University of São Paulo
"Luiz de Queiroz" College of Agriculture

Unraveling the molecular mechanisms involved in plant growth promotion and soil bacterial community modulation by *Pantoea agglomerans* 33.1

Aline Aparecida Oliveira Ferreira

Dissertation presented to obtain the degree of Master in Sciences. Area: Agricultural Microbiology

Piracicaba
2023

Aline Aparecida Oliveira Ferreira
Bachelor in Agronomy

Unraveling the molecular mechanisms involved in plant growth promotion and soil bacterial community modulation by *Pantoea agglomerans* 33.1

Advisor:
Profª. Drª. **MARIA CAROLINA QUECINE VERDI**

Dissertation presented to obtain the degree of Master in Sciences. Area: Agricultural Microbiology

Piracicaba
2023

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Ferreira, Aline Aparecida Oliveira

Unraveling the molecular mechanisms involved in plant growth promotion and soil bacterial community modulation by *Pantoea agglomerans* 33.1 / Aline Aparecida Oliveira Ferreira. - - Piracicaba, 2023.

102 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. *Pantoea agglomerans* 33.1 2. PGPB 3. Fósforo 4. PSS 5. CRISPR-Cas9 I. Título

Aos meus avós Geni Gabriel e José C. Galvão.

DEDICO

AGRADEDIMENTOS

Agradeço a minha amada família. A todos os Oliveira, Arruda, Gabriel e Galvão, presentes fisicamente ou não. Todos tiveram um papel primordial para cada passo da minha vida e é muito gratificante saber que posso contar com eles.

Agradeço especialmente a minha mãe Angela Oliveira, que nunca mediu esforços para me proporcionar uma boa educação, com muito amor e dedicação. Mãe que sempre foi amiga também, e é meu parâmetro para tudo o que pretendo ser na vida.

Agradeço ao Lucas Antunes pela quase uma década de parceria, paciência e amor. Obrigada por nunca ter me deixado desistir.

Sou imensamente grata à minha orientadora Maria Carolina Quecine, por ter me recebido em seu grupo e sempre me apoiar em todos os âmbitos. Agradeço pelos conselhos, pela paciência de sempre, e pelas dezenas de elefantes que tirou das minhas costas ao longo do mestrado. Espero que você saiba o quão incrível é.

Agradeço especialmente a Maria Letícia Bonatelli, minha co-orientadora, por ser uma das pessoas mais fantásticas que conheci. Ela participou de todas as etapas desse mestrado, e não mediu esforços para me ensinar e amparar nas questões acadêmicas e da vida. Ela mais do que gentilmente me recebeu em Leipzig, abriu sua casa para mim, me integrou na UFZ e na cidade. Jamais esquecerei a gigante que ela é.

A Manuella Dourado, minha primeira orientadora, responsável por despertar em mim o amor pela pesquisa e pelas PGPB, e por me abrir diversas portas, inclusive essa.

Ao Helmholtz Centre for Environmetal Research (UFZ), especialmente à Sabine Kleinsteuber e Ute Lohse, pela oportunidade de estágio e pela hospitalidade com que me receberam.

A Gladys Castillo e Jéssica Ferrarezi. Duas amigas que me apoiaram desde o início do mestrado e para além, e que sempre fizeram muito por mim. Obrigada pela irmandade!

Aos amigos do CRISPR: Ariane Hannes, Bruna Moura, Everthon Figueredo e Maurício Rosa. Desbravar o mundo da edição gênica não foi trivial, mas vocês tornaram a jornada mais fácil.

A FAPESP (Processos: 2021/04834-0; 2022/07666-3) e Capes (Código de Financiamento 001) pelo apoio financeiro primordial para minha formação e manutenção.

A todos os queridos amigos do Lab. de Genética de Microrganismos “Prof. João Lúcio de Azevedo”, pela amizade e apoio diários.

Agradeço a todos os funcionários da ESALQ pelo cuidado conosco, especialmente ao Zezo (pelo zelo com nosso Lab. e amparo aos alunos), a Solizete (pelas dezenas de dúvidas gentilmente respondidas), e a Eliana (pelas orientações sobre a formatação dessa dissertação e correção).

Aos friends and teachers Neil Oliveira e Rodrigo Morais pela ajud(ON)a com o inglês.

A Thainá Gallioti pelo suporte psicológico primordial nos últimos tempos.

Ao Laboratório de Microbiologia de Solo-ESALQ/USP, nas pessoas do Prof. Fernando Andreote e Denise Mescolotti, pela cessão das AMF e apoio técnico em sua manipulação.

As Mendelindas, que fizeram do nosso apartamento um lar: Aline, Carol, Camila, Jéssica, Mariana e Merilyn.

Em tudo isso tem a mão de Deus!

Mas é preciso ter manha, é preciso ter graça

É preciso ter sonho sempre

Quem traz na pele essa marca

Possui a estranha mania de ter fé na vida

Milton Nascimento

CONTENTS

RESUMO

Desvendando o papel dos mecanismos moleculares envolvidos na promoção de crescimento vegetal e modulação da comunidade bacteriana do solo por *Pantoea agglomerans* 33.1

O cenário de aumento populacional e de preocupação com questões socioambientais e ecológicas, tem levado a uma demanda por maior produção agrícola, com menor impacto ambiental, e menor expansão de áreas cultivadas. Nesse sentido, adotar ferramentas alternativas que impulsionem a produção de alimentos tem se tornado uma necessidade. Assim, a utilização de microrganismos benéficos tem sido cada vez mais explorada dentro da agricultura. *Pantoea agglomerans* 33.1 é uma bactéria que estabelece interação simbiótica com a planta hospedeira, sobretudo cana-de-açúcar, favorecendo seu desenvolvimento por meio da produção de fitormônios e disponibilização de nutrientes. Com o objetivo de aprofundar o conhecimento a respeito da positiva interação entre a linhagem 33.1 e as plantas, nosso grupo têm estudado os diferentes mecanismos dessa relação há mais de uma década. Visando gerar mais conhecimento sobre a biologia de 33.1 e entender seus mecanismos de promoção do crescimento vegetal, especialmente através da solubilização de fósforo (P), o presente estudo teve por objetivos *i)* avaliar a interação da linhagem 33.1 com diferentes fungos micorrízicos arbusculares (FMA) na dinâmica de disponibilização de P, na promoção de crescimento de cana-de-açúcar, e na modulação da comunidade bacteriana do solo; e *ii)* obter, de maneira inédita, o genoma completo da linhagem 33.1, e a partir dele, identificar genes envolvidos na promoção de crescimento vegetal, bem como os genes de Sistemas de Secreção de Proteínas (SSP), e avaliar seu efeito na promoção de crescimento, mediante nocaute gênico via CRISPR-Cas9. Nossos resultados demonstraram a sinergia e efeito "helper" da linhagem 33.1 com *Rhizophagus intraradices*, com resultados promissores nas frações lábeis de P no solo e na atividade enzimática de fitase e fosfatase. O genoma completo dessa linhagem foi obtido com boa cobertura e qualidade, permitindo a identificação de quatro diferentes plasmídeos e fornecendo preciosas informações sobre sua biologia. Foram anotados os genes envolvidos nas principais vias de promoção de crescimento, especialmente os relacionados ao metabolismo de P. Os SSP T1SS, T5aSS, T5bSS, T6SS, as vias Tat e Sec, e sistemas acessórios foram localizados em seu genoma, permitindo a seleção dos genes *bepC* e *prn* para nocaute. Apesar da construção do sistema dois-plasmídeos para nocaute ter sido adequadamente realizada, com genes clonados nos devidos plasmídeos e estes inseridos na linhagem 33.1, não obtivemos, até o momento, o nocaute de genes de SSP. Ainda que a edição não tenha ocorrido com a técnica utilizada, abordagens alternativas estão sendo avaliadas e grandes avanços foram feitos quanto ao entendimento das interações de 33.1 com outros microrganismos, com o meio, e com a planta hospedeira, bem como o possível papel dos SSP nessa relação.

Palavras-chave: *Pantoea agglomerans* 33.1, Fósforo, Fungos micorrízicos arbusculares, Sistemas de secreção de proteínas

ABSTRACT

Unraveling the molecular mechanisms involved in plant growth promotion and soil bacterial community modulation by *Pantoea agglomerans* 33.1

The scenario of population growth and concern for socio-environmental and ecological issues has led to a demand for increased agricultural production with less environmental impact and reduced expansion of cultivated areas. In this sense, adopting alternative tools that boost food production has become a necessity. Thus, the use of beneficial microorganisms has been increasingly explored in agriculture. *Pantoea agglomerans* 33.1 is a bacterium that establishes a symbiotic interaction with the host plant, especially sugarcane, favoring its development through the production of phytohormones and the provision of nutrients. To deepen the knowledge regarding the positive interaction between strain 33.1 and plants, our group has been studying the different mechanisms of this relationship for over a decade. Aiming to generate more knowledge about the biology of 33.1 and understand its mechanisms of promoting plant growth, especially through phosphorus (P) solubilization, the present study had the objectives *i)* to evaluate the interaction of strain 33.1 with different arbuscular mycorrhizal fungi (AMF) in the dynamics of P availability, in the promotion of sugarcane growth, and the modulation of the soil bacterial community; and *ii)* to obtain, in an unprecedented way, the complete genome of 33.1, and from it, identify genes involved in promoting plant growth, as well as the Protein Secretion Systems (PSS) genes, and evaluate their effect on growth promotion, through gene knockout via CRISPR-Cas9. Our results demonstrated the synergy and "helper" effect of strain 33.1 with *Rhizophagus intraradices*, with promising results in the labile P fractions in the soil and enzymatic activity of phytase and phosphatase. The complete genome of 33.1 was obtained with good coverage and quality, allowing the identification of four different plasmids and providing valuable information about its biology. Genes involved in the main growth promotion pathways were annotated, especially those related to P metabolism. The T1SS, T5aSS, T5bSS, T6SS PSS, the Tat and Sec pathways, and accessory systems were identified in its genome, allowing the selection of *bepC* and *prn* genes for knockout. Despite the successful construction of the two-plasmid system for knockout, with genes cloned in the respective plasmids and inserted into strain 33.1, we have not yet achieved the knockout of PSS genes. Although editing did not occur with the technique used, alternative approaches are being evaluated, and significant progress has been made in understanding the interactions of 33.1 with other microorganisms, the environment, and the host plant, as well as the possible role of PSS in this relationship.

Keywords: *Pantoea agglomerans* 33.1, Phosphorus, Arbuscular mycorrhizal fungi, Protein secretion systems

1. INTRODUCTION

According to the United Nations, it is estimated that the world population will reach 8.5 billion inhabitants by 2030 and 9.7 billion by 2050 (United Nations, 2022). Due to the populational growth, concerns about several socio-environmental issues involved in the demands of foods and goods are also increasing. The big challenge is to increase food production using the smallest possible area and with minimal environmental impact.

The high agricultural productivity with reduced environmental impact requires the use of innovative tools. The use of Plant Growth-Promoting Bacteria (PGPB) is one relevant alternative. PGPB contribute to the protection, development, and productivity of plants through various mechanisms, such as the production of phytohormones, nutrient fixation and uptake, antibiotic production, among others (Glick, 2012).

The *Pantoea agglomerans* 33.1 is a PGPB, isolated from *Eucalyptus grandis* (Procópio et al., 2009), and has been studied by our group for at least a decade. The enhancement of development and productivity in crops, especially sugarcane (Quecine, 2010; Quecine et al.; 2012), caused by the 33.1, is due to its ability to produce phytohormones and enzymes, and especially to fix and capture nutrients. In this context, the 33.1 stands out as a phosphate-solubilizing bacterium, by its potential to solubilize inorganic phosphorus through the production and secretion of organic acids, and to mineralize organic phosphorus through enzyme production, for example (data not published).

A deep understanding of the biology of 33.1 and its interactions with the environment becomes essential considering its potential agricultural applications. For this purpose, several studies have been conducted by our group to increase understanding of the molecular mechanisms used by the 33.1 to promote the plant growth.

Due to its participation in various vital processes for plants, such as photosynthesis and biosynthesis of macromolecules, for example, phosphorus (P) is among the most important elements for plant development (Rawat et al., 2020), and it is estimated that approximately 90% of the global demand for P comes from food production (Cordell, Drangert, and White, 2009). Low P concentrations affect plant growth and metabolism, and P availability in the agricultural system can lead to significant constraints on food production (Raghothama, 1999).

The transformations and interactions of P with other elements in the soil have been known for decades and are significant challenges for agricultural production. Goldstein (1992) reported that up to 75% of applied phosphate fertilizers could be reprecipitated, leading farmers to apply up to 4 times the recommended fertilizer dose. This scenario represents a significant environmental and productivity issues, as rock phosphate, which comprises a significant portion of the phosphate source used in fertilizers, is a non-renewable resource (Hawkesford et al., 2012), with global reserve depletion estimated between 50 and 100 years (Cordell, Drangert, and White, 2009).

Protein Secretion Systems (PSS) can be defined as machinery that allows the transport of proteins through the bacterial outer membrane (Denise, Abby, and Rocha, 2020). The secreted substrates have a wide range of functions and are often associated with pathogenicity due to their potential to secrete toxins, enabling bacterial modulation and interactions with the environment and other organisms.

Despite the extensive research on the use of PSS in pathogenicity, studies evaluating their potential in beneficial interactions are scarce. There are few studies related to the use of PSS by PGPB and their importance in beneficial bacterial interactions with crops. The importance of certain PSS in secreting crucial compounds used by bacteria in the processes of P solubilization and mineralization has been understanding (Liu et al., 2020).

It is important to consider the scenarios: *i)* the relevance of supplementation for crops using sustainable alternatives, *ii)* the need to understand the mechanisms used by the PGPB to promote plant growth, especially through the dynamics of P solubilization and mineralization, as well as its interactions with other solubilizing microorganisms, and *iii)* the increasing interest and research related to the use of PSS in beneficial bacteria/host interactions. Thus, we hypothesized that *i)* the PGPB *P. agglomerans* 33.1 interacts, in synergy, with arbuscular mycorrhizal fungi (AMF), to promote plant growth and to modulate the soil bacterial community; and *ii)* the PSS from 33.1 has a key role in 33.1/plant interaction and in the host growth promotion.

Therefore, the objective of this research was to investigate the dynamics of sugarcane growth promotion and P availability by the 33.1, as well as the modulation of the soil bacterial community when this strain is associated (or not) with AMF (Chapter 1). We also performed the whole-genome sequencing of 33.1, identifying genes related to plant growth promotion, predicting its PSS, and finally, evaluating whether the PSS of this strain

would be used to achieve plant growth promotion, through the knockout of PSS genes using the CRISPR-Cas9 technology (Chapter 2).

1.1. The genus *Pantoea*

The genus *Pantoea* was proposed by Gavini et al. (1989), to include the previously classified bacterium in the complex *Erwinia herbicola- Enterobacter agglomerans* as *Pantoea agglomerans*, and to describe the new species *Pantoea dispersa*. Members of this genus are Gram-negative, non-encapsulated, non-spore-forming, rod-shaped, and yellow-pigmented bacteria (Gavini et al., 1989; Walterson and Stavrinides, 2015; Tambong, 2019).

Pantoea belongs to the *Erwiniaceae* family and currently harbors 17 species: *Pantoea agglomerans*, *Pantoea eucalyptii*, *Pantoea vagans*, *Pantoea conspicua*, *Pantoea deleyi*, *Pantoea anthophila*, *Pantoea brenneri*, *Pantoea ananatis*, *Pantoea allii*, *Pantoea stewartii*, *Pantoea cypripedii*, *Pantoea dispersa*, *Pantoea séptica*, *Pantoea wallisii*, *Pantoea eucrina*, *Pantoea rodasii,* and *Pantoea rwandensis* (Tambong, 2019).

This genus is widespread and ubiquitous and can be found in several niches such as soil, plants, sediments, aquatic environments, animals, and even humans (Brady et al., 2008; Rezzonico et al., 2009). The complexity of *Pantoea* members is reflected in their ability to assume different roles, including as pathogens, commensals, or symbionts (Shariati et al., 2017).

The most extensively studied species from this genus is *P. agglomerans.* When associated with the plant, either in the rhizosphere or endosphere, *P. agglomerans* can adopt mechanisms to enhance plant growth promotion, such as acquisition and uptake of nutrients (e.g., nitrogen fixation), production and regulation of phytohormones, and production of antibiotics, among others (Glick, 2012; Dutkiewicz et al., 2016, Lorenzi et al., 2022).

P. agglomerans has already been isolated as an endophyte in several crops, including citrus, sweet potato, and corn (Teixeira, Melo, and Vieira, 2005). In addition, many strains of *P. agglomerans* have been reported to contribute to the biocontrol of pathogens in more than 20 crops. Examples include the biocontrol of *Meloidogyne incognita* in tomatoes (Munif, Hallmann, and Sikora, 2001) and *Botrytis cinerea* in apples, pears, and lentils (Nunes et al., 2002; Huang and Erickson, 2002). Furthermore, this species shows

potential for important applications in bioremediation (Jacobucci, Oriani, and Durrant, 2009).

This promising interaction between *P. agglomerans* and plants has been explored in commercial products in Brazil and abroad. In New Zealand, Blossom Bless (Grochem) uses *P. agglomerans* P10c to protect pip fruit flowers against the fire blight infection (*Erwinia amylovora*). Equivalent to Bloomtime (Nap), using *P. agglomerans* E325 in Canada. In Brazil, the commercial product in development, Sprinter (Bionat) utilizes 33.1, an inoculant recommended for enhancing plant growth in various crops, such as sugarcane, corn, soybean, and common bean. This strain is known for its ability to solubilize P (Quecine et al., 2012).

33.1 is an endophytic bacterium isolated from *Eucalyptus grandis* (Procópio et al., 2009) and well described as a growth promoter, especially to sugarcane (Quecine, 2010). Quecine et al. (2012) demonstrated that inoculation with 33.1 can increase the shoot dry weight of sugarcane plants by over 30%. Quecine et al. (2014) also showed that 33.1, when transformed with a plasmid that harbors the *Cry* gene (in this case, the pJTT vector), becomes able to control sugarcane borer (*Diatraea saccharalis*).

The biotechnological potential of 33.1 is justified by its ability to produce indole-3-acetic acid (IAA), fix nitrogen, solubilize phosphates, and produce several enzymes such as lipase, pectinase, endoglucanase, and esterase, in addition to stimulate the production of proteins such as chitinase and cellulase by the host plant (Quecine, 2010; Quecine et al., 2012), thus demonstrating its importance and potential in a more productive and sustainable agricultural scenario.

1.2. Protein Secretion Systems (PSS)

Protein secretion is an essential process for bacterial development and survival. It allows the secretion of different substrates, such as small molecules, proteins, and even DNA. These substrates can be secreted into the environment, into a host - which can be a eukaryotic cell or even another prokaryotic cell - or into other cellular compartments (Costa et al., 2015; Green and Mecsas, 2016).

The different substrates secreted by the bacterium are essential in its diverse physiological processes, such as adhesion, adaptation, and survival (Costa et al., 2015).

Therefore, protein secretion represents an important mechanism for bacterial interactions and development. Prokaryotes use different methods to secrete their substrates, but frequently this process is mediated by Protein Secretion Systems (PSS).

For Gram-negative bacteria, PSS can be defined as protein transport machinery across the outer membrane (Denise, Abby, and Costa, 2020). Currently, at least six PSS have been identified in Gram-negative bacteria (Types I to VI, also described as T1SS to T6SS) (Table 1).

Table 1. Description of PSS in Gram-negative bacteria.

Secretion System	Characteristics
Type I Secretion System (T1SS)	-Bacterial T1SS mediates the secretion of a wide range of protein substrates from the cytoplasm to the extracellular environment; -The secreted products, which vary in size and function, are associated with nutrition and virulence; -T1SS can promote antibacterial competition in a contact-dependent manner; -T1SS consists of three essential structural components: an ABC transporter, a membrane fusion protein, and an outer membrane factor.
Type II Secretion System (T2SS)	-Due to its channel is just found in the outer membrane, proteins transported through this system need to be initially delivered to the periplasm via the Sec or Tat secretion pathways, which transport them across the inner membrane; -T2SSs can secrete a wide variety of substrates outside the bacterial cell, some of which contribute to the virulence of pathogens; -In some bacterial species, T2SS is necessary for the secretion of multiple substrates, while in others, it is used only to transport a single protein; -T2SS is complex and composed of at least 15 proteins.
Type III Secretion System (T3SS)	-Found in many pathogenic and symbiotic Gram-negative bacteria; -Secretes a wide variety of protein substrates across the inner and outer bacterial membranes in a single step; -Facilitates the transfer of bacterial effector proteins into the cytoplasm or

	plasma membrane of target cells; -T3SS consists of 9 essential proteins.
Type IV Secretion System (T4SS)	-T4SSs secrete substrates in a single step, facilitating the transfer of their substrates to the host/target cell; -T4SSs are related to bacterial DNA conjugation systems and can secrete a variety of substrates, including proteins and protein-protein and DNA-protein complexes; -Due to their ability to transfer DNA and proteins, T4SSs can be involved in DNA conjugative transfer, DNA uptake and release, and translocation of effector proteins or DNA/protein complexes directly into recipient cells; -The model T4SS (from *Agrobacterium tumefaciens*) consists of 12 proteins.
Type V Secretion System (T5SS)	-T5SS is found in the outer membrane, is dependent on Sec Pathway, and performs protein transport in two steps; -T5SS is the only system where the substrate and its secretion pore fuse to form a polypeptide; -The best-known T5SS substrates are virulence proteins, serving as toxins and receptor-binding proteins.
Type VI Secretion System (T6SS)	-T6SSs translocate proteins into a variety of recipient cells, including eukaryotic target cells and, most commonly, other bacteria, in a single step; -It translocates toxic effector proteins into eukaryotic and prokaryotic cells and plays a fundamental role in bacterial pathogenesis and competition; -The model T6SS consists of 13 essential components.

Adapted from Costa et al., 2015; Green and Mecsas, 2016.

PSS represent a diversity of characteristics and substrates: some of them are conserved across multiple species, while others are specific to only a few species. Furthermore, depending on the PSS, a broad range of substrates can be secreted, or only a limited number of proteins (Green and Mecsas, 2016). The substrates can remain associated with the outer membrane, be released into the extracellular space, or injected into a target cell, depending on the characteristics of the secretion system (Costa et al., 2015).

The secreted substrates vary depending on the type of secretion system and the bacterial characteristics. As example, initially related to secreting RTX proteins (repeats-in-

toxins), it is known the T1SS can secrete a diverse range of substances, either in size -from proteins with 17kDa in *Serratia marcescens* (Létoffé, Ghigo, and Wandersman, 1994), until proteins up to 900 kDa in *Pseudomonas fluorecens* (Hinsa et al., 2003)- or in function- this system often secretes bacteriocins, proteins related to nutrient acquisition and adhesins (Maphosa, Moleleki, and Motaung, 2023).

In Gram-negative bacteria, there are two categories of secretion machinery (Figure 1): *i)* those that enable single-step secretion: where the machinery is designed to traverse both membranes (inner and outer membranes), transporting the protein directly from the cytoplasm, thus the secretion occurs in a single step (T1SS, T3SS, T4SS, and T6SS), and *ii)* those where secretion occurs in two steps: where the machinery is designed to traverse only the outer membrane, transporting the protein from the periplasm. These systems typically rely on the Sec and Tat mechanisms that mediate the transport of the substrate from the cytoplasm to the periplasm (T2SS and T5SS) (Costa et al., 2015; Green and Mecsas, 2016; Denise, Abby, and Costa, 2020). Additionally, T3SS, T4SS and T6SS, due to their characteristics, can inject the substrates directly into the host cell (Tseng, Tyler, and Setubal, 2009; Costa et al., 2015; Green and Mecsas, 2016).

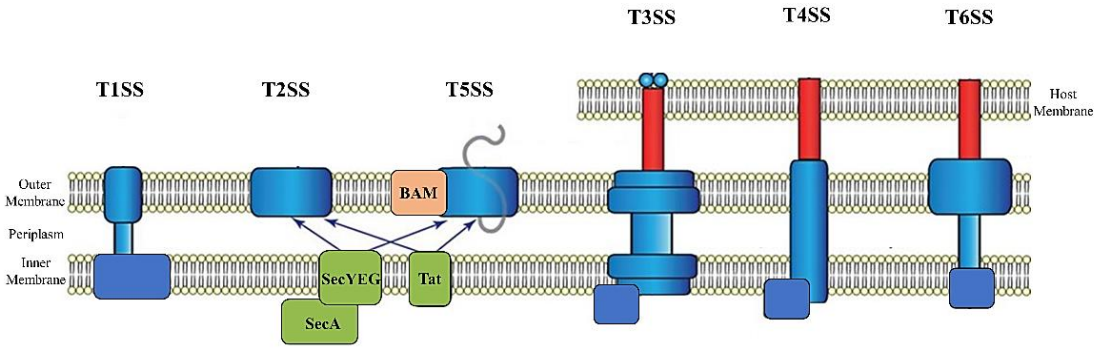

Figure 1. Representation of PSS of Gram-negative bacterium. In this picture are represented the type I secretion system (T1SS), the systems that rely on the Sec and Tat Pathways (T2SS and T5SS) and the Bam Complex (T5SS), all of them without direct interaction with the host. Additionally, systems able to interact with the target/host cell (T3SS, T4SS, and T6SS) are represented (Adapted from Tseng, Tyler, and Setubal, 2009).

To mediate the transport of the substrate from the cytoplasm to the periplasm, two main secretion pathways can be utilized: the General Secretion Pathway (known as the Sec pathway) and the Twin-Arginine Translocation Pathway (known as the Tat pathway). The Sec pathway catalyzes the transmembrane translocation of unfolded proteins, while the Tat

pathway translocate folded proteins (Natale, Brüser, and Driessen, 2008). The conformation of these secretion mechanisms differs between Gram-negative and Gram-positive bacteria.

The Sec pathway, which was the first secretion pathway discovered in bacteria (Beckwith, 2013), consists of the SecYEG channel (composed of three membrane proteins: SecY, SecE, and SecG), the ATPase SecA, which acts as a motor for protein translocation across the membrane, and the chaperone SecB, which prevents protein folding and enables them to pass through the SecYEG complex (Natale, Brüser, and Driessen, 2008; Beckwith, 2013).

In contrast, the Tat pathway is found in bacteria, archaea, and plant chloroplasts and allows the transport of fully folded proteins across the membrane (Patel, Smith, and Robinson, 2014). The Tat pathway requires the subunits TatA, TatB, and TatC. In *Escherichia coli*, TatB and TatC bind to the signal peptide of Tat-secreted proteins, recruiting TatA to form a channel across the membrane, enabling the transport of folded proteins to the periplasm (Green and Mecsas, 2016).

Besides the Tat and Sec Pathways, another complex may be important for protein transport: the Bam (β-barrel Assembly Machinery) Complex. Composed of BamA (outer membrane protein), BamB, BamC, BamD, and BamE (lipoproteins), the Bam Complex plays an essential role in the folding of Outer Membrane Proteins (OMP) into a β-barrel conformation and in their insertion into the outer membrane (Van Ulsen et al., 2014). For Gram-negative T5SS, for example, this activity is vital.

Many pathogens utilize PSS to secrete virulence factors from the bacterial cytosol into host cells or the environment. In other cases, pathogenic bacteria can use PSS to manipulate the host and establish a niche for multiplication and colonization (Green and Mecsas, 2016). In an agricultural context, it is known, for example, that without the T3SS, the pathogen cannot suppress the plant's basal defenses, reproduce, cause lesions in hosts, or induce a hypersensitive response in non-hosts (Alfano and Collmer, 2004).

