Universidade de Aveiro Departamento de Biologia
2015

Helena de Fátima Silva Lopes

Microbioma do *Pinus pinaster*

Microbiome of *Pinus pinaster*

DECLARAÇÃO

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

Universidade de Aveiro Departamento de Biologia
2015

Helena de Fátima Silva Lopes

Microbioma do *Pinus pinaster*

Microbiome of *Pinus pinaster*

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Microbiologia, realizada sob a orientação científica da Doutora Anabela de Oliveira Pereira, Investigadora em pós-doutoramento do CESAM, Departamento de Biologia da Universidade de Aveiro e co-orientação da Doutora Isabel da Silva Henriques, Investigadora auxiliar do iBiMED, Departamento de Biologia da Universidade de Aveiro.

Apoio financeiro do FEDER no âmbito do programa COMPETE e da FCT-Fundação para a Ciência e a Tecnologia através do projeto MicroNema (PTDC/BIA-MIC/3768/2012 e FCOMP-01-0124-FEDER-028368).

o júri

presidente

Prof. Doutora Sónia Alexandra Leite Velho Mendo Barroso
Professora auxiliar com agregação do Departamento de biologia da Universidade de Aveiro

Doutora Jacinta Maria Marques de Oliveira
Investigadora em pós-doutoramento do CESAM, Departamento de Biologia da Universidade de Aveiro

Doutora Anabela de Oliveira Pereira
Investigadora em pós-doutoramento do CESAM, Departamento de Biologia da Universidade de Aveiro

agradecimentos

Agradeço à Universidade de Aveiro, à FCT pelo apoio financeiro e ao professor António Correia por me receber no projeto.
Gostaria de agradecer à Anabela Pereira e Isabel Henriques por me aceitarem no laboratório, pela orientação e inspiração.
As minhas colegas do projeto MicroNema, Marta Alves e Patrícia Matos, as quais considero muito e que foram um componente essencial na minha jornada.
As colegas de laboratório que me receberam tão bem, me ensinaram muito sobre técnicas e sobre a vida.
As minhas colegas do curso de microbiologia pelo apoio nos momentos de aperto e pelo companheirismo durante essa jornada. Especialmente à Sara Peixoto que sempre esteve do meu lado.
As colegas de casa Magda e Nicole por facilitarem minha vida e tornarem os dias mais divertidos.
Ao Steven pela companhia nas madrugadas de escrita.
À Flávia por me convencer a vir pra Portugal e por ter parado de fumar.
Aos amigos que ficaram no Brasil e me apoiam a distância.
À minha Avó, tios, tias e primos que me deram suporte e cuidaram de mim.
Aos meus pais e irmãs que me deram essa oportunidade.

"Who can say if I've been changed for the better?
But because I knew you
I have been changed for good". Musical Wicked

palavras-chave

Comunidades bacterianas; Bactérias endofíticas; *Pinus pinaster*; DGGE; Pirosequenciação; Microbioma.

resumo

Pinus pinaster possui uma grande importância ambiental, histórica e económica em Portugal, conquistando o terceiro lugar em ocupação florestal. Entretanto sua área está a diminuir devido a modificações na utilização dos terrenos, fogos e doenças como a da murchidão do pinheiro (DMP). Nativa da América do Norte, a DMP vem se espalhando em uma escala global, já está presente na Ásia e mais recentemente na Europa, sendo detectada em 1999 em Portugal, representando assim uma ameaça a florestas do mundo inteiro. Existem evidências que apontam para a importância da comunidade bacteriana no desenvolvimento da DMP e poucos estudos abordam o microbioma do *P. pinaster*, com a maioria utilizando metodologias baseadas em cultivo. Este trabalho tem como objetivo revelar a comunidade bacteriana de árvores de *P. pinaster* saudáveis utilizando técnicas independentes de cultivo. Para observar a comunidade bacteriana do *P. Pinaster* saudável, árvores entre 20 e 30 anos de idade foram recolhida de 4 localidades portuguesas: Comporta, Góis, Ilha da Madeira e Vouzela. O tronco dessas árvores foi cortado em discos e o serrim do centro desses discos foi utilizado para extração de DNA. Todas as amostras foram testadas para a presença do nemátodo sendo que este estava ausente em todas as amostras. Foi feito DGGE com o objectivo de comparar a comunidade bacteriana de diferentes amostras, e esta foi identificada utilizando-se pirosequenciação. Tendo em conta os perfis de DGGE, foi feita a clonagem e sequenciação de uma das amostras, e verificou-se que as bandas mais intensas, e presentes em todas as amostras, eram material genético proveniente de cloroplastos. Os resultados de pirosequenciação revelaram uma grande quantidade de *reads* proveniente de cloroplastos que foram eliminadas do estudo. A análise dos resultados da pirosequenciação permitiu identificar seis filos: Proteobacteria, Acidobacteria, Firmicutes, Bacteriodetes, Armatimonadetes e Actinobacteria. Proteobacteria foi o filo mais comum e dele Gammaproteobacteria foi a classe mais abundante. As amostras se agruparam de acordo com o local de origem e as comunidades dos locais se diferenciavam significativamente. Apenas duas OTUs eram compartilhadas por todos os locais. Apesar do baixo número de *reads,* a estrutura da comunidade bacteriana foi caracterizada e essa informação pode ser agora utilizada em futuros estudos.

keywords

Bacterial communities; Bacterial endophytes; *Pinus pinaster;* DGGE; pyrosequencing; Microbiome.

abstract

Pinus pinaster has a great environmental, historical and economical importance to Portugal. It represents the third place in occupation in forestry area. However, it has been losing area due to the modifications on land use, fires and diseases such as the pine wilt disease (PWD). Native from the North America the PWD has been spreading worldwide, it has spread through Asia and more recently was identified in Europe being detected in Portugal in 1999, representing a major threat to forests. Recently, the importance of the bacteria community to the PWD has been accessed and few studies address the microbiome of *P. pinaster*. Most of the existing studies uses culture-dependent techniques. This work aimed to reveal the bacterial community of healthy *P. pinaste*r trees using culture-independent techniques. To observe the bacterial communities of healthy *P. pinaster,* stem samples from trees aged between 20 and 30 years were collected from 4 Portuguese locations: Comporta, Góis, Madeira Island and Vouzela. The trunks were cut into disks and the sawdust of the center of these disks were used to extract DNA. Samples were tested for the presence of the PWD nematode and all were negative. DGGE analysis was performed to compare the bacterial community of different samples and pyrosequencing was used to identify the community. After analyzing the DGGE profiles, a sample was cloned and sequenced, and the results showed that the most common and intense bands belonged to chloroplast genetic material. Pyrosequencing results had a great amount of reads belonging to chloroplast and they were eliminated from the study. With the pyrosequencing, six plyla were detected: Proteobacteria, Acidobacteria, Firmicutes, Bacteriodetes, Armatimonadetes and Actinobacteria. Proteobacteria was the most common, and from this plylum, Gammaproteobacteria was the most abundant class. Samples grouped by location and the location community differ significantly, only two OTUs were shared by all locations. Despite the low read number the bacterial community was characterized, and this information can be used for future studies.

Table of contents

LIST OF FIGURES 2
LIST OF TABLES 4
INTRODUCTION 6
1. Pinus Pinaster 6
2. Microbiota of plants 8
2.1. Bacterial habitats in plants 9
2.2. Factors that modulate bacterial community 10
2.3. Host colonization 11
2.4. Bacterial phytopathogens 12
2.5. Beneficial bacteria 14
2.6. Endophytic bacteria in *Pinus* 15
3. Pine wilt disease 16
4. Objectives 18
MATERIAL AND METHODS 20
1. Sampling 20
1.1. Sampling sites 20
1.2. Sampling methodology 21
2. Genomic DNA extraction 21
3. Pine wilt disease nematode molecular detection 23
4. Denaturing Gradient Gel Electrophoresis (DGGE) analysis 23
5. Cloning and sequencing for DGGE bands phylogenetic affiliation 25
6. Barcoded 454 pyrosequencing 26
7. Statistical Analysis 28
RESULTS 29
1. DGGE analysis of asymptomatic tree samples 29
2. Cloning and sequencing 31
3. Analysis of pyrosequencing data 33
4. Description of the community 35
DISCUSSION 41
CONCLUSION 45
REFERENCES 46

LIST OF FIGURES

Figure 1 - Map of the natural and naturalized distribution of *Pinus pinaster*. Dark blue marks indicate the area where the species is distributed (*In* Alía & Martín 2003). 7
Figure 2 – Intervention Zones, areas where the presence of the PWD nematode is known or the nematode has an imminent risk of establishing itself and disperse, in continental Portugal (Adapted from Icnf.pt, 2015)..... 8
Figure 3 - Plant-associated microbes and their functions along the Poplar tree (*Populus sp.*) (*In* Schadt & Hacquard, 2015) 9
Figure 4 – Pine wilt disease cycle. The vector transmit the nematode to the tree during feeding and oviposition. Inside the host the nematode feed on the cells of resin canals and reproduce. The nematode migrates along the tree through resin canals, which is also colonized by bacteria. The insect vector larvae is infected by the nematode and when the adult insect emerge, it already carries the nematode, being able to spread the disease. 18
Figure 5 – Map of Portugal with numbered red dots marking the sampling locations. 1 – Madeira; 2 – Vouzela; 3 – Góis; 4 – Comporta. 21
Figure 6 - DGGE profiles from the midsection trunk of *P. pinaster* tree from the four locations. Each lane of the DGGE gel represents one tree sample named with a letter corresponding to the location (C – Comporta, G – Góis, M – Madeira and V – Vouzela) and numbered according storage order at the lab. Two lanes were loaded with a DGGE maker for internal normalization and as an indication of the quality of the analysis (Henriques *et al.*, 2006). 29
Figure 7 – Dendrogram of the DGGE profiles of the V3 region of the 16S rRNA gene from tree samples collected in the 4 locations (C – Comporta, G – Góis, M – Madeira and V – Vouzela). Patterns were created with UPGMA method and Pearson correlation. 30
Figure 8 – A and B are agarose gels showing PCR products obtained by electrophoresis and C a DGGE gel. Numbers 1 to 60 are the clone samples, L – Ladder, Cn– Negative control, CNn – Nested negative control, AZ – Blue colony sample (positive control) and C1 – Comporta sample (Donor sample). A) First PCR of the clones with the plasmid primers (T7F/13MR). The ladder used was Gene Ruler DNA ladder mix (Thermo Scientific- Fermentas, Burlington, Canada). B) Second PCR of the clones for DGGE analysis using primers for the V3 region of the16S rRNA gene with a GC clamp (338F – GC/518R). The ladder used was Gene Ruler DNA ladder mix (Thermo Scientific- Fermentas, Burlington, Canada). C) DGGE of the clones with each lane representing a clone sample. Two lanes were loaded with a DGGE marker for internal normalization and as an indication of the quality of the analysis (Henriques *et al.*, 2006)..... 32
Figure 9 – Rarefaction curves for all samples selected for pyrosequencing. The vertical line indicates the sample size of the smallest sample. 34
Figure 10 - Relative abundance of bacterial phyla for samples from all locations, Comporta (C), Góis (G), Madeira (M) and Vouzela (V). The relative abundance was calculated as the percentage of sequences belonging to a particular lineage of the total sequences recovered for each sample. 36
Figure 11 - Relative abundance of bacterial families from the phylum Proteobacteria from samples of all locations, Comporta (C), Góis (G), Madeira (M) and Vouzela (V). The relative abundance was calculated as the percentage of sequences belonging to a particular family of the total sequences for the Proteobacteria phylum. 37
Figure 12 – Venn diagram showing the number of shared and unique OTUs among locations (Madeira, Vouzela, Góis and Comporta). The diagram was obtained using an online tool (http://bioinformatics.psb.ugent.be/webtools/Venn/). 38
Figure 13 – Dendrogram based on the percentage of similarity between locations from the pyrosequencing data. Vouzela (V), Madeira (M), Góis (G) and Comporta (C). 40

LIST OF TABLES

Table 1– Factors involved in modulating the plant's bacterial community. 11
Table 2 – Rank of the top phytopatogenic bacteria. (Adapted from Mansfield *et al*., 2012). 14
Table 3 – Sawdust DNA extraction protocol. 22
Table 4 – BLAST results for the sequencing of selected clones. 33
Table 5 – Total sequences and OTUs obtained for all samples after quality treatment and the removal of chloroplast-affiliated reads. OTUs were defined with a 97 % of similarity. 34
Table 6 – OTUs affiliated at the species level; the number of reads in each sample is presented. 35
Table 7 – Heat map with the families of OTUs that represent more than 10 % of the reads of each sample, with the exception of OTU_15 (*) since class was the lower phylogenetic taxa that this OTU could be assigned to. 39
Table 8 – Values for the diversity indexes calculated for the pyrosequencing data. 40

INTRODUCTION

1. *Pinus Pinaster*

Forestry plays an important role in Portugal economy, being the main use of the soil in continental Portugal, forests occupy 35.4 % of Portuguese territory (ICNF, 2013). In the European context, forestry activities in Portugal have a great impact on national economy, since its products, such as paper, cardboard, paper pulp, cork, wood, resinous products and furniture represent 10 % of national export trade. It has a great impact in Portuguese GPD (Gross Domestic Product) which is the value of goods and services produced by labor and property in a region (PEFC, 2015).

With over 100 species, the genus *Pinus* is native to the northern hemisphere and distributed in Europe, Asia, North Africa, North America, and Central America (Keeley *et al.*, 2012). Belonging to the genus *Pinus, Pinus pinaster* is a species that naturally occurs in southwest Europe and west Mediterranean, mostly in coastal or near coastal areas (Figure 1) (Farjon & Filer, 2013).

