Spring 2017

Molecular mechanisms of salt and osmotic stress tolerance in Frankia strains isolated from Casuarina trees

REDIET OSHONE
University of New Hampshire, Durham

Molecular Mechanisms of Salt and Osmotic Stress Tolerance in *Frankia* Strains Isolated from *Casuarina* trees

BY

Rediet Tamire Oshone

B.Sc., Mekelle University, 2005

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirement for the Degree of

Doctor of Philosophy

In

Genetics

May, 2017

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Genetics by:

Dissertation Director, Louis S. Tisa, Professor of Microbiology and Genetics

Estelle Hrabak, Associate Professor of Genetics and Molecular Biology

W. Kelly Thomas, Professor of Genetics

Feixia Chu, Associate Professor of Biochemistry, Molecular, and Cellular Biology

Antony Champion, Plant Molecular Biology, Institut de recherche pour le développement (IRD), France

On April 03, 2017

Original approval signatures are on file with the University of New Hampshire Graduate School.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Louis Tisa, for providing me with the resources and funding required for the completion of my research. The continued financial support by Dr. Tisa obviated the need to financially support myself and allowed me to solely focus on my research. Dr. Tisa provided me with invaluable guidance and advice and was actively engaged in the entire process of my research- thank you Dr. Tisa! I would also like to extend my deepest gratitude to my dissertation committee members, Dr. Estelle Hrabak, Dr. Feixia Chu, Dr. W. Kelly Thomas, and Dr. Antony Champion, for the valuable comments, suggestions, and guidance they provided. The UNH graduate school, which awarded me a generous merit-based graduate fellowship when I first joined the graduate school in August 2011, deserves my heartfelt appreciation and gratitude. The generous award made my settlement in the USA smooth and easier.

Last but not least, I would like to thank my family for always being there for me and for giving me unconditional support. The selfless devotion and sacrifice of my mother to ensure a better future for me and my siblings deserves the highest appreciation. I'm also grateful to my father for instilling in me the love of academia and science. His passion for education and the extraordinary commitment he showed to make environment conducive for my academic progress was undoubtedly a key positive force in my academic journey. I was raised to be competitive from an early age and settling for second best was never an option. Although it seemed a little strenuous, or even

unnecessary at times, the high expectations my parents placed on me molded my work ethic, discipline, and outlook on life in general; I'm forever grateful for that.

The following funding sources supported the research in this dissertation:

New Hampshire Agricultural Experimental Station (NHAES)

National Institute of Food and Agriculture - USDA

Academy of Applied Sciences – REAP

USAID Borlaug Fellowship

Hatch NH637

Summer TA fellowship

NH Graduate School Travel Grants

MCBS Travel Grants

Zsigray Travel Grants

Contents

LIST OF FIGURES

LIST OF TABLES

Abstract

Molecular mechanisms of salt and osmotic stress tolerance in *Frankia* strains isolated from *Casuarina* trees

by

Rediet Tamire Oshone

University of New Hampshire, May 2017

Globally, 20% of total cultivated and 33% of irrigated agricultural lands are affected by high salinity. By 2050, more than 50% of the arable land will be salinized. The hyper-ionic and hyper-osmotic stresses associated with salt-affected soils threaten the ability of cells to maintain optimal turgor pressure and intracellular ionic concentration for growth and functioning. The nitrogen-fixing soil actinobacterium *Frankia* shows marked variability in its tolerance to salinity. When in a symbiotic association with actinorhizal plants, *Frankia* enhances the tolerance of the plants to a range of abiotic stresses, including salinity. The *Casuarina-Frankia* association has been used to reclaim salt affected soils worldwide. Optimizing the use of the *Casuarina-Frankia* association for saline soil reclamation requires identifying salt-tolerant symbionts, unlocking the molecular mechanism behind the tolerance, and ultimately developing *Frankia* strains that combine the best symbiotic characteristics with high level of salt tolerance.

In this study, *Frankia* strains were screened for salt and osmotic stress tolerance under nitrogen-proficient and nitrogen-deficient conditions. Salt-tolerant and salt-sensitive strains were identified and the effect of salt and osmotic stress on the physiology of the strains and on their symbiotic performance was assessed. Tolerant

strains were sequenced and comparative genomics, transcriptome profiling, proteomics, and physiological analysis were employed to identify potential mechanisms and candidate genes responsible for the contrasting phenotypes. An expression vector that stably replicates in *Frankia* was developed and used to constitutively express some of the candidate genes in the salt-sensitive strain.

Salt-tolerant *Frankia* strains (CcI6 and Allo2) that could withstand up to 1000 mM NaCl and a salt-sensitive *Frankia* strain (CcI3) which could withstand only up to 475 mM NaCl were identified. Comparative genomic analysis showed that all of the *Casuarina* isolates belonged to the same species (*Frankia casuarinae* at *p=0.05* level). Pangenome analysis revealed a high abundance of singletons among all *Casuarina* isolates. The two salt-tolerant strains contained 153 shared unique genes (the majority of which code for hypothetical proteins) that were not found in the salt-sensitive strain. Transcriptome, proteome, and physiological analysis of the salt-tolerant and sensitive strains revealed vast differences in salt stress response with regards to cellular functions such as transcriptional regulation, cell envelope remodeling, osmolyte biosynthesis, and signal transduction. Among the 153 genes shared only between the salt-tolerant strains, seven, including a zinc peptidase, were responsive to salt stress. Constitutive expression of the zinc peptidase gene in the salt-sensitive strain (CcI3) led to increased salt-tolerance.

The comprehensive approach we took to analyze the complex trait of salt stress tolerance led to important findings that shape our understanding of the salt stress response. The tools and the data we generated in this study willaaaAAAaa serve as a

springboard for future work in the area or in the broader field of *Frankia* genetics in general.

Introduction

Soil salinization

Soil salinization is a worldwide problem that is intensifying because of the effects of climate change. Globally, 20% of total cultivated and 33% of irrigated agricultural lands are affected by high salinity [1]. Salinized areas are expanding at an alarmingly rate of 10% per annum due to a variety of factors, including, but not limited to, low precipitation, high surface evaporation, weathering of native rocks, irrigation with saline water, and poor cultural practices [1]. Every day for more than 20 years, an average of 2,000 hectares of irrigated land in arid and semi-arid areas across 75 countries have been degraded by salt [2]. By 2050, more than 50% of the arable land is predicted to be salinized [3]. The global loss in crop production due to salt-induced degradation in irrigated areas is estimated at US$ 27.3 billion per year [2]. A saline soil is defined a soil with electrical conductivity (EC) of the saturation extract (ECe) greater than 4 dS m^{-1} (approximately 40 mM NaCl) at 25°C and with an exchangeable sodium of less than 15%. The yield of most crop plants is reduced at a salinity level of 4 dS m^{-1} or even lower [3, 4].

Salt stress affects crop growth in the form of osmotic stress followed by ion toxicity [5, 6]. The osmotic effect of salt stress, which follows immediately after the application of salt, results in stomatal closure, reduced water uptake by the root system, limited cell expansion and cell division, decreased photosynthetic activity, nutrient imbalance, and reduced ability to detoxify reactive oxygen species [5, 6, 7, 8]. During long-term salinity stress, plants also experience a hyperionic stress. The accumulation of both Na^+ and Cl^- inside cells causes severe ion imbalance and physiological

disorders [5, 9]. Oxidative damage to various cellular components such as proteins, lipids, and DNA from salinity-induced reactive oxygen species (ROS) [singlet oxygen, superoxide, hydroxyl radical, and hydrogen peroxide] formation can have detrimental effects [10, 11]. High Na^+ concentration also inhibits uptake of K^+, NO^-, PO_4^{3-} ions, which leads to decreased growth or death [9, 12].

Feeding the burgeoning global population, which is projected to reach almost 10 billion by 2050, requires slowing down the rate of soil salinization and reclaiming land already affected by salinity. Various methods are used to reclaim salt-affected soils. One effective method commonly used in the reclamation of salt-affected soils involves initiating plant succession using fast-growing, nitrogen-fixing actinorhizal trees such as the *Casuarina* [13]. Actinorhizal trees form a nitrogen-fixing, mutually beneficial association (also referred to as the actinorhizal symbiosis) with *Frankia*, a sporulating filamentous actinomycete that can fix nitrogen and supply it to the plant [14]. The symbiosis with *Frankia* contributes towards the ability of actinorhizal plants to thrive in and ameliorate harsh environmental conditions. A brief overview the actinorhizal symbiosis and its role in stress adaptation by actinorhizal plants is provided in the next subsection.

The actinorhizal symbiosis

Actinorhizal plants are defined by their ability to form root nodules when in symbiosis with *Frankia*, and comprise of more than 220 dicotyledonous species classified into 25 genera and 8 Angiosperm families, including Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Elaeagnaceae, Myricaceae, Rhamnaceae and Rosaceae [14, 15]. In the nodule, *Frankia* fixes nitrogen, converts atmospheric N_2 to NH_3, a form that is

transported to the host plant and used to meet most of the nitrogen requirements of the plant [16]. The *Frankia*, in return, benefits from the association in the form of reduced carbon [17].

Actinorhizal plants, along with legumes (family Fabaceae), which form a nitrogen-fixing association with rhizobia (Gram negative) members of the alpha-subgroup of the phylum proteobacteria, belong to the 'nitrogen-fixing Clade' within the Rosid I lineage described previously [18]. Despite sharing some common features and a similar outcome with the rhizobium-legume symbiosis, the actinorhizal symbiosis is remarkably distinct [15]. Unlike the legumes, where the nodules represent stem-like organs with a peripheral vascular system and infected cells in the central tissue, the architecture of actinorhizal nodules consists of multiple modified lateral roots with central vascular tissue and infected cells in the expanded cortex. Actinorhizal nodule primordia, like lateral root primordia, appear in the root pericycle [17] whereas legume nodule primordia appear in the root cortex. Another distinct feature of the actinorhizal symbiosis is the wide distribution of the plant symbionts in eight botanical families. In the rhizobium-legume symbiosis, all the plant symbionts belong to just one family, the Fabaceae. Actinorhizal plants inhabit all continents except Antarctica. In contrast to the legumes, actinorhizal plants are mainly woody shrubs and trees, the genus *Datisca* being the exception [17]. In terms of nitrogen fixed per hectare of land per year, the actinorhizal symbiosis is comparable to the rhizobium-legume symbiosis [14]. Actinorhizal plants prevent soil erosion, replenish nutrients, enhance soil fertility and crop yield, help to restore a healthy soil microbiome in contaminated areas, and provide fuel wood, rendering them invaluable in sustainable agroforestry [19, 20]. Additionally,

actinorhizal trees have food and medicinal value [21, 22, 23], play a role in land reclamation, serve as wind breaks, and prevent coastal erosion [24, 25, 26]. Plants of Casuarinaceae and Betulaceae families are the most widely planted around the world for the rehabilitation of degraded lands [27].

The microsymbiont in actinorhizal symbiosis, *Frankia*, is a Gram-positive, aerobic and heterotrophic bacterium that can fix nitrogen under both free-living conditions and inside symbiotic root nodules [28]. During free living nitrogen fixation and under symbiotic conditions with most hosts, *Frankia* differentiates into vesicle cells, which are unique structures protected by multi-layered, hopanoid lipid envelopes [29]. The oxygen-liable nitrogenase enzyme catalyzes the conversion atmospheric di-nitrogen into ammonia inside the vesicles. This is different from the case in rhizobia, where leghemoglobin supplied by the host plant controls oxygen levels [30]. Based on 16S rDNA phylogeny, 4 lineages of *Frankia* are identified. Lineage I strains form nodules on members of the Betulaceae, Casuarinaceae and Myricaceae. Lineage II strains associate with the Datiscaceae, Coriariaceae, actinorhizal members of the Rosaceae, and Ceanothus of the Rhamnaceae. Lineage III *Frankia* nodulate the Myricaceae, Rhamnaceae, Elaeagnaceae, and Gymnostoma of the Casuarinaceae. Lineage IV includes non-symbiotic *Frankia* strains that are unable to re-infect their host plant and/or fix nitrogen [31]. A concatenated phylogenetic tree of the core genome based on Manhattan distance supports the lineage classification based on the 16S rDNA (Fig 1) [32]. Our understanding of *Frankia* was dramatically improved after the sequencing of the first three *Frankia* genomes in 2007, which revealed new potential with respect to metabolic diversity, natural product biosynthesis and stress tolerance [33, 34, 35]. In the

decade that elapsed since 2007, more than three dozen *Frankia* genomes have been sequenced and annotated, further providing a rich database for the genetic, genomic, transcriptomic, and proteomic study of the bacterium. The current state of the *Frankia* genomes and some of the intriguing insights they provided into the eccentric, symbiotic lifestyle of this microorganism are succinctly discussed below.

Status of *Frankia* genomes

Since 2007, more than 35 *Frankia* isolates from all four lineages have been sequenced. Two of the genomes come from uncultured *Frankia*. Seven of the sequenced genomes have been assembled to a single scaffold while the rest are at varying degrees of completion ranging from 2 to 2738 scaffolds. The sizes of the genomes vary from 5.0 Mb for *Frankia* sp. strain CeD to 11.2 Mb for *Frankia* sp. strain. BMG5.36. The sequenced *Frankia* strains include *Casuarina* isolates from a broad range of geographic locations [32, 33, 36 - 41]. Complete genomes provide a level of contiguity and error checking not possible with draft genomes [42]. Nevertheless, the finishing step involved in closing a genome is very labor-intensive and costly. The advent of next generation sequencing and the concomitant increase in computational capability allowed for a rapid and cost-effective production of draft genomes, which usually provide sufficient information for analysis.

After the sequencing of the first three *Frankia* genomes [33], a hypothesis was proposed that genome size was related to host specificity and biogeography ranges. Subsequent sequencing of additional genomes from all four lineages of *Frankia* rendered support to this hypothesis [32, 43]. The size variations in *Frankia* genomes have been suggested to reflect differences in saprotrophic potential [31]. Genome

elasticity as measured by the distribution and numbers of mobile genetic elements, including horizontal transferred genes (HTG) and insertion sequences (IS), is not correlated with genome size [32]. HTG number is not correlated with IS number and genome size. Examination of the relationship between sequenced *Frankia* genomes and their corresponding GC contents revealed three different clusters (Fig 2) [32]. Cluster 1 is characterized by the lowest GC content and the smallest genome size. Included in this cluster are lineages Ic and II. The second cluster, which includes lineages III and IV, is characterized by the largest size genome and a GC content that is in-between the other two clusters. The last cluster has the highest GC content and an intermediate-size genome and is comprised of members of lineages Ia, III and IV.

In rhizobia, genes involved in symbiosis are often clustered on large plasmids (pSym), or are located within symbiosis islands in the genome, suggesting that they may not be essential under most circumstances and they could be acquired through horizontal

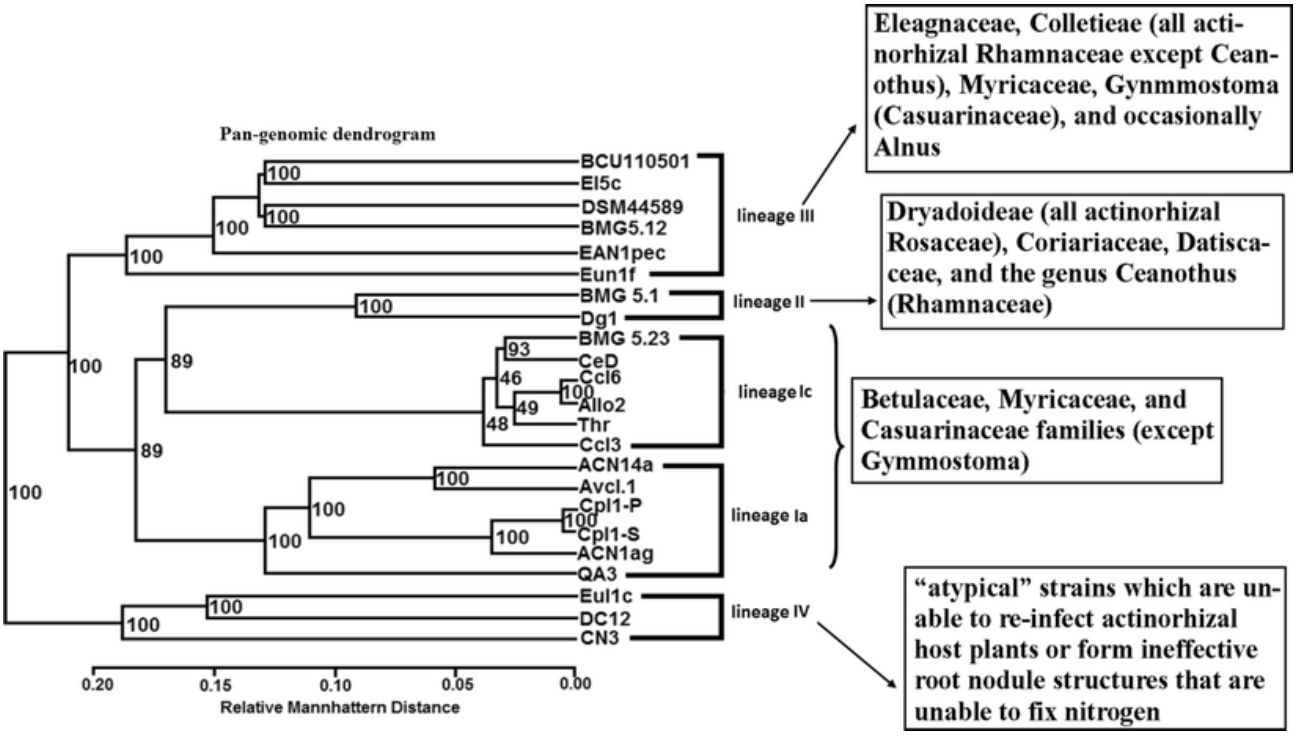

Figure 1. A concatenated phylogenetic tree of the core genome based on Manhattan distance and bootstrap value of 1000. The bootstrap values have been represented as percentages. For each lineage, the host plants are given within the box. The core genome consisted of 1421 genes. Adapted from (Tisa et al. 2016).

gene transfer [44]. In *Frankia*, however, the symbiotic genes are not clustered, but are spread through the *Frankia* genomes [34].

In legume/rhizobia symbioses, lipochitooligosaccharide signal molecules known as Nod-factors are involved in the recognition between the microsymbiont and the plant host. The *nod* genes encoding the proteins involved in the synthesis of the Nod factors are activated when the bacteria perceive flavonoids that are secreted by the host plant root [45]. The perception of Nod factors by host plant through LysM-receptor-like kinases (LysM-RLKs) induces a signal transduction cascade that is required for infection and nodule organogenesis [46]. This molecular dialog takes place in most legume–rhizobium interactions. Analysis of the *Frankia* genomes for common *nod* genes failed to reveal their presence. With the exception of *Frankia datiscae* Dg1 [47] and a recent assemblage of three lineage II strains from California (Dg2) [48], other *Frankia* strains do not contain the canonical *nod* genes (*NodA*-acyl transferase, *NodB*-chitin deacetylase, *NodC*-chitin synthase), suggesting that the majority of *Frankia* strains use novel signaling compounds during the infection of actinorhizal plants. DNA samples used for the sequencing of *Frankia datiscae* Dg1 and the Dg2 assemblage genomes were extracted from vesicle clusters isolated from the root nodules of *Datisica glomerate* [47, 48].

The availability of several sequenced *Frankia* genomes has allowed for the use of genome mining, comparative genomics, transcriptomics and proteomics approaches to cast light on the stress tolerance, secondary metabolism biosynthesis, symbiosis and nitrogen fixation mechanisms of *Frankia* [26, 33, 49-56]. Our laboratory has been undertaking the sequencing of numerous *Frankia* genomes, some of which were from

salt tolerant *Casuarina* isolates and were sequenced with the goal of understanding the genetic basis for the observed difference in salt tolerance between different *Casuarina* isolates. Sequencing of the genomes allowed for comparative genomics, transcriptomics, and proteomic analysis of salt tolerance in *Casuarina* isolates.

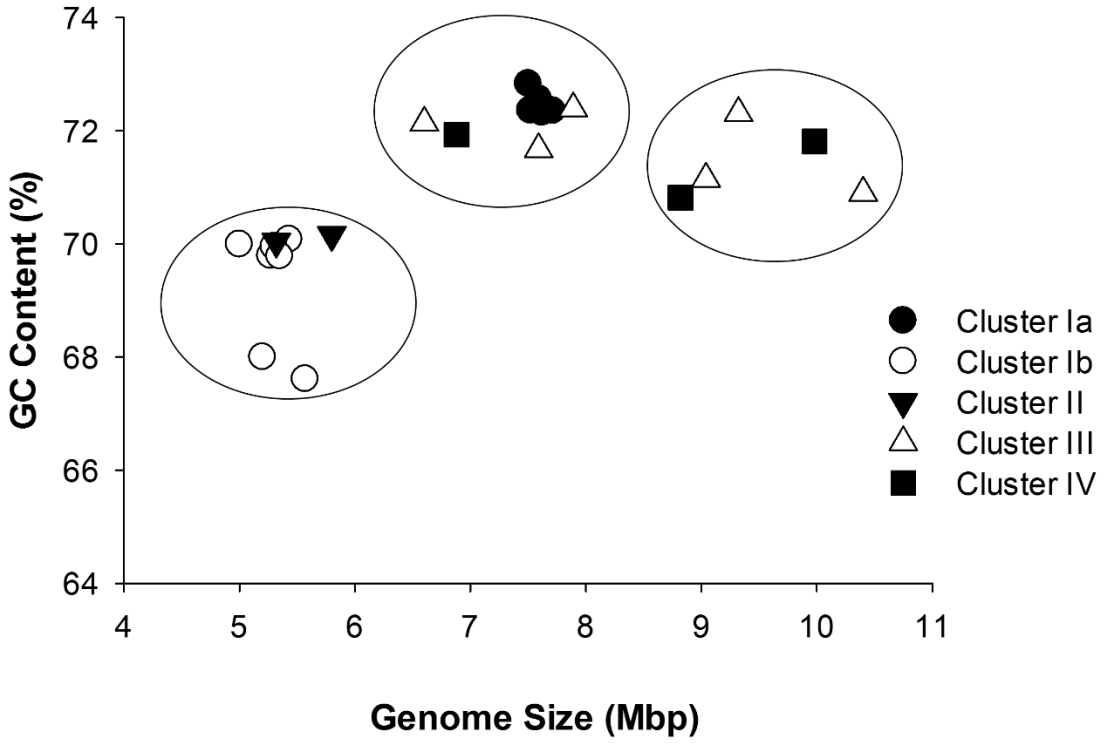

Figure 2 Analysis of *Frankia* genomes sizes and GC content. *Frankia* genomes were clustered in three groups. Adapted from (Tisa et al. 2016).

Actinorhizal symbiosis and environmental stress

The symbiotic association with *Frankia* provides actinorhizal plants with multifaceted benefits, including the ability to colonize harsh environmental terrains, such as soils characterized by extreme pH, high levels of toxic heavy metals, low water availability, poor drainage, and low nutrient content. [17, 57, 58]. The benefit from actinorhizal symbiosis in the form of fixed nitrogen translates into increased tolerance to environmental stresses by the plant as many stress coping mechanisms heavily rely on protein biosynthesis and hence the availability of usable form of nitrogen. Although nitrogen (N_2) is by far the most abundant gas in the atmosphere (80%), only certain prokaryotes can form the nitrogenase enzyme complex to convert atmospheric dinitrogen into ammonia, a form that can be used by plants. In addition to the symbiosis with *Frankia*, actinorhizal plants also take part in mycorrhizal associations, giving them an additional edge to colonize marginal soils [59]. Consequently, the actinorhizal plant-*Frankia* system is widely used for reclaiming lands affected by abiotic stresses [60]. By promoting pedogenetic processes that lead to the formation of a more favorable soil for the installation of other plant species, actinorhizal plants play a vital ecological role in the revegetation of different landscapes to prevent desertification [24]. As expected based on their ability to thrive and form effective symbiosis under harsh environmental conditions, many *Frankia* strains are resistant to extreme environmental conditions, including elevated level of several heavy metals [61] and soil salinity [62].

The success of the actinorhizal symbiosis in conferring tolerance to environmental stress is nowhere more evident than in the salt tolerance of *Casuarina*–*Frankia* associations. *Casuarina* trees are able to grow under very saline conditions

and have been extensively used in the reclamation of degraded lands worldwide [27, 63]. The ability of *Casuarina* plants to thrive in and alter degraded lands is harnessed in Africa where the Casuarinaceae are extensively planted to reclaim salt-affected soils and to stabilize coastal and desert dunes [64]. The *Casusarina-Frankia* association and its role in the salt tolerance of the plant and the microsymbiont is covered in depth in the following subsection. Some of the information presented comes from published works by our group [65, 66].

The *Casuarina* – *Frankia* association and salt tolerance

Originating from Australia, South-East Asia, Malaysia, Melanesian, and Polynesian regions of the Pacific, New Guinea, the Casuarinaceae are widely distributed in tropical areas [27]. This family is comprised of 96 species distributed in 4 genera. In intercropping cultivation with crops, *Casuarina* trees improve soil fertility and crop yield [67]. In several countries in Africa and Asia, *Casuarina* trees have been widely used to fix sand dunes and serve as wind breaks to protect cultivated crops [67, 68]. *Casuarina* trees are also used as poles, source of smokeless fuelwood with a high calorific value, raw material for construction [69] and in the production of paper pulp wood [7].

Because of their ability to tolerate a range of stresses such as salinity, drought, toxic heavy metals, and flooding [70], *Casuarina* trees are widely used in the reclamation of degraded lands [71]. The growth and stress tolerance of *Casuarina* trees improves when the plants are in a symbiotic association with *Frankia* [65, 70, 72]. *Casuarina* trees are particularly outstanding among actinorhizal plants in terms of their salt tolerance [70]. *Casuarina* trees are able to grow under saline conditions and have

been used as a green barrier [63, 73, 74]. Some *Casuarina* species are found growing naturally near brackish waters and swamps or in saline soils [75]. In hydroponic medium supplemented with adequate nitrogen, *Casuarina glauca, Casuarina obesa*, and *Casuarina equisetifolia* var. incana will withstand up to 500 mM NaCl [63]. Although generally tolerant to salt stress, *Casuarina* trees have a high degree of diversity in salt tolerance. This is evident in *C. equisetifolia* clones which show a marked variability in salt tolerance [76–54]. Therefore, saline land reclamation efforts using *Casuarina* trees should involve screening processes to identify salt-tolerant clones that can grow in degraded lands [27]. *Casuarina* strains are absent or rare in soil without the host plant, which indicates the requirement of inoculation of the host plant before transplantation for rehabilitation of saline soil [28].

Like their plant partners, *Frankia* strains isolated from *Casuarina* and *Allocasuarina* are more NaCl tolerant than other *Frankia* strains isolated from actinorhizal plant species not normally growing under sodic conditions [77]. *Frankia* sp. strain CcI6 isolated from the nodules of *Casuarina cunninghamiana* trees growing in Egyptian soil is highly NaCl tolerant, exhibiting a minimum inhibitory concentration (MIC) value of 1000 mM [78]. *Frankia* sp. strain CcO1, also from *C. cunninghamiana*, tolerates up to 500 mM NaCl [13]. Similarly, *Frankia* sp. strain Ceq1 isolated from *C. equisetifolia* is able to withstand up to 500 mM NaCl [70]. However, a huge variation in terms of salt tolerance exists among the different *Casuarina* isolates starting from a NaCl MIC value as low as 100 mM [13]. *Casuarina* isolates also differ in terms of the concentration of NaCl they can tolerate to initiate the symbiosis formation with their plant partner [65].

Given the diversity of the microsymbiont and the host in terms of salt tolerance, identification and molecular characterization of salt tolerant *Casuarina* species and associated *Frankia* are imperative for the successful utilization of *Casuarina* tress in saline soil reclamation efforts. This dissertation work is therefore pivoted around unlocking the molecular mechanisms of salt and osmotic stress tolerance in *Frankia* strains isolated from *Casuarina* trees. The effect of actinorhizal symbiosis on the salt tolerance of *Casuarina* trees is also examined. The goal of the research is discussed in detail in a separate subsection. Any attempt to understand the mechanism of salt tolerance in *Frankia* is incomplete without data mining for classical salt tolerance genes and pathways. A brief overview of classical salt and osmotic stress coping mechanisms employed by microorganisms exposed to a hyperosmotic environment is provided below.

Salt and osmotic stress tolerance of mechanisms of microorganisms

Microbes use a myriad of well-characterized mechanisms to adapt to fluctuations in osmolarity or salt. Classical salt/osmotic stress mechanisms employed by microorganisms include: (1) reestablishing osmotic balance by accumulating low molecular weight organic compatible solutes [79-81], (2) exclusion of Na^+ ion from cells via the action of a Na^+/ H^+ antiporter and Na^+ -ATPase [80], (3) altering membrane composition through changes in fatty acid saturation or phospholipid composition to better cope with the changed turgor pressure [82], (4) reactive oxygen species scavenging to prevent the oxidative degradation of lipids, also known as lipid peroxidation [80], and (5) restoration of the native folding of proteins through the actions of molecular chaperons [83].

Bacterial cells must maintain high turgor pressure to sustain cell volume and to allow for cell growth. Turgor pressure is established within cells according to the Morse equation, $P = RT\ (C_{in} - C_{out})$, where C_{in} is the osmolarity of the cytoplasm, C_{out} is the osmolarity of the extracellular medium, R is the gas constant, and T is the temperature [84]. The turgor pressure typically maintained is 3 -10 atmospheres for gram negative bacteria and about 20 atmospheres (ten times the pressure in a fully inflated automobile tire) for gram positive bacteria [85]. Under hyperosmotic environments, water fluxes out of the cell and the hydrostatic pressure exerted against the cell membrane is reduced, causing a threat to the cell's ability to maintain adequate turgor pressure. The immediate response of the bacteria upon losing water and turgor pressure is restoring rehydration by the rapid uptake of K^+, which is counterbalanced by endogenously synthesized glutamate [86]. Among the monovalent cations, K^+ perturbs cellular functions the least, but it nevertheless lacks the compatibility of organic osmolytes. Long term osmoadaptation response calls for the replacement of the K^+ with so-called organic osmolytes (compatible solutes). Compatible solutes encompass a restricted range of highly water soluble, osmotically active, low molecular weight amino acids and their derivatives, sugars or sugar alcohols, and other alcohols [87, 88]. Commonly employed compatible solutes include the sugar trehalose, the amino acids proline, serine and glutamate [89], quaternary ammonium compounds such as glycine betaine and proline betaine [90], polyamines, and organic solutes [89]. The strikingly limited number of compatible solutes used in all forms of life from bacteria to higher organisms is reflective of the challenge in finding solutes that are compatible with cellular functions. The osmotic function of a compatible solute depends on the degree of methylation and

length of the hydrocarbon chain [91]. Evolutionary selection on a compatible solute depends on the osmotic function as well as on other secondary functions such as its contribution towards heat and cold tolerance [92]. Accumulation of compatible solutes helps to avoid external osmolality-triggered water fluxes along the osmotic gradient causing either swelling in hypotonic environments or plasmolysis under hypertonic environments. The mechanism of action of compatible solutes involves changing the structure of the solvent and/or subtle changes in the dynamic properties of the protein as opposed to changing the structure of the protein itself. Osmotic adaptation using compatible solutes is characterized by a minimal requirement for genetic change and a high degree of flexibility in allowing organisms to adapt to wide ranges of external osmolarity [93].

Salt stress can upset the delicate balance between different cellular processes. The uncoupling of different pathways leads to the transfer of high energy electron to molecular oxygen (O_2), causing formation of reactive oxygen species, ROS [94]. ROS cause oxidative damage to proteins, DNA and lipids [95]. Oxidative stress results in the oxidative degradation of lipid membranes, also referred to as lipid peroxidation. The cell needs to be equipped with a mechanism of getting rid of the over 200 types of aldehydes, many of which are highly reactive and toxic, generated from lipid peroxidation [96]. Aldehyde dehydrogenases (ALDHs) counteract the effects of oxidative stress and the accompanying aldehyde production by metabolizing endogenous and exogenous aldehydes and converting them into TCA cycle intermediates.

Cells growing in high salt medium also face a loss of intracellular water, which creates a high ionic strength environment inside the cell. Proteins risk permanent unfolding in the resulting intracellular environment. Misfolding and aggregation of protein molecules are major threats to the survival of the cell. To cope with the problem, cells rely on a system composed of molecular chaperones and proteases [83, 97].

The deleterious effects of salt stress are first experienced by the cell membrane, which separates the interior of the cell from the outside environment [98]. Various kinds of stresses including, but not limited to, heat shock, cold shock, osmotic shock and salinity stress cause disruption of membranes, thereby affecting membrane-linked physiological processes such as transport, enzyme activities, and signal transduction. Maintaining correct fluidity of the bilayer over a wide range of salinity determines the extent of cell survival during salt stress [99]. Management of the lipid profile in response to salinity involves induction of fatty acid desaturases, which help to synthesize unsaturated fatty acids from saturated fatty acids [100]. The inherent salt tolerance of the organism dictates such responses and is a key factor accounting for the disparity in salt tolerance between organisms [101]. In addition to the common mechanisms outlined above, bacteria employ various unique mechanisms to adapt to high salt stress.

Research background and objectives

The primary goal of the research was to determine the role of actinorhizal symbiosis in the salt stress tolerance of *Casuarina* plants and to unlock the molecular mechanisms of salt/osmotic stress tolerance of the microsymbiont. *Casuarina* isolates show marked differences in salt and osmotic stress tolerance. Given the availability of

more than three dozen *Frankia* genomes, several of which represent strains isolated from *Casuarina* trees, it is of great interest to the scientific community to tie the observed difference in the salt tolerance of *Casuarina* isolates to their genetic make-up. Understanding of the molecular mechanisms of salt/osmotic stress tolerance in *Frankia* allows for the development of strains that combine the best symbiotic and salt tolerance attributes. We have taken a comprehensive approach to address the research question. The salt stress tolerance levels for several *Frankia* strains isolated from *Casuarina* trees were assayed and two salt-stress tolerant strains (*Frankia* sp. strain CcI6 and Allo2) which could withstand up to 1000 mM NaCl and one relatively salt-sensitive strain (*Frankia casuarinae* strain CcI3), which could withstand only 475 mM NaCl were identified. Comparative genomics were used to identify strain-specific genes and pathways potentially involved in salt and osmotic stress tolerance. Transcriptome and proteomic profiling of the salt-tolerant and salt-sensitive strains was carried out to identify strain-specific genes that are responsive to salt and osmotic stress. Physiological analysis was undertaken to confirm the accumulation of osmolytes predicted based on the transcriptome and proteome analysis. An expression vector that could stably replicate in *Frankia* was identified. Cloning of some of the salt-responsive genes from the salt-tolerant-strain (*Frankia* sp. strain CcI6) and expressing it in the salt-sensitive-strain (*Frankia casuarinae* strain CcI3), revealed novel genes that confer salt stress tolerance.

The research was part of an international collaborative endeavor to unlock the mechanism of salt and osmotic stress tolerance in *Casuarina glauca* and its micro-symbiont *Frankia* using new sequencing technologies, proteomics, mutational analysis,

and physiological approaches. Collaborators included: UNH (USA), LCM IRD (Senegal), IRD (France), and IFGTB (India). While our work focused on the salt tolerance mechanisms of the microsymbiont, the work of our collaborators was focused on demystifying the salt tolerance mechanism of *Casuarina glauca* under symbiotic and non-symbiotic conditions and assessing the performance of the symbiosis under salt-stressed conditions in semi-axenic and field conditions.

Materials and Methods

Frankia Growth Media and Culture Conditions

Frankia strains used in this study are listed in Table 1. *Frankia* stock cultures were grown and maintained in basal MP [102, 103] growth medium supplemented with 5.0 mM NH_4CI and the appropriate carbon source (Table 1). Basal MP growth medium consisted of MOPS-phosphate buffer (50 mM MOPS, 10 mM K_2HPO_4, pH 6.8) supplemented with 1 mM Na_2MoO_4, 2 mM $MgSO_4$, 20 µM $FeCl_3$ with 100 M nitrilotriacetic acid (NTA), and modified trace salts solution [102]. For experimental conditions, *Frankia* cultures were grown in MP or BAP growth media [104] as described previously. The composition of BAP growth medium is: 0.01 M KH_2PO_4, 0.01 M K_2HPO_4, 5 mM propionate, 0.135 mM $CaCl_2$, , 0.045 mM $MgSO_4*7H_2O$, 0.1% (v/v) Fe-EDTA stock (195 mM $FeNa_2EDTA$), 0.1% (v/v) Oligoelements (0.115 mM H_3BO_3, 0.08 µM $CUSO_4*5H_2O$, 2.3 µM $MnCl_2*4H_2O$, 0.19 µM $ZnSO_4$, 0.026 µM $Na_2MoO_4*2H_2O$), 0.1% (v/v) Wolff's vitamins (0.059 mM pyridoxine HCl, 0.036 mM ρ-aminobenzoic acid, 0.024 mM lipoic acid, 0.04 mM nicotinic acid, 0.013 mM riboflavin, 0.016 mM thiamine HCl, 0.01 mM calcium DL-pantothenate, 0.008 mM biotin, 0.004 mM folic acid, and 0.0737 µM vitamin B12), pH 6.7 supplemented with 2% (v/v) Mes-Tris (0.5 M 2[N-

Morpholino] ethan sulfonic acid, adjusted to pH 6.8 with Tris). For growth under nitrogen-deficient conditions, N_2 was the sole nitrogen source.

Table 3. *Frankia* strains used in this study

***Frankia* strain**	**Carbon source**	**Relevant phenotype**	**Source/Reference**
Frankia casuarinae strain CcI3	5 mM propionate	Kanamycin resistant (50 µg/ml)	[105, 106]
Frankia sp. strain CcI6	5 mM propionate	Kanamycin resistant (50 µg/ml)	[107]
Frankia sp. strain Allo2	5 mM propionate	Kanamycin resistant (50 µg/ml)	[108]
Frankia sp. strain CeD	5 mM propionate		[109]
Frankia sp. strain BR	5 mM propionate		[110]
Frankia sp. strain BMG5.23	5 mM propionate		[111]
Frankia sp. strain Thr	5 mM propionate	Kanamycin resistant (50 µg/ml)	[112]
Frankia sp. strain CgI82	5 mM propionate		Lab stock
Frankia alni strain ACN14a	20 mM succinate		[105, 113]
Frankia sp. strain EAN1pec	20 mM succinate		[114]
Frankia sp. strain DC12	20 mM glucose		[115, 116]
Frankia inefficax strain EUI1c	20 mM glucose		[117, 118]
Frankia casuarinae strain CcI3/pHTK1	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-22605	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-13590	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-17915	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-21730	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-19875	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-20860	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study
Frankia casuarinae strain CcI3/pHTK1-RO-17580	5 mM propionate	Kanamycin resistant (50 µg/ml) Tetracycline resistant (30 µg/ml)	This study

Table 4. *E. coli* strains used in this study

***E. coli* strain**	**Relevant genotype or description**	**Source or reference (s)**
E. coli DH5α	F– Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) *U169 rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *supE44* λ– *thi-1 gyrA96 relA1*	Lab stock
E. coli DH5αλpir	F– Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) *U169 rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *supE44* λ– *thi-1 gyrA96 relA1*/*pir*	Lab stock
E. coli BW29427	*dap* auxotroph, *tra pir*	Lab stock
E. coli S17-1	*pro thi hsdR*$^+$ Tp^r Sm^r; chromosome::RP4-2 Tc::Mu-Kan::Tn7	Lab stock
E. coli S17-1λpir	*pro thi hsdR*$^+$ Tp^r Sm^r; chromosome::RP4-2 Tc::Mu-Kan::Tn7/λpir	Lab stock
E. coli HB101	F$^-$ *Δ(gpt-proA)62leuB6glnV44 ara-14 galK2 lacY1 Δ(mcrC-mrr) rpsL20* (Str^R) *xyl-5 mtl-1 recA13 thi-1*	Lab stock
E. coli Top10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen (Carlsbad, CA)
E. coli BW25113	K-12 derivative, *araBAD rhaBAD*	[119]
E. coli ET12567	*dam dcm hsdS cat tet*	[120]

Plant growth media

For germination and growth before transplanting, plants were grown in commercially available 1/4 strength Hoagland's Modified Basal Salt Solution (1/4 HS; MP Biomedicals, Solon, Ohio) containing 0.25 mM $(NH_4)_3PO_4$, 115 µM H_3BO_3, 1 mM $Ca(NO_3)_2$, 0.08 µM $CuSO_4*5H_2O$, 22.53 µM Na_2EDTA, 22.5 µM $FeSO_4*7H_2O$, 0.5 mM $MgSO_4$, 2.3 µM $MnCl_2*4H_2O$, 0.0275 µM MoO_3, 1.5 mM KNO_3, 0.19 µM $ZnSO_4*7H_2O$, pH 5.5. After transplanting, plants were grown in BD medium [121]. The composition of the BD medium is: 1000 µM $CaCl_2$, 250 µM $MgSO_4$, 500 µM KH_2PO_4, 10 µM Fe EDTA, 250 µM K_2SO_4, 1 µM $MnSO_4$, 0.5 µM $ZnSO_4$, 0.2 µM $CuSO_4$, 0.1 µM $CoSO_4$, and 0.1 µM Na_2MoO_4. For nitrogen-sufficient conditions, plants were grown in BD medium supplemented with 5 mM KNO_3.

