



UNIVERSIDADE D
COIMBRA

Jorge Miguel Pereira da Silva

**O MICROBIOMA DE LESÕES NA CASCA
NO DECLÍNIO DO SOBREIRO**

Tese no âmbito do Mestrado em Biologia orientada pelo Professor Doutor Jorge Manuel Pataca Leal Canhoto, Diretor do Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra e pela Doutora Maria da Conceição Venâncio Egas, Investigadora do Centro de Neurociências e Biologia Celular da Universidade de Coimbra e apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Setembro de 2019

Faculdade de Ciências e Tecnologia
da Universidade de Coimbra

O MICROBIOMA DE LESÕES NA CASCA NO DECLÍNIO DO SOBREIRO

Jorge Miguel Pereira da Silva

Tese no âmbito do Mestrado em Biologia orientada pelo Professor Doutor Jorge Manuel Pataca Leal Canhoto, Diretor do Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra e pela Doutora Maria da Conceição Venâncio Egas, Investigadora do Centro de Neurociências e Biologia Celular da Universidade de Coimbra e apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

setembro de 2019



UNIVERSIDADE D
COIMBRA

À memória de minha mãe.

Agradecimentos

Tenho para mim que, toda a tela do trabalho de investigação, que antevê a composição de uma dissertação de mestrado como a aqui exposta, está longe de ser apenas o resultado do esforço e dedicação do seu proponente. Os condicionalismos da memória tornam o espaço de agradecimentos um local de injustiça, que esquece ou não dá o devido valor a todos quantos para ele contribuíram.

Ainda assim, arrisco-me a sublinhar, em primeiro lugar os contributos incalculáveis da Doutora Maria da Conceição Venâncio Egas, Investigadora do Centro de Neurociências e Biologia Celular da Universidade de Coimbra, e do Professor Doutor Jorge Manuel Pataca Leal Canhoto, Diretor do Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra, cujo esforço de orientação será para sempre uma referência, quer para a vida pessoal, quer académica.

Da mesma forma, este projeto não poderia ser concretizado sem a receptividade e abertura da GenoInseq, a Unidade de Sequenciação Avançada do Biocant Park, na pessoa da sua diretora, Doutora Conceição Egas, em consentir a utilização dos seus recursos e acompanhamento pelo seu corpo de investigadores e funcionários, em particular, Diogo Pinho e Inês Gregório, num espírito de aprendizagem que espero ter retribuído com empenho e dedicação.

Gostaria também de deixar uma palavra a todos os docentes, não só da Universidade de Coimbra, mas também da Universidade do Algarve, que fizeram parte do meu crescimento e percurso académico, pelo que a presente dissertação de mestrado reflete também o seu cunho. Em especial à Professora Doura Isabel Luci Pisa Mata da Conceição, do Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra, pela disponibilidade, estima e todo o apoio demonstrado ao longo do último ano. A todos eles muito obrigado.

Uma última palavra de profundo apreço para todos aqueles que torceram e me acompanharam nesta tese, toda a minha família e amigos.

Não sou capaz de colocar por palavras todo o sentimento de dívida que sinto perante aqueles que lutaram dia a dia pela minha educação e formação desde muito cedo, aos meus avós, agradeço o mundo, *ad aeternum*.

Barcelos, 18 de agosto de 2019.

Este trabalho foi efetuado na *Genoinseq*, a Unidade de Sequenciação Avançada do *Biocant Park* e do Centro de Neurociências e Biologia Celular da Universidade de Coimbra (CNC), sob a supervisão de Doutora Conceição Egas, Investigadora do Centro de Neurociências e Biologia Celular da Universidade de Coimbra, contando com o suporte de Diogo Pinho e toda a equipa de sequenciação e bioinformática.



Resumo

Durante as últimas décadas, têm ocorrido eventos de mortalidade de sobreiros dentro de várias florestas de sobreiro na bacia mediterrânica, o que levou a uma perda progressiva dos montados, comprometendo a biodiversidade única que estas florestas acarretam. A mortalidade do sobreiro é um problema cada vez mais recorrente em Portugal e tem gerado grande preocupação em relação à sustentabilidade dos montados.

Uma das principais enfermidades que tem afetado os carvalhos é o Declínio Agudo do Carvalho, uma doença relativamente nova e bastante agressiva, encontrada sobretudo na Grã-Bretanha, que provoca a morte de carvalhos. Em árvores sintomáticas desta doença surgem lesões na casca externa, por onde escorre um exsudato negro característico, cuja origem é atribuída a um grupo de bactérias, onde se incluem as espécies *Gibbsiella quercinecans*, *Brenneria goodwinii*, *Lonsdalea quercina* ssp. *britannica*, *Brenneria roseae* ssp. *roseae*, *Rahnella victoriana*, *Rahnella variigena* e *Gibbsiella greigii*.

Com o propósito de perceber se estas lesões, observadas e analisadas em Portugal, correspondem àquelas descritas para o AOD na Grã-Bretanha, foi realizado um estudo detalhado do microbioma bacteriano e fúngico das lesões na casca de sobreiros com diferentes estados de desfolha, em montados com diferentes percentagens de declínio. Para isso, os exsudatos foram recolhidos, o seu DNA total foi extraído, quantificado e purificado, e as sequências das regiões hipervariáveis 3-4 do gene RNA Ribossomal 16S, para bactérias, e do espaçador interno transcrito 2, para fungos, foram amplificados por PCR e sequenciadas na plataforma de sequenciação de última geração *MiSeq Illumina*[®].

Os resultados foram analisados bioinformaticamente e estatisticamente recorrendo aos softwares *QIIME*[™] e *MicrobiomeAnalyst*, em termos de alfa e beta diversidade, revelando uma elevada biodiversidade dentro das amostras recolhidas. De uma forma geral, as comunidades microbianas encontradas enquadram-se com a literatura do declínio agudo do carvalho, sendo semelhantes àquelas encontradas em outras árvores.

A heterogeneidade das comunidades bacterianas e fúngicas revelou não ser provocada pelos fatores ambientais identificados. No entanto, as amostras mostraram um agrupamento de comunidades semelhantes, de acordo com fatores ainda não identificados, e a presença de microrganismos emergentes, o que poderá revelar-se um novo caminho a explorar.

PALAVRAS-CHAVE: Sobreiro, Exsudatos, Declínio Agudo do Carvalho, Sequenciação de DNA, Microbioma.

Abstract

Over the last few decades, cork oak mortality events have occurred within several Mediterranean basin forests, which has led to a progressive loss of montados, compromising the unique biodiversity that these forests entail. Cork oak mortality is an increasingly recurring problem in Portugal and has generated great concern regarding the sustainability of montados.

One of the main illnesses that has affected oaks is the Acute Oak Decline, a relatively new and aggressive disease, found mainly in Britain, which leads to death. In symptomatic trees of this condition, lesions appear on the outer bark, through which a characteristic black exudate flows, whose origin is credited to a group of bacteria, including the species *Gibbsiella quercinecans*, *Brenneria goodwinii*, *Lonsdalea quercina* ssp. *britannica*, *Brenneria roseae* ssp. *roseae*, *Rahnella victoriana*, *Rahnella variigena* and *Gibbsiella greigii*.

In order to understand if these lesions, observed and analyzed in Portugal, correspond to those described for AOD in Great Britain, a detailed study of the bacterial and fungal microbiome of cork oak bark lesions from trees with different defoliation states and from montados with different percentages of decline was carried out. For this, the exudates were collected, their total DNA was extracted, quantified and purified, and the sequences of the 3-4 hypervariable regions of the 16S Ribosomal RNA gene, for bacteria, and the internal transcribed spacer 2, for fungi, were amplified by PCR and sequenced on the latest generation sequencing platform MiSeq Illumina®.

Results were analyzed bioinformatically and statistically using QIIME™ and MicrobiomeAnalyst softwares, in terms of alpha and beta diversity, revealing a high biodiversity within the collected samples. In general, the microbial communities found fit the acute oak decline literature and are similar to those found in other trees.

The bacterial and fungal community's heterogeneity revealed not to be caused by the sampled environmental factors. However, the samples showed a cluster of similar communities according to factors not yet identified, and the presence of emerging microorganisms, which may prove to be a new exploration pathway.

KEYWORDS: Cork oak, Exudates, Acute Oak Decline, DNA Sequencing, Microbiome.

Índice

Dedicatória.....	iii
Agradecimentos.....	v
Resumo.....	ix
<i>Abstract</i>	xi
Índice de tabelas	xv
Índice de figuras	xvii
Lista de abreviaturas	xix
1. Introdução.....	21
1.1. <i>Quercus suber</i> L.....	21
1.2. Ameaças.....	23
1.3. Declínio do sobreiro	24
1.3.1. Declínio Crónico do Carvalho	25
1.3.2. Declínio Agudo do Carvalho	27
1.4. Objetivos.....	31
2. Estudo do Microbioma.....	33
2.1. Recolha de amostras	34
2.2. Protocolos de DNA.....	35
2.3. Sequenciação de última geração.....	37
2.4. Processamento bioinformático dos dados de sequenciação	41
2.5. Análise estatística dos dados de sequenciação	43
3. Materiais e métodos	45
3.1. Recolha de amostras	45
3.2. Protocolos de DNA.....	47
3.2.1. Extração.....	47
3.2.2. Quantificação.....	47
3.2.3. Amplificação	48
3.2.4. Purificação	49
3.3. Sequenciação	49

3.3.1. Indexação.....	49
3.3.2. Normalização, <i>pooling</i> e quantificação da biblioteca	49
3.3.3. Sequenciação	50
3.4. Processamento bioinformático dos dados de sequenciação.....	50
3.5. Análise estatística dos dados de sequenciação.....	50
3.5.1. Filtração dos dados.....	51
3.5.2. Análise dos dados.....	51
3.5.3. Análise do microbioma no <i>MicrobiomeAnalyst</i>	53
4. Resultados e Discussão	55
4.1. Extração	55
4.2. Quantificação	55
4.3. Amplificação.....	56
4.4. Sequenciação	58
4.5. Controlo do <i>kit</i>	60
4.6. Análise dos dados.....	64
4.7. Descrição do microbioma	65
4.7.1. Alfa diversidade.....	66
4.7.2. Diversidade taxonómica.....	68
4.7.3. Microbioma nuclear.....	73
4.7.4. Agrupamento das amostras.....	74
4.7.5. Beta diversidade	77
4.7.6. Alfa diversidade dos grupos	79
4.7.7. Diversidade taxonómicas dos grupos.....	81
4.8. Influência das variáveis ambientais no microbioma.....	83
5. Conclusão	89
Bibliografia.....	91

Índice de tabelas

Tabela 1. Caracterização das amostras recolhidas através dos esfregaços dos exsudados dos sobreiros.....	46
Tabela 2. Condições das diferentes etapas de PCR para amplificação das regiões V3-V4 e ITS2.....	48
Tabela 3. Condições das diferentes etapas de PCR para amplificação dos índices e adaptadores de sequenciação.	49
Tabela 4. Opções disponíveis no MicrobiomeAnalyst para normalização dos dados.	52
Tabela 5. Tabela de metadados criada para análise do microbioma das lesões na casca do sobreiro.	53
Tabela 6. Quantificação do DNA extraído por <i>NanoDrop</i> [™] e por <i>Qubit</i> [®] <i>dsDNA HS</i>	55
Tabela 7. Dados relativos à sequenciação de bactérias e fungos em termos de número de <i>singletons</i> , número de <i>reads</i> e <i>Good's coverage</i>	59
Tabela 8. Tabela resumo do <i>upload</i> para o <i>MicrobiomeAnalyst</i> dos dados de sequenciação de bactérias e fungos para análise do controlo do <i>kit</i>	61
Tabela 9. Métricas de alfa diversidade relativas à sequenciação de bactérias dos esfregaços recolhidos, calculadas através dos estimadores de riqueza e ou variância <i>Chao1</i> , <i>Simpson</i> e <i>Shannon</i>	67
Tabela 10. Métricas de alfa diversidade relativas à sequenciação de fungos dos esfregaços recolhidos, calculadas através dos estimadores de riqueza e ou variância <i>Chao1</i> , <i>Simpson</i> e <i>Shannon</i>	67
Tabela 11. Classes de bactérias e fungos detetadas nas amostras com prevalência superior a 50%.	73
Tabela 12. Comparação entre os <i>clusters</i> de bactérias (B1-B4) e de fungos (F1-F3), e os dados ambientais recolhidos.	76

Índice de figuras

Figura 1. Fotografias e imagens alusivas ao sobreiro	21
Figura 2. Dados e estatísticas sobre o sobreiro e a cortiça	22
Figura 3. Recortes de plataformas digitais sobre o sobreiro e o estado do montado	24
Figura 4. Declínio crónico do carvalho.....	25
Figura 5. A espiral do declínio da árvore de Manion	26
Figura 6. Declínio agudo do carvalho	28
Figura 7. Linha cronológica com alguns dos principais eventos ligados ao Declínio Agudo do Carvalho.	30
Figura 8. Limites geográficos das zonas de intervenção florestal	34
Figura 9. Amplificação da região de interesse para sequenciação.....	36
Figura 10. Plataforma de sequenciação de última geração <i>MiSeq Illumina</i> [®]	38
Figura 11. Fluxo de trabalho de sequenciação de última geração da <i>Illumina</i> [®]	40
Figura 12. Esquema sumário da sequenciação.. ..	42
Figura 13. Classes de deterioração foliar refletindo a percentagem de desfolhamento da árvore	46
Figura 14. Eletroforese dos produtos de PCR resultantes da amplificação da região V3-V3 do 16s rRNA.....	56
Figura 15. Eletroforese dos produtos de PCR resultantes da nova amplificação da região V3-V3 do 16s rRNA	57
Figura 16. Eletroforese dos produtos de PCR resultantes da amplificação da região ITS2	57
Figura 17. Curvas de rarefação de bactérias e de fungos	58
Figura 18. Dendrogramas das amostras de bactérias e de fungos	61
Figura 19. Análise de beta diversidade de bactérias e fungos.....	62
Figura 20. Abundância relativa das famílias de bactérias encontradas no controlo do <i>kit</i>	63
Figura 21. Abundância relativa das famílias de fungos encontradas no controlo do <i>kit</i>	64
Figura 22. Histogramas e gráficos quantil-quantil dos dados de bactérias e fungos, antes e após transformação logarítmica	65
Figura 23. Bibliotecas de bactérias e fungos do microbioma das lesões na casaca no sobreiro	66
Figura 24. Gráfico de barras empilhadas da abundância relativa de bactérias a nível do filo.	68

Figura 25. Gráfico de barras empilhadas da abundância relativa de bactérias a nível da família, obtido no <i>MicrobiomeAnalyst</i>	69
Figura 26. Gráfico de barras empilhadas da abundância relativa de fungos a nível do filo	71
Figura 27. Gráfico de barras empilhadas da abundância relativa de fungos a nível da família	72
Figura 28. Classes de bactérias e fungos detetadas nas amostras com pelo menos 95% de prevalência e 0,10% de abundância relativa.	73
Figura 29. <i>Heatmap</i> do <i>clustering</i> hierárquico das amostras de fungos a nível taxonómico das classes.	74
Figura 30. <i>Heatmap</i> do <i>clustering</i> hierárquico das amostras de bactérias a nível das classes.	75
Figura 31. Dendrogramas das amostras de bactérias e de fungos.	76
Figura 32. Análise de beta diversidade de bactérias	77
Figura 33. Análise de beta diversidade de fungos	78
Figura 34. <i>Boxplots</i> das diferentes métricas de alfa diversidade entre os <i>clusters</i> de bactérias, a nível da família.....	79
Figura 35. <i>Boxplots</i> das diferentes métricas de alfa diversidade entre os <i>clusters</i> de fungos a nível da família.....	80
Figura 36. Gráfico de barras empilhadas da abundância relativa dos clusters de bactérias a nível das famílias	81
Figura 37. Gráfico de barras empilhadas da abundância relativa dos clusters de fungos a nível das famílias	82
Figura 40. Análise de beta diversidade de bactérias da região do Divor	84
Figura 41. Análise de beta diversidade de fungos da região do Divor	84
Figura 42. Métricas de alfa diversidade entre as bactérias das duas regiões, a nível da família	85
Figura 43. Métricas de alfa diversidade entre os fungos das duas regiões, a nível da família	85
Figura 44. Gráfico de barras empilhadas da abundância relativa de bactérias agrupadas por região a nível das famílias	86
Figura 45. Gráfico de barras empilhadas da abundância relativa de fungos agrupadas por região a nível das famílias.....	87

Lista de abreviaturas

%ID	Porcentagem de Identidade
16S rRNA	Gene RNA Ribossomal 16S
A	Adenosina
ANOVA	Análise de Variância (<i>Analysis of variance</i>)
ANOSIM	Análise de Semelhanças (<i>Analysis of similarities</i>)
AOD	Declínio Agudo do Carvalho (<i>Acute Oak Decline</i>)
APCOR	Associação Portuguesa de Cortiça
BIOM	Matriz Biológica de Observação (<i>Biological Observation Matrix</i>)
bp	Pares de bases (<i>Base Pairs</i>)
C	Citosina
CLR	<i>Centered log ratio</i>
CNC	Centro de Neurociências e Biologia Celular
COD	Declínio Crônico do Carvalho (<i>Chronic Oak Decline</i>)
CSS	<i>Cumulative Sum Scaling</i>
DNA	Ácido Desoxirribonucleico (<i>Deoxyribonucleic acid</i>)
dNTPs	Desoxirribonucleótido trifosfato (<i>Deoxynucleotide triphosphates</i>)
FAO	Organização das Nações Unidas para a Alimentação e a Agricultura (<i>Food and Agriculture Organization</i>)
G	Guanina
GSC	Consórcio de Padrões Genômicos (<i>Genomic Standards Consortium</i>)
ICNF	Instituto da Conservação da Natureza e das Florestas
ITS2	Espaçador Interno Transcrito 2 (<i>Internal Transcribed Spacer 2</i>)
NGS	Sequenciação de Última Geração (<i>Next-Generation Sequencing</i>)
nMDS	<i>Non-Metric Multidimensional Scaling</i>
OTU	Unidade Taxonômica Operacional (<i>Operational Taxonomic Unit</i>)
PCoA	Análise de Coordenadas Principais (<i>Principal Coordinate Analysis</i>)
PCR	Reação em Cadeia da Polimerase (<i>Polymerase Chain Reaction</i>)
PERMANOVA	Análise de Variância Multivariada Permutacional (<i>Permutational Multivariate Analysis of Variance</i>)
PERMDISP	Análise de dispersões permutacionais (Permutational analysis of multivariate dispersions)
PGF	Plano de Gestão Florestal
RLE	Expressão logarítmica relativa (<i>Relative Log Expressions</i>)

RNA	Ácido Ribonucleico (<i>Ribonucleic acid</i>)
T	Timina
TMM	<i>Trimmed mean of M-values</i>
TSS	<i>Total sum scaling</i>
UQ	<i>Upper-quantile normalization</i>
V3	Região Hipervariável 3
V4	Região Hipervariável 4
ZIF	Zona de Intervenção Florestal

1. Introdução

1.1. *Quercus suber* L.