P. pinaster is a resinous tree that when mature can reach 20-30 m presenting a thick cracked bark, dark purple outside and dark red inside, with persistent and very long needles grouped in sets of two or three with a basal sheath (Bajaj, 2013). This plant can occur in many environments. From areas at sea level to 2100 m of elevation, areas with more than 1400 mm of annual rainfall or with dry months, in acid, basic and sandy soils, and soils poor in nutrients, where other commercial species would not grow (Alía & Martín, 2003) and it is resistant to fire regimes (Keeley, 2012).

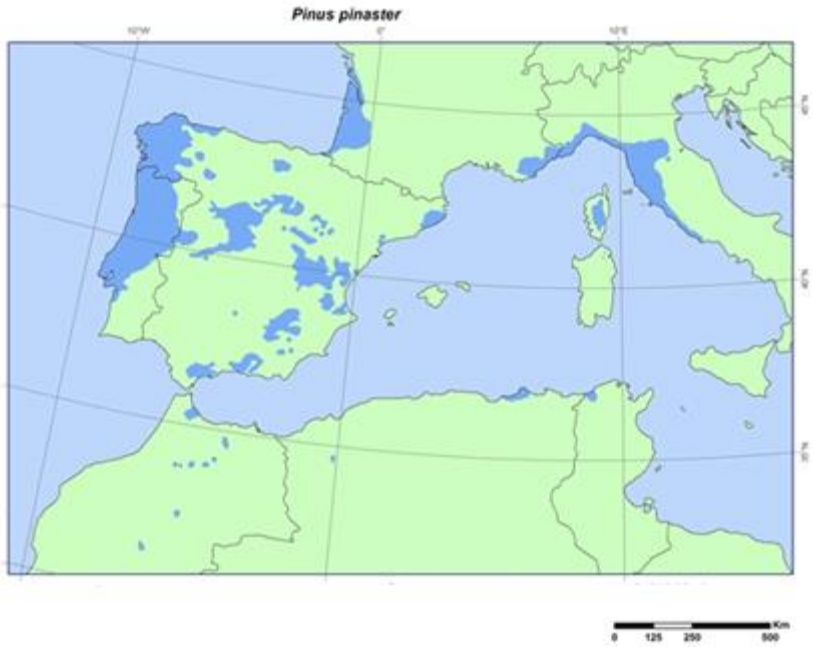

Figure 1 - Map of the natural and naturalized distribution of *Pinus pinaster*. Dark blue marks indicate the area where the species is distributed (*In* Alía & Martín 2003).

The maritime pine (*P. pinaster*) is the third tree species in forest occupation in the Portuguese territory with 714 thousand ha (23 % of the total area), after eucalyptus and cork tree. The *P. pinaster* area decreased between 1995 and 2010 in 263000 ha due to changes in land uses: transformation into bush land and pasture, increase urban area and reforestation with eucalyptus and other arboreal species (ICNF, 2013). Another cause for the decreasing of *P. pinaster* forest area is the pine wilt disease, which is affecting gravely the national wild pinewood (Figure 2) causing the cut of trees due to phytosanitary impositions (ICNF, 2013). The pine wilt disease (PWD) is a serious threat to conifers forest of the world and it has been causing serious environmental and economic losses (Nascimento *et al*., 2015).

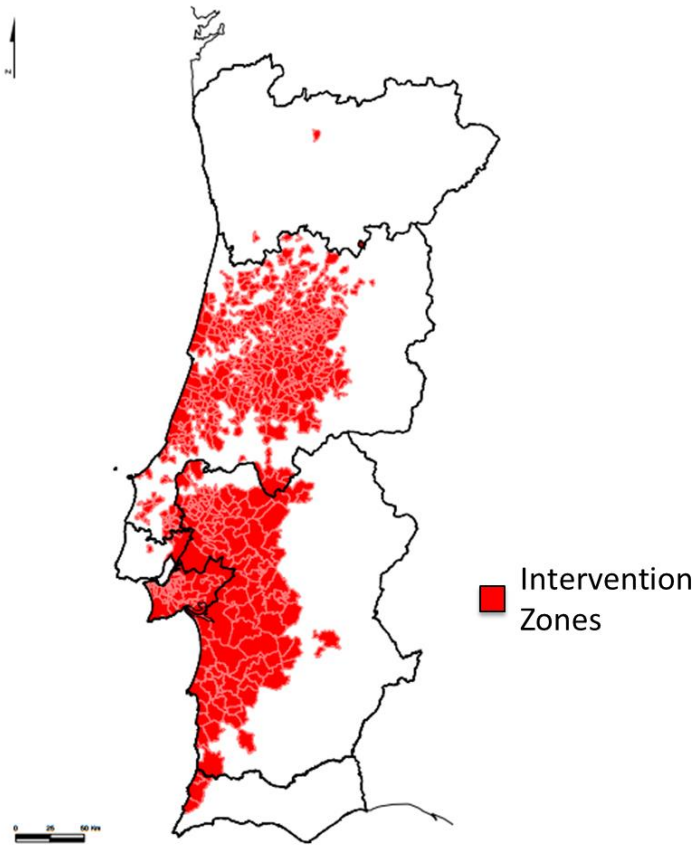

Figure 2 – Intervention Zones, areas where the presence of the PWD nematode is known or the nematode has an imminent risk of establishing itself and disperse, in continental Portugal (Adapted from Icnf.pt, 2015).

2. Microbiota of plants

Microorganisms are present in many environments and are able to of colonize different habitats, including living organisms. They interact with the host in many ways and, among those interactions, diseases are the most studied (Borer *et al*., 2013).

Plants are able to produce their own energy through photosynthesis, which makes them attractive to plant-associated heterotrophic microorganisms (pathogens or beneficial) as a nutrient source (Hardoim *et al*., 2008). Microorganisms provide host benefits that expand plants ability to adapt to the environment (Figure 3) *(*Bulgarelli *et al*., 2013).

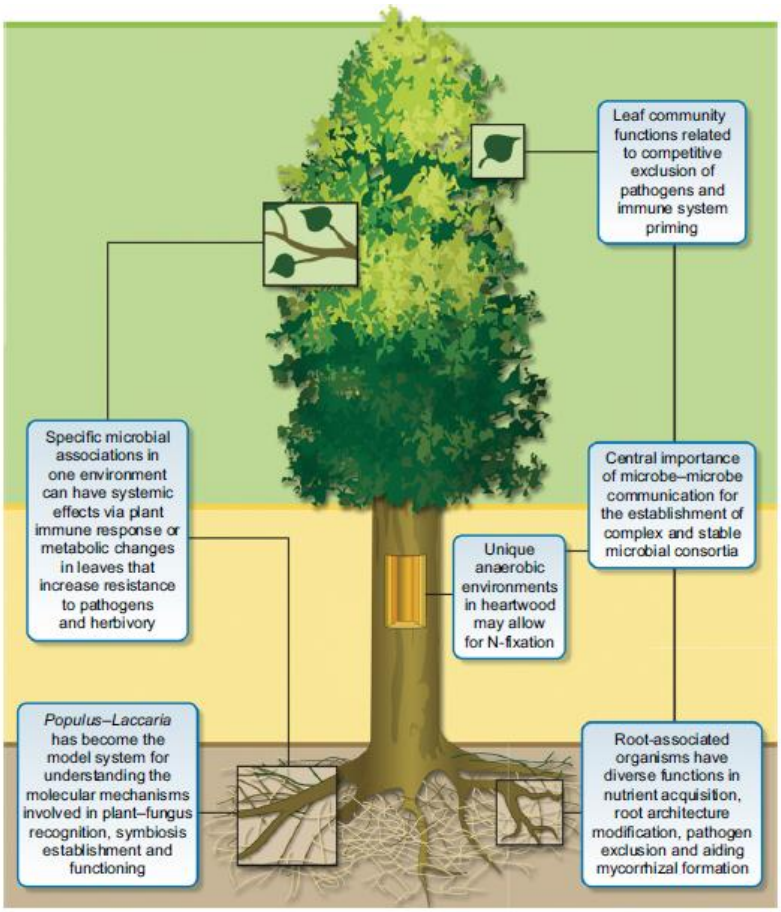

Figure 3 - Plant-associated microbes and their functions along the Poplar tree (*Populus sp.*) (*In* Schadt & Hacquard, 2015) .

The microbiota is able to provide many benefits to the plant such as nitrogen fixation (N), phosphorus (P) solubilization, siderophore production, volatile and antimicrobial compounds production and inducing systemic resistance *(*Bulgarelli *et al*., 2013). All of them very important in promoting plant growth.

2.1. Bacterial habitats in plants

The classification of the habitats in the plant varies depending on the author. Turner and collaborators (2013) use a definition of phyllosphere as plants aerial surfaces, rhizosphere as root surface and adjacent soil and endosphere as internal tissues, which corroborate with Izhaki and collaborators (2013). Some authors may consider phyllosphere as all aboveground tissues and rhizosphere all belowground tissues as seen in Bodenhausen and collaborators (2014) and Müller & Ruppel (2014). The habitants of

rhizosphere and phyllosphere living on surfaces are considered epiphytes, whereas inside all plant tissue are endophytes (Turner *et al*., 2013).

2.2. Factors that modulate bacterial community

The structure of the bacterial community varies along the plant. A study with species of coniferous and deciduous trees with different samples of plant parts presented variances in the bacterial diversity in belowground and aboveground internal tissues (Izumi *et al*., 2008). Population density of bacterial endophytes also seems to be higher on root and lower on stem, decreasing closer the apex (Mocali *et al*., 2003). This shows that the niche occupied by the microorganisms affects the structure of the community.

Several other factors can affect the structure of the bacterial community. With a synthetic bacterial community and mutants of *Arabidopsis thaliana,* Bodenhausen and collaborators (2014) examined the effect of host genotype on the phyllosphere community composition and abundance. These authors concluded that some mutations such as in the ethylene signalling and cutin biosynthesis, have great impact on bacterial community (Bodenhausen *et al*., 2014).

Like the host genotype, species may be a modulating factor of the community. Leaf surface bacterial community presented a variation when comparing plant species of the same location (Izhaki *et al*., 2013). The age of ginseng trees influences the variation of endophytes isolated from stems. This effect happens due the different maturation stages of the plant, the type and amount of root exudates in each stage, placing age as another factor that modulates bacterial community (Vendan *et al*., 2010).

The factors involved in the modulation of the community structure still need further research to elucidate how they contribute to the microbial community. Due to the divergence in concepts and methodologies, it is difficult to compare results. However, it is safe to say that the factors mentioned (Table 1) play a significant role in the bacterial community modulation.

Table 1– Factors involved in modulating the plant's bacterial community.

Factor	Plant species	Plant part	Method of analysis	Reference
Genotype	Mutants of *Arabidopsis thaliana*	Aerial tissues and surfaces	Culture-independent	Bodenhausen *et al*., 2014
Host specificity	*Nicotiana glauca* *Amygdalus communis* *Citrus paradisi*	Leaves surface	Culture-dependent and independent	Izhaki *et al*., 2013
Age	*Panax ginseng*	Stem	Culture-dependent	Vendan *et al*., 2010
Above and belowground	*Pinus sylvestris* *Betula penula* *Sorbus aucuparia*	Leaves, stems and roots	Culture-dependent and independent	Izumi *et al*., 2008

2.3. Host colonization

From the possible origins of endophytic bacteria, soil is the most accepted. Isolates of culturable diazotrophic and heterotrophic bacteria from *Pinus sylvestris* aerial parts were related to species commonly found in the soil, suggesting that bacteria are capable of systemic colonization of plants through the roots (Bal *et al*., 2012). The density of the endophytes is generally lower in the aerial parts compared to the roots and this may suggest an upward movement from soil (Turner *et al*., 2013).

Roots exudates and mucilage attracts the growth promoting bacteria from the soil (Santi *et al*., 2013) and cracks on root junctions may be the main entrance points for colonizing microorganisms (Hardoim *et al*., 2008). After penetration, microorganisms may colonize intercellular spaces of root cortex, reaching the xylem, and spreading to stems and leaves (Santi *et al.*, 2013).

Endophytic bacterial communities can have other origins than the rhizosphere, like the phyllosphere (Compant *et al*., 2005). Other entering points for colonization, such as wounds caused by pathogens or predators and the stomata (Hardoim *et al*., 2008) may represent alternatives to explain how microorganism enter the host.

After entering the plant, microorganisms activate the immune response of the host. The expression of defense genes by plants can differentiate between phytopathogens

and endophytic bacteria, and this genomic response is different depending on the host and bacterial genotype (Turner *et al*., 2013). In addition, a set of environmental and genetic bacterial factors will determine if bacteria are capable of becoming endophytic, since it is necessary to have the ability to adapt to the habitat change from the exterior to the endosphere (Hardoim *et al.,* 2008).

Once inside the plant, the microorganisms can settle in different types of habitats. Studies with root endosphere revealed that the endophytic bacteria live in the intercellular apoplast and in dead or dying cells and are often found in the xylem vessels (Turner *et al*., 2013). In woody plants such as *Populus*, many conifers and other forest trees, the saturated xylem and dead parenchyma cells promote an anaerobic environment that favors fermentation and methanogenesis, with numerous diazotrophics being isolated from this habitat (Hacquard & Schadt, 2015).

2.4. Bacterial phytopathogens

Plant nutrients are the target for phytopathogenic bacteria and its availability is an important factor in determining pathogens colonization of plants niches (Fatima & Senthil-Kumar, 2015). Bacterial pathogens have to modulate their metabolism according to the nutrients available, and some may secrete molecules that affect plant cells and enhance the availability of nutrients (Fatima & Senthil-Kumar, 2015) allowing the pathogen to successfully colonize plants.

Pathogens invade the plants, usually through injuries or natural openings, overcoming physical defense mechanisms such as wax layers, rigid cell walls and cuticular lipids (Muthamilarasan & Prasad, 2013). After penetration, the plant defense response divides into two interconnected mechanisms: microbial associated molecular pattern triggered-immunity and effector triggered-immunity (Muthamilarasan & Prasad, 2013) in an attempt to avoid invasion.