Seed Sterilization and Plant Growth Conditions

Casuarina cunninghamiana seeds used in this study were obtained from Sheffield's Seed Company, Locke, New York, or F. W. Schumacher Company, Sandwich, Massachusetts. Before use, seeds were soaked in sterile tap water overnight at room temperature (RT). After decanting the water, seeds were surface-sterilized by suspending in 15 mL of fresh 30% hydrogen peroxide containing two drops of Tween 20 and agitating at room temperature for 5 minutes. The sterilized seeds were washed six times with 10 ml sterilized deionized H_2O (sdH_2O) to remove the hydrogen peroxide and the detergent. Sterilized seeds were aseptically sown on 100 ml perlite (Scotts Company, Marysville, Ohio) wetted to field capacity with 45 ml of ¼ HS medium in Magenta GA-7 boxes (Magenta Corp., Chicago, Illinois). Germination took place at 25°C with a 16 h light period and 8 h dark period.

Salt sensitivity assay for *Frankia*

The salt tolerance levels of *Frankia* strains were determined by measuring the total cellular protein content and/or total cellular dry weight after growth under salt or osmotic stress. For total cellular protein determination, a 24-well growth assay was used as described previously [78]. Briefly, cells were grown in propionate basal medium with or without 5 mM NH_4Cl containing different concentrations of NaCl or sucrose. For strains ACN14a, EuI1c, DC12 and EAN1pec, propionate was replaced with the appropriate carbon source [Table 1]. The inoculum was adjusted to 40 µg/ml of total protein and the plates were incubated at 28°C for 14 days. Growth was measured by total cellular protein content as described below. Growth yield was determined by subtracting the protein content of the inoculum.

For total cellular dry weight determination, *Frankia* strains were inoculated into 25 ml of basal growth medium containing different concentrations of NaCl or sucrose. The inoculum concentration was adjusted to 40 µg/ml protein. The *Frankia* cells were grown for 14 days at 28 °C. Growth was measured by total cellular dry weight as described below. Growth yield was determined by subtracting the dry weight of the inoculum.

To evaluate the levels of tolerance, the following two parameters were used: maximum tolerable concentration (MTC) and minimum inhibitory concentration (MIC). The MTC is the highest concentration of salt which does not affect the growth, while MIC is the lowest concentration of salt that inhibits growth.

Protein content and dry weight determination

Total cellular protein content was measured by the bicinchoninic acid (BCA) method [122] per the manufacturer's specifications (Pierce, Rockford, lL, USA) and

bovine serum albumin was used as a standard. Cells solubilized in 1 N NaOH were boiled at 95°C for 10 minutes and centrifuged at 13,000g for 5 min. The supernatant containing solubilized proteins was used for quantification. Triplicate measurements were made for each sample. Total cellular dry weight was determined using tared polycarbonate membranes [123].

Vesicle induction and nitrogenase activity

To determine vesicle production and nitrogenase activity, *Frankia* cells were harvested after 14 days of growth in medium supplemented with 5 mM NH_4Cl and washed three times with buffer containing 20 mM morpholinepropanesulphonic acid (MOPS) and 10 mM KH_2PO_4 at pH 6.8. The washed cells were inoculated into growth medium lacking an external nitrogen source and containing different concentrations of NaCl or sucrose. The cultures were incubated at 30°C for 4 days. The vesicle numbers were determined as described previously [124]. The activity of the nitrogenase enzyme was determined by the acetylene reduction assay as described previously [124].

Effect of salt treatment on the nodulation of *Casuarina cunninghamiana*

Two weeks after germination, five *Casuarina cunninghamiana* seedlings were transplanted into Magenta boxes containing 50 ml of BD medium supplemented with KNO_3 (5 mM) as a nitrogen source at pH 6.7 [121] and were incubated for one month in a growth chamber at 25°C with a 16 h light period. After a month of growth, different levels of salt stress were applied to the plants. Final concentrations of NaCl tested were 0, 50, 100, 150, and 200 mM. For concentrations above 50 mM, the stress was applied gradually through the weekly increment of 50 mM concentration of NaCl. One week after the desired NaCl concentrations were reached for all treatments, the plants were

inoculated with either *Frankia* sp. strain CcI6 or *Frankia casuarinae* strain CcI3. Control plants were not inoculated with *Frankia*. Prior to inoculation with *Frankia*, the plants were incubated in BD medium without supplemental nitrogen for two weeks. To prepare *Frankia* for inoculation, two week old *Frankia* culture grown in MP medium supplemented with 5 mM NH_4Cl and 5 mM propionate was harvested and washed twice with 20 mL of BD medium to remove all traces of nitrogen. The washed cells were resuspended in 50 ml of BD medium and incubated at 28°C for one week. The *Frankia* culture was harvested and adjusted to 60 µg/ml protein in fresh BD medium containing the appropriate NaCl concentration. Spent plant growth medium was decanted from each Magenta box and replaced with 50 ml of the *Frankia* suspension. The inoculated plants were incubated at 25°C with 16 h light period and plant growth medium was replaced weekly to avoid nutrient depletion and pH drift. For two months after inoculation, frequency of nodulation and number of nodules per plant were followed. Each treatment condition consisted of 15 plants. The experiment was laid out in a 3 X 5 factorial randomized complete block design (RCBD). Statistical analysis was completed using ANOVA on JMP software (JMP, Cary, North Carolina).

The effect of symbiosis with *Frankia* on the salt tolerance of *Casuarina cunninghamiana*

Two weeks after germination, *Casuarina cunninghamiana* seedlings were transplanted into Magenta boxes containing 50 ml of BD medium supplemented with KNO_3 (5 mM) as a nitrogen source at pH 6.7 [121] and were incubated for one month in a growth chamber at 25°C with a 16 h light period. Subsequently, the plants were inoculated with *Frankia casuarinae* strain CcI3 or *Frankia* sp. strain CcI6 as described above. Control

plants were not inoculated with *Frankia*. After one month of growth, different levels of salt stress were applied to the plants. Final concentrations of NaCl tested were 0, 50, 100, 150, and 200 mM. Each treatment condition consisted of 15 plants. For concentrations above 50 mM, the stress was applied gradually through the weekly increment of 50 mM concentration of NaCl. After the desired NaCl concentration was achieved, the plants were allowed to grow for additional two months with a weekly replacement of the depleted nutrient medium. Growth measurements were taken on plant height and root and shoot dry weight. For root and shoot dry weight determination, plants were harvested, the respective parts were cut and washed in deionized water, surface-wiped with blotting paper and dried at 80°C for 48 hours before measurement.

Genome sequencing of *Frankia* strains

High-quality *Frankia* gDNA for Illumina sequencing was extracted using the cetyltrimethylammonium bromide (CTAB) method [125]. Briefly, 6 ml of seven day old *Frankia* culture was harvested by centrifugation at 16,000 x g for 10 min and resuspended in 567 µl of TE buffer [10 mM Tris-HCl, pH 8.0, 1 mM Ethylenediaminetetraacetic acid (EDTA)]. The resuspended cells were heated at 80°C for 20 min. To degrade the peptidoglycan layer, 11 µl of 50 mg/ml of lysozyme was added to the cells and incubated at 37 for 60 min. Thirty microliters of 10% sodium dodecyl sulfate (SDS) and 5 µl of 20 mg/ml proteinase K were added to the cells, mixed gently, and incubated at 65°C for 20 min. One hundred microliters of 5 M NaCl and 80 µl of CTAB/NaCl (10% CTAB, 700 mM NaCl) solution were added to the mixture, mixed by inversion, and incubated at 65°C for 10 min. An equal volume of chloroform:isoamyl alcohol (24:1) was added to the mixture and centrifuged at 16,000 x g for 10 min.

Isopropanol (800 µl) was added to the aqueous phase to precipitate the gDNA at -20°C overnight. The gDNA was collected by centrifugation at 16,000 x g for 20 min at 4°C. The pelleted gDNA was washed with 1 ml 70% ethanol. After drying, the pelleted gDNA was dissolved in 50 µl of nuclease free water (or TE buffer). The DNA was quantified using the Nanodrop 2000c spectrophotometer (Thermo Scientific, Wilmington, Delaware) and the Qubit dsDNA BR Assay (Invitorgen, Carlsbad, California). The integrity of the gDNA was verified by electrophoresis on a 1% agarose gel and checking for a band of about 40 kb. The gDNA sample was sent to the Hubbard Genome Center at the University of New Hampshire to generate standard Illumina shotgun library. Sequencing was undertaken using the Illumina HiSeq 2000 platform. All raw Illumina sequence data was processed through CLC Genomics Workbench to trim on quality and adapters. Adapter trimming was carried out by searching for adapter sequences on the forward and the reverse strands. The ends of reads were also trimmed on quality using the prediction reliability calculated by the base-caller on a limit of 0.05. Reads were further trimmed based on the presence of 2 or more ambiguous characters. Assembly of the genomes was undertaken via an integrated approach utilizing CLC Genomics Workbench and ALLPATHS-LG [126]. Briefly, the assembly process encompassed the following steps. Illumina reads that passed the quality control and the trimming steps were *de novo* assembled using CLC Genomics Workbench (the different versions of CLC used with the different genomes are provided in Table 3). A sequence read simulation tool called wgsim was used to generate 3 million, 150 bp simulated paired-end reads (with 1-3 kb distance between the outer end of the pairs) using contigs created by CLC Genomics Workbench as inputs. Trimmed Illumina reads along with the

simulated paired reads generated by CLC Genomics Workbench were assembled using ALLPATHS-LG [the versions used are given in table 3]. The following parameters were used for *de novo* assembly by CLC Genomics Workbench: for the de Bruijn Graph, word size and bubble size were calculated automatically. The minimum effective contig length was set at 200 bp (when 100 bp short reads are assembled) or at 300 bp (when 150 bp short reads where assembled). Paired distance for the paired end sequence data was calculated by the assembly program and scaffolding was performed. For mapping the reads back to the assembled contigs, the following default parameters were used: mismatch cost (2), Insertion cost (3), Deletion cost (3), Length fraction (0.5), and similarity fraction (0.8). For generating simulated reads using contigs assembled by CLC Genomics Workbench as inputs, the following wgsim code was used: wgsim -e 0 -d -N -1 150 -2 150 -r 0 -R 0 -X 0. In the code, d represented the distance between the outer ends of the paired reads, N represented the number of simulated read pairs to be generated, -1 represents the length of the first read, -2 represents the length of the second read, -e represents the base error rate, -R represents the fraction of indels, -X represents the probability that an indel is extended. The assembled genomes were annotated via the Integrated Microbial Genomes (IMG) platform developed by the Joint Genome Institute, Walnut Creek, CA [127]. Annotations were manually checked to ensure conformity to NCBI standards and the draft genomes were submitted to the NCBI database.

Table 3. Different versions of CLC and ALLPATHS-LG assembly programs used for the different *Casuarina* isolates sequenced by Tisa lab, UNH, USA.

Frankia strain	Version of CLC Genomics Workbench used for assembly	Version of ALLPATHS-LG	Illumina libraries generated	References
Frankia sp. strain CcI6	6.5.1	r41043	Nextera	[38]
Frankia sp. strain CeD	8.0.1	r41043	Nextera	[39]
Frankia sp. strain Allo2	8.0.1	r41043	Nextera	[41]
Frankia sp. strain BMG5.23	6.5.1	r41043	Nextera	[36]
Frankia sp. strain Br	8.5.0	r41043	Nextera	[40]
Frankia sp. strain Thr	6.5.1	r41043	Nextera	[37]

Accession numbers

For bioinformatics analysis, genome sequences and their annotations including amino acid and nucleotide FASTA files were obtained from the NCBI database (http://www.ncbi.nlm.nih.gov) under GenBank accession numbers [NZ_AYTZ00000000.1, NZ_LRTJ00000000.1, NZ_JENI00000000.1, NZ_JPGU00000000.1, NZ_JDWE00000000.1, NZ_JPHT00000000.1, NC_007777.1, NC_008278.1, NC_008578.1].

Pan genome analysis

The web platform OrthoVenn [128] was used to identify orthologous gene clusters. OrthoVenn uses a modified version of the heuristic approach named OrthoMCL [129] to identify ortholog groups. An E-value cut off of $1e^{-5}$ was used for all-to-all protein similarity comparisons. An inflation value of 1.5 was used for the generation of orthologous clusters using the Markov Cluster Algorithm [130]. To determine single copy orthologs among the most tolerant, moderately tolerant and most

sensitive strains, a modified Lerat method was used [131]. As a stringent criterion for homology, only gene pairs representing a bit score value equal to or higher than 30% of the maximal possible bit score value were considered homologous genes. VennDiagram in R was used to construct the four-way Venn diagram between salt tolerant (2 strains), moderately salt tolerant (1 strain), and the relatively salt sensitive strain.

Average nucleotide identity, average amino acid identity and average genomic distance

The average nucleotide identity (ANI) and average amino acid identity (AAI) between strains was estimated by using reciprocal best hits (two-way ANI or two-way AAI), as previously described [132]. Genome to genome distance was calculated using the web platform called GGDC 2.1 according to the standard operating procedure previously described [133]. GGDC 2.1 BLAST+ was chosen as the alignment method for finding intergenomic matches.

Phylogenetic Analysis

A concatenated maximum parsimony phylogenetic tree was generated from 394 conserved single copy pan-orthologous genes determined by a modified Lerat method [131]. The rationale for including only single-copy genes representing species divergences was to minimize potential errors caused by gene duplication. The tree was calculated by determining the ratio of the bit score to the maximal bit score (i.e., protein match against itself). As a stringent criterion for homology, two genes are considered homologous only if the bit score value for the pair is at least 30% of the maximal bit

score. The 30% cutoff maximized the number of families containing genes, and is optimal for the interspecific identification of homologous sequences [134].

RNA-Seq sample preparation and data analysis

To analyze gene expression of the salt-tolerant strain under salt stressed conditions, RNA-Seq analysis was performed on *Frankia casuarinae* strain CcI3 (a salt-sensitive strain), *Frankia* sp. strain CeD (a moderately salt-tolerant strain), and *Frankia* sp. strain CcI6 (a salt-tolerant strain). Cultures were grown for 7 days at 28°C in a 5 mM propionate, 5 mM NH_4Cl basal growth medium [117] alone or supplemented with 200 mM NaCl or sucrose. The bacteria were harvested and the pellets were frozen at -80°C. Total RNA was extracted using a modified RNeasy Midi kit (Qiagen Sciences, Valencia, CA). Frozen bacterial pellets were resuspended in 0.5 ml TE buffer, pH 8, supplemented with 5 mg/ml lysozyme and incubated at room temperature for 10 minutes. RLT buffer (2 ml) supplemented with 1µl/ml β-Mercaptoethanol (β-ME) was added to each sample and the pellets were homogenized using glass tissue grinders. Subsequently, the RNeasy midi kit procedure was followed as per the manufacturer's recommendation with one major modification: after addition of ethanol to the lysate, the RNeasy mini kit procedure, instead of the RNeasy midi kit procedure, was used. RNA samples were treated with DNase I (New England Biolabs, Ipswich, Massachusetts) per the manufacturer's instructions. RNA was quantified using Qubit RNA assay (Invitrogen) and Nanodrop 2000c spectrophotometer (Thermo Scientific, Wilmington, Delaware) according to manufacturers' specifications. The quality of each RNA sample was determined using the Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA) according to the Prokaryote Total RNA Nano protocol. RNA quality was represented by RNA integrity

number (RIN value), which ranged from 1 to 10, with 10 representing the most intact RNA. Samples with RIN value greater than or equal to 8 were used for downstream analysis. Ribosomal RNA was removed from 2-4 ng of total RNA by the use of the MicrobeExpress kit (Ambion, Foster City, CA) according to the manufacturer's specifications. The MEGAclear kit (Life Technologies, Carlsbad, CA) was used to remove tRNA according to manufacturer's specifications. cDNA libraries were prepared using the TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA) as described by the manufacturer. The cDNA library was verified for appropriate fragment size (approximately 250 bp) on an Agilent 2100 Bioanalyzer (Agilent, Santa Clara,CA) according to the DNA 1000 protocol described by the manufacturer. The Qubit dsDNA BR Assay (Invitorgen, Carlsbad, California) was used to determine the cDNA concentration of each library according to manufacturer's recommendations. Libraries were normalized to 10 nM with 10 mM Tris-HCl, pH 8.5, supplemented with 0.1% Tween 20. Illumina sequencing was carried out at Hubbard Genome Center at the University of New Hampshire on an Illumina HiSeq 2500 platform. Reads were separated on adapter assignment and pre-processed through CASAVA 1.8.3. The resulting FASTQ files of sequence reads were processed using CLC Genomics Workbench 9.0 (CLC bio, Cambridge, MD). Adapters were trimmed from reads by searching on the forward and reverse strands. The ends of reads were quality trimmed based on quality scores from a base-caller algorithm using a limit value of 0.05. The high quality trimmed reads were mapped to Frankia sp. strain CeD, *Frankia* casuarinae strain CcI3 or *Frankia* sp. strain CcI6 gene regions. Reads mapping to rRNA operons were excluded from downstream analysis. Mapping parameters were as follows: The

maximum number of mismatches allowed was 2. The minimum length fraction was set so that at least 50% of the read length aligns to the reference sequence. The minimum fraction of identity between the read and the reference sequence was set at 80%. A read that matched to more than 10 distinct places in the reference was not mapped. If the read matched to multiple distinct places, but below 10 different locations, it was randomly assigned to one of the distinct places. After mapping, the gene expression level for each gene was expressed in terms of the unique number of reads mapping to that gene. All RNAseq experiments were normalized by the total number of reads in each library. A gene was expressed if it had at least one unique sequence read aligned with it. To determine differential gene expression, statistical analysis on proportions was carried out. Two-sided p-value for multiple biological replicates was computed using Baggerley's test [135].

Proteome analysis of salt-stressed *Frankia*

Cultures were grown for 7 days at 28°C in 5 mM propionate, 5 mM NH_4Cl basal growth medium [102] alone or supplemented with 200 mM NaCl or sucrose. *Frankia* mycelium from 50 mL culture was harvested by centrifugation at 3,400 x g for 20 min and resuspended in 2 mL of lysis buffer [10 mM Tris–HCl (pH 7.4), 1 mg/mL $MgCl_2$, 50 µg/mL DNase, 50 µg/mL RNAse, and 50 µg/mL lysozyme]. A French pressure cell was used to lyse the cells at 137,895 kPa. Lysed samples were centrifuged at 16,200 x g and the supernatant fluid containing soluble proteins was collected. Protein samples were quantified using the Bradford assay [136]. One milligram of soluble protein was precipitated with 10% (v/v) trichloroacetic acid in acetone solution containing 20 mM dithiothreitol (DTT) overnight at -20°C. The samples were centrifuged at 16,200 x g for

30 min at 4°C to pellet the proteins. Traces of TCA were removed from the pellet by washing with 20 mM DTT in acetone. The protein pellet was dissolved in 300 µL of rehydration buffer [7 M urea, 2 M thiourea, 5% (*m/v*) DTT, 2% (*v/v*) Triton X-100, 2% (*v/v*) immobiline pH gradient (IPG) buffer, 0.02% bromophenol blue]. The solubilized proteins were used to rehydrate an 11 cm Immobiline™ DryStrip pH gradient strip (pH 4–7) (GE Healthcare Biosciences, Pittsburgh, Pennsylvania). The protein sample suspended in rehydration buffer was evenly distributed along the lanes of a strip holder and the Immobiline strips were rehydrated with the gel side facing down. DryStrip cover fluid (GE Healthcare Biosciences) was used to cover the strips to prevent evaporation and crystallization of the urea. Rehydration proceeded for 12 h. The rehydrated strips were moved to an Isoletric focusing (IEF) tray and were positioned with the gel side facing down. Wet ProteomIQ™ IPG Wicks (Proteome Systems, Woburn, Massachusetts) were placed at the anode and the cathode to collect excess salt and other contaminants. The strips and the wicks were covered with 50 mL of DryStrip cover fluid. The IEF tray was placed in an IsoelectrIQ 2 unit (Proteome Systems, Woburn, Massachusetts) and isoelectric focusing was performed under the following settings: 100 – 10,000 V gradient for 8 hours and 10,000 V constant for 8 hours. Strips were stored gel side up at -80°C until the second-dimension protein separation step. Strips were washed in 1 x sodium dodecyl sulfate (SDS) running buffer [0.2 M glycine, 25 mM Tris-base, 0.1% (m/v) SDS] and reduction of the proteins was undertaken by incubating the strips in SDS equilibration buffer [50 mM Tris–HCl (pH 8.8), 6 M urea, 30% (v/v) glycerol, 2% (m/v) SDS, trace bromophenol blue] supplemented with 65 mM DTT for 20 min. Alkylation reaction was performed by washing the strips in deionized

water followed by incubation in the same SDS equilibration buffer, supplemented with 135 mM iodoacetamide (IAA), instead of DTT, for 20 min. The strips were rinsed in 1x SDS running buffer and loaded on a 12% SDS-Polyacrylamide gel with the plastic backing against one of the glass plates. Strips were sealed in place with 1% (m/v) agarose in TAE buffer (40 mM Tris-Acetate, 1 mM EDTA, pH 8) supplemented with trace amounts of bromophenol blue for tracking purposes. Proteins were separated by electrophoresis at 100 V for 30 min followed by 200 V for 5 h. Protein spots were visualized by incubating the gels in Coomassie Blue stain [0.1% Coomassie Brilliant Blue R-250, 10% (v/v) glacial acetic acid, 50% (v/v) methanol] followed by destaining with a destain solution [10% (v/v) ethanol, 5% (v/v) glacial acetic acid]. Differentially expressed spots were excised from the gel and placed in a 0.5 mL Eppendorf tube. Gel pieces were washed three times by adding fresh 50 µL of 25 mM NH_4HCO_3/50% acetonitrile (ACN) each time and vortexing for 15 min. The gel pieces were incubated in 25 µL DTT solution (10 mM DTT in 25 mM NH_4HCO_3) at 56°C for 1 h. The supernatant was discarded and 100 µL of IAA solution (55 mM IAA in 25 mM NH_4HCO_3) was added to the samples. The samples were incubated at room temperature in the dark for 45 minutes. The supernatant was discarded and samples were washed by adding 100 µl of 25 mM NH_4HCO_3 and vortexing for 10 minutes. The supernatant was discarded and gel pieces were dehydrated by adding 100 µl of 25 mM NH_4HCO_3/50% ACN solution and vortexing for 10 minutes. The dehydration step was repeated twice. Gel pieces were completely dried under a speed vacuum for 20 min. Five microliters of trypsin solution (40 ng/µL in 25 mM NH_4HCO_3) was added to the gel pieces. An additional 30 µl of 25 mM NH_4HCO_3 was added to cover the gel pieces. The trypsin digestion took place at

37°C for 4 h. The supernatant from the digestion was transferred to a new tube with 5 µl of extraction buffer [50% (v/v) ACN/3%(v/v) acetic acid]. Gel pieces were extracted twice by adding 35 µl of extraction solution and vortexing for 20 min. Extracts were combined and dried in a vacuum centrifuge. Extracted peptides were resuspended in 7 µl of 0.1% formic acid. Samples were analyzed using liquid chromatography – mass spectrometry (LC–MS) and LC–MS/MS analysis. Aliquot of the digestion mixture (1 µL) was used for LC separation using the Ultimate 3000 RSLCnano UHPLC system with an autosampler (Dionex Corporation, Sunnyvale, California). The eluent was ionized by a nanoelectrospray ionization source of an LTQ Orbitrap XL mass spectrometer (Thermo Scientific, Waltham, Massachusetts). LC–MS data were acquired in an information-dependent acquisition mode, cycling between a MS scan (m/z 310–2000) acquired in the Orbitrap, followed by low-energy collision-induced dissociation (CID) analysis in the linear ion trap. The centroid peak lists of the CID spectra were generated by PAVA [137] and searched against a database that consisted of the National Center for Biotechnology Information (NCBI) protein database, to which a randomized version had been concatenated, using Batch-Tag, a program in the University of California-San Francisco Protein Prospector version 5.10.15. A precursor mass tolerance of 15 ppm and a fragment mass tolerance of 0.5 Da were used for the protein database search. Protein hits were reported with the following parameters: a Protein Prospector protein score of ≥22, peptide score ≥15, and E value for protein ≤0.01 [138]. This set of protein identification parameters threshold did not return any substantial false-positive protein hits from the randomized half of the concatenated database. Test samples were compared with corresponding control samples using the Search Compare program.

Unique and overexpressed proteins found in the test samples were noted and subjected to genetic studies previously mentioned.

Quantitative reverse transcription PCR (qRT-PCR)

The same RNA samples used for RNA sequencing were also used for qRT-PCR validation of the RNA-Seq results. The RNA (400 ng) was transcribed into cDNA using the GoScript™ Reverse Transcriptase (Promega, Madison, Wisconsin) according to the manufacturer's instructions. The cDNA was quantified by a Nanodrop 2000c spectrophotometer, diluted to 10 ng/µL working stocks in RNase-free H2O, and stored at –20°C. Amplification and detection of gene expression were performed using Agilent MP3000 qPCR system (Agilent Technologies, Santa Clara, California). The primers used for these experiments are listed in Table 4. Each primer sequence was blasted against the respective *Frankia* genome to ensure specificity to the target gene. Standard curves were generated using the genomic DNA and each primer set to test primer efficiency before use. Primer sets with efficiency values 85%- 115% were used for the experiments. The *rpsO* (CCI6_RS04555) gene was used as the normalizer for all qRT-PCR experiments. The qRT-PCRs were done using 50 ng template cDNA, primer mix (0.3 µM), and SYBR Green PCR Master Mix (Applied Biosystems, Carlsbad, California) in a 25 µL total reaction volume. The following thermal cycler parameters were used: (i) 95 °C for 15 min; (ii) 40 cycles of 95 °C for 15 s and 60 °C for 30 s; and (iii) 1 thermal disassociation cycle of 95 °C for 60 s, 55 °C for 30 s, and incremental increases in temperature to 95 °C for 30 s. Reactions were performed in triplicate. The ΔΔCt method [139] was used to calculate relative expression (fold changes).

Table 4. Primer sets used in this study.

Locus tag	Forward primer	Reverse primer
Primers used for qPCR validation of RNAseq data		
CCI6_RS21730	TGC ACT TCT ATC GGC AAC C	GAA GTA CTG CGA CAG GAA GAA G
CCI6_RS17915	TAG AGT TCC GTC CAG GTC TT	GGA CCG CAC AAC AGT CTT TA
CCI6_RS19875	ACA CTC AAC GCA CGA ATC A	GTG TTG ATG CGG GTT ATC ATT TC
CCI6_RS17580	AAG AAC TGG CGA ATC CTC AC	CGG ATT GCT GTT CGT TGA TTT
CCI6_RS01600	CGA CAT CAA GAT CGA CCA CTA C	TTG GAC TTT CCG CCG TTT
CCI6_RS18570	CCG GCA CTT CAC CTT CAT	CCG GAA GTG CGC GAT AA
CCI6_RS06495	CGT CGC AAC CTC TAC ATC TAC	CGG GAT GAA CTG GAT GAC AA
CCI6_RS02325	CAA CGG GCA GGT GAT CTA TT	GAA TCC GTC AAC ACG CTC T
CCI6_RS12340	GCA GAA CCA GCT CTT CCC	AAC GGC TGGAAC CAG AAC
CCI6_RS08505	GGG TGA AGG GTG ATC CTT ATG	GTT GAT CAT GGA TGG CAG GTA
CCI6_RS19950	ATA CGC TTC TGC TCG TGA AC	CCG GCA CGA TCT GTG TAA ATA
Primers used for amplifying *pR, pL, rrnbT1* containing region from pASV2		
PR_PL_rrnbT1	ACGTGAGCTCCCAGGCATCAAATAAAACG	ATCAGGTACCACTAGTATTAGGGCCCTGAAAAGTTCTTCTCCTTTACTC
For cloning into pHTK1-RO		
CCI6_RS22605	ATATGGGCCCCAGACGAGTCGGCGGCGACCA	TCATACTAGTAAGGCCTCGGCCCTCGCGAGA
CCI6_RS17915	ACATACTAGTAACGAGGCTCGGCATGGACA	ATAT ACTAGTAATCAGCTCGCCGTCGGTGA
CCI6_RS21730	ACTAGTGATCGTCACCTGCGCTCAGCT	ACTTGGTACCTCGACATTACCGCCAGCCTCA
CCI6_RS13590	ACTAGGGCCCTATAAATGCGACGCCACCAA	ATCTACTAGT TTGTGGCCACGACCCGCATA
CCI6_RS19875	ATATGGGCCCACCCGCATCCGCCGGCTGAT	ATATACTAGTTTATGGTGACGTCTTCGTTA
CCI6_RS20860	ATATGGGCCCGCGGCGACCGGTTTGCTCGT	TATTACTAGTGCGACCACTCGTCGCCGGTAA
CCI6_RS17580	ATATGGGCCCCCCGGTCGGCTTCACGTCGCT	TTAAACTAGTGGTTCTTCGCCTGGTACAAC

Amino Acid analysis

Approximately 40 mg of fresh *Frankia* mycelia were suspended in 800 µl of 5% (v/v) ice-cold perchloric acid (PCA) and stored at -20°C until needed. Before analysis, the frozen samples were lysed by thawing and freezing three times. The samples were pelleted by centrifugation at 13500 x g for 10 min and the supernatant extract was used for dansylation of both amino acids and polyamines according to the procedure described previously [140] with a few modifications. Briefly, 100 µl of saturated sodium carbonate solution and 100 µl of dansyl chloride (20 mg/ml in acetone) were added to microfuge tubes containing 20 µl aliquots of a mix of two internal standards (0.1 mM heplanediamine for polyamines and 1 mM α-methyl-DL-phenylalanine for amino acids in 5% PCA) and 100 µl aliquots of sample extracts or 100 µl of a mix of 23 amino acids and three common polyamines standards in 5% perchloric acid. The tubes were vortexed and incubated in a water bath at 60ºC for 1 h. Subsequently, 50 µl of L-asparagine (100 mg/ml in water) was added as a termination agent. Incubation proceeded at 60ºC for additional 30 mins. Acetone was evaporated from the tubes by centrifuging in a speed vacuum for 5 min. Toluene (400 µl) was added to each tube, vortexed for 1 min and centrifuged at 13500 x g for 1 min. The toluene phase (200 µl) containing only polyamines was transferred to a new microfuge tube and the toluene was completely evaporated under vacuum for 15 min. The dry dansyl polyamines were dissolved in 730 µl of methanol. To each tube, 135 µl of the toluene-extracted aqueous fraction was added along with 135 µl of 2.9 N acetic acid to bring the total volume to 1 ml. Excess carbonate was removed by adding acetic acid and letting the resulting CO_2 escape for 10 min. The entire sample was filtered through a 0.45 µm nylon filter into

auto-sampler vials. The blank HPLC runs were conducted using dansylated 5% PCA. The separation of polyamines was carried out using a Perkin-Elmer series 200 pump and autosampler fitted with 20 µl loop (10 µl injection volume); a Perkin-Elmer Pecosphere: 3 x 3 CR C_{18}, 33 mm x 4.6 mm i.d. cartridge column (3 pm particle size) and a fluorescence detector (LS-1, Perkin-Elmer). The excitation and emission wavelengths were set at 340 and 510nm, respectively. A TotalChrom HPLC software package (Perkin-Elmer) was used to interpret the data. HPLC conditions for the simultaneous separation and quantitation amino acids and polyamines were as previously described [140].

Cloning and expression of salt tolerance genes in strain CcI3

A DNA fragment containing the ribosomal RNA *rrnB T1* transcriptional terminator and the λ-bacteriophage *pR* and *pL* promoters was amplified from pASV2 [141] using primers so that the amplified fragment contained *SacI* and *KpnI/SpeI/ApaI* cut sites (Table 4). The resulting fragment was digested with *SacI* + *KpnI* and ligated to pHTK1 [142] digested with the same restriction endonucleases, generating pHTK1RO. Tolerant-strain-specific genes that showed differential expression under salt stress were amplified using PCR and cloned into the pHTK1RO plasmid downstream of the constitutive λ-bacteriophage *pR* and *pL* promoters positioned in tandem. The positon of the *rrnB* T1 transcriptional terminator upstream of the *pR* and *pL* promoters prevented read-through transcription into the cloned gene. The primers used in the amplification of the salt-responsive genes are given in Table 4. The construct containing the desired fragment was used to transform *E. coli BW29427*, a diaminopimelic acid (DAP) auxotroph. The construct was introduced into *Frankia casuarinae* strain CcI3 through

filter mating with *E. coli* BW29427. Briefly, *Frankia casuarinae* strain CcI3 culture (50 mL) in the logarithmic phase (7 days old) was harvested and washed once with MP medium. The *Frankia* cells were resuspended in 1 mL of MP medium supplemented with 5 mM NH_4Cl + 5 mM sodium propionate and mixed with 600 µL of 0.6 OD *E. coli* BW29427. The *Frankia* – *E. coli* mixture was filtered through 0.22 µM of nitrocellulose filter under the application of vacuum pressure. The nitrocellulose filter was transferred to LB agar plate supplemented with 300 µM DAP and incubated at 28ºC for 24 hrs. The filter (with the conjugants on it) was suspended in 3 mL MPN medium supplemented with 5 mM sodium propionate and vortexed until the pellet separated from the filter. The *Frankia–E. coli* cake was homogenized in a glass homogenizer. The homogeneous *Frankia–E. coli* mixture (2 ml) was mixed with 2 ml of 0.8% agar containing 20 µg/ml tetracycline and 50 µg/ml kanamycin and overlaid on 2% MPN agar plates containing 20 µg/ml tetracycline and 50 µg/ml kanamycin. Different dilutions of the *Frankia–E. coli* mixture (10^{-1} - 10^{-6}) were plated. As a control, empty pHTK1RO vector was conjugated following similar procedures. Recipient *Frankia casuarinae* strain CcI3 was also plated as a control. After a month, single colonies of *Frankia* were used to start a liquid culture. Transformants were confirmed through PCR and restriction digestion analysis of the extracted plasmid. The transformed strain CcI3 and the two controls (recipient *Frankia*, transformed with empty vector) were tested for salt tolerance as described above. qRT-PCR was used to confirm expression of the cloned genes.

Table 5. Plasmids used in this study

Plasmid	Characteristics	Source or reference
pGEM-5Zf (+)	*Bla, lacZ*	Promega, Madison, WI
pHTK1	*tetA*	
pCR®2.1-TOPO	*Bla*, *neo*	Invitrogen (Carlsbad, CA)
pIJ790	λ-RED (gam, bet, exo), cat, araC, rep101ts	[143]
pUZ8002	*tra*, *neo*, RP4, helping plasmid for conjugation	[144]
pRK2013	*neo*, *tra,*	[145]
pHTK1-RO	*tetA*, *mob*	This study
pHTK1-RO-22605	*tetA*, *mob, CCI6_RS22605*	This study
pHTK1-RO-17915	*tetA, mob, CCI6_RS17915*	This study
pHTK1-RO-21730	*tetA, mob, CCI6_RS21730*	This study
pHTK1-RO-13590	*tetA, mob, CCI6_RS13590*	This study
pHTK1-RO-19875	*tetA, mob, CCI6_RS19875*	This study
pHTK1-RO-20860	*tetA, mob, CCI6_RS20860*	This study
pHTK1-RO-17580	*tetA, mob, CCI6_RS17580*	This study

Results

Comparative Genomic analysis of salt tolerance in *Casuarina* isolates

Before employing comparative genomics to determine the molecular mechanisms of salt and osmotic stress tolerance in *Casuarina* isolates, several *Frankia* strains were screened for salt and osmotic stress tolerance under nitrogen-proficient and nitrogen-deficient conditions. The effect of salt and osmotic stress on the physiology of the strains and on their symbiotic performance was assessed. Comparative genomic analysis of the salt-tolerant and salt-sensitive strains revealed key genetic differences that account for the observed disparity in phenotype.

Casuarina isolates manifest diverse salt tolerance levels

The salt tolerance levels for *Frankia* strains isolated from *Casuarina* hosts were measured and compared to the levels found for *Frankia* strains isolated from non-*Casuarina* hosts. Strain CcI6 and Allo2 were highly salt-tolerant and exhibited a NaCl MIC value of up to 1000 mM (Figure 3A). Strain CcI3 was the least salt-tolerant having a MIC value around 475 mM. In general, *Casuarina* isolates had a higher level of salt tolerance compared to other *Frankia* isolates, but this higher level of tolerance was not extended to osmotic stress (Figure 3A). Here, the other *Frankia* isolates exhibited higher levels of tolerance to osmotic stress compared to the *Casuarina* isolates.