O sobreiro, *Quercus suber* L. (figura 1A-D), é uma angiospérmica dicotiledónea, um carvalho esclerófilo perene, da família Fagaceae. Nativo da região mediterrânica ocidental (figura 1E) (Moricca et al., 2016), é aqui que encontra as condições ideais para o seu crescimento, quer vegetativamente, quer semeado ou plantado. Com a particularidade de ter sido a base da descoberta da célula, em 1665, por Robert Hooke, do tronco do sobreiro é, a cada nove anos, retirada a cortiça (figura 1B). Esta técnica, que não danifica a árvore, o chamado descortiçamento (figura 1C), representa o trabalho agrícola mais bem pago do mundo. Principal diferença em relação aos outros carvalhos, a cortiça, composta sobretudo por suberina, é uma matéria-prima natural com propriedades singulares que lhe conferem um carácter inigualável, sendo simultaneamente biodegradável, renovável e reciclável.

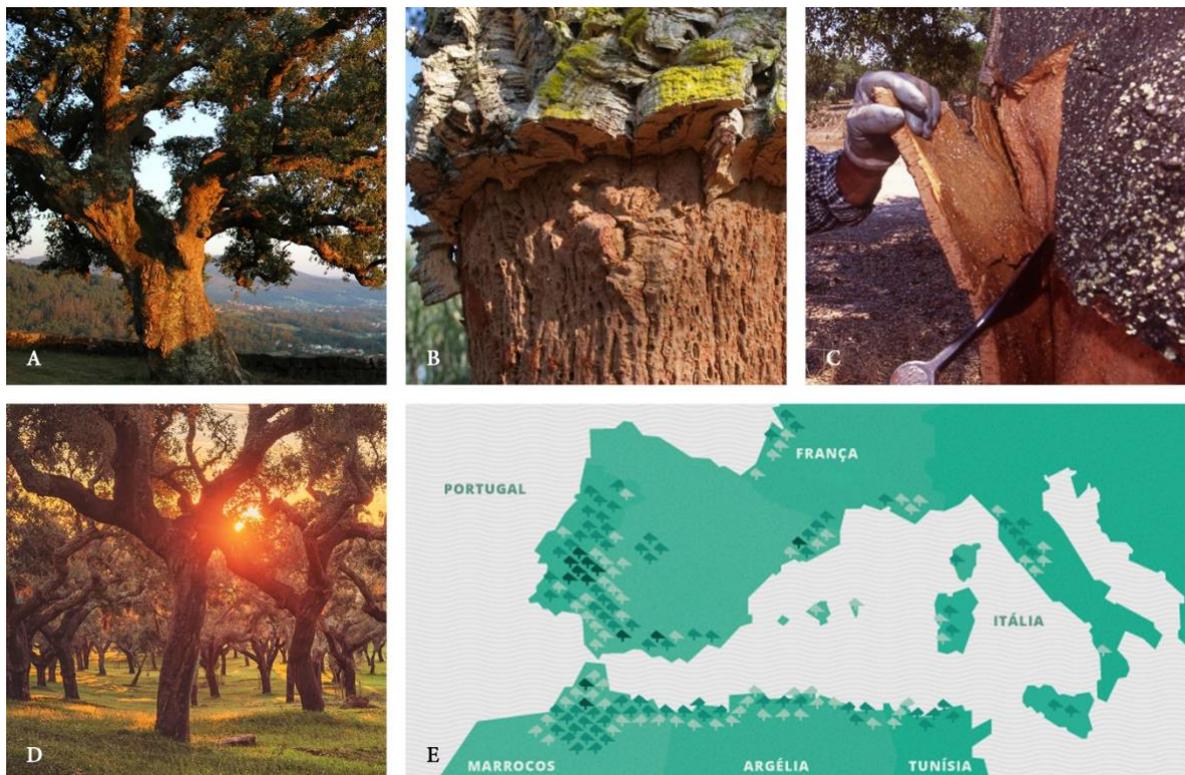


Figura 1. Fotografias e imagens alusivas ao sobreiro (*Quercus suber* L.). **A.** Fotografia de um sobreiro adulto (Elentir, 2011). **B.** Pormenor de um sobreiro após descortiçamento, onde se pode observar um corte transversal da cortiça (Muniz, s.d.). **C.** Fotografia onde se capta o momento de descortiçamento de um sobreiro (APCOR, s.d. b). **D.** Panorama típico de um montado (Haker, 2013). **E.** Distribuição e ocupação do *Quercus suber* L. em território nacional e na região mediterrânica (adaptado) (FAO, 2010).

1. Introdução

Em Portugal, destaca-se no panorama florestal, estando protegido por lei desde o século XIII e sendo inclusivamente considerado a árvore nacional desde 2011, pelo Projeto de Resolução da Assembleia da República nº 123/XII, de 22 de dezembro de 2011. Portugal detém a liderança mundial sobre a área de sobreiro, com 740 mil *ha* dos 2,2 milhões *ha* existentes no planeta, sobre a produção de cortiça, com 63% do volume exportado mundialmente e, ainda, sobre o processamento de cortiça, com 74% de toda a cortiça processada (Pereira-Leal et al., 2014).

No passado, os carvalhos, principalmente o sobreiro, dominavam a floresta autóctone portuguesa, mas a sua área diminuiu rapidamente com a atividade humana. Ainda assim, esta folhosa histórica, existente em todo o território nacional, representa 23% da floresta portuguesa (ICNF, 2013), na qual forma povoamentos característicos, o montado de sobreiro (figura 1D), onde o sobreiro coexiste com culturas agrícolas ou de pastagem (Camilo-Alves et al., 2017; Costa et al., 2010). Estes e outros dados estatísticos sobre o sobreiro e o montado, e a indústria e o comércio da cortiça, podem ser encontrados na figura 2, construída a partir de dados provenientes da Associação Portuguesa de Cortiça (APCOR).

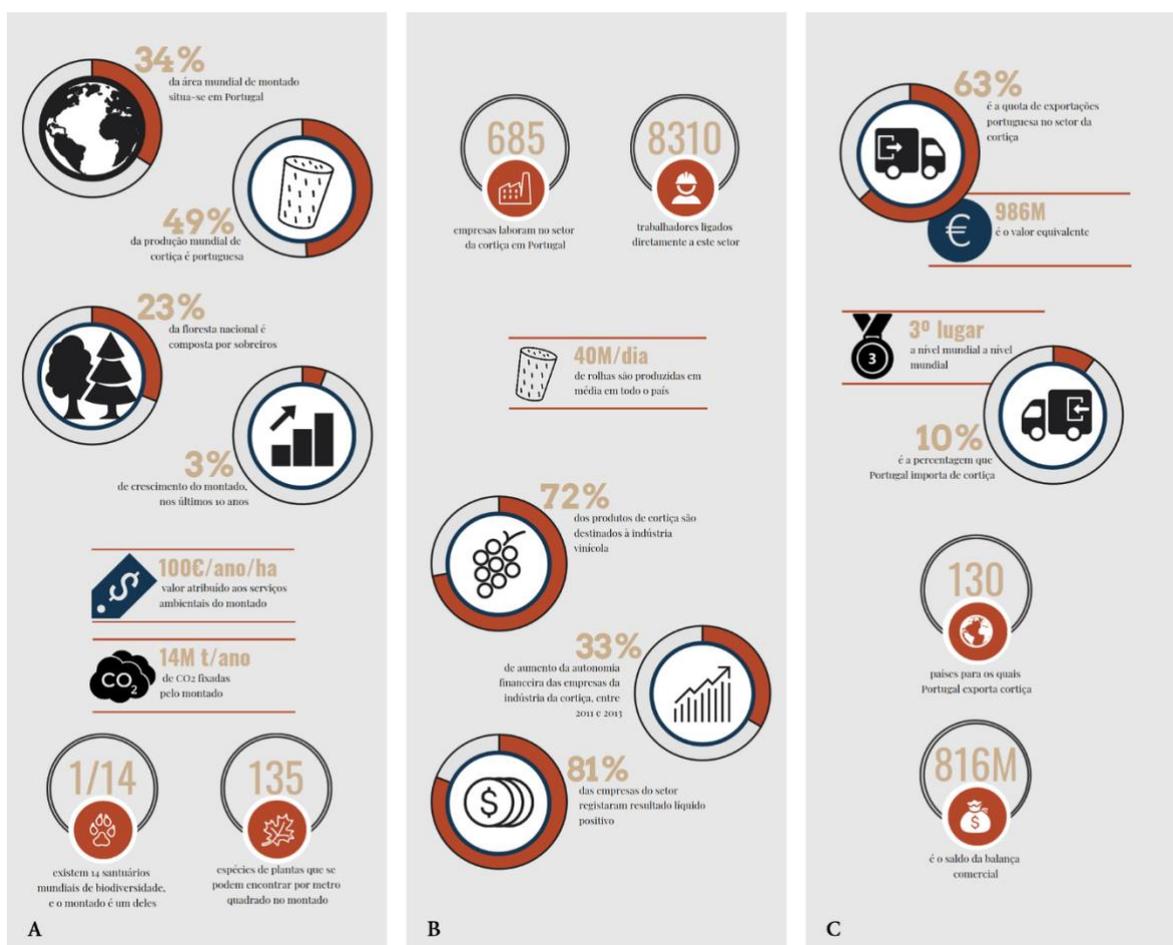


Figura 2. Dados e estatísticas sobre o sobreiro e a cortiça (adaptado) (APCOR, s.d. a). **A.** Informação diversa sobre o montado e os seus serviços. **B.** Alguns números relativos à indústria da cortiça. **C.** Factos sobre o comércio externo do setor da cortiça e o seu rendimento.

Unanimemente consagrado em Assembleia da República, o sobreiro representa uma grande importância económica, social e ambiental, pois, para além de suportar uma das maiores indústrias do país, a fileira da cortiça, que envolve um número substancial de empresas e de postos de trabalho (Kim et al., 2017), assume-se também como um contributo fundamental contra o êxodo rural e desertificação do interior português, simultaneamente preservando a elevada biodiversidade rica em endemismos que este tipo de florestas heterogêneas abriga (Kim et al., 2017; Moricca et al., 2016).

1.2. Ameaças

A progressiva redução do preço de mercado da cortiça está a contribuir para o crescente abandono das florestas de sobreiro (figura 3A) e, em sentido oposto, a alta pressão antropogénica através da exploração excessiva de recursos florestais, sobretudo sobrepastoreio e corte irregular de lenha (figura 3B), têm contribuído para a degradação de ecossistemas de cortiça dentro de algumas áreas (Kim et al., 2017; Moricca et al., 2016). Aliado à falta de uma gestão apropriada (figura 3C), faz com que estes frágeis ecossistemas sejam rapidamente invadidos e cobertos por espécies mediterrâneas de matos que, por sua vez, promovem um aumento do risco de incêndios e perda de habitats e biodiversidade (Moricca et al., 2016).

Apesar da sua resiliência na superação de episódios de perturbação, as árvores enfrentam stresses que desencadeiam o seu declínio (figura 3D). A atual preocupação global com doenças e pragas de árvores tem vindo a aumentar, uma vez que ameaçam não só estes, bem como outros ecossistemas e os serviços essenciais fornecidos pelos mesmos (figura 3E).

Durante as últimas décadas, têm ocorrido eventos de mortalidade de sobreiros, com um aumento da sua frequência, dentro de várias florestas de sobreiro na bacia mediterrânica, o que levou a uma perda progressiva dos montados, comprometendo o valor socioeconómico da cortiça (figura 3F), bem como a biodiversidade única das florestas mediterrânicas (Moricca et al., 2016).

A mortalidade do sobreiro é um problema cada vez mais recorrente no sudoeste de Portugal (Costa et al., 2010), e tem gerado grande preocupação em relação à sustentabilidade destas florestas (figura 3G, H). É fulcral que se mantenham como fonte de rendimento para agricultores, através das atividades agrícola, silvícola, e pastorícia (Moricca et al., 2016) e, sobretudo, que se assegure o fornecimento de matérias-primas para a indústria da cortiça (Costa et al., 2010).

Adicionalmente, está previsto um agravamento deste cenário (figura 3I), antevendo-se, inclusive, que as doenças do carvalho se tornem mais severas (figura 3J) e se

1. Introdução

expandam para norte e leste europeus, dentro do próximo século (Pereira-Leal et al., 2014).

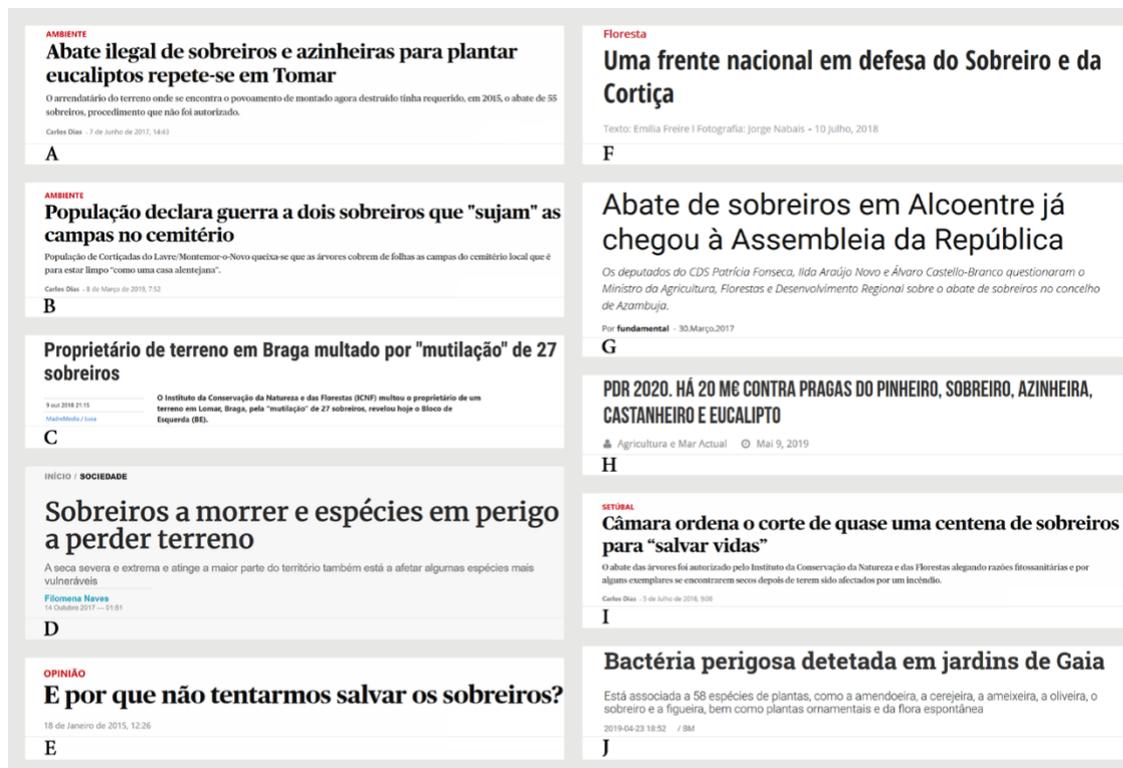


Figura 3. Recortes de plataformas digitais sobre o sobreiro e o estado atual do montado. **A.** Notícia do jornal *Público* sobre o corte ilícito de sobreiros em Tomar (Dias, 2017). **B.** Notícia do jornal *Público* acerca da intenção de derrubar sobreiros em Montemor-o-Novo por parte da população (Dias, 2019). **C.** Notícia do meio de comunicação social *Sapo 24* relativa a coima aplicada a cidadão bracarense por gestão inapropriada de terrenos (Lusa, 2018). **D.** Notícia do *Diário de Notícias* sobre a perda de biodiversidade no montado alentejano provocada pela seca (Neves, 2017). **E.** Artigo de opinião do jornal *Público* sobre o sobreiro como um importante ativo natural (Pinto, 2015). **F.** Publicação da revista *Vida Rural* sobre a cortiça, o sobreiro e o futuro do montado (Freire, 2018). **G.** Notícia do jornal digital *Fundamental* sobre as preocupações colocadas ao Ministro da Agricultura, Florestas e Desenvolvimento Regional, devido ao abate de sobreiros no concelho de Azambuja (Silva, 2017). **H.** Notícia na revista digital *Agricultura e Mar Actual* relativa aos apoios disponibilizados pelo Governo para reforçar a defesa da floresta contra danos causados por agentes bióticos (Sá, 2019). **I.** Notícia do jornal *Público* sobre o abate de árvores em Azeitão, autorizado por razões fitossanitárias (Dias, 2018). **J.** Reportagem do canal de informação geral *tvi24* sobre a identificação de uma bactéria patogénica em jardins particulares e espaços públicos de Vila Nova de Gaia (Mendes, 2019).

1.3. Declínio do sobreiro

O sobreiro, em particular, sofre de algumas doenças importantes, tendo sido relatadas na literatura pelo menos 100 espécies de fungos patogénicas (Moricca et al., 2016), bem como outras espécies de bactérias, nesta árvore.

Uma das principais enfermidades que tem afetado os sobreiros é o declínio do carvalho. Esta doença é considerada uma patologia multifatorial, na qual diversos fatores abióticos e bióticos, extremamente variáveis, estão envolvidos e em interação, como a

seca, a geada, as pragas de insetos, e os patógenos (Camilo-Alves et al., 2017; Moricca et al., 2016).

O declínio do sobreiro está inserido numa ampla síndrome de declínio de carvalhos, onde se reconhecem condições distintas e identificáveis com base em sintomas inconfundíveis, o Declínio Crónico do Carvalho (COD) e o Declínio Agudo do Carvalho (AOD). A Morte Súbita do Carvalho, uma doença onde há morte repentina da árvore, caracterizada pela secagem rápida da copa, embora apresente um nome sugestivo ao grupo de doenças do declínio de carvalho, não faz parte deste, sendo causada por uma infeção pelo fungo *Phytophthora ramorum*, em várias espécies vegetais (Harmon e Harmon, 2014).