However, pathogens also have mechanisms to bypass host defense responses, such as production of effector proteins that are delivery inside the cytosol of the host cell via type III secretion system (TTSS) or production of phytotoxins such as coronatine, that reopens the stomata (Kim *et al.,* 2008).

Only nine genera were known to have plant pathogenic bacteria in 1978, however, due to the advances in DNA technology and classification methods there are now 39 genera that plant pathogenic bacteria can belong (Bull *et al*., 2014). Mansfield and collaborators (2012) surveyed plant pathologists about their opinion on the most scientifically/economically important bacterial pathogens and ranked them in a top 10 list. The genera *Pseudomonas, Ralstonia, Agrobacterium, Xanthomonas, Erwinia, Xylella, Dickeya* and *Pectobacterium* made the top 10 list (Table 2) and the species *Clavibacter michiganensis* (*michiganensis* and *sepedonicus*), *Pseudomonas savastanoi* and *Candidatus liberibacter* (pv *asiaticus*) recived honorable mentions for their scientifically and economical importance (Mansfield *et al*., 2012).

Table 2 – Rank of the top phytopatogenic bacteria. (Adapted from Mansfield *et al*., 2012).

Rank	Bacterial pathogen	Author of bacterial description	Importance
1	*Pseudomonas syringae* pathovars	John Mansfield	Scientifically Impact on the understanding of microbial pathogenicity and economically important plant diseases
2	*Ralstonia solanacearum*	Stéphane Genin	Economic importance worldwide, affecting a broad host range
3	*Agrobacterium tumefaciens*	Shimpei Magori, Vitaly Citovsky	Primarily for the role in scientific breakthroughs and applications
4	*Xanthomonas oryzae* pv. *oryzae*	Malinee Sriariyanum, Pamela Ronald	Economically important pathology because of the host target
5	*Xanthomonas campestris* pathovars	Max Dow	Cause disease in a range of crops worldwide
6	*Xanthomonas axonopodis* pv. *manihotis*	Valérie Verdier	Economically important due to the host target
7	*Erwinia amylovora*	Steven V. Beer	Important in scientific history and economy causing disease in ornamental and fruit trees and bushes
8	*Xylella fastidiosa*	Marcos A. Machado	Related to several important crop and trees disease, being the first bacterial phytopathogen to have its genome sequenced
9	*Dickeya* (dadantii and solani)	Ian Toth	Economically important diseases
10	*Pectobacterium carotovorum* (and *P. atrosepticum*)	George Salmond	Economically important, scientific milestones and involvement in treatment of some leukaemias

2.5. Beneficial bacteria

Endophytes are possessors of some plant growth promoting characteristics, such as hormone production, phosphate-solubilization and nitrogen fixation (Turner *et al*., 2013). Endophytic bacteria isolated from Ginseng (*Panax giseng*) stems were screened for plant growth promotion traits such as N_2 fixation, phosphorous solubilization, IAA

production and siderophore secretion, and although isolates exhibiting all those traits were rare, most of them were positive for at least one (Vendan *et al*., 2010).

Studies with ^{15}N showed that most foliar N was actively fixed in the wetwood (saturated xylem tissues) (Hacquard & Schadt, 2015). The N fixation appears to be influencieted by bacterial strains, plant genotype, growth stage, and environmental conditions (Santi *et al*., 2013). Endophytic diazotrophs habitat provides more appropriated conditions for nitrogen fixation and distribution on the host plant (Santi *et al*., 2013). *Azoarcus* spp., *Herbaspirillum seropedicae* and *Glucenobacter* are diazotrophic rhizobacteria recognized as endophytes (Santi *et al.*, 2013).

In plant-bacteria interaction, secondary metabolites can be used for nutrient uptake, modulation of plant hormones and stress tolerance (Brader *et al.*, 2014). There are evidences that endophytic bacteria have a high potential to produce a range of metabolites with pharmaceutical interest, such as multicyclic indolesesquiterpenes from *Streptomyces* spp., endophytes from mangrove tree (Brader *et al.*, 2014).

The manipulation of the plant microbiome has potential to increase production and reduce the incidence of plant disease, therefore decreasing the utilization of chemicals, resulting in an environmental friendly agriculture (Turner *et al*., 2013).

2.6. Endophytic bacteria in *Pinus*

There are few studies about the microbiota of coniferous trees and most of them focus on the isolation or/and application of diazothrophic bacteria. Studies that focus on the complete microbial community are rare, especially for the *Pinus* genus.

As mentioned before, bacterial community can be affected by several factors. Analysis of *Pinus flexilis* and *Pinus engelmannii* needles using pyrosequencing show a low intra and inter individual variability in the structure of endophytic community structure (Carrel & Frank, 2014). These authors believed that the consistence found in the conifer needles is a result of a consistent abiotic factor and reflects the bacterial ability to survive in the conifers needles or a relevant conifer-bacteria partnership.

Izumi and collaborators (2008) used DGGE to profile the community of *Pinus sylvestris* and found differences between endosphere and phyllosphere in the same individual. There is only a few studies on *Pinus,* the methodology, the plant species

studied and the plant tissue sampled in the studies mentioned here are different making hard to make comparisons.

In the needles of *P. flexilis* and *P. engelmannii* was found a set of bacteria belonging to Proteobacteria and Acidobacteria was found, and both species shared the dominant phylotype, *Gluconacetobacter diazotrophicus* and *Gluconacetobacter liquefaciens* (Carrel & Frank, 2014). The high relative abundance of few phylotypes may be relate to the significant association of these bacteria and conifers, especially from the family *Acetobacteraceae*, which have species with documented functions on N fixation, phytohormone production and pathogen antagonism (Carrel & Frank, 2014).

3. *Pine wilt disease*

PWD is an important disease caused by the nematode *Bursaphelenchus xylophilus*. The infection cycle of PWD depends on three organisms: the pine (host), the nematode (pathogen), and *Monochamus* beetles (vector).

Native from North America, the PWD spread through Asia (Japan, China, Taiwan, and Korea) and more recently in Europe (Nascimento *et al.*, 2015). This disease was first detected in Portugal in 1999 in a *P. pinaster* tree (Mota *et al*., 1999). Until recently *P. pinaster* was the only species associated with PWD in Portugal, but now it has also been reported in *P. nigra* (Inácio *et al*., 2014). This disease constitutes a major economic and environmental threat for the forests ecosystems worldwide (Vicente *et al*., 2012).

The disease affects mainly trees of the *Pinus* genus but other conifers can also act as hosts (primarily *Larix*, *Abies* and *Picea*). Not all *Pinus* species are susceptible to the PWD. The species *P. bungeana*, *P. densiflora*, *P. luchuensis*, *P. massoniana* and *P.thunbergii*, and the European species: *P. nigra*, *P. sylvestris* and *P. pinaster*, are the only ones known to die by the PWD as mature trees in the field (EPPO, 2014).

The beetles from the *Monochamus* genus are the vector for the PWD nematode worldwide. *Monochamus alternatus* was reported to be the most effective vector in Japan, *Monochamus carolinensis* is the primary vector in North America and *Monochamus galloprovincialis* in Europe (Akbulut & Stamps, 2012).

Monochamus beetles feeding wounds and oviposition provide entry portals for the nematode who feeds on the epithelial cells that line the resin channels consequently disturbing the sap flow and causing withering of the tree (Figure 4) (Futai, 2013). The nematodes that develop in the tree infect the pulp of the *Monochamus* beetles and when adults emerge they are already infected with the nematode and able to disperse the disease to healthy trees (Nascimento *et al.,* 2015).

There still is debate in the role of bacteria in the development of PWD. Mostly cultivable methods have been applied to unveil the bacterial community (Nascimento *et al*., 2015). It is known that there are bacteria present in the cuticle surface of the nematode and an accumulation of bacteria on the nematode while inside the infected tree (Zhao *et al*., 2014). The bacterial community may be involved in the development of the PWD, with the production of enzymes to digest cellulose, xenobiotic detoxification and protection to oxidative stress (Nascimento *et al*., 2015). Nevertheless, the precise role of those bacteria is still obscure.

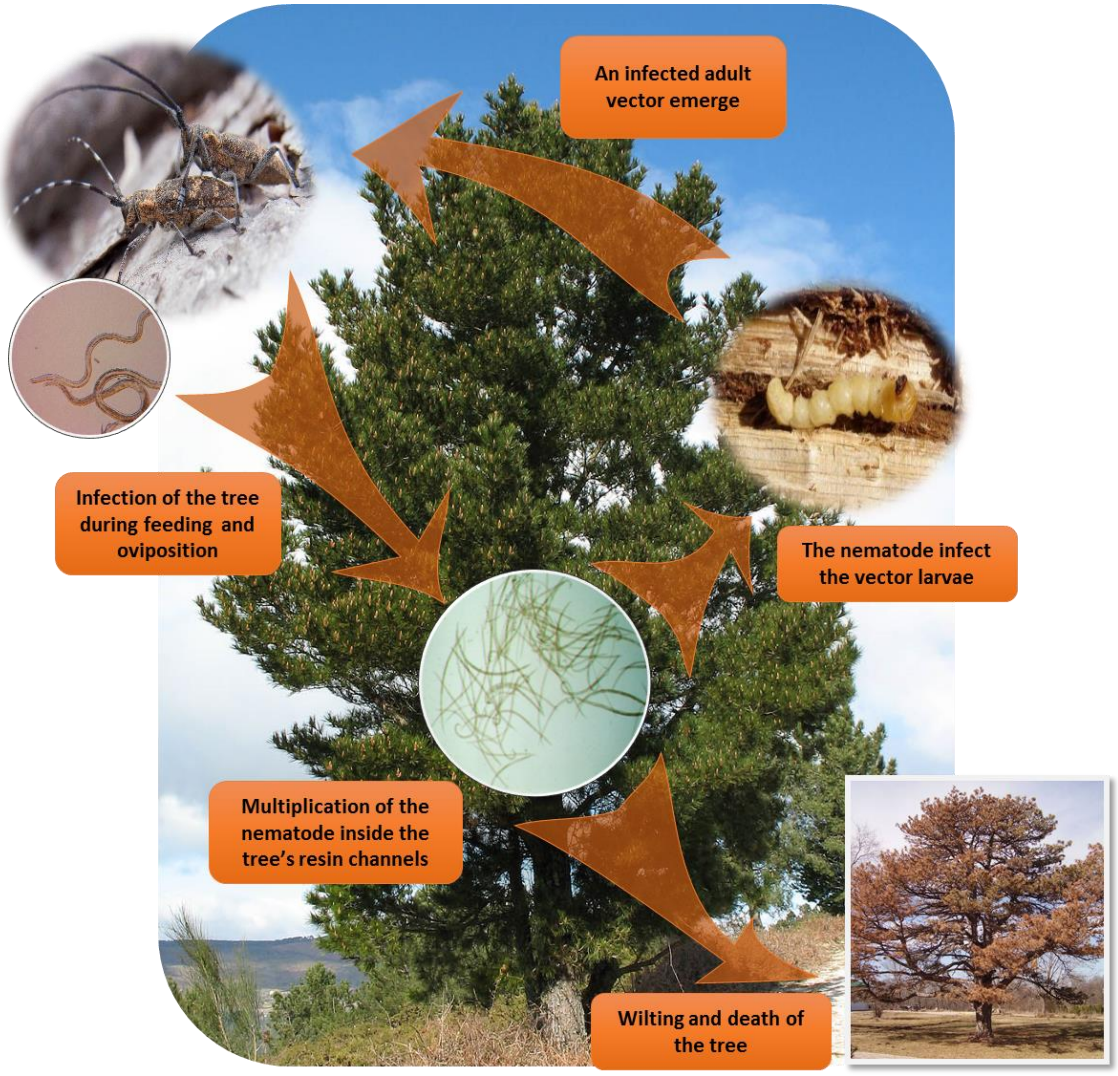

Figure 4 – Pine wilt disease cycle. The vector transmit the nematode to the tree during feeding and oviposition. Inside the host the nematode feed on the cells of resin canals and reproduce. The nematode migrates along the tree through resin canals, which is also colonized by bacteria. The insect vector larvae is infected by the nematode and when the adult insect emerge, it already carries the nematode, being able to spread the disease.

4. *Objectives*

Considering the economic and ecological importance of *Pinus pinaster* and the implications of the dispersal PWD, it is imperative to perform studies that elucidate the mechanisms implicated in the PWD nematode infection. Some authors show that the bacterial community may have implications in the development of the disease, but the precise role of these microorganisms is not yet enlightened.

This work is embedded in a bigger project named MicroNema, which aims to elucidate the interactions of bacterial communities and the PWD through advance technology. This project aims to answer some questions about the role of the bacterial

community and the PWD that still remains: Which differences exists in the bacterial community between healthy and symptomatic trees? Do the vector, host and agent share common bacteria? Are the bacteria associated with the nematode and the vector present in healthy trees? It is out of the scope of the present work to answer those questions but this study gathered crucial knowledge of the healthy trees, especially answering the question about what is the microbiome of healthy *P. pinaster*, enabling the possibility of comparison in further studies.

Since the bacterial community has an important role in the improvement of plant growth and protection among other characteristics, the study of the tree microbiome may reveal important species of bacteria with biotechnological and environmental application. Accordingly, the main objective of the present work is to determine the bacterial community of health *P. pinaster* from different locations of Portugal.

MATERIAL AND METHODS

1. Sampling

1.1. Sampling sites

Samples were taken from four sites: Comporta, Góis, Madeira and Vouzela (Figure 5). All of these sites are in the ICNF list of intervention locations, which means that the PWD nematode is present or in eminent risk of stablishing itself, with Vouzela entering the list after the time it was sampled (ICNF, 2015).

