the supernatant and total cell lysates. Transcriptional analysis showed a down-regulation of expression levels of levansucrase and glycosyl hydrolase genes in the *lscR*-deficient mutant. Ultimately, a direct binding of LscR to the promoter region of levansucrase was demonstrated using electrophoretic mobility shift assays allowing to conclude that a bacteriophage-derived regulator dictates expression of bacterial genes involved in *in planta* fitness.

Keywords: Plant pathogen, Bacterial blight, *Pseudomonas syringae*, Levansucrase, Prophage-borne transcriptional regulator, LscR.

Introduction

Pseudomonas syringae is a phytopathogenic Gamma-proteobacterium that induces a variety of diseases on various agronomically important crops and wild plant species (O'Brien *et al.*, 2011). *P. syringae* are classified into more than 50 different pathovars according to distinct host specificities (Young *et al.*, 1996). The virulence of *P. syringae* is triggered under cold and humid conditions (Dunleavy, 1988) where symptoms of infections are characterized by water-soaked lesions that develop into necrotic leaf spots enclosed by cholortic halos (Li & Ullrich, 2001). Aside of well-described pathogenicity factors such as type III secretion systems and its effectors (Lindeberg *et al.*, 2008, Xin & He, 2013), several virulence factors of *P. syringae* such as the production of major exopolysaccharides (EPS) alginate and levan (Osman *et al.*, 1986), play an important role during early stages of infection (Osman *et al.*, 1986, Huynh *et al.*, 1989, Denny, 1995, Kasapis *et al.*, 1994).

The bacterial blight pathogen of soybean, *P. syringae* pv. glycinea PG4180, carries a nonsense mutation in the *algT* gene which causes absence of alginate production (Schenk *et al.*, 2006). Thus it is a suitable organism to study levan regulation as levan is its major EPS. Levan is a β-(2,6)-linked fructosyl homopolymer with extensive branching through β-(2,1) linkages (Hettwer *et al.*, 1998) and is synthesized by the activity of levansucrase (EC 2.4.1.10, Lsc) (Hettwer *et al.*, 1995). In PG4180, Lsc is encoded by three alleles: *lscA* and *lscC* are located on the chromosome, while *lscB* is located on a native plasmid (Li & Ullrich, 2001). Interestingly, only *lscB* and *lscC* are expressed in *P. syringae* and show 98% identity towards each other in their nucleotide sequences (Schenk *et al.*, 2006). In contrast, *lscA* is transcriptionally inactive and differed by approximately 14 % in its nucleotide sequence as compared to *lscB* and *lscC* (Li & Ullrich, 2001, Srivastava *et al.*, 2012). In addition to both *lscB* or *lscC* variants being present in many different *P. syringae* pathovars, all *lscA* variants lack a ~450-bp prophage-borne DNA region located upstream of the *lscB* and *lscC* genes (Srivastava *et al.*, 2012). This region, in addition to a 48-bp DNA encoding the common 5'-region sequences of all *lscB/C*, was termed phage-associated promoter element (PAPE) (Srivastava *et al.*, 2012). The importance of the PAPE was revealed by Khandekar *et al.* (2014) who demonstrated that this 450-bp upstream region, which carries the translational start sequence of the gene variants *lscB/lscC* and a prophage-associated *com* gene, was needed for expression of *lsc*. The 48-bp 5'-region of *lsc* genes was shown to increase the level of *lsc* expression

(Khandekar *et al.*, 2014). Moreover, the PAPE was also found upstream of several genes encoding glycosyl hydrolases in *P. syringae* strains PG4180, B728a, and 1448A (Srivastava *et al.*, 2012).

Li *et al.* (2006) presented evidence for the temperature-dependent expression and secretion of Lsc in *P. syringae*: maximum production occurred at 18°C, when the pathogen is highly virulent, and is minimal at 28°C, the bacterium's most favorable growth temperature. Expression and regulation of Lsc were previously studied in several Gram-negative bacteria such as *Erwinia amylovora* (Bereswill & Geider, 1997), *Zymomonas mobilis* (Song *et al.*, 1999), *Rahnella aquatilis* (Seo *et al.*, 2002), and *Gluconacetobacter diazotrophicus* (Martinez-Fleites *et al.*, 2005). However, results show different mechanisms controlling these processes. For instance, a single *lsc* gene in the plant pathogen *E. amylovora* was shown to be plant-inducible (Zhao *et al.*, 2005), and its expression was down-regulated by overexpression of two regulatory proteins, RcsA and RcsB (Bereswill & Geider, 1997). In the bacterium *R. aquatilis*, maximum expression of the homologous *lsc* gene *lsrA* was found to be growth-phase dependent and controlled by two regulators, an activator (LsrS) and a repressor (LsrR), located upstream of the *lsc* gene (Seo *et al.*, 2002). In the plant endophyte *G. diazotrophicus*, the gene encoding Lsc is located in an operon with a levanase gene where both are co-transcribed (Martinez-Fleites *et al.*, 2005). Interestingly, in *Z. mobilis* the lack of a phosphoenolpyruvate-dependent carbohydrate phosphotransferase (PTS) system was probably the reason for transcription of Lsc-encoding gene *levU* as a bicistronic mRNA along with a sucrase gene (Song *et al.*, 1999). Moreover, both genes were significantly induced by sucrose (Song *et al.*, 1999).

The aim of the current study was to identify a regulatory protein in *P. syringae* that is responsible for the expression of Lsc genes. A screening of a *P. syringae* genomic cosmid library in the heterologous host *P. putida* KT2440 was conducted. A prophage-borne transcriptional regulator, LscR, needed for expression of Lsc was identified and found to bind to the *lsc* upstream sequence. A *lscR*-deficient mutant was generated and its levan-negative phenotype confirmed by Western blotting and zymographic detection. Furthermore, expression of genes downstream of the PAPE element was down-regulated in the *P. syringae lscR*-deficient mutant as compared to the wild type when tested by qRT-PCR analysis.

Experimental Procedures

Bacterial Strains, plasmids, and growth conditions

Bacterial strains and plasmids used in this study are listed in **Table 1**. *Escherichia coli* DH5α was maintained at 37°C on Luria-Bertani (LB) medium (Sambrook & Russell, 2001). *E. coli* ST18 was grown on LB agar plates containing 50 μg ml^{-1} 5-aminolevulinic acid. *P. syringae* and *P. putida* were routinely maintained at 28°C or 18°C on mannitol-glutamate (MG) medium (Keane P, 1970) or King's medium B (KB) (King *et al.*, 1954). For culturing *P. syringae* in liquid medium, Hoitink-Sinden medium supplemented with L-arabinose (HSA) was used (Bender *et al.*, 1993). Phenotypic assessment of *P. syringae* and *P. putida* transconjugants was done by streaking the cells on MG agar medium containing 5% sucrose as described by Li and Ullrich (2001). Bacterial growth was monitored by measuring the optical density at 600 nm (OD_{600}). Antibiotics were added to media at the following concentrations (μg ml^{-1}): ampicillin, 50; kanamycin, 25; tetracycline, 25.

Table 1: Bacterial strains and plasmids used in this study

Strain	Relevant characteristics	Reference
Bacteria		
Escherichia coli		
DH5α	supE44 D*lac*U169 (F80 *lacZ*DM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	Sambrook and Russell (2001)
ST18	λ*pir* Δ*hemA pro thi hsdR*$^+$ Tp^r Sm^r chromosome::RP4-2 Tc::Mu-Km::Tn7	Thoma and Schobert (2009)
Pseudomonas syringae pv. glycinea		
PG4180	Wild type, levan positive	Bender *et al.* (1993)
PG4180.*lscR*::Km	Km^r, *lscR*-deficient mutant of PG4180, levan negative	This study
Pseudomonas putida		
KT2440	Wild type, levan negative	Ramos-Diaz and Ramos (1998)

Plasmids		
pBBR1MCS-2	Km^r, broad-host-range cloning vector	Kovach *et al.* (1995a)
pLB7.2	Ap^r, contains *lscB* on 7.2-kb *Eco*RV fragment in pBluescript	(Li & Ullrich, 2001)
plscB	Km^r, contains *lscB* on 7.2-kb *Eco*RV fragment in pBBR1MCS-2	This study
pBBR1MCS-3	Tc^r, broad-host-range cloning vector	Kovach *et al.* (1995a)
pBluescript II SK	Ap^r, broad host range cloning vector	Stratagene
pBluescript. lscR	Ap^r, contains *lscR* on a 552-bp *Cla*I fragment in pBluescript II SK	This study
p1350	Tc^r, contains *lscR* and homologue of PSPPH 0652 on a 1350-bp *Sac*I fragment in pBBR1MCS-3	This study
plscR	Tc^r, contains *lscR* gene in pBBR1MCS-3	This study
p0652	Tc^r, contains homologue of PSPPH 0652 gene in pBBR1MCS-3	This study
pRK7813	Tc^r, cosmid vector used for generation of genomic libraries	Jones and Gutterson (1987)
Cos-245	Tc^r, genomic library clones of PG4180; approx. 26069 bp insert in pRK7813, promote *lscB* expression in *P. putida* carrying *lscB*	This study
Cos-532	Tc^r, genomic library clones of PG4180; approx. 24069 bp insert in pRK7813, promote *lscB* expression in *P. putida* carrying *lscB*	This study
Cos-671	Tc^r, genomic library clones of PG4180; approx. 27293 bp insert in pRK7813,	This study

	promote *lscB* expression in *P. putida* carrying *lscB*	
Cos-532::Tn	Km^r, Tc^r, Cos-532 with entranceposon insertion in *lscR*; does not promote *lscB* expression in *P. putida* carrying *lscB*	This study
pFKm	Source of Km cassette flanked with FRT sequences	Choi *et al.* (2008)
pGEM-T Easy	Ap^r, vector for cloning of PCR products	Promega
pGEM.lscR-down	Ap^r, contains 593-bp fragment downstream of *lscR* gene	This study
pGEM.lscR-down-Km	Km^r, Ap^r, 1,230-bp *Kpn*I fragment containing a Km^r cassette flanked with FRT sequences ligated into 593-bp pGEM.lscR-down	This study
pGEM.lscR-Up	Ap^r, contains 547-bp fragment upstream of *lscR*	This study
pGEM.lscR-Km	Km^r, Ap^r, carries mutagenic construct cassette for mutagenesis of *lscR* of PG4180	This study
pEX18Tc	Tc^r, *oriT*+ *sacB*+, MCS from pUC18	Hoang *et al.* (1998)
pEX18Tc.lscR-Km	Tc^r, Km^r, carries mutagenic construct cassette for mutagenesis of *lscR* of PG4180	This study
pRK2013	Km^r, helper plasmid	(Figurski and Helinski (1979))
pMal	Ap^r, encodes maltose binding protein in pMal-C2 vector	New England Biolabs
pMal-LscR	Ap^r, *lscR* gene in fusion with maltose binding protein in pMal-C2 vector	This study

Ap = Ampicillin; Km = Kanamycin; Sm = Streptomycin; Tc = Tetracycline

PCR amplifications, DNA manipulations and cloning

All primers used are listed in Table 2. Primers were designed based on *Pseudomonas syringae* pv. phaseolicola 1448A genome sequences available from NCBI (GenBankCP000058.1). Routine molecular methods were performed using standard protocols (Sambrook & Russell, 2001). Nucleotide sequencing was carried out commercially (Eurofins MWG Operon Ebersberg, Germany). PCR reactions were carried out using the DreamTaq DNA polymerase (Thermo Fisher Scientific, Schwerte, Germany) in accordance with the manufacturer's instructions. Phusion DNA polymerase (Thermo Fisher Scientific) was used for high fidelity PCR reactions. Restriction enzymes and T4 DNA ligase were purchased from Thermo Fisher Scientific and reactions were performed as per the manufacturer's instructions. DNA purifications were either performed using the GeneJET PCR purification kit (Thermo Fisher Scientific) or the GeneJET Gel extraction kit (Thermo Fisher Scientific) following the manufacturer's instructions.

Visualization of levan formation

Bacterial cells grown on MG agar plates containing 5% sucrose were used for the qualitative visualization of Lsc activity, which led to levan formation in form of a mucoid, dome-shaped colony morphology.

Screening for transcriptional activators of lsc expression using genomic library of P. syringae in the heterologous host P. putida

From plasmid pLB7.2 (Li & Ullrich, 2001), a 7.2-kb *Eco*RV fragment containing *lscB* from *P. syringae* was cloned into the *Eco*RV-treated vector pBBR1MCS-2 generating plasmid plscB. Subsequently, *P. putida* strain KT2440 cells were transformed with plscB using electroporation (resistance 200Ω, capacitance 25µF, 2,5kV) and selected on MG plates containing kanamycin at 28°C. A pooled genomic cosmid library of *P. syringae* (Hettwer *et al.*, 1998) was introduced by triparental conjugation into levan-negative *P. putida* KT2440 carrying plasmid plscB. Transconjugants were selected on MG plates supplemented with kanamycin and tetracycline. Approximately 2,000 of the resulting transconjugants were patched onto MG agar plates supplemented with 5% sucrose and incubated at 18°C for 10 days. Bacterial colonies exhibiting a dome-shaped morphology were considered levan-positive and their protein extracts further tested

for Lsc by Western blot analysis. Three individual transconjugants, termed TCos-245, TCos-532, and TCos-671, showed a levan-positive and Lsc-positive phenotype. The corresponding cosmids were isolated from the transformants, treated with restriction endonucleases, and the resulting digestion patterns compared.

Transposon mutagenesis of cosmid DNA

Transposon *in vitro* mutagenesis was performed using the TGS II Template Generation System (Finnzymes, Espoo, Finland). Briefly, the isolated genomic library clone Cos-532 was mutagenized by insertion of the entranceposon encoding for kanamycin resistance and then electroporated into *E. coli*. Resulting transformants were pooled and their cosmids being conjugated into *P. putida* (plscB). 100 transconjugants were patched onto MG agar supplemented with 5% sucrose, incubated at 18°C for 10 days, and screened for lack of levan formation. One Tn-mutagenized cosmid Cos-532::Tn did not lead to levan formation in the respective transformant. The cosmid was isolated from this colony and further subjected to DNA extraction and endonuclease digestion. The resulted digestion pattern was compared to the respectively treated DNA of Cos-532. Cos-532::Tn was sequenced using primers SeqW and SeqE (**Table 2**) to obtain nucleotide sequence information for the entranceposon insertion site. The entranceposon-carrying a 267-bp ORF was termed *lscR*. From Cos-532, a 1,350-bp DNA fragment containing *lscR* was removed using endonuclease *Sac*I and cloned into the *Sac*I-treated broad host range vector pBBR1MCS-3 (Kovach *et al.*, 1995b) in opposite orientation to the vector-borne P_{lac} promoter yielding plasmid p1350. Subsequently, this plasmid was conjugated into *P. putida* (plscB). The resulting transconjugant, *P. putida* (plscB, p1350), was grown on sucrose-containing MG agar to verify its levan positive phenotype. Its cellular protein extract was further subjected to Western blot analysis using polyclonal antibodies against Lsc.

Generation of a lscR-deficient mutant in P. syringae

The *P. syringae* mutant *lscR*::Km was generated by insertion of a kanamycin resistant cassette in the *lscR* gene using the broad-host-range Flp-FRT recombination system (Hoang *et al.*, 1998). Briefly, two ~500-bp fragments flanking the *lscR* gene were amplified from PG4180 genomic DNA using two pairs of primers: PSPPH_0650_R (*Hind*III)/PSPPH_0651_F (*Hind*III-*Kpn*I) and PSPPH_0651_R (*Hind*III)/PSPPH_0652_F (**Table 2**). PCR products were cloned into pGEM-T

Easy (Promega, Mannheim, Germany), yielding plasmids pGEM.lscR-down and pGEM.lscR-Up respectively (**Table 1**). All inserts were verified by nucleotide sequencing. A 1,230-bp *Kpn*I fragment containing a Km^r cassette flanked with Flp-*FRT* sites was removed from plasmid pFKm (Choi *et al.*, 2008) and ligated into *Kpn*I-digested pGEM.lscR-down, yielding pGEM.lscR-down-Km. An 1,828-bp *Hind*III fragment digested from pGEM.lscR-down-Km was ligated into *Hind*III digested pGEM.lscR-Up, yielding plasmid pGEM.lscR-Km. Finally, a 2,400-bp *Eco*RI fragment was removed from pGEM.lscR-Km and ligated into *Eco*RI-digested plasmid pEX18Tc (Hoang *et al.*, 1998), yielding the *lscR* gene replacement plasmid pEX18TC.lscR-Km. This plasmid was mobilized into *P. syringae* by tri-parental mating using helper plasmid pRK2013 (Figurski & Helinski, 1979). To counter-select against mutations from single crossovers, single colonies were picked on KB plates containing 5% sucrose. Levan-deficient mutants were screened on MG medium agar plates supplemented with kanamycin and were subsequently confirmed for the genotype by PCR using primers lscR_F1/lscR_R1, and PSPPH_0650_1436F/PSPPH_0653_213R (**Table 2**). Amplified PCR products were verified by sequencing.

Complementation of lscR-deficient mutant in P. syringae

From plasmid pBluescript.lscR, a 552-bp *Sac*I DNA fragment containing *lscR* with its native promoter was cloned into the *Sac*I-treated vector pBBR1MCS-3 generating plscR. Subsequently, *P. syringae* mutant cells were transformed with plscR using electroporation and selected on MG plates containing tetracycline at 28°C. Similar experiment was performed using empty pBBR1MCS-3 plasmid as a negative control. Transformants were then subjected to plasmid extraction (Kado & Liu, 1981) followed by restriction digest analysis. Moreover, they were grown on MG agar plates supplemented with 5% sucrose and monitored for levan formation at 18°C.

Immunological and enzymatic detection of levansucrase

Extracellular proteins in the 50-fold concentrated cell-free supernatant and of total cell lysates from *P. syringae* and its *lscR*-deficient mutant were obtained as described previously (Li & Ullrich, 2001). For immunological detection of Lsc, total proteins were separated on a 12.5% SDS-PAGE and the resulting protein fractions were transferred onto a Hybond-PVDF membrane (Amersham, Buckinghamshire, England) for Western blot experiments using polyclonal antibodies raised against purified Lsc as reported earlier (Smirnova *et al.*, 2001). Zymographic detection of Lsc was

done as described earlier (Li *et al.*, 2006). Briefly, total proteins were separated on a 12.5% native polyacrylamide gel electrophoresis (PAGE) under non-denaturing conditions. Following electrophoresis, gels were incubated in sterile water containing 5% sucrose (Li *et al.*, 2006). All experiments were conducted in triplicates.

Purification of LscR

Fusion protein purification was conducted via maltose-binding protein tagging of the over-expressed proteins with vector pMal-c2 and amylose resin beads (pMAL™ Protein Fusion & Purification System, New England Biolabs, Massachusetts, USA) following the manufacturer's instructions. Briefly, a 267-bp fragment carrying the *lscR* gene was amplified from PG4180 genomic DNA using primers lscR_F (*Sal*I)/lscR_R (*Pst*I) (**Table 2**). PCR product was cloned into pMal (New England Biolabs), yielding plasmid pMal-LscR. *E. coli* cultures harboring pMal or pMal-LscR constructs, respectively, were grown at 37°C in LB medium with 0.2% glucose until they reached an OD_{600} of 0.5. The cultures were then induced with 0.3 mM IPTG and grown for 4 hours at 37°C followed by overnight growth at 18°C. The cells were harvested by centrifugation at 4,000 x g for 20 minutes and resuspended in chilled sonication buffer (20 mM HEPES, 200 mM NaCl, 0.2 mM DTT, 1 mM EDTA, 10% glycerol). Cell suspensions were sonicated on ice (10 x 1 sec, 4 sec break, 3 repeats; amplitude 70%; Active Motif® sonicator, Regensburg, Germany) and cell debris was centrifuged off afterwards (30 min, 4,000 x g, 4°C). Supernatants were stored at -20°C. 2 mL of amylose resin (New England Biolabs) were added to 2.5 x 10 cm poly-prep chromatography columns (Biorad, München, Germany). Affinity chromatography was carried out following manufacturer's instruction (New England Biolabs). Protein content was determined via bicinchoninic acid test (Thermo Fisher Scientific) according to manufacturer's instruction. The purity of eluted samples was analyzed by 12.5% SDS-PAGE followed by Coomassie Brilliant Blue staining.

Electrophoretic mobility shift assay

Bait DNA fragments used for the electrophoretic mobility shift assay (EMSA) were PCR amplified using Cy5-labeled primers to perform a non-radioactive EMSA. The *lscB* upstream region (500 bp) was PCR-amplified using primers lscB_rv_Cy5/lscB_PG-500_fwd while the coding region of *lscB* (458 bp) used as a control was PCR-amplified using primers lscB_458R/lscB_F_Cy5 (**Table**

2). Approximately 0.16 pmol of Cy5-labeled DNA was mixed with increasing concentrations of MBP or MBP-tagged LscR protein, respectively, in a binding buffer reaction (50 mM Tris–HCl, pH 7.5; 1 mM DTT; 500 mM MgCl2; 100 mM EDTA; 10 mM NaCl; 5% glycerol) (Pletzer & Weingart, 2014). MBP was used as a control to rule out any interference of the MBP fusion to DNA binding. Incubation was done at room temperature for 30 min. The total reaction was run on a 1% agarose gel electrophoresis in 1x Tris-acetate-EDTA (TAE) buffer. After electrophoresis, fluorescence signals of the labeled DNA were visualized using a FLA-3000 phosphoimager (Raytest, Straubenhardt, Germany).

Analysis of lsc and glycosyl hydrolase gene expression by quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR)

Cell cultures of *P. syringae* and its *lscR*-deficient mutant were grown in HSA medium at 18°C and harvested during early logarithmic growth at an OD_{600} of 0.5. Total RNA was isolated by acid phenol/chloroform extraction (Mansfield, 2009). Yield and purity of RNA were checked by gel electrophoresis and determination of the A260/A280 and A260/A230 ratios using a Nanodrop ND-2000 spectrophotometer (Thermo Fischer Scientific). Total RNA samples were treated with TURBO DNA-free (Ambion life technologies, Carlsbad, USA) to remove remaining traces of genomic DNA as described by the manufacturer's recommendation. High quality RNA was reverse transcribed and amplified with the QuantiTect SYBR Green One-step RT-PCR Kit (Qiagen). Template RNA (5 ng) was used in a standard 25µl qRT-PCR reaction with the following primers for *lsc* (lsc_qRT_F/lsc_qRT_R) and *glycosyl hydrolase* gene (PSPPH_0655_qRT_F2/PSPPH_0655_qRT_R2). 23srRNA-qRT-F and 23srRNA-qRT-R primers were used for the reference gene (**Table 2**). For analysis, a Mastercycler ep realplex2 gradient S instrument (Eppendorf) was used. Cycling parameters included a 15 min initial denaturation at 95°C to activate the DNA polymerase followed by 40 cycles consisting of 15 s at 95°C, 30 s at 55°C and 30 s at 72°C and a final extension for 15 min at 72°C. The final step consisted of 1 min at 95 °C and 30 s at 55 °C. A melting curve analysis with a temperature ramp from 55°C to 95°C in 20 min was performed at the end of each run to determine specificity of amplified qPCR products. Reactions were performed in technical and biological triplicates. Reactions with no addition of reverse transcriptase served as negative controls and proved lack of DNA contamination. Due to very high identity at the nucleotide sequence (~98%) between *lscB* and *lscC*,

it was not possible to design primers discriminating between these two mRNAs, hence expression profile of *lsc* is always referred as a combination of both genes. Quantification of mRNA transcripts was performed by the comparative C_t method described by Pletzer and Weingart (2014).

Statistical analysis

All assays were performed in three biological triplicates. Data were expressed as mean ± standard deviation (SD). Statistical analysis was performed using R (R Development Core Team, 2011). Levels of significance were calculated by One-Way ANOVA and $p < 0.005$ was considered statistically significant.

Table 2: Oligonucleotide primers used in this study

Primers	**Nucleotide sequence (5'-3') ***
PSPPH_0650_R (*Hind*III)	GACAAGCTTCGCAGCAGTGTGCCTGTG
PSPPH_0651_F (*Hind*III-*Kpn*I)	GACAAGCTTGGTACCCACCCCTTCGTGTAC
PSPPH_0651_R (*Hind*III)	GACAAGCTTGCGAAGTTGATCCAG
PSPPH_0652_F	CGGCATGCCACGTCCAAG
lscR_F1	GCAAACCACAGGGAGTCATGG
lscR_R1	TCACGATGGCGCTCAGGTAT
PSPPH_0650_1436F	CCCACATCTCACCGTCGGTA
PSPPH_0653_213R	GCCAGAGCATTCCTGAAGGC
lscB_PG-500_fwd	GATGAGCTCAGTCGCAATTAATGCGAG
lscB_rv_Cy5	AATTGATACCTTTAAATAGCTTTGGGAG
lscB_458R	ATTCAACAGTATGGGTGTGCCC
lscB_F_Cy5	CTAGCAGCTCTGCTGTAAGCCAG
lscR_F (*Sal*I)	CAGATGCTGGTCGACATGCAAACCACAGGGAG
lscR_R (*Pst*I)	CAGATGCTGCTGCAGTCAGGCTGCCCTGTTG
PSPPH_0655_qRT_F2	CCAATAGCAGTTCTGCTGTCAGC
PSPPH_0655_qRT_R2	CGATTGGGCCACTGATAATGAG
lsc_qRT_F	TCGGTTATCCTGACCCTGAC
lsc_qRT_R	CCATGACGATCTTCCCAGTC
23srRNA-qRT-F	GTTGGGTGACTGCGTACCTT
23srRNA-qRT-R	ACTCCGATCCACAGGTCATC
SeqW	GGTGGCTGGAGTTAGACATC
SeqE	CGACACACTCCAATCTTTCC

*Restriction sites in the primers are underlined: AAGCTT–*Hind*III, GGTACC–*Kpn*I, CTGCAG–*Pst*I, GTCGAC–*Sal*I.

Results

Heterologous screening for transcriptional activator(s) of lscB expression

In order to find possible genes allowing for the expression of *lsc*, we screened a genomic cosmid library of *P. syringae* in the heterologous host, *P. putida* strain KT2440, which lacks *lsc* genes in its genome and does not form levan when grown on a sucrose-containing medium. Consequently, *lscB* from *P. syringae* along with a 680-bp upstream sequence was introduced to *P. putida*. The respective transconjugant, termed *P. putida* (plscB), exhibited a levan-negative phenotype on sucrose-containing agar (**Fig. 1AB**). A pooled genomic cosmid library of *P. syringae* was conjugated into *P. putida* (plscB), and transformants grown on sucrose-containing agar plates. Three individual transconjugant colonies termed TCos-245, TCos-532, and TCos-671, showed a levan-positive phenotype, i.e. a dome-shaped colony morphology and translucent slime formation, indicating expression of *lscB* (Supporting Information **Fig. S1A**). Subsequently, cosmids Cos-245, Cos-532, and Cos-671, were isolated from those three transformants. In order to rule out that the levan-positive phenotype was derived from the cosmids itself, we re-introduced them to the wild type of *P. putida* to test the new transformants' phenotype. None of the cosmid clones induced levan formation in *P. putida* wild type verifying that the cosmids did not carry any *lsc* genes or other polysaccharide synthesis-encoding gene(s) (**Fig. 1C**). In contrast, when re-introduced into *P. putida* (plscB), typical levan formation was observed (**Fig. 1DE**) and presence of Lsc was confirmed by Western blotting (data not shown). The corresponding cosmid clones were isolated and treated with restriction endonucleases to determine their genetic similarities. The cosmid inserts exhibited several common restriction fragments although they were not identical with each other suggesting that they contained partially overlapping genomic DNA regions (Supporting Information **Fig. S1B**). One of these cosmids, Cos-532, was selected for further analysis.

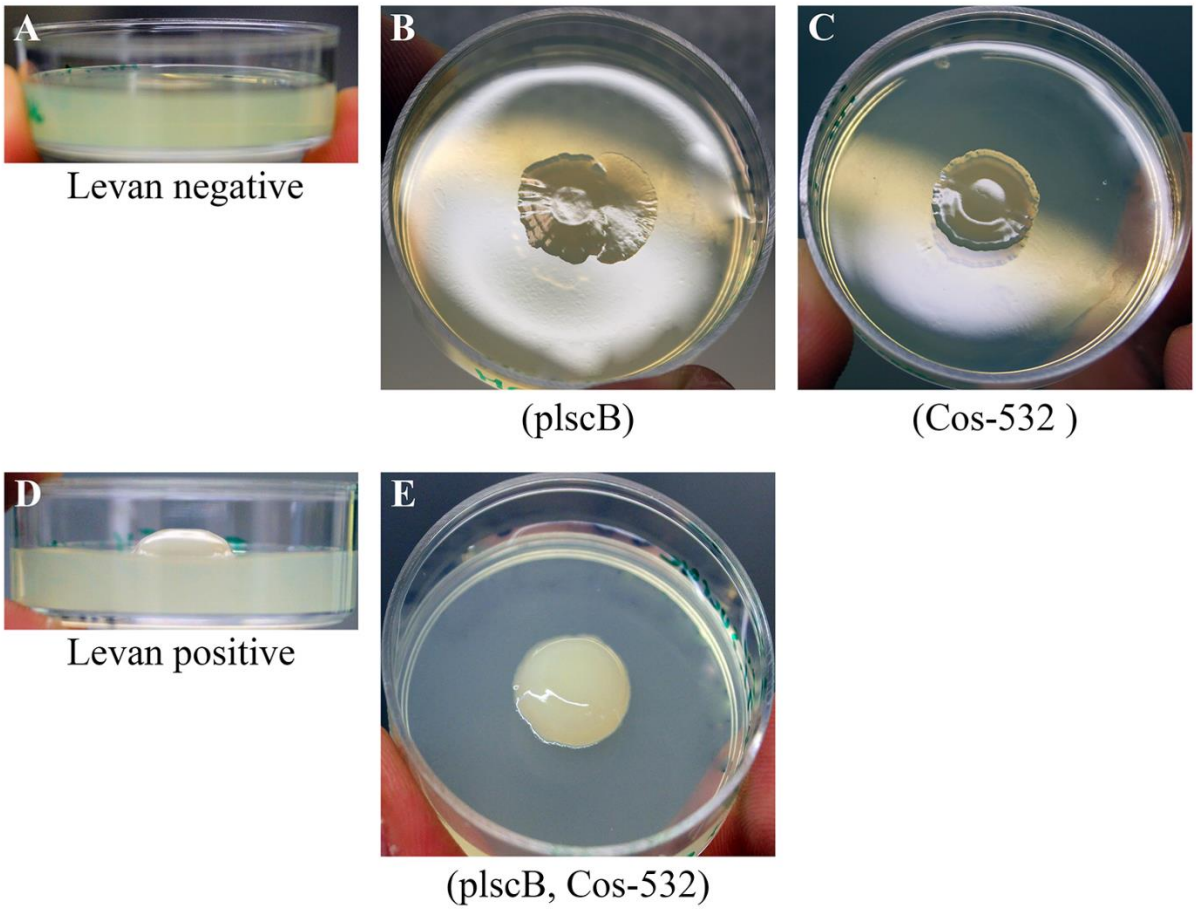


Figure 1: Levan phenotype observed in *P. putida* transformants grown on MG agar plates supplemented with 5% sucrose at 18°C. **A)** Side view of agar plate that shows a levan negative phenotype. **B)** *P. putida* carrying plasmid plscB and showing a levan negative phenotype. **C)** *P. putida* carrying cosmid Cos-532 and showing a levan negative phenotype. **D)** Side view of agar plate that show a levan positive phenotype characterized by presence of a dome-shaped structure. **E)** *P. putida* carrying both plscB and cosmid Cos-532 and showing a levan positive phenotype.