Despite their importance for pathogenic bacteria, some studies suggest that the same systems can benefit the host plant. According to Chang, Desveaux, and Creason (2014), the T3SS machinery could stimulate plant immunity based on research conducted with *Pseudomonas syringae* mutants. Tseng, Tyler, and Setubal (2009) report that although initially discovered in pathogenic bacteria, T3SS systems have been identified in nitrogen-fixing bacteria.

According to Becker et al. (2018), the colonization and plant growth promotion in several species by the endophytic bacterium *Kosakonia radicincitans* may be facilitated by genes that contribute to its motility (two flagellar systems) and high competitiveness (presence of three T6SS). Bernal, Llamas, and Filloux (2018) suggested that this secretion system (T6SS), although related to pathogenicity, is also widely distributed in Gram-negative commensal and symbiotic bacteria.

T1SS and T2SS have been described as responsible to secrete organic acids and enzymes into the environment, especially those related to the phosphate solubilization and mineralization processes in *Burkholderia multivorans* WS-FJ9 (Liu et al., 2020). Moreover, Lucero et al. (2022) have observed that the absence of T6SS interferes in the colonization process of *Enterobacter* sp. J49 -a phosphate solubilizing bacterium- in peanuts plants. According to Allsopp et al. (2020), it is evident that T6SS provides an advantage in several model systems, particularly in the agricultural context.

1.3. Phosphate-solubilizing microorganisms

P is one of the most important nutrients for plant development, participating in processes such as photosynthesis, cell division, and biosynthesis of macromolecules, among others (Rawat et al., 2020). In the soil, this nutrient can be found in inorganic (Pi), organic (Po), and microbial (Pmic) pools (Rawat et al., 2020; Bergkemper et al., 2016).

Among 35% to 70% of the total soil P is in Pi form (Ducousso-Détrez et al., 2022). However, the major fractions of Pi are absorbed or fixed on soil particles, making less than 1% of Pi available for plants (Wang et al., 2022). In soil, Po is originated from vegetal waste, microbial biomass, and products of its decomposition (Martinazzo et al., 2007), and is estimated that organic fraction constitutes between 30% to 65% of the total P in the soils (Ducousso-Détrez et al., 2022). On the other hand, Pmic is variable and represents a dynamic source of P in the system (Bini and Lopez, 2016).

Through the solubilization and mineralization processes the inaccessible Pi and Po, respectively, can be converted into available forms of P (Khan, Zaidi and Ahmad, 2014). Several microorganisms, called Phosphate Solubilizing Microorganisms (PSM), can convert insoluble P into available P to plants (Rawat, et al., 2020).

PSM can adopt different mechanisms to provide soluble P to plants, depending on its source. These microorganisms can solubilize Pi through processes such as the production of organic acids; production of inorganic acids and H_2S; proton extrusion; pathways of direct oxidation, exopolysaccharides (EPSs) and siderophore production (Rawat et al., 2020).

Among the several mechanisms to solubilize P used for PSM, the secretion of low molecular mass organic acids is the well-recognized and widely accepted way of P-solubilization (Khan et al., 2014).

On the other hand, enzymes produced and secreted by PSM play a crucial role in the mineralization process, facilitating the dissolution of Po. According to Rawat et al. (2020), three important enzymes are involved in P-mineralization: *i)* non-specific acid phosphatases (NSAPs) or phosphomonoesterases, that can be classified into two groups (alkaline or acid phosphatases); *ii)* phytases, and *iii)* phosphonatases/Carbon-Phosphorus (C-P) lyases.

REFERENCES

Alfano, J. R., & Collmer, A. (2004). Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu. Rev. Phytopathol.*, *42*, 385-414.

Allsopp, L. P., Bernal, P., Nolan, L. M., & Filloux, A. (2020). Causalities of war: the connection between type VI secretion system and microbiota. *Cellular microbiology*, *22*(3), e13153.

Becker, M., Patz, S., Becker, Y., Berger, B., Drungowski, M., Bunk, B., ... & Ruppel, S. (2018). Comparative genomics reveal a flagellar system, a type VI secretion system and plant growth-promoting gene clusters unique to the endophytic bacterium Kosakonia radicincitans. *Frontiers in microbiology*, *9*, 1997.

Beckwith, J. (2013). The Sec-dependent pathway. *Research in microbiology*, *164*(6), 497-504.

Bergkemper, F., Bünemann, E. K., Hauenstein, S., Heuck, C., Kandeler, E., Krüger, J., ... & Schulz, S. (2016). An inter-laboratory comparison of gaseous and liquid fumigation based methods for measuring microbial phosphorus (Pmic) in forest soils with differing P stocks. *Journal of microbiological methods*, *128*, 66-68.

Bernal, P., Llamas, M. A., & Filloux, A. (2018). Type VI secretion systems in plant-associated bacteria. *Environmental microbiology*, *20*(1), 1-15.

Bini, D., & Lopez, M. V. (2016). Transformações Microbianas do Fósforo. In Cardoso, E. J. B. N., & Andreote, F. D., *Microbiologia do Solo,* 2ª ed., (pp. 149-166), Piracicaba: ESALQ.

Brady, C., Cleenwerck, I., Venter, S., Vancanneyt, M., Swings, J., & Coutinho, T. (2008). Phylogeny and identification of Pantoea species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA). *Systematic and Applied Microbiology*, *31*(6-8), 447-460.

Chang, J. H., Desveaux, D., & Creason, A. L. (2014). The ABCs and 123s of bacterial secretion systems in plant pathogenesis. *Annual review of phytopathology*, *52*, 317-345.

Cordell, D., Drangert, J. O., & White, S. (2009). The story of phosphorus: global food security and food for thought. *Global environmental change*, *19*(2), 292-305.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*, *13*(6), 343-359.

Denise, R., Abby, S. S., & Rocha, E. P. (2020). The evolution of protein secretion systems by co-option and tinkering of cellular machineries. *Trends in Microbiology*, *28*(5), 372-386.

Ducousso-Détrez, A., Fontaine, J., Lounès-Hadj Sahraoui, A., & Hijri, M. (2022). Diversity of phosphate chemical forms in soils and their contributions on soil microbial community structure changes. *Microorganisms*, *10*(3), 609.

Dutkiewicz, J., Mackiewicz, B., Lemieszek, M. K., Golec, M., & Milanowski, J. (2016). Pantoea agglomerans: a mysterious bacterium of evil and good. Part IV. Beneficial effects. *Annals of Agricultural and Environmental Medicine*, *23*(2).

Gavini, F., Mergaert, J., Beji, A. M. O. R., Mielcarek, C., Izard, D., Kersters, K., & De Ley, J. (1989). Transfer of Enterobacter agglomerans (Beijerinck 1888) Ewing and Fife 1972 to Pantoea gen. nov. as Pantoea agglomerans comb, nov. and Description of Pantoea dispersa sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *39*(3), 337-345.

Glick, B. R. (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica*, *2012*.

Goldstein, A. H. (1992, July). Phosphate starvation inducible enzymes and proteins in higher plants. In *Society for Experimental Biology Seminar Series* (Vol. 49, pp. 25-44).

Green, E. R., & Mecsas, J. (2016). Bacterial secretion systems: an overview. *Virulence mechanisms of bacterial pathogens*, 213-239.

Hawkesford, M., Horst, W., Kichey, T., Lambers, H., Schjoerring, J., Møller, I. S., & White, P. (2012). Functions of macronutrients. In *Marschner's mineral nutrition of higher plants* (pp. 135-189). Academic press.

Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L., & O'Toole, G. A. (2003). Transition from reversible to irreversible attachment during biofilm formation by Pseudomonas fluorescens WCS365 requires an ABC transporter and a large secreted protein. *Molecular microbiology*, *49*(4), 905-918.

Huang, H. C., & Erickson, R. S. (2002). Biological control of Botrytis stem and blossom blight of lentil. *Plant Pathology*, *11*(1), 8.

Jacobucci, D. F. C., Oriani, M. R. D. G., & Durrant, L. R. (2009). Reducing COD level on oily effluent by utilizing biosurfactant-producing bacteria. *Brazilian Archives of Biology and Technology*, *52*, 1037-1042.

Khan, M. S., Zaidi, A., & Ahmad, E. (2014). Mechanism of phosphate solubilization and physiological functions of phosphate-solubilizing microorganisms. *Phosphate solubilizing microorganisms: principles and application of microphos technology*, 31-62.

Letoffe, S., Ghigo, J. M., & Wandersman, C. (1994). Secretion of the Serratia marcescens HasA protein by an ABC transporter. *Journal of bacteriology*, *176*(17), 5372-5377.

Liu, Y. Q., Wang, Y. H., Kong, W. L., Liu, W. H., Xie, X. L., & Wu, X. Q. (2020). Identification, cloning and expression patterns of the genes related to phosphate solubilization in Burkholderia multivorans WS-FJ9 under different soluble phosphate levels. *AMB Express*, *10*(1), 1-11.

Lorenzi, A. S., Bonatelli, M. L., Chia, M. A., Peressim, L., & Quecine, M. C. (2022). Opposite sides of Pantoea agglomerans and its associated commercial outlook. *Microorganisms*, *10*(10), 2072.

Lucero, C. T., Lorda, G. S., Ludueña, L. M., Nievas, F., Bogino, P. C., Angelini, J., ... & Taurian, T. (2022). Participation of type VI secretion system in plant colonization of phosphate solubilizing bacteria. *Rhizosphere*, *24*, 100582.

Maphosa, S., Moleleki, L. N., & Motaung, T. E. (2023). Bacterial secretion system functions: evidence of interactions and downstream implications. *Microbiology*, *169*(4).

Martinazzo, R., Santos, D. R. D., Gatiboni, L. C., Brunetto, G., & Kaminski, J. (2007). Fósforo microbiano do solo sob sistema plantio direto em resposta à adição de fosfato solúvel. *Revista Brasileira de Ciência do solo*, *31*, 563-570.

Munif, A., Hallmann, J., & Sikora, R. A. (2001). Induced systemic resistance of selected endophytic bacteria against Meloidogyne incognita on tomato. *Mededelingen (Rijksuniversiteit te Gent. Fakulteit van de Landbouwkundige en Toegepaste Biologische Wetenschappen)*, *66*(2b), 663-669.

Natale, P., Brüser, T., & Driessen, A. J. (2008). Sec-and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1778*(9), 1735-1756.

Nunes, C., Usall, J., Teixido, N., Torres, R., & Vinas, I. (2002). Control of Penicillium expansum and Botrytis cinerea on apples and pears with the combination of Candida sake and Pantoea agglomerans. *Journal of Food Protection*, *65*(1), 178-184.

Patel, R., Smith, S. M., & Robinson, C. (2014). Protein transport by the bacterial Tat pathway. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1843*(8), 1620-1628.

Procópio, R. E. L., Araújo, W. L., Maccheroni Jr, W., Azevedo, J. L. (2009). Characterization of an endophytic bacterial community associated with Eucalyptus spp. *Genetics and Molecular Research*, *8*(4), 1408-1422.

Quecine, M. C. (2010). *Aspectos biotecnológicos da interação entre bactérias e cana-de-açúcar (Saccharum sp., L.)* (Doctoral dissertation, Universidade de São Paulo).

Quecine, M. C., Araújo, W. L., Rossetto, P. B., Ferreira, A., Tsui, S., Lacava, P. T., ... & Pizzirani-Kleiner, A. A. (2012). Sugarcane growth promotion by the endophytic bacterium Pantoea agglomerans 33.1. *Applied and Environmental Microbiology*, *78*(21), 7511-7518.

Quecine, M. C., Araújo, W. L. D., Tsui, S., Parra, J. R. P., Azevedo, J. L. D., & Pizzirani-Kleiner, A. A. (2014). Control of Diatraea saccharalis by the endophytic Pantoea agglomerans 33.1 expressing cry1Ac7. *Archives of microbiology*, *196*, 227-234.

Rawat, P., Das, S., Shankhdhar, D., & Shankhdhar, S. C. (2021). Phosphate-solubilizing microorganisms: mechanism and their role in phosphate solubilization and uptake. *Journal of Soil Science and Plant Nutrition*, *21*, 49-68.

Raghothama, K. G. (1999). Phosphate acquisition. *Annual review of plant biology*, *50*(1), 665-693.

Rezzonico, F., Smits, T. H., Montesinos, E., Frey, J. E., & Duffy, B. (2009). Genotypic comparison of Pantoea agglomerans plant and clinical strains. *BMC microbiology*, *9*(1), 1-18.

Shariati J, V., Malboobi, M. A., Tabrizi, Z., Tavakol, E., Owlia, P., & Safari, M. (2017). Comprehensive genomic analysis of a plant growth-promoting rhizobacterium Pantoea agglomerans strain P5. *Scientific reports*, *7*(1), 1-12.

Sun, L., Cheng, L., Ma, Y., Lei, P., Wang, R., Gu, Y., ... & Xu, H. (2022). Exopolysaccharides from Pantoea alhagi NX-11 specifically improve its root colonization and rice salt resistance. *International Journal of Biological Macromolecules*, *209*, 396-404.

Tambong, J. T. (2019). Taxogenomics and systematics of the genus Pantoea. *Frontiers in microbiology*, *10*, 2463.

Teixeira, M. A., de MELO, I. S., & Vieira, R. F. (2005). Diversidade de bactérias endofíticas na cultura da mandioca. EMBRAPA Meio Ambiente. *Boletim de Pesquisa e Desenvolvimento,* 33.

Tseng, T. T., Tyler, B. M., & Setubal, J. C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology*, *9*(1), 1-9.

United Nations, Department of Economic and Social Affairs, Population Division (2022). *World Population Prospects 2022: Summary of Results*. UN DESA/POP/2022/TR/NO. 3.

Van Ulsen, P., ur Rahman, S., Jong, W. S., Daleke-Schermerhorn, M. H., & Luirink, J. (2014). Type V secretion: from biogenesis to biotechnology. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1843*(8), 1592-1611.

Walterson, A. M., & Stavrinides, J. (2015). Pantoea: insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS microbiology reviews*, *39*(6), 968-984.

Wang, Y. P., Huang, Y., Augusto, L., Goll, D. S., Helfenstein, J., & Hou, E. (2022). Toward a global model for soil inorganic phosphorus dynamics: Dependence of exchange kinetics and soil bioavailability on soil physicochemical properties. *Global Biogeochemical Cycles*, *36*(3), e2021GB007061.

2. *P. agglomerans* 33.1 A PHOSPHATE-SOLUBILIZING BACTERIUM AND ITS INTERACTIONS WITH ARBUSCULAR MYCORRHIZAL FUNGI: FROM PLANT GROWTH PROMOTION TO SOIL BACTERIAL COMMUNITY MODULATION

Aline Aparecida Oliveira Ferreira[1,2], Gladys Angelica Apaza-Castillo[1], Jéssica Aparecida Ferrarezi[1], Carolina Alessandra Hayashibara[1], Renan Fantine[1], Guilherme Hosaka[1], Maria Letícia Bonatelli[2], Maria Carolina Quecine[1*]

[1]Genetics of Microorganisms Laboratory "Prof. João Lúcio de Azevedo", Department of Genetics, Luiz de Queiroz College of Agriculture, Piracicaba, Brazil

[2]Department of Environmental Microbiology, Helmholtz Centre for Environmental Research GmbH – UFZ, Leipzig, Germany

* Correspondence:

Maria Carolina Quecine

mquecine@usp.br

Abstract

It is known that several microorganisms can promote plant growth through a range of mechanisms. The availability of P to plants is among these mechanisms. In this context, Phosphate solubilizing microorganisms (PSM) are environmentally friendly alternatives for supplying available P to plants, and consequently, increasing agricultural productivity. *Pantoea agglomerans* 33.1 is a bacterium known for its ability to promote plant growth, especially sugarcane, and its capacity to solubilize P. On the other hand, Arbuscular Mycorrhizal Fungi (AMF) represent symbiotic interactions between soil fungi and the roots of the host plant. This interaction is widely known for its role in making P available to plants. Considering the role of 33.1 and different AMF in the supply of P to the plant, understanding the biocompatibility between them, as well as their impact on the modulation of the soil bacterial community, is essential for obtaining a microbial consortium that can be used as a biofertilizer in agriculture. Thus, our research aimed to assess the effects of individual inoculation and the consortia among 33.1 and the AMF *Rhizophagus clarus, Rhizophagus intraradices,* and *Dentiscutata heterogama,* on the promotion of sugarcane growth. To understand the mechanisms in plant growth promotion underlying individual or consortia inoculation, we evaluated the height and weight of plants, the enzymatic activity of key enzymes for P assimilation, and we assessed P fractions found in the soil near the plant roots. Furthermore, we evaluated how the soil bacterial community composition was affected by PSM inoculation through 16S rRNA amplicon sequencing. Our results demonstrated the potential of AMF in making P available and promoting sugarcane growth. The strain 33.1 exhibited potential as a helper microorganism, enhancing the results achieved by applying AMF individually. Additionally, the synergy between 33.1 and AMF1 (*R. intraradices*) was demonstrated in promising results related to P fractionation and enzymatic activity in soil. This positive interaction could prove highly advantageous in agriculture by boosting the availability of P to plants and enhancing their development. These findings provide insight into the interaction preferences of the inoculated microorganisms, a perspective to be considered in strategies to improve the efficiency of P use.

Keywords: *P. agglomerans* 33.1, Arbuscular Mycorrhizal Fungi, Phosphorus, Plant Growth Promotion, Soil Bacterial Community.

2.1. Introduction

Beneficial microorganisms that can hydrolyze organic and inorganic insoluble P into a P-soluble form, which are efficiently assimilated by plants, are known PSM (Phosphate Solubilizing Microorganisms) (Kalayu, 2019). It is known that one of the most important mechanisms used by bacteria to promote plant growth is the supply of necessary and lacking nutrients and resources to plants, such as nitrogen and P (Glick, 2012).

Despite its vital role in plant development, the major fraction of P in the soil is fixed, with only a small fraction available for plant uptake (Khan, Zaidi, and Ahmad, 2014). Due to PSM' potential, they are described as eco-friendly and cost-effective approaches to overcome the P unavailability and uptake by plants (Kalayu, 2019).

Mycorrhizae comprise an important group of PSM. Mycorrhizae are the symbiotic relationship of plants with basidiomycetes fungi (Manivel et al., 2023). Arbuscular mycorrhiza fungi (AMF) are essential components of the soil ecosystem, due to their ability to regulate plant growth and nutrient intake, besides of stabilizing the soil structure (Gayathri et al., 2023). In the dynamics of P in the soil, AMF facilitate its uptake by plants by increasing the roots' absorbing surface area and mobilizing sparsely available P (Wang et al., 2017).

In an AMF colonized plant, P may be absorbed through two different pathways: in the direct pathway, P is absorbed directly in the soil-root interface via root epidermis and root hairs; on the other hand, in the indirect pathway, external AMF hyphae in the soil absorb P and translocate it quickly to AMF complex within the roots (Smith, Smith, and Jakobsen, 2003; Wang et al., 2017).

Besides their potential to provide P to the host plant and to enhance soil quality, AMF also interact with the soil bacterial community, especially with PGPB (Miransari, 2010). PGPB can increase AMF development and P uptake by plants, and the interaction between them has been reported as beneficial to plant growth (Richardson et al., 2009).

Considering the complexity of the soil system and the fact that high percentage of bacteria are not yet culturable *in vitro* (Vartoukian, Palmer, and Wade, 2010), techniques that allow the access of these microorganisms have become increasingly popular. It is of particular interesting in our work to evaluate how microbes that can promote plant growth can also interact with the surrounding microbial community.

It is possible to study microbial communities using several approaches. Nowadays, metataxonomic is one of the most common and it involves, among other techniques, the amplification and sequencing of marker genes. Usually regions of the ribosomal RNA (rRNA) gene are used because they are highly conserved across taxa (Breitwieser, Lu, and Salzberg, 2015), e.g, 16S rRNA gene for bacteria and archaea, and ITS for fungi.

To uncover the composition of prokaryotic communities, the amplicon sequencing of the 16S rRNA gene can be considered a benchmark, justified by its low cost, easy availability, the practicality of extraction and preparation kits, and extensive databases (Starke, Pylro, and Morais, 2021).

P. agglomerans 33.1 is a PGPB isolated from *Eucalyptus grandis* (Procópio et al., 2009) that can promote the growth of sugarcane (Quecine 2010; Quecine et al., 2012), soybean, common bean and corn (unpublished data). 33.1 can also be considered a PSM, since it can increase the P availability through the processes of solubilization and mineralization (Quecine et al., 2012). Since the interaction between PGPB and AMF can be beneficial to increase P assimilation and plant growth, the present study aimed to assess the effects of the interaction between 33.1 and different AMF on the sugarcane plant growth, in the dynamics of P availability, as well as the modulation of the soil bacterial community, through metataxonomics- amplicon sequencing.

2.2. Material and Methods

2.2.1. Microbial strains

In this study, different mycorrhizae related to P solubilization were aplied in co-inoculation with 33.1: *Rhizophagus intraradices*, from commercial product Rootella® BR, *Rhizophagus clarus* and *Dentiscutata heterogama* from Soil Microbiology Laboratory, (ESALQ/USP). *R. clarus* and *D. heterogama* have been previously identified as colonizers of sugarcane roots (Ferreira, 2016).

2.2.2. Greenhouse and plant growth promotion assay

The interactions among strain 33.1 and AMF as consortia, and their effect on plant growth promotion and soil bacterial community, were evaluated in a co-inoculation experiment.

The experiment was conducted under controlled conditions in the greenhouse of the Department of Genetics at ESALQ/USP, using soil collected from a pasture area at a depth of 0-20 cm. A sample of the soil was sent to fertility analysis at the Department of Soil Sciences (ESALQ/USP) (Table 2). The soil with medium sandy texture was initially classified as deficient in in P (<7 mg dm^{-3}) and potassium (K) (0.9 mmolc dm^{-3}) for sugarcane crop (Table 2), and then corrected with dolomitic limestone (RPTN 70%), left to rest for 60 days to raise the base saturation (BS%) to 60% and balance the pH in the range of 5.5-6.

Following the recommendations from Boletim 100 (van Raij, 1996) for high sugarcane yields, 35 mg kg^{-1} of urea (~30 kg ha^{-1} of N) and 100 mg kg^{-1} of KCl (~120 kg ha^{-1} of K_2O) were incorporated into the soil. Additionally, filter cake was applied at a rate of 5 g kg^{-1} (~10 t ha^{-1}) to increase the soil organic matter. As a source of Pi, 45 mg kg^{-1} of reactive Bayovar natural phosphate (28.9% P_2O_5) was incorporated, which corresponds to 50% (~90 kg ha^{-1} of P_2O_5) of the recommended dose for high sugarcane yields.

Table 2. Chemical characterization of the soil used in the greenhouse bioassay.

Ph	OM	P	S	Ca	Mg	K	Al	H+Al	SB	CEC	BS	M
-	g dm^{-3}	mg dm^{-3}		mmolc dm^{-3}							%	
4,47	11,3	<7	7,7	6,9	4,1	0,9	4,8	25,2	11,9	37,1	32	29

pH: hydrogen potential; OM: organic matter; P: phosphorus; S: sulfur; Ca: calcium; Mg: magnesium; K: potassium; Al: aluminum; H+Al: potential acidity; SB: sum of exchangeable bases; CEC: cation exchange capacity; BS: base saturation; m: aluminum saturation.

The experiment was a randomized block design with treatments arranged in a 3x2 factorial scheme, with 6 replicates per treatment. The factors evaluated were (1) the AMF and (2) the strain 33.1. The AMF evaluated were *R. intraradices* (AMF1)- from Rootella® BR (NovaTero), *R. clarus* (AMF2), and *D. heterogama* (AMF3)- both kindly provided by Soil Microbiology Laboratory, (ESALQ/USP). Hence, the treatments were evaluated with and without inoculation of strain 33.1, as well as with and without inoculation of AMF.

Clones of sugarcane (*Saccharum* sp.) genotype SP80-3280 were germinated in a mycorrhiza-free Basaplant substrate (Base Substratos), using stalks acquired from the Brazilian Sugarcane Genotype Panel, located at the Experimental Station of the Sugarcane Breeding Program (Araras, São Paulo). Plants measuring 15-20 cm in height were

transplanted into pots containing 10 kg of soil. Each pot contained one plant, representing an experimental unit.

The inoculants were applied in the furrow simultaneously with planting, using 2 mL of bacterial inoculum at 10^7 CFU mL^{-1} (OD_{600}= 0.1) grown in LB medium at 28°C with 150 rpm agitation for 24 hours, or 2 mL of the sterile medium in the non-bacterial inoculation treatment. AMF spores of *R. clarus* and *D. heterogama* were recovered from soil substrate using the wet sieving and decanting technique (Gerdemann and Nicolson, 1963), followed by centrifugation in a 70% sucrose solution. As for *R. intraradices* spores, they were directly obtained from the Rootella® BR commercial formulation. In the AMF treatments, 50 spores per plant with viable morphological characteristics were applied.

Sugarcane plants were sampled 120 days after inoculation (dai) and their height and weight were measured. Bulk and rhizosphere soils were also sampled and kept at -80°C for downstream analysis. The results obtained for the evaluated parameters were then subjected to analysis of variance (ANOVA) with comparison of treatment means by Tukey's Test ($p<0.05$).

2.2.3. Dynamics of P availability

The dynamics of P in the soil was evaluated through the fractionation of P in the rhizospheric soil, and the enzymatic activity (for phytase and phosphatases) of the rhizosphere microbiome.

The chemical analysis of P fractionation was carried out by the Laboratory of Research in Soil Fertility, Fertilization and Nutrition, (ESALQ/USP), following the methodology proposed by Hedley, Stewart, and Chauhan (1982), with modifications by Condron, Goh, and Newman (1985). The analysis was performed in the rhizospheric soil collected 120 dai, with 0.5 g of soil subjected to successive extractions with anion exchange resin (P_{RTA}), $NaHCO_3$ 0,5 mol L^{-1} (P_{BIC}), NaOH 0,1 mol L^{-1} ($P_{Hid\text{-}0,1}$), HCl 1,0 mol L–1 (P_{HCl}), and NaOH 0,5 mol L^{-1} ($P_{Hid\text{-}0,5}$). After extraction, the soil was subjected to digestion with H_2SO_4 + H_2O_2+ $MgCl_2$ (Brookes e Powlson, 1981; Hedley et al., 1982) to obtain the P-residual (P_{RES}).

The Pi content was obtained by acidifying the alkaline extracts ($NaHCO_3$ and NaOH) according to the method proposed by Dick and Tabatabai (1977). To determine the total P (P_T) of these fractions, an aliquot of each extract was subjected to digestion with H_2SO_4 +

$(NH_4)2_S2O_8$ in an autoclave at 103 kPa, 121°C for 2 hours. The Pi concentrations of the acidic extracts (RTA e HCl) including the P_T from the previous digestions were quantified by the method of Murphy and Riley (1962). The Po of the fractions was estimated by the difference between P_T and Pi. The determined P fractions were grouped according to the lability of P in the soil, with labile P = P_{RTA} + Pi_{BIC}; labile Po = Po_{BIC}; moderately labile Pi = $Pi_{Hid-0,1}$ + Pi_{HCl}; moderately labile Po = $Po_{Hid-0,1}$; e non- labile P = $P_{Hid-0,5}$ + P_{RES}.

The alkaline and acid phosphatase enzymes were determined following the method described by Tabatabai and Bremner (1969), with some modifications. For this, 1 g of soil was mixed with 4 mL of MUB buffer (pH 6.5 for acid phosphatase and pH 11 for alkaline phosphatase) and 1 mL of 0.05 M p-nitrophenyl phosphate solution (p-NPP, Sigma-Aldrich) in the corresponding MUB buffer (added to all except the control). After incubation at 37°C for 1 hour, 1 mL of 0.5 M $CaCl_2$ and 4 mL of 0.5 M NaOH were added, and the mixture was vigorously shaken. In this last step, 1 mL of 0.05 M p-NPP solution was added to the control.