Salt tolerance is highly dependent on the external supply of nitrogen

Under nitrogen-deficient conditions, the level of salt tolerance by *Frankia* strains (including the most salt-tolerant isolates) dramatically decreased (Figure 3B). For

strains CcI6 and Allo2, the NaCl MIC value went from 1000 mM under nitrogen-sufficient

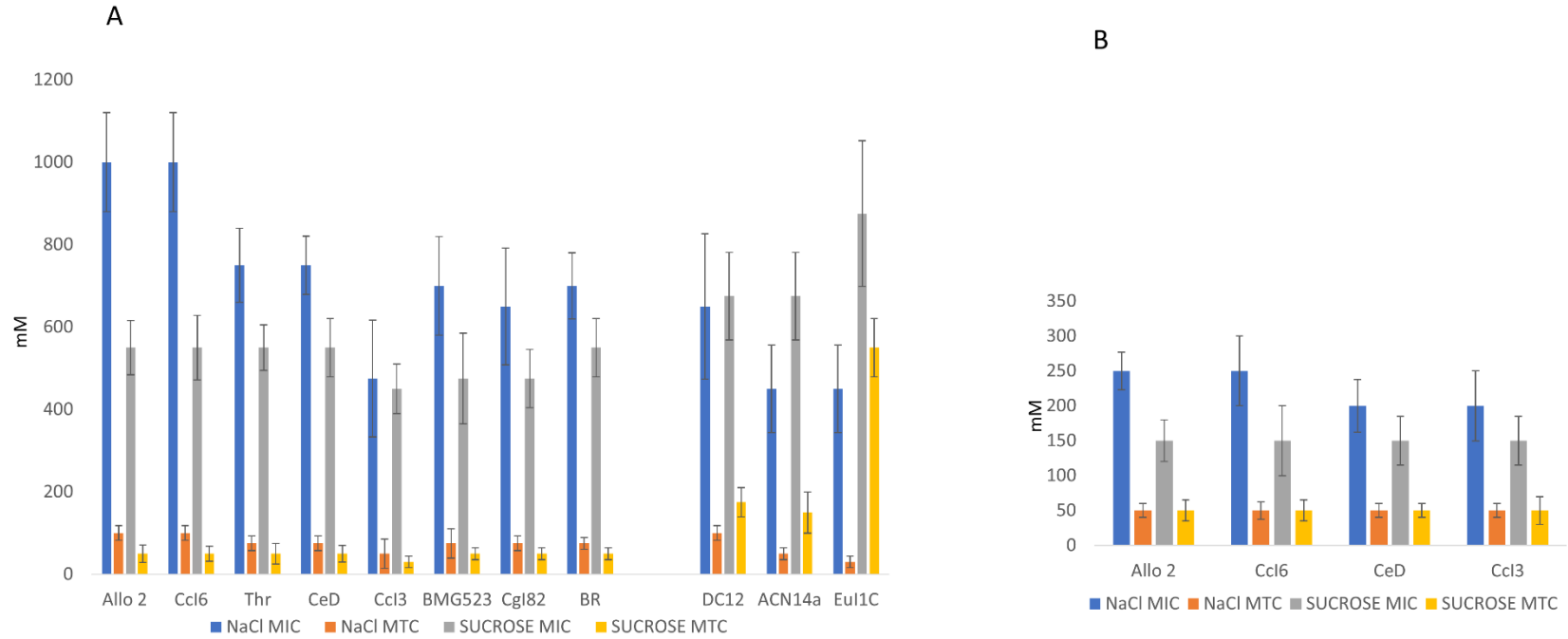

Figure 3. Salt sensitivity assay. (A) The minimum inhibitory concentration (MIC) and maximum tolerance concentration (MTC) values for the different *Frankia* strains exposed to salt (NaCl) and osmotic stress (imposed by sucrose treatment) under nitrogen-sufficient (NH_4Cl) conditions. For *Casuarina* isolates (left of the bar graph), the MIC and MTC values are expressed as the average values calculated from the dry weight and the BCA protein assays. For the non-Casuarina isolates (DC12, ACN14a, and Eul1c), only the dry weight measurements were used to determine the MIC and the MTC values as natural pigments produced by these strains interfered with the BCA protein assay. (B) The MIC and MTC values for salt-tolerant and salt-sensitive *Casuarina* isolates under nitrogen-deficient (N_2) conditions.

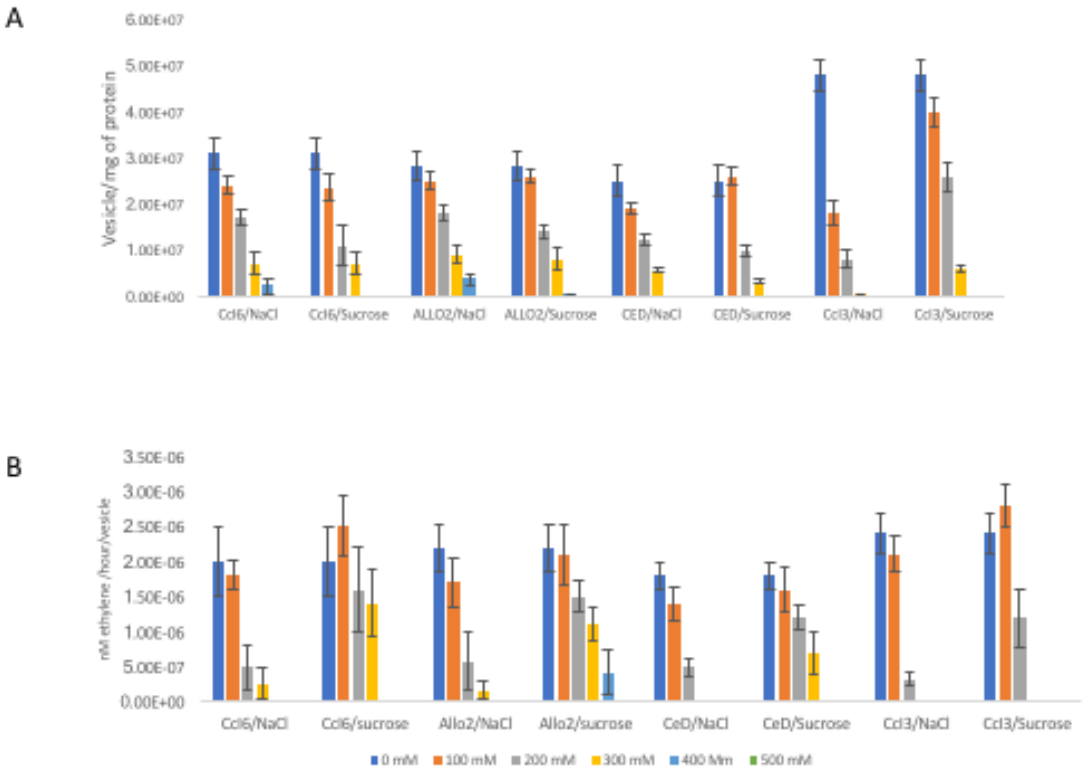

Figure 4. The effect of salt stress on the vesicle formation and nitrogenase activity by salt-tolerant and salt-sensitive *Casuarina isolates*. Cultures were grown under nitrogen-deficient conditions with various degrees of salt or osmotic stress. (A) vesicle production. (B) nitrogenase activity expressed on a per-vesicle basis.

conditions to 250 mM under nitrogen-deficient conditions. The observed differences in salt tolerance levels between the salt-tolerant and salt-sensitive strains also substantially decreased under nitrogen deficient conditions (Fig 3B). Vesicle formation and nitrogenase activities were also affected by salt stress. Nitrogenase activity was more severely affected (Fig 4). When the strains were grown in a medium containing greater than 200 mM NaCl, nitrogenase activity was drastically reduced and barely detectable (Fig 4).

In vitro level of salt tolerance was not correlated with symbiotic performance under salt stress

Salt stress affected nodulation of *Casuarina cunninghamiana* by strains CcI3 and CcI6 (Table 6). The number of nodulating plants as well as the number of nodules per plant were affected by salt stress. At 200 mM NaCl concentration, the salt-sensitive strain (CcI3) was unable to induce nodule formation on *Casuarina cunninghamiana*. Strain CcI6, however, induced the formation of few nodules at 200 mM NaCl. At the control condition (0 mM NaCl), strain CcI3 and strain CcI6 induced comparable numbers of nodules, however, at NaCl concentrations above 50 mM, strain CcI6 performed better at inducing nodules. Under control conditions (no salt stress), both strains induced nodulation on all of the experimental plants within 30 days after inoculation. However, after the application of 100 mM NaCl, the percentages of nodulating plants were reduced to 53% and 60% for plants inoculated with strains CcI3 and CcI6, respectively.

Examination of the effect of salt stress on the growth of *Casuarina cunninghamiana* revealed that pre-inoculation with *Frankia* increases plant height, root and shoot dry weight at all levels of NaCl tested (Table 7). Nevertheless, at all levels of salt stress

tested, the strain of *Frankia* used for inoculation did not make a difference on the growth parameters measured. Control as well as pre-inoculated plants did not survive at NaCl concentrations above 200 mM.

Table 6. The effect of salinity on the nodulation of *Casuarina cunninghamiana*. Fifteen plants per treatment were used to generate the mean values presented. Two parameters were measured: nodule number per plant and nodulation frequency. Statistical analysis was done separately for each time period after inoculation (2 – 8 weeks). For each salt concentration, different uppercase letters (B–C) indicate significant differences between strains (CcI3 Vs CcI6). For each strain used for inoculation, different small case letters (a–e) indicate significant differences between NaCl treatments (0 mM Vs 50 mM Vs 100 mM Vs 150 mM Vs 200 mM) according to the least significant difference (LSD) test at $p < 0.05$.

Frankia strains	Treatment NaCl (mM)	Nodule number per plant				Nodulation frequency (%)			
		2 weeks	4 weeks	6 weeks	8 weeks	2 weeks	4 weeks	6 weeks	8 weeks
CcI3	0	3.2^{Ba}	9.6^{Ba}	15^{Ba}	18^{Ba}	26.7	100.0	100.0	100.0
	50	2.6^{Ba}	7.1^{Bb}	10.6^{Bb}	12^{Bb}	26.7	60.0	70.0	70.0
	100	0.07^{Bc}	2.1^{Bc}	2.6^{Bc}	3.2^{Bc}	6.7	46.7	46.7	53.3
	150	0	0	0.07^{Bde}	0.13^{Bde}	0	0	6.7	6.7
	200	0	0	0	0	0	0	0	0
CcI6	0	2.8^{Ba}	8.8^{Ba}	14.0^{Ba}	16.0^{Ba}	26.7	100.0	100.0	100.0
	50	2.3^{Bb}	8.4^{Ca}	9.8^{Bb}	14.0^{Ba}	40.0	80.0	80.0	80.0
	100	0.13^{Bc}	2.8^{Cc}	3.5^{Cc}	4.0^{Cc}	6.7	40.0	53.3	60.3
	150	0	0.13^{Bde}	0.4^{Cd}	0.8^{Cd}	0	6.7	13.0	13.0
	200	0	0	0.07^{Be}	0.13^{Be}	0	0	6.7	6.7

Table 7. The effect of salinity on the growth of pre-inoculated *Casuarina cunninghamiana.* Fifteen plants per treatment were used to generate the mean values presented in the table. Three parameters were measured: mean plant height, root dry weight, and shoot dry weight. Statistical analysis was done for each one of the three parameters separately. For each salt concentration, different uppercase letters (A–C) indicate significant difference between strains (Control Vs CcI3 Vs CcI6). For each strain used for inoculation, different lowercase letters (a–e) indicate significant differences between NaCl treatments (0 mM Vs 50 mM Vs 100 mM Vs 150 mM Vs 200 mM) according to the least significant difference (LSD) test at $P < 0.05$.

Frankia strains	Treatment NaCl (mM)	Mean Plant height (cm) at harvest	Root dry weight (grams)	Shoot dry weight (grams)
CcI3	0	43.3^{Ba}	0.30^{Ba}	0.66^{Ba}
	50	34.2^{Bb}	0.28^{Ba}	0.60^{Bb}
	100	27.3^{Bc}	0.24^{Bd}	0.50^{Bc}
	150	26.1^{Bc}	0.22^{Bde}	0.40^{Bd}
	200	22.5^{Be}	0.21^{Be}	0.34^{Be}
Cci6	0	40.5^{Ba}	0.32^{Ba}	0.62^{Ba}
	50	36.8^{Bb}	0.27^{Bb}	0.58^{Bb}
	100	28.6^{Bc}	0.22^{Bc}	0.52^{Bc}
	150	26.2^{Bc}	0.23^{Bc}	0.44^{Bd}
	200	21.8^{Be}	0.19^{Be}	0.36^{Be}
Control	0	18.2^{Aa}	0.16^{Aa}	0.36^{Aa}
	50	18.3^{Aa}	0.13^{Ab}	0.32^{Ab}
	100	17.2^{Aad}	0.13^{Ab}	0.30^{Ab}
	150	16.2^{Ade}	0.11^{Abe}	0.25^{Ae}
	200	15.0^{Ae}	0.09^{Ae}	0.23^{Ae}

Genomic characteristics of *Frankia* strains isolated from *Casuarina* trees

After screening for salt stress tolerance, several salt-tolerant *Frankia* strains were sequenced and annotated with the goal of identifying genetic differences with other *Frankia* strains, including strain CcI3, a salt-sensitive strain for which a complete genome is available. Table 8 presents the genomic features of strain CcI3 and six *Casuarina* isolates sequenced in this study. The size and G + C content of the genomes ranged from approximately 5 to 5.6 MB and 69.3 to 70.1%, respectively. Strain CcI3 had the highest number of coding sequences (CDSs), 4327, while strain CeD had the lowest number of CDSs (3807).

We were interested in getting an overall picture of phylogeny of the *Casuarina* isolates and examined multiple genes that are found in common to these genomes. Orthologs found among the *Frankia* genomes and *Acidothermus cellulotycus* (the outgroup) were determined by a modified Lerat method [131]. Figure 5A shows a maximum-parsimony concatenated tree generated from 394 conserved orthologs (panorthologs). As expected, the eight *Casuarina* strains grouped together and were distinct from the closely related Cluster-1a strain ACN14a, which was isolated from *Alnus* trees. Strains CcI6 and Allo2 showed close similarity and grouped together, while strains CeD and BMG5.23 showed the least similarity with other strains (Fig 5A).

Average nucleotide identity, average amino acid identity, and genome to genome distance

The average nucleotide identity between any pair of *Frankia* strains isolated from *Casuarina* trees was greater than 99%. In contrast, the average nucleotide identity between any *Casuarina* isolate and the closely-related strain ACN14a was less than

85% (Figure 5B). The average amino acid identity between all pair of *Casuarina* isolates was greater than 98%, while the average amino acid identity between any *Casuarina* isolate and the closely-related *Alnus* strain was less than 77% (Fig 5B). The genome to genome distance (multiplied by 1000) between the *Casuarina* isolates fell in the narrow range of 1.9 (CcI6 - Allo2) to 8.1 *(*Allo2 - BMG5.23). The genome-to –genome distance between any one of the *Casuarina* isolates and the closely-related strain ACN14a was in the range between 171.9 and 173.8 (Figure 5B). Based on DNA-DNA hybridization (DDH) prediction by GGD 2.1, at *p=0.05* level, any two *Casuarina* isolates have at least 70% DDH value, the cutoff point for species delineation. On the other hand, in a pairwise comparison with strain ACN14a, none of the *Casuarina* isolates had DDH value greater than or equal to 70% at p = 0.05 level.

Table 8. Genomic features of *Frankia* strains isolated from *Casuarina* trees

	Frankia sp. strains isolated from *Casuarina* trees						
	CcI6	Allo2	CeD	Thr	Br	CcI3	BMG5.23
Chromosome size(Mb)	5.58	5.35	5.00	5.31	5.23	5.43	5.27
GC %	69.3	70.0	70.0	70.0	70.0	70.1	69.9
N_{50} (bp)	103, 000	96,900	73,600	71,600	60,200	543,3628	64,900
CDS	4,280	4,224	3,857	4,209	4,220	4,327	4,114
rRNA	9	8	6	5	4	6	9
tRNAs	45	45	45	45	45	45	48
#Scaffolds	136	110	120	169	180	1	166
#Contigs	155	133	154	184	180	1	191
Reference	[38]	[41]	[39]	[37]	[40]	[33]	[36]

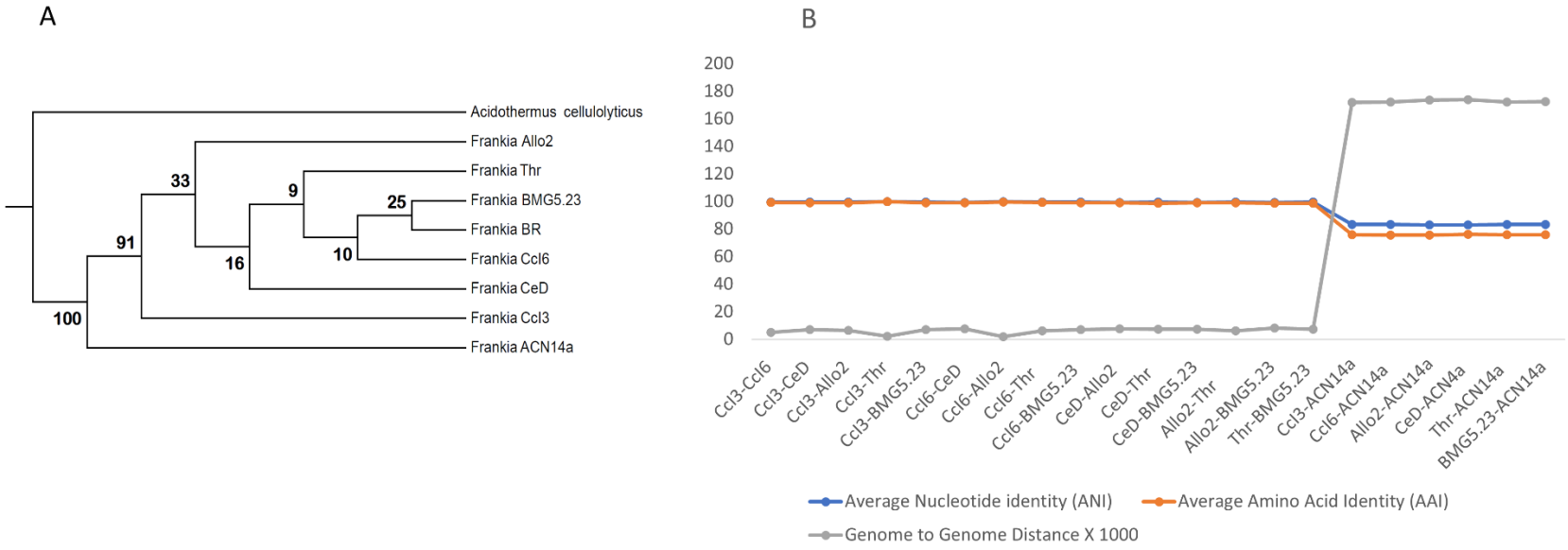

Figure 5. Genomic taxonomy of *Frankia* strains isolated from *Casuarina* spps. (A) Concatenated phylogenetic affiliation of 394 maximum-parsimony trees for amino acid sequences of orthologs among all of the genomes including *Casuarina* isolates, *F. alni* strain ACN14a isolated from *Alnus* and *Acidothermus celloluyticus,* which was used as an outgroup. The numbers on the branches represent the percent confidence of speciation of a given branch. (B) AAI, ANI, and genome to genome distance (multiplied by 1000) values for the different *Casuarina* isolates and *F. alni* strain ACN14a isolated from *Alnus* which is included for comparison. Genome to genome distance was determined by GGDC as a function of sum of all identities found in HSPs divided by overall HSP length. GGDC2 BLAST+ was used as the alignment method for finding intergenomic matches.

Pan-genome analysis reveals a high abundance of singletons among all of the strains

Pan-genome analysis was performed by orthologous clustering using OrthoVenn. OrthoVenn utilizes OrthoMCL to perform an all-against-all BLASTP alignment and identify putative orthology and in paralog relationships with the Inparanoid algorithm. The OrthoVenn analysis of the six *Casuarina* isolates (strain CcI3 and five strains sequenced in this study) revealed 3,278 pan-orthologous gene clusters, of which 3,246 were single copy pan orthologous gene clusters (Figure 6A). Pair-wise comparison of the genomes showed that strains CcI6 and Allo2 shared the highest number of unique clusters (132), not found in the other strains. No genome had more than 2 clusters unique to itself, but all had many singletons (Figure 6B) suggesting that there was insufficient time for gene duplication events to occur after the appearance of singletons. Among the six *Casuarina* isolates, strain BMG5.23 had the highest number of singletons (160). About 30% of the singletons in any one strain were hypothetical proteins. Singletons generally occurred dispersed within the genome, suggesting they are acquired independently. However, in strain BMG5.23, singletons seem to be clustered in the same region. The singletons in strain BMG5.23 had varying GC contents and the fact that at least some of them co-occur in the same region suggests that there might be hot spots for insertion.

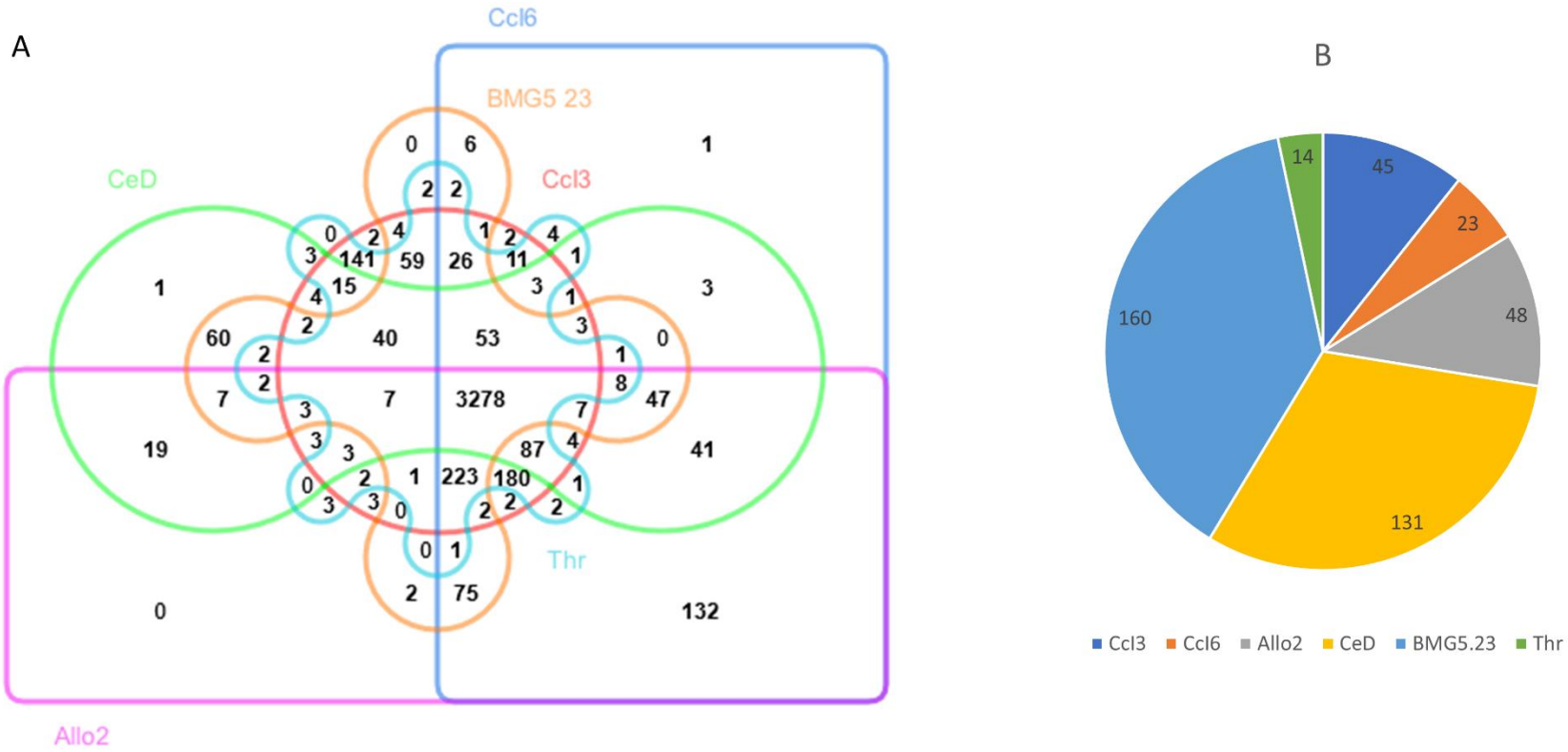

Figure 6. Pangenome overview of six *Frankia* strains isolated from *Casuarina* trees. (A) Six-way Venn diagram showing shared and specific gene clusters among the *Casuarina* isolates as determined by OrthoVenn. An E-value cutoff of 1 e^{-5} was used for protein similarity search and inflation value of 1.5 was used for the generation of orthologous clusters using the Markov Cluster Algorithm. (B) Number of singletons identified in each *Casuarina* isolate.

The two salt-tolerant strains contain many hypothetical proteins absent in the other strains

Comparison of the single copy orthologous gene clusters between the salt-sensitive isolate (CcI3), the moderately salt-tolerant strain (CeD), and the two highly salt-tolerant strains (CcI6 and Allo2) was performed using the Lerat program. All four strains shared 2919 single copy core genes (Figure 7A). The two highly salt-tolerant strains contained 153 single copy orthologous genes that were not shared with the moderately-tolerant and salt-sensitive strains. Both highly salt-tolerant strains and the moderately-tolerant strain shared 88 single copy genes that were not present in the salt-sensitive strain. Fig 7B shows the distribution of genes found in the two highly salt-tolerant strains into cluster of orthologous groups of proteins (COG) functional categories. Among the 153 unique genes found in the two highly salt-tolerant strains, 114 of the genes were annotated as hypothetical proteins. However, re-annotation using the RPSBLAST program on COG database (prokaryotic proteins) revealed only 99 hypothetical proteins. The three COG categories that were highly represented among the unique genes found only in the two highly salt-tolerant strains were: (COG R) general function prediction only; (COG L) DNA replication, recombination, and repair; and (COG M) cell wall/membrane envelope biogenesis. Tolerant-strain-specific genes assigned to cell wall/membrane biogenesis (COG M) category included genes encoding glycosyl transferases, proteins involved in cellulose synthesis, D-alanine:D-alanine ligase (Ddl), and a predicted nucleoside-diphosphate-sugar epimerase. Glycosyl transferases allow for a more flexible response to environmental stress. In tobacco, ectopic expression of a glycosyl transferase (UGT85A5) leads to enhanced salt

tolerance [146]. Ddl is involved in the D-alanine branch of peptidoglycan biosynthesis. Mutation in a d-alanine–d-alanine ligase of *Azospirillum brasilense* Cd result in an overproduction of exopolysaccharides and a decreased tolerance to saline stress [147].

Another COG functional category that was represented among the genes found only in the tolerant strains was coenzyme metabolism. Two genes encoding hypothetical proteins (CCI6_RS06345, CCI6_RS15790) with ubiquinone synthesis-related methyl transferase domains and a hypothetical protein with a geranylgeranyl pyrophosphate synthase domain (CCI6_RS02885) were among the tolerant strain-specific genes assigned into the coenzyme metabolism functional category. Geranylgeranyl pyrophosphate (GGPS) synthase catalyzes formation of geranylgeranyl pyrophosphate (GGP), which is a key step in biosynthetic pathway of carotenoids and many other terpenes [148].

Among the 153 unique genes found only in the two highly salt-tolerant strains, genes encoding D-alanine:D-alanine ligase (Ddl) [CCI6_RS21780], anti-sigma regulatory factor (CCI6_RS20000), carbamoyl transferase NodU family (CCI6_RS17910), and D-glycero-beta-D-manno-heptose 1-phosphate adenylyltransferase (CCI6_RS17935), and a transcriptional regulator (CCI6_RS02080) had homology hits (identity > 65%) in the highly salt-tolerant actinomycete *Nocardiopsis halotolerans*.

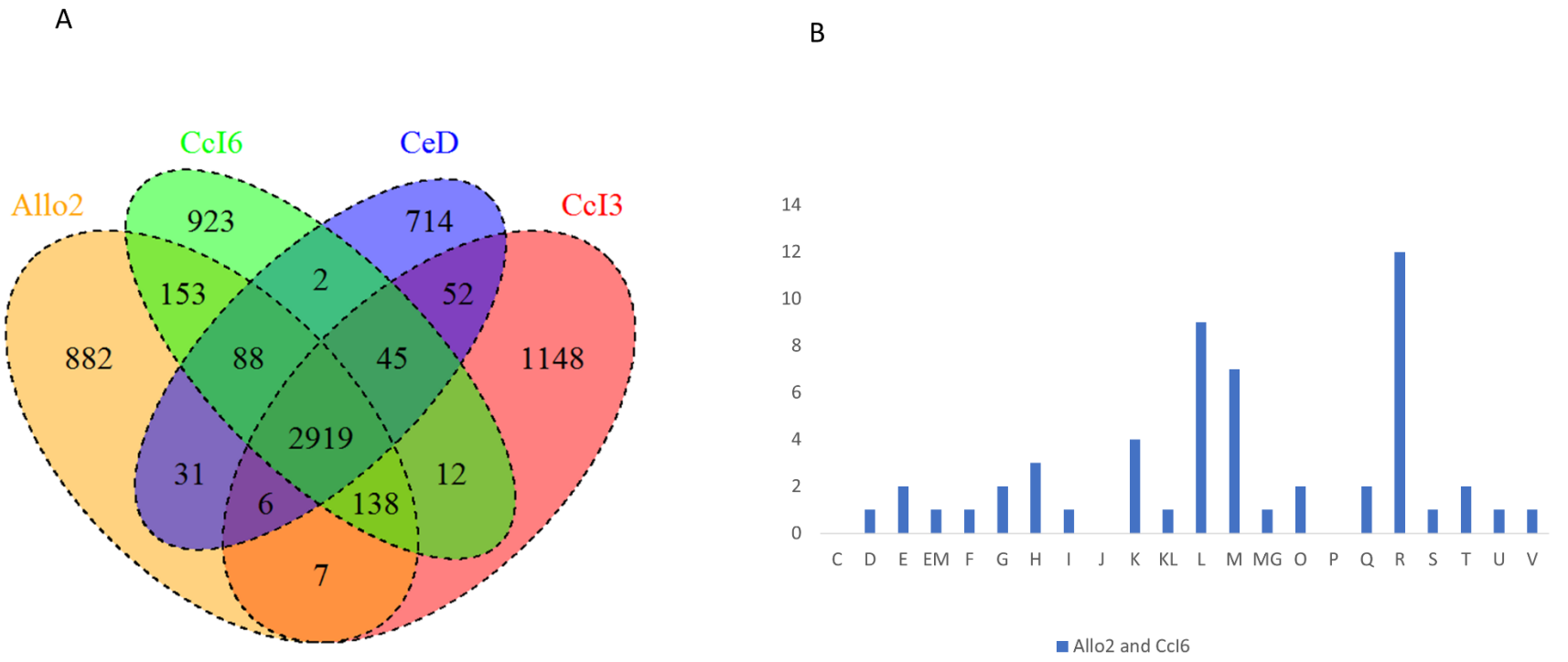

Figure 7. Pangenome analysis of single-copy genes (A) Shared and specific single-copy orthologous CDSs among the highly salt-tolerant (Allo2 and CcI6), the moderately salt-tolerant (CeD), and the salt sensitive (CcI3) strains. (B) Distribution of the 153 single-copy genes specific to the two highly salt-tolerant strains into functional COG categories: C, energy production and conversion; D, cell division and chromosome partitioning; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme metabolism; I, lipid transport and metabolism; J, translation, ribosomal structure and biogenesis; K, transcription; L, DNA replication, recombination and repair; M, cell wall/membrane biogenesis; O, posttranslational modification, protein turnover, chaperones; P, inorganic ion transport and metabolism; Q, secondary metabolite biosynthesis, transport and catabolism; R, general function prediction only; S, function unknown; T, signal transduction mechanisms; U, intracellular trafficking and secretion; and V, defense mechanisms. Proteins that could be classified to more than one category are represented by two letters.

Tolerant and sensitive strains contain the same set of classical salt-tolerance genes

Many of the known salt tolerance mechanisms are also osmotic stress response mechanisms or verse versa. The *Casuarina* genomes were data mined for the presence of these known osmotic/salt tolerance mechanisms.

All of the *Casuarina* genomes had a *kdpFABCDE* operon encoding the membrane-associated P-type ATPase, Kdp-ATPase (*kdpFABC*), involved in K^+ uptake and a two-component regulatory system (*kdpDE*), which regulates the expression of *kdpFABC* [149]. The Kdp system plays a role in ion homeostasis and adaptation to osmotic stress. Both tolerant and sensitive strains also contained the Trk system, which is the predominant uptake system in medium containing more than 1 mM K^+. All of the *Casuarina* genomes lacked mechanisms for *de novo* synthesis or uptake of glycine betaine. However, all of the *Casuarina* isolates possessed the ability for the biosynthesis of the important osmo-protectant proline. Three pathways for the synthesis of trehalose, an effective osmolyte, were also present in all the *Casuarina* genomes. The first pathway (the TreY-TreZ pathway) synthesizes trehalose from glycogen-like alpha (1-->4)-linked glucose polymers. The second pathway (the TreS pathway) synthesizes trehalose from maltose, while the third pathway, the OtsA-OtsB pathway, utilizes glucose-6-phosphate and UDP-glucose to synthesize trehalose through a two-step enzymatic process involving trehalose-6-phosphate synthase (OtsA) and trehalose-6-phosphate phosphatase (OtsB). All *Casuarina* isolates contained the *asnO–ngg* cluster putatively involved in the synthesis of N-acetylglutaminylglutamine

amide (NAGGN), a dipeptide identified as an osmolyte in a few bacteria. The first step of the reaction involves the N-acetylation of a glutamine residue and the subsequent dipeptide bond formation between this residue and a second L-glutamine residue in a reaction catalyzed by Ngg. In the second step of the reaction catalyzed by AsnO, an amide group is transferred from a free L-glutamine molecule to the second L-glutamine residue of NAGG to produce NAGGN. Just like most other genomes containing the *asnO–ngg* cluster, the genomes of all *Casuarina* isolates encode a dipeptidase immediately downstream of the *ngg* gene. A possible role for such peptidase could be balancing of the NAGGN pool during adaptation to osmotic fluctuations. The identities between AsnO from *Casuarina* isolates and AsnO found within other bacterial species was high (greater that 60%) whereas the putative Ngg protein from *Casuarina* isolates had low identity (< 15 %) with Ngg proteins identified in other species. In contrast to the asnO-ngg organization found within other genomes, the *asnO* and the *ngg* genes in the *Casuarina* isolates were not contiguous, but had the dipeptidase gene between them.

Transcriptome analysis of salt stress tolerance in *Casuarina* isolates

After identifying genetic differences between salt-tolerant and salt-sensitive *Frankia* strains through comparative genomics, we became interested in identifying differences in the salt stress response at the transcription level. Focus was placed on the expression pattern of tolerant-strain-specific genes. To do this analysis, the transcriptome profiles of strain CcI6 (one of the two highly salt-tolerant strains) and strain CcI3 (the salt-sensitive strain) under control conditions (no salt stress) were compared to the respective profiles under salt and osmotic stress conditions. Strains CcI6 and CcI3 were exposed to either no stress, salt stress or osmotic stress for seven

days and the transcriptome profile was analyzed using RNA-Seq. Two biological and independent experiments were performed for each condition. After sequencing of the libraries, 17 million, 9.5 million, and 19.8 million reads were obtained under the control, salt stress, and osmotic stress conditions, respectively for strain CcI6. For strain CcI3, 165 million, 160 million, and 170 million reads were obtained under the control, salt stress, and osmotic stress conditions CLC Genomics Workbench 9.0.1 was used to map the reads to the respective genomes (the annotated *Frankia* sp. strain CcI6 genome or the completed *Frankia casuarinae* strain CcI3 genome). For strain CcI6, an average of 2.14 million, 0.85 million, and 2.5 million reads could be unambiguously mapped in pairs for the reference condition, for the salt stress condition, and for the osmotic stress condition, respectively. For strain CcI3, an average of 66 million, 60 million, and 70 million reads could be unambiguously mapped in pairs for the reference condition, for the salt stress condition, and for the osmotic stress condition, respectively.

Salt stress-induced changes in the transcriptome differ between tolerant and the sensitive strains

Transcriptome analysis of strain CcI3 revealed that a total of 214 and 226 genes were up-regulated and 303 and 165 genes were downregulated under salt and osmotic stress, respectively. For strain CcI3, 163 and 37 genes were up-regulated and 111 and 78 genes were downregulated under salt and osmotic stress, respectively (Fig 8). The complete list of differentially expressed genes for strains CcI3 and CcI6 are provided in Tables A1-A8. In strain CcI3, a total of 179 genes were only upregulated under salt stress, while 191 genes were only upregulated under osmotic stress. In strain CcI3, 155 genes were only upregulated under salt stress, while 29 genes were only upregulated

under osmotic stress. For the salt-tolerant strain (CcI6), 35 upregulated and 65 downregulated genes were found in common with both salt stress and osmotic stress conditions (Fig 8). The corresponding numbers for the salt sensitive strain (CcI3) were 8 and 27, respectively. In strain CcI6, among the genes upregulated under both salt and osmotic stress were hypothetical proteins, proteins involved in cell wall/membrane biogenesis functions and the following transport proteins: an ABC-type Fe3+ hydroxamate transport system, periplasmic component (CCI6_RS09145), and ABC-type Fe3+ siderophore transport system, permease component (CCI6_RS09155). This result suggests that increased iron uptake is part of the general response to salt and osmotic stresses. Increased iron uptake under salt stress was previously reported for *Bacillus subtilis* [150]. In strain CcI3, genes upregulated under both salt and osmotic stress conditions included: Francci3_4204 (AMP-dependent synthetase and ligase), Francci3_4205 (2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase), Francci3_4206 (3-dehydroquinate synthase), Francci3_2603 (cobyric acid synthase), Francci3_4365 (transglycosylase-like), Francci3_4194 (methyltransferase type 12), Francci3_2676 (helicase-like), and Francci3_4196 (hypothetical protein).

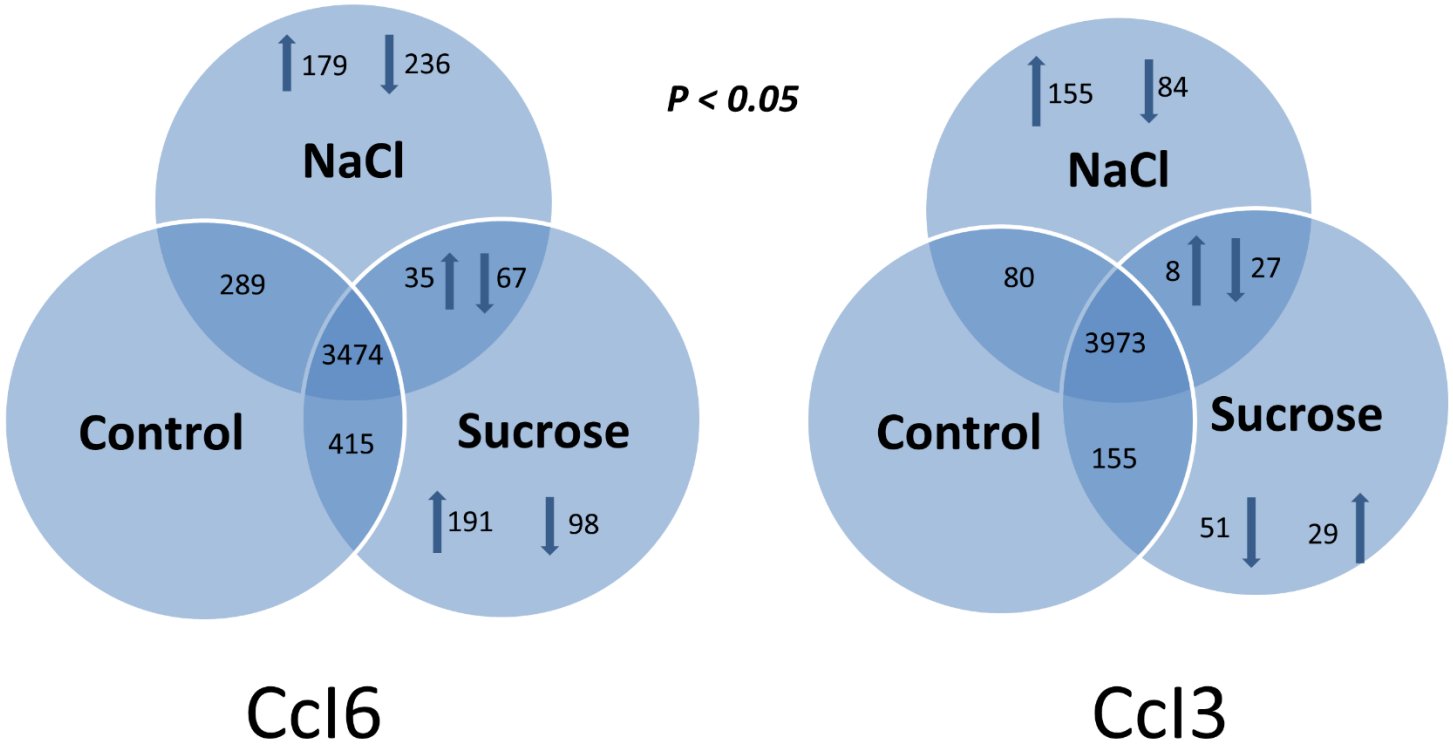

Figure 8. Global gene expression responses following salt and osmotic stress. Venn diagram showing the extent of overlap between genes differentially expressed under salt and osmotic stress in the salt-tolerant strain *Frankia* sp. CcI6 (A) and in the salt-sensitive strain *Frankia casuarinae* strain CcI3 (B). The arrows indicate the number of upregulated and downregulated genes under salt and osmotic stress. The intersection indicates the number of differentially expressed genes under both conditions.