Identification of the transcriptional activator lscR

Cos-532 was subjected to an *in vitro* transposon mutagenesis to find the gene(s) responsible for the *lsc* expression phenotype. The randomly mutagenized Cos-532 pool was introduced into *P. putida* (plscB) and transconjugants were screened for deficiency in levan formation. We found that one mutagenized cosmid, Cos-532::Tn, showed a levan-negative phenotype (**Fig. 2A**). This mutant cosmid was isolated from the respective transformant. The resulting digestion pattern of Cos-532::Tn was compared with that of the respectively treated DNA of Cos-532. We found a difference showing a possible insertion of the transposon in one of the DNA fragments (**Fig. 2B**). Sequencing results revealed that the transposon of Cos-532::Tn was inserted to a 267-bp ORF, which was found to be 100% identical in its nucleotide sequence to its homolog from *P. syingae* pv. phaseolicola 1448A and previously annotated as a prophage-borne transcriptional activator PSPPH_0651 (Winsor *et al.*, 2011) (**Fig. 2D**). Due to its potential function, we propose to designate this ORF as *lscR* for levansucrase regulator. A cloned 1,350-bp *Sac*I fragment of Cos-532 termed p1350, carrying two ORFs homologous to the *P. syingae* 1448A genes PSPPH_0651 [*lscR*] and PSPPH_0652, was sufficient to induce levan formation in *P. putida* (plscB) when grown on 5% sucrose (data not shown). This observation was also confirmed by Western blot analysis using cell lysates from *P. putida* (plscB, p1350), *P. putida* (plscB), and *P. putida* (p1350), were a signal was observed only when both *lscR* and *lscB* were present (**Fig. 2C**). Moreover, a smaller sub-clone of p1350 termed plscR, which consists of 379-bp and carrying a homolog of *P. syingae* 1448A's PSPPH 0651 in opposite orientation to the vector-borne P_{lac} promoter, also rendered *P. putida* (plscB) levan-positive. However, a 417-bp sub-clone of p1350 termed p0652, carrying only the homolog from *P. syingae* 1448A annotated as PSPPH 0652, was found to be levan-negative in *P. putida* (plscB) (data not shown) demonstrating that *lscR* was the only gene responsible for the heterologous expression of *lscB*.

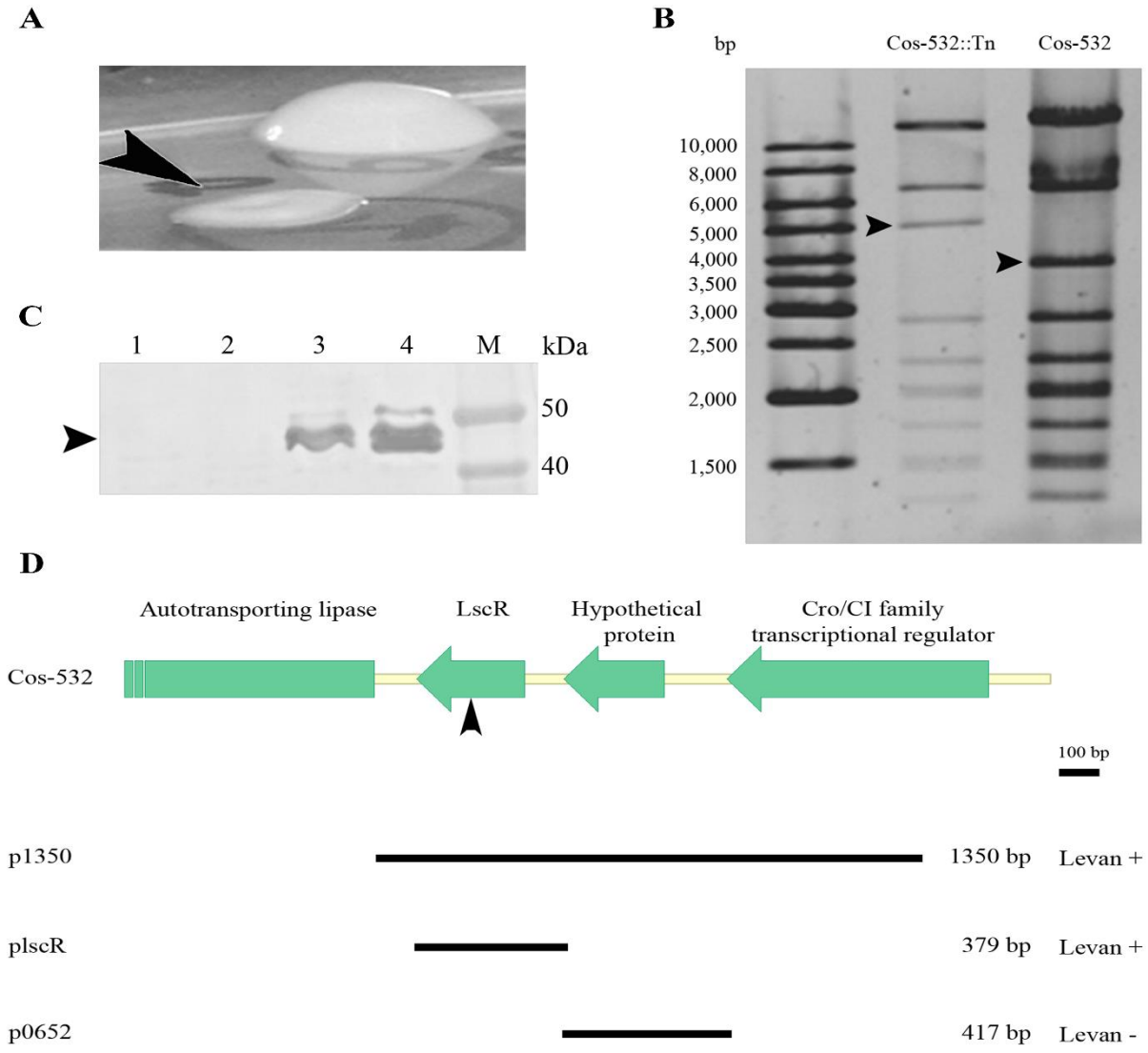


Figure 2: Identification of lscR on cosmid Cos-532 by in vitro transposon mutagenesis. **A)** Growth of *P. putida* (plscB) transconjugants harboring Cos-532 on sucrose-containing MG agar after in vitro mutagenesis. The arrow indicates the levan-minus transconjugant carrying mutagenized cosmid, Cos-532::Tn. **B)** Electrophoretic comparison of cosmids, Cos-532 and Cos-532::Tn, after treatment with endonuclease *Pst*I. The arrows indicate the fragment containing the entranceposon insertion in Cos-532::Tn. **C)** Western blot analysis of cell lysates using LscB-specific antibodies. Cells were grown to an OD_{600} of 1.6 at 18°C and subjected to total protein extraction. Lanes: (1) *P. putida* WT, (2) *P. putida* (plscB), (3) *P. putida* (plscB, p1350), (4) *P. syringae* PG4180.WT. **D)** Genetic map of the lscR harboring DNA region common to the three identified cosmids. Sub-cloned fragments and their impact on Lsc expression *P. putida* (plscB) are indicated for p1350, pLscR and p0652. The arrow marks the entraceposon insertion in Cos-532::Tn.

Phenotypic characterization of lscR-deficient mutant of P. syringae

Our results demonstrated an essential role of *lscR* in levansucrase expression in the heterologous host *P. putida* KT2440. In order to investigate this regulatory gene's function in *P. syringae*, an *lscR*-deficient mutant was generated and genetically verified. Since sucrose leads to a levan-positive growth phenotype in *P. syringae*, the wild type and its *lscR*-deficient mutant were grown on MG agar plates supplemented with this disaccharide. We found that the *lscR*-deficient mutant exhibited a clear levan-deficient phenotype (**Fig. 3A**) suggesting that an activator of levansucrase expression in *P. syringae* had been knocked-out in this mutant. Complementing the mutant with plasmid plscR restored the levan-positive phenotype (Supporting Informations **Fig. S2**).

Detection of Lsc in cell lysate and supernatant of P. syringae and its lscR-deficient mutant

In order to investigate the impact of the *lscR* gene product on levan formation, zymograms and Western blot experiments were applied to protein extracts from total cell lysates and cell-free supernatants of wild type *P. syringae* and its *lscR*-deficient mutant. Protein bands representing Lsc were detected in zymograms by a whitish swelling of the gel matrix that corresponded to levan formation. Levan formation was only observed in cell lysate and supernatant of *P. syringae* wild type but not in the corresponding samples of the *lscR* mutant (**Fig. 3B**). To differentiate between enzymatic activity and secretion of non-active Lsc, Western blot experiments were performed on cell lysates and cell-free supernatants. Results obtained were similar to those observed in zymographic analysis. Signals specific for Lsc (~47.6 kDa) could only be detected in *P. syringae* wild type samples but not in *P. syringae lscR*::Km (**Fig. 3C**).

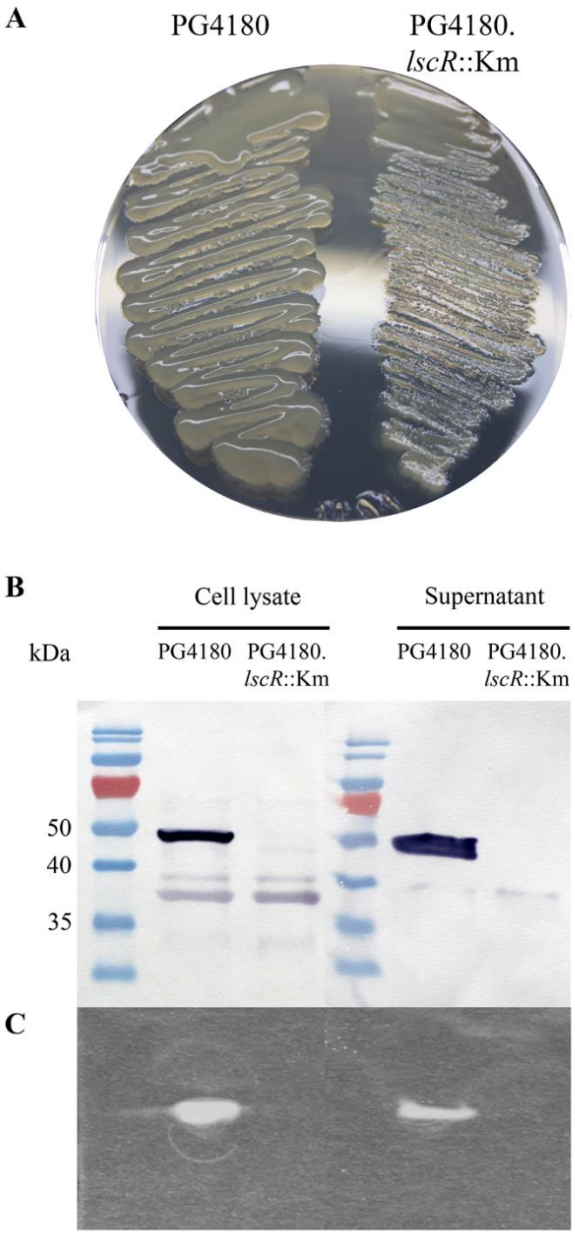


Figure 3: Phenotypic characterization of *P. syringae* and its *lscR*-deficient mutant. **A)** Levan formation by *P. syringae* and its *lscR*-deficient mutant. Bacteria were streaked on MG agar plates containing 5% sucrose and incubated at 18°C for 10 days. Unlike *lscR*-deficient mutant, *P. syringae* wild type clearly show levan formation characterized by a white dome-shaped colony morphology. **B)** Detection of levansucrase by western blot analysis. The dark bands (~47.6 kDa) resembles levansucrase. **C)** Zymogram with protein samples from total cell lysate and cell free supernatant were separated on a 12.5% native-PAGE and incubated in sterile water containing 5% sucrose. White bands indicate levan formation as a result of sucrose utilization by levansucrase.

Binding of LscR to the upstream sequence of lscB/C

The PAPE sequence upstream of *lscB/C* was previously shown to be essential for levansucrase expression in *P. syringae* (Srivastava *et al.*, 2012, Khandekar *et al.*, 2014). Thus, it could represent a potential target sequence for LscR binding leading to gene expression. To confirm this binding *in vitro*, an electrophoretic mobility shift assay (EMSA) was performed. DNA fragments used in the EMSA were the upstream region of *lscB/C* (500 bp) and the coding region of *lscB* (458 bp) as a control. The DNA fragments were incubated with increasing amounts of purified proteins. No interactions were detected between MBP and the upstream region of *lscB* (**Fig. 4A**) or between MBP-LscR and the coding region of *lscB/C* (**Fig. 4C**). Purified MBP-LscR showed binding to the upstream sequence of *lscB/C* with increasing concentrations observable as a smear signal (**Fig. 4B**). These results suggested that the LscR portion of MBP-LscR but not MBP itself bound to the PAPE upstream region of *lscB*. The results furthermore demonstrated that this binding was specific since MBP-LscR did not show binding to the coding sequence of *lscB*.

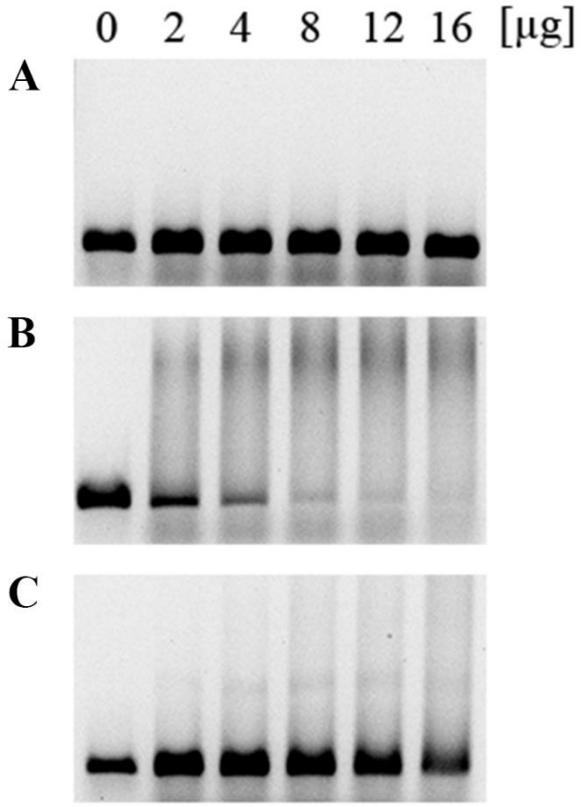


Figure 4: Electrophoretic mobility shift analysis of LscR interaction with cy5-labeled DNA fragments. **A)** Purified unfused MBP with 500 bp upstream region of *lscB*. **B)** Purified MBP-LscR with 500 bp upstream region of *lscB*. **C)** Purified MBP-LscR with 458 bp coding region of *lscB*. Approximately 100 ng of DNA were incubated with increasing amounts of proteins (indicated on top of the lanes). The DNA-protein complexes were separated on 1% agarose gels.

Transcriptional analysis of lsc and glycosyl hydrolase gene in P. syringae and its lscR-deficient mutant

Interestingly, the gene encoding for a glycosyl hydrolase in *P. syringae* was found to harbor a PAPE sequence in its upstream region, which is very similar to the one present upstream of *lscB* and *lscC* (Srivastava *et al.*, 2012). Excitingly, this gene is located in close proximity to the gene coding for LscR within the bacteriophage-derived sequence identified in Cos-532 (**Table 1**). In order to analyze the effect of the mutation of *lscR* on the expression level of *lsc* and the *glycosyl hydrolase* gene, transcriptional analyses were conducted using qRT-PCR. As shown in **Fig. 5**, the *lsc* and glycosyl hydrolase genes were down regulated 11- and 53-folds, respectively, in the *lscR*-deficient mutant indicating a similar or common type of regulation potentially dictated by LscR. Furthermore, our results showed that *lsc* genes were ~15-fold more expressed as compared to the glycosyl hydrolase gene in *P. syringae* (data not shown) suggesting either that the stability of these transcripts seem to vary significantly or that additional regulatory factors contributed to diverging levels of expression.

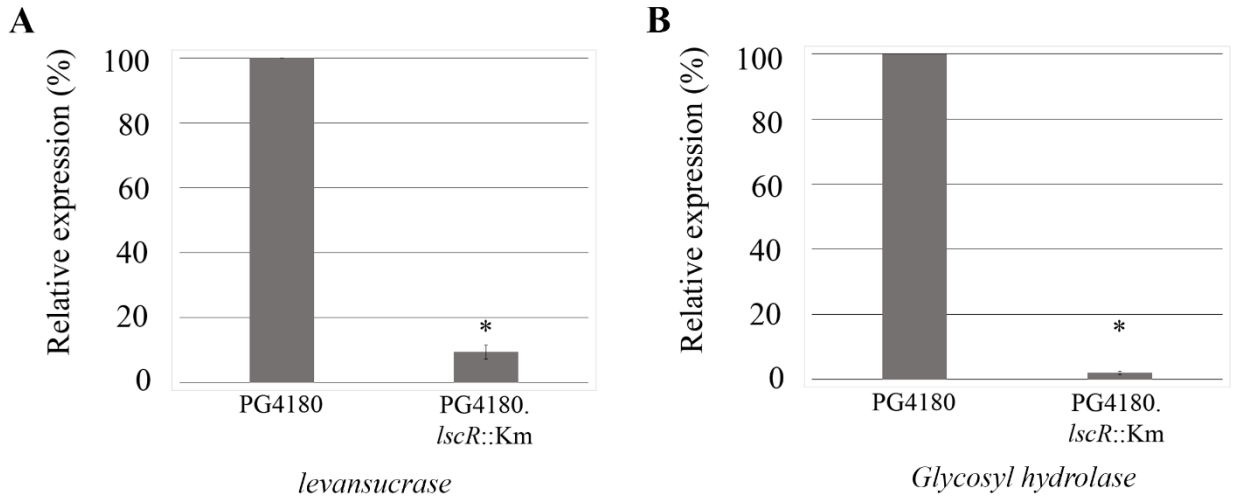


Figure 5: Quantitative Reverse Transcriptase PCR analysis of **A)** *levansucrase* and **B)** *glycosyl hydrolase* gene expression in *P. syringae* and PG4180.*lscR*::Km. Cells were grown at 18°C in HS + arabinose. Relative mRNA levels were related to the mean value determined for the signals of *P. syringae* wild type at an OD_{600} of 0.5, which was defined as 100%. Data shows the mean values and error bars represent standard deviation of the mean of three biological replicates (n=3). Statistically significant differences ($p < 0.005$) were calculated by one-way ANOVA-analysis and are marked with asterisks.

Discussion

Environmental conditions during the epiphytic growth phase of *Pseudomonas syringae* play a crucial role for the infection cycle as they lead to an abundant population needed for the invasion into the leaf tissue (Beattie & Lindow, 1995). However, leaf surfaces often represent stressful environments to bacteria due to nutrient limitation, fluctuation in temperature and water availability, or UV irradiation (Monier & Lindow, 2003). Similar to other EPS, levan may mask and protect plant-pathogenic bacteria during early stages of infection (Kasapis *et al.*, 1994). In addition, it can act as a barrier against plant defense compounds (Király *et al.*, 1997) or as a source of carbon compounds which are often limiting factors for epiphytic bacteria (Mercier & Lindow, 2000). Activation of *lsc* gene expression at 18°C indicates a strong link with the global low-temperature induction of pathogenicity and virulence in *P. syringae* (Dunleavy, 1988, Smirnova *et al.*, 2001). Therefore, formation of Lsc is likely to be regulated at the transcriptional level to avoid energy-demanding gene product synthesis under unfavorable conditions. However, previously identified regulatory systems required for the expression of other cold-induced genes in *P. syringae* did not show any influence on *lsc* expression. These systems included the two-component system CorRP/S (Ullrich *et al.*, 1995) and the alternative sigma factor, AlgT (Schenk *et al.*, 2008). Moreover, no influence on levansucrase regulation was found for the two-component system, GacS/GacA (unpublished data), which controls the synthesis of secondary metabolites and extracellular enzymes and is involved in pathogenicity, tolerance to stress, and ecological fitness in several *Pseudomonas* species including *P. syringae* (Gaffney *et al.*, 1994, Altier *et al.*, 2000). Our recent study by Mehmood *et al.* (2015) suggested that the hexose metabolism repressor, HexR, co-regulates expression of *lsc* with genes involved in the central hexose metabolism. However, the *P. syringae hexR* mutant only showed slight upregulation of *lsc* genes (Mehmood *et al.*, 2015) indicating that this type of regulation is rather of peripheral impact on levan formation.

For the first time, we herein provide evidence for a transcriptional activator, LscR, which plays an instrumental role in regulation of levan formation in the plant pathogen *P. syringae*. Screening for transcriptional activators of *lsc* expression was performed in the heterologous host *P. putida* using a genomic library of *P. syringae*. The co-presence of the *lscB* gene with cosmid Cos-532 resulted in a levan-positive phenotype suggesting that this cosmid might potentially encode for a transcriptional activator. *In vitro* transposon mutagenesis revealed that such an activator could be

encoded by an ORF which has a homolog in *P. syringae* 1448A annotated as a prophage-borne transcriptional activator PSPPH_0651 (Winsor *et al.*, 2011). As a consequence, the ORF PSPPH_0651 homolog in the genome of strain PG4180 was designated *lscR*. Interestingly, *lscR* is conserved only in genomes of *P. syringae* species but absent in any other organism's genome including any other *Pseudomonas* species. Furthermore, it is noteworthy that *P. syringae* is the only known bacterial species which possesses multiple highly homologous *lsc* genes some of which have nearly identical upstream sequences. The herein reported discovery of LscR as a transcriptional activator for both homologs and its phylogenetic restriction to one species with multiple pathovars sheds light on the evolutionary distribution of levan formation in *P. syringae*.

Previous studies had demonstrated that the upstream PAPE region of *lscB* and *lscC* in *P. syringae* was highly conserved (Srivastava *et al.*, 2012, Khandekar *et al.*, 2014). Interestingly, this region was also found to be present upstream of a gene encoding for a glycosyl hydrolase in *P. syringae* (Srivastava *et al.*, 2012), which is located in close proximity to the *lscR* gene within a prophage-like DNA region as found in the current study. We hypothesized that regulation of the *glycosyl hydrolase* gene is also under control of LscR. To test this, the expression levels of *lsc* and *glycosyl hydrolase* genes in both *P. syringae* wild type and its *lscR*-deficient mutant were analyzed and turned out to be strictly LscR-dependent, thus confirming that LscR is responsible for the activation of expression of *lsc* and *glycosyl hydrolase* genes. In the *P. syringae* wild type, the expression of the *glycosyl hydrolase* gene was shown to be lower than that of *lsc*. A similar pattern of expression was observed earlier by Srivastava *et al.* (2012). Accordingly, it became of particular interest to determine whether LscR controls expression of levansucrase by direct binding to its upstream PAPE sequence. The results of the DNA-binding assay conducted herein showed that LscR is able to bind specifically to *lsc* upstream PAPE indicating that the PAPE sequence carries a site of transcription regulation where LscR might bind and subsequently induce gene expression. Additional investigation of the exact binding site of LscR will need to be conducted to provide further understanding of this regulation and whether other genes are also under control of LscR.

Nucleotide sequencing of the *lscR*-carrying insert of cosmid Cos-532 revealed that *lscR* is located three ORFs downstream of the *glycosyl hydrolase* gene harboring the PAPE upstream sequence. Interestingly, both *lscR* and the PAPE sequence were found to be located in a region characterized by bacteriophage-associated genes (data not shown). Moreover, information from the

www.pseudomonas.com website showed similar results for *lscR* and PAPE homologs in *P. syringae* strains B728a and 1448A, suggesting a prophage insertion in this region. Bacteriophages are very abundant (Ceyssens & Lavigne, 2010), are considered reservoirs for genetic diversity, and can integrate their DNA into their bacterial hosts (Suttle, 2005) with their insertion location varying but sometimes bound to specific integration sites (Casjens, 2003). While replication of prophage DNA might reduce the fitness of its host due to its high metabolic cost (Canchaya *et al.*, 2003), they often also supply their hosts with important agents that are needed to survive in specific ecological niches (Srividhya *et al.*, 2007). In fact, bacterial pathogens encode several virulence factors, such as toxins, pili (fimbriae), adhesins and secretion systems, which are of prophage origin (Wagner & Waldor, 2002). A recent study by Hockett *et al.* (2015) revealed that *P. syringae* carries a genomic region between *trpE* and *trpG* that contains a fragment of a bacteriophage origin. This fragment, referred to as R-type syringacin, was shown to encode only for a tail of a bacteriophage responsible for killing activity in *P. syringae* B728a. Interestingly, LscR and the PAPE sequence were found to be located in the same prophage region encoding for R-type syringacin. To determine whether LscR plays any role in syringacin production, we tested *P. syringae* PG4180 and its *lscR*-deficient mutant for production of R-type syringacin. Results showed no bacteriocin activity against a panel of *P. syringae* strains that have diverse sensitivity spectra (Kevin Hockett - personal communication) suggesting that LscR is not involved in bacteriocin synthesis or regulation. Although this region was conserved among other *P. syringae* strains, essential genes needed for producing an active R-type syringacin were found missing in *P. syringae* pv. phaseolicola 1448A (Hockett *et al.*, 2015) and *P. syringae* pv. glycinea PG4180 (current study) thus supporting our hypothesis that LscR is not involved in syringacin synthesis. Instead, the results of the current study imply that a phage-borne regulatory protein, LscR, got evolutionary involved in the expression of the bacterial fitness genes, *lscB* and *lscC*, via the genetic recombination of a 500-bp PAPE sequence upstream of these target genes. To our knowledge this is the first report for the employment of a phage-borne regulator in the expression of levansucrase genes in any bacterial species. Since the Lsc-substrate, sucrose, is an abundant transport sugar of host plants, our results demonstrated that bacteriophages can shape the environmental potential of plant pathogenic bacteria not only via supplying additional genes but also by supplying associated regulatory DNA elements therefore efficiently altering the bacterial fitness during the infection cycle.

Acknowledgments

The authors would like to thank Kevin L. Hockett and David A. Baltrus for their help in testing the bacteriocin activity and sensitivity of our bacterial strain and its mutant. This work was supported by a doctoral stipend from the Deutscher Akademischer Austauschdienst. The authors declare no conflict of interest for this study.

References

Altier, C., M. Suyemoto & S.D. Lawhon, (2000) Regulation of *Salmonella enterica* serovar typhimurium invasion genes by csrA. *Infect Immun* **68**: 6790-6797.

Beattie, G.A. & S.E. Lindow, (1995) The secret life of foliar bacterial pathogens on leaves. *Annual review of phytopathology* **33**: 145-172.

Bender, C.L., H. Liyanage, D. Palmer, M. Ullrich, S. Young & R. Mitchell, (1993) Characterization of the Genes-Controlling the Biosynthesis of the Polyketide Phytotoxin Coronatine Including Conjugation between Coronafacic and Coronamic Acid. *Gene* **133**: 31-38.

Bereswill, S. & K. Geider, (1997) Characterization of the rcsB gene from *Erwinia amylovora* and its influence on exoploysaccharide synthesis and virulence of the fire blight pathogen. *J Bacteriol* **179**: 1354-1361.

Canchaya, C., C. Proux, G. Fournous, A. Bruttin & H. Brussow, (2003) Prophage genomics. *Microbiology and molecular biology reviews : MMBR* **67**: 238-276, table of contents.

Casjens, S., (2003) Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol* **49**: 277-300.

Ceyssens, P.J. & R. Lavigne, (2010) Bacteriophages of *Pseudomonas*. *Future Microbiol* **5**: 1041-1055.

Choi, K.H., T. Mima, Y. Casart, D. Rholl, A. Kumar, I.R. Beacham & H.P. Schweizer, (2008) Genetic tools for select-agent-compliant manipulation of *Burkholderia pseudomallei*. *Appl Environ Microbiol* **74**: 1064-1075.