The reaction was centrifuged at 10,000 g for 5 minutes to remove soil particles, and the absorbance of the supernatant was determined by spectrophotometry at 420 nm. The standard curve of p-nitrophenol (p-NP, Sigma-Aldrich) was used to extrapolate the concentrations of p-NP from the filtrates using serially diluted solutions (10, 20, 30, 40, and 50 µg mL^{-1} of p-NP). The values were expressed as µg of p-NP g^{-1} of soil $hour^{-1}$.

Phytase activity measurement was performed following the method of Ames (1966). For this, 1 g of soil was mixed with 5 mL of 1 mM sodium phytate (Sigma-Aldrich) in 0.2 M sodium acetate buffer, pH 4.5. In the control samples, 5 mL of 0.2 M sodium acetate buffer without sodium phytate was added. After incubation at 37°C for 1 hour, the reaction was stopped by adding 0.5 mL of 10% trichloroacetic acid and then centrifuged at 10,000 g for 5 minutes to remove soil particles. An aliquot of 300 µL of the supernatant was combined with 1200 µL of a solution containing 0.02 M ammonium molybdate, 0.01 M ammonium metavanadate, and 0.18 M nitric acid, and the absorbance of this mixture was determined by spectrophotometry at 420 nm. Pi concentrations were extrapolated using serial diluted solutions (0; 0.01; 0.05; 0.1 and 0.2 µg mL^{-1}) of KH_2PO_4 (Sigma-Aldrich), and the values were expressed as µg of P released g^{-1} of soil $hour^{-1}$.

The results obtained for the P fractionation and enzymatic activity were then subjected to analysis of variance (ANOVA) with comparison of treatment means by Tukey's Test ($p<0.05$).

2.2.4. Soil bacterial community modulation

2.2.4.1. Initial preparation

The DNA from soil samples (rhizosphere and bulk soil) was extracted using the DNeasy PowerSoil Pro Kit (Qiagen) following the manufacturer's instructions. The concentration and quality of DNA samples were analyzed by quantification in Biodrop (Biochrom) and in 1.2% agarose gel, stained with SYBR Green (Sigma).

Sequencing was performed at Helmholtz Centre for Environmental Research (UFZ). To prepare the samples, the quality of the extracted DNA from soil samples was verified again by 1.5% agarose gel stained with ethidium bromide. Then, its quantification was evaluated by Qubit 3.0 Fluorometer (Invitrogen) using the Broad Range Kit (Invitrogen), following the manufacturer's instructions.

Samples with a concentration lower than 5 ng/µL were concentrated using SpeedVac (Thermo Scientific), followed by a new quantification using Qubit, as described above. Thus, all DNA samples had their concentration normalized to 5 ng/µL.

2.2.4.2. 16S rRNA amplification

PCR was performed to amplify the regions V3 and V4 of the 16S rRNA gene. The primers used were 341F (5'CCTACGGGNGGCWGCAG3') and 785R (5' GACTACHVGGGTATCTAATCC3') (Klindworth et al., 2013) with overhang adapters attached.

The PCR reaction was performed using 2X MyTaq Mix (Bioline); 1 µM forward primer; 1 µM reverse primer; 5 ng/µL environmental DNA, in a final volume of 25 µL. The cycles used were 95°C for 3 minutes; 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds; 72°C for 5 minutes. Amplicons were observed in a 1.5% agarose gel, stained with 0.5 µg/mL ethidium bromide.

2.2.4.3. PCR Clean-Up

For purification of the amplicons, the NGS Clean Up and Size Selection Kit (NucleoMag) was used. Firstly, the magnetic beads from Kit were vortexed for 30 seconds to make sure that the beads were evenly dispersed. Then, 20 µL of magnetic beads were added

to each well of the Amplicon PCR sample. The entire volume was pipetted up and down 10 times and incubated at room temperature without shaking for 5 minutes. Samples were placed on a magnetic stand for 2 minutes.

With the amplicon still on the magnetic stand, the supernatant was removed and discarded. Then, with the samples still on the magnetic stand, the beads were washed with freshly prepared 80% ethanol following the workflow: *i)* 200 µL of freshly prepared 80% ethanol were added to each sample; *ii)* samples were incubated on the magnetic stand for 30 seconds; *iii)* the supernatant was removed and discarded. This step was performed twice, whereas, in the last step, all excess ethanol was removed.

Thus, beads were air-dried for 10 minutes, still on the magnetic stands. After this step, samples were removed from the magnetic stand and 52.5 µL of 10 mM Tris pH 8.5 were added to each one, and the entire volume was mixed up and down 10 times until beads are fully resuspended. Samples were incubated at room temperature for 2 minutes and replaced on the magnetic stand for 2 minutes. Finally, 50 µL of the supernatant from the purified Amplicon PCR was transferred to new and free DNA tubes and stored at -20°C.

2.2.4.4. Index PCR and purification

This step attached dual indices and Illumina sequencing adapters using the Nextera XT Index Kit (Illumina). The reaction was performed using 25 µL of 2x My Taq Mix (Bioline); 10 µL PCR water; 5 µL Index 1; 5 µL Index 2, and 5 µL of purified Amplicon PCR product, in a final volume of 50 µL. The PCR cycles were 95°C for 3 minutes; 8 cycles of:95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds; and 72°C for 5 minutes.

Then PCR products were purified as described previously, with some changes: at first, 56 µL of magnetic beads were added to each sample of the Index PCR. After the double washing with ethanol, 27.5 µL of 10 mM Tris pH 8.5 were added to each sample, and 25 µL of the supernatant from the purified Index PCR was transferred to new and free DNA tubes and stored at -20°C.

2.2.4.5. Library quantification, normalization and pooling

After the second clean-up, samples were quantified again using Qubit 3.0 Fluorometer (Invitrogen). The Broad Range Kit (Invitrogen) was used, following the manufacturer's instructions.

According to the quantification results, libraries were diluted to 4nM using 10mM Tris pH 8.5. The dilution calculation was based on the formula below.

$$\frac{(\text{concentration in ng/µl})}{(660\ \text{g/mol} \times \text{average library size})} \times 10^6 = \text{concentration in nM}$$

With libraries normalized to 4nM, a pre-pool was obtained. Libraries built with the same Index 1 primers were pooled in a single pool. Thus, a new quantification was performed using the High Sensitivity Kit (Invitrogen) on Qubit 3.0 Fluorometer (Invitrogen). A final pool was obtained by adding amounts of each sample in a way that each one had 4nM in the final pool (to make sure that all libraries had enough concentration to sequencing).

2.2.4.6. Library denaturing and Miseq sample loading

In preparation for cluster generation and sequencing, pooled libraries were denatured with NaOH, diluted with the hybridization buffer, and then heat denatured before MiSeq sequencing. The PhiX Control Kit v3 (Illumina) was also used.

For DNA denaturing, 5 µL of 4 nM pooled library and 5 µL of freshly diluted 0.2 N NaOH were combined in a microcentrifuge tube. The mix library-NaOH was briefly vortexed and then, centrifuged at 280× g at 20°C for 1 minute.

To denature the DNA into single strands, this mix was incubated for 5 minutes at room temperature. After, 990 µL of pre-chilled HT1 (hybridization buffer) were added to a tube containing 10 µL of denatured DNA, resulting in a 20pM denatured library. So, the denatured DNA was kept on ice until the final dilution.

For final dilution, the denatured library was diluted with pre-chilled HT1 to reach the concentration of 4pM, and after some inversions to mix, the denatured and diluted DNA was kept on the ice again.

The PhiX control was prepared according to the manufacturer's instructions, to ensure the same loading concentration as the Amplicon library. Finally, in microcentrifuge tubes, 30 µL of denatured and diluted PhiX control, and 570 µL denatured and diluted amplicon library were mixed and the combined library and PhiX control tubes were incubated at 96°C for 2 minutes, using a heat block. After the incubation, the tubes were inverted 2 times to mix, and immediately were placed in the ice-water bath and kept there for 5 minutes. So, immediately, the mixture was loaded onto the MiSeq v3 reagent cartridge.

2.2.4.7. Amplicon sequencing and data analysis

For library sequencing, the Illumina MiSeq (2x300bp) platform was used. The reads were demultiplexed and converted to FASTQ format. Finally, their adapters were removed using Illumina algorithms.

The demultiplexed sequences were analyzed using the software dada2 v.1.24.0 (Callahan et al., 2016) that generated the Amplicon Sequence Variants (ASV). Paired-end reads have their primers removed, were trimmed by length and quality, and merged. Chloroplasts and mitochondria sequences were also removed. The ASVs were taxonomically classified using the database of 16S by Silva v. 138.1 (Quast et al., 2012).

For analysis of diversity, the package phyloseq v. 1.40.0 (McMurdie and Holmes, 2013) and vegan v. 2.6-4 (Oksanen et al., 2007) were used on the software R (Team, R. C., 2013). The alpha-diversity was estimated using Shannon and Simpson indices. Beta-diversity was calculated using Bray-Curtis disimilarity. Permanova was performed to evaluate the variances among the variables.

2.3. Results

2.3.1. Greenhouse and plant growth promotion assay

The results of the consortia among 33.1 and different AMF in sugarcane growth promotion demonstrated that the mean values of plant height (Figure 2) and shoot dry weight (Figure 3) were not statistically different between plants inoculated and non-inoculated with strain 33.1. Among the AMF, the performance of AMF3 (*D. heterogama*) stands out in the evaluated parameters.

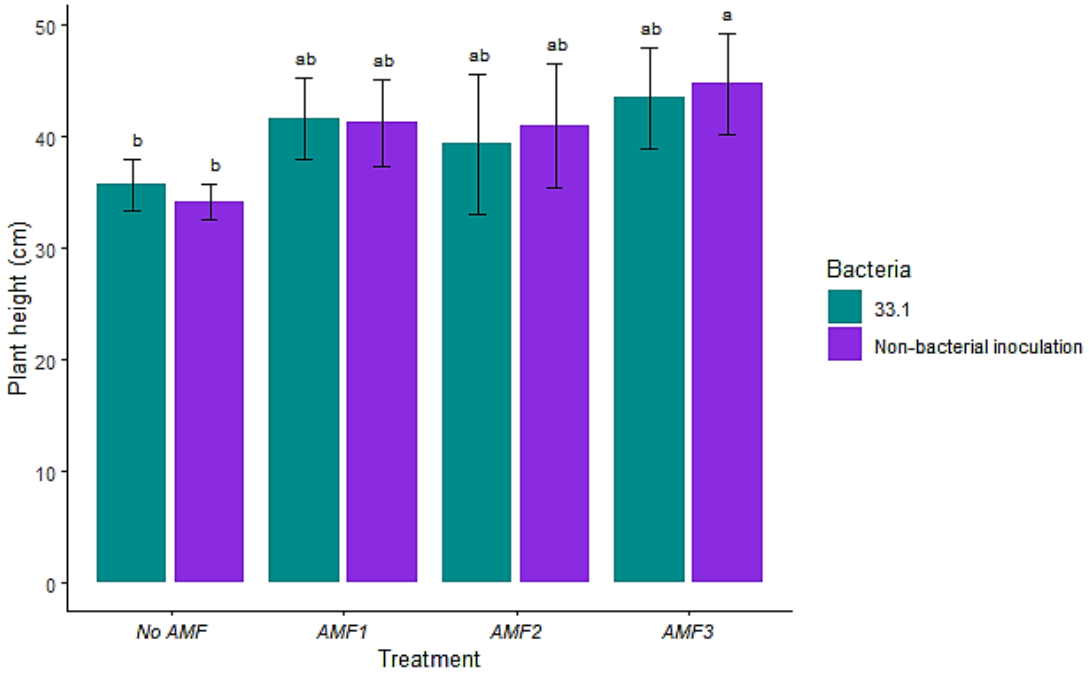

Figure 2. Sugarcane height at 120 dai. The treatments plotted are, from left to right: No AMF, AMF1 (*R. intraradices*), AMF2 (*R. clarus*), and AMF3 (*D. heterogama*). Treatments containing different letters assigned by the Tukey Test indicate that there are statistically significant differences between means ($p < 0.05$).

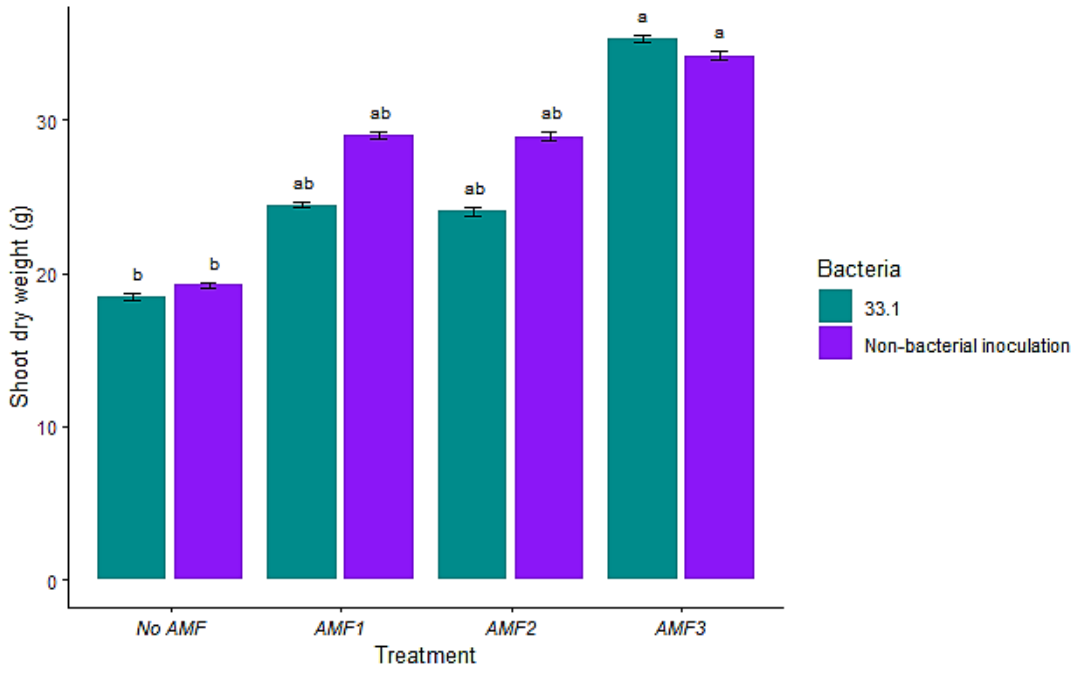

Figure 3. Sugarcane shoot dry weight at 120 dai. The treatments plotted are, from left to right: No AMF, AMF1 (*R. intraradices*), AMF2 (*R. clarus*), and AMF3 (*D. heterogama*). Treatments containing different letters assigned by the Tukey Test indicate that there are statistically significant differences between means ($p < 0.05$).

2.3.2. Dynamics of P availability

In general, Pi fractions increased in treatments containing only AMF, while organic fractions were not significantly affected by them. In this scenario, it is also observed that the application of these mycorrhizae also led to an increase in the non-labile Pi fraction (Table 3). No significant differences were observed in the treatments under inoculation with strain 33.1, or the consortium of 33.1 and AMF2, when compared to the non-inoculated treatment (Table 3).

Table 3. P fractionation in soil under different treatments, 120 dai. The colored lines highlight treatments with different inoculations that showed a significant difference (p<.0.05) in P fractionation compared to the treatment that did not receive any type of inoculation (in light blue). The evaluated treatments were 33.1 (*P. agglomerans* 33.1), AMF1 (*R. intraradices*), AMF2 (*R. clarus*), and AMF3 (*D. heterogama*) and their combination as a consortium.

Treatments	Labile Pi	Labile Po	Moderately labile Pi	Moderately labile Po	Non-labile Pi	Non-labile Po	Total Pi	Total Po	Total P
No inoculation	5.1 d	4.17 a	17.8 d	91.3 ab	72.4 c	24.4 a	95.2 d	119.8 b	215.1 b
33.1	8.2 bcd	9.7 a	23.1 cd	94.2 ab	77.5 bc	24.1 a	108.7 cd	128 ab	236.7 b
AMF1	28.1 ab	10.8 a	71.4 a	106.5 ab	95.6 ab	27.4 a	195.1 ab	144.8 ab	339.9 a
AMF1 and 33.1	50.2 a	8.7 a	73.3 a	112.5 a	102.7 a	40.1 a	226.1 a	161.2 a	387.3 a
AMF2	31.2 ab	10.1 a	48.2 ab	107.2 ab	98.6 a	42.2 a	178 ab	159.5 ab	337.4 a
AMF2 and 33.1	7.11 cd	9.9 a	20.7 cd	106.5 ab	86.9 abc	28.7 a	114.7 cd	145.2 ab	259.8 b
AMF3	24.2 abc	8.7 a	30.1 bc	82 b	92.3 ab	34.5 a	146.6 bc	125.2 ab	271.8 b
AMF3 and 33.1	9.3 bcd	9.7 a	22.4 cd	104.5 ab	84.4 abc	30.8 a	116.1 cd	145 ab	261.1 b

On the other hand, strain 33.1 seems to have an additive effect when associated with AMF1 in the solubilization of Pi fractions, especially in labile-Pi: the increase in these fractions, already significant in the presence of AMF1, was reinforced in consortium with 33.1 (Table 3).

Regarding the enzyme production related to the mineralization of P, it is worth noting the additive effect demonstrated by 33.1 in some treatments (as shown in Figure 4). When assessed individually, neither 33.1 nor the AMF exhibited significantly higher phytase and phosphatase activities in comparison to the control treatment. However, some combinations of microbes increase the enzymes production, such as phytase activity in the consortium of 33.1 and AMF1, acid phosphatase in the consortium of 33.1 and AMF 2 and 3, and alkaline phosphatase when 33.1 and AMF1 worked together.

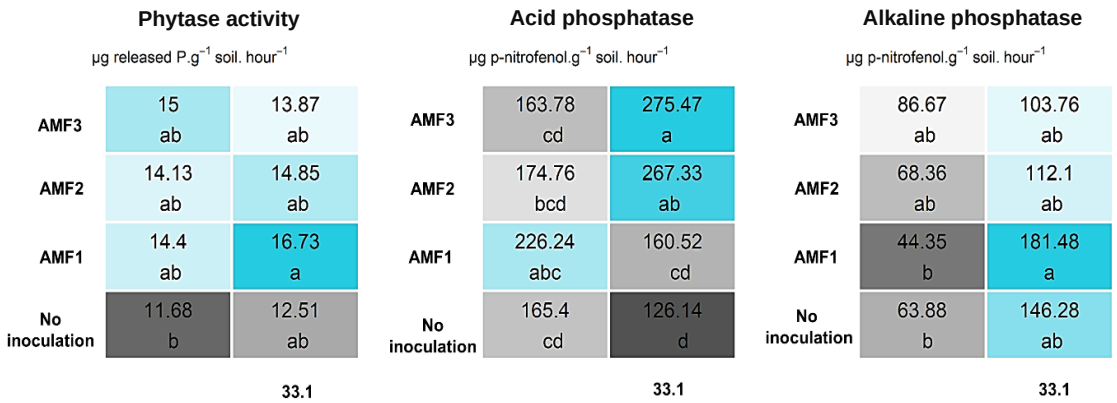

Figure 4. Heatmap for enzyme activity of (from left to right): phytase, acid phosphatase and alkaline phosphatase, in the evaluated treatments. The quadrants on the left indicate inoculations of the respective AMF, while the quadrants on the right indicate their values when in consortia with 33.1. The evaluated treatments were 33.1 (*P. agglomerans* 33.1), AMF1 (*R. intraradices*), AMF2 (*R. clarus*), and AMF3 (*D. heterogama*) and their combination as a consortium. Treatments containing different letters assigned by the Tukey Test indicate that there are statistically significant differences between them ($p < 0.05$).

2.3.3. Soil bacterial community modulation

The microbial community from bulk soil and rhizosphere from plants inoculated with the PSMs were accessed. Simpson and Shannon indices for alpha-diversity (Figure 5) did not show clear differences in diversity within the analyzed treatments. On the other hand, samples from the rhizosphere tend to show lower diversity compared to those obtained from the bulk soil.

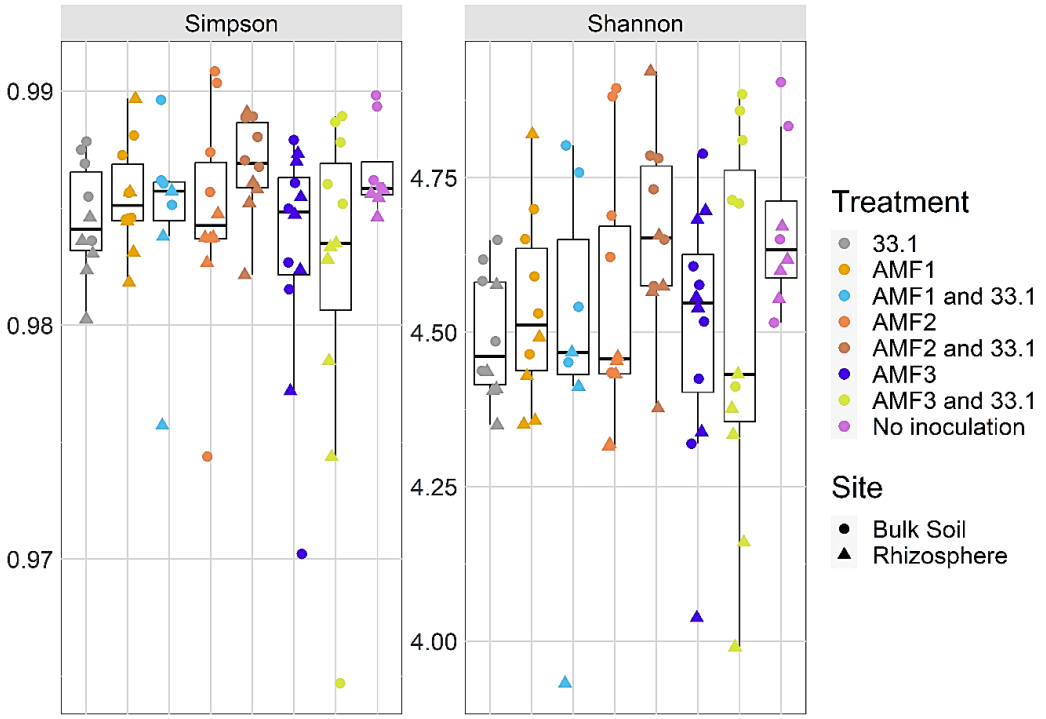

Figure 5. Simpson (at left) and Shannon (at right) indices for richness and evenness. Both plots measure alpha-diversity in samples from different treatments (33.1 -*P. agglomerans* 33.1, AMF1 -*R. intraradices*, AMF2 -*R. clarus*, and AMF3 -*D. heterogama*, and their combination as a consortium), and from different sites of collection.

Considering beta-diversity, it was possible to observe the separation between the different treatments along the NMDS, especially those that received AMF3 (Figure 6). Furthermore, the differentiation between bulk soil and rhizosphere was evident, demonstrating that the collection site was a factor that contributed to the differences in diversity between them.

The results of the Permutation Based Analysis of Variance (PERMANOVA) demonstrated a significant difference in diversity between the collection sites (p=0.001) and among the evaluated treatments (p=0.002). However, the presence or absence of strain 33.1 did not prove to be a significant factor in this separation (p=0.451).

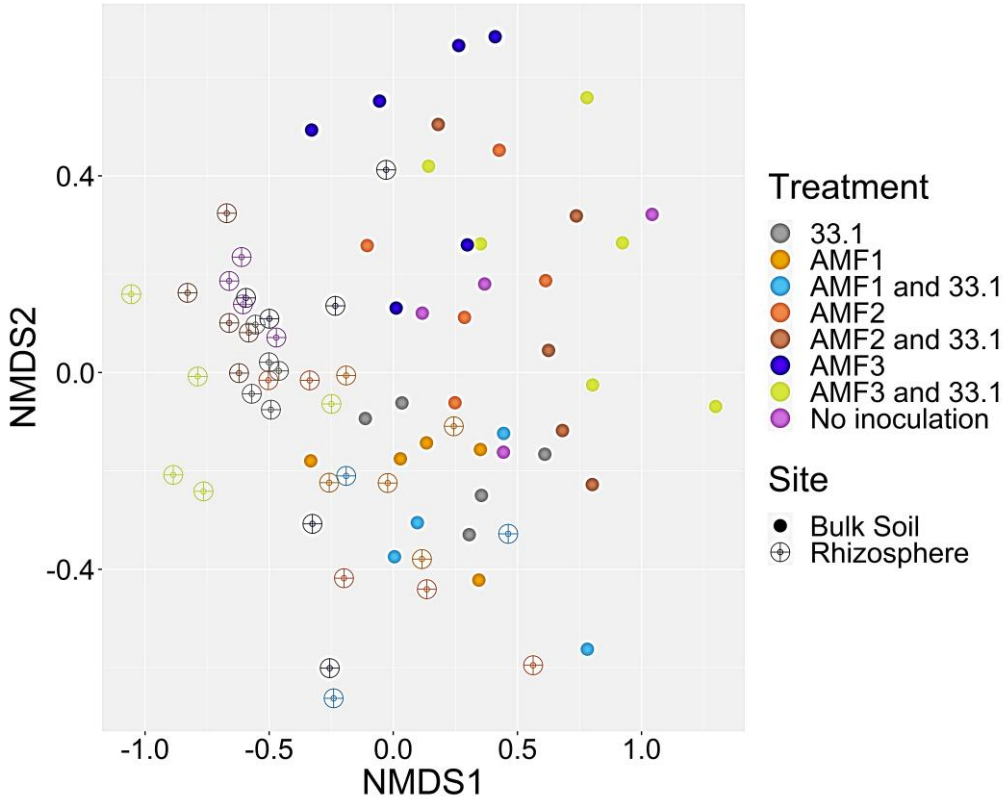

Figure 6. Non-metric Multi-dimensional Scaling (NMDS) plot for beta-diversity among the different treatments (33.1 -*P. agglomerans* 33.1, AMF1 -*R. intraradices*, AMF2 -*R. clarus*, and AMF3 -*D. heterogama*, and their combination as a consortium), and sites of collection. The distance between points in the NMDS indicates the similarity (closest points) or distinction (most separated points) among the corresponding samples in terms of beta-diversity.

As expected, the composition of the soil bacterial community was also altered depending on the sampled location (Figures 7 and 8). The reduction in the abundance of members of Cyanobacteriota in the rhizosphere, when compared with the bulk soil, is noteworthy.

