universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

"ECF sigma factors: How endophytes sense the plants"

verfasst von / submitted by

Raheleh Sheibani Tezerji, MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2016/Vienna, 2016

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 685 437
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Biologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Thomas Rattei

Table of Content

Acknowledgments

I firmly believe that this work would have never been possible without the profound support and contribution of others. First of all, I would like to express my gratitude to my supervisors Prof. Dr. Thomas Rattei and Dr. Birgit Mitter, who have guided me over the last five years and gave me wonderful opportunities in the course of my thesis. I am very grateful to Dr. Mitter, who gave me the opportunity to start my PhD project at Austrian institute of Technology (AIT) and for her scientific advice and support during all stages of my research process. A special appreciation to Prof. Dr. Rattei, who gave me the opportunity to start my PhD study at University of Vienna, and he taught me how to develop new scientific ideas and given me the support needed to pursue my research unimpeded.

Special thanks go out to Prof. Dr. Angela Sessitsch, Prof. Dr. Heinz Himmelbauer and Dr. Stefanie Wienkoop, who kindly accept to review my dissertation as examiners.

In my journey, I have been fortunate enough to collaborate with several amazing people who show me how informative, exciting, appealing and fun research can be.

A special appreciation goes to the contributions of my co-authors at Austrian Institute of Technology (AIT) and Vienna University, Prof. Dr. Angela Sessitsch, Dr. Friedericke Trognitz, Dr. Marc-André Jehl, Dr Muhammad Naveed, Alexandra Petric, Dr. Stéphane Compant, Dr. Livio Antonielli, Nikolaus Pfaffenbichler and Teresa Berninger.

I acknowledge the support and collaborations of my colleagues at University of Vienna. Special thanks to Dr. Marc-André Jehl with his support for eukaryote like domain analysis and Thomas Weinmaier for providing the gene prediction and annotation workflow (ConsPred).

I would like to thank the people that worked with me during their master studies or internship, specially, Maryam Mashaghi Tabari, Johanna Jahn and Alba Nicolas for your help in the lab work. I am also thankful for the help of all technician staff in AIT specially, Milica Pastar for her help during my lab work. A special thank goes out to the office in AIT to Iris Hagenauer and University of Vienna to Gabriele Kaindl who

helped me with all non-scientific matters. I am also grateful to Florian Goldenberg for his help and support during my thesis.

This thesis would not have been possible without the support from Austrian Science Fund (FWF). Therefore, I would like to thank them for their financial support of the project (Grant No.: P22867-B16).

I have been really lucky to have the support and friendship of my great “outside” friends, specially, Laleh Rafati and Elena Krstevska. Thank you for all your enthusiasm and for uncomplainingly listening to my PhD-related stories.

My deepest appreciation and thanks to my family for their unconditional love, understanding and support. Mom, Hamid, Saeed, Khatereh and Bahareh; you all have been the main source of encouragement in my life, and I can never thank you enough. I also dedicate my thesis to the memory of my father, who would have been happy to see me following my dreams.

Chapter I

Introduction

Introduction

Endophytic bacteria are microorganisms that span all or part of their life cycle inside the tissues of living plants and cause unapparent and asymptomatic infections [1]. For over hundred years, endophytic bacterial communities are recognized that colonize the internal tissue of their host plant. It is well known that bacterial endophytes actively grow within and interact with their host plant. A positive impact of several endophytic bacteria on a broad range of plants and crops has been verified [2,3]. However, the mechanisms of plant-bacteria communication in response to different environmental conditions are still poorly understood.

In the following sections, I will first give an overview on the endophytic bacterial colonization and life style inside plant tissue. Second, I discuss how high throughput sequencing analysis assists in discovering the beneficial traits of endophytic bacteria.

Variety in life style of endophytic bacteria inside host plant

The interaction between endophytic bacteria and plants -depending on their host and ecological niches- can show differences in life styles such as endophytic, saprophytic, epiphytic and pathogenic [15].

Colonisation of plants by endophytic bacteria

More than 200 endophytic bacterial genera from 16 phyla are identified. Among these, three phyla are more studied belonging to Actinobacteria, Proteobacteria and Firmicutes [4]. They include members of *Burkholderia* [5], *Bacillus* [6], *Herbaspirillum* [7], *Azoarcus* [8], *Gluconobacter* [9], *Enterobacter* [10], *Pseudomonas* [11], *Stenotrophomonas* [12], *Streptomyces* [13] and *Serratia* [12]. Figure 1 shows an example of endophytic bacteria (*Burkholderia phytofirmance* PsJN [5]) inside the plant tissue (grapevine) using fluorescence *in situ* hybridization methods (FISH) [14].

The origin of endophytic bacteria is rhizosphere or phyllosphere. They can enter plants through tissue wounds, stomata and lenticels [16]. Specific endophytic strains are capable of colonizing and surviving in reproductive plant organs such as flowers, ovaries and fruits. They can also be transmitted through the seeds [17]. Compant et al. suggested that in grapevine, *B. phytofirmance* PSJN penetrate plant through the transpiration stream from root xylem vessels. In this way, bacteria remain invisible to the plant immune system [14]. Mobility, penetration capability, and capacity for adjustment of metabolism help the process of bacterial colonisation inside the plant

tissue [16]. The colonization of endophytic bacteria inside the host plants can be considered as a sign of healthy plant system as it can promote the growth of their host plant [10].

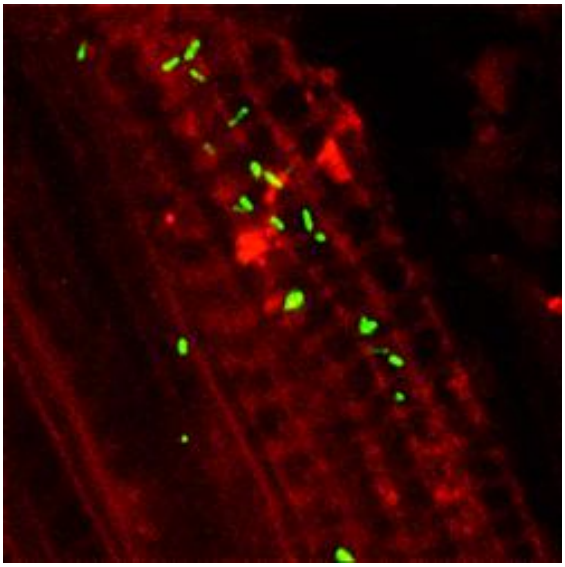


Figure 1. *Burkholderia phytofirmans* PsJN inside grapevine endorhiza visualized by FISH (fluorescence in situ hybridization) combined with confocal laser scanning microscopy using specific probes [14].

The establishment of endophytic populations inside plant tissue depends on the recognition of signal molecules via extracytoplasmatic function (ECF) sigma factors and two-component system [16]. One way how bacteria sense and react to the extracellular environment is the so-called cell surface signalling (CSS), employing alternative sigma factors. This signal transduction system consists of an outer membrane receptor, an inner-membrane bound sigma factor regulator (anti-sigma factor) and bound to that an extracytoplasmatic function (ECF) group IV sigma factor. Upon signal recognition the ECF sigma factor is released and activates expression of its target genes. The ECF subfamily is the largest group among the s70 family and its members are involved in a wide range of environmental responses, such as metal homeostasis, starvation and resistance to antimicrobial peptides, being also required for pathogenesis in some cases [18], and in the symbiotic efficiency in the plant symbiont *Bradyrhizobium japonicum* USDA110 [19].

Growth promoting effect of endophytic bacteria on the host plant

After establishment inside the plant tissue, endophytic bacteria apply different strategies to overcome plant defence response and adapt to the new environment [20]. Thereby, bacterial endophytes show a range of properties that are playing an

important role in plant growth promotion, including improvement of nutrient uptake, increasing stress tolerance, antagonizing plant pathogens and induction of systemic resistance [21–25].

In the following, the major mechanisms of endophytic bacteria to stimulate plant growth are discussed. For many of these mechanisms experimental studies were published that report support by *in planta* experiments. The main plant growth promoting mechanisms of endophytic bacteria are illustrated in Figure 2.

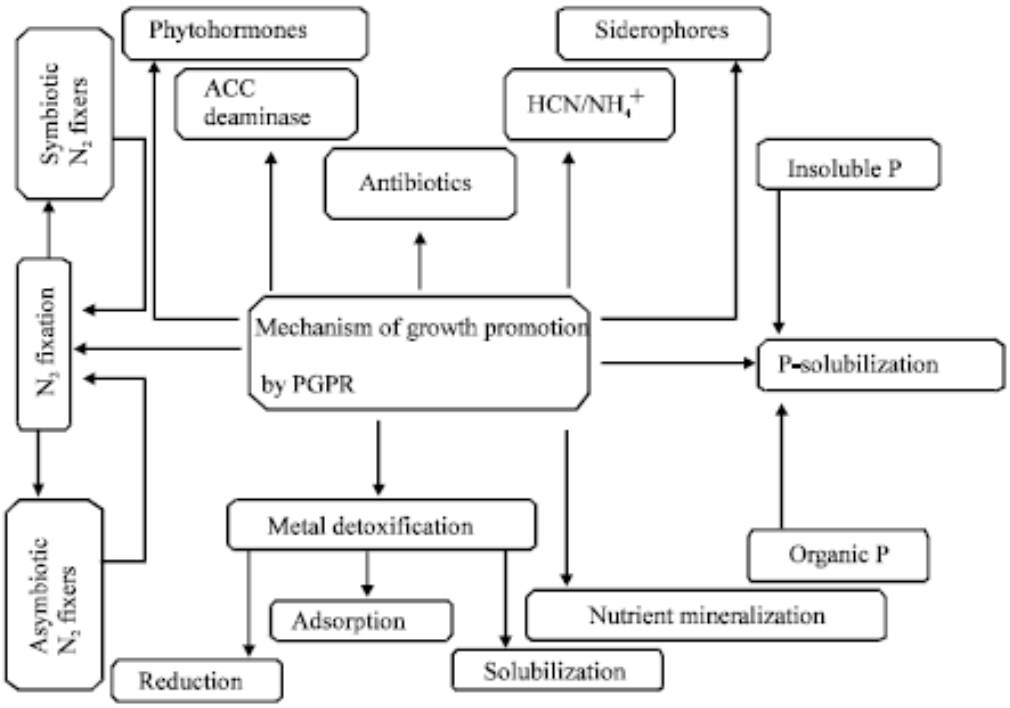


Figure 2. Illustration of the main plant growth promotion effects (PGP) mediated by endophytic plant-beneficial bacteria [28].

Endophytic bacterial colonization inside plant tissue directly starts the plant growth stimulation by secretion of phytohormones such as indole-3-acetic acid (IAA) [26,27]. This phytohormone leads to increased uptake of nutrients from the recipient environment [28]. The effects of IAA on plant cell division, extension, differentiation, photosynthesis, pigment formation, biosynthesis of various metabolites, and resistance to stressful conditions are already studied [29]. Furthermore, IAA increases the rate of xylem and root development besides playing a role in seed stimulation and tuber germination. This hormone has several plant growth promoting effects such as controls of vegetative growth, initiation of root formation, and mediation of the response to gravity and light [30,31]. Plant growth promotion and root nodulation are both affected by IAA [29].

Many IAA-producing endophytes possess 1-aminocyclopropane-1-carboxylate (ACC)-deaminase activity which is involved in lowering the level of plant ethylene [32]. Elevated levels of ethylene caused by some stress conditions are known to inhibit root elongation and lateral root emergence [33]. According to the model proposed by Glick et al., bacterial IAA can induce ACC synthase of plants resulting in the production of ACC, the ethylene precursor. Some bacteria can use ACC as a nutrient source and thereby decrease the synthesis of ethylene in plants [34]. ACC-deaminase activity has been widely studied for plant growth promoting endophytic strains of *Burkholderia* [35,36], *Herbaspirillum* [37] and *Pseudomonas* [32]. In *B. phytofirmans* PsJN, a high level of ACC reduces the level of the inhibitory hormone ethylene to promote the growth of potato and grapevine [14].

Plant growth promotion by endophytic bacteria is continued along nitrogen fixation [38], phosphate solubilization and the production of siderophore, ammonia and antibiotic [29,39,40]. The host plant obtains nutrients through the biological nitrogen fixation (BNF) process that in several crops (such as maize, rice, sugar cane and wheat) is dependent on the symbiosis with endophytic bacteria. Some members of the endophytic bacteria such as *Burkholderia*, *Gluconobacter*, *Azoarcus*, *Herbaspirillum* and *Klebsiella* are known to be included in the nitrogen fixation process [41,42]. James and colleagues suggested that the plants can obtain up to 70% of their required nitrogen by association with endophytic diazotrophs in the biological nitrogen fixation process [41].

Endophytic batcteria are also indirectly capable to reduce microbial populations that are harmful to the plant. Thereby, they act as agents of biological control through competition, antibiosis, or systemic resistance induction [43]. As they colonize the same ecological niche as plant pathogens and can compete with pathogens, endophytes are particularly interesting candidates for biological controls [44,45]. Furthermore, endophytic bacteria have been reported to carry genes encoding the degradation of xenobiotics and therefore have the potential to improve phytoremediation applications [2,46]. These studies show the importance of biotechnological applications of endophytic bacteria for improving phytoremediation, biofuel production and plant fortification [5]. Yousaf and colleagues investigated that by phytoremediation of inorganic pollutants endophytic bacteria can reduce the

phytotoxicity and increase the mobilization and accumulation of heavy metals in aboveground plant biomass [46].

The positive effects of endophytic bacteria on plants growth under abiotic stress have been addressed in several studies confirming their important role in tolerance to abiotic stress [47]. Plant growth promoting endophytic bacteria *B. phytofirmans* PsJN was used to investigate the effects of drought stress on growth, physiology and yield of crop in wheat (*Triticum aestivum* L.) and two maize cultivars (cvs. Mazurka and Kaleo) [48,49]. Enhancement of tolerance to high temperatures in sorghum seedlings [50], the enhancement of salt tolerance in rice [51], and mitigation of the negative effects of drought in the common bean [52]. Furthermore, the impact of endophytic bacteria under stress condition has been reviewed in several studies [53,54].

Unraveling endophyte-plant interaction using next generation sequencing

Since 1990s, when the first bacterial genomes were sequenced, genome analysis became an effective means for improving our understanding of the biology, genetic, evolution, epidemiology and pathogenesis of bacterial strains [55]. Genome analysis provides detailed information on the differences in the host interaction strategies of endophytic bacteria and reveals the traits related to plant growth promoting bacteria and the functionality of endophytic bacteria in host plants. The intent of this part is to investigate how high throughput sequencing analysis assists to discover the beneficial traits of endophytic bacteria inside plant tissue as they are reflected in their genome sequences. These traits include genes for motility, colonization, type IV pili, flagella, diverse secretion systems, phytohormone synthesis and inhibition, bacterial volatiles, and antimicrobials [56].

In Table 1, the endophytic bacteria with completely sequenced genomes are summarized. The fully sequenced endophytic bacterial genomes from class Alphaproteobacteria include *M. populi* BJ001 [57], *G. diazotrophicus* PaI5 [9], *Azospirillum* sp. B510 [58] and *A. lipoferum* 4B [59]. The endophytic bacterial genomes of *Azoarcus* sp. BH72 [60], *H. seropedicae* SmR1 [7], *B. phytofirmans* PsJN [5], *Burkholderia* spp. KJ006 [61] and *V. paradoxus* EPS [62] belong to Betaproteobacteria. The endophytic bacteria that are classified as Gamaproteobacteria include *P. stuzeri* A1501 [63], *K. pneumoniae* 342 [64], *S. proteamaculans* 568 [11], *S. maltophilia* R551-3 [11], *P. putida* W619 [11], *Enterobacter* sp. 638 [10] and *E. cloacae* ENHKU01 [65].

Table1. Bacterial endophytes with completely sequenced genomes.

Endophyte Species	Class	Genome size (Mb)	Host plant	Accession Number	Reference
Azospirillum lipoferum 4B	Alphaproteobacteria	6.85 (1 Chr., 6 Pl.)	Rice, maize, wheat	FQ311868-FQ311874	Wisniewski-Dyé et al., 2011
Azospirillum sp. B510		7.6 (1 Chr., 6 Pl.)	Rice	AP010946–AP0109452	Kaneko et al., 2010
Gluconacetobacter diazotrophicus PaI5		3.9 (1 Chr., 2 Pl.)	Sugarcane, rice, coffee, tea	AM889285–AM889287	Bertalan et al., 2009
Methylobacterium populi BJ001		5.8 (1 Chr., 2 Pl.)	Poplar	CP001029-CP001031	Van Aken et al., 2004
Azoarcus sp. BH72	Betaproteobacteria	4.37 (1 Chr.)	Rice	AM406670	Krause et al., 2006
Burkholderia phytofirmans PsJN		8.2 (2 Chr., 1 Pl.)	Potato, tomato, maize, barley, onion, canola, grapevine	CP001052-CP001054	Weilharter et al., 2011
Burkholderia spp. KJ006		6.6 (3 Chr., 1 Pl.)	Rice	CP003514-CP003517	Kwak et al., 2012
Variovorax paradoxus EPS		6.55 (1 Chr.)	Potato	CP002417	Han et al., 2013
Herbaspirillum seropedicae SmR1		5.51 (1 Chr.)	Sugarcane, rice	CP002039	Pedrosa et al., 2011
Enterobacter cloacae ENHKU01	Gammaproteobacteria	4.7 (1 Chr.)	Pepper	CP003737	Liu et al., 2012
Enterobacter sp. 638		4.67 (1 Chr., 1 Pl.)	Poplar	CP000653-CP000654	Taghavi et al., 2010
Klebsiella pneumoniae 342		5.9 (1 Chr., 2 Pl.)	Maize, wheat	CP000964–CP000966	Fouts et al., 2008
Pseudomonas putida W619		5.77 (1 Chr.)	Poplar	CP000949	Taghavi et al., 2009
Pseudomonas stuzeri A1501		4.5 (1 Chr.)	Rice	CP000304	Yan et al., 2008
Serratia proteamaculans 568		5.5 (1 Chr., 1 Pl.)	Soybean	CP000826-CP000827	Taghavi et al., 2009
Stenotrophomonas		4.57	Poplar	CP001111	Taghavi et al.,

maltophilia R551-3		(1 Chr.)			2009
Nostoc azollae 0708	Cyanobacteria	5.4 (1 Chr.,2 Pl.)	*Azolla filiculoides* (Water Fern)	CP002059- CP002061	Ran et al., 2010

After a well designed and effective experiment that is followed by bacterial genome sequencing, suitable bioinformatics analysis and tools are required to answer the biological hypotheses. The genome analysis is initiated with pre-processing of the reads considering quality control, trimming of the reads based on their length or removing the adapter sequences [67]. In order to reconstruct a bacterial genome, de novo assembly and/or reference-based-assembly will be applied in absence of a reference genome. Otherwise reads will be aligned to an existing bacterial reference genome [68–70]. This process is followed by genome annotation to extract the biological information that can be inferred from sequence similarity (e.g. genes, functions, pathways) and the genome quality will be improved by gap filling [71].

Comparative genomics focuses on analysing functions and structures of genomes (e.g. finding genes, testing functional divergence, detecting specialized islands, identifying the genetic variants and understanding important biological pathways associated with special conditions) to increase our understanding of biological systems [71]. For example, comparative genomic analysis of the *B. phytofirmans* PsJN and eight other endophytic bacteria revealed the mechanisms that endopyhtic bacteria use for colonization inside plant. The study investigated the capability of *B. phytofirmans* PsJN to harbour a high number of cell surface signalling and secretion systems that allow it to interact with a variety of host plant species and adapt to different environments [16].

Metagenomic analysis enables rapid estimation of the organismal composition, diversity and metabolic potential encoded in genetic material obtained from microbial communities in several habitats [72,73]. Endophytic bacterial communities in *Aloe Vera* were evaluated by this approach that can lead us to better understanding of endophyte-host plant interaction [74]. Metagenomic studies of the most abundant endophytic bacteria of rice verified traits which are shared among bacterial endophytes and might be important for their interactions with plants [75]. These traits

include secretion systems, enzymes involved in ROS degradation, cellulolytic and pectinolytic enzymes, flagellins receptors and transporters for iron uptake, quorum sensing systems, metabolic pathways for degradation of plant compounds, and several plant-growth promoting and biocontrol traits (such as ACC-deaminase activity, biological nitrogen fixation, production of antimicrobial compounds, phytohormones and volatiles) [4].

Applying transcriptome analysis can link the genomic potential with function and therefore give a deeper insight into endophyte-plant interaction. After bacterial RNA-sequencing, pre-processed reads with high quality can be assembled (de novo or reference-based) and/or can be mapped to the reference genome sequence. To quantify the activity of transcriptional features, the abundance of the transcripts must be measured [76]. Normalization of the read count data obtained from RNA-seq is necessary to remove possible biases in the data and to quantify the differential expression of transcriptional features within or between different samples and experiments [76,77]. There are different approaches available for read count normalization such as Upper Quantile [78], TMM [79], RPKM [80], FPKM [81] or DESeq [82]. One of the common normalization approaches is to calculate the RPKM value: reads per kilo base per million mapped sequence reads (RPKM) value as gene expression measure [80]. Another option is to normalize read counts using expression levels of housekeeping genes. This method is often used for normalizing qRT-PCR expression measures [77]. To compute differential expression levels between two experimental conditions given the expression level of the considered features, several packages are implemented in R/Bioconductor such as NoiSeq [83] and DeSeq2 [84]. These packages perform a statistical test for the null hypothesis, which is that the expression of the gene in the two conditions is the same or not. This analyzes eventually points towards the determination of differentially expressed genes under certain experimental conditions [77].

The differentially expressed candidate genes are likely to be involved in the beneficial plant-endophyte interactions. They can be further investigated for their phenotypic effects and functional characterization of endophyte-plant interaction such as resistance, virulence and stress tolerance [85].

So far, the gene expression patterns of plants and their bacterial endophytes have only been addressed regarding the host plant side [86,87]. However, the patterns of

gene expression of beneficial interactions between endophytes and plants are tightly linked. Therefore it is important to analyse both of them, together inside the planta. Transcriptomic analysis of a simultaneous transcriptional profiling approach of endophytic bacteria *in planta* could offer valuable insights.

Also the molecular basis of endophyte-host interactions is far from being well understood. The genomes of many important endophytic bacterial strains are still not completely sequenced. We still have only a limited knowledge of the bacterial traits determining endophytic lifestyle and internal colonization of host plants [2,16,88]. In addition, there are few mechanisms experimentally characterized *in planta* and the functional context of the majority of endophyte genes is unknown [89].

In this thesis, I investigate the genetics of plant-endophytic bacteria inside the plant environment on the levels of DNA and RNA, while addressing the experimental challenges of investigating endophytic bacterial cells during colonization of the host plant tissue. To get a deeper insight into the functionality of endophytes and their roles in interaction with the host plant, I applied high throughput sequencing analysis along with extensive experimental methods. The application of comparative genomics and transcriptomics approaches to study the function and lifestyle of endophytic bacteria *in planta* under external stress conditions are among the main aims of this thesis.

References

1 Wilson D. Endophyte: The Evolution of a Term, and Clarification of Its Use and Definition. *Oikos* 1995;**73**:274–6. doi:10.2307/3545919

2 Ryan RP, Germaine K, Franks A, *et al.* Bacterial endophytes: recent developments and applications. *FEMS microbiology letters* 2008;**278**:1–9. doi:10.1111/j.1574-6968.2007.00918.x

3 Hardoim PR, Hardoim CCP, van Overbeek LS, *et al.* Dynamics of Seed-Borne Rice Endophytes on Early Plant Growth Stages. *PloS one* 2012;**7**:e30438. doi:10.1371/journal.pone.0030438

4 Malfanova NV, Natuurwetenschappen F der W en. Endophytic bacteria with plant growth promoting and biocontrol abilities.

2013.https://openaccess.leidenuniv.nl/handle/1887/20732 (accessed 31 Jan2016).

5 Weilharter A, Mitter B, Shin MV, *et al.* Complete Genome Sequence of the Plant Growth-Promoting Endophyte Burkholderia phytofirmans Strain PsJN. *Journal of Bacteriology* 2011;**193**:3383–4. doi:10.1128/JB.05055-11

6 Deng Y, Zhu Y, Wang P, *et al.* Complete Genome Sequence of Bacillus subtilis BSn5, an Endophytic Bacterium of Amorphophallus konjac with Antimicrobial Activity for the Plant Pathogen Erwinia carotovora subsp. carotovora. *Journal of Bacteriology* 2011;**193**:2070–1. doi:10.1128/JB.00129-11

7 Pedrosa FO, Monteiro RA, Wassem R, *et al.* Genome of Herbaspirillum seropedicae Strain SmR1, a Specialized Diazotrophic Endophyte of Tropical Grasses. *PLoS Genetics* 2011;**7**:e1002064. doi:10.1371/journal.pgen.1002064

8 Krause A, Ramakumar A, Bartels D, *et al.* Complete genome of the mutualistic, N2-fixing grass endophyte Azoarcus sp. strain BH72. *Nat Biotech* 2006;**24**.

9 Bertalan M, Albano R, Pádua V, *et al.* Complete genome sequence of the sugarcane nitrogen-fixing endophyte Gluconacetobacter diazotrophicus Pal5. *BMC Genomics* 2009;**10**:1–17. doi:10.1186/1471-2164-10-450

10 Taghavi S, van der Lelie D, Hoffman A, *et al.* Genome sequence of the plant growth promoting endophytic bacterium Enterobacter sp. 638. *PLoS genetics* 2010;**6**:e1000943. doi:10.1371/journal.pgen.1000943

11 Taghavi S, Garafola C, Monchy S, *et al.* Genome Survey and Characterization of Endophytic Bacteria Exhibiting a Beneficial Effect on Growth and Development of Poplar Trees. *Applied and Environmental Microbiology* 2009;**75**:748–57. doi:10.1128/AEM.02239-08

12 Ryan RP, Monchy S, Cardinale M, *et al.* The versatility and adaptation of bacteria from the genus Stenotrophomonas. *Nat Rev Micro* 2009;**7**:514–25. doi:10.1038/nrmicro2163

13 Suzuki T, Shimizu M, Meguro A, *et al.* Visualization of Infection of an Endophytic Actinomycete *Streptomyces galbus* in Leaves of Tissue-cultured Rhododendron. *Actinomycetologica* 2005;**19**:7–12. doi:10.3209/saj.19.7

14 Compant S, Duffy B, Nowak J, *et al.* Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Applied and Environmental Microbiology* 2005;**71**:4951–9. doi:10.1128/AEM.71.9.4951-4959.2005

15 Coutinho TA, Venter SN. Pantoea ananatis: an unconventional plant pathogen. *Molecular plant pathology* 2009;**10**:325–35. doi:10.1111/j.1364-3703.2009.00542.x

16 Mitter B, Petric A, Shin MW, *et al.* Comparative genome analysis of Burkholderia phytofirmans PsJN reveals a wide spectrum of endophytic lifestyles based on interaction strategies with host plants. *Frontiers in plant science* 2013;**4**:120. doi:10.3389/fpls.2013.00120

17 Compant S, Kaplan H, Sessitsch A, *et al.* Endophytic colonization of Vitis vinifera L. by Burkholderia phytofirmans strain PsJN: from the rhizosphere to inflorescence tissues. *FEMS microbiology ecology* 2008;**63**:84–93. doi:10.1111/j.1574-6941.2007.00410.x

18 Helmann JD. The extracytoplasmic function (ECF) sigma factors. *Advances in microbial physiology* 2002;**46**:47–110.

19 Gourion B, Sulser S, Frunzke J, *et al.* The PhyR-sigma(EcfG) signalling cascade is involved in stress response and symbiotic efficiency in Bradyrhizobium japonicum. *Molecular microbiology* 2009;**73**:291–305. doi:10.1111/j.1365-2958.2009.06769.x

20 Zeidler D, Zähringer U, Gerber I, *et al.* Innate immunity in Arabidopsis thaliana: Lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. *Proceedings of the National Academy of Sciences of the United States of America* 2004;**101**:15811–6. doi:10.1073/pnas.0404536101

21 Sturz AV, Christie BR, Nowak J. Bacterial Endophytes: Potential Role in Developing Sustainable Systems of Crop Production. *Critical Reviews in Plant Sciences* 2000;**19**:1–30. doi:10.1080/07352680091139169

22 Sessitsch A, Reiter B, Berg G. Endophytic bacterial communities of field-grown potato plants and their plant-growth-promoting and antagonistic abilities. *Canadian journal of microbiology* 2004;**50**:239–49. doi:10.1139/w03-118

23 SASAKI K, IKEDA S, EDA S, *et al.* Impact of plant genotype and nitrogen level on rice growth response to inoculation with Azospirillum sp. strain B510 under paddy field conditions. *Soil Science & Plant Nutrition* 2010;**56**:636–44. doi:10.1111/j.1747-0765.2010.00499.x

24 Ambrosini A, Beneduzi A, Stefanski T, *et al.* Screening of plant growth promoting Rhizobacteria isolated from sunflower (Helianthus annuus L.). *Plant and Soil* 2012;**356**:245–64. doi:10.1007/s11104-011-1079-1

25 Arruda L, Beneduzi A, Martins A, *et al.* Screening of rhizobacteria isolated from maize (Zea mays L.) in Rio Grande do Sul State (South Brazil) and analysis of their potential to improve plant growth. *Applied Soil Ecology* 2013;**63**:15–22. doi:10.1016/j.apsoil.2012.09.001

26 Anwar MS, Siddique MT, Verma A, *et al.* Multitrait plant growth promoting (PGP) rhizobacterial isolates from Brassica juncea rhizosphere : Keratin degradation and growth promotion. *Communicative & Integrative Biology* 2014;**7**:e27683. doi:10.4161/cib.27683

27 Spaepen S, Das F, Luyten E, *et al.* Indole-3-acetic acid-regulated genes in Rhizobium etli CNPAF512. *FEMS Microbiology Letters* 2009;**291**:195–200. doi:10.1111/j.1574-6968.2008.01453.x

28 Ahemad M, . MSK. Functional Aspects of Plant Growth Promoting Rhizobacteria: Recent Advancements. *Insight Microbiology* 2011;**1**:39–54. doi:10.5567/IMICRO-IK.2011.39.54

29 Glick BR. Plant Growth-Promoting Bacteria: Mechanisms and Applications. *Scientifica* 2012;**2012**:15. doi:10.6064/2012/963401

30 Spaepen S, Vanderleyden J. Auxin and Plant-Microbe Interactions. *Cold Spring Harbor Perspectives in Biology* 2011;**3**:a001438. doi:10.1101/cshperspect.a001438

31 Tsavkelova EA, Klimova SI, Cherdyntseva TA, *et al.* [Microbial producers of plant growth stimulators and their practical use: A review]. *Prikl Biokhim Mikrobiol* 2006;**42**:133–43.

32 Long HH, Schmidt DD, Baldwin IT. Native Bacterial Endophytes Promote Host Growth in a Species-Specific Manner; Phytohormone Manipulations Do Not Result in Common Growth Responses. *PLoS ONE* 2008;**3**:e2702. doi:10.1371/journal.pone.0002702

33 Ivanchenko MG, Muday GK, Dubrovsky JG. Ethylene–auxin interactions regulate lateral root initiation and emergence in Arabidopsis thaliana. *The Plant Journal* 2008;**55**:335–47. doi:10.1111/j.1365-313X.2008.03528.x

34 Glick BR. Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiological Research* 2014;**169**:30–9. doi:http://dx.doi.org/10.1016/j.micres.2013.09.009

35 Sun Y, Cheng Z, Glick BR. The presence of a 1-aminocyclopropane-1-carboxylate (ACC) deaminase deletion mutation alters the physiology of the endophytic plant growth-promoting bacterium Burkholderia phytofirmans PsJN. *FEMS Microbiol Lett* 2009;**296**:131–6. doi:10.1111/j.1574-6968.2009.01625.x

36 Gasser I, Cardinale M, Müller H, *et al.* Analysis of the endophytic lifestyle and plant growth promotion of Burkholderia terricola ZR2-12. *Plant Soil* 2011;**347**:125–36. doi:10.1007/s11104-011-0833-8

37 Rothballer M, Eckert B, Schmid M, *et al.* Endophytic root colonization of gramineous plants by Herbaspirillum frisingense. *FEMS Microbiol Ecol* 2008;**66**:85–95. doi:10.1111/j.1574-6941.2008.00582.x

38 Pedraza RO. Recent advances in nitrogen-fixing acetic acid bacteria. *Int J Food Microbiol* 2008;**125**:25–35. doi:10.1016/j.ijfoodmicro.2007.11.079

39 Yang J, Kloepper JW, Ryu CM. Rhizosphere bacteria help plants tolerate abiotic stress. *Trends in plant science* 2009;**14**:1–4. doi:10.1016/j.tplants.2008.10.004

40 Lopez BR, Bashan Y, Bacilio M. Endophytic bacteria of Mammillaria fraileana, an endemic rock-colonizing cactus of the southern Sonoran Desert. *Arch Microbiol* 2011;**193**:527–41. doi:10.1007/s00203-011-0695-8

41 James EK. Nitrogen fixation in endophytic and associative symbiosis. *Field Crops Research* 2000;**65**:197–209. doi:http://dx.doi.org/10.1016/S0378-4290(99)00087-8

42 Sevilla M, Burris RH, Gunapala N, *et al.* Comparison of benefit to sugarcane plant growth and 15N2 incorporation following inoculation of sterile plants with Acetobacter diazotrophicus wildtype and Nif- mutant strains. *Molecular plant-microbe interactions : MPMI* 2001;**3**. doi:10.1094/MPMI.2001.14.3.358

43 Quecine MC, Araujo WL, Rossetto PB, *et al.* Sugarcane growth promotion by the endophytic bacterium Pantoea agglomerans 33.1. *Applied and Environmental Microbiology* 2012;**78**:7511–8. doi:10.1128/AEM.00836-12

44 Berg G, Hallmann J. Control of plant pathogenic fungi with bacterial endophytes. Microbial Root Endophytes (Schulz BJE, Boyle CJC & Sieber TN, eds). *Springer Verlag, Berlin* 2006;:53–69.

45 Kloepper J, Ryu C-M. Bacterial Endophytes as Elicitors of Induced Systemic Resistance. In: Schulz BE, Boyle CC, Sieber T, eds. *Microbial Root Endophytes*. Springer Berlin Heidelberg 2006. 33–52.

46 Yousaf S, Afzal M, Anees M, *et al.* Ecology and Functional Potential of Endophytes in Bioremediation: A Molecular Perspective. In: Verma VC, Gange AC, eds. *Advances in Endophytic Research*. Springer India 2014. 301–20.http://link.springer.com/chapter/10.1007/978-81-322-1575-2_16 (accessed 18 Feb2016).

47 Souza R de, Meyer J, Schoenfeld R, *et al.* Characterization of plant growth-promoting bacteria associated with rice cropped in iron-stressed soils. *Ann Microbiol* 2014;**65**:951–64. doi:10.1007/s13213-014-0939-3

48 Naveed M, Mitter B, Reichenauer TG, *et al.* Increased drought stress resilience of maize through endophytic colonization by Burkholderia phytofirmans PsJN and Enterobacter sp. FD17. *Environmental and Experimental Botany* 2014;**97**:30–9. doi:10.1016/j.envexpbot.2013.09.014

49 Naveed M, Hussain MB, Zahir ZA, *et al.* Drought stress amelioration in wheat through inoculation with Burkholderia phytofirmans strain PsJN. *Plant Growth Regulation* 2013;**73**:121–31. doi:10.1007/s10725-013-9874-8

50 Ali SZ, Sandhya V, Grover M, *et al.* Pseudomonas sp. strain AKM-P6 enhances tolerance of sorghum seedlings to elevated temperatures. *Biol Fertil Soils* 2009;**46**:45–55. doi:10.1007/s00374-009-0404-9

51 Bal HB, Das S, Dangar TK, *et al.* ACC deaminase and IAA producing growth promoting bacteria from the rhizosphere soil of tropical rice plants. *J Basic Microbiol* 2013;**53**:972–84. doi:10.1002/jobm.201200445

52 Figueiredo MVB, Burity HA, Martínez CR, *et al.* Alleviation of drought stress in the common bean (Phaseolus vulgaris L.) by co-inoculation with Paenibacillus polymyxa and Rhizobium tropici. *Applied Soil Ecology* 2008;**40**:182–8. doi:10.1016/j.apsoil.2008.04.005

53 Dimkpa C, Weinand T, Asch F. Plant–rhizobacteria interactions alleviate abiotic stress conditions. *Plant, Cell & Environment* 2009;**32**:1682–94. doi:10.1111/j.1365-3040.2009.02028.x

54 Grover M, Ali SZ, Sandhya V, *et al.* Role of microorganisms in adaptation of agriculture crops to abiotic stresses. *World J Microbiol Biotechnol* 2010;**27**:1231–40. doi:10.1007/s11274-010-0572-7

55 McAdam PR, Richardson EJ, Fitzgerald JR. High-throughput sequencing for the study of bacterial pathogen biology. *Curr Opin Microbiol* 2014;**19**:106–13. doi:10.1016/j.mib.2014.06.002

56 Frank AC. The Genomes of Endophytic Bacteria. In: Pirttilä AM, Frank AC, eds. *Endophytes of Forest Trees*. Springer Netherlands 2011. 107–36.