Denny, T.P., (1995) Involvement of bacterial polysaccharides in plant pathogenesis. *Annual review of phytopathology* **33**: 173-197.

Dunleavy, J., (1988) Bacterial, fungal, and viral diseases affecting soybean leaves In: Soybean Diseases of the North Central Region. T.D. Wyllie, Scott, D. H. (ed). American Phytopathological Society, pp. 149

Figurski, D.H. & D.R. Helinski, (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proceedings of the National Academy of Sciences of the United States of America* **76**: 1648-1652.

Gaffney, T.D., S.T. Lam, J. Ligon, K. Gates, A. Frazelle, J. Di Maio, S. Hill, S. Goodwin, N. Torkewitz, A.M. Allshouse & et al., (1994) Global regulation of expression of antifungal factors by a *Pseudomonas fluorescens* biological control strain. *Mol Plant Microbe Interact* **7**: 455-463.

Hettwer, U., M. Gross & K. Rudolph, (1995) Purification and characterization of an extracellular levansucrase from *Pseudomonas syringae* pv. phaseolicola. *J Bacteriol* **177**: 2834-2839.

Hettwer, U., F.R. Jaeckel, J. Boch, M. Meyer, K. Rudolph & M.S. Ullrich, (1998) Cloning, nucleotide sequence, and expression in *Escherichia coli* of levansucrase genes from the plant pathogens *Pseudomonas syringae* pv. glycinea and *P. syringae* pv. phaseolicola. *Appl Environ Microbiol* **64**: 3180-3187.

Hoang, T.T., R.R. Karkhoff-Schweizer, A.J. Kutchma & H.P. Schweizer, (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.

Hockett, K.L., T. Renner & D.A. Baltrus, (2015) Independent Co-Option of a Tailed Bacteriophage into a Killing Complex in *Pseudomonas*. *mBio* **6**: e00452.

Huynh, T.V., D. Dahlbeck & B.J. Staskawicz, (1989) Bacterial blight of soybean: regulation of a pathogen gene determining host cultivar specificity. *Science (New York, N.Y.)* **245**: 1374-1377.

Jones, J.D.G. & N. Gutterson, (1987) An Efficient Mobilizable Cosmid Vector, Prk7813, and Its Use in a Rapid Method for Marker Exchange in *Pseudomonas-Fluorescens* Strain Hv37a. *Gene* **61**: 299-306.

Kado, C.I. & S.T. Liu, (1981) Rapid procedure for detection and isolation of large and small plasmids. *J Bacteriol* **145**: 1365-1373.

Kasapis, S., E.R. Morris, M. Gross & K. Rudolph, (1994) Solution properties of levan polysaccharide from *Pseudomonas syringae* pv. phaseolicola, and its possible primary role as a blocker of recognition during pathogenesis. *Carbohydrate Polymers* **23**: 55-64.

Keane P, K.A., New P, (1970) Crown Gall of Stone Fruit II. Identification and Nomenclature of *Agrobacterium* Isolates. *Aust J Biol Sci*.

Khandekar, S., A. Srivastava, D. Pletzer, A. Stahl & M.S. Ullrich, (2014) The conserved upstream region of lscB/C determines expression of different levansucrase genes in plant pathogen *Pseudomonas syringae*. *BMC Microbiol* **14**: 79.

King, E.O., M.K. Ward & D.E. Raney, (1954) Two simple media for the demonstration of pyocyanin and fluorescin. *The Journal of laboratory and clinical medicine* **44**: 301-307.

Király, Z., H.M. El-Zahaby & Z. Klement, (1997) Role of Extracellular Polysaccharide (EPS) Slime of Plant Pathogenic Bacteria in Protecting Cells to Reactive Oxygen Species. *Journal of Phytopathology* **145**: 59-68.

Kovach, M.E., P.H. Elzer, D.S. Hill, G.T. Robertson, M.A. Farris, R.M. Roop, 2nd & K.M. Peterson, (1995a) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.

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

Li, H., A. Schenk, A. Srivastava, D. Zhurina & M.S. Ullrich, (2006) Thermo-responsive expression and differential secretion of the extracellular enzyme levansucrase in the plant pathogenic bacterium *Pseudomonas syringae* pv. glycinea. *FEMS microbiology letters* **265**: 178-185.

Li, H. & M.S. Ullrich, (2001) Characterization and mutational analysis of three allelic lsc genes encoding levansucrase in *Pseudomonas syringae*. *J Bacteriol* **183**: 3282-3292.

Lindeberg, M., C.R. Myers, A. Collmer & D.J. Schneider, (2008) Roadmap to New Virulence Determinants in *Pseudomonas syringae*: Insights from Comparative Genomics and Genome Organization. *Molecular Plant-Microbe Interactions* **21**: 685-700.

Mansfield, J.W., (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: an overview of 25 years of progress in our understanding of plant innate immunity. *Molecular plant pathology* **10**: 721-734.

Martinez-Fleites, C., M. Ortiz-Lombardia, T. Pons, N. Tarbouriech, E.J. Taylor, J.G. Arrieta, L. Hernandez & G.J. Davies, (2005) Crystal structure of levansucrase from the Gram-negative bacterium *Gluconacetobacter diazotrophicus*. *The Biochemical journal* **390**: 19-27.

Mehmood, A., K. Abdallah, S. Khandekar, D. Zhurina, A. Srivastava, N. Al-Karablieh, G. Alfaro-Espinoza, D. Pletzer & M.S. Ullrich, (2015) Expression of extra-cellular levansucrase in *Pseudomonas syringae* is controlled by the in planta fitness-promoting metabolic repressor HexR. *BMC Microbiol* **15**: 48.

Mercier, J. & S.E. Lindow, (2000) Role of leaf surface sugars in colonization of plants by bacterial epiphytes. *Appl Environ Microbiol* **66**: 369-374.

Monier, J.M. & S.E. Lindow, (2003) *Pseudomonas syringae* Responds to the Environment on Leaves by Cell Size Reduction. *Phytopathology* **93**: 1209-1216.

O'Brien, H.E., D. Desveaux & D.S. Guttman, (2011) Next-generation genomics of *Pseudomonas syringae*. *Curr Opin Microbiol* **14**: 24-30.

Osman, S.F., W.F. Fett & M.L. Fishman, (1986) Exopolysaccharides of the phytopathogen *Pseudomonas syringae* pv. glycinea. *J Bacteriol* **166**: 66-71.

Pletzer, D. & H. Weingart, (2014) Characterization of AcrD, a resistance-nodulation-cell division-type multidrug efflux pump from the fire blight pathogen *Erwinia amylovora*. *BMC Microbiol* **14**: 13.

Ramos-Diaz, M.A. & J.L. Ramos, (1998) Combined physical and genetic map of the *Pseudomonas putida* KT2440 chromosome. *J Bacteriol* **180**: 6352-6363.

Sambrook, J. & D. Russell, (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Schenk, A., M. Berger, L.M. Keith, C.L. Bender, G. Muskhelishvili & M.S. Ullrich, (2006) The algT gene of *Pseudomonas syringae* pv. glycinea and new insights into the transcriptional organization of the algT-muc gene cluster. *J Bacteriol* **188**: 8013-8021.

Schenk, A., H. Weingart & M.S. Ullrich, (2008) Extraction of high-quality bacterial RNA from infected leaf tissue for bacterial in planta gene expression analysis by multiplexed fluorescent Northern hybridization. *Molecular plant pathology* **9**: 227-235.

Seo, J.W., K.H. Jang, S.A. Kang, K.B. Song, E.K. Jang, B.S. Park, C.H. Kim & S.K. Rhee, (2002) Molecular characterization of the growth phase-dependent expression of the lsrA gene, encoding levansucrase of *Rahnella aquatilis*. *J Bacteriol* **184**: 5862-5870.

Smirnova, A., H. Li, H. Weingart, S. Aufhammer, A. Burse, K. Finis, A. Schenk & M.S. Ullrich, (2001) Thermoregulated expression of virulence factors in plant-associated bacteria. *Archives of microbiology* **176**: 393-399.

Song, K.B., J.W. Seo & S.K. Rhee, (1999) Transcriptional analysis of levU operon encoding saccharolytic enzymes and two apparent genes involved in amino acid biosynthesis in *Zymomonas mobilis*. *Gene* **232**: 107-114.

Srivastava, A., N. Al-Karablieh, S. Khandekar, A. Sharmin, H. Weingart & M.S. Ullrich, (2012) Genomic Distribution and Divergence of Levansucrase-Coding Genes in *Pseudomonas syringae*. *Genes (Basel)* **3**: 115-137.

Srividhya, K.V., V. Alaguraj, G. Poornima, D. Kumar, G.P. Singh, L. Raghavenderan, A.V. Katta, P. Mehta & S. Krishnaswamy, (2007) Identification of prophages in bacterial genomes by dinucleotide relative abundance difference. *PLoS One* **2**: e1193.

Suttle, C.A., (2005) Viruses in the sea. *Nature* **437**: 356-361.

Thoma, S. & M. Schobert, (2009) An improved *Escherichia coli* donor strain for diparental mating. *FEMS microbiology letters* **294**: 127-132.

Ullrich, M., A. Penaloza-Vazquez, A.M. Bailey & C.L. Bender, (1995) A modified two-component regulatory system is involved in temperature-dependent biosynthesis of the *Pseudomonas syringae* phytotoxin coronatine. *J Bacteriol* **177**: 6160-6169.

Wagner, P.L. & M.K. Waldor, (2002) Bacteriophage control of bacterial virulence. *Infect Immun* **70**: 3985-3993.

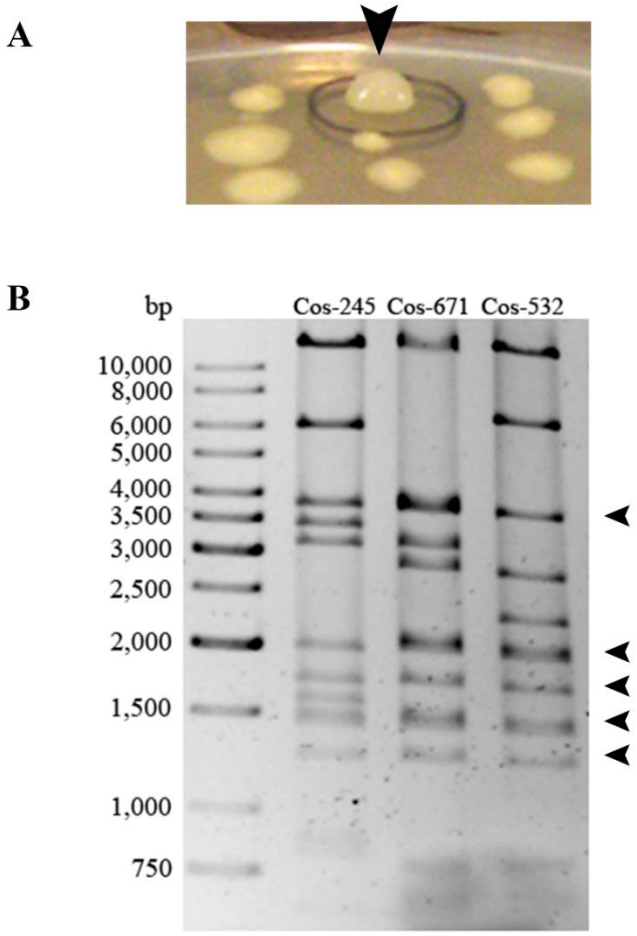
Winsor, G.L., D.K. Lam, L. Fleming, R. Lo, M.D. Whiteside, N.Y. Yu, R.E. Hancock & F.S. Brinkman, (2011) Pseudomonas Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Res* **39**: D596-600.

Xin, X.F. & S.Y. He, (2013) *Pseudomonas syringae* pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. *Annual review of phytopathology* **51**: 473-498.

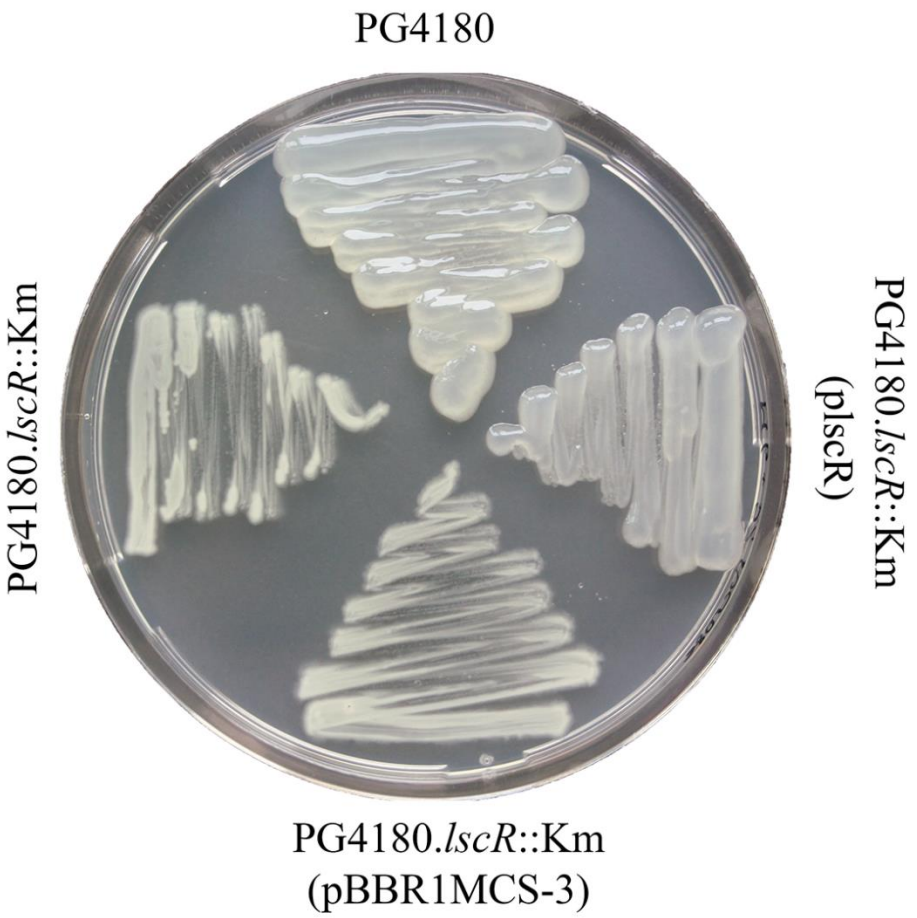
Young, J.M., G.S. Saddler, Y. Takikawa, S.H.d. Boer, L. Vauterin, L. Gardan, R.I. Gvozdyak & D.E. Stead, (1996) Names of plant pathogenic bacteria 1864-1995. *Review of Plant Pathology* **75**: 721-763.

Zhao, Y., S.E. Blumer & G.W. Sundin, (2005) Identification of *Erwinia amylovora* Genes Induced during Infection of Immature Pear Tissue. *Journal of Bacteriology* **187**: 8088-8103.

Supporting Information



Supporting Information Figure S1: Heterologous screening for levan-promoting gene(s) in a genomic cosmid library of *P. syringae* in *P. putida*. **A)** Growth of *P. putida* (plscB) transconjugants harboring PG4180 genomic library cosmids on sucrose-containing MG agar. The arrow indicates the dome-shaped levan-positive phenotype of *P. putida* (plscB, Cos-532). **B)** Electrophoretic comparison of expression-promoting cosmids, Cos-245, Cos-671, and Cos-532, after treatment with endonuclease *Pst*I. The arrows indicate fragments common to all three cosmids.



Supporting Information Figure S2: Phenotype of PG4180.*lscR*::Km complemented with plscR construct and grown on MG agar plates supplemented with 5% sucrose at 18°C. PG4180.*lscR*::Km and PG4180.*lscR*::Km (pBBR1MCS-3) served as negative controls. PG4180 wild type served as a positive control.

3.2 Analysis of the prophage-derived genomic region flanked by *trpE* and *trpG* in *P. syringae* pv. glycinea PG4180

Khaled Abdallah and Matthias Ullrich

Summary

Fitness and virulence factors encoded on genes derived from prophages are found to be integrated in several bacterial genomes as well as in plant pathogens such as *Pseudomonas syringae*. In this study, we found that the phytopathogen *P. syringae* pv. glycinea PG4180 harbors a ~25 kb fragment of a bacteriophage origin and flanked by anthranilate synthase encoding genes, *trpE* and *trpG*. The sequence, order, and function of the harbored genes were determined using nucleotide sequencing and BLAST-P. We found genes needed for phage tail morphogenesis but those that are essential for producing a functional and complete phage were absent. Although this region was previously found to be involved in bacteriocin production in other *P. sryingae* strains, no bacteriocin activity was observed in PG4180. However, this defective prophage was found to harbor a prophage assosciated promoter element and *lscR* gene previously found to play an essential role for levansucrase expression in PG4180. Comparative alignment with other *Pseudomonas* strains revealed that the possible source of this region is a phage of *Myoviridae* family, specifically phage MU or SfV. Moreover, the mosaic component of the aligned prophages can result from horizontal exchanges events which are frequent between the dsDNA tailed phages.

Keywords: *Pseudomonas syringae*, Levansucrase, Bacteriophage, Bacteriocins, Prophage-borne transcriptional regulator, LscR, Prophage assosciated promoter element, PAPE.

Introduction

Bacterial pathogens constitute a major part of the diseases affecting crop production (Varani *et al.*, 2013). These pathogens have developed several strategies, including the use of mobile genetic elements (MGE) such as bacteriophages, to interact with the host plants (Varani *et al.*, 2013). Studies on Gram negative bacterial pathogens revealed that the type III secretion systems (TTSS), a main virulence machinery needed for transferring bacterial proteins to eukaryotic host cells, are encoded on plasmids and pathogenicity islands (Hueck, 1998, Jackson *et al.*, 1999, Mansfield, 2009).

Bacteriophages, also referred to as phages, are viruses that infect and replicate inside bacterial hosts (Davies *et al.*, 2016). They play a vital role in the virulence and evolution of many pathogens (Brussow *et al.*, 2004) and are considered to be the most abundant form of life (Roux *et al.*, 2015, Frampton *et al.*, 2012). Bacteriophages are structurally composed of a head that carries a nucleic acid genome, mostly double stranded DNA, capsuled in a protein capsid. Moreover, some phages carry a tail formed of proteins or lipids (De Paepe & Taddei, 2006). The order *Caudovirales*, which constitutes 95% of known phages (Maniloff & Ackermann, 1998), can be further subdivided to three main phage families according to their morphological characteristics, specifically, tail shape (Yu *et al.*, 2016). *Myoviridae* phages are characterized by a long, helical and contractile tails. *Siphoviridae* phages have long and flexible tails while *Podoviridae* phages have short tails (Deveau *et al.*, 2006, Knezevic *et al.*, 2009).

Phages are classified into two types: Virulent phages which undergo a lytic life cycle that lead to killing of host cells (Davies *et al.*, 2016) whereas temperate phages are in a dormant state known as a lysogenic state (Frost *et al.*, 2005). Throughout lysogeny, the phage DNA is integrated into the host genome via recombination to become a prophage (Juhala *et al.*, 2000). Bacterial whole-genome sequencing shows that prophage DNA can highly contribute to the composition of the host genome content (Casjens, 2003). One example is *Escherichia coli* O157:H7 strain Sakai, a food pathogen, which harbors 18 prophage genome elements that corresponds to 16% of the total genome content (Canchaya *et al.*, 2003).

There are several ways by which temperate phages contribute to bacterial virulence and fitness (Matos *et al.*, 2013). They can introduce additional genes, referred to as morons or lysogenic genes,

not essential for viral replication but beneficial for pathogenic bacteria (Davies *et al.*, 2016, Hacker & Carniel, 2001). Lysogenic genes can encode virulence factors that enhance bacterial fitness in an environmental niche and broaden its host range by allowing avoidance of host immune defenses or offering new methods to penetrate host structural barriers (Miao & Miller, 1999). They can encode toxins such as ADP-ribosyltransferase in *V. cholerae* (Chinnapen *et al.*, 2007) and *Pseudomonas aeruginosa* (Sun & Barbieri, 2003), type III effector proteins in *Salmonella enterica* (Figueroa-Bossi & Bossi, 1999, Mirold *et al.*, 1999), and detoxifying enzymes, as Tpx and SodC in *E. coli* O157 (Kim *et al.*, 2006). Other ways by which phages affect bacterial fitness include disruption of bacterial genes such as inactivation of beta-toxin-encoding gene in *Staphylococcus aureus* (Coleman *et al.*, 1991), genome rearrangement, lysis of competing strains and protection of host cells from foreign lytic phages (Brussow *et al.*, 2004).

Previous studies on the bacterial blight pathogen of soybean, *Pseudomonas syringae* pv. *glycinea* PG4180, show that it carried a prophage associated promoter element (PAPE) (Srivastava *et al.*, 2012, Khandekar *et al.*, 2014) and a transcriptional activator (LscR) essential for the expression of levansucrase (Abdallah and Ullrich 2016, unpublished). Interestingly, both PAPE and *lscR* were of a prophage origin. In the current study, the region surrounding the above mentioned genes was subjected to sequencing analysis and open reading frames were annotated using BLAST-N. GC profile and codon usage analysis was applied and the possible function of this region was investigated. Sequencing results revealed a prophage region, spanning between *trpE* and *trpG*, which harbor genes present in similar prophage regions of the following Pseudomonas strains: *P. savastanoi* pv. phaseolicola 1448a, *P. syringae* pv. syringae B728a, and *P. syringae* pv. syringae UMAF0158. Thus, a comparative map of the gene content between the above *Pseudomonas* strains was also generated.

Experimental Procedures

Bacterial Strains, plasmids, and growth conditions

Bacterial strains and plasmids used in this study are listed in **Table 1**. *Escherichia coli* DH5α was maintained at 37°C on Luria-Bertani (LB) medium (Sambrook & Russell, 2001). Bacterial growth was continuously monitored by measuring the optical density at 600 nm (OD_{600}). Antibiotics were added to media at the following concentrations (μg/ml): tetracycline, 25.

Table 3: Bacterial strains and plasmids used in this study

Strain	Relevant characteristics	References
Bacteria		
Escherichia coli		
DH5α	supE44 D*lac*U169 (F80 *lacZ*DM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	Sambrook and Russell (2001)
Plasmids		
pRK7813	Tc^r, cosmid vector used for generation of genomic libraries	Jones and Gutterson (1987)
Cos-532	Tc^r, genomic library clones of PG4180; approx. 24069 bp insert in pRK7813, promote *lscB* expression in *P. putida* carrying *lscB*	This study

Cosmid isolation and sequencing

All primers used are listed in **Supplementary Table 1**. First set of primers (Blue) were designed based on *Pseudomonas syringae* pv. phaseolicola 1448A genome sequences available from NCBI (GenBank CP000058.1). Second set (Orange) was designed based on the reads from sequencing results. Routine molecular methods were performed using standard protocols (Sambrook et al., 1989). Nucleotide sequencing was carried out commercially (Eurofins MWG Operon, Ebersberg, Germany). DNA purifications were performed using the GeneJET PCR purification kit (Thermo Fisher Scientific, Schwerte, Germany) following the manufacturer's instructions.

Bioinformatics analysis

Vector NTI Advance 10.1.1 (Invitrogen Corporation, USA) was used for generation of the genetic maps. ISfinder online tool was used for identifying transposase-like sequences (Siguier *et al.*, 2006). The sequence manipulation suite: Codon usage was used for codon usage analysis (Stothard, 2000). GC-profiles were created using a website based tool http://tubic.tju.edu.cn/GC-Profile/ (Zhang *et al.*, 2005, Gao & Zhang, 2006). Usually for an AT-rich region, -z' curve is decreasing whereas for a GC-rich genome, -z' curve is increasing. BLAST-N and BLAST-P programs were used for online sequencing analysis (Mount, 2007). The Pseudomonas Genome database (www.pseudomonas.com) was used for gene identification and comparison of *Pseudomonas* genes and proteins (Winsor *et al.*, 2011, Winsor *et al.*, 2016).

Results and Discussion

Sequencing of *P. syringae* pv. glycinea PG4180 prophage-like region

Almost 2% of the *P. syringae* pv. syringae and 4% of *P. syringae* pv. phaseolicola genomes are related to prophage and IS elements (Lindeberg *et al.*, 2008). In the previous study regarding LscR, we showed that *P. syringae* pv. glycinea PG4180 harbors a prophage-borne region carrying genes involved in regulation of levansucrase expression. To investigate this region, genome sequencing of cosmid Cos-532 was performed. Results revealed a total of 25 genes, including PAPE and *lscR*, residing on a ~25 kb sequence flanked by *trpE* and *trpG.* All of the genes were of prophage origin and their function was identified using BLAST-P (**Fig. 1**).

Majority of the identified genes, such as those encoding for tail tape measure protein, baseplate and tail fiber, are known to be involved in tail morphogenesis (Shan *et al.*, 2014). However, no functions could be assigned to few gene products. A gene encoding for autotransporting lipase was found to be disrupted by ISPsy1 Transposase, an insertion sequence previously found to be located on a 90-kb p4180 plasmid encoding the phytotoxin coronatine in PG4180 (Alarcon-Chaidez *et al.*, 1999). Plasmids are well-known factors of virulence evolution in *P. syringae* because of their ability to hold several pathogenicity genes and perform horizontal transfer within strains (Vivian *et al.*, 2001). However, no clear evidence indicates that p4180 plasmid is a major contributor to the identified prophage region. Another gene encoding a tail fiber protein did not have the full gene sequence when compared to its homolog in *P. syringae* pv. syringae B301D. No integrases were found in the sequenced region although most temperate phages contain an integrase gene. However, presence of integrases is neither sufficient nor necessary to prove existence of a prophage (Casjens, 2003).

In phage genomes, gene clustering and ordering occurs with respect to their function. Proteins interacting with DNA are frequently found close to their DNA targets (Casjens, 2003). This might explain the close proximity of LscR to its target PAPE sequence in the identified prophage region. In addition, genes involved in tail morphogenesis were found organized in the following order: tail sheath followed by tapemeasure protein, tail baseplate proteins and tail fibre. Interestingly, this gene order is conserved as previously described in tailed prophages (Casjens, 2003, Casjens & Thuman-Commike, 2011)

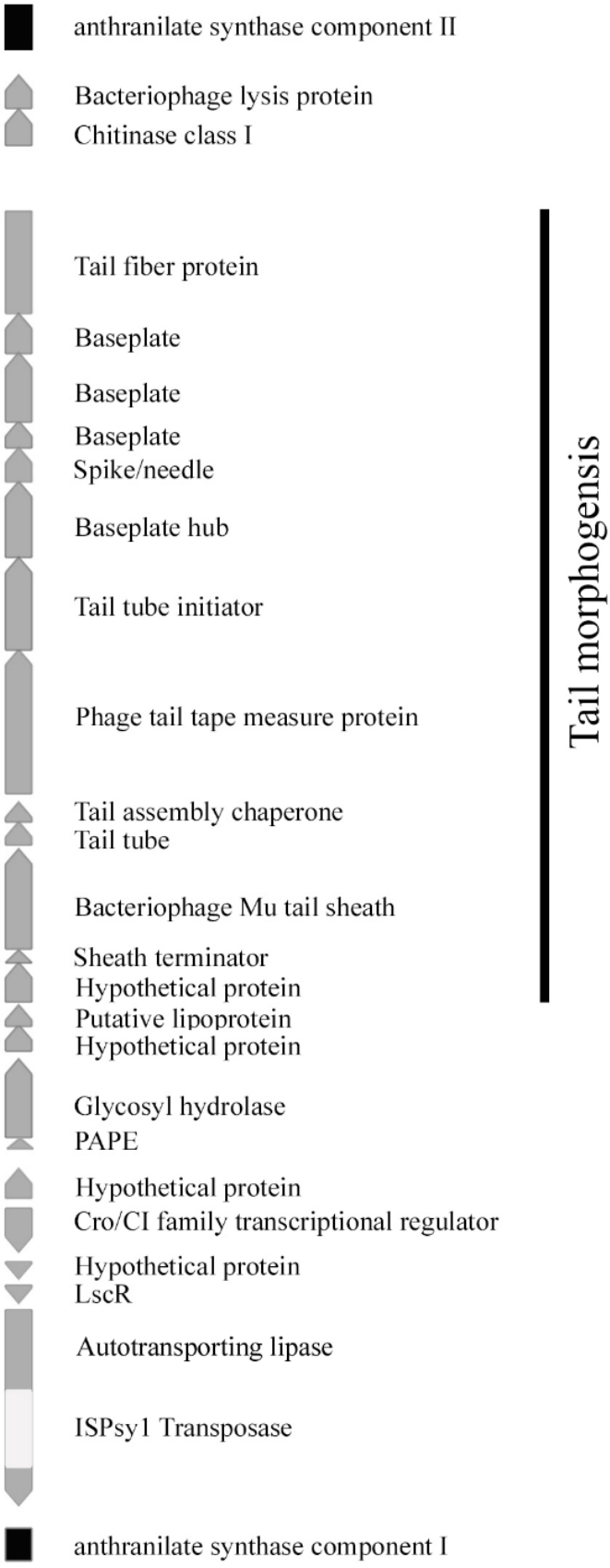


Figure 1: The genomic organization of a prophage-like DNA element flanked by *trpE* and *trpG* in *P. syringae* pv. glycinea PG4180. Genes of prophage origin are colored in grey, and those of *trp* are colored in black. ISPsy1 Transposase gene (light grey) is inserted in a gene encoding for autotransporting lipase. The product name is stated below each gene.