1.3.1. Declínio Crónico do Carvalho

O COD, também conhecido como *dieback*, caracteriza-se por um declínio progressivo com perda gradual do vigor da folhagem, a partir das extremidades para o centro, que conduz à morte da árvore (figura 4C). Começou a surgir na literatura no início do século XX, mas, desde os anos 20 até ao final da década de 80, apenas era mencionada esporadicamente, tendo-se agravado nos últimos tempos.

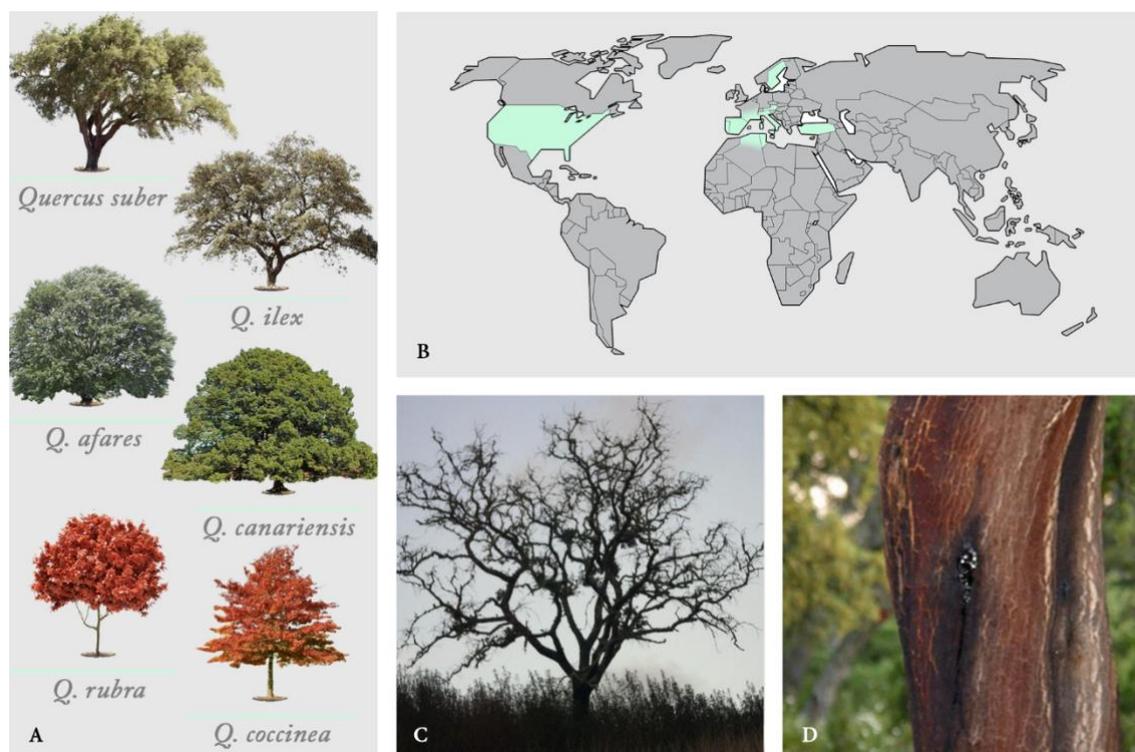


Figura 4. Declínio crónico do carvalho **A.** Espécies de carvalhos afetadas pela condição. **B.** Incidência geográfica da doença. **C.** Carvalho sintomático (Kaltenbach, 2019). **D.** Sangramento de um exsudato negro através de uma lesão na casca de um carvalho (Moricca et al, 2016).

Árvores afetadas, predominantemente carvalhos mediterrâneos (*Q. suber* L. e *Q. ilex* L.), do norte de África (*Q. afares* Pomel e *Q. canariensis* Willd.) e da América do Norte

1. Introdução

(*Q. rubra* L. e *Q. coccinea* Münchh.) (figura 4A), desenvolvem sintomas ao longo de décadas, como danos e morte da copa, murchidão de folhas, podridão e perda de raízes e, em alguns casos, apresentam um fluido negro a escorrer pela casca (figura 4D) (Costa et al., 2010; Denman et al., 2010; Smahi et al., 2017).

Apesar da grande distribuição geográfica do COD (figura 4B), o seu processo básico e sintomas são similares em todos os lugares onde ocorre esta patologia: as árvores são afetadas por agentes causadores de stresse, as folhas secam, tornam-se suscetíveis a organismos nocivos, são atacadas por eles, declinam e acabam por morrer (Gottschalk e Wargo, 1996). De um modo geral, crê-se que há a atuação sequencial ou simultânea de fatores abióticos que levam a um aumento do stresse hídrico e a uma diminuição da capacidade de defesa da árvore, estabelecendo condições favoráveis para a infecção por agentes bióticos, aos quais seriam normalmente capazes de resistir (Camilo-Alves et al., 2017; Gottschalk e Wargo, 1996; Smahi et al., 2017).

Efetivamente, e ao contrário das doenças de árvores mais comuns, o COD não é causado por uma única peste ou patógeno. É uma síndrome complexa, que envolve os efeitos sequenciais, combinados e cumulativos de uma série de agentes bióticos e abióticos, seguindo o modelo de espiral desenvolvido por Manion em 1981 (figura 5).

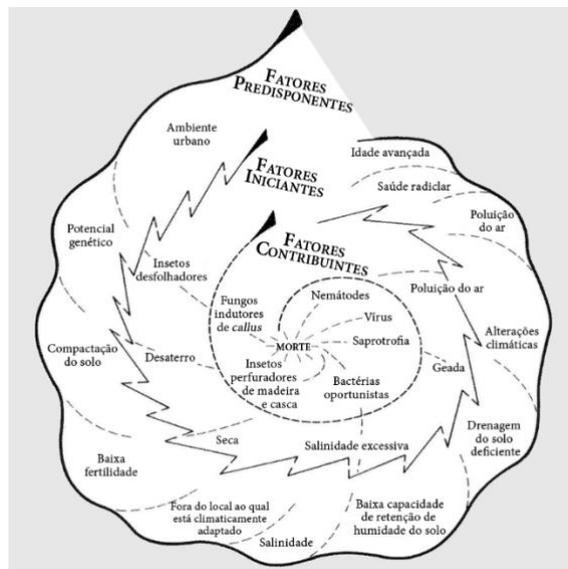


Figura 5. A espiral do declínio da árvore de Manion (adaptado) (Manion, 1981).

É, portanto, difícil explicar de forma clara a origem e as causas deste fenómeno, dado tratar-se de um processo gradual e multifatorial (Costa et al., 2010; Smahi et al., 2017), mas, ainda assim, têm sido apontados como motivos principais, os fenómenos ambientais e as condições edafoclimáticas, como secas severas, cheias prolongadas, flutuações da disponibilidade de água do solo e invernos rigorosos que, para além de afetarem diretamente a árvore, providenciam condições mais favoráveis para a

sobrevivência e desenvolvimento de patógenos (Camilo-Alves et al., 2017; Costa et al., 2010; Kim et al., 2017). Para além destes, foi exposto que o sobrepastoreio e o impacto de atividades humanas desempenham também um papel nesta doença (Kim et al., 2017).

Mesmo que o fator que despoleta o COD seja o stresse abiótico, são agentes bióticos secundários que provocam a seca das folhas e morte das árvores afetadas. Geralmente é causado por insetos ou fungos que atacam árvores enfraquecidas, mas que raramente o conseguiriam fazer com sucesso em árvores saudáveis (Gottschalk e Wargo, 1996).

Entre estes organismos, que desempenham um importante papel ecológico ao matar árvores enfraquecidas, são destacados os fungos patogénicos *Diplodia corticola*, *Biscogniauxia mediterranea*, *Collybia fusipes*, *Armillaria spp.* e *Phytophthora spp.* (Denman et al., 2010; Kaltenbach; Moricca et al., 2016). A todos eles tem sido atribuído um papel proeminente na mortalidade de carvalhos em declínio, nos mais variados contextos, sendo associados ao declínio geral da floresta de carvalhos na Península Ibérica, no sul da França, na Itália, Áustria, Turquia e Suécia e no oeste da Argélia (figura 4B) (Kaltenbach; Kim et al., 2017; Moricca et al., 2016; Smahi et al., 2017).

O primeiro fungo mencionado foi considerado o patógeno mais virulento do sobreiro, com ataques descritos em Itália, Marrocos, Portugal, Espanha e Tunísia. Já o segundo, tem emergido como um fungo patógeno oportunista e potencialmente invasivo, cuja frequência dos seus ataques em carvalhos tem aumentado significativamente ao longo das últimas décadas, dentro da área mediterrânica (Moricca et al., 2016). Os restantes, tratam-se de fungos que invadem sobretudo as raízes de carvalhos, causando podridão radicular (Denman et al., 2010; Moricca et al., 2016), particularmente, o *Phytophthora cinnamomi* considerado um dos fitopatogénicos mais devastadores e uma das piores espécies invasoras a nível mundial (Moricca et al., 2016; Smahi et al., 2017).

Além destes, outros fungos patogénicos novos e emergentes, que incluem os agentes indutores de formação de *callus* *Coryneum modonium*, *Neofusicoccum parvum*, e vários fungos do género *Raffaelea*, têm contribuído para o alargamento da lista de fungos invasores de sobreiros (Denman et al., 2014; Moricca et al., 2016).

1.3.2. Declínio Agudo do Carvalho

O AOD (figura 6A) é uma doença relativamente nova que afeta ambas as espécies de carvalho nativas da Grã-Bretanha, *Q. robur* L. e *Q. petraea* (Matt.) Liebl. (figura 6B) (Denman et al., 2014). Mais agressiva que a anterior, os sintomas desenvolvem-se mais rapidamente e resultam em maiores taxas de mortalidade, quatro ou cinco anos após o aparecimento dos primeiros indícios (Denman et al., 2010, 2014).

1. Introdução

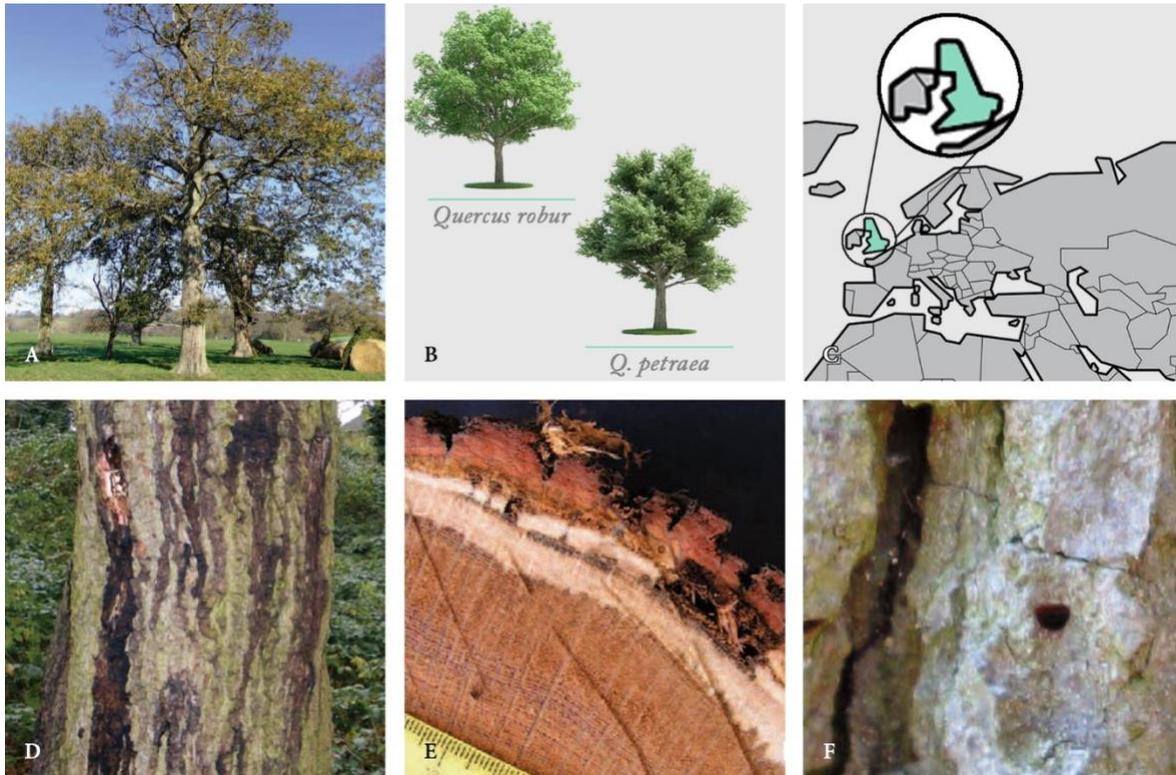


Figura 6. Declínio agudo do carvalho **A.** Sobreiro sintomático (Denman et al., 2010). **B.** Espécies de carvalhos afetadas pela condição. **C.** Distribuição geográfica da patologia. **D.** Sangramentos no tronco de uma árvore doente (Denman et al., 2014). **E.** Corte transversal que mostra uma cavidade que irá dar origem a um ponto de sangramento (Denman et al., 2010). **F.** Buracos em forma de “D” provocada pela larva de *Agrilus biguttatus* (Denman et al., 2010).

Esta doença, encontrada sobretudo nos *Midlands* de Inglaterra e País-de-Gales, mas também na Escócia (figura 6C), é difícil de ser diagnosticada, ainda que apresente uma sintomatologia característica, como os sangramentos (figura 6D), necroses e fendas na casca, galerias larvares e buracos de saída em forma de “D” (figura 6F), provocados pelas larvas oportunistas de *Agrilus biguttatus*, muitas vezes associadas a esta doença (Denman et al., 2010, 2014).

Surge em árvores normalmente com mais de 50 anos, é altamente contagiosa, e pensa-se que seja causada por bactérias patogénicas, que afetam, não só a folhagem, mas sobretudo o tronco dos carvalhos, que passam a apresentar um grande número de sangramentos, onde um líquido negro pegajoso é libertado por manchas necróticas sob as placas de cortiça (Denman et al., 2010).

No tronco de árvores sintomáticas, surgem rasgões, tipicamente longitudinais, com cinco a dez centímetros de comprimento, que se formam nas fendas entre as placas de cortiça (Denman et al., 2010, 2014). Nestes locais, a casca interna rompe-se, deixando uma cavidade (figura 6E) cheia de líquido, onde irá surgir um ponto de sangramento. Sangramentos estes, geralmente encontrados um a dois metros acima do solo, mas que podem também ocorrer às dezenas ao longo do tronco até ao topo da copa das árvores (Denman et al., 2010, 2014).

Para além destes exsudatos característicos, costuma também ser visível a formação de *callus*, uma massa desorganizada de parênquima que surge por tentativa da planta cobrir a lesão (Denman et al., 2010, 2014). Paralelamente, espécies de besouros oportunistas de *Agilus biguttatus* exploram e aproveitam-se do estado enfraquecido dos carvalhos, deixando visíveis orifícios de saída em forma de “D”, provocados pelas suas larvas após se alimentarem da casca interna da árvore durante a fase de pupa (Denman et al., 2010, 2014).

Ambos os declínios, COD e AOD, provocam sintomas semelhantes, principalmente porque o primeiro pode eventualmente apresentar fluídos a escorrer pelo tronco – aqui incluindo o efeito das doenças causadas pelos fungos que tiram partido do estado enfraquecido da árvore – e porque o segundo pode também apresentar *dieback* da canópia. No entanto, os sangramentos do AOD são bastante singulares, e os danos da folhagem são mais severos e podem aparecer somente quando a árvore está perto da morte (Denman et al., 2010). Assim, e ao contrário do COD, não é a folhagem, nem a raiz, que é principalmente afetada, ao invés, é o tronco da árvore o principal alvo.

A origem da doença também é bastante diferente entre ambas. Enquanto que a primeira surge fortemente associada a agentes abióticos causadores de stresse, neste caso crê-se que são bactérias patogénicas as entidades responsáveis por provocar as lesões necróticas sobre a casca e conduzir aos sangramentos no tronco (Denman et al., 2010).

O primeiro episódio de AOD na Grã-Bretanha, aconteceu nos anos 20, e a sua causa foi na altura atribuída a efeitos colaterais de desfolhas sucessivas provocadas pela lagarta *Tortrix viridana*, seguidas por ataques de *Erysiphe alphitoides*, fungo que provoca o oídio em carvalhos (Brady et al., 2010; Denman e Webber, 2009). A primeira evidência convincente de sangramentos no tronco induzidos por bactérias, apenas foi reportada em 1997, na casca interna do *Q. ilex*, em Espanha (Brady et al., 2010; Denman e Webber, 2009). No entanto, somente quando esta doença teve novo surto na Grã-Bretanha, em 2009, é que foi descrita e distinguida dos restantes declínios do carvalho pela primeira vez (Denman e Webber, 2009).

Embora o distúrbio tivesse ocorrido no *Q. ilex*, a descoberta de 1997 impulsionou a ideia de que bactérias poderiam desempenhar um papel importante nos sangramentos do tronco de carvalhos afetados pelo AOD. Simultaneamente, deixou patente que o AOD, poderia também ser encontrado em diferentes espécies de carvalhos europeus, como por exemplo no já mencionado *Q. ilex* (Brady et al., 2010; Denman e Webber, 2009), mas também no *Q. cerris* (Sapp et al., 2016) ou no *Quercus pyrenaica* (Denman et al., 2016), existindo, inclusive, casos reportados no Médio Oriente e no continente americano

1. Introdução

(Denman et al., 2018). Um epítome destes, e de outros acontecimentos ligados ao AOD pode ser observado na figura 7.



Figura 7. Linha cronológica com alguns dos principais eventos ligados ao Declínio Agudo do Carvalho.

De facto, têm sido isoladas bactérias a partir de carvalhos com sintomas de AOD de uma forma recorrente. *Gibbsiella quercinecans* (Brady et al., 2010), *Brenneria goodwinii* (Denman et al., 2012), *Lonsdalea quercina* ssp. *britannica* (Brady et al., 2012), *Brenneria roseae* ssp. *roseae* (Brady et al., 2014a), *Rahnella victoriana* (Brady et al., 2014b), *Rahnella variigena* (Brady et al., 2014b) e *Gibbsiella greigii* (Brady et al., 2014c), estão entre as treze espécies e dois géneros de bactérias (Denman et al., 2016) que foram descobertas pela primeira vez, ou renomeadas, após terem sido identificadas em tecidos de árvores com esta condição. No mesmo *modus operandi*, insere-se uma *Pseudomonas* ainda por nomear (Denman et al., 2016; Sapp et al., 2016), e várias outras novas espécies que estão a ser formalmente descritas (Denman et al., 2018).