The functional category analysis of differentially expressed genes in strain CcI6 revealed that of the total 214 genes upregulated under salt stress, 18, 8, and 80 represented COG R (general function prediction only), COG S (function unknown), and genes that are not in COGs, respectively (Table 9). On the other hand, for strain CcI3, out of the 163 genes upregulated under salt stress, 16, 3, and 36 represented COG R (general function prediction only), COG S (function unknown), and genes that are not in COGs, respectively. For both strain CcI3 and CcI6, some of the other COG categories that were highly-represented among the genes upregulated under salt stress included COG M (cell wall and membrane biogenesis), COG E (amino acid transport and metabolism), COG H (coenzyme transport and metabolism), and COG I (lipid transport and metabolism). Strain CcI6 had slightly more COG M genes upregulated under salt stress while strain CcI3 had slightly more COG I genes upregulated under salt stress. For strain CcI6, the COG categories highly represented among the genes upregulated under osmotic stress include COG R (general function prediction only), COG S (function unknown), COG M (wall and membrane biogenesis), COG E (amino acid transport and metabolism), COG H (coenzyme transport and metabolism), COG I (lipid transport and metabolism), COG C (energy production and conversion), and COG P (inorganic ion transport and metabolism).

Table 9. Functional categories of genes differentially expressed under salt and osmotic stress conditions in strains CcI3 and CcI6

COG category	Differentially expressed genes p < 0.05							
	NaCl				Sucrose			
	Upregulated		Downregulated		Upregulated		Downregulated	
	CcI6	CcI3	CcI6	CcI3	CcI6	CcI3	CcI6	CcI3
Amino acid transport and metabolism	12	14	13	2	12	1	10	3
Carbohydrate transport and metabolism	6	10	10	4	6	2	6	3
Cell Wall/membrane biogenesis	13	7	15	3	11	1	6	-
Energy production and conversion	8	8	16	1	15		9	6
Inorganic ion transport and metabolism	5	8	4	4	10	1	4	6
Lipid transport and metabolism	7	10	13	3	7	3	10	2
Posttranslational modification, protein turn over, chaperons	8	2	5	7	4	1	4	5
Replication, recombination, and repair	7	5	14	15	5	7	7	6
Signal transduction mechanism	3	4	14	4	5	-	4	3
Transcription	6	5	20	6	2	-	8	12
Coenzyme metabolism	9	9	9	-	7	2	7	-
Translation	5	2	11	3	6	-	6	-
General function prediction only	18	16	19	8	26	4	9	8
Function unknown	8	3	19	6	9	-	8	1
Not in COGs	80	36	101	40	88	10	56	21

Comparison of the salt and osmotic stress responsive genes in CcI3 and CcI6 revealed that 16 and 6 genes showed a similar pattern of upregulation across the two strains under salt and osmotic stress conditions, respectively (Fig 9). Similarly, 20 and 12 genes showed a similar pattern of downregulation across the two strains under salt and osmotic stress conditions, respectively. Some of the COG categories represented among genes that showed the same pattern of differential expression across strains were COG K (transcription), COG M (cell wall and membrane biogenesis), COG E (amino acid transport and metabolism, COG I (lipid transport and metabolism), and COG C (energy production and conversion).

Because the RNA-seq data were used for downstream analyses, we validated the data set by performing quantitative reverse transcription PCR (qRT-PCR). The primer sets used for the qRT-PCR validation of the RNA-seq data are given in table 4. A high degree of correlation ($R = 0.95$) was observed between the normalized values of the fold change from the qPCR data and the normalized fold change values from the RNA-Seq data.

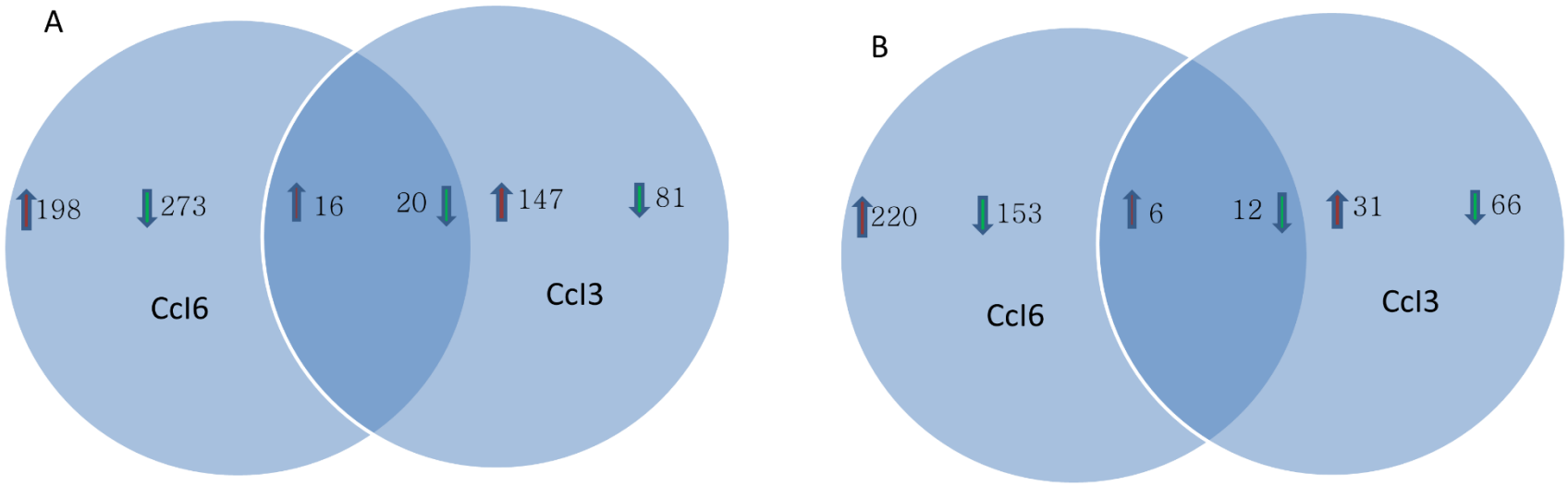

Figure 9. The overlap in response to salt/osmotic stress between strains CcI3 and CcI6. (A) Genes with similar or deferring patterns of expression in strains CcI3 and CcI6 under salt stress. (B) Genes with similar or deferring patterns of expression in strains CcI3 and CcI6 under osmotic stress.

Many hypothetical proteins that are unique to the tolerant strain were upregulated under salt stress

Under osmotic stress, an acetyl transferase with general function prediction only and seven hypothetical proteins that are unique to the two salt-tolerant strains were upregulated. Under salt stress, a zinc peptidase, ADP-heptose:LPS heptosyltransferase, a major facilitator superfamily (MFS) transporter, and five hypothetical proteins unique to the tolerant strains were upregulated (Table 10). One hypothetical protein (CCI6_RS13590) was upregulated under both salt and osmotic stress conditions.

Versatile responses of transcription factors

Not surprisingly, COG K (transcription) genes were highly represented in the transcriptome of both the salt-tolerant and sensitive strains under salt stress. In strain CcI6, six genes (CCI6_RS12535, CCI6_RS20460, CCI6_RS18600, CCI6_RS15305, CCI6_RS01570, CCI6_RS12900, CCI6_RS02550) encoding transcriptional regulators from the GntR, TetR, LysR, and the Crp/Fnr families and six genes (CCI6_RS00475, CCI6_RS02550, CCI6_RS10900, CCI6_RS11405, CCI6_RS17055, CCI6_RS21525) encoding transcriptional regulators from Crp/Fnr and LuxR families were upregulated under salt and osmotic challenge, respectively. In strain CcI6, one transcriptional regulator belonging to the Crp/Fnr family (CCI6_RS02550) was upregulated under both salt and osmotic stress conditions. In strain CcI3, two WhiB transcription factors (Francci3_3790, Francci3_1482), an XRE family transcriptional regulator (Francci3_0523), a putative transcriptional regulator (Francci3_1195) and three other genes (Francci3_0210, Francci3_1767, and Francci3_1469) encoding transcriptional

factors from MarR, LacI, and MerR families were upregulated under salt stress. Transcriptional factors belonging to GntR, TetR, LysR, and the Crp/Fnr families have been implicated previously in several stress responses including heat and osmotic shock [151]. In strain CcI6, a sigma factor with sigF domain (CCI6_RS19210) and an extracytoplasmic stress sigma factor (CCI6_RS15595) were upregulated and downregulated, respectively under salt stress. The corresponding homologues of the two genes in strain CcI3 (100% identity) showed the same pattern of expression under salt stress.

Table 10. Genes unique to the tolerant strain with increased expression under salt stress

Salt Stress			Osmotic Stress		
Locus Tag	Protein product	COG category	Locus Tag	Protein product	COG category
CCI6_RS13590	Hypothetical Protein	-	CCI6_RS13590	Hypothetical Protein	-
CCI6_RS22605	Predicted Zn peptidase	E	CCI6_RS17905	Acetyltransferase (isoleucine patch superfamily)	R
CCI6_RS17915	ADP-heptose:LPS heptosyltransferase	M	CCI6_RS13555	Hypothetical Protein	-
CCI6_RS19875	Hypothetical Protein	-	CCI6_RS15755	Hypothetical Protein	-
CCI6_RS20860	Hypothetical Protein	-	CCI6_RS21770	Hypothetical Protein	-
CCI6_RS17580	Hypothetical Protein	-	CCI6_RS13535	Hypothetical Protein	-
CCI6_RS21730	MFS transporter	-	CCI6_RS02120	Hypothetical Protein	-

Salt stress upregulated several genes involved in peptidoglycan modification

In the salt-tolerant strain, two genes encoding polysaccharide deacetylases (CCI6_RS03540, CCI6_RS11155) were upregulated under salt stress, but were unchanged under osmotic stress. In *Bacillus anthracis*, a polysaccharide deacetylase plays a role in the adaptation of the bacteria to a high salt environment [152]. Under salt

stress only, four glycosyl transferases (CCI6_RS10910, CCI6_RS07965, CCI6_RS21895, CCI6_RS01640) associated with cell wall/membrane biogenesis showed more than four-fold increase in the salt-tolerant strain (CcI6). Similarly, under osmotic stress, four glycosyl transferases associated with cell wall/membrane biogenesis (CCI6_RS02325, CCI6_RS10910, CCI6_RS11195, CCI6_RS11220) were upregulated. CCI6_RS11220 showed a 9.5-fold increase under osmotic stress. One of the above glycosyl transferase genes (CCI6_RS10910) showed a statistically significant up-regulation under both salt and osmotic stress. In the salt-sensitive strain (CcI3), two genes (Francci3_4365, Francci3_4366) encoding transglycosylase proteins were upregulated under salt stress. One of the two genes (Francci3_4365) was also upregulated under osmotic stress.

In strain CcI6, three nucleoside diphosphate sugar epimerase genes (CCI6_RS19225, CCI6_RS08005, CCI6_RS04525) were upregulated under salt stress, while two nucleoside diphosphate sugar epimerases (CCI6_RS00960, CCI6_RS08005) were upregulated under osmotic stress. In contrast, in strain CcI3, no epimerase associated with cell wall and membrane biogenesis was differentially expressed. Taken together the results suggest that under salt stress conditions, cell-wall-related alterations are more prominent in the salt-tolerant strain (CcI6) as opposed to in the salt-sensitive strain (CcI3).

Modulation of membrane composition

In the salt-tolerant strain, two genes encoding acyl-acyl carrier protein (ACP) desaturases (CCI6_RS10965, CCI6_RS10965), were upregulated under salt stress, but not under osmotic stress. In addition to desaturases, two other genes [acyl-CoA

dehydrogenase (CCI6_RS05135) and ABC-type branched-chain amino acid transport systems, periplasmic component (CCI6_RS10620)] that help determine membrane fluidity were upregulated under salt stress. Furthermore, salt stress caused the upregulation of the *ubiE* (CCI6_RS17660) gene encoding a multispecies ubiquinone biosynthesis protein. Ubiquinone accumulation has been shown to increase salt tolerance in *E. coli* through mechanical stabilization of the membrane [153]. In the salt-sensitive strain, four genes coding enoyl-CoA hydratase/isomerases (Francci3_2457, Francci3_2456, Francci3_2455, Francci3_1675) were upregulated under salt stress. Taken together, these results suggest that *Frankia* membrane fluidity was altered by salt stress.

Osmolytes are strain and condition-specific

In the salt-tolerant strain (CcI6), trehalose synthase gene (CCI6_RS13215) was upregulated under salt stress, while glutamate synthase (CCI6_RS10225) and threonine synthase (CCI6_RS08750) genes were upregulated under osmotic stress. Trehalose synthase (TreS, EC 5.4.99.16) catalyzes the intramolecular rearrangement of the α-1,4-linkage of maltose to the α-1,1-linkage of trehalose [154]. Glutamate synthase catalyzes the production of L-glutamate from L-glutamine and 2-oxoglutarate. Threonine synthase catalyzes the reversible reaction between O-phospho-L-homoserine and H_2O to produce L-threonine and phosphate. In the salt-sensitive strain, all the genes involved in the synthesis of ornithine from glutamate (Francci3_3172, Francci3_3173, Francci3_3174, Francci3_3175) were upregulated under salt stress. The *proA* gene involved in the biosynthesis of proline, a widely used osmolyte, was also upregulated under salt stress. Additionally, the transcriptome data showed that the TreY-TreZ

pathway for the synthesis of trehalose from glycogen-like alpha (1-->4)-linked glucose polymers was upregulated, although the change didn't reach the p = 0.05 cutoff for statistical significance. However, qRT-PCR analysis on the three genes involved in the pathway (*treZ, treX and treY*) revealed a statistically significant upregulation under salt and osmotic stress conditions. The regulation of intracellular osmolality by the transport or biosynthesis of compatible solutes is believed to be the principal osmoprotection response in bacteria.

Phosphatidylinositol-3-phosphate mediated signaling may be important during salt stress

In the salt-tolerant strain, a gene encoding an ATP-binding protein (CCI6_RS18715) and a phosphatidylinositol-3-phosphate phosphatase (CCI6_RS01730) gene were upregulated under salt stress, but not under osmotic stress. In mycobacteria, brief phosphatidylinositol 3-phosphate (PI3P) synthesis takes place upon exposure to salt stress [155]. PI3P phosphatase plays a role in the recycling of PI3P. *De novo* PI3P synthesis is rare among bacteria and occurs only in the actinomycetes which possess a novel phosphatidylinositol (PI) biosynthesis pathway. The genes for this pathway are present in the sequenced *Frankia* genomes. In the salt-sensitive strain, several genes involved in signal transduction, including a cyclic nucleotide-binding domain-containing protein (Francci3_4230), a two-component transcriptional regulator (Francci3_4261), a diguanylate cyclase/phosphodiesterase with PAS/PAC sensor (Francci3_1004), and a histidine kinase (Francci3_0470) were upregulated under salt stress.

Proteomics analysis reveals additional functions that might be involved in salt stress tolerance

The proteome profile of the salt-tolerant (CcI6) and salt sensitive (CcI3) strains exposed to salt or osmotic stress was examined by the use of 2D SDS-PAGE (Figure 10). For both strains, prominent changes in protein abundance were readily noticeable under the three different conditions (no stress, salt stress and osmotic stress) and occurred in multiple replicates.

Differential protein spots were screened and those that showed the most prominent differences between the stress conditions and control gels were targeted for further analysis. For the salt-tolerant strain, 16 differentially expressed spots for all three conditions were analyzed and 19 proteins were identified. For the salt-sensitive strain, 14 differentially expressed spots for all three conditions were analyzed and 8 proteins were identified. For strain CcI6, eleven and three spots were upregulated and downregulated, respectively, under salt stress, while five and three spots were upregulated and downregulated, respectively, under osmotic stress. Out of the 19 salt-responsive strain CcI6 proteins, 18 were assigned into COG functional categories, including energy production and conversion (COG C, 4 proteins), transcription (COG K, 3 proteins), amino acid transport and metabolism (COG E, 2 proteins), post translational modification, protein turn over, chaperon functions (COG O, 2 proteins), carbohydrate transport and metabolism (COG G, 2 proteins) [Table 11].

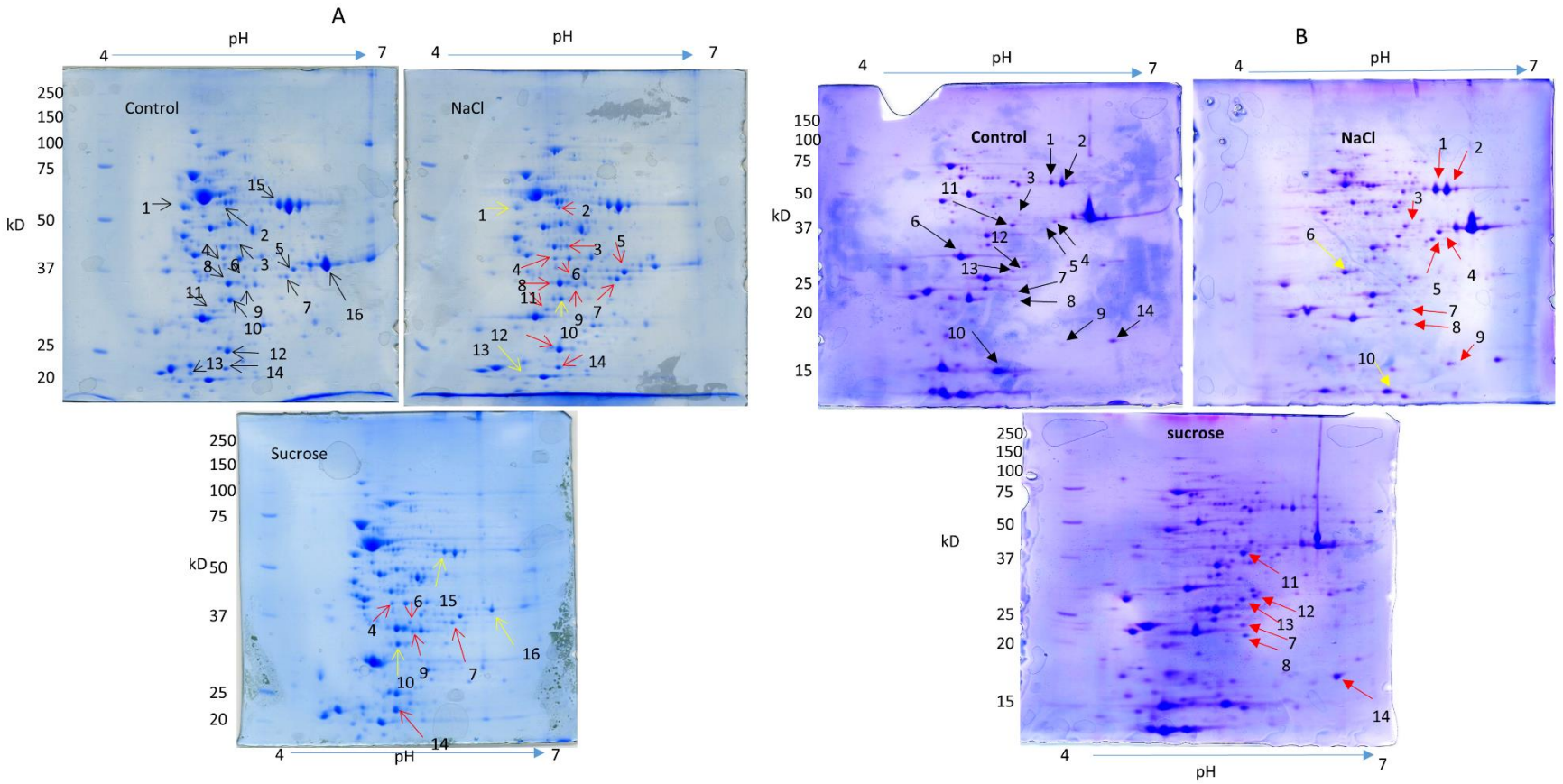

Fig 10**.** Two-dimensional polyacrylamide gel electrophoresis (PAGE) analysis of *salt stress response.* (A) 2D-gel profile of *Frankia sp.* strain *CcI6* under control (no stress), 200 mM NaCl, and 200 mM sucrose stress conditions. (B) 2D-gel profile of *Frankia casurainae* strain CcI3 under control (no stress), 200 mM NaCl, and 200 mM sucrose stress conditions. Red arrows indicate that proteins are upregulated relative to the control, while yellow arrows indicate down regulated proteins relative to the control. The corresponding number spots were in-gel digested with trypsin and analyzed by liquid chromatography-mass spectrometry (LC-MS) and LC-MS/MS for protein identification.

Among the proteins upregulated in the salt-tolerant strain was a cysteine synthase (CysK) which was identified from a downregulated spot under salt stress (spot no. 10) but had 2.5 times more peptide counts in the NaCl-treated samples. The most abundant protein from the same spot (spot no. 10) was an electron transfer flavoprotein alpha subunit apoprotein (CCI6_RS13315), which, as expected, had a lower peptide count in the NaCl-treated samples. Several pyridoxal phosphate-binding proteins, including cysteine synthase, are differentially expressed under salt stress in wheat chloroplasts and help in the synthesis of cysteine as a protective measure against toxic ions [156]. Transcriptomics analysis of the salt-tolerant strain (CcI6) revealed that CysK and CysM were upregulated under salt stress. CysK and CysM are pyridoxal phosphate-dependent enzymes. Our proteomics result showed that pyridoxal phosphate synthase yaaD subunit (WP_011435808.1) was upregulated under salt stress in strain CcI6. Another protein from the COG E category that was upregulated under salt stress in strain CcI6 was the glutamine synthetase (WP_011437527.1). Amino acids including proline and glutamine, serve as protective osmolytes under salt stress [81]. The transcriptome analysis of CcI6 showed that glutamine synthetase was upregulated 1.4-fold times, but the change was not statistically significant. A GroEL chaperon (WP_011438752.1) and an ATP-dependent Clp protease proteolytic subunit ClpP (WP_011435649.1) were among the COG O proteins upregulated under salt stress in strain CcI6. In *B. subtilis*, the synthesis of ClpP protein increases during heat shock, salt and oxidative stress, glucose and oxygen deprivation [157]. Our transcriptome data also showed that the GroEL chaperon (CCI6_RS08745) and ClpP (CCI6_RS22940) are upregulated in salt-stressed CcI6. Another ClpP protein (ClpP_2) [CCI6_RS22535] was

upregulated under osmotic stress in strain CcI6 based on the RNAseq data, but this protein was not identified in our proteomics data. GroEL is one of the heat shock proteins and shows upregulation under salt stress in *Lactococcus lactis*, suggesting there is overlap between salt and heat stress responses [158]. In strain CcI6, among differentially expressed proteins identified in the COG K functional category, a DNA-directed RNA polymerase sigma subunit (WP_023840564.1) was upregulated under salt stress. The transcriptome analysis did not reveal upregulation of the sigma factor, suggesting posttranscriptional regulation of the transcript. An extracytoplasmic function (ECF) family RNA polymerase, sigma subunit (WP_035729933.1) was downregulated under salt stress in strain CcI6. Our transcriptome analysis also showed downregulation of the ECF family sigma factor. The ECF sigma factors are small divergent group of regulatory proteins that control the transcription of genes associated with response to extracytoplasmic stress conditions and some aspect of the cell surface or transport [159]. Two proteins, glyceraldehyde-3-phosphate dehydrogenase and fructose-bisphosphate aldolase, belonging to the carbohydrate transport and metabolism functional category (COG G), were upregulated under salt stress in strain CcI6. Overexpression of glyceraldehyde-3-phosphate dehydrogenase in rice plants improved salt tolerance [160]. Similarly, the overexpression of fructose-bisphosphate aldolase in *Brassica napus* led to increased salt stress tolerance [161]. Another salt-tolerant strain, Allo2, showed similar changes in the proteome prolife under salt and osmotic stress conditions (Table A9, Fig A1).

For strain CcI3, seven and three proteins were identified from upregulated spots under salt and osmotic stress conditions, respectively. All of the proteins upregulated

under osmotic stress were also upregulated under salt stress. CcI3 proteins upregulated under salt stress include glyceraldehyde-3-phosphate dehydrogenase (COG G), fructose-bisphosphate aldolase (GOC G), ribosome-recycling factor (COG J), aldolase (COG G), XRE family transcriptional regulator (COG K), 6,7-dimethyl-8-ribityllumazine synthase (COG H), FMN reductase (COG R), and acetyl-CoA acetyltransferase (COG I). Two of the proteins upregulated under salt stress in CcI3, glyceraldehyde-3-phosphate dehydrogenase and fructose-bisphosphate aldolase, show a similar pattern of upregulation in all three strains (CcI3, CcI6, and Allo2).

Table 11. Proteins differentially expressed under stress conditions in strains CcI6 and CcI3. The identified proteins were classified by COG functional categories. More than one protein per spot has been identified for some spots. Upregulated proteins are shown by the upward pointing arrow (↑) whereas downregulated proteins are shown by the downward pointing arrow (↓). No change (N/C) indicates that a spot was not picked for that particular condition because it showed similar intensity as the control.

	SPOT #	Acc. No	Locus Tag	Protein Name	MW (Da)	PI	NaCl	Sucrose
	[C] Energy production and Conversion							
CcI6	10	563313506	CCI6_RS13315	electron transfer flavoprotein alpha subunit apoprotein	32842.70	5.01	↓	↓
	15	563312797	CCI6_RS16290	NAD-dependent aldehyde dehydrogenase	54456.60	5.40	N/C	↓
	1	563313455	CCI6_RS13080	ATP synthase F1 subcomplex beta subunit	56735.20	4.56	↓	N/C
	8	563314955	CCI6_RS05745	malate dehydrogenase (NAD)	34399.10	4.96	↑	N/C
	[E] Amino acid transport and metabolism							
CcI6	2	563312117	CCI6_RS19490	L-glutamine synthetase	53786.20	4.97	↑	N/C
	10	563315075	CCI6_RS05150	cysteine synthase (CysK)	32443.60	4.95	↑	↓
	[G] Carbohydrate transport and metabolism							
CcI6	5	563312326	CCI6_RS18410	glyceraldehyde-3-phosphate dehydrogenase (NAD+)	35515.80	5.76	↑	N/C
	7	563314408	CCI6_RS08915	fructose-bisphosphate aldolase	36894.20	5.35	↑	↑
CcI3	4	WP_011436076.1	FRANCCI3_RS08225	glyceraldehyde-3-phosphate dehydrogenase (NAD+)	35515.80	5.76	↑	N/C
	5	WP_011438718.1	FRANCCI3_RS22085	fructose-bisphosphate aldolase	36894.20	5.35	↑	N/C
	11	WP_011437192.1	FRANCCI3_RS14110	aldolase	38526.7	5.3	N/C	↑
	[H] Coenzyme metabolism							
CcI6	3	563314085	CCI6_RS10325	methionine adenosyltransferase	42893.60	5.05	↑	N/C
	9	563313788	CCI6_RS11580	pyridoxal phosphate synthase yaaD subunit	32577.80	5.34	↑	↑

CcI3	9	WP_011437570.1	FRANCCI3_RS16030	6,7-dimethyl-8-ribityllumazine synthase	16137.8	5.5	↑	N/C
	[I] Lipid transport and metabolism							
CcI3	3	WP_011438063		acetyl-CoA acetyltransferase	39714.5	5.3	↑	N/C
	[J] Translation							
CcI3	8	WP_011437955.1	FRANCCI3_RS18065	ribosome-recycling factor	20830.7	5.4	↑	↑
	[K] Transcription							
CcI6	3	563314632	CCI6_RS07960	DNA-directed RNA polymerase, sigma subunit (sigma70/sigma32)	44399.30	5.10	↑	N/C
	12	563312999	CCI6_RS15595	RNA polymerase, sigma 24 factor	29218.90	5.49	↓	N/C
	13	563314238	CCI6_RS09105	Transcriptional regulator Crp/Fnr	51393.10	4.97	↓	N/C
CcI3	14	WP_011437589.1	FRANCCI3_RS16125	XRE family transcriptional regulator	17993.4	5.7	↑	N/C
	[M] Cell wall/membrane/envelop biogenesis							
CcI6	6	563313716	CCI6_RS12035	UDP-glucose pyrophosphorylase	34856.10	5.04	↑	↑
	11	563315562	CCI6_RS03270	Nucleoside-diphosphate-sugar pyrophosphorylase family protein	31492.90	4.92	↑	N/C
	[O] Post-translational modification, protein turnover, chaperone functions							
CcI6	2	563314374	CCI6_RS08745	chaperonin GroEL	56735.20	4.72	↑	N/C
	14	563311297	CCI6_RS22940	ATP-dependent Clp protease proteolytic subunit ClpP	23039.30	4.79	↑	↑
	[R] General Functional Prediction only							
CcI6	16	563312796	CCI6_RS16285	Zn-dependent alcohol dehydrogenase	35062.50	5.65	N/C	↓
CcI3	7	WP_011434678.1	FRANCCI3_RS01050	FMN reductase	21222.1	5.3	↑	↑
	Not assigned to COG categories							
CcI6	4	563313603	CCI6_RS12255	Nitroreductase	37351.40	4.90	↑	↑

Amino acid analysis confirms predictions of the transcriptome analysis

The transcriptome analysis of strains CcI3 and CcI6 revealed the upregulation of several genes involved in amino acid biosynthesis under salt and osmotic stress conditions, suggesting that amino acids accumulate as osmoprotectants to increase intracellular solute concentration and prevent plasmolysis. To confirm the predictions of the transcriptome results, *Casuarina* isolates, including the two salt-tolerant strains (Allo2 and CcI6), the salt-sensitive strain (CcI3), and the moderately salt-tolerant strain (CeD), were exposed to salt and osmotic stress for seven days and the amino acid profiles were compared to that of the control condition (no stress). The results revealed the strain-specific accumulation of certain amino acids under salt and osmotic stress (Fig 11). Ornithine, GABA, proline, methionine, and glutamine accumulated in strain CcI3 under salt and osmotic stress conditions. Serine was accumulated only under osmotic stress conditions. Transcriptome analysis of strain CcI3 had shown the upregulation under stress of several genes involved in amino acid (proline, ornithine, and serine) biosynthesis. Cysteine accumulated in the two salt-tolerant strains (CcI6 and Allo2) under salt and osmotic stress conditions. The result was in agreement with the transcriptome (strain CcI6) and proteomic (strain CcI6 and Allo2) analysis of the two strains under salt and osmotic stress conditions. Interestingly, proline, the amino acid commonly accumulated under salt stress in a broad spectrum of organisms ranging from bacteria to plants, did not show substantial accumulation in the salt-tolerant strains. Selected amino acids are accumulated by bacteria to counteract the effects of salt and osmotic stress [162]

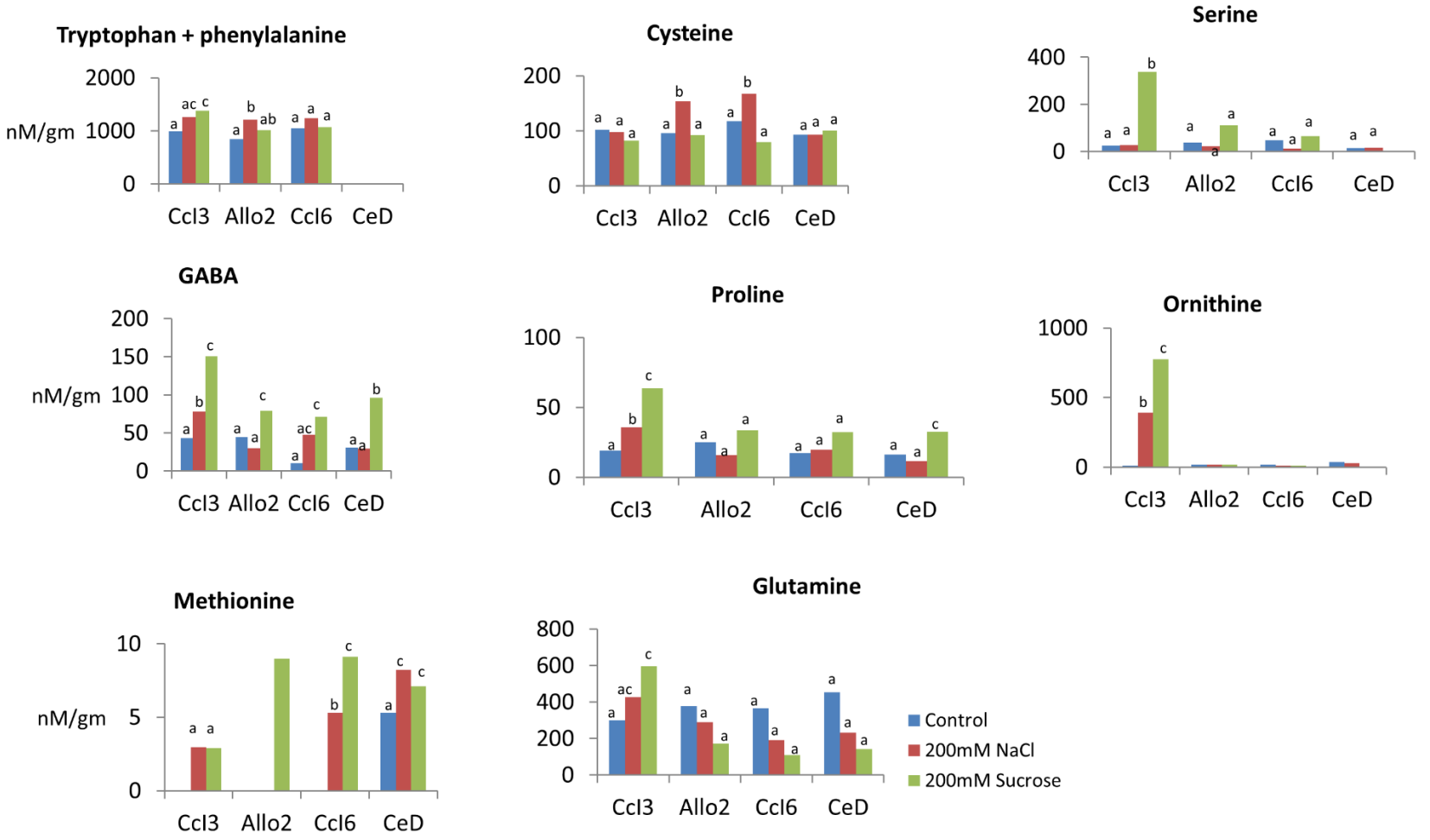

Figure 11. Changes in the amino acid profiles of *Casuarina* isolates exposed to salt and osmotic stress for seven days. Within a single strain, different superscript letters indicate statistically significant differences between treatments (a-c).

Constitutive expression of a zinc peptidase from strain CcI6 led to improved salt stress tolerance in strain CcI3

Through comparative genomic analysis, we identified 153 genes that were present only in the salt-tolerant strains. It was predicted that the observed difference in salt stress tolerance between tolerant and sensitive strains lies in those 153 genes. The expression pattern of the tolerant-strain-specific and other genes under salt and osmotic stress conditions was analyzed through RNA-Seq. Among the 153 tolerant-strain-specific genes, nine and seven were upregulated under salt and osmotic stress conditions, respectively, and became candidates for cloning and expression in the salt-sensitive strain. In order to clone the genes and express them in the salt-sensitive strain, a plasmid (pHTK1) that stably replicates in *Frankia* was identified. pHTK1 had a broad host range pBBR1 origin of replication, *mob* gene, a gene encoding the REP protein needed for replication, a *tetA* gene for selection, and a GFP gene. The pHTK1 was introduced into *Frankia* via conjugation with a DAP^- *E. coli* Bw29427 strain expressing the *tra* genes. After propagating the conjugants for several months in a basal growth medium containing 30 µg/ml of tetracycline, the stable maintenance of pHTK1 inside *Frankia* was confirmed through PCR, restriction analysis, and fluorescent microscopy (Fig 12). The pHTK1 plasmid was modified and the resulting pHTK1RO plasmid, which was constructed so that it contained the phage *pR* and *pL* promoters immediately upstream of the cloning site, was used for the constitutive expression of the candidate tolerant-strain-specific genes in strain CcI3.

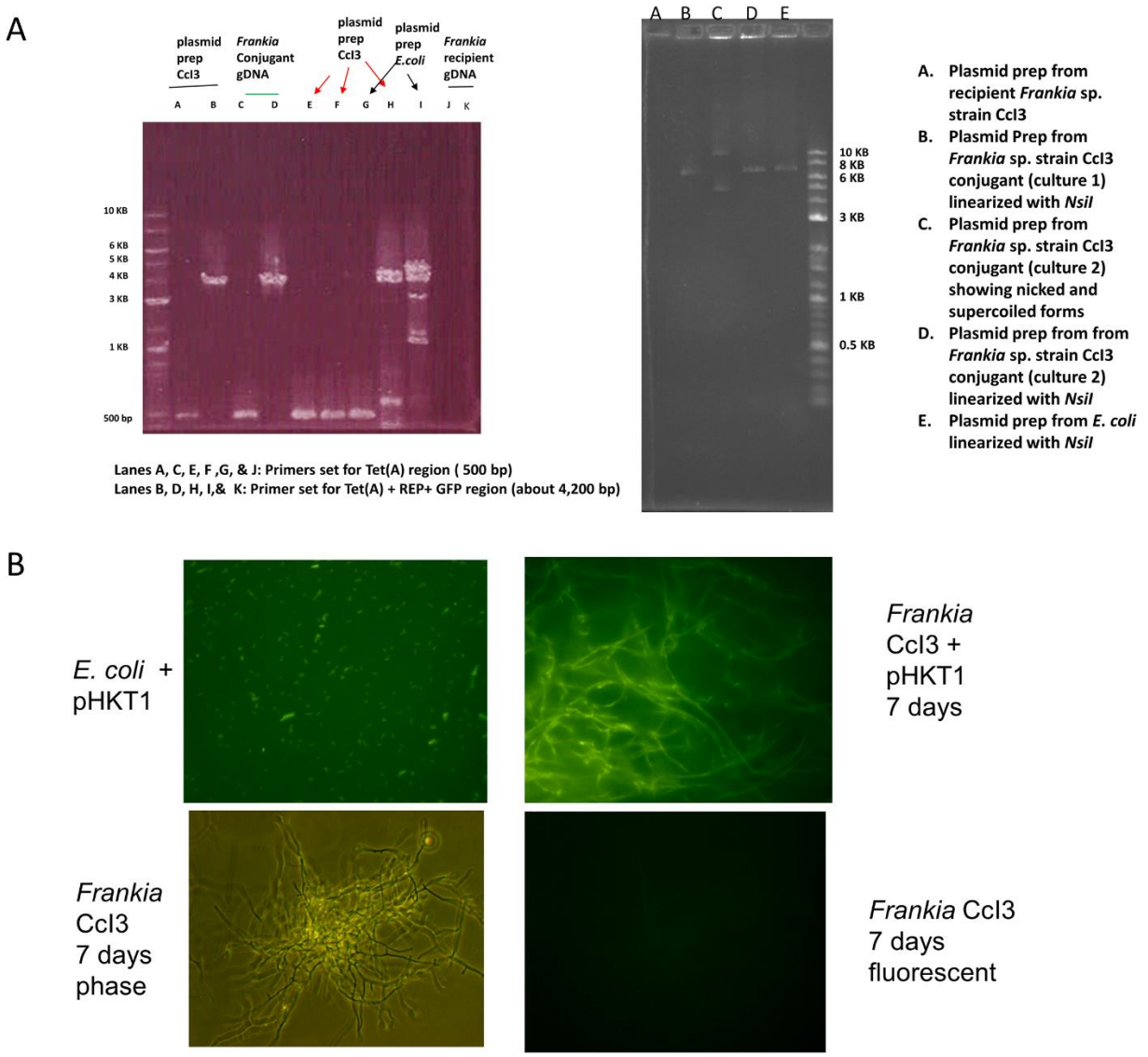

Figure 12. The pHTK1 plasmid is stably maintained in *Frankia*. (A) PCR amplification of regions of the pHTK1 plasmid as well as restriction analysis of plasmid isolated from conjugants propagated for several months in a basal growth medium supplemented with 30 µg/ml of tetracycline confirmed stable maintenance of the plasmid. (B) Fluorescent microscopy showed that the GFP gene on pHTK1 is expressed in *Frankia.*

Salt sensitivity assay on the conjugants showed that, out of the seven candidate genes separately expressed in strain CcI3, a gene encoding zinc peptidase (CCI6_RS22605) caused a reproducible, statistically significant ($p < 0.05$), increase in the MIC, but not MTC, value for NaCl (Fig 13). The NaCl MIC value for the CcI3 conjugants expressing the zinc peptidase gene was 650 mM as opposed to 475 mM for the recipient strain CcI3. The increased salt stress tolerance of conjugants did not extend to osmotic stress tolerance. Transformation with the empty pHTK1RO vector did not change the salt tolerance of CcI3. The MIC value observed for the conjugants expressing the zinc peptidase gene was less than the NaCl MIC values observed for strains CcI6 and Allo2 (1000 mM). Two other genes, MFS transporter (CCI6_RS21730) and a hypothetical protein (CCI6_RS13590) also increased the MIC values of NaCl for strain CcI3 but the results were variable and were not statistically significant at $p < 0.05$ (Fig 13).