What function does the described prophage have in the above *Pseudomonas* strains? Does it carry any important function for the host? After infection, phages can either undergo a lytic or a lysogenic cycle in the bacterial cell (Varani *et al.*, 2013). Temperate prophages in a lysogenic state can be further characterized to four types: satellite and defective prophages, gene transfer agents and bacteriocins (Casjens, 2003). Several studies in *P. aeruginosa* have shown that the prophage region between *trpE* and *trpG* encode for bacteriocins (Shinomiya *et al.*, 1983, Kageyama, 1985). Bacteriocins produced by some bacteria resemble phage tails (Nakayama *et al.*, 2000) and are mainly protein in nature (Daw & Falkiner, 1996). They are used to kill or inhibit species and strains closely related to the producing strain (Riley & Wertz, 2002). A recent study by Hockett *et al.* (2015) showed that the bacteriophage-derived region located between *trpE* and *trpG* also encodes for a R-type syringacin responsible for the killing activity in *P. syringae* pv. glycinea B728a. R-type syringacins are retractile bacteriocins that are produced by *Pseudomonas syringae*. The R-type syringacin was found to be distributed throughout many *Pseudomonas* strains, including *P. putida, P. fluorescens* and *P. syringae* (Mavrodi *et al.*, 2009, Loper *et al.*, 2012). This prophage region, also located between *trpE* and *trpG* among most of *P. syringae* strains, was shown to be derived from bacteriophages with a Mu-like tail of *Myoviridae* family, specifically bacteriophages Mu and SfV. Although genes responsible for replication and head morphogenesis in phages MU or SfV were found missing, the preserved genes resemble those involved in tail morphogenesis or regulation (Hockett *et al.*, 2015).

Thus we hypothesized that the R-type syringacin of PG4180 could have a similar function to that observed in B728a. To confirm this hypothesis, we tested PG4180 for bacteriocin activity against a range of *P. syringae* strains that have different sensitivity spectra. However, no killing activity was observed which could be explained by the lack of a right tester strain or because PG4180 doesn't produce an active syringacin (Kevin Hockett – Personal communication). Further analysis of the alignment map revealed that B728a carried two genes encoding for chaperone and receptor binding protein that were absent in PG4180. More importantly, these two genes were found to be essential for the killing activity produced by B728a (Hockett *et al.*, 2015). Therefore, it is more likely that PG4180 produces a defective R-type syringacin.

What is the possible function then for this prophage region in *Pseudomonas*? The presence of the prophage region in addition to levansucrase genes was investigated against the draft genome

sequences of 62 *P. syringae* type and pathotype strains (Thakur *et al.*, 2016). Results revealed that this prophage region was absent in 8 strains (ICMP 8947, ICMP 4455, ICMP 3507, ICMP 4331, ICMP 16945, ICMP 9756, ICMP 7848, and ICMP 6289). Interestingly, these 8 strains also lacked any levansucrase genes (data not shown). This finding support our previous hypothesis regarding the evolutionary role of *lscR* gene in expression of the bacterial fitness gene levansucrase. However, more investigation is needed to determine whether the prophage region have other possible functions.

Function of PG4180 prophage-like region and its homology to other *Pseudomonas* strains

Prophages located at similar loci in different genomes can descend from a distinct ancestral prophage or from several unrelated integrations at the same loci (Bobay *et al.*, 2014). Therefore, prophages can be a combination of orthologous and nonorthologous prophages in a certain chromosomal locus (Bobay *et al.*, 2013). The nucleotide sequence of the identified prophage region in PG4180 was compared against other *Pseudomonas* strains using BLAST-N. Results showed that this region seems to be chimeric, harboring genes homologous to those found in the following *Pseudomonas* strains: *Pseudomonas savastanoi* pv. phaseolicola 1448a, *Pseudomonas syringae* pv. syringae B728a, and *Pseudomonas syringae* pv. syringae UMAF0158. Interestingly, all these strains showed an insertion between *trpE* and *trpG* similar to that of PG4180 thus hinting to a common prophage origin. To further investigate this, the prophage region from all the above strains was aligned with that of PG4180 based on nucleotide sequence similarity and highly related sequences were shown by shading (**Fig. 2**).

As expected, the four *Pseudomonas* strains shared a large part of the prophage with some exceptions. ISPsy1 Transposase from p4180 plasmid was only found disrupting the autotransporting lipase in PG4180. This result suggests that the inserted IS element was introduced separately and that probably doesn't play a role in structure and function of this prophage region. Another gene encoding a tail fiber protein was found unique to PG4180 although it shows an incomplete nucleotide sequence. *P. savastanoi* pv. phaseolicola 1448a carries five ORFs downstream of the *trpG* gene and encoding proteins involved in tail morphogenesis but absent from the other *Pseudomonas* strains. In addition, only B728a carried a transposase and an integrase upstream of *trpE*. UMAF0158 carried a gene encoding a protein of unknown function instead of a

glycosyl hydrolase gene usually present at this locus in PG4180, B728a and 1448A. Moreover, it carried two unique genes, downstream of *trpE*, encoding proteins involved in tail morphogenesis.

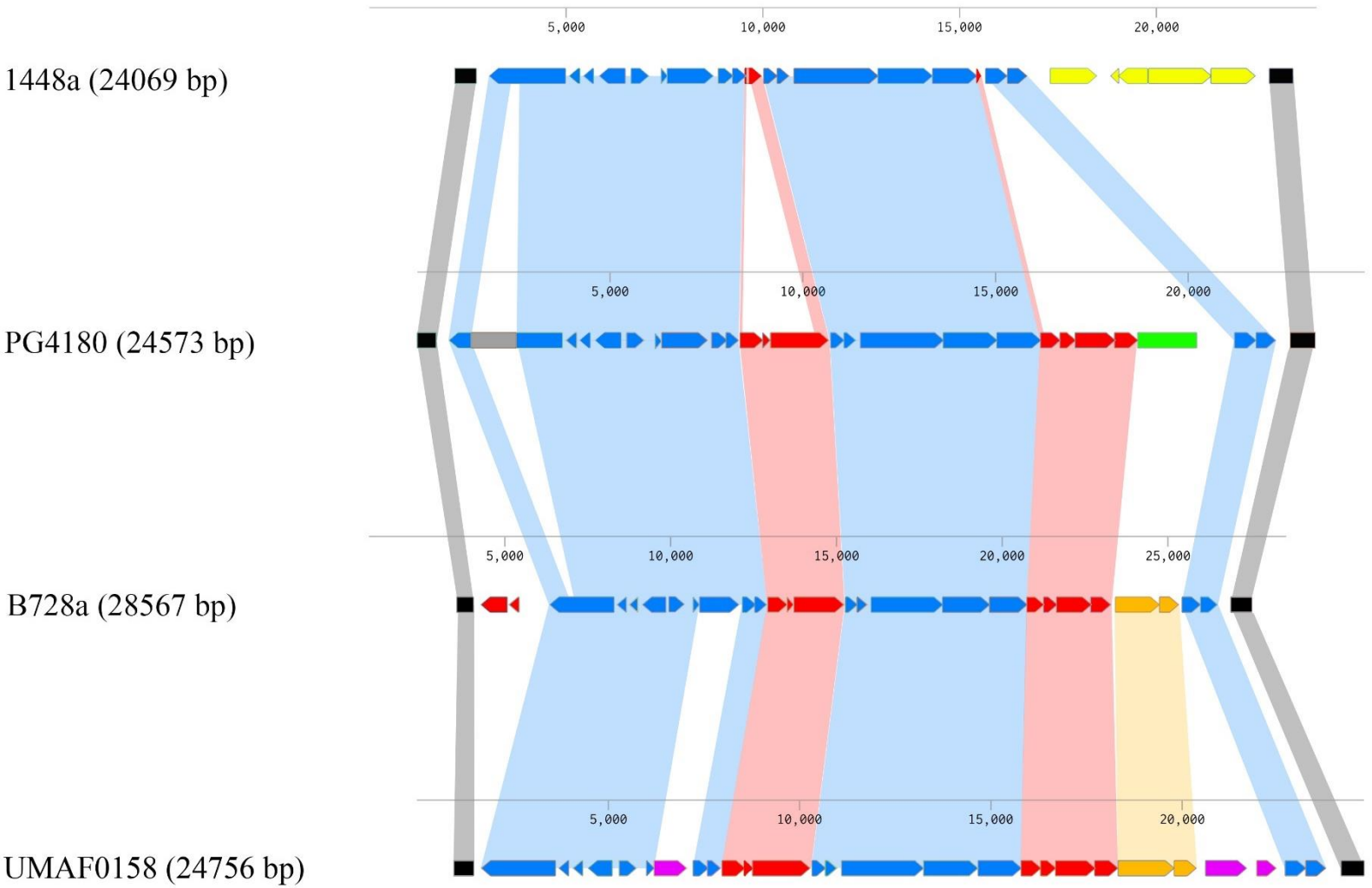


Figure 2: Comparative alignment of prophage-like regions flanked by *trpE* and *trpG* from different bacterial plant pathogens. Genes sharing nucleotide identity are linked by shadings. Availability of genes in the above *Pseudomonas* strains is color coded. Black and blue: present in all the compared *Pseudomonas* strains. Red: genes present in PG4180, B728a and UMAF0158 but largely reduced in 1448a. Orange: genes present in B728a and UMAF0158 only. Pink: genes present only in UMAF0158. Turquoise: genes present only in B728a. Green: genes present only in PG4180. Yellow: genes present only in 1448a.

The GC profile of prophage-like regions flanked by *trpE* and *trpG* in four *Pseudomonas* plant pathogens

GC profile analysis was applied on the four *Pseudomonas* strains to understand the chimeric nature of the prophage region. The -z' curves of the four prophage regions are shown in **Fig. 3**. A drop (or a rise) in the -z' curves designates decrease (or increase) in GC content (Gao & Zhang, 2006). In these curves, many regions have fluctuations, indicating the GC content is not homogenous. Some regions show high GC contents such as those highlighted in red, while other regions highlighted in yellow, green, orange, turquoise and pink show a low GC content. Also, a sharp turning point in the -z' curve was observed, such as regions highlighted in yellow, orange and pink, indicating a change from a GC-rich to a GC-poor region or vice versa (Chen, 2006). BLAST-P analysis of the highlighted areas in **Fig. 3** showed that the majority of proteins are similar to proteins found in prophages of other *Pseudomonas* strains (data not shown). This shows that the analyzed prophage regions are variable and it is more likely that they were obtained from unrelated prophages and/or recombination events. This finding is consistent with what was previously described regarding the evolution of dsDNA tailed bacteriophages (Hendrix *et al.*, 2000). Moreover, the fact that a lower %GC content was found in some of the highlighted regions suggests transfer of genetic information via horizontal recombination (Ronning *et al.*, 2010). Codon usage analysis was also applied to support the variability of the prophages, however the data obtained were not significantly different. In addition, PG4180 genome is not fully sequenced which doesn't allow to apply the proper codon usage analysis.

Although determination of GC content and codon usage analysis could support in identifying prophage sequences in their host's genome, however this form of analysis has not proceeded to the stage that it can clearly recognize prophage sequences (Blaisdell *et al.*, 1996). Thus, the identification of prophages in bacterial genomes should be mainly based on gene similairty to well-known phage genes (Casjens, 2003).

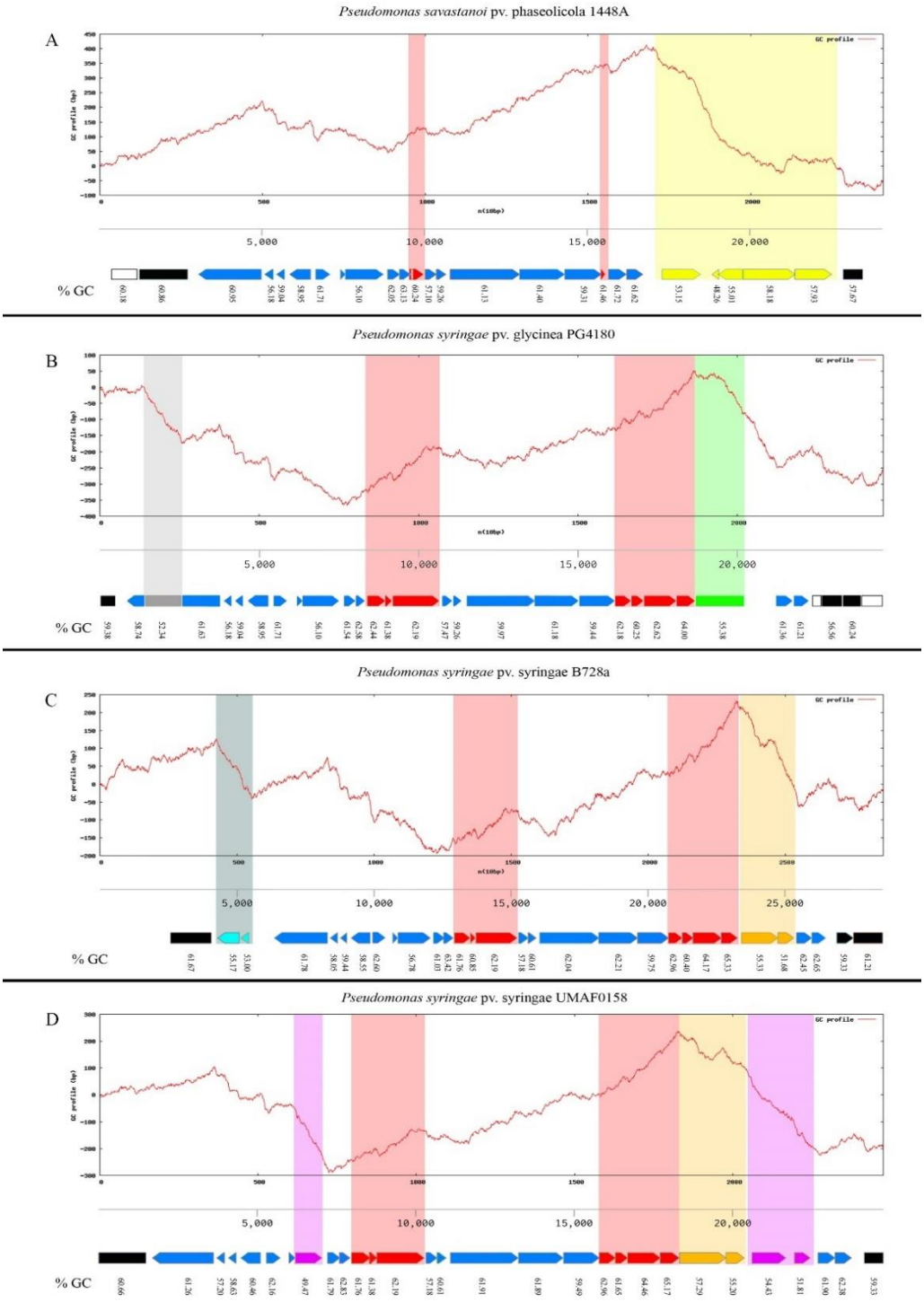


Figure 3: The -z' curves for prophage-like region of four plant pathogens. A) *Pseudomonas savastanoi* pv. phaseolicola 1448a. B) *Pseudomonas syringae* pv. glycinea 1448a. C) *Pseudomonas syringae* pv. syringae 1448a. D) *Pseudomonas syringae* UMAF0158. Shaded regions represent genes which are not common in all the above *Pseudomonas* strains.

Acknowledgments

This work was supported by a doctoral stipend from the Deutscher Akademischer Austauschdienst.

The authors declare no conflict of interest for this study.

References

Alarcon-Chaidez, F.J., A. Penaloza-Vazquez, M. Ullrich & C.L. Bender, (1999) Characterization of plasmids encoding the phytotoxin coronatine in Pseudomonas syringae. *Plasmid* **42**: 210-220.

Blaisdell, B.E., A.M. Campbell & S. Karlin, (1996) Similarities and dissimilarities of phage genomes. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 5854-5859.

Bobay, L.-M., M. Touchon & E.P.C. Rocha, (2014) Pervasive domestication of defective prophages by bacteria. *Proceedings of the National Academy of Sciences* **111**: 12127-12132.

Bobay, L.M., E.P. Rocha & M. Touchon, (2013) The adaptation of temperate bacteriophages to their host genomes. *Molecular Biology and Evolution* **30**: 737-751.

Brussow, H., C. Canchaya & W.D. Hardt, (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiology and Molecular Biology Reviews : MMBR* **68**: 560-602, table of contents.

Canchaya, C., C. Proux, G. Fournous, A. Bruttin & H. Brussow, (2003) Prophage genomics. *Microbiology and Molecular Biology Reviews : MMBR* **67**: 238-276, table of contents.

Casjens, S., (2003) Prophages and bacterial genomics: what have we learned so far? *Molecular Microbiology* **49**.

Casjens, S.R. & P.A. Thuman-Commike, (2011) Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. *Virology* **411**: 393-415.

Chen, L.L., (2006) Identification of genomic islands in six plant pathogens. *Gene* **374**: 134-141.

Chinnapen, D.J.F., H. Chinnapen, D. Saslowsky & W.I. Lencer, (2007) Rafting with cholera toxin: endocytosis and tra/cking from plasma membrane to ER. *FEMS Microbiology Letters* **266**: 129-137.

Coleman, D., J. Knights, R. Russell, D. Shanley, T.H. Birkbeck, G. Dougan & I. Charles, (1991) Insertional inactivation of the Staphylococcus aureus beta-toxin by bacteriophage phi 13 occurs by site- and orientation-specific integration of the phi 13 genome. *Molecular Microbiology* **5**: 933-939.

Davies, E.V., C. Winstanley, J.L. Fothergill & C.E. James, (2016) The role of temperate bacteriophages in bacterial infection. *FEMS Microbiology Letters* **363**.

Daw, M.A. & F.R. Falkiner, (1996) Bacteriocins: nature, function and structure. *Micron (Oxford, England : 1993)* **27**: 467-479.

De Paepe, M. & F. Taddei, (2006) Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biology* **4**: e193.

Deveau, H., S.J. Labrie, M.C. Chopin & S. Moineau, (2006) Biodiversity and classification of lactococcal phages. *Applied Environmental Microbiology* **72**: 4338-4346.

Figueroa-Bossi, N. & L. Bossi, (1999) Inducible prophages contribute to Salmonella virulence in mice. *Molecular Microbiology* **33**: 167-176.

Frampton, R.A., A.R. Pitman & P.C. Fineran, (2012) Advances in bacteriophage-mediated control of plant pathogens. *International Journal of Microbiology* **2012**: 326452.

Frost, L.S., R. Leplae, A.O. Summers & A. Toussaint, (2005) Mobile genetic elements: the agents of open source evolution. *Nature Reviews. Microbiology* **3**: 722-732.

Gao, F. & C.T. Zhang, (2006) GC-Profile: a web-based tool for visualizing and analyzing the variation of GC content in genomic sequences. *Nucleic Acids Research* **34**: W686-691.

Hacker, J. & E. Carniel, (2001) Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO reports* **2**: 376-381.

Hendrix, R.W., J.G. Lawrence, G.F. Hatfull & S. Casjens, (2000) The origins and ongoing evolution of viruses. *Trends Microbiol* **8**: 504-508.

Hockett, K.L., T. Renner & D.A. Baltrus, (2015) Independent Co-Option of a Tailed Bacteriophage into a Killing Complex in Pseudomonas. *mBio* **6**: e00452.

Hueck, C.J., (1998) Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants. *Microbiology and Molecular Biology Reviews* **62**: 379-433.

Jackson, R.W., E. Athanassopoulos, G. Tsiamis, J.W. Mansfield, A. Sesma, D.L. Arnold, M.J. Gibbon, J. Murillo, J.D. Taylor & A. Vivian, (1999) Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen Pseudomonas syringae pathovar phaseolicola. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 10875-10880.

Jones, J.D.G. & N. Gutterson, (1987) An Efficient Mobilizable Cosmid Vector, Prk7813, and Its Use in a Rapid Method for Marker Exchange in Pseudomonas-Fluorescens Strain Hv37a. *Gene* **61**: 299-306.

Juhala, R.J., M.E. Ford, R.L. Duda, A. Youlton, G.F. Hatfull & R.W. Hendrix, (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *Journal of Molecular Biology* **299**: 27-51.

Kageyama, M., (1985) [Pyocins and their genetic determinants]. *Tanpakushitsu kakusan koso. Protein, Nucleic acid, Enzyme* **30**: 722-733.

Khandekar, S., A. Srivastava, D. Pletzer, A. Stahl & M.S. Ullrich, (2014) The conserved upstream region of lscB/C determines expression of different levansucrase genes in plant pathogen Pseudomonas syringae. *BMC Microbiology* **14**: 79.

Kim, Y.H., Y. Lee, S. Kim, J. Yeom, S. Yeom, B. Seok Kim, S. Oh, S. Park, C.O. Jeon & W. Park, (2006) The role of periplasmic antioxidant enzymes (superoxide dismutase and thiol peroxidase) of the Shiga toxin-producing Escherichia coli O157:H7 in the formation of biofilms. *Proteomics* **6**: 6181-6193.

Knezevic, P., R. Kostanjsek, D. Obreht & O. Petrovic, (2009) Isolation of Pseudomonas aeruginosa specific phages with broad activity spectra. *Current microbiology* **59**: 173-180.

Lindeberg, M., C.R. Myers, A. Collmer & D.J. Schneider, (2008) Roadmap to new virulence determinants in Pseudomonas syringae: insights from comparative genomics and genome organization. *Molecular Plant-Microbe Interactions : MPMI* **21**: 685-700.

Loper, J.E., K.A. Hassan, D.V. Mavrodi, E.W. Davis, II, C.K. Lim, B.T. Shaffer, L.D.H. Elbourne, V.O. Stockwell, S.L. Hartney, K. Breakwell, M.D. Henkels, S.G. Tetu, L.I. Rangel, T.A. Kidarsa, N.L. Wilson, J.E. van de Mortel, C. Song, R. Blumhagen, D. Radune, J.B. Hostetler, L.M. Brinkac, A.S. Durkin, D.A. Kluepfel, W.P. Wechter, A.J. Anderson, Y.C. Kim, L.S. Pierson, III, E.A. Pierson,

S.E. Lindow, D.Y. Kobayashi, J.M. Raaijmakers, D.M. Weller, L.S. Thomashow, A.E. Allen & I.T. Paulsen, (2012) Comparative Genomics of Plant-Associated <italic>Pseudomonas</italic> spp.: Insights into Diversity and Inheritance of Traits Involved in Multitrophic Interactions. *PLoS Genetics* **8**: e1002784.

Maniloff, J. & H.W. Ackermann, (1998) Taxonomy of bacterial viruses: establishment of tailed virus genera and the order *Caudovirales*. *Archives of virology* **143**: 2051-2063.

Mansfield, J.W., (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: an overview of 25 years of progress in our understanding of plant innate immunity. *Molecular Plant Pathology* **10**: 721-734.

Matos, R.C., N. Lapaque, L. Rigottier-Gois, L. Debarbieux, T. Meylheuc, B. Gonzalez-Zorn, F. Repoila, M.d.F. Lopes & P. Serror, (2013) <italic>Enterococcus faecalis</italic> Prophage Dynamics and Contributions to Pathogenic Traits. *PLoS Genetics* **9**: e1003539.

Mavrodi, D.V., J.E. Loper, I.T. Paulsen & L.S. Thomashow, (2009) Mobile genetic elements in the genome of the beneficial rhizobacterium Pseudomonas fluorescens Pf-5. *BMC Microbiology* **9**: 1-18.

Miao, E.A. & S.I. Miller, (1999) Bacteriophages in the evolution of pathogen–host interactions. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 9452-9454.

Mirold, S., W. Rabsch, M. Rohde, S. Stender, H. Tschape, H. Russmann, E. Igwe & W.D. Hardt, (1999) Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic Salmonella typhimurium strain. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 9845-9850.

Mount, D.W., (2007) Using the Basic Local Alignment Search Tool (BLAST). In Bioinformatics: Sequence and Genome Analysis. 2nd edition. *Cold Spring Har Protoc* **7**.

Nakayama, K., K. Takashima, H. Ishihara, T. Shinomiya, M. Kageyama, S. Kanaya, M. Ohnishi, T. Murata, H. Mori & T. Hayashi, (2000) The R-type pyocin of Pseudomonas aeruginosa is related to P2 phage, and the F-type is related to lambda phage. *Molecular Microbiology* **38**: 213-231.

Riley, M.A. & J.E. Wertz, (2002) Bacteriocins: evolution, ecology, and application. *Annual Review of Microbiology* **56**: 117-137.

Ronning, C.M., L. Losada, L. Brinkac, J. Inman, R.L. Ulrich, M. Schell, W.C. Nierman & D. DeShazer, (2010) Genetic and phenotypic diversity in Burkholderia: contributions by prophage and phage-like elements. *BMC Microbiology* **10**: 1-17.

Roux, S., F. Enault, B.L. Hurwitz & M.B. Sullivan, (2015) VirSorter: mining viral signal from microbial genomic data. *PeerJ* **3**: e985.

Sambrook, J. & D. Russell, (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Shan, J., S. Korbsrisate, P. Withatanung, N.L. Adler, M.R.J. Clokie & E.E. Galyov, (2014) Temperature dependent bacteriophages of a tropical bacterial pathogen. *Frontiers in Microbiology* **5**: 599.

Shinomiya, T., S. Shiga & M. Kageyama, (1983) Genetic determinant of pyocin R2 in Pseudomonas aeruginosa PAO. I. Localization of the pyocin R2 gene cluster between the trpCD and trpE genes. *Molecular & general genetics : MGG* **189**: 375-381.

Siguier, P., J. Perochon, L. Lestrade, J. Mahillon & M. Chandler, (2006) ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Research* **34**: D32-36.

Srivastava, A., N. Al-Karablieh, S. Khandekar, A. Sharmin, H. Weingart & M.S. Ullrich, (2012) Genomic Distribution and Divergence of Levansucrase-Coding Genes in Pseudomonas syringae. *Genes (Basel)* **3**: 115-137.

Stothard, P., (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* **28**: 1102, 1104.

Sun, J. & J.T. Barbieri, (2003) Pseudomonas aeruginosa ExoT ADP-ribosylates CT10 regulator of kinase (Crk) proteins. *The Journal of Biological Chemistry* **278**: 32794-32800.

Thakur, S., B.S. Weir & D.S. Guttman, (2016) Phytopathogen Genome Announcement: Draft Genome Sequences of 62 Pseudomonas syringae Type and Pathotype Strains. *Molecular Plant-Microbe Interactions : MPMI* **29**: 243-246.

Varani, A.M., C.B. Monteiro-Vitorello, H.I. Nakaya & M.A. Van Sluys, (2013) The role of prophage in plant-pathogenic bacteria. *Annual Review of Phytopathology* **51**: 429-451.

Vivian, A., J. Murillo & R.W. Jackson, (2001) The roles of plasmids in phytopathogenic bacteria: mobile arsenals? *Microbiology (Reading, England)* **147**: 763-780.

Winsor, G.L., E.J. Griffiths, R. Lo, B.K. Dhillon, J.A. Shay & F.S. Brinkman, (2016) Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic Acids Research* **44**: D646-653.

Winsor, G.L., D.K. Lam, L. Fleming, R. Lo, M.D. Whiteside, N.Y. Yu, R.E. Hancock & F.S. Brinkman, (2011) Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. *Nucleic Acids Research* **39**: D596-600.

Yu, J.G., J.A. Lim, Y.R. Song, S. Heu, G.H. Kim, Y.J. Koh & C.S. Oh, (2016) Isolation and Characterization of Bacteriophages Against Pseudomonas syringae pv. actinidiae Causing Bacterial Canker Disease in Kiwifruit. *Journal of Microbiology and Biotechnology* **26**: 385-393.

Zhang, C.T., F. Gao & R. Zhang, (2005) Segmentation algorithm for DNA sequences. *Physical review. E, Statistical, nonlinear, and soft matter physics* **72**: 041917.