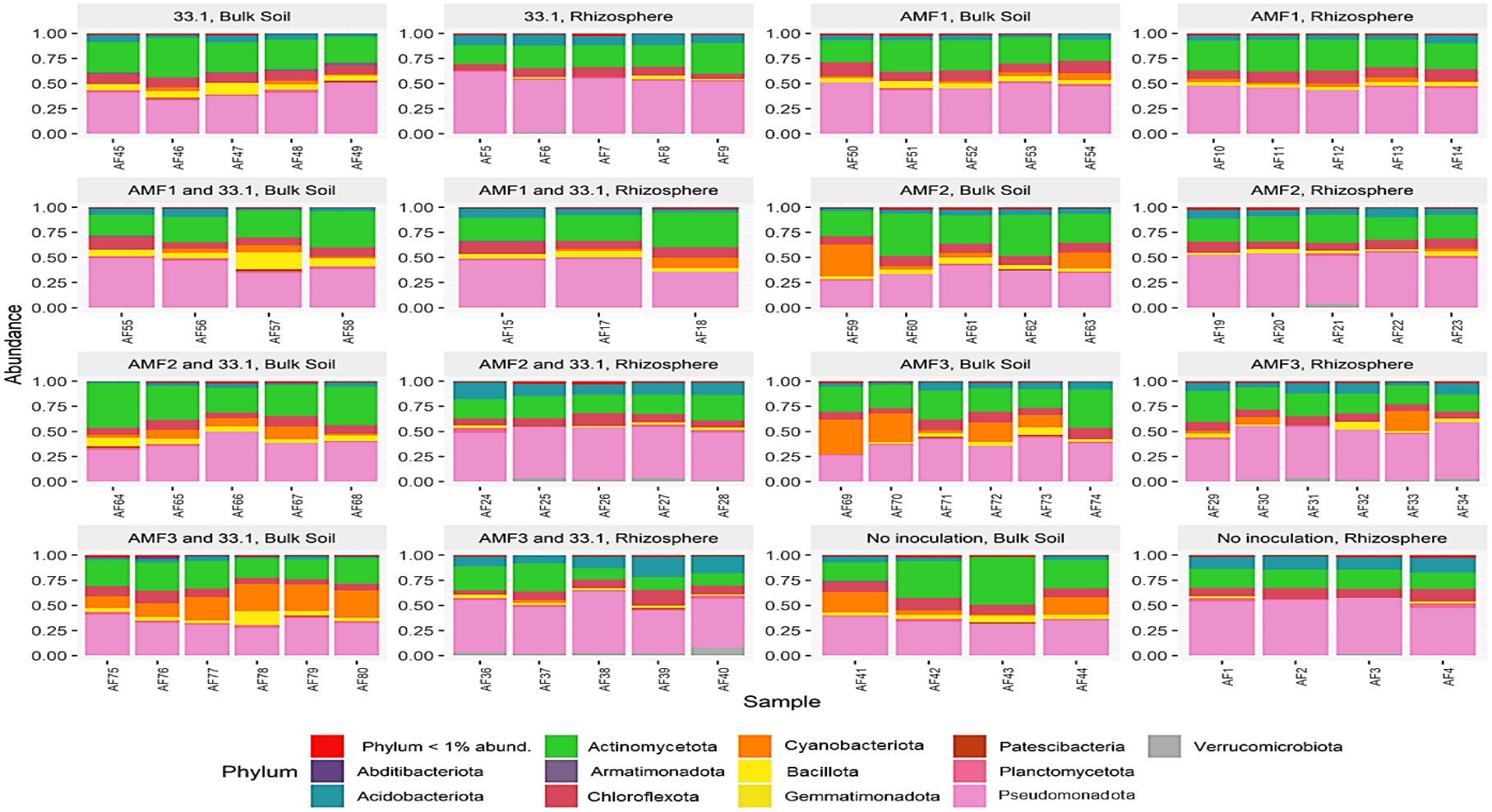

Figure 7. Top thirteen most abundant phylum (based on their relative abundance) in each treatment and site of collection. The bars contained in the representations of each collection site per treatment demonstrate the taxonomic composition of each replicate sampled.

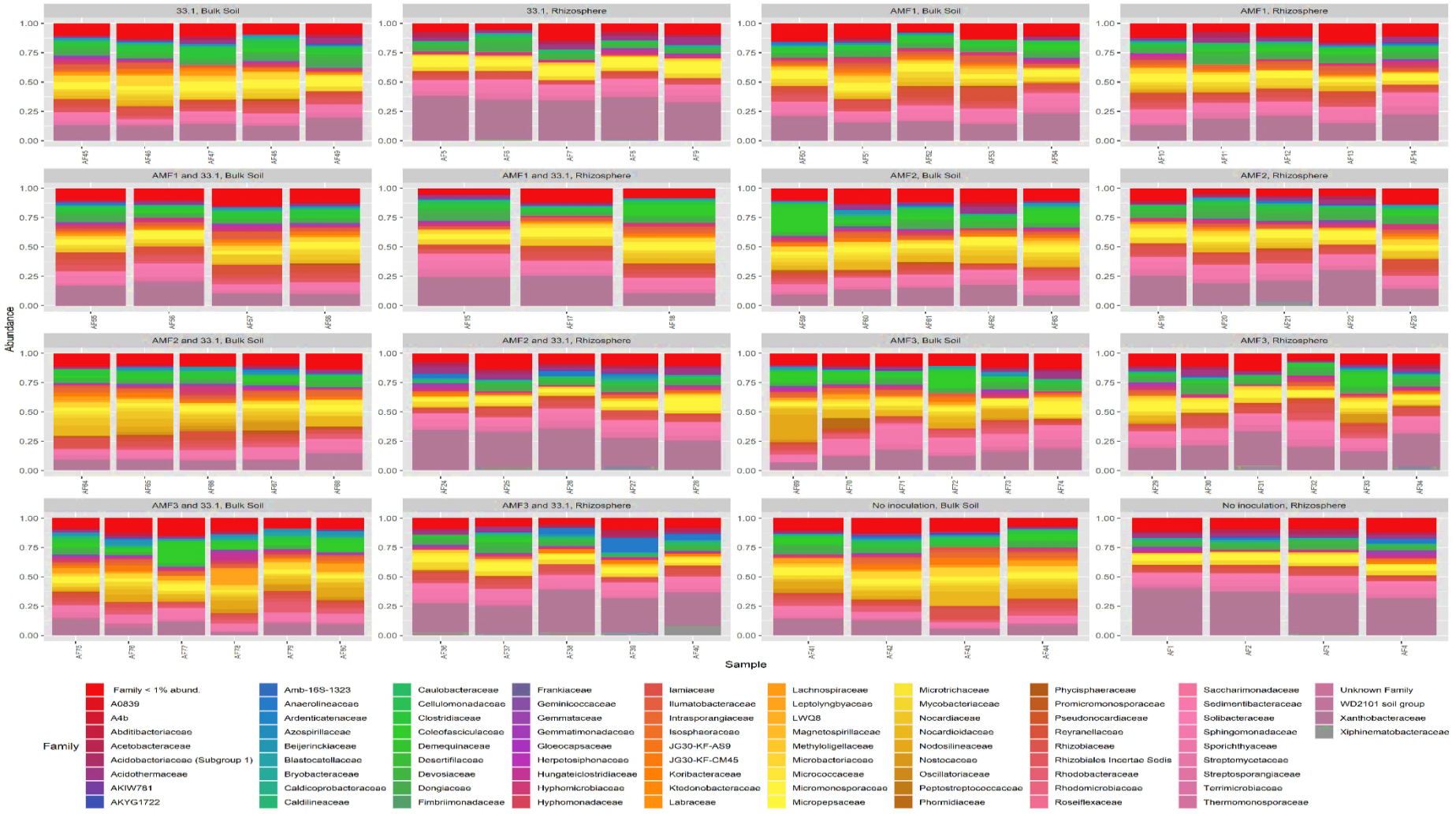

Figure 8. Most abundant families (based on their relative abundance) in each treatment and site of collection. The bars contained in the representations of each collection site per treatment demonstrate the taxonomic composition of each replicate sampled.

Differentially abundant families ($p < 0.05$) in the rhizosphere, compared to the bulk soil, were also evaluated (Figure 9). An increase in various members of Pseudomonadota, Actinomycetota, and Chlorofexota, especially families related to plant growth promotion, such as *Rhizobiaceae*, *Devosiaceae*, and *Streptomycetaceae*, can be observed. While some families were favored in the rhizosphere, others were reduced in this niche.

As mentioned, it was observed a decrease in Cyanobacteriota members in the rhizosphere, especially in the *Leptolyngbyaceae* and *Nodosilineacea* families. Interestingly, reductions in families with a speculated role in pathogenesis, such as *Nocardiaceae* and *Microbacteriaceae*, were also observed.

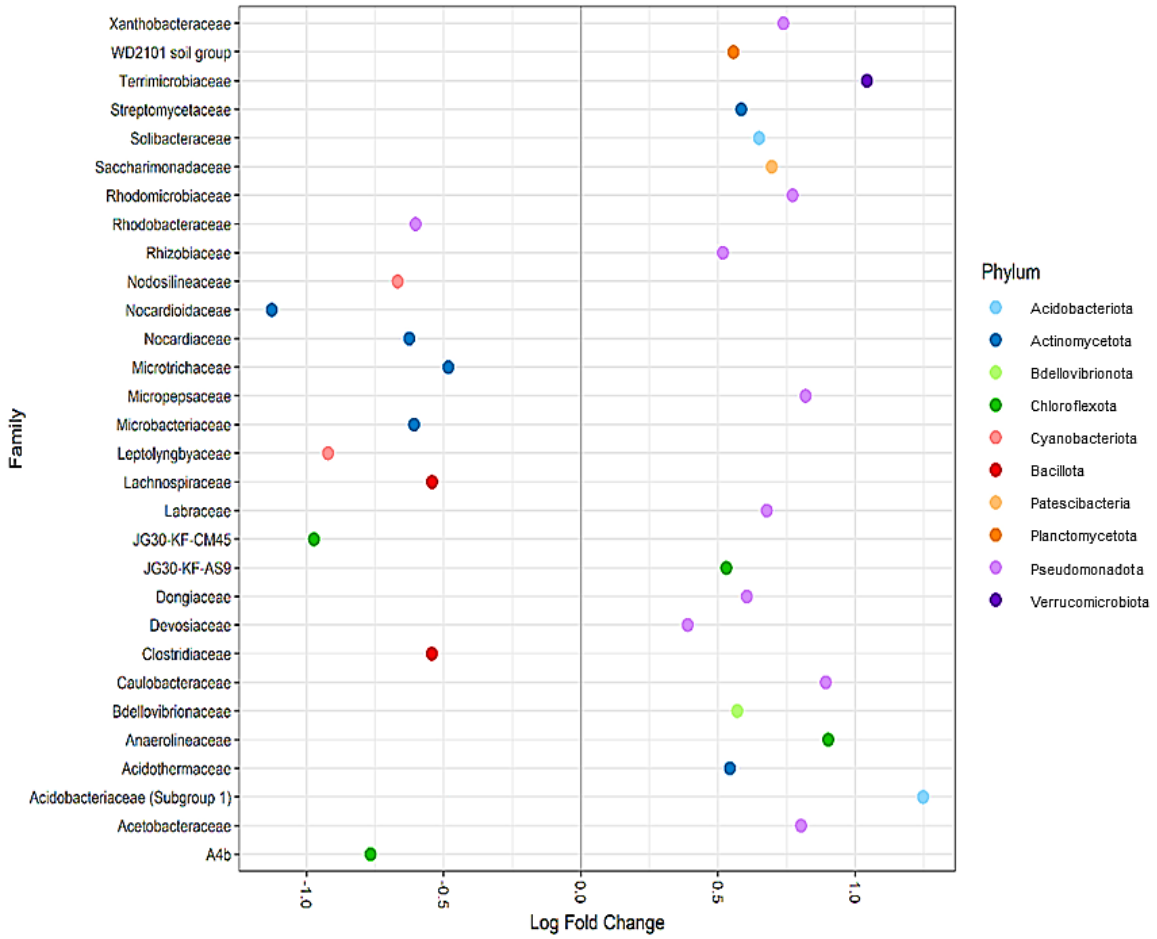

Figure 9. Log2 fold change in the composition of taxa present in the soil bacterial community analyzed in the rhizosphere, when compared with bulk soil. The values on the left demonstrate a reduction in abundance of the indicated taxa, whereas values on the right indicate an increase in abundance of the taxa.

To evaluate the effects of inoculation with 33.1 in the soil bacterial community modulation, differentially abundant soil bacterial genera ($p < 0.05$) were also assessed by Log2 Fold Change (Figure 10). Among the genera differentially reduced in the presence of

33.1, the genus *Afipia*, composed of opportunistic pathogens (even for humans), can be mentioned. Interestingly, the complex of nitrogen-fixing bacteria (N), *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*, also had its abundance reduced under the inoculation of 33.1. On the other hand, the genus with potential for bioremediation, *Rhodococcus*, was favored with the application of 33.1.

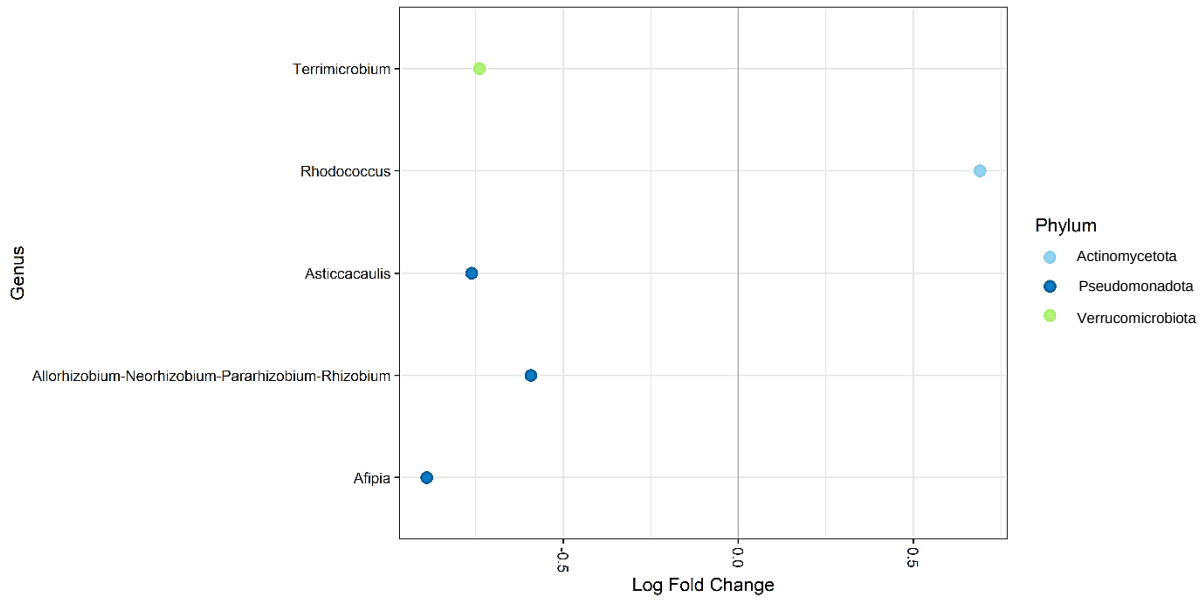

Figure 10. Log2 fold change in the composition of genera/phyla of the soil bacterial community as a function of 33.1 inoculation. The values on the left demonstrate a reduction in the abundance of the indicated families. Values on the right indicate an increase in the abundance of the families mentioned.

2.4. Discussion

Microorganisms can contribute to plant growth and crop production in several mechanisms: nutrient uptake, modulation of phytohormone levels, induction of systemic resistance, protection against pathogen attacks through antimicrobial production, among other mechanisms (Glick, 2012). *P. agglomerans* 33.1 has been described as a PGPB (Quecine et al., 2012). As mentioned earlier, arbuscular mycorrhiza can contribute to P solubilization processes. In this sense, our study aimed to evaluate the interaction of 33.1 with several AMF at multiple levels: in promoting plant growth, supplying available P, and modulating the soil bacterial community.

Regarding growth promotion, although the results for sugarcane shoot dry mass and plant height in treatments with co-inoculation of 33.1 and some AMF did not reach

significantly higher values than those obtained in single inoculations of each microorganism, in general, the application of 33.1 or AMF individually was able to enhance plant development compared to those without any inoculation.

Interactions between AMF and bacteria can be governed by a series of factors, including the species and strains interacting, characteristics of AMF mycelium exudates, soil structure, competition for nutrients, and plant root architecture (Bianciotto et al., 1996; Johansson, Paul, and Finlay, 2004; Sangwan and Prasanna, 2021). In this context, a specific group of bacteria stands out, known as Mycorrhiza Helper Bacteria or Mycorrhization Helper Bacteria (abbreviated as MHB in both cases), which, through several distinct mechanisms, promote the establishment of AMF symbiosis (Mycorrhization Helper Bacteria) or assist in already established mycorrhizal functions (Mycorrhiza Helper Bacteria) (Sangwan and Prasanna, 2021).

Considering the overall scenario that the results obtained in the P fractionation and enzymatic activity of AMF1 were particularly potentiated when in consortium with 33.1, a synergy between them is suggested. In this context, 33.1 can play the role of MHB when associated with AMF1. A possible reason for this putative additive effect is the specificity of MHB to AMF. According to Garbaye and Dupponois (1992), the bacteria that enhance mycorrhiza formation by some fungi can inhibit the establishment of symbiosis by others. Thus, Hameeda et al. (2007), and Sangwan and Prasanna (2021) suggest that the specificity of this whole interaction is governed by the fungus, not the plant.

Regarding the diversity and modulation of the soil bacterial community under influence of 33.1 and different AMF, various alterations were observed in the rhizosphere microbiome composition compared to the bulk soil. The alterations occur especially due to the physicochemical contrasts that lead to niche differentiation, well-documented in the literature (Trivedi et al., 2020; Ling, Wang, Kuzyakov, 2022). Additionally, the diversity observed in the rhizosphere was lower than that in the bulk soil. This condition is common and expected, as increased substrate availability leads to a decrease in bacterial diversity (Ling, Wang, Kuzyakov, 2022).

In the rhizosphere, there was an increase in the abundance of groups with well-described beneficial plant/microorganism interactions, such as *Rhizobiaceae*, *Devosiaceae*, and *Streptomycetaceae*. Interestingly, these microorganisms have been studied due to their positive interactions with AMF, acting as MHB and enhancing plant nutrition and

productivity (Xavier and Germida, 2003; Dobo, 2022; Karimi and Noori, 2022; Zhang, 2023). Additionally, *Acetobactereaceae*, also more abundant in the rhizosphere, has been increasingly studied for its potential to promote plant growth through nitrogen fixation and exopolysaccharide (EPS) production (Reis and Teixeira, 2015; Wünsche and Schmid, 2023). It is interesting to speculate that the families more abundant in the rhizosphere might be a result of the direct selection of the plant to those more favorable microorganisms.

Considering the presence of 33.1, inoculated alone and in the presence of different AMF, the genus *Rhodococcus* (family *Nocardiaceae*) was the only one whose abundance in the soil bacterial community increased due to the inoculation this strain. Interestingly, *Rhodococcus* isolates have also been described to act as MHB (Poole et al., 2001; Frey-Klett, Garbaye, and Tarkka, 2007; Deveau and Labbé, 2016; Sangwan and Prasanna, 2021).

An important factor that may be crucial for establishing the beneficial MHB-AMF interaction is the bacterial Protein Secretion Systems (PSS). This is because some MHB can use their PSS as tools for manipulation and interaction with the environment, injecting effectors directly into the partner AMF (Deveau and Labbé, 2016). Some groups have already demonstrated the importance of the T3SS of different MHB in mycorrhization (Warmink and van Elsas, 2008), and mutations in this PSS can lead to the loss of its "helper" effect (Viollet et al., 2011; Cusano et al., 2010).

Our results clearly demonstrated the potential of different AMF in promoting sugarcane growth and providing P to plants through Pi solubilization processes and enzyme production. Our results also demonstrated that 33.1 can enhance the enzymatic activity of the consortium, as well as the P labiality, in the interaction with specific AMF. This is highlighted by the synergistic interaction between 33.1 and AMF1, which can beneficiate agriculture by facilitating P availability and promoting plant development.

Finally, we demonstrated the potential role of strain 33.1 as an MHB, acting in synergy with AMF1 (*R. intraradices*), with a particular performance in making P available to sugarcane. However, a deeper understanding of the molecular mechanisms involved in this synergy, how it occurs, and its effects on the biology of 33.1 and AMF1 is still necessary. Thus, this study provided important insights into the interaction between these different PSM, serving as a starting point for more comprehensive future research aimed at creating a microbial consortium that represents gains for agricultural production.

REFERENCES

Ames, B. N. (1966). [10] Assay of inorganic phosphate, total phosphate and phosphatases. In *Methods in enzymology* (Vol. 8, pp. 115-118). Academic Press.

B. van RAIJ. (1996). *Recomendações de adubação e calagem para o Estado de São Paulo* (Vol. 100, p. 285p). Campinas: IAC.

Bianciotto, V., Minerdi, D., Perotto, S., & Bonfante, P. (1996). Cellular interactions between arbuscular mycorrhizal fungi and rhizosphere bacteria. *Protoplasma*, *193*, 123-131.

Breitwieser, F. P., Lu, J., & Salzberg, S. L. (2019). A review of methods and databases for metagenomic classification and assembly. *Briefings in bioinformatics*, *20*(4), 1125-1136.

Brookes, P. C., & Powlson, D. S. (1981). Preventing phosphorus losses during perchloric acid digestion of sodium bicarbonate soil extracts. *Journal of the Science of Food and Agriculture*, *32*(7), 671-674.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods*, *13*(7), 581-583.

Condron, L. M., Goh, K. M., & Newman, R. H. (1985). Nature and distribution of soil phosphorus as revealed by a sequential extraction method followed by 31P nuclear magnetic resonance analysis. *Journal of Soil Science*, *36*(2), 199-207.

Cusano, A. M., Burlinson, P., Deveau, A., Vion, P., Uroz, S., Preston, G. M., & Frey-Klett, P. (2011). Pseudomonas fluorescens BBc6R8 type III secretion mutants no longer promote ectomycorrhizal symbiosis. *Environmental microbiology reports*, *3*(2), 203-210.

Deveau, A., & Labbé, J. (2016). Mycorrhiza helper bacteria. *Molecular mycorrhizal symbiosis*, 437-450.

Dick, W. A., & Tabatabai, M. A. (1977). An alkaline oxidation method for determination of total phosphorus in soils. *Soil Science Society of America Journal*, *41*(3), 511-514.

Ferreira, D. A. (2016). *Interações entre fungos micorrízicos arbusculares e a microbiota de solos* (Doctoral dissertation, Universidade de São Paulo).

Frey-Klett, P., Garbaye, J. A., & Tarkka, M. (2007). The mycorrhiza helper bacteria revisited. *New phytologist*, *176*(1), 22-36.

Gayathri, G., Gomathi, S., Ambikapathy, V., & Panneerselvam, A. (2023). Improvement of soil fertility through plant microbial interaction. In *Plant-Microbe Interaction-Recent Advances in Molecular and Biochemical Approaches* (pp. 281-301). Academic Press.

Garbaye, J., & Duponnois, R. (1993). Specificity and function of mycorrhization helper bacteria (MHB) associated with the Pseudotsuga menziesii-Laccaria laccata symbiosis. *Symbiosis*.

Glick, B. R. (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica*, *2012*.

Hameeda, B., Harini, G., Rupela, O. P., & Reddy, G. (2007). Effect of composts or vermicomposts on sorghum growth and mycorrhizal colonization. *African Journal of Biotechnology*, *6*(1), 9.

Hedley, M. J., Stewart, J. W. B., & Chauhan, B. (1982). Changes in inorganic and organic soil phosphorus fractions induced by cultivation practices and by laboratory incubations. *Soil Science Society of America Journal*, *46*(5), 970-976.

Johansson, J. F., Paul, L. R., & Finlay, R. D. (2004). Microbial interactions in the mycorrhizosphere and their significance for sustainable agriculture. *FEMS microbiology ecology*, *48*(1), 1-13.

Kalayu, G. (2019). Phosphate solubilizing microorganisms: promising approach as biofertilizers. *International Journal of Agronomy*, *2019*, 1-7.

Khan, M. S., Zaidi, A., & Ahmad, E. (2014). Mechanism of phosphate solubilization and physiological functions of phosphate-solubilizing microorganisms. *Phosphate solubilizing microorganisms: principles and application of microphos technology*, 31-62.

Ling, N., Wang, T., & Kuzyakov, Y. (2022). Rhizosphere bacteriome structure and functions. *Nature communications*, *13*(1), 836.

Manivel, T., Sandhiya, T., Deepika, S., Selvakumar, S. V., Karnan, T. M., Adeyemi, D. E., & Thanapaul, R. J. R. S. (2023). Chemical communication between plant roots and microbes within the rhizosphere. *Plant-Microbe Interaction-Recent Advances in Molecular and Biochemical Approaches*, 141-164.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4), e61217.

Miransari, M. (2010). Contribution of arbuscular mycorrhizal symbiosis to plant growth under different types of soil stress. *Plant Biology*, *12*(4), 563-569.

Murphy, J. A. M. E. S., & Riley, J. P. (1962). A modified single solution method for the determination of phosphate in natural waters. *Analytica chimica acta*, *27*, 31-36.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. A. S. S. (2007). The vegan package. *Community ecology package*, *10*(631-637), 719.

Poole, E. J., Bending, G. D., Whipps, J. M., & Read, D. J. (2001). Bacteria associated with Pinus sylvestris–Lactarius rufus ectomycorrhizas and their effects on mycorrhiza formation in vitro. *New Phytologist*, *151*(3), 743-751.

Procópio, R. E. L., Araújo, W. L., Maccheroni Jr, W., Azevedo, J. L. (2009). Characterization of an endophytic bacterial community associated with Eucalyptus spp. Genetics and Molecular Research, 8(4), 1408-1422.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research*, *41*(D1), D590-D596.

Quecine, M. C., Araújo, W. L., Rossetto, P. B., Ferreira, A., Tsui, S., Lacava, P. T., ... & Pizzirani-Kleiner, A. A. (2012). Sugarcane growth promotion by the endophytic bacterium Pantoea agglomerans 33.1. *Applied and Environmental Microbiology*, *78*(21), 7511-7518.

R Core Team, R. (2013). R: A language and environment for statistical computing.

Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and biophysical research communications*, *469*(4), 967-977.

Reis, V. M., & Teixeira, K. R. D. S. (2015). Nitrogen fixing bacteria in the family Acetobacteraceae and their role in agriculture. *Journal of basic microbiology*, *55*(8), 931-949.

Richardson, A. E., Barea, J. M., McNeill, A. M., & Prigent-Combaret, C. (2009). Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms.

Sangwan, S., & Prasanna, R. (2022). Mycorrhizae helper bacteria: unlocking their potential as bioenhancers of plant–arbuscular mycorrhizal fungal associations. *Microbial ecology*, *84*(1), 1-10.

Smith, S. E., Smith, F. A., & Jakobsen, I. (2003). Mycorrhizal fungi can dominate phosphate supply to plants irrespective of growth responses. *Plant physiology*, *133*(1), 16-20.

Starke, R., Pylro, V. S., & Morais, D. K. (2021). 16S rRNA gene copy number normalization does not provide more reliable conclusions in metataxonomic surveys. *Microbial ecology*, *81*, 535-539.

Tabatabai, M. A., & Bremner, J. M. (1969). Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. *Soil biology and biochemistry*, *1*(4), 301-307.

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: from community assembly to plant health. *Nature reviews microbiology*, *18*(11), 607-621.

Vartoukian, S. R., Palmer, R. M., & Wade, W. G. (2010). Strategies for culture of ‘unculturable’bacteria. *FEMS microbiology letters*, *309*(1), 1-7.

Viollet, A., Corberand, T., Mougel, C., Robin, A., Lemanceau, P., & Mazurier, S. (2011). Fluorescent pseudomonads harboring type III secretion genes are enriched in the mycorrhizosphere of Medicago truncatula. *FEMS microbiology ecology*, *75*(3), 457-467.

Wang, W., Shi, J., Xie, Q., Jiang, Y., Yu, N., & Wang, E. (2017). Nutrient exchange and regulation in arbuscular mycorrhizal symbiosis. *Molecular plant*, *10*(9), 1147-1158.