57 Van Aken B. Methylobacterium populi sp. nov., a novel aerobic, pink-pigmented, facultatively methylotrophic, methane-utilizing bacterium isolated from poplar trees (Populus deltoidesxnigra DN34). *International journal of systematic and evolutionary microbiology* 2004;**54**:1191–6. doi:10.1099/ijs.0.02796-0

58 Kaneko T, Minamisawa K, Isawa T, *et al.* Complete Genomic Structure of the Cultivated Rice Endophyte Azospirillum sp. B510. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes* 2010;**17**:37–50. doi:10.1093/dnares/dsp026

59 Wisniewski-Dyé F, Borziak K, Khalsa-Moyers G, *et al.* Azospirillum Genomes Reveal Transition of Bacteria from Aquatic to Terrestrial Environments. *PLoS Genetics* 2011;**7**:e1002430. doi:10.1371/journal.pgen.1002430

60 Krause A, Ramakumar A, Bartels D, *et al.* Complete genome of the mutualistic, N2-fixing grass endophyte Azoarcus sp. strain BH72. *Nat Biotech* 2006;**24**.http://dx.doi.org/10.1038/nbt1243

61 Kwak M-J, Song JY, Kim S-Y, *et al.* Complete Genome Sequence of the Endophytic Bacterium Burkholderia sp. Strain KJ006. *Journal of Bacteriology* 2012;**194**:4432–3. doi:10.1128/JB.00821-12

62 Han J-I, Spain JC, Leadbetter JR, *et al.* Genome of the Root-Associated Plant Growth-Promoting Bacterium Variovorax paradoxus Strain EPS. *Genome Announcements* 2013;**1**:e00843–13 – e00843–13. doi:10.1128/genomeA.00843-13

63 Yan Y, Yang J, Dou Y, *et al.* Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated Pseudomonas stutzeri A1501. *Proceedings of the National Academy of Sciences* 2008;**105**:7564–9. doi:10.1073/pnas.0801093105

64 Fouts DE, Tyler HL, Boy RT, *et al.* Complete genome sequence of the N2-fixing broad host range endophyte Klebsiella pneumoniae 342 and virulence predictions verified in mice. *PLoS Genet* 2008;**47**. doi:10.1371/journal.pgen.1000141

65 Liu W-Y, Chung KM-K, Wong C-F, *et al.* Complete Genome Sequence of the Endophytic Enterobacter cloacae subsp. cloacae Strain ENHKU01. *Journal of Bacteriology* 2012;**194**:5965–5965. doi:10.1128/JB.01394-12

66 Santoyo G, Moreno-Hagelsieb G, del Carmen Orozco-Mosqueda M, *et al.* Plant growth-promoting bacterial endophytes. *Microbiological Research* 2016;**183**:92–9. doi:10.1016/j.micres.2015.11.008

67 Hasman H, Saputra D, Sicheritz-Ponten T, *et al.* Rapid Whole-Genome Sequencing for Detection and Characterization of Microorganisms Directly from Clinical Samples. *Journal of Clinical Microbiology* 2014;**52**:139–46. doi:10.1128/JCM.02452-13

68 Koren S, Treangen TJ, Hill CM, *et al.* Automated ensemble assembly and validation of microbial genomes. *BMC bioinformatics* 2014;**15**:1–9. doi:10.1186/1471-2105-15-126

69 Seemann T. Prokka: rapid prokaryotic genome annotation. In: *Bioinformatics*. 2014.

70 Compeau PEC, Pevzner PA, Tesler G. How to apply de Bruijn graphs to genome assembly. *Nat Biotechnol* 2011;**29**:987–91. doi:10.1038/nbt.2023

71 Edwards DJ, Holt KE. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microbial Informatics and Experimentation* 2013;**3**:2–2. doi:10.1186/2042-5783-3-2

72 Simon C, Daniel R. Metagenomic Analyses: Past and Future Trends. *Applied and Environmental Microbiology* 2011;**77**:1153–61. doi:10.1128/AEM.02345-10

73 Oulas A, Pavloudi C, Polymenakou P, *et al.* Metagenomics: Tools and Insights for Analyzing Next-Generation Sequencing Data Derived from Biodiversity Studies. *Bioinformatics and Biology Insights* 2015;**9**:75–88. doi:10.4137/BBI.S12462

74 Akinsanya MA, Goh JK, Lim SP, *et al.* Metagenomics study of endophytic bacteria in Aloe vera using next-generation technology. *Genom Data* 2015;**6**:159–63. doi:10.1016/j.gdata.2015.09.004

75 Sessitsch A, Hardoim P, Doring J, *et al.* Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Molecular plant-microbe interactions : MPMI* 2012;**25**:28–36. doi:10.1094/mpmi-08-11-0204

76 Creecy JP, Conway T. Quantitative bacterial transcriptomics with RNA-seq. *Current Opinion in Microbiology* 2015;**23**:133–40. doi:10.1016/j.mib.2014.11.011

77 McClure R, Balasubramanian D, Sun Y, *et al.* Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Research* Published Online First: 28 May 2013. doi:10.1093/nar/gkt444

78 Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 2010;**11**:R25–R25. doi:10.1186/gb-2010-11-3-r25

79 Bullard JH, Purdom E, Hansen KD, *et al.* Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC bioinformatics* 2010;**11**:94. doi:10.1186/1471-2105-11-94

80 Mortazavi A, Williams BA, McCue K, *et al.* Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Meth* 2008;**5**:621–8. doi:10.1038/nmeth.1226

81 Trapnell C, Williams BA, Pertea G, *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech* 2010;**28**:511–5. doi:10.1038/nbt.1621

82 Anders S, Huber W. Differential expression analysis for sequence count data. *Genome biology* 2010;**11**:R106. doi:10.1186/gb-2010-11-10-r106

83 Tarazona S, Garcia-Alcalde F, Dopazo J, *et al.* Differential expression in RNA-seq: a matter of depth. *Genome research* 2011;**21**:2213–23. doi:10.1101/gr.124321.111

84 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014;**15**:550. doi:10.1186/s13059-014-0550-8

85 Schenk PM, Carvalhais LC, Kazan K. Unraveling plant–microbe interactions: can multi-species transcriptomics help? *Trends in Biotechnology* 2012;**30**:177–84. doi:http://dx.doi.org/10.1016/j.tibtech.2011.11.002

86 Liu J, Wang E, Ren D, *et al.* Mixture of endophytic Agrobacterium and Sinorhizobium meliloti strains could induce nonspecific nodulation on some woody legumes. *Archives of microbiology* 2010;**192**:229–34. doi:10.1007/s00203-010-0543-2

87 Delmotte N, Ahrens CH, Knief C, *et al.* An integrated proteomics and transcriptomics reference data set provides new insights into the Bradyrhizobium japonicum bacteroid metabolism in soybean root nodules. *Proteomics* 2010;**10**:1391–400. doi:10.1002/pmic.200900710

88 Compant S, Mitter B, Colli-Mull JG, *et al.* Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microbial ecology* 2011;**62**:188–97. doi:10.1007/s00248-011-9883-y

89 Suarez-Moreno ZR, Caballero-Mellado J, Coutinho BG, *et al.* Common features of environmental and potentially beneficial plant-associated Burkholderia. *Microbial ecology* 2012;**63**:249–66. doi:10.1007/s00248-011-9929-1main aim of this study. These analyses can help us to test endophytic bacterial ability to promote plant growth and tolerate the external stress such as drought stress.

Chapter II

Synopsis of the Publications/Manuscripts

Chapter III gain insight into the whole transcriptome sequencing of *B. phytofirmans* PsJN inside cv. Bionta potato plants (*in vitro*) as they were exposed to drought stress condition in different time points that alter their environment. RNA-seq was used to analyze *in planta* gene activity, gene expression patern and the response of strain PsJN to plant stress. The transcriptome of PsJN colonizing *in vitro* potato plants showed a broad array of functionalities encoded on the genome of strain PsJN. High throuput RNA-seq analysis identified transcripts up-regulated in response to plant drought stress were mainly involved in transcriptional regulation, cellular homeostasis and the detoxification of reactive oxygen species, indicating oxidative stress response in PsJN. This analysis confirmed that the genes with modulated expression included genes for extracytoplasmatic function (ECF) group IV sigma factors. These cell surface signaling elements allow bacteria to sense changing environmental conditions and to adjust their metabolism accordingly. TaqMan-qPCR was performed to identify ECF sigma factors in PsJN that were activated in response to plant stress. Six ECF sigma factor genes were expressed in PsJN colonizing potato plants. This study indicates that endophytic *B. phytofirmans* PsJN cells are active inside plants. Moreover, the activity of strain PsJN is affected by plant drought stress; it senses plant stress signals and adjusts its gene expression accordingly.

Authors names: **Raheleh Sheibani-Tezerji**, Thomas Rattei, Angela Sessitsch, Friederike Trognitz, and Birgit Mitter

Manuscript Title: **Transcriptome profiling of the endophyte *Burkholderia phytofirmans* PsJN indicates sensing of the plant environment and drought stress**

Reference: mBio 2015; 6(5): e00621-15.

Contributions: B. Mitter defined the research project. Experiments were designed and performed by R. Sheibani-Tezerji under supervission of B. Mitter except for plant colonization that was done by F. Trognitz. R. Sheibani-Tezerji performed the bioinformatic analysis under supervission of T. Rattei. R. Sheibani-Tezerji designed figures. B. Mitter and R. Sheibani-Tezerji wrote the manuscript that was revised by T. Rattei and A. Sessitsch.

Chapter IV provides a detailed description of comparative genome analysis of three closely related *P. ananatis* strains, which were isolated from maize seeds of healthy plants. Plant inoculation experiments such as testing the effect of endophytic strains and functional characterization of them on maize seed revealed that each of these strains exhibited a different phenotype ranging from pathogenic (S7), commensal (S8), to a beneficial, growth-promoting effect (S6) in maize. We performed a comparative genomics analysis starting with genome assembly, genome annotation, phylogenetic analysis, plasmid sequence alignment analysis for each strain following by identification of orthologous and eukaryotic-like protein domains of the three strains, in order to find genetic determinants responsible for the differences observed. Our findings showed that these three *P. ananatis* strains colonizing the same ecological niche but showing distinct interaction strategies with the host plant. Contrasting the genomes of three strains of *P. ananatis* revealed that they are highly similar. However, genomic differences in genes encoding protein secretion systems and putative effectors, and transposase/integrases/phage related genes could be observed that indicate molecular mechanisms for the different phenotypes.

Authors names: **Raheleh Sheibani-Tezerji**, Muhammad Naveed, Marc-André Jehl, Angela Sessitsch, Thomas Rattei and Birgit Mitter

Manuscript title: **Comparative genome analysis of closely related *Pantoea ananatis* seed endophyte strains having different effects on the host plant**

Reference: Frontiers in Microbiology 2015; 6: 440

Contributions: B. Mitter and A. Sessitsch defined the research topic. M. Naveed performed the laboratory work. R. Sheibani-Tezerji performed all the bio-informatic analyses (except than for ELD analysis that was done by M.-A. Jehl) under the supervision of B. Mitter and T. Rattei. R. Sheibani-Tezerji drafted the manuscript with contributions of M. Naveed and M.-A. Jehl. The paper was revised by B. Mitter, T. Rattei and A. Sessitsch.

Chapter V describes the means of introducing new microbes into seeds to modify the plant microbiome and traits in defined ways. A completely new approach is applied to study the roles of members of seed microbiomes, and changing the microbiome of elite crop seed embryos, both monocots and dicots. Selected microbes are introduced into the parent plant before seed development occurs (by spraying the bacteria on the parent flowers, which enter the plant and colonize the emerging seeds) and these become incorporated into the seed microbiome. Furthermore, the internally colonized seeds were planted, which the bacteria get activated and colonize the offspring generation plant from the first day that can have effects on the growth regulation of crops. The new approach has many advantages over mixing seeds with the microbes before planting, and should have significant impact on fundamental endophyte-plant interaction studies as well as plant-microbe optimization in agriculture.

Authors names: Birgit Mitter, Nikolaus Pfaffenbichler, Richard Flavell, Stéphane Compant, Livio Antonielli, Alexandra Petric, Teresa Berninger, Naveed Muhammad, **Raheleh Sheibani-Tezerji**, Geoffrey von Maltzahn and Angela Sessitsch

Manuscript title: **Modification of a plant microbiome and traits by adding new microbes to seeds during seed production**

Reference: Under submission

Contributions: B. Mitter and A.Sessitsch conceived and designed the research. N. Pfaffenbichler., L. Antonielli. and B. Mitter analyzed data, N. Pfaffenbichler., performed DNA isolation, PCR and qPCR, plant experiments and designed figures, S. Compant and T. Berninger designed FISH probe and performed FISH experiments, A. Petric performed Illumina amplicon sequencing, L. Antonielli. analyzed sequencing data, **R. Sheibani-Tezerji** developed the qPCR system, B. Mitter, R. Flavell and G.v.Maltzahn wrote the paper.

Chapter III

Transcriptome profiling of the endophyte *Burkholderia phytofirmans* PsJN indicates sensing of the plant environment and drought stress

mBio 2015

6(5): e00621-15

Transcriptome profiling of the endophyte *Burkholderia phytofirmans* PsJN indicates sensing of the plant environment and drought stress

Raheleh Sheibani-Tezerji[a,b], Thomas Rattei[b], Angela Sessitsch[a], Friederike Trognitz[a], and Birgit Mitter[a]

[a]Health & Environment Department, Bioresources Unit, AIT Austrian Institute of Technology GmbH, Konrad-Lorenz-Strasse 24, 3430 Tulln, Austria
[b]University of Vienna, Department of Microbiology and Ecosystem Science, Division of Computational System Biology, Althanstrasse 14, 1090 Vienna, Austria

Address correspondence to Birgit Mitter, birgit.mitter@ait.ac.at

Running title: *B. phytofirmans* PsJN – transcriptional response to plant stress

Keywords: *Burkholderia phytofirmans* PsJN, endophyte, plant-microbe interaction, ECF sigma factor, RNAseq, transcriptome, drought stress, potato

ABSTRACT

It is widely accepted that bacterial endophytes actively colonize plants, interact with their host, and frequently show beneficial effects on plant growth and health. However, the mechanisms of plant-endophyte communication and bacterial adaption to the plant environment are still poorly understood. Here, whole-transcriptome sequencing of *B. phytofirmans* PsJN colonizing potato (Solanum tuberosum L.) plants was used to analyze *in planta* gene activity and the response of strain PsJN to plant stress. The transcriptome of PsJN colonizing *in vitro* potato plants showed a broad array of functionalities encoded in the genome of strain PsJN. Transcripts upregulated in response to plant drought stress were mainly involved in transcriptional regulation, cellular homeostasis, and the detoxification of reactive oxygen species, indicating an oxidative stress response in PsJN. Genes with modulated expression included genes for extracytoplasmatic function (ECF) group IV sigma factors. These cell surface signaling elements allow bacteria to sense changing environmental conditions and to adjust their metabolism accordingly. TaqMan quantitative PCR (TaqMan-qPCR) was performed to identify ECF sigma factors in PsJN that were activated in response to plant stress. Six ECF sigma factor genes were expressed in PsJN colonizing potato plants. The expression of one ECF sigma factor was upregulated whereas that of another one was downregulated in a plant genotype-specific manner when the plants were stressed. Collectively, our study results indicate that endophytic *B. phytofirmans* PsJN cells are active inside plants. Moreover, the activity of strain PsJN is affected by plant drought stress; it senses plant stress signals and adjusts its gene expression accordingly.

Importance

In recent years, plant growth-promoting endophytes have received steadily growing interest as an inexpensive alternative to resource-consuming agrochemicals in sustainable agriculture. Even though promising effects are recurrently observed under controlled conditions, these are rarely reproducible in the field or show undesirably strong variations. Obviously, a better understanding of endophyte activities in plants and the influence of plant physiology on these activities is needed to develop more-successful application strategies. So far, research has focused mainly on analyzing the plant response to bacterial inoculants. This prompted us to study the gene expression of the endophyte *Burkholderia phytofirmans* PsJN in potato plants. We found that endophytic PsJN cells express a wide array of genes

and pathways, pointing to high metabolic activity inside plants. Moreover, the strain senses changes in the plant physiology due to plant stress and adjusts its gene expression pattern to cope with and adapt to the altered conditions.

INTRODUCTION

Burkholderia phytofirmans PsJN (1) is a naturally occurring endophyte isolated from onion roots (2) that is able to establish nonpathogenic relationships with a wide range of plant species, both monocotyledons and dicotyledons (3, 4). Numerous studies have reported beneficial effects of strain PsJN on host plants such as increased plant growth (for a review, see reference 5) and enhanced biotic and abiotic stress tolerance (4, 5).

There are increasing efforts to understand the plant physiological and genetic response to inoculation with *B. phytofirmans* PsJN. Bordiec and colleagues (6) compared local defense reactions in grapevine cell cultures inoculated with either strain PsJN or the nonhost pathogen Pseudomonas syringae pv. pisi. The authors found that strain PsJN does not induce the defense events commonly found after pathogen attack in plants. Infection with the pathogen led to a two-phase oxidative burst and a hypersensitive response (HR)-like reaction, whereas no oxidative burst or cell death was observed in cells treated with strain PsJN (6). Theocharis et al. (7) demonstrated that inoculation of Chardonnay grapevine plantlets with *B. phytofirmans* PsJN speeds up the plant response to chilling and plant adaption to cold temperatures. Numbers of cold stress-related gene transcripts and metabolites increased earlier and faster and reached higher levels in bacterized plantlets than in control plants. Fernandez et al. (8) demonstrated that the higher tolerance of chilling of PsJN-colonized grapevine plantlets could be related to changes in plant photosynthesis and sugar metabolism. More recently, Poupin and colleagues (9) studied the response of Arabidopsis thaliana to inoculation with *B. phytofirmans* PsJN. The bacterium affected the whole life cycle of Arabidopsis plants, increasing plant growth, especially at the early stage of ontogeny, and speeding up maturity. These physiological changes correlated with altered expression of plant growth regulator genes; i.e., genes involved in auxin and gibberellin pathways were induced in PsJN-inoculated plants, and flowering-control genes were induced earlier in PsJN-inoculated plants than in control plants.

Whereas the plant response to beneficial bacteria has been described in several studies, very little is known about bacterial physiology, response, and adaptation processes *in planta*. Efforts to characterize the effects of the plant environment on endophytic bacteria have been rare (10–12), and information on *in planta* gene expression of *B. phytofirmans* PsJN is missing. For example, how does *B. phytofirmans* PsJN recognize the plant environment? Does the bacterium respond to changing physiological conditions, e.g., due to plant stress, in plants?

Therefore, the aim of this study was to investigate gene expression patterns of *B. phytofirmans* PsJN cells colonizing potato (Solanum tuberosum L.) plants and, furthermore, to assess the effect of plant drought stress on the transcriptome of strain PsJN. *In vitro*-grown potato plants of two varieties (Russet Burbank and Bionta) were inoculated with *B. phytofirmans* PsJN, and drought stress was induced by application of polyethylene glycol (PEG). Bacterial transcriptomes of cells colonizing potato plants (cv. Bionta) were analyzed under nonstressed conditions (control) and at three different time points after drought stress induction by direct short-read deep sequencing (Illumina RNA-seq). Differentially expressed genes included genes for extracytoplasmatic function (ECF) group IV sigma factors. TaqMan quantitative PCR (TaqMan-qPCR) assays were performed to quantitatively assess ECF sigma factor activation in *B. phytofirmans* PsJN colonizing potato plants of two varieties (Russet Burbank and Bionta) in response to plant stress.

RESULTS

Detection of *B. phytofirmans* PsJN in plants. Six weeks after inoculation with *B. phytofirmans* PsJN, potato plants of two varieties (Bionta and Russet Burbank) showed increased shoot and root length in response to colonization by PsJN (see Fig. S1 in the supplemental material). Application of PEG caused visible symptoms of water deficiency in potato plants (see Fig. S2). PCR amplification with primers targeting bacterial 16S rRNA genes and plant 18S rRNA genes resulted in two bands in all inoculated plants, the mitochondrial band and a band of about 720 bp representing the bacterial 16S rRNA gene. No amplification of the bacterial fragment was found in control plants (see Fig. S3). Isolation and sequencing of the bacterial bands confirmed the presence of *B. phytofirmans* PsJN in all inoculated potato plants.

Transcriptome sequencing. Sequencing of cDNA samples yielded 35.7 to 42.2 million reads of cDNA, corresponding to over 2 billion nucleotides of cDNA per sample (Table 1). Around 10% of the reads were removed by initial quality filters to trim poor-quality data. After poly(A) tail removal and length trimming, more than 37% of the reads were removed, mainly because the bacterial mRNA molecules were poly(A) tailed during cDNA library preparation. Around 61% to 63% of total reads were considered for further analysis. By removal of rRNA sequences, the data set was further reduced by 5%. Of the remaining sequences, 0.3% to 1.95% of the reads mapped to the genome of PsJN (Table 1). Normalization of transcript levels in control and stressed plants was done by RPKM (reads per kilobase per million mapped) normalization using NOISeq. For verification of this procedure, we used the expression data of selected housekeeping genes which were shown by qPCR to be stably expressed (see Fig. S4 in the supplemental material).

TABLE 1 Statistics of cDNA sequencing reads and their alignment on the *B. phytofirmans* PsJN genome.

Sample	Total nucleotide	RNA-seq reads	Trimmed reads[1]	rRNA reads	Total reads[2]	Mapped_Ps[3]
Control	1,865,634,927	36,581,077	20,759,363	1141647	19,617716	93404 (0.48%)
Stress 1	2,156,419,995	42,282,745	24,171,466	1358716	22,812750	382477 (1.70%)
Stress 6	1,825,100,280	35,786,280	20,023,318	1036065	18,987253	56604 (0.30%)
Stress	1,926,749,451	37,779,401	20,048,099	1266245	18,781854	64047

Transcriptomic profile of *B. phytofirmans* PsJN in potato plants. A total number of 4,591 different transcripts of *B. phytofirmans* PsJN were detected in PsJN-colonized potato plants. The expressed genes were evenly distributed on both chromosomes, and 76 of 167 genes carried on the plasmid were active in this experiment (Fig. 1). Among the latter, we found genes for type II B. phytofirmans 7353 (Bphyt_7353) and type IV (Bphyt_7341, Bphyt_7342, Bphyt_7347, Bphyt_7350, and Bphyt_7351) secretion system proteins.

Figure 2 shows a comparison of the genome of *B. phytofirmans* PsJN and the transcriptome of PsJN colonizing *in vitro* potato plants on the basis of the relative distributions of clusters of orthologous group (COG) categories. Overall, the functional categories of the *in planta* transcriptome and the genome of strain PsJN were highly similar. Few differences were found in the following COG categories: cell motility, defense mechanisms, and extracellular structures. The relative abundance of cell motility- and defense mechanism-related genes was reduced from 2% in the

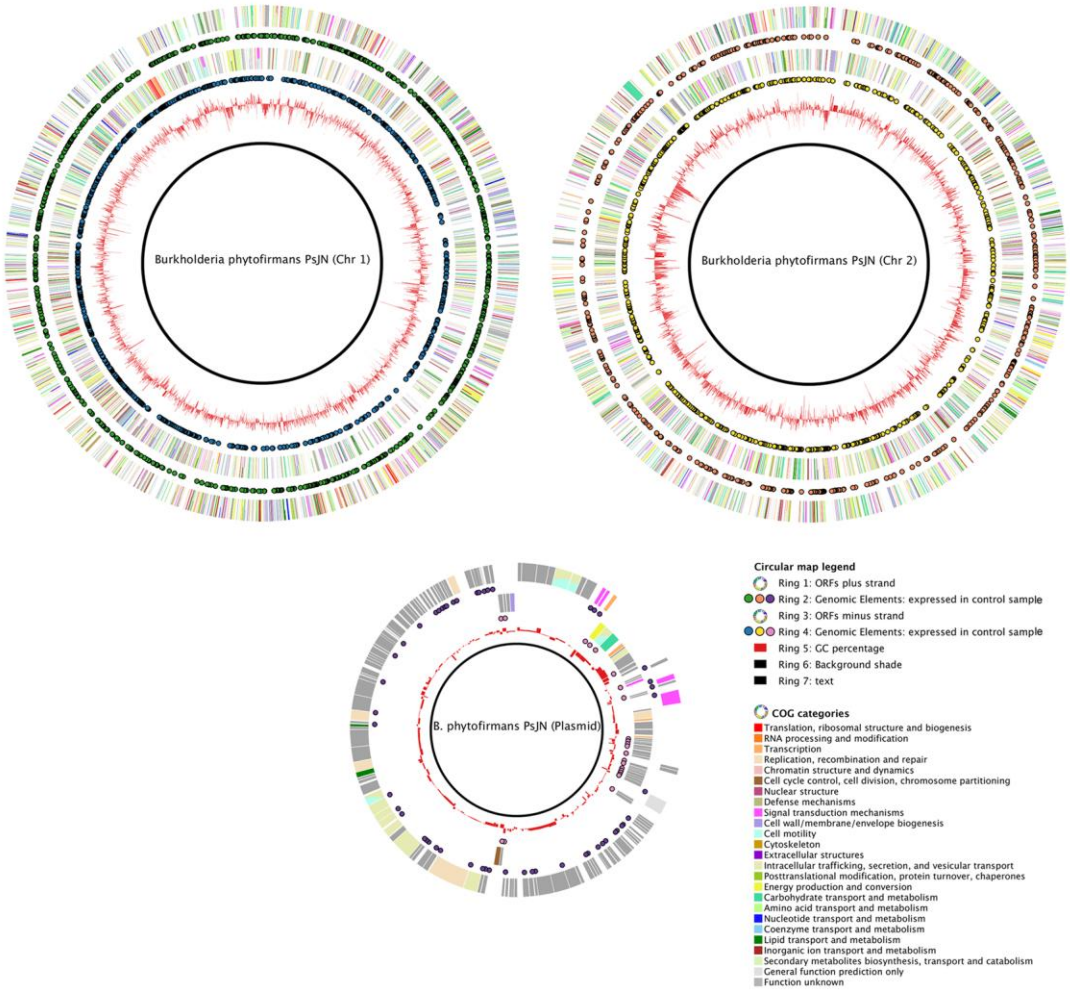


FIG 1. Circular maps representing the two chromosomes and the plasmid of *Burkholderia phytofirmans* PsJN. The following rings are included in each map: open reading frames (ORFs) on the plus (rings 1) and minus (rings 3) strands of the genome of strain PsJN (color by COG categories). Transcripts expressed on the plus (rings 2) and minus (rings 4) strands of the genome of strain PsJN during colonization of unstressed in vitro potato plants are indicated. The images were generated with a microbial genome context viewer (MgcV; http://mgcv.cmbi.ru.nl/) (53).

genome to 1% in the transcriptome. Genes encoding proteins in the COG category of “extracellular structures” covered 1% of the genes in the genome of PsJN but were not found in the transcriptome.

Expressed genes with an RPKM value of ≥55 were grouped in 74 clusters representing 354 functional classes by gene ontology (GO) analysis (see Table S1 in the supplemental material). The majority of functions were related to transcription regulation, general metabolism (sugars, amino acids, lipids, and nucleotides), energy production, and cellular homeostasis. Furthermore, we found a high number of transcripts for signal transduction mechanisms such as two-component systems and

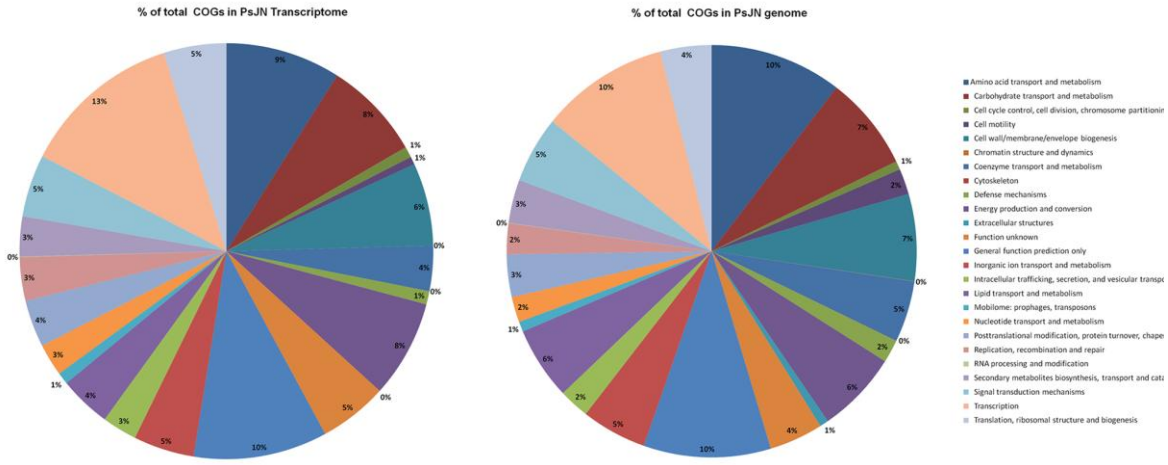


FIG 2. Relative distributions of functional COG categories in the *in planta* transcriptome and genome of *B. phytofirmans* PsJN.

extracytoplasmatic (ECF) sigma factor genes (Bphyt_1327, Bphyt_1666, Bphyt_1784, Bphyt_2906, Bphyt_3189, Bphyt_4397, Bphyt_4574, Bphyt_4980, Bphyt_5021, Bphyt_5131, and Bphyt_5142).

GO functions that were found in the PsJN transcriptome included functions generally considered important for endophytic plant colonization and plant growth promotion such as cell motility and chemotaxis, cellular iron homeostasis, and photosynthesis (see Table S1 in the supplemental material). Cellular iron homeostasis was represented mainly by ferritin (Bphyt_0714, Bphyt_2657, and Bphyt_5727) and bacterioferritin (Bphyt_1412 and Bphyt_2740) genes. By analyzing genes represented by the GO function “photosynthesis,” we found NADH dehydrogenase subunit A, B, C, and D (Bphyt_1343 to Bphyt_1346) and polyprenyl synthetase (Bphyt_3450) genes, which do not clearly indicate putative photosynthetic activity. Furthermore, we found expression of an N-acyl homoserine lactone (AHL) synthase gene (bpI.2; Bpyht_4275), a quinolinate phosphoribosyl transferase gene

(Bphyt_3152), indole-3-acetic acid (IAA) synthesis genes such as those encoding nitrile hydrolase (Bphyt_7182 and Bphyt_7181), and a gene encoding an IAM hydrolase of the indole-3-acetamide (IAM) pathway as well as aldehyde dehydrogenase (Bphyt_5803) of the tryptophan side-chain oxidase pathway bypassing IPyA. Transcripts of IAA degradation (aromatic ring hydroxylating dioxygenase; Bphyt_2156) and 1-aminocyclopropane-1-carboxylate (ACC) deaminase (Bphyt_5397) genes were not detected.

Transcriptional response of *B. phytofirmans* PsJN to plant stress.

(i) One hour after stress induction. Analysis of differentially expressed genes in *B. phytofirmans* PsJN colonizing nonstressed and drought-stressed potato plants identified 194 genes with modified expression 1 h after stress induction (see Table S2 in the supplemental material). Among these, 137 genes were upregulated and 57 genes were downregulated (Fig. 3).

For a better understanding of the genetic traits involved in the response of *B. phytofirmans* PsJN to drought stress of its host plant, the differentially expressed genes were affiliated to biological processes using gene ontology (GO) analysis (see Table S3 in the supplemental material). The differentially expressed genes represented eight different GO biological processes at 1 h after inducing drought stress. Genes that were upregulated in the transcriptome of *B. phytofirmans* PsJN belonged mostly to the following functional categories: cellular homeostasis, homeostatic process, and cell redox homeostasis. Among them were genes such as those encoding bacterioferritins, glutaredoxin, redoxin domain protein, thioredoxin, and RNA polymerase factor sigma 70 (ECF sigma factor; Bphyt_1327) (see Table S3). The downregulated genes represented regulation of transcription and DNA-dependent functions (see Table S3) such as those encoding various types of transcriptional regulators.

(ii) Six hours after stress induction. A total number of 354 genes were differentially expressed compared to the control at this time point, with 229 genes being upregulated and 125 genes being downregulated (Fig. 3). The complete list of differentially expressed genes is available (see Table S2 in the supplemental material). GO analysis grouped these genes into 79 different biological processes (see Table S3).

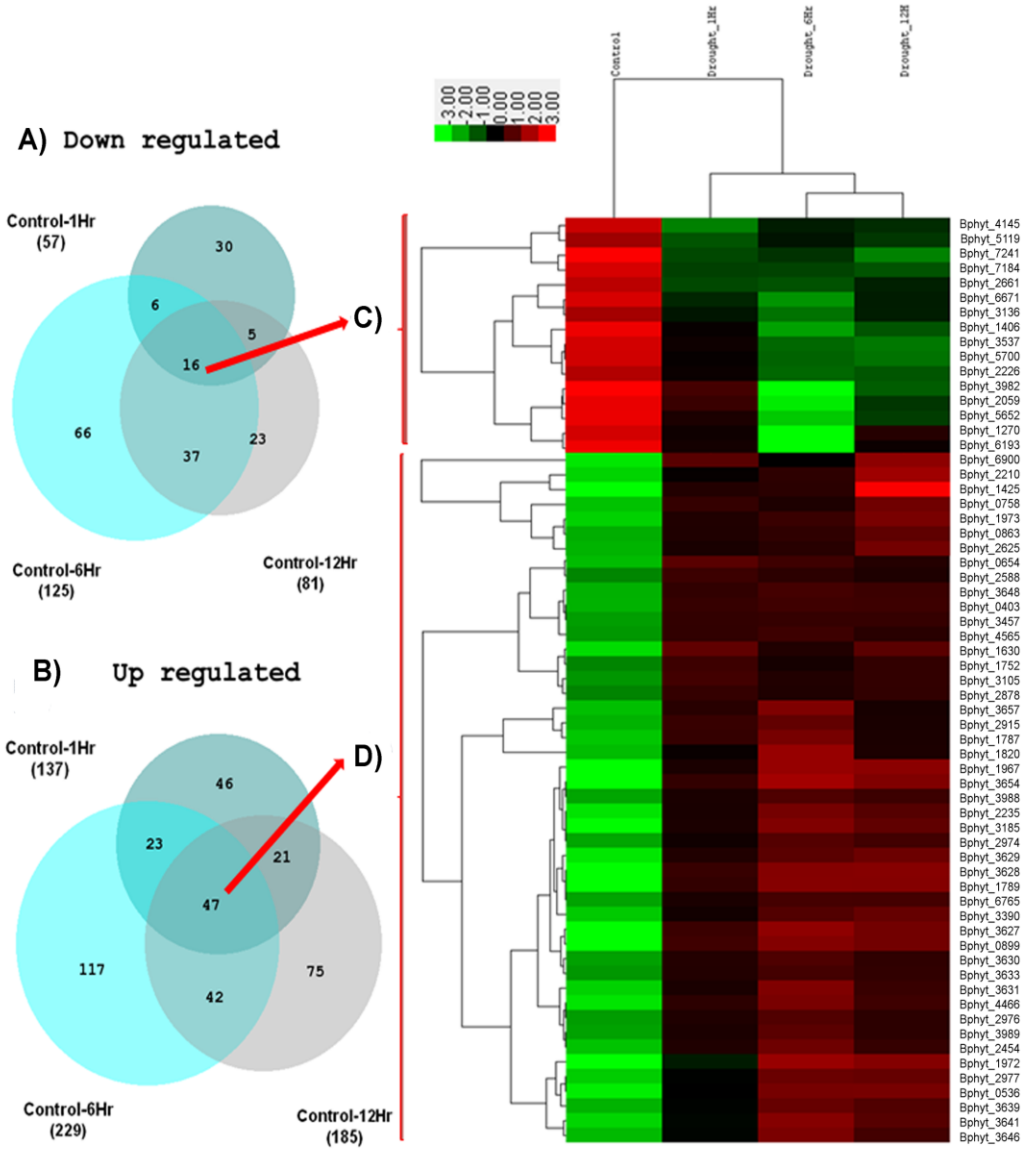


FIG 3. (Left) Venn diagrams illustrating numbers of downregulated (A) and upregulated (B) differentially expressed (DE) genes either shared or unique at all time points after drought stress induction from pairwise comparisons of control samples at each of the drought stress time points (control at 1 h [1Hr], control at 6 h [6Hr], and control at 12 h [12Hr]). The total numbers of up- or downregulated DE genes in each time point are indicated in parentheses. (Right) Hierarchical clustering heat map of expression changes for downregulated (C) and upregulated (D) differentially expressed genes that are common across all time points (16 genes in downregulated genes and 47 genes in upregulated genes). The heat map was constructed based on RPKM expression values. Rows correspond to differentially expressed genes, and columns represent control and drought-stressed samples at the indicated time points (in hours) of drought stress. Green and red boxes represent differentially expressed genes that decreased and increased their expression levels at the indicated time points of drought stress, respectively. The identifiers (ID) and descriptions of differentially expressed genes are listed at the right in the table.

The induced functions with the highest enrichment score were oxidative

phosphorylation, hydrogen transport, proton transport, ATP synthesis-coupled proton transport, energy-coupled proton transport, ion transmembrane transport, ATP biosynthetic and metabolic processes, purine nucleoside triphosphate biosynthetic processes, and ribonucleoside triphosphate biosynthetic and metabolic processes. The most enriched genes were those encoding ATP synthase C and gamma chain, ATP synthase subunit a/b/alpha/beta, sulfate adenylyltransferase, cytochrome o ubiquinol oxidase subunit III, succinate dehydrogenase, and acetolactate synthase (see Table S2).

(iii) Twelve hours after stress induction. At this time point, potato plants treated with PEG showed severe wilting and 266 genes were differentially expressed in *B. phytofirmans* PsJN (Fig. 3). The complete list of differentially expressed genes is available (see Table S2 in the supplemental material). Twelve hours after stress induction, 185 genes were upregulated and 81 genes were downregulated. These genes were grouped by GO analysis into 52 different biological processes (see Table S3). The GO functions showing highest enrichment at this time point were those corresponding to positive regulation of the cellular biosynthetic process, transcription, the macromolecule biosynthetic process, gene expression, the nitrogen compound metabolic process, and the macromolecule metabolic process. Peptidylprolyl isomerase FK506-binding protein (FKBP), UTP-glucose-1-phosphate uridylyltransferase, and glucose-6-phosphate dehydrogenase genes were upregulated at this time point under conditions of drought stress. Genes encoding histone family protein, transcriptional regulator GntR, transposase mutator type, peroxidases, and catalase/peroxidase (HPI) were downregulated at this time point (see Table S3).

(iv) Transcriptional response of *B. phytofirmans* PsJN to plant stress at all three time points. As shown by comparisons of the genes that are expressed in control plants and stressed plants, 47 genes were upregulated and 16 genes were downregulated in *B. phytofirmans* PsJN in response to plant stress at all three time points (Fig. 3). These genes were subjected to further hierarchical clustered analysis and were classified into four groups by Cluster 3.0 (13). The functions of these genes correspond to the functions with the highest enrichment score at all time points obtained using David (Database for Annotation, Visualization, and Integrated Discovery) software.

The first group consisted of two oxidoreductase activity-related genes (glutaredoxin and alkyl hydroperoxide reductase subunit) that are involved in cellular homeostasis and cell redox homeostasis. The second group represented genes involved in regulation of transcription activity and consisted of, among others, those encoding a GntR family transcriptional regulator and a cold-shock DNA-binding protein. The third group consisted of UTP-glucose-1-phosphate uridylyltransferase and glyceraldehyde-3-phosphate dehydrogenase, which showed similar (upregulated) expression patterns at all time points. The proteins encoded by those genes represent glucose metabolic processes. In the fourth group, we found an elongation factor, Tu, and ribosomal protein S7 (Fig. 3C and D).