Supplementary Table 1: Oligonucleotide primers used in this study

Primers	Nucleotide sequence (5'-3')
0671_F	ACACGTATCAGAACGCCGAGAC
0671_R	GAGGTTGCAGGACAGTATCGTACC
0670_F	CCATTGCCAAGGAGTGGAAGGA
0670_R	ATCCGTCGTTGTTGCTGAGGAC
0669_0668F	CGATTTGGATGCTGCATTGTGG
0669_0668 R	GCAACCCCTTCCTCGACCTATG
0667_0666 F	GGCTGTCAGTCCTCCTTGCCAG
0667_0666 R	CCCGCTGTTATCTCACGAAGCT
0666_0663 F	GCCAATTGGGAGCGTGGTCAGC
0666_0663 R	GCCGGAGGGAACATCAGTGCGC
0663_0662 F	AGAGCGAGCGGCACCAGCAGTT
0663_0662 R	GGCTTGTGAAACCGATGACCGG
0662_0661 F	GTGATGGGCCTGTCACCGCT
0662_0661 R	AATCCTGTTGGACGTGCCGC
0661_0660 F	GGGAGCTGGCAATGAGTGAAGT
0661_0660 R	ATCAGAAGACCTGCCGCCAG
0660_0655 F	AGCAATCAGGCGGAGACCGT
0660_0655 R	GCCGGAAGGTCAGTTGCGAA
0655 F	TTCCAGAAGCAACAACGCCG
0655 R	GGGCTCCGTTGTTGGTTCAG
0654_0652 F	ACAGGGAATCGATGTACCAGGAC
0654_0652 R	TGACCGGAGTTCAACACGAG
0652_0650 F	GTAGCCTCTTCCATCAGCATCG
0652_0650 R	GGCCGAAGGTATCAGCGTGA
0650_trpE F	AGACCTGCTGTCCGATCACAAGG
0650_trpE R	CAGTGGCAGAGCGTGCTTGC
0650down_F	GTGCATGGCGACCGCTGGAGCG
p4180A_R	TGCCGGAGCGTATCGGCAG

p4180A_F	GTAGCTGCCGCCGATGGTCAG
0651_middle_F	CAGTGAAGGGCACGGCGTGGCG
0658_F	GGTCGAGAGTGCTGCATCA
0659_F	CGGTGTGGACATGACCATCAC
0662_F	CAGGCGCTGCGCAGCGCCCTGT
11+12+9_F	CAGCCATCTGCATGGCGGTG
0666_F	CACCGCTGGCGATGGAATG
TrpG_F	GCATCGTGGTCTCGCCTGGC
TrpG_R	GCCAGGCGAGACCACGATGC
Cosmidseq_1C	CACGATAGGACGCAGAGTG
Cosmidseq_2C	CGTGCGACGCTCCGCGTCGCAT
Cosmidseq_3C	CGAGACCACGTCCTTGCTGAC
Cosmidseq_4C	CAGCGGTGCCTACGCCAAT
Cosmidseq_5C	CGAGATGGACTGCGTTGAC
Cosmidseq_6C	GCAGCAGTTCCAGCAGATG
Cosmidseq_7C	GATGACCAACCGTGTCGATCG
Cosmidseq_8C	GTGATGGCCACAGTGGACG
Cosmidseq_9C	CATCGCAGCTGCTGGCTG
Cosmidseq_10C	CAGGCAGGGGCTCGATGG
Cosmidseq_11C	GATGTCCAGGGAGCTGGAC
Cosmidseq_12C	GCTCCATGCTGGCCTCGA
Cosmidseq_13C	GACTGGATCGTCGACCAG
Cosmidseq_14C	GTCATGCGTGAGCTGGCA
Cosmidseq_15C_F	GCTGCTGGCGAGCAGTTG
Cosmidseq_15C_R	CGCACATAATGCAGCTGC
Cosmidseq_16C_R	CGACATGACGGCATGCTCCAG
Cosmidseq_16C_F	GCTGGACCTGCACAACCGTG
Cosmidseq_17C	GCCAGCTACTGGGGAGAGC
Cosmidseq_18C	CGACAGTGCCTGTGCTGG
Cosmidseq_19C	GACTCCGCTACCTGCTGGCT
Cosmidseq_20C	GTTGATAGGAGGAGGGAAT

3.3 Expression of extra-cellular levansucrase in *Pseudomonas syringae* is controlled by the *in planta* fitness-promoting metabolic repressor HexR

Summary

Pseudomonas syringae PG4180 causes bacterial blight on soybean plants and enters the leaf tissue through stomata or open wounds, where it encounters a sucrose-rich milieu. Sucrose is utilized by invading bacteria via the secreted enzyme, levansucrase (Lsc), liberating glucose and forming the polyfructan levan. *P. syringae* PG4180 possesses two functional *lsc* alleles transcribed at virulence-promoting low temperatures. We hypothesized that transcription of *lsc* is controlled by the hexose metabolism repressor, HexR, since potential HexR binding sites were identified upstream of both *lsc* genes. A *hexR* mutant of PG4180 was significantly growth-impaired when incubated with sucrose or glucose as sole carbon source, but exhibited wild type growth when arabinose was provided. Analyses of *lsc* expression resulted in higher transcript and protein levels in the *hexR* mutant as compared to the wild type. The *hexR* mutant's ability to multiply *in planta* was reduced. HexR did not seem to impact hypersensitive response (*hrp*) gene expression as evidenced by the *hexR* mutant's unaltered hypersensitive response in tobacco and its unmodified protein secretion pattern as compared to the wild type under *hrp*-inducing conditions. Our data suggested a co-regulation of genes involved in extra-cellular sugar acquisition with those involved in intra-cellular energy-providing metabolic pathways in *P. syringae*.

Keywords: Plant pathogen, Bacterial blight, Soybean, *Pseudomonas syringae*, Levansucrase, HexR

Introduction

Fructan or glucan polymers are formed wherever microbes encounter sucrose-rich conditions, would it be in association with plants, in the oral cavity, in food manufacturing, or during bio-fuel production processes (Srivastava *et al.*, 2009). When plant-borne sucrose is present, the soybean-infecting bacterial blight pathogen, *Pseudomonas syringae* pv. glycinea, uses levansucrase (Lsc) to synthesize the extra-cellular high-molecular fructofuranan, levan, thereby releasing glucose for primary metabolism. Three levansucrase-encoding genes, *lscA*, *lscB*, and *lscC*, were identified in *P. syringae* pv. glycinea PG4180, from which only *lscB* and *lscC* are expressed, as a mutant lacking *lscB* and *lscC* but possessing *lscA* is levan-deficient (Li & Ullrich, 2001). Furthermore, quantitative expression analysis of *lsc* genes by quantitative Reverse Transcriptase (qRT)-PCR showed that *lscB* and *lscC* are actively expressed. However, *lscA* is not being expressed due to an altered upstream region of *lscA* which does not seem to promote *lsc* expression (Khandekar *et al.*, 2014). Both enzymes are synthesized maximally at 18ºC *in vitro* and *in planta* and their expression is optimal at the early logarithmic growth stage (Li *et al.*, 2006, Schenk *et al.*, 2008).

Bacterial communities growing epiphytically on plants are primarily affected by carbon availability as supported by the finding that very low sugar concentrations are sufficient to support the growth of 10^7 to 10^8 cells per leaf (Lindow & Brandl, 2003). Stomatal openings and wounds provide the site of entry for *P. syringae*. Under favorable micro-environmental conditions, the bacterial cells live endophytically and subsequently initiate the infection process via production of the phytotoxin coronatine (Budde & Ullrich, 2000, Melotto *et al.*, 2008) and attachment to plant cell surfaces. The infection process is fostered by low environmental temperatures such as 18-20°C as opposed to the optimal growth temperature of *P. syringae*, 28°C (Dunleavy, 1988, Smirnova *et al.*, 2001). A complex sequence of events mediated by injection of bacterial hypersensitive reaction and pathogenicity (Hrp) effector proteins into plant cells (Mansfield, 2009) ultimately activates plant-borne K^+ efflux and H^+ influx, which increases the apoplastic pH from 5.5 to 7.5 (Atkinson & Baker, 1987). Subsequently, this high extra-cellular pH induces efflux of the dominant photo assimilate, sucrose, from plant cells (Atkinson & Baker, 1987). Apoplastic sucrose ranging in concentrations from 20 µM to 1-5 mM is hydrolyzed by either plant-borne invertases or by extra-cellular microbial enzymes, e.g. Lsc (Roitsch & Gonzalez, 2004, Biemelt & Sonnewald, 2006).

For glucose metabolism, metabolic pathway structures vary among bacterial species with different ecological niches (Papp *et al.*, 2009). In contrast to enterobacteria (Lessie & Phibbs, 1984), pseudomonads utilize the Entner-Doudoroff (ED) pathway due to lack of 6-phosphofructokinase and hence do not catabolize sugars via the Embden-Meyerhof-Parnas pathway (Entner & Doudoroff, 1952, Portais & Delort, 2002). The ED pathway can be linear, alternative, modified with non-phosphorylated intermediates, or cyclic (Conway, 1992) and was first described in *Pseudomonas saccharophila*, which now belongs to the order *Burkholderiales* (Entner & Doudoroff, 1952, Xie & Yokota, 2005).

The genes required for glucose metabolism in *Pseudomonas putida* KT2440 are organized in several operons (Kim *et al.*, 2008). The first operon consists of the *zwf*, *pgl*, and *eda* genes coding for glucose-6-phosphate dehydrogenase, 6-phosphogluconolactonase (both enzymes of the glucose phosphorylative pathway), and Eda (an enzyme of the Entner-Doudoroff pathway) (Daddaoua *et al.*, 2009, del Castillo *et al.*, 2008). Divergently directed to this operon is the gene encoding for the hexose metabolic repressor, *hexR* (Kim *et al.*, 2008). The second operon consists of the *edd* and *glk* genes that encode 6-phosphogluconate dehydratase (the first enzyme of the Entner-Doudoroff pathway) and glucokinase (an enzyme of the glucose phosphorylative pathway), respectively (del Castillo *et al.*, 2008, Daddaoua *et al.*, 2009). Divergently directed to this operon is the gene *gap1* which codes for glyceraldehydes 3-phosphoste dehydrogenase (Kim *et al.*, 2008). The transcription of these genes and operons is negatively regulated by HexR. Two monomers of HexR bind to the promoter regions of *edd, zwf,* and *gap-1* genes by recognizing a palindromic sequence TTGTN7–8ACAA (del Castillo *et al.*, 2008, Kim *et al.*, 2008, Petruschka *et al.*, 2002, Leyn *et al.*, 2011). In the current study, it was hypothesized that in *P. syringae* not only genes involved in cellular glucose metabolism but also genes encoding extra-cellular Lsc were controlled by HexR. In turn, this might have consequences for our understanding about what determines bacterial *in planta* fitness and potentially virulence.

In order to address this hypothesis, a *hexR* mutant was generated in PG4180 and tested for its growth in minimal medium containing glucose, sucrose, or arabinose as sole carbon source. Analyses of *lsc* expression by qRT-PCR and Western blotting were conducted for the PG4180 wild type and its *hexR* mutant. The mutant was compared to the wild type in terms of its *in planta* fitness.

Furthermore, hypersensitive response (HR) reactions and profiles of secreted proteins were compared for the wild type and the *hexR* mutant when grown under *hrp*-inducing conditions.

Material and Methods

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in **Table. 1**. *Escherichia coli* was maintained at 37°C on Lysis Broth medium (Sambrook *et al.*, 1989). *P. syringae* was routinely maintained at 28°C on mannitol-glutamate (MG) medium (Keane *et al.*, 1970). For liquid cultures at 18°C, bacteria were grown in 200 ml of Hoitink-Sinden (HS) medium (Palmer & Bender, 1993) supplemented with various carbon sources in 1-liter Erlenmeyer flasks. 113 mM of glucose (HS + glucose) was replaced by 57 mM of sucrose in HS + sucrose medium while HS + arabinose medium had the following constituents: 0.8 mM $MgSO_4$, 30 mM KH_2PO_4, 16 mM K_2HPO_4, 16 mM KNO_3, 20 µM $FeCl_3$, 133 mM L-arabinose. Bacterial growth was continuously monitored by measuring the optical density at 600 nm (OD_{600}). Antibiotics were added to media at the following concentrations (µg/ml): ampicillin, 50; spectinomycin, 25; kanamycin, 25; tetracycline, 25; gentamicin, 2.

Table 1: Bacterial strains and plasmids used in this study.

Strain	**Relevant characteristics[a]**	**Reference / source**
Escherichia coli		
DH5α	*supE*44 Δ*lac*U169 (Φ80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	(Sambrook *et al.*, 1989)
Pseudomonas syringae pv. glycinea		
PG4180	wild type, levan+	(Bender *et al.*, 1993)
hexR	Km^r, *hexR* mutant of PG4180, levan++	This study
Plasmid		
pRK2013	Km^r, helper plasmid	(Figurski & Helinski, 1979)
pBBR1MCS	Cm^r, broad-host-range cloning vector	(Kovach *et al.*, 1994)
pBBR1MCS-3	Tc^r, broad-host-range cloning vector	(Kovach *et al.*, 1995)

pGEM-T Easy	Ap^r, vector for cloning of PCR products	Promega
pEX18Ap	Ap^r, oriT+ sacB+ gene replacement vector	(Hoang *et al.*, 1998)
pFKm	Source of Km cassette flanked with *FRT* sequences	(Choi *et al.*, 2008)
pGEM.hexR1	Ap^r, contains 456-bp upstream region of *hexR*	This study
pGEM.hexR2	Ap^r, contains 360-bp downstream region of *hexR*	This study
pGEM.hexR1-Km	Ap^r, Km , 1230-bp KpnI fragment containing Km-*FRT* cassette of pFKm1 cloned into pGEM.hexR1	This study
pGEM.hexR-Km	Ap^r, Km, 360-bp BamHI-SpeI fragment of pGEM.hexR2 cloned into pGEM.hexR1-Km	This study
pEX.hexR-Km	Ap^r, Km, 2046-bp EcoRI fragment of pGEM.hexR-Km cloned into pEX18Ap	This study

[a]Ap, ampicillin; Cm, chloramphenicol; Gm, gentamicin; Km, kanamycin; Sp, spectinomycin; Tc, tetracycline

Generation of *hexR* mutant in PG4180

A *P. syringae* PG4180 *hexR* mutant was generated using the broad-host-range Flp-FRT recombination system (Hoang *et al.*, 1998). Two fragments flanking the *hexR* gene were amplified from PG4180 genomic DNA using two pairs of primers: HexR_1f/HexR_1r and Hex_2f/HexR_2r (**Table. 2**). PCR products were cloned into pGEM-T Easy (Promega) yielding plasmids pGEM.HexR1 and pGEM.HexR2 (**Table. 1**). A 1,230-bp *Kpn*I fragment containing a Km^R cassette flanked with FRT sequences was removed from plasmid pFKm (Choi *et al.*, 2008) and ligated into *Kpn*I-digested pGEM.HexR1, yielding pGEM.HexR1-Km. A 360-bp *Spe*I-*Bam*HI fragment digested from pGEM.HexR2 was ligated into *Spe*I-*Bam*HI-digested pGEM.HexR1-Km, yielding plasmid pGEM.HexR-Km. Finally, a 2,046-bp *Eco*RI fragment was removed from pGEM.HexR-Km and ligated into *Eco*RI-digested plasmid pEX18Ap (Hoang *et al.*, 1998), yielding the *hexR* gene replacement plasmid pEX.HexR-Km. This plasmid was mobilized into *P. syringae* PG4180 by tri-parental mating. Putative mutants were screened on MG medium agar plates supplemented

with kanamycin and were subsequently confirmed for the genotype by PCR using primers hex_up_fwd and hex_down_rev (**Table. 2**).

Table 2: Oligonucleotide primers used in this study.

Primers	**Nucleotide sequence (5'-3')**
HR_1f	GGATCC GTTCAACTCATCGAGTC
HR_1r	CAGATGCGACTGTTCGTC
HR_2f	GACCCCCGGATCAGTGCCAG
HR_2r	GGATCC GGTACC CAGCCGCTATCCGATCGAG
lscB/C_RT_fwd	TCGGTTATCCTGACCCTGAC
lscB/C_RT_rev	CCATGACGATCTTCCCAGTC
lscB_hex_fwd	CGCAATTAATGCGAGCCCGCAGG
lscB_hex_rev	TTGCATTGGTCTCCTTGTGCTTC
hex_up_fwd	CGAGCAAGTCGCACCG
hex_down_rev	GAAGTCGACATGCAGGTAG
hexR_fwd	GAATTCATGGACCGCGTAAGAAACC
hexR_rev	CTGCAGTCAGCCTTGATCCTCGATC
hexR_verif_fwd	CTCAACCCGCAGATGGCAA
hexR_verif_rev	CGATGACCTCGCGGATCAT

[a] Restriction sites incorporated in primers are underlined**.**

Verification of the *hexR* mutant's phenotype by Reverse-Transcriptase polymerase chain reaction (RT-PCR)

Template-specific primers were designed for *hexR* and *lscB/C* genes of *P. syringae* PG4180. Bacterial cells were grown in HS + arabinose medium and harvested at an OD600 of 0.5. RNA was extracted by acid phenol/chloroform extraction method (Schenk *et al.*, 2008). RT-PCR was performed on total RNA using RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, Schwerte, Germany). The hexR_fwd, hexR_rev, hexR-verif_fwd, hexR_verif_rev and lscB/C primers were used to check for presence of a *hexR and lscB/C* mRNA by PCR using cDNA as template. Regular PCR with the same primer-pairs and genomic DNA as template were used as control. The thermocycler program was as follows: 1 cycle of 95°C for 60 s; 30 cycles of 95°C for 10 s, 60°C for 30 s, 72°C for 30 s; 1 cycle of 72°C for 5 min. The results were analyzed by 1% agarose gel electrophoresis.

Analysis of *lsc* expression by quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR)

Bacterial cells were grown in HS + arabinose medium at 18°C. When cultures reached respective OD_{600} values, total RNA was isolated by acid phenol/chloroform extraction, and samples were normalized by multiplexed fluorescent Northern hybridization and 23S rRNA transcript amount comparison as described previously (Schenk *et al.*, 2008). Yield and purity of RNA were determined by measuring absorption at 260 and 280 nm. Total RNA samples were treated with TURBO DNA-free (Applied Biosystems, Darmstadt, Germany) to remove remaining traces of genomic DNA as described by the manufacturer's recommendation.

SYBR green-based qRT-PCR was performed with 1 ng normalized RNA template and 200 nM primers (lscBC_RT_fwd, lscBC_RT_rev) using the QuantiTect SYBR Green one-step RT-PCR Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The thermocycler program comprised an initial step of 95°C for 15 min followed by 40 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s. Reactions were performed in technical duplicates and biological triplicates with a Mastercycler® realplex2 real-time PCR system (Eppendorf, Hamburg, Germany) as described by the manufacturer using their universal program. Reactions with no addition of reverse transcriptase served as negative controls and proved lack of DNA contamination. Specificity of

amplification was assessed by analyzing the melting curve of the amplification product. Due to very high sequence identity between *lscB* and *lscC* it was not possible to design primers discriminating between these two mRNAs, thereby expression profile of *lsc* is always referred as a combination of both genes.

Immunological detection of Lsc

Generation and concentration of cell-free supernatants of *P. syringae* cells and the use of polyclonal antibodies were carried out as described previously (Li *et al.*, 2006). Cultures were grown in HS+ arabinose medium. Equal aliquots of protein fractions were loaded (10 µg/lane) and separated by 10% SDS-PAGE. Electrophoresis, electro-blotting on nitrocellulose membranes, and immunodetection were conducted by standard procedures (Sambrook *et al.*, 1989).

Growth of PG4180 wild type and *hexR* mutant *in planta*

In planta growth of PG4180 and its *hexR* mutant was evaluated on soybean plants. Soybean seedlings were germinated and grown in an environmentally controlled chamber for approximately three weeks prior to the growth assays. PG4180 wild type and *hexR* mutant were incubated for 48 hours at 28 °C on MG agar plates. Cells were suspended in distilled water, adjusted to an OD_{600} of 0.1 (corresponding to approximately 10^7 CFU/ml) and applied to the leaves with an airbrush (~8 psi) until the leaf surfaces were uniformly wet. Subsequently, humid environment was achieved by enclosing the inoculated plants with a clear plastic bag for overnight. Inoculated plants were grown in a plant growth chamber (19-21°C) with a 12-hr light period. Survival and growth of bacterial strains was monitored by removing random leaf samples at 1-14 days post inoculation. At days 1, 3, 5, 7, 9, 11 and 14 after inoculations, two individual leaves were randomly excised from plants corresponding to each inoculums and their weight was measured. Epiphytic bacteria were isolated by placing leaves in a 50-ml falcon tubes containing 20 ml of external wash buffer (0.1M potassium phosphate, 0.1% bactopeptone and pH=7.0) and the tubes were sonicated for 7 minutes (O'Brien R & Lindow, 1988). Leaves were then removed and macerated in 20 ml of isotonic solution (0.9% NaCl) using sterile mortar and pestle. Bacterial counts (CFU/g fresh weight) were determined by plating dilutions from external wash buffer and leaf homogenate onto MG agar and counting of fluorescent colonies after incubation at 28°C for 96 h.

Hypersensitive Response assay on tobacco

Tobacco plants (*Nicotiana tabacum* cv. Petit Havana SR1) were grown in a light chamber at 20 to 25°C, 60% humidity, with a 12-h photoperiod (15,000 lux). PG4180 wild type and *hexR* mutant were incubated for 48 h at 28°C on MG agar. Cells were suspended in sterile 0.9 % NaCl, adjusted to an OD_{600} of 0.1 (corresponding to approximately 10^7 CFU/ml). Tobacco plants were inoculated with bacterial suspensions by syringe injection of leaf veins of the third and fourth leaf. As negative control sterile 0.9 % NaCl was used. Infiltrated areas were monitored for development of the hypersensitive reaction in form of necrosis after 24 and 48 h.

Extra-cellular protein pool preparation and SDS-PAGE

For sample preparation of secreted proteins, bacteria were grown in two consecutive overnight pre-cultures in King's B broth (King *et al.*, 1954) at 28°C. From the first pre-culture, the cell suspension was adjusted to an OD_{600} of 0.1 with 50 ml of fresh King's B broth and incubated at 28°C. Bacterial cultures were harvested at an OD_{600} of 1.0 and centrifuged at 4,000 rpm for 30 min. The bacterial cells were washed two times with Hrp-inducing medium (He *et al.*, 1993, Huynh *et al.*, 1989) or with HS+ glucose medium (Palmer & Bender, 1993), respectively. Subsequently, the cell pellets were resuspended in 50 ml of Hrp-inducing medium or HS+ glucose medium at 28°C and incubated shaking at 250 rpm for 4 h. Next, 25 ml of the bacterial cultures were centrifuged at 4,000 rpm for 30 min. Cell-free supernatant samples were prepared by filter-sterilization through 0.2-µm pore size membrane filters and concentrated 50-fold using 10-kDa millipore filters (Amicon, Billerica, USA). Extra-cellular protein samples were stabilized in 50mM Tris-HCl (pH 8.8). Total protein amounts were determined using a Nanodrop apparatus (Thermo Fisher Scientific, Langenselbold, Germany), 10 µl of protein samples were separated by 12.5 % SDS-PAGE and subsequently visualized by silver staining according to established procedures (Sambrook & Russell, 2001).

Results

Generation and genotypic characterization of the *P. syringae hexR* mutant

del Castillo *et al.* (2008) and Daddaoua *et al.* (2009) had previously shown that genes encoding enzymes of the phosphorylative branch and the ED pathway of glucose catabolism in *P. putida* were regulated by the hexose metabolism repressor, HexR. In *P. syringae*, extra-cellular Lsc releases glucose by sucrose hydrolysis. In order to investigate the impact of HexR on expression of *lscB* and *lscC* genes, a *hexR*-deficient mutant of *P. syringae* PG4180 was generated by insertion of a kanamycin resistance cassette in the *hexR* gene as a result of homologous recombination. The resulting mutant genotype was verified by growth on kanamycin-containing medium and RT-PCR analysis (**Supplementary Fig. 1**).

In silico analysis of HexR binding site upstream of *lsc* genes

Analysis of the nucleotide sequences upstream of *lscB* and *lscC,* which are almost identical at nucleotide level (Hettwer *et al.*, 1998, Li & Ullrich, 2001), revealed the presence of a sequence (nucleotides +113 to +127 with respect to the corresponding translational start site) with high similarity to the conserved motif TTGTN$_{7–8}$ACAA previously shown to be the DNA binding site of HexR in *P. putida* (Petruschka *et al.*, 2002, del Castillo *et al.*, 2008, Daddaoua *et al.*, 2009) (**Fig. 1**). This finding suggested a putative binding of HexR of *P. syringae* to the upstream sequence of both *lsc* genes.

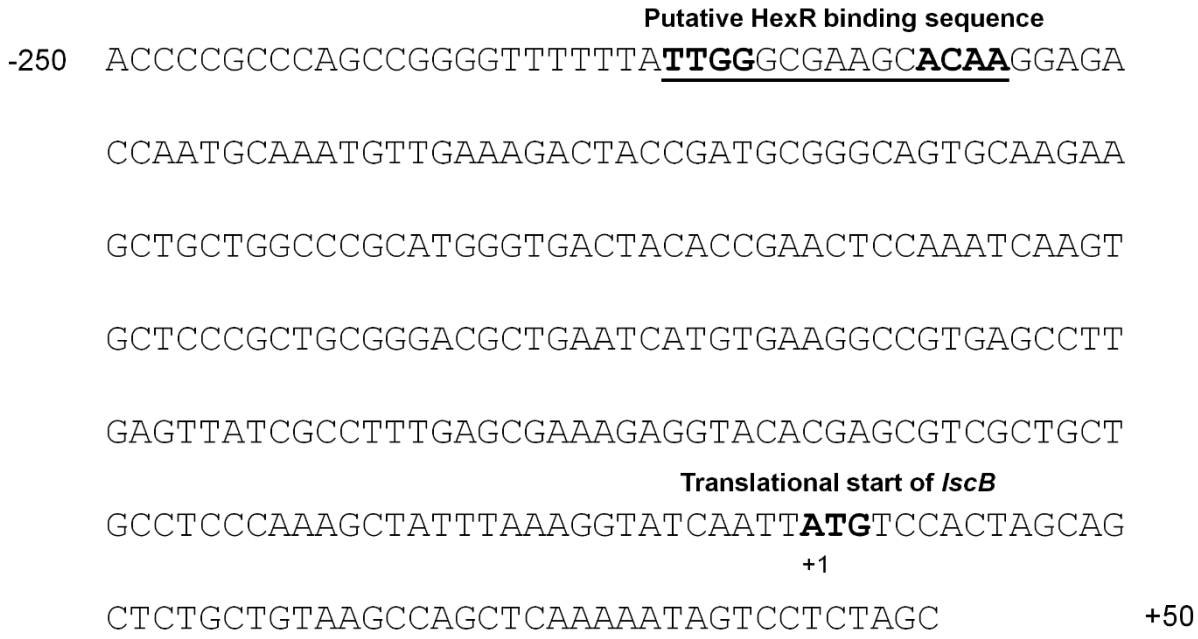


Figure 1. Nucleotide sequence of the upstream region of *lscB*. Sequence contains a putative HexR binding site (underlined) similar to that described by Daddaoua *et al.* (2009). +1 represents the translational start site of *lscB*.

Comparison of different HexR protein sequences

Multiple sequence alignments of several HexR protein sequences derived from Lsc-producing *P. syringae* strains revealed that those proteins have a high pair-wise identity scores of >97 % (data not shown). PFAM analysis (Finn *et al.*, 2010) of these proteins showed that there is a DNA binding domain (PFAM01418) and a sugar isomerase (SIS) binding domain (PFAM 01380) as described for *P. putida* KT2440 by Daddaoua *et al.* (2009) (**Supplementary Fig. 2**) suggesting a similar way of enzymatic activity. A consensus sequence obtained from multiple alignments of HexR sequences from several *P. syringae* strains revealed 91 % identity at the amino acid sequence level to that of *P. putida* (**Supplementary Fig. 3**).

In vitro and *in planta* growth of *P. syringae* PG4180 and its *hexR* mutant

PG4180 wild type and its *hexR* mutant were comparatively grown in minimal media containing different carbon sources at 18°C (**Fig. 2**). In contrast to the wild type, the *hexR* mutant did not grow significantly in HS medium containing 20 g/ml (113 mM) of glucose. Replacing glucose by 10 g/ml (57 mM) of sucrose, thereby providing an equal total number of carbon atoms did not change the weak growth phenotype of the *hexR* mutant. However, growth of the wild type was unaffected by this change of carbon source.

When glucose was replaced by an equal amount of 20 g/ml (133 mM) arabinose as the sole carbon source, growth of the *hexR* mutant was not distinguishable from that of the wild type indicating that HexR might not be involved in regulation of pathways utilizing arabinose, and that the *hexR* mutant growth phenotype was restricted to glucose utilization (**Fig. 2**). For the wild type, sucrose apparently allowed for a faster adaptation as sucrose-supplemented cultures grew faster during early logarithmic growth. No significant difference was observed for the *hexR* mutant grown on glucose- or sucrose-supplemented medium. Arabinose allowed for the most efficient growth independent of the genotypes studied. These results indicated a potential HexR-mediated regulatory link of intra-cellular hexose metabolism with synthesis of Lsc since sucrose is the source of glucose via the enzymatic activity of the later enzyme.

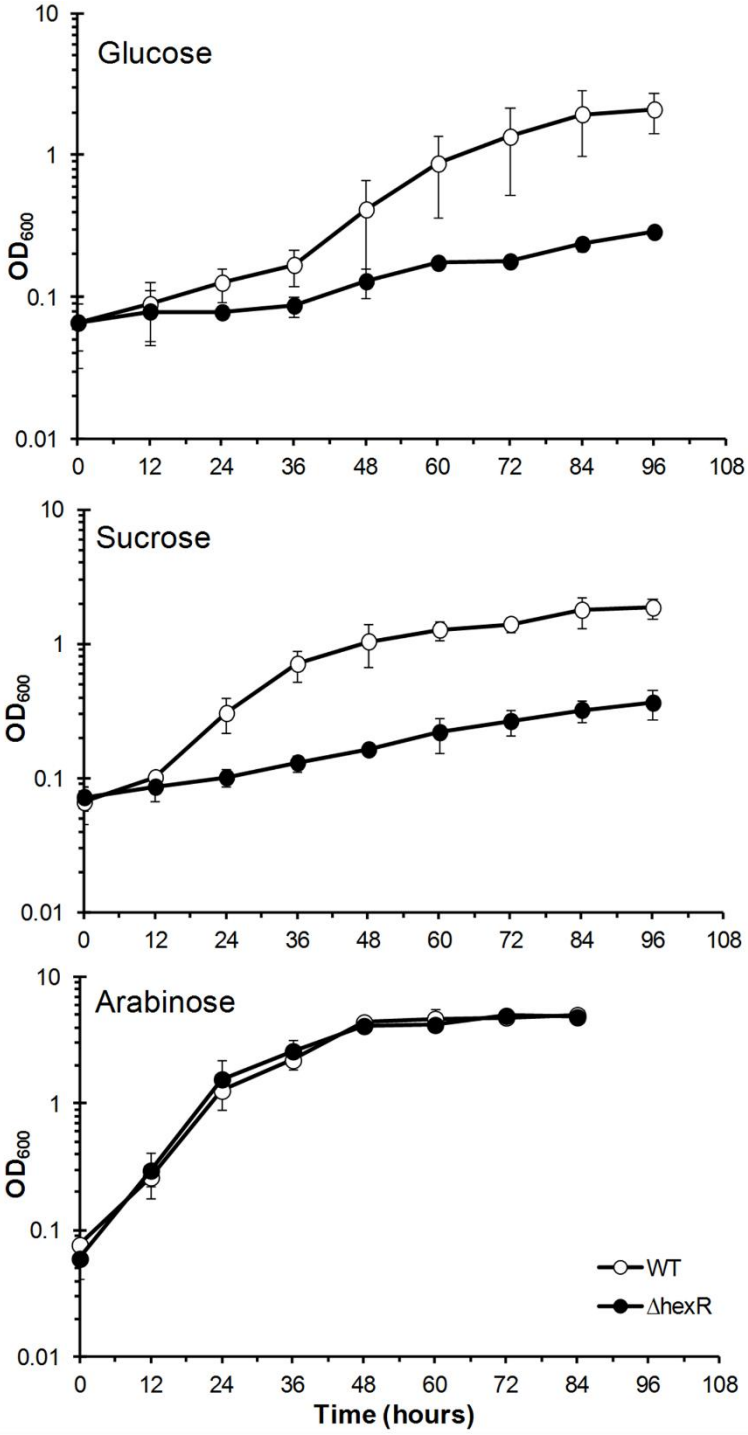


Figure 2. Growth curve of PG4180 wild type (WT) (o) and its *hexR* mutant (●). Cultures are grown in HS minimal media supplemented with glucose, sucrose, or arabinose as sole carbon source at 18°C. Error bars represent standard deviation of the mean of three biological replicates (n=3).