Warmink, J. A., & Van Elsas, J. D. (2008). Selection of bacterial populations in the mycosphere of Laccaria proxima: is type III secretion involved?. *The ISME journal*, *2*(8), 887-900.

Wünsche, J., & Schmid, J. (2023). Acetobacteraceae as exopolysaccharide producers: Current state of knowledge and further perspectives. *Frontiers in Bioengineering and Biotechnology*, *11*, 1166618.

3. REVEALING THE GENETIC BASIS OF PLANT GROWTH PROMOTION BY *Pantoea agglomerans* 33.1 FOCUSING ON ITS PROTEIN SECRETION SYSTEMS

Aline Aparecida Oliveira Ferreira[1], Ariane Vitória Silvestre Hannes[1], Maurício Santos Rosa[1], Maria Letícia Bonatelli[2], Maria Carolina Quecine[1*]

[1]Genetics of Microorganisms Laboratory "Prof. João Lúcio de Azevedo", Department of Genetics, Luiz de Queiroz College of Agriculture, Piracicaba, Brazil

[2]Department of Environmental Microbiology, Helmholtz Centre for Environmental Research GmbH – UFZ, Leipzig, Germany

* Correspondence:

Maria Carolina Quecine

mquecine@usp.br

Abstract

Protein Secretion Systems (PSS) are essential molecular machineries to bacterial survival and biotic and abiotic interactions. The PSS have been reported as an important system in beneficial plant/bacteria interactions, delivering beneficial substrates to the host. *Pantoea agglomerans* 33.1 is a phosphate solubilizing bacterium, with the capacity to promote plant growth described for several crops. Considering the agricultural potential of 33.1, a comprehensive understanding of its mechanisms to promote plant growth is vital. Thus, this research aimed to investigate the role of PSS in plant growth promotion by 33.1 through gene knockout via CRISPR-Cas9. To achieve this knowledge, the whole-genome of 33.1 was sequenced and assembled, genes related to the main growth-promoting pathways were annotated, and those related to PSS were mining and manually annotated, and compared to those found in other *P. agglomerans* strains. Based on this annotation, two genes of different PSS were selected for knockout through CRISPR-Cas9, employing the two-plasmid strategy. We obtained a total genome of 4,844,324 base-pairs (bp) divided into the chromosome and four plasmids, that harbor genes of processes related to plant growth promotion, such as P metabolism and indole-3-acetic-acid (IAA) production. The classical secretion systems 1, 5a, 5b, and 6, the secretion pathways Sec and Tat, and the appendages and accessories Bam Complex and T4P were found in 33.1. Our PSS prediction is in line with what was observed in the comparative genomics results, where systems 1 and 5 were shown to be the most conserved among the different strains of *P. agglomerans*. Concerning the gene editing, we obtained the 33.1.:pCas strain and the *bepC* and *prn* genes were properly cloned into the pTarget series. Despite the two-plasmid system required for CRISPR-Cas9 have been correctly assembled, the system failed to knock out the Outer Membrane Protein (OMP) and the Autotransporter (AT) of the 33.1, from T1SS and T5aSS, respectively. Regardless of the unsuccessful efforts, our research provides important insights about the 33.1 biology and topics to be improved for knockout assays that will be continued by our group, until the mutants are obtained.

Keywords: Protein Secretion Systems, *Pantoea agglomerans* 33.1, Phosphate-solubilizing Bacterium, Knockout, CRISPR-Cas9.

3.1. Introduction

Bacteria, beneficial or not, are known to employ several mechanisms to interact with their environment and hosts. Among these interactions, mechanisms involved in the plant growth promotion have been extensively studied.

Protein Secretion Systems (PSS) is an important tool used in bacterial interactions and widely studied due to its role in pathogenesis (Alfano and Collmer, 2004; Shyntum et al., 2015; Bernal, Llmas, and Filloux, 2018). PSS also play a vital role in modulating biotic interactions of bacteria, whether mutualistic or pathogenic, during the interaction with diverse hosts (Tseng, Tyler, and Setubal, 2009).

Despite the relevance of PSS on bacterial fitness, studies about their effects in beneficial interactions are scarce. Previously associated with pathogenesis, PSS can perform important functions in beneficial bacteria-plant interactions, especially in agricultural contexts. For instance, Lucero et al. (2022) conducted research showing a reduction in endophytic and epiphytic colonization ability in peanuts by defective mutants for the T6SS of the PSM *Enterobacter* sp. J49.

The CRISPR-Cas9 system (Clustered Regularly Interspaced Short Palindromic Repeats), is a natural defense mechanism against phage infection and plasmid transfer found in many bacteria and most archaea. It has been adapted as an RNA-guided DNA targeting tool, to achieve gene editing and other applications, such as transcriptional perturbation, epigenetic modulation, and genome imaging (Makarova et al., 2011; Jiang and Doudna, 2017). This technique aiming genome editing involves three molecules: a nuclease (usually the Cas9 wild type of *Streptococcus pyogenes*), a guide RNA- gRNA (usually improved as a single-guide RNA- sgRNA), and the target (usually the DNA) (Vieira et al., 2016).

According to Barrangou and Doudna (2016), the Cas9 endonuclease initially identifies specific sequence complementary of gRNA and its protospacer-adjacent motif (PAM), which flanks the target DNA site. If the first 12 bp (“seed sequence”) of the gRNA match the target DNA strand, RNA strand invasion accompanies local DNA unwinding to form an R-loop. This process results in a precise cleavage of each DNA strand by the Cas9’s domains RuvC and HNH, generating a blunt double-strand DNA break at three base pairs upstream of the 3' edge of the PAM sequence (Garneau et al., 2010; Sternberg, Haurwitz, and Doudna, 2012).

After the cleavage process performed by Cas9, the resulting blunt ends must be repaired by the cell, and, in this case, two repair mechanisms can be utilized: the Non-Homologous End-Joining (NHEJ) or the Homology Directed Repair (HDR) (Vieira et al., 2016). In the absence of a DNA template, the NHEJ can produce specific mutations such as insertions or deletions (indels) at the cleavage site (Barrangou and Doudna, 2016; Cui et al., 2019). On the other hand, when a DNA template (native or engineered) is provided, HDR replaces the targeted allele with an alternative sequence through recombination (Barrangou and Doudna, 2016; Cui et al., 2019). HDR-based approaches have demonstrated the ability to enhance the efficiency and precision of gene editing, especially in prokaryotes (Jiang and Doudna, 2017; Cui et al., 2019).

The gene editing achieved by CRISPR-Cas9 presents broad applicability, contributing from basic research (e.g., gene function identification) to applied studies, such as those conducted in agriculture, aiding in the generation of genetic variability, combating pests, and obtaining disease-resistant plants (Pereira, 2016). For this reason, CRISPR-Cas9 represents a milestone in modern science, with new CRISPR-based techniques emerging and existing ones being improved.

P. agglomerans 33.1 is a known PGPB that has been studied by our research group for over 10 years, aiming understand its processes for promoting plant growth and mechanisms to interact with the environment. Thus, our research aimed to increase the understanding of the role of PSS in beneficial interactions between bacteria and plants, with an emphasis on plant growth-promotion by *P. agglomerans* 33.1, through the knockout of PSS genes using the CRISPR-Cas9 system.

3.2. Material and Methods

3.2.1. Bacterial DNA extraction

For the extraction of bacterial DNA, strain 33.1 was cultured in LB medium at 28°C under agitation (150 rpm). Following the Sambrook and Russell (2001) Phenol-Chloroform protocol, genomic DNA was extracted, and its integrity was verified using 1% agarose gel stained with SYBR Green (Sigma), as well as pulsed-field gel electrophoresis, carried out in 2% agarose gel and stained with SYBR Green (Sigma). Subsequently, the extracted DNA was quantified using Biodrop (Biochrom).

3.2.2. Sequencing and assembly of 33.1 genome

To obtain a complete and high-quality sequenced genome, a hybrid approach of genome sequencing and assembly was employed, integrating data obtained from second and third-generation sequencing platforms. Thus, genome sequencing was performed using Pacific Biosciences (PacBio) - CLR and Illumina NextSeq 550 platforms.

Both the sequencing of the 33.1 genome and library preparation were conducted at SENAI/CETIQT in Rio de Janeiro-RJ (senaicetiqt.com/). From the sequencing, the genome was assembled using the sequences generated by PacBio and the Microbial Assembly Pipeline of the SMRTLink package v.9.0 (Pacific Biosciences). Employing the raw data obtained by Illumina sequencing, the assembled genome was polished using Pilon software v.1.23 (Walker et al., 2014), and the genomic completeness was evaluated using BUSCO v.4.1.3 (Simão et al., 2015).

3.2.3. Genome annotation and prediction of PSS of the strain 33.1

The annotation of genome of the strain 33.1 was performed using the Prokaryotic Genome Annotation Pipeline (PGAP) from the National Center for Biotechnology Information (NCBI) (Tatusova et al., 2016), and additionally using Prokka v.1.14.6 (Seemann, 2014). In both cases, Artemis v. 18.1.0 (Carver et al., 2012) was utilized for annotation visualization.

To explore the biological information that sequencing can offer and considering the potential of strain 33.1 as a PGPB, BlastKOALA v. 2.2 (Kanehisa, Sato, and Morishima, 2016), and RAST v. 2.0 (Overbeek et al., 2014) servers were used to investigate processes related to plant growth promotion.

Additionally, Antibiotic Resistance Genes (ARGs) were predicted using RGI-Resistance Gene Identifier from the Comprehensive Antibiotic Resistance Database-CARD (Alcock et al., 2020), considering more than 70% identity with well-characterized ARGs in the database.

Using the genome annotation of strain 33.1 (RefSeq GCF_020149765.1), and to perform a highly accurate prediction of PSS and their components in 33.1, three different tools were employed. The automatic annotator TXSScan Galaxy v. 1.0.5 (Abby et al., 2016) was used, in addition to RAST v. 2.0 (Overbeek et al., 2014), and BlastKOALA v. 2.2 (Kanehisa, Sato and Morishima, 2016). The components predicted for each platform were thoroughly compared against each other and confronted with information available in the literature.

3.2.4. Comparative genomics of *P. agglomerans* PSS

Genomes of *P. agglomerans* with distinct functions (PGPB, biocontrol, bioremediation, and even phytopathogens and opportunistic pathogens) were obtained from the NCBI Assembly Database and annotated using Prokka v.1.14.6 (Seemann, 2014) (Table 4). Exploratory analyses of comparative genomics were performed also. Initially, a manual curation between 33.1 and these strains allowed the observation and comparison of characteristics of the selected strains, such as genome size, number of genes and proteins, number of plasmids, RNAs, among others.

Table 4. *P. agglomerans* strains used for comparative genomics analysis.

Strain	Reference	Role	RefSeq
L15	Rekosz-Burlaga, et al. (2014)	Biocontrol	GCF_003860325.1
UAEU18	Alkaabi, et al. (2020)	PGP	GCF_010523255.1
P5	Shariati, et al. (2017)	PGP	GCF_002157425.2

C1	Luziatelli, et al. (2020)	PGP	GCF_009759885.1
P10c	Smits, et al. (2015)	Biocontrol	GCF_001288285.1
ANP8	Noori, et al. (2021)	PGP	GCF_017315165.1
R190	Lim, et al. (2014)	Biocontrol	GCF_000731125.1
LMAE-2	Corsini, et al. (2016)	Bioremedition	GCF_000814075.1
4	Town, et al. (2016)	Biocontrol	GCF_000743785.2
E325	Pusey, et al. (2011)	Biocontrol	GCF_014353865.1
KM1	Guevarra, et al. (2021)	Opportunistic pathogen	GCF_012241415.1
BD 1274	Moloto, et al. (2020)	Phytopathogen	GCF_003369505.1
824-1	Nissan, et al. (2018)	Phytopathogen	GCF_001661985.2
4188	Nissan, et al. (2018)	Phytopathogen	GCF_001662025.2
Tx10	Smith, et al. (2013)	Pathogen	GCF_000475055.1
DAPP-PG 734	Moretti, et al. (2014)	Phytopathogen	GCF_000710215.1

With the previous annotation, it was possible to use the TXSScan-Galaxy v.1.0.5 (Abby et al., 2016) to predict the PSS of these strains and compare them with those present in 33.1.

3.2.5. Selection of target genes of PSS for knockout

To select candidate genes related to PSS for knockout, the following information was considered: *i)* the structure of each system, as well as genes related to the assembly of secretion machinery; *ii)* the completeness of the systems found in 33.1; *iii)* abundance of secretion systems, and *iv)* copy number for the same gene.

Based on the observed genes and the presented inferences, two genes from different secretion systems were selected for the knockout assays: *bepC* (from T1SS) and *prn* (T5aSS).

3.2.6. Bacterial strains, plasmids, and culture conditions for knockout

To knockout 33.1 PSS genes the CRISPR-Cas9 system was applied. *E. coli* DH5α was used to cloning purposes.

In the application of the CRISPR-Cas9 technique, the two-plasmids approach was performed, separating Cas9 and sgRNA (targeting the *bepC* and *prn* genes) into distinct plasmids. Both plasmids used (pCas and pTargetF) harbor several essential sequences for the system's functionality. The pCas (*Addgene* #62225) carries out the gene encoding the nuclease Cas9, from *Streptococcus pyogenes* MGAS5005; the λ-Red gene for homologous recombination of the donor DNA cassette; a temperature-sensitive replicon; the kanamycin resistance gene (*KanR*), and a small sgRNA with a $lacI^{q-}P_{trc}$ promoter guiding the pMB1 replication of pTarget (Jiang et al., 2015).

On the other hand, pTargetF (*Addgene* #62226) harbors the sgRNA and N20 sequence; the pJ23119 promoter; the spectinomycin resistance gene (*aadA*), and multiple restriction sites (Jiang et al., 2015).

3.2.7. Transformation of the strain 33.1 with the pCas

Electrocompetent 33.1 cells were obtained from a pre-inoculum in 1 mL of SOB medium (20 g peptone; 0.5 g NaCl; 5 g yeast extract, in 1000 mL distilled water, pH 7), which was transferred to another 100 mL of the same medium and incubated at 28°C, 200 rpm, until reaching an OD_{600} of 0.7. Subsequently, successive washes and resuspension of the bacterial pellet in MilliQ water and 10% glycerol were performed (two washes in water, one wash in 10% glycerol). Finally, the pellet was resuspended in 10% glycerol, and the OD_{600} was measured, ranging from 0.1-0.15. The entire process was carried out under refrigeration. The cells were then stored at -80°C.

For transformation, 200 ng of the pCas were added to 50 µL of electrocompetent cells of the strain 33.1 and electroporated using a Gene Pulser capacitor (Bio-Rad) at 25 uF, 25 Kv, 400 Ω resistor. Then, 1 mL of SOB medium was added to the transformed culture and incubated at 28°C, 180 rpm, for 3 hours to allow bacterial recovery. Subsequently, aliquots of 100 and 900 µL of the culture were plated on LB medium containing kanamycin (50

mg/mL) and incubated at 28°C for 18 hours for subsequent determination of transformation efficiency.

To select 33.1 transformants with the pCas (33.1.:pCas), the plasmids of the transformed strains were extracted using the QIAprep Spin MiniPrep Kit (Qiagen) according to the manufacturer's instructions, verified on a 1.2% agarose gel stained with SYBR Green (Sigma), and quantified using Biodrop (Biochrom). Using 20 ng of these plasmids as templates, a PCR reaction was performed employing primers designed for the detection of pCas (APPENDIX A).

The PCR reaction consisted of 1x Sigma Buffer; 0.2 mM of each dNTP; 0.2 µM of each forward and reverse primer; 1.6 mM $MgCl_2$; 0.06 U of Taq Polymerase (Sigma); in a final volume of 25 µL. The thermocycler cycles consisted of 1 initial denaturation cycle at 94°C for 5 minutes; 35 cycles of annealing at 95°C/30 sec, 55°C/30 sec, 72°C/30 sec; and a final extension at 72°C for 7 minutes. The amplicons were visualized on a 1.2% agarose gel stained with SYBR Green (Sigma) and sent for Sanger Sequencing at the Center for Nuclear Energy in Agriculture (CENA-USP), along with the cited primers for pCas detection (APPENDIX A). Then, using MUSCLE (Edgar, 2004) the sequencing-derived sequences were aligned with the pCas sequences to confirm its insertion. Sequences were analyzed using Jalview v. 2.11.2.7 (Waterhouse et al., 2009).

3.2.8. Obtaining and selecting N20 sequences for sgRNA

The N20 sequences required for the construction of sgRNA sequences were obtained based on the sequence of the selected genes (*bepC* and *prn*) and targeted to the respective genes in the genome of strain 33.1, flanked by a PAM sequence (Jiang et al., 2015).

To generate the N20 sequences, the CRISPR RGEN Tools Cas-Designer (www.rgenome.net/cas-designer/) (Bae et al., 2014) and CCTOP (https://cctop.cos.uni-heidelberg.de/) (Stemmer et al., 2015; Labuhn et al., 2018) tools were used, employing the genome of 33.1 as a reference. Obtained sequences from each of them were selected based on the criteria established by the tools themselves: CG content between 40-65%; out-of-frame > 66%; high CRISPRater index (for the CCTOP platform), and 0 off-targets. The

sequences in common between the platforms and with the best attributes were selected for the construction of pTarget series plasmids.

3.2.9. Construction of $pTarget_{bepC}$ and $pTarget_{prn}$ series

To construct the $pTarget_{bepC}$ and $pTarget_{prn}$, the gap-repair cloning (GRC) technique in *E. coli* DH5α was optimized. This technique generates a circular DNA through homologous recombination between two or more linear DNA fragments (Bessa et al., 2012). Using the pTargetF as a template, two fragments were constructed (Figure 11, APPENDIX A and APPENDIX B). Fragment A contained the replication origin, and the promoter j23119 and had 1,034 bp, while fragment B contains the N20 sequence, sgRNA, and the spectinomycin resistance gene, and had 1,123 bp.

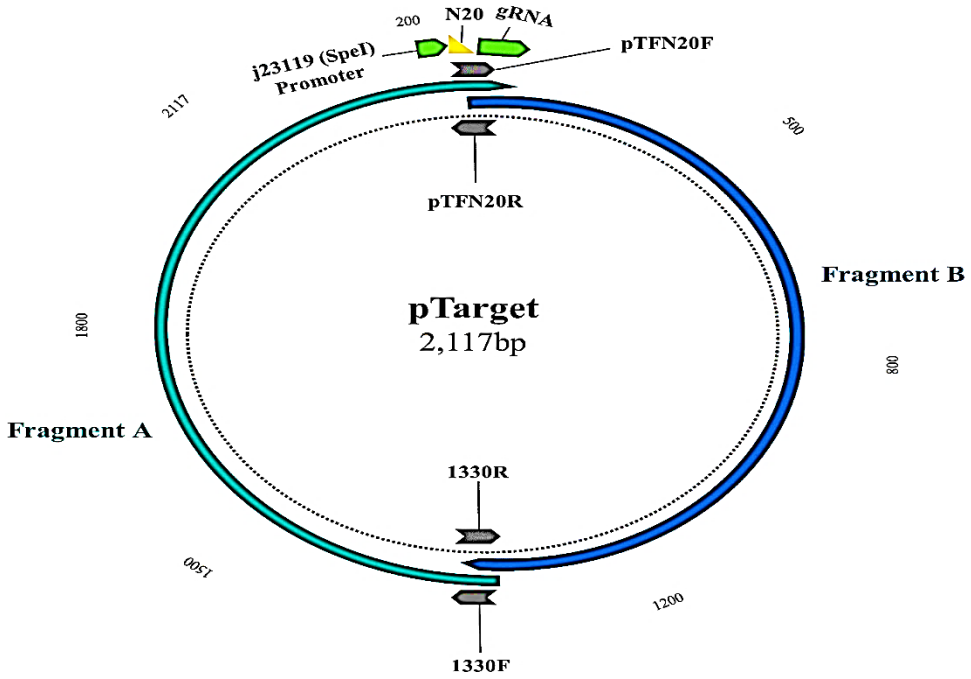

Figure 11. Assembly for pTarget series plasmid. The pTarget -series *bepc* and *prn*- was based on the building of two different fragments (A and B), using specific primers (pTFN20R+1330F, for fragment A, and pTFN20F+1330R, for fragment B), on the pTargetF as a template. Fragment A harbors the replication origin and the promoter, and fragment B the N20 sequence in the sgRNA, and the spectinomycin resistance gene.

To construct the pTarget series, the first step of the gap-repair protocol involved the backbone preparation by amplification. To do that, primers called pTFN20 series (APPENDIX A) were designed to harbor the N20 sequence and to contain regions of the promoter

pij23119 (reverse primers) or of the sgRNA (forward primers). So, using the pTargetF as a template, the pTFN20 primers, and primers named as 1330 series (designed to allow the fragments assembly with similar sizes), fragments A and B were amplified using a High-Fidelity DNA Polymerase with the following reaction: 1x Phusion HF Buffer; 0.2 mM of each dNTP; 0.2 µM of forward and reverse primer; 3% DMSO; 0.02 U/µL of Phusion High-Fidelity DNA Polymerase (Thermo Scientific); and 10ng of pTargetF, in a final volume of 50 µL. The thermocycler program consisted of an initial denaturation cycle at 98°C for 30 seconds, followed by 35 cycles of denaturation at 98°C for 10 seconds, annealing at 60°C for 30 seconds, extension at 72°C for 30 seconds, and a final extension cycle at 72°C for 10 minutes. The amplicons were visualized on a 1.2% agarose gel stained with SYBR Green (Sigma).

To optimize the plasmid template removal protocol, different strategies were employed: *i)* excision and purification of bands from agarose gel with expected size, *ii)* digestion with DpnI, and *iii)* a hybrid approach involving treatment of excised and purified gel fragments with DpnI. However, this latter approach resulted in significant DNA losses. Therefore, the optimization of the pTargetF template removal protocol, resulting in samples with satisfactory concentrations and integrity, is described as follows.

Fragments A and B were treated with DpnI (New England Biolabs) to remove the template plasmid. The reaction mix contained 1x NEBuffer, 20 U of the enzyme, and 2 µg of the amplified product from the Phusion High-Fidelity reaction. Incubation was carried out at 37°C for 75 minutes, followed by enzyme denaturation at 80°C for 20 minutes. The treated fragments were verified on a 1.2% agarose gel stained with SYBR Green (Sigma).

The fragments were then purified using the illustra™ GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) according to the manufacturer's instructions, and verified on a 1.2% agarose gel stained with SYBR Green (Sigma).

To recombine the fragments A and B, the transformation process was performed as described in section 3.2.7, using 200 ng of each purified fragment, and 50 µL of DH5α electrocompetent cells previously obtained. The transformed culture was incubated with 1000 mL of LB medium at 37°C for 1 hour, under 180 rpm. Aliquots of 100 and 900 µL of transformants were plated on LB medium containing spectinomycin (100 mg/mL), and the plates were incubated at 37°C for 24 hours.

The transformants were identified by colony PCR using designed primers (APPENDIX A) for the detection of the N20 sequence insertion in the plasmid. The PCR reaction

consisted of 1x Platus KCl buffer; 0.2 mM of each dNTP; 0.2 µM of forward and reverse primers; 1.6 mM MgCl2, and 0.025 U of Platus DNA Polymerase (Sinapse), in a final volume of 25 µL. The thermocycler program included an initial denaturation cycle at 95°C for 10 minutes, followed by 35 cycles at 95°C/30 sec, 60°C/30 sec, 72°C/2 min, for annealing. A final extension cycle was performed at 72°C for 7 minutes. The amplicons were visualized on a 1.2% agarose gel stained with SYBR Green (Sigma).

The positive clones had their pTarget series plasmids extracted following the QIAprep Spin MiniPrep Kit (Qiagen) instructions, verified on a 1.2% agarose gel stained with SYBR Green (Sigma), and quantified using Biodrop (Biochrom). Then, the plasmids were submitted for Sanger Sequencing at Center for Nuclear Energy in Agriculture (CENA-USP), along with multiple primers, to sequence their whole length. Using MUSCLE (Edgar, 2004) and T-Coffee (Notredame et al., 2000) the sequences obtained from sequencing were aligned with the pTargetF sequences to confirm the N20's insertion and the correct assembly of the plasmid. Jalview v. 2.11.2.7 (Waterhouse et al., 2009) was employed to view the results.

3.2.10. Construction of donor cassettes for homologous recombination

Downstream and upstream regions of the N20 sequences (excluding the N20s) in each interest gene (*bepC* and *prn*) in the genome of 33.1 were used to draw the needed primers to build the donor cassettes (APPENDIX A). To obtain the DNA, the overlap PCR technique was performed. This technique is based on the assembly of PCR fragments in a single DNA sequence.

To do that, the reverse complement of the reverse upstream primer was inserted into the beginning of the forward downstream primer, and vice versa (Choi and Schweizer, 2005), as shown in Figure 12.

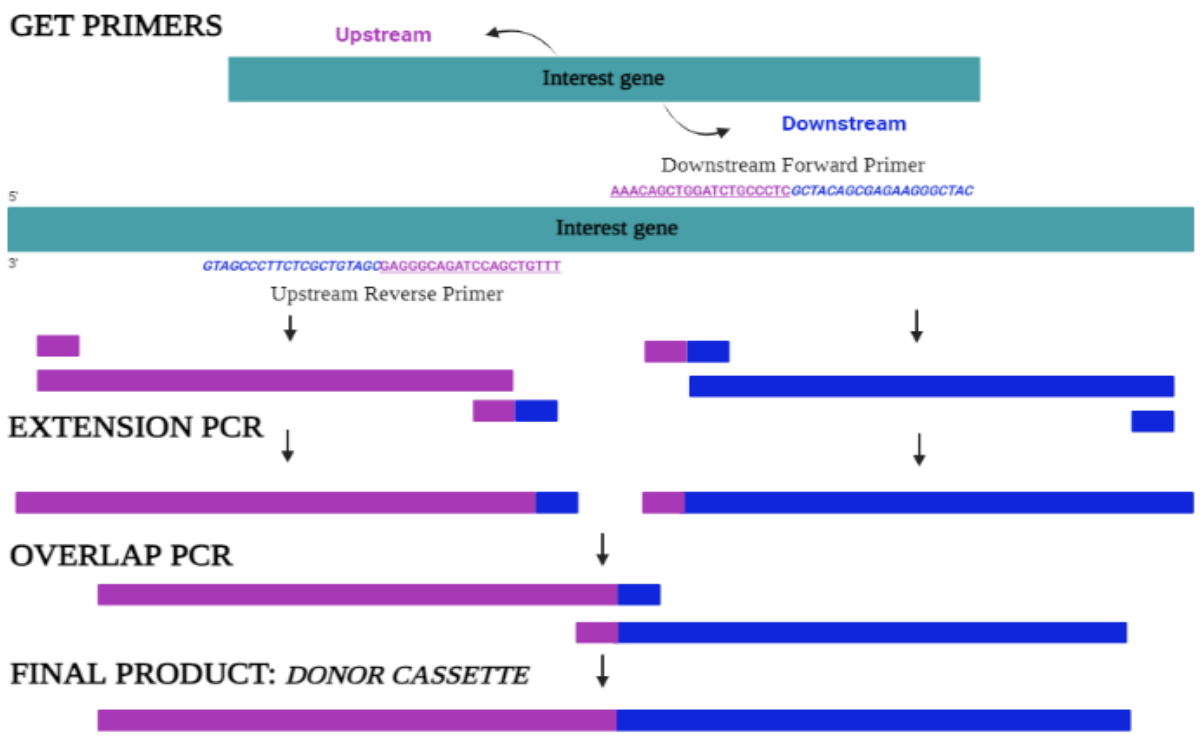

Figure 12. **Overlap PCR workflow.** The steps to obtain the donor cassettes are started by drawing the primers, followed by an extension PCR to attach the overhangs. Finally, the final product is assembled by the overhangs overlap.