Desde então, estas bactérias foram, por mais do que uma vez, identificadas nas imediações de lesões necróticas de árvores sintomáticas (Brady et al., 2010, 2014a, 2014c, 2014b, 2012; Broberg et al., 2018; Costa et al., 2010; Denman et al., 2018, 2012, 2014, 2016; Doonan et al., 2015, 2019; Sapp et al., 2016). Crê-se que estas espécies desempenham um papel fundamental no AOD, com a formação de necrose nos tecidos, provavelmente via enzimas necrogénicas ou sistemas de secreção associados a proteínas efetoras (Doonan et al., 2015).

Estes diagnósticos iniciais, revelaram alta incidência de bactérias através de técnicas de isolamento e cultura de microrganismos, mas neles, estava também presente uma vasta panóplia de fungos, muitos dos quais saprófitas generalistas (Denman et al., 2016). No entanto, estudos mais recentes apontam que fungos, animais (com a exceção do *A. biguttatus*) e vírus, não estão significativamente associados ao AOD (Broberg et al., 2018; Denman et al., 2018).

Todos estes dados apontam para a necessidade de considerar o AOD uma preocupação mundial que, devido à falta de evidência empíricas e à sua complexa

natureza, continua a ludibriar investigadores na tentativa de identificar os agentes causadores das lesões e a extensão real deste problema, o que constitui uma barreira para o desenvolvimento de estratégias de gestão informadas para esta doença (Denman et al., 2018).

1.4. Objetivos

É com estas premissas que este projeto visa principalmente caracterizar as comunidades bacterianas e fúngicas de lesões na casca de sobreiros em declínio, descrevendo extensivamente o microbioma dos seus exsudatos negros. Informação que tem o propósito de permitir perceber se estas lesões, observadas e analisadas em Portugal, são semelhantes às descritas no AOD na Grã-Bretanha.

Detetar e identificar as bactérias e os fungos mais relevantes nas comunidades destas características lesões no tronco, poderá vir a desvendar potenciais microrganismos patogénicos envolvidos na sua formação, podendo estes revelarem-se como potenciais micróbios de interesse biotecnológico, por exemplo, com enzimas capazes de degradarem paredes celulares vegetais.

Ao entender o cenário de declínio deste fenómeno e as suas causas, novas direções de estudos de interesse abrir-se-ão. Poderão ser desenvolvidos novos diagnósticos baseados em DNA que contribuam para uma mais rápida e precisa deteção desta doença, e estratégias de controlo e de gestão apropriadas para os agentes mais deletérios, de modo a evitar ou minimizar esta rápida devastação de extensos ecossistemas de sobreiros.

2. Estudo do Microbioma

Os relatos da ocorrência do declínio de carvalho datam, portanto, desde décadas atrás, mas a preocupação sobre o aumento da deterioração da saúde destas espécies só se intensificou em toda a região europeia nos últimos 50 anos e tem continuado ao longo do novo milénio (Camilo-Alves et al., 2017; Denman et al., 2014). Apesar desta apreensão e do aparente aumento da incidência da perturbação, avanços no conhecimento sobre as causas e impactos desta condição têm sido lentos e isso deve-se parcialmente à natureza complexa do problema (Denman et al., 2014).

Além disso, as árvores não podem ser consideradas entidades independentes, é necessária uma visão holística para perceber que, uma vez que estas abrigam uma panóplia de microrganismos, tanto dentro, como fora dos seus tecidos, estes são parte integrante de si. Sobretudo fungos e bactérias estão envolvidos em funções vitais da árvore como a nutrição, saúde e resiliência, e a resistência a stresse biótico e abiótico, influenciando o crescimento e a sobrevivência da planta (Vandenkoornhuysen et al., 2015).

Nos últimos anos, o estudo de microbiomas, ou seja, a análise da diversidade de microrganismos associados à planta dentro do chamado microbioma vegetal, progrediu significativamente. Simultaneamente, descrições morfológicas detalhadas e modernas, acopladas com dados de sequenciação de DNA atualmente acessíveis em bases de dados públicas, estão a facilitar a identificação dos *taxa*. Assim, e uma vez que foram feitos poucos progressos na aplicação dos últimos avanços na sequenciação para caracterizar esta doença, é imprescindível realizar um estudo do microbioma destas lesões para perceber a natureza desta condição (Moricca et al., 2016).

Conduzir um estudo desta natureza não é trivial. Procedimentos experimentais, fatores ambientais e métodos de análise podem afetar os resultados (Knight et al., 2018). Apesar dos estudos de microbioma estarem a avançar rapidamente, não existe uma padronização dos métodos a utilizar, ainda que possam ser obtidos resultados bastante satisfatórios com as práticas atuais (Knight et al., 2018). Neste sentido, é necessário fazer escolhas e tomar decisões ao longo de todo o estudo, documentando cada uma delas, com particular atenção para o desenho experimental e estudo estatístico, de modo a trabalhar a grande quantidade de dados obtida com a elevada variedade de ferramentas computacionais disponível, e possibilitar a reprodutibilidade deste estudo.

Para documentar e descrever detalhadamente a composição das comunidades microbianas dos sobreiros em declínio e diferenças entre estados progressivos de doença socorremo-nos, portanto, de um estudo do microbioma.

2.1. Recolha de amostras

Na literatura, os microrganismos são tratados de acordo com o local da árvore em que ocorrem maioritariamente, sejam eles folhas, tronco ou raízes. Estes locais são abrigo de uma enorme diversidade microbiana que desempenha os mais variados papéis biogeoquímicos na saúde da árvore. Além disso, a recolha das amostras deve ser pensada de modo a que gere dados significativos que permitam a análise da estrutura e composição da comunidade, informação que constituirá as covariáveis em torno das amostras e que deverá ser posteriormente utilizada na análise de dados (Goodrich et al., 2014; Knight et al., 2018).

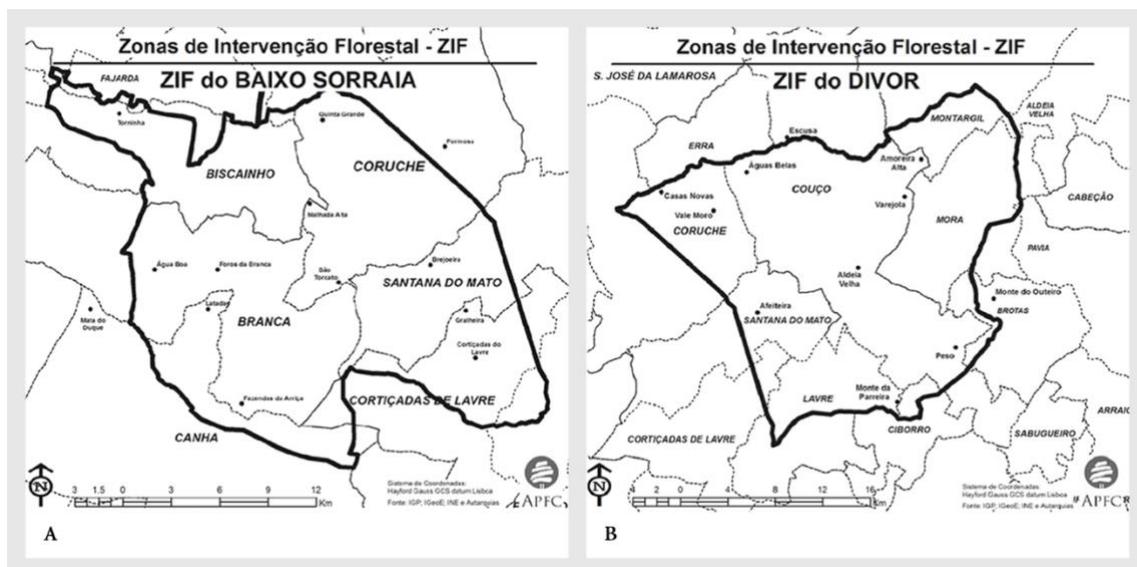


Figura 8. Limites geográficos das zonas de intervenção florestal. **A.** ZIF Baixo Sorraia, englobando vários prédios rústicos das freguesias de Coruche, Fajarda, Biscainho, Branca, Santana do Mato, Canha e Cortiçadas de Lavre, dos concelhos de Coruche, Montemor-o-Novo e Montijo (APFC, 2016a). **B.** ZIF Divor, englobando vários prédios rústicos das freguesias de Coruche, Couço, Montargil, Mora, Brotas, Caborro, Lavre, Cortiçadas de Lavre e Santana do Mato (APFC, 2016b).

Assim, e uma vez que uma das sintomatologias característica do declínio dos carvalhos, principalmente do AOD nos sobreiros, é a presença de um fluído negro a escorrer por lesões na cortiça, neste estudo são recolhidas amostras dos exsudatos presentes na casca externa de sobreiros, em duas zonas de intervenção florestal (ZIF), ZIF 240/10 Divor e ZIF 241/10 Baixo Sorraia (figura 8).

Segundo o Instituto da Conservação da Natureza e das Florestas (ICNF), uma ZIF é uma área territorial contínua e delimitada, maioritariamente constituída por espaços florestais, que está submetida a um Plano de Gestão Florestal (PGF) e que cumpre o estabelecido nos planos municipais de defesa da floresta contra incêndios, sendo administrada por uma única entidade (Guiomar et al., 2011).

As ZIF assumem particular relevância nos objetivos da política florestal nacional, assentes na dinamização da gestão florestal privada e na sua associação para uma gestão

comum, que visa garantir uma coordenação eficiente dos espaços florestais à escala da paisagem e uma aplicação coerente dos apoios públicos ao desenvolvimento florestal (Guiomar et al., 2011).

Dada a importância destas áreas, as ZIF constituem-se como um importante alvo deste tipo de estudos, uma vez que estes podem ajudar a entender algumas das causas do declínio da biodiversidade destes locais. Perceber que comunidades este exsudato abriga, pode revelar pormenores interessantes sobre os agentes causadores desta doença e quem sabe conduzir a estratégias de preservação a ser incluídas, por exemplo, num PGF destas ZIF.

2.2. Protocolos de DNA

É graças a avanços científicos, por nomes desde Johann Friedrich Miescher a James D. Watson e Francis Crick, passando por Charles Darwin e Gregory Mendel, que hoje sabemos que o DNA está presente praticamente em todas as células, de todos os organismos vivos, quer sejam animais, plantas, bactérias ou fungos. As últimas décadas têm sido marcadas por um grande número de estudos focados na compreensão da presença e distribuição destes microrganismos numa panóplia de ambientes apoiados, quer em métodos de cultura, quer em métodos moleculares por extração de DNA ambiental (Herrera e Cockell, 2007).

Independentemente da finalidade da investigação, estes convergem na motivação de explorar a diversidade microbiana das comunidades. Desta forma, podemos afirmar que a extração de DNA é muitas vezes o passo inicial na grande maioria dos processos laboratoriais de biologia molecular, e de todas as suas ómicas, assumindo-se como uma etapa fundamental para se obter alta eficiência de amplificação nos protocolos que usam a reação em cadeia da polimerase (PCR) (Oliveira et al., 2007). Seguindo esta tendência, o estudo do microbioma das lesões na casca do sobreiro pressupõe uma extração do DNA total para que este seja amplificado, purificado e posteriormente sequenciado.

Foram estabelecidos diversos métodos para isolar material genético a partir de materiais biológicos, este processo de extração pode ser realizado em qualquer tipo de organismo e através de diferentes tipos de amostra. Existem inclusive no mercado, diversos *kits* para extração de DNA, sendo imprescindível escolher o ideal para atender aos requisitos do estudo, tendo em conta o rendimento, degradação do material genético durante a extração, a sua eficiência em termos de custo e tempo, e simplicidade da metodologia. Dependendo do tipo de amostra, podem ser testados, adaptados e otimizados vários protocolos de forma a obter DNA de boa qualidade, mas,

2. Estudo do microbioma

independentemente disso, o objetivo da extração é proporcionar o material genético purificado pronto para ser utilizado em diferentes aplicações.

Entre várias fases de protocolos moleculares, é importante obter uma leitura rápida e precisa da concentração e pureza do DNA. Após extração do material genético, este deve ser quantificado por espectrofotometria e/ou fluorimetria para verificar a possibilidade de degradação do material genético, mas também para perceber a qualidade do produto de DNA resultante. Nos ácidos nucleicos, a absorvância máxima é de 260 *nm*, para proteína este valor fixa-se a 280 *nm*, sendo que a proporção de ácido nucleico/proteína (260/280) é geralmente usada como um indicador da pureza das amostras de DNA pelos métodos de espectrofotometria. Já os fluorímetros, fazem uso da quantificação fluorescente do DNA, através de um sinal de fluorescência que é medido quando um corante se liga a ácidos nucleicos.

O passo seguinte passa por amplificar as regiões do DNA que se pretendem sequenciar (figura 9), de modo a que apenas um subconjunto de genes ou regiões do genoma seja sequenciado para determinar relações filogenéticas.

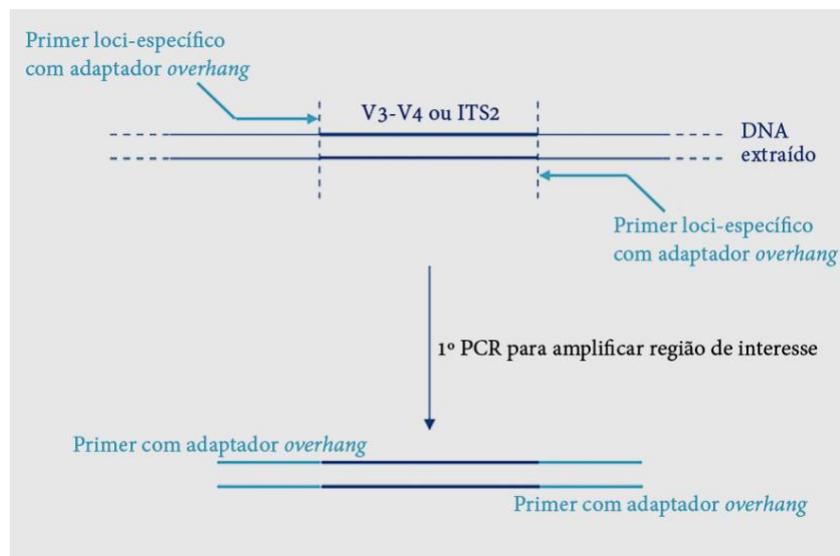


Figura 9. Amplificação da região de interesse para sequenciação. Este primeiro protocolo de PCR permite amplificar uma cadeia molde a partir das amostras de DNA, utilizando *primers* específicos para a região de interesse. A estes *primers*, são adicionadas sequências *overhang*, que são necessários para passos a formação da biblioteca de sequenciação.

(adaptado) (Wang, 2017)

Para isso, faz-se uso de genes marcador, genes que tipicamente possuem uma região polimórfica altamente variável, usada para identificação detalhada, que é flanqueada por regiões altamente conservadas, usadas como local de ligação de *primers* de PCR, o que permite que estes sejam utilizados para identificar a filogenia microbiana de amostras (Knight et al., 2018; de Muinck et al., 2017).

Para identificar e comparar bactérias presentes em amostras recolhidas, podem ser utilizadas a região hipervariável V3-V4 do gene RNA Ribossomal 16S gene (16S rRNA), e para fungos, a sub-região espaçadora interna transcrita 2 (ITS2) do DNA ribossomal nuclear. O 16S rRNA, com aproximadamente 1,5 mil pares de bases (bp), contém nove regiões variáveis intercaladas com regiões conservadas, que são frequentemente usadas em classificação filogenética de populações, fazendo com que seja o gene marcador mais amplamente utilizado em conjuntos de dados de referência usados para bactérias e Archaea (Goodrich et al., 2014; Knight et al., 2018; de Muinck et al., 2017). Enquanto que, o ITS2 é geralmente usado para obter a resolução taxonómica e filogenética de fungos (Goodrich et al., 2014; Knight et al., 2018; de Muinck et al., 2017).

Finalmente, segue-se a purificação dos produtos de PCR, que visa a limpeza da reação de amplificação. Esta é realizada de modo a remover *primers* que não reagiram, ou que hibridizaram entre si, estruturas que são indesejáveis na sequenciação, e podem resultar no fracasso de todos os ensaios a jusante deste.

2.3. Sequenciação de última geração

Desde há décadas que é reconhecido que o número de microrganismos observado, por recolha e coloração direta de amostras clínicas ou ambientais, excede em larga escala o número daqueles que podem ser cultivados em meio laboratorial a partir das mesmas amostras, aquilo que é reconhecido como “a grande anomalia da contagem de placas” (Fadrosh et al., 2014). Foi exatamente o desenvolvimento de métodos de amplificação e sequenciação independentes de meio de cultura, para detetar organismos delicados ou mesmo não-cultiváveis, que precipitou uma revolução em biologia (Fadrosh et al., 2014), desde a exploração de ramos profundos da Árvore da Vida, até ao delinear das fronteiras entre espécies, populações e comunidades (Shokralla et al., 2015).

Os avanços combinados de várias áreas tecnológicas culminaram naquilo que, há quase uma década, se descreve como sequenciação de última geração (NGS), a forma coloquial de descrever métodos de sequenciação altamente paralelos ou de alto rendimento, que produzem dados à escala do genoma ou para além dele (Levy e Myers, 2016). Desde o seu aparecimento no mercado em 2005, que estas tornaram amplamente disponível e rotineiro aquilo que vinha sendo feito até então, e a um custo decrescente de milhares de milhões de euros para milhares de euros, em menos de 25 anos (Levy e Myers, 2016).

A NGS assumiu-se como a tecnologia de sequenciação mais desenvolvida, sendo utilizada para identificar bactérias e fungos cultiváveis e não cultiváveis, permitindo, num único teste, analisar um ou vários genes ou mesmo o exoma completo, numa fração do

2. Estudo do microbioma

tempo e custo que as tecnologias anteriores permitiam. A introdução destas tecnologias levou mesmo a que o número de aplicações e os métodos que tiram partido deste poder de sequenciação aumentasse exponencialmente, tendo sido criadas dezenas de empresas, tecnologias e protocolos de sequenciação, o que culminou numa explosão da bioinformática como uma importante disciplina científica (Levy e Myers, 2016).

Nos últimos anos, estas tecnologias foram usadas e testadas num amplo mercado e numa ampla variedade de projetos de investigação, ajudando na maturação de métodos e instrumentos, e possibilitando uma diversidade de publicações, protocolos e aplicativos de tecnologia de sequenciação (Levy e Myers, 2016). Os resultados obtidos por estes métodos são distintos, todos têm pontos contra e a favor da sua utilização, e em função do custo, robustez, resolução e dificuldade do método, bem como das amostras e objetivos de um estudo desta índole, a NGS de genes marcador através da plataforma *MiSeq* *Illumina*[®] assume particular destaque (figura 10A).