Herdade da Comporta, Comporta (38º22'48.67"N / 8º47'25.00"W) belongs to the Setubal district in the municipality of Alcácer do Sal and it is located in the base of Tróia Peninsula on the southern margin of the Sado River. It has an elevation of 24m (79ft) and a warm and temperate climate. Annual rainfall average is 614 mm with rainfall higher in the winter than in summer. Annual average temperature is 17.2 °C (Pt.climate-data.org, 2015).

Serra da Lousã, Góis (40º09'07.3"N 8º07'34.1"W) belongs to the Coimbra district in the center of continental Portugal. Góis has a warm and temperate climate; the annual average temperature is 15.7 °C and annual rainfall of 958 mm (Pt.climate-data.org, 2015).

Vouzela (40°38'02.4"N 8°11'25.4"W) belongs to the district of Viseu on the center of continental Portugal. With a warm and temperate climate, the annual average temperature is 14.5 °C and average rainfall of 1112 mm (Pt.climate-data.org, 2015).

Prazeres is located in the Madeira Island (32º45'45.5"N 17º11'47.9"W). Madeira is an island on the southwest of the Portuguese seaboard. Prazeres has a warm and temperate climate with the annual temperature of 15.8 °C and 665 mm of rainfall (Pt.climate-data.org, 2015).

Figure 5 – Map of Portugal with numbered red dots marking the sampling locations. 1 – Madeira; 2 – Vouzela; 3 – Góis; 4 – Comporta.

1.2. Sampling methodology

A total of 24 healthy trees (no symptom of wilt), 6 for each site, with approximately 20 to 30 years old were collected in the season of spring, with the exception of Vouzela samples that were collected in October. Trees were felled and their trunk cut into disks. Two disks from the midsection of each tree were processed in the laboratory. The disks were used to obtain sawdust for the DNA extraction.

Sawdust from the wood disk was obtained using a drill. The drill was sterilized with ethanol (90 %) between each sample. The disk was drilled in several points into its length to ensure that the bacterial community from the trunk was well represented. The sawdust obtained was stored in plastic bags and frozen in -20°C until the DNA extraction.

2. Genomic DNA extraction

The total genomic DNA from the sawdust samples was extracted using the PowerSoil® DNA Isolation Kit (MO BIO laboratories, inc., CA, USA) following the manufacturer's instructions (Table 3) with some alterations (underlined).

Table 3 – Sawdust DNA extraction protocol.

DNA Extraction protocol
1. 0.13 g of sawdust and 200 µl of TE buffer (100 mM Tris/ 10 mM ETDA) were added to the PowerBead tube provided. 2. The tube was gently vortex and 60 µl of **Solution C1** was added, the tubes were then vortex on maximum speed for 10 minutes on a tube holder adapter for vortexes. 3. PowerBead tubes were centrifuged at 10.000 x g for 30 seconds at room temperature. 4. The supernatant was transferred to the 2 ml Collection Tube and 250 µl of **Solution C2** was added. 5. The tubes were vortex for 5 seconds and incubated at 4°C for 5 minutes. The tubes were centrifuged at room temperature for 1 minute at 10,000 x g. 6. Up to 600 µl of the supernatant was transferred to a clean 2 ml Collection Tube and 200 µl of **Solution C3** was added. 7. The tubes were vortex briefly and incubated at 4°C for 5 minutes. After that, tubes were centrifuged at room temperature for 1 minute at 10,000 x g. 8. Up to 750 µl of supernatant was transferred to a clean 2 ml Collection Tube and 1200 µl of **Solution C4** was added and vortex for 5 seconds. 9. Approximately 675 µl of the mix was loaded onto a spin filter and centrifuged for 1 minute at 10,000 x g. The flow through was discarded. The procedure was repeated until the mix was finished. 10. 500 µl of **Solution C5** was added and the tubes were centrifuged for 30 seconds at 10,000 x g and the flow through was discarded. The tubes were centrifuged again for 1 minute at 10,000 x g. 11. The spin filter was placed in a clean 2 ml Collection Tube and 50 µl of **Solution C6** was added to the center of the filter membrane. 12. After centrifuging for 30 seconds at 10,000 x g the spin filter was discarded and the DNA was stored in -20°C.

3. *Pine wilt disease nematode molecular detection*

The PCR reaction mixture for the molecular detection of *B. xylophilus* contained 12.5 µl NZYTaq 2x Green Master Mix (NZYTech, Lisboa, Portugal), 10 µl of ultrapure water and 0.75 µl of each of the primers. The primers used were specific for the pine wilt nematode. The final volume was 24 µl to which 1 µl of the sample was added. As for positive control for this reaction DNA extracted from the pine wilt nematode was used. Sterile dH_2O replaced DNA template for the negative control.

The procedure was performed as described in Cardoso and collaborators (2012). The amplification program consisted in an initial denaturation at 94°C for 5 minutes, followed by 15 amplification cycles of denaturation at 94°C for 30 seconds, hybridization at 49°C for 1 minute and extension 72°C for 2 minutes, and a final extension step of 72°C for 5 minutes. The resulting PCR products (5 µl) were analysed in a 1.5 % agarose gel. The Gene Ruler DNA ladder mix (2 µl) (Thermo Scientific- Fermentas, Burlington, Canada) was loaded in the first and last wells of the gel for reference. After the run at 80 V for 1 hour and 10 minutes, the gel was placed in ethidium bromide (0.5 µg/ml) for 10 minutes and subsequently in water for 15 minutes. Molecular Imager® Gel Doc™ XR+ System (Bio-Rad Laboratories, Hercules, California, USA) was used for visualization under UV light.

Negative samples or with faint bands were re-amplified. The product of the first PCR was diluted 1:10 and amplified in the same conditions as the first PCR.

4. *Denaturing Gradient Gel Electrophoresis (DGGE) analysis*

DGGE analysis was conducted targeting the V3 region of the 16S rRNA gene. In order to increase sensitivity, a nested PCR technique was applied.

For the first PCR reaction, the universal primer: 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′) were used. The reaction mixture for this PCR contained 6.25 µl NZYTaq 2x Green Master Mix (NZYTech, Lisboa, Portugal) (0.2 U/µL Taq DNA polymerase, 200 µM of dNTPs 2.5 mM, MgCl2), 16.25 µl of ultrapure water and 0.75 µl of each primer. The final volume was 24 µl to which was added 1 µl of the sample. DNA from *Eschericha coli* ATCC 25922, was used for positive control, sterile dH_2O replacing DNA template was used as negative control.

The program for amplification consisted in an initial denaturation at 94°C for 3 minutes, followed by 30 amplification cycles of denaturation at 94°C for 1 minute, hybridization at 52°C for 1 minute and extension 72°C for 2 minutes, and a final extension step of 72°C for 10 minutes.

The resulting PCR products (5 µl) were analysed by electrophoresis on a 1 % agarose gel. Gene Ruler DNA ladder mix (2 µl) (Thermo Scientific- Fermentas, Burlington, Canada) was loaded in the first and last well for reference ladder. After the run at 80 V for 1 hour and 20 minutes the gel was placed in ethidium bromide 0.5 µg/ml for 10 minutes and then washed in distilled water for 15 minutes. The visualization was made using the Molecular Imager® Gel Doc™ XR+ System (Bio-Rad Laboratories, Hercules, California, USA) under UV light.

The second PCR targeted the V3 region of the 16S rRNA gene using the product of the first PCR as a template. In this reaction the primers used were 338F (5′-GACTCCTACGGGAGGCAGCAG-3′) and 518R (5′-ATTACCGCGGCTGCTGG-3′) with a GC clamp attached to the forward primer.

The reaction mixture for this second PCR contained 6.25 µl NZYTaq 2x Green Master Mix (NZYTech, Lisboa, Portugal) (0.2 U/µL Taq DNA polymerase, 200 µM of dNTPs 2.5 mM, MgCl2), 16.25 µl of ultrapure water and 0.75 µl of each primer. The final volume was 24 µl to which was added 1 µl of the first PCR product. DNA from a control strain, *Escherichia coli* ATCC 25922, was used for positive control and a tube with sterile dH_2O replacing DNA template was used as negative control.

The program for amplification on this second PCR was consisted in an initial denaturation at 94°C for 5 minutes, followed by 35 amplification cycles of denaturation at 92°C for 30 seconds, hybridization at 55°C for 30 seconds and extension 72°C for 30 seconds, and a final extension step of 72°C for 30 minutes.

The resulting products (5 µl) from the second PCR were analysed by electrophoresis on a 1 % agarose gel. Gene Ruler DNA ladder mix (2 µl) (Thermo Scientific- Fermentas, Burlington, Canada) was loaded in the first and last well for reference ladder. After the run at 80 V for 1 hour and 20 minutes the gel was placed in ethidium bromide 0.5 µg/ml for 10 minutes and then washed in distilled water for 15 minutes. The visualization was made using the Molecular Imager® Gel Doc™ XR+ System (Bio-Rad Laboratories, Hercules, California, USA) under UV light.

PCR products were loaded into a 8 % polyacrylamide (37.5:1, acrylamide/bisacrylamide) gel with linear denaturing gradient ranging from 35 % to 60 % (100 % corresponds to 7 M Urea and 40 % formamide). Two lanes were loaded with a DGGE maker for internal normalization and as an indication of the quality of the analysis (Henriques *et al*., 2006).

Electrophoresis was conducted in a DCode System (Bio-Rad) at 20 V for 15 minutes and 75 V for 16 hours at the temperature of 60 °C. The gel was place in ethidium bromide (0.5 µg/ml) for 10 minutes and then washed in distilled water for 15 minutes. The visualization was made using the Molecular Imager® Gel Doc^{TM} XR+ System (Bio-Rad Laboratories, Hercules, California, USA) under UV light.

DGGE patterns were analysed using Bionumerics Software (Applied Maths, Belgium). Cluster analysis of DGGE profiles was performed using the UPGMA method (Unweighted Pair Group Method with Arithmetic mean) applying Pearson correlation coefficient.

5. *Cloning and sequencing for DGGE bands phylogenetic affiliation*

The DGGE profiles of the samples looked very similar, therefore the sample C1 was selected for the cloning and sequencing test. To obtain a PCR product for cloning, a nested PCR was conducted with the same conditions as the nested PCR for the DGGE.

TA Cloning® Kit (Invitrogen, California, USA) was used according to the manufacturer's instruction to transform competent *Escherichia Coli* TOP10F´cells (Invitrogen, California, USA) using pCR®2.1 vectors. From the transformation culture, 60 positive clones (white colonies) were cultivated again to make sure that they were true positive and 36 were selected for further analysis.

Cells from each clone were resuspended in 20 µl of dH_2O and used as template in PCR reactions. A PCR was conducted using a set of primers for the insert, T7 (5'-TAATACGACTCACTATAGGG-3') as the forward primer and M13 (5'-CAGGAAACAGCTATGAC-3') as the reverse. Reaction mixture for this PCR contained 6.25 µl NZYTaq 2x Green Master Mix (NZYTech, Lisboa, Portugal) (0.2 U/µL Taq DNA polymerase, 200 µM of dNTPs 2.5 mM, MgCl2), 14.25 µl of ultrapure water and 0.75 µl of each primer. The final volume was 22 µl to which was added 3 µl of the sample (cell suspension). DNA from a blue colony was used for positive control and a tube with sterile dH_2O replacing DNA template was used as negative control. Amplification program

consisted in an initial denaturation at 94°C for 5 minutes, followed by 30 amplification cycles of denaturation at 94°C for 30 seconds, hybridization at 55°C for 30 seconds and extension 72°C for 1 minute, and a final extension step of 72°C for 5 minutes.

Samples were selected by analysing the PCR products by electrophoresis on a 1 % agarose gel. Clone samples that resulted in a PCR product with double bands or bands too big or too small compared to the expected band size (410 bp) were excluded.

Subsequently, a DGGE analysis was performed to the selected clones through amplification of the 16S rDNA V3 region (with the 338F_GC and 518R primer pair) using the conditions described previously for the DGGE.

The reaction mixture contained 6.25 µl NZYTaq 2x Green Master Mix (NZYTech, Lisboa, Portugal) (0.2 U/µL Taq DNA polymerase, 200 µM of dNTPs 2.5 mM, MgCl2), 16.25 µl of ultrapure water and 0.75 µl of each primer. To the reaction mixture 1µl of the first PCR product was added as template. DGGE proceeded using the same parameters mentioned in the DGGE analysis.

The profiles from the clone samples were compared with the profile of the donor sample for the selection of representative number of clones with the expected bands (the most intense ones). Based on the DGGE profile 6 samples were sent for sequencing at GATC Biotech (Cologne, Germany). Samples were analysed using the online tool NCBI BLAST: Basic Local Alignment Search Tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

6. *Barcoded 454 pyrosequencing*

Before sending the samples for pyrosequencing a PCR using the universal set of primers: 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′) was performed. The reaction mixture contained 6.25 µl NZYTaq 2x Green Master Mix (NZYTech, Lisboa, Portugal) (0.2 U/µL Taq DNA polymerase, 200 µM of dNTPs 2.5 mM, $MgCl_2$), 16.25 µl of ultrapure water and 0.75 µl of each primer. The final volume was 24 µl to which was added 1 µl of the sample. DNA from a control strain, *Eschericha coli* ATCC 25922, was used for positive control and a tube with sterile dH_2O replacing DNA template was used as negative control.

The program for amplification consisted in an initial denaturation at 94°C for 3 minutes, followed by 30 amplification cycles of denaturation at 94°C for 1 minute,

hybridization at 52°C for 1 minute and extension 72°C for 2 minutes, and a final extension step of 72°C for 10 minutes.