The overlap PCR consisted of two PCR rounds. The first one, the primers were used to amplify the interest region in the 33.1 genome and to attach the overhangs in the amplicon. The first round was performed as follows: 1x Sinapse Buffer; 0.2 mM of each dNTP; 0.2 µM of each forward and reverse primers; 1.6 mM $MgCl_2$; 0.06 U of Taq Polymerase (Sinapse), and 20 ng of 33.1 DNA, in a final volume of 50 µL. The cycles consisted of 1 initial denaturation cycle at 94°C for 5 minutes; 35 cycles of annealing at 95°C/30 sec, 55°C/30 sec, 72°C/1 min; and a final extension at 72°C for 7 minutes. The amplicons were visualized on a 1.2% agarose gel stained with SYBR Green (Sigma).

The second round overlapped the overhangs of the upstream and downstream sequences. Thus, 25 ng of each purified 1° round PCR product (upstream and downstream amplicons) was used as a template in this step. The PCR reaction was performed as described in the first round. Primers were not added to the reaction, and after three annealing cycles in the thermocycler, the equipment was paused and 0.2 µM of forward upstream and reverse downstream primers were added to the reaction. Finally, the cycles

were resumed until the end of the PCR program, and the overlapped fragments were gel-purified (illustra™ GFX PCR DNA and Gel Band Purification Kit- GE Healthcare).

3.2.11. Knockout of *bepC* and *prn* genes from PSS

Electrocompetent cells of the 33.1.:pCas were obtained from the culture grown in LB medium containing kanamycin (50 mg/mL) and L-arabinose (10 mM), to induce the λ-Red genes. Electrocompetent cell was obtained as described in section 3.2.7.

Then, 50 µL of electrocompetent cells were gently mixed with 100 ng of the respective pTarget series ($pTarget_{bepC}$ or $pTarget_{prn}$) and 400 ng of the respective donor cassette. Electroporation was conducted as described in item 3.2.7. Then, 1000 mL of LB medium supplemented with 10mM of L-arabinose was added to the electroporated cells, which were recovered at 28°C, under agitation at 180rpm. After 3 hours of recovery, aliquots of 100 and 900 µL of cells were plated on LB medium containing kanamycin (50 mg/mL), spectinomycin (100 mg/µL), and L-arabinose (10 mM), and incubated at 28°C for 24 hours.

Colony PCR was performed to confirm the knockout of the interest genes. The PCR reaction was conducted as follows: 1x Platus KCl buffer; 0.2 mM of each dNTP; 0.2 µM of forward and reverse primers; 1.6 mM MgCl2, and 0.025 U of Platus DNA Polymerase (Sinapse), in a final volume of 25 µL. The thermocycler program included an initial denaturation cycle at 95°C for 10 minutes; followed by 35 cycles at 95°C/30 sec, 60°C/30 sec, 72°C/1 min, for annealing. A final extension cycle was performed at 72°C for 7 minutes. The amplicons were visualized on a 1.2% agarose gel stained with SYBR Green (Sigma).

To confirm the deleted region in the *bepC* and *prn*, the forward upstream and reverse downstream primers from the donor cassette for the respective gene, used in the PCR, were sent along with the purified amplicons (illustra™ GFX PCR DNA and Gel Band Purification Kit- GE Healthcare) for Sanger Sequencing at Center for Nuclear Energy in Agriculture (CENA-USP). Sanger's results were aligned with the sequences of 33.1 wild type using T-Coffee (Notredame et al., 2000).

3.3. Results

3.3.1. Sequencing and assembly of strain 33.1's genome

The genome sequencing and assembly revealed that 33.1 harbors a circular chromosome of 4,087,626 bp and four plasmids of 527,897 bp, 207,704 bp, 16,637 bp, and 4,460 bp. A completeness of 99% was reported, considering the Enterobacterales class (0.2% duplicated genes), as determined by the BUSCO software (Simão et al., 2015). Notably, among the features observed in the sequencing and assembly of strain 33.1, the assembled genome exhibited good contiguity and completeness (Table 5).

Table 5. Genomic characteristics of 33.1 generated with PGAP.

Characteristics	Values
Total genome size (chromosome + plasmids) in bp	4,844,324
Replicons	5
% GC	58.5
Genes (total)	4,648
CDS (total)	4,534
rRNA	8, 7, 7 (5S, 16S, 23S)
tRNA	80
ncRNAs	12
Pseudo genes (total)	94

3.3.2. Genome annotation and prediction of PSS of the strain 33.1

Functional annotation of genes related to plant growth promotion enabled the localization of genes and clusters associated with P solubilization, siderophore and IAA production, as well as genes related to the intercellular communication process (quorum sensing) among others (Figure 13), supporting the previous report of strain 33.1 as a PGPB.

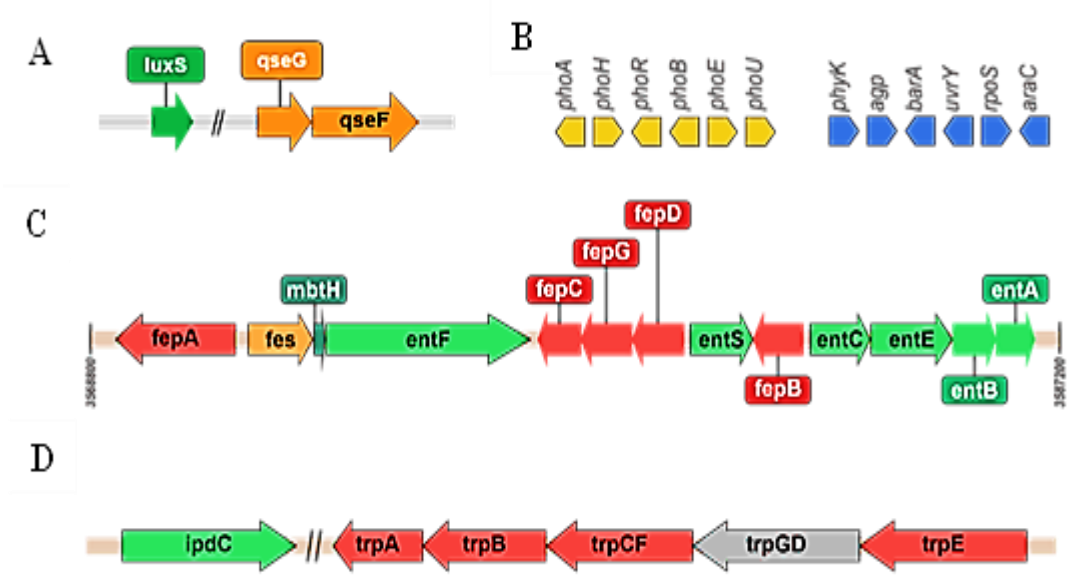

Figure 13. Gene clusters of processes related to plant growth promotion in strain 33.1, highlighting A) genes involved in the quorum sensing mechanism; B) genes involved in phosphatase activities (in yellow) and phytase activities (in blue); C) genes related to iron acquisition and metabolism - siderophore enterobactin; D) genes related to IAA biosynthesis.

Antibiotic resistance genes in 33.1 were identified using the RGI tool from the CARD database (Alcock et al., 2020), against its reference sequences. Six ARGs were identified in the chromosome, and 1 ARG was found in the plasmid p33.1_1 of 33.1. The identified ARGs in strain 33.1 provide resistance to several classes of antibiotics, such as macrolides, fluoroquinolones, elfamycins, aminoglycosides, carbapenems, cephalosporins, penams, penems, peptide antibiotics, diaminopyrimidines, phenicols, tetracyclines, rifamycins, as well as disinfectants and antiseptics (Table 6). However, these results are based on *in silico* analysis, so *in vitro* test to assess the sensitivity of 33.1 to these antibiotics should be performed.Antibiotic resistance genes found in 33.1's genome.

Table 6. AGRs located in 33.1 genome.

Antibiotic Resistance Ontology	Gene	Gene name	Gene Family	Antibiotic Class	Resistance Mechanism	% Identity	Location	Locus tag
ARO:3000518	*CRP*	cAMP-activated global transcriptional regulator CRP	Resistance-nodulation-division (RND) antibiotic efflux pump	Penam, macrolide, fluoroquinolone	Antibiotic efflux	99	Chromosome	LB453_03795
ARO:3003369	*tuf*	Elongation factor Tu	Elfamycin-resistant EF-Tu	Elfamycin	Alteration of antibiotic target	92	Chromosome	LB453_03920 LB453_20760
ARO:3004597	*emrB*	multidrug efflux MFS transporter permease subunit EmrB	Major facilitator superfamily (MFS) antibiotic efflux pump	Macrolide, fluoroquinolone, aminoglycoside, carbapenem, cephalosporin, penam, peptide antibiotic, penem	Antibiotic efflux	87	Chromosome	LB453_06620
ARO:3005069	*csrA*	carbon storage regulator CsrA	Resistance-nodulation-division (RND) antibiotic efflux pump	Fluoroquinolone, diaminopyrimidine, phenicol	Antibiotic efflux	85	Chromosome	LB453_06435
ARO:3004612	*ampH*	D-alanyl-D-alanine-carboxypeptidase/endopeptidase AmpH	AmpC beta-lactamase	Cephalosporin, penam	Antibiotic inactivation	73	Chromosome	LB453_07710
ARO: 3004583	*mdtI*	Multidrug/spermidine efflux SMR transporter subunit MdtI	Major facilitator superfamily (MFS) antibiotic efflux pump	Macrolide, aminoglycoside, cephalosporin, tetracycline, peptide antibiotic, rifamycin, disinfectants, and antiseptics	Antibiotic efflux	72	Plasmid p33.1_1	LB453_00800

The results obtained from different platforms to predict the 33.1 PSS (TXSScan-Galaxy v.1.0.5 (Abby et al., 2016); RAST v. 2.0 (Overbeek et al., 2014) and BlastKOALA v. 2.2 (Kanehisa, Sato and Morishima, 2016)) were manually curated and compared with the literature. This manual validation allowed the identification of the PSS, their components, and clusters present in 33.1.

It was observed that the chromosome of 33.1 harbors a T1SS, four T5SS (with the "a – autotransporter" and "b - two-partner secretion" subgroups located, the last in three clusters), and an incomplete T6SS. Additionally, a T5aSS was found on plasmid p33.1_1. Thus, it was possible to locate the PSS present in 33.1, the localization of their respective components, and the completeness of each system (Table 7).

The completeness of these systems is based on the structure of each machinery already reported in the literature. The T1SS is composed of 3 essential components: an Outer Membrane Protein/Factor (OMP or OMF), present in the outer membrane; an ABC transporter, present in the inner membrane, and a Membrane Fusion Protein (MFP), that spans the inner membrane and binds to the outer membrane (Green and Mecsas, 2016). The automatic annotation performed by Prokka v.1.14.6 (Seemann, 2014) allowed the identification of these components in the 33.1's chromosome as *bepC*, *ltxB*, and *prsE*, respectively.

The T5SS is subdivided into 5 groups ("a" to "e") (Van Ulsen et al., 2014), where subgroup 5a comprises autotransporters, while 5b encompasses a system known as TPS (Two-Partner Secretion), formed by the components TpsA and TpsB. In the case of 33.1, the autotransporters of the T5aSS system were found in the chromosome (identified as *prn*) and on plasmid p33.1_1. As mentioned, the TPS system (T5bSS) was found in three clusters in 33.1's genome.

The T6SS is formed by 13 essential components (Zoued et al., 2014). The obtained results allowed the identification of 6 (*cplV*, *hcp*, *tssB*, *tssC*, *tssL*, and *tssM*) of these components in 33.1, suggesting that it is incomplete.

In addition, secretion pathways components were found in the genome of 33.1, such as the Tat (Twin-Arginine Translocation) and Sec (general secretion) pathways (Table 8). Genes encoding proteins that facilitate the folding and transport of OMP, such as the Bam complex were also found (Table 9). Moreover, 33.1 also harbors important appendages related to protein secretion, such as the T4P (type IV pili)- sometimes classified as a

subgroup of T2SS (Table 9). Figure 14 shows a schematic representation of the components of each classical secretion system and secretion pathways found in 33.1 and their position in the secretory machinery.

Table 7. Classical PSS located in 33.1, their components, positions and sizes.

SYSTEM		COMPONENT	ORIGIN	LOCUS TAG	POSITION (nucleotides)	SIZE (amino acids)
T1SS		OMP	Chromosome	LB453_05960	724084...725424	447
		ABC transporter	Chromosome	LB453_05965	725421...727601	727
		MFP	Chromosome	LB453_05970	727598...728818	407
T5SS	**T5aSS**	Autotransporter	Chromosome	LB453_17740	3137709...3138596	296
	T5aSS	Autotransporter	Plasmid p33.1_1	LB453_00055	7875...9854	660
	T5bSS	TpsA	Chromosome	LB453_06730	898710...904178	1823
		TpsB	Chromosome	LB453_06720	896473...898167	565
	T5bSS	TpsA	Chromosome	LB453_11595	1882614...1894358	3915
		TpsB	Chromosome	LB453_11590	1880955...1882601	549
	T5bSS	TpsA	Chromosome	LB453_17675	3114846...3125468	3541
		TpsB	Chromosome	LB453_17680	3125570...3127234	555
T6SS		Hcp	Chromosome	LB453_09735	1513339...1513818	160
		TssM	Chromosome	LB453_13515	2275006...2277120	705
		TssL	Chromosome	LB453_13510	2273789...2275006	406
		TssC	Chromosome	LB453_08045	1152495...1153244	250
		TssB	Chromosome	LB453_08040	1151970...1152464	165
		CplV	Chromosome	LB453_06565	857865...860438	858

Table 8. Secretion Pathways located in 33.1, their location and sizes.

PATHWAY	COMPONENT	ORIGIN	LOCUS TAG	POSITION (nucleotides)	SIZE (amino acids)
Tat	tatA	Chromosome	LB453_20890	3824079...3824333	85
	tatB	Chromosome	LB453_20885	3823542...3824075	188
	tatC	Chromosome	LB453_20880	3822781...3823539	253
Sec	secY	Chromosome	LB453_09735	338098...339429	444
	secE	Chromosome	LB453_20755	3799070...3799453	128
	secG	Chromosome	LB453_19660	3558978...3559313	112
	secA	Chromosome	LB453_18695	3348740...3351445	907
	secB	Chromosome	LB453_21500	3964451...3964915	155

Table 9. Accessories and "helpers" for PSS functioning located in 33.1, their location and sizes.

PATHWAY	COMPONENT	ORIGIN	LOCUS TAG	POSITION (nucleotides)	SIZE (amino acids)
BAM	bamA	Chromosome	LB453_18335	3264775...3267186	804
	bamB	Chromosome	LB453_07430	1027925...1029106	394
	bamC	Chromosome	LB453_07630	1071238...1072269	344
	bamD	Chromosome	LB453_06550	855168...855902	245
	bamE	Chromosome	LB453_07165	966654...967001	116
T4P	pilA	Chromosome	LB453_18655	3342337...3342786	150
	pilB	Chromosome	LB453_18660	3342773...3344164	464
	pilC	Chromosome	LB453_18665	3344151...3345350	400
	pilD	Chromosome	LB453_06390	829311...830084	258
	pilM	Chromosome	LB453_03685	282076...282891	272
	pilN	Chromosome	LB453_03690	282891...283433	181
	pilQ	Chromosome	LB453_03705	284232...285515	428
	pilT	Chromosome	LB453_05900	693136...694188	351

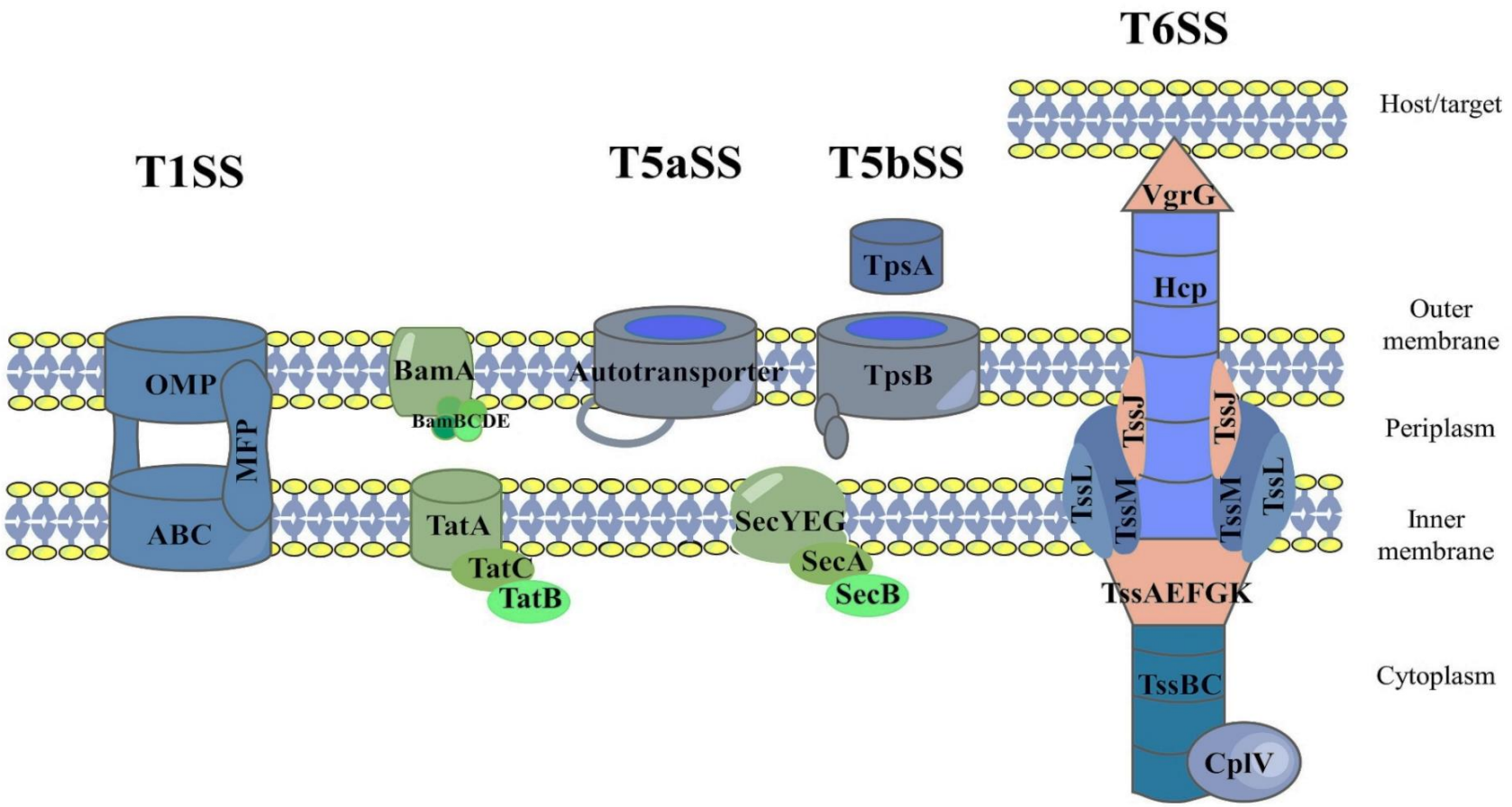

Figure 14. Hypothetical structure of T1SS, T5aSS, T5bSS and T6SS in 33.1. All essential genes for T1SS and T5SS are indicated in different shades of blue. Proteins in orange (in T6SS) were not localized in 33.1. The core components of the Tat and Sec pathways, and Bam Complex are demonstrated in different shades of green.

3.3.3. Comparative genomics of *P. agglomerans* PSS

Using the 33.1 genome annotation was possible to compare its characteristics to other *P. agglomerans* strains available on NCBI. Initially, the general genomic characteristics of these strains were evaluated, demonstrating the similarity among them (Table 10).

On the other hand, the PSS genes prediction allowed the comparison among the PSS located in 33.1 and in other *P. agglomerans* strains, and to observe the most abundant ones (Figure 15).

Table 10. Genome information of *Pantoea agglomerans* strains retrieved from NCBI.

Strain	Contigs, Scaffolds or chromosome	Plasmids	%GC	Total genome size	Genes	Proteins	tRNA	rRNA
33.1	1	4	58.55	4,844,324	4,655	4,436	80	22
L15	1	3	55.12	4,858,869	4,547	4,382	81	22
UAEU18	1	3	55.16	4,825,350	4,499	4,324	80	22
P5	127	0	54.90	5,074,260	4,788	4,611	53	8
C1	21	0	55.2	4,846,162	4,567	4,380	71	24
P10c	16	2	55.08	4,775,916	4,471	4,345	67	7
ANP8	1390	0	55	5,035,017	4,792	4,495	75	67
R190	2	3	55.1	5,002,566	4,682	4,483	79	24
LMAE-2	155	0	55.1	4,981,165	4,737	4,473	80	51
4	4	0	55.2	4,827,890	4,503	4,348	73	19
E325	162	0	55.2	4,786,783	4,557	4,262	73	35
KM1	46	3	55.1	4,995,756	4,660	4,508	64	2
BD 1274	246	0	55	4,968,508	4,703	4,545	39	11
824-1	1	4	54.8	5,034,205	4,775	4,561	77	22
4188	1	3	54.94	5,041,798	4,756	4,524	77	22
Tx10	25	0	55.1	4,856,993	4,554	4,404	72	22
DAPP-PG 734	195	0	54.7	5,365,929	5,183	4,918	72	36

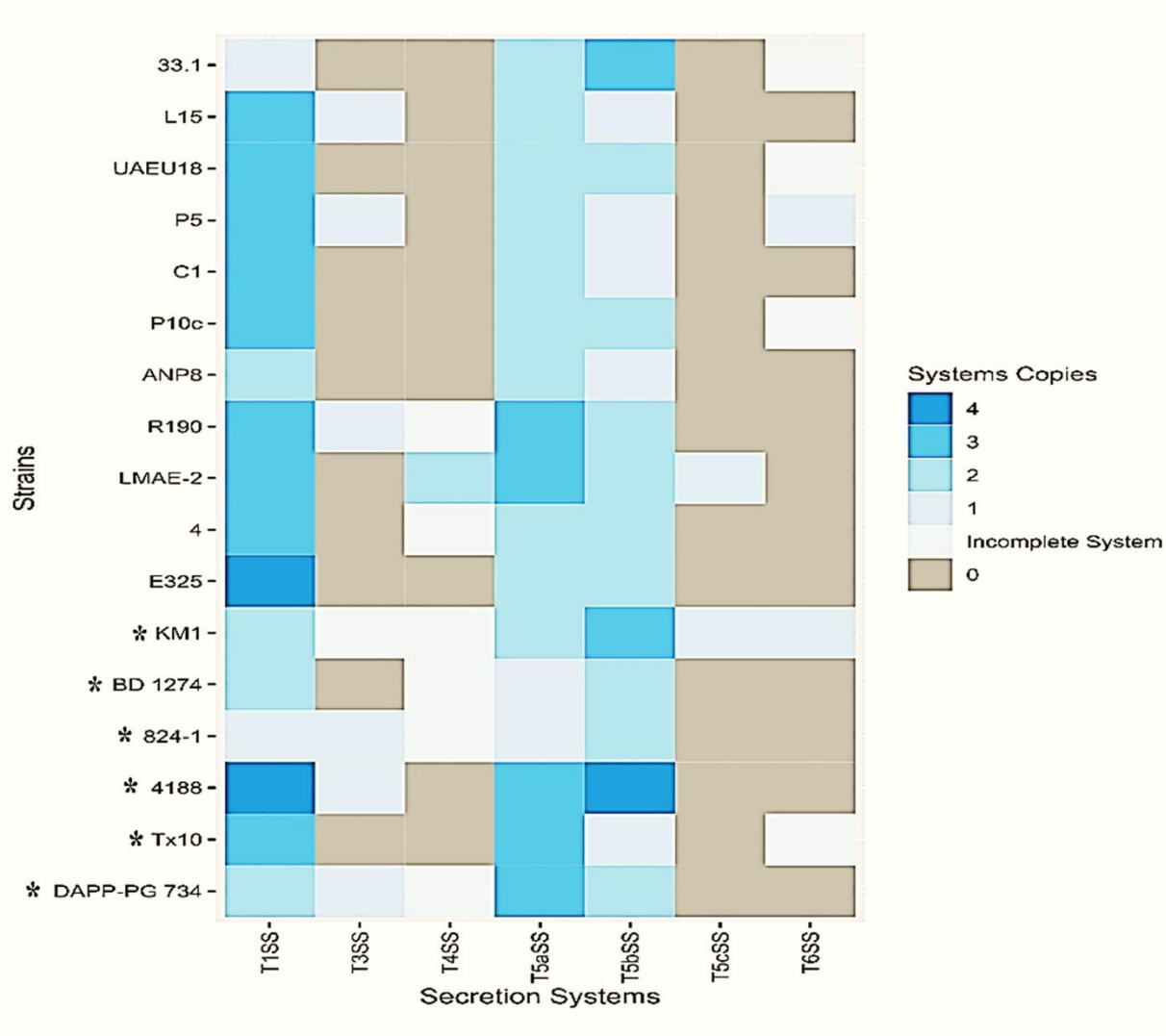

Figure 15. **Putative secretion systems in different *P. agglomerans* strains.** Blue represents gradual copy numbers of each PSS, where light blue represents incomplete systems and darker blue until 4 copies to the same PSS. Grey represents that systems were not found. Strains with an asterisk* (KM1, BD 1274, 824-1, 4188, Tx10, and DAPP-PG 734) exemplify the harmful strains.

The results allow us to observe the characteristics in common in the *P. agglomerans* strains, especially the PSS widely dispersed among them. In all evaluated strains, T1SS and T5SS (especially subtypes “a” and “b”), regardless of the beneficial or deleterious function of the microorganism, were found. Similarly, considering that these are strains of the same species, divided into groups according to the same function, the presence of T3SS, T4SS, and T6SS, was not homogeneous among them, demonstrating specificity according to each

strain. Furthermore, the proposed methodology was not able to find the T2SS in none of the studied stains.

3.3.4. Transformation of strain 33.1 with pCas

The observed results indicated the transformation of strain 33.1 with the pCas. After 18 hours of incubation, kanamycin resistant colonies were observed, allowing the calculation of transformation efficiency in 1 x 10^5 transformants per ug of DNA (Figure 16).

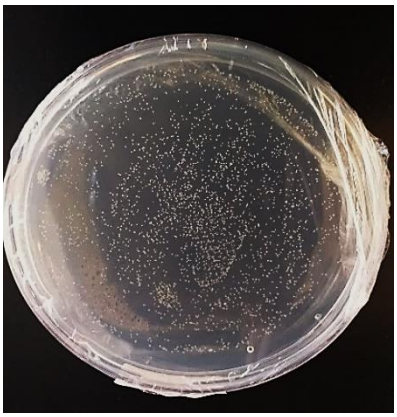

Figure 16. 33.1.:pCas clones after plating 100 µL of transformed culture and 18 h of incubation at 28°C.

To confirm the transformation, a PCR reaction using specific primers designed to detect the pCas was performed. Both the transformed bacteria and the positive control showed amplification of a fragment of approximately 600 bp, as expected (Figure 17).