To analyze the effect of a *hexR* mutation on growth *in planta*, PG4180 wild type and its *hexR* mutant were spray-inoculated onto soybean leaves with suspensions adjusted to 1 x 10^7 CFU/ml. Subsequently, the plants were kept in a plant growth chamber and bacteria were recovered between days 1 to 14 post inoculation (**Fig. 3**). Results showed that the total population number of bacteria was not significantly different between plants inoculated with PG4180 wild type or its *hexR* mutant. However, a clear trend was observed for the percentage of bacteria that entered the interior of the leaf tissue. Except for days 1 and 9, plants inoculated with PG4180 wild type showed a higher percentage of internalized bacterial population when compared to plants subjected to the *hexR* mutant. These results might indicate an important role of HexR for the *in planta* fitness of *P. syringae*.

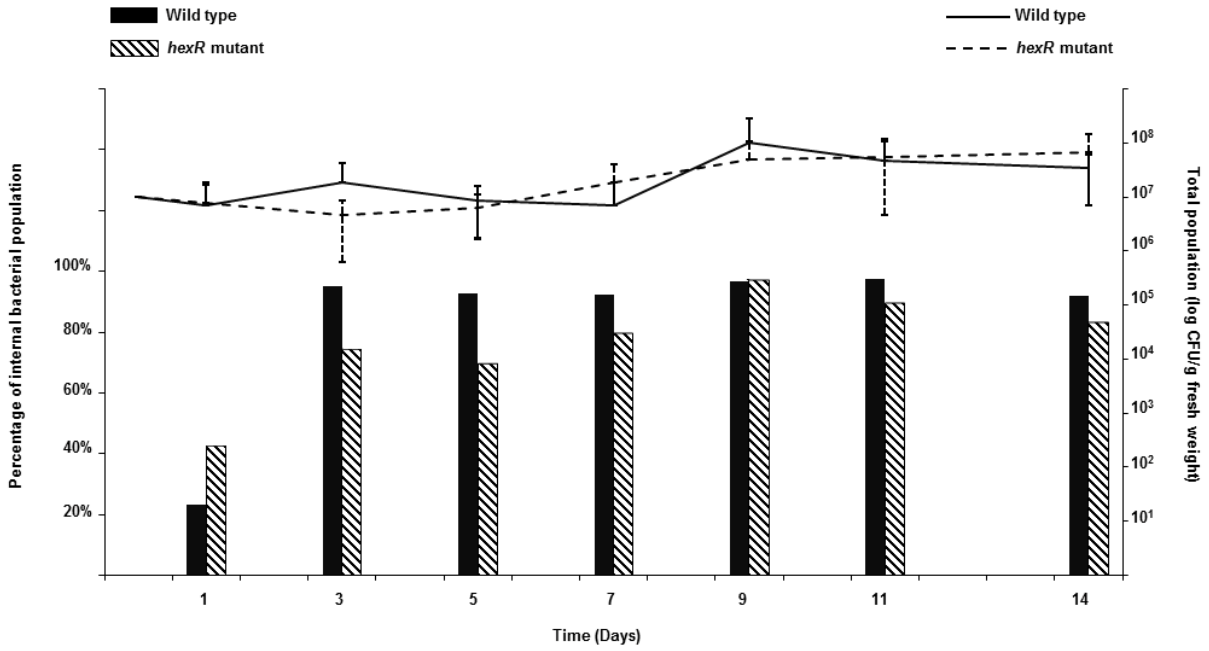


Figure 3. Total population of PG4180 wild type (——) and its *hexR* mutant (- - - - -) in soybean leaves. Columns represent the percentage of internalized PG4180 wild type (Black) and its *hexR* mutant (Black stripes). Bacterial suspensions were spray-inoculated on leaves of soybean plants grown in a greenhouse at 19-21°C. Data represent the mean values from five independent experiments with each two leaf samples. Error bars represent standard deviation of the mean of five biological replicates (n=5).

Transcriptional analyses of *lsc* genes

To analyze the effect of the *hexR* mutation on *lsc* gene expression, a growth-phase dependent transcriptional analysis was conducted using qRT-PCR. Cells of PG4180 wild type and the *hexR* mutant were grown in minimal medium containing arabinose at 18°C since levan production was shown to be maximal at this temperature (Li *et al.*, 2006). Transcription of *lsc* genes in wild type and *hexR* mutant was highest during the early exponential growth phase and significantly decreased during further growth (**Fig. 4**). Expression of *lsc* in the *hexR* mutant showed a tendency of being higher than the wild type in the early and mid-logarithmic growth phase. Interestingly, a significantly higher expression ($p < 0.01$) of *lsc* was observed in the *hexR* mutant as compared to that of the wild type at an OD_{600} of 2.0 referring to late-logarithmic to stationary phase.

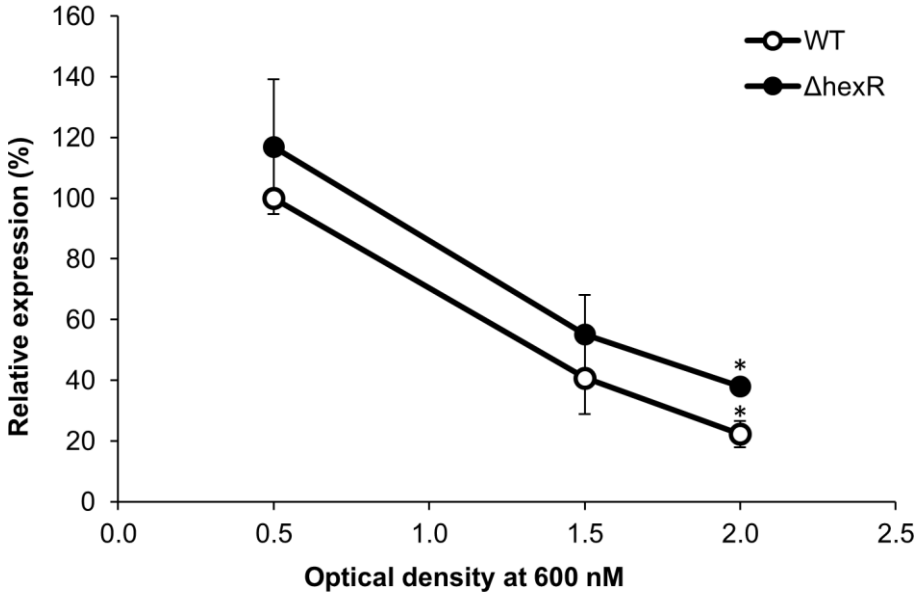


Figure 4. Quantitative Reverse Transcriptase PCR analysis of growth phase-dependent *lsc* gene expression. PG4180 and its *hexR* mutant were grown at 18°C in HS + arabinose. Relative mRNA levels were related to the mean value determined for the signals of PG4180 wild type at an OD_{600} of 0.5, which was defined as 100%. Data show the means and standard errors of three biological replicates (n=3) (*=P<0,005).

Abundance of Lsc in PG4180 wild type and *hexR* mutant

To qualitatively assess levan formation, PG4180 wild type and its *hexR* mutant were grown on sucrose-containing MG agar plates resulting in indistinguishable levels of levan formation for both (Data not shown). The accumulation of Lsc in extracellular fractions of the wild type and the *hexR* mutant of PG4180 was tested using immunological detection. Protein samples were obtained from cell-free culture supernatants of bacterial cultures grown to late exponential phase at 18°C. Comparison of wild type and *hexR* mutant showed a slightly higher Lsc accumulation in the mutant's culture supernatant (**Fig. 5**). These results further supported the hypothesis that HexR might repress *lsc* gene expression, resulting in more extra-cellular accumulation of its gene product in the *hexR* mutant.

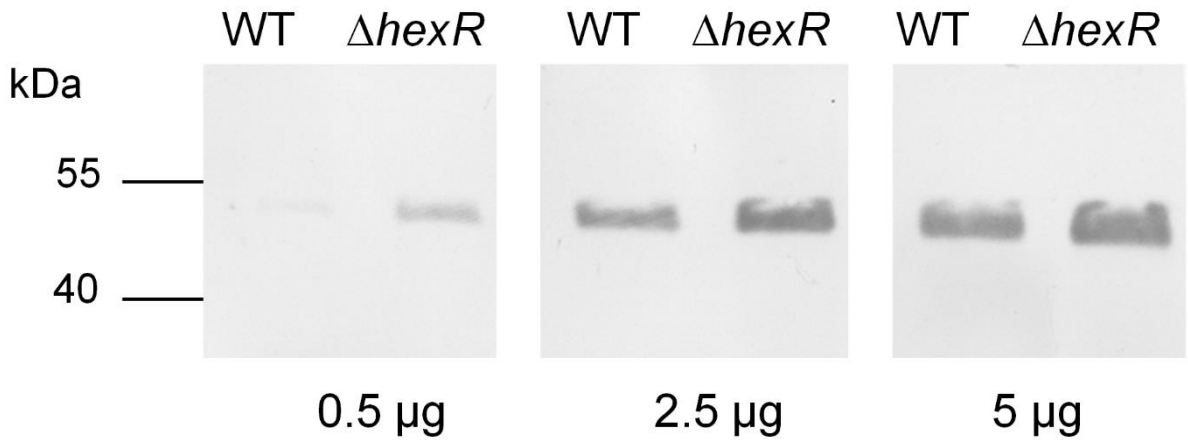


Figure 5. Qualitative Western blot analysis of extra-cellular Lsc in cell-free supernatant of PG4180 wild type (WT) and its *hexR* mutant. Cultures are grown in HS medium supplemented with arabinose as sole carbon source at 18°C. 0.5, 2.5, 5 µg of protein samples per lane were electrophoretically separated, transferred to a polyvinylidene fluoride membrane, and hybridized with Lsc-specific polyclonal antibodies.

Lack of an effect of HexR on protein secretion and hypersensitive response

To further investigate the decreased *in planta* fitness of the *hexR* mutant of *P. syringae* as compared to its wild type, hypersensitive response (HR) assays were performed for both on the tobacco plant *Nicotiana tabacum* (**Fig. 6**). The resulting HR after 24 hours on leaves inoculated with the mutant was indistinguishable from that induced by the wild type.

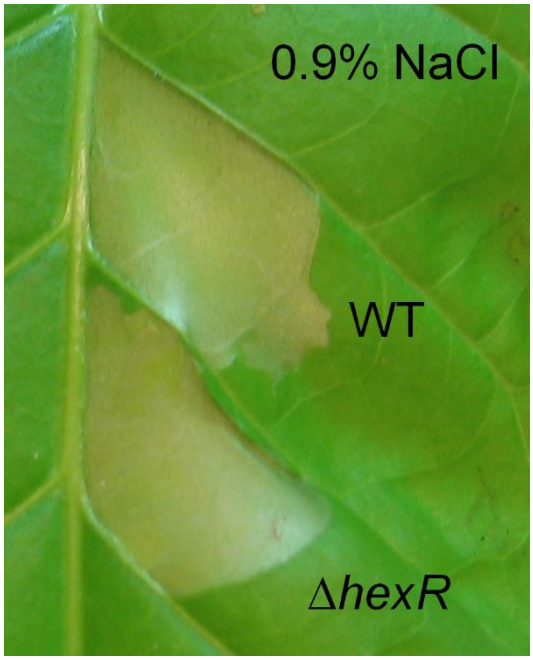


Figure 6. **Hypersensitive response assay on tobacco.** Typical hypersensitive response reactions elicited on tobacco plants as non-host defense responses by PG4180 wild type and its *hexR* mutant. Sterile 0.9% sodium chloride solution was used as a negative control.

In addition, extra-cellular protein profiles were determined for PG4180 wild type and its *hexR* mutant incubated in *hrp* gene-inducing IM medium or in HS+ glucose medium (**Fig. 7**). After electrophoretic separation of extra-cellular proteins, nearly identical proteins profiles were observed for PG4180 wild type and its *hexR* mutant. In summary, these results indicated that a mutation of *hexR* does not influence the ability of *P. syringae* to induce an HR or alter its protein secretion pattern.

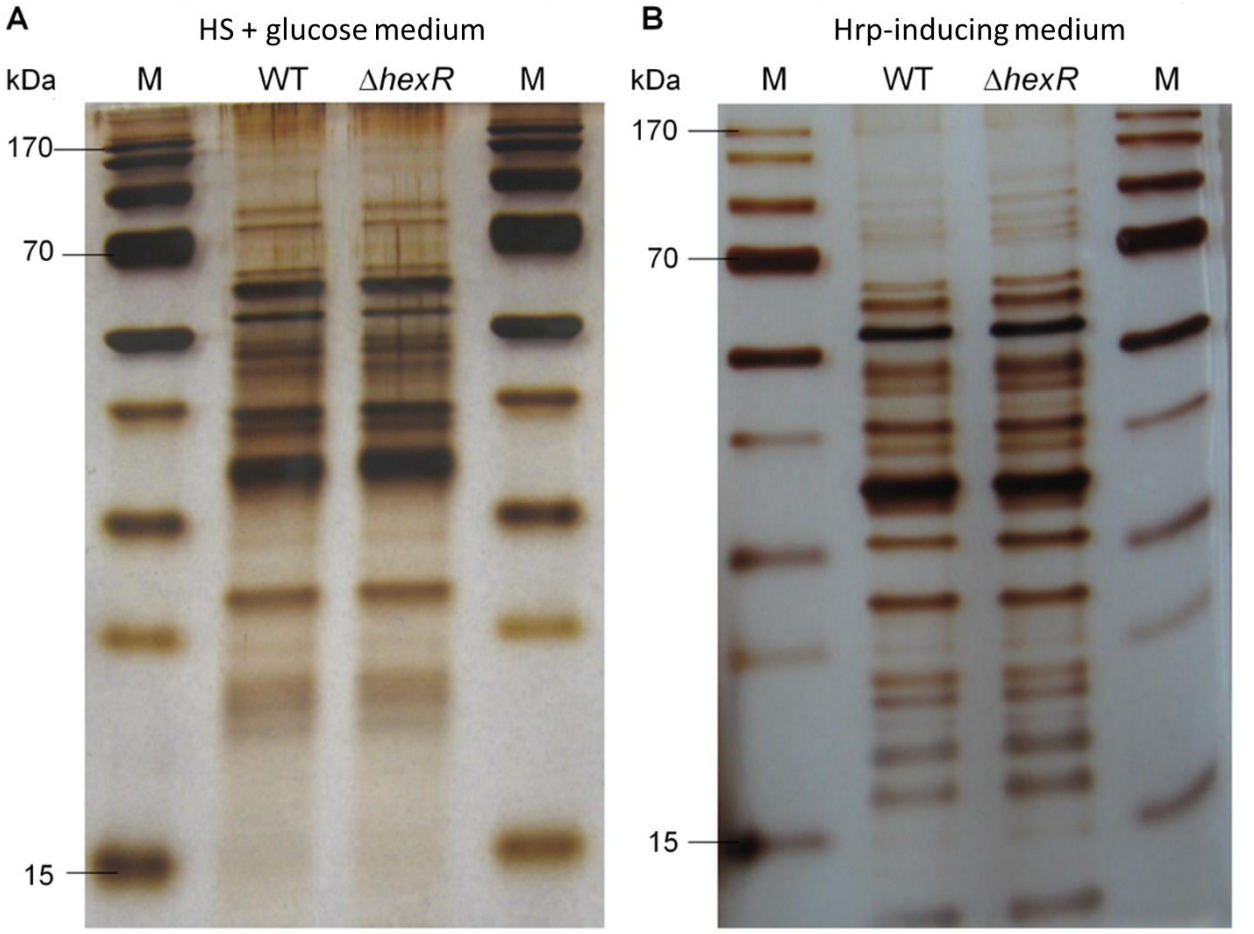


Figure 7. Extra-cellular protein profiles for PG4180 wild type and its *hexR* mutant. Bacterial cultures grown in **(A)** Hrp-inducing medium or in **(B)** HS+ glucose medium. Protein samples were separated by 12.5 % SDS-PAGE and subsequently visualized via silver staining.

Discussion

This study revealed that expression of genes encoding for an extra-cellular protein appear to be co-regulated with genes required for central hexose metabolism in a Gram-negative bacterium. Complementing previous studies on the global hexose metabolism repressor, HexR, in *P. putida* (Petruschka *et al.*, 2002, del Castillo *et al.*, 2008, Kim *et al.*, 2008, Daddaoua *et al.*, 2009), our results suggested that involvement of HexR in regulation of *lsc* expression might be a selective adaptation of the plant pathogen, *P. syringae*, to its well-studied infection cycle (Dulla *et al.*, 2005, Morris *et al.*, 2008). Once Lsc is secreted, cellular resources needed for its synthesis such as amino acyl residues are not available for the cell anymore. Consequently, *P. syringae* might repress Lsc synthesis in coordination with hexose utilization when sufficient levels of intra-cellular glucose are available to balance the cell's energy demands.

In close proximity and upstream of the translation start site (TSS) of *P. syringae lsc* genes, palindromic sequences were identified, which resemble HexR binding sites previously predicted for *P. putida* (Daddaoua *et al.*, 2009). Repressors such as HexR were suggested to bind to inverted repeats that partially or fully overlap RNA-polymerase binding sites (Rojo, 1999).

Sucrose is the most abundant plant storage sugar (Mercier & Lindow, 2000). Bacterial *in planta* and *in vitro* growth analyses indicated that the substrate of Lsc, sucrose, and in consequence its enzymatic product, glucose, seem to be major nutrient sources for *P. syringae* during *in planta* growth. This is in line with previous findings reporting high molecular abundances of sucrose in bean plant's apoplastic fluids (Atkinson & Baker, 1987). Consequently, expression and secretion of Lsc might be a fitness factor for the *in planta* life of *P. syringae*. Along the course of the experiment, the total populations of PG4180 wild type and its *hexR* mutant varied but were not significantly different from each other. This variation reflects a feature that is usually observed during assessment of bacterial populations (Mercier & Lindow, 2000). Kinkel *et al.* (1995) showed that two leaves of the same plant species might vary by over 10-fold in their total epiphytic bacteria. Moreover, a study by Hirano and Upper (1990) also demonstrated that regardless of the geographic area, plant species, or time scale tested, variations in population sizes of *P. syringae* are common. Many factors such as availability of nutrients, bacterial immigration (Kinkel *et al.*, 1995), and

ability of bacteria to tolerate environmental stresses on leaf surfaces (Beattie & Lindow, 1995) contribute to the observed variations.

A tendency is seen regarding the percentage of internalized bacteria in plants, although it is not significantly different. PG4180 wild type shows higher percentage of internalized bacteria as opposed to those inoculated with the hexR mutant. On days 3, 5, and 7 post inoculation, the wild type had 20, 23 and 12 % more internalized bacteria, respectively, than the *hexR* mutant. The high local abundance of sugars, mainly sucrose and glucose, in the apoplast of leave tissue (Benkeblia *et al.*, 2007) could potentially be responsible for the weaker multiplication of the *hexR* mutant inside the plant tissue. Consequently, it is hypothesized that HexR might play an important role in survival of PG4180 inside soybean leaves. In turn, a higher percentage of external *hexR* mutant cells were surviving on the leaf surface (data not shown). Although our study did not provide direct evidence for this, an increase in levan formation by the *hexR* mutation might have protected the mutant from damage by UV light or desiccation leading to higher survival rates on the leaf surface.

Finally, one may hypothesize that production of the levan exopolymer would rather be a 'shunt' product during the release of glucose from sucrose which, in turn, could be the actual major function of Lsc in the sugar metabolic pathway (**Fig. 8**). Genes encoding Lsc might be part of the HexR regulon of *P. syringae* in contrast to the situation in *P. putida* (del Castillo *et al.*, 2008), which is neither phytopathogenic nor harboring any *lsc* genes. However, other bacterial species, which possess similar enzymes for cleavage of sucrose to obtain readily usable glucose, could show a similar HexR-mediated regulation. Therefore, it is suggestive to screen the most important oral cavity inhabiting bacterial species (Bergeron & Burne, 2001) as well as bacteria, which cause mucus formation in sucrose-based food manufacturing (Bekers *et al.*, 2003) or bio-fuel production (Lee *et al.*, 1980) for presence of this regulatory linkage. The use of arabinose as sole carbon source had no effect on the growth phenotype of the *hexR* mutant. This was not surprising since the assimilatory pathways of glucose and arabinose are independent. L-arabinose is converted to α–keto-glutarate in *Pseudomonas* which can directly be utilized in the Tricarboxylic acid cycle (TCA cycle) independent of HexR regulation (Weimberg & Doudoroff, 1955, Palleroni *et al.*, 1956).

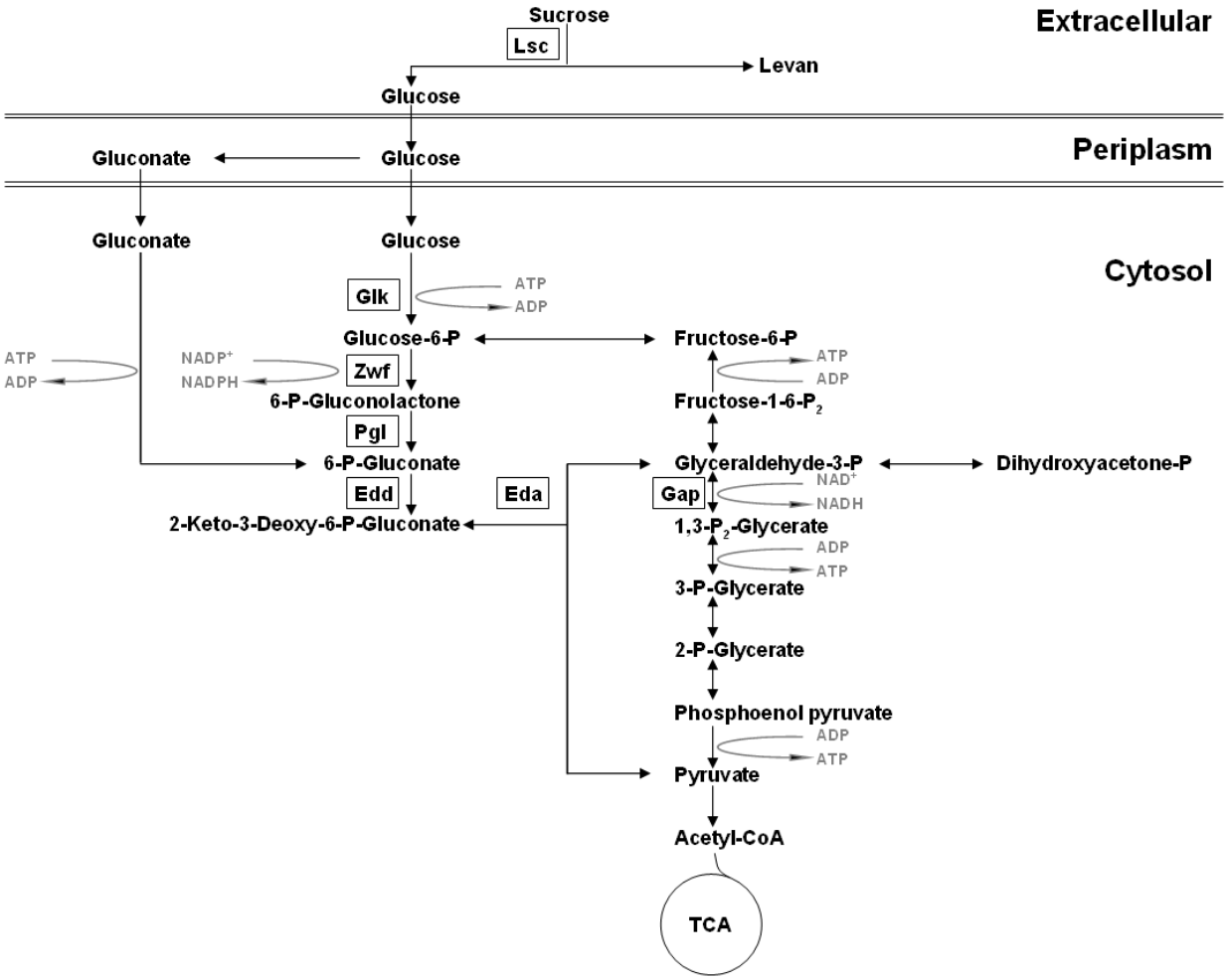


Figure 8. Schematic presentation of putative sucrose utilization pathway in *P. syringae* PG4180. Enzymes shown in blocks are presumed to be repressed by HexR in *P. putida* (del Castillo *et al.*, 2008, Daddaoua *et al.*, 2009) or in *P. syringae* (present study). Lsc, levansucrase; Glk, glucose kinase; Zwf, glucose-6-phosphate dehydrogenase; Pgl, 6-phosphogluconolactonase; Edd, 6-phosphogluconate dehydratase; Eda, 2-keto-3-deoxy-6-phosphogluconate aldolase; Gap, glyceraldehyde 3-phosphate-dehydrogenase; TCA, tricarboxylic acid cycle.

Currently, there is no plausible explanation for the lack of significant growth of the *hexR* mutant when supplemented with glucose or sucrose as sole carbon source. However, one might speculate on an up-regulation of genes normally repressed by HexR such as *glk*, *zwf-1*, *pgl*, *edd*, *eda*, and *gap-1* as previously shown in *P. putida* (del Castillo *et al.*, 2008). The effect could be potentiated in *P. syringae* by an increased expression of *lsc*, whose gene product in turn provides even more glucose. Derepressed glucose consumption in the *hexR* mutant might cause a surplus production of

NADPH, NADH, and ATP (**Fig. 8**). In turn, this may lead to an imbalance of cellular redox homeostasis thus alleviating cellular 'reductive stress' or even inducing 'energy spilling', respectively.

In *E. coli*, the redox potential influences the synthesis of fermentation products, which are formed to recycle and reoxidize NADH (Berrios-Rivera *et al.*, 2002). When *E. coli* cells are aerobically challenged with high glucose concentrations, they undergo a so-called 'acetate switch', which decelerates growth (Wolfe, 2005). A metabolic flux analysis of *E. coli* predicted that excess of carbon and energy might cause over-flow metabolism, which results in less efficient carbon utilization and decreased growth (Schuetz *et al.*, 2007). In *Streptococcus bovis*, excess ATP generation can cause 'energy spilling' by futile cycling of protons through the membrane, which leads to lesser biomass production (Bond & Russell, 2000, Russell, 2007). Whether reductive stress or 'energy spilling' take place in a glucose-exposed *hexR* mutant of *P. syringae* remains to be analyzed in future studies.

Expression of *lsc* genes was higher in the *hexR* mutant as compared to the wild type during late logarithmic growth. This result is in accordance with previous results of micro-array analyses of glucose metabolic genes in *P. putida* KT2440 where genes *glk* and *zwf-1* showed a ~two-fold increased expression while genes *pgl, edd, eda,* and *gap-1* exhibited a four- to six-fold increased expression in a respective *hexR* mutant (del Castillo *et al.*, 2008). Why *lsc* genes were only moderately up-regulated in the *hexR* mutant of *P. syringae* might be explained by the peripheral role of these genes in glucose metabolism.

The HR test is a classical assay to qualitatively show pathogenicity of a plant-associated microbe. The extracellular protein profiles and HR assay conducted with the wild type and the *hexR* mutant, respectively, suggested that HexR does not influence *P. syringae*'s ability to cause a HR on non-host plants. Furthermore, our results suggested that the secretion of *hrp*-associated proteins was not affected when *hexR* was mutated. It is therefore tempting to speculate that the significantly reduced ability of the *hexR* mutant to survive inside of the plant is not due to altered *hrp* gene expression but is rather due to a distorted sugar metabolism.

Data of this study prompt the question whether HexR-controlled genes such as *edd*, *eda, glk, pgl, zwf*-1, or *gap*-1 (del Castillo *et al.*, 2008) are indeed co-regulated with *lsc* genes in *P. syringae*. The

current study revealed exciting options for an in-depth analysis of intra-cellular and extra-cellular hexose metabolism in the plant pathogen *P. syringae* and may allow us to better understand the potentially complex interplay of factors and parameters contributing to epiphytic or pathogenic behavior of this organism, respectively.

Acknowledgments

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (UL 169/5-1). The authors are grateful to Marc Auchter, Ramesh Mavathur, Petra Dangel, Helge Weingart, and Georgi Muskhelishvili for excellent technical assistance and valuable scientific advises.

References

Atkinson, M.M. & C.J. Baker, (1987) Alteration of plasmalemma sucrose transport in Phaseolus vulgaris by *Pseudomonas syringae* pv. syringae and its association with K+/H+ exchange. *Phytopathology* **77**: 1573-1578.

Beattie, G.A. & S.E. Lindow, (1995) The secret life of foliar bacterial pathogens on leaves. *Annual Review of Phytopathology* **33**: 145-172.