Samples were prepared for 454 pyrosequencing by PCR amplification of the V3-4 hypervariable region with fusion primers containing the Roche-454 Titanium sequencing adapters, an eight-base barcode sequencing in fusion primer A, the forward primer 5'- ACTCCTACGGGAGGCAG -3' and the reverse primer 5'- TACNVRRGTHTCTAATYC -3' (Wang & Qian, 2009). The PCR reaction occurred in 40 µl reaction with Advantage Taq (Clontech) using 0.2 µM of each primer, 0.2 mM dNTPs, 1X polymerase mix and 6 % of DMSO. The amplification program was 94°C for 4 minutes for initial denaturation, followed by 25 cycles of 94°C for 30 seconds, 44°C for 45 seconds and 68°C for 1 minute, and with a final elongation step at 68°C for 10 minutes. The amplicons were quantified bv fluorimetry with PicoGreen (Invitrogen, CA, USA), pooled in equimolar concentrations and sequencing in the direction with GS 454 FLX Titanium chemistry, according to manufacturer's instruction (Roche, 454 Life Sciences, Brandford, CT, USA) at Biocant (Cantanhede, Portugal). Each DNA sequence was afterwards traced back to its original sample through barcode analysis.

Sequences were processed using both UPARSE (Edgar, 2013) and QIIME (Caporaso *et al*., 2011) pipelines on a computer using the Linux© operating system.

In the UPARSE workflow, barcodes were striped and reads were quality filtered to a maximum expected error of 1.0, trimmed to 350 bp, dereplicated (identical reads were merged) and singletons were discarded. Operational Taxonomic Units (OTUs) were defined at 97 % similarity using the UPARSE-OTU algorithm that simultaneously identifies and discards chimeras. Taxonomy assignment was made through QIIME using Uclust as assignment method and Greengenes reference databases.

Richness index (S), Shannon index of diversity (H) (Shannon, 1948) and the equitability index (E) (Pielou, 1075) were calculated using PRIMER software (Anderson *et al*., 2008) for each sample as follows:

$$H=-\sum(n_i/N)\log(n_i/N),$$

$$E=H/\log S,$$

n_i is the OTU abundance, S is the number OTUs (used to indicate the number of species) and N is the sum of all reads for a given sample (used as estimates of species abundance).

7. Statistical Analysis

Statistical analysis was used to determine if the sampling location had significant influence ($P<0.05$) in the composition of bacterial communities. Statistical significance factor was evaluated through PERMANOVA based on 9999 permutations using PRIMER v6 software. PERMANOVA was performed on Bray-Curtis distance matrixes constructed from the abundance tables. Square root transformation was previously applied to each abundance table (the OTU abundance table samples was rarefied to the lowest number of reads obtained in the samples).

RESULTS

None of the 24 trees sampled for this study were positive for the molecular detection of the PWD nematode.

1. *DGGE analysis of asymptomatic tree samples*

The DGGE profiles obtained to compare the bacterial community of the *P. pinaster* trees are presented in figure 6. It is possible to observe that the profiles were in general similar to each other.

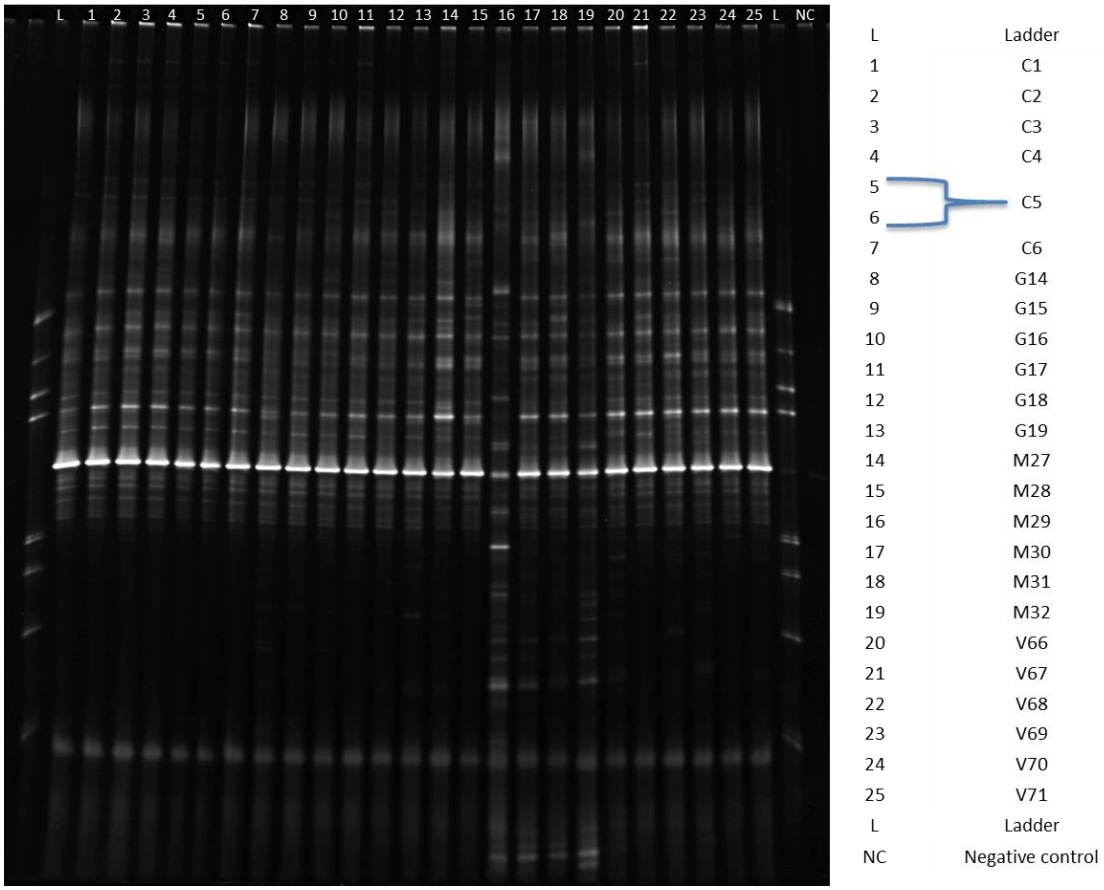

Figure 6 - DGGE profiles from the midsection trunk of *P. pinaster* tree from the four locations. Each lane of the DGGE gel represents one tree sample named with a letter corresponding to the location (C – Comporta, G – Góis, M – Madeira and V – Vouzela) and numbered according storage order at the lab. Two lanes were loaded with a DGGE maker for internal normalization and as an indication of the quality of the analysis (Henriques *et al*., 2006).

To observe the relationship between samples a dendrogram was created (Figure 7) using the Bionumerics software. The data was transformed into a similarity matrix, and UPGMA method clustered the samples according to their fingerprint profile.

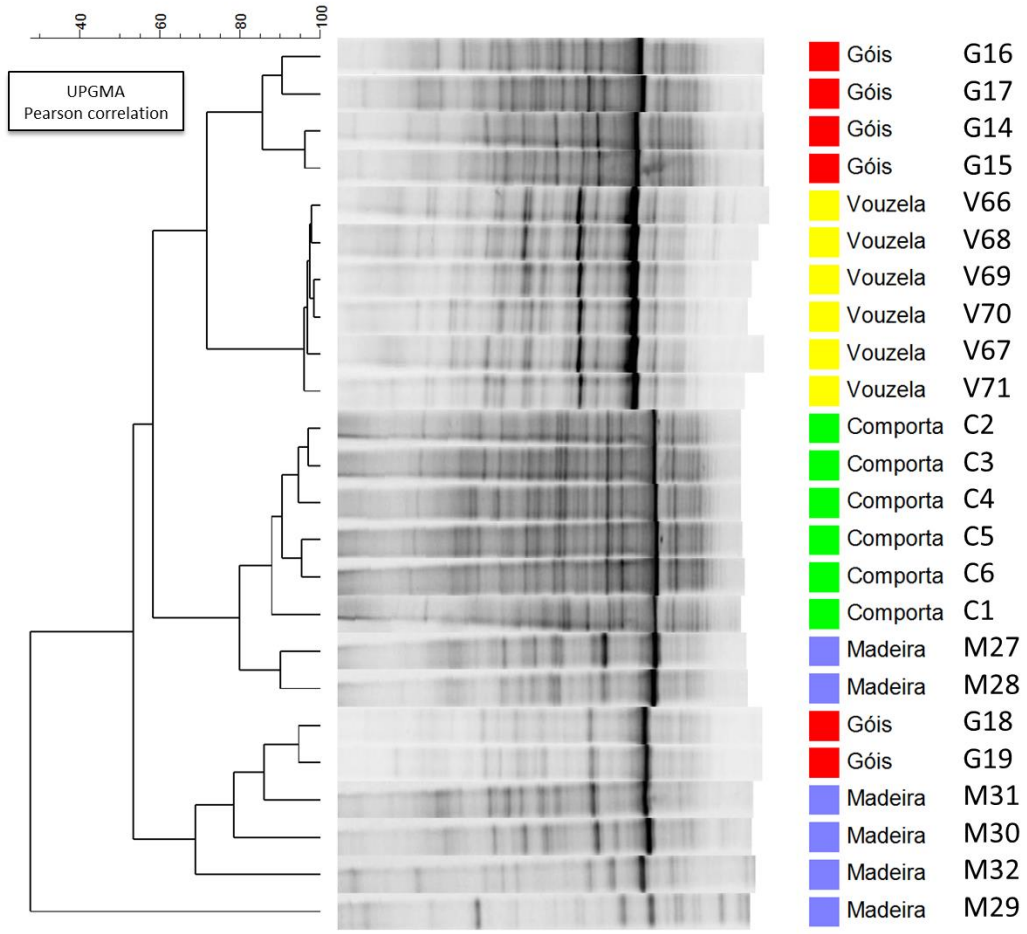

Figure 7 – Dendrogram of the DGGE profiles of the V3 region of the 16S rRNA gene from tree samples collected in the 4 locations (C – Comporta, G – Góis, M – Madeira and V – Vouzela). Patterns were created with UPGMA method and Pearson correlation.

It is possible to observe in the dendrogram that most samples cluster by location, with some exceptions. M29, a sample from Madeira, differ from the other Madeira samples, having less than 30 % of similarity with all samples. Two Góis samples clustered with Madeira samples having similarity over 90 % and two Madeira samples clustered with Comporta samples having a little more than 80 % of similarity. In this analysis, Vouzela and Góis clustered together and have more than 70 % of similarity.

PERMANOVA was conducted with Primer v6 software using Bray Curtis similarity matrix from the DGGE abundance table, and the samples arranged by location. The result

was p=0.0001 ($p<0.05$) showing that the location significantly influences the bacterial community of the trees.

2. *Cloning and sequencing*

Due to the presence of intense bands in all samples, cloning and sequencing techniques were used to determine the phylogenetic affiliation of dominant bands in DGGE profiles. The bands profile on the DGGE appeared to have very similar prominent bands that were present in all samples lanes. Therefore, tree sample C1 was chosen as the donor for the production of the clones. Positive clones (36) were used for the first PCR (Figure 8–A). The second PCR (Figure 8– B) for the DGGE analysis used 29 clones since 7 clones from the previous PCR were discarded for not having the expected band size.

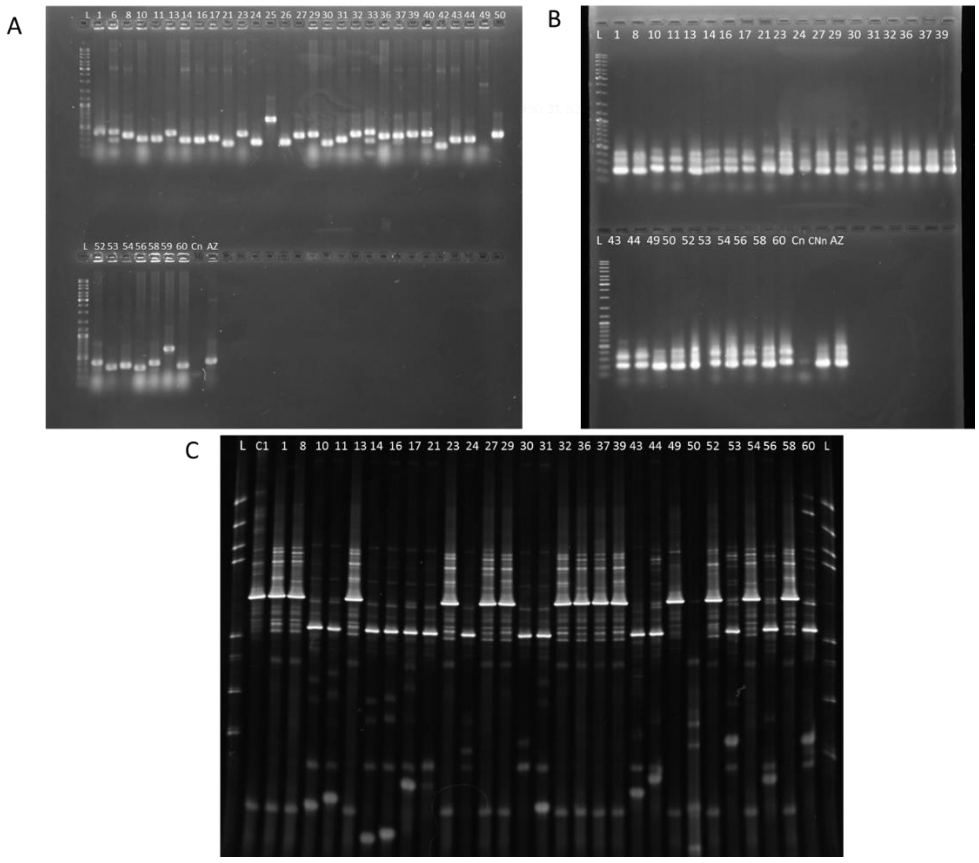

Figure 8 – A and B are agarose gels showing PCR products obtained by electrophoresis and C a DGGE gel. Numbers 1 to 60 are the clone samples, L – Ladder, Cn– Negative control, CNn – Nested negative control, AZ – Blue colony sample (positive control) and C1 – Comporta sample (Donor sample). A) First PCR of the clones with the plasmid primers (T7F/13MR). The ladder used was Gene Ruler DNA ladder mix (Thermo Scientific- Fermentas, Burlington, Canada). B) Second PCR of the clones for DGGE analysis using primers for the V3 region of the16S rRNA gene with a GC clamp (338F – GC/518R). The ladder used was Gene Ruler DNA ladder mix (Thermo Scientific- Fermentas, Burlington, Canada). C) DGGE of the clones with each lane representing a clone sample. Two lanes were loaded with a DGGE marker for internal normalization and as an indication of the quality of the analysis (Henriques *et al.*, 2006).