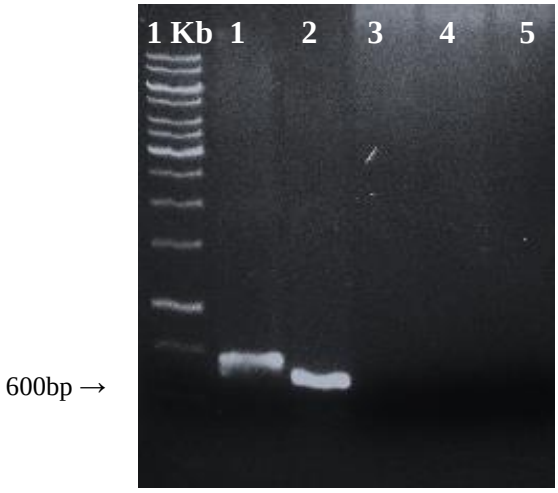

Figure 17. Confirmation of transformation of strain 33.1 with pCas. From left to right: 1 kb molecular marker (Thermo); 1- plasmid extracted from the 33.1.:pCas; 2- positive control: pCas used in transformation; 3- negative control: genomic DNA from 33.1; 4- negative control: PCR product without DNA; 5- SYBR Green (Sigma) without DNA sample.

In addition to PCR confirmation, the resulting amplicons were subjected to Sanger sequencing. The obtained sequences were compared to the known sequence of the pCas, and this comparison confirmed that the plasmid was correctly inserted into 33.1 (Figure 18).

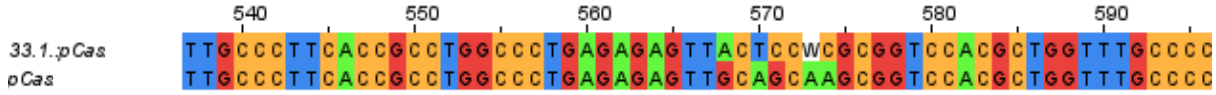

Figure 18. Partial sequence alignment of the pCas extracted from the transformed 33.1 (33.1.:pCas), against the pCas sequence (pCas). The same colors between aligned sequences represent identical sequences.

3.3.5. Obtaining and selecting N20 sequences for sgRNA

The CCTOP and CRISPR Rgen Tools allowed obtaining multiple N20 sequences for sgRNA targeting the *bepC* and *prn* genes (Figure 19). For the *bepC* gene, CCTOP generated 194 N20 sequences, while Rgen Tools generated 183. Additionally, 50 sequences met the specified parameters (CRISPRater - CCTOP's proprietary algorithm for determining sequence efficiency - >0.74; CG content between 40-65%; off-targets > 66%, and 0 mismatches/off-targets) in the first platform, and 55 in the second platform, with 12 sequences observed in common.

For the *prn* gene, CCTOP generated 103 N20 sequences, whereas Rgen Tools generated 104 sequences. A total of 26 sequences met the specified parameters in the first tool, and 20 in the second, with 5 sequences observed in common.

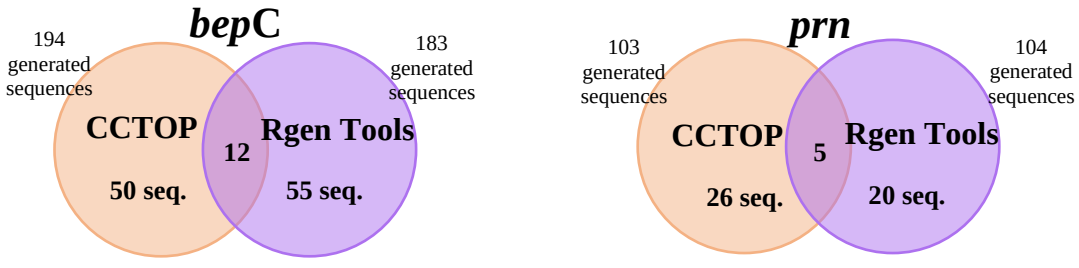

Figure 19. The N20 sequences generated by CCTOP (in pink) and Rgen Tools (in purple) for the *bepC* gene (left) and *prn* gene (right). Within each circle, the number of sequences obtained with the specified parameters is based on the total number of generated sequences (described outside the circles). The intersection points between the circles represent the sequences generated by both platforms that meet the determined parameters.

Two N20 sequences targeting each of the respective genes, generated by both platforms, were selected for assembling the sgRNAs contained in the pTarget$_{bepC}$ and pTarget$_{prn}$ (Table 11).

Table 11. N20 sequences targeting the *bepC* and *prn* genes of 33.1 generated by CCTOP and Rgen Tools. Right-pointing arrows indicate the 5'-3' direction.

Target gene	Sequence	PAM	Sense	GC%	*Out-of-frames%*	CRISPRater
bepC	CGACTCCCTGACCATTAATC	AGG	→	50,0	71,7	0,74
	AGTCGAGATGCCGCTCTATC	AGG	→	55,0	77,5	0,74
prn	TACGCGCATTCGTCCGCACG	CGG	→	65,0	79,9	0,74
	TATCGCGCGGCAAAAAGAGG	TGG	→	55,0	73,3	0,78

3.3.6. Construction of pTarget$_{bepC}$ and pTarget$_{prn}$ series

The pTarget(s) assembly was divided into two steps: in the first, using pTargetF as a template, two fragments (A and B) were obtained by amplification, and one of them contained the N20 sequence (fragment B). Agarose gel electrophoresis allowed the observation of the amplicons presenting the expected sizes for A and B (Figure 20).

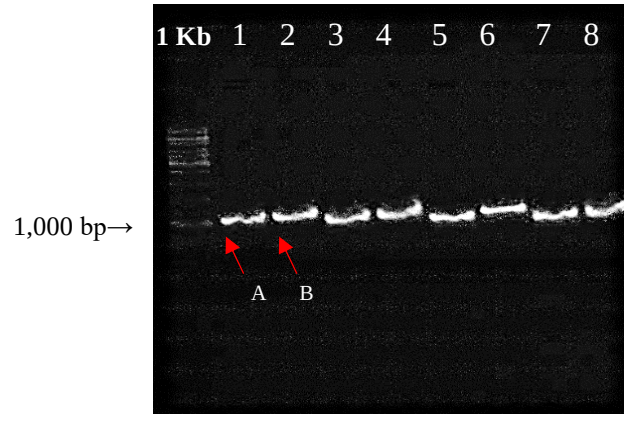

Figure 20. Fragments A and B for pTarget series assembly. From left to right: 1 kb molecular marker (Thermo); 1 and 3- fragment A for pTarget$_{bepC}$; 2 and 4- fragment B for pTarget$_{bepC}$; 5 and 7- fragment A for pTarget$_{prn}$; 6 and 8- fragment B for pTarget$_{prn}$. All fragments "A" had approximately 1,034 bp, as well as the "B" fragments have approximately 1,123 bp.

In the second step, after the plasmid template removal, fragments were recombined by transformation in *E. coli* DH5α. The possible clones were observed in the plates containing spectinomycin (Figure 21), and identified by PCR.

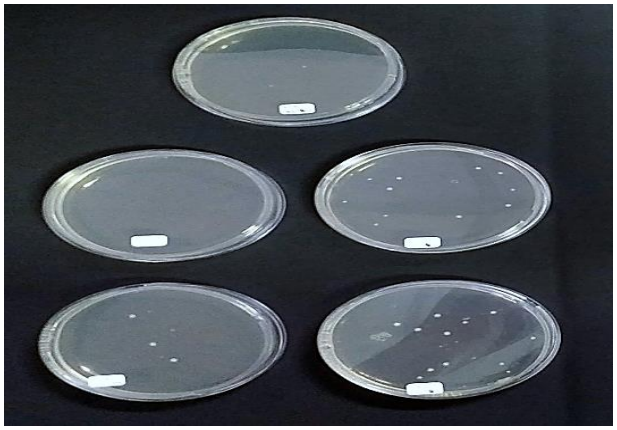

Figure 21. Possible pTarget$_{bepC}$ clones plated at different concentrations.

Sanger Sequencing of the pTarget series extracted from clones, using multiple primers, was able to cover the entire extent of the assembled plasmids, particularly the insertion region of the N20 sequences (Figure 22), demonstrating that they were correctly assembled.

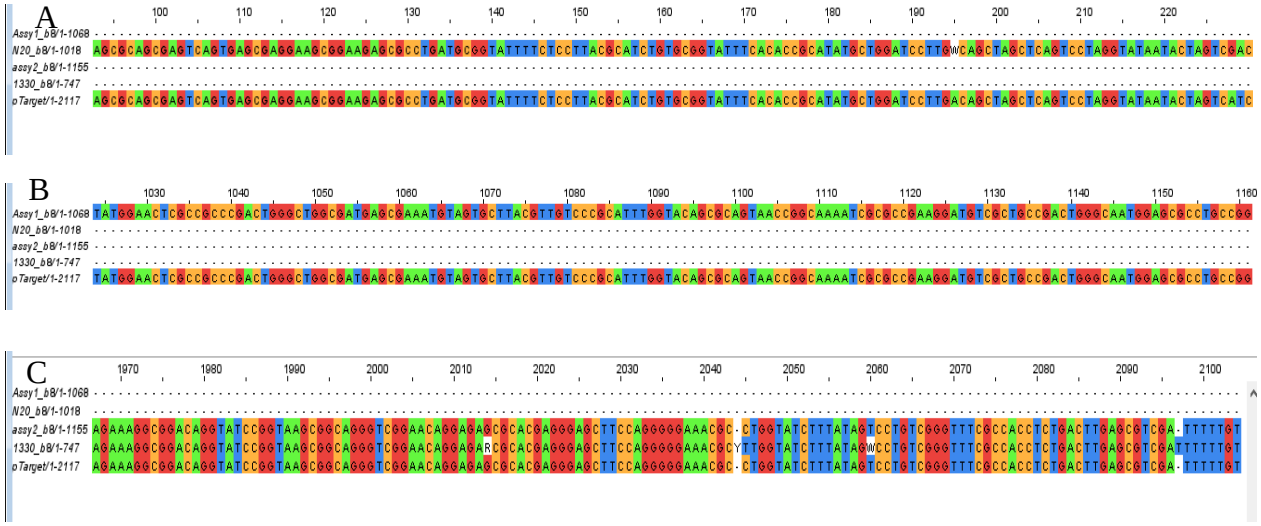

Figure 22. Alignment among the pTarget$_{prn}$ sequence and different forward primers, designed to cover its entire length. "A" represents primer started at 46 bp, "B" primer started at 842 bp, and "C" primers designed to cover from 1330 bp and 1548 bp. The same colors between aligned sequences represent identical sequences.

In the cloning of the pTarget$_{bepC}$, 34 potential clones were obtained, 26 were confirmed as false positives, and the remaining 8 were positive for cloning (2 for N20 option: CGACTCCCTGACCATTAATC, and 6 for N20 option: AGTCGAGATGCCGCTCTATC). As for the pTarget$_{prn}$, out of the 4 possible clones obtained, 3 resulted in false positives, and 1 had the cloning of N20 option TATCGCGCGGCAAAAAGAGG confirmed.

3.3.7. Construction of donor cassettes for homologous recombination

During the construction of the donor cassettes for *bepC* and *prn*, incompatibility of certain primer sets was observed, sometimes failing to amplify the upstream and downstream regions of the respective genes, or failing to allow the joining of the up and down fragments into a final cassette. Even after numerous optimization attempts, another difficulty observed was the formation of nonspecific bands during the overlap process (Figure 23), as seen by Choi and Schweizer (2005), which necessitated the excision and purification of the bands of the expected size (Figure 24).

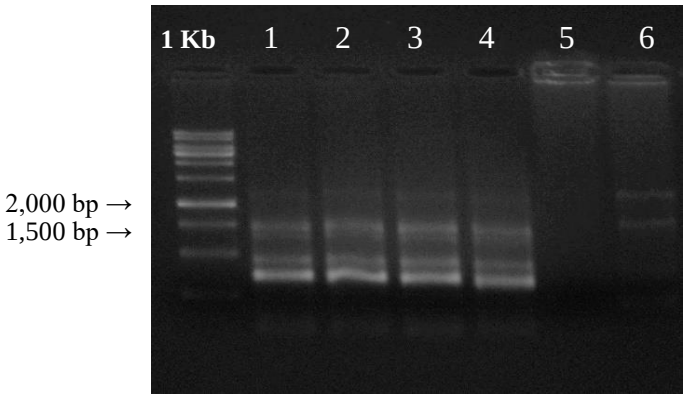

Figure 23. Nonspecific bands obtained after the second PCR round to assemble the *prn* cassettes. From left to right 1 kb molecular marker (Thermo); 1 to 4- *prn* donor cassettes; 5- negative control: PCR product without DNA; 6- positive control: intact DNA from 33.1.

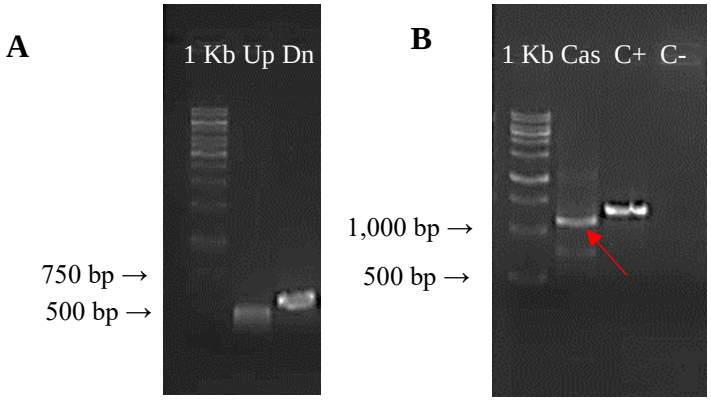

Figure 24. Construction of the donor cassette for the *bepC* gene. A) Obtaining the Upstream and Downstream fragments. From left to right: 1 kb molecular marker (Thermo); "Up" for the upstream fragment, approximately 400 bp, and "Dn" for the downstream fragment, approximately 600bp. B) Obtaining the *bepC* donor cassette (represented by the red arrow). From left to right: 1 kb molecular marker (Sinapse); "Cas" for *bepC* donor cassette, approximately 1,000 bp, discounting the deleted area; "C+" for the positive control, using intact DNA from the 33.1 as a template, approximately 1,200 bp, and "C-" for the negative control, without DNA sample.

The overlap PCR technique proved effective in joining multiple fragments into one (Figure 25). For the *bepC* donor cassette, upstream and downstream fragments of 394 bp and 597 bp were constructed, respectively. The final fragment consisted of 991 bp, with an excepted deletion of 206 bp compared to the intact gene fraction (1,197 bp). As for the *prn* donor cassette, fragments of 552 bp and 590 bp for the upstream and downstream regions were constructed, respectively. The constructed fragment, totaling 1,142 bp, featured a 695 bp deletion compared to the intact prn gene and adjacent regions (1,837 bp).

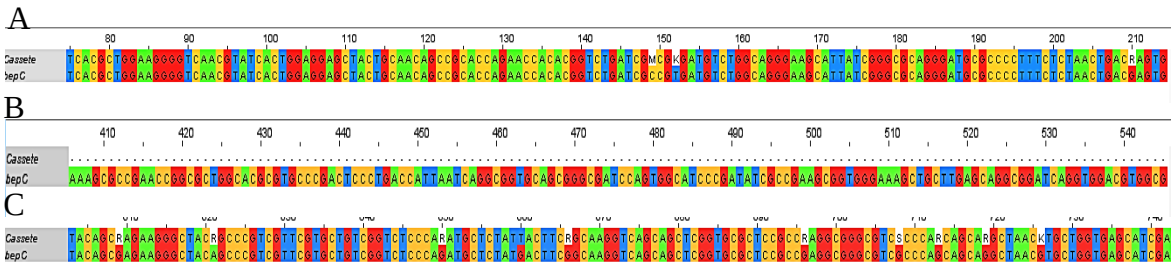

Figure 25. Partial alignment among the sequence of *bepC* donor cassette and the sequence of the intact *bepC* gene. "A" represents the upstream region, "B" the deleted region, and "C" downstream region. The same colors between aligned sequences represent identical sequences.

3.3.8. Knockout of *bepC* and *prn* genes from PSS of strain 33.1

Approximately 24 hours after plating, colonies with apparent resistance to the antibiotics kanamycin (50 mg/µL) and spectinomycin (100 mg/µL) were observed, suggesting the insertion of the pTarget series into the 33.1.:pCas cells (Figure 26).

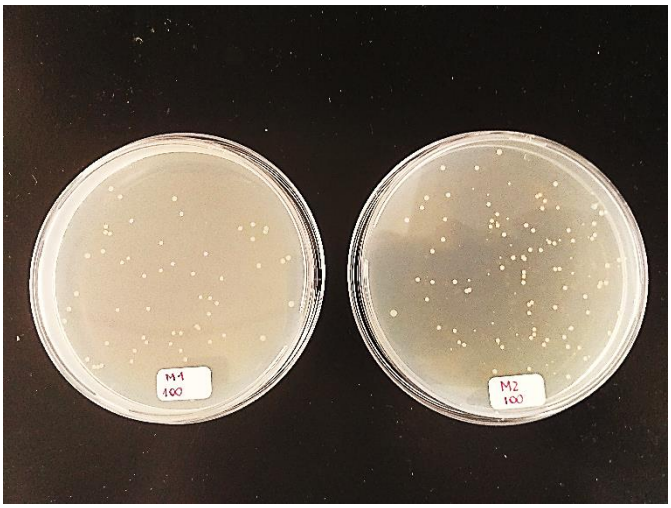

Figure 26. Possible *bepC* knockouts in LB medium containing kanamycin (50 mg/mL), spectinomycin (100 mg/mL), and L-arabinose (10 mM), after 24 h of incubation.

With no clear differences in amplicon size between the potential mutants and the 33.1 wild-type, the sequencing of these amplicons did not show any deletions in the *bepC* or *prn* gene regions, suggesting the ineffectiveness of gene knockout using the CRISPR-Cas9 (Figure 27).

```
33.1      CGCCCCTTTCTCTAACTGACGAGTGGAAAACGATGGATCGGATGCAATTAAGCGCAATATATCGCCGTTCTT
33.1_1    CGCCCCTTTCTCTAACTGACGAGTGGAAAACGATGGATCGGATGCAATTAAGCGCAATATATCGCCGTTCTT
cons      ************************************************************************

33.1      TTAAGCTCAGCGTTCTCTGTGCTGGCATCGCTTTTTTTACAACTTCTTCCGCTCTGGCGGCCAGCAGCGATC
33.1_1    TTAAGCTCAGCGTTCTCTGTGCTGGCATCGCTTTTTTTACAACTTCTTCCGCTCTGGCGGCCAGCAGCGATC
cons      ************************************************************************

33.1      TGATGCAGACGGCCAGCCAGATCACCGCCACCGGCCTGGCAAAACAGCAGGATCTGCCCTCTTTAACCGGCG
33.1_1    TGATGCAGACGGCCAGCCAGATCACCGCCACCGGCCTGGCAAAACAGCAGGATCTGCCCTCTTTAACCGGCG
cons      ************************************************************************

33.1      AAAGCGCCGAACCGGCGCTGGCACGCGTGCCCGACTCCCTGACCATTAATCAGGCGGTGCAGCGGGCGATCC
33.1_1    AAAGCGCCGAACCGGCGCTGGCACGCGTGCCCGACTCCCTGACCATTAATCAGGCGGTGCAGCGGGCGATCC
cons      ************************************************************************

33.1      AGTGGCATCCCGATATCGCCGAAGCGGTGGGAAAGCTGCTTGAGCAGGCGGATCAGGTGGACGTGGCGAAAG
33.1_1    AGTGGCATCCCGATATCGCCGAAGCGGTGGGAAAGCTGCTTGAGCAGGCGGATCAGGTGGACGTGGCGAAAG
cons      ************************************************************************
```

Figure 27. T-Coffee alignment between the *bepC* region for 33.1 wild-types (33.1) and the same region of the DNA of the colonies obtained after the knockout process (33.1_1). There is no evidence of deletions.

3.4. Discussion

The low availability of whole-genomes deposited in major databases, especially for non-model organisms, represents a challenge in reference genome assembly and comparative genomic studies. This information gap was observed for *P. agglomerans*, such that as of October 2023, only 21 complete genomes (including that of 33.1) have been deposited on the NCBI. Thus, the whole-genome of 33.1 with good quality and completeness not only allowed the mining of its plant growth-promoting genes, and especially of PSS genes, but will also provide necessary material for more comprehensive studies on the biology of this strain, as well as in understanding the molecular diversity of this species.

Among the annotated genes in the genome of 33.1 are those related to important processes well described in growth promotion, such as those involved in P metabolism (especially those conferring phytase and phosphatase activity), IAA and siderophore enterobactin production, quorum sensing, among others. Considering the more than 4 thousand genes of 33.1, several others associated with processes governing the bacteria/plant interaction can be functionally identified and deeply studied by our group.

Considering that PGPB typically harbor ARGs, the interest in the impact of such genes, which can be horizontally transferred to other microorganisms in the native soil community, to the plant, or to other niches, has increased (Ramakrishna, Yadav, and Li, 2019; Mahdi et al., 2022). However, despite the growing interest in this topic, the relevance of ARGs in this scenario remains unclear. In this context, the *in silico* identification of ARGs in 33.1 represents a starting point in investigations on this subject, and will be experimentally validated in the future.

Among the different strategies that can be employed to understand the role of a specific gene in a biological process, one is the perturbation of the gene's expression (e.g., knockout or silencing), followed by the observation of the consequences of this perturbation in the organism and/or its interactions with the environment (Passos et al., 2016). In recent years, the CRISPR system has gained prominence in the context of gene editing for understanding gene functions. In this research, CRISPR-Cas9 was employed for gene knockout of protein secretion systems (PSS) in the *P. agglomerans* 33.1.

Based on the whole genome sequencing for 33.1, its PSS genes were predicted. Among the genes encoding secretory proteins, two genes, *bepC* (from T1SS) and *prn* (from T5aSS), were selected for knockout using the CRISPR-Cas9 system. The optimization of a gene knockout protocol was conducted to contribute to the understanding of the PSS' function in plant growth promotion by the *P. agglomerans* 33.1, as presented in the previous sections.

Despite the increasing interest in PSS subject, there is much less available information about secreted proteins and their importance in beneficial plant/bacteria interactions than on processes related to pathogenicity (Molina, Ramos, and Espinoza-Urgel, 2006). Thus, the role of PSS in the beneficial plant/bacteria interaction is not well-understood (Dias et al., 2019). Indeed, the current knowledge in PSS and their mechanisms and features is still limited (Hui et al., 2021).

Therefore, this lack of information drives several difficulties in the PSS' *in silico* prediction. Some obstacles in the automatic annotation of PPS genes include: components of different secretion systems have sequences with significant similarity to each other (e.g. components of T2SS and T4SS); others are functionally equivalent but have little similarity; some PSS have not been well characterized yet, and/or domains that detect their components, deposited in databases (Romine, 2011).

In addition to the characteristics of the secretory machinery, computational resources (or the lack of them) can represent a barrier in PSS prediction. There is a clear interest in the systems with direct contact with the host (e.g. T3SS, T4SS, and T6SS), and this interest reflects in many tools developed for them. On the other hand, for other PSS (e.g. T1SS and T2SS) the shortage of tools remains (Hui et al., 2021).

To perform a comprehensive annotation for PSS genes and appendages, several tools were used in our study, in addition to manual validation. Even with some similar results, the prediction was not unanimous among the different platforms. Interestingly, despite being a conserved system among Gram-negative bacteria, none of the employed tools were able to find the classical T2SS (T2aSS), neither in 33.1 nor strains analyzed in comparative genomics assays, and only the subtype T2bSS (T4P) was found in 33.1.

Fimbriae and pili systems are essential for bacterial attachment and invasion, biofilm formation, and cell motility, as well as the transport of protein across membranes (Waksman and Hultgren, 2009). Due to their characteristics, these appendages are

occasionally classified as independent protein secretion systems. Desvaux et al. (2009) proposed that de Chaperone-Usher (CU) Fimbrial Biosynthesis Pathway was referred to as the type VII secretion system (T7SS) for Gram-negative bacteria. Still according to Desvaux et al. (2009), despite the CU Pathway presenting some similarity in structure and mechanisms to the T5SS, it should not be considered a subgroup of the T5SS, because these similarities are based on “analogy rather than homology”. However, one of the most cited reviews about PSS (Green and Mecsas, 2016), categorize the CU Pathway belongs to T5SS class.

T2SS and some systems for pili biosynthesis can share some similarities. It is the case of T4P (type IV pili system) and the Tad (tight-adherent pili system). Homology among T2SS, T4P, and Tad has been demonstrated through phylogenetic assays, suggesting a common ancestor among them (Desvaux et al., 2009; Tseng, Tyler, and Setubal, 2009). For Desvaux et al. (2009), and Hui et al. (2021), they should be considered subgroups of T2SS.

All these lack information and tools to improve the PSS’s studies, homologies among the systems, and reclassifications, making it difficult to understand the PSS assembly and mechanisms. However, we can find the classicals T1SS, T5aSS, T5bSS (in three clusters), T6SS (in an incomplete configuration); the Sec and Tat Pathways, and the accessories systems Bam and T4P, in the 33.1’s genome.

While some secretion systems are widely found in bacteria, others have been identified in a limited number of bacterial groups (Green and Mecsas, 2016). Comparing PSS from different *P. agglomerans* strains, we observed that T1SS and T5SS are ubiquitous among them. This conclusion agrees with Abby et al. (2016), which showed that T1SS and T5SS are the most widespread and abundant in bacterial genomes. According to Dias et al. (2019) a large amount of these clusters can indicate their importance in processes such as plant colonization and interactions with the microbiome.

The *in silico* prediction for PSS of 33.1 did not identify all 13 essential genes for the functioning of the classic T6SS (also known as T6SS^{i}). T6SS can be divided into subgroups based on their genetic organization and functioning. Even systems with fewer components or variations from the canonical T6SS^{i} can remain functional (Bernal, Llamas, and Filloux, 2018). Considering that, according to Bernal, Llamas, and Filloux (2018), the subgroups T6SS^{ii}, T6SS^{iii}, and T6SS^{iv} are absent in plant-related bacteria, and the classic T6SS is found primarily in Pseudomonodota (the phylum to which *Pantoea* belongs), it can be inferred that

the T6SS found in strain 33.1 belongs to the classic subgroup (T6SS or T6SS^{i}). Thus, due to its incomplete organization, it may not be functional.

So, the *bepC* and *prn* genes, from T1SS and T5aSS, respectively, were selected to knockout. It is important to highlight that reports in the literature indicate the use of T1SS in important processes for the PGPB/plant interaction. Among them is the secretion of the rhizobial protein NodO, which may extend nodulation capacity (Economou et al., 1990), and the secretion of organic acids and enzymes related to P solubilization and mineralization (Liu et al., 2020).

The components OMP (corresponding to the *bepC* gene) of T1SS, and AT (corresponding to the *prn* gene) of T5aSS, were chosen for gene knockout due to the evidence that their homologous or the PSS to which they belong may contribute to beneficial activities performed by microorganisms such as *R. leguminosarum*, *S. meliloti*, and *P. putida* (Economou et al., 1990; Cosme et al., 2008; Molina, Ramos, and Espinosa-Urgel, 2006).

Although it has not been demonstrated the impact of T5aSS on growth promotion, other subgroups of T5SS, such as T5bSS, have been described as necessary for the colonization of plants, such as corn (Molina, Ramos, and Espinosa-Urgel, 2006). Molina, Ramos and Espinosa-Urgel (2006) have observed that T5bSS mutants can have a reduced ability to colonize corn seeds, and these results indicated that a complete TPS system is necessary for the effective colonization of corn seeds.

Another important factor in the selection of the genes for being knocked out was the copy number of each gene. For example, the gene annotated as *cdiA* (corresponding to the TpsA component of T5bSS - this gene may be related to communication processes and inhibition of the growth of other surrounding bacteria) present in 10 copies in the genome, certainly performs an important biological functions, but due to its high number of copies, observing the knockout effect on the phenotype could be hampered. Thus, the *bepC* and *prn* genes, selected for knockout purposes, are present in a single copy in the genome.