The roles of differentially expressed genes in cellular metabolic pathways were analyzed using the KEGG database. Seven pathways were found to be differentially expressed under conditions of plant stress. Genes corresponding to four pathways, namely, those corresponding to oxidative phosphorylation, sulfur metabolism, pentose and glucuronate interconversions, and aminosugar and nucleotide sugar metabolism, were upregulated. The genes corresponding to the KEGG pathway for glutathione metabolism, a two-component system, and the pentose phosphate pathway were downregulated.

Furthermore, KEGG analysis revealed that the oxidative phosphorylation pathway was the only metabolic pathway that was activated under conditions of drought stress at all three different time points (see Fig. S5 in the supplemental material). This pathway was activated at 1 h after drought stress induction with upregulation of ATP synthase subunit delta. After the organism had been maintained under conditions of drought stress for 6 h, the number of expressed genes reached 16 and included genes for cytochrome o ubiquinol oxidase subunit III, succinate dehydrogenase hydrophobic membrane anchor protein (SdhC), NADH-quinone oxidoreductase, and several ATP synthase subunits and NADH dehydrogenase subunits. This pathway was still active after 12 h, and the genes expressed included those encoding protoheme IX farnesyltransferase, succinate dehydrogenase, cytochrome b556 subunit (SdhD), NADH-quinone oxidoreductase, and several ATP synthase subunits and NADH dehydrogenase subunits.

Expression of ECF sigma factor genes in *B. phytofirmans* PsJN colonizing potato plants. Expression of extracytoplasmatic function (ECF) sigma factor genes

in *B. phytofirmans* PsJN colonizing potato plants (cv. Bionat and Russet Burbank) with and without drought stress was tested by real-time qPCR with cDNA of four biological replicates per treatment. Transcripts of six ECF sigma factor genes (ECF_164, ECF_886, ECF_429, ECF_718, and ECF_474) were detected in PsJN-colonized plants (see Fig. S6 in the supplemental material). The number of expressed ECF sigma factor genes in stressed plants was different from the number in control plants, and the numbers varied over time in stressed plants. ECF_164, ECF_429, and ECF_886 were expressed in control plants of both cultivars and were also active in stressed plants at all time points. The number of expressed ECF sigma factor genes increased under conditions of drought stress to a maximum of five genes after 1 h in Bionta and after 6 h in Russet Burbank (see Fig. S6). Transcripts of ECF_718 were detected in both cultivars, whereas ECF_474 was found in Russet Burbank and ECF_230 in Bionta only.

Differential expression of ECF sigma factor genes in *B. phytofirmans* PsJN colonizing potato plants. The expression levels of ECF sigma factor genes were normalized to those of the most stably expressed control gene, Bphyt_2615 (glutamine synthetase). Relative expression levels of ECF sigma factor genes under conditions of drought stress are shown in Fig. 4. ECF_886 was significantly

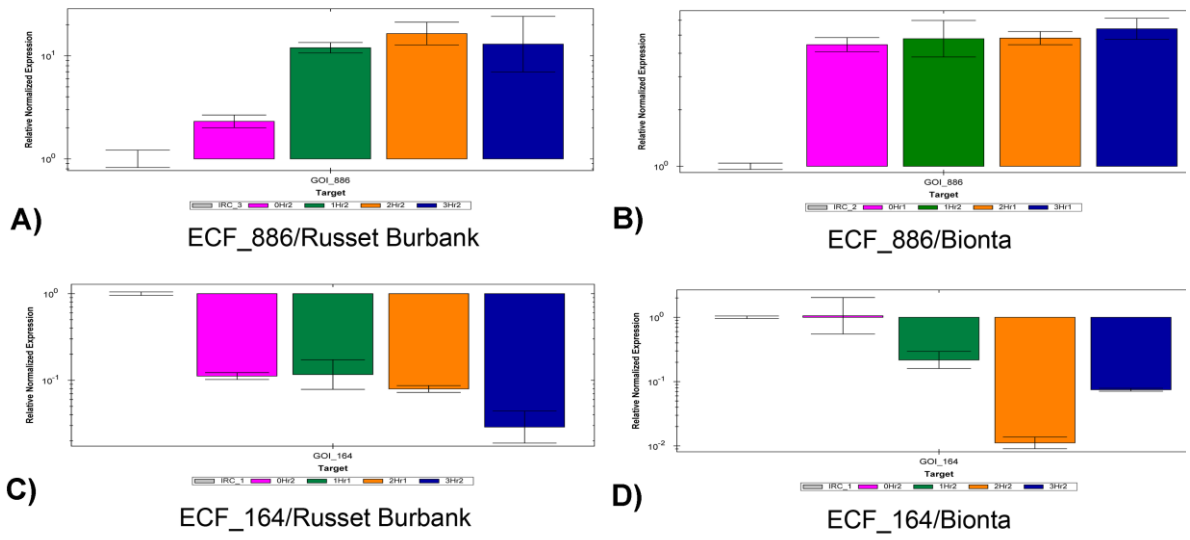


FIG 4 Relative expression levels of ECF sigma factors with modulated expression in *B. phytofirmans* PsJN in response to plant drought stress. A) ECF_886 (Bphyt_1327) in Russet Burbank; B) ECF_886 (Bphyt_1327) in Bionta; C) ECF_164 (Bphyt_4063) in Russet Burbank and D) ECF_164 (Bphyt_4063) in Bionta

upregulated in *B. phytofirmans* PsJN colonizing cv. Russet Burbank potato plants under conditions of drought stress. The expression of ECF_886 reached a maximum (3×) after 6 h of drought stress. In cultivar Bionta, ECF_886 expression did not change significantly after stress induction compared to control results. Expression of ECF_164 was relatively constant in Russet Burbank but was clearly downregulated in Bionta under conditions of plant drought stress. The transcript levels of the other expressed ECF sigma factor genes (ECF_429, ECF_718, ECF_230, and ECF_474) remained relatively constant throughout the different time points of drought stress in both cultivars.

DISCUSSION

The genome of *B. phytofirmans* PsJN encodes 7,405 genes, 4,591 (62%) of which were expressed in PsJN colonizing cv. Bionta potato plants *in vitro*. The active genes were evenly distributed across both chromosomes and the plasmid. Interestingly, about 46% of the coding sequences (CDS) located on the plasmid were expressed by PsJN inside potato plants. The genome of *B. phytofirmans* PsJN shows a high degree of similarity to that of *B. xenovorans* LB400, but the pBPHYT01 plasmid is different from the megaplasmid of *B. xenovorans*. In *B. xenovorans* LB400, the small chromosome is the “lifestyle-determining” replicon, whereas strain-specific functions are encoded on the megaplasmid (14).

In *B. phytofirmans* PsJN, only 29% of the CDS on pBPHYT01 could be functionally described (15); consequently, the majority of plasmid-carried genes that are active in PsJN colonizing cv. Bionta potato plants are of unknown function. Several reports of studies have proposed that plasmids are genetic determinants of functional diversification and niche adaptation (16, 17). We can only speculate on the role of pBPHYT01 in the endophytic lifestyle of *B. phytofirmans* PsJN. Experiments designed to test plasmid-free PsJN for the ability to colonize plants and to establish endophytic population could give further insights.

The transcriptome profile of *B. phytofirmans* PsJN colonizing potato plants indicates that the bacterium is metabolically active in plants. The majority of expressed traits were related to transcription regulation, general metabolism (sugars, amino acids, lipids, and nucleotides), energy production, and cellular homeostasis. The overall patterns of functions encoded on the genome of *B. phytofirmans* PsJN and

expressed in *in vitro* potato plants were highly similar. We conclude from this that the plant interior as a habitat for bacteria does not require very selected and specialized functionalities. Only a few differences in the COG patterns were found, with cell motility and defense mechanisms being less represented in the transcriptome than in the genome of *B. phytofirmans* PsJN. Cell motility might be important in plant invasion and in the spreading of endophytic microorganisms throughout plant organs and tissues (18). Recently, Balsanelli and colleagues showed that motility-related functions such as chemotaxis and type VI pilus functions play an important role in the initial contact with plants and the epiphytic colonization of maize roots by Herbaspirillum seropedicae (19). Our data indicate that active movement is less important once a bacterial population is successfully established inside plants. Interestingly, defense-related traits seem to play a minor role in the endophytic life of *B. phytofirmans* PsJN also. A possible explanation is that the plant endosphere is a protected niche allowing endophytes to escape the high competitive pressure in the rhizosphere and soil.

B. phytofirmans PsJN stimulates plant growth in many of its host plants. Metabolic properties suggested to be involved in this activity include the production and degradation of auxin phytohormone indole-3-acetic acid (IAA) (20), ACC deaminase activity (21), quinolinate phosphoribosyl transferase (QPRTase) or nicotinate-nucleotide pyrophosphorylase activity (22), and siderophore production (1). In our experiment, *in vitro* potato plants colonized by *B. phytofirmans* PsJN showed increased growth compared to an untreated control. By showing the expression of quinolinate phosphoribosyl transferase and indole-3-acetic acid (IAA) synthesis genes of strain PsJN colonizing potato plantlets, the results of the present study support previous reports of the importance of these functions for the beneficial interaction between PsJN and plants.

Plants colonized by *B. phytofirmans* PsJN often show increased tolerance of abiotic stress such as chilling (23) and drought (4). One of the main ambitions of this study was to elucidate whether and how endophytic *B. phytofirmans* PsJN is affected by host plant drought stress. Analysis of differentially expressed genes in *B. phytofirmans* PsJN colonizing nonstressed and drought-stressed potato plants showed that 137, 229, and 185 genes were upregulated and 57, 125, and 81 genes were downregulated in response to host plant drought stress at 1, 6, and 12 h after

PEG application. This indicates that *B. phytofirmans* PsJN adjusts gene expression to physiological conditions that have been altered in host plants due to plant stress responses. Genes that were significantly upregulated in *B. phytofirmans* PsJN in response to host plant stress are mostly involved in transcription regulation, cellular homeostasis, and cell redox homeostasis, indicating a stress response in *B. phytofirmans* PsJN. Drought stress provokes an oxidative burst in plants as a primary immune response. This increase in the production of reactive oxygen species (ROS) serves on the one hand as an alarm signal that triggers acclimation and defense reactions and is kept in tight balance by the antioxidant system in plants (24). If, on the other hand, the drought stress continues for a certain period of time, the oxidative burst may lead to extensive cellular damage and finally to cell death (24). Upregulation of genes involved in detoxification of ROS in strain PsJN colonizing potato plants suffering from drought stress led us to the assumption that endophytic *B. phytofirmans* PsJN senses and is affected by plant-produced ROS. We propose the following scenario. Water limitation leads to ROS burst in plants. Endophytic PsJN cells respond with the expression of genes involved in the defense against oxidative stress in order to prevent cell damage. ROS scavenging by endophytes is also very important during the early stage of plant colonization, as previously shown for Gluconacetobacter diazotrophicus PAL5 (11). Whether bacterial antioxidant activity may help to maintain the redox balance in plants and thus reduce the effects of drought stress on plants remains elusive and merits further investigation.

Oxidative phosphorylation was found to be activated in *B. phytofirmans* PsJN over time during plant drought stress. In the process of cellular respiration, aerobic bacteria pass electrons from oxidizable substances to molecular oxygen via the so-called electron transport chain. The released energy is used to produce energy-rich ATP from ADP by phosphorylation. Oxidative phosphorylation generates the energy needed for almost all vital processes (25). Apart from this, pentose and glucuronate interconversions and amino sugar and nucleotide sugar metabolism were also activated. Upregulation of genes involved in these processes in *B. phytofirmans* PsJN colonizing drought-stressed potato plants could indicate activation of bacterial growth. At least for Epichloe endophytes, it is well documented that the mutualistic interaction of fungi and host plant is tightly regulated. Perturbations of this balance result in a switch from restricted to proliferative growth of the endophyte inside the plant and a breakdown from mutualistic to pathogenic behavior (26). Furthermore,

anarchic proliferation of otherwise asymptomatic bacterial endophytes is a common phenomenon in *in vitro* plant propagation when cultures are under stress (27). We therefore quantified PsJN in the control and stressed potato plants by performing qPCR with the selected housekeeping genes used for data normalization but did not find a significant increase in copy numbers over time under conditions of drought stress (data not shown). It is possible that the time span (12 h) was too short to observe a significant increase in cell numbers, but it is also likely that the increase in metabolic activity in *B. phytofirmans* PsJN under conditions of host plant drought stress was not coupled with proliferated growth.

One way that bacteria sense and react to the extracellular environment is by the so-called cell surface signaling-employing extracytoplasmatic function (ECF) sigma factors (28). This signal transduction system consists of an outer membrane receptor, an inner membrane-bound sigma factor regulator (anti-sigma factor), and, bound to that, an ECF sigma factor. In the absence of a signal, the anti-sigma factor tightly binds the ECF sigma factor, thereby keeping it in its inactive state. The anti-sigma factor is proteolytically degraded in the presence of a stimulus. As a result, the sigma factor is released and activates expression of its target genes (28). The ECF subfamily is the largest group in the sigma 70 family, and its members are involved in a wide range of environmental responses, such as metal homeostasis, starvation, and resistance to antimicrobial peptides, and are also required for pathogenesis in some cases (28). ECF sigma factors may also play a role in the establishment of plant-microbe interactions. Gourion and colleagues (29) showed that an extracytoplasmatic sigma factor is involved in symbiotic efficiency in the plant symbiont Bradyrhizobium japonicum USDA110. The genome of *B. phytofirmans* PsJN harbors eighteen different CDS putatively coding for extracytoplasmatic function (ECF) sigma factors. Analysis of the *in planta* transcriptome of *B. phytofirmans* PsJN revealed the expression of eleven extracytoplasmatic sigma factor genes, and the expression of one of these genes was upregulated upon plant stress induction. TaqMan-quantitative PCR experiments were performed to quantitatively assess ECF sigma factor activation in *B. phytofirmans* PsJN in response to plant stress. Six ECF sigma factor genes were expressed in *B. phytofirmans* PsJN colonizing drought-stressed potato plants. One of these genes (Bphyt_1327; ECF_886) was significantly upregulated in response to plant stress. This gene has orthologs in 36 other sequenced *Burkholderia* strains (Burkholderia

Genome Database [30]) with similar genetic neighborhoods. Little is known of the function of this ECF sigma factor in *Burkholderiaceae*; only the ortholog in *Burkholderia cenocepacia*, EcfD, was found to be involved in the response to chlorhexidine (31). The biological role of Bphyt_1327 in *B. phytofirmans* PsJN remains unclear and requires further investigation.

Another ECF sigma factor (Bphyt_4063; ECF_164) was downregulated in *B. phytofirmans* PsJN. It is orthologous to EcfI in *Burkholderia cenocepacia* and can be found in 31 other *Burkholderia* genomes in similar genetic neighborhoods. In *B. cenocepacia*, EcfI is involved in the synthesis of ornibactin siderophores and, thus, iron transport (32). Downregulation of this ECF sigma factor in *B. phytofirmans* PsJN colonizing drought-stressed potato plants could be directly linked to oxidative stress response. In bacteria, the regulation of iron homeostasis is coordinated with defense against oxidative stress (33). In many bacteria, Fur-like transcription regulators act as kind of master regulator in this regulatory network. Fur-like proteins control iron supply in dependence on the redox status of cells, either directly by repressing iron acquisition genes or indirectly by repression of other regulators such as ECF sigma factors (32, 33). Consequently, it seems likely that Bphyt_4063 in *B. phytofirmans* PsJN is involved in the regulation of iron uptake.

Interestingly, we found differences in the intensities of ECF sigma factor activation in PsJN in the two potato cultivars (Bionta and Russet Burbank). In agreement, it is well known that the plant growth-promoting effect of PsJN in potato is cultivar dependent (34, 35). In previous studies, the intensity of the effects correlated with the PsJN titer in the plants, which was found to be much higher in cultivars showing a greater increase in growth (35). We therefore quantified PsJN in the potato plants used in this study by performing qPCR with the selected housekeeping genes used for data normalization but did not find significant differences in copy numbers in cv. Bionta and Russet Burbank (data not shown). Moreover, the differences in response may have been due to differences in DNA methylation in plants, which was found to be enhanced in poorly responsive potato cultivars (34). Together with the previous observations, our findings indicate that the plant genotype-dependent plant growth-promoting effect of *B. phytofirmans* PsJN is accompanied by differences in the responsiveness of the strain to plant physiology.

Summarizing our data, one of the main outcomes of this study is that endophyte-plant interactions are not a one-way relationship in which the plant responds to the endophyte but represent a complex interplay in which each partner is affected by the other. This may hold true, and may become even more complicated under natural conditions, when plants are colonized by a rich microbial community consisting of bacteria, fungi, and viruses. The term “hologenome” has been introduced to describe the sum of the genetic information corresponding to an organism and its microbiota, which function in consortium (36). Our findings very much support this theory, and it is obvious that we need to develop a better understanding of the plant phenotype as an outcome of the interplay between inoculants and the host plants and their endogenous microbiota to be able to fully explain beneficial plant-microbe interactions.

Our data provide novel insights into the response of the plant growth-promoting endophyte *B. phytofirmans* PsJN to the plant host but also raise many issues such as those concerning the role of the plasmid in the endophytic lifestyle of strain PsJN and whether endophytes are involved in maintaining redox and energy balance in plants.

MATERIAL and METHODS

Bacterial strains and growth conditions. In this study, the plant growth-promoting rhizosphere bacterium and endophyte *Burkholderia phytofirmans* PsJN (= LMG 22146 T = CCUG 49060 T) was used (1). The bacterial strain was grown by loop inoculation of a single colony in LB broth. The bacterial culture was incubated at 28 ± 2°C for 2 days at 180 rpm in a shaking incubator.

Plant experiment. Two potato varieties (Solanum tuberosum cv. Russet Burbank and Bionta) were grown *in vitro* in a growth chamber at 20°C with a 16-h-light/8-h-dark-photoperiod cycle. Four-week-old elongated apical shoots (10 cm in length) were used directly as explants. They were inoculated with *B. phytofirmans* PsJN by dipping for some seconds in a bacterial culture (1.3 × 108 CFU/ml). Inoculated plants were grown in 10 ml solid Murashige-Skoog (MS) medium containing 8% Duchefa Daishin agar (pH 5.8) and 20% saccharose for 4 weeks in glass tubes (2.5-cm diameter). Rooted potato plantlets were transferred into custom-tailored glass tubes with a narrow neck allowing hydroponic plant culturing. The plantlets were grown in 15 ml liquid MS medium containing 20% saccharose for two more weeks. Drought stress was induced by adding polyethylene glycol (PEG) (molecular weight [MW],

6,000) to reach a final concentration of 45%. Control plants were not treated. Shoots of stressed and control plantlets of each variety (6 replicates per treatment) were harvested at 1, 6, and 12 h after PEG application, immediately frozen in liquid nitrogen, and stored at −80°C for further analysis. To evaluate the presence of *B. phytofirmans* PsJN in inoculated potato plantlets, 16S rRNA gene PCR was performed using universal primers 799F [5′-AAC(AC)GGATTAGATACCC(GT)-3′] (37) and 1520R (5′-AAGGAGGTGATCCAGCCGCA-3′) (38). Amplification with this primer pair allows exclusion of chloroplast 16S rRNA gene-based amplicons but results in coamplification of plant mitochondrial small-subunit rRNA gene fragments and the bacterial 16S rRNA gene (see Fig. S1 in the supplemental material). Each PCR (50 µl) contained 30 to 50 ng/µl bacterial or potato DNA as the template, 1× PCR buffer, 2.5 mM MgCl2, 200 µM deoxynucleoside triphosphate (dNTP) mix (Thermo Scientific), a 150 nM concentration of each forward and reverse primers, 2.5 U of Firepol DNA polymerase (Solis Biodyne, Estonia), and PCR-grade water. Cycling conditions were as follows: initial denaturation for 5 min at 95°C; 34 cycles of 30 s at 95°C, 1 min at 60°C, 2 min at 72°C; and final elongation for 4 min at 72°C. Amplified PCR products (5 µl) were separated by electrophoresis (80 V) on 1% (wt/vol) agarose gels. Agarose gels were stained with ethidium bromide. Bacterial PCR amplicons were sequenced, making use of the sequencing service at LGC Genomics (Germany). Retrieved sequences were visualized and aligned with ClustalW as implemented in BioEdit v7.1.3 (39). For identification, sequences were subjected to BLAST analysis with the NCBI database.

Total RNA isolation and cDNA synthesis from plant tissue and bacterial cells. Frozen plant material (100 mg) was prechilled with liquid nitrogen in 2-ml Safe-Lock tubes (Greiner Bio-One, Germany) and homogenized by the use of a ball mill MM301 mixer (Retsch GmbH & Co., Germany) at 30 Hz for 2 min using a single steel ball (5-mm diameter). Afterward, the material was immediately subjected to RNA isolation as described by Chang et al. (40). Extraction of total RNA from pure bacterial cultures was done using TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions, and the RNA was used as a control in qPCR experiments. RNA samples were treated with DNase (Turbo DNA-free kit; Ambion, USA) according to the manufacturer's protocol for purification from DNA contamination. RNA samples were tested for contaminating DNA by 16S rRNA gene PCR. RNA was analyzed at a 260-nm/280-nm ratio using a NanoDrop 1000

spectrophotometer (Thermo Scientific, USA), and the integrity was checked by electrophoresis in a 1% agarose gel. Purified RNA samples were reverse transcribed to cDNA with an iScript cDNA synthesis kit (BioRad Inc., USA) using random hexamers according to the manufacturer's instructions.

Transcriptome sequencing. In order to get a comprehensive image of the total transcriptome of *B. phytofirmans* PsJN colonizing potato plants (cv. Bionta), the total RNAs of four biological replicates for each treatment and control samples were pooled at equal concentrations to obtain approximately 48 µg RNA per treatment. rRNA depletion, cDNA library preparation, and sequencing on an Illumina HiSeq 2000 system with a 50-bp single-end read length were outsourced to Vertis Biotechnologie AG (Germany). In brief, plant rRNA molecules were depleted from the total RNA using a Ribo-Zero rRNA removal kit (Epicentre, USA). Plant mRNA molecules were removed by the use of oligo(dT) magnetic beads. The bacterial RNA samples were poly(A) tailed using poly(A) polymerase, and the RNA species which carried a 5′ monophosphate were degraded with Terminator exonuclease (Epicentre, USA). First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and Moloney murine leukemia virus (MMLV) reverse transcriptase. The resulting cDNA was PCR amplified to about 20 to 30 ng/µl using high-fidelity DNA polymerase. The cDNA was then purified using an Agencourt AMPure XP kit (Beckman Coulter Genomics, USA) and analyzed by capillary electrophoresis. The raw data are available in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3524.

Pre-processing of sequencing data and mapping of reads. RNA-seq reads were subjected to quality filtering using Prinseq (41) on the basis of the criterion of a minimum read length of 40, and good-quality reads were obtained with Q20 (sequencing error rate lower than 1%) for all reads. The poly(A) tails with a minimum length of 10 bp were removed by the use of Prinseq . Reads longer than 30 bp were considered for further analysis. Before read alignment, rRNA fragments were filtered from the transcriptomics data by the use of SortMeRNA software and the default rRNA database included in the software package (42). Reads were aligned to the genome of *B. phytofirmans* PsJN, and gene annotations were obtained from the NCBI database of bacterial genomes

(ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/Burkholderia_phytofirmans_PsJN_uid58729/).

Transcriptome analysis. RNA-seq reads were aligned to the genome of *B. phytofirmans* PsJN using the Burrows-Wheeler Alignment Tool (BWA v0.6.2) (43). Transcript abundance was calculated using in-house Python scripts. Differential gene expression levels were analyzed and visualized with NOIseq (44). Gene expression levels were normalized using the number of reads per kilobase of coding sequence per million mapped reads (RPKM) in *B. phytofirmans* PsJN. As there was no replicate available for our data set, NOISeq-sim was used with the highest threshold (q = 0.9) to compute the probability of differential expression of genes under different conditions. To determine the variations of the differentially expressed genes which were expressed across all the time points from control compared to drought-stressed samples in *B. phytofirmans* PsJN, hierarchical cluster analysis was performed using Cluster 3.0 (13). The clustering results were visualized using TreeView (http://jtreeview.sourceforge.net/). Functional annotation of target genes based on Gene Ontology terms was performed using David v6.7 software (45) and all PsJN genome data available at NCBI RefSeq database (46) as the background. Venn diagrams were drawn using BioVenn software (47). Functional COG (clusters of orthologous groups) categories of the transcripts expressed in the control were listed using the COG database (48) and compared to the COG categories of the *B. phytofirmans* PsJN genome obtained from Integrated Microbial Genomes (IMG) systems (http://img.jgi.doe.gov).

Design of primers and probes for amplification of ECF sigma factor genes. Oligonucleotides and probes for amplification of extracytoplasmatic function (ECF) sigma factor genes (see Table S4 in the supplemental material) were designed on the basis of the genome sequence of *B. phytofirmans* PsJN (GenBank project accession no. CP001052, CP001053, and CP001054) by making use of the ARB software package with its subfunction “Probe design,” version ARBuntu 2.0 (49). Probes were labeled with the reporter dye 6-carboxyfluorescein (FAM) at the 5′ end and with black hole quencher 1 (BHQ-1) fluorophore at the 3′ end.

Test for specificity of primers and probes for ECF sigma factor genes. Primer specificity was checked by PCR amplification using genomic DNA of *B. phytofirmans* PsJN. DNA was extracted from bacterial cell pellets using a FastDNA Spin kit for soil

(MP Biomedicals, LLC). DNA concentrations were measured with a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE). PCR amplification was performed in a T-gradient PCR thermocycler (Biometra, Germany). PCR and PCR-amplicon sequencing as well as sequence analysis and identification were done as described above. Three primer sets (Bphyt_5131, Bphyt_7017, and Bphyt_1666) were not specific for their target ECF sigma factor genes and were removed from further analysis.

Selection of housekeeping genes for normalization of expression data. Primers and probes were designed for 12 (see Table S4 in the supplemental material) gene candidates using the ARB software package (49), and probes were labeled with the 6-FAM reporter dye at the 5′ end and with black hole quencher 1 (BHQ-1) at the 3′ end. The primers were designed to be specific for *B. phytofirmans* PsJN and to avoid amplification of host plant material. The specificity of primers was tested by PCR, using genomic DNA of *B. phytofirmans* PsJN and potato plants as the template in separate reactions. The analysis of PCR amplification, sequencing, and sequences was performed as described above. Expression of all housekeeping genes was checked by TaqMan-qPCR using cDNA synthesized from total RNA isolated from stressed and control plants. The expression stability of selected housekeeping genes under conditions of drought stress and the variation in quantitative PCR efficiency in *B. phytofirmans* PsJN were calculated using qBasePLUS software with the geNormPLUS algorithm implemented (50, 51). Glutamine synthetase (Bphyt_2615) was the candidate housekeeping gene product that showed the most stable expression at different time points of drought stress and in both potato varieties (see Fig. S4).

TaqMan Real-Time PCR assays. Quantitative PCR was carried out using the TaqMan-qPCR assay and a Bio-Rad CFX-96 real-time detection system (Bio-Rad, Hercules, CA). Triplicate qPCR reactions were performed with 1 µl of cDNA as the template, 1× BioRad SsoFast probe mix (BioRad Inc., Hercules, CA), 10 µM of forward and reverse primers, and 5 µM probe in a final volume of 10 µl. Cycling conditions were as follows: a hot start at 95°C for 2 min, 69 cycles of denaturation at 95°C for 5 s, and 20 s of elongation at 60°C. In each run, 2 negative controls were used, one as a no-template control performed with PCR-grade water instead of cDNA and another that included cDNA from non-PsJN-inoculated control plant

samples. Four biological replicates of both potato varieties were tested. Data were analyzed with Bio-Rad CFX Manager software (version 3.0). Based on the Pfaffl equation (52) implemented in this software, normalized relative quantity (NRQ) values of ECF sigma factor genes in comparison to the most stably expressed reference gene (Bphyt_2615) were determined.

ACKNOWLEDGEMENTS

This work was supported by a grant provided by the Austrian Science Fund (FWF): P22867-B16.

Supplementary Material

The supplementary tables for this article can be found online at: http://mbio.asm.org/content/6/5/e00621-15.full#sec-24

Table S1 Examples of biological functions expressed in *B. phytofirmans* PsJN colonizing *in vitro* potato plants (cv. Bionta).

Table S2 Complete list of genes, that are differentially expressed in *B. phytofirmans* PsJN after A) one hour, B) six hours and C) twelve hours under host plant drought stress.

Table S3 Examples of biological functions expressed in *B. phytofirmans* PsJN under drought stress at three different time points.

Table S4 Oligonucleotide primer and probe sequences (5' → 3') for A) ECF sigma factor genes and B) housekeeping genes of *B. phytofirmans* PsJN (FWD= Forward primer; RWD= Reverse primer; FAM= Reporter dye 6-carboxyfluorescein; BHQ-1= black hole quencher 1 fluorophore).

FIG S1. *In vitro* potato plants before stress application. Six weeks after inoculation with *B. phytofirmans* PsJN potato plants of both cultivars (Russet Burbank and Bionta) showed increased shoot and root length in response to colonization by strain PsJN. +, plants inoculated with *B. phytofirmans* PsJN, -, control plants treated with sterile growth media.

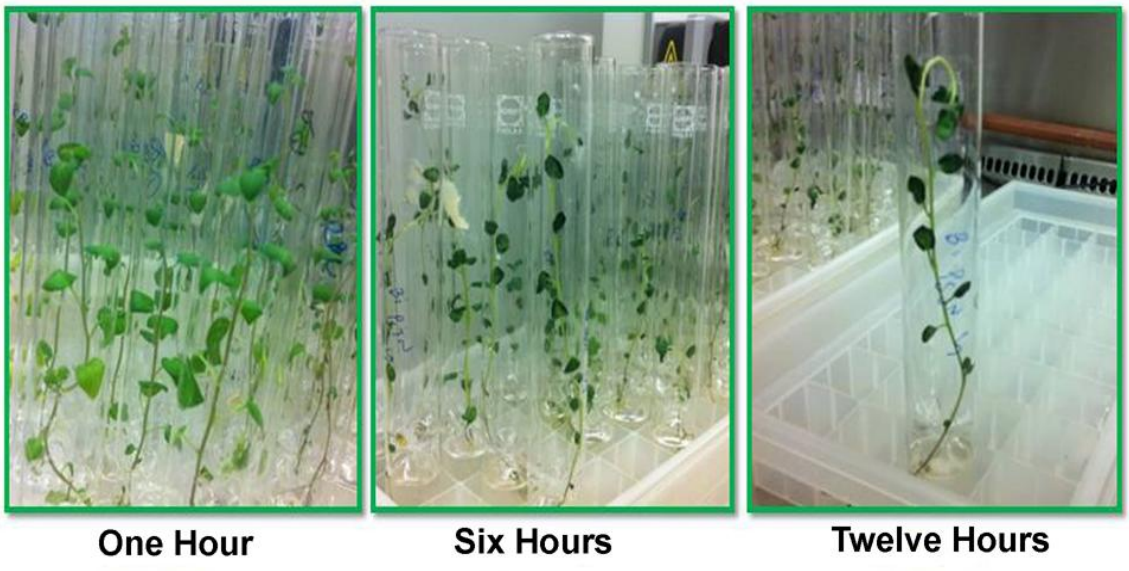


FIG S2. Application of polyethylene glycol (PEG) caused visible symptoms of water deficiency in *in vitro* potato plants at one, six and twelve hours after stress induction.

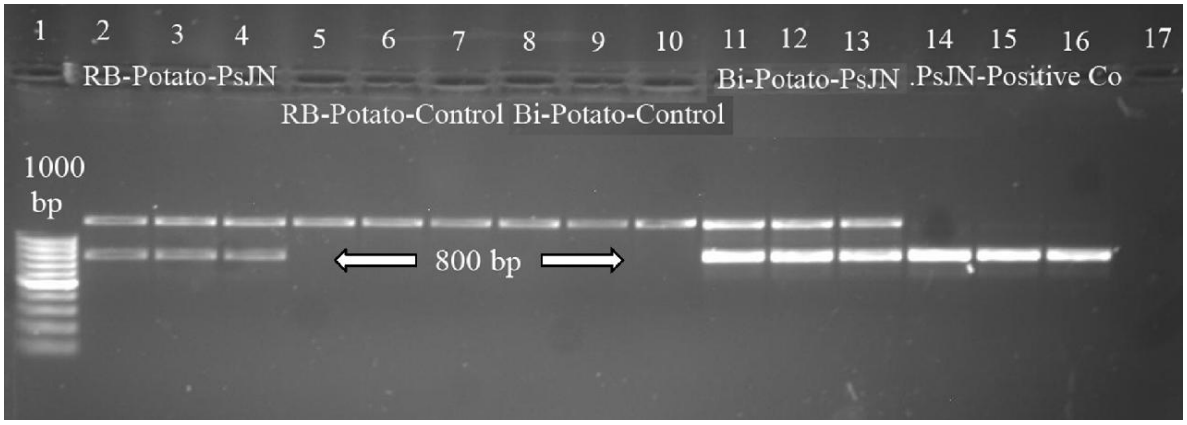


FIG S3. Detection of *B. phytofirmans* PsJN in inoculated potato plants. Lanes 1: molecular marker (GeneRuler 100 bp); Lanes 2,3, 4, 11, 12 and 13: PCR products from PsJN-inoculated plant samples (Russet Burbank (RB) and Bionta (Bi) varieties); Lanes 5 to 10: PCR products from non-inoculated Russet Burbank (RB) and Bionta (Bi) plant samples; Lanes 14 to 16: PCR products of genomic DNA of *B. phytofirmans* PsJN; Lanes 17: PCR-negative control.

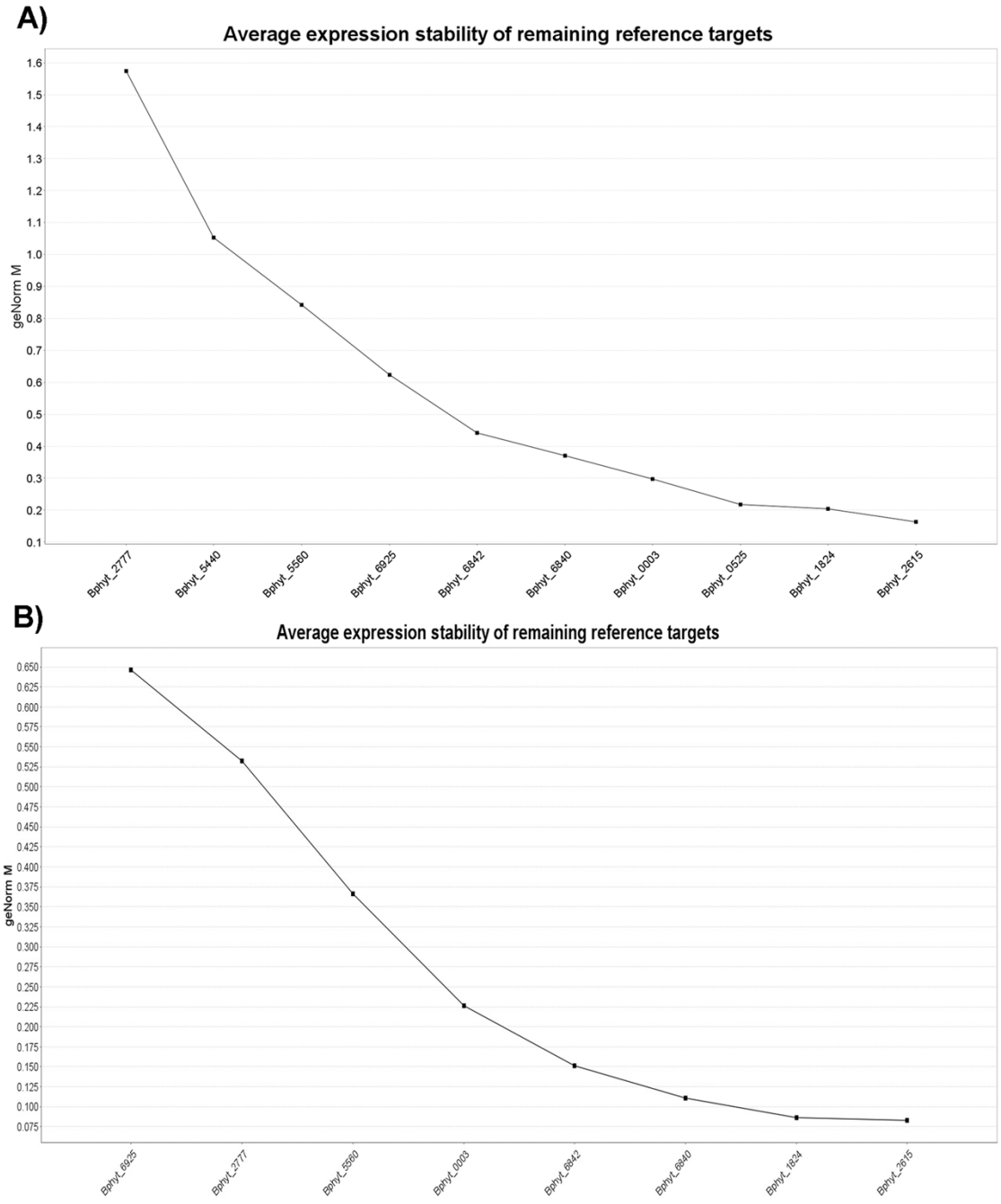


FIG S4. Evaluation of the gene expression stability of housekeeping genes (geNorm M value) from the most unstable genes (high M value in the Left side) to the most stable ones (Low M value in the right side) in A) Russet Burbank and B) Bionta. In both potato cultivars glutamine synthetase (Bphyt_2615) was the candidate gene with the most stable expression at the different time points tested with an average geNorm M<0.2 (lowest M value).