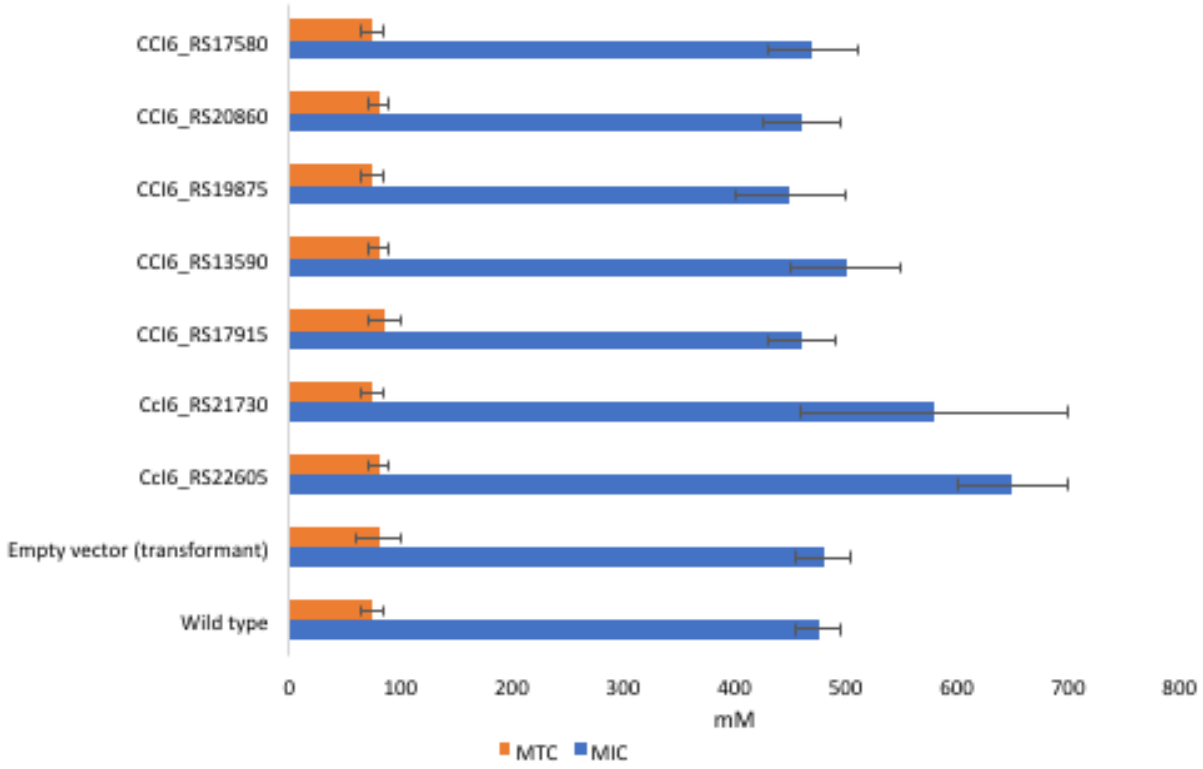

Figure 13. MIC and MTC values of NaCl for strain CcI3 conjugants expressing candidate salt tolerance genes from stain CcI6.

Discussion

Salt tolerance mechanisms depend on the supply of external nitrogen

The salt tolerance level of strains Allo2 and CcI6 (1000 mM) was comparable to the 3 - 15% NaCl tolerance reported for moderately halophilic gram positive bacteria [163]. The salt tolerance level of strain CcI3 (475 mM) was about half of the tolerance level of the two salt-tolerant strains, but was only slightly less than the reported 3% (528 mM) salt tolerance level for *Rhizobium meliloti*, a salt-tolerant rhizobium species [164]. Non-*Casuarina* isolates in general, and strain EuI1C in particular, outperformed the *Casuarina* isolates under osmotic stress imposed by sucrose. The results suggest that *Casuarina* isolates have developed a mechanism to specifically cope with the toxic effects of Na+ and Cl- ions. The more drastic effect of sucrose compared to NaCl treatment is also observed in the comprehensive salt tolerance analysis of *C. crescentus* [162]. One hypothesis is that, given the low carbon source concentration in the natural environment, the bacterium could have been selected for limited or no capability to tolerate high concentration of carbohydrates, as the reduction/absence of the molecular response system obviates the energy demand related to the maintenance of the genetic information and expression of these systems [162]. The difference in salt tolerance between salt-tolerant and salt-sensitive *Casuarina* isolates dissipated under nitrogen-deficient (N_2) conditions, suggesting salt stress response is reliant on the supply of nitrogen sources. The result is expected as many of the osmotic adjustments that take place -from protective osmolyte synthesis to cell envelope remodeling- heavly rely on nitrogen containing compounds. Under nitrogen-deficient conditions, survival of *Frankia* is dependent on the activities of the nitrogenase enzyme, which reduces

atmospheric dinitrogen into ammonia. For all *Casuarina* isolates, nitrogenase activity per milligram of protein and vesicle number per milligram of protein decreased with increasing levels of salinity, causing a compound effect on nitrogen fixation. The fact that nitrogenase activity was affected more severely than overall growth with increasing concentrations of NaCl suggests that under nitrogen-deficient conditions, nitrogenase activity is the limiting factor determining salt tolerance. The salt sensitivity of the nitrogen fixation process and the dependence of salt tolerance mechanisms on the availability of nitrogen also explains why there was no correlation between in vitro salt stress tolerance and symbiotic performance under salt stress conditions.

Comparative genomics suggests a shared salt tolerance mechanism

Based on a cutoff value of 95% used for species delineation [132], the ANI values (>99%) and the AAI values (>98%) observed between any two pairs of *Casuarina* isolates indicate that all *Casuarina* isolates belong to the same species, *Frankia casuarinae*, and are distinct from the closely related cluster Ia isolate *Frankia* alni strain ACN14a. The concatenated phylogenetic affiliation of 394 maximum-parsimony trees based on amino acid sequences also reveals that *Casuarina* isolates group together and are distinct from the closely related cluster Ia isolate strain ACN14a. *In silco* DNA-DNA hybridization confirmed that all *Casuarina* isolates belong to the same species at $p = 0.05$ level, however, no two strains were in the same subspecies at $p = 0.05$ level. The two highly salt-tolerant strains (Allo2 and CcI6) had the lowest genome-to-genome distance, the highest ANI and AAI, suggesting the salt tolerance mechanism of the two strains is shared. The result was further confirmed by the proteomic analysis of the two strains under salt stress which revealed a similar pattern

of differentially expressed proteins. A shared salt-tolerance mechanism with a common origin for the two strains is further supported by pangenome analysis of the *Casuarina* isolates which revealed hundreds of single copy genes that are exclusively shared between the two salt-tolerant strains. Transcriptome analysis revealed some of these tolerant strain-specific genes are responsive to salt and osmotic stress.

Differences between RNA-Seq and Proteome results

After identifying genetic differences between the salt-tolerant and the salt-sensitive strains, we proceeded with transcriptomics and proteomics to determine if the genetic difference includes genes that are responsive to salt and osmotic stress. Because of factors such as half-lives and post transcription machinery, the correlation between mRNA and protein expressions can be low. Therefore, joint analysis of transcriptomic and proteomic data provides useful insight that is otherwise impossible to obtain from individual analysis of mRNA or protein expressions [165]. Qualitative comparison of our proteomics data with the transcriptome data revealed that only 30% of differentially expressed proteins showed the same pattern of expression at the transcriptome level. The remaining proteins were identified only through proteomic analysis. Our results suggest that both transcriptional and posttranscriptional controls are involved in the regulation of genes under hyperosmotic stress in strains CcI6 and CcI3.

Unraveling the enigma: the underlying cause for phenotypic difference

After sequencing salt-tolerant *Casuarina* isolates and identifying their close similarity with the salt-sensitive strain, it was proposed that the observed difference in phenotype boils down to a handful of genes, likely with stress signaling or transcriptional

regulation roles. The hypothesis was further supported by the transcriptome and proteome analysis of strains CcI6 and CcI3 which showed vast differences in transcriptome profile changes under salt stress. The the same osmoadaptation genes behaved differently from one strain to the other under salt and osmotic stress conditions. The comprehensive approach we took to address the research problem helped us to untangle the complex salt stress response and identify at least one of the predicted handful of genes that make the difference between the salt-tolerant and the salt sensitive strains.

All *Casuarina* isolates had the same set of classical genes involved in salt and osmotic stress tolerance, suggesting the observed difference in tolerance was due to previously unknown or less characterized mechanisms. All *Casuarina* isolates lacked the BCCT family transporters, which are present in the closely related strain ACN14a, and the ability to synthesize or acquire glycine betaine (or the precursor choline) from the environment. Strains CcI6 and Allo2 manifest high salt tolerance in a minimal growth medium confirming the idea that the ability to acquire glycine betaine from the environment is a not a key factor in the salt tolerance of *Casuarina* isolates.

The majority of the hundreds of tolerant-strain-specific genes code for hypothetical proteins, suggesting the novelty of the mechanism responsible for the tolerance. The remaining include genes involved in replication, recombination and repair, and cell wall/membrane biogenesis. This would indicate that the ability to maintain the integrity of the genetic material, the replication process, and the cell envelope are all important for salt tolerance. The presence of unique genes in the tolerant strains that are involved in cell wall and membrane biogenesis suggests that

preexisting differences in membrane/cell wall composition and structure might be contributing factors to the observed difference in salt tolerance levels between salt-sensitive and salt-tolerant strains.

Some of the salt-tolerant strain-specific genes had homology with the highly salt-tolerant actinomycete *Nocardiopsis halotolerans*. Enzymes coded by halotolerant bacteria are compatible with high levels of osmolytes accumulated under salt stress condition. Osmoprotectants could inhibit or promote enzyme activity depending on the enzyme [166]. The presence of these genes in the salt-tolerant *Casuarina* isolates suggests that osmolyte-friendly enzymes are one of the several advantages the tolerant strains have over the salt-sensitive strain.

Among the tolerant-strain-specific genes, only a small fraction was responsive to salt and osmotic stress. Among those responsive to salt/osmotic stress, the majority were hypothetical proteins, suggesting that novel mechanisms are responsible for the observed difference in salt-tolerance between strain CcI3 and the two salt-tolerant strains (CcI6 and Allo2). None of the gene products unique to the tolerant strains were identified from our proteomics analysis probably because they were present in amounts below the detection threshold for Coomassie brilliant blue R-250 staining, approximately 100 ng. A gene encoding zinc peptidase was among the salt-tolerant strain-specific genes that were responsive to salt stress. Transcriptome analysis of salt-tolerant and salt-sensitive varieties of rice had a zinc peptidase gene as one of the 50 top responsive genes [167]. Constitutive expression of the zinc peptidase gene in the salt-sensitive strain led to increased salt stress tolerance. To our knowledge, this is the first time that a zinc peptidase gene was shown to confer salt stress tolerance. The zinc peptidase

had a DNA binding domain, suggesting it might play a role in response regulation. Salt stress and the accompanying oxidative stress leads to structural changes that compromise the function of proteins. Misfolded and aggregated proteins are degraded by proteases. Proteases are increasingly being associated with salt tolerance and sensitivity to abiotic stress [168].

It all starts at transcription

The salt-tolerant and sensitive strains showed marked differences in their global response to salt and osmotic stress. Examination of the transcriptome of the two strains revealed that most of the salt-responsive genes behaved in a strain-specific manner, suggesting that part of the difference in phenotype was due to differences in transcriptional regulation under stress. Sensing of the salt-stress stimulus and coupling the signal with a change in transcription is an important part of the salt stress response. The transcriptome analysis of the tolerant and sensitive strains showed the strain-specific expression of several transcriptional factors. The use of alternative sigma factors and transcription regulators creates flexibility in adaptation to environmental stress [169]. Some of the salt and/or osmotic-stress-responsive transcriptional regulators identified in strains CcI6 (the GntR, TetR, LysR, Crp/Fnr, and LuxR families) and CcI3 (WhiB, MarR, LacI, MerR, and XRE families) were previously implicated in multidrug resistance, biosynthesis of antibiotics, osmotic stress response, and pathogenicity of Gram-negative and Gram-positive bacteria. The proteomics analysis did not reveal any transcriptional factors probably because they are low abundance proteins that fell below the detection range. The only sigma factor (CcI6_RS19210) upregulated in strain CcI6 under salt stress had a SigF domain and showed a similar

pattern of upregulation in strain CcI3. We were not able to detect this sigma factor from the proteomics analysis. The sigF regulon in *Mycobacterium smegmatis* mediates stationary phase adaptation and general stress response [170]. It was suggested that *M. smegmatis* SigF regulates the biosynthesis of the osmoprotectant trehalose, as well as an uptake system for osmoregulatory compounds. Our transcriptome analysis showed that trehalose synthase gene is upregulated under salt stress in strain CcI6. The expression of the genes for regulatory proteins SigH3, PhoP, WhiB1, and WhiB4 was suggested as being under direct control of SigF. The WhiB genes showed upregulation in strain CcI3. In both strains CcI3 and CcI6, an extracytoplasmic stress sigma factor (CCI6_RS15595) was downregulated under salt stress. The result is in contrast with the previously reported auto-upregulation of ECF sigma factors in response to extracytoplasmic stress conditions, including salt stress [171]. ECF sigma factors recognize promoter elements with an 'AAC' motif in the -35 regions and are usually co-transcribed with a transmembrane anti-sigma factor with an extracytoplasmic sensory domain and an intracellular inhibitory domain. In strains CcI6 and CcI3, the down-regulated ECF sigma factor lies upstream of a mycothiol system anti-sigma-R factor, suggesting that the ECF sigma factor is a homologue of sigmaR. In *Streptomyces coelicolor* A3(2), sigmaR is regulated by the cognate anti-sigma-R factor (RsrA), which loses affinity for sigmaR following oxidative stress that introduces intramolecular disulfide bond formation in RsrA [172]. ECF sigma factors can be regulated at the transcriptional, translational, and posttranslational levels [171]. The transcriptional control of ECF factors can involve a hierarchical regulatory cascade of sigma factors. The most important regulation of ECF sigma factors involves the

reversible binding of the sigma factor to an anti-sigma factor, holding it in an inactive complex as long as the cognate environmental stimulus is absent. Changes in environmental conditions sensed by the anti-sigma factor lead to release of the sigma factor, which will subsequently bind to the RNA polymerase core enzyme and initiate transcription [173]. Proteomics, but not transcriptomic, analysis of strain CcI6 showed that a σ^{32} factor involved in the cytoplasmic heat shock response is up regulated under salt stress. This agrees with the result from the transcriptome analysis where the small heat-shock protein Hsp20 was upregulated under salt stress. Overall, while the same set of sigma factors were differentially expressed in both the salt-tolerant and the salt-sensitive strains under salt stress, there was a huge variation between the two strains in terms of the number and types of differentially expressed transcriptional factors under salt and osmotic stress conditions. Unlocking the mechanism of salt tolerance in *Casuarina* isolates ultimately requires identifying the factor(s) causing the differential expression of different sets of transcriptional regulators in the two strains. Nevertheless, knowledge of the strain-specific-response of transcriptional regulators helps to answer lower level questions, such as why certain osmolytes (for example ornithine in strain CcI3) were accumulated in a strain-specific manner even though both strains (CcI6 and CcI3) possess the same sets of genes for the biosynthesis of ornithine.

Potential mechanisms of tolerance

The transcriptome analysis of CcI6 and CcI3 showed a clear overlap between salt and osmotic stress responses, but most of the responses were condition-specific. This result partly explains why there is a huge disparity between the salt and osmotic stress tolerance levels among the *Casuarina* isolates. For both strains, and for CcI6 in

particular, a significant number of the genes differentially expressed under salt, osmotic stress, or under both conditions code for hypothetical proteins. In addition to hypothetical proteins, many genes involved in cell wall/membrane biogenesis, compatible solute biosynthesis, and signal transduction were differentially expressed under salt and/or osmotic stress conditions.

Modification of the cell wall

The composition of the cell envelope plays an important role in osmoadaptation [174]. Cell envelope-related changes triggered by salt stress include alterations in the structure and composition of the peptidoglycan layer [175], changes in membrane and/or periplasmic protein composition, lipid composition, periplasmic glucan levels, and capsular polysaccharide biosynthesis [176]. The peptidoglycan layer is a key structural component that imparts structural strength to most bacterial cells and counteracts the effects of osmotic pressure of the cytoplasm. Based on results from expression studies, adaptation to salt stress involves modification of the cell wall to create a diffusion barrier and reduce the influx of inorganic ions into the periplasm [177]. The peptidoglycan is comprised of a linear chain of alternating residues of β-(1,4) linked N-acetylglucosamine and N-acetylmuramic acid cross-linked by peptide side chains. Glycosyl transferases catalyze the polymerization of the peptidoglycan by the addition of disaccharide pentapeptide subunits onto the growing glycan chain. Subsequent cross-linking of the peptide side chains takes place through the activity of transpeptidases [175]. Salt-stress induced alterations of the peptidoglycan layer involve several enzymes including glycosyl transferases, polysaccharide deacetylases and sugar epimerases. Our transcriptome data showed that more glycosyl transferase coding genes (COG M) were

upregulated under salt and osmotic stress conditions in strain CcI6 as opposed to in strain CcI3. In addition to glycosyl transferases, polysaccharide deacetylases involved in cell wall/membrane biogenesis was also upregulated in strain CcI6 in response to salt stress. The polysaccharide deacetylases, which include chitin deacetylases, acetylxylan esterases, xylanases, rhizobial NodB chitooligosaccharide deacetylases, and PG deacetylases, catalyze the hydrolysis of either the N-linked acetyl group from GlcNAc residues (chitin deacetylase, NodB, and peptidoglycan GlcNAc deacetylase) or O-linked acetyl groups from O-acetylxylose residues (acetylxylan esterase, and xylanase) [150]. In *Bacillus anthracis*, one of the polysaccharide deacetylases, BA0330, which is a lipoprotein, plays a structural role in stabilizing the membrane and is important for the adaptation of the bacterium to grow in the presence of a high concentration of salt. The polysaccharide deacetylases upregulated in strain CcI6 under salt stress, however, are not lipoproteins and do not have a transmembrane domain as predicted by LipoP 1.0. The polysaccharide deacetylases upregulated in strain CcI6 contained the catalytic NodB homology domain of the carbohydrate esterase family and had a conserved domain with the bacterial peptidoglycan N-glucose amine deacetylase. Another gene product involved in peptidoglycan layer modification is the sugar epimerase. Several nucleoside diphosphate sugar epimerases were among the genes upregulated under salt and osmotic stress in the salt-tolerant strain. Simple sugars such as glucose, galactose, xylose, arabinose, rhamnose, glucuronic acid, galacuronic acid, mannose, and fructose need to be activated to serve as building blocks for cell components. Nucleotide sugars, which are activated forms of simple sugars, serve as biosynthetic substrates in the synthesis of polysaccharides [178]. Different nucleotide sugars can be

modified at their glycosyl moieties by nucleotide sugar interconversion enzymes to generate different sugars [179]. Different nucleoside-diphosphate (NDP) sugars are used by glycosyl transferases to assemble cell wall polysaccharides with varying compositions [180]. Overall, more genes involved in the modification of the peptidoglycan layer were upregulated in the salt-tolerant strain compared to in the salt-sensitive strain, suggesting the ability to significantly alter the peptidoglycan layer under salt stress is important for tolerance.

Changes in membrane fluidity

Regulating the fluidity of the membrane in response to osmotic stress is an important aspect of cell envelope remodeling during salt stress [181]. Regulation of membrane fluidity mainly involves changes in the fatty acid composition of the membrane by varying the length of acyl chains, number of double bonds or branching of acyl chains by methyl groups [182]. Acyl-acyl carrier protein (ACP) desaturases, ABC-type branched-chain amino acid transport systems, Acyl CoA dehydrogenases, and ubquinone directly or indirectly play a role in modulating the fluidity of the membrane. ACP desaturases catalyze the conversion of saturated fatty acids into unsaturated fatty acids by the introduction of at least one double bond. The Unsaturated fatty acids (UFAs) determine the fluidity and function of biological membranes. Branched chain amino acid transport systems indirectly determine the composition of the cell membrane. Branched chain amino acids (leucine, isoleucine and valine) undergo oxidative deamination and decarboxylation reaction and subsequently serve as precursors in anteiso-branched chain fatty acid synthesis [183]. Acyl CoA dehydrogenase, which catalyze the desaturation of various CoA-conjugated fatty acids,

have been implicated in salt stress tolerance [184, 185] possibly by altering the fluidity of the cell membrane. Biological processes such as selective uptake, transport and signaling performed by membrane spanning intrinsic proteins are affected by the degree of unsaturation and composition of the membrane. Membrane fluidity and composition determines the conformation and geometry of proteins as well as the electrostatic interactions between the protein surface charged domains and the lipid head groups [186]. Adaptation to salt stress involves the active role of membrane bound transporters (such as Na+/H+ antiporter systems), the activity of which is enhanced with increased membrane unsaturation [187]. UFAs homeostasis in many organisms is achieved by feedback regulation of fatty acid desaturase gene transcription through signaling pathways that are governed by sensors embedded in cellular membranes [188]. In addition to membrane desaturation, the accumulation of ubiquinone in the membrane could affect membrane structure and hence salt tolerance [152]. Ubiquinone accumulation has been shown to increase salt tolerance in *E. coli* [152]. The role of ubiquinone in osmotic tolerance results from mechanical stabilization of the cytoplasmic membrane. Presence of ubiquinone in the membrane helps to prevent a drastic volume decrease upon osmotic shock. Deletion in one of the genes involved in ubiquinone biosynthesis, ubiG, had more severe impact than deletion of any one of the genes involved in the synthesis of the osmoprotectant trehalose in *E. coli* [152]. Ubiquinone accumulation led to enhanced salt tolerance during fermentative anaerobic growth, where respiration is inactive, ruling out the possible explanation that the osmotic effect of ubiquinone stems from its known role as a membrane-localized electron carrier in the respiratory chain [152]. The fact that more genes associated with maintaining the fluidity

of the cell membrane were upregulated in the salt-tolerant strain as opposed to in the salt-sensitive strain under salt stress suggests that the ability to maintain the fluidity of the membrane is vital for salt stress tolerance.

Compatible solutes

The transcriptome, proteome, and physiological analysis showed a strain-specific accumulation of osmolytes (mainly amino acids) under salt and osmotic stress conditions. Trehalose synthesis was upregulated in both the tolerant and the sensitive strains although the pathway involved was different for the two; the TreY-TreZ pathway for strain CcI3 and the TreS pathway for strain CcI6. Both the salt-tolerant and the salt-sensitive stains had the same set of canonical osmolyte biosynthesis genes, suggesting that an acquired ability to synthesize a specific osmolyte was not the underlying cause for the difference in salt tolerance. Observed differences between CcI6 and CcI3 in the expression and accumulation of osmolytes are likely ascribed to differences in transcriptional and post transcriptional regulations.

The sugar trehalose and amino acids are among the commonly used compatible solutes [89]. Osmolytes can either be synthesized by the cell or transported into the cell from the medium. Through mechanisms that involve strong exclusion of the protective osmolyte from the protein surface, protective osmolytes push the equilibrium of protein folding towards the native form [189]. Under hyperosmotic conditions, water fluxes out of the cell, causing a loss in turgor pressure and an increase in the cellular concentration of all cellular constituents, including inorganic salts [190]. The resulting changes in cellular volume and turgor pressure exert strong mechanical forces on the cytoplasmic membrane and associated proteins and, if high enough, could lead to

growth arrest and even death of the bacterium [191]. Many inorganic salts are required for cellular functions, however, at concentrations above typically found in cells, they could interfere with the functions of proteins. To restore the cellular volume, cells take in a large quantity of inorganic salts followed by the osmotic uptake of water. The inorganic salts are subsequently replaced by compatible solutes, which are highly soluble neutral molecules [192] with little interference in protein structure and function. Due to their unique properties, compatible solutes can be accumulated to intracellular concentrations greater than 1 M without causing significant disruption in vital cellular processes [193]. Apart from serving as osmotic balancers, compatible solutes also play a key role in stabilizing enzyme function by providing protection against salinity, high temperature, freeze-thaw treatment and even drying [194]. The role of osmolytes transcends maintaining cell turgor by increasing intracellular osmolality. Molecular dynamics studies have demonstrated an interaction between the osmolyte trehalose and the membrane lipid head groups, although the observed resistance of membranes to strong osmotic stress could not fully be ascribed to the interaction [195]. In bacteria and eukaryotes, osmolytes have a neutral charge or are modified to have a neutral charge, whereas in archae the same osmolytes tend to be modified to have negative charge. The increased expression of the trehalose synthase gene only under chronic salt stress and the increased in expression of the threonine and glutamate synthase genes only under osmotics stress (induced by sucrose) suggests that the preferred protective osmolyte depends on the external solute causing the stress.

Signal transduction

Signal transduction is an important component of the salt stress response and coupled with transcriptional regulation we believe it lies at core of the observed difference in salt tolerance. The transcriptome analysis revealed that different sets of signal transduction genes were upregulated in the salt-tolerant and in the salt sensitive strains, however none of those genes were among the genes unique to the salt-tolerant strains. Nevertheless, the zinc peptidase gene that we determined to confer salt tolerance to CcI3 had a DNA binding domain and probably plays a role as a response regulator.

In strain CcI6 the upregulation of the gene coding for PI3P phosphatase, under salt stress indicates PI3P-mediated signaling is important during salt stress phosphoinositides serve a fundamental role in regulating membrane dynamics and intracellular signaling. Actinomycetes are the only group of bacteria which possess a novel PI biosynthesis pathway [196]. In actinomycetes, PI is a major component of the plasma membrane and acts as a membrane anchor for abundant glycolipids, such as PI mannosides (PIMs) and lipoarabinomannan [196]. In *Mycobacterium smegmatis*, PI3P production is induced in response to salt stress, indicating that eukaryote-like lipid-mediated signaling may occur in some bacteria [196]. By regulating the recycling of PI3P, PI3P phosphatase could play a role in PI3P mediated signaling.

Limitations of the methods used in the study

After identifying salt-tolerant (Allo2 and CcI6) and salt-sensitive (CcI3) strains, we used comparative genomics, transcriptomics, and proteomics to identify candidate genes responsible for the observed difference in phenotype. While the RNA-Seq analysis we

used to identify salt stress-responsive genes in the salt-tolerant and in the salt-sensitive strains helped us to identify hundreds of salt and/or osmotic stress-responsive genes, the possibility of bias during library construction meant that some salt-responsive genes were overlooked during the process. After obtaining reads from our RNA-Seq experiments, our mapping was restricted to protein coding sequences. By doing that, we likely missed some salt-responsive, Nonprotein-coding genes, particularly those encoding regulatory RNAs. After identifying tolerant strain-specific, salt-responsive genes, we developed an expression vector that could stably replicate in *Frankia* and expressed the candidate genes in the salt-sensitive strain (CcI3). Through this method, we determined that a gene encoding zinc peptidase (CcI6_RS22605) confers salt-tolerance to CcI3. While we confirmed that the gene ((CcI6_RS22605) was expressed in CcI3, we nevertheless did not examine the expression at the protein level. Tagging of the protein product (zinc peptidase) and studying the expression at the protein level, and analysis of the interaction of the enzyme with other proteins will help to further elucidate the mechanisms involved in the salt stress response. Site-specific mutagenesis of the gene encoding zinc peptidase in CcI6 or Allo2 will provide direct evidence in support of the role of the gene in the salt tolerance of the two strains.

Conclusion

In summary, the comprehensive approach we took to address the complex research question provided answers to some of the questions we originally had, but raised many more unanswered questions. Salt-tolerant and salt-sensitive strains showed very little overlap in their salt stress responses, even though they overall had a high degree of genetic similarity. While genes directly involved in cell envelop

remodeling or osmolyte biosynthesis are unquestionably important for salt stress tolerance, in the case of the *Casuarina* isolates, how the osmoadaptation genes are regulated seems to be more important. Through the comprehensive approach we took, we were able to explain at least some of the factors for the observed difference in salt tolerance between the *Casuarina* isolates. The genetic tools and the data we generated in this study will serve as a springboard for future work in the area or in the broader field of *Frankia* genetics.

APPENDIX A

Differential gene expression in *Casuarina* isolates exposed to salt and osmotic stress

The list of differentially expressed genes in *Frankia* casuarinae and *Frankia* sp. strain CcI6 exposed to salt and osmotic stress is given in Tables A1-A8.

Table A1. List of genes upregulated under salt stress in strain CcI3 based on RNA-sq

Feature ID	Description
aceE	#N/A
argC	#N/A
argD	#N/A
argJ	#N/A
carB	#N/A
Francci3_0040	hypothetical protein
Francci3_0081	citrate synthase 2
Francci3_0082	phosphoserine aminotransferase
Francci3_0178	MscS mechanosensitive ion channel
Francci3_0209	NADPH-dependent FMN reductase
Francci3_0210	MarR family transcriptional regulator
Francci3_0285	NLP/P60
Francci3_0466	tetratricopeptide TPR_2
Francci3_0469	phosphate uptake regulator, PhoU
Francci3_0470	histidine kinase
Francci3_0487	uroporphyrinogen-III synthase / uroporphyrinogen-III C-methyltransferase
Francci3_0523	XRE family transcriptional regulator
Francci3_0525	ferredoxin--nitrite reductase
Francci3_0777	hypothetical protein
Francci3_0816	MMPL
Francci3_0821	amine oxidase
Francci3_0822	polyprenyl synthetase
Francci3_0823	squalene cyclase
Francci3_0933	hypothetical protein
Francci3_1003	hypothetical protein
Francci3_1004	diguanylate cyclase/phosphodiesterase with PAS/PAC sensor(s)
Francci3_1060	hypothetical protein
Francci3_1063	carnitine O-acetyltransferase
Francci3_1064	hypothetical protein
Francci3_1065	hypothetical protein
Francci3_1082	hypothetical protein

Francci3_1153	HNH endonuclease
Francci3_1195	putative transcriptional regulator
Francci3_1216	radical SAM family protein
Francci3_1218	ribonuclease
Francci3_1241	beta-lactamase-like
Francci3_1312	aconitase
Francci3_1313	inositol-1(or 4)-monophosphatase
Francci3_1332	hypothetical protein
Francci3_1341	sulfate adenylyltransferase subunit 2
Francci3_1342	sulfate adenylyltransferase subunit 1
Francci3_1350	glycogen debranching protein GlgX
Francci3_1356	hypothetical protein
Francci3_1408	hypothetical protein
Francci3_1414	cell cycle protein
Francci3_1418	cell division protein FtsZ
Francci3_1440	hypothetical protein
Francci3_1461	NLP/P60
Francci3_1469	MerR family transcriptional regulator
Francci3_1472	daunorubicin resistance ABC transporter ATP-binding subunit
Francci3_1482	transcription factor WhiB
Francci3_1505	metallophosphoesterase
Francci3_1675	enoyl-CoA hydratase
Francci3_1678	pyruvate kinase
Francci3_1679	hypothetical protein
Francci3_1736	type II secretion system protein E
Francci3_1748	GTP cyclohydrolase
Francci3_1749	NUDIX hydrolase
Francci3_1750	3-polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases-like
Francci3_1751	phosphoribosyltransferase
Francci3_1752	GMP reductase
Francci3_1753	radical SAM family protein
Francci3_1754	PfkB
Francci3_1755	putative 6-pyruvoyl tetrahydropterin synthase
Francci3_1765	sulfatase
Francci3_1766	hypothetical protein
Francci3_1767	LacI family transcription regulator
Francci3_1768	major facilitator transporter
Francci3_1769	cytochrome P450
Francci3_1804	ribonucleotide reductase large subunit
Francci3_1805	hypothetical protein
Francci3_1854	GCN5-related N-acetyltransferase
Francci3_1864	transposase IS66
Francci3_1945	(NiFe) hydrogenase maturation protein HypF
Francci3_1998	O-succinylhomoserine sulfhydrylase

Francci3_2208	YceI
Francci3_2312	hypothetical protein
Francci3_2403	amidase
Francci3_2408	SNF2-related
Francci3_2453	prephenate dehydrogenase
Francci3_2454	4-hydroxyphenylpyruvate dioxygenase
Francci3_2455	enoyl-CoA hydratase
Francci3_2456	enoyl-CoA hydratase/isomerase
Francci3_2457	enoyl-CoA hydratase/isomerase
Francci3_2458	chalcone and stilbene synthases-like
Francci3_2462	ABC transporter related
Francci3_2464	chorismate mutase
Francci3_2465	tryptophan halogenase
Francci3_2603	cobyric acid synthase
Francci3_2605	DSH-like
Francci3_2676	helicase-like
Francci3_2804	citrate lyase
Francci3_2816	hypothetical protein
Francci3_2872	delta-1-piperideine-6-carboxylate dehydrogenase
Francci3_3006	methylmalonyl-CoA mutase
Francci3_3022	imidazole glycerol phosphate synthase subunit hisF
Francci3_3023	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase
Francci3_3031	O-methyltransferase family protein
Francci3_3055	hypothetical protein
Francci3_3073	oxidoreductase-like
Francci3_3089	putative esterase/lipase
Francci3_3134	dihydrolipoamide dehydrogenase
Francci3_3139	hypothetical protein
Francci3_3173	acetylglutamate kinase
Francci3_3184	ATPases involved in chromosome partitioning-like
Francci3_3191	primosome assembly protein PriA
Francci3_3194	DNA-directed RNA polymerase subunit omega
Francci3_3229	Male sterility-like
Francci3_3242	hypothetical protein
Francci3_3245	carbonate dehydratase
Francci3_3356	epoxide hydrolase-like
Francci3_3359	EmrB/QacA family drug resistance transporter
Francci3_3405	twin-arginine translocation pathway signal
Francci3_3409	hypothetical protein
Francci3_3410	serine/threonine protein kinase
Francci3_3411	forkhead-associated
Francci3_3414	hypothetical protein
Francci3_3467	succinate dehydrogenase/fumarate reductase iron-sulfur subunit

Francci3_3481	S-malonyltransferase
Francci3_3493	hypothetical protein
Francci3_3507	ribonucleoside diphosphate reductase, B12-dependent
Francci3_3523	metal dependent phosphohydrolase
Francci3_3543	beta-lactamase-like
Francci3_3544	dihydrodipicolinate synthase
Francci3_3663	acetyl-coenzyme A synthetase
Francci3_3667	hypothetical protein
Francci3_3671	hypothetical protein
Francci3_3684	hypothetical protein
Francci3_3754	anti-sigma factor
Francci3_3755	RNA polymerase sigma factor
Francci3_3773	histidinol-phosphate phosphatase
Francci3_3790	transcription factor WhiB
Francci3_3791	hypothetical protein
Francci3_3836	leucyl aminopeptidase
Francci3_3920	Ppx/GppA phosphatase
Francci3_3926	transcription-repair coupling factor
Francci3_3942	serine O-acetyltransferase
Francci3_3943	putative acyl carrier protein
Francci3_3944	FkbH domain-containing protein
Francci3_4035	hypothetical protein
Francci3_4194	methyltransferase type 12
Francci3_4195	hypothetical protein
Francci3_4196	hypothetical protein
Francci3_4197	hypothetical protein
Francci3_4198	hypothetical protein
Francci3_4199	hypothetical protein
Francci3_4200	hypothetical protein
Francci3_4201	ABC transporter related
Francci3_4204	AMP-dependent synthetase and ligase
Francci3_4205	2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase
Francci3_4206	3-dehydroquinate synthase
Francci3_4207	salicylate 1-monooxygenase
Francci3_4230	cyclic nucleotide-binding domain-containing protein
Francci3_4261	two component transcriptional regulator
Francci3_4263	phosphate ABC transporter permease
Francci3_4264	phosphate ABC transporter permease
Francci3_4265	periplasmic phosphate binding protein
Francci3_4365	transglycosylase-like
Francci3_4366	transglycosylase-like
Francci3_4377	L-aspartate oxidase
Francci3_4402	hypothetical protein
gatA	#N/A
mhpA	#N/A

pgk	#N/A
proA	#N/A
sdhA_2	#N/A
tpiA	#N/A

Table A2. List of genes downregulated in strain CcI3 under salt stress based on RNA-seq

Feature ID	description
Francci3_0068	sigma-24
Francci3_0107	hypothetical protein
Francci3_0117	IS630 family transposase
Francci3_0118	IS630 family transposase
Francci3_0126	hypothetical protein
Francci3_0145	SARP family transcriptional regulator
Francci3_0150	putative 3-carboxy-cis,cis-muconate cycloisomerase
Francci3_0153	hypothetical protein
Francci3_0223	DegT/DnrJ/EryC1/StrS aminotransferase
Francci3_0224	NAD-dependent epimerase/dehydratase
Francci3_0238	hypothetical protein
Francci3_0247	hypothetical protein
Francci3_0251	putative sulfonate binding protein precursor
Francci3_0351	short-chain dehydrogenase/reductase SDR
Francci3_0352	luciferase-like
Francci3_0423	NLP/P60
Francci3_0447	thioredoxin-related
Francci3_0514	hypothetical protein
Francci3_0575	pseudouridine synthase
Francci3_0668	hypothetical protein
Francci3_0705	hypothetical protein
Francci3_0804	hypothetical protein
Francci3_0805	glyoxalase/bleomycin resistance protein/dioxygenase
Francci3_0808	hypothetical protein
Francci3_0873	hypothetical protein
Francci3_0875	antibiotic biosynthesis monooxygenase
Francci3_0890	hypothetical protein
Francci3_0908	XRE family transcriptional regulator
Francci3_0909	hypothetical protein
Francci3_0944	short-chain dehydrogenase/reductase SDR
Francci3_0969	hypothetical protein
Francci3_0996	hypothetical protein
Francci3_1017	hypothetical protein
Francci3_1045	hypothetical protein
Francci3_1046	putative DNA-binding protein
Francci3_1047	endopeptidase La
Francci3_1068	hypothetical protein