The BepC, as a protein of the TolC family, plays an important role in the export of several molecules, from small agents (e.g. detergents, solvents, heavy metals and antibiotics), to large proteins, such as enzymes and toxins (e.g cyclolysin), interacting with a specific inner membrane translocase in each case (Koronakis et al., 2000; Koronakis, Eswaran, and Hughes, 2004). As an OMP, BepC is a membrane protein with a β-barrel

structure, that presents small loops between strands on the periplasmic side and large, extended them on the extracellular side (Rollauer, et al. 2015).

Due to its key location, exposed to the exterior of the bacterial cell, and its role as the first point of interaction between bacteria and their surroundings, OMPs are essential for several functions such as serving as adhesion factors, channels for nutrient uptake, siderophore receptors, and enzymes like proteases and lipases (Rollauer, et al. 2015).

According to our annotation, the autotransporter from T5aSS encoded in the 33.1 chromosome is the *prn* gene. Pertactin (*prn*) is an AT protein conserved in the *Bordetella brochiseptica* group, and a putative virulence factor of this genus (Inatsuka, et al., 2010). Despite the alignment of the sequence of 33.1's putative *prn* with other reviewed *prn* sequences available on UniProt (The UniProt Consortia, 2023) presents an identity lower than 30%, and there are no reports about *Pantoea* spp. harboring Pertactin genes, for the organization of this dissertation, we will keep Prokkas's annotation with the gene selected for knockout purposes named as a putative *prn*.

The CRISPR-Cas9 system for genome editing has two essential components: the cas9 and the sgRNA. The endonuclease cas9 uses the sgRNA to form base pairs with the DNA target sequences, allowing Cas9 to create a site-specific double-break in the DNA (Doudna and Charpentier, 2014). In the two-plasmid system, used in our research, the Cas9 and the sgRNA were separated into two different plasmids: pCas and pTargetF, respectively.

Despite the whole pCas-pTargetF system being properly assembled, as demonstrated in our results, the knockout process has failed. Several studies have shown the effectiveness of the pCas-pTargetF system for gene knockout (Jiang et al., 2015; Bai et al., 2020; Yang et al., 2021; Figueredo et al., 2023). Even for others *Pantoea* spp. this approach has been successfully applied, as demonstrated by Sun et al. (2022), where the role of exopolysaccharides of *Pantoea alhagi* NX-11 in its root colonization ability and in enhancing rice salt resistance was evaluated, through knockout using CRISPR-Cas9 system, specifically the pCas-pTarget approach.

Considering the scenario where cells obtained after knockout attempts showed resistance to the antibiotics spectinomycin and kanamycin (not observed in the wild-type strain), but the cleavage of the target genes did not occur, it is possible to hypothesize that some factors may have led to the failure of the knockout of PSS genes in the 33.1 strain.

The first of these is related to the expression of Cas9. Javaid and Choi (2021) reviewed several factors that could affect the efficiency of the CRISPR-Cas9 system and discussed how the level of Cas9 could influence the specificity and kinetics of gene editing. According to them, the expressing promoter, codon optimization, and positional effect, can directly affect the level of Cas9 expression and targeting efficiency. At this point, Cas9 may not be adequately expressed, and considering its role as a “molecular scissor”, without its active presence, there would be no DNA cleavage.

Li et al. (2021) demonstrated that the pCas-pTargetF system failed to edit *E. coli* BL21 (DE3), and suggested that the sgRNA present in pCas, which is specific to the pMB1 replicon of pTargetF, may show higher leaky expression in some strains, in this case, BL21 (DE3), which could result in pTargetF being cut by Cas9 itself, preventing the recovery of transformed cells. Thus, it would be possible to infer that this potential undesired cleavage of the pTargetF could lead to a series of consequences, such as the loss of functional sgRNA, resulting in the CRISPR-Cas9 system's inability to target the desired site for genome editing.

Other possible reason for the failure of knockout experiments could be the essential nature PSS genes. The essentiality of some genes is related with the impact on the organism's fitness and its importance in core cellular processes (Jordan et al., 2002; Peters et al., 2016). According to Jordan et al. (2002), a gene is considered essential if its knockout results in lethality or unfeasibility. As mentioned before, PSS play crucial roles in bacterial functioning and interactions with the environment. While previous studies have successfully deleted PSS genes or entire systems in other strains, Nicholson and Champion (2022) argue that some PSS are indispensable for bacterial physiology, and specialized or general PSS, whether essential or not, represent physiological niches necessary for overall bacterial survival.

Our research allowed the whole-genome sequencing of strain 33.1, assembled with good quality, which will enable increasingly comprehensive studies on its growth-promoting mechanisms and interactions. Furthermore, its complete genome has provided us the fundamental tool for the exploration of its PSS. From these findings, essential and accessory genes for the functioning of the secretory machinery have been identified in strain 33.1, aligning with those observed in other strains of the same species. Additionally, a knockout protocol developed for the model organism *E. coli* MG1655 (Jiang et al., 2015) was optimized for our strain. Sequences were cloned into plasmids, which were properly constructed and

inserted into 33.1, and donor cassettes for homologous recombination were also properly constructed.

Although gene knockout of PSS in the 33.1 was not achievable with the applied methodology, progress has been made toward understanding their function in plant growth promotion, even in a scenario of scarcity of information and specific tools. At this point, it is important to explore gaps and improve upon them, considering alterations in the plasmid system used for CRISPR-Cas9 knockout, as suggested by Li et al. (2021), or alternative approaches such as the λ-Red recombination technique already established for the 33.1 (Factor, 2023). Important experimental insights may be explored in the future regarding the essentiality and significance of PSS genes in plant growth promotion by 33.1.

REFERENCES

Abby, S. S., Cury, J., Guglielmini, J., Néron, B., Touchon, M., & Rocha, E. P. (2016). Identification of protein secretion systems in bacterial genomes. *Scientific reports*, *6*(1), 23080.

Alcock, B. P., Raphenya, A. R., Lau, T. T., Tsang, K. K., Bouchard, M., Edalatmand, A., ... & McArthur, A. G. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic acids research, 48(D1), D517-D525.

Alfano, J. R., & Collmer, A. (2004). Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu. Rev. Phytopathol.*, *42*, 385-414.

Alkaabi, A. S., Sudalaimuthuasari, N., Kundu, B., AlMaskari, R. S., Salha, Y., Hazzouri, K. M., ... & Amiri, K. M. (2020). Complete genome sequence of the plant growth-promoting bacterium Pantoea agglomerans strain UAEU18, isolated from date palm rhizosphere soil in the United Arab Emirates. *Microbiology resource announcements*, *9*(17), 10-1128.

Bae, S., Park, J., & Kim, J. S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*, *30*(10), 1473-1475.

Bai, L., Gao, M., Cheng, X., Kang, G., Cao, X., & Huang, H. (2020). Engineered butyrate-producing bacteria prevents high fat diet-induced obesity in mice. *Microbial Cell Factories*, *19*(1), 1-13.

Barrangou, R., & Doudna, J. A. (2016). Applications of CRISPR technologies in research and beyond. *Nature biotechnology*, *34*(9), 933-941.

Bernal, P., Llamas, M. A., & Filloux, A. (2018). Type VI secretion systems in plant-associated bacteria. *Environmental microbiology*, *20*(1), 1-15.

Bessa, D., Pereira, F., Moreira, R., Johansson, B., & Queirós, O. (2012). Improved gap repair cloning in yeast: treatment of the gapped vector with Taq DNA polymerase avoids vector self-ligation. *Yeast*, *29*(10), 419-423.

Carver, T., Harris, S. R., Berriman, M., Parkhill, J., & McQuillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*, *28*(4), 464-469.

Choi, K. H., & Schweizer, H. P. (2005). An improved method for rapid generation of unmarked Pseudomonas aeruginosa deletion mutants. *BMC microbiology*, *5*(1), 1-11.

Corsini, G., Valdés, N., Pradel, P., Tello, M., Cottet, L., Muiño, L., ... & Gonzalez, A. R. (2016). Draft genome sequence of a copper-resistant marine bacterium, Pantoea agglomerans strain LMAE-2, a bacterial strain with potential use in bioremediation. *Genome Announcements*, *4*(3), 10-1128.

Cosme, A. M., Becker, A., Santos, M. R., Sharypova, L. A., Santos, P. M., & Moreira, L. M. (2008). The outer membrane protein TolC from Sinorhizobium meliloti affects protein secretion, polysaccharide biosynthesis, antimicrobial resistance, and symbiosis. *Molecular plant-microbe interactions*, *21*(7), 947-957.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*, *13*(6), 343-359.

Cui, Y., Dong, H., Ma, Y., & Zhang, D. (2019). Strategies for applying nonhomologous end joining-mediated genome editing in prokaryotes. *ACS Synthetic Biology*, *8*(10), 2194-2202.

Desvaux, M., Hébraud, M., Talon, R., & Henderson, I. R. (2009). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends in microbiology*, *17*(4), 139-145.

Dias, G. M., de Sousa Pires, A., Grilo, V. S., Castro, M. R., de Figueiredo Vilela, L., & Neves, B. C. (2019). Comparative genomics of Paraburkholderia kururiensis and its potential in bioremediation, biofertilization, and biocontrol of plant pathogens. *MicrobiologyOpen*, *8*(8), e00801.

Dobo, B. (2022). Effect of arbuscular mycorrhizal fungi (AMF) and rhizobium inoculation on growth and yield of Glycine max L. varieties. International Journal of Agronomy, 2022.

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096.

Economou, A., Hamilton, W. D., Johnston, A. W., & Downie, J. A. (1990). The Rhizobium nodulation gene nodO encodes a Ca2 (+)-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins. *The EMBO journal*, *9*(2), 349-354.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, *32*(5), 1792-1797.

Factor, B. G. (2022). *Silenciamento gênico de Diatraea saccharalis utilizando a técnica de RNA de interferência (RNAi)* (Doctoral dissertation, Universidade de São Paulo).

Figueredo, E. F., da Cruz, T. A., de Almeida, J. R., Batista, B. D., Marcon, J., de Andrade, P. A. M., ... & Quecine, M. C. (2023). The key role of indole-3-acetic acid biosynthesis by Bacillus thuringiensis RZ2MS9 in promoting maize growth revealed by the ipdC gene knockout mediated by the CRISPR-Cas9 system. *Microbiological Research*, *266*, 127218.

Garneau, J. E., Dupuis, M. È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., ... & Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67-71.

Glick, B. R. (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica*, *2012*.

Green, E. R., & Mecsas, J. (2016). Bacterial secretion systems: an overview. *Virulence mechanisms of bacterial pathogens*, 213-239.

Guevarra, R. B., Magez, S., Peeters, E., Chung, M. S., Kim, K. H., & Radwanska, M. (2021). Comprehensive genomic analysis reveals virulence factors and antibiotic resistance genes in Pantoea agglomerans KM1, a potential opportunistic pathogen. *PLoS One*, *16*(1), e0239792.

Hui, X., Chen, Z., Zhang, J., Lu, M., Cai, X., Deng, Y., ... & Wang, Y. (2021). Computational prediction of secreted proteins in gram-negative bacteria. *Computational and Structural Biotechnology Journal*, *19*, 1806-1828.

Inatsuka, C. S., Xu, Q., Vujkovic-Cvijin, I., Wong, S., Stibitz, S., Miller, J. F., & Cotter, P. A. (2010). Pertactin is required for Bordetella species to resist neutrophil-mediated clearance. *Infection and immunity*, *78*(7), 2901-2909.

Javaid, N., & Choi, S. (2021). CRISPR/Cas system and factors affecting its precision and efficiency. *Frontiers in cell and developmental biology*, *9*, 761709.

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., & Yang, S. (2015). Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Applied and environmental microbiology*, *81*(7), 2506-2514.

Jiang, F., & Doudna, J. A. (2017). CRISPR–Cas9 structures and mechanisms. *Annual review of biophysics*, *46*, 505-529.

Jordan, I. K., Rogozin, I. B., Wolf, Y. I., & Koonin, E. V. (2002). Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome research*, *12*(6), 962-968.

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology*, *428*(4), 726-731.

Karimi, R., & Noori, A. (2022). Streptomyces rimosus rhizobacteria and Glomus mosseae mycorrhizal fungus inoculation alleviate salinity stress in grapevine through morphophysiological changes and nutritional balance. Scientia Horticulturae, 305, 111433.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., & Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*, *405*(6789), 914-919.

Koronakis, V., Eswaran, J., & Hughes, C. (2004). Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annual review of biochemistry*, *73*(1), 467-489.

Labuhn, M., Adams, F. F., Ng, M., Knoess, S., Schambach, A., Charpentier, E. M., ... & Heckl, D. (2018). Refined sgRNA efficacy prediction improves large-and small-scale CRISPR–Cas9 applications. *Nucleic acids research*, *46*(3), 1375-1385.

Li, Q., Sun, B., Chen, J., Zhang, Y., Jiang, Y., & Yang, S. (2021). A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in Escherichia coli. *Acta Biochimica et Biophysica Sinica*, *53*(5), 620-627.

Lim, J. A., Lee, D. H., Kim, B. Y., & Heu, S. (2014). Draft genome sequence of Pantoea agglomerans R190, a producer of antibiotics against phytopathogens and foodborne pathogens. *Journal of biotechnology*, *188*, 7-8.

Lucero, C. T., Lorda, G. S., Ludueña, L. M., Nievas, F., Bogino, P. C., Angelini, J., ... & Taurian, T. (2022). Participation of type VI secretion system in plant colonization of phosphate solubilizing bacteria. *Rhizosphere*, *24*, 100582.

Luziatelli, F., Ficca, A. G., Cardarelli, M., Melini, F., Cavalieri, A., & Ruzzi, M. (2020). Genome sequencing of Pantoea agglomerans C1 provides insights into molecular and genetic mechanisms of plant growth-promotion and tolerance to heavy metals. *Microorganisms*, *8*(2), 153.

Mahdi, I., Fahsi, N., Hijri, M., & Sobeh, M. (2022). Antibiotic resistance in plant growth promoting bacteria: A comprehensive review and future perspectives to mitigate potential gene invasion risks. Frontiers in Microbiology, 13, 999988.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., ... & Koonin, E. V. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, *9*(6), 467-477.

Molina, M. A., Ramos, J. L., & Espinosa-Urgel, M. (2006). A two-partner secretion system is involved in seed and root colonization and iron uptake by Pseudomonas putida KT2440. *Environmental Microbiology*, *8*(4), 639-647.

Moloto, V. M., Goszczynska, T., Hassen, A. I., Pierneef, R., & Coutinho, T. (2020). Draft genome sequences of Pantoea agglomerans strains BD1274 and BD1212, isolated from onion seeds, reveal major differences in pathogenicity and functional genes. *Microbiology Resource Announcements*, *9*(45), 10-1128.

Moretti, C., Cortese, C., Passos da Silva, D., Venturi, V., Torelli, E., Firrao, G., & Buonaurio, R. (2014). Draft genome sequence of a hypersensitive reaction-inducing Pantoea agglomerans strain isolated from olive knots caused by Pseudomonas savastanoi pv. savastanoi. *Genome announcements*, *2*(4), 10-1128.

Nicholson, K. R., & Champion, P. A. (2022). Bacterial secretion systems: networks of pathogenic regulation and adaptation in mycobacteria and beyond. *PLoS Pathogens*, *18*(7), e1010610.

Nissan, G., Gershovits, M., Morozov, M., Chalupowicz, L., Sessa, G., Manulis-Sasson, S., ... & Pupko, T. (2018). Revealing the inventory of type III effectors in Pantoea agglomerans gall-forming pathovars using draft genome sequences and a machine-learning approach. *Molecular plant pathology*, *19*(2), 381-392.

Noori, F., Etesami, H., Noori, S., Forouzan, E., Jouzani, G. S., & Malboobi, M. A. (2021). Whole genome sequence of Pantoea agglomerans ANP8, a salinity and drought stress–resistant bacterium isolated from alfalfa (Medicago sativa L.) root nodules. *Biotechnology Reports*, *29*, e00600.

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of molecular biology*, *302*(1), 205-217.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., ... & Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic acids research*, *42*(D1), D206-D214.

Passos, G. A., Hernandez, C. A. S., Sousa, L. C., Felício, R. F. M & Souza, T. A. J. (2016). Visão Geral do Mecanismo Básico de Ação. In Pereira, T. C*. Introdução à Técnica de CRISPR*, (pp. 39-47), Ribeirão Preto: Sociedade Brasileira de Genética.

Pereira, T. C*.* (2016). *Introdução à Técnica de CRISPR*, p. 250, Ribeirão Preto: Sociedade Brasileira de Genética.

Peters, J. M., Colavin, A., Shi, H., Czarny, T. L., Larson, M. H., Wong, S., ... & Gross, C. A. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*, *165*(6), 1493-1506.

Pusey, P. L., Stockwell, V. O., Reardon, C. L., Smits, T. H. M., & Duffy, B. (2011). Antibiosis activity of Pantoea agglomerans biocontrol strain E325 against Erwinia amylovora on apple flower stigmas. *Phytopathology*, *101*(10), 1234-1241.

Ramakrishna, W., Yadav, R., & Li, K. (2019). Plant growth promoting bacteria in agriculture: Two sides of a coin. Applied Soil Ecology, 138, 10-18.

Rekosz-Burlaga, H., Borys, M., & Goryluk-Salmonowicz, A. (2014). Cultivable microorganisms inhabiting the aerial parts of Hypericum perforatum. *Acta Scientiarum Polonorum Hortorum Cultus*, *13*(5), 117-129.

Rollauer, S. E., Sooreshjani, M. A., Noinaj, N., & Buchanan, S. K. (2015). Outer membrane protein biogenesis in Gram-negative bacteria. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1679), 20150023.

Romine, M. F. (2011). Genome-wide protein localization prediction strategies for gram negative bacteria. *BMC genomics*, *12*, 1-13.

Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. 3rd Edition, Vol. 1, Cold Spring Harbor Laboratory Press, New York.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068-2069.

Shariati J, V., Malboobi, M. A., Tabrizi, Z., Tavakol, E., Owlia, P., & Safari, M. (2017). Comprehensive genomic analysis of a plant growth-promoting rhizobacterium Pantoea agglomerans strain P5. Scientific reports, 7(1), 15610.

Shyntum, D. Y., Theron, J., Venter, S. N., Moleleki, L. N., Toth, I. K., & Coutinho, T. A. (2015). Pantoea ananatis utilizes a type VI secretion system for pathogenesis and bacterial competition. *Molecular Plant-Microbe Interactions*, *28*(4), 420-431.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19), 3210-3212.

Smith, D. D., Kirzinger, M. W., & Stavrinides, J. (2013). Draft genome sequence of the antibiotic-producing cystic fibrosis isolate Pantoea agglomerans Tx10. *Genome announcements*, *1*(5), 10-1128.

Smits, T. H., Rezzonico, F., Blom, J., Goesmann, A., Abelli, A., Kron Morelli, R., ... & Duffy, B. (2015). Draft genome sequence of the commercial biocontrol strain Pantoea agglomerans P10c. *Genome announcements*, *3*(6), 10-1128.

Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J., & Mateo, J. L. (2015). CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PloS one*, *10*(4), e0124633.

Sternberg, S. H., Haurwitz, R. E., & Doudna, J. A. (2012). Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *Rna*, *18*(4), 661-672.

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., ... & Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic acids research*, *44*(14), 6614-6624.

Tseng, T. T., Tyler, B. M., & Setubal, J. C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology*, *9*(1), 1-9.

Town, J., Links, M., & Dumonceaux, T. J. (2016). High-quality draft genome sequences of Pantoea agglomerans isolates exhibiting antagonistic interactions with wheat seed-associated fungi. *Genome Announcements*, *4*(3), 10-1128.

Van Ulsen, P., ur Rahman, S., Jong, W. S., Daleke-Schermerhorn, M. H., & Luirink, J. (2014). Type V secretion: from biogenesis to biotechnology. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1843*(8), 1592-1611.

Vo, J. L., Martínez Ortiz, G. C., Subedi, P., Keerthikumar, S., Mathivanan, S., Paxman, J. J., & Heras, B. (2017). Autotransporter adhesins in Escherichia coli pathogenesis. *Proteomics*, *17*(23-24), 1600431.

Vieira, G. V., Cecílio, N. T., Arruda, L. M., & Sales, K. U. (2016). Visão Geral do Mecanismo Básico de Ação. In Pereira, T. C*. Introdução à Técnica de CRISPR*, (pp. 39-47), Ribeirão Preto: Sociedade Brasileira de Genética.

Waksman, G., & Hultgren, S. J. (2009). Structural biology of the chaperone–usher pathway of pilus biogenesis. *Nature Reviews Microbiology*, *7*(11), 765-774.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., ... & Earl, A. M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one, 9(11), e112963.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189-1191.

Xavier, L. J., & Germida, J. J. (2003). Selective interactions between arbuscular mycorrhizal fungi and Rhizobium leguminosarum bv. viceae enhance pea yield and nutrition. Biology and fertility of soils, 37, 261-267.

Yang, L., Mu, X., Nie, Y., & Xu, Y. (2021). Improving the production of NAD+ via multi-strategy metabolic engineering in Escherichia coli. *Metabolic Engineering*, *64*, 122-133.

Zhang, C. (2023). Exploring the role of microbial interactions in soil and rhizosphere and their effects on litter decomposition, mycorrhizal associations, and plant growth (Doctoral dissertation, Utrecht University).

Zoued, A., Brunet, Y. R., Durand, E., Aschtgen, M. S., Logger, L., Douzi, B., ... & Cascales, E. (2014). Architecture and assembly of the Type VI secretion system. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1843*(8), 1664-1673.

APPENDICES

APPENDIX A. Primers used in this study.

Primer	Function	Sense	Sequence
PcasIF	pCas detection	Forward	CGGTGCCACTTTTTCAAGTT
PcasIR	pCas detection	Reverse	GTTTGAGACGAGTCGCTTCC
pTFN20 bepC- F1	sgRNA assembly- contains the N20 sequence- *bepC* (option 1)	Forward	CGACTCCCTGACCATTAATCGTTTTAGAGCTAG
pTFN20 bepC- R1	sgRNA assembly- contains the N20 sequence- *bepC* (option 1)	Reverse	GATATTTGGTCAGGGAGTCGACTAGTATTATACCTAGG
pTFN20 bepC- F2	sgRNA assembly- contains the N20 sequence- *bepC* (option 2)	Forward	AGTCGAGATGCCGCTCTATCGTTTTAGAGCTAG
pTFN20 bepC- R2	sgRNA assembly- contains the N20 sequence- *bepC* (option 2)	Reverse	GATAGAGCGGCATCTCGACTACTAGTATTATACCTAGG
pTFN20 prn- F1	sgRNA assembly- contains the N20 sequence- *prn* (option 1)	Forward	TACGCGCATTCGTCCGCACGGTTTTAGAGCTAG
pTFN20 prn- R1	sgRNA assembly- contains the N20 sequence- *prn* (option 1)	Reverse	CGTGCGGACGAATGCGCGTAACTAGTATTATACCTAGG
pTFN20 prn- F2	sgRNA assembly- contains the N20 sequence- *prn* (option 2)	Forward	TATCGCGCGGCAAAAAGAGGGTTTTAGAGCTAG

pTFN20 prn- R2	sgRNA assembly- contains the N20 sequence- *prn* (option 2)	Reverse	CCTCTTTTTGCCGCGCGATAACTAGTATTATACCTAGG
1330F	pTarget assembly (pairs with the respective pTFN20 primer for plasmid construction)	Forward	CATAAGTTCCTATTCCGAAGTTCCG
1330R	pTarget assembly (pairs with the respective pTFN20 primer for plasmid construction)	Reverse	CTTCGGAATAGGAACTTATGAGCTC
pTargetN20F	Detection of insertion of the N20 sequence	Forward	TACCGCCTTTGAGTGAGCTG
pTargetN20R	Detection of insertion of the N20 sequence	Reverse	CGCTTCCCTCATGACATTGC
pTargetAssy1F	Confirmation of correct pTarget assembly	Forward	GGCAGCGCAATGACATTCTT
pTargetAssy1R	Confirmation of correct pTarget assembly	Reverse	GGCGGTGCTACAGAGTTCTT
pTargetAssy2F	Confirmation of correct pTarget assembly	Forward	GCCGGATCAAGAGCTACCAA
pTargetAssy2R	Confirmation of correct pTarget assembly	Reverse	CGTTGCTGGCGTTTTTCCAT
BEPCcass1UpF	Assembly of donor cassette- gene *bepC*	Forward	TACCGCAGATGGCAATACCG
BEPCcass1UpR	Assembly of donor cassette - gene *bepC*	Reverse	GTAGCCCTTCTCGCTGTAGCGAGGGCAGATCCAGCTGTTT
BEPCcass1DnF	Assembly of donor cassette - gene *bepC*	Forward	AAACAGCTGGATCTGCCCTCGCTACAGCGAGAAGGGCTAC
BEPCcass1DnR	Assembly of donor cassette gene *bepC*	Reverse	TAAGGTAGTGCGTGACCTGC
BEPCcass2UpF	Assembly of donor cassette - gene *bepC*	Forward	TCGGTCTCCCAGATGCTCTA

BEPCcass2UpR	Assembly of donor cassette - gene *bepC*	Reverse	GGCTCTGCGATTCGTTTAGCTAAGGTAGTGCGTGACCTGC
BEPCcass2DnF	Assembly of donor cassette - gene *bepC*	Forward	GCAGGTCACGCACTACCTTAGCTAAACGAATCGCAGAGCC
BECPcass2DnR	Assembly of donor cassette - gene *bepC*	Reverse	AGCGGCGATCAGGATAATGG
PRNcassUpF	Assembly of donor cassette - gene *prn*	Forward	TTTAGCCGCGACGATCTCAG
PRNcassUpR	Assembly of donor cassette - gene *prn*	Reverse	ACTTCAGACGCATCCAGGTCGTAAAGCCCGTACTGCCAGG
PRNcassDnF	Assembly of donor cassette - gene *prn*	Forward	CCTGGCAGTACGGGCTTTACGACCTGGATGCGTCTGAAGT
PRNcassDnR	Assembly of donor cassette - gene *prn*	Reverse	ATCAGCGGCCCTTTCTTCAT

APPENDIX B. **Sequences of each fragment used to construct the pTargetF series.** In this construction pTargetF were divided into two fragments. Yellow sequences demonstrate 1330F-R series primers, and blue sequence indicates the promoter pij23119. Additionally, green sequences represent the site to insert the N20 sequences provided by platforms, and the pink represents the sgRNA sequence.

Fragment A

CATAAGTTCCTATTCCGAAGTTCCGCGAACGCGTAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGT
CCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA
CCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGT
ATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCT
ATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGC
CGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGCTGGATCCTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTC
ATCGCCGCAGCGGTTTCAG

Fragment B

CATCGCCGCAGCGGTTTCAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGAATTCTCTAGAGTCGACCTGCAGAAGCTTAGATCTATTACC
CTGTTATCCCTACTCGAGTTCATGTGCAGCTCCATAAGCAAAAGGGGATGATAAGTTTATCACCACCGACTATTTGCAACAGTGCCGTTGATCGTGCTATGATCGACTGATGTCATCAGCGGTGGAGTGCAATGTCATGAGGGAA
GCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGG
TTACGGTGACCGTAAGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTGGCG
TTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCA

GCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCG
CAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCA
GTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAAGATGCCGCTCGCCAGTCGATTGGCT**GAGCTCATAAGTTCCTATTCCGAAG**