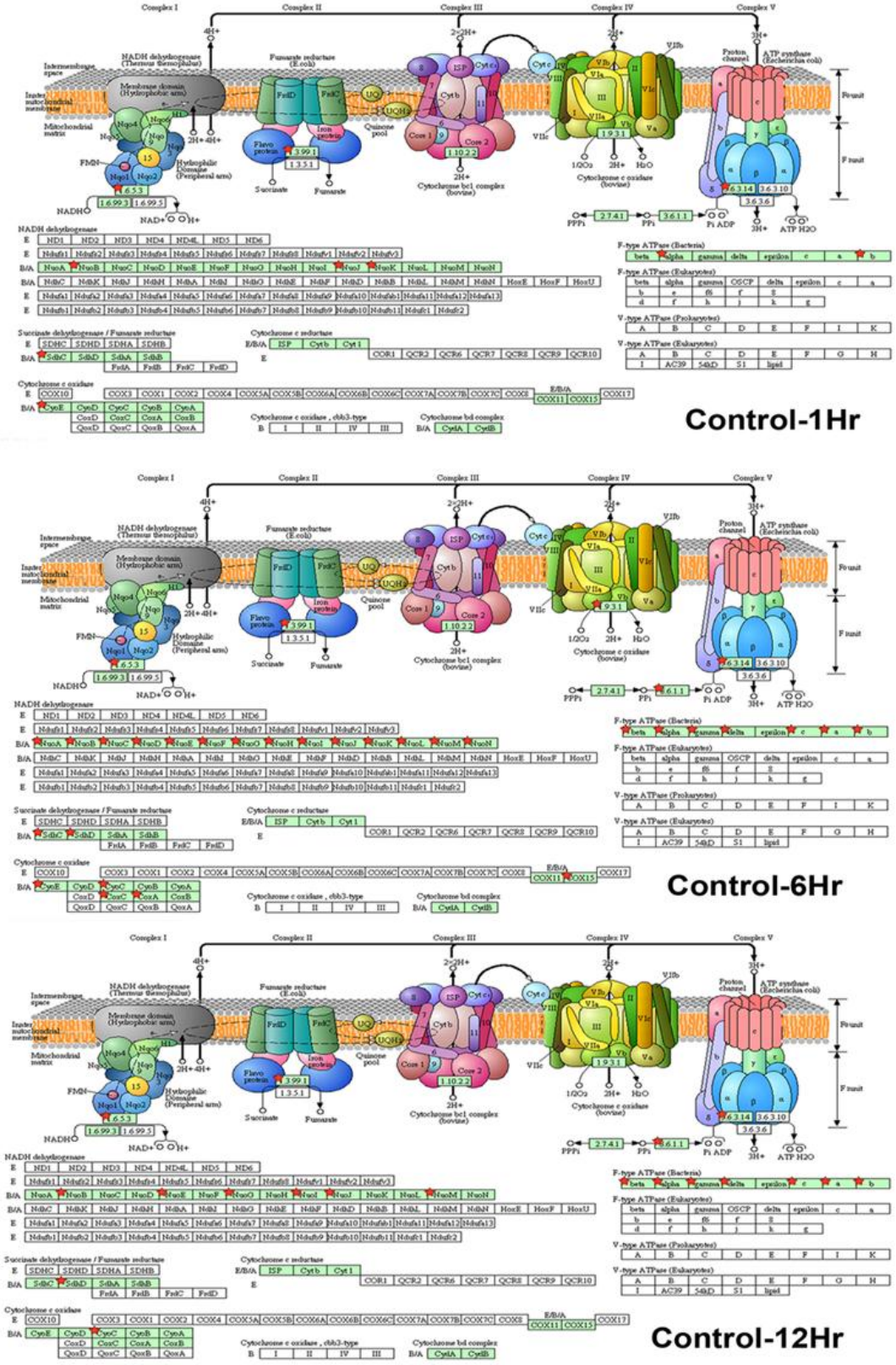


FIG S5. Activation of oxidative phosphorylation in *B. phytofirmans* PsJN colonizing drought stressed potato plants over time. Genes that are up-regulated at a certain time point are labeled with a red asterisk.

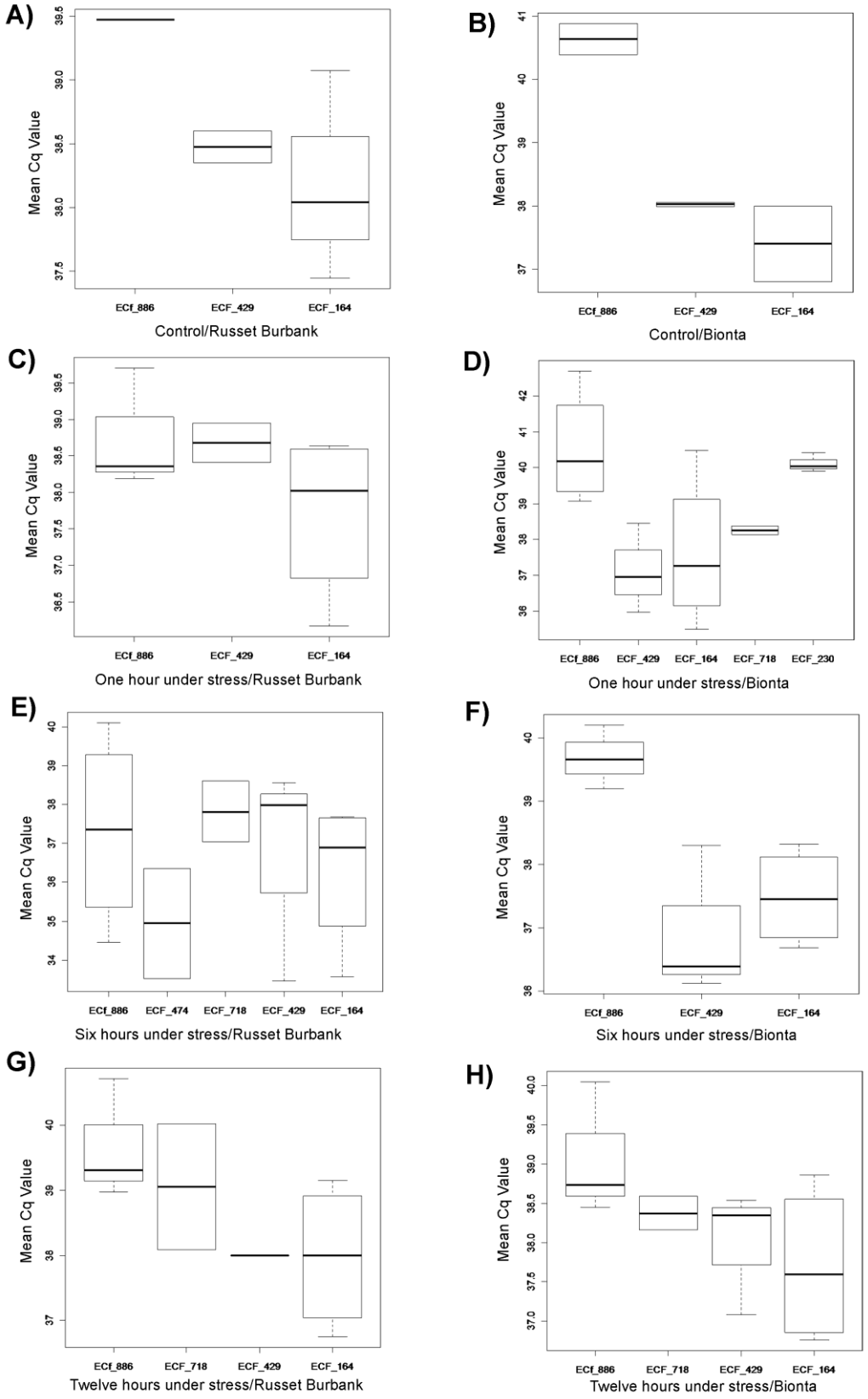


FIG S6. Range of mean Cq values for expressed ECF sigma factor genes in *B. phytofirmans* PsJN colonizing potato plants under drought stress at different time points in cultivar Russet Burbank (A, C, E and G) and in potato cultivar Bionta (B, D, F and H). Each box indicates the 25% and 75% percentiles of mean Cq values. Whiskers represent the maximum and minimum of Cq values. The median is depicted as a line across the box.

REFERENCES

1. **Sessitsch, A., T. Coenye, A. V. Sturz, P. Vandamme, E. Ait Barka, J. F. Salles, J. D. van Elsas D. Faure, B. Reiter, B. R. Glick, G. Wang-Pruski, and J. Nowak.** 2005. *Burkholderia phytofirmans* sp. nov., a novel plant-associated bacterium with plant beneficial properties. Int. J. Syst. Evol. Bacteriol. **55**:1187-1192.
   doi:10.1099/ijs.0.63149-0.
2. **Frommel, M. I., J. Nowak, and G. Lazarovits.** 1991. Growth enhancement and developmental modifications of *in vitro* grown potato (*Solanum tuberosum* ssp. *tuberosum*) as affected by a nonfluorescent *Pseudomonas* sp. Plant Physiol. **96**:928–936.
3. **Compant, S., J. Nowak, T. Coenye, C. Clément, and E. Ait Barka.** 2008. Diversity and occurrence of *Burkholderia* spp. in the natural environment. FEMS Microbiol. Rev. **32**:607-626.
   doi:10.1111/j.1574-6976.2008.00113.x.
4. **Naveed, M., B. Mitter, T. G. Reichenauer, W. Krzysztof, and A. Sessitsch.** 2014. Increased drought stress resilience of maize through endophytic colonization by *Burkholderia phytofirmans* PsJN and *Enterobacter* sp. FD17. Environmental and Experimental Botany **97**:30-39.
5. **Mitter B., A. Weilharter, P. S. G. Chain, F. Trognitz, J. Nowak, S. Compant, and A. Sessitsch.** 2013. Genome analysis, ecology and plant growth promotion of the endophyte Burkholderia phytofirmans strain PsJN. *In* F. de Bruijn (ed.) Molecular microbial ecology of the rhizosphere, Wiley-Blackwell publishing.
6. **Bordiec, S., S. Paquis, H. Lacroix, S. Dhondt, E. Ait Barka, S. Kauffmann, P. Jeandet, F. Mazeyrat-Gourbeyre, C. Clément, F. Baillieul, and S. Dorey.** 2011. Comparative analysis of defence responses induced by the endophytic plant growth-promoting rhizobacterium *Burkholderia phytofirmans* strain PsJN and the non-host bacterium *Pseudomonas syringae* pv. *pisi* in grapevine cell suspensions. J. Exp. Bot. **62**:595-602.
7. **Theocharis, A., S. Bordiec, O. Fernandez, S. Paquis, and S. Dhondt-Cordelier.** 2012. *Burkholderia phytofirmans* PsJN primes *Vitis vinifera* L. and confers a better tolerance to low nonfreezing temperatures. Mol. Plant-Microb. Interact. **25**:241-249.

doi:10.1094/MPMI-05-11-0124.

8. **Fernandez, O., A. Theocharis, S. Bordiec, R. Feil, L. Jasquens, C. Clement, F. Fontaine, and E. Ait Barka.** 2012. *Burkholderia phytofirmans* PsJN acclimates grapevine to cold by modulating carbohydrate metabolism. Mol. Plant-Microb. Interact. **25**:496-504.
http://dx.doi.org/10.1094/MPMI -09-11-0245.
9. **Poupin M. J., T. Timmermann, A. Vega, A. Zuñiga, and B. González.** 2013. Effects of the plant growth-promoting bacterium *Burkholderia phytofirmans* PsJN throughout the life cycle of *Arabidopsis thaliana*. PLoS ONE. **8**:e69435.
doi:10.1371/journal.pone.0069435.
10. **Shidore, T., T. Dinse, J. Öhrlein, A. Becker, and B. Reinhold-Hurek.** 2012. Transcriptomic analysis of response to exudates reveals genes required for rhizosphere competence of the endophyte *Azoarcus* sp. strain BH72. Environ. Microbiol. **14**:2775-2787.
doi:10.1111/j.1462-2920.2012.02777.x.
11. **Alquéres, S., C. Meneses, L. Rouws, M. Rothballer, I. Baldani, M. Schmid, and A. Hartmann.** 2013. The bacterial superoxide dismutase and glutathione reductase are crutial for endophytic colonization of rice roots by *Gluconacetobacter diazotrophicus* PAL5. Mol. Plant-Microb. Interact. **26**:937-945.
http://dx.doi.org/10.1094/MPMI-12-12-0286-R.
12. **Drogue, B., H. Sanguin, S. Borland, C. Prigent-Combaret, and F. Wisniewski-Dyé.** 2014. Genome wide profiling of *Azospirillum lipoferum* 4B gene expression during interaction with rice roots. FEMS Microbiol. Ecol. **87**:543-555.
doi:10.1111/1574-6941.12244.
13. **De Hoon, M. J. L., S. Imoto, J. Nolan, and S. Miyano.** 2004. Open source clustering software. Bioinformatics **20**:1453–1454.
doi:10.1093/bioinformatics/bth078.
14. **Chain, P. S., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agulló, V. L. Reyes, L. Hauser, M. Córdova, L. Gómez, M. González, M. Land, V. Lao, F. Larimer, J. J. LiPuma, E. Mahenthiralingam, S. A. Malfatti, Z. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul T. V. Tsoi, L. E. Ulrich, I. B. Zhulin, and A. M. Tiedje.** 2006.

Burkholderia xenovorans LB400 harbors a multi-replicon, 9.79-Mbp genome shaped for versatility. Proc. Natl. Acad. Sci. USA **103**:15280-15287.
doi: 10.1073/pnas.0606924103.

15. **Mitter, B., A. Petric, M. W. Shin, P. S. G. Chain, L. Hauberg-Lotte, B. Reinhold-Hurek, J. Nowak, and A. Sessitsch.** 2013. Bacterial genomes reveal a wide spectrum of endophytic lifestyles based on interaction strategies with host plants. Frontiers in Plant Sciences Vol. **4**:Article 120.
doi: 10.3389/fpls.2013.00120.
16. **De Maayer, P., W. Y. Chan, E. Rubagotti, S. N. Venter, I. K. Toth, P. R. Birch, and T. A. Coutinho.** 2014. Analysis of the Pantoea ananatis pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts. BMC Genomics **15**:404.
doi: 10.1186/1471-2164-15-404.
17. **Heuer, H., and K. Smalla.** 2012. Plasmids foster diversification and adaption of bacterial populations in soil. FEMS Microbial. Rev. **16**:1083-1104.
doi:10.1111/j.1574-6976.2012.00337.x.
18. **Compant, S., C. Clement, and A. Sessitsch.** 2010. Bacterial colonization in the rhizo- and endosphere of plants by plant growth-promoting bacteria. Soil Biol. Biochem. **42**: 669-678.
doi:10.1016/j.soilbio.2009.11.024.
19. **Balsanelli, E., M. Z. Tadra-Sfeir, H. Faoro, V. C. S. Pankievicz, V. A. de Baura, F. O. Pedrosa, E. M. de Souza, R. Dixon and R. A. Monteiro** (2015). Molecular adaptations of *Herbaspirillum seropedicae* during colonization of the maize rhizosphere. Environ. Microbiol.
doi:10.1111/1462-2920.12887.
20. **Zúñiga, A., M. J. Poupin, R. Donoso, T. Ledger, N. Guiliani, R. A. Gutiérrez, and B. González.** 2013. Quorum sensing and indole-3-acetic acid degradation play a role in colonization and plant growth promotion of *Arabidopsis thaliana* by *Burkholderia phytofirmans* PsJN. Mol. Plant-Microb. Interact. **26**:546–553.
http://dx.doi.org/10.1094/MPMI-10-12-0241-R.
21. **Sun, Y., Z. Cheng, and B. R. Glick.** 2009. The presence of a 1-aminocyclopropane-1-carboxylate (ACC) deaminase deletion mutation alters

the physiology of the endophytic plant growth-promoting bacterium *Burkholderia phytofirmans* PsJN. FEMS Microbiol. Letters **296**:131-136.
doi:10.1111/j.1574-6968.2009.01625.x.

22. **Wang, K., K. Conn, and G. Lazarovits.** 2006. Involvement of quinolinate phosphoribosyl transferase in promotion of potato growth by a *Burkholderia* strain. Appl. Environ. Microbiol. **72**:760–768.
doi:10.1128/AEM.72.1.760–768.2006.
23. **Ait Barka, E., J. Nowak, and C. Clément.** 2006. Enhancement of chilling resistance of inoculated grapevine plantlets with a plant growth promoting rhizobacterium, *Burkholderia phytofirmans* PsJN. Appl. Environ. Microbiol. **72**:7246-7252.
doi:10.1128/AEM.01047-06.
24. **Dat, J., S. Vandenabeele, E. Vranová, M. Van Montagu, D. Inzé, and F. Van Breusegem.** 2000. Dual action of the active oxygen species during plant stress responses. Cell Mol Life Sci. **57**:779-95.
25. **Stryer, Lubert.** 1988. Biochemistry. 3rd ed., W. H. Freeman Company, New York, NY.
26. **Eaton, C. J., M. P. Cox, and B. Scott.** 2011. What triggers grass endophytes to switch from mutualism to pathogenism? Plant Sci. **180**:190-195.
27. **Fki, L., O. Nasir, M. Chkir, M. Maalej, H. M. Shumacher, and N. Drira.** 2014. Abstr. Risk assessment of endophytes. WG meeting, COST Action FA 1103 “Endophytes in Biotechnology and Agriculture” abstr. Risk evaluation of endophytic bacteria´s anarchic proliferation within *in vitro* tissue cultures.
doi:http://dx.doi.org/10.1104/pp.96.3.928.
28. **Helman, J. D.** 2002. The extracytoplasmic function (ECF) sigma factors. Adv. Microb. Physiol. **46**:47-110.
29. **Gourion, B., S. Sulser, J. Frunzke, A. Francez-Charlot, P. Stiefel, G. Pessi, J. A. Vorholt, and H. M. Fischer.** 2009. The PhyR-sigma(EcfG) signalling cascade is involved in stress response and symbiotic efficiency in *Bradyrhizobium japonicum*. Mol. Microbiol. **73**:291-305.
doi: 10.1111/j.1365-2958.2009.06769.x.
30. **Winsor, G. L., B. Khaira, T. Van Rossum, R. Lo, M. D. Whiteside, and F. S. Brinkman.** 2008. The Burkholderia Genome Database: facilitating flexible queries and comparative analyses. Bioinformatics 24:2803-2804.

doi:10.1093/bioinformatics/btn524.

31. **Coenye, T., H. Van Acker, E. Peeters, A. Sass, S. Buroni, G. Riccardi, and E. Mahenthiralingam.** 2011. Molecular mechanisms of chlorhexidine tolerance in *Burkholderia cenocepacia* biofilms. Antimicrobial Agents and Chemotherapy. **55**:1912-1919.
doi:10.1128/AAC.01571-10.
32. **Agnoli, K., C. A. Lowe, K. L. Farmer, S. I. Husnain, and M. S. Thomas.** 2006. The ornibactin biosynthesis and transport genes of *Burkholderia cenocepacia* are regulated by an extracytoplasmic function sigma factor which is a part of the Fur regulon. J. Bacteriol. **188**:3631-3644.
doi:10.1128/JB.188.10.3631–3644.2006.
33. **Cornelis, P., Q. Wei, S. C. Andrews, and T. Vinckx.** 2011. Iron homeostasis and management of oxidative stress response in bacteria. Metallomics **3**:540–549.
doi:10.1039/c1mt00022e. doi:10.1093/jxb/erq291.
34. **Da, K., J. Nowak, and B. Flinn. 2012.** Potato cytosine methylation and gene expression changes induced by a beneficial bacterial endophyte, *Burkholderia phytofirmans* strain PsJN. Plant Physiol. Biochem. **50**:24-34.
35. **Nowak, J., R. E. Veilleux, J. Nowak, S. Turgeon.** 2007. Priming for transplant stress resistance in in vitro propagation via plantlet bacterization**.** Acta Hort. **748**:65e76.
36. **Zilber-Rosenberg, I., and E. Rosenberg.** 2008. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. FEMS Microbiol. Rev. **32**:723–735.
37. **Chelius, M. K., and E. W. Triplett.** 2001. The diversity of archaea and bacteria in association with the roots of *Zea mays* L. Microb. Ecol. **41**:252-263.
doi:10.1007/s002480000087.
38. **Edwards, U., T. Rogall, H. Blocker, M. Emde, and E. C. Bottger.** 1989. Isolation and direct complete nucleotide determination of entire genes - Characterisation of a gene coding for 16S-ribosomal RNA. Nucleic Acids Res. **17**:7843-7853. doi:10.1093/nar/17.19.7843.
39. **Hall, T. A.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. **41**:95-98.

40. **Chang, S., J. Puryear, and J. Cainey.** 1993. Simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. **11**:113-116.
41. **Schmieder R., and R. Edwards.** 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics **27**:863-864.
doi:10.1093/bioinformatics/btr026.
42. **Kopylova, E., L. Noé, H. Touzet.** 2012. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics, Oxford University Press (OUP): Policy B - Oxford Open Option B, **28**:pp.3211-3217.
43. **Li, H., and R. Durbin.** 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics **25**:1754–1760.
doi:10.1093/bioinformatics/btp324.
44. **Tarazona, S., F. Garcia-Alcalde, J. Dopazo, A. Ferrer, and A. Conesa.** 2011. Differential expression in RNA-seq: a matter of depth. Genome Res. **21**:2213-2223.
doi:10.1101/gr.124321.111.
45. **Huang, D. W., B. T. Sherman, and R. A. Lempicki.** 2009. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. **4**:44-57.
doi:10.1038/nprot.2008.211.
46. **Tatusova, T., S. Ciufo, S. Federhen, B. Fedorov, R. McVeigh, K. O´Neill, I. Tolstoy, and L. Zastavsky.** 2015. Update on RefSeq microbial genomes resources. Nucleic Acids Res. **43**: D599-605
doi: 10.1093/nar/gku1062.
47. **Hulsen, T., J. de Vlieg, and W. Alkema.** 2008. BioVenn - A web application for the comparison and visualization of biological lists using areaproportional Venn diagrams. BMC Genomics **9**:488.
doi:10.1186/1471-2164-9-488.
48. **Tatusov, R. L., D. A. Natale, I. V. Garkavtsev, T. A. Tatusova, U. T. Shankavaram, B. S. Rao, B. Kiryutin, M. Y. Galperin, N. D. Fedorova, and E. V. Koonin.** 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29**:22-28.
doi: 10.1093/nar/29.1.22.
49. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, A. Yadhukumar,**

A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Förster, I. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lüssmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K.-H. Schleifer. 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32**:1363–1371.

doi: 10.1093/nar/gkh293.

50. **Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman.** 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. **3**:research0034-research0034.11

doi:10.1186/gb-2002-3-7-research0034.

51. **Hellemans, J., G. Mortier, A. De Paepe, F. Speleman, and J. Vandesompele.** 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. **8**:R19.

doi:10.1186/gb-2007-8-2-r19.

52. **Pfaffl, M. W.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29**:e45.

doi: 10.1093/nar/29.9.e45.

53. **Overmars, L., R. Kerkhoven, and C. Franke.** 2013. MGcV: the microbial genomic context viewer for comparative genome analysis. BMC Genomics **14**:209.

doi: 10.1186/1471-2164-14-209.

Chapter IV

Comparative genome analysis of closely related *Pantoea ananatis* seed endophyte strains having different effects on the host plant

Frontiers in Microbiology 2015

6: 440

Comparative genome analysis of closely related *Pantoea ananatis* seed endophyte strains having different effects on the host plant

Raheleh Sheibani-Tezerji[1,2], Muhammad Naveed[1], Marc-André Jehl[2], Angela Sessitsch[1], Thomas Rattei[2*] and Birgit Mitter[1*]

[1]Health & Environment Department, Bioresources Unit, AIT Austrian Institute of Technology GmbH, Konrad-Lorenz-Strasse 24, 3430 Tulln, Austria
[2]University of Vienna, Department of Microbiology and Ecosystem Science, Division of Computational System Biology, Althanstrasse 14, 1090 Vienna, Austria

Number of Figures: 5
Number of Tables: 6

***Corresponding authors:**
Thomas Rattei
Department of Microbiology and Ecosystem Science
Division of Computational Systems Biology
University of Vienna
Althanstrasse 14
1090 Vienna, Austria
Email: thomas.rattei@univie.ac.at

Birgit Mitter
AIT Austrian Institute of Technology GmbH
Health & Environment Department, Unit of Bioresources
Konrad Lorenz Strasse 24
3430 Tulln, Austria
Email: birgit.mitter@ait.ac.at

Key words: seed endophyte, *Pantoea ananatis*, comparative genomics, plant growth promotion

Abstract

The seed as a habitat for microorganisms is as yet under-explored and has quite distinct characteristics as compared to other vegetative plant tissues. In this study, we investigated three closely related *P. ananatis* strains (named S6, S7, and S8), which were isolated from maize seeds of healthy plants. Plant inoculation experiments revealed that each of these strains exhibited a different phenotype ranging from weak pathogenic (S7), commensal (S8), to a beneficial, growth-promoting effect (S6) in maize. We performed a comparative genomics analysis in order to find genetic determinants responsible for the differences observed. Recent studies provided exciting insight into the genetic drivers of niche adaption and functional diversification of the genus *Pantoea*.

However, we report here for the first time on the analysis of *P. ananatis* strains colonizing the same ecological niche but showing distinct interaction strategies with the host plant. Our comparative analysis revealed that genomes of these three strains are highly similar. However, genomic differences in genes encoding protein secretion systems and putative effectors, and transposase/integrases/phage related genes could be observed.

Introduction

Bacterial endophytes have been defined as “bacteria, which for all or part of their life cycle invade the tissues of living plants and cause unapparent and asymptomatic infections entirely within plant tissues, but cause no symptoms of disease” (Wilson, 1995). Based on this definition, endophytes are clearly distinct from plant pathogens. However, bacteria can exist in plants in quiescence but proliferate and become detrimental to the host under certain conditions such as plant growth perturbations (Kloepper et al., 2013). Moreover, plant-pathogen interactions are often plant species specific and bacteria that are pathogenic to one plant species can exhibit an endophytic lifestyle in other plants (Bashan et al., 1982). On the other hand, it has been shown that plants respond differently to endophytes and plant pathogens (Bordiec et al., 2011). A promising approach in revealing differences in the host interaction strategies of pathogens and plant beneficial bacteria might be the comparison of functionalities and gene content of closely related bacterial strains that show different modes of interaction with host plants. Genome sequencing provides detailed information on the genes present in bacteria and offers a basis for

comparative genomics that aids in revealing differences in the host interaction strategies of pathogens and plant growth promoting bacteria.

The seed as a habitat for microorganisms is under-explored, although the first report of bacteria colonizing seeds dates back to the 1970s (Mundt and Hinkle, 1976). Only few studies have been performed on seed endophytes (Compant et al., 2011; Johnston-Monje and Raizada, 2011; Hardoim et al., 2012) and the origin of endophytes is under debate. A few studies suggest that at least some bacterial endophytes are vertically transmitted (Johnston-Monje and Raizada, 2011). The seed has quite distinct characteristics as compared to other vegetative plant tissues and one would expect that it also harbors distinct microbial communities. Based on cultivation-based analysis it has been reported that Gammaproteobacteria represent the most abundant class of maize seed endophytes, comprising mostly *Pantoea* and *Enterobacter* (Johnston-Monje and Raizada, 2011). Similarly, Rijavec et al. (2007) identified *Pantoea* as a major genus among endophytes isolated from maize seeds.

Pantoea ananatis is a bacterial species that was originally discovered in pineapple in the Philippines, in 1928 (Serrano, 1928). Members of this species have been shown to infect many mono- and dicotyledonous plant species, such as onion, rice, melon, sudan grass, tomato, and sorghum (Stall et al., 1969; Wells et al., 1987; Gitaitis and Gay, 1997; Azad et al., 2000; Cother et al., 2004; Cota et al., 2010). In maize *P. ananatis* is the causing agent of the foliar disease termed maize white spot disease (Paccola-Meirelles et al., 2001). *P. ananatis* strains display a wide range of ecological versatility, as they are commonly recovered from water, soil, insects, and plants (De Maayer et al., 2014). Depending on their host and ecological niches, *P. ananatis* strains can show different life styles such as mutualistic, saprophytic and pathogenic life styles (Coutinho and Venter, 2009). De Maayer et al. (2012a) showed that the Large Pantoea Plasmid (LLP-1) plays a crucial role in niche adaption and functional diversification of the genus *Pantoea*. By analyzing the pan-genome of eight sequenced *P. ananatis* strains De Maayer et al. (2014) identified a large number of proteins in this species with orthologs restricted to bacteria associated either with plants, animals or insects. The mechanisms of the diverse interactions between *P. ananatis* and the host are still poorly understood and only little is known on the genetic traits underlying plant pathogenic or beneficial activity. Shyntum et al. (2014) showed that type IV section system could play a role in pathogenicity and niche

adaptation. Genome analysis of the plant growth promoting strain *P. ananatis* B1-9 that has been isolated from the rhizosphere of green onions in Korea indicates that the strain lacks traits related to pathogenicity. Furthermore, it harbors genes that are putatively involved in plant growth stimulation and yield improvement (Kim et al., 2012).

In this work, we studied three endophytic *P. ananatis* strains (S6, S7, S8) isolated from maize seeds. Although they were isolated from seeds of healthy plants, they showed distinct characteristics in regard to plant growth and health. Strain S6 exhibited clear beneficial effects on maize growth, whereas S8 had hardly any effect and is considered as neutral and S7 caused disease symptoms known from *P. ananatis* infections. Therefore, this closely related group of strains represents a promising model to unravel genetic determinants in *P. ananatis* responsible for beneficial and pathogenic effects. Consequently, we functionally characterized the strains by testing for various known plant growth-promoting characteristics as well as for their effect on plant growth, and performed a comparative genome analysis to elucidate genetic features determining the type of plant-microbe interaction.

Material and Methods

Maize Varieties and Seed Surface Sterilization

Seeds of the maize cultivars (Helmi, Morignon, Pelicon, and Peso) were obtained from local farmers in Seibersdorf, Austria. Maize seeds with no cracks or other visible deformations were surface-sterilized with 70% ethanol for 3 min and 5% sodium hypochlorite for 5 min, and followed by repeated washing with sterile distilled water (3 times for 1 min). The efficacy of surface sterilization was checked by plating 3–5 seeds and aliquots of the final rinse onto 10% tryptic soy agar plates, and incubated for 3 days at 28 ± 1°C. The medium was checked daily for bacterial or fungal growth.

Isolation of Endophytic Bacteria from Maize Seeds

Seed-borne bacteria were isolated following the procedure described by Rijavec et al. (2007) with some modifications. For isolation, 50 surface-sterilized seeds of each cultivar were crushed and blended aseptically in 90 mL of half strength nutrient broth (Difco, Detroit, Michigan) for 5 min. The blend was then incubated at room temperature for 4 h on a rotary shaker (VWR International GmbH, Austria) at 100 r min^{-1}. Half strength nutrient broth containing 200 mg/L cycloheximide was inoculated with a series of the incubation mixture (10:1 mL ratio) and further incubated for 4

days on a rotary shaker at room temperature. Aliquots were taken from Erlenmeyer flasks with observed microbial growth and plated onto R2A (Difco, Detroit, Michigan). Plates were incubated at 28°C for 24–48 h. One hundred colonies were picked, and pure cultures were obtained by further streaking on agar plates. Single colonies were picked, inoculated in LB broth and incubated with shaking at 28°C for 24–48 h. Bacterial strains were preserved at −80°C as saturated cultures containing 20% (w/v) glycerol.

Partial 16S rRNA Gene Sequencing

For phylogenetic identification of maize seed endophytes we performed partial 16S rDNA (V1 to V3) PCR and sequencing as described by Reiter and Sessitsch (2006). Sequencing was performed by LGC Genomics (Berlin, Germany).

Preparation of Inoculum

Inoculum of the selected strains (S6, S7, S8) were prepared in 100 mL 10% tryptic soy broth in 250 mL Erlenmeyer flasks and incubated at 28 ± 2^{o}C for 48 h in an orbital shaking incubator (VWR International, GmbH) at 180 r min^{-1}. The optical density of the broth was adjusted to 0.5 measured at λ 600 nm using spectrophotometer (Gene Quant Pro, Gemini BV, The Netherlands) to obtain an uniform population of bacteria (10^8 - 10^9 colony-forming units (CFU) mL^{-1}) in the broth at the time of inoculation.

Testing the Effect of Endophytic Strains on Maize Under Axenic Conditions

Seeds were surface-sterilized by dipping them in 70% ethanol for 3 min and then in a 5% sodium hypochlorite solution for 5 min and subsequently thoroughly washing with sterilized distilled water. The efficacy of surface sterilization was checked by plating seeds, and aliquots of the final rinse onto 10% tryptic soy agar. Samples were considered to be successfully sterilized when no colonies were observed on the tryptic soy agar plates after inoculation for 3 days at 28°C. Surface-disinfected seeds of three maize cultivars (DaSilvie, Kaleo, and Mazurka) were immersed in the bacterial suspensions for 1 h. For the uninoculated control, sterilized tryptic soy broth was used for the seed treatment. Fifteen seeds per treatment were planted in plastic trays with sterilized compost (Blumenerde, COMPO SANA®) and trays were arranged using a randomized design with 3 replications resulting in total number of 45 seeds per treatment. The experiment was conducted for 24 days and data of shoot and root length as well as biomass were recorded.

Functional Characterization of Seed Endophytic Bacteria

Phenotypic, Physiological and Biochemical Characterization

Color and shape of bacterial colonies, growth behavior in different pHs, salt concentrations and C sources as well as aggregate and biofilm formation and motility were tested following the procedures described by Naveed et al., (2014). Biochemical testing of oxidase, catalase, gelatin hydrolysis and casein hydrolysis activity of the selected strains was performed according to Naveed et al., (2014).

Plant Growth Promoting Activities

Strains were tested for activities known to be involved in plant growth regulation and/or rhizosphere competence such as ACC-deaminase activity, auxin production, phosphate solubilization (organic/inorganic P) and siderophore production as well as ammonia, hydrogen cyanide and PHB production as described by Naveed et al. (2014).

Cell Wall-degrading Activities

Bacterial cell wall hydrolyzing activities such as amylase, cellulase, chitinase, lipase, pectinase, phosphatase, protease, and xylanase were screened on diagnostic plates as described by Naveed et al. (2014).

Antibiotic Resistance of the Isolates

Antibiotic resistance was tested individually on tryptic soy agar plates containing the antibiotics ampicillin, cycloheximide, gentamycin, kanamycin, chloramphenicol, rifampicin, spectinomycin, streptomycin or tetracycline respectively at the following concentrations: 25, 50, 75, 100 µg ml^{-1}. The plates were incubated at 28 ± 2°C for 5 days and resistance was observed in terms of bacterial growth.

Antagonistic Activities Against Plant Pathogens

The antagonistic activities of bacterial isolates were screened against plant pathogenic fungi (Fusarium caulimons, Fusarium graminarium, Fusarium oxysporum, Fusarium solani, Rhizoctonia solani, Thielaviopsis basicola) and oomycetes Phytophthora infestans, Phytophthora citricola, Phytophthora cominarum). Antagonistic activity of the bacterial isolates against fungi and oomycetes was tested by the dual culture technique on potato dextrose agar (PDA) and yeast malt agar (YMA) media as described by Naveed et al. (2014).

Statistical Analyses

The data of plant growth parameters and colonization were subjected to analyses of variance (ANOVA). The means were compared with Least Significant Difference (LSD) testing ($p < 0.05$) to detect statistical significance among treatments (Steel et al., 1997). Statistical analyses were conducted using SPSS software version 19 (IBM SPSS Statistics 19, USA).

Genomic DNA Isolation for Sequencing

For DNA isolation, the bacterial strains were grown by loop-inoculating one single colony in 5 mL LB broth. The bacterial cultures were incubated at 28 ± 2°C overnight at 180 rpm in a shaking incubator. The overnight cultures were used to inoculate 50 mL fresh LB broth and again incubated at 28 ± 2°C overnight at 180 rpm in a shaking incubator. Bacterial cells were harvested by centrifugation at 4700 rpm for 10 min at 4°C. DNA was extracted from bacterial cell pellets according to the following protocol: The cell pellet was washed with 5 mL lysis buffer (0.1 M NaCl; 0.05 M EDTA, pH 8.0), resuspended in 4 mL lysis buffer containing lysozyme (20 mg mL-1; Roche Diagnostics, Mannheim, Germany) and incubated at 37°C for 20 min. Then 300 µl of 10% sarkosyl was added and placed on ice for 5 min. DNA was extracted with phenol-chloroform-isoamylalcohol (25:24:1, Fluka, Sigma-Aldrich Co.) and re-extracted with chloroform (1:1, Merck, Darmstadt, Germany) followed by precipitation with 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volume of ice-cold absolute ethanol (Merck, Darmstadt, Germany) at −20°C overnight. DNA pellets were washed with 1 ml of 70% ethanol and dissolved in 100 µL TE buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA, pH 8·0). DNA was treated with RNase A (final concentration 0.2 gl −1; Invitrogen, Carlsbad, CA) for 90 min at 37°C. DNA quality was analyzed by electrophoresis (80 V) on 0.8% (w/v) agarose gels stained with ethidium bromide. DNA concentration was measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

Genome Sequencing, Assembly and Genome Alignment of *P. ananatis* S6, S7 and S8 strains

Genome sequencing of the three strains of *P. ananatis* (S6, S7, and S8) was done by GATC Biotech AG (Konstanz, Germany) using a Roche/454 GS-FLX system. After sequencing, pairwise analysis of average nucleotide identity (ANI) was performed between the *P. ananatis* strains with closed genome sequences and the strains S6, S7, and S8 draft genomes individually as described previously (Goris et al., 2007).

The raw reads from sequencing projects have been deposited at the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena/data/view/) under the following project accession numbers: *P. ananatis* S6, PRJEB7511; *P. ananatis* S7, PRJEB7512, and *P. ananatis* S8, PRJEB7513. Genome assemblies are available in ENA under accession numbers CVNF01000001 to CVNF01000077 for *P. ananatis* S6, CVNG01000001 to CVNG01000071 for *P. ananatis* S7 and CVNH01000001 to CVNH01000061 for *P. ananatis* S8. The contigs were assembled using AMOScmp comparative assembler (Pop et al., 2004) and the Roche GS de novo assembler package (Newbler v2.6) in the 454 GS-FLXTM system (http://www.454.com/), indipendently. For AMOScmp assembly, four complete genomes of *P. ananatis* strains (*P. ananatis* AJ13355, *P. ananatis* LMG20103, *P. ananatis* LMG5342 and *P. ananatis* PA13) were used as potential reference genomes. As the assemblies based on *P. ananatis* AJ13355 resulted into highest coverage and mapping quality, this genome was used as reference genome for AMOScmp. The result of quality and coverage control of the assembly of each *P. ananatis* genome sequence was calculated using Qualimap v.1.0 (Garcia-Alcalde et al., 2012). In repetitive regions, such as rRNA operons, the assembly was further evaluated based on the read coverage distribution. Whole genome comparisons between *P. ananatis* S6, S7, and S8 strains were performed using Mauve v.2.3.1 (Darling et al., 2004). In Mauve, the Progressive Mauve algorithm was used to order the contigs against *P. ananatis* AJ13355 as reference genome. Genome assemblies are accessible via http://fileshare.csb.univie.ac.at/pantoea/.