Figura 10. A. Plataforma de sequenciação de última geração *MiSeq* *Illumina*[®]. **B.** Tecnologia de sequenciação emparelhada, que permite que ambas as extremidades do fragmento de DNA sejam sequenciadas. Como a distância entre cada leitura emparelhada é conhecida, os algoritmos de alinhamento podem usar essas informações para mapear as leituras em regiões repetitivas do genoma com maior precisão (adaptado) (Illumina, 2017).

Atualmente, com sequenciação de fragmentos com comprimento até 2 x 300 bp, alto rendimento e baixo custo, a plataforma *MiSeq* *Illumina*[®] é um dos sistemas mais usados mundialmente (Reuter et al., 2015; Schirmer et al., 2015). Esta tecnologia, estabeleceu-se como uma poderosa ferramenta de sequenciação, para uma caracterização rápida, adaptável e profunda da composição da comunidade microbiana (Fadrosh et al., 2014).

O incrível aumento na produção de dados e a redução no custo por base sequenciada, foram impulsionados principalmente pelo aumento da paralelização em

tecnologias como esta (Levy e Myers, 2016). Isto é, a *Illumina*[®] compensa os seus curtos comprimentos de sequências, com a sequenciação emparelhada (*Paired-End Sequencing*) (Figura 10B) de ambas as extremidades dos fragmentos de DNA de uma biblioteca, e o alinhamento das sequências *forward* e *reverse* como pares de sequências (Illumina, 2017). Cada extremidade da mesma molécula de DNA é sequenciada até ao comprimento total e, para além de serem produzidas o dobro do número de sequências pelo mesmo tempo e esforço de preparação, a informação final emparelhada melhora muito a taxa de alinhamento em comparação com a de sequências únicas (*single reads*) (Levy e Myers, 2016).

O campo da genómica contribuiu com dezenas de métodos para a sequenciação nas plataformas *Illumina*[®] (Levy e Myers, 2016), que de uma forma geral diferem em detalhes. Estes detalhes possibilitam estratégias específicas capazes de determinar a qualidade e a quantidade dos dados de sequenciação resultantes, e a utilidade da plataforma para aplicações específicas. Normalmente, todos eles seguem um paradigma geral semelhante, que começa pela preparação do DNA, envolve amplificação clonal e é finalizado por ciclos de sequenciação paralela massiva (Reuter et al., 2015). Neste sentido, o fluxo de trabalho de sequenciação da *Illumina*[®] é sempre composto por quatro etapas básicas: preparação da biblioteca (figura 11A), formação de *clusters* (figura 11B), sequenciação (figura 11C), e alinhamento e análise de dados (figura 11D).

Realizada de forma a organizar as amostras para sequenciação, são preparadas bibliotecas (figura 11A) para a amplificação dos genes marcador, uma de 16S rRNA para bactérias, e outra de ITS2 para fungos. Para que todas as amostras sejam sequenciadas em conjunto na plataforma, é necessário atribuir-lhes índices que permitam a sua identificação durante e após a sequenciação.

As bibliotecas são carregadas numa célula de fluxo, onde irá ocorrer a sequenciação. Para que isto seja possível, são-lhes também adicionados adaptadores que permitem esta ligação por complementaridade com a sua superfície de oligonucleótidos existentes na célula de fluxo.

Porém, ainda antes de serem introduzidas na plataforma, as bibliotecas indexadas são desnaturadas, normalizadas, agrupadas (*pooling*) e quantificadas. Isto permite que um grande número de bibliotecas seja sequenciado durante um único evento de sequenciação, reduzindo drasticamente o tempo necessário para realizar estudos que envolvam várias amostras (Illumina, 2017), possibilitando simultaneamente, uma triagem e identificação *in silico* a jusante.

2. Estudo do microbioma

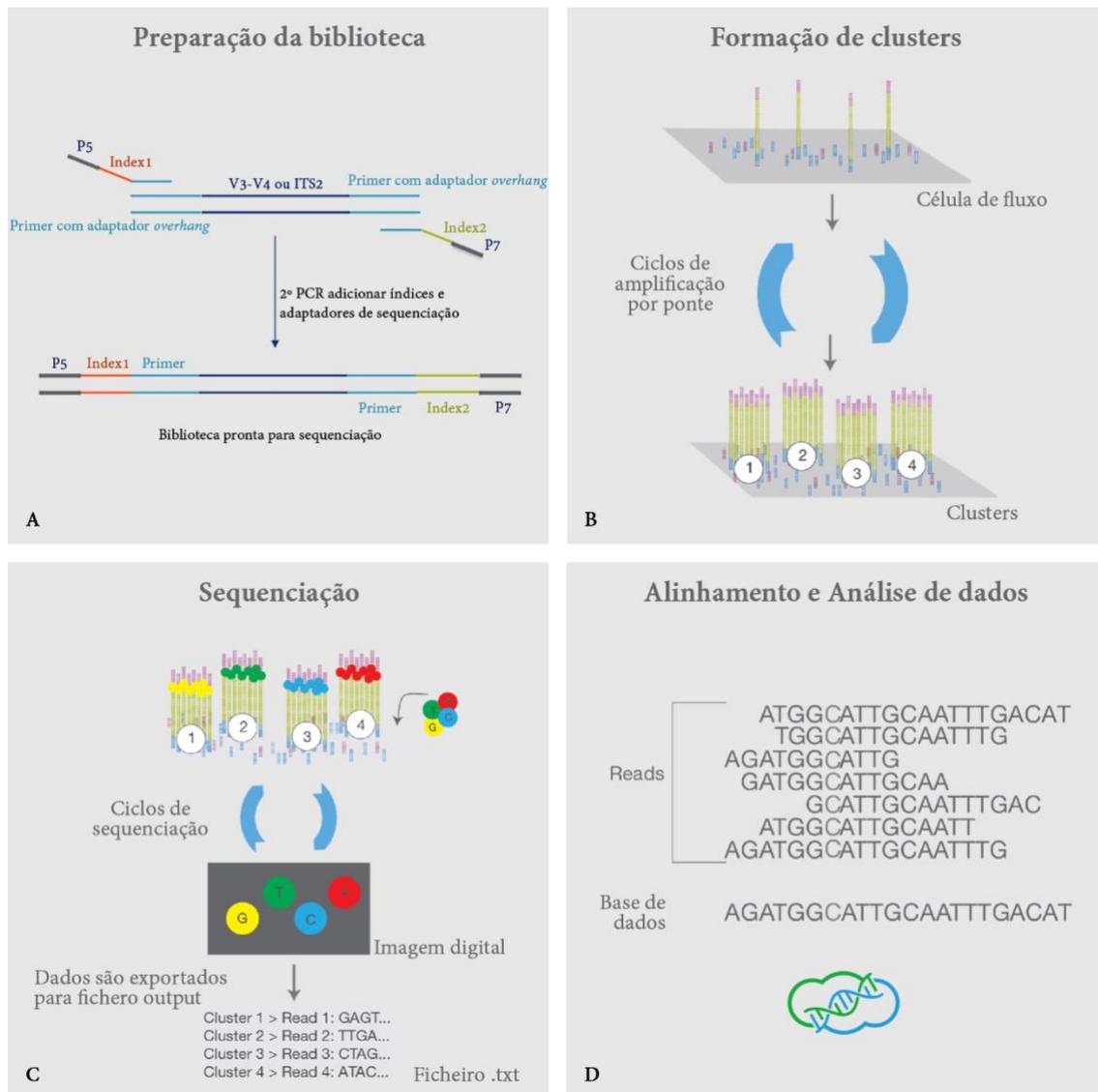


Figura 11. Fluxo de trabalho de sequenciamento de última geração da *Illumina*[®], constituído por quatro etapas (adaptado) (Illumina, 2017). **A.** Preparação da biblioteca de sequenciamento por indexação. São adicionadas, através de uma nova reação de PCR, índices (Index 1 e 2) e adaptadores de sequenciamento (P5 e P7) em ambas as extremidades da região alvo amplificada. Estas novas sequências, vão hibridizar com as sequências *overhang* que foram adicionadas aos primers no primeiro PCR, aquando da amplificação da região específica. Após normalização e *pooling* das amostras, estas encontram-se preparadas para sequenciamento (adaptado) (Wang, 2017). **B.** A biblioteca é carregada numa célula de fluxo, cuja superfície vai capturar os fragmentos de DNA por complementaridade com as regiões adaptadoras introduzidas anteriormente. Cada molécula de DNA é amplificada num *cluster* clonal através de amplificação por ponte. **C.** O sistema *MiSeq Illumina*[®] adota uma abordagem patenteada de sequenciamento por síntese. Este método incorpora dNTPs marcados com fluorescência durante ciclos sequenciais de síntese de DNA. No final de cada ciclo, os nucleótidos incorporados são identificados por excitação do fluoróforo, permitindo que apenas uma base seja lida de cada vez. Este ciclo repete-se "n" vezes para criar uma sequência de "n" bases. **D.** As *reads* são alinhadas com bases de dados através de *software* informático. Depois do alinhamento, as diferenças entre o genoma de referência e as *reads* recém sequenciadas podem ser identificadas.

Projetada para apresentar o DNA de uma maneira que facilite o acesso a enzimas, ao mesmo tempo que assegura alta estabilidade da molécula (Illumina, 2018), na formação de *clusters* (figura 11B) cada molécula do fragmento de DNA é amplificada

isotermicamente numa célula de fluxo. Esta possui uma lâmina de vidro com faixas, cada uma revestida por uma camada composta por dois tipos de oligonucleótidos (P5 e P7), que vão capturar as moléculas de DNA introduzidas. Uma polimerase cria uma cópia dos fragmentos de DNA, a molécula resultante é desnaturada, a cópia dobra e liga-se também à célula de fluxo por hibridização, formando uma ponte de cadeia dupla. Este processo é repetido várias vezes, por ciclos de desnaturação e extensão, semelhantes a uma PCR, e ocorre simultaneamente para milhões de clusters, resultando numa amplificação clonal de todos os fragmentos através de amplificação por ponte.

O sistema *MiSeq Illumina*[®] adota uma abordagem de sequenciação por síntese (figura 11B), um método patenteado que utiliza desoxirribonucleótidos trifosfatos (dNTPs) terminadores reversíveis marcados com fluorescência (Quail et al., 2012). Neste processo, a DNA polimerase catalisa a incorporação de dNTPs de A, T, C e G marcados com fluorescência numa cadeia molde de DNA, durante ciclos sequenciais de síntese. Devido ao bloqueio da extremidade 3' do dNTP incorporado, ocorre a adição de uma só base em cada passo da síntese da nova cadeia. No final de cada ciclo, os nucleótidos incorporados são identificados por excitação de fluoróforo, e o terminador é removido para que comece o ciclo seguinte (Illumina, 2017, 2018; Reuter et al., 2015; Schirmer et al., 2015).

Todo este processo gera milhões de leituras que representam todas as amostras de DNA recolhidas. As bibliotecas agrupadas são separadas com base nos índices únicos introduzidos durante a preparação da amostra. Uma vez que se trata de uma plataforma de sequenciação *paired-end* (figura 10B), as leituras *forward* e *reverse* são emparelhadas, criando sequências únicas (figura 11D).

2.4. Processamento bioinformático dos dados de sequenciação

Os dados da sequenciação dos genes marcador, provenientes de comunidades microbianas complexas, apresentam-se como desafios bioinformáticos, estatísticos e computacionais (Goodrich et al., 2014). Os pacotes de *software* para a análise de amplicões sequenciados vieram dar solução a este problema, sendo o mais utilizado o *QIIME*[™]. Estes é uma coleção de *scripts* de linha de comando, projetada para transformar dados de sequenciação brutos e metadados, em gráficos e estatísticas com qualidade de publicação (Caporaso et al., 2010).

A análise dos genes marcador começa com a desmultiplexação (figura 12) das leituras sequenciais, onde cada sequência é atribuída à sua amostra de origem com base nos índices utilizados na preparação da biblioteca. Neste passo, também se executa uma filtração de qualidade, onde erros de sequenciação e sequências que não correspondam a nenhum índice são excluídos, o mesmo acontece com sequências que não possuem os

2. Estudo do microbioma

limites mínimos de qualidade exigidos, que são específicos para a plataforma usada (Goodrich et al., 2014; Knight et al., 2018).

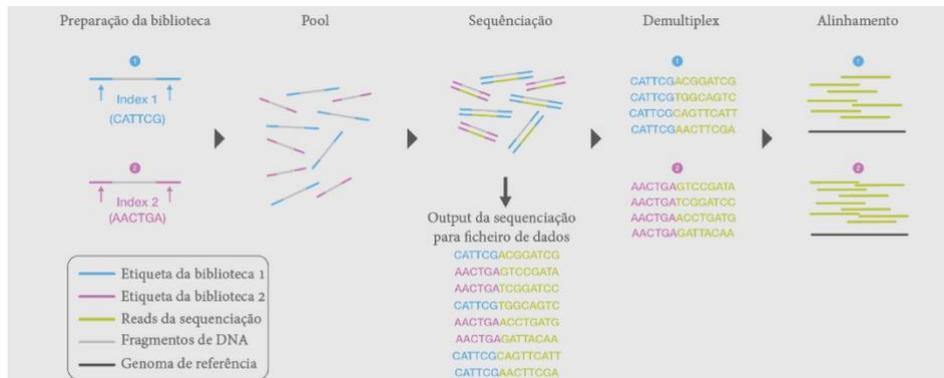


Figura 12. Esquema sumário da sequencição. Durante a sequencição foram adicionas seqüências exclusivas a ambas as extremidades das bibliotecas. As bibliotecas que foram agrupadas na mesma *pool* e carregadas na mesma célula de fluxo, foram sequenciadas numa única operação, resultando num único ficheiro *output*. Um algoritmo de desmultiplexação vai separar os *reads* de acordo com os seus índices. Cada conjunto de *reads* é alinhado contra a base de dados correspondente (adaptado) (Illumina, 2017).

Uma vez retiradas as seqüências ruído, as restantes são agrupadas, em função da identidade das seqüências (%ID), em Unidades Taxonómicas Operacionais (OTUs) que fornecem um nome de trabalho para grupos de microrganismos relacionados (Goodrich et al., 2014). A %ID é utilizada para representar diferentes níveis taxonómicos, sendo, portanto, uma estimativa aproximada do grau de variabilidade da seqüência dentro do gene marcador, do comprimento do amplicão, e dos *taxa* em questão (Goodrich et al., 2014). Estas OTUs são alinhadas e cruzadas com bases de dados (figura 12) para lhes atribuir valor taxonómico, usando para tal a *GreenGenes* para bactérias (McDonald et al., 2012), e a *Unite*, para fungos (Nilsson et al., 2019). Muitas das OTUs não têm identificação taxonómica completa até ao nível de espécie, uma vez que não há confiança total no grupo filogenético em que se encaixa a OTU, ou esta corresponde a uma ramificação da filogenia que não possui informação taxonómica suficiente (Goodrich et al., 2014), o que constitui uma limitação deste tipo de ensaios.

O *output* do processamento bioinformático é geralmente um ficheiro no formato matriz biológica de observação (.BIOM). Este tipo de ficheiros foi desenhado para ser um formato de uso geral para representar amostras biológicas, através de tabelas de observação de contingências. O .BIOM é um padrão reconhecido pelo Projeto Microbioma da Terra, apoiado pelo Consórcio de Padrões Genómicos (GSC). Em investigação de genes marcador, onde este formato encontra o seu uso primário, este tipo de ficheiro representa extensas tabelas OTU, onde são apresentadas as observações, neste caso as OTUs, e a matriz que contém as contagens correspondentes ao número de vezes que cada OTU é observada em cada amostra.

É também nesta fase do estudo que se averigua o sucesso da sequenciação. Em estudos ecológicos, as curvas de rarefação de espécies são usadas para indicar se a amostragem realizada foi suficiente para detetar o número de espécies total da comunidade. Por sua vez, em estudos do microbioma, estas curvas são usadas para verificar se o esforço de sequenciação foi suficiente para cobrir toda a comunidade, ou se, com um aumento do número de sequências obtidas, o número de OTUs encontrado iria aumentar.

2.5. Análise estatística dos dados de sequenciação

Atualmente, toda a pesquisa científica é fundamentada por princípios estatísticos. Em Biologia, o uso de estatística consagra o estudo quantitativo de fenómenos coletivos e observáveis, acoplando-se a métodos e técnicas para colheita, organização, análise e interpretação de dados, para fundamentar conclusões. São realizados ensaios experimentais e/ou observacionais, levando à recolha de dados numéricos, com o propósito de responder a uma questão científica.

A análise de dados de sequenciação provenientes de estudos de microbioma não é exceção, muito pelo contrário, faz uso de um número substancial de ferramentas estatísticas, acopladas aos *softwares* bioinformáticos. Estes *softwares* permitem manipular uma grande quantidade de dados proveniente da sequenciação, simultaneamente dinamizando a aplicação de métodos estatísticos, de modo a obter um estudo coerente e assertivo. Exemplo disso, o *MicrobiomeAnalyst* (MA), uma ferramenta *on-line* abrangente para tratamento estatístico, gráfico e meta-análise de dados de microbiomas, permite desempenhar a análise cabal dos dados brutos do microbioma, para explorar e caracterizar comunidades (Dhariwal, 2017).

O MA possibilita, a um nível potencial totalmente novo, conduzir uma análise abrangente de conjuntos de dados gerados a partir do estudo de microbiomas, focando-se na análise estatística e exploração de dados, enquanto realiza uma análise abrangente da composição e diversidade das comunidades microbianas.

A diversidade de microbiomas é tipicamente descrita em termos de diversidade alfa e beta, isto é, diversidade dentro da amostra e entre amostras, respetivamente. A diversidade alfa quantifica a diversidade de características em amostras individuais e pode ser comparada entre grupos de amostras (Knight et al., 2018), enquanto que a diversidade beta fornecem uma medida do grau em que as amostras diferem umas das outras e pode revelar aspetos da ecologia microbiana que não são aparentes, observando a composição de amostras individuais (Goodrich et al., 2014).

2. Estudo do microbioma

Estas análises permitem perceber que e quantas espécies diferentes estão presentes nas amostras, qual a sua diversidade filogenética, quão semelhantes são as amostras/grupo de amostras, e ainda que fatores ambientais (local, vitalidade, por exemplo) estão associados a diferenças na composição microbiana e biodiversidade.