The analysis of the DGGE gel (Figure 8–C) showed some similar profiles, from those, 6 samples were selected for sequencing. Good quality sequences were analysed using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). All samples aligned with the *P. pinaster* chloroplast sequences with similarity ranging from 99 % to 100 % (Table 4).

Table 4 – BLAST results for the sequencing of selected clones.

Sample		Accession number	Similarity (%)
1	*Pinus pinaster* chloroplast, partial genome	FJ899583.2	100
36	*Pinus pinaster* chloroplast, partial genome	FJ899583.2	100
17	*Pinus pinaster* chloroplast, partial genome	FJ899583.2	100
43	*Pinus pinaster* chloroplast, partial genome	FJ899583.2	99

3. Analysis of pyrosequencing data

Seven samples, at least one per location, were selected for pyrosequencing based on the DGGE analysis: one for Comporta and two for the other locations (Góis, Madeira and Vouzela). All of these samples were negative in the molecular screening for *B. xylophilus*.

UPARSE pipeline was used for the quality treatment and removal of chimeras and singletons: barcodes were striped and reads were quality filtered to a maximum expected error of 1.0, trimmed to 350 bp, dereplicated (identical reads were merged) and singletons and chimeras were discarded. This process resulted in 44020 reads that were clustered into 158 operational taxonomic units (OTUs) with a 97 % of similarity match. This percentage is traditionally considered adequate for species assignment.

QIIME pipeline was used to taxonomically assign an identification to the OTUs. From 158 OTUs, 70 were assigned as chloroplast and were excluded leaving 88 OTUs with 4373 reads. The final number of reads from each sample differs greatly between locations and between samples from the same location (Table 5), ranging from 48 to 2927. The number of OTUs recovered ranged from 12 to 52 among samples.

Table 5 – Total sequences and OTUs obtained for all samples after quality treatment and the removal of chloroplast-affiliated reads. OTUs were defined with a 97 % of similarity.

Sampling location	Samples ID	Number of Sequences	Number of OTUs
Comporta	C1	48	13
Góis	G16	165	12
	G19	563	14
Madeira	M27	48	18
	M32	2927	52
Vouzela	V66	426	40
	V68	165	35

Rarefaction curves did not reach an asymptote for all the samples, suggesting that the community associated with some of the samples may be richer and were not fully characterized (Figure 9). The samples G16 and G19 from Góis, V66 from Vouzela and M32 from Madeira, were the only ones to reach an asymptote.

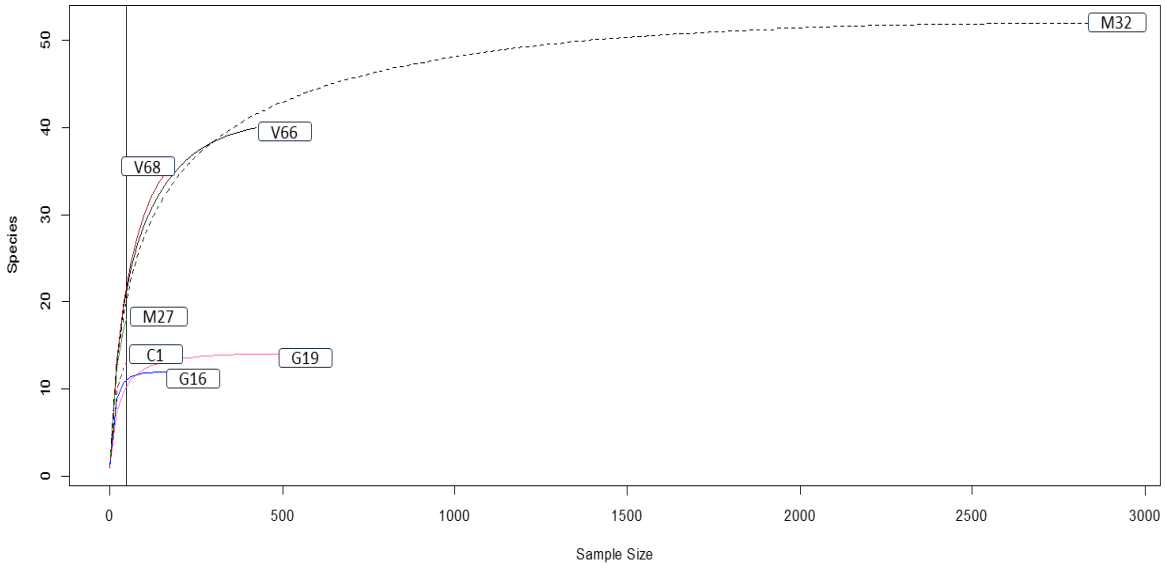

Figure 9 – Rarefaction curves for all samples selected for pyrosequencing. The vertical line indicates the sample size of the smallest sample.

4. Description of the community

From the results of the pyrosequencing data analysis, bacterial communities from *Pinus pinaster* comprised 6 phyla, 11 classes, 18 orders and 25 families. Only 21 OTUs were affiliated at the genus level and 9 OTUs were affiliated at species level with the identification of 6 different species (Table 6). Three OTUs could not be assigned to any organism in the database, one belonging to a Madeira sample, one to Vouzela samples and another one present in samples from both locations.

Table 6 – OTUs affiliated at the species level; the number of reads in each sample is presented.

C1	G16	G19	M27	M32	V66	V68	Taxon
4	0	17	0	6	0	0	*Acinetobacter johnsonii*
0	0	0	0	4	0	6	*Candidatus Solibacter* (1)
0	0	0	0	0	13	0	*Ochrobactrum intermedium*
3	0	0	0	4	0	0	*Paracoccus marcusii*
0	0	0	0	8	0	0	*Sphingomonas wittichii*
0	0	28	0	0	0	0	*Staphylococcus epidermidis*

(1) The designation Candidatus is not a rank but a status (Murray & Stackebrandt, 1995)

The most abundant phylum was Proteobacteria, represented by 52 OTUs and present in all samples, followed by Acidobacteria with 24 OTUs and Firmicutes with 7 OTUs. Acidobacteria was present in all samples, with the exception of Góis. Firmicutes was present in samples of Góis, Comporta, and in one sample of Vouzela.

The remaining phyla were represented by only one OTU each. The OTU belonging to the phylum Bacteroidetes was found in all samples of Góis and in one sample from Madeira. The phylum Armatimonadetes was present only in one sample of Vouzela and the phylum Actinobacteria was present only in Góis samples.

Proteobacteria was dominant in Góis and Comporta samples, followed by Firmicutes. Acidobacteria was more abundant in Madeira and Vouzela (Figure 10). Firmicutes was not present in Madeira samples and Acidobacteria was not present in Góis.

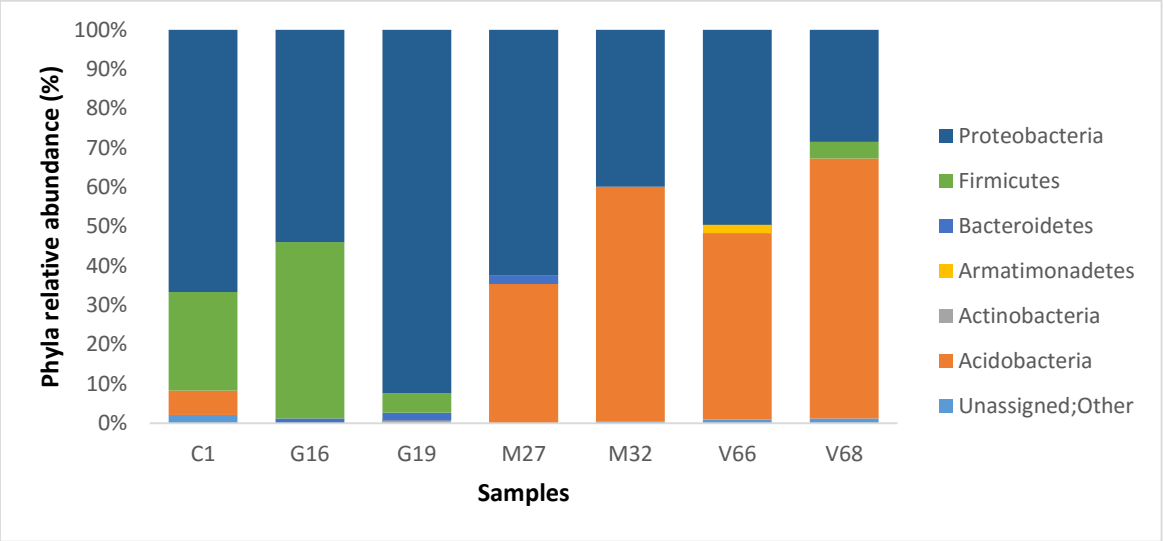

Figure 10 - Relative abundance of bacterial phyla for samples from all locations, Comporta (C), Góis (G), Madeira (M) and Vouzela (V). The relative abundance was calculated as the percentage of sequences belonging to a particular lineage of the total sequences recovered for each sample.

From the phylum Proteobacteria, the class Gammaproteobacteria was the most abundant (Figure 11), despite having differences in relative abundance between samples. The family Pseudomonadaceae was the most common being present in all samples, ranging from 0.5 % (sample M32) to 55 % (sample G19) of the samples.

The class Alphaproteobacteria presented a distribution with a smaller range and variated between 7 (Góis 16) and 38 % (Madeira 32). In samples from Madeira and Vouzela the family Acetobacteriaceae had a greatest representation. In the sample M32 in particular, the families Acetobacteriaceae and Methylocystaceae had a similar abundance, dominating the sample. The sample V66 was the most diverse in relation to of family diversity (Figure 11).

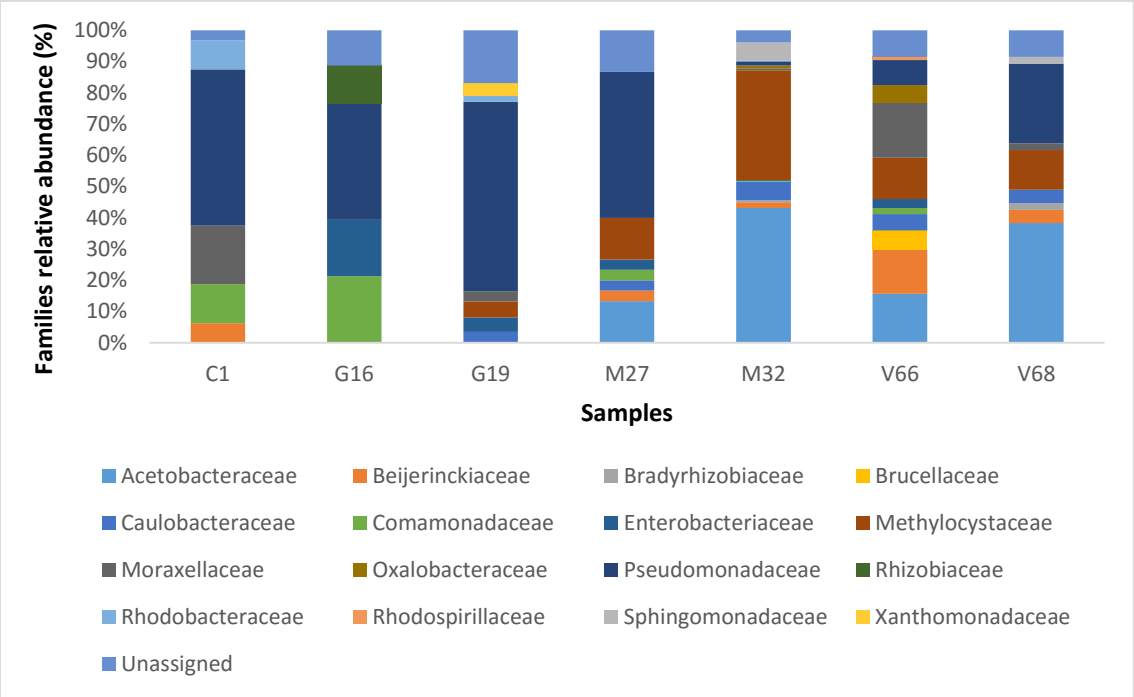

Figure 11 - Relative abundance of bacterial families from the phylum Proteobacteria from samples of all locations, Comporta (C), Góis (G), Madeira (M) and Vouzela (V). The relative abundance was calculated as the percentage of sequences belonging to a particular family of the total sequences for the Proteobacteria phylum.

The samples from Madeira and Vouzela presented a high percentage of the phylum Acidobacteria (Figure 10). From this phylum, only two families were represented: Acidobacteriaceae and Solibacteriaceae. The abundance of the Acidobacteriaceae family was the main reason for the high percentage of Acidobacteria, representing 35 to 58 % of the total number of reads per sample.

The Venn diagram (Figure 12) was used to indicate the common OTUs among all locations. Only two OTUs were shared between all the locations, and both were affiliated to Proteobacteria. Both OTUs belonged to the class Gammaproteobacteria, and one was identified as belonging to the genus *Pseudomonas*, however the other could not be assigned to a lower taxonomic unit than class. Vouzela and Madeira were the locations that had more OTUs in common and had the higher number of exclusive OTUs.