Francci3_1187	hypothetical protein
Francci3_1235	transposase, IS4
Francci3_1247	hypothetical protein
Francci3_1320	nucleic acid binding, OB-fold, tRNA/helicase-type
Francci3_1321	hypothetical protein
Francci3_1376	(p)ppGpp synthetase I
Francci3_1507	hypothetical protein
Francci3_1572	parallel beta-helix repeat-containing protein
Francci3_1593	AMP-dependent synthetase and ligase
Francci3_1616	hypothetical protein
Francci3_1628	integral membrane protein TerC
Francci3_1676	ABC transporter related
Francci3_1924	hypothetical protein
Francci3_1936	hypothetical protein
Francci3_1989	putative PAS/PAC sensor protein
Francci3_2055	haloacid dehalogenase-like hydrolase
Francci3_2109	hypothetical protein
Francci3_2112	pyridoxamine 5'-phosphate oxidase-related, FMN-binding
Francci3_2163	short-chain dehydrogenase/reductase SDR
Francci3_2183	transposase, IS4
Francci3_2289	major facilitator transporter
Francci3_2326	major intrinsic protein
Francci3_2327	glyoxalase/bleomycin resistance protein/dioxygenase
Francci3_2359	hypothetical protein
Francci3_2368	zinc/iron permease
Francci3_2373	recombinase
Francci3_2436	protein-L-isoaspartate(D-aspartate) O-methyltransferase
Francci3_2468	carbon starvation protein CstA
Francci3_2470	hypothetical protein
Francci3_2522	transposase IS66
Francci3_2716	hypothetical protein
Francci3_2717	ArsR family transcriptional regulator
Francci3_2718	putative integral membrane protein
Francci3_2743	hypothetical protein
Francci3_2746	ribonuclease BN
Francci3_2762	putative methyltransferase
Francci3_2807	aminotransferase, class I and II
Francci3_2808	hypothetical protein
Francci3_2809	hypothetical protein
Francci3_2857	protein of unknown function, ATP binding
Francci3_2883	hypothetical protein
Francci3_2960	Pirin-like
Francci3_3062	hypothetical protein
Francci3_3063	CoA-binding

Francci3_3164	hypothetical protein
Francci3_3247	hypothetical protein
Francci3_3262	hypothetical protein
Francci3_3263	integrase
Francci3_3291	AbrB family transcriptional regulator
Francci3_3292	PilT protein-like
Francci3_3299	hypothetical protein
Francci3_3314	hypothetical protein
Francci3_3322	putative ATP-binding protein
Francci3_3766	hypothetical protein
Francci3_3767	anti-sigma factor
Francci3_3768	RNA polymerase sigma factor RpoE
Francci3_3895	peptide methionine sulfoxide reductase
Francci3_3917	hypothetical protein
Francci3_3986	ABC transporter related
Francci3_3989	recombinase
Francci3_4004	major facilitator transporter
Francci3_4073	hypothetical protein
Francci3_4112	IS630 family transposase
Francci3_4113	IS630 family transposase
Francci3_4154	EmrB/QacA family drug resistance transporter
Francci3_4182	hypothetical protein
Francci3_4190	hypothetical protein
Francci3_4226	hypothetical protein
Francci3_4227	transposase, IS4
Francci3_4464	cytochrome P450
Francci3_4467	protein-L-isoaspartate(D-aspartate) O-methyltransferase
Francci3_4536	thioredoxin reductase
Francci3_4537	thioredoxin

Table A3. List of genes upregulated under osmotic stress in strain CcI3 based on RNA-seq

Feature ID	Description
Francci3_0027	hypothetical protein
Francci3_0505	transposase IS66
Francci3_0881	transposase IS66
Francci3_1947	hydrogenase expression/formation protein HypD
Francci3_2051	transposase IS66
Francci3_2073	transposase IS66
Francci3_2078	transposase IS66
Francci3_2289	major facilitator transporter
Francci3_2450	amino acid adenylation
Francci3_2459	amino acid adenylation
Francci3_2514	zinc-binding alcohol dehydrogenase
Francci3_2515	hypothetical protein
Francci3_2516	amidohydrolase 2
Francci3_2517	monooxygenase component MmoB/DmpM
Francci3_2518	methane/phenol/toluene hydroxylase
Francci3_2519	oxidoreductase FAD-binding region
Francci3_2520	methane monooxygenase
Francci3_2603	cobyric acid synthase
Francci3_2674	N-6 DNA methylase
Francci3_2675	hypothetical protein
Francci3_2676	helicase-like
Francci3_2736	hypothetical protein
Francci3_2897	small multidrug resistance protein
Francci3_4000	transposase IS66
Francci3_4026	AMP-dependent synthetase and ligase
Francci3_4144	short-chain dehydrogenase-like
Francci3_4145	transposase IS66
Francci3_4159	hypothetical protein
Francci3_4160	hypothetical protein
Francci3_4161	hypothetical protein
Francci3_4194	methyltransferase type 12
Francci3_4196	hypothetical protein
Francci3_4204	AMP-dependent synthetase and ligase
Francci3_4205	2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase
Francci3_4206	3-dehydroquinate synthase
Francci3_4365	transglycosylase-like
groEL_3	#N/A

Table A4. List of genes downregulated under osmotic stress in strain CcI3 based on RNA-seq

Feature ID	
Francci3_0035	putative integral membrane protein
Francci3_0118	IS630 family transposase
Francci3_0127	XRE family transcriptional regulator
Francci3_0222	alpha/beta hydrolase fold
Francci3_0226	hypothetical protein
Francci3_0227	hypothetical protein
Francci3_0352	luciferase-like
Francci3_0398	putative DNA-binding protein
Francci3_0447	thioredoxin-related
Francci3_0557	heat shock protein HtpX
Francci3_0827	radical SAM family protein
Francci3_0943	FAD linked oxidase-like
Francci3_0949	hypothetical protein
Francci3_0987	ArsR family transcriptional regulator
Francci3_0991	acyl transferase region
Francci3_0995	sodium/hydrogen exchanger
Francci3_0996	hypothetical protein
Francci3_0997	alpha/beta hydrolase fold
Francci3_0998	hypothetical protein
Francci3_0999	crotonyl-CoA reductase
Francci3_1000	3-hydroxyacyl-CoA dehydrogenase
Francci3_1066	hemerythrin HHE cation binding region
Francci3_1177	putative secreted protein
Francci3_1247	hypothetical protein
Francci3_1320	nucleic acid binding, OB-fold, tRNA/helicase-type
Francci3_1359	sporulation and cell division protein SsgA
Francci3_1507	hypothetical protein
Francci3_1531	hypothetical protein
Francci3_1532	periplasmic binding protein/LacI transcriptional regulator
Francci3_1533	phosphoenolpyruvate phosphomutase
Francci3_1544	alcohol dehydrogenase GroES-like protein
Francci3_1545	hypothetical protein
Francci3_1592	extracellular solute-binding protein
Francci3_1627	putative transmembrane alanine and leucine rich protein
Francci3_1628	integral membrane protein TerC
Francci3_1659	putative transcriptional regulator
Francci3_1660	FeS assembly protein SufB
Francci3_1661	FeS assembly protein SufD

Francci3_1662	Rieske (2Fe-2S) protein
Francci3_1914	BFD-like (2Fe-2S)-binding region
Francci3_1924	hypothetical protein
Francci3_1989	putative PAS/PAC sensor protein
Francci3_2027	pyruvate:ferredoxin (flavodoxin) oxidoreductase
Francci3_2082	TetR family transcriptional regulator
Francci3_2183	transposase, IS4
Francci3_2257	hypothetical protein
Francci3_2326	major intrinsic protein
Francci3_2327	glyoxalase/bleomycin resistance protein/dioxygenase
Francci3_2368	zinc/iron permease
Francci3_2717	ArsR family transcriptional regulator
Francci3_2746	ribonuclease BN
Francci3_2807	aminotransferase, class I and II
Francci3_2808	hypothetical protein
Francci3_2809	hypothetical protein
Francci3_2862	hypothetical protein
Francci3_2873	L-lysine aminotransferase
Francci3_2888	haloacid dehalogenase-like hydrolase
Francci3_2943	hypothetical protein
Francci3_2990	hypothetical protein
Francci3_3125	TetR family transcriptional regulator
Francci3_3126	hypothetical protein
Francci3_3164	hypothetical protein
Francci3_3252	hypothetical protein
Francci3_3253	two component LuxR family transcriptional regulator
Francci3_3254	putative transcriptional regulator
Francci3_3261	hypothetical protein
Francci3_3262	hypothetical protein
Francci3_3403	iron permease FTR1
Francci3_3511	small GTP-binding protein domain-containing protein
Francci3_3768	RNA polymerase sigma factor RpoE
Francci3_4003	BadM/Rrf2 family transcriptional regulator
Francci3_4051	putative serine/threonine kinase anti-sigma factor
Francci3_4073	hypothetical protein
Francci3_4113	IS630 family transposase
Francci3_4181	hypothetical protein
Francci3_4182	hypothetical protein
Francci3_4227	transposase, IS4
Francci3_4536	thioredoxin reductase

Table A5. List of genes upregulated under salt stress in strain CcI6 based on RNA-seq

Gene ID	Description
CCI6_RS00160	inosine 5-monophosphate dehydrogenase
CCI6_RS00205	ribosomal-protein-alanine N-acetyltransferase RimI
CCI6_RS00330	preprotein translocase subunit SecY
CCI6_RS00475	DNA-directed RNA polymerase subunit beta
CCI6_RS00495	50S ribosomal protein L11
CCI6_RS00935	
CCI6_RS00940	thioredoxin family protein
CCI6_RS01525	acetyl-/propionyl-CoA carboxylase subunit alpha
CCI6_RS01555	hypothetical protein
CCI6_RS01570	transcriptional regulator
CCI6_RS01640	glycosyl transferase family 1
CCI6_RS01730	phosphatidylinositol-3-phosphate phosphatase
CCI6_RS01875	hypothetical protein
CCI6_RS01925	histidine kinase
CCI6_RS01945	hypothetical protein
CCI6_RS01965	peptidase S15
CCI6_RS02035	SsrA-binding protein
CCI6_RS02115	hypothetical protein
CCI6_RS02250	chromosome partitioning protein
CCI6_RS02325	glycosyl transferase
CCI6_RS02330	
CCI6_RS02375	hypothetical protein
CCI6_RS02420	hypothetical protein
CCI6_RS02550	Crp/Fnr family transcriptional regulator
CCI6_RS02565	coenzyme F420-reducing hydrogenase subunit alpha
CCI6_RS02600	nitrogenase molybdenum-iron protein alpha chain
CCI6_RS02810	glycosyl hydrolase family 15
CCI6_RS02820	endoribonuclease L-PSP
CCI6_RS02950	hypothetical protein
CCI6_RS02955	hypothetical protein
CCI6_RS03270	glucose-1-phosphate cytidylyltransferase
CCI6_RS03475	adenylate kinase
CCI6_RS03540	polysaccharide deacetylase
CCI6_RS03865	taurine dioxygenase
CCI6_RS03870	nitrate ABC transporter substrate-binding protein
CCI6_RS03980	hypothetical protein
CCI6_RS04030	hypothetical protein
CCI6_RS04150	type I restriction-modification system methyltransferase subunit
CCI6_RS04505	histidine--tRNA ligase
CCI6_RS04525	nucleoside-diphosphate sugar epimerase

CCI6_RS04715	signal peptidase I
CCI6_RS05110	hypothetical protein
CCI6_RS05135	acyl-CoA dehydrogenase
CCI6_RS05150	cysteine synthase A
CCI6_RS05375	2-phosphoglycerate kinase
CCI6_RS05400	hypothetical protein
CCI6_RS05575	site-specific DNA-methyltransferase
CCI6_RS05860	hypothetical protein
CCI6_RS05975	ATPase
CCI6_RS06240	FMN reductase
CCI6_RS06605	nucleotidyltransferase
CCI6_RS06650	methyltransferase, FxLD system
CCI6_RS06730	
CCI6_RS06895	glucans biosynthesis protein C
CCI6_RS06950	hypothetical protein
CCI6_RS07140	metallophosphoesterase
CCI6_RS07280	ABC transporter
CCI6_RS07410	Zn-dependent hydrolase
CCI6_RS07440	cytotoxic translational repressor of toxin-antitoxin stability system
CCI6_RS07540	membrane protein
CCI6_RS07600	rod shape-determining protein MreC
CCI6_RS07660	hypothetical protein
CCI6_RS07915	isoprenyl transferase
CCI6_RS07965	glycosyl transferase family 1
CCI6_RS07970	hypothetical protein
CCI6_RS08005	epimerase
CCI6_RS08015	hypothetical protein
CCI6_RS08090	type I-E CRISPR-associated protein Cse1/CasA
CCI6_RS08215	hypothetical protein
CCI6_RS08355	hypothetical protein
CCI6_RS08395	
CCI6_RS08450	amidophosphoribosyltransferase
CCI6_RS08505	hypothetical protein
CCI6_RS08595	haloacid dehalogenase
CCI6_RS08725	hypothetical protein
CcI6_RS08745	molecular chaperone GroEL
CCI6_RS08915	class II fructose-bisphosphate aldolase
CCI6_RS09145	ABC transporter substrate-binding protein
CCI6_RS09155	ABC transporter permease
CCI6_RS09180	23S rRNA (guanosine(2251)-2\'-O)-methyltransferase RlmB
CCI6_RS09240	DNA integrity scanning protein DisA
CCI6_RS09525	haloacid dehalogenase
CCI6_RS09910	hypothetical protein
CCI6_RS10260	methylmalonyl-CoA mutase

CCI6_RS10340	ferredoxin
CCI6_RS10380	ABC transporter
CCI6_RS10580	hypothetical protein, partial
CCI6_RS10620	branched-chain amino acid ABC transporter substrate-binding protein
CCI6_RS10625	multidrug ABC transporter ATP-binding protein
CCI6_RS10640	hypothetical protein
CCI6_RS10740	hypothetical protein
CCI6_RS10750	luciferase-like protein
CCI6_RS10760	hypothetical protein
CCI6_RS10910	glycosyl transferase
CCI6_RS10925	pyruvate carboxylase
CCI6_RS10965	acyl-ACP desaturase
CCI6_RS11030	hypothetical protein
CCI6_RS11155	polysaccharide deacetylase
CCI6_RS11270	glucose-1-phosphate cytidylyltransferase
CCI6_RS11495	hypothetical protein
CCI6_RS11580	pyridoxal biosynthesis lyase PdxS
CCI6_RS11600	Holliday junction DNA helicase RuvA
CCI6_RS11670	Mg-chelatase subunit ChlD
CCI6_RS11890	hypothetical protein
CCI6_RS11955	tRNA-specific adenosine deaminase
CCI6_RS12030	molybdopterin molybdenumtransferase MoeA
CCI6_RS12035	UDP-glucose pyrophosphorylase
CCI6_RS12055	hypothetical protein
CCI6_RS12100	diaminopimelate decarboxylase
CCI6_RS12135	RND superfamily drug exporter
CCI6_RS12160	glycine dehydrogenase
CCI6_RS12175	hypothetical protein
CCI6_RS12185	hypothetical protein
CCI6_RS12225	
CCI6_RS12310	pyridine nucleotide-disulfide oxidoreductase
CCI6_RS12470	hypothetical protein
CCI6_RS12500	hypothetical protein
CCI6_RS12535	transcriptional regulator
CCI6_RS12800	proteasome accessory factor PafA2
CCI6_RS12900	putative transcriptional regulator
CCI6_RS13215	trehalose synthase
CCI6_RS13230	hypothetical protein
CCI6_RS13590	hypothetical protein
CCI6_RS13650	oxidoreductase
CCI6_RS13655	hypothetical protein
CCI6_RS13965	hypothetical protein
CCI6_RS13985	amino acid/amide ABC transporter substrate-binding protein, HAAT family
CCI6_RS14005	hypothetical protein

CCI6_RS14305	non-ribosomal peptide synthetase
CCI6_RS14525	6-phosphofructokinase
CCI6_RS14550	
CCI6_RS14740	phosphodiesterase/alkaline phosphatase D
CCI6_RS14820	hypothetical protein
CCI6_RS14850	hypothetical protein
CCI6_RS14960	hypothetical protein
CCI6_RS15145	phosphoserine phosphatase
CCI6_RS15305	TetR family transcriptional regulator
CCI6_RS15330	hypothetical protein
CCI6_RS15530	hypothetical protein
CCI6_RS15620	IS110 family transposase
CCI6_RS15625	hypothetical protein
CCI6_RS15785	hypothetical protein
CCI6_RS15820	radical SAM protein
CCI6_RS15880	hypothetical protein
CCI6_RS15990	hypothetical protein
CCI6_RS16015	hypothetical protein
CCI6_RS16210	quercetin 2,3-dioxygenase
CCI6_RS16255	prephenate dehydratase
CCI6_RS16295	hypothetical protein
CCI6_RS16550	riboflavin synthase subunit alpha
CCI6_RS16730	peroxidase
CCI6_RS16810	sporulation protein SsgA
CCI6_RS17225	glucose-6-phosphate dehydrogenase
CCI6_RS17310	membrane protein
CCI6_RS17520	hypothetical protein
CCI6_RS17580	hypothetical protein
CCI6_RS17660	ubiquinone biosynthesis protein UbiE
CCI6_RS17915	ADP-heptose--LPS heptosyltransferase
CCI6_RS18130	DUF1298 domain-containing protein
CCI6_RS18210	
CCI6_RS18295	DNA-binding protein
CCI6_RS18410	type I glyceraldehyde-3-phosphate dehydrogenase
CCI6_RS18525	hypothetical protein
CCI6_RS18550	heat-shock protein Hsp20
CCI6_RS18600	LysR family transcriptional regulator
CCI6_RS18715	ATP-binding protein
CCI6_RS18745	hypothetical protein
CCI6_RS18765	ADP-ribose pyrophosphatase
CCI6_RS18975	hypothetical protein
CCI6_RS18980	Sec-independent protein translocase TatB
CCI6_RS19085	pyoverdine biosynthesis protein
CCI6_RS19210	RNA polymerase sigma factor
CCI6_RS19225	dTDP-4-dehydrorhamnose 3,5-epimerase
CCI6_RS19280	aminotransferase

CCI6_RS19305	FAD-dependent oxidoreductase
CCI6_RS19510	lipoyl synthase
CCI6_RS19585	phage terminase large subunit
CCI6_RS19875	hypothetical protein
CCI6_RS19940	hypothetical protein
CCI6_RS19950	aspartate aminotransferase family protein
CCI6_RS19985	
CCI6_RS20030	methyltransferase, FxLD system
CCI6_RS20095	hypothetical protein
CCI6_RS20100	DDE transposase
CCI6_RS20335	disulfide bond formation protein DsbA
CCI6_RS20385	hypothetical protein
CCI6_RS20460	transcriptional regulator, GntR family, partial
CCI6_RS20525	hydantoinase
CCI6_RS20740	hypothetical protein
CCI6_RS20860	hypothetical protein
CCI6_RS20960	hypothetical protein
CCI6_RS21235	aldo/keto reductase
CCI6_RS21440	pterin-4-alpha-carbinolamine dehydratase
CCI6_RS21635	type I restriction-modification system methyltransferase subunit
CCI6_RS21670	acyltransferase
CCI6_RS21730	MFS transporter
CCI6_RS21895	glycosyl transferase family 1
CCI6_RS22005	
CCI6_RS22150	hypothetical protein

Table A6. List of genes downregulated under salt stress in strain CcI6 according to RNA-seq analysis

Gene ID	Description
CCI6_RS00055	hypothetical protein
CCI6_RS00100	hypothetical protein
CCI6_RS00240	NAD(P)H-hydrate dehydratase
CCI6_RS00250	PadR family transcriptional regulator
CCI6_RS00290	50S ribosomal protein L17
CCI6_RS00335	50S ribosomal protein L15
CCI6_RS00485	50S ribosomal protein L10
CCI6_RS00560	orotate phosphoribosyltransferase
CCI6_RS00590	2-oxoacid ferredoxin oxidoreductase subunit beta
CCI6_RS00605	NADH-quinone oxidoreductase subunit N
CCI6_RS00635	NADH-quinone oxidoreductase subunit H
CCI6_RS00705	serine/threonine protein kinase
CCI6_RS00810	cytochrome C biogenesis protein CcdA
CCI6_RS00865	hypothetical protein
CCI6_RS00965	hypothetical protein
CCI6_RS00985	hypothetical protein
CCI6_RS01050	transcriptional regulator
CCI6_RS01130	hypothetical protein
CCI6_RS01145	anti-sigma factor antagonist
CCI6_RS01185	shikimate dehydrogenase
CCI6_RS01195	putative phosphatase
CCI6_RS01235	hypothetical protein
CCI6_RS01245	glycoside hydrolase
CCI6_RS01435	succinate dehydrogenase, cytochrome b556 subunit
CCI6_RS01480	phosphoglucomutase
CCI6_RS01485	purine-nucleoside phosphorylase
CCI6_RS01595	transcriptional regulator
CCI6_RS01615	NAD(P)-dependent oxidoreductase
CCI6_RS01620	transcriptional regulator, TetR family
CCI6_RS01705	glycosyl transferase family 1
CCI6_RS01710	glycosyl transferase family 1
CCI6_RS01830	ATPase AAA
CCI6_RS01885	hypothetical protein
CCI6_RS02010	uroporphyrinogen-III C-methyltransferase
CCI6_RS02075	
CCI6_RS02265	N-acetylmuramoyl-L-alanine amidase
CCI6_RS02305	MFS transporter
CCI6_RS02345	30S ribosomal protein S6
CCI6_RS02495	guanylate kinase
CCI6_RS02515	hypothetical protein
CCI6_RS02545	hypothetical protein
CCI6_RS02650	heme biosynthesis protein HemY
CCI6_RS02665	2-oxoglutarate ferredoxin oxidoreductase subunit alpha

CCI6_RS02735	adenosine deaminase
CCI6_RS02850	cell division protein FtsI
CCI6_RS02930	helix-turn-helix transcriptional regulator
CCI6_RS02965	hypothetical protein
CCI6_RS03100	ribose-phosphate pyrophosphokinase
CCI6_RS03115	fatty acid desaturase
CCI6_RS03200	hypothetical protein
CCI6_RS03205	putative enzyme involved in methoxymalonyl-ACP biosynthesis
CCI6_RS03580	NUDIX hydrolase
CCI6_RS03635	
CCI6_RS03660	TIGR03085 family protein
CCI6_RS03670	serine/threonine phosphatase
CCI6_RS03695	glycoside hydrolase
CCI6_RS03740	hypothetical protein
CCI6_RS03745	hypothetical protein
CCI6_RS03900	hypothetical protein
CCI6_RS04105	hypothetical protein
CCI6_RS04110	hypothetical protein
CCI6_RS04250	amidase
CCI6_RS04260	hypothetical protein
CCI6_RS04270	glycosyl transferase family 1
CCI6_RS04340	hypothetical protein
CCI6_RS04445	DNA starvation/stationary phase protection protein
CCI6_RS04490	acetate kinase
CCI6_RS04530	hypothetical protein
CCI6_RS04540	4-hydroxy-tetrahydrodipicolinate reductase
CCI6_RS04615	
CCI6_RS04645	1-deoxy-D-xylulose-5-phosphate reductoisomerase
CCI6_RS04740	30S ribosomal protein S16
CCI6_RS04895	hypothetical protein
CCI6_RS04935	branched chain amino acid aminotransferase
CCI6_RS05020	hypothetical protein
CCI6_RS05060	alpha/beta hydrolase
CCI6_RS05340	PMT family glycosyltransferase 4-amino-4-deoxy-L-arabinose transferase
CCI6_RS05440	two-component sensor histidine kinase
CCI6_RS05705	resolvase
CCI6_RS05745	malate dehydrogenase
CCI6_RS05865	LysR family transcriptional regulator
CCI6_RS05925	protein-tyrosine-phosphatase
CCI6_RS05940	polyketide synthase regulator
CCI6_RS06050	acetyltransferase
CCI6_RS06090	LexA repressor
CCI6_RS06100	AraC family transcriptional regulator
CCI6_RS06105	glyoxalase

CCI6_RS06130	tRNA (N6-isopentenyl adenosine(37)-C2)-methylthiotransferase MiaB
CCI6_RS06235	LLM class F420-dependent oxidoreductase
CCI6_RS06340	hypothetical protein
CCI6_RS06365	hypothetical protein, partial
CCI6_RS06645	hypothetical protein
CCI6_RS06750	
CCI6_RS06875	hypothetical protein
CCI6_RS06930	exodeoxyribonuclease III
CCI6_RS06980	dihydroorotate oxidase
CCI6_RS07125	stress response protein, TerZ- and CABP1
CCI6_RS07275	hypothetical protein
CCI6_RS07480	hypothetical protein
CCI6_RS07625	radical SAM protein
CCI6_RS07690	VWA domain-containing protein
CCI6_RS07745	plasmid partitioning protein
CCI6_RS07795	
CCI6_RS07835	competence protein ComEC
CCI6_RS07910	DNA repair protein RecO
CCI6_RS07955	DNA primase
CCI6_RS07990	
CCI6_RS08035	hypothetical protein
CCI6_RS08055	glycosyltransferase WbuB
CCI6_RS08065	oxidoreductase
CCI6_RS08185	hypothetical protein
CCI6_RS08195	twin-arginine translocation pathway signal protein
CCI6_RS08210	anti-sigma factor antagonist
CCI6_RS08220	hypothetical protein
CCI6_RS08340	hypothetical protein
CCI6_RS08380	TIR domain-containing protein
CCI6_RS08430	hypothetical protein
CCI6_RS08435	DUF2530 domain-containing protein
CCI6_RS08475	DNA-binding protein
CCI6_RS08545	
CCI6_RS08605	cold-shock protein
CCI6_RS08655	phosphoribosylformylglycinamidine synthase II
CCI6_RS08660	phosphoribosylformylglycinamidine synthase I
CCI6_RS08770	hypothetical protein
CCI6_RS08775	hypothetical protein
CCI6_RS08805	serine/threonine protein phosphatase
CCI6_RS08825	Fe-S oxidoreductase
CCI6_RS08835	type III pantothenate kinase
CCI6_RS08850	Aspartate 1-decarboxylase 2
CCI6_RS08980	nuclease
CCI6_RS09030	hypothetical protein
CCI6_RS09045	hypothetical protein

CCI6_RS09085	ATP-binding protein
CCI6_RS09105	Crp/Fnr family transcriptional regulator
CCI6_RS09140	MFS transporter
CCI6_RS09170	formylmethionine deformylase
CCI6_RS09195	hypothetical protein
CCI6_RS09220	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
CCI6_RS09230	hypothetical protein
CCI6_RS09410	secretion system protein
CCI6_RS09440	anti-sigma factor
CCI6_RS09570	hypothetical protein
CCI6_RS09655	acetyltransferase
CCI6_RS09720	CoA ester lyase
CCI6_RS09740	
CCI6_RS09770	hypothetical protein
CCI6_RS09885	McrBC 5-methylcytosine restriction system component-like protein
CCI6_RS09940	actinorhodin polyketide beta-ketoacyl synthase
CCI6_RS09945	3-oxoacyl-ACP synthase
CCI6_RS10085	MarR family transcriptional regulator
CCI6_RS10115	membrane protein
CCI6_RS10255	hypothetical protein
CCI6_RS10290	membrane protein
CCI6_RS10315	hypothetical protein
CCI6_RS10430	allophanate hydrolase
CCI6_RS10440	hypothetical protein
CCI6_RS10445	transcriptional regulator
CCI6_RS10520	hypothetical protein
CCI6_RS10600	hypothetical protein
CCI6_RS10645	cystathionine beta-synthase
CCI6_RS10690	hypothetical protein
CCI6_RS10840	ferredoxin subunit of nitrite reductase and ring-hydroxylating dioxygenase
CCI6_RS10865	VWA domain-containing protein
CCI6_RS10950	transcriptional regulator
CCI6_RS11025	GGDEF domain-containing protein
CCI6_RS11255	sugar nucleotide processing enzyme
CCI6_RS11355	acetyltransferase
CCI6_RS11385	type 12 methyltransferase
CCI6_RS11435	sulfate adenylyltransferase subunit 2
CCI6_RS11445	hypothetical protein
CCI6_RS11850	cell division protein FtsZ, partial
CCI6_RS11975	transcriptional regulator
CCI6_RS12255	nitroreductase
CCI6_RS12280	FAD/FMN-dependent dehydrogenase
CCI6_RS12300	acyl-CoA dehydrogenase
CCI6_RS12330	hypothetical protein

CCI6_RS12350	alpha/beta hydrolase
CCI6_RS12380	
CCI6_RS12425	electron transporter SenC
CCI6_RS12480	hypothetical protein
CCI6_RS12665	precorrin-6A synthase (deacetylating)
CCI6_RS12710	hypothetical protein
CCI6_RS12775	RecB family exonuclease
CCI6_RS12805	prokaryotic ubiquitin-like protein Pup
CCI6_RS12835	
CCI6_RS12905	WYL domain-containing protein
CCI6_RS13080	ATP synthase subunit beta
CCI6_RS13100	cob(I)yrinic acid a,c-diamide adenosyltransferase
CCI6_RS13155	hypothetical protein
CCI6_RS13175	hypothetical protein
CCI6_RS13185	
CCI6_RS13190	two-component sensor histidine kinase
CCI6_RS13265	hypothetical protein
CCI6_RS13385	hypothetical protein
CCI6_RS13400	IS66 family transposase
CCI6_RS13495	hypothetical protein
CCI6_RS13670	dipeptidyl aminopeptidase/acylaminoacyl peptidase
CCI6_RS13680	carboxylate--amine ligase
CCI6_RS13730	hypothetical protein
CCI6_RS13755	hypothetical protein
CCI6_RS13795	ABC transporter permease
CCI6_RS13820	transcriptional regulator, TetR family
CCI6_RS13835	squalene synthase HpnD
CCI6_RS13840	amine oxidase
CCI6_RS13950	hypothetical protein
CCI6_RS13980	hypothetical protein
CCI6_RS14040	
CCI6_RS14485	hypothetical protein
CCI6_RS14785	hypothetical protein
CCI6_RS14935	aminopeptidase N
CCI6_RS14970	transposase
CCI6_RS14995	acyl-CoA dehydrogenase
CCI6_RS15310	hypothetical protein
CCI6_RS15355	RecB family exonuclease
CCI6_RS15365	hypothetical protein
CCI6_RS15375	chromosome partitioning protein
CCI6_RS15410	Non-homologous end joining protein Ku
CCI6_RS15420	adenylyltransferase/sulfurtransferase MoeZ
CCI6_RS15595	RNA polymerase sigma24 factor
CCI6_RS15595	RNA polymerase sigma24 factor
CCI6_RS15685	
CCI6_RS15760	hypothetical protein

CCI6_RS15770	ATP-binding protein
CCI6_RS15795	hypothetical protein
CCI6_RS15970	heat-shock protein Hsp20
CCI6_RS16005	hypothetical protein
CCI6_RS16060	epoxide hydrolase
CCI6_RS16090	hypothetical protein
CCI6_RS16160	methyltransferase
CCI6_RS16270	NAD(P) transhydrogenase subunit beta
CCI6_RS16305	potassium transporter TrkA
CCI6_RS16350	2-nitropropane dioxygenase
CCI6_RS16450	3-dehydroquinate synthase
CCI6_RS16480	dihydroorotase
CCI6_RS16590	RNA methyltransferase
CCI6_RS16635	arginine repressor
CCI6_RS16645	argininosuccinate lyase
CCI6_RS16740	ATPase
CCI6_RS16750	threonine--tRNA ligase
CCI6_RS16800	aldo/keto reductase
CCI6_RS16935	hypothetical protein
CCI6_RS17060	hypothetical protein
CCI6_RS17095	hypothetical protein
CCI6_RS17125	glucose-1-phosphate adenylyltransferase
CCI6_RS17185	ABC transporter
CCI6_RS17230	glucose-6-phosphate dehydrogenase
CCI6_RS17270	isopentenyl-diphosphate Delta-isomerase
CCI6_RS17275	biotin synthase BioB
CCI6_RS17425	hypothetical protein
CCI6_RS17650	hypothetical protein
CCI6_RS17735	glucokinase
CCI6_RS17890	transcriptional regulator
CCI6_RS18005	polysaccharide deacetylase familiy protein
CCI6_RS18040	hypothetical protein
CCI6_RS18240	uroporphyrinogen decarboxylase
CCI6_RS18275	hypothetical protein
CCI6_RS18370	UDP-N-acetyl-D-glucosamine dehydrogenase
CCI6_RS18450	glycosyl transferase
CCI6_RS18480	peptidase M4 family protein
CCI6_RS18515	hypothetical protein
CCI6_RS18590	AMP-dependent synthetase
CCI6_RS18705	hypothetical protein
CCI6_RS19000	aldehyde dehydrogenase
CCI6_RS19065	alpha/beta hydrolase
CCI6_RS19185	signal transduction histidine kinase
CCI6_RS19220	acetyltransferase
CCI6_RS19470	NAD+ synthase
CCI6_RS19495	RDD family protein

CCI6_RS19550	transcriptional regulator
CCI6_RS19630	hypothetical protein
CCI6_RS19640	hypothetical protein
CCI6_RS19735	site-specific recombinase DNA invertase Pin
CCI6_RS19780	FAD-binding monooxygenase
CCI6_RS19825	hypothetical protein
CCI6_RS20080	methylenetetrahydrofolate reductase
CCI6_RS20140	urate oxidase
CCI6_RS20175	hypothetical protein
CCI6_RS20245	hypothetical protein
CCI6_RS20270	NUDIX hydrolase
CCI6_RS20455	glutamate--tRNA ligase
CCI6_RS20610	hypothetical protein
CCI6_RS20650	hypothetical protein, partial
CCI6_RS20795	hypothetical protein
CCI6_RS20965	
CCI6_RS21000	phage shock protein A
CCI6_RS21005	hypothetical protein
CCI6_RS21045	hydantoinase/oxoprolinase
CCI6_RS21240	hypothetical protein
CCI6_RS21545	type VI secretion protein
CCI6_RS21675	hypothetical protein
CCI6_RS21815	glutamate racemase
CCI6_RS21940	hypothetical protein
CCI6_RS22000	transposase
CCI6_RS22200	hemolysin
CCI6_RS22385	hypothetical protein
CCI6_RS22490	NADPH:quinone reductase
CCI6_RS22545	hypothetical protein
CCI6_RS22565	hypothetical protein
CCI6_RS22615	site-specific recombinase XerD
CCI6_RS22875	DDE transposase family protein
CCI6_RS22915	hypothetical protein
CCI6_RS23195	
thrS	threonine--tRNA ligase
valS	valine--tRNA ligase
Gene ID	Description
CCI6_RS00055	hypothetical protein
CCI6_RS00100	hypothetical protein
CCI6_RS00240	NAD(P)H-hydrate dehydratase
CCI6_RS00250	PadR family transcriptional regulator
CCI6_RS00290	50S ribosomal protein L17
CCI6_RS00335	50S ribosomal protein L15
CCI6_RS00485	50S ribosomal protein L10
CCI6_RS00560	orotate phosphoribosyltransferase
CCI6_RS00590	2-oxoacid ferredoxin oxidoreductase subunit beta

CCI6_RS00605	NADH-quinone oxidoreductase subunit N
CCI6_RS00635	NADH-quinone oxidoreductase subunit H
CCI6_RS00705	serine/threonine protein kinase
CCI6_RS00810	cytochrome C biogenesis protein CcdA
CCI6_RS00865	hypothetical protein

Table A7. List of genes upregulated under osmotic stress in strain CcI6 based on RNA-seq

Gene ID	Description
CCI6_RS00005	hypothetical protein
CCI6_RS00035	phosphoesterase
CCI6_RS00090	succinyl-CoA ligase subunit beta
CCI6_RS00140	twin-arginine translocation pathway signal protein
CCI6_RS00205	ribosomal-protein-alanine N-acetyltransferase RimI
CCI6_RS00300	30S ribosomal protein S4
CCI6_RS00325	adenylate kinase
CCI6_RS00410	30S ribosomal protein S19
CCI6_RS00475	DNA-directed RNA polymerase subunit beta
CCI6_RS00505	preprotein translocase subunit SecE
CCI6_RS00525	hypothetical protein
CCI6_RS00595	hypothetical protein
CCI6_RS00610	NADH-quinone oxidoreductase subunit M
CCI6_RS00630	NADH-quinone oxidoreductase subunit I
CCI6_RS00635	NADH-quinone oxidoreductase subunit H
CCI6_RS00705	serine/threonine protein kinase
CCI6_RS00850	methyltransferase
CCI6_RS00900	methionine--tRNA ligase
CCI6_RS00910	heavy metal transport/detoxification protein
CCI6_RS00960	nucleoside-diphosphate sugar epimerase
CCI6_RS01075	hypothetical protein
CCI6_RS01160	hypothetical protein
CCI6_RS01215	hypothetical protein
CCI6_RS01245	glycoside hydrolase
CCI6_RS01285	Kef-type K+ transport system, predicted NAD-binding component
CCI6_RS01385	
CCI6_RS01515	NAD(P)H-quinone dehydrogenase
CCI6_RS01775	mannose-6-phosphate isomerase
CCI6_RS01810	RDD family protein
CCI6_RS01930	pilus assembly protein CpaF
CCI6_RS01980	hypothetical protein
CCI6_RS02120	hypothetical protein
CCI6_RS02275	thioredoxin-disulfide reductase
CCI6_RS02325	glycosyl transferase
CCI6_RS02360	50S ribosomal protein L9
CCI6_RS02550	Crp/Fnr family transcriptional regulator
CCI6_RS02640	NifZ protein
CCI6_RS02670	2-oxoacid ferredoxin oxidoreductase subunit beta

CCI6_RS02750	hypothetical protein
CCI6_RS02845	serine/threonine protein kinase
CCI6_RS02955	hypothetical protein
CCI6_RS02985	molybdate ABC transporter permease
CCI6_RS03245	hypothetical protein
CCI6_RS03285	hypothetical protein
CCI6_RS03485	NUDIX hydrolase
CCI6_RS03675	hypothetical protein
CCI6_RS03910	molecular chaperone HtpG
CCI6_RS04080	transcriptional regulator
CCI6_RS04090	hypothetical protein
CCI6_RS04120	ABC transporter
CCI6_RS04165	(2Fe-2S)-binding protein
CCI6_RS04460	CDP-alcohol phosphatidyltransferase
CCI6_RS04570	exopolyphosphatase-like enzyme
CCI6_RS04620	sulfate ABC transporter ATP-binding protein
CCI6_RS04795	hypothetical protein
CCI6_RS04975	acetolactate synthase small subunit
CCI6_RS05060	alpha/beta hydrolase
CCI6_RS05355	hypothetical protein
CCI6_RS05425	hypothetical protein, partial
CCI6_RS05570	hypothetical protein
CCI6_RS05635	hypothetical protein
CCI6_RS05830	hypothetical protein
CCI6_RS05860	hypothetical protein
CCI6_RS05870	succinate dehydrogenase
CCI6_RS05965	propionyl-CoA carboxylase subunit beta
CCI6_RS05975	ATPase
CCI6_RS06010	hypothetical protein
CCI6_RS06085	peptidoglycan-binding protein LysM
CCI6_RS06140	ABC transporter substrate-binding protein
CCI6_RS06405	hypothetical protein
CCI6_RS06485	K+-transporting ATPase subunit F
CCI6_RS06615	hypothetical protein
CCI6_RS06760	
CCI6_RS06915	hypothetical protein
CCI6_RS07025	acetyl-CoA carboxylase carboxyltransferase subunit alpha/beta
CCI6_RS07105	hypothetical protein
CCI6_RS07130	protein disulfide-isomerase
CCI6_RS07155	nucleic acid-binding protein
CCI6_RS07420	serine/threonine protein kinase
CCI6_RS07470	hypothetical protein
CCI6_RS07600	rod shape-determining protein MreC
CCI6_RS07615	rod shape-determining protein RodA
CCI6_RS07635	hypothetical protein