3. Materiais e métodos

Neste estudo, propomo-nos a analisar os exsudatos negros das lesões na casca externa do sobreiro, através da extração, amplificação e sequenciação de última geração dos amplicões 16S rRNA e ITS2, usando a plataforma *MiSeq Illumina*[®]. Os dados das comunidades bacterianas e fúngicas daí resultantes, serão analisadas bioinformaticamente através do *pipeline QIIME*[™] (Caporaso et al., 2010), e gráfica e estatisticamente recorrendo à ferramenta online *MicrobiomeAnalyst* (Dhariwal et al., 2017).

Para a recolha de material genético proveniente de exsudatos de sobreiros, dada a ausência de protocolos específicos para este tipo de amostras, foi necessário testar e otimizar procedimentos experimentais existentes para outros fins, de modo a adaptá-los aos objetivos propostos para este estudo. Foi com esta finalidade que, inicialmente, se seguiu o desenho experimental descrito neste capítulo para um reduzido número de amostras, recolhidas em 2017 no Baixo Sorraia, e apenas posteriormente para as restantes.

A análise de dados provenientes da extração e sequenciação das primeiras amostras, permitiu uma caracterização detalhada do microbioma bacteriano e fúngico de lesões na casca de sobreiros em declínio, ainda que limitada a um reduzido número de indivíduos. Esses resultados preliminares, permitiram validar todos as estratégias definidas ao longo de todo o desenho experimental, desde a recolha de amostras, à validação da extração de DNA, desde a amplificação e sequenciação dos genes marcadores escolhidos, através de *primers* específicos para cada uma das comunidades, até à ferramenta de análise dos dados de sequenciação.

Os procedimentos aqui apresentados, foram adotados de forma consistente, mantendo as condições para ambos os grupos de amostras, para que não resultassem em diferentes perfis de diversidade e para que se evitassem flutuações que se pudessem refletir na modelagem estatística.

3.1. Recolha de amostras

A recolha de exsudatos decorreu nos sobreiros amostrados para o projeto “*Cork oak microbiome: Deciphering the rhizosphere microbial community of Quercus suber*”, em curso na *Genoinseq*. Em todos os sobreiros amostrados para a rizosfera e que apresentaram lesões na cortiça e exsudatos foi recolhida uma amostra através do

esfregação de uma zaragatoa nas feridas da casca externa das árvores com escorrência húmida.



Figura 13. Classes de deterioração foliar refletindo a percentagem de desfolhamento da árvore (Michel e Seidling, 2016).

Os sobreiros amostrados foram georreferenciados, e foi feita uma avaliação visual da condição das suas copas, agrupando-as em classes com diferente deterioração foliar (Michel e Seidling, 2016) (figura 13). A classe 0 corresponde à ausência de dados na foliação e só na classe 1 começam a ser visíveis alguns danos. A classe 2 já representa uma preocupação moderada em relação à condição da árvore, na classe 3 a árvore apresenta danos severos e, por fim, a última classe corresponde a árvores que apresentam mais de 90% de desfolhamento, com danos muito severos ou morte da mesma.

Tabela 1. Caracterização das amostras recolhidas pelos dos esfregaços dos exsudados dos sobreiros.

ID	ZIF	Região	Condição da Copa	Número da árvore	Ano de amostragem
Esfregaço 1	Baixo Sorraia	Vitalidade	3	103	2017
Esfregaço 2	Baixo Sorraia	Vitalidade	2	102	2017
Esfregaço 3	Baixo Sorraia	Mortalidade	3	23	2017
Esfregaço 4	Baixo Sorraia	Vitalidade	3	63	2017
Esfregaço 5	Divor	Mortalidade	0	130	2018
Esfregaço 6	Divor	Mortalidade	0	140	2018
Esfregaço 7	Divor	Vitalidade	0	180	2018
Esfregaço 8	Divor	Vitalidade	0	190	2018
Esfregaço 9	Divor	Vitalidade	1	161	2018
Esfregaço 10	Divor	Vitalidade	1	201	2018
Esfregaço 11	Divor	Mortalidade	2	112	2018
Esfregaço 12	Divor	Mortalidade	2	122	2018
Esfregaço 13	Divor	Mortalidade	2	132	2018
Esfregaço 14	Divor	Mortalidade	2	142	2018
Esfregaço 15	Divor	Vitalidade	2	162	2018
Esfregaço 16	Divor	Vitalidade	2	183	2018
Esfregaço 17	Divor	Vitalidade	2	202	2018
Esfregaço 18	Divor	Mortalidade	3	113	2018
Esfregaço 19	Divor	Mortalidade	3	123	2018
Esfregaço 20	Divor	Mortalidade	3	143	2018
Esfregaço 21	Divor	Mortalidade	3	153	2018

Além disso, foi necessário ter atenção a condições com algum grau de influência na composição da microbiota, tal como as condições ambientais, que podem diferir e ter

efeito nas diferentes colónias que as plantas abrigam (Goodrich et al., 2014). Assim, para além da óbvia recolha de amostras em árvores com diferentes estados declinantes, foram, não só recolhidas amostras em duas ZIF diferentes (Baixo Sorraia e Divor), mas em cada uma destas áreas, foram delineadas duas regiões diferentes em relação à taxa de mortalidade (tabela 1). A taxa de mortalidade foi determinada com base na avaliação temporal (2010-2015) do número de árvores em declínio. Para a região de vitalidade foi selecionada um montado com 0-1% de árvores em declínio, enquanto que, para a de mortalidade foi escolhido um montado com 9-22% de declínio (APFC 2016a, APFC2016b).

As amostras, uma vez que são perecíveis, devem ser acondicionadas de forma adequada até o momento de serem submetidas à extração do DNA (Oliveira et al., 2007), pelo que as zaragatoas recolhidas em 2017, e posteriormente em 2018, foram armazenadas a -20°C até à extração do material genético para sequenciação.

3.2. Protocolos de DNA

3.2.1. Extração

As amostras tiveram o seu DNA total extraído utilizando o *kit* de DNA *NucleoSpin® Soil* (Macherey-Nagel, Düren, Alemanha). Este *kit* é indicado para o isolamento de DNA genómico de microrganismos, sejam eles bactérias, Gram-positivas e Gram-negativas, Archaea, fungos ou algas, em amostras de solo, lodo e sedimento, todavia provou-se igualmente eficaz na aplicação ao material recolhido neste estudo.

A extração do DNA das zaragatoas com o *kit* foi realizada segundo as instruções e recomendações do fabricante, à exceção do passo de preparação da amostra, onde os esfregaços foram embebidos em 1000 µL de tampão *SL1* e 150 µL de potenciador *SX*. Foi ainda efetuado um controlo negativo para a extração.

3.2.2. Quantificação

O DNA extraído de cada amostra foi quantificado por espectrofotometria no *NanoDrop™* (ThermoFisher Scientific, Waltham, Estados Unidos da América) (tabela 2). No entanto, algumas amostras não apresentavam curvas de absorvância aceitáveis para DNA “puro”, isto é, a relação entre a absorvância a 260 nm e a 280 nm não se aproximava dos valores de ~1,6 indicados pelo fabricante.

Assim, foi necessário recorrer ao *kit* de ensaio *Qubit® dsDNA HS* (ThermoFisher Scientific, Waltham, Estados Unidos da América), e a um fluorímetro *Qubit®*, para proceder a uma quantificação do DNA mais precisa (tabela 2). Seguindo as instruções e recomendações do fabricante, foi possível medir com exatidão a concentração do DNA isolada de cada amostra, sendo possível garantir que o DNA com que trabalhamos é de boa qualidade.

3.2.3. Amplificação

O protocolo utilizado seguiu um método otimizado para preparar amostras para sequenciação da região V3-V4 do gene 16S rRNA de bactérias, mas que através do uso de diferentes *primers* específicos se adapta perfeitamente para sequenciar outras regiões, tal como a região hipervariável de ITS2 de fungos. Deste modo, para cada amostra é necessária uma preparação para obter o microbioma bacteriano e outra preparação para obter o microbioma fúngico.

Para amplificar o gene de interesse para as bactérias, foram usados o *primer forward* Bakt_341F 5'-CCTACGGGNGGCWGCAG-3' e o *primer reverse* Bakt_805R 5'-GACTACHVGGGTATCTAATCC-3' (Herlemann et al., 2011; Klindworth et al., 2013). Para fungos, para a amplificação da região hipervariável de ITS2, utilizaram-se um conjunto de *primers forward* ITS3NGS1_F 5'-CATCGATGAAGAACGCAG-3', ITS3NGS2_F 5'-CAACGATGAAGAACGCAG-3', ITS3NGS3_F 5'-CACCGATGAAGAACGCAG-3', ITS3NGS4_F 5'-CATCGATGAAGAACGTAG-3', ITS3NGS5_F 5'-CATCGATGAAGAACGTGG-3', e ITS3NGS10_F 5'-CATCGATGAAGAACGCTG-3', e o *primer reverse* ITS4NGS001_R 5'-TCCTSCGCTTATTGATATGC-3'. (Tedersoo et al., 2014). Os adaptadores *overhang Illumina*[®], foram adicionados nas extremidades dos *primers* acima referidos: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[primer]-3', para sequencias *forward* e 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[primer]-3', para sequencias *reverse*.

Deste modo, quer para as bactérias, quer para fungos, as reações de PCR foram realizadas, usando os respetivos *primers*, *KAPA HiFi HotStart PCR* (KAPABiosystems, Cidade do Cabo, África do Sul) de acordo as instruções e recomendações do fabricante, ajustando, quando necessário, a concentração dos reagentes do PCR, para obter amplificação das amostras. As condições utilizadas no termociclador para as duas reações estão disponíveis na tabela 2.

Tabela 2. Condições das diferentes etapas de PCR para amplificação das regiões V3-V4 e ITS2.

	V3-V4			ITS2		
	Temperatura	Duração	N.º de ciclos	Temperatura	Duração	N.º de ciclos
Desnaturação inicial	95°C	3min	1	95°C	3min	1
Desnaturação	98°C	20sec	30	98°C	20sec	25
<i>Annealing</i>	55°C	30sec				
Elongação	72°C	30sec				
Elongação final	72°C	5min	1	72°C	5min	1
<i>Hold</i>	10°C	∞	-	10°C	∞	-

No final dos PCR, o sucesso da amplificação foi confirmado por eletroforeses em gel de agarose a 1% para visualizar as bandas da região de interesse. Na amplificação foram incluídos controlos positivo e negativo, e sempre que uma das amostras não

apresentou banda, as concentrações dos reagentes do PCR foram ajustadas, a amplificação repetida, e o produto do PCR corrido novamente em gel.

3.2.4. Purificação

As amostras foram purificadas com *KAPA Pure Beads* (KAPABiosystems, Cidade do Cabo, África do Sul), segundo as instruções e recomendações do fabricante. A suspensão destas esferas paramagnéticas no tampão otimizado, incluídos no *kit*, permitem a purificação e preparação do DNA para sequenciação. Uma vez mais, realizou-se eletroforese das amostras para conferir o sucesso do procedimento.

3.3. Sequenciação

3.3.1. Indexação

A indexação foi feita por PCR utilizando índices *Nextera[®] XT* (*Illumina[®]*, San Diego, Califórnia, Estados Unidos da América) (tabela 3).

Tabela 3. Condições das diferentes etapas de PCR para amplificação dos índices e adaptadores de sequenciação.

	Temperatura	Duração	N.º de ciclos
Desnaturação inicial	95°C	3min	1
Desnaturação	95°C	30sec	
<i>Annealing</i>	50°C	30sec	8
Elongação	72°C	30sec	
Elongação final	72°C	5min	1
<i>Hold</i>	4°C	∞	-

Uma vez mais, os produtos do PCR foram purificados, desta vez utilizando o *Agencourt AMPure XP* (*Beckman Coulter[™]*, Indianapolis, Estados Unidos da América). Este sistema de purificação utiliza, também, uma tecnologia de esferas paramagnéticas para purificação de alto rendimento.

3.3.2. Normalização, pooling e quantificação da biblioteca

Para que todas as amostras fossem sequenciadas em conjunto na plataforma, para além da atribuição dos índices no passo de indexação, que tornam possível a sua identificação, foi necessário fazer com que todas as amostras tivessem a mesma molaridade, leia-se, assegurar uma distribuição de *reads* uniforme para todas as amostras durante a sequenciação.

Para isso, as amostras foram normalizadas usando o *kit SequalPrep[™]* (*ThermoFisher Scientific*, Waltham, EUA), de acordo com as instruções e recomendações do fabricante, para que ficassem com a mesma concentração. Naturalmente, este passo cria diferenças volumétricas entre as amostras, tornando necessário o seu agrupamento (*pooling*) em função do volume das soluções. As *pools* foram organizadas tendo em conta

a região amplificada e a qualidade da amplificação de cada amostra. Assim, foi garantido que todas as amostras tiveram igual hipótese durante a sequenciação e não foram sub ou sobrevalorizadas.

3.3.3. Sequenciação

Com tudo preparado para proceder à sequenciado no sequenciador *MiSeq Illumina*[®] (Illumina, San Diego, Califórnia, Estados Unidos da América), a biblioteca combinada das amostras e do controlo, foi carregada para o *MiSeq reagent cartridge* incluído no *MiSeq Reagent Kit v3 (Illumina*[®], San Diego, Califórnia, Estados Unidos da América), de acordo com as instruções e recomendações do fabricante. Na sequenciação é adicionada uma biblioteca pré-preparada *PhiX (Illumina*[®]) para controlo da sequenciação.

3.4. Processamento bioinformático dos dados de sequenciação

Os resultados obtidos após sequenciação de última geração dos amplicões 16S e ITS2 das amostras, foram importados para o *QIIME*[™] (Caporaso et al., 2010), onde sofreram desmultiplexação e filtração de qualidade. No *QIIME*[™] foram também geradas as curvas de rarefação para análise do esforço de sequenciação da plataforma *MiSeq Illumina*[®].

Finalmente, recorrendo às bases de dados *GreenGenes*, para bactérias, e *Unite*, para fungos, foi também realizada a atribuição taxonómica a cada OTU. Assim, foi gerado um ficheiro no formato .BIOM para bactérias, e outro para fungos, onde consta o número de *reads* sequenciados para cada OTU, em cada uma das amostras, com a respetiva classificação taxonómica que lhes foi atribuída.

3.5. Análise estatística dos dados de sequenciação

O resultado do processamento bioinformático dos dados de sequenciação foi transformado numa tabela de *Microsoft*[®] *Office Excel*[®] (*Microsoft*[®], Redmond, WA, Estados Unidos) para que as tabelas OTU pudessem ser formatadas e trabalhadas, antes de serem carregadas no *MicrobiomeAnalyst* (Dhariwal et al., 2017) para análise.

A sequenciação genética confirmou uma suspeita que tinha surgido com a eletroforese realizada no final da extração e purificação de DNA. Houve presença de microrganismos na amostra controlo, isto é, o controlo negativo do *kit* de extração de DNA sofreu uma contaminação. Foi necessário saber se esta foi introduzida pelo *kit*, cuja ocorrência significaria uma conseqüente contaminação de todas as amostras, ou, se foi o manuseamento das amostras em laboratório que, por engano, fez com que este evento de contaminação tivesse acontecido.

Antes de avançar para a análise do microbioma foi então necessário analisar a amostra controlo, para perceber o que pode ter ocorrido, mas sobretudo para remover, de toda a análise, as OTUs dominantes na amostra controlo. Isto, permitiu salvaguardar a validade da análise, removendo OTUs presentes nas várias amostras, que possam ter sido introduzidas por uma qualquer contaminação.

3.5.1. Filtração dos dados

Após a remoção de todas as OTUs possivelmente introduzidas pela contaminação, o primeiro passo da preparação das tabelas OTU para *upload* no MA passou por eliminar todas e quaisquer OTUs que não correspondessem a bactérias ou fungos, na respetiva tabela OTU. Na tabela bacteriana, foram removidas as OTUs correspondentes a DNA de Archaea, cloroplastidial e mitocondrial, que também foram amplificadas e sequenciadas, mas não correspondem a bactérias. OTUs que foram classificadas taxonomicamente como “*Unassigned Kingdom*”, isto é, não tiveram atribuição taxonómica a nível do reino, quando cruzadas com as bases de dados, foram também eliminadas, quer para bactérias, quer para fungos, uma vez que não podemos garantir que se trate de informação genética proveniente das regiões amplificadas.

Para remover elementos de baixa qualidade e/ou não-informativos, de modo a melhorar a análise estatística a jusante deste passo, foram eliminadas todas as OTUs de baixa prevalência e de baixa abundância. Ainda que o MA permita fazer esta filtração de dados, esta foi feita diretamente nas tabelas OTU, para ter maior controlo sobre as OTUs removidas. Assim, todas as OTUs presentes em apenas uma das amostras, foram eliminadas, sendo que o mesmo foi levado a cabo para todas as aquelas cujo número total de *reads* era inferior a 20.

3.5.2. Análise dos dados

Sempre que se coloca a necessidade de acompanhar uma análise pelo respetivo teste estatístico, o MA dá-nos essa opção. Seja da composição e estrutura da comunidade, seja da alfa ou beta diversidade ou uma análise comparativa, o MA tem disponível uma seleção de métodos paramétricos e não-paramétricos para testar a significância dessa análise. Torna-se necessário explorar o conjunto de dados que temos à disposição para saber que tipo de métodos estatísticos têm de ser aplicados.

Os métodos paramétricos baseiam-se na suposição de que os dados observados na amostra são provenientes de uma população com distribuição de forma teórica conhecida (Paes, 2009). Esta suposição é assumida para a maioria dos métodos estatísticos mais utilizados, os testes paramétricos, como o teste *t de Student*, análise de variância (ANOVA), regressão linear e intervalos de confiança (Paes, 2009).

No entanto, os dados provenientes do estudo de microbiomas, devido à sua natureza, praticamente nunca seguem uma distribuição normal (Thorsen et al., 2016; Weiss et al., 2017). Para assegurar esta premissa, foram analisadas as distribuições dos histogramas dos dados e o declive dos seus gráficos quartil-quartil, e utilizados os testes de normalidade de *Kolmogorov-Smirnov* (Massey Jr., 1951), *Shapiro-Wilk* (Srivastava e Hui, 1987) e *D'Agostino & Pearson* (D'agostino et al., 2012), realizados no *software GraphPad Prism 8* (GraphPad Software, Califórnia, Estados Unidos da América). A utilização de testes não-paramétricos, como *Mann-Whitney* (Nachar, 2008) e *Kruskal-Wallis* (Vargha e Delaney, 1998), tem sido amplamente adotada em ecologia, pois estes não fazem suposições sobre a forma matemática da distribuição da abundância (Paes, 2009; Weiss et al., 2017).