Bekers, M., D. Upite, E. Kaminska, M. Grube, J. Laukevics, I. Vina, A. Vigants & P. Zikmanis, (2003) Fructan Biosynthesis by Intra- and Extracellular Zymomonas mobilis Levansucrase after Simultaneous Production of Ethanol and Levan. *Acta Biotechnologica* **23**: 85-93.

Bender, C.L., H. Liyanage, D. Palmer, M. Ullrich, S. Young & R. Mitchell, (1993) Characterization of the genes controlling the biosynthesis of the polyketide phytotoxin coronatine including conjugation between coronafacic and coronamic acid. *Gene* **133**: 31-38.

Benkeblia, N., T. Shinano & M. Osaki, (2007) Metabolite profiling and assessment of metabolome compartmentation of soybean leaves using non-aqueous fractionation and GC-MS analysis. *Metabolomics* **3**: 297-305.

Bergeron, L.J. & R.A. Burne, (2001) Roles of fructosyltransferase and levanase-sucrase of Actinomyces naeslundii in fructan and sucrose metabolism. *Infection and Immunity* **69**: 5395-5402.

Berrios-Rivera, S.J., G.N. Bennett & K.Y. San, (2002) Metabolic engineering of Escherichia coli: increase of NADH availability by overexpressing an NAD(+)-dependent formate dehydrogenase. *Metabolic Engineering* **4**: 217-229.

Biemelt, S. & U. Sonnewald, (2006) Plant-microbe interactions to probe regulation of plant carbon metabolism. *Journal of Plant Physiology* **163**: 307-318.

Bond, D.R. & J.B. Russell, (2000) Protonmotive force regulates the membrane conductance of Streptococcus bovis in a non-ohmic fashion. *Microbiology* **146 (Pt 3)**: 687-694.

Budde, I.P. & M.S. Ullrich, (2000) Interactions of Pseudomonas syringae pv. glycinea with host and nonhost plants in relation to temperature and phytotoxin synthesis. *Molecular Plant-Microbe Interactions : MPMI* **13**: 951-961.

Choi, K.H., T. Mima, Y. Casart, D. Rholl, A. Kumar, I.R. Beacham & H.P. Schweizer, (2008) Genetic tools for select-agent-compliant manipulation of Burkholderia pseudomallei. *Applied and Environmental Microbiology* **74**: 1064-1075.

Conway, T., (1992) The Entner-Doudoroff pathway: history, physiology and molecular biology. *FEMS Microbiology Reviews* **9**: 1-27.

Daddaoua, A., T. Krell & J.L. Ramos, (2009) Regulation of glucose metabolism in Pseudomonas: the phosphorylative branch and entner-doudoroff enzymes are regulated by a repressor containing a sugar isomerase domain. *The Journal of Biological Chemistry* **284**: 21360-21368.

del Castillo, T., E. Duque & J.L. Ramos, (2008) A set of activators and repressors control peripheral glucose pathways in Pseudomonas putida to yield a common central intermediate. *Journal of Bacteriology* **190**: 2331-2339.

Dulla, G., M. Marco, B. Quinones & S. Lindow, (2005) A closer look at Pseudomonas syringae as a leaf colonist—the pathogen P. syringae thrives on healthy plants by employing quorum sensing, virulence factors, and other traits. *ASM News* **71**: 469–475.

Dunleavy, J.M., (1988) Bacterial, fungal, and viral diseases affecting soybean leaves. In Soybean Diseases of the North Central Region *American Phytopathological Society*: 40-46.

Entner, N. & M. Doudoroff, (1952) Glucose and gluconic acid oxidation of Pseudomonas saccharophila. *The Journal of Biological Chemistry* **196**: 853-862.

Figurski, D.H. & D.R. Helinski, (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proceedings of the National Academy of Sciences of the United States of America* **76**: 1648-1652.

Finn, R.D., J. Mistry, J. Tate, P. Coggill, A. Heger, J.E. Pollington, O.L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E.L. Sonnhammer, S.R. Eddy & A. Bateman, (2010) The Pfam protein families database. *Nucleic Acids Research* **38**: D211-222.

He, S.Y., H.-C. Huang & A. Collmer, (1993) Pseudomonas syringae pv. syringae harpinPss: A protein that is secreted via the hrp pathway and elicits the hypersensitive response in plants. *Cell* **73**: 1255-1266.

Hettwer, U., F.R. Jaeckel, J. Boch, M. Meyer, K. Rudolph & M.S. Ullrich, (1998) Cloning, nucleotide sequence, and expression in Escherichia coli of levansucrase genes from the plant pathogens Pseudomonas syringae pv. glycinea and P. syringae pv. phaseolicola. *Applied and Environmental Microbiology* **64**: 3180-3187.

Hirano, S.S. & C.D. Upper, (1990) Population Biology and Epidemiology of Pseudomonas-Syringae. *Annual Review of Phytopathology* **28**: 155-177.

Hoang, T.T., R.R. Karkhoff-Schweizer, A.J. Kutchma & H.P. Schweizer, (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.

Huynh, T., D. Dahlbeck & B. Staskawicz, (1989) Bacterial blight of soybean: regulation of a pathogen gene determining host cultivar specificity. *Science* **245**: 1374-1377.

Keane, P., A. Kerr & P. New, (1970) Crown Gall of Stone Fruit II. Identification and Nomenclature of Agrobacterium Isolates. *Australian Journal of Biological Sciences* **23**: 585-596.

Khandekar, S., A. Srivastava, D. Pletzer, A. Stahl & M.S. Ullrich, (2014) The conserved upstream region of lscB/C determines expression of different levansucrase genes in plant pathogen Pseudomonas syringae. *BMC Microbiology* **14**: 79.

Kim, J., C.O. Jeon & W. Park, (2008) Dual regulation of zwf-1 by both 2-keto-3-deoxy-6-phosphogluconate and oxidative stress in Pseudomonas putida. *Microbiology (Reading, England)* **154**: 3905-3916.

King, E.O., M.K. Ward & D.E. Raney, (1954) Two simple media for the demonstration of pyocyanin and fluorescin. *The Journal of Laboratory and Clinical Medicine* **44**: 301-307.

Kinkel, L.L., M. Wilson & S.E. Lindow, (1995) Effect of sampling scale on the assessment of epiphytic bacterial populations. *Microbial Ecology* **29**: 283-297.

Kovach, M.E., P.H. Elzer, D.S. Hill, G.T. Robertson, M.A. Farris, R.M. Roop, 2nd & K.M. Peterson, (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.

Kovach, M.E., R.W. Phillips, P.H. Elzer, R.M. Roop, 2nd & K.M. Peterson, (1994) pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* **16**: 800-802.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson & D.G. Higgins, (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Lee, K.J., M. Lefebvre, D.E. Tribe & P.L. Rogers, (1980) High productivity ethanol fermentations with Zymomonas mobilis using continuous cell recycle. *Biotechnology Letters* **2**: 487-492.

Lessie, T.G. & P.V. Phibbs, (1984) Alternative Pathways of Carbohydrate Utilization in Pseudomonads. *Annual Review of Microbiology* **38**: 359-388.

Leyn, S.A., X. Li, Q. Zheng, P.S. Novichkov, S. Reed, M.F. Romine, J.K. Fredrickson, C. Yang, A.L. Osterman & D.A. Rodionov, (2011) Control of proteobacterial central carbon metabolism by the HexR transcriptional regulator: a case study in Shewanella oneidensis. *The Journal of Biological Chemistry* **286**: 35782-35794.

Li, H., A. Schenk, A. Srivastava, D. Zhurina & M.S. Ullrich, (2006) Thermo-responsive expression and differential secretion of the extracellular enzyme levansucrase in the plant pathogenic bacterium Pseudomonas syringae pv. glycinea. *FEMS Microbiology Letters* **265**: 178-185.

Li, H. & M.S. Ullrich, (2001) Characterization and mutational analysis of three allelic lsc genes encoding levansucrase in Pseudomonas syringae. *Journal of Bacteriology* **183**: 3282-3292.

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

Mansfield, J.W., (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: an overview of 25 years of progress in our understanding of plant innate immunity. *Molecular Plant Pathology* **10**: 721-734.

Melotto, M., W. Underwood & S.Y. He, (2008) Role of stomata in plant innate immunity and foliar bacterial diseases. *Annual Review of Phytopathology* **46**: 101-122.

Mercier, J. & S.E. Lindow, (2000) Role of leaf surface sugars in colonization of plants by bacterial epiphytes. *Applied and Environmental Microbiology* **66**: 369-374.

Morris, C.E., D.C. Sands, B.A. Vinatzer, C. Glaux, C. Guilbaud, A. Buffiere, S. Yan, H. Dominguez & B.M. Thompson, (2008) The life history of the plant pathogen Pseudomonas syringae is linked to the water cycle. *The ISME journal* **2**: 321-334.

O'Brien R, D. & S.E. Lindow, (1988) Effect of Plant Species and Environmental Conditions on Ice Nucleation Activity of Pseudomonas syringae on Leaves. *Applied and Environmental Microbiology* **54**: 2281-2286.

Palleroni, N.J., R. Contopoulou & M. Doudoroff, (1956) Metabolism of carbohydrates by Pseudomonas saccharophila. II. Nature of the kinase reaction involving fructose. *Journal of Bacteriology* **71**: 202-207.

Palmer, D.A. & C.L. Bender, (1993) Effects of Environmental and Nutritional Factors on Production of the Polyketide Phytotoxin Coronatine by Pseudomonas syringae pv. Glycinea. *Applied and Environmental Microbiology* **59**: 1619-1626.

Papp, B., B. Teusink & R.A. Notebaart, (2009) A critical view of metabolic network adaptations. *HFSP journal* **3**: 24-35.

Petruschka, L., K. Adolf, G. Burchhardt, J. Dernedde, J. Jurgensen & H. Herrmann, (2002) Analysis of the zwf-pgl-eda-operon in Pseudomonas putida strains H and KT2440. *FEMS Microbiology Letters* **215**: 89-95.

Portais, J.C. & A.M. Delort, (2002) Carbohydrate cycling in micro-organisms: what can (13)C-NMR tell us? FEMS Microbiology Review **4**:375-402.

Roitsch, T. & M.C. Gonzalez, (2004) Function and regulation of plant invertases: sweet sensations. *Trends in Plant Science* **9**: 606-613.

Rojo, F., (1999) Repression of transcription initiation in bacteria. *Journal of Bacteriology* **181**: 2987-2991.

Russell, J.B., (2007) The energy spilling reactions of bacteria and other organisms. *Journal of Molecular Microbiology and Biotechnology* **13**: 1-11.

Sambrook, J., E.F. Fritsch & T. Maniatis, (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Sambrook, J. & D.W. Russell, (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Schenk, A., H. Weingart & M.S. Ullrich, (2008) Extraction of high-quality bacterial RNA from infected leaf tissue for bacterial in planta gene expression analysis by multiplexed fluorescent Northern hybridization. *Molecular Plant Pathology* **9**: 227-235.

Schuetz, R., L. Kuepfer & U. Sauer, (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Molecular Systems Biology* **3**: 119.

Smirnova, A., H. Li, H. Weingart, S. Aufhammer, A. Burse, K. Finis, A. Schenk & M.S. Ullrich, (2001) Thermoregulated expression of virulence factors in plant-associated bacteria. *Archives of Microbiology* **176**: 393-399.

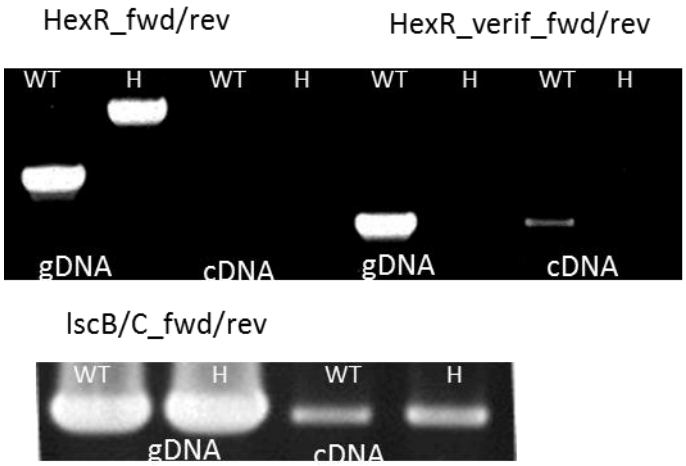
Srivastava, A., D. Zhurina & M.S. Ullrich, (2009) Levansucrase and Levan Formation in Pseudomonas syringae and Related. *Bacterial Polysaccharides: Current Innovations and Future Trends*: 213.

Weimberg, R. & M. Doudoroff, (1955) The oxidation of L-arabinose by Pseudomonas saccharophila. *The Journal of Biological Chemistry* **217**: 607-624.

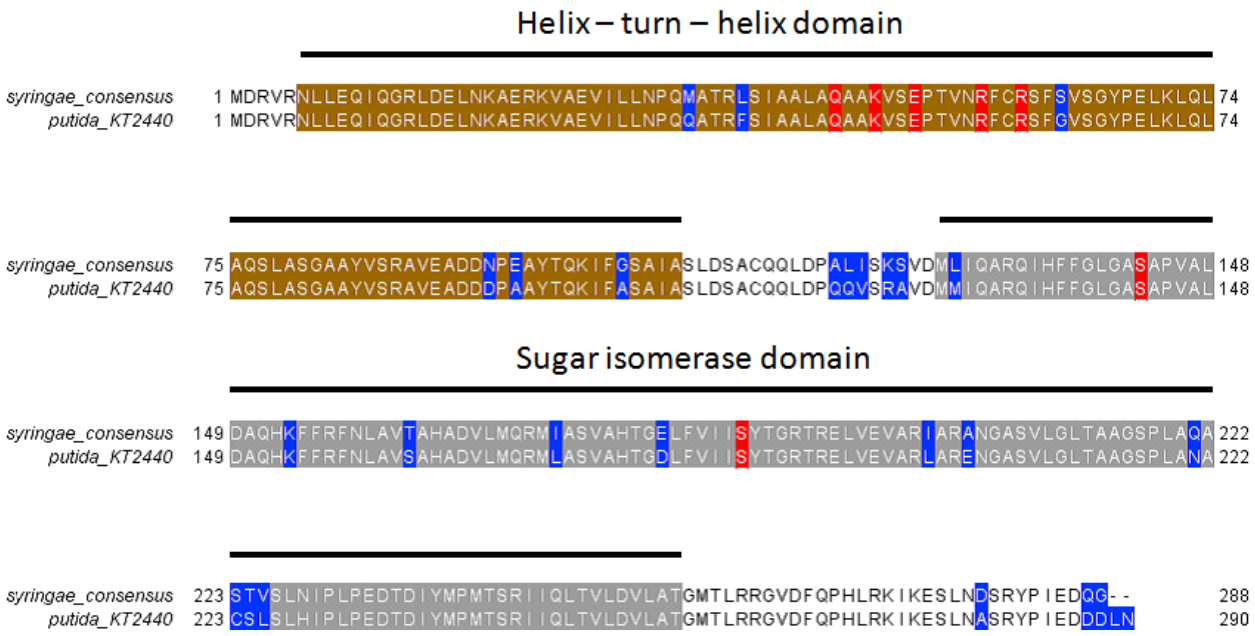
Wolfe, A.J., (2005) The acetate switch. *Microbiology and Molecular Biology Reviews : MMBR* **69**: 12-50.

Xie, C.H. & A. Yokota, (2005) Reclassification of Alcaligenes latus strains IAM 12599T and IAM 12664 and Pseudomonas saccharophila as Azohydromonas lata gen. nov., comb. nov., Azohydromonas australica sp. nov. and Pelomonas saccharophila gen. nov., comb. nov., respectively. *International Journal of Systematic and Evolutionary Microbiology*. **55**.

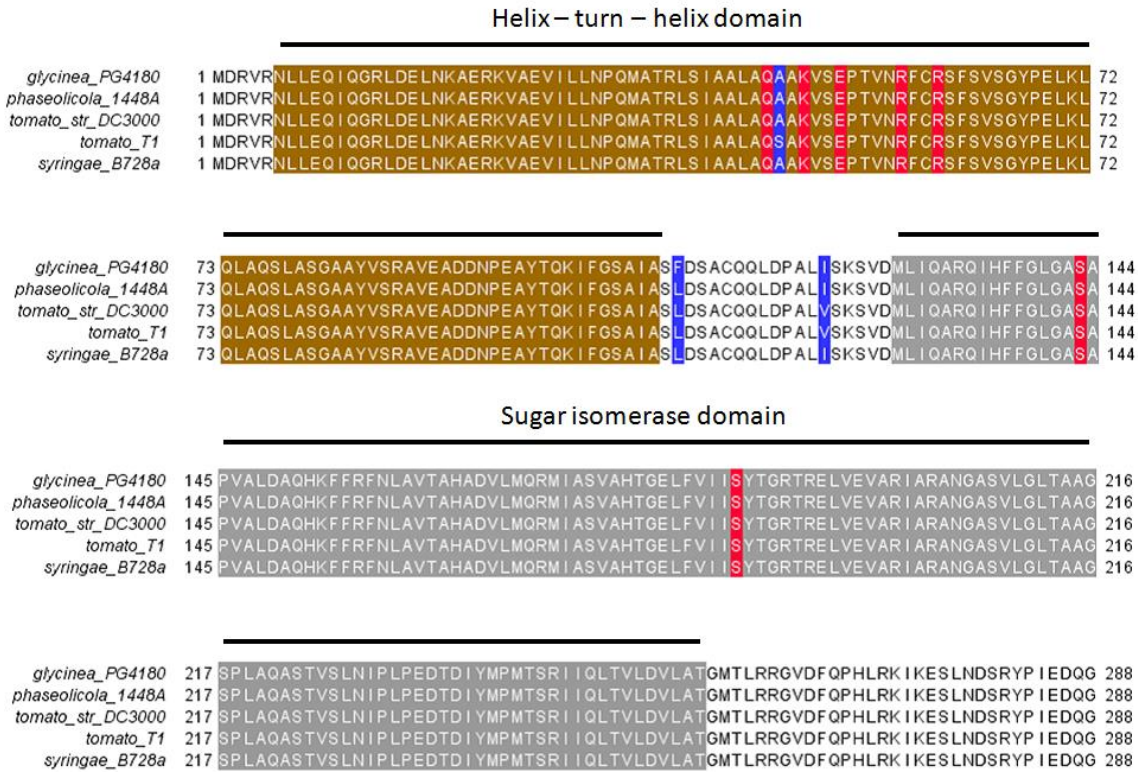
Supporting Information



Supporting Information Figure 1. Verification of *hexR* null mutant phenotype by PCR amplification. The total RNA was extracted by phenol chloroform method followed by cDNA generation. PCR amplification of *hexR* fragment on total cDNA and genomic DNA (gDNA) using *hexR* specific primers. The quality of total cDNA and genomic DNA were checked by performing PCR amplification of *lscB/C* gene which signified correct amplicon. H= *hexR* mutant, WT= PG4180 wild type.



Supporting Information Figure 2. Sequence alignment of a *P. syringae* HexR consensus sequence with that of *P. putida* KT2440 HexR (Daddaoua *et al.*, 2009). Mismatched residues are marked in blue. Residues marked in red are the predicted DNA recognition (Q46, K49, E52, R57, R60) and effector recognition (S143, S187) residues of HexR, respectively (Daddaoua *et al.*, 2009).



Supporting Information Figure 3. Multiple sequence alignment of HexR amino acid sequences from levan-producing *Pseudomonas syringae* strains using ClustalW (Larkin *et al.*, 2007). PFAM comparison revealed two domains: Helix-turn-helix domain from residues 6-108 (PFAM PF01418) shown in pale grey and SIS domain from residues 128-256 (PFAM PF01380) shown in dark grey. Mismatched residues are marked in blue. Residues marked in red are the predicted DNA recognition (Q46, K49, E52, R57, R60) and effector recognition (S143, S187) residues of HexR, respectively (Daddaoua *et al.*, 2009).

4 Discussion

Plant leaves are natural habitats which can be colonized by many microorganisms (Hirano & Upper, 2000). Among these microorganisms is *Pseudomonas syringae*, a plant-pathogenic bacterium that frequently forms large epiphytic populations on plant leaves (Yu *et al.*, 2013). The leaf surface and the apoplast are stressful environments for foliar bacterial pathogens, and therefore *P. syringae* must establish fitness strategies to adapt and survive the host environment (Rico *et al.*, 2011). These include a tolerance strategy to endure the environmental stresses on the leaf surface and an avoidance strategy to escape, modify or resist stress in the endophytic sites (Beattie & Lindow, 1995).

Role of bacteriophage-derived region in *P. syringae* pv. glycinea PG4180

In the present study, a prophage region inserted between *trpE* and *trpG* of *P. syringae* pv. glycinea PG4180 was identified. Majority of the recognized genes encode proteins involved in phage tail morphogenesis. However, genes that are essential for producing a functional bacteriophage were absent. Moreover, our results revealed that this region carries genes of a prophage origin that are essential for expression of levansucrase.

The fact that prophages can enhance bacterial pathogenesis and fitness has been studied in many plant pathogens such as *Pseudomonas* (Varani *et al.*, 2013). They can introduce additional genes, not essential for viral replication (Davies *et al.*, 2016) but beneficial for enhancing bacterial fitness in an environmental niche (Miao & Miller, 1999). A recent study by Hockett *et al.* (2015) revealed that the bacteriophage-derived region located between *trpE* and *trpG* encodes for a phage tail-like bacteriocin in *P. syringae* pv. glycinea B728a. Bacteriocins can be found in many proteo-bacterial genera, such as *P. putida* and *P. syringae*, and signify a second class of antibacterial defense weapons (Ghequire & De Mot, 2014). They are used to kill or inhibit species and strains closely related to the producing strain (Riley & Wertz, 2002) via depolarizing the bacterial membrane potential (Ge *et al.*, 2015). In order to understand why *P. syringae* pv. glycinea PG4180 spend energy on the production of bacteriocins, it is important to look back at the initial stages of plant-bacterial interaction, when bacteria reside on plant leaves. The leaf surface is a stressful environment for bacterial populations as they encounter frequent limitations in water and nutrient

availability, fluctuations in temperature, and exposure to solar irradiation (Monier & Lindow, 2003). Moreover, several studies showed that the composition and size of epiphytic populations can be largely affected by availability of nutrients (Wilson & Lindow, 1994, Mercier & Lindow, 2000b). Gene expression studies revealed that *P. syringae* can adapt to these environmental stresses by active exploration of the leaf surface via chemosensing, chemotaxis and flagellar motility (Yu *et al.*, 2013). However, *Pseudomonas* species are not the only colonizers of the leaf environment. Other δ-proteobacteria species isolated from soybean leaves include *Erwinia* sp., *Klebsiella pneumoniae, Enterobacteriaceae* and many others (Kuklinsky-Sobral *et al.*, 2004). All of the above mentioned factors drive intermicrobial competition for niches that are rich in resources such as water and nutrients (Wilson & Lindow, 1994, Lindow & Brandl, 2003). Therefore, bacteriocins may play a vital role in assisting *P. syringae* to resolve this competition between other microbes which might also explain the wide distribution of the prophage region in 54 out of 62 recently published draft genome sequences of *P. syringae* type and pathotype strains (Thakur *et al.*, 2016).

However, no killing activity was observed for the prophage region identified in *P. syringae* pv. glycinea PG4180. This can be explained by the lack of a right tester strain or because PG4180 doesn't produce an active syringacin from the prophage region that is located between *trpE* and *trpG* (Kevin Hockett – Personal communication). The latter hypothesis is supported by the fact that *P. syringae* pv. glycinea PG4180 lacks two genes encoding for chaperone and receptor binding protein that are essential for the production of an active R-type syringacin in *P. syringae* pv. syringae B728a (Hockett *et al.*, 2015). Interestingly, previous work by Hockett *et al.* (2015) showed that *P. syringae* pv. syringae B728a carry two additional non R-type bacteriocins. Thus we can speculate that *P. syringae* pv. glycinea PG4180 might harbor additional bacteriocins which can replace the defective one. To further investigate this, additional experiments and analysis are needed to identify other bacteriocins from *P. syringae* pv. glycinea PG4180. It is also important to obtain the full genome sequence of *P. syringae* pv. glycinea PG4180. Although this step might be costly, it will allow us to use the web-based programs BAGEL3 (van Heel *et al.*, 2013)and BLAST-P query to predict different classes of bacteriocins. An alternative method could be applied by inducing bacterial cultures with mitomycin C as reported by Hockett *et al.* (2015). However, this would require to identify a strain that exhibits unambiguous sensitivity to the putative PG4180 tailocins.

Molecular analysis of regulatory proteins involved in the expression of levansucrase in *P. syringae* pv. glycinea PG4180

Before going into details of this study and all the known factors affecting the expression of levansucrase, it is worthwhile to investigate the possible roles of levan in the infection process of PG4180. As an important prerequisite for disease development, *P. syringae* must establish a large epiphytic population on leaves of host plants to successfully invade into the leaf apoplast (Monier & Lindow, 2003). One of the mechanisms that allow plant pathogenic bacteria to survive the harsh and unstable environment on the leaf surface during early stages of tissue colonization is the production of exopolysaccharides (Xin & He, 2013). Since levan is the major exopolysaccharide in PG4180, its role to improve bacterial survival and growth on surfaces via the formation of biofilms was investigated. Formation of biofilms could be significant at some point as they might concentrate nutrients and provide protection from UV irradiation (Zago & Chugani, 2009). However, a study by Laue *et al.* (2006) showed that levan was not important for biofilm development. Hence what benefits do levan contribute for bacterial survival and fitness during the infection process?

It could thus be assumed that levan does not contribute mainly in the epiphytic stage of infection but it rather plays a major role when bacteria are in the plant apoplast. Global transcriptome profiling of *P. syringae* B728a showed that genes involved in the production and regulation of levan were five-fold more induced in the plant apoplast compared to its leaf surface (Yu *et al.*, 2013). Once inside the apoplast, bacteria will encounter a nutrient-deficient environment that is protected by a range of plant defenses which could alter the successful colonization of this niche (Li & Wang, 2012). Thus, one possible function of levan is that it can act as an energy storage complex to be utilized during starvation periods of *P. syringae* (Laue *et al.*, 2006). This is supported by the finding that sucrose assimilation was not induced in the plant apoplast and that *P. syringae* preferentially consumed other sugars and organic acids (Rico & Preston, 2008). Moreover, the role of levan in the survival of PG4180 was investigated by comparing the wild type to a *lscB/lscC* double mutant in terms of its *in planta* fitness. Results showed that the mutant ability to survive in the plant tissue was significantly impaired (Mehmood *et al.*, 2015) thus hinting to a possible additional role of levan in providing protection from plant defenses and reducing desiccation. The latter is supported by the fact that transcripts of genes required for synthesis and regulation of

exopolysaccharide were increased 3.8-fold by osmotic stress (Yu *et al.*, 2013). As an outlook, it would be interesting to perform transcriptional analysis of levansucrase genes using RNA extracted from *P. syringae* isolated from apoplast and compare it to those present on the leaf surface.

Prophage-borne transcriptional regulator LscR

Lsc of PG4180 is encoded by three alleles *lscA, lscB* and *lscC* (Li & Ullrich, 2001). Only *lscB* and *lscC* are expressed in *P. syringae* (Schenk *et al.*, 2006) while *lscA* is transcriptionally inactive as it lacks an upstream phage-associated promoter element (PAPE) (Srivastava *et al.*, 2012, Khandekar *et al.*, 2014). In the present study, the genome of *P. syringae* pv. glycinea PG4180 was screened for transcriptional regulators of levansucrase expression. Results indicated that a prophage-borne transcriptional regulator LscR induces expression of *lsc* by binding to its upstream region. To our knowledge, this is the first time an activator for Lsc expression is reported.

Although replication of prophage DNA might reduce the fitness of its host due to its high metabolic cost (Canchaya *et al.*, 2003a), they often also supply their hosts with important agents that are needed to survive in specific ecological niches (Srividhya *et al.*, 2007). Previous studies demonstrated that the upstream PAPE region of *lscB* and *lscC* in *P. syringae* is highly conserved and found to be present upstream of a gene encoding for glycosyl hydrolase (Srivastava *et al.*, 2012, Khandekar *et al.*, 2014). Interestingly, we found that the gene encoding for glycosyl hydrolase and its PAPE were located in close proximity to the *lscR* gene within a prophage-like DNA region known to encode for bacteriocins in B728a (Hockett *et al.*, 2015). In addition, we were able to confirm by DNA-binding assays and transcriptional analyses that both *lsc* and *glycosyl hydrolase* are regulated by LscR. All these findings strengthen the previously suggested hypothesis regarding evolution of Lsc in *P. syringae* (Srivastava *et al.*, 2012). Briefly, we speculate that the *lscA* gene could have been obtained from the plant pathogen *E. amylovora* but remained inactive as its native promoter was not suitable for expression in PG4180. Following infection by a bacteriophage, which carries a *glycosyl hydrolase* gene under control of LscR, PAPE was inserted upstream of *lscA* followed by gene duplication, resulting in *lscB* and *lscC.* These recombination events allowed PG4180 to control expression of levansucrase.

A PG4180 *lscR*-deficient mutant exhibited a levan-negative phenotype when grown on a sucrose-rich medium and was confirmed by zymographic analysis and Western blots. Consequently, we

can conclude that LscR mediates expression of levansucrase. But does LscR have other functions? Protein blast for LscR revealed ~99% protein identity to a gene in *P. syringae* B728a encoding for a transcriptional regulator PrtN, an activator of bacteriocins produced by *P. aeruginosa* upon DNA damage (Matsui *et al.*, 1993). Up to date, only four research papers were published regarding the role of PrtN and were limited to *P. aeruginosa* (Shinomiya *et al.*, 1983, Matsui *et al.*, 1993, Michel-Briand & Baysse, 2002, Wu & Jin, 2005). The function description in *P. syringae* B728a is based on functional predictions from Interpro (Winsor *et al.*, 2016). PrtN is known to bind a consensus sequence of a P box, ATTGnn(n)GTnn(n), that is repeated twice in R-type bacteriocins (Sano *et al.*, 1993). Analysis of the PAPE sequence showed that *P. syringae* might carry a P-box, however we found a major difference in the number of repeats between ATTG and GT. In *P. aeruginosa*, the nucleotide repeats were around 10 (Matsui *et al.*, 1993) while in PG4180 it was approximately 27 nucleotides (**Fig. 7**).