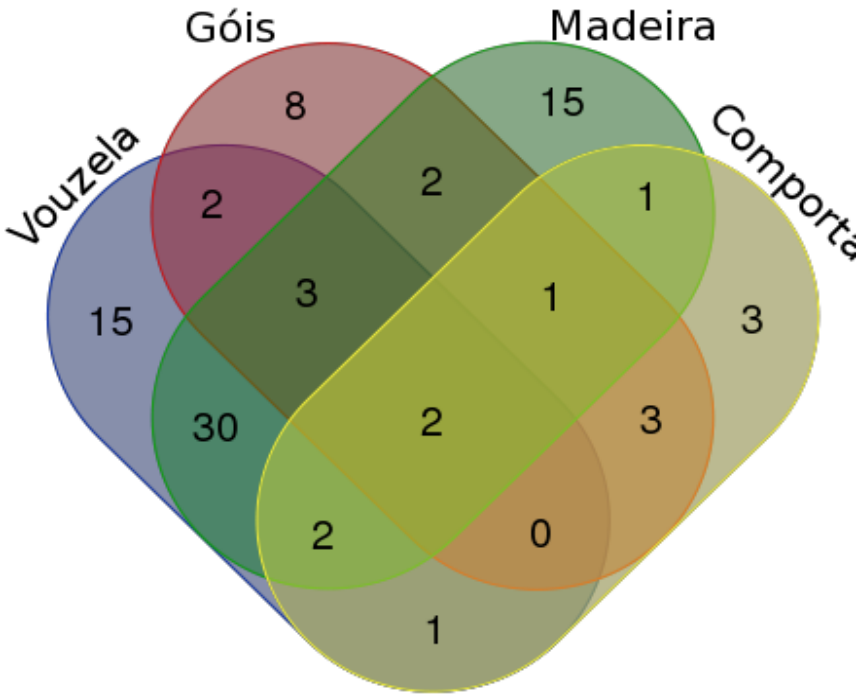

Figure 12 – Venn diagram showing the number of shared and unique OTUs among locations (Madeira, Vouzela, Góis and Comporta). The diagram was obtained using an online tool (http://bioinformatics.psb.ugent.be/webtools/Venn/).

A heat map (Table 7) was obtained using the OTUs with representation higher than 10 % for each sample. The common OTUs among all samples, OTU_3 and OTU_15, can be observed in the heat map along with their abundance in number of reads. The OTUs from the family Acidobacteriaceae were common only between Madeira and Vouzela. Methylocystaceae was common between all samples with the exception of Comporta, and had a high number of reads on the Madeira sample M32. Within the OTUs with more than 10 % of representation, four of them could be assigned to a genus. The OTU_57 is from the genus *Paenibacillus*, OTU_49 to *Bacillus* and both OTU_3 and OTU_143 to *Pseudomonas*.

Table 7 – Heat map with the families of OTUs that represent more than 10 % of the reads of each sample, with the exception of OTU_15 (*) since class was the lower phylogenetic taxa that this OTU could be assigned to.

Family	OTUs	C1	G16	G19	M27	M32	V66	V68
Pseudomonadaceae	OTU_3	16	24	260	6	14	11	10
*Gammaproteobacteria	OTU_15	1	10	88	4	9	3	2
Methylocystaceae	OTU_2	0	0	21	4	331	11	0
Acidobacteriaceae	OTU_21	0	0	0	1	70	55	11
Acidobacteriaceae	OTU_51	0	0	0	6	38	17	4
Acidobacteriaceae	OTU_145	0	0	0	0	130	56	19
Solibacteraceae	OTU_32	0	0	0	0	19	0	30
Acidobacteriaceae	OTU_4	0	0	0	2	321	0	0
Pseudomonadaceae	OTU_143	0	0	0	6	0	0	0
Paenibacillaceae	OTU_57	1	37	0	0	0	0	0
Bacillaceae	OTU_49	4	29	0	0	0	0	0

The statistical analysis of the pyrosequencing data was perfomed using Primer v6 software. Data was transformed using square root, and PERMANOVA was calculated after the construction of a Bray-Curtis similarity matrix, to evaluate if the bacterial community differed according to the location. The result was $p=0.013$ ($p<0.05$), indicating that sampling location significantly influences *P. pinaster* bacterial composition.

Diversity (Shannon index), richness and equitability (evenness) were calculated for each sample (Table 8). Equitability values ranged from 0.71 to 0.93 (close to 1), displaying an almost uniform distribution of the OTUs abundance within each sample. Diversity ranged from 1.87 to 3.16 and richness from 12 to 52 OTUs.

Table 8 – Values for the diversity indexes calculated for the pyrosequencing data.

	Richness	Equitability	Diversity
V66	40	0.85	3.14
V68	35	0.86	3.07
M27	18	0.93	2.69
G19	14	0.71	1.87
G16	12	0.90	2.24
M32	52	0.80	3.16
C1	13	0.87	2.22

A dendrogram (Figure 13) was constructed based on the pyrosequencing data using Primer 6 and it shows that the samples are divided in two major groups, one with Madeira and Vouzela and other with Góis and Comporta.

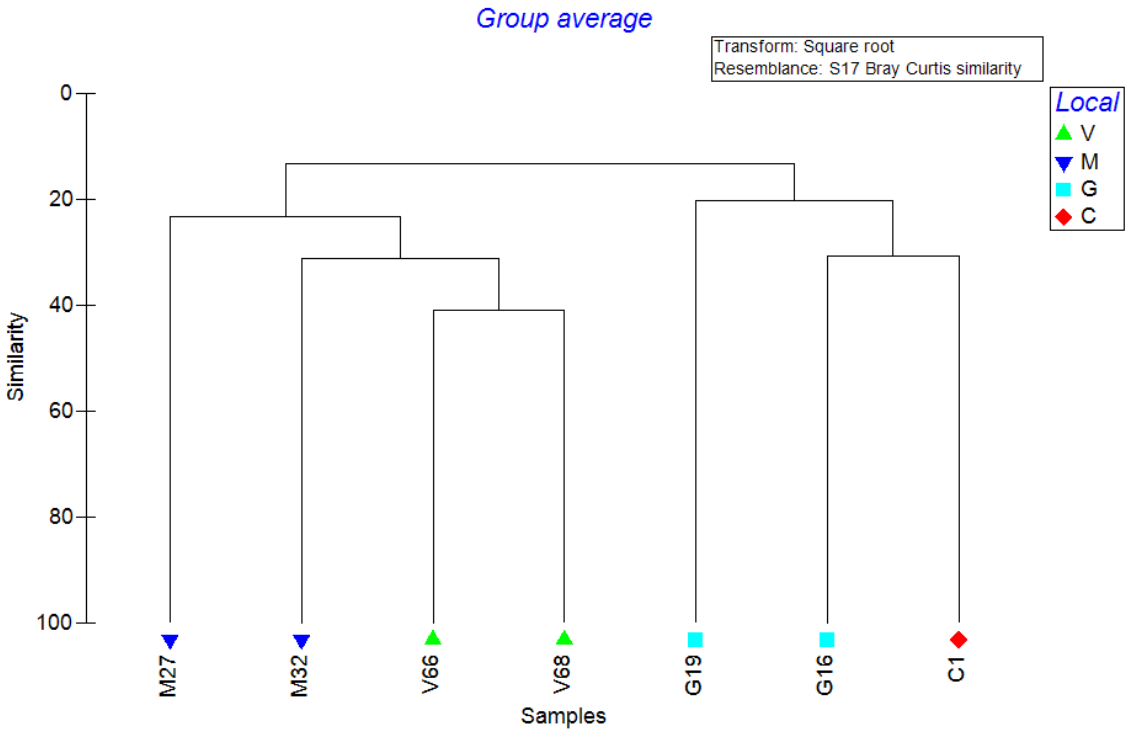

Figure 13 – Dendrogram based on the percentage of similarity between locations from the pyrosequencing data. Vouzela (V), Madeira (M), Góis (G) and Comporta (C).

DISCUSSION

The knowledge of the bacterial community present in *P. pinaster* is important especially because of the increasing dispersion of pine wilt disease. This information can be useful for future studies about the influence of bacteria in the development of the disease since there is not much information on the *P. pinaster* microbiome and most of the studies until now used a culture dependent approach.

The pyrosequencing data analysis revealed that Proteobacteria and Acidobacteria are the phyla that consistently colonize asymptomatic *P. pinaster* trees. Both phyla had great representation in terms of OTUs numbers. The phylum Proteobacteria had the larger number of reads and OTUs, and it was present in every sample. Based on that, it is probable that the phylum Proteobacteria includes species that play an important role for the plant or are well adapted to life as an endophyte. Proteobacteria was the main phylum of endophytes on the leaf of tomatoes (*Solanum lycopersicum*) analysed by pyrosequencing (Romero *et al*., 2014) and also in *P. flexis* and *P. engelmannii* leaves together with Acidobacteria (Carrell & Frank, 2014).

In the Proteobacteria phylum, Alphaproteobacteria is the most common class in the leaf surface of various *Pinus* species (Redford *et al.*, 2010), and also in the endophytic community of *P. flexis* and *P. engelmannii* leaves (Carrell & Frank, 2014). In this study, the samples of *P. pinaster* presented Gammaproteobacteria as the most common class in the majority of the samples, followed by Alphaproteobacteria. A study with *P. pinaster* contaminated and non-contaminated with pine wilt disease from another region of Portugal had a similar result (Proença & Morais, 2015). Gammaproteobacteria was also the dominant phylum in soybean stems (Okubo *et al*., 2009) and in *Stellera chamaejasme* stems (Jin *et al*., 2014). This suggest that this class may have phylotypes well adapted for the stem environment.

The 97 % similarity cut-off for the OTUs clustering was enough for the identification up until the species level of some OTUs, but is not possible to guarantee that this information as correctly assigned due to the size and quality of the reads obtained by pyrosequencing. Despite that, the species were analysed according to information available in the literature.

Acinetobacter johnsonnii was isolated also as an endophyte from tomatoes (*Solanum lycopersicum*) (Barretti *et al*., 2009) and it presented a growth promotion

characteristic in sugar beets (Shi *et al.*, 2010). *Candidatus solibacter* was found in glacier fore field vegetated soil, degrading complex organic compounds from the soil (Rime *et al*., 2015). The designation "Candidatus" is not a rank but a status that is given to organisms that can be recognized by their molecular structure but could not be assigned to a known genus (Murray & Stackebrandt, 1995). *Ochrobactrum intermedium* was isolated from water and soil, and was able to promote plant growth in soils contaminated with Chromium (Cr) or Lead (Pb) (Waranusantigul *et al*., 2011; Faisal & Hasnain, 2006). *Paracoccus marcusii* is known for the production of carotenoids and use of nitrate as electron acceptor (Harker *et al.*, 1998) and was found in Artic marine sediments (Cha *et al*., 2015). *Sphigomonas wittichii* has a gene cluster for the degradation of the pant hormone indole-3 acetic acid and was isolated from water (Leveau & Gerards, 2008). *Staphylococcus epidermidis* is known for causing opportunistic infections and colonization of medical devices, is abundant on the skin and that is why is the most common cause of contamination of clinical specimens (Büttner *et al.*, 2015). The presence of *S. epidermidis* is restricted to the sample G19 and it has a high number of reads compared to the other species found, this may indicate that this sample was imprudently handled in some step of the experiment.

The bacterial community varied between samples and the statistical analysis showed that they were significantly different between sampling locations. This observation implies that other factors such as soil, climate and location may have a stronger influence in the community of *P. pinaster* than the tree genotype. Several studies demonstrate the influence of the environmental conditions on the endophytic community (Rosenblueth & Martínez-Romero, 2006). In olive trees, community changed based in the geographical origin of each sample (Müller *et al*., 2015). *P. flexis* and *P. engelmannii* samples from the same region, despite being different species, have a consistent bacterial community (Carrell & Frank, 2014).

In the Venn diagram only two OTUs were common between all samples. Both of them belonging to the Gammaproteobacteria class, one of them was assigned as belonging to the genus *Pseudomonas,* and showed a high abundance in the heat map for the sample G19. The genus *Pseudomonas* is well studied for its plant growth promotion ability, namely the production of antimicrobial compounds and phosphorous solubilisation ability, yielding the genus a role in the biological control products (Bulgarelli *et al*., 2013). However, phytopathogens from this genus also exist such as

Pseudomonas syringae causing leaf spots, blights and wilts (Mercado-Blanco & Bakker, 2007).

Comparing the dendrogram from the DGGE and pyrosequencing, it is possible to observe that the samples group by sampling location. Nevertheless, the sampling locations are grouped in different ways, with Vouzela grouping with Góis in the DGGE and with Madeira in the pyrosequencing. The DGGE profiles look similar to each other, suggesting that the community is similar throughout the locations. However, the results of the cloning approach showed that the strong bands observed on the DGGE are result from the amplification of chloroplast DNA. Consequently, it is not advisable to take conclusions based only in the DGGE analysis.

The results of the cloning were correlated with the pyrosequencing analysis, where the reads eliminated as chloroplasts during quality control were approximately 90 % of the total number of the initial reads. The chloroplasts DNA interference seems to be recurrent problem with the molecular analyses in plant related studies. Some authors propose different sets of primers that supposedly exclude chloroplasts DNA (Redford *et al.*, 2010). Others aside from primers may use enzymatic digestion to ensure efficiency (Shen & Fulthorpe, 2015).

Not all samples reached an asymptote in the rarefaction curves, indicating that the diversity was not deeply characterized. One sample from Góis reached an asymptote but in a very low point compared to the other samples, Vouzela (V66) and Madeira (M32) indicating low richness and diversity. This can be confirmed with the indexes, as Góis samples have a low richness and diversity compared to V66 and M32.

The low number of reads caused by the exclusion of the chloroplasts assigned OTUs may have compromised the access to all the phylotypes present in the community. The community that inhabits the stem of the tree may be in a low density, which difficult its detection. There is a low density of bacteria in the stem when contrasted to the endophytes bacteria of the root (Mocali *et al.*, 2003). According to some authors diluted samples, with density lower than $< 10^5$ bacteria per ml (DNA <1 pg/µl) are prone to have deviations on the pyrosequencing bacterial community profile when compared to the original sample, probably because of the increase of the interference of contaminants (Biesbroek *et al.*, 2012). Endophytes rarely exceed 10^8 colony forming unit (CFU) per gram of fresh weight since a high concentration may result in an elicitation of the defense response, and can often be lower than 10^3 CFU per gram of fresh weight depending on age and genotype (Turner *et al.*, 2013).