As características únicas dos dados metagenómicos tornaram também inadequado que lhes sejam aplicados diretamente métodos desenvolvidos noutros campos da biologia, para realizar análises comparativas e diferenciais (Thorsen et al., 2016; Weiss et al., 2017). Portanto, antes de qualquer comparação significativa, foi necessário normalizar as diferenças na profundidade de sequenciação e na dispersão, presentes nestes dados, através de *data scaling* e *data transformation*.

Os dados, provenientes da abundância de ambas as comunidades microbianas, foram então normalizados devido às diferenças no número de *reads* entre as amostras, à heterogeneidade das suas distribuições e ao elevado número de zeros, para que fosse possível interpretá-los biológica e estatisticamente. Vários métodos de normalização, que têm sido usualmente aplicados a diferentes conjuntos de dados, estão disponíveis no MA (tabela 4). Seguindo recomendações presentes na literatura (Dhariwal, 2017; Dhariwal et al., 2017; Thorsen et al., 2016; Weiss et al., 2017), ao longo deste estudo foi adotado o *data scaling* por *Cumulative Sum Scaling* (CSS) (Qiu e Hawkins, 2003), à exceção da análise diferencial das amostras com o DESeq2 (Love et al., 2014), para a qual foi adotada a transformação de dados por *Trimmed mean of M-values* (TMM) (Fraiman e Muniz, 2001).

Tabela 4. Opções disponíveis no MicrobiomeAnalyst para normalização dos dados.

<i>Data rarefying</i>	<i>Do not rarefy my data</i> <i>Rarefy to the minimum library size</i>
<i>Data scaling</i>	<i>Do not scale my data</i> <i>Total sum scaling (TSS)</i> <i>Cumulative sum scaling (CSS)</i> <i>Upper-quantile normalization (UQ)</i>
<i>Data transformation</i>	<i>Do not transform my data</i> <i>Relative log expression (RLE)</i> <i>Trimmed mean of M-values (TMM)</i> <i>Centered log ratio (CLR)</i>

Outra função disponível no MA é a rarefação dos dados, no entanto, esta não se justifica, uma vez que não existem diferenças na profundidade de sequenciação entre as

amostras na ordem das dezenas, além de que esta prática é extremamente desaconselhada (McMurdie e Holmes, 2014).

3.5.3. Análise do microbioma no MicrobiomeAnalyst

Para fazer o *upload* para o MA, as tabelas OTU filtradas foram editadas para que preenchessem os critérios exigidos. A estas, foi anexada uma tabela com os metadados recolhidos para esta análise (tabela 5), que incluem as duas variáveis ambientais que estes dados permitem analisar, a região de onde as amostras foram recolhidas (vitalidade/mortalidade) e a condição foliar dos sobreiros amostrados (0-1, 2, 3).

Tabela 5. Tabela de metadados criada para análise do microbioma das lesões na casca do sobreiro.

ID	Região	Condição da Copa
Esfregaço 1	Vitalidade	3
Esfregaço 2	Vitalidade	2
Esfregaço 3	Mortalidade	3
Esfregaço 4	Vitalidade	3
Esfregaço 5	Mortalidade	0-1
Esfregaço 6	Mortalidade	0-1
Esfregaço 7	Vitalidade	0-1
Esfregaço 8	Vitalidade	0-1
Esfregaço 9	Vitalidade	0-1
Esfregaço 10	Vitalidade	0-1
Esfregaço 11	Mortalidade	2
Esfregaço 12	Mortalidade	2
Esfregaço 13	Mortalidade	2
Esfregaço 14	Mortalidade	2
Esfregaço 15	Vitalidade	2
Esfregaço 16	Vitalidade	2
Esfregaço 17	Vitalidade	2
Esfregaço 18	Mortalidade	3
Esfregaço 19	Mortalidade	3
Esfregaço 20	Mortalidade	3
Esfregaço 21	Mortalidade	3

Para a condição da copa, uma vez que as classe 0 e 1 correspondem a estados de deterioração foliar muito próximos e como unindo estas duas classes numa só, é possível obter um número de amostras comparável com as outras duas classes (6 amostras da união das classes 0 e 1 vs. 7 amostras da classe 2 vs. 8 amostras da classe 3), foi criada a classe 0-1 que engloba as condições de copa de classe 0 e de classe 1.

A maioria das ferramentas que o MA providencia para análise de genes marcador é útil para responder ao objetivo deste trabalho e foi empregue segundo as recomendações dos desenvolvedores e informação presente na literatura. Outras, devido à natureza dos dados de sequenciação e ao objetivo deste projeto, são redundantes ou não podem ser aplicadas neste estudo.

4. Resultados e Discussão

4.1. Extração

A extração do DNA correu de acordo com o esperado. Para os esfregaços 7 e 10 foi necessário eluir o DNA em menor volume para se conseguir extrair DNA mensurável. O esfregaço 18 apresentava uma cor acastanhada, um forte indício da presença de substâncias húmicas que podem atuar como inibidores de PCR. Estas substâncias são produzidas por bactérias, fungos e protozoários durante a degradação de matéria orgânica como o material vegetal e, devido às suas propriedades químicas, a maioria dos métodos de purificação não as conseguem distinguir das moléculas de DNA atuando, por isso, como inibidores de PCR extremamente potentes.

4.2. Quantificação

Na tabela 6 são apresentados os resultados de quantificação de DNA das amostras. O esfregaço 17, apresentou valores abaixo do limite de deteção do fluorímetro, pelo que só através de eletroforese se pôde confirmar a presença de DNA.

Tabela 6. Quantificação do DNA extraído por *NanoDrop*TM e por *Qubit*[®] *dsDNA HS*.

ID	<i>NanoDrop</i> TM			<i>Qubit</i> [®] <i>dsDNA HS</i> DNA (ng/μL)
	DNA (ng/μL)	A260/280	A260/280	
Esfregaço 5	25,2	1,7	0,58	20,4
Esfregaço 6	17,2	1,95	1,06	14,3
Esfregaço 7	27,0	1,4	0,64	8,36
Esfregaço 8	38,8	1,48	0,54	0,893
Esfregaço 9	107,5	1,52	0,77	35,1
Esfregaço 10	14,1	1,63	0,63	0,849
Esfregaço 11	5,0	1,47	0,43	0,307
Esfregaço 12	14,8	1,76	1,13	11,3
Esfregaço 13	24,8	1,7	1,07	14,6
Esfregaço 14	153,5	1,59	0,87	32,5
Esfregaço 15	51,3	1,22	0,56	11,3
Esfregaço 16	8,9	1,51	0,35	0,241
Esfregaço 17	4,2	1,53	0,36	Abaixo do limite de deteção
Esfregaço 18	264,3	1,15	0,55	4,59
Esfregaço 19	38,7	1,6	1,26	31,2
Esfregaço 20	16,3	1,3	0,52	5,47
Esfregaço 21	5,4	2,08	0,49	1,31
Controlo do <i>kit</i>	4,2	1,27	0,18	Abaixo do limite de deteção

4. Resultados e Discussão

Os esfregaços 7 e 10, que tinham sofrido perda de material, e o esfregaço 18, que tinha uma coloração acastanhada, não apresentaram quaisquer anomalias na quantificação de DNA. Mas, estranhamente, o controlo de extração do *kit* de DNA *NucleoSpin® Soil*, doravante referido apenas como controlo do *kit*, acusou presença de DNA (tal como no esfregaço 17, em valores abaixo do limite de deteção), sugerindo uma contaminação do próprio *kit* ou uma contaminação decorrente do procedimento laboratorial e/ou manipulação de amostras.

4.3. Amplificação

A eletroforese da amplificação das amostras após extração e quantificação (figura 14), mostrou que praticamente todas tiveram amplificação de DNA na região V3-V4 do 16rRNA, com alguns casos particulares. O esfregaço 17 apresentou uma banda muito tênue e o esfregaço 18 não apresentou qualquer banda. De lembrar que, o primeiro tinha apresentado uma concentração demasiado baixa durante a quantificação, e a extração de DNA do segundo resultou numa cor acastanhada, na qual se suspeitava a presença de inibidores de PCR. Aqui, foi também confirmada a presença de DNA no controlo do *kit*, pois este apresentou uma banda, ainda que subtil.



Figura 14. Eletroforese dos produtos de PCR resultantes da amplificação da região V3-V3 do 16s rRNA. O esfregaço 17 (Swab17) e o controlo do *kit* apresentam bandas muito tênues. O esfregaço 18 (Swab18) não possui qualquer banda.

Marcador: NzyTech Ladder I (200 a 1800 bp)

Para estas amostras em questão (a vermelho na figura 14), foi realizado nova amplificação e respetiva eletroforese (figura 15). Para o esfregaço 17, foi utilizado um maior volume de DNA na preparação do PCR, uma vez que estava a ocorrer amplificação, embora em pequena quantidade. Já para o esfregaço 18, foram utilizadas diferentes diluições (1:10, 1:100) do DNA extraído, juntamente com uma amostra de DNA controlo (C++), para testar a presença dos inibidores já mencionados. Isto porque poderia não estar a ser visível banda por duas situações, pela existência de inibidores ou pela inexistência de DNA amplificado. Assim, a eletroforese poderia revelar três cenários: o C++ e o esfregaço 18 não apresentarem banda, indicando presença de inibidores de PCR; o C++ apresentar banda e o esfregaço 18 não, revelando que o esfregaço 18 não possui DNA;

ambos apresentarem banda, o que significaria que o esfregaço 18 possui inibidores, mas que uma redução da sua concentração permite que o DNA seja amplificado.

Também o controlo do *kit* foi novamente amplificado e visualizado no gel para confirmar a presença de banda.



Figura 15. Eletroforese dos produtos de PCR resultantes da nova amplificação da região V3-V3 do 16s rRNA para o esfregaço 17 com um maior volume de DNA usado no PCR (Swab17+DNA), o esfregaço 18 ao qual se adicionou DNA controlo (Swab18 (1:1) C++), o esfregaço 18 em diluição 1:10 (Swab18 (1:10)), o esfregaço 18 em diluição 1:10 ao qual se adicionou DNA controlo (Swab18 (1:10) C++), o esfregaço 18 em diluição 1:100 (Swab18 (1:100)), o esfregaço 18 em diluição 1:100 ao qual se adicionou DNA controlo (Swab18 (1:100) C++), e para o controlo do *kit*. O esfregaço 17 com mais DNA e o controlo apresentam bandas ténues, o esfregaço 18 diluído 1:100 apresenta banda. Marcador: NzyTech Ladder I (200 a 1800 bp)

Utilizando uma maior concentração de DNA no PCR do esfregaço 17 e uma diluição 1:100 no esfregaço 18, já foi possível observar banda na eletroforese. Estes dados indicam que a amplificação anterior não estava a ser possível pela pouca quantidade de material genético do esfregaço 17 e devido há existência de inibidores no esfregaço 18. Diminuir a concentração da amostra do esfregaço 18, e consequentemente dos inibidores, já foi possível realizar o PCR e visualizar a banda na eletroforese. O controlo do *kit* continua a apresentar banda, o que confirma a contaminação.

Em relação a fungos, foi igualmente realizado o PCR para todas as amostras, mas, para os esfregaços 17 e 18, que tinham encontrado dificuldades de amplificação de DNA bacteriano, foram utilizadas as soluções que as permitiram amplificar, isto é, o PCR do esfregaço 17 foi feito com mais DNA, e o esfregaço 18 foi diluído 1:100.

Na eletroforese (figura 16), é possível verificar que todas as amostras foram amplificadas com sucesso, à exceção do esfregaço 17 que não apresentou banda, mesmo nestas condições. O controlo do *kit*, ao contrário do sucedido para a região bacteriana V3V4, também não apresentou banda na região ITS2.



Figura 16. Eletroforese dos produtos de PCR resultantes da amplificação da região ITS2. O esfregaço 17 (Swab17) e o controlo do *kit* não apresentaram banda. Marcador: NzyTech Ladder I (200 a 1800 bp)

4. Resultados e Discussão

Para tentar que o esfregaço 17 fosse amplificado, foi realizada uma pPCR desta amostra, onde se utilizaram três quantidades diferentes de DNA. No entanto, mesmo nestas condições, não foi possível visualizar qualquer banda na eletroforese.

4.4. Sequenciação

Todas as amostras foram então sequenciadas a nível das comunidades bacterianas e fúngicas, incluindo o controlo do *kit* e a amostra 17, mesmo não tendo apresentado DNA amplificado na região ITS2.

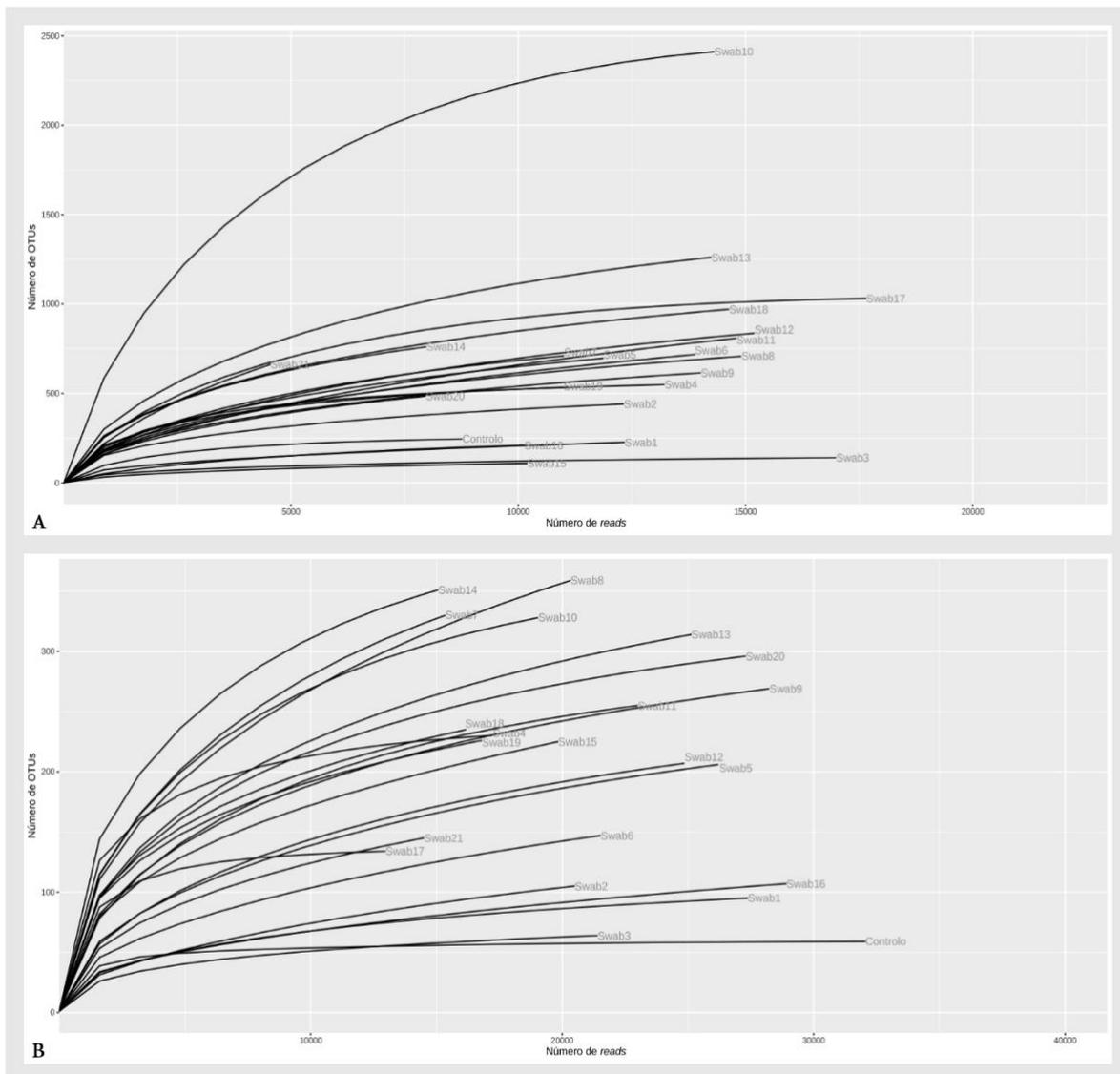


Figura 17. Curvas de rarefação, indicando a relação entre o número de OTUs e o número de *reads* obtidas na sequenciação de bactérias (A) e de fungos (B).

Nas curvas de rarefação (figura 17), obtidas pela relação entre o tamanho da sequenciação (número de *reads*) e a riqueza das amostras (número de OTUs), verifica-se que com o aumento do número de seqüências, o número de OTUs encontradas aumenta, até que a curva atinge a saturação, o *plateau*. Quando se atinge este ponto com declive de

assimptota zero, significa que toda a diversidade microbiana presente na amostra foi caracterizada (Chao e Chiu, 2016; Gotelli e Chao, 2013; Gotelli e Colwell, 2011), isto é, com o aumento do número de sequências, o número de OTUs já não sofre grandes alterações.

Deste modo, verifica-se que em todas as amostras este ponto foi alcançado, pelo que se pode afirmar que a sequenciação foi bem-sucedida e suficiente para analisar estes microbiomas. Estes gráficos mostram também a grande heterogeneidade da diversidade existente entre as amostras quanto ao número de *reads* e OTUs. Ainda que estes gráficos correspondam aos dados pré-filtrados, ditam a necessidade da normalização para prosseguir com a análise.

A cobertura de sequenciação (*Good's coverage*) de cada amostra, outro indicador da qualidade da sequenciação, está disponível na tabela 7 e é definida como:

$$goods = 1 - \left(\frac{no.singleton}{\sum no.seqs} \right) \quad (1)$$

Onde o termo *singleton* é usado para OTUs que estão apenas representadas por uma única *read* numa única amostra.

Tabela 7. Dados relativos à sequenciação de bactérias e fungos em termos de número de *singletons*, número de *reads* e *Good's coverage*.