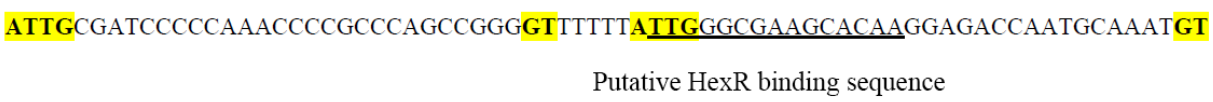


Figure 7. Nucleotide sequence of the upstream region of *lscB*. Sequence of a P box, ATTGnn(n)GTnn(n) are in bold and highlighted. Sequence contains a putative HexR binding site (underlined).

Interestingly, these repeats harbored a putative binding sequence for the protein HexR previously found to slightly repress levansucrase expression. Since the investigated prophage region from PG4180 lacks two genes essential for bacteriocin synthesis, it would be very difficult to test this hypothesis. Therefore, one possible suggestion is to knock out *lscR* gene in *P. syringae* B728a and test for bacteriocin synthesis. We might further identify additional functions for LscR by performing total proteome analysis of PG4180 wild type against its *lscR*-deficient mutant. Finally, it would be interesting to perform *in planta* assays to determine the effect of LscR on fitness and survival of epiphytic versus endophytic populations.

Unpublished data from the PhD thesis of Zhurina (2009) and Khandekar (2013) revealed the involvement of two histone-nucleoid structuring (H-NS)-like proteins, MvaT and MvaU, in expression of levansucrase. Although these data were not reproducible, it was believed that H-NS-like proteins contribute to the temperature-dependent repression of levansucrase. Li *et al.* (2006) presented evidence for the temperature-dependent expression and secretion of Lsc in *P. syringae*: maximum production occurred at 18°C, when the pathogen is highly virulent, and minimum production occurred at 28°C, the bacterium's optimal growth temperature. Moreover, another study by (Hockett *et al.*, 2013) showed that LscR and MvaU homologues in B728a are not thermo-regulated while MvaT was two-folds downregulated at 30 °C compared to 20 °C. Thus we can formulate the following hypothesis for regulation of levansucrase. At the bacterial optimum growth temperature, MvaU blocks transcription of *lsc* by binding to its upstream region and inhibits binding of LscR. Once temperature drops to 18 °C, more MvaT proteins are produced which in turn bind to MvaU and expose binding sites in the *lsc* upstream region for LscR (**Fig. 8**).

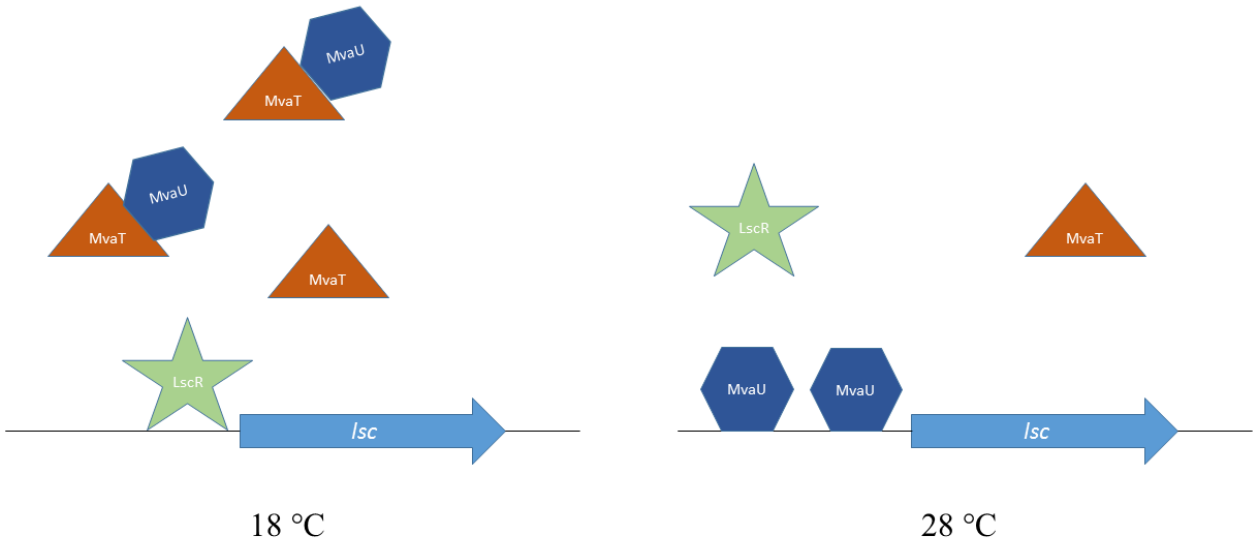


Figure 8: Schematic representation of the hypothesized mechanism for the temperature dependant regulation of *levansucrase* in *P. syringae*.

To further confirm this, we recommend performing EMSA again and use MvaT/MvaU to determine whether they compete over the binding site of LscR. Furthermore, we recommend to optimize the conditions for DNA affinity chromatography performed by Zhurina (2009) to reproduce the isolation of MvaT and MvaU. Finally, it is important to note that the *lscR*-deletion mutant generated by Zhurina (2009) and also analyzed by Khandekar (2013) was not a true mutant (data not shown). This explains their observation regarding absence of a levan-negative phenotype in PG4180 *lscR*-deficient mutant. In addition, qRT-PCR was applied to determine whether *lscR* expression was also thermo-dependent. Our results oppose those found by Zhurina (2009) as expression levels of *lscR* at 18 °C and 28 °C were not significantly different (**Fig. 9**).

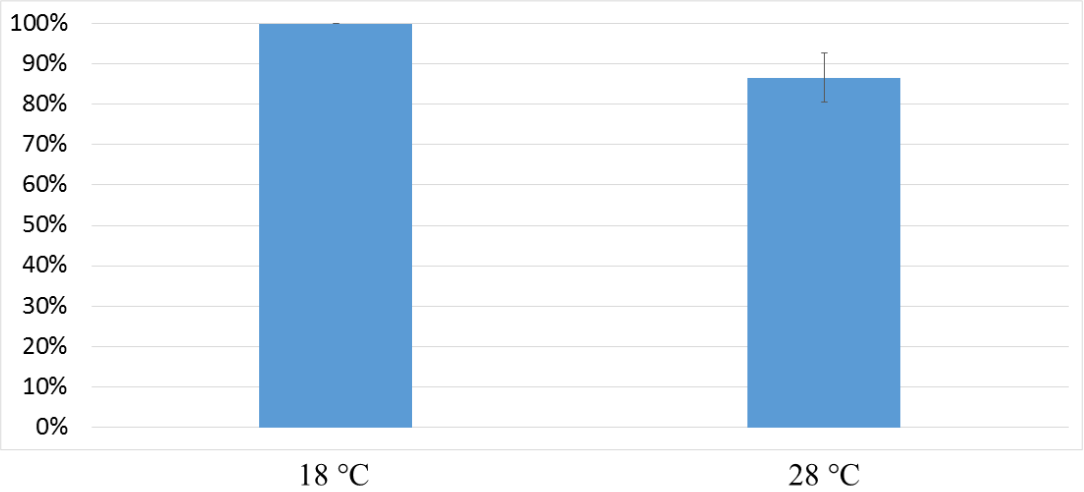


Figure 9: Quantitative Reverse Transcriptase PCR analysis of *lscR* gene expression in *P. syringae* pv. glycinea PG4180. Cells were grown at 18°C and 28°C in HS + arabinose. Relative mRNA levels were related to the mean value determined for the signals at 18°C of *P. syringae* wild type at an OD_{600} of 0.5, which was defined as 100%. Data shows the mean values and error bars represent standard deviation of the mean of three biological replicates (n=3). Statistically significant differences ($p < 0.005$) were calculated by one-way ANOVA-analysis and are marked with asterisks.

Hexose metabolism repressor, HexR

Bacterial communities growing epiphytically on plants are mainly affected by nutrient availability on leaf surfaces (Lindow & Brandl, 2003). For the glucose metabolism, pseudomonads utilize the Entner-Doudoroff (ED) pathway due to lack of 6-phosphofructokinase and hence do not catabolize sugars via the Embden-Meyerhof-Parnas pathway (Entner & Doudoroff, 1952, Portais & Delort, 2002). Our results show that the expression of genes encoding for an extra-cellular protein appears to be co-regulated with genes required for central hexose. Results also suggested that the involvement of HexR in regulation of *lsc* expression might be a selective adaptation of *P. syringae* (Morris *et al.*, 2008). Once Lsc is secreted, cellular resources needed for its synthesis such as amino acyl residues are not available for the cell anymore. Consequently, *P. syringae* might repress Lsc synthesis in coordination with hexose utilization when sufficient levels of intra-cellular glucose are available to balance the cell's energy demands. HexR might play a role in survival of PG4180 inside soybean leaves where sucrose is considered the most abundant plant storage sugar (Mercier & Lindow, 2000b). Although *in planta* assays did not show significant difference in the total bacterial populations between PG4180 wild type and its *hexR* mutant, there were a tendency for wild type endophytic bacteria to survive better than the *hexR* mutant. This could be explained by the fact that more levan is produced that can supply the invading bacteria with more protection from plant defenses or desiccation.

5 References

Ackermann, H.W., (2007) 5500 Phages examined in the electron microscope. *Archives of Virology* **152**: 227-243.

Agrios, G.N., (2009) Plant Pathogens and Disease: General Introduction. In: Encyclopedia of Microbiology (Third Edition). S. Editor-in-Chief: Moselio (ed). Oxford: Academic Press, pp. 613-646.

Azadbakht, L., R. Shakerhosseini, S. Atabak, M. Jamshidian, Y. Mehrabi & A. Esmaill-Zadeh, (2003) Beneficiary effect of dietary soy protein on lowering plasma levels of lipid and improving kidney function in type II diabetes with nephropathy. *European Journal of Clinical Nutrition* **57**: 1292-1294.

Beattie, G.A. & S.E. Lindow, (1995) The secret life of foliar bacterial pathogens on leaves. *Annual Review of Phytopathology* **33**: 145-172.

Canchaya, C., C. Proux, G. Fournous, A. Bruttin & H. Brussow, (2003a) Prophage genomics. *Microbiology and Molecular Biology Reviews: MMBR* **67**: 238-276, table of contents.

Canchaya, C., C. Proux, G. Fournous, A. Bruttin & H. Brüssow, (2003b) Prophage Genomics. *Microbiology and Molecular Biology Reviews* **67**: 238-276.

Casjens, S., (2003) Prophages and bacterial genomics: what have we learned so far? *Molecular Microbiology* **49**: 277-300.

Ceyssens, P.J. & R. Lavigne, (2010) Bacteriophages of Pseudomonas. *Future Microbiology* **5**: 1041-1055.

Daddaoua, A., T. Krell, C. Alfonso, B. Morel & J.L. Ramos, (2010) Compartmentalized glucose metabolism in Pseudomonas putida is controlled by the PtxS repressor. *Journal of Bacteriology* **192**: 4357-4366.

Daddaoua, A., T. Krell & J.L. Ramos, (2009) Regulation of glucose metabolism in Pseudomonas: the phosphorylative branch and entner-doudoroff enzymes are regulated by a repressor containing a sugar isomerase domain. *The Journal of Biological Chemistry* **284**: 21360-21368.

Davies, E.V., C. Winstanley, J.L. Fothergill & C.E. James, (2016) The role of temperate bacteriophages in bacterial infection. *FEMS Microbiology Letters* **363**: fnw015.

del Castillo, T., E. Duque & J.L. Ramos, (2008) A set of activators and repressors control peripheral glucose pathways in Pseudomonas putida to yield a common central intermediate. *Journal of Bacteriology* **190**: 2331-2339.

Entner, N. & M. Doudoroff, (1952) Glucose and gluconic acid oxidation of Pseudomonas saccharophila. *The Journal of Biological Chemistry* **196**: 853-862.

Friedman, L. & R. Kolter, (2004a) Genes involved in matrix formation in Pseudomonas aeruginosa PA14 biofilms. *Molecular Microbiology* **51**: 675-690.

Friedman, L. & R. Kolter, (2004b) Two genetic loci produce distinct carbohydrate-rich structural components of the Pseudomonas aeruginosa biofilm matrix. *Journal of Bacteriology* **186**: 4457-4465.

Ge, P., D. Scholl & P.G. Leiman, (2015) Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nature Structural and Molecular Biology* **22**: 377-382.

Ghequire, M.G. & R. De Mot, (2014) Ribosomally encoded antibacterial proteins and peptides from Pseudomonas. *FEMS Microbiology Reviews* **38**: 523-568.

Green, S., D.J. Studholme, B.E. Laue, F. Dorati, H. Lovell, D. Arnold, J.E. Cottrell, S. Bridgett, M. Blaxter, E. Huitema, R. Thwaites, P.M. Sharp, R.W. Jackson & S. Kamoun, (2010) Comparative genome analysis provides insights into the evolution and adaptation of Pseudomonas syringae pv. aesculi on Aesculus hippocastanum. *PLoS ONE* **5**: e10224.

Gross, M. & K. Rudolph, (1987) Studies on the Extracellular Polysaccharides (EPS) Produced in vitro by Pseudomonas phaseolicola I. Indications for a Polysaccharide Resembling Alginic Acid in Seven P. syringae Pathovars. *Journal of Phytopathology* **118**: 276-287.

Hartman, G.L., E.D. West & T.K. Herman, (2011) Crops that feed the World 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. *Food Security* **3**: 5-17.

Hettwer, U., M. Gross & K. Rudolph, (1995) Purification and characterization of an extracellular levansucrase from Pseudomonas syringae pv. phaseolicola. *Journal of Bacteriology* **177**: 2834-2839.

Hettwer, U., F.R. Jaeckel, J. Boch, M. Meyer, K. Rudolph & M.S. Ullrich, (1998) Cloning, nucleotide sequence, and expression in Escherichia coli of levansucrase genes from the plant pathogens Pseudomonas syringae pv. glycinea and P. syringae pv. phaseolicola. *Applied and Environmental Microbiology* **64**: 3180-3187.

Hirano, S.S. & C.D. Upper, (2000) Bacteria in the leaf ecosystem with emphasis on Pseudomonas syringae-a pathogen, ice nucleus, and epiphyte. *Microbiology and Molecular Biology Reviews: MMBR* **64**: 624-653.

Hockett, K.L., A.Y. Burch & S.E. Lindow, (2013) Thermo-regulation of genes mediating motility and plant interactions in Pseudomonas syringae. *PLoS ONE* **8**: e59850.

Hockett, K.L., T. Renner & D.A. Baltrus, (2015) Independent Co-Option of a Tailed Bacteriophage into a Killing Complex in Pseudomonas. *MBio* **6**: e00452.

Ignjatov, M., J. Balaž, M. Milošević, M. Vidić & T. Popović, (2008) Studies on Plant Pathogenic Bacterium Causal Agent of Soybean Bacterial Spots (Pseudomonas syringaepv. glycinea (Coerper) Young et al.). In: Pseudomonas syringae Pathovars and Related Pathogens – Identification, Epidemiology and Genomics. M.B. Fatmi, A. Collmer, N. Iacobellis, J. Mansfield, J. Murillo, N. Schaad & M. Ullrich (eds). Springer Netherlands, pp. 419-426.

Khandekar, S., (2013) Expression and regulation of levansucrase in Pseudomonas syringae. In: School of Engineering and Science. Germany: Jacobs University Bremen, pp.

Khandekar, S., A. Srivastava, D. Pletzer, A. Stahl & M.S. Ullrich, (2014) The conserved upstream region of lscB/C determines expression of different levansucrase genes in plant pathogen Pseudomonas syringae. *BMC Microbiology* **14**: 79.

Kim, J., C.O. Jeon & W. Park, (2008) Dual regulation of zwf-1 by both 2-keto-3-deoxy-6-phosphogluconate and oxidative stress in Pseudomonas putida. *Microbiology* **154**: 3905-3916.

Knezevic, P., R. Kostanjsek, D. Obreht & O. Petrovic, (2009) Isolation of Pseudomonas aeruginosa specific phages with broad activity spectra. *Current Microbiology* **59**: 173-180.

Kuklinsky-Sobral, J., W.L. Araujo, R. Mendes, I.O. Geraldi, A.A. Pizzirani-Kleiner & J.L. Azevedo, (2004) Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion. *Environmental Microbiology* **6**: 1244-1251.

Kulminskaya, A.A., M. Arand, E.V. Eneyskaya, D.R. Ivanen, K.A. Shabalin, S.M. Shishlyannikov, A.N. Saveliev, O.S. Korneeva & K.N. Neustroev, (2003) Biochemical characterization of Aspergillus awamori exoinulinase: substrate binding characteristics and regioselectivity of hydrolysis. *Biochimica et Biophysica Acta* **1650**: 22-29.

Lal, R., (2009) Soil degradation as a reason for inadequate human nutrition. *Food Security* **1**: 45-57.

Laue, H., A. Schenk, H. Li, L. Lambertsen, T.R. Neu, S. Molin & M.S. Ullrich, (2006) Contribution of alginate and levan production to biofilm formation by Pseudomonas syringae. *Microbiology* **152**: 2909-2918.

Li, H., A. Schenk, A. Srivastava, D. Zhurina & M.S. Ullrich, (2006) Thermo-responsive expression and differential secretion of the extracellular enzyme levansucrase in the plant

pathogenic bacterium Pseudomonas syringae pv. glycinea. *FEMS Microbiology Letters* **265**: 178-185.

Li, H. & M.S. Ullrich, (2001) Characterization and mutational analysis of three allelic lsc genes encoding levansucrase in Pseudomonas syringae. *Journal of Bacteriology* **183**: 3282-3292.

Li, J. & N. Wang, (2012) The gpsX gene encoding a glycosyltransferase is important for polysaccharide production and required for full virulence in Xanthomonas citri subsp. citri. *BMC Microbiology* **12**: 31.

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

Mann, E.E. & D.J. Wozniak, (2012) Pseudomonas biofilm matrix composition and niche biology. *FEMS Microbiology Reviews* **36**: 893-916.

Mansfield, J., S. Genin, S. Magori, V. Citovsky, M. Sriariyanum, P. Ronald, M. Dow, V. Verdier, S.V. Beer, M.A. Machado, I. Toth, G. Salmond & G.D. Foster, (2012) Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology* **13**: 614-629.

Matsui, H., Y. Sano, H. Ishihara & T. Shinomiya, (1993) Regulation of pyocin genes in Pseudomonas aeruginosa by positive (prtN) and negative (prtR) regulatory genes. *Journal of Bacteriology* **175**: 1257-1263.

Matsukawa, M. & E.P. Greenberg, (2004) Putative exopolysaccharide synthesis genes influence Pseudomonas aeruginosa biofilm development. *Journal of Bacteriology* **186**: 4449-4456.

Mehmood, A., K. Abdallah, S. Khandekar, D. Zhurina, A. Srivastava, N. Al-Karablieh, G. Alfaro-Espinoza, D. Pletzer & M.S. Ullrich, (2015) Expression of extra-cellular levansucrase in Pseudomonas syringae is controlled by the in planta fitness-promoting metabolic repressor HexR. *BMC Microbiology* **15**: 48.

Mercier, J. & S.E. Lindow, (2000a) Role of Leaf Surface Sugars in Colonization of Plants by Bacterial Epiphytes. *Applied and Environmental Microbiology* **66**: 369-374.

Mercier, J. & S.E. Lindow, (2000b) Role of leaf surface sugars in colonization of plants by bacterial epiphytes. *Applied and Environmental Microbiology* **66**: 369-374.

Miao, E.A. & S.I. Miller, (1999) Bacteriophages in the evolution of pathogen-host interactions. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 9452-9454.

Michel-Briand, Y. & C. Baysse, (2002) The pyocins of Pseudomonas aeruginosa. *Biochimie* **84**: 499-510.

Mindich, L., X. Qiao, J. Qiao, S. Onodera, M. Romantschuk & D. Hoogstraten, (1999) Isolation of additional bacteriophages with genomes of segmented double-stranded RNA. *Journal of Bacteriology* **181**: 4505-4508.

Monier, J.M. & S.E. Lindow, (2003) Pseudomonas syringae Responds to the Environment on Leaves by Cell Size Reduction. *Phytopathology* **93**: 1209-1216.

Mooij, M.J., E. Drenkard, M.A. Llamas, C.M. Vandenbroucke-Grauls, P.H. Savelkoul, F.M. Ausubel & W. Bitter, (2007) Characterization of the integrated filamentous phage Pf5 and its involvement in small-colony formation. *Microbiology* **153**: 1790-1798.

Morgan, G.J., G.F. Hatfull, S. Casjens & R.W. Hendrix, (2002) Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in Haemophilus, Neisseria and Deinococcus. *Journal of Molecular Biology* **317**: 337-359.

Morris, C.E., D.C. Sands, B.A. Vinatzer, C. Glaux, C. Guilbaud, A. Buffiere, S. Yan, H. Dominguez & B.M. Thompson, (2008) The life history of the plant pathogen Pseudomonas syringae is linked to the water cycle. *The ISME journal* **2**: 321-334.

Murray-Kolb, L.E., R. Welch, E.C. Theil & J.L. Beard, (2003) Women with low iron stores absorb iron from soybeans. *The American Journal of Clinical Nutrition* **77**: 180-184.

O'Brien, H.E., D. Desveaux & D.S. Guttman, (2011) Next-generation genomics of Pseudomonas syringae. *Current Opinion in Microbiology* **14**: 24-30.

Olsthoorn, R.C., G. Garde, T. Dayhuff, J.F. Atkins & J. Van Duin, (1995) Nucleotide sequence of a single-stranded RNA phage from Pseudomonas aeruginosa: kinship to coliphages and conservation of regulatory RNA structures. *Virology* **206**: 611-625.

Oner, E.T., L. Hernandez & J. Combie, (2016) Review of Levan polysaccharide: From a century of past experiences to future prospects. *Biotechnology Advances* **34**: 827-844.

Portais, J.C. & A.M. Delort, (2002) Carbohydrate cycling in micro-organisms: what can (13)C-NMR tell us? *FEMS Microbiology Reviews* **26**: 375-402.

Qiao, X., Y. Sun, J. Qiao, F. Di Sanzo & L. Mindich, (2010) Characterization of Phi2954, a newly isolated bacteriophage containing three dsRNA genomic segments. *BMC Microbiology* **10**: 55.

Rico, A., S.L. McCraw & G.M. Preston, (2011) The metabolic interface between Pseudomonas syringae and plant cells. *Current Opinion in Microbiology* **14**: 31-38.

Rico, A. & G.M. Preston, (2008) Pseudomonas syringae pv. tomato DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. *Molecular Plant-Microbe Interactions : MPMI* **21**: 269-282.

Riley, M.A. & J.E. Wertz, (2002) Bacteriocins: evolution, ecology, and application. *Annual Review of Microbiology* **56**: 117-137.

Rosell, M.S., P.N. Appleby, E.A. Spencer & T.J. Key, (2004) Soy intake and blood cholesterol concentrations: a cross-sectional study of 1033 pre- and postmenopausal women in the Oxford arm of the European Prospective Investigation into Cancer and Nutrition. *The American Journal of Clinical Nutrition* **80**: 1391-1396.

Sano, Y., H. Matsui, M. Kobayashi & M. Kageyama, (1993) Molecular structures and functions of pyocins S1 and S2 in Pseudomonas aeruginosa. *Journal of Bacteriology* **175**: 2907-2916.

Schenk, A., M. Berger, L.M. Keith, C.L. Bender, G. Muskhelishvili & M.S. Ullrich, (2006) The algT gene of Pseudomonas syringae pv. glycinea and new insights into the transcriptional organization of the algT-muc gene cluster. *Journal of Bacteriology* **188**: 8013-8021.

Schnider-Keel, U., K.B. Lejbolle, E. Baehler, D. Haas & C. Keel, (2001) The sigma factor AlgU (AlgT) controls exopolysaccharide production and tolerance towards desiccation and osmotic stress in the biocontrol agent Pseudomonas fluorescens CHA0. *Applied and Environmental Microbiology* **67**: 5683-5693.

Shenge, K.C., R.B. Mabagala, C.N. Mortensen, D. Stephan & K. Wydra, (2007) First report of bacterial speck of tomato caused by Pseudomonas syringae pv. tomato in Tanzania. *Plant Disease* **91**: 462-462.

Shinomiya, T., S. Shiga & M. Kageyama, (1983) Genetic determinant of pyocin R2 in Pseudomonas aeruginosa PAO. I. Localization of the pyocin R2 gene cluster between the trpCD and trpE genes. *Molecular & General Genetics : MGG* **189**: 375-381.

Srivastava, A., N. Al-Karablieh, S. Khandekar, A. Sharmin, H. Weingart & M.S. Ullrich, (2012) Genomic Distribution and Divergence of Levansucrase-Coding Genes in Pseudomonas syringae. *Genes (Basel)* **3**: 115-137.

Srividhya, K.V., V. Alaguraj, G. Poornima, D. Kumar, G.P. Singh, L. Raghavenderan, A.V. Katta, P. Mehta & S. Krishnaswamy, (2007) Identification of prophages in bacterial genomes by dinucleotide relative abundance difference. *PLoS ONE* **2**: e1193.

Strange, R.N. & P.R. Scott, (2005) Plant disease: a threat to global food security. *Annual Review of Phytopathology* **43**: 83-116.

Sutherland, I., (2001a) Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* **147**: 3-9.

Sutherland, I.W., (2001b) The biofilm matrix--an immobilized but dynamic microbial environment. *Trends in Microbiology* **9**: 222-227.

Tampakaki, A.P., E. Hatziloukas & N.J. Panopoulos, (2009) Plant Pathogens, Bacterial. In: Encyclopedia of Microbiology (Third Edition). S. Editor-in-Chief: Moselio (ed). Oxford: Academic Press, pp. 655-677.

Thakur, S., B.S. Weir & D.S. Guttman, (2016) Phytopathogen Genome Announcement: Draft Genome Sequences of 62 Pseudomonas syringae Type and Pathotype Strains. *Molecular Plant-Microbe Interactions : MPMI* **29**: 243-246.

van Heel, A.J., A. de Jong, M. Montalban-Lopez, J. Kok & O.P. Kuipers, (2013) BAGEL3: Automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Research* **41**: W448-453.

Varani, A.M., C.B. Monteiro-Vitorello, H.I. Nakaya & M.A. Van Sluys, (2013) The role of prophage in plant-pathogenic bacteria. *Annual Review of Phytopathology* **51**: 429-451.

Villegas, R., Y.T. Gao, G. Yang, H.L. Li, T.A. Elasy, W. Zheng & X.O. Shu, (2008) Legume and soy food intake and the incidence of type 2 diabetes in the Shanghai Women's Health Study. *The American Journal of Clinical Nutrition* **87**: 162-167.

Wagner, P.L. & M.K. Waldor, (2002) Bacteriophage control of bacterial virulence. *Infection and Immunity* **70**: 3985-3993.

Wilson, M. & S.E. Lindow, (1994) Coexistence among Epiphytic Bacterial Populations Mediated through Nutritional Resource Partitioning. *Applied and Environmental Microbiology* **60**: 4468-4477.

Winsor, G.L., E.J. Griffiths, R. Lo, B.K. Dhillon, J.A. Shay & F.S. Brinkman, (2016) Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic Acids Research* **44**: D646-653.

Winsor, G.L., D.K. Lam, L. Fleming, R. Lo, M.D. Whiteside, N.Y. Yu, R.E. Hancock & F.S. Brinkman, (2011) Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. *Nucleic Acids Research* **39**: D596-600.

Wu, W. & S. Jin, (2005) PtrB of Pseudomonas aeruginosa suppresses the type III secretion system under the stress of DNA damage. *Journal of Bacteriology* **187**: 6058-6068.

Xin, X.F. & S.Y. He, (2013) Pseudomonas syringae pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. *Annual Review of Phytopathology* **51**: 473-498.

Yu, X., S.P. Lund, R.A. Scott, J.W. Greenwald, A.H. Records, D. Nettleton, S.E. Lindow, D.C. Gross & G.A. Beattie, (2013) Transcriptional responses of Pseudomonas syringae to growth in epiphytic versus apoplastic leaf sites. *Proceedings of the National Academy of Sciences of the United States of America* **110**: E425-434.

Zago, A. & S. Chugani, (2009) Pseudomonas. In: Encyclopedia of Microbiology (Third Edition). S. Editor-in-Chief: Moselio (ed). Oxford: Academic Press, pp. 245-260.

Zhurina, D., (2009) Identification and partial characterization of transcriptional regulators involved in temperature-dependent expression of levansucrase in *Pseudomonas syringae*. In: School of Engineering and Science. Germany: Jacobs University Bremen, pp.

Statutory Declaration

(Declaration on Authorship of a Dissertation)

I, **Khaled Abdallah** hereby declare, under penalty of perjury, that I am aware of the consequences of a deliberately or negligently wrongly submitted affidavit, in particular the punitive provisions of § 156 and § 161 of the Criminal Code (up to 1-year imprisonment or a fine at delivering a negligent or 3 years or a fine at a knowingly false affidavit).

Furthermore, I declare that I have written this PhD thesis independently, unless where clearly stated otherwise. I have used only the sources, the data and the support that I have clearly mentioned.

This PhD thesis has not been submitted for the conferral of a degree elsewhere.

Bremen	**17.10.2016**
Place	Date

Signature