Overview of *P. ananatis* Genomes Used in the Current Study

Five complete genomes of *P. ananatis* strains with different life styles and environmental origin were used in the comparative genomics and phylogenetic analysis. *P. ananatis* PA13 (accession numbers CP003085 and CP003086) is known as a pathogen of rice causing grain and sheath rot (Choi et al., 2012). *P. ananatis* AJ13355 (accession numbers AP012032 and AP012033) shows saprophytic life style and was isolated from soil (Hara et al., 2012). *P. ananatis* LMG20103 (accession number CP001875) is a pathogenic strain causing the severe blight and dieback of *Eucalyptus* (De Maayer et al., 2010). *P. ananatis* LMG5342 (accession numbers HE617160 and HE617161) is an opportunistic human pathogen reported from clinical isolations (De Maayer et al., 2012b). *P. vagans* C9-1 (accession numbers CP002206,

CP001893, CP001894, and CP001895) is known as a common plant epiphyte (Smits et al., 2010).

Phylogenetic Analysis

We constructed a phylogenetic tree for *P. ananatis* S6, S7, S8 and the *Pantoea* genomes mentioned above. *P. vagans* C9-1 was included as outgroup. Mauve v2.3.1 (Darling et al., 2004) was used to identify specific and shared SNPs between all compared genomes. The alignments of the genomes were checked manually to eliminate possible false positive SNPs in less conserved regions, particularly if they occur in direct neighborhood of insertions and deletions. The obtained SNPs were filtered based on the position of phylogenetic markers of *P. ananatis* AJ13355 as reference [identified by AMPHORA2; Wu and Scott (2012)] to get the core SNPs of the genome sequences of *P. ananatis* strains. Afterwards, the phylogenetic tree was computed with Geneious 8.0 (Kearse et al., 2012) using 1000 runs for bootstrapping.

Genome Annotation

Gene prediction and annotation were obtained from the in-house ConsPred workflow. ConsPred consists of two phases: ab initio as well as homology-based predictions. Ab initio predictions are followed by Genemark.hmm (Lukashin and Borodovsky, 1998), Glimmer (Delcher et al., 2007), Prodigal (Hyatt et al., 2010), Critica (Badger and Olsen, 1999) and additional homology based information derived from a BLAST search against the NCBI non-redundant sequence database (NR) (Sayers et al., 2012). Protein domains were predicted by InterProScan (Zdobnov and Apweiler, 2001). For protein sequences without significant hits in NR, functional annotation of protein-coding genes was obtained by a similarity search against the UniProt/SwissProt database (Uniprot consortium, 2009). Non protein-coding elements such as tRNA and rRNA were predicted using tRNAScan and RNAmmer tools, respectively (Lowe and Eddy, 1997; Lagesen et al., 2007). Non-coding RNA genes (ncRNAs) were identified and annotated by a search against RFAM database (Griffiths-Jones et al., 2005). To check for the completeness of housekeeping genes in the genomes of strains S6, S7 and S8 we used AMPHORA2 (Wu and Scott, 2012) with 31 bacterial phylogenetic marker genes for inferring phylogenetic information.

Plasmid Sequence Alignment Analysis

To identify the plasmid sequences within the assembled contigs we compared the plasmid sequence of the closest reference genome (*P. ananatis* AJ13355) to the

assembly of *P. ananatis* S6, S7 and S8 strains using Mauve v2.3.1 (Darling et al., 2004).

To visualize the coverage of the plasmids in the draft genome sequences, the plasmid sequence of *P. ananatis* AJ13355 were used as reference for comparative circular alignments of the three *P. ananatis* S6, S7, and S8 strains using the BLAST Ring Image Generator (Stothard and Wishart, 2005; Alikhan et al., 2011).

Comparative Genome Analyses

Identification of Orthologous Groups

Paralogous and orthologous clusters were identified using OrthoMCL (Li et al., 2003) using the predicted proteomes of seven *P. ananatis* strains (*P. ananatis* AJ13355, *P. ananatis* LMG20103, *P. ananatis* LMG5342, *P. ananatis* PA13 and *P. ananatis* S6, S7, and S8 strains) which initially required an all-vs.-all blastp (E-value cut-off of 1 × 10−5). Then the mcl clustering algorithm was used to deduce the relationship between genes.

Identification of Eukaryotic-like Protein Domains

To identify eukaryotic-like protein domains (ELDs) in protein sequences in the genomes of strains S6, S7, and S8, those genomes were included in the individual ELD calculation procedure of the Effective web-portal (Jehl et al., 2011). The approach detects protein domains that are present in eukaryotic organisms and significantly enriched in pathogenic and symbiotic compared to non-pathogenic, non-host-associated bacteria. Using default settings, all eukaryotic-like protein domains with an enrichment score greater or equal to 4 were considered for comparison regarding functional differences in *P. ananatis* strains of diverse phenotype.

Results

Selection of Strains and Effects of Maize Seed Endophytes on Maize Seedling Growth

In a previous study, we isolated 90 bacterial strains from seeds of healthy maize plants grown at organic farming fields in Austria. Thirty-seven of these strains shared highest 16S rDNA sequence homology with *P. ananatis* strains (data not shown). Ten strains were randomly selected and tested for effects on seedling growth of maize grown in sterile hydroponic cultures (for a description see Naveed et al., 2014). Along with strains that did not influence maize seedling growth we found one strain with clear detrimental effect and other strains that promoted maize seedling

growth (data not shown). One representative of each group was selected and further tested on maize grown in compost. Strain S6 significantly increased seedling growth in all three maize cultivars tested compared to the control (Figure S1; Table 1). Depending on the plant variety root- and shoot-dry biomass was increased up to 47 and 41%, respectively. Root and shoot length was increased up to 57 and 41%, respectively. Strain S8 showed positive effects on plant growth in cultivar DaSilvie only but did not significantly affect growth of the cultivars Kolea and Mazurka (Table 1). In contrast, strain S7 had a negative effect on seedling growth in all the maize cultivars with the effect being significant in DaSilvie and Kolea and less pronounced in Mazurka (Table 1; Figure S1). Apart from reduced biomass S7 treated plants showed white streaks on leaves (Figure S2).

Table 1. Effect of inoculation with seed-associated endophytic bacteria on root/shoot length and biomass of maize seedlings

Strains	**DaSilvie**	**Kaleo**	**Mazurka**	**DaSilvie**	**Kaleo**	**Mazurka**
	Root length (cm)			**Shoot length (cm)**		
Control	16.67 fgh[a]	15.67 gh	19.00 ef	27.67 bcd	24.67 ef	25.33 def
P. ananatis S6	25.67 ab	24.50 bc	27.33 a	34.66 a	34.97 a	34.83 a
P. ananatis S7	16.33 fgh	15.00 h	17.67 fgh	26.67 cde	26.33 cde	23.33 f
P. ananatis S8	20.67 de	18.00 fg	21.38 de	30.00 b	28.33 bc	27.00 bcd
	Root dry biomass (mg)			**Shoot dry biomass (mg)**		
Control	20.98 cde	20.49 de	22.78 bcd	229.31 e	224.38 e	248.75 bcd
P. ananatis S6	29.57 a	30.09 a	30.51 a	323.11 a	324.00 a	330.42 a
P. ananatis S7	18.78 ef	17.23 f	21.24 cde	199.29 f	191.18 f	232.45 de
P. ananatis S8	24.02 b	22.03 bcd	23.39 bc	256.45 bc	241.25 cde	262.89 b

[a]Means sharing same letter(s) do not differ significantly at *P = 0.05*

Functional Characterization of Maize Seed Isolates Based on *in vitro* Assays

A range of activities known to contribute to plant growth promotion, stress tolerance or biocontrol was tested. The results of functional characterization are summarized in Table 2. All strains exhibited ACC-deaminase activity and showed auxin, NH_3 and siderophore production (qualitative). All three strains showed P-solubilization and were able to produce AHL and PHB. S6, S7 and S8 behaved similar in tests for motility and chemotaxis as well as the biochemical characters mentioned in Table 2. No strain produced EPS in our assays. Lipase, pectinase, phosphatase and xylanase activity was detected in all strains, whereas none of the strains showed amylase, cellulose, chitinase or protease activity. Strain S6 showed *in vitro* antagonistic activity against all bacterial pathogens tested but *F. solani.* Strain S7 inhibited growth of *F.*

oxysporum, *T. basicola* and *P. citricola* in our assays, whereas strain S8 negatively affected growth of *F. graminarium*, *F. oxysporum*, *R. solani* and *P. citricola.*

Table 2. Physico-chemical and growth-promoting characteristics of maize seed-borne endophytic bacteria

Characteristics	***P. ananatis* S6**	***P. ananatis* S7**	***P. anantis* S8**
Phenotypic characterization			
Colony color	Yellow	Yellow	Yellow
Colony morphology	Round	Round	Round
Bacterial growth conditions			
Temperature			
4^{o}C	+	+	+
42^{o}C	-	-	-
NaCl			
2%	+	+	+
6%	+	+	+
pH:			
5	+	+	+
12	+	+	+
Motility / Chemotaxis[a]			
Swimming	+	+	+
Swarming	++	+	+
Twitching	+	+	+
Biofilm formation			
OD (600 nm)	0.95±0.08	0.89±0.07	0.92±0.06
Biofilm (595 nm)	0.08±0.01	0.07±0.01	0.06±0.01
Aggregate stability (%)	32.61±2.13	28.61±1.93	30.61±2.01
Biochemical characterization[b]			
Catalase	+	+	+
Oxidase	-	-	-
Casein	-	-	-
Gelatin	3.5±0.15	2.9±0.10	3.2±0.12
Methanol	-	-	-
Ethanol	-	-	-
Growth promoting characterization[a]			
ACC-deaminase activity	+	+	+
Auxin production (IAA equivalent µg mL^{-1})			
Without L-TRP	0.87±0.55	0.68±0.52	0.78±0.54
With L-TRP	32.67±3.17	27.45±2.89	30.89±3.17
P-solubilization (Inorganic/organic P)			

$Ca_3(PO_4)_2$	1.6±0.10	1.2±0.14	1.4±0.14
$CaHPO_4$	1.5±0.08	1.0±0.06	1.2±0.08
Ca-Phytate	2.5±0.11	2.0±0.10	2.3±0.11
Na-Phytate	1.4±0.06	0.9±0.02	1.0±0.06
Exopolysaccharide	-	-	-
HCN production	-	-	-
NH_3 production	+	+	+
Siderophore	-	-	-
AHL	+	+	+
PHB	+	+	+
Enzyme hydrolyzing activity[a] (colon diameter cm)			
Amylase	-	-	-
Cellulase	-	-	-
Chitinase	-	-	-
Lipase	2.2±0.09	1.8±0.08	2.0±0.09
Pectinase	1.5±0.11	1.2±0.04	1.0±0.05
Phosphatase	1.6±0.08	1.3±0.07	1.0±0.08
Protease	-	-	-
Xylanase	1.3±0.09	0.8±0.02	1.0±0.06
Antibiotic resistance (µg mL^{-1})			
Ampicillin	-	-	-
Gentamycin	-	-	-
Kanamycin	-	-	-
Chloramphenicol	-	-	-
Rifampicin	-	-	-
Spectinomycin	-	-	-
Streptomycin	-	-	-
Tetracycline	-	-	-
Anti-fungal activity (colon diameter cm)			
F. caulimons	2.0±0.05	-	-
F. graminarium	1.2±0.04	-	1.0±0.04
F. oxysporum	1.0±0.03	1.0±0.03	1.0±0.03
F. solani	-	-	-
R. solani	1.8±0.07	-	1.5±0.06
T. basicola	1.2±0.05	1.2±0.05	-
Anti-oomycete activity			
P. infestans	3.4±0.11	-	-
P. citricola	3.5±0.09	3.0±0.09	3.0±0.12
P. cominarum	2.8±0.08	-	-

[a]Results in characterization table are of 4-6 replicates

[b] -, absent; +, present

[c] **+,** low efficiency; ++, medium efficiency; +++, high efficiency

Genome Sequences of *P. ananatis* Strains S6, S7 and S8

Genomic DNA of strains S6, S7 and S8 was sequenced and the generated raw reads represented 230, 76 and 79 million bases respectively (Table 3). The number of sequenced reads varied from 570,490 in strain S6 with an average length of 406 bp to 174,500 and 179,051 in S7 and S8 respectively, with an average length of 441 bp in both strains.

Table 3. Genome characteristics of sequencing and assembly of three strains of *P. ananatis* S6, S7 and S8.

Sequencing statistics						
Species	**Strain**	**Total nucleotides (bp)**		**Total reads**	**Average length**	
P. ananatis	S6	231,806,398		570,490	406	
P. ananatis	S7	76,917,000		174,500	441	
P. ananatis	S8	79,039,900		179,051	441	
Comparative assembly statistics (AMOScmp)						
Species	**Strain**	**# Contigs**	**N50**	**Total size**	**Assembly Score**	**Average Coverage**
P. ananatis	S6	93	127341	4361793	5972420241	43.08
P. ananatis	S7	92	134747	4553649	6669462411	13.62
P. ananatis	S8	63	178470	4618012	13082168280	14.38

The pairwise comparison of average nucleotide identity (ANI) of the draft genomes of strains S6, S7 and S8 with the *P. ananatis* AJ13355 genome showed that the similarity of the analyzed strains and strain AJ13355 exceeds 99% (Supplementary Table 1).

Comparative sequence assembly was performed by AMOScmp program (Pop et al., 2004) as a conservative method that uses the most similar available complete genome sequence as a reference to assemble the 454 reads (Table 3). The *P. ananatis* S6, S7 and S8 draft genomes consist of 93, 92 and 63 contigs, respectively, and range from 4.3 to 4.6 Mb in length.

De novo assembly resulted in almost the same coverage assembly but less assembly score and N50 value in comparison to AMOScmp assembler.

The comparison of draft genome assembly for *P. ananatis* S6, S7 and S8 against *P. ananatis* AJ13355 as reference genome are shown in Figure 1, illustrating a higher degree of genome conservation among the strains S6, S7, S8, as compared to *P. ananatis* AJ13355 (Figure 1). Phylogenetic analysis revealed a close relationship between strains S6, S7,S8 and the other four genomes of *Pantoea ananatis* in comparison to *P. vagans* (Figure 2).

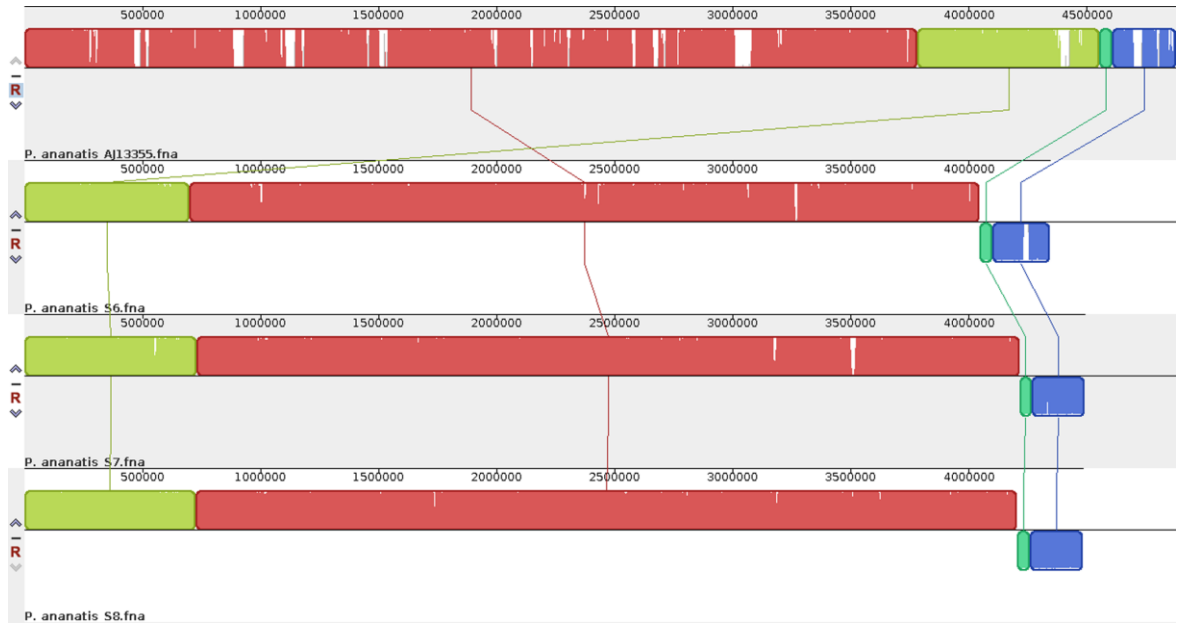


Figure 1. Genome-scale comparison for draft genome sequences of the three *P. ananatis* strains (S6, S7 and S8) and complete genome sequence of *P. ananatis* AJ13355. Homologous DNA regions among the strains are marked by the same coloured blocks, while gaps correspond to non-homologous regions. The figure was generated using nucleotide sequences of the genomes using Mauve v2.3.1.

Genome Annotation of *P. ananatis* S6, S7, and S8 Strains

The genome annotation of *P. annanatis* S6, S7 and S8 resulted in different numbers of protein-coding genes. The genome of strain S6 consists of 4.375 predicted coding sequences (CDSs), while S7 and S8 contain 4.516 and 4.528 predicted CDSs, respectively (Table 4). Seven 16S rRNA, seven 23S rRNA and eight 5S rRNA genes are encoded in each of the *P. annanatis* strains. In total all tRNA genes for 33 different anticodons were found in all *P. annanatis* strains.

The results of annotation analysis of three *P. ananatis* S6, S7 and S8 strains and reannotation of *P. ananais* strains (*P. ananatis* AJ13355, *P. ananatis* LMG20103, *P. ananatis* LMG5342 and *P. ananatis* PA13) are summarized in Table 4.

The fact that tRNA genes for all essential amino acids, the 16S rRNA gene and 31 housekeeping genes were found in the draft genomes of strains S6, S7 and S8 indicates that the genomes are close to complete. Moreover, the overall pattern of distribution of housekeeping genes and the gene copy number are identical to other members of the *Enterobacteriaceae* family.

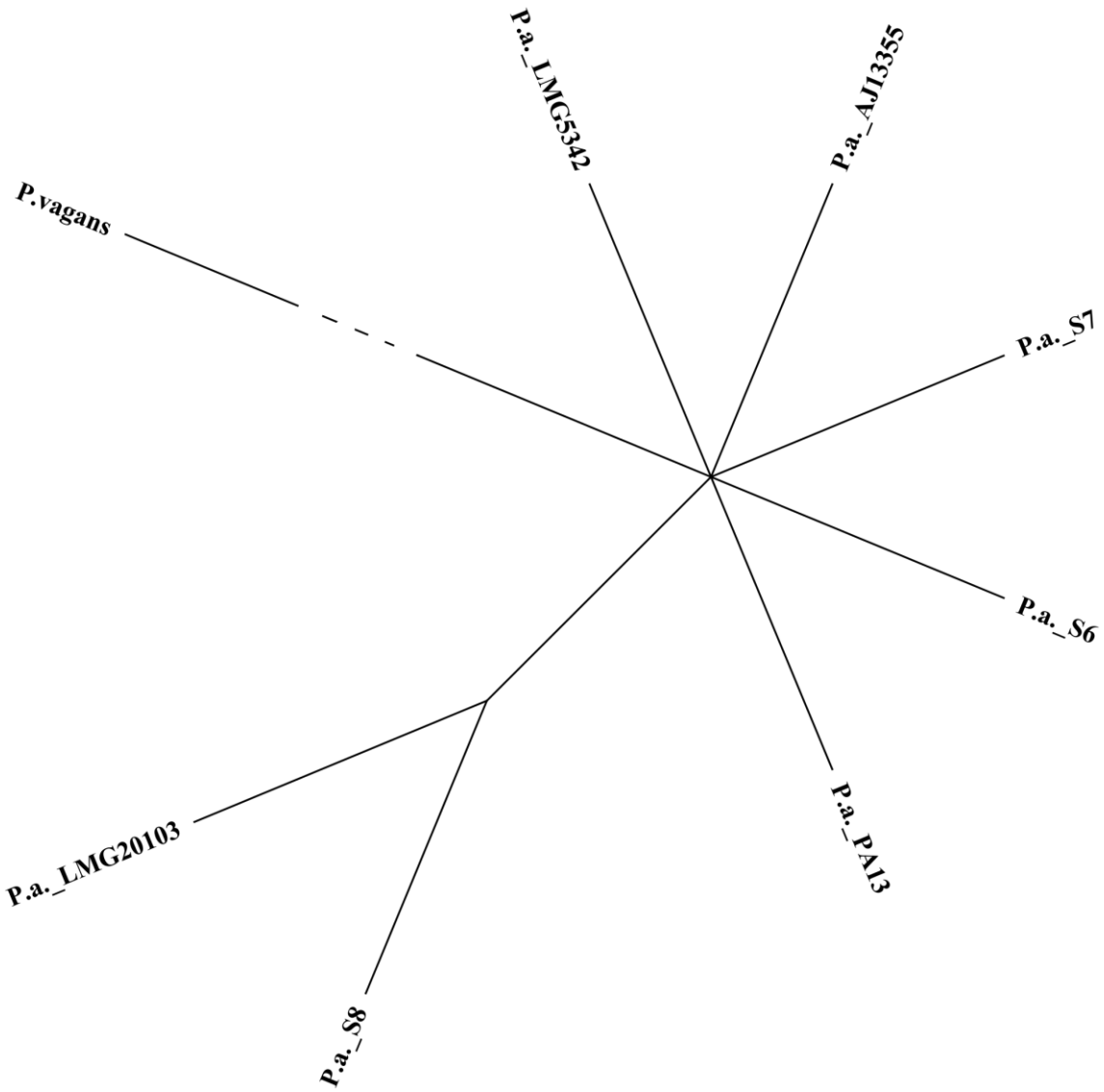


Figure 2. Phylogenetic trees of three strains of *Pantoea ananatis* S6, S7 and S8 with four closed genomes from *Pantoea* genus. *Pantoea vagans* was included as an out group (edge has been shortened).

To verify the sequence quality generated by 454 sequencing technology we identified putative pseudogenes represented by frameshifts in the draft genomes of *P. ananatis*. The low number of pseudogenes in the genomes of strains S6, S7 and S8 (11, 13 and 11 respectively) indicated that the genome draft has sufficient quality for further comparative genomics analysis (Table 4).

Table 4. Comparison of (A) draft genome annotation of three *P. ananatis* S6, S7 and S8 strains and (B) re-annotation of four complete genome of *P. ananatis* strains.

	Species	Strain	GC content	# CDS	tRNA	rRNA			ncRNA	Pseudo-genes
						5S	16S	23S		
(A)	*P. ananatis*	S6	54%	4375	69	8	7	7	144	11
	P. ananatis	S7	54%	4516	68	8	7	7	143	13
	P. ananatis	S8	54%	4528	68	8	7	7	142	11
(B)	*P. ananatis*	AJ13355	54%	4977	78	8	7	7	167	21
	P. ananatis	LMG5342	53%	5010	77	8	7	7	154	12
	P. ananatis	LMG20103	54%	4715	70	8	7	7	154	24
	P. ananatis	PA13	54%	5038	83	8	7	7	167	13

Plasmid Sequence Alignment Analysis

Five, six and seven contigs in *P. ananatis* S6, S7 and S8 genome sequences, respectively, were homologous with the plasmid sequence of *P. ananatis* AJ13355 (Supplementary Table 2). In total, 287, 271 and 276 genes were identified in the plasmid contigs of strains S6, S7 and S8. The core factors identified on the large universal *Pantoea* plasmid LPP-1 (De Maayer et al., 2012a) such as genes coding for thiamine biosynthesis proteins (*thiOSF*), pigment biosynthetic protiens (*crtEXYIBZ*), arbutin/cellobiose/salicin transport and catabolism components (*ascBFG*), malate:quinone oxidoreductase (*mqo*), 1,3-diaminopropane production (dat, ddc) and branched-chain amino acid transport protein (*azlDC*) are present on the plasmid sequences of *P. ananatis* S6, S7 and S8 (Supplementary Table 1).

Comparative circular blast alignments of the plasmid sequences in Figure 3 shows high homology between plasmid sequences of P. ananatis S6, S7 and S8 as compared to the P. Ananatis AJ13355 plasmid sequence (Stothard and Wishart, 2005; Alikhan et al., 2011) (Figure 3).

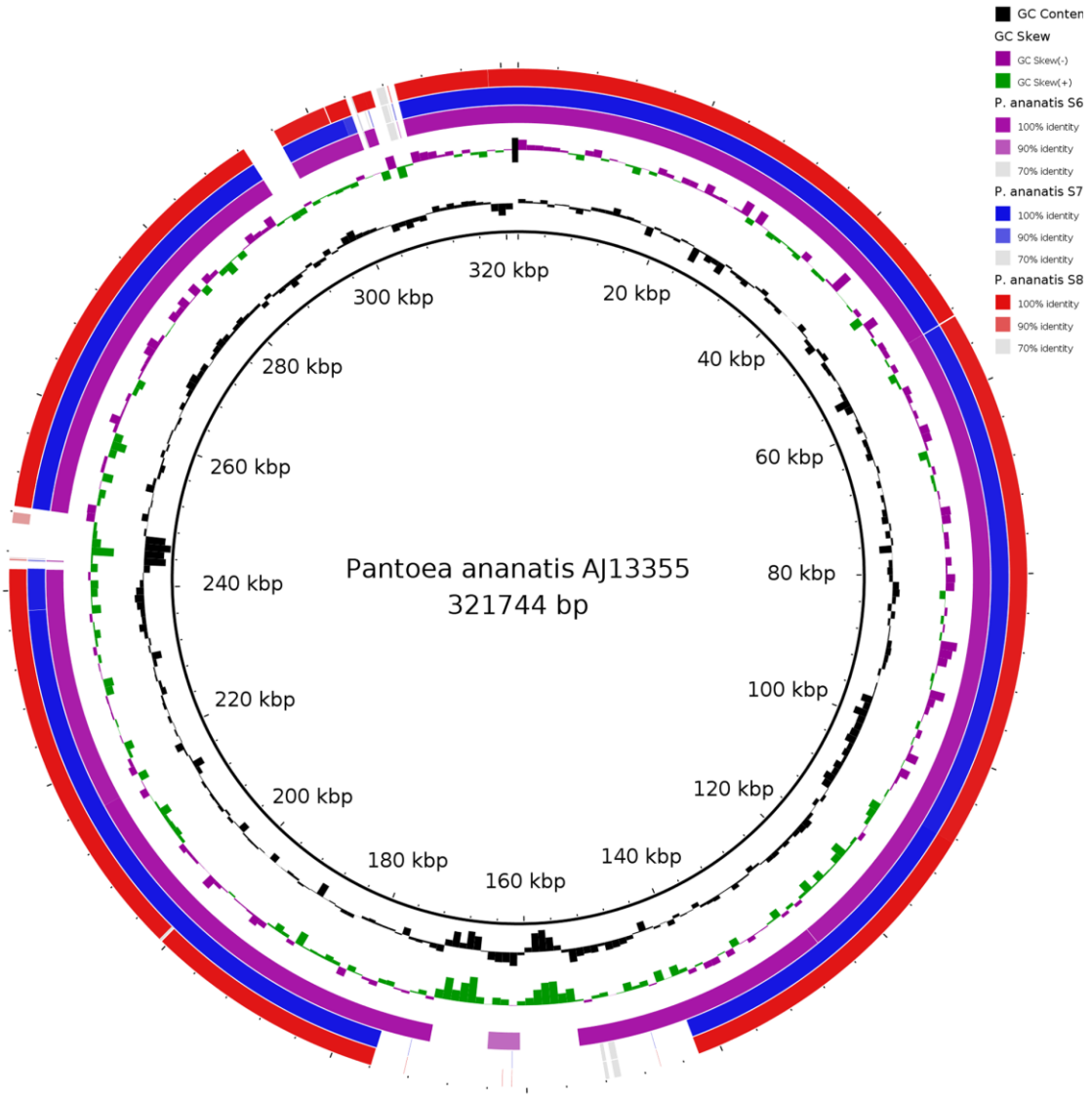


Figure 3. Comparison of the circular genome map of plasmid sequences of three *P. ananatis* S6, S7 and S8 genome structures with the known *P. ananatis* AJ13355 plasmid sequence as reference genome using blast ring image generator (BRIG). The inner circle shows the scale (bp). The first and the second rings show the GC content (black) and GC skew (purple/green), respectively, with respect to the reference genome. The 3rd, 4th and 5th rings show BLAST comparisons of *P. ananatis* strains S6, S7 and S8 plasmid sequences, respectively.

Comparative Genomics Analysis

To identify the core *P. ananatis* genome, we clustered orthologous groups from genes predicted in the seven *P. ananatis* genomes of this study (*P. ananatis* AJ13355, *P. ananatis* LMG20103, *P. ananatis* LMG5342, *P. ananatis* PA13 and strains S6, S7, S8) using OrthoMCL (Li et al., 2003).

Of the total 33,159 protein-coding genes in all *P. ananatis* strains, 31,987 genes clustered into 4,959 gene families. Out of these, 27,578 genes representing 3,785 gene families, were common to all *P. ananatis* strains, hereafter referred to as the core *P. ananatis* proteome (Figure 4).

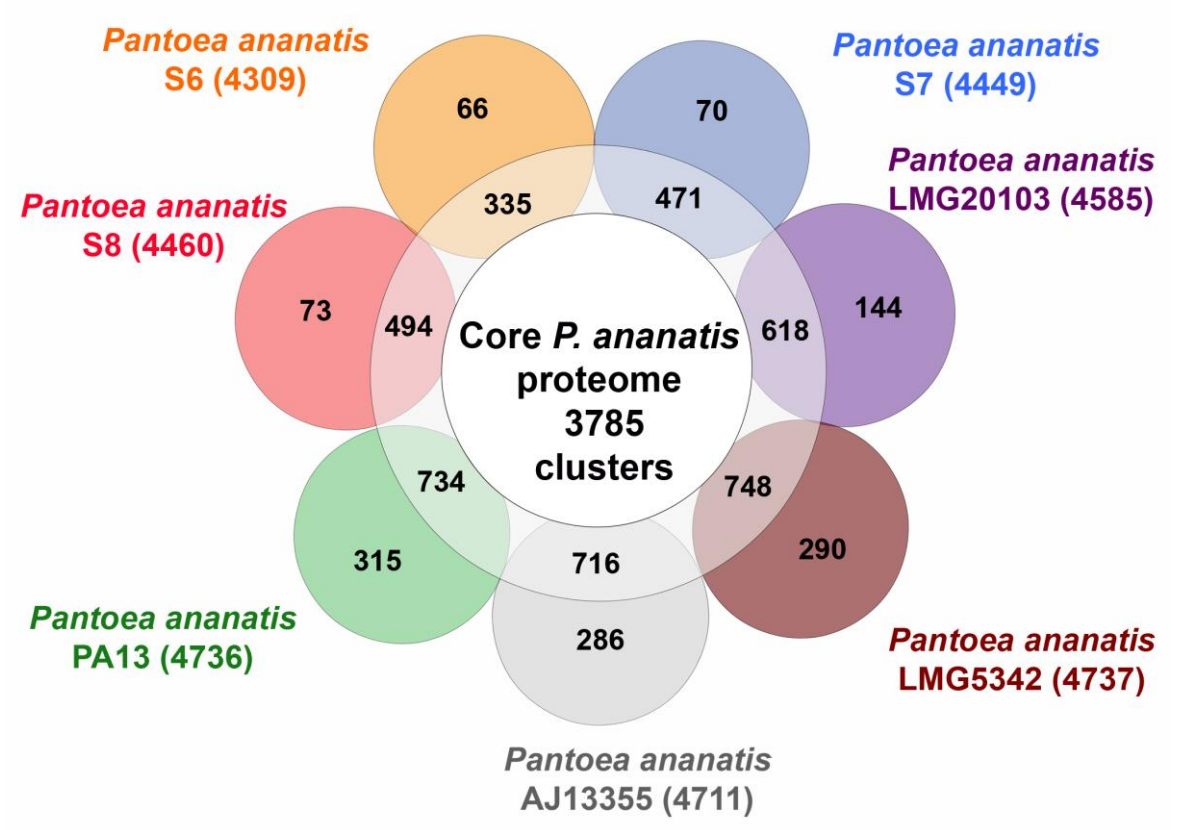


Figure 4. Clusters of orthologous gene families in seven *P. ananatis* strains identified by OrthoMCL. The inner circle shows the core proteome shared between all strains. The numbers of gene clusters shared between specific strains are shown in the ring. The specific proteins for each strain are indicated in each of the outer circles. The numbers outside the Venn diagram show the total number of genes (in parentheses) for each strain.

Fifty-three clusters were shared between *P. ananatis* S6 and S7. *P. ananatis* S7 and S8 have 207 clusters in common while *P. ananatis* S6 and S8 shared 79 clusters

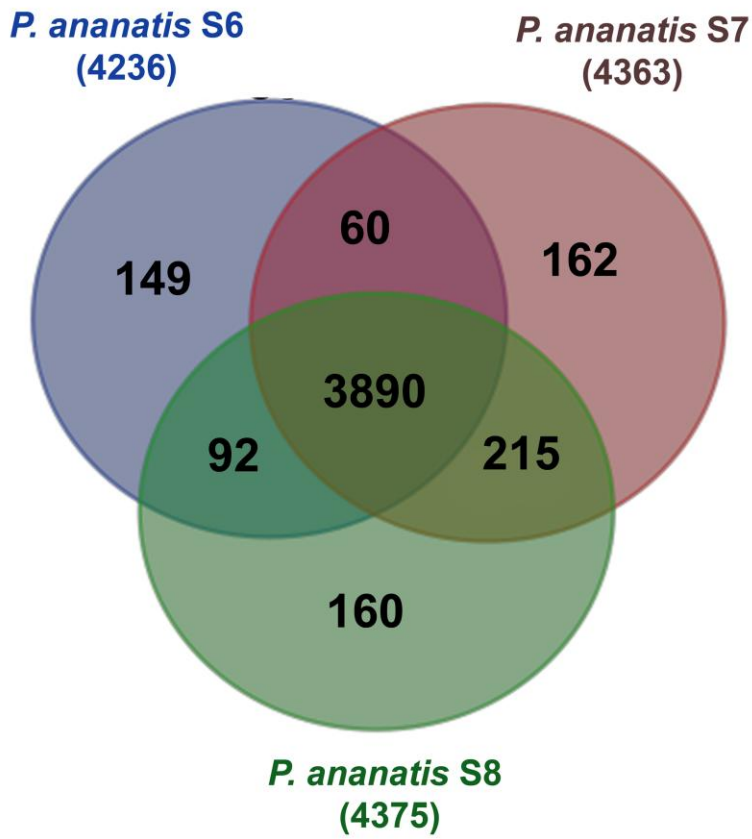


Figure 5. Venn diagram of OrthoMCL cluster distribution across three *P. ananatis* S6, S7 and S8 strains identified by OrthoMCL. The number of core proteome clusters, gene families shared between the species and the specific proteins for each strain is indicated in each of the components. The numbers outside the Venn diagram show the total number of genes (in parentheses) for each strain.

(Figure 5).

Gene Functional Classification of *P. ananatis* strains

To understand the functions of shared and specific genes between the *P. ananatis* strains, we analysed the functional categories of the respective *P. ananatis* gene clusters based on the NOG annotations (Jensen et al., 2008).

As expected, the core *P. ananatis* genes were categorized in functions involved in metabolism, cellular processes and signalling activity, information storage and processing (Supplementary Tables 3 and 5). The beneficial *P. ananatis* S6 specific genes encode proteins with putative functions in metabolism, signal transduction and

information storage and processing. Whereas, pathogenic *P. ananatis* S7 specific genes were specifically involved in cell cycle control, cell division, chromosome partitioning and amino acid transport. The commensal *P. ananatis* S8 specific genes were responsible for transcription and amino acid transport (Supplementary Tables 4, 6).

Functional Annotation of *P. ananatis* strains

Functional annotations of orthologous groups on the predicted proteomes of S6, S7 and S8 and the published *P. ananatis* genomes revealed functions that were common within all the genomes. This analysis also indicated gene families that

Figure 6. Functional COG categories in the genomes of the three *P. ananatis* strains S6, S7, and S8. (A) Comparison of the COG categories in the genomes of the three *P. ananatis* strains S6, S7, and S8. (B) The COG categories that present in two of the *P. ananatis* strains but are absent in the third strain (S6, S7, or S8). (C) The COG categories existing in only one of the *P. ananatis* strains (S6, S7, or S8).

cause differences among the strains on the functional level. The distribution of genes in COG functional categories is shown in Figure 6.

Type IV pilus biogenesis proteins such as *PilNQCWTZ*, type IV pilus secretin *PilQ*, pili assembly chaperone and prepilin type IV endopeptidase were found in the core proteome of *P. ananatis* strains (Supplementary Table 5). Interestingly, two genes related to pili assembly chaperon and fimbrial-type adhesion (the uncharacterized fimbrial chaperone *YhcA* and *F17a-A* fimbrial protein) were found in all *P. ananatis* strains but missing from beneficial *P. ananatis* S6 and two other genes related to the pili (fimbrial chaperone *YfcS* and chaperone protein *PapD*) were absent in pathogenic *P. ananatis* S7 but found in all other *P. ananatis* strains (Supplementary Table 7; Figure 6B).

Transposases related proteins such as tyrosine recombinase *XerD*, tyrosine recombinase *XerC* and site-specific recombinase *XerD* were identified in the core proteome of all *P. ananatis* strains. The only difference seen was the*YhgA*-like transposase that was found only in the commensal *P. ananatis* S8 (Supplementary Tables 5, 6; Figures 6A, C).

Virulence associated genes on mobile genetic elements showed that phage/bacteriophage related proteins such as bacteriophage P2 (*GpU)*, bacteriophage tail protein *Gp41* and phage tail tape measure protein are present in *P. ananatis* S7 and S8 and all other *P. ananatis* strains but do not have orthologs in the beneficial *P. ananatis* strain S6. The bacteriophage T7, Gp4, DNA primase/helicase is presented only in commensal *P. ananatis* S8 strain. Orthologous for integrases were not found in the beneficial strain S6 but were presented in the other strains (Supplementary Tables 6, 7; Figures 6B, C).

The chemotaxis related proteins such as chemotaxis methyl-accepting receptor (*CheR*) and chemotaxis proteins (*CheVWY*) were identified in the core proteome of the *P. ananatis* strains. The methyl-accepting chemotaxis signalling proteinI *TSR* is missing in the beneficial *P. ananatis* S6 strain but this strain contains the methyl-accepting chemotaxis signalling protein (*MCP*) which has the same activity in transducing the signal to downstream signalling proteins in the cytoplasm (Supplementary Tables 5, 6; Figures 6A, C).