CCI6_RS07650	GTPase Obg
CCI6_RS07745	plasmid partitioning protein
CCI6_RS07770	
CCI6_RS08005	epimerase
CCI6_RS08090	type I-E CRISPR-associated protein Cse1/CasA
CCI6_RS08255	hypothetical protein
CCI6_RS08260	hypothetical protein
CCI6_RS08320	hypothetical protein
CCI6_RS08335	large mechanosensitive ion channel protein MscL
CCI6_RS08375	pyridoxine/pyridoxamine 5\'-phosphate oxidase
CCI6_RS08430	hypothetical protein
CCI6_RS08750	threonine synthase
CCI6_RS08995	hypothetical protein
CCI6_RS09110	hypothetical protein
CCI6_RS09145	ABC transporter substrate-binding protein
CCI6_RS09155	ABC transporter permease
CCI6_RS09260	phosphate ABC transporter, permease protein PstA
CCI6_RS09280	hypothetical protein
CCI6_RS09330	ATPase
CCI6_RS09565	hypothetical protein
CCI6_RS09625	pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase component subunit alpha
CCI6_RS09860	ABC transporter substrate-binding protein
CCI6_RS09995	enoyl-CoA hydratase
CCI6_RS10090	hydrolase or acyltransferase of alpha/beta superfamily
CCI6_RS10225	glutamate synthase
CCI6_RS10340	ferredoxin
CCI6_RS10700	hypothetical protein
CCI6_RS10740	hypothetical protein
CCI6_RS10755	short-chain dehydrogenase
CCI6_RS10805	branched chain amino acid ABC transporter substrate-binding protein
CCI6_RS10900	DNA-binding response regulator
CCI6_RS10910	glycosyl transferase
CCI6_RS10965	acyl-ACP desaturase
CCI6_RS10990	hypothetical protein
CCI6_RS11040	hypothetical protein
CCI6_RS11060	phosphopantetheine-binding protein
CCI6_RS11125	hypothetical protein
CCI6_RS11160	glycosyl transferase
CCI6_RS11195	glycosyl transferase
CCI6_RS11220	glycosyl transferase
CCI6_RS11335	3-hydroxybutyryl-CoA dehydrogenase
CCI6_RS11400	membrane protein
CCI6_RS11405	helix-turn-helix transcriptional regulator

CCI6_RS11580	pyridoxal biosynthesis lyase PdxS
CCI6_RS11990	site-specific recombinase DNA invertase Pin
CCI6_RS12035	UDP-glucose pyrophosphorylase
CCI6_RS12060	hypothetical protein
CCI6_RS12160	glycine dehydrogenase
CCI6_RS12165	hypothetical protein
CCI6_RS12225	
CCI6_RS12255	nitroreductase
CCI6_RS12535	transcriptional regulator
CCI6_RS12560	transcriptional regulator
CCI6_RS12630	cytochrome P450
CCI6_RS12745	carboxylate--amine ligase
CCI6_RS12845	ATPase
CCI6_RS12940	ABC transporter ATP-binding protein
CCI6_RS13050	ATP synthase subunit A
CCI6_RS13555	hypothetical protein
CCI6_RS13590	hypothetical protein
CCI6_RS13650	oxidoreductase
CCI6_RS13885	urease accessory protein UreD
CCI6_RS14275	dimethylmenaquinone methyltransferase
CCI6_RS14320	hypothetical protein
CCI6_RS14425	menaquinol-cytochrome c reductase cytochrome b subunit
CCI6_RS14435	glyoxalase
CCI6_RS14670	hypothetical protein
CCI6_RS14905	hypothetical protein
CCI6_RS14930	hypothetical protein
CCI6_RS15030	ABC transporter permease
CCI6_RS15755	hypothetical protein
CCI6_RS16040	hypothetical protein
CCI6_RS16145	hypothetical protein
CCI6_RS16380	membrane protein
CCI6_RS16425	aminodeoxychorismate lyase
CCI6_RS16455	elongation factor P
CCI6_RS16505	hypothetical protein
CCI6_RS16675	hypothetical protein
CCI6_RS16810	sporulation protein SsgA
CCI6_RS16960	ABC transporter
CCI6_RS16970	hypothetical protein
CCI6_RS17055	LuxR family transcriptional regulator
CCI6_RS17465	hemerythrin
CCI6_RS17545	death-on-curing protein
CCI6_RS17635	MFS transporter
CCI6_RS17670	B12-binding domain-containing radical SAM protein
CCI6_RS17795	ribonucleotide reductase

CCI6_RS17890	transcriptional regulator
CCI6_RS17905	acetyltransferase (isoleucine patch superfamily)
CCI6_RS18285	1-deoxy-D-xylulose-5-phosphate synthase
CCI6_RS18295	DNA-binding protein
CCI6_RS18305	portal protein
CCI6_RS18315	nucleotide-binding protein
CCI6_RS18425	RNase adaptor protein RapZ
CCI6_RS18445	Zn-dependent hydrolase
CCI6_RS18480	peptidase M4 family protein
CCI6_RS18555	hypothetical protein
CCI6_RS18645	Ni,Fe-hydrogenase maturation factor
CCI6_RS18720	hypothetical protein
CCI6_RS18740	hypothetical protein
CCI6_RS18745	hypothetical protein
CCI6_RS18760	arylsulfatase regulator (Fe-S oxidoreductase)
CCI6_RS18855	hypothetical protein
CCI6_RS18980	Sec-independent protein translocase TatB
CCI6_RS19025	ABC transporter substrate-binding protein
CCI6_RS19090	pyoverdine biosynthesis
CCI6_RS19135	hypothetical protein
CCI6_RS19330	radical SAM family protein
CCI6_RS19505	hypothetical protein
CCI6_RS19680	site-specific integrase
CCI6_RS19880	hypothetical protein
CCI6_RS19925	salicylate 1-monooxygenase
CCI6_RS19940	hypothetical protein
CCI6_RS19950	aspartate aminotransferase family protein
CCI6_RS20190	dCMP deaminase
CCI6_RS20335	disulfide bond formation protein DsbA
CCI6_RS20420	
CCI6_RS20595	hypothetical protein
CCI6_RS20735	GNAT family N-acetyltransferase
CCI6_RS20750	
CCI6_RS21155	hypothetical protein
CCI6_RS21235	aldo/keto reductase
CCI6_RS21305	hypothetical protein
CCI6_RS21525	LuxR family transcriptional regulator, partial
CCI6_RS21685	hypothetical protein
CCI6_RS21770	DnaK antisense family putative NAD-specific glutamate dehydrogenase
CCI6_RS21870	hypothetical protein
CCI6_RS21895	glycosyl transferase family 1
CCI6_RS22005	
CCI6_RS22030	MFS transporter
CCI6_RS22130	multidrug ABC transporter ATPase/permease
CCI6_RS22620	endonuclease

CCI6_RS22660	hypothetical protein
CCI6_RS22705	hypothetical protein
CCI6_RS22765	hypothetical protein
CCI6_RS22820	hypothetical protein
CCI6_RS22850	
CCI6_RS22910	hypothetical protein
clpP_2	ATP-dependent Clp protease proteolytic subunit
guaA	GMP synthetase
prfB	peptide chain release factor 2
pyrH	UMP kinase
sufC	ABC transporter ATP-binding protein
Gene ID	Description
CCI6_RS00005	hypothetical protein
CCI6_RS00035	phosphoesterase
CCI6_RS00090	succinyl-CoA ligase subunit beta
CCI6_RS00140	twin-arginine translocation pathway signal protein
CCI6_RS00205	ribosomal-protein-alanine N-acetyltransferase RimI
CCI6_RS00300	30S ribosomal protein S4
CCI6_RS00325	adenylate kinase
CCI6_RS00410	30S ribosomal protein S19
CCI6_RS00475	DNA-directed RNA polymerase subunit beta
CCI6_RS00505	preprotein translocase subunit SecE
CCI6_RS00525	hypothetical protein
CCI6_RS00595	hypothetical protein
CCI6_RS00610	NADH-quinone oxidoreductase subunit M
CCI6_RS00630	NADH-quinone oxidoreductase subunit I
CCI6_RS00635	NADH-quinone oxidoreductase subunit H
CCI6_RS00705	serine/threonine protein kinase
CCI6_RS00850	methyltransferase
CCI6_RS00900	methionine--tRNA ligase
CCI6_RS00910	heavy metal transport/detoxification protein
CCI6_RS00960	nucleoside-diphosphate sugar epimerase
CCI6_RS01075	hypothetical protein
CCI6_RS01160	hypothetical protein
CCI6_RS01215	hypothetical protein
CCI6_RS01245	glycoside hydrolase
CCI6_RS01285	Kef-type K+ transport system, predicted NAD-binding component
CCI6_RS01385	
CCI6_RS01515	NAD(P)H-quinone dehydrogenase
CCI6_RS01775	mannose-6-phosphate isomerase
CCI6_RS01810	RDD family protein
CCI6_RS01930	pilus assembly protein CpaF
CCI6_RS01980	hypothetical protein
CCI6_RS02120	hypothetical protein
CCI6_RS02275	thioredoxin-disulfide reductase

CCI6_RS02325	glycosyl transferase
CCI6_RS02360	50S ribosomal protein L9
CCI6_RS02550	Crp/Fnr family transcriptional regulator
CCI6_RS02640	NifZ protein
CCI6_RS02670	2-oxoacid ferredoxin oxidoreductase subunit beta
CCI6_RS02750	hypothetical protein
CCI6_RS02845	serine/threonine protein kinase
CCI6_RS02955	hypothetical protein
CCI6_RS02985	molybdate ABC transporter permease
CCI6_RS03245	hypothetical protein
CCI6_RS03285	hypothetical protein
CCI6_RS03485	NUDIX hydrolase
CCI6_RS03675	hypothetical protein
CCI6_RS03910	molecular chaperone HtpG
CCI6_RS04080	transcriptional regulator
CCI6_RS04090	hypothetical protein
CCI6_RS04120	ABC transporter
CCI6_RS04165	(2Fe-2S)-binding protein
CCI6_RS04460	CDP-alcohol phosphatidyltransferase
CCI6_RS04570	exopolyphosphatase-like enzyme
CCI6_RS04620	sulfate ABC transporter ATP-binding protein
CCI6_RS04795	hypothetical protein
CCI6_RS04975	acetolactate synthase small subunit
CCI6_RS05060	alpha/beta hydrolase
CCI6_RS05355	hypothetical protein
CCI6_RS05425	hypothetical protein, partial
CCI6_RS05570	hypothetical protein
CCI6_RS05635	hypothetical protein
CCI6_RS05830	hypothetical protein
CCI6_RS05860	hypothetical protein
CCI6_RS05870	succinate dehydrogenase
CCI6_RS05965	propionyl-CoA carboxylase subunit beta
CCI6_RS05975	ATPase
CCI6_RS06010	hypothetical protein
CCI6_RS06085	peptidoglycan-binding protein LysM
CCI6_RS06140	ABC transporter substrate-binding protein
CCI6_RS06405	hypothetical protein
CCI6_RS06485	K+-transporting ATPase subunit F
CCI6_RS06615	hypothetical protein

Table A8. List of genes downregulated under osmotic stress in strain CcI6 based on RNA-seq

Gene ID	Description
CCI6_RS00485	50S ribosomal protein L10
CCI6_RS00560	orotate phosphoribosyltransferase
CCI6_RS00730	hypothetical protein
CCI6_RS00830	glutamate-1-semialdehyde 2,1-aminomutase
CCI6_RS01360	
CCI6_RS01615	NAD(P)-dependent oxidoreductase
CCI6_RS01705	glycosyl transferase family 1
CCI6_RS01740	coenzyme F420-0:L-glutamate ligase
CCI6_RS01760	7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase
CCI6_RS01850	amidophosphoribosyltransferase
CCI6_RS01885	hypothetical protein
CCI6_RS01910	
CCI6_RS02160	hypothetical protein
CCI6_RS02185	hypothetical protein
CCI6_RS02380	dihydroxy-acid dehydratase
CCI6_RS02655	thiamine biosynthesis protein ThiF
CCI6_RS02660	cysteine desulfurase
CCI6_RS02715	hypothetical protein
CCI6_RS02905	hypothetical protein
CCI6_RS02920	hypothetical protein
CCI6_RS03040	hypothetical protein
CCI6_RS03090	16S rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase
CCI6_RS03120	glycerophosphoryl diester phosphodiesterase
CCI6_RS03205	putative enzyme involved in methoxymalonyl-ACP biosynthesis
CCI6_RS03355	hypothetical protein
CCI6_RS03385	PadR family transcriptional regulator
CCI6_RS03580	NUDIX hydrolase
CCI6_RS03660	TIGR03085 family protein
CCI6_RS03745	hypothetical protein
CCI6_RS03820	aspartate-semialdehyde dehydrogenase
CCI6_RS03855	hypothetical protein
CCI6_RS03900	hypothetical protein
CCI6_RS03980	hypothetical protein
CCI6_RS04110	hypothetical protein
CCI6_RS04470	hypothetical protein
CCI6_RS04530	hypothetical protein
CCI6_RS04540	4-hydroxy-tetrahydrodipicolinate reductase
CCI6_RS05010	amino acid-binding protein
CCI6_RS05075	ABC-type antimicrobial peptide transport system,

	permease component
CCI6_RS05115	histidine kinase
CCI6_RS05170	ABC transporter substrate-binding protein
CCI6_RS05215	FAD/FMN-dependent dehydrogenase
CCI6_RS05330	xanthine dehydrogenase
CCI6_RS05510	hypothetical protein
CCI6_RS05705	resolvase
CCI6_RS05800	galactose-1-phosphate uridylyltransferase
CCI6_RS05895	amino acid dehydrogenase
CCI6_RS06130	tRNA (N6-isopentenyl adenosine(37)-C2)-methylthiotransferase MiaB
CCI6_RS06175	DEAD/DEAH box helicase
CCI6_RS06340	hypothetical protein
CCI6_RS06365	hypothetical protein, partial
CCI6_RS06780	serine phosphatase
CCI6_RS07020	acetyl/propionyl-CoA carboxylase subuit alpha
CCI6_RS07275	hypothetical protein
CCI6_RS07400	hypothetical protein
CCI6_RS07560	
CCI6_RS07655	glutamate 5-kinase
CCI6_RS07830	DNA-binding protein
CCI6_RS07845	hypothetical protein
CCI6_RS07950	deoxyguanosinetriphosphate triphosphohydrolase
CCI6_RS07970	hypothetical protein
CCI6_RS07990	
CCI6_RS08195	twin-arginine translocation pathway signal protein
CCI6_RS08435	DUF2530 domain-containing protein
CCI6_RS08650	hypothetical protein
CCI6_RS08805	serine/threonine protein phosphatase
CCI6_RS08870	adenylosuccinate lyase
CCI6_RS08890	hypothetical protein
CCI6_RS09115	nitroreductase
CCI6_RS09140	MFS transporter
CCI6_RS09220	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
CCI6_RS09275	hypothetical protein
CCI6_RS09410	secretion system protein
CCI6_RS09420	hypothetical protein
CCI6_RS09740	
CCI6_RS09770	hypothetical protein
CCI6_RS10480	hypothetical protein
CCI6_RS10520	hypothetical protein
CCI6_RS10785	ABC transporter ATP-binding protein
CCI6_RS10840	ferredoxin subunit of nitrite reductase and ring-hydroxylating dioxygenase
CCI6_RS10885	hypothetical protein
CCI6_RS10905	two-component sensor histidine kinase

CCI6_RS10950	transcriptional regulator
CCI6_RS11355	acetyltransferase
CCI6_RS11385	type 12 methyltransferase
CCI6_RS11465	malto-oligosyltrehalose trehalohydrolase
CCI6_RS11685	phytoene dehydrogenase
CCI6_RS11710	phosphatidylglycerophosphate synthase
CCI6_RS11735	phosphoglycerate dehydrogenase-like oxidoreductase
CCI6_RS11975	transcriptional regulator
CCI6_RS12050	hypothetical protein
CCI6_RS12070	hypothetical protein
CCI6_RS12190	hypothetical protein
CCI6_RS12255	nitroreductase
CCI6_RS12330	hypothetical protein
CCI6_RS12710	hypothetical protein
CCI6_RS12775	RecB family exonuclease
CCI6_RS12835	
CCI6_RS12905	WYL domain-containing protein
CCI6_RS13065	ATP synthase subunit delta
CCI6_RS13175	hypothetical protein
CCI6_RS13190	two-component sensor histidine kinase
CCI6_RS13465	WD40 repeat-containing protein
CCI6_RS13680	carboxylate--amine ligase
CCI6_RS13840	amine oxidase
CCI6_RS13930	allantoinase
CCI6_RS13950	hypothetical protein
CCI6_RS13980	hypothetical protein
CCI6_RS14465	long-chain-fatty-acid--CoA ligase
CCI6_RS14485	hypothetical protein
CCI6_RS14515	hypothetical protein
CCI6_RS14580	oxidoreductase
CCI6_RS14620	MBL fold hydrolase
CCI6_RS14675	hypothetical protein
CCI6_RS15035	carbohydrate ABC transporter ATP-binding protein, CUT1 family
CCI6_RS15070	hypothetical protein
CCI6_RS15195	hypothetical protein
CCI6_RS15340	magnesium transporter
CCI6_RS15555	gamma-aminobutyraldehyde dehydrogenase
CCI6_RS15940	MaoC family dehydratase
CCI6_RS15950	transporter
CCI6_RS16305	potassium transporter TrkA
CCI6_RS16345	beta-hydroxydecanoyl-ACP dehydratase
CCI6_RS16355	acetyl-CoA carboxylase carboxyltransferase subunits alpha/beta
CCI6_RS16530	replication restart DNA helicase PriA
CCI6_RS16590	RNA methyltransferase

CCI6_RS16635	arginine repressor
CCI6_RS16750	threonine--tRNA ligase
CCI6_RS16755	hypothetical protein
CCI6_RS17120	glycosyl transferase family 1
CCI6_RS17215	protoheme IX farnesyltransferase
CCI6_RS17230	glucose-6-phosphate dehydrogenase
CCI6_RS17270	isopentenyl-diphosphate Delta-isomerase
CCI6_RS17300	hypothetical protein, partial
CCI6_RS17385	hypothetical protein
CCI6_RS17430	NHL repeat-containing protein
CCI6_RS17735	glucokinase
CCI6_RS18005	polysaccharide deacetylase familiy protein
CCI6_RS18150	pyruvate dehydrogenase (acetyl-transferring) E1 component subunit alpha
CCI6_RS18270	3-hydroxyacyl-CoA dehydrogenase
CCI6_RS18290	23S rRNA methyltransferase
CCI6_RS18665	
CCI6_RS18695	cold-shock protein
CCI6_RS18945	leucyl aminopeptidase
CCI6_RS19065	alpha/beta hydrolase
CCI6_RS19410	DNA repair protein RecN
CCI6_RS19470	NAD+ synthase
CCI6_RS19480	glutamine-synthetase adenylyltransferase
CCI6_RS19495	RDD family protein
CCI6_RS19630	hypothetical protein
CCI6_RS19640	hypothetical protein
CCI6_RS19815	ABC transporter ATP-binding protein
CCI6_RS20020	NAD-dependent epimerase
CCI6_RS20475	monooxygenase
CCI6_RS20610	hypothetical protein
CCI6_RS20795	hypothetical protein
CCI6_RS21005	hypothetical protein
CCI6_RS21240	hypothetical protein
CCI6_RS21530	
CCI6_RS21730	MFS transporter
CCI6_RS22000	transposase
CCI6_RS22490	NADPH:quinone reductase
CCI6_RS22545	hypothetical protein
cobN	cobaltochelatase subunit CobN
truB	tRNA pseudouridine(55) synthase TruB
Gene ID	Description
CCI6_RS00485	50S ribosomal protein L10
CCI6_RS00560	orotate phosphoribosyltransferase
CCI6_RS00730	hypothetical protein
CCI6_RS00830	glutamate-1-semialdehyde 2,1-aminomutase
CCI6_RS01360	

CCI6_RS01615	NAD(P)-dependent oxidoreductase
CCI6_RS01705	glycosyl transferase family 1
CCI6_RS01740	coenzyme F420-0:L-glutamate ligase
CCI6_RS01760	7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase
CCI6_RS01850	amidophosphoribosyltransferase
CCI6_RS01885	hypothetical protein
CCI6_RS01910	
CCI6_RS02160	hypothetical protein
CCI6_RS02185	hypothetical protein
CCI6_RS02380	dihydroxy-acid dehydratase
CCI6_RS02655	thiamine biosynthesis protein ThiF
CCI6_RS02660	cysteine desulfurase
CCI6_RS02715	hypothetical protein
CCI6_RS02905	hypothetical protein
CCI6_RS02920	hypothetical protein
CCI6_RS03040	hypothetical protein
CCI6_RS03090	16S rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase
CCI6_RS03120	glycerophosphoryl diester phosphodiesterase
CCI6_RS03205	putative enzyme involved in methoxymalonyl-ACP biosynthesis
CCI6_RS03355	hypothetical protein
CCI6_RS03385	PadR family transcriptional regulator
CCI6_RS03580	NUDIX hydrolase
CCI6_RS03660	TIGR03085 family protein
CCI6_RS03745	hypothetical protein
CCI6_RS03820	aspartate-semialdehyde dehydrogenase
CCI6_RS03855	hypothetical protein
CCI6_RS03900	hypothetical protein
CCI6_RS03980	hypothetical protein

APPENDIX B

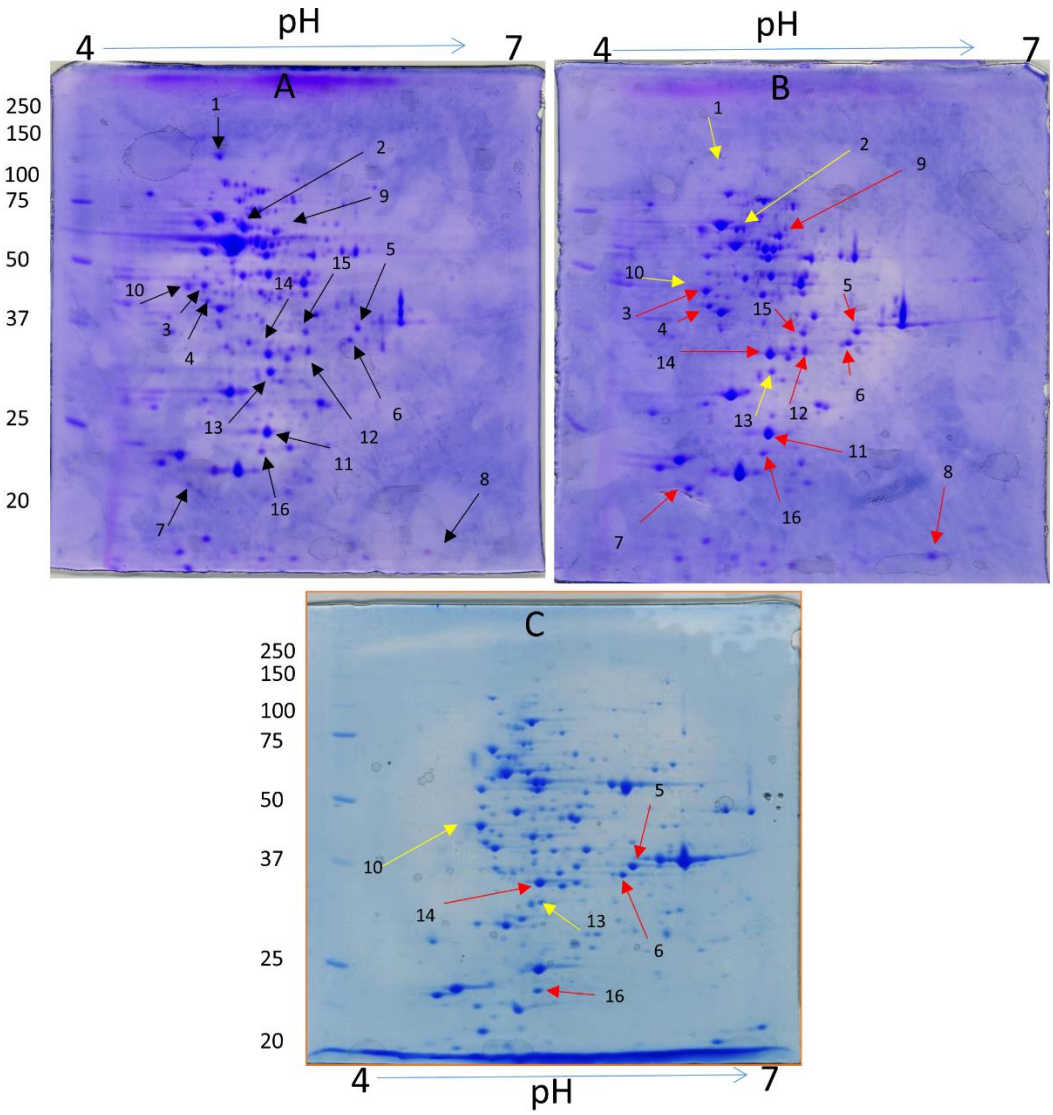

Fig A1. Two-dimensional polyacrylamide gel electrophoresis (PAGE) analysis of *Frankia* sp. strain Allo2 under control (no stress) conditions (A), 200 mM NaCl (B), and 200 mM sucrose (C). Red arrows indicate that proteins are upregulated relative to the control, while yellow arrows indicate down regulated proteins relative to the control. The corresponding number spots were in-gel digested with trypsin and analyzed by liquid chromatography-mass spectrometry (LC-MS) and LC-MS/MS for protein identification

Table A9. *Frankia* sp. strain Allo2 proteins differentially expressed under stress conditions. The identified proteins were classified by COG functional categories. Upregulated proteins are shown by the upward pointing arrow (↑) whereas downregulated proteins are shown by the downward pointing arrow (↓). No change (N/C) indicates that a spot was not picked for that particular condition because it showed similar intensity as the control.

SPOT #	Acc. No	Locus Tag	Protein Name	MW (Da)	PI	NaCl	Sucrose
[C] Energy production and Conversion							
1	WP_035732921.1	ALLO2_RS15565	aconitate hydratase	98737.30	4.80	↓	N/C
4	WP_035729778.1	ALLO2_RS03755	malate dehydrogenase	41530.50	4.80	↑	N/C
14	WP_011437820.1	ALLO2_RS10070	malate dehydrogenase (NAD)	34399.10	4.96	↑	↑
13	WP_035734595.1	ALLO2_RS22325	electron transfer flavoprotein alpha subunit apoprotein	32842.70	5.01	↓	↓
[E] Amino acid transport and metabolism							
13	WP_035732776.1	ALLO2_RS14775	cysteine synthase (CysK)	32443.60	4.95	↑	↑
[G] Carbohydrate transport and metabolism							
3	WP_035729979.1	ALLO2_RS04610	enolase	44751.80	4.60	↑	N/C
5	WP_023841695.1	ALLO2_RS05865	glyceraldehyde-3-phosphate dehydrogenase	35515.80	5.70	↑	↑
6	WP_035729149.1	ALLO2_RS01575	fructose-bisphosphate aldolase	36894.20	5.35	↑	↑
[H] Coenzyme metabolism							
12	WP_035732888.1	ALLO2_RS15275	pyridoxal phosphate synthase yaaD subunit	32577.80	5.34	↑	N/C
[k] Transcription							
10	WP_011435063.1	ALLO2_RS03340	DNA-directed RNA polymerase subunit alpha	37859.00	4.60	↓	↓
[I] Lipid transport and metabolism							
11	WP_035732830.1	ALLO2_RS14885	short chain enoyl-CoA hydratase	26841.10	4.98	↑	N/C
[**M] Cell wall/membrane/envelop biogenesis**							
15	WP_023840985.1		UDP-glucose pyrophosphorylase	34856.10	5.04	↑	N/C
[O] Post-translational modification, protein turnover, chaperone functions							
2	WP_011435086.1	ALLO2_RS03455	molecular chaperone GroEL	57446.70	4.90	↓	N/C
7	WP_011435509.1	ALLO2_RS18350	glutathione peroxidase	19567.10	4.70	↑	N/C
8	WP_011438798.1	ALLO2_RS01170	peptidyl-prolyl cis-trans isomerase	19104.60	6.10	↑	N/C
16	WP_035730024.1	ALLO2_RS05025	ATP-dependent Clp protease proteolytic subunit ClpP	23039.30	4.79	↑	↑
Not assigned to COG categories							
9	WP_035733723.1	ALLO2_RS18330	carnitine O-acetyltransferase	67431.10	5.20	↑	N/C

APPENDIX C

Developing a genetic system for *Frankia*

Introduction

The genetics of *Frankia* is at a nascent stage and routes for introducing defined changes into the *Frankia* chromosome are missing. Developing efficient tools for the manipulation of the *Frankia* genome will allow for the analysis of genes involved in secondary metabolite biosynthesis, symbiosis, and other fascinating aspects of the life of *Frankia*. For this project, developing site-specific gene disruption methods will help us to confirm the roles of some of the candidate genes identified from our transcriptomic, proteomic, and comparative genomic analysis. With this goal, we attempted to develop several site-specific gene disruption mechanisms to introduce defined changes to the *Frankia* genome. Here, we describe our attempt to develop a λ Red–mediated gene disruption (based on the recombination of the bacteriophage lambda) system for *Frankia*. Use of the λ Red–mediated system will help to bypass two major hurdles: (1) degradation by the r*ecBCD* exonuclease when using linear DNA for recombination in bacteria, and (2) the need for a long stretch of DNA for homologous recombination to occur when the λ Red system is not employed.

Method

Because of the challenges associated with making λ Red –proficient *Frankia*, a two-step strategy was designed for *Frankia* as previously described [197]. As a proof of concept, a gene coding for the UPRTase enzyme (Francci3_3203) was targeted. Knockout of the gene allows *Frankia casuarinae* strain CcI3 to grow in medium containing 5- fluorouracil (5-FU). The first step involved Red-mediated recombination in a λ Red –proficient *E. coli* containing the genomic region of interest on a plasmid. The gene of interest was

targeted by a selectable marker that has been generated by PCR with primers containing 39 nt homology arms [143]. The second step involved the genetic exchange in the *Frankia* itself and was achieved though homologous recombination. Briefly, the origin of transfer (oriT) was amplified from pTNR-oriT [198] using primers so that the amplified fragment contained ApaI and HindIII sites. The amplified product was digested with ApaI + NcoI and ligated into pGEM-5Zf(+) (Promega) digested with the same enzymes. A 4.7 kb region containing the UPRTase coding gene (Francci3_3203) and 2 kb flanking regions on either side was amplified from the genomic DNA of CcI3 with primers so that the amplified product contained NcoI and NdeI sites. The construct was electroporated into a λ Red –proficient *E. coli* BW25113 containing the pIJ790 vector. Positive transformants were selected on LB agar plates containing ampicillin (50 μg/ml) and chloramphenicol (20 μg/ml). The tetracycline resistance gene (*tetA*) from the pHTK1 plasmid was amplified with primers so that the amplified product contained 60 bp arms homologous to the regions flanking the coding sequence of the target gene (Francci3_3203). The linear PCR product was electroporated into the λ Red –proficient *E. coli* BW25113 containing the construct. The λ Red recombinase system was activated by 1 M L-arabinose. The Amp^R and Tet^R mutagenized constructs were introduced into *E. coli* ET12567 containing the pUZ8002 plasmid by electroporation and then transferred to *Frankia* through filter mating. Tet^R and 5-FU resistant (50 μg/ml) *Frankia* conjugants were screened for ampicillin sensitivity, indicating a double-crossover homologous recombination.

Table A10. Primers used in the λ Red recombination system

Forward	Reverse
Primers used to amplify oriT from pHTK1	
5′ ATA TTC GGG CCC TCG CGG ACG TGC TCA TAG TC 3′	5′ ATT CAT CTC CAT GGC TCG CCT GTC CCC TCA GTT CA 3′
Primers used to amplify Francci3_3203 along with flanking region	
5′ ATA TTA ACC ATG GGC GGC CCA CTC CGT CTC CTT GT 3′	5′ ATC ATT ATC ATA TGT CGT CGG CTT CCA GCA CGT CAA 3'
Primers used to amplify *tetA* gene with 39 BP homologous arms	
5′ TAA GGA CCG TCC CGT GAG GCG GGG AAG GAG GGA TGC CGC ATT GCG CGC TTG GCG TAA TCA 3′	5′ CGA ACC CGG CTT GTG CTC CCC ACA CGG CAC TTC CCC CGA CGG CAG CGC GAC AAC AAT T 3′

Results and future directions

A double-crossover homologous recombination frequency of 1.4 x 10^{-6} was obtained (Table A11). Mutants need to be further confirmed through sequencing and PCR.

Table A11. Frequency of double-crossover homologous recombination in *Frankia casuarinae* strain CcI3.

	5-FU^R, Tet^R, Amp^s	Tet^R, 5-FU^R	5-FU^R	Tet^R
Efficiency	1.4 x 10^{-6} ± 4.0 x 10^{-7}	10^{-6} ± 8.0 x 10^{-7}	10^{-5} ± 3.0 x 10^{-6}	2.8 x 10^{-6} ± 2 x 10^{-6}
Remark	double crossover homologous recombinants			

APPENDIX D

List of publications from the dissertation work

Peer-reviewed journal articles

1. Ngom M, Gray K, Diagne N, **Oshone R**, Fardoux J, Gherbi H, Hocher V, SVISTOONOFF S, Laplaze L, Tisa L, Sy MO and Champion A. 2016. Symbiotic performance of diverse *Frankia* strains on salt-stressed *Casuarina glauca* and *Casuarina equisetifolia* plants. Front. Plant Sci. 7:1331. doi: 10.3389/fpls.2016.01331
2. **Oshone R**, Ngom M, Abebe-Akele F, Simpson S, Morris K, Sy MO, Champion A, Thomas WK, Tisa LS. 2016. Permanent draft genome sequence of *Frankia* sp. strain Allo2, a salt-tolerant nitrogen-fixing actinobacterium isolated from the root nodules of *Allocasuarina*. Genome Announc 4(3):e00388-16 doi:10.1128/genomeA.00388-16
3. Ngom M, **Oshone R**, Diagne N, Cissoko M, Svistoonoff S, Tisa LS, Laplaze L, Sy MO, Champion A. 2016. Tolerance to environmental stress by the nitrogen-fixing actinobacterium *Frankia* and its role in actinorhizal plants adaptation. Symbiosis.
4. Bose D, Sarkar I, Labar R, **Oshone R**, Ghazal S, Morris K, Abebe-Akele F, Thomas WK, Tisa LS, Sen A. 2016. Comparative genomics of *Prauserella* sp. Am3, an actinobacterium isolated from root nodules of *Alnus nepalensis* in India. Symbiosis.
5. Ghazal S, **Oshone R**, Simpson S, Morris K, Abebe-Akele F, Thomas WK, Khalil KM, Tisa LS. 2016. Draft genome sequence of *Photorhabdus luminescens* subsp. *laumondii* HP88, an entomopathogenic bacterium isolated from nematodes. Genome Announc 4(2):e00154-16. doi:10.1128/genomeA.00154-16
6. Tisa LS, **Oshone R**, Sarkar I, Ktari A, Sen A, Gtari M. 2016. Genomic approaches toward understanding the actinorhizal symbiosis: an update on the status of the *Frankia* genomes. Symbiosis.
7. **Oshone R**, Hurst SG, IV, Abebe-Akele F, Simpson S, Morris K, Thomas WK, Tisa LS. 2016. Permanent draft genome sequences for two variants of *Frankia* sp. strain CpI1, the first *Frankia* strain isolated from root nodules of *Comptonia peregrina*. Genome Announc 4(1):e01588-15. doi:10.1128/genomeA.01588-15.
8. Swanson E, **Oshone R**, Simpson S, Morris K, Abebe-Akele F, Thomas WK, Tisa LS. 2015. Permanent draft genome sequence of *Frankia* sp. strain AvcI1, a nitrogen-fixing actinobacterium isolated from the root nodules of *Alnus viridis* subsp. *crispa* grown in Canada. Genome Announc 3(6):e01511-15. doi:10.1128/genomeA.01511-15
9. Swanson E, **Oshone R**, Simpson S, Morris K, Abebe-Akele F, Thomas WK, Tisa LS. 2015. Permanent draft genome sequence of *Frankia* sp. strain ACN1ag, a nitrogenfixingactinobacterium isolated from the root nodules of *Alnus glutinosa*. Genome Announc 3(6):e01483-15. doi:10.1128/genomeA.01483-15.
10. Tisa LS, Beauchemin N, Cantor MN, Furnholm T, Ghodhbane-Gtari F, Goodwin L, Copeland A, Gtari M, Huntemann M, Ivanova N, Kyrpides N, Markowitz V, Mavrommatis K, Mikhailova N, Nouioui I, **Oshone R**, Ovchinnikova G, Pagani I, Palaniappan K, Pati A, Sen A, Shapiro N, Szeto E, Wall L, Wishart J, Woyke T. 2015. Draft genome sequence of *Frankia* sp. strain DC12, an atypical,

noninfective, ineffective isolate from *Datisca cannabina*. Genome Announc 3(4):e00889-15. doi:10.1128/genomeA.00889-15.

11. Ghodhbane-Gtari F, Hurst SG, IV, **Oshone R**, Morris K, Abebe-Akele F, Thomas WK, Ktari A, Salem K, Gtari M, Tisa LS. 2014. Draft genome sequence of *Frankia* sp. Strain BMG5.23, a salt-tolerant nitrogen-fixing actinobacterium isolated from the root nodules of *Casuarina glauca* grown in Tunisia. Genome Announc. 2(3):e00520-14. doi:10.1128/genomeA.00520-14
12. Hurst SG IV, **Oshone R**, Ghodhbane-Gtari F, Morris K, Abebe-Akele F, Thomas WK, Ktari A, Salem K, Mansour S, Gtari M, Tisa LS. 2014. Draft genome sequence of *Frankia* sp. strain Thr, a nitrogen-fixing actinobacterium isolated from the root nodules of *Casuarina cunninghamiana* grown in Egypt. Genome Announc. 2(3):e00493-14. doi:10.1128/genomeA.00493-14
13. Mansour SR, **Oshone R**, Hurst SG, IV, Morris K, Thomas WK, Tisa LS. 2014. Draft genome sequence of *Frankia* sp. strain CcI6, a salt-tolerant nitrogen-fixing actinobacterium isolated from the root nodule of *Casuarina cunninghamiana*. Genome Announc. 2(1):e01205-13. doi:10.1128/genomeA.01205-13.

14. **Oshone R**, Mansour SR, Tisa LS. 2013. Effect of salt stress on the physiology of *Frankia* sp. strain CcI6. J. Biosci. 38:699–702.

Articles submitted for publication

1. Oshone R, Ngom M, Chu F, Mansour SR, Sy MO, Champion A, Tisa L. 2017. A comprehensive genomic, transcriptomic, and proteomic approach towards understanding the molecular mechanisms of salt tolerance in *Frankia* strains isolated from *Casuarina* trees. BMC Genomics.

References

1. Shrivastava P, Kumar R. Soil salinity. A serious environmental issue and plant growth promoting bacteria as one of the tools for its alleviation. Saudi Journal of Biological Sciences. 2015; 22:123–131.

2. Qadir M, Quillérou E, Nangia V, Murtaza G, Singh M, Thomas R, et al. Economics of salt-induced land degradation and restoration. Natural Resources Forum. 2014;38:282–95.

3. Jamil A, Riaz S, Ashraf M, Foolad MR. Gene Expression Profiling of Plants under Salt Stress. Critical Reviews in Plant Sciences. 2011;30:435–458.

4. Munns R. Genes and salt tolerance: bringing them together. New Phytol. 2005;167:645–663.

5. Gupta B, Huang B. Mechanism of Salinity Tolerance in Plants: Physiological, Biochemical, and Molecular Characterization. International Journal of Genomics. 2014;2014:1–18.

6. Carillo P, Grazia M, Pontecorvo G, Fuggi A, Woodrow P. Salinity Stress and Salt Tolerance. In: Shanker A, Venkateswarlu B, editors. Abiotic stress in plants: mechanisms and adaptations. Rijeka: InTech; 2011. p. 21-39

7. Munns R, Tester M. Mechanisms of Salinity Tolerance. Annual Review of Plant Biology. 2008;59:651–81.

8. Rahnama A, James RA, Poustini K, Munns R. Stomatal conductance as a screen for osmotic stress tolerance in durum wheat growing in saline soil. Functional Plant Biology. 2010;37:255.