Nevertheless, the community structure did not differ from other findings showing that there may be a common bacterial community structure that is transversal to the host species. Therefore, the results can be considered a representation of the microbiome of *P. pinaster* despite the low number of reads. The community may be not deeply characterized but it is enough to be used in further studies and comparisons.

CONCLUSION

In the study of the microbiome of *P. pinaster*, six phyla were found. The most common was Proteobacteria, especially the Gammaproteobacteria class. Only two OTUs were common between all samples, one from the Gammaproteobacteria class and the other was assigned to the *Pseudomonas* genus. All samples significantly differ between locations. Despite the low number of reads, the community structure was characterized and this information can be now useful to further studies especially on the PWD.

REFERENCES

Andersson, F. (2005). Coniferous Forests. Elsevier. Retrieved from https://books.google.pt/books?id=_YoO8xKYSJAC

Akbulut, S., & Stamps, W. T. (2012). Insect vectors of the pinewood nematode: a review of the biology and ecology of *Monochamus* species. *Forest Pathology*, 42(2), 89–99.

Alía, R., & Martín, S. (2003). EUFORGEN: Technical guidelines for genetic conservation and use for Maritime pine (*Pinus pinaster*). *International Plant Genetic Resources Institute*, 6.

Bajaj, Y. P. S. (2013). Trees III. Springer Berlin Heidelberg. Retrieved from https://books.google.pt/books?id=ybj0CAAAQBAJ

Bal, A., Anand, R., Berge, O., & Chanway, C. (2012). Isolation and identification of diazotrophic bacteria from internal tissues of *Pinus contorta* and *Thuja plicata*. *Canadian Journal of Forest Research*, 813, 807–813.

Barretti, P. B., Romeiro, R. da S., Mizubuti, E. S. G., & de Souza, J. T. (2009). Seleção de bactérias endofíticas de tomateiro como potenciais agentes de biocontrole e de promoção de crescimento. *Ciência E Agrotecnologia*, 2038–2044.

Biesbroek, G., Sanders, E. A. M., Roeselers, G., Wang, X., Caspers, M. P. M., Trzciński, K., Bogaert & Keijser, B. J. F. (2012). Deep sequencing analyses of low density microbial communities: working at the boundary of accurate microbiota detection. *PloS One*, 7(3), e32942.

Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., & Vorholt, J. A. (2014). A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genetics*, 10(4), e1004283.

Borer, E. T., Kinkel, L. L., May, G., & Seabloom, E. W. (2013). The world within: Quantifying the determinants and outcomes of a host's microbiome. *Basic and Applied Ecology*, 14(7), 533–539.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., Themaat, E. V. L., & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant Biology*, 64, 807–38.

Bull, C. T., Coutinho, T. A., Denny, T. P., Firrao, G., Saux, M. F., Li, X., Saddler, G. S., Scortichini, M., Stead, D. E. & Takikawa, Y. (2014). List of new names of plant pathogenic bacteria 2011-2012. *Journal of Plant Pathology*, 96, 223–226.

Büttner, H., Mack, D., & Rohde, H. (2015). Structural basis of *Staphylococcus epidermidis* biofilm formation: mechanisms and molecular interactions. *Frontiers in Cellular and Infection Microbiology*, 5, 1–15.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, K., Gordon, J. I., Huttley, G., Kelley, S. T., Knights, D., Jeremy, E., Ley, R. E., Lozupone, C., Mcdonald, D., Muegge, B. D., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J. & Walters, W. A. (2011). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7(5), 335–336.

Cardoso, J. M. S., Fonseca, L., & Abrantes, I. (2012). Direct molecular detection of the pinewood nematode, *Bursaphelenchus xylophilus*, from pine wood, bark and insect vector. *European Journal of Plant Pathology*, 133(2), 419–425.

Carrell, A. a., & Frank, A. C. (2014). Pinus flexilis and Picea engelmannii share a simple and consistent needle endophyte microbiota with a potential role in nitrogen fixation. *Frontiers in Microbiology*, 5, 1–11.

Cha, I.-T., Song, E.-J., Seok, Y. J., Lee, H., Park, I., Lee, Y. K., Roh, W. S., Choi, H., Nam, Y. & Seo, M.-J. (2015). Draft genome sequence of a denitrifying bacterium *Paracoccus marcusii* PAMC 22219 isolated from Arctic marine sediment. *Marine Genomics*, 21, 27–9.

Compant, S., Duffy, B., Nowak, J., Cle, C., & Barka, E. A. (2005). Use of plant growth-promoting bacteria for biocontrol of plant diseases : Principles, mechanisms of action and future prospects. *Applied and Environmental Microbiology*, 71(9), 4951–4959.

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10(10), 996–998.

EPPO, (2014). Bursaphelenchus xylophilus. Data Sheets on Quarentine Pests. Retrived from http://www.eppo.int/QUARANTINE/data_sheets/nematodes/BURSXY_ds.pdf

Faisal, M., & Hasnain, S. (2006). Growth stimulatory effect of *Ochrobactrum intermedium* and *Bacillus cereus* on *Vigna radiata* plants. *Letters in Applied Microbiology*, 43, 461–466.

Farjon, A., & Filer, D. (2013). An atlas of the world's conifers: an analysis of their distribution, biogeography, diversity and conservation status. *Brill Academic Publication*, Leiden. 512 p.

Fatima, U., & Senthil-Kumar, M. (2015). Plant and pathogen nutrient acquisition strategies. *Frontiers in Plant Science*, 6, 1–12.

Futai, K. (2013). Pine Wood Nematode, Bursaphelenchus xylophilus. *Annual Review of Phytopathology*, 51(1), 61–83.

Hardoim, P. R., van Overbeek, L. S., & Elsas, J. D. (2008). Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology*, 16(10),

463–471.

Harker, M., Hirschberg, J., & Oren, A. (1998). *Paracoccus narcusii* sp. nov., an orange Gram-negative coccus. *International Journal of Systematic Bacteriology*, (48), 543–548.

Henriques, I., Alves, A., Tacão, M., Almeida, A., Cunha, A. & Correia, A. (2006). Seasonal and spatial variability of free-living bacterial community composition along an estuarine gradient (Ria deAveiro, Portugal). E*stuarine, Coastal Shelf Science*. 68, 139–148.

ICNF, (2013). Instituto de Conservação da Natureza e da Floresta. IFN6 - Áreas dos usos do solo e das espécies florestais de Portugal continental em 1995, 2005 e 2010, 34.

ICNF, (2015). Instituto de Conservação da Natureza e da Floresta. Freguesias classificadas como Local de Intervenção (LI). Retrived from http://www.icnf.pt/portal/florestas/prag-doe/ag-bn/nmp/resource/fich/dl95-freg/LI_NMP_Out2015_v2.pdf.

Icnf.pt, (2015). Informação geográfica — ICNF. [online] Available at: http://www.icnf.pt/portal/florestas/prag-doe/ag-bn/nmp/infgeo [Accessed 21 Dec. 2015].

Inácio, M., Nóbrega, F., Vieira, P., Bonif, L., Naves, P., Sousa, E., & Mota, M. (2014). First detection of Bursaphelenchus xylophilus associated with *Pinus nigra* in Portugal and in Europe. *Forest Pathology*, 45, 235–238.

Izhaki, I., Fridman, S., Gerchman, Y., & Halpern, M. (2013). Variability of bacterial community composition on leaves between and within plant species. *Current Microbiology*, 66(3), 227–235.

Izumi, H., Anderson, I. C., Killham, K., & Moore, E. R. B. (2008). Diversity of predominant endophytic bacteria in European deciduous and coniferous trees. *Canadian Journal of Microbiology*, 54(3), 173–179.

Jin, H., Yang, X. Y., Yan, Z. Q., Liu, Q., Li, X. Z., Chen, J. X., Zhang, D. H., Zeng, L. M. & Qin, B. (2014). Characterization of rhizosphere and endophytic bacterial communities from leaves, stems and roots of medicinal *Stellera chamaejasme* L. *Systematic and Applied Microbiology*, 37(5), 376–385.

Keeley, J. E. (2012). Ecology and evolution of pine life histories. *Annals of Forest Science*, 69(4), 445–453.

Kim, M. G., Kim, W. J., Lee, J. R., Jung, Y. J., & Lee, S. Y. (2008). Host immunity-suppressive molecular weapons of phytopathogenic bacteria. *Journal of Plant Biology*, 51, 233–239.

Leveau, J. H. J., & Gerards, S. (2008). Discovery of a bacterial gene cluster for catabolism of the plant hormone indole 3-acetic acid. *FEMS Microbiology Ecology*, 65(2), 238–

250.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M., Toth, I., Salmond, G. & Foster, G. D. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*, 13(6), 614–629.

Mercado-Blanco, J., & Bakker, P. A. H. M. (2007). Interactions between plants and beneficial *Pseudomonas* spp.: Exploiting bacterial traits for crop protection. Antonie van Leeuwenhoek, *International Journal of General and Molecular Microbiology*, 92(4), 367–389.

Mocali, S., Bertelli, E., Di Cello, F., Mengoni, A., Sfalanga, A., Viliani, F., Caciotti, A., Tegli, S., Surico, G. & Fani, R. (2003). Fluctuation of bacteria isolated from elm tissues during different seasons and from different plant organs. *Research in Microbiology*, 154(2), 105–114.

Mota, M., Braasch, H., Bravo, M. A., Penas, A. C., Burgermeister, W., Metge, K., & Sousa, E. (1999). First report of *Bursaphelenchus xylophilus* in Portugal and in Europe. *Nematology*, 1, 727–734.

Müller, H., Berg, C., Landa, B. B., Auerbach, A., Moissl-Eichinger, C., & Berg, G. (2015). Plant genotype-specific archaeal and bacterial endophytes but similar *Bacillus* antagonists colonize Mediterranean olive trees. *Frontiers in Microbiology*, 6, 138.

Müller, T., & Ruppel, S. (2014). Progress in cultivation-independent phyllosphere microbiology. *FEMS Microbiology Ecology*, 87, 2–17.

Murray, R. G. E., & Stackebrandt, E. (1995). Taxonomic Note : Implementation of the provisional status *Candidatus* for incompletely described procaryotes. *International Journal of Systematic Bacteriology*, 45(1), 186–187.

Muthamilarasan, M., & Prasad, M. (2013). Plant innate immunity: An updated insight into defense mechanism. *Journal of Biosciences*, 38(2), 433–449.

Nascimento, F. X., Hasegawa, K., Mota, M., & Vicente, C. S. L. (2015). Bacterial role in pine wilt disease development - review and future perspectives. *Environmental Microbiology Reports*, 7(1), 51–63.

Proença, D. N., & Morais, P. V. (2015). O papel das bactérias na Doença da Murchidão do Pinheiro. *Magazine Da SPM*, 2015(4), 1–6.

pt.climate-data.org, (2015). Dados climáticos para cidades mundiais - Climate-Data.org. [online] Available at: http://pt.climate-data.org/ [Accessed 30 Oct. 2015].

Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y., & Fierer, N. (2010). The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental Microbiology*, 12(11), 2885–2893.

Rime, T., Hartmann, M., Brunner, I., Widmer, F., Zeyer, J., & Frey, B. (2015). Vertical distribution of the soil microbiota along a successional gradient in a glacier forefield. *Molecular Ecology*, 24(5), 1091–1108.

Romero, F. M., Marina, M., & Pieckenstain, F. L. (2014). The communities of tomato (*Solanum lycopersicum* L) leaf endophytic bacteria, analyzed by 16S-ribosomal RNA gene pyrosequencing. *FEMS Microbiology Letters*, 351(2), 187–194.

Santi, C., Bogusz, D., & Franche, C. (2013). Biological nitrogen fixation in non-legume plants. *Annals of Botany*, 111(5), 743–767.

Schadt, C. W., & Ephane Hacquard, S. (2015). Tansley insight towards a holistic understanding of the beneficial interactions across the *Populus* microbiome. *New Phytologist*, 205,1424–1430.

Shannon, C. E. (1948). The mathematical theory of communication. 1963. MD *Computing Computers in Medical Practice*.

Shen, S. Y., & Fulthorpe, R. (2015). Seasonal variation of bacterial endophytes in urban trees. *Frontiers in Microbiology*, 6, 1–13.

Shi, Y., Lou, K., & Li, C. (2010). Growth and photosynthetic efficiency promotion of sugar beet (Beta vulgaris L.) by endophytic bacteria. *Photosynthesis Research*, 105(1), 5–13.

Turner, T. R., James, E. K., & Poole, P. S. (2013). The Plant Microbiome. *Genome Biology*, 6(14), 209.

Vendan, R. T., Yu, Y. J., Lee, S. H., & Rhee, Y. H. (2010). Diversity of endophytic bacteria in ginseng and their potential for plant growth promotion. *Journal of Microbiology*, 48(5), 559–565.

Vicente, C., Espada, M., Vieira, P., & Mota, M. (2012). Pine Wilt Disease: a threat to European forestry. *European Journal of Plant Pathology*, 133(1), 89–99.

Wang, Y., & Qian, P.-Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE*, 4(10), e7401.

Waranusantigul, P., Lee, H., Kruatrachue, M., Pokethitiyook, P., & Auesukaree, C. (2011). Isolation and characterization of lead-tolerant *Ochrobactrum intermedium* and its role in enhancing lead accumulation by *Eucalyptus camaldulensis*. *Chemosphere*, 85(4), 584–590.

Zhao, L., Mota, M., Vieira, P., Butcher, R. A., & Su, J. (2014). Interspecific communication between pinewood nematode, its insect vector, and associated microbes. *Trends in Parasitology*, 30(6), 299–308.