The orthologous groups that are related to flagellar structures in the core *P. ananatis* proteome consists of flagella basal body P-ring formation proteins FlgAC; flagellar assembly proteins *FliH*; flagellar basal body rod protein components *FlaE, FlgJ*; flagellar hook-basal body complex proteins *FliELK*, *FlgCK* and flagellar biosynthesis proteins *FlhAQRO*. Other main flagellar related proteins are *FliJ*, *FlhE* flagellar motor protein MotA/MotB and *FliNGMSTZ* identified in the core proteome of *P. ananatis* strains (Supplementary Table 5; Figure 6A).

Gene families for T6SS loci were found on the core proteome of all seven *P. ananatis* strains investigated in this study. These common genes encoding *DotU* (COG2885), ATPase *ClpV1* (COG0542), FHA domain-containing protein (COG3456), *IcmF* (COG3523), lipoprotein *SciN* (COG3521), lysozyme-related protein (*impF*) (COG3518), OmpA/MotB domain (COG3455), T6SS RhsGE-associated *Vgr* family subset (COG3501), T6SS-associated *BMAA0400* (COG3913), T6SS -associated *ImpA* (COG3515) and *Hcp1* (COG3157) (Supplementary Table 5; Figure 6A). Our analysis showed also that the outer component of the T6SS, which consist of two proteins, *VgrG* (COG3501) and *Hcp* (COG3157) have also been identified as secreted effectors of the T6SS in some of the *P. ananatis* strains (Supplementary Table 7; Figure 6B). The effector protein genes *hcp*1, *hcp*1_2 and *hcp*1_3 loci are presented in six *P. ananatis* strains but absent in the pathogenic strain *P. ananatis* S7 (Table 5). The *HcpC* as major exported protein is missing in commensal *P. ananatis* S8 and pathogenic strains *P. ananatis* S7 and LMG5342, however it was present in the beneficial S6 strains, *P. ananatis* AJ13355 and pathogenic strains of *P. ananatis* PA13 and *P. ananatis* LMG20103 (Supplementary Table 7; Figure 6B).

Table 5. Hemolysin co-regulated effector proteins (Hcp) presented in the Type VI secretion system identified in orthologous clusters of *P. ananatis* strains.

P. ananatis **Strains**	**T6SS hemolysin co-regulated effector proteins (Hcp)**			
	Hcp1 (PAGR_1583)*	**Hcp1_2 (PAGR_1584)***	**Hcp1_3 (IPR008514)**	**HcpC (PAGR_3636)***
S6	BN1182_BN_00010	BN1182_BN_00910	BN1182_BN_00920	BN1182_CY_00040
S7	-	-	-	-
S8	BN1184_BC_00200	BN1184_BC_01090	BN1184_BC_01100	-
AJ13355	PantAJ13_A_20550	PantAJ13_A_21490	PantAJ13_A_21500	PantAJ13_B_01630

PA13	PantPA13_B_18870	PantPA13_B_18000	PantPA13_B_17990	PantPA13_B_41060
LMG20103	PantLMG20_A_26140	PantLMG20_A_27020	PantLMG20_A_27030	PantLMG20_A_46190
LMG5342	PantLMG53_A_18720	PantLMG53_A_17820	PantLMG53_A_17810	-

Hcp locus tag PAGR- are reported in Shyntum et al., 2014.

Eukaryotic-like Protein Domains in *P. ananatis* Strains

We identified eukaryotic-like protein domains (ELDs) in strains S6, S7 and S8 by applying the prediction framework of the Effective web-portal (Jehl et al., 2011). The prediction assigns a eukaryotic-like domain enrichment score (ELD score) to each protein domain, reflecting the maximal enrichment of that domain in any pathogen or symbiont compared to the background frequency of the protein domain in non-pathogenic, non-host-associated bacteria. A high ELD score equals strong enrichment of the protein domain in pathogenic/symbiotic bacteria and suggests an important functional role of the secreted protein in the interaction with the host cell. All ELDs with a significant ELD score greater or equal to 4 were considered to investigate the genomic variance of *P. ananatis* strains S6, S7 and S8 that cause different phenotypes in the host plant.

In summary, 29 different ELDs were predicted (Table 6). The majority, i.e. 26 ELDs are shared between all three genomes, supporting the assumption of a high average functional similarity of effector proteins. One eukaryotic-like protein domain, the tRNA delta-isopentenylpyrophosphate (IPP) transferase domain (PF01715) was exclusively found in the genome of the beneficial maize seed strain *P. ananatis* S6. IPP transferases are involved in the modification of tRNAs and convert A (37) to isopentenyl A (37). Another one was unique in the pathogenic strain S7 and contains the C terminal part of a GMP synthase (PF00958). This enzyme belongs to the family of ligases and is involved the biosynthesis of the nucleic acid guanine. A eukaryotic-like domain containing the signature of the collagen-binding domain of bacterial collagenases (PF12904) was found in S7, S8 and all other *Pantoea ananatis* genomes but was absent in S6.

Table 6. Differences of eukaryotic-like protein domain (ELD) enrichment in *P. ananatis* strains of diverse phenotype.

	Pfam ID	**Domain description**	**ELD Score***
Only in pathogenic *P. ananatis* S7	PF00958	GMP synthase C terminal domain	7

Only in beneficial *P. ananatis* S6	PF01715	IPP transferase	5
missing only in beneficial *P. ananatis* S6	PF12904	Putative collagen-binding domain of a collagenase	6
Shared in all *P. ananatis* S6, S7 and S8 strains	PF14328	Domain of unknown function (DUF4385)	-
	PF14145	YrhK-like protein	-
	PF13718	GNAT acetyltransferase 2	-
	PF13347	MFS/sugar transport protein	-
	PF10685	Stress-induced bacterial acidophilic repeat motif	-
	PF09825	Biotin-protein ligase N terminal	-
	PF09330	D-lactate dehydrogenase membrane binding	-
	PF08351	Domain of unknown function (DUF1726)	-
	PF08125	Mannitol dehydrogenase C-terminal domain	-
	PF07798	Protein of unknown function (DUF1640)	-
	PF07350	Protein of unknown function (DUF1479)	-
	PF06500	Alpha/beta hydrolase-unknown function- DUF1100	-
	PF05870	Phenolic acid decarboxylase (PAD)	-
	PF05704	Capsular polysaccharide synthesis protein	-
	PF05433	Glycine zipper 2TM domain	-
	PF05127	Helicase	-
	PF03825	Nucleoside H+ symporter	-
	PF02551	Acyl-CoA thioesterase	-
	PF01306	LacY proton/sugar symporter	-
	PF01232	Mannitol dehydrogenase Rossmann domain	-
	PF01204	Trehalase	-
	PF01116	Fructose-bisphosphate aldolase class-II	-
	PF00625	Guanylate kinase	-
	PF00328	Histidine phosphatase superfamily (branch 2)	-
	PF00294	pfkB family carbohydrate kinase	-
	PF00070	Pyridine nucleotide-disulphide oxidoreductase	-

*The domains without score have different scores for each *P. ananatis* strains.

Discussion

The genus *Pantoea* comprises bacteria that are frequently associated with eukaryotic hosts such as plants but strains, even those belonging to the same species (such as *P. ananatis*), have different type of interactions with their host ranging from pathogenicity to mutualism (De Maayer et al., 2014). In our study we showed that genetically closely related *P. ananatis* strains with different effects on plant growth colonize maize seeds.

In our study, the maize seed endophyte *P. ananatis* S6 showed clear beneficial effects on maize growth, while strain S7 induced weak pathogenicity symptoms. *P. ananatis* S8 had hardly any effect and can be considered as commensal. The pan genome of eight *P. ananatis* genomes indicated as open pan genome that they can colonize and exploit several different environmental niches by De Maayer et al. (2014). As three *P. ananatis* strains (S6, S7, and S8) are also capable to colonize inside maize seeds and interact with their host, we can expect that the pan genome of these strains can be defined as open pan genome.

Our comparative analysis showed that an average of 85–87% of CDSs predicted for each individual strain of *P. ananatis* S6, S7, and S8 have orthologs encoded by the genomes of the other strains (*P. ananatis* AJ13355, *P. ananatis* LMG20103, *P. ananatis* LMG5342 and *P. ananatis* PA13). These results suggest that the core genomes of strains S6, S7, and S8 strains are highly conserved (Figure 1). Despite the overall high degree of similarity between the core genomes of the three maize seed endophytes, we found differences in transposase/integrases/phage related genes, type VI secretion system, and eukaryotic-like protein domains. Similarly, the analysis of the open pan-genome of eight sequenced genomes of *P. ananatis* indicated that between 89.3 and 95.7% of the proteins are common between all strains and they are important for metabolism and cellular processes (De Maayer et al., 2014).

Genes of the accessory genome of selected *P. ananatis* strains analyzed by De Maayer et al. (2014) encoded mainly poorly characterized proteins including transposases, integrases, and mobile genetic elements. The role of horizontal gene transfer in the diversification of *P. ananatis* strains was suggested (De Maayer et al., 2014). Similarly, phage related genes were reported to have a significant role in transferring pathogenicity factors to their bacterial host and thereby to affect bacterial evolution (Lima-Mendez et al., 2008). Due to the differences found in regard mobile genetic elements such as integrase genes, transposase genes and phage related genes, our study confirms a potential role of these elements in the diversification of related strains colonizing the same ecological niche. An over-representation of transposase genes and mobile elements also indicates the genomes' potential for acquisition of novel functions. The reduced number of mobile elements in *P. ananatis*

S6 on the other hand could indicate high stability of its genome, implying good adaption to the habitat.

De Maayer et al. (2012a) proposed the Large Pantoea Plasmid (LLP-1) as genetic determinant of niche adaption and functional diversification of the genus *Pantoea*. All three maize seed endophytes S6, S7, and S8 contain a LLP-1 plasmid and no differences in LPP-1 related genes were found between the genomes of these strains and the core genome of *P. ananatis*. Our analysis revealed further that genes encoding the pigment biosyntetic (CrtEXYIBZ) and thiamine biosynthesis (ThiOSGF) proteins are present on the plasmid of *P. ananatis* S6, S7, and S8 (Supplementary Table 1). These genes are among those genes identified by De Maayer et al. (2014) to be specific for plant-associated bacteria (PAB) among *P. ananatis*. In addition, the core proteome of the maize seed *P. ananatis* strains contains PAB-specific CDs with prediction functions in metabolism and transport of carbohydrates, iron uptake and metabolism, and carbon, nitrogen and energy sources (De Maayer et al., 2014). In conclusion, our findings support the concept of functional diversification of the species *P. ananatis* proposed by De Maayer et al. (2014).

The T6SS is one of the most studied secretion system in *P. ananatis* (Coutinho and Venter, 2009; De Maayer et al., 2011; Shyntum et al., 2014). Three T6SS loci (T6SS-1, -2, and –3) have been described in *P. ananatis* strains, translocating effectors into the host plant (De Maayer et al., 2011; Shyntum et al., 2014). The T6SS-1 locus is found on the genomes of all *P. ananatis* strains, while T6SS-2 is restricted to pathogenic strains of *P. ananatis*. The presence of T6SS-1 and T6SS-2 in both pathogenic and non-pathogenic *P. ananatis* strains support the idea that the T6SS itself is not necessarily a determinant of pathogenicity and could play a role in competition against other microorganisms, fitness or niche adaptation (Weber et al., 2009; English et al., 2012; Shyntum et al., 2014). T6SS-3 was found to be mainly restricted to *P. ananatis* AJ13355, *P. ananatis* LMG 20103, and *P. ananatis* PA4 (De Maayer et al., 2014).

Beside the T6SS loci related genes VgrG and Hcp genes are present in the maize seed *P. ananatis* strains. The VgrG genes were found in S6, S7, and S8, whereas differences were seen in the presence of hemolysin co-regulated effector proteins (Hcp) between these three strains. A recent study showed that three hcp genes exist in *P. ananatis* strains comprising hcp-1, hcp-2 (having homologs in all sequenced

strains of *P. ananatis*) and hcp-3 genes (found in *P. ananatis* PA13) (Shyntum et al., 2014). The hcp-3 gene is highly divergent from hcp-1, hcp-2, and the T6SS associated hcp genes (Shyntum et al., 2014). The plant-beneficial strain S6 has orthologs with all hcp genes identified in the orthologous gene families, while plant-pathogenic strain S7 has no orthologs for hcp genes. HcpC is presented in all *P. ananatis* strains but it is missing from S7 and S8 strains. This Hcp protein is located on the plasmid sequence of this strain. Paralogs of hcp influence bacterial motility, protease production and biofilm formation (Sha et al., 2013). A potential role of Hcp and VgrG proteins in host interaction is not described. As all hcp genes are present in other *P. ananatis* strains (ranging from pathogenic to saprophytic life style), the hcp genes in in the beneficial *P. ananatis* strain S6 might not be responsible for the differences in the phenotype of plant-microbe interaction of the three maize seed strains S6, S7, and S8.

The analysis of effector candidates containing eukaryotic-like protein domains (ELDs) revealed varying molecular repertoire in the genomes of the three maize seed *P. ananatis* strains. The plant-beneficial strain S6 carries a gene for a tRNA delta-isopentenylpyrophosphate (IPP) transferase domain which is not present in the strains S7 and S8. In *E. coli* this enzyme is involved in increasing spontaneous mutation frequency when cells need to adapt to environmental stress (Connolly and Winkler, 1989). Moreover, tRNA modifications mediated by tRNA delta-isopentenylpyrophosphate (IPP) transferase are required for virulence in Shigella flexneri by regulating posttranscriptional expression of the regulatory gene virF (Durand et al., 1997). The collagen-binding domain of bacterial collagenases is missing in the beneficial *P. ananatis* S6, although present in S7 and S8. This domain is a major component of the extracellular matrix (ECM) and plays a role in cell attachment, haemostasis, differentiation and bacterial adhesion in human and plant pathogens (Foster and Hook, 1998). Interestingly, in Yersinia enterocolitica it is a part of the pathogenic bacterial strategy for avoiding host response (Nummelin et al., 2004). The GMP synthase domain exclusively found in the pathogenic *P. ananatis* strain S7 is known to play an important role in cell-to-cell signaling in regulation of virulence in the plant pathogen Xanthomonas campestris (Ryan et al., 2006a). This domain is also involved in aggregative behavior, adhesion, biofilm formation, and the virulence of animal and plant pathogens (Ryan et al., 2006b). The role of these EDLs

in the interaction of the three maize seed strains *P. ananatis* S7, S8, and S9 with maize plants remains unclear and merits further investigation.

Overall, our study showed that groups of bacterial endophytes with highly related genotypes but different phenotypes in terms of effects on host plants may exist in the same ecological niche. It can be expected that seed endophytes colonize, at least to a certain extent, plants derived from these seeds. Consequently, both, potential plant pathogens and mutualistic endophytes, may be together transmitted to the developing plant.

To predict the phenotype of plant-microbe interactions from traits manifested on the genome of bacteria is an attractive idea and would very much facilitate efforts in selecting microbial inoculants for improved plant production in sustainable agriculture. However, given the high genomic similarity between strains showing distinct phenotypes in regard to their interaction with plants, we conclude that plant pathogenicity and mutualism in *P. ananatis* may be based on rather subtle differences, e.g., on the expression of genes leading to plant defense reactions. Plant-bacteria interactions, irrespectively of whether pathogenic or beneficial must be considered as a multi-dimensional system and the expression of pathogenic or beneficial effects might depend on a multitude of parameters such as the plant/bacterial physiology, environmental conditions and/or a very fine-tuned interaction between bacterial elicitors and plant response.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

This work was supported by grants provided by the FWF (National Science Foundation, grant no. P22867-B16 and P26203-B22). The maize seeds of cultivars Kaleo and Mazurka used in this experiment were kindly provided by DOW AgroSciences, Neusiedl am See, Austria and cultivar DaSilvie was kindly provided by Saatbau Linz, Austria.

Supplementary Material

The supplementary tables for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00440/abstract

Supplementary Table 1. Proportion (%) of sequence similarity between seven *P. ananatis* strains. Proportions are given for the strains in each row in relation to the comparator strains in the top row.

Supplementary Table 2. The Plasmid contigs and coding sequences of three *P. ananatis* S6 (A), S7 (B) and S8 (C) strains.

Supplementary Table 3. Functional classification of proteins encoded in the core *P. ananatis* proteome.

Supplementary Table 4. Functional classification of proteins encoded by genes specific for *P. ananatis* strains.

Supplementary Table 5. Functional annotation of proteins encoded in the core *P. ananatis* proteome.

Supplementary Table 6. Functional annotation of proteins encoded by genes specific for *P. ananatis* strains.

Supplementary Table 7. Functional annotation of proteins encoded by genes shared between specific *P. ananatis* strains.

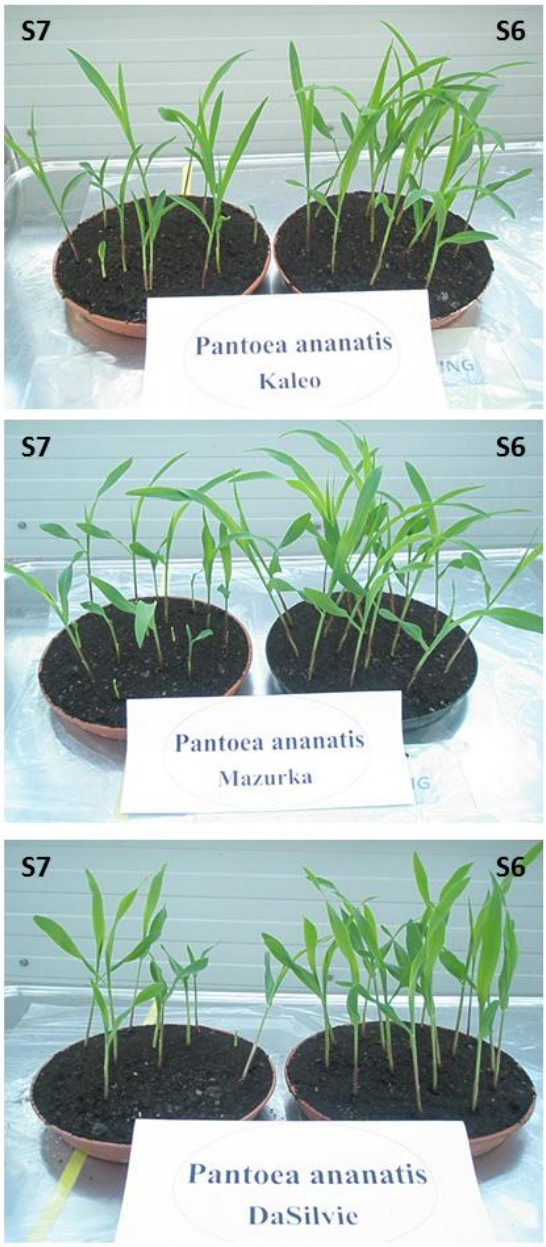


Figure S1. Effect of seed endophytes p. ananatis S6 and S7 on germination and seedling growth of three different maize hybrids (Kaleo, Mazurka and DaSilvie).

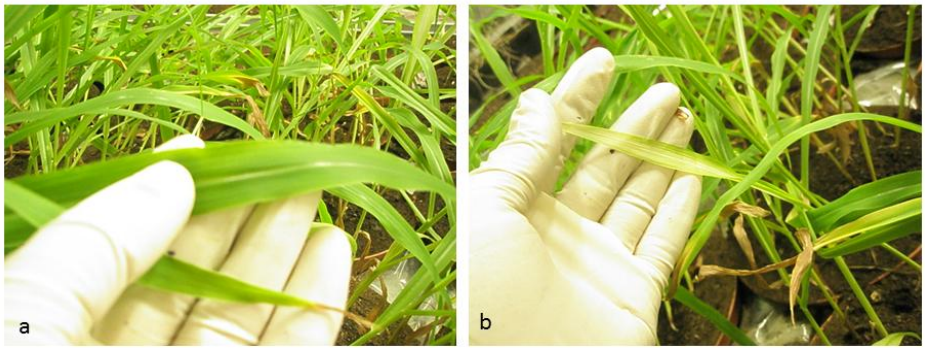


Figure S2. Seed inoculation with P. ananatis S7 causes white streaks on leaves of maze plants. Leaves of maize hybrids Mazurka (a) and DaSilvie (b) are shown.

References

Alikhan, N.-F., Petty, N., Ben Zakour, N., and Beatson, S. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12**,** 402.

Azad, H., Holmes, G., and Cooksey, D. (2000). A new leaf blotch disease of sudangrass caused by *Pantoea ananas* and *Pantoea stewartii*. *Plant Dis* 84**,** 973 - 979.

Badger, J.H., and Olsen, G.J. (1999). CRITICA: coding region identification tool invoking comparative analysis. *Mol Biol Evol* 16**,** 512-524.

Bashan, Y., Diab, S., and Okon, Y. (1982). Survival ofXanthomonas campestris pv.vesictoria in pepper seeds and roots in symptomless and dry leaves in non-host plants and in the soil. *Plant and Soil* 68**,** 161-170. doi: 10.1007/BF02373702.

Bordiec, S., Paquis, S., Lacroix, H., Dhondt, S., Ait Barka, E., Kauffmann, S., Jeandet, P., Mazeyrat-Gourbeyre, F., Clement, C., Baillieul, F., and Dorey, S. (2011). Comparative analysis of defence responses induced by the endophytic plant growth-promoting rhizobacterium Burkholderia phytofirmans strain PsJN and the non-host bacterium Pseudomonas syringae pv. pisi in grapevine cell suspensions. *J Exp Bot* 62**,** 595-603. doi: 10.1093/jxb/erq291.

Choi, O., Lim, J.Y., Seo, Y.S., Hwang, I., and Kim, J. (2012). Complete genome sequence of the rice pathogen *Pantoea ananatis* strain PA13. *J Bacteriol* 194**,** 531. doi: 10.1128/JB.06450-11.

Compant, S., Mitter, B., Colli-Mull, J.G., Gangl, H., and Sessitsch, A. (2011). Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microb Ecol* 62**,** 188-197. doi: 10.1007/s00248-011-9883-y.

Connolly, D.M., and Winkler, M.E. (1989). Genetic and physiological relationships among the miaA gene, 2-methylthio-N6-(delta 2-isopentenyl)-adenosine tRNA modification, and spontaneous mutagenesis in Escherichia coli K-12. *J Bacteriol* 171**,** 3233-3246.

Cota, L., Costa, R., Silva, D., Parreira, D., Lana, U., and Casela, C. (2010). First report of pathogenicity of *Pantoea ananatis* in sorghum (Sorghum bicolor) in Brazil. *Australas Plant Dis Notes* 5**,** 120 - 122.

Cother, E., Reinke, R., Mckenzie, C., Lanoiselet, V., and Noble, D. (2004). An unusual stem necrosis of rice caused by *Pantoea ananas* and the first record of this pathogen on rice in Australia. *Australas Plant Pathol* 33**,** 495 - 503.

Coutinho, T.A., and Venter, S.N. (2009). *Pantoea ananatis*: an unconventional plant pathogen. *Mol Plant Pathol* 10**,** 325-335. doi: 10.1111/j.1364-3703.2009.00542.x.

Darling, A., Mau, B., Blattner, F., and Perna, N. (2004). Mauve: multiple alignment of conserved genomic sequences with rearrangements. *Genome Res* 14**,** 1394 - 1403.

Delcher, A.L., Bratke, K.A., Powers, E.C., and Salzberg, S.L. (2007). Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23**,** 673-679. doi: 10.1093/bioinformatics/btm009.

De Maayer, P., Chan, W., Blom, J., Venter, S., Duffy, B., Smith, T., and Coutinho, T. (2012a). The large universal Pantoea plasmid LPP-1 plays a major role in biological and ecological diversification. *BMC Genomics* 13**,** 625.

De Maayer, P., Chan, W., Rezzonico, F., Buhlmann, A., Venter, S., Blom, J., Goesmann, A., Frey, J., Smiths, T., Duffy, B., and Coutinho, T. (2012b). Complete genome sequence of clinical isolate *Pantoea ananatis* LMG5342. *J Bacteriol* 194**,** 1615 - 1616.

De Maayer, P., Chan, W., Venter, S., Toth, I., Birch, P., Joubert, F., and Coutinho, T. (2010). Genome sequence of *Pantoea ananatis* LMG20103, the causative agent of Eucalyptus blight and dieback. *J Bacteriol* 192**,** 2936 - 2937.

De Maayer, P., Chan, W.Y., Rubagotti, E., Venter, S.N., Toth, I.K., Birch, P.R., and Coutinho, T.A. (2014). Analysis of the *Pantoea ananatis* pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts. *BMC Genomics* 15**,** 404. doi: 10.1186/1471-2164-15-404.

De Maayer, P., Venter, S., Kamber, T., Duffy, B., Coutinho, T., and Smits, T. (2011). Comparative genomics of the type VI secretion systems of *Pantoea* and *Erwinia* species reveals the presence of putative effector islands that may be translocated by the VgrG and Hcp proteins. *BMC Genomics* 12**,** e576.

Durand, J.M., Bjork, G.R., Kuwae, A., Yoshikawa, M., and Sasakawa, C. (1997). The modified nucleoside 2-methylthio-N6-isopentenyladenosine in tRNA of Shigella flexneri is required for expression of virulence genes. *J Bacteriol* 179**,** 5777-5782.

English, G., Trunk, K., Rao, V., Srikannathasan, V., Hunter, W., and Coulthurst, S. (2012). New secreted toxin and immunity proteins encoded within the Type VI secretion system gene cluster of Serratia marcescens. *Mol Microbiol* 86**,** 921 - 936.

Foster, T.J., and Hook, M. (1998). Surface protein adhesins of Staphylococcus aureus. *Trends Microbiol* 6**,** 484-488.

Garcia-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T.F., and Conesa, A. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28**,** 2678-2679. doi: 10.1093/bioinformatics/bts503.

Gitaitis, R.D., and Gay, J.D. (1997). First Report of a Leaf Blight, Seed Stalk Rot, and Bulb Decay of Onion by *Pantoea ananas* in Georgia. *Plant Disease* 81**,** 1096-1096. doi: 10.1094/PDIS.1997.81.9.1096C.

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57**,** 81-91. doi: 10.1099/ijs.0.64483-0.

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R., and Bateman, A. (2005). Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Research* 33**,** D121-D124. doi: 10.1093/nar/gki081.

Hara, Y., Kadotani, N., Izui, H., Katashkina, J., Kuvaeva, T., Andreeva, I., Golubeva, L., Malko, D., Makeev, V., Mashko, S., and Kozlov, Y. (2012). The complete genome sequence of *Pantoea ananatis* AJ13355, an organism with great biotechnological potential. *Appl Microbiol Biotechnol* 93**,** 331 - 341.

Hardoim, P.R., Hardoim, C.C., Van Overbeek, L.S., and Van Elsas, J.D. (2012). Dynamics of seed-borne rice endophytes on early plant growth stages. *PLoS One* 7**,** e30438. doi: 10.1371/journal.pone.0030438.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11**,** 119. doi: 10.1186/1471-2105-11-119.

Jehl, M.A., Arnold, R., and Rattei, T. (2011). Effective--a database of predicted secreted bacterial proteins. *Nucleic Acids Res* 39**,** D591-595. doi: 10.1093/nar/gkq1154.

Jensen, L.J., Julien, P., Kuhn, M., Von Mering, C., Muller, J., Doerks, T., and Bork, P. (2008). eggNOG: automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res* 36**,** D250-254. doi: 10.1093/nar/gkm796.

Johnston-Monje, D., and Raizada, M.N. (2011). Conservation and diversity of seed associated endophytes in Zea across boundaries of evolution, ethnography and ecology. *PLoS One* 6**,** e20396. doi: 10.1371/journal.pone.0020396.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28**,** 1647-1649. doi: 10.1093/bioinformatics/bts199.

Kim, H.J., Lee, J.H., Kang, B.R., Rong, X., Mcspadden Gardener, B.B., Ji, H.J., Park, C.S., and Kim, Y.C. (2012). Draft genome sequence of Pantoea ananatis B1-9, a nonpathogenic plant growth-promoting bacterium. *J Bacteriol* 194**,** 729. doi: 10.1128/JB.06484-11.

Kloepper, J.W., Mcinroy, J.A., Liu, K., and Hu, C.-H. (2013). Symptoms of Fern Distortion Syndrome Resulting from Inoculation with Opportunistic Endophytic Fluorescent *Pseudomonas* spp. *PLoS ONE* 8**,** e58531. doi: 10.1371/journal.pone.0058531.

Lagesen, K., Hallin, P., Rodland, E.A., Staerfeldt, H.H., Rognes, T., and Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35**,** 3100-3108. doi: 10.1093/nar/gkm160.

Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13**,** 2178-2189. doi: 10.1101/gr.1224503.

Lima-Mendez, G., Van Helden, J., Toussaint, A., and Leplae, R. (2008). Prophinder: a computational tool for prophage prediction in prokaryotic genomes. *Bioinformatics* 24**,** 863 - 865.

Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Research* 25**,** 0955-0964. doi: 10.1093/nar/25.5.0955.

Lukashin, A., and Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. 26, 1107–1115. doi: 10.1093/nar/26.4.1107

Mundt, J.O., and Hinkle, N.F. (1976). Bacteria within ovules and seeds. *Appl Environ Microbiol* 32**,** 694-698.

Naveed, M., Hussain, M.B., Zahir, Z.A., Mitter, B., and Sessitsch, A. (2013). Drought stress amelioration in wheat through inoculation with Burkholderia phytofirmans strain PsJN. *Plant Growth Regulation* 73**,** 121-131. doi: 10.1007/s10725-013-9874-8.

Nummelin, H., Merckel, M.C., Leo, J.C., Lankinen, H., Skurnik, M., and Goldman, A. (2004). The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel β-roll. *The EMBO Journal* 23**,** 701-711. doi: 10.1038/sj.emboj.7600100.

Paccola-Meirelles, L.D., Ferreira, A.S., Meirelles, W.F., Marriel, I.E. and Casela, C.R. (2001). Detection of a bacterium associated with a leaf spot disease of maize in Brazil. Journal *Phytopathology* 149, 275-279.

Pop, M., Phillippy, A., Delcher, A.L., and Salzberg, S.L. (2004). Comparative genome assembly. *Brief Bioinform* 5**,** 237-248.

Reiter, B., and Sessitsch, A. (2006). Bacterial endophytes of the wildflower Crocus albiflorus analyzed by characterization of isolates and by a cultivation-independent approach. *Canadian Journal of Microbiology* 52**,** 140-149. doi: 10.1139/w05-109.

Rijavec, T., Lapanje, A., Dermastia, M., and Rupnik, M. (2007). Isolation of bacterial endophytes from germinated maize kernels. *Can J Microbiol* 53**,** 802-808. doi: 10.1139/w07-048.

Ryan, R.P., Fouhy, Y., Lucey, J.F., Crossman, L.C., Spiro, S., He, Y.-W., Zhang, L.-H., Heeb, S., Cámara, M., Williams, P., and Dow, J.M. (2006a). Cell–cell signaling in Xanthomonas campestris involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proceedings of the National Academy of Sciences* 103**,** 6712-6717. doi: 10.1073/pnas.0600345103.

Ryan, R.P., Fouhy, Y., Lucey, J.F., and Dow, J.M. (2006b). Cyclic Di-GMP Signaling in Bacteria: Recent Advances and New Puzzles. *Journal of Bacteriology* 188**,** 8327-8334. doi: 10.1128/jb.01079-06.

Sayers, E.W., Barrett, T., Benson, D.A., Bolton, E., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., Dicuccio, M., Federhen, S., Feolo, M., Fingerman, I.M., Geer, L.Y., Helmberg, W., Kapustin, Y., Krasnov, S., Landsman, D., Lipman, D.J., Lu, Z., Madden, T.L., Madej, T., Maglott, D.R.,

Marchler-Bauer, A., Miller, V., Karsch-Mizrachi, I., Ostell, J., Panchenko, A., Phan, L., Pruitt, K.D., Schuler, G.D., Sequeira, E., Sherry, S.T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T.A., Wagner, L., Wang, Y., Wilbur, W.J., Yaschenko, E., and Ye, J. (2012). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 40**,** D13-25. doi: 10.1093/nar/gkr1184.

Serrano, F. (1928). Bacterial fruitlet brown-rot of pineapple in the Philippines. *Phil J Sci* 36**,** 271 - 305.

Sha, J., Rosenzweig, J., Kozlova, E., Wang, S., Erova, T., Kirtley, M., Van Lier, C., and Chopra, A. (2013). Evaluation of the roles played by Hcp and VgrG type 6 secretion system effectors in Aeromonas hydrophila SSU pathogenesis. *Microbiology* 159**,** 1120-1135. doi: 10.1099/mic.0.063495-0.

Shyntum, D., Venter, S., Moleleki, L., Toth, I., and Coutinho, T. (2014). Comparative genomics of type VI secretion systems in strains of Pantoea ananatis from different environments. *BMC Genomics* 15**,** 163.

Smits, T., Rezzonico, F., Kamber, T., Goesmann, A., Ishimaru, C., Stockwell, V., Frey, J., and Duffy, B. (2010). Genome sequence of the biocontrol agent Pantoea vagans C9-1. *J Bacteriol* 192**,** 6486 - 6487.

Stall, R.E., Alexander, L.J., and Hall, C.B. (1969). Effect of tobacco mosaic virus and bacterial infections on occurrence of graywall of tomato (Erwinia ananas). *Fla State Hort. Soc. Proc* 81**,** 157–161.

Steel, R.G.D., Torrie, J.H., and Dicky, D.A. (1997). Principles and procedures of statistics: A biometrical approach. . *McGraw-Hill Book Int. Co., Singapore.* 3rd edition.

Stothard, P., and Wishart, D.S. (2005). Circular genome visualization and exploration using CGView. *Bioinformatics* 21**,** 537-539. doi: 10.1093/bioinformatics/bti054.

Uniprot consortium (2009). The Universal Protein Resource (UniProt) *Nucleic Acids Res* 37**,** D169-174. doi: 10.1093/nar/gkn664.

Weber, B., Hasic, M., Chen, C., Wai, S., and Milton, D. (2009). Type VI secretion modulates quorum sensing and stress response in Vibrio anguillarum. *Environ Microbiol* 11**,** 3018 - 3028.

Wells, J., Sheng, W., Ceponis, M., and Chen, T. (1987). Isolation and characterization of strains of Erwinia ananas from honeydew melons. *Phytopathol* 77**,** 511 - 514.

Wilson, D. (1995). Endophyte: The Evolution of a Term, and Clarification of Its Use and Definition. *Oikos* 73**,** 274-276.

Wu, M., and Scott, A.J. (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* 28**,** 1033-1034. doi: 10.1093/bioinformatics/bts079.

Zdobnov, E., and Apweiler, R. (2001). InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17**,** 847 - 848.

Chapter V

Modification of a plant microbiome and traits by adding new microbes to seeds during seed production

Under submission

Letter for

NATURE

Modification of a plant microbiome and traits by adding new microbes to seeds during seed production.

Birgit Mitter[1*], Nikolaus Pfaffenbichler[1], Richard Flavell[2], Stéphane Compant[1], Livio Antonielli[1], Alexandra Petric[1], Teresa Berninger[1], Naveed Muhammad[1†], Raheleh Sheibani-Tezerji[1], Geoffrey von Maltzahn[2] and Angela Sessitsch[1]

[1] Health & Environment Department, Bioresources Unit, AIT Austrian Institute of Technology GmbH, Konrad-Lorenz-Strasse 24, 3430 Tulln, Austria

[2] Indigo, 101 Main Street, 14th floor, Cambridge, MA 02142, USA

[†]Present Address: Muhammad Naveed, Institute of Soil and Environmental Sciences, University of Agriculture, Faisalabad, Faisalabad, Pakistan

Running title: Modification of seed microbiomes by introduction of bacteria into the flowers of parent plants.

Keywords: seed, endophyte, microbiome, *Burkholderia phytofirmans* PsJN, EndoSeed, flowers, strain delivery

The microbial component of dry seeds, inherited between plant generations together with the plant genomes, is very important for germination, plant performance and survival. Internal plant microbiomes are complex communities of endophytes that interact with plant cells and the roles of individual members in programming plant growth are poorly understood. Optimization of the microbiome composition needs to be addressed in plant breeding so that seeds planted are optimal for food, feed and fiber production in the relevant environments. Here we describe a new, far reaching, approach for changing the microbiome of elite crop seed embryos, of both monocots and dicots and for revealing the roles of members of seed microbiomes in plant growth. Selected microbes are introduced into the parent plant before seed development occurs and these become incorporated into the seed microbiome thereby achieving vertical inheritance to the offspring generation. We illustrate both the introduction of a bacterium, not usually found in seeds, into the seeds of multiple plant species and the consequential modifications to seed microbiome composition and growth traits in wheat, thereby proving the role of the seed-based microbiome in determining plant traits. The new approach has many advantages over mixing seeds with the microbes before planting, especially in relation to microbe stability, and should have significant impact on fundamental plant-endophyte association studies as well as plant-microbe optimization in agriculture.

Plant internal microbiomes are complex communities of archaea, bacteria and fungi, which live as endophytes in all plants[1,2]. Their importance to plant growth and survival has recently been recognized much more extensively, following the series of revelations in humans about the far reaching importance of the microbiomes for healthy growth[3-8]. All organs of plants have been found to host a microbiome. However, it is likely that it is the pre-existing microbiome of the planted seed that provides the foundation for successful plant growth, before being augmented by microbes from the soil. Seed microbiomes have not been studied extensively or defined until recently[9,10]. The microbiome typically found in seeds consists of a limited range of microbial species[11]. It appears to have evolved by co-selection with the plant species, providing important traits for plant survival[12-14]. Its genes presumably complement those encoded in plant chromosomes and hence plant traits and evolution are determined by both plant and microbial genomes[2,15,16]. The many

roles of plant microbiomes imply that plant breeders need to embrace creation of the best plant-microbiome associations for optimum plant performance[17]. We describe, here, the means of introducing new microbes into seeds to modify the plant microbiome and plant traits in defined ways (Extended Data Fig. 1). The bacteria are sprayed on the parent flowers, enter the plant, colonize the emerging seeds and are thus inherited between the generations. After the internally colonized seeds are planted, the bacteria become active, proliferate and colonize the new generation plant, thereby unfolding effects on growth regulation from the first day of germination of the new crop generation.