9. Subramanian S, Souleimanov A, Smith DL. Proteomic Studies on the Effects of Lipo-Chitooligosaccharide and Thuricin 17 under Unstressed and Salt Stressed Conditions in *Arabidopsis thaliana.* Frontiers in Plant Science. 2016;7:1314

10. Mahajan S, Tuteja N. Cold, salinity and drought stresses: An overview. Archives of Biochemistry and Biophysics. 2005;444:139–58.

11. Prasad MNV, Ahmad P. Abiotic stress responses in plants: metabolism, productivity and sustainability. New York: Springer; 2012.

12. James RA, Blake C, Byrt CS, Munns R. Major genes for Na exclusion, Nax1 and Nax2 (wheat HKT1;4 and HKT1;5), decrease Na accumulation in bread wheat

leaves under saline and waterlogged conditions. Journal of Experimental Botany. 2011;62:2939–47.

13. Haffez FY, Hameed S, Malik KA. *Frankia* and *Rhizobium* strains as inoculum for fast growing trees in saline environment. Pak. J. Bot. 1999; 31:173-182.

14. Wall LG. The actinorhizal symbiosis. J Plant Growth Regul. 2000;19:167–182.

15. Benson DR, Silvester WB. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiological Reviews. 1993;57:293-319.

16. Benson DR, Vanden Heuvel BD, Potter D. Actinorhizal symbioses: Diversity and biogeography. In: Gillings M, editors. Plant microbiology. Oxford: BIOS Scientific Publishers; 2004.

17. Santi C, Bogusz D, Franche C. Nitrogen fixation in non-legumes. Ann Bot. 2013 111:743–767

18. Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, et al. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proceedings of the National Academy of Sciences. 1995;92:2647–51.

19. Budowski G, Russo R. Nitrogen-fixing trees and nitrogen fixation in sustainable agriculture: Research challenges. Soil Biology and Biochemistry. 1997;29:767–70.

20. Izquierdo I, Caravaca F, Alguacil M, Hernández G, Roldán A. Use of microbiological indicators for evaluating success in soil restoration after revegetation of a mining area under subtropical conditions. Applied Soil Ecology. 2005;30:3–10.

21. Abid H, Ali J, Hussain A, Afridi SR. Production and quality evaluation of sea buckthorn (Hippophae rhamnoides L.) vinegar using *Accetobacter acceti.* Pakistan Journal of Biochemestry and Molecular Biology. 2010;43:185–188

22. Becking JH. N2-fixing tropical non-legumes. In: Dommergues YR, Diem HG, editors. Microbiology of Tropical Soil and Plant Productivity. The Hague: Martinus Nijhoff; 1982. p. 109-146.

23. Kalia RK, Singh R, Rai MK, Mishra GP, Singh SR, Dhawan AK. Biotechnological interventions in sea buckthorn (*Hippophae L*.): current status and future prospects. Trees. 2011;25:559–75.

24. Moiroud ACA. Diversité et écologie des plantes actinorhiziennes. Acta Botanica Gallica. 1996;143:651–61.

25. Dommergues Y. Contribution of actinorhizal plants to tropical soil productivity and rehabilitation. Soil Biology and Biochemistry. 1997;29:931–41.

26. Perrine-Walker F, Doumas P, Lucas M, Vaissayre V, Beauchemin NJ, Band LR, et al. Auxin Carriers Localization Drives Auxin Accumulation in Plant Cells Infected by *Frankia* in *Casuarina glauca* Actinorhizal Nodules. Plant Physiology. 2010;154:1372–80.

27. Diagne N, Arumugam K, Ngom M, Nambiar-Veetil M, Franche C, Narayanan KK, et al. Use of *Frankia* and Actinorhizal Plants for Degraded Lands Reclamation. BioMed Research International. 2013;2013:1–9.

28. Simonet P, Navarro E, Rouvier C, Reddell P, Zimpfer J, Dommergues Y, et al. Co-evolution between *Frankia* populations and host plants in the family Casuarinaceae and consequent patterns of global dispersal. Environmental Microbiology. 1999;1:525–33.

29. Parsons R, Silvester WB, Harris S, Gruijters WTM, Bullivant S. *Frankia* Vesicles Provide Inducible and Absolute Oxygen Protection for Nitrogenase. Plant Physiology. 1987;83:728–31.

30. Ott T, Dongen JTV, Gu¨nther C, Krusell L, Desbrosses G, Vigeolas H, et al. Symbiotic Leghemoglobins Are Crucial for Nitrogen Fixation in Legume Root Nodules but Not for General Plant Growth and Development. Current Biology. 2005;15:531–5.

31. Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson J, et al. Molecular Phylogeny of the Genus *Frankia* and Related Genera and Emendation of the Family Frankiaceae. International Journal of Systematic Bacteriology. 1996;46:1–9.

32. Tisa LS, Oshone R, Sarkar I, Ktari A, Sen A, Gtari M. Genomic approaches toward understanding the actinorhizal symbiosis: an update on the status of the *Frankia* genomes. Symbiosis. 2016;70:5–16.

33. Normand P, Lapierre P, Tisa LS, Gogarten JP, Alloisio N, Bagnarol E, et al. Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. Genome Research. 2007;17:7–15.

34. Normand P, Queiroux C, Tisa LS, Benson DR, Rouy Z, Cruveiller S, et al. Exploring the genomes of *Frankia*. Physiologia Plantarum. 2007;130:331–43.

35. Udwary DW, Gontang EA, Jones AC, Jones CS, Schultz AW, Winter JM, et al. Significant Natural Product Biosynthetic Potential of Actinorhizal Symbionts of the Genus *Frankia*, as Revealed by Comparative Genomic and Proteomic Analyses. Applied and Environmental Microbiology. 2011;77:3617–25.

36. Ghodhbane-Gtari F, Hurst SG, Oshone R, Morris K, Abebe-Akele F, Thomas WK, et al. Draft Genome Sequence of *Frankia* sp. Strain BMG5.23, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of Casuarina glauca Grown in Tunisia. Genome Announcements. 2014;2.

37. Hurst SG, Oshone R, Ghodhbane-Gtari F, Morris K, Abebe-Akele F, Thomas WK, et al. Draft Genome Sequence of *Frankia* sp. Strain Thr, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of Casuarina cunninghamiana Grown in Egypt. Genome Announcements. 2014;2.

38. Mansour SR, Oshone R, Hurst SG, Morris K, Thomas WK, Tisa LS. Draft Genome Sequence of *Frankia* sp. Strain CcI6, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodule of Casuarina cunninghamiana. Genome Announcements. 2014;2.

39. Ngom M, Oshone R, Hurst SG, IV, Abebe-Akele F, Simpson S, Morris K, Sy MO, Champion A, Thomas WK, Tisa LS. Permanent draft genome sequence for *Frankia* sp. strain CeD, a nitrogen-fixing actinobacterium isolated from the root nodules of Casuarina equistifolia grown in Senegal. Genome Announcements. 2016;4.

40. D'Angelo T, Oshone R, Abebe-Akele F, Simpson S, Morris K, Thomas WK, et al. Permanent Draft Genome Sequence of *Frankia* sp. Strain BR, a Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Casuarina equisetifolia*. Genome Announcements. 2016;4.

41. Oshone R, Ngom M, Abebe-Akele F, Simpson S, Morris K, Sy MO, et al. Permanent Draft Genome Sequence of *Frankia* sp. Strain Allo2, a Salt-Tolerant Nitrogen-Fixing Actinobacterium Isolated from the Root Nodules of *Allocasuarina*. Genome Announcements. 2016;4.

42. Mardis E. What is Finished, and Why Does it Matter. Genome Research. 2002;12:669–71

43. Tisa LS, Beauchemin N, Gtari M, Sen A, Wall LG. What stories can the Frankia genomes start to tell us? Journal of Biosciences. 2013;38:719–26.

44. Maclean AM, Finan TM, Sadowsky MJ. Genomes of the Symbiotic Nitrogen-Fixing Bacteria of Legumes. Plant Physiology. 2007;144:615–22.

45. Zuanazzi JCAAS, Clergeot PH, Quirion J-C, Husson H-P, Kondorosi A, Ratet P. Production of *Sinorhizobium meliloti* nod Gene Activator and Repressor Flavonoids from *Medicago sativa* Roots. Molecular Plant-Microbe Interactions. 1998;11:784–94.

46. Oldroyd GED. Speak, friend, and enter: signaling systems that promote beneficial symbiotic associations in plants. Nature Reviews Microbiology. 2013;11:252–63.

47. Persson T, Benson DR, Normand P, Heuvel BV, Pujic P, Chertkov O, et al. Genome Sequence of "*Candidatus Frankia datiscae*" Dg1, the Uncultured Microsymbiont from Nitrogen-Fixing Root Nodules of the Dicot *Datisca glomerata.* Journal of Bacteriology. 2011;193:7017–8.

48. Nguyen TV, Wibberg D, Battenberg K, Blom J, Heuvel BV, Berry AM, et al. An assemblage of *Frankia* Cluster II strains from California contains the canonical nod genes and also the sulfotransferase gene *nodH*. BMC Genomics. 2016;17.

49. Mastronunzio JE, Tisa LS, Normand P, Benson DR. Comparative secretome analysis suggests low plant cell wall degrading capacity in *Frankia* symbionts. BMC Genomics. 2008;9:47.

50. Sen A, Sur S, Bothra AK, Benson DR, Normand P, Tisa LS. The implication of life style on codon usage patterns and predicted highly expressed genes for three Frankia genomes. Antonie van Leeuwenhoek. 2008;93:335–46.

51. Udwary DW, Gontang EA, Jones AC, Jones CS, Schultz AW, Winter JM, et al. Significant Natural Product Biosynthetic Potential of Actinorhizal Symbionts of the Genus *Frankia*, as Revealed by Comparative Genomic and Proteomic Analyses. Applied and Environmental Microbiology. 2011;77:3617–25.

52. Alloisio N, Queiroux C, Fournier P, Pujic P, Normand P, Vallenet D, et al. The *Frankia alni* Symbiotic Transcriptome. Molecular Plant-Microbe Interactions. 2010;23:593–607.

53. Alloisio N, Félix S, Maréchal J, Pujic P, Rouy Z, Vallenet D, et al. Frankia alni proteome under nitrogen-fixing and nitrogen-replete conditions. Physiologia Plantarum. 2007;130:440–53.

54. Popovici J, Comte G, Bagnarol E, Alloisio N, Fournier P, Bellvert F, et al. Differential Effects of Rare Specific Flavonoids on Compatible and Incompatible Strains in the *Myrica gale-Frankia* Actinorhizal Symbiosis. Applied and Environmental Microbiology. 2010;76:2451–60.

55. Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M. Phylogeny of members of the Frankia genus based on gyrB, nifH and glnII sequences. Antonie van Leeuwenhoek. 2011;100:579–87.

56. Bagnarol E, Popovici J, Alloisio N, Maréchal J, Pujic P, Normand P, et al. Differential *Frankia* protein patterns induced by phenolic extracts from Myricaceae seeds. Physiologia Plantarum. 2007;130:380–90.

57. Echbab H, Arahou M, Ducousso M, Nourissier-Mountou S, Duponnois R, Lahlou H, et al. Successful nodulation of *Casuarina* by *Frankia* in axenic conditions. Journal of Applied Microbiology. 2007;103:1728–37.

58. Schwencke J, Carú M. Advances in Actinorhizal Symbiosis: Host Plant- *Frankia* Interactions, Biology, and Applications in Arid Land Reclamation. A Review. Arid Land Research and Management. 2001;15:285–327.

59. Dawson JO. Ecology of actinorhizal plants. In: Pawlowski K, Newton WE. Nitrogen-fixing Actinorhizal Symbioses. Dordrecht: Springer; 2008. p. 199-234.

60. Salt DE, Blaylock M, Kumar NP, Dushenkov V, Ensley BD, Chet I, et al. Phytoremediation: A Novel Strategy for the Removal of Toxic Metals from the Environment Using Plants. Bio/Technology. 1995;13:468–74.

61. Richards JW, Krumholz GD, Chval MS, Tisa LS. Heavy Metal Resistance Patterns of *Frankia* Strains. Applied and Environmental Microbiology. 2002;68:923–7.

62. Srivastava A, Singh SS, Mishra AK. Sodium transport and mechanism(s) of sodium tolerance in *Frankia* strains. Journal of Basic Microbiology. 2012;53:163–74.

63. El-Lakany MH, Luard EJ. Comparative salt tolerance of selected *Casuarina* species. Aust. For. Res. 1982; 13: 11-20.

64. Diem H, Dommergues Y. Current and Potential Uses and Management of Casuarinaceae in the Tropics and Subtropics. The Biology of *Frankia* and Actinorhizal Plants. 1990;317–42.

65. Ngom M, Gray K, Diagne N, Oshone R, Fardoux J, Gherbi H, et al. Symbiotic Performance of Diverse *Frankia* Strains on Salt-Stressed *Casuarina glauca* and *Casuarina equisetifolia* Plants. Frontiers in Plant Science. 2016;7.

66. Ngom M, Oshone R, Diagne N, Cissoko M, Svistoonoff S, Tisa LS, et al. Tolerance to environmental stress by the nitrogen-fixing actinobacterium *Frankia* and its role in actinorhizal plants adaptation. Symbiosis. 2016;70:17–29.

67. Zhong C, Zhang Y. Introduction and management of *Casuarina* tree species in China. China Forestry Science and Technology. 2003;17:3–5.

68. Sayed WF. Improving *Casuarina* growth and symbiosis with *Frankia* under different soil and environmental conditions—review. Folia Microbiologica. 2011;56:1–9.

69. Monneveux P, Nemmar M. Contribution à l'étude de la résistance à la sécheresse chez le blé tendre (*Triticum aestivum L*.) et chez le blé dur (*Triticum durum Desf*.): étude de l'accumulation de la proline au cours du cycle de développement. Agronomie. 1986; 6:583–590.

70. Tani C, Sasakawa H. Salt tolerance of *Casuarina equisetifolia* and *Frankia* Ceq1 strain isolated from the root nodules of *C. equisetifolia.* Soil Science and Plant Nutrition. 2003;49:215–22.

71. Andéké-Lingui MA, Dommergues. Coastal sand dune stabilization in Senegal. In: Midgley SJ, Turnbull JW, and Johnston RD, editors. *Casuarina* ecology, management and utilization: proceedings of an international workshop; Canberra, 17th -21st August, 1981. Melbourne: CSIRO; 1983. p. 158-166.

72. Shah SK, Shah RP, Xu HL, Aryal UK. Biofertilizers: An Alternative Source of Nutrients for Sustainable Production of Tree Crops. Journal of Sustainable Agriculture. 2007;29:85–95.

73. Girgis MGZ, Ishac YZ, Diem HG, Dommergues YR. Selection of salt tolerant *Casuarina glauca* and *Frankia*. Acta Oecologica. 1992; 13:443–451.

74. Mailly D, Ndiaye P, Margolis HA, Pineau M. Fixation des dunes et reboisement avec le filao (*Casuarina equisetifolia*) dans la zone du littoral nord du Sénégal. The Forestry Chronicle. 1994; 70:282–290.

75. Doran JC, Hall N: Notes on fifteen Australian *Casuarina* species. In: Midgeiey SJ, Tumbuil JW, Johnson RD, editors. *Casuarin*a ecology, management and utilization: proceedings of an international workshop, Canberra, Australia. 17-21 August 1981. Melbourne: CSIRO; 1983. p.19-52.

76. Tripathi SB, Mathish NV, Gurumurthi K. Changes in protein profiles in Casuarina equisetifolia during salt stress. In: Gurumurthi K. Proceedings of the 5th Annual Workshop on Casuarina. Rajahmundry: Regional Forest Research Center; 2001. p. 126–134

77. Dawson JO, Gibson AH. Sensitivity of selected Frankia isolates from Casuarina, Allocasuarina and North American host plants to sodium chloride. Physiologia Plantarum. 1987;70:272–8.

78. Oshone R, Mansour SR, Tisa LS. Effect of salt stress on the physiology of *Frankia* sp strain CcI6. Journal of Biosciences. 2013;38:699–702.

79. Brown A. Microbial water stress. Bacteriol. Rev. 1976;40:803–846

80. Waditee R, Hibino T, Nakamura T, Incharoensakdi A, Takabe T. Overexpression of a Na /H antiporter confers salt tolerance on a freshwater cyanobacterium, making it capable of growth in sea water. Proceedings of the National Academy of Sciences. 2002;99:4109–14.

81. Shamseldin A, Nyalwidhe J, Werner D. A Proteomic Approach Towards the Analysis of Salt Tolerance in *Rhizobium etli* and *Sinorhizobium meliloti* Strains. Current Microbiology. 2006;52:333–9.

82. Romantsov T, Guan Z, Wood JM. Cardiolipin and the osmotic stress responses of bacteria. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2009;1788:2092–100.

83. Brígido C, Alexandre A, Oliveira S. Transcriptional analysis of major chaperone genes in salt-tolerant and salt-sensitive mesorhizobia. Microbiological Research. 2012;167:623–9.

84. Rojas E, Theriot JA, Huang KC. Response of *Escherichia coli* growth rate to osmotic shock. Proceedings of the National Academy of Sciences. 2014;111:7807–12.

85. Kempf B, Bremer E. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. Archives of Microbiology. 1998;170:319–30.

86. Mclaggan D, Naprstek J, Buurman ET, Epstein W. Interdependence of K+ and glutamate accumulation during osmotic adaptation of *Escherichia coli*. J. Biol. Chem. 1994;269:1911–1917

87. Ventosa A, Nieto JJ, Oren A. Biology of moderately halophilic aerobic bacteria. Microbiol. Mol. Biol. Rev. 1998;62: 504-44.

88. Smith LT, Smith GM. An osmoregulated dipeptide in stressed *Rhizobium meliloti*. Journal of Bacteriology. 1989;171:4714–7.

89. TeChien C, Maundu J, Cavaness J, Dandurand LM, Orser CS. Characterization of salt-tolerant and salt-sensitive mutants of *Rhizobium leguminosarum* biovar viciae strain C1204b. FEMS Microbiology Letters. 1992;90:135–40.

90. Tsuzuki M, Moskvin OV, Kuribayashi M, Sato K, Retamal S, Abo M, et al. Salt Stress-Induced Changes in the Transcriptome, Compatible Solutes, and Membrane Lipids in the Facultatively Phototrophic Bacterium *Rhodobacter sphaeroides.* Applied and Environmental Microbiology. 2011;77:7551–9

91. Peddie B. Relationship between osmoprotection and the structure and intracellular accumulation of betaines by *Escherichia coli*. FEMS Microbiology Letters. 1994;120:125–31.

92. Ko R, Smith LT, Smith GM. Glycine betaine confers enhanced osmotolerance and cryotolerance on *Listeria monocytogenes*. Journal of Bacteriology. 1994;176:426–31.

93. Sleator R. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. FEMS Microbiology Reviews. 2001;25.

94. Miller G, Suzuki N, Ciftci-Yilmaz S, Mittler R. Reactive oxygen species homeostasis and signaling during drought and salinity stresses. Plant, Cell & Environment. 2010;33:453–67.

95. Apel K, Hirt H. REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. Annual Review of Plant Biology. 2004;55:373–99.

96. Singh S, Brocker C, Koppaka V, Chen Y, Jackson BC, Matsumoto A, et al. Aldehyde dehydrogenases in cellular responses to oxidative/electrophilic stress. Free Radical Biology and Medicine. 2013;56:89–101.

97. Susin MF, Baldini RL, Gueiros-Filho F, Gomes SL. GroES/GroEL and DnaK/DnaJ Have Distinct Roles in Stress Responses and during Cell Cycle Progression in Caulobacter crescentus. Journal of Bacteriology. 2006;188:8044–53.

98. Bajji M, Kinet JM, Lutts S. The use of the electrolyte leakage method for assessing cell membrane stability as a water stress tolerance test in durum wheat. Plant Growth Regul. 2001;36:61–70

99. Turk M, Méjanelle L, Šentjurc M, Grimalt JO, Gunde-Cimerman N, Plemenitaš A. Salt-induced changes in lipid composition and membrane fluidity of halophilic yeast-like melanized fungi. Extremophiles. 2003;8:53–61

100. Tasaka Y, Gombos Z, Nishiyama Y, Mohanty P, Ohba T, Ohki K, Murata N. Targeted mutagenesis of acyl-lipid desaturases in Synechocystis: evidence for the important roles of polyunsaturated membrane lipids in growth, respiration and photosynthesis. EMBO J. 1996;15:6416–6425

101. Srivastava A, Singh SS, Mishra AK. Modulation in fatty acid composition influences salinity stress tolerance in *Frankia* strains. Annals of Microbiology. 2013;64:1315–23.

102. Tisa L, Mcbride M, Ensign JC. Studies of growth and morphology of *Frankia* strains EAN1 pec , EuI1 c , CpI1, and ACN1 AG. Canadian Journal of Botany. 1983;61:2768–73.

103. Tisa LS, Chval MS, Krumholz GD, Richards J. Antibiotic resistance patterns of *Frankia* strains. Canadian Journal of Botany. 1999;77:1257–60.

104. Peret B, Swarup R, Jansen L, Devos G, Auguy F, Collin M, et al. Auxin Influx Activity Is Associated with *Frankia* Infection during Actinorhizal Nodule Formation in *Casuarina glauca*. Plant Physiology. 2007;144:1852–62.

105. Proposal of a type strain for Frankia alni (Woronin 1866) Von Tubeuf 1895, emended description of *Frankia alni*, and recognition of *Frankia casuarinae* sp. nov. and *Frankia elaeagni* sp. nov. International Journal of Systematic and Evolutionary Microbiology. 2016;66:5201–10.

106. Zhang Z, Lopez MF, Torrey JG. A comparison of cultural characteristics and infectivity of *Frankia* isolates from root nodules of *Casuarina* species. Frankia Symbioses. 1984;78:79–90.

107. Mansour SR, Moussa LAA. Role of Gamma-radiation on spore germination and infectivity of *Frankia* strains CeI523 and CcI6 isolated from Egyptian Casuarina. Isotope Rad. Res. 2005;37:1023–1038

108. Girgis MGZ, Schwencke J. Differentiation of *Frankia* strains by their electrophoretic patterns of intracellular esterases and aminopeptidases. Journal of General Microbiology. 1993;139:2225–32.

109. Diem HG, Dommergues Y. The isolation of Frankia from nodules of *Casuarina*. Can J Bot. 1983;61:2822–2825.

110. Müller A, Benoist P, Diem HG, Schwencke J. Age-dependent changes in extracellular proteins, aminopeptidase and proteinase activities in *Frankia* isolate BR. J Gen Microbiol. 1991;137:2787–2796.

111. Ghodhbane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. 16S-23S rRNA intergenic spacer region variability in the genus *Frankia*. Microb. Ecol. 2010; 60:487–495.

112. Girgis MGZ, Ishac YZ, El-Haddad M, Saleh EA, Diem HG, Dommergues YR. First report on isolation and culture of effective *Casuarina*-compatible strains of

Frankia from Egypt. In: El-Lakany MH, Turnbull JW, Brewbaker JL, editors. Advances in *Casuarina* research and utilization: proceedings of the second International *Casuarina* Workshop, Cairo, Egypt, January 15- 20, 1990. Cairo, Egypt: Desert Development Center, American University in Cairo; 1990. p 156–164.

113. Normand P, Lalonde M. Evaluation of *Frankia* strains isolated from provenances of two *Alnus* species. Canadian Journal of Microbiology. 1982;28:1133–42.

114. Lalonde M, Calvert HE, Pine S, Calvert HE, Pine S. Isolation and use of *Frankia* strains in actinorhizae formation. In: Gibsen AH. Current perspectives in nitrogen fixation: proceedings of the Fourth International Symposium on Nitrogen Fixation held in Canberra, Australia, 1 to 5 Dec. 1980. Amsterdam: Elsevier; 1981. pp. 296–299

115. Hameed S, Hafeez FY, Mirza MS, Malik KA, Akkermans ADL. Confirmation of an isolate from *Datisca cannabina* as atypical *Frankia* strain using PCR amplified 16 rRNA sequence analysis. Pak J Bot. 1994;26:247–251.

116. Hafeez F. Nitrogen fixation and nodulation in Datisca cannabina L. and Alnus nitida Endl. PhD thesis Quaid-e-Azam University, Islamabad, Pakistan.1983.

117. Baker D, Newcomb W, Torrey JG. Characterization of an ineffective actinorhizal microsymbiont, *Frankia* sp. EuI1 (Actinomycetales). Canadian Journal of Microbiology. 1980;26:1072–89.

118. Nouioui I, Ghodhbane-Gtari F, Montero-Calasanz MDC, Rohde M, Tisa LS, Gtari M, et al. *Frankia inefficax* sp. nov., an actinobacterial endophyte inducing ineffective, non nitrogen-fixing, root nodules on its actinorhizal host plants. Antonie van Leeuwenhoek. 2016;110:313–20.

119. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences. 2000;97:6640–5.

120. Bradford M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Analytical Biochemistry. 1976;72:248–54.

121. Broughton WJ, Dilworth MJ. Control of leghaemoglobin synthesis in snake beans. Biochemical Journal. 1971;125:1075-1080.

122. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, et al. Measurement of protein using bicinchoninic acid. Anal. Biochem. 1985;150:76 –85.

123. Furnholm T, Beauchemin N, Tisa LS. Development of a semi-high-throughput growth assay for the filamentous actinobacteria *Frankia*. Archives of Microbiology. 2011;194:13–20.

124. Tisa LS, Ensign JC. Comparative physiology of nitrogenase activity and vesicle development for *Frankia* strains CpI1, ACN1 ag , EAN1pec and EUN1f. Archives of Microbiology. 1987;147:383–8.

125. Murray M, Thompson W. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research. 1980;8:4321–6.

126. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proceedings of the National Academy of Sciences. 2010;108:1513–8.

127. Markowitz VM, Mavromatis K, Ivanova NN, Chen I-MA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics. 2009;25:2271–8.

128. Wang Y, Coleman-Derr D, Chen G, Gu YQ. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Research. 2015;43.

129. Li L. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Research. 2003;13:2178–89

130. Enright AJ. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Research. 2002;30:1575–84.

131. Lerat E, Daubin V, Ochman H, Moran NA. Evolutionary Origins of Genomic Repertoires in Bacteria. PLoS Biology. 2005;3.

132. Goris J, Klappenbach JA, Vandamme P, Coenye T, Konstantinidis KT, Tiedje JM. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. International Journal of Systematic and Evolutionary Microbiology. 2007;57:81–91.

133. Auch AF, Klenk H-P, Göker M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Standards in Genomic Sciences. 2010;2:142–8.

134. Lerat E, Daubin V, Moran NA. From Gene Trees to Organismal Phylogeny in Prokaryotes:The Case of the γ-Proteobacteria. PLoS Biology. 2003;1.

135. Baggerly KA, Deng L, Morris JS, Aldaz CM. Differential expression in SAGE: accounting for normal between-library variation. Bioinformatics. 2003;19:1477–83.

136. Bradford M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Analytical Biochemistry. 1976;72:248–54.

137. Guan S, Burlingame AL. Data Processing Algorithms for Analysis of High Resolution MSMS Spectra of Peptides with Complex Patterns of Posttranslational Modifications. Molecular & Cellular Proteomics. 2009;9:804–10.

138. Chalkley RJ. Comprehensive Analysis of a Multidimensional Liquid Chromatography Mass Spectrometry Dataset Acquired on a Quadrupole Selecting, Quadrupole Collision Cell, Time-of-flight Mass Spectrometer: II. New Developments in Protein Prospector Allow for Reliable and Comprehensive Automatic Analysis of Large Datasets. Molecular & Cellular Proteomics. 2005;4:1194–204.

139. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nature Protocols. 2008;3:1101–8.

140. Minocha R, Long S. Simultaneous separation and quantitation of amino acids and polyamines of forest tree tissues and cell cultures within a single high-performance liquid chromatography run using dansyl derivatization. Journal of Chromatography A. 2004;1035:63–73.

141. Suarez A, Güttler A, Strätz M, Staendner LH, Timmis KN, Guzmán CA. Green fluorescent protein-based reporter systems for genetic analysis of bacteria including monocopy applications. Gene. 1997;196:69–74.

142. Tomlin KL, Clark SR, Ceri H. Green and red fluorescent protein vectors for use in biofilm studies of the intrinsically resistant *Burkholderia cepacia* complex. Journal of Microbiological Methods. 2004;57:95–106.

143. Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proceedings of the National Academy of Sciences. 2003;100:1541–6.

144. Paget MS, Chamberlin L, Atrih A, Foster, SJ, and Buttner, MJ. Evidence that the extracytoplasmic function sigma factor *sigmaE* is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). J. Bacteriol.181 (1), 204 –211.

145. Figueroa-Bossi N, Uzzau S, Maloriol D, Bossi L. Variable assortment of prophages provides a transferable repertoire of pathogenic determinants in Salmonella. Molecular Microbiology. 2001;39:260–72.

146. Kim S-K. Cloning and Characterization of the UDP Glucose/Galactose Epimerases of *Oryza sativa*. Journal of the Korean Society for Applied Biological Chemistry. 2009;52:315–20.

147. Jofré E, Fischer S, Príncipe A, Castro M, Ferrari W, Lagares A, et al. Mutation in a d-alanine-d-alanine ligase of *Azospirillum brasilense* Cd results in an overproduction of exopolysaccharides and a decreased tolerance to saline stress. FEMS Microbiology Letters. 2008;290:236–46.

148. Chen W, He S, Liu D, Patil GB, Zhai H, Wang F, et al. A Sweetpotato Geranylgeranyl Pyrophosphate Synthase Gene, IbGGPS, Increases Carotenoid Content and Enhances Osmotic Stress Tolerance in Arabidopsis thaliana. Plos One. 2015;10.

149. Polarek JW, Williams G, Epstein W. The products of the kdpDE operon are required for expression of the Kdp ATPase of *Escherichia coli*. Journal of Bacteriology. 1992;174:2145–51.

150. Hoffmann T, Schutz A, Brosius M, Volker A, Volker U, Bremer E. High-Salinity-Induced Iron Limitation in *Bacillus subtilis*. Journal of Bacteriology. 2002;184:718–27.

151. Ramos JL, Martinez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang X, et al. The TetR Family of Transcriptional Repressors. Microbiology and Molecular Biology Reviews. 2005;69:326–56.

152. Arnaouteli S, Giastas P, Andreou A, Tzanodaskalaki M, Aldridge C, Tzartos SJ, et al. Two Putative Polysaccharide Deacetylases Are Required for Osmotic Stability and Cell Shape Maintenance in *Bacillus anthracis*. Journal of Biological Chemistry. 2015;290:13465–78.

153. Sévin DC, Sauer U. Ubiquinone accumulation improves osmotic-stress tolerance in Escherichia coli. Nature Chemical Biology. 2014;10:266–72.

154. Burg MB, Ferraris JD. Intracellular Organic Osmolytes: Function and Regulation. Journal of Biological Chemistry. 2008;283:7309–13.

155. Morita YS, Yamaryo-Botte Y, Miyanagi K, Callaghan JM, Patterson JH, Crellin PK, et al. Stress-induced Synthesis of Phosphatidylinositol 3-Phosphate in *Mycobacteria*. Journal of Biological Chemistry. 2010;285:16643–50.

156. Youssefian S, Nakamura M, Sano H. Tobacco plants transformed with the O-acetylserine (thiol) lyase gene of wheat are resistant to toxic levels of hydrogen sulphide gas. The Plant Journal. 1993;4:759–69.

157. Volker U, Engelmann S, Maul B, Riethdorf S, Volker A, Schmid R, et al. Analysis of the induction of general stress proteins of *Bacillus subtilis*. Microbiology. 1994;140:741–52.

158. Kędzierska S, Staniszewska M, Węgrzyn A, Taylor A. The role of DnaK/DnaJ and GroEL/GroES systems in the removal of endogenous proteins aggregated by heat-shock from Escherichia coli cells. FEBS Letters. 1999;446:331–7.

159. Heimann JD. The extracytoplasmic function (ECF) sigma factors. Advances in Microbial Physiology. 2002;:47–110.

160. Cho J-I, Lim H-M, Siddiqui ZS, Park S-H, Kim A-R, Kwon T-R, et al. Over-expression of PsGPD, a mushroom glyceraldehyde-3-phosphate dehydrogenase gene, enhances salt tolerance in rice plants. Biotechnology Letters. 2014;36:1641–8.

161. Zeng Y, Tan X, Zhang L, Long H, Wang B, Li Z, et al. A fructose-1,6-biphosphate aldolase gene from *Camellia oleifera*: molecular characterization and impact on salt stress tolerance. Molecular Breeding. 2015;35.

162. Kohler C, Lourenço RF, Bernhardt J, Albrecht D, Schüler J, Hecker M, et al. A comprehensive genomic, transcriptomic and proteomic analysis of a hyperosmotic stress sensitive α-proteobacterium. BMC Microbiology. 2015;15.

163. Ventosa A, Márquez MC, Garabito MJ, Arahal DR. Moderately halophilic gram-positive bacterial diversity in hypersaline environments. Extremophiles. 1998;2:297–304.

164. Mohammad RM, Campbell WF. The search for salt-tolerant alfalfa/*Rhizobium*. Utah Science. 1985;46:36-37.

165. Haider S, Pal R. Integrated Analysis of Transcriptomic and Proteomic Data. Current Genomics. 2013;14:91–110.

166. Manetas Y, Petropoulou Y, Karabourniotis G. Compatible solutes and their effects on phosphoenolpyruvate carboxylase of C4-halophytes. Plant, Cell and Environment. 1986;9:145–51.

167. Pandit A, Rai V, Sharma TR, Sharma PC, Singh NK. Differentially expressed genes in sensitive and tolerant rice varieties in response to salt-stress. Journal of Plant Biochemistry and Biotechnology. 2011;20:149–54.

168. Kidric M, Kos J, Sabotic J. Proteases and their endogenous inhibitors in the plant response to abiotic stress. Bot Serbica. 2014;38:139–158

169. Rushton PJ, Somssich IE. Transcriptional control of plant genes responsive to pathogens. Current Opinion in Plant Biology. 1998;1:311–5.

170. Humpel A, Gebhard S, Cook GM, Berney M. The SigF Regulon in Mycobacterium smegmatis Reveals Roles in Adaptation to Stationary Phase, Heat, and Oxidative Stress. Journal of Bacteriology. 2010;192:2491–502.

171. Helmann JD. Anti-sigma factors. Current Opinion in Microbiology. 1999;2:135–41

172. Paget MSB, Molle V, Cohen G, Aharonowitz Y, Buttner MJ. Defining the disulphide stress response in Streptomyces coelicolor A3(2): identification of the *sigmaR* regulon. Molecular Microbiology. 2001;42:1007–20.

173. Hughes KT, Mathee K. The Anti-Sigma Factors. Annual Review of Microbiology. 1998;52:231–86.

174. López C, Heras H, Garda H, Ruzal S, Sánchez-Rivas C, Rivas E. Biochemical and biophysical studies of *Bacillus subtilis* envelopes under hyperosmotic stress. International Journal of Food Microbiology. 2000;55:137–42.

175. Piuri M, Sanchez-Rivas C, Ruzal S. Cell wall modifications during osmotic stress in Lactobacillus casei. Journal of Applied Microbiology. 2005;98:84–95.

176. Miller KJ, Wood JM. Osmoadaptation By Rhizosphere Bacteria. Annual Review of Microbiology. 1996;50:101–36.

177. Prã-ncipe A, Jofrã© E, Alvarez F, Mori G. Role of a serine-type d-alanyl-d-alanine carboxypeptidase on the survival ofOchrobactrumsp. 11a under ionic and hyperosmotic stress. FEMS Microbiology Letters. 2009;295:261–73.

178. Seifert GJ. Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. Current Opinion in Plant Biology. 2004;7:277–84.

179. Feingold DS, Avigad G. Sugar nucleotide transformations in plants. In: Hatch MD, Stumpf PK, editors. The biochemistry of plants: a comprehensive treatise. New York, N.Y. u.a.: Acad. Pr.; 1981. p. 101–170

180. Murata N, Wada H. Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria. Biochemical Journal. 1995;308:1–8.

181. Klein W, Weber MHW, Marahiel MA. Cold shock response of *Bacillus subtilis*: isoleucine-dependant switch in the fatty acid branching pattern for membrane adaptation to low temperatures. J Bacteriol 181:5341–5349

182. Beck HC, Hansen AM, Lauritsen FR. Catabolism of leucine to branched-chain fatty acids in *Staphylococcus xylosus*. J Appl Microbiol. 2004;96:1185–1193

183. Nagarajan T, Vanderleyden J, Tripathi AK. Identification of salt stress inducible genes that control cell envelope related functions in A*zospirillum brasilense* Sp7. Molecular Genetics and Genomics. 2007;278:43–51.

184. Pumirat P, Cuccui J, Stabler RA, Stevens JM, Muangsombut V, Singsuksawat E, et al. Global transcriptional profiling of *Burkholderia pseudomallei* under salt stress reveals differential effects on the Bsa type III secretion system. BMC Microbiology. 2010;10:171.

185. Israelachvili JN. Aggregation of amphiphatic molecules into soft structures. In: Israelachvili JN, editor. Intramolecular and surface forces, with applications to colloidal and biological systems. London: Academic; p. 251–285

186. Kamada Y, Jung US, Piotrowski J, Levin DE. The protein kinase C-activated MAP kinase pathway of Saccharomyces cerevisiae mediates a novel aspect of the heat shock response. Genes & Development. 1995;9:1559–71.

187. Aguilar PS, Mendoza DD. Control of fatty acid desaturation: a mechanism conserved from bacteria to humans. Molecular Microbiology. 2006;62:1507–14.

188. Avonce N, Mendoza-Vargas A, Morett E, Iturriaga G. Insights on the evolution of trehalose biosynthesis. BMC Evolutionary Biology. BMC Evolutionary Biology. 2006;6:109.

189. Pereira CS, Hünenberger PH. Effect of Trehalose on a Phospholipid Membrane under Mechanical Stress. Biophysical Journal. 2008;95:3525–34.

190. Bartlett D. Pressure effects on in vivo microbial processes. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology. 2002;1595:367–81.

191. Galinski E. Osmoadaptation in Bacteria. Advances in Microbial Physiology Volume 37 Advances in Microbial Physiology. 1995;:273–328.

192. Jr MR, Courtenay ES, Cayley D, Guttman HJ. Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. Trends in Biochemical Sciences. 1998;23:143–8.

193. Lippert K, Galinski E. Enzyme stabilization be ectoine-type compatible solutes: protection against heating, freezing and drying. Applied Microbiology and Biotechnology. 1992;37.

194. Reina-Bueno M, Argandoña M, Salvador M, Rodríguez-Moya J, Iglesias-Guerra F, Csonka LN, et al. Role of Trehalose in Salinity and Temperature Tolerance in the Model Halophilic Bacterium *Chromohalobacter salexigens.* PLoS ONE. 2012;7.

195. Morita YS, Yamaryo-Botte Y, Miyanagi K, Callaghan JM, Patterson JH, Crellin PK, et al. Stress-induced Synthesis of Phosphatidylinositol 3-Phosphate in *Mycobacteria*. Journal of Biological Chemistry. 2010;285:16643–50.

196. Brennan P. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. Tuberculosis. 2003;83:91–7.

197. Perez-Pantoja D, Ledger T, Pieper DH, Gonzalez B. Efficient Turnover of Chlorocatechols Is Essential for Growth of *Ralstonia eutropha* JMP134(pJP4) in 3-Chlorobenzoic Acid. Journal of Bacteriology. 2003;185:1534–42.

198. Sallam KI, Tamura N, Tamura T. A multipurpose transposon-based vector system mediates protein expression in Rhodococcus erythropolis. Gene. 2007;386:173–82.