We demonstrated the feasibility of modifying seed microbiomes in a targeted, directed way by using the bacterium *Burkholderia phytofirmans* PsJN – a well-studied endophyte and a powerful plant growth regulator, able to establish populations in a broad range of genetically unrelated plants[18,19]. Initially, a variant of strain PsJN chromosomally tagged with the beta-glucuronidase gene for detection and monitoring of the strain by color formation[20] was either applied on seeds or sprayed on female flowers of maize (*Zea mays* L. cvs. Peso and Morignon) in the greenhouse. At maturity, we detected gus stained PsJN cells in maize seeds of plants that had been sprayed with the bacterium (Fig. 1) at viable population densities that ranged from 10^2-10^5 CFU g^{-1} fresh weight but no gus stained cells were recovered from control

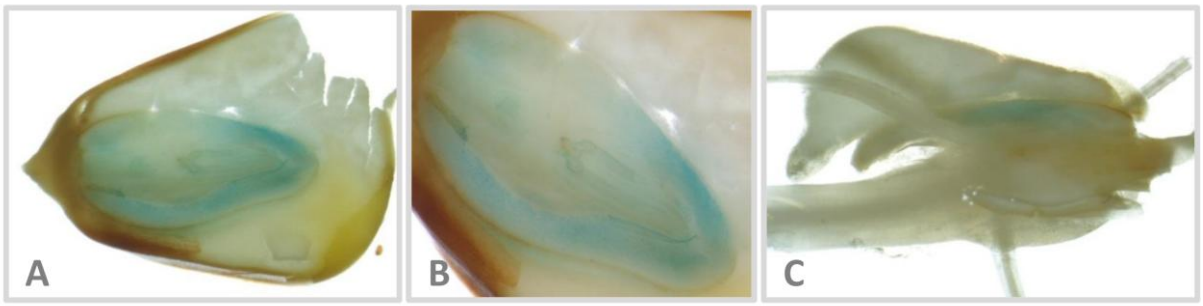


Figure 1: Light microscopy images of a mature seed colonized by *Burkholderia phytofirmans* strain PsJN::gusA. The blue is due to gus stained bacterial cells. Strain PsJN is present inside the embryo (a, b) and in radicals (c). PsJN starts moving from embryo to germinated parts (c). In the second picture (b) we present a zoom-in of the first photograph (a).

seeds from flowers sprayed with a solution lacking PsJN. Strain PsJN was not recovered from next generation seeds when the bacterium was applied on seed. After 12 months of storage of PsJN-colonized seeds, we still recovered about 100 viable cells per g maize seeds illustrating the stability of the PsJN in seeds.

This finding prompted us to test whether strain PsJN colonizes also seeds of dicotyledonous plants when it is sprayed on flowers and we performed pot experiments in the greenhouse with soy (*Glycine max* L. cvs. Essor and Merlin) and pepper (*Capsicum annuum* L. cv. Feher). *B. phytofirmans* PsJN was localized inside soy and pepper seeds by fluorescence *in situ* hybridization (FISH) using a specific probe targeting the 23S rRNA gene of *B. phytofirmans* and general probes for bacteria. Yellow fluorescent PsJN cells were found inside the embryo of soy along with other bacteria. Especially, *B. phytofirmans* was detected in the cotyledon part of the embryo together with other bacteria (green fluorescent) (Fig.2b-d), which also colonized the seed coat (Fig. 2a), while in control seeds only the native bacteria were present (Fig. 2d). By using the NONEUB probe (does not target bacterial sequences), neither *B. phytofirmans* nor other bacteria could be visualized inside seed tissues (Fig. 2e-h); only few natural green/blue-cyan-autofluorescent microbes could be seen inside the embryo of seeds colonized by PsJN and in control seeds (Fig. 2e-h). The number of PsJN bacteria detected in soy seeds (tested by strain-specific qPCR) ranged from about 360 to about 4500 genome equivalents per seed. Similar results were obtained with pepper plants, and PsJN was detected within the embryo with other bacteria that were also detected on the seed coat (Extended Data Fig.2).

The next step in our study was to test whether we could modify the microbiome of seeds during seed production in the field. We planted wheat (*Triticum aestivum* cv. Trappe) in an experimental field in Tulln (Austria). At flowering we applied *B. phytofirmans* PsJN only or a mixture of *B. phytofirmans* PsJN and *Paenibacillus* sp. S10, respectively. S10 had been isolated from surface-sterilized maize seeds. At seed maturity we found strain PsJN to be effectively introduced into the seeds (Fig. 3d) – 21 out of 24 seeds were tested positive in PsJN-specific qPCR assays. This means that up to 92% of wheat seeds became colonized by strain PsJN after spraying of young parent flowers. The number of genome equivalents per seed varied strongly and reached up to 7000 in some samples (Fig. 3c). Simultaneous application of *B. phytofirmans* PsJN with another bacterial strain (*Paenibacillus* sp. S10) was less efficient. Only 13% of the seeds harbored strain PsJN and the number of genome equivalents per seed reached only 62 in assayed samples, perhaps due to competition between PsJN and S10 (Fig. 3c,d).

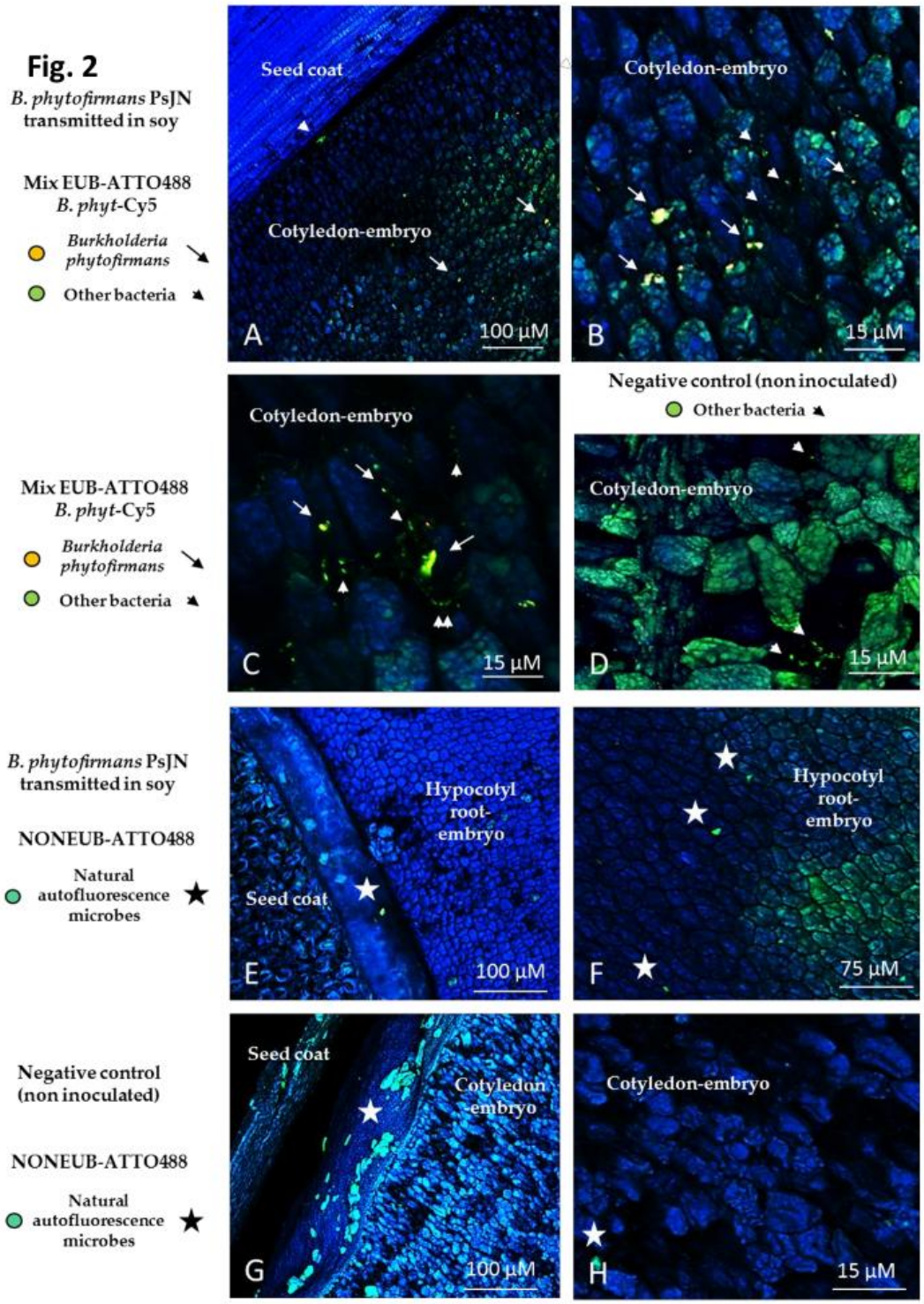


Figure 2: Visualization of *B. phytofirmans* PsJN in seeds of Glycine max L. (soy) by DOPE-FISH/CSLM microscopy. The mixEUB and *B. phytofirmans* probes were applied on seed from parent plants sprayed or not sprayed with strain PsJN. The presence of *B. phytofirmans* inside the embryo along with other microbes is shown (a-c). *B. phytofirmans* cells were not detected in seeds of non-inoculated plants (d) and the negative control using the NONEUB probe does not show bacteria except for a few autofluorescent microbes (e-h).

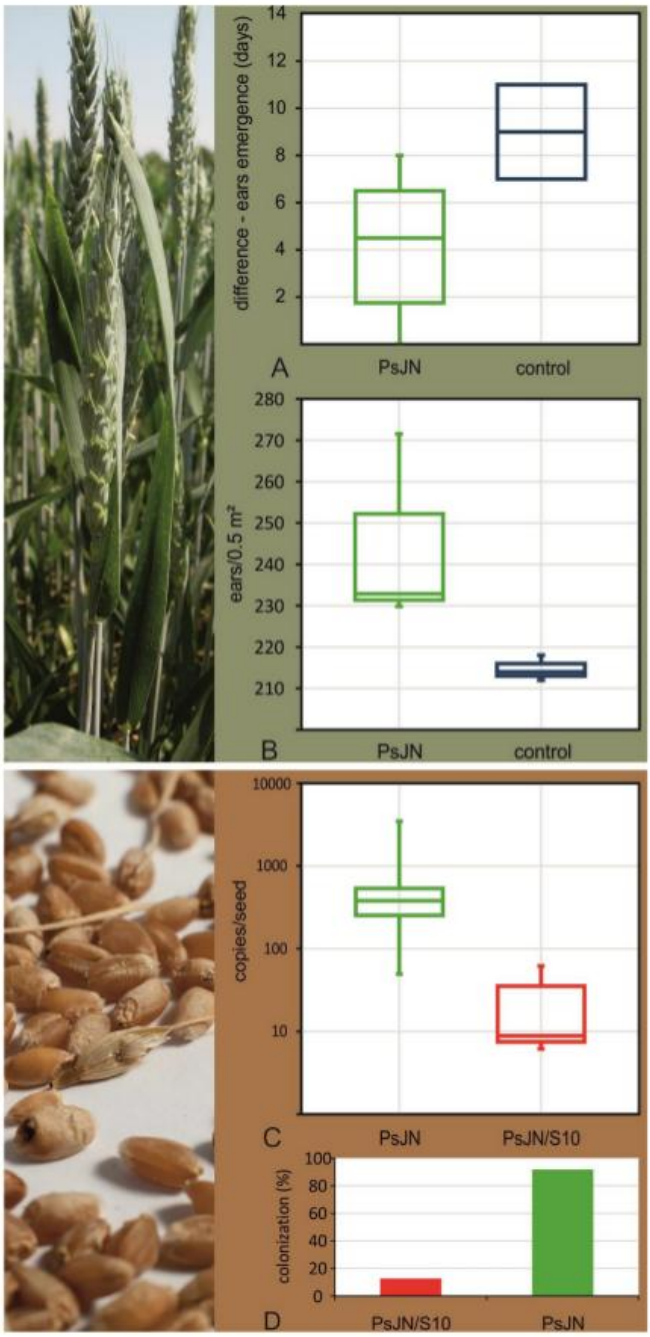


Figure 3: Differences in ear emergence times in wheat plants (*Triticum aestivum* cv. Trappe) growing from seeds colonized by *B. phytofirmans* PsJN and control seeds (a) observed in greenhouse pot experiments and in the field (b). a: the significant delay in flower formation in control plants is shown relative to the PsJN-seed offspring plants. b: a significantly higher number of ears per square meter in the plots was observed for PsJN-plants as compared to control plants. All wheat plants tested belonged to the F1 generation, derived from parent plants which were sprayed with a suspension of *B. phytofirmans* PsJN or sterile buffer (control). c-d: Efficiency of introducing *B. pyhtofirmans* PsJN into wheat (*Triticum aestivum* cv. Trappe) seed determined by strain specific *Taqman*-qPCR. c: number of PsJN genome equivalents detected in single wheat seeds and d: percentage of colonized wheat seeds (n=24).

One of the main purposes of modulating seed microbiomes is to achieve growth modulation of the resulting plant. Therefore, we compared the growth and development of wheat plants growing from seeds internally colonized by *B. phytofirmans* PsJN in pot experiments in the greenhouse as well as in the field with plants growing from non-colonized control seeds. Seeds were stored for two and seven months at room temperature before being planted in the greenhouse and field, respectively. In our greenhouse experiments plants emerging from PsJN-colonized seeds showed significant alterations in spike onset, which started an average of five days earlier in PsJN-plants than in plants emerging from control seeds (Fig. 3a). Similar effects were observed in the field. At a given time point, the number of ears per square meter in the field was significantly higher in plants emerging from PsJN-seeds as compared to the control plants owing to the earlier flowering (Fig. 3b). Further, whatever the variation in numbers of PsJN in the planted seeds, the effects on the crop were relatively uniform and essentially non-overlapping with the controls. These effects on flowering were not unexpected as it is known that *B. phytofirmans* PsJN speeds up maturity in many of its host plants and an earlier start in flower formation is often observed in PsJN inoculated plants[21,22]. In *Arabidopsis thaliana* alterations in anthesis correlated with an earlier induction of flowering control genes in PsJN-inoculated plants as compared to control plants[22].

Colonization of offspring plants by seed-born PsJN was tested by strain-specific qPCR. In field grown wheat plants we detected PsJN in root and shoot tissue with an average of 269 and 388 genome equivalents per gram plant tissue. Passage of strain PsJN from colonized seeds to the next generation of seeds was tested for pepper, soy and wheat but PsJN was not found in any of the seeds.

For a comprehensive assessment of the effects of incorporating selected bacteria into wheat seed on the bacteria microbiome we performed culture-independent community analysis of single seeds by Illumina 16S rRNA gene-amplicon sequencing. Samples of all three treatments - control seeds (plants were sprayed with sterile buffer), PsJN-seeds (plants were sprayed with strain PsJN) and PsJN+S10-seeds (plants were sprayed with a mixture of PsJN and S10) were used for sequencing in nine replicates (sequencing statistics are given in Extended Data Table 1).

OTU abundance data were clustered at each major taxonomic level, that is, phylum, class and order (Extended Data Fig. 3). The seed bacterial communities were dominated by *Proteobacteria*, which made up 92%, 90% and 98% of the OTUs in the control seeds, PsJN- seeds and PsJN+S10-seeds, respectively. While the species richness and diversity (□-diversity, Fig. 4a) were not affected by the introduction of strain PsJN or the mixture of PsJN and S10 into spring wheat seeds, the sequencing results showed a clear effect on the community structure (β-diversity, Fig. 4b). The three treatments differed mainly in the abundance of certain groups (Fig. 4c).

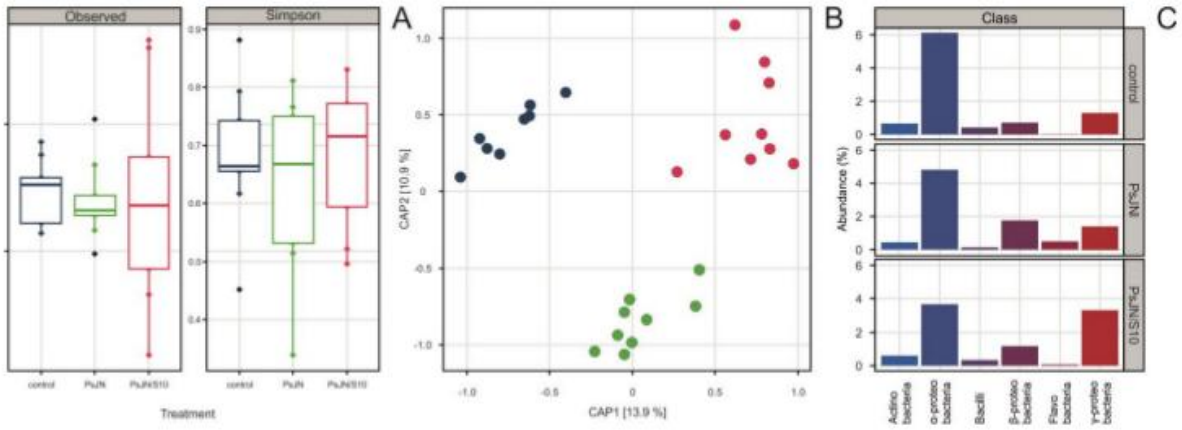


Figure 4: Seed endophyte community profiling based on 16S rRNA gene V5-V7 sequences. (**a**) Alpha diversity within subject by treatment, as measured counting the observed OTU richness (Observed) and calculating the Simpson's diversity index (Simpson). A permutation ANOVA (RVAideMemoire R package) showed (p > 0.05, perm = 9999) that neither the richness nor the diversity values were significantly different when grouped by treatment. (**b**) Bray–Curtis beta diversity among subjects as depicted by a Constrained Analysis of Principal Coordinates (CAP) constrained to the treatments. Control samples are shown in blue, PsJN treated samples in green and PsJN/S10 treated samples in red (**a** and **b**). A permutation test assessed the significance of the treatments on CAP (p < 0.001, perm = 9999) (vegan R package) and permutation pairwise comparisons between the treatments on the Bray-Curtis dissimilarity matrix confirmed that each group of treated samples is different from all the others (q < 0.001 FDR corrected, perm = 9999) (RVAideMemoire R package). (**c**) Proportional abundance barplot of the most variant OTUs by treatment summarized at class level. The OTUs were determined after differential OTU abundance analysis via permutation ANOVA (q < 0.001 FDR corrected, perm = 9999) (RVAideMemoire R package) and plotted using the phyloseq R package.

Besides the expected increase in □-*Proteobacteria* by the introduction of strain PsJN (4% in control seeds, 39% in PsJN-seeds and 9% in PsJN+S10-seeds), the *Flavobacteria* and □-*Proteobacteria* were enriched in PsJN- and PsJN+S10-seeds, respectively whereas OTUs belonging to □-*Proteobacteria* decreased upon

introduction of PsJN and PsJN+S10 (Fig. 4c). The *Flavobacteria* constituted less than 0.4% in control seeds but were enriched to 6% in PsJN-seeds and the □-*Proteobacteria* were enriched from 21% in control seeds to 80% in PsJN+S10-seeds. In contrast, instead of making up 67% of the bacteria in control seeds, the □-*Proteobacteria* share was reduced to 39% in PsJN-seeds and 9% in PsJN+S10-seeds (Fig. 4c). Due to these differences, we postulate the existence of interactions between bacteria taxa present in the three different seed types (Extended Data Fig. 4). This analysis confirmed a negative correlation between the occurrence and abundance of □- and □-*Proteobacteria* within the seed samples. A positive correlation was found between OTUs most probably representing PsJN and the occurrence and abundance of *Variovorax*- (OTU2281), *Chryseobacterium*- (OTU3949), Paracoccus- (OTU2798) and *Sphingomonas*- (OTU79) specific OTUs. A clear negative correlation was also found between PsJN-OTUs and the occurrence and abundance of a sequence (OTU52) most probably representing *Enhydrobacter* sp. We assume that all these modifications to seed microbiome composition were determined in the parent plant in and around the tissues that give rise to gametes and seeds during grain filling. The approach clearly opens up ways of discovering links between inherited microbiome constituents and plant traits. Such effects can be via the added microbe alone or via the observed community effects.

The use of microbial inoculants in crop production and protection is a rapidly growing area in agricultural technology. However, to realize large-scale implementation of microbial strains in agricultural practice there needs to be successful delivery of beneficial microbes into the plant at a large scale. Such strategies are largely missing (especially for gram-negative bacteria) and this constitutes a bottle neck in application. Addition of microbes to seeds at time of planting is possible but this study breaks new ground both by targeting plant reproductive organs as entry ports and, importantly, using the plant seed as a protective carrier with a long shelf life for microbial inoculants. It also achieves a restructuring of the seed microbiome before sowing as opposed to during germination. We assume that the ability of bacteria to enter plant tissues, propagate and join the other microbes destined to populate seeds is responsible for the effectiveness of the seed modification described here for the first time.

Apart from application in agriculture the approach presented here opens up new avenues in seed endophyte research by enabling studies with single strains and their effects on plant traits. This will in future facilitate the answering of questions on the roles and fate of seed endophytes during seed formation, dormancy and germination as well as the role of seed endophytes in the development of plant endophyte communities. Moreover, possible multi- generational heritability of seed endophytes – although not found in our experiments – could be studied in detail. The relative ease of introducing bacteria into plant seed by applying them on flowers of parent plants indicates that at least a part of the seed microbiome may routinely derive from flower or pollen colonizing microorganisms[23] and the air or insects visiting the plant during flowering[24] might be important sources for seed endophytes. This aspect has been hardly tackled in studies on the ecology of seed endophytes so far and merits greater attention.

Acknowledgements

The authors thank Theresa Ringwald and Marija Karakaš for technical support during the greenhouse and field experiments. The work was supported in part by a grant of the Austrian Science Foundation (FWF P22867-B16) provided to B.M.

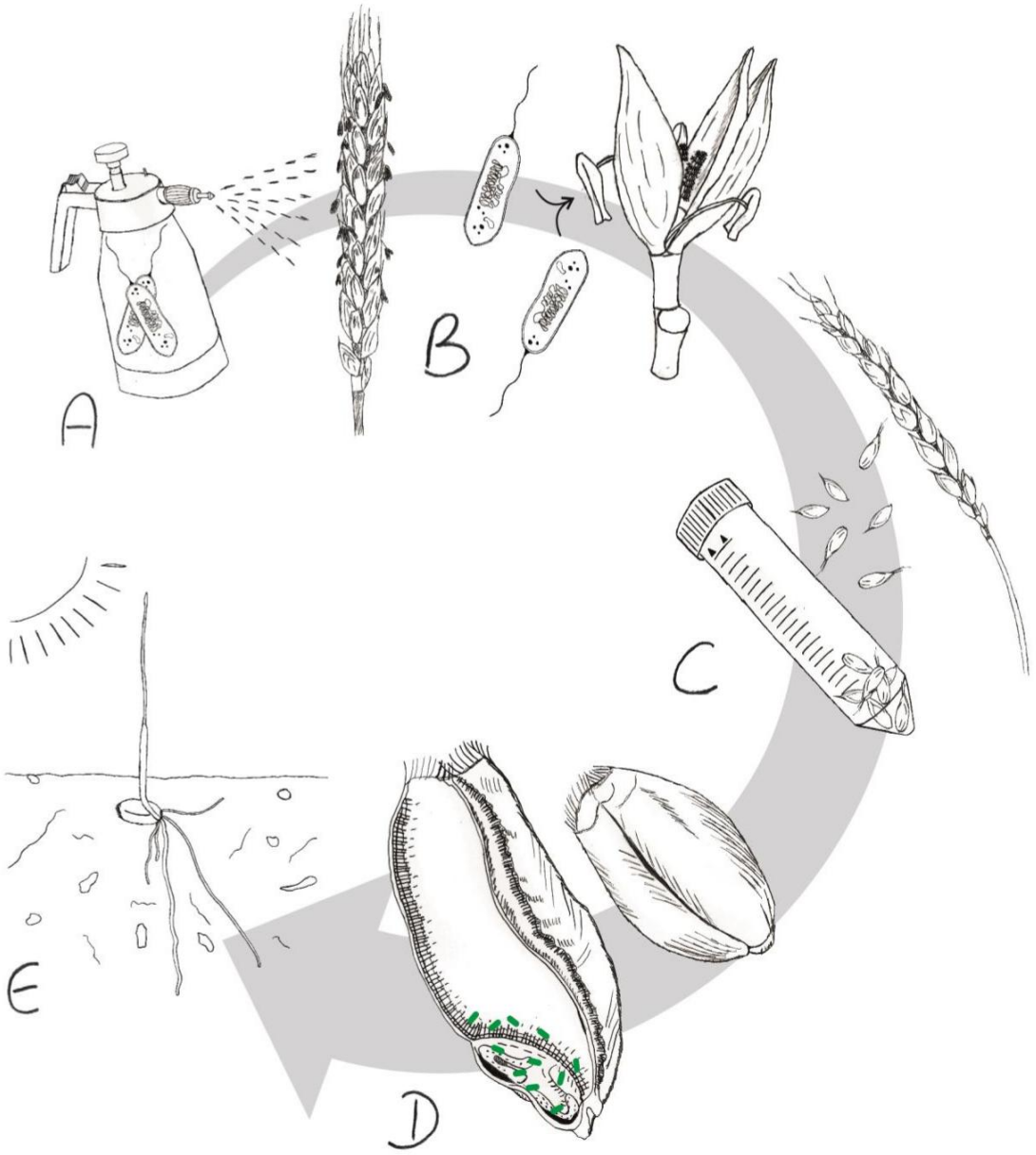
Author Contributions

B.M. and A.S. conceived and designed the research, N.P., L.A. and B.M. analyzed data, N.P., performed DNA isolation, PCR and qPCR, plant experiments and designed figures, S.C. and T.B. designed FISH probe and performed FISH experiments, A.P. performed Illumina amplicon sequencing, L.A. analyzed sequencing data, R.S.-T. developed the qPCR system, B.M., R.F. and G.v.M. wrote the paper. All authors read and approved the final manuscript.

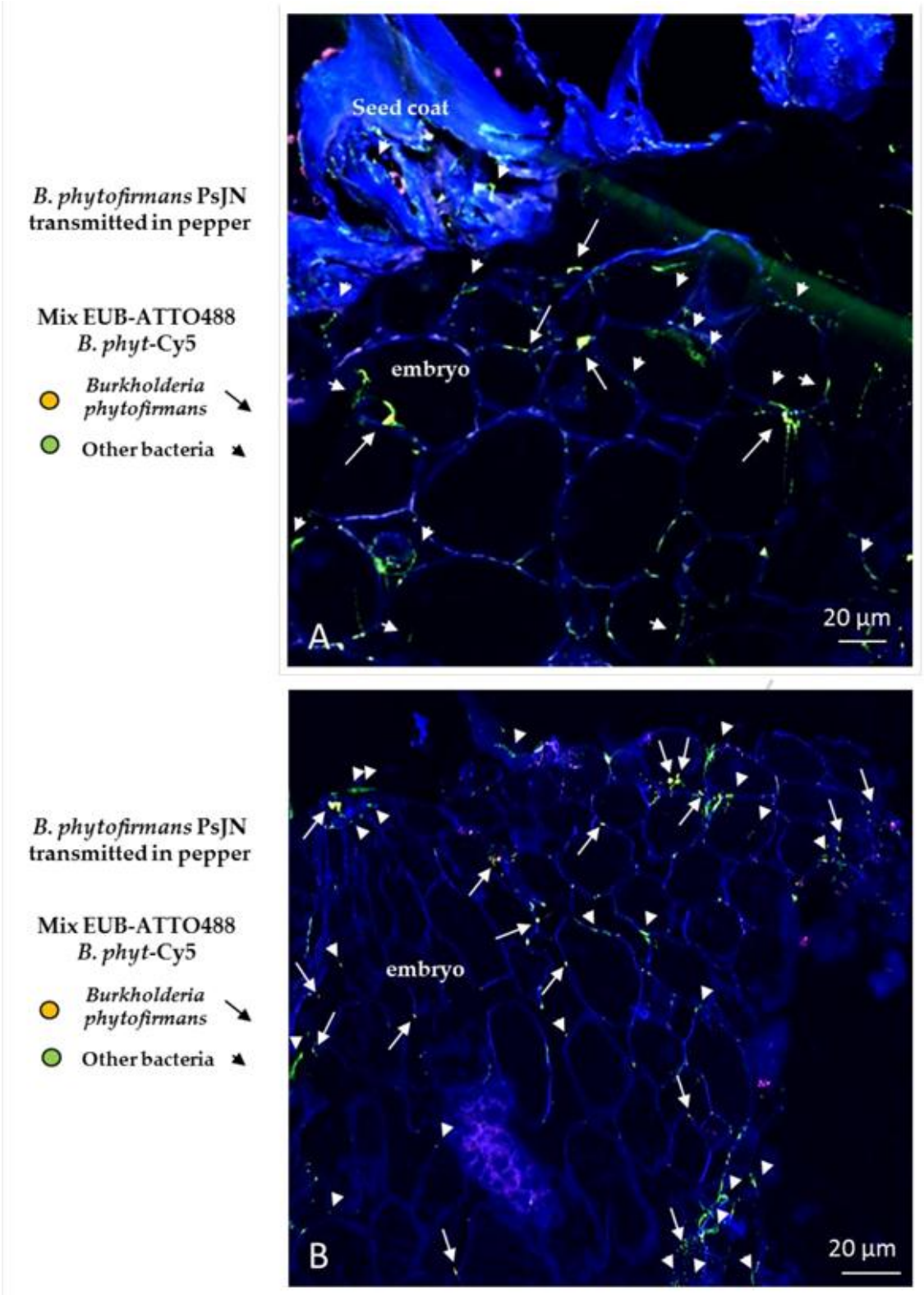
Author Information

Community sequencing data are made available at NCBI SRA database under the accession SRP067570 and the BioProject number PRJNA305879. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to birgit.mitter@ait.ac.at.

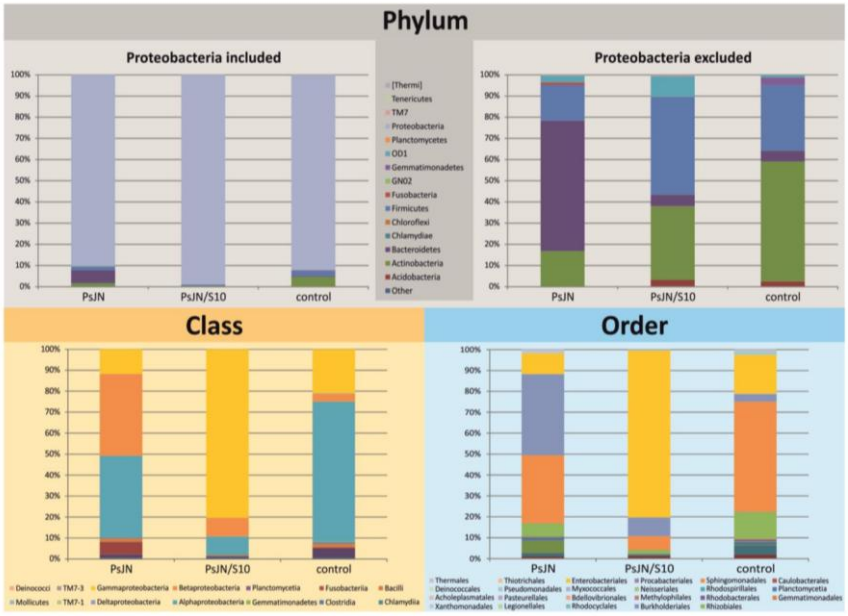
Extended Data Figures



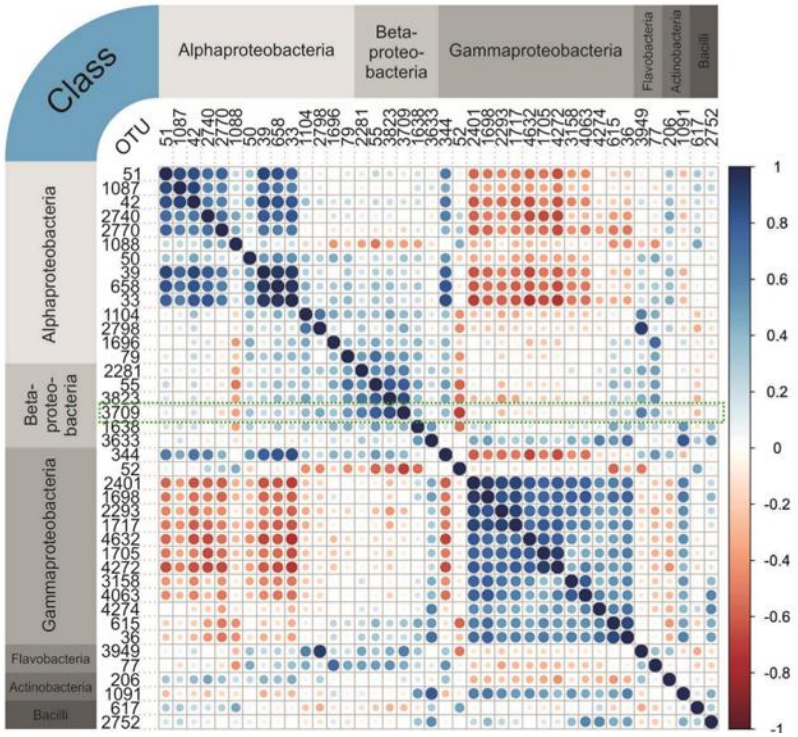
Extended Data Fig. 1: Illustration of the method to introduce plant beneficial bacteria into plant seed. (**A**) Plant flowers are sprayed with a bacterial suspension; (**B**) the bacteria colonize flowers and the developing fruit; (**C**) mature seeds are collected and (**D**) endophytes stay viable during seed storage; (**E**) endophytes proliferate during germination and colonize the offspring plant generation.



Extended Data Fig. 2: Visualization of *B. phytofirmans* PsJN in seeds of *Capsicum annuum* L. (pepper) by DOPE-FISH/CSLM microscopy showing the presence of *B. phytofirmans* (yellow) inside the embryo together with other bacteria (green) (**a, b**).



Extended Data Fig. 3: OTU abundance data clustered at the following taxonomic level, (**a**) phylum, (**b**) phylum with *Proteobacteria*-specific OTUs being excluded (to give more details on the less abundant bacteria groups), (**c**) class and (**d**) order.



Extended Data Fig. 4: Pearson linear correlation of OTUs selected after differential abundance analysis across the treatments (q < 0.05 FDR corrected, perm = 9999). Correlation matrix patterns were reordered using a hierarchical clustering and displayed by a corrplot (corrplot R package) from dark blue (positive correlation) to deep red (negative correlation), fading color and size out when getting not correlated. Correlations of OTU 3709, representing *B. phytofirmans* PsJN are framed in green.

Extended Data Table 1 16S rRNA gene amplicon sequencing statistics

	control seeds	PsJN-seeds	PsJN+S10-seeds
total read count	1,089,120.00		
average read counts	11,203.33 ± 7,752.89	32,191 ± 20,791.57	77,619 ± 48,485.75
average read length	367.18 ± 31.55 bp		
average OTUs	89.44 ± 12.53	127.22 ± 37.18	172.56 ± 68.05
Alpha-diversity			
observed OTUs	74.2 ± 12.30	65.58 ± 15.70	69.54 ± 41.85
Simpson's diversity index	0.69 ± 0.12	0.64 ± 0.15	0.69 ± 0.12

The average read counts, average read length, average number of OTUs and alpha-diversity characteristics are shown as mean ± s.d. of nine replicates per treatment (control, seeds emerging from plants sprayed with PsJN only or a mixture of PsJN and S10.

1. Hardoim, P. R. *et al.* The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol. Mol. Biol. Rev.* **79**, 293-320 (2015).
2. Turner, T. R., James, E. K. & Poole, P. S. The plant microbiome. *Genome Biology* **14**, 209 (2013).
3. The Human Microbiome Project Consortium . Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
4. Turnbaugh, P. J. *et al*. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-1031 (2006).
5. Turnbaugh, P. J. *et al*. A core gut microbiome in obese and lean twins. *Nature* **457**, 480-484 (2009).
6. Ding1, T. & Schloss, P. D. Dynamics and associations of microbial community types across the human body. *Nature* **509**, 357–360 (2014).
7. Qin, N. *et al.* Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59–64 (2014).
8. Zhang, D. *et al.* Neutrophil ageing is regulated by the microbiome. *Nature* **25**, 528–532 (2015).
9. Barret, M. *et al*. Emergence shapes the structure of the seed microbiota. *Appl. Environ. Microbiol.* **81**, 1257–1266 (2015).
10. Klaedtke S. *et al.* Terroir is a key driver of seed-associated microbial assemblages. *Environ. Microbiol.* DOI:10.1111/1462-2920.12977 (2015).
11. Truyens, S., Weyens, N., Cuypers, A. & Vangronsveld, J. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environ. Microbiol. Reports* **7**, 40-50 (2014).
12. Puente, M. E., Li, C. Y. & Bashan, Y. Endophytic bacteria in cacti seeds can improve the development of cactus seedlings. *Environ. Exp. Bot.* **66,** 402–408 (2009).
13. Johnston-Monje, D. & Raizada, M. N. Conservation and diversity of seed associated endophytes in Zea across boundaries of evolution, ethnography and ecology. *PLoS ONE* **6**, e20396 (2011).
14. Hardoim, P. R., Hardoim, C. C. P., van Overbeek, L. S. & van Elsas, J. D. Dynamics of seed-borne rice endophytes on early plant growth stages. *PLoS ONE* **7**, e30438 (2012).

15. Bouffaud, M.-L., Poirier, M.-A., Muller, D. & Moënne-Loccoz, Y. Root microbiome relates to plant host evolution in maize and other *Poaceae*. *Environ. Microbiol.* **16**, 2804–2814 (2014).
16. Delaux, P.-M. *et al*. Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. *PLoS Genet.* **10**, e1004487 (2014).
17. Sessitsch, A. & Mitter, B. 21st century agriculture: integration of plant microbiomes for improved crop production and food security. *Microb. Biotech.* **8**, 32-33 (2014).
18. Sessitsch, A. *et al. Burkholderia phytofirmans* sp. nov., a novel plant-associated bacterium with plant beneficial properties. *Int J. Syst. Evol. Bacteriol*. **55**, 1187-1192 (2005).
19. Compant, S., Nowak, J., Coenye, T., Clément, C. & Ait Barka, E. Diversity and occurrence of *Burkholderia* spp. in the natural environment. *FEMS Microbiol. Rev.* **32**, 607-626 (2008).
20. Compant, S. *et al.* Endophytic colonization of *Vitis vinifera* L. by plant growth promoting bacterium *Burkholderia* sp. strain PsJN. *Appl. Environ. Microbiol.* **71**, 1685-1693 (2005).
21. Wanga, B., Seilera, J. R. & Meib, C. *Burkholderia phytofirmans* strain PsJN advanced development and altered leaf level physiology of switchgrass. *Biomass & Bioenergy* **83**, 493–500 (2015).
22. Poupin M. J. *et al.* Effects of the plant growth-promoting bacterium *Burkholderia phytofirmans* PsJN throughout the life cycle of *Arabidopsis thaliana*. *PLoS ONE* **8**, e69435 (2013).
23. Aleklett, K., Hart, M. & Shade, A. The microbial ecology of flowers: an emerging frontier in phyllosphere research. *Botany* **92**, 1–14 (2014).
24. Ushio, M. *et al.* Microbial communities on flower surfaces act as signatures as pollinator visitation. *Scientific Reports* 5, 8695 (2015).

METHODS

Bacterial strain and inoculum preparation. In this study *Burkholderia phytofirmans* PsJN (=LMG 22146T)[18] and variants of PsJN chromosomally tagged with the beta-glucuronidase gene[20] were used. The bacterial strain was grown by loop-inoculating one single colony in LB broth (PsJN wild-type) and LB broth amended with spectinomycin [100 µg mL^{-1}] (PsJN::*gus*A110). Bacterial cultures were incubated at 28±2^0C for two days at 180 rpm in a shaking incubator. Cells were harvested by centrifugation and the bacterial pellet re-suspended by vortexing in 20 mL sterile PBS (0.2 g/L KCl, 1.44 g/L Na2HPO4 and 0.24 g /L KH2PO4, in dH2O, pH 7.4). The concentration of the suspensions was measured with a NanoDrop ND-1000 spectrophotometer (Wilmington, DE, USA) and adjusted to 3x10^8 CFU/mL.

Introducing *B. phytofirmans* PsJN into seeds of maize, pepper and soy. Seeds of pepper (*Capsicum anuum* cultivar Feher) and soybean (*Glycine max* L. cultivars Merlin and Essor) were sown into potting soil in plastic trays kept in a greenhouse chamber. Ten days after sowing seedlings were individually potted into 1L (soy) and 3L (pepper) pots containing potting soil. Soy plants were watered automatically twice a week by flooding for 10 min and fertilized once with liquid fertilizer suspension Wuxal Super 3% (Aglukon) (NPK + trace elements). Pepper plants were watered daily and fertilized every 4 weeks with Wuxal Super 0.5 %. The frequency of fertilizer application was increased to a weekly treatment upon fruit set of the pepper plants. Maize husbandry was done as described elsewhere[25]. Specific inoculation of flowers was conducted when the plants reached growth stage 61 – 63 on the BBCH scale (for pepper and soy: first flower open – third flower open; for maize: flowering, anthesis)[26]. A suspension of *B. phytofirmans* PsJN and its variant *B. phytofirmans* PsJN::*gusA*, respectively and buffer only for the control were added to pump spray bottles previously sterilized with 70% ethanol. Flowers were sprayed and a filter paper was used to shield the surrounding plant parts such as leaves and stem from drift liquid and to prevent surplus inoculum dripping on the soil. The treated inflorescences/flowers were marked with a twist tie to allow for later identification. The inoculum was prepared as described above.

Introducing *B. phytofirmans* PsJN into spring wheat seeds under field conditions. The production of seeds internally colonized by *B. phytofirmans* PsJN under field conditions was tested with *Triticum aestivum* L (cultivar Trappe). Ten by

1.3 m plots were planted on March 13, 2014 with spring wheat at a density of 180 kg/ha in a field located in Tulln, Austria. Plants were sprayed with herbicide once (1.25 l/ha Andiamo Maxx) and fertilized twice. NPK-Fertilizer 16:6:18+5S was applied at a concentration of 300 kg/ha and N-Fertilizer 27% was applied at a concentration of 220 kg/ha. At flowering time, each plot was sprayed twice (one week interval) with a suspension of *B. phytofirmans* PsJN or a mixture of *B. phytofirmans* PsJN and *Paenibacillus* sp. S10, respectively. The bacterial inoculant used for spraying was prepared as follows: endophytes were streaked on large (diameter: 14.5 cm) 20% tryptic soy agar plates, grown at 28°C for 2 days, scraped from the plates and suspended in 2L of 1x PBS supplemented with 20g zeolite (used as a carrier) and 200µL Silwet L-77 (final OD_{600} of about 0.1). Suspensions were added to spraying bottles and inflorescences in each plot were sprayed as uniformly as possible with 1L of the corresponding treatment. Negative control plots were sprayed with 1x PBS containing zeolite and Silwet L-77.

Detection and quantification of *B. phytofirmans* PsJN in plant seed tissue by GUS-staining and viable cell counting. Endophytic colonization of root, stem and leaves of maize plants by the *gus*A-labeled variant of *B. phytofirmans* PsJN was determined by plate counting and colonies were identified by comparison of the 16S-23S rRNA intergenic spacer region DNA fragment pattern with pure culture *B. phytofirmans* PsJN as described elsewhere[25]. Gus-staining of plant tissue was performed as following: The plant material was cut with a sterile scalpel and subsequently incubated in GUS-staining solution (1mM EDTA, 5mM potassium ferricyanide, 5mM potassium ferrocyanide, 100mM sodium phosphate, pH 7.0, 1% Triton-X-100, 0.1 mg/mL X-Gluc pre-dissolved in 5µL/mg N,N-dimethylformamide, 0.1% IPTG) directly after harvesting at 37°C for 20 hours. Afterwards, destaining was done by rinsing the samples with 70% ethanol. The ethanol was then discarded and the samples fixed in paraformaldehyde solution (4% paraformaldehyde dissolved in PBS at 60°C with constant stirring until clarifying of the solution) overnight at 4°C. Finally, the fixed samples were rinsed 3 times in PBS and stored in the last rinse at 4°C until further processing.

Detection of PsJN in seeds and green parts of plants using DOPE-FISH. For microscopy analysis, plant samples were cut into 0.5-cm long sections. Samples were then fixed overnight at 4°C in a paraformaldehyde solution (4% in PBS, pH 7.2),

and rinsed twice in PBS. Treatment with a lysozyme solution (1 mg mL-1 in PBS) was then applied to the samples for 10 min at 37°C before being dehydrated in an ethanol series (25, 50, 75 and 99.9%; 15 min each step). Fluorescence in situ hybridization using double labeling of oligonucleotide probes (DOPE-FISH) was carried out using probes from Eurofins (Germany) labeled at both the 5' and 3' positions. A probe mixture targeting eubacteria, EUBmix (equivalent mixture of EUB338, EUB338II, EUB338III) coupled with a ATTO488 fluorochrome27,28, and a probe for B. phytofirmans coupled with Cy5 were used. NONEUB probe29, coupled with Cy5 or ATTO488 was used independently as a negative control.

B. phytofirmans-specific probes were designed targeting the 16S rRNA and 23S rRNA of *B. phytofirmans* PsJN by making use of the Biosearch Technologies' Stellaris FISH Probe Designer software (www.biosearchtech.com/stellarisdesigner/) and the sequence of chromosome 1 of *B. phytofirmans* PsJN (GenBank project accession CP001053.1). Specificity of suggested probes was tested by alignment to the nucleotide database of NCBI and by using the Probe Match and Evaluation Tool (http://www.arb-silva.de/search/testprobe/) provided by SILVA[30] as well as probeCheck[31]. The nucleotide sequences targeting the 23S rDNA were identified to serve as potential probes hybridizing exclusively with PsJN and were further revised concerning the presence of complementary nucleotide sequences within the stretch of the probe, which can lead to loop formation and is therefore not desirable. The sequences were discarded if they displayed this property. Candidate probes were then analyzed regarding suitable reaction settings (reaction temperature, Na-concentration, probe concentration) for efficient hybridization by help of the web tool mathFISH[32] and tested on pure culture of *B. phytofirmans* PsJN at different temperatures of hybridization and formamide concentrations. The 23S rDNA probe 5´-CTCTCCTACCATGCACATAAA3´ was selected for further experiments. Fluorescence *in situ* hybridization was carried out at 46°C for 2 h with 10–20 µL solution (containing 20 mM Tris–HCl pH 8.0, 0.01% w/v SDS, 0.9 M NaCl, 10% formamide, and 10 ng μL^{-1} of each probe) applied to each plant sample placed on slides in a 50-mL moist chamber (also housing a piece of tissue imbibed with 5 mL hybridization buffer). Washing was conducted at 48°C for 30 min with a post-FISH pre-warmed solution containing 20 mM Tris–HCl pH 8.0, 0.01% (w/v) SDS, and NaCl at a concentration corresponding to the formamide concentration. Samples were then rinsed with distilled water before air drying for at least 1 day in the dark.

The samples were then observed under a confocal microscope (Olympus Fluoview FV1000 with multiline laser FV5-LAMAR-2 HeNe(G)laser FV10-LAHEG230-2). X, Y, Z pictures were taken at 405, 488, 633 nm and then merged (RGB) using Image J software. Z Project Stacks was then used to create pictures as described elswhere[33].

DNA isolation. Plant material was surface-sterilized as described earlier[25]. Single surface-sterilized seeds were aseptically peeled using a scalpel, cut in pieces and crushed using a sterile mortar. Vegetative plant material was cut in pieces. All types of plant material were homogenized for 30s in lysing matrix E (MPbio DNA isolation kit from soil) using a bead beater (FastPrep FP 120, Bio101, Savant Instruments, Inc., Holbrook, NY). DNA was then extracted with the MPbio DNA isolation kit from soil (MP Biomedicals, Solon, OH, USA) according to protocol provided by the manufacturer. DNA (5 µl) was separated and visually tested for quality by electrophoresis (80 V) on 1% (w/v) agarose gels stained with ethidium bromide. DNA concentration was measured using a NanoDrop ND-1000 spectrophotometer.

Quantification of B. phytofirmans PsJN in plant tissue using qPCR. Quantification of B. phytofirmans PsJN in seeds and vegetative plant tissue was performed by qPCR using a Taqman probe and a Biorad CFX96 real-time detection system (Bio-Rad, Hercules, CA, USA). The probe and primers were designed in a previous study34 to match the gene for transcription termination factor rho (Bphyt_1824) in the genome of strain PsJN. qPCR reactions contained (10µl total volume): 1x SsoFast Probes, 0,5 µM of each primer, 0,35 µM probe and 5 - 100 ng DNA. The qPCR was run at the following settings: hot start at 95°C for 2 min, 40 cycle denaturation at 95°C for 5 sec and hybridization and elongation for 20 sec at 59 °C. For qPCR standard preparation, chromosomal DNA of B. phytofirmans PsJN was isolated as described above. DNA concentration was determined using a NanoDrop ND-1000 spectrophotometer and doing five replicate measurements. The mean value was used for further calculations. The number of DNA copies was calculated as follows:

$$numberofcopies = \frac{DNAquantity\left(\frac{g}{\mu l}\right)}{fragmentlength * 660\ g/mol} * 6{,}022 * 10^{23}$$

Fragment length is 8214658 bp (size of PsJN genome)[35]. A dilution series was prepared to generate a standard curve. Unknown starting quantity of DNA copy numbers in the samples could be calculated based on the standard curve from the dilution series of known concentrations, which produced an r^2 value of 0.990. All data analysis was performed with Bio-Rad CFX Manager 3.0.

Testing of the effects of *B. phytofirmans* PsJN incorporated in seeds on the development of offspring plants. (1) Greenhouse experiments with spring wheat seeds: Plant nursery was done as described above for pepper and soy. On day 17 after seed sowing, 6 plants per treatment were potted individually in pots with a diameter of 15 cm, containing potting soil. Plant height was measured once a week and from day 48 onwards tillering was also counted. The appearance of the first spike per plant was documented till day 73. (2) Field testing of spring wheat seeds: The performance of seeds internally colonized by *B. phytofirmans* PsJN under field conditions was tested with *Triticum aestivum* L. (cultivar Trappe). Plots were planted on March 18, 2014 and field management was done as described above. Regular ratings of germination, plant height, tillering and spike counting were performed and plant colonization by strain PsJN was tested by qPCR as described above.

Analysis of microbial communities of spring wheat endoseed prepared in the field. Genomic DNA was isolated using FastDNA SPIN Kit for soil as described above and concentration was adjusted to 5ng/µl. A nested PCR approach was used to amplify bacterial 16S rDNA. The first amplification was performed with primers 799for (5´-AACMGGATTAGATACCCKG-3´) and 1392rev (5´-ACGGGCGGTGTGTRC-3´)[35] with the following reaction parameters: 25 µl reaction volume contained 200 nM of each primer, 300 µM dNTPs, 0.5 units KAPA HiFi DNA polymerase (Kapa Biosystems, Boston, MA, USA), 1x buffer, 5 ng template DNA. The amplification conditions were as follows: initial denaturation for 5 min at 95°C, 25x 30 sec at 95°C, 30 sec at 52°C and 30 sec at 72°C, and a final elongation for 5 min at 72°C. PCR amplification was performed in a peqSTAR thermocycler (peQlab, Erlangen, Germany). Amplicons were subjected to electrophoresis (100V for 1 h) in 2% (w/v) TBE agarose gels (Biozym Biotech Trading, Vienna, Austria). Amplification with the primer pair 799F and 1392R allows exclusion of the chloroplast 16S rDNA and results in co-amplification of bacterial and mitochondrial ribosomal genes, 600 bp and 1000 bp amplicon size respectively. The band containing the PCR-product of bacterial 16S rDNA was excised. The gel pieces were put in a filter tip that was placed in a fresh tube and DNA was collected by centrifugation for 2 min at 1000

rpm. The second amplification was performed with the primers 799 for_illumina and 1175R1_illumina, using amplification reaction procedures as described above. PCR amplicons were subjected to electrophoresis and the 500bp bands were excised and DNA collected as described above. Index PCR was performed with Nextera XT Index Kit (Illumina Inc., San Diego, USA) according to the manufacturer´s protocol and resulting amplicons were purified using AMPure XP beads (New England Biolabs, Ipswich, MA, USA) following the manufacturer's protocol. Amplicon concentration was measured using a Nanodrop and about 10 ng per sample were pooled. DNA quality and quantity of the pooled library were tested with an Agilent 2100 Bioanalyzer. The final amplicon size was about 570 bp including the adapter, sequencing primer binding site and index on both sides. The library denaturing, addition of internal control DNA (PhiX, Illumina) and sample loading were done according to the Illumina protocol. Sequencing was performed on a MiSeq desktop sequencer (Illumina Inc., San Diego, USA).

16S rRNA gene sequencing processing. MiSeq raw data quality was checked in FASTQC[36] and reads were screened for PhiX contamination using Bowtie 2.2.6[37]. A Bayesian clustering for error correction[38,39] was applied before merging the PE reads using PEAR 0.9.6[40] ($p<0.001$). Forward and reverse primers were then stripped from merged reads employing Cutadapt 1.8.3[41] and quality filtering performed in USEARCH v8.0.1517[42,43] (maximum expected error=0.5). Filtered reads were labelled according to the sample name of origin and combined in QIIME[44]. Sequences were dereplicated, sorted and clustered at 97% of similarity using VSEARCH 1.1.1[45]. Chimeras were checked adopting both a de-novo and a reference based approach, as routine of the above mentioned tool. The RDP classifier training set v15 (09/2015) was used as a reference database. METAXA2[46] was used to target the extraction and to verify the 16S V7-V9 region of the representative sequences. An optimal global alignment was applied afterwards in VSEARCH and a BIOM table generated. Taxonomy assignment was performed employing the naïve Bayesian RDP classifier[47] with a minimum confidence of 0.6 and a customized version of the Greengenes database[48] (08/2013), including the PsJN and S10 strain sequences and taxonomy. **16S rRNA gene-based microbial community analysis and statistics.** An OTU-based analysis was performed in QIIME to calculate the richness and diversity after multiple rarefaction. The observed OTUs were counted and the diversity within each individual sample was estimated using the Simpson's diversity

index. Richness and diversity values were compared between the control and the treatments by means of permutation ANOVA and permutational pairwise comparisons in the RVAideMemoire R package[49]. The resulting P values were adjusted by False Discovery Rate (FDR). Richness and diversity value boxplots were then plotted via ggplot2[50] package in R.

A data-driven adaptive method for selecting normalization scale quantile was conducted on the BIOM table and data normalized by scaling counts by the nth percentile of each sample's nonzero count distribution in the metagenomeSeq Bioconductor package[51,52]. The resulting normalized BIOM table was used for the beta-diversity analysis. Multivariate analysis of community structure and diversity was performed according to the recommendations by Anderson and Willis[53]: 1) unconstrained ordination offered by Principal Coordinate Analysis (PCoA), 2) constrained multidimensional scaling using Constrained Analysis of Principal Coordinates (CAP) as re-implemented in the vegan R package[54], 3) permutation test for assessing the significance of the constraints and permutational multivariate analysis of variance (PERMANOVA), and 4) individuation and correlation of OTUs responsible for shaping the diversity structure.

In more details, the differences between bacterial communities were investigated using the Bray–Curtis dissimilarity distance and the ordination methods applied to the same distance matrices. All the ordination analyses were computed and CAP plotted in phyloseq[55] (points 1 and 2). The significance of the treatment grouping factor used as constraint in the CAP was assessed via the permutation test[56] in the vegan R package. The null hypothesis of no differences between *a priori* defined groups was investigated recurring to the PERMANOVA approach[57], implemented in vegan as the ADONIS function and applied to the Bray–Curtis dissimilarity distances.

Permutational pairwise comparisons between the treatments were carried out in the RVAideMemoire R package and P values adjusted by FDR (point 3). A permutation ANOVA was applied for differential OTU abundance calculation among the treatments and P values corrected by FDR. Significantly different OTUs were further processed via Pearson's correlation[58] and plotted using the corrplot R package[59] (point 4). Quantitative differences of the PsJN OTUs across the treatments were plotted as boxplots in the microbiome R package[60].

25. Naveed, M. *et al.* Increased drought stress resilience of maize through endophytic colonization by *Burkholderia phytofirmans* PsJN and *Enterobacter* sp. FD17. *Environ. Exp. Bot.* **97**, 30-39 (2014).
26. Bleiholder, H. *et al.* In: Meier U. (ed.) Growth stages of mono- and dicotyledonous plants. Federal Biological Research Center for Agriculture and Forestry, 2nd edition (2001).
27. Amann, R. *In situ* identification of microorganisms by whole cell hybridization with rRNA targeted nucleic acid probes. *Meth. Mol. Ecol. Manual* **3.3.6**, 1 – 15 (1995).
28. Daims, H. *et al.* Fluorescence in situ hybridization for the detection of prokaryotes. In: Osborn, A.M., Smith, C.J. (ed.): *Molecular Microbial Ecology*. New York, Taylor & Francis (2005).
29. Wallner, G. *et al.* Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**, 136–1434 (1993).
30. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590 – 596 (2012).
31. Loy, A., Maixner, F., Wagner, M., & Horn, M. probeBase – an online resource for rRNA-targeted oligonucleotide probes: new features 2007. *Nucleic Acid Res.* **35**, 800 – 804 (2007).
32. Yilmaz, L. S., Okten, H. E., Noguera, D. R. MathFish, a web tool that uses thermodynamics-based mathematical models for *in silico* evaluation of oligonucleotide probes for fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **77**, 1118 – 1122 (2011).
33. Campisano, A. *et al.* Interkingdom transfer of the acne-causing agent, *Propionibacterium acnes*, from human to grapevine. *Mol. Biol. Evol.* **31**, 1059-1065 (2014).
34. Sheibani-Tezerji, R. *et al.* Transcriptome profiling of the endophyte *Burkholderia phytofirmans* PsJN indicates sensing of the plant environment and drought stress. *mBio* **6**, e00621-15 (2015).
35. Weilharter, A. et al. Complete genome sequence of the plant-growth promoting endophyte *Burkholderia phytofirmans* strain PsJN. *J. Bacteriol.* **193**, 3383–3384 (2011).

36. Chelius, M.K., & Triplett, E.W. The diversity of archaea and bacteria in association with the roots of *Zea mays* L. *Microb. Ecol.* **41**, 252–263 (2001).
37. Andrews, S. FastQC: a quality control tool for high throughput sequence data . (2010). at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
38. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
39. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* **14**, 7 (2013).
40. Schirmer, M. *et al.* Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* gku1341 (2015).
41. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).
42. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
43. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996–998 (2013).
44. Edgar, R. C. & Flyvbjerg, H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* btv401 doi:10.1093/bioinformatics/btv401 (2015).
45. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
46. Rognes, T. VSEARCH: versatile open-source tool for metagenomics. (2015). at https://github.com/torognes/vsearch
47. Bengtsson-Palme, J. *et al.* metaxa2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Mol. Ecol. Resour.* **15**, 1403–1414 (2015).
48. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
49. McDonald, D. *et al*. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).

50. Hervé, M. RVAideMemoire: Diverse Basic Statistical and Graphical Functions. (2015) at https://cran.r-project.org/web/packages/RVAideMemoire/index.html
51. Wickham, H. & Chang, W. ggplot2: An Implementation of the Grammar of Graphics. (2015) at https://cran.r-project.org/web/packages/ggplot2/index.html
52. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nature Methods* **10**, 1200–1202 (2013).
53. McMurdie, P. J. & Holmes, S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput. Biol.* **10**, e1003531 (2014).
54. Anderson, M. J. & Willis, T. J. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* **84**, 511–525 (2003).
55. Oksanen, J., *et al.* vegan: Community Ecology Package. (2015) at https://cran.r-project.org/web/packages/vegan/index.html
56. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).
57. Legendre, L. Numerical ecology. (Elsevier Science, 1998).
58. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **26**, 32–46 (2001).
59. Jr, F. E. H., *et al.* Hmisc: Harrell Miscellaneous. (2015) at https://cran.r-project.org/web/packages/Hmisc/index.html
60. Wei, T. corrplot: Visualization of a correlation matrix. (2013) at https://cran.r-project.org/web/packages/corrplot/index.html
61. Lathi, L. & Saloja, J. microbiome R package. (2014) at <http://microbiome.github.com>

Chapter VI

Conclusions & Future Perspectives

Conclusions and future perspectives

The interaction of endophytic bacteria with their host plant is in the center of attention due to its effect on plant growth promotion and its biocontrol abilities. In this thesis, the fate of endophytic bacterial strains inside the host plant tissue in normal as well as under stressed conditions is studied *in planta* for the first time. I investigated the establishment, genetic potential, molecular and functional mechanisms of inoculated endophytes after entering the host plant tissue. Thereby, I used comparative genomics approaches and transcriptome analysis in combination with extensive experimental methods.

Foremost, I monitored the endophytic strain *B. phytofirmans* PsJN that colonized inside cv. Bionta potato plants (*in vitro*) as they were exposed to drought stress in different time points. Our findings show that the bacterium is metabolically active in plants. Eleven out of eighteen extracytoplasmatic function (ECF) sigma factors are expressed *in planta* transcriptome of *B. phytofirmans* PsJN. In addition, Six ECF sigma factor genes were expressed in *B. phytofirmans* PsJN colonizing drought stressed potato plants. These cell surface signaling elements allow bacteria to sense changing environmental conditions and to adjust their metabolism accordingly.

Furthermore, the transcriptome of *B. phytofirmans* PsJN colonizing potato plants provided the possibility to study endophyte-host interactions on the RNA level in more detail. The transcriptomic analysis reveals that cell motility and defense mechanisms show minor changes inside plants. This is in accordance with recent findings that show these mechanisms to be important in the spreading of endophytic cells throughout plant organs and tissues [1,2]. My results indicate that active movement is less important once a bacterial population is successfully established inside plants. Interestingly, also defense related traits seem to play a minor role in the endophytic life of *B. phytofirmans* PsJN. A possible explanation is that the plant endosphere is a protected niche allowing endophytes to escape the high competitive pressure in rhizosphere and soil. In addition, numerous important traits related to plant growth promotion detected in the transcriptome of *B. phytofirmans* PsJN *in planta* under stress condition in comparison with control samples. These traits include genes encoding quinolinate phosphoribosyl transferase, indole-3-acetic acid (IAA) synthesis, ACC deaminase activity and siderophore production genes. Furthermore, my analysis showed that the majority of expressed traits in the new

habitat are related to transcription regulation, general metabolism (sugars, amino acids, lipids, and nucleotides), energy production, and cellular homeostasis. Previous studies suggested that plasmids are genetic determinants of functional diversification and niche adaptation [3,4], while our analysis on the plasmid of *strain* PsJN enabled speculation about its ability to colonize plants and to establish endophytic population.

One of the main interest of this thesis is to investigate whether and how endophytic *B. phytofirmans* PsJN is affected by host plant drought stress. Transcriptome analysis of *B. phytofirmans* PsJN colonizing non-stressed and drought-stressed potato plants indicates the alteration of bacterial gene expression under drought stress condition. This analysis shows that *B. phytofirmans* PsJN adjusts gene expression to altered physiological conditions in host plants due to plant stress response. Our results show that an increase in bacterial activity upon host plant drought stress might help to maintain the redox and energy balance inside plant tissue, thus reducing the effects of drought stress on plants. Overall, transcriptomics analysis indicates that the bacterium is metabolically active inside host plant environment without major changes in the main functions encoded on the genome of *B. phytofirmans* PsJN.

Furthermore, three closely related strains of *P. ananatis* with different lifestyles, colonizing maize seeds, were analyzed to gain a comprehensive understanding of plant-endophyte interaction by comparative genomics approaches. These strains of *P. ananatis* were shown to have different types of interactions with their host ranging from pathogenicity to mutualism [3]. Genome comparison of closely related strains with different lifestyles is a well-established way to reveal the link between presence and absence of genes with lifestyle factors such as niche specificity or host range [5]. In my thesis, the maize seed endophyte *P. ananatis* S6 shows strong beneficial effects on maize growth, while strain S7 induced weak pathogenicity symptoms. *P. ananatis* S8 had hardly any effect and can be considered as commensal. My comparative genomics analysis indicates that the core genomes of *P. ananatis* strains S6, S7 and S8 are highly conserved (85-87%). Despite the overall high degree of similarity between the core genomes of these three maize seed endophytes, I was able to detect several important differences in transposase/integrases/phage related genes, type VI secretion system, and eukaryotic-like protein domains.

By contrasting the genomes of three *P. ananatis* strains, I found differences in mobile genetic elements such as integrase genes, transposase genes and phage related genes. That confirms a potential role of these elements in the diversification of related strains colonizing the same ecological niche. An over-representation of transposase genes and mobile elements also indicates the genomes potential for acquisition of novel functions. The reduced number of mobile elements in *P. ananatis* S6 on the other hand could indicate high stability of its genome, implying good adaption to the habitat. Another interesting observation is that the type VI secretion system elements are present in both pathogenic and non-pathogenic *P. ananatis* strains. This finding supports the idea that the type VI secretion system itself is not necessarily a determinant of pathogenicity and could play a role in competition against other microorganisms, fitness or niche adaptation [6–8]. In addition, comparing the genes of Large Pantoea Plasmid (LLP-1) in three strains of *P. ananatis* S6, S7 and S8 in my thesis supports the concept of niche adaption and functional diversification of the genus *Pantoea* which was introduced before [3,9]. Overall, the comparative genomics analysis detected the factors and traits involved in the genotype and phenotype differences between *P. ananatis* strains S6, S7 and S8 colonized maize seed.

The use of microbial inoculants in crop production, growth and protection is increasing in agricultural studies. I conducted an approach to introduce a beneficial microorganism which was not usually found in seeds into the seeds of multiple plant species. The introduced bacteria are capable of subsequential modifications to seed microbiome and growth traits in crop seed embryos of both monocots and dicots. To reach this aim, the bacterium *B. phytofirmans* PsJN is sprayed on the parent flowers (maize, soy and pepper). My findings show that the PsJN strain entered the plants and was detected in the cotyledon part of the embryo and also on the seed coat together with other bacteria. These results show that the PsJN strain colonized the emerging seeds and was thus inherited between the generations. Furthermore, *B. phytofirmans* PsJN only or a mixture of *B. phytofirmans* PsJN and *Paenibacillus* sp. S10, respectively, was applied on a wheat field to test the possibility of modulation of seed microbiome during seed production in the field. The results show that at seed maturity strain PsJN was effectively introduced into the wheat seeds (up to 92%) while the inoculation with S10 strain was less efficient (only 113%). Additionally, culture-independent community analysis of single seeds from control seeds, PsJN-

seeds and PsJN+S10-seeds showed that three treatment were differing mainly in the abundance of certain groups of bacteria (α-, β-, γ-*Proteobacteria*). My analysis illustrates several positive and negative correlations between OUTs representing PsJN and OTUs from other taxas (α-, β-, γ-*Proteobacteria*). These findings show that the seed microbiome composition in the parent plants changed, either via the added microbe alone or via the observed community effects. In overall, this novel approach for altering the seed microbiome to achieve growth modulation in plants has significant impact on follow-up seed-endophyte studies, as well as on the endophyte-plant optimization in agriculture.

The genomic information and factors that were obtained in this study by application of genomics and transcriptomics approaches provide novel insights into endophyte-plant interactions and enable us to monitor endophytic bacteria upon different conditions and lifestyles. In future, the development of dual RNA-seq might have the potential to determine the transcriptome of endophytic bacteria and their host, simultaneously.

In conclusion, I could trace the fate of endophytic bacterial strains that colonize inside the host plant tissue. The applied approaches gave insights into the gene expression pattern of endophytic bacteria and their metabolic and functional potentials in interaction with their host plant. Due to the low DNA concentration of endophytes inside the plant tissue, it is difficult to detect the bacteria inside the host plant using experimental methods. Succeeding inoculation assays with other endophytic strains *in planta* might further increase our understanding of the complex interactions between endophytic bacteria and their host plant. The findings of this thesis will facilitate the efforts of selecting microbial inoculants for improved plant production in sustainable agriculture in future.

References

1 Compant S, Mitter B, Colli-Mull JG, *et al.* Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. *Microbial ecology* 2011;**62**:188–97. doi:10.1007/s00248-011-9883-y

2 Balsanelli E, Tadra-Sfeir MZ, Faoro H, *et al.* Molecular adaptations of Herbaspirillum seropedicae during colonization of the maize rhizosphere:H. seropedicae transcript profiles in maize rhizosphere. *Environmental Microbiology* 2015;:n/a – n/a. doi:10.1111/1462-2920.12887

3 De Maayer P, Chan WY, Rubagotti E, *et al.* Analysis of the Pantoea ananatis pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts. *BMC genomics* 2014;**15**:404. doi:10.1186/1471-2164-15-404

4 Heuer H, Smalla K. Plasmids foster diversification and adaptation of bacterial populations in soil. *FEMS Microbiol Rev* 2012;**36**:1083–104. doi:10.1111/j.1574-6976.2012.00337.x

5 Frank AC. The Genomes of Endophytic Bacteria. In: Pirttilä AM, Frank AC, eds. *Endophytes of Forest Trees*. Springer Netherlands 2011. 107–36.http://dx.doi.org/10.1007/978-94-007-1599-8_7

6 Weber B, Hasic M, Chen C, *et al.* Type VI secretion modulates quorum sensing and stress response in Vibrio anguillarum. *Environmental microbiology* 2009;**11**:3018–28.

7 English G, Trunk K, Rao VA, *et al.* New secreted toxin and immunity proteins encoded within the Type VI secretion system gene cluster of Serratia marcescens. *Molecular microbiology* 2012;**86**:921–36.

8 Shyntum D, Venter S, Moleleki L, *et al.* Comparative genomics of type VI secretion systems in strains of Pantoea ananatis from different environments. *BMC genomics* 2014;**15**:163.

9 De Maayer P, Chan WY, Blom J, *et al.* The large universal Pantoea plasmid LPP-1

plays a major role in biological and ecological diversification. *BMC genomics* 2012;**13**:625.

Chapter VII

Summary & Zusammenfassung

Summary

Endophytic bacteria have a large spectrum of effects towards increasing their host plant productivity. Therefore, they are in the center of attention for having a sustainable agriculture in a less-stable environment. The welfare of interaction between endophytic bacteria and their host plant effects plant growth, nutrient uptake, protection against pathogens and biotic/abiotic stress tolerance.

The primary objective of my study is to explore the machnisms of endophyte-plant communication and adaptation of endophytic bacteria to the host plant environment. Whole transcriptome sequencing of *Burkholderia phytofirmans* PsJN colonizing potato plants (*Solanum tuberosum* L.) enabled the analysis of *in planta* gene activity and the response of strain PsJN to plant stress using high throughput RNA-seq analysis and TaqMan-qPCR. Regarding the broad array of functions encoded on the genome of strain PsJN, my analysis shows that transcripts up-regulated in response to plant drought stress are mainly involved in transcriptional regulation, cellular homeostasis and the detoxification of reactive oxygen species, indicating oxidative stress response in PsJN. Endophytes react to changes in the plant/seed physiology due to plant stress and adjust its gene expression pattern to cope with and adapt to the altered conditions.

Furthermore, genetic attributes that could explain the phenotypic differences between closely related endophytic, commensal and pathogenic bacteria were identified regarding their interaction with the host plant. Here, I investigate three novel closely related *Pantoea ananatis* strains (named S6, S7 and S8) and their effect on the maize host plant. Although they were isolated from seeds of healthy plants, they showed distinct characteristics in regard to plant growth and health ranging from pathogenic (S7), commensal (S8) to a beneficial, growth promoting effect on maize (S6). Despite the high similarity in the genomes of three strains, my comparative analysis indicates that genomic differences mainly exist in cell surface components, motility related proteins, type VI secretion system, transposase/integrase/phage related genes and eukaryotic-like domain containing proteins.

Apart from the agricultural importance of endophytic bacterial interaction with plants, they also play important roles in increasing crop productivity and protection in endophyte-seed interaction. For revealing the roles of members of seed microbiomes in plant growth, I describe a novel approach to modulate the microbiome of elite crop seed embryos, of both monocots and dicots. A selected microbe (*B. phytofirmans* PsJN) is introduced into the parent plants (maize, soy, pepper and wheat) before seed development occurs. These become incorporated into the seed microbiome, thereby achieving vertical inheritance to the offspring generation. Using this approach in future will help us to track the faith of seed

endophytes during seed dormancy and germination and unpuzzle the role of seed endophytes in the development of the plant endophyte microbiota.

In my thesis, I showed a two-sided interaction between endophytic bacteria with their host (plant and/or seed), which will opens up new avenues of thinking about this complex communication inside the plant environment.

Zusammenfassung

Endophytische Bakterien haben ein breites Spektrum an Wirkungen auf ihre Wirtspflanze um deren Produktivität zu steigern. Daher sind sie im Mittelpunkt der Aufmerksamkeit im Bestreben eine nachhaltigen Landwirtschaft in einer sich rapide ändernden Umgebung zu schaffen. Die Interaktion zwischen endophytischen Bakterien und ihren Wirtspflanzen hat Auswirkungen auf das Pflanzenwachstum, die Nährstoffaufnahme, Schutz gegen Krankheitserreger und auf die biotischen / abiotische Stresstoleranz

Das primäre Ziel meiner Doktorarbeit ist es, die Mechanismen der Kommunikation zwischen Endophyten und ihrer Wirtspflanze sowie die Anpassung von endophytischen Bakterien an ihre Umgebung in der Wirtspflanze zu erforschen. Whole Transkriptom-Sequenzierung von *Burkholderia phytofirmans* PsJN in kolonisierten Kartoffelpflanzen (Solanum tuberosum L.) ermöglichte die Analyse der *in planta* Genaktivität und die Reaktion von PsJN auf Stressfaktoren der Wirtspflanze mittels Hoch-Durchsatz RNA-seq Analyse und TaqMan-qPCR. Im Hinblick auf die breite Palette an Funktionen welche auf dem Genom von PsJN codiert sind, zeigt meine Analyse, dass während der Austrocknung der Pflanze hochregulierte Transkripte hauptsächlich beteiligt sind an der Transkriptionsregulation, zellulärer Homöostase und der Entgiftung von reaktiven Sauerstoff. Dies weist auf die oxidative Stressreaktion in PsJN hin. Endophyten reagieren auf Veränderungen in der Physiologie von Pflanze bzw. Samen aufgrund von Stressreaktionen und passen sich durch veränderung der Genexpressionsmuster an die veränderten Bedingungen an.

Darüber hinaus habe ich genetischen Eigenschaften identifiziert, die die phänotypischen Unterschiede zwischen nahe verwandten endophytisch, commensal und pathogene Bakterien in Bezug auf ihre Interaktion mit der Wirtspflanze erklären könnten. Hier untersuchte ich drei kürzlich entdeckte, eng verwandte *Pantoea ananatis* Stämme (genannt S6, S7 und S8) sowie deren Wirkung auf die Mais-Wirtspflanze. Obwohl die Stämme aus Samen von gesunden Pflanzen isoliert wurden, zeigten sie unterschiedliche Merkmale in Bezug auf das Pflanzenwachstum und die Gesundheit der Maispflanzen, von einer pathogenen (S7) zu einer neutralen (S8) bis hin zu einer vorteilhaften, wachstumsfördernden Wirkung (S6). Trotz der hohen Ähnlichkeit der Genome der drei Stämme zeigt meine vergleichende Analyse, dass genomische Unterschiede bestehen, vor allem in der Komponenten der Zelloberfläche, in Motilität verwandten Proteinen, Proteinen des Typ VI-Sekretionssystems, Transposon / Integrase / Phagen-verwandte Gene und Proteinen, welche eukaryotische Domänen enthaltenden.

Neben der landwirtschaftlichen Bedeutung der Interaktion zwischen endophytischen Bakterien und Pflanzen, spielt auch die Schutzwirkung durch die Interaktion zwischen

Endophyt und Pflanzensamen eine wichtige Rolle bei der Steigerung der Ernteerträge. In meiner Arbeit beschreibe ich einen neuartigen Ansatz zur Modulierung der Mikrobiome in Saatgut, gleichermaßen für Monokotyledonen und Dikotyledonen. Ein ausgewählter Stamm (*B. phytofirmans* PsJN) wird in die Hostpflanzen (Mais, Soja, Paprika und Weizen) eingeführt bevor die Samenentwicklung eintritt. So wird der Endophyt in die Samen eingebaut, wodurch eine vertikale Vererbung auf die Nachkommen Generation erreicht wird. Dieser Ansatz wird in Zukunft dabei helfen, die Entwicklungsprozesse von Endophyten während Samenruhe und Keimung zu untersuchen und die Rolle von Samen-Endophyten in der Entwicklung der Mikrobiota zwischen Pflanze und Endophyt aufzuklären.

In meiner Dissertation habe ich eine gegenseitige Interaktion zwischen endophytischen Bakterien mit ihrem Wirt (Pflanze bzw. Samen) gezeigt. Dies kann neue Wege eröffnen um diese komplexen Wechselbeziehungen besser zu verstehen.