ID	Bactérias			Fungos		
	<i>singletons</i>	<i>reads</i>	<i>Good's coverage</i>	<i>singletons</i>	<i>reads</i>	<i>Good's coverage</i>
Esfregação1	103	12345	99,17%	28	27390	99,90%
Esfregação2	138	12328	98,88%	47	20502	99,77%
Esfregação3	33	16998	99,81%	18	21414	99,92%
Esfregação4	84	13220	99,36%	21	17201	99,88%
Esfregação5	272	11878	97,71%	72	26214	99,73%
Esfregação6	327	13895	97,65%	63	21531	99,71%
Esfregação7	262	11010	97,62%	114	15361	99,26%
Esfregação8	271	14910	98,18%	135	20343	99,34%
Esfregação9	221	14021	98,42%	77	28242	99,73%
Esfregação10	312	14318	97,82%	78	19051	99,59%
Esfregação11	401	14800	97,29%	62	23010	99,73%
Esfregação12	336	15196	97,79%	70	24860	99,72%
Esfregação13	376	14252	97,36%	95	25142	99,62%
Esfregação14	271	7980	96,60%	89	15076	99,41%
Esfregação15	41	10203	99,60%	89	19841	99,55%
Esfregação16	76	10134	99,25%	45	28927	99,84%
Esfregação17	75	17662	99,58%	8	12976	99,94%
Esfregação18	320	14645	97,81%	73	16182	99,55%
Esfregação19	101	11001	99,08%	68	16805	99,60%
Esfregação20	202	7965	97,46%	70	27302	99,74%
Esfregação21	373	4550	91,80%	58	14510	99,60%
Controlo do kit	44	8769	99,50%	2	32058	99,99%

4. Resultados e Discussão

Todos os valores indicados são superiores a 95%, à exceção do esfregaço 21 para bactérias, mas mesmo neste, a percentagem de cobertura da sequenciação é superior a 90%, pelo que é perfeitamente aceitável. Na prática, estes valores significam que, se uma amostra tem 95% de *Good's coverage*, então 5% dos *reads* dessa amostra são de OTUs que aparecem apenas uma vez nessa amostra. Estes valores apontam para a mesma conclusão, a sequenciação foi bem-sucedida e suficiente para analisar os microbiomas bacteriano e fúngico.

4.5. Controlo do *kit*

Durante o controlo realizado à extração de DNA, por quantificação e mais tarde por eletroforese, foi detetada a presença de material genético, no que seria o controlo negativo do *kit* de extração de DNA. Estes dados remetem-nos para um erro de manipulação das amostras durante a extração de DNA.

Mesmo antes de introduzir os dados da sequenciação no MA para analisar a amostra controlo do *kit* e tentar descobrir alguma relação entre esta e as amostras, foi possível observar que as amostras 17 e controlo apresentaram exatamente a mesma concentração de DNA (4,2 ng/μL) por quantificação no *NanoDrop*[™], e que foram as únicas a apresentar uma concentração demasiado baixa de DNA na quantificação pelo *Qubit*[®] *dsDNA HS*, uma evidência de que algo poderia ter acontecido durante o trabalho laboratorial, possivelmente uma troca de material entre estas duas amostras.

Houve ainda outra situação a considerar, pois para fungos (figura 18), ao contrário da situação para bactérias (figura 17), não foi obtida nenhuma banda na eletroforese do PCR da amostra do controlo do *kit*. No entanto, a sequenciação resultou num elevado número de *reads* para essa mesma amostra (tabela 7). A ausência de banda na região ITS2 não implica a não existência de material genético de outras regiões do DNA, pelo que o elevado número de *reads* neste caso pode resultar da fraca qualidade do material genético, ou da existência de material de outras proveniências. De qualquer forma, uma vez detetada a contaminação bacteriana, foi também necessário realizar a análise da amostra controlo em termos de fungos, independentemente de ter apresentado, ou não, banda na região alvo da amplificação.

Mais do que procurar um possível alvo/foco de contaminação ou motivo da mesma, pretendeu-se procurar quais as amostras mais possivelmente afetadas por esta ocorrência, e sobretudo, identificar quais as OTUs que devem ser eliminadas da análise do microbioma destas lesões.

A tabela 8 resume a informação da sequenciação de bactérias e fungos que entrou para análise no MA.

Tabela 8. Tabela resumo do *upload* para o *MicrobiomeAnalyst* dos dados de sequenciação de bactérias e fungos para análise do controlo do *kit*.

	Bactérias	Fungos
Base de dados para classificação das OTU	<i>Greengenes</i>	<i>Unite</i>
Número de OTU	863	285
Número de fatores experimentais	2	2
Número total de <i>reads</i>	213549	348253
Número médio de <i>reads por amostra</i>	9706	15829
Número máximo de <i>reads por amostra</i>	15052	31396
Número mínimo de <i>reads por amostra</i>	3342	2974
Número de amostras	22	22
Correspondência das amostras entre tabelas de metadata e OTU	Sim	Sim

A primeira abordagem no MA, foi recorrer à análise de clusters para agrupar as amostras em função das suas dissimilaridades, de forma a encontrar possíveis relações entre o controlo do *kit* e outras amostras, minimizando o esforço de procura de eventuais contaminações entre amostras. O diagrama criado com as relações hierárquicas entre as amostras, a nível das OTU de bactérias (figura 18A), mostra que, tal como suspeitado, o esfregaço 17 é aquele que possui uma relação mais próxima com a amostra controlo do *kit*. Para fungos (figura 18B), esta análise mostrou uma maior proximidade entre o esfregaço 6 e o controlo do *kit*, surgindo, logo de seguida, uma relação próxima destas com o esfregaço 17.

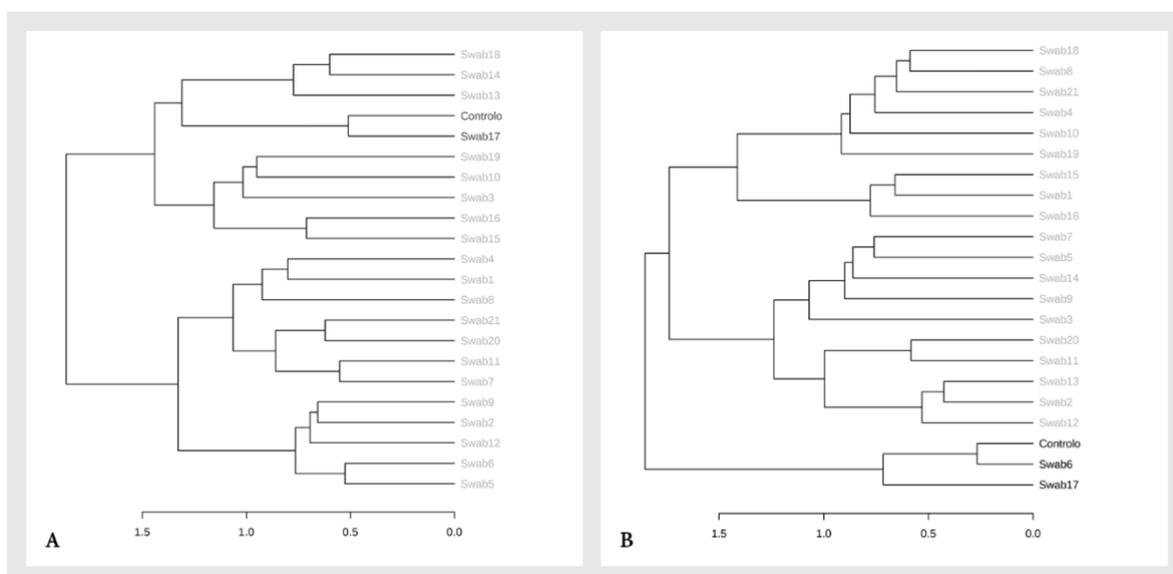


Figura 18. Dendrogramas das amostras de bactérias (A) e de fungos (B), gerado no *MicrobiomeAnalyst* a nível OTU, utilizando o índice de *Bray-Curtis* como medida de distância e o método de *Ward* como algoritmo de *clustering*. Destacadas as amostras hierarquicamente mais próximas do controlo do *kit*.

No mesmo sentido, procedeu-se à análise da beta diversidade das amostras, que permite a comparação explícita das comunidades microbianas em função da sua composição. Uma vez que esta fornece uma medida da distância entre as amostras,

4. Resultados e Discussão

possibilita, tal como anteriormente, identificar as amostras mais semelhantes ao controlo do *kit*.

A beta diversidade das amostras, foi representada usando os métodos de ordenação *non-metric multidimensional scaling* (nMDS) (resultados não exibidos) e análise de coordenadas principais (PCoA) (figura 19). Estes métodos resumem os dados da abundância da comunidade ao produzirem um espaço de ordenação de baixa dimensão (poucos eixos), onde existe uma relação direta entre a distância dos pontos no referencial e a semelhança da comunidade das amostras.

Em ambas as situações podemos observar o mesmo que anteriormente, há de facto grande semelhança entre as amostras mencionadas e o controlo do *kit*, traduzida pela proximidade dos pontos das respetivas amostras, em ambos os gráficos. Esta semelhança é validada pelos *p-values* significativos das análises de variância multivariada permutacional (PERMANOVA) das duas situações, que apontam para diferenças consideráveis do esfregaço 17 e do controlo do *kit*, em relação às restantes amostras (figura 19A), e dos esfregaços 6 e 17 e do controlo do *kit* em comparação com as demais (figura 19B).

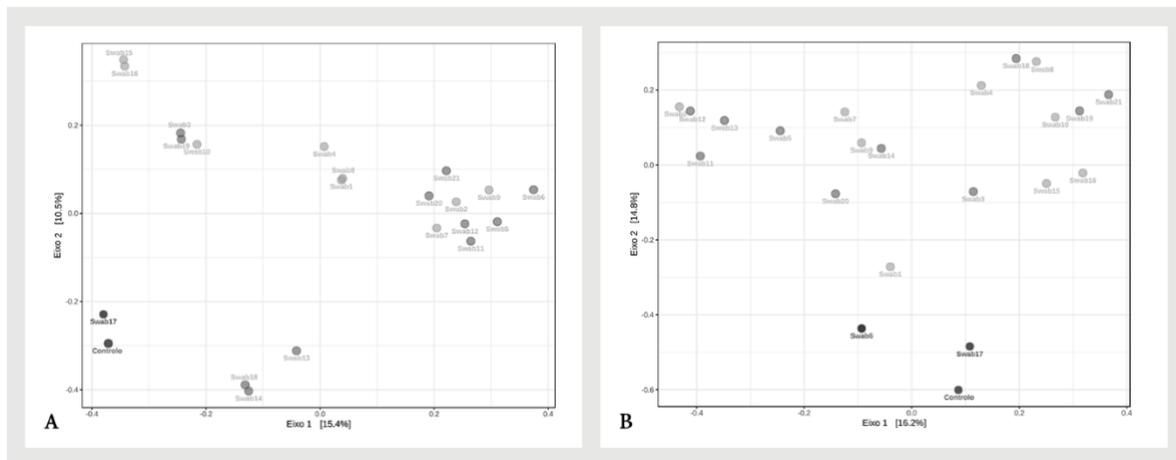


Figura 19. Análise de beta diversidade de bactérias (A) e fungos (B), através do método de ordenação PCoA gerado no *MicrobiomeAnalyst* a nível OTU, utilizando o índice de *Bray-Curtis* como medida de distância. Destacadas as amostras hierarquicamente mais próximas do controlo do *kit*. Teste estatístico entre o grupo de amostras destacadas e as restantes: **A.** PERMANOVA: $F\text{-value}=2,140$; $R^2=0,097$; $p\text{-value}<0,004$. **B.** PERMANOVA: $F\text{-value}=3,573$; $R^2=0,15158$; $p\text{-value}<0,001$.

A priori, os dados destas duas análises, sobretudo em relação a bactérias, apontam uma provável contaminação do controlo negativo da extração, sem qualquer material genético, com material proveniente do esfregaço 17, por contacto com uma ponta utilizada nessa amostra, ou qualquer outro erro protocolar durante a extração de DNA. Isto fez com que a amostra que serviu como controlo do *kit* de extração, e consequentemente como branco durante a quantificação de DNA, apresentasse material genético. No entanto, a diferença dos resultados para os fungos, que embora possam ser resultantes da fraca

qualidade do DNA fúngico presente no controlo do *kit*, faz com que esta conclusão seja apenas um forte indício.

De facto, mesmo provando que os componentes do *kit* de sequenciação utilizado não estavam contaminados (ao repetir o procedimento de obtenção deste controlo, não foi observada qualquer presença de DNA), é-nos impossível afirmar que apenas houve contacto e conseqüente contaminação entre duas amostras.

A forma mais correta para salvaguardar a veracidade de todos os dados e avançar com segurança para o estudo do microbioma é realizar uma análise da composição da amostra controlo, e retirar da análise do microbioma todas e quaisquer OTU dominantes nessa amostra, que também estejam presentes com alguma representatividade noutros esfregaços.

A amostra controlo é composta sobretudo por bactérias (figura 20) das famílias *Leuconostocaceae* e *Moraxellaceae*, e, em menor quantidade, por *Enterobacteriaceae*. Explorando esta informação, descobrimos que estas duas primeiras famílias correspondiam, praticamente na totalidade, aos géneros *Leuconostoc* e *Acinetobacter*, respetivamente. Verificando a tabela OTU, concluiu-se que estas três famílias estavam representadas por três OTU da base de dados *GreenGenes*, respetivamente, ID#3513545 (1337 reads), ID#4449456 (697 reads) e ID#812351 (185 reads).

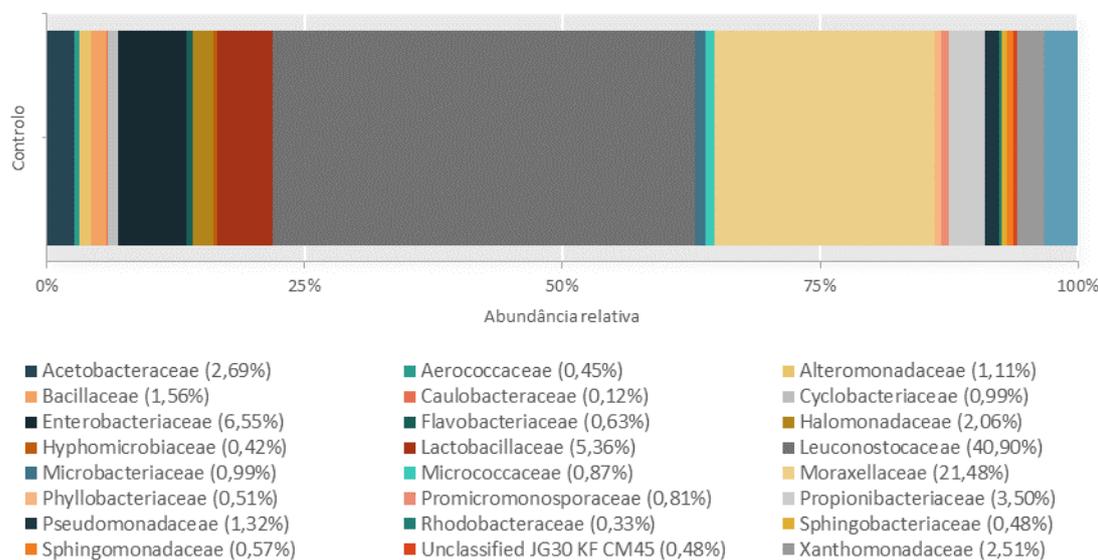


Figura 20. Abundância relativa das famílias de bactérias encontradas no controlo do *kit*.

Uma vez que estas OTU estavam presentes noutras amostras com alguma abundância, e pelos motivos já expostos, estas foram retiradas da tabela final, e, portanto, excluídas da análise do microbioma.

Procedendo da mesma forma, em termos de fungos (figura 21), o domínio da família *Saccharomycetaceae* é acentuado, havendo também a presença de *Saccharomycodaceae*. A primeira, representada pela espécie *Saccharomyces cerevisiae*,

4. Resultados e Discussão

e a segunda, pelo género *Hanseniaspora*. Cada uma destas, estava representada por duas OTU da base de dados *UNITE*, respetivamente, SH004658.07FU_AB018043_refs e SH492529.07FU_KP196597_reps_singleton, e SH002002.07FU_AB501148_refs e New.ReferenceOTU14, as quais foram também eliminadas.

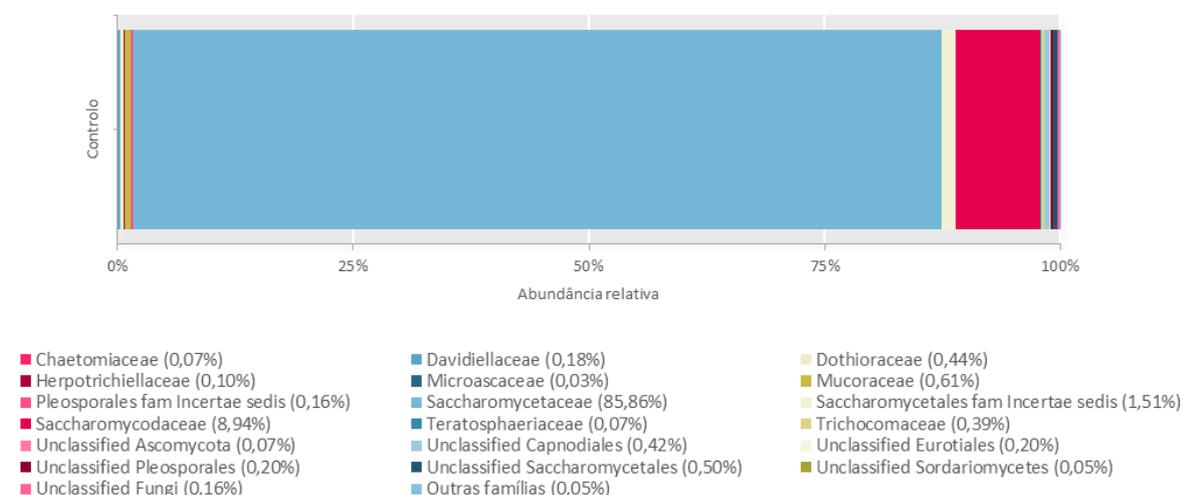


Figura 21. Abundância relativa das famílias de fungos encontradas no controlo do *kit*.

A remoção de OTUs derivada da contaminação que ocorreu, teve uma influência bastante reduzida no número de *reads* sequenciado. Apenas 0,57% de *reads* de bactérias e 3,34% de *reads* de fungos foram removidos por possivelmente serem resultado de contaminação, sendo que as amostras mais afetadas por esta remoção foram os esfregaços 6 e 17, dado a sua proximidade à amostra controlo do *kit*.

4.6. Análise dos dados

Após a filtração dos OTUs, foram consideradas 860 OTUs correspondendo a 208703 *reads*, para análise da comunidade bacteriana, e 281 OTUs correspondendo a 302108 *reads*, para a análise da comunidade fúngica.

Através da análise dos histogramas e gráficos quartil-quartil (figura 22), realizados no *GraphPad Prism 8*, obteve-se a garantia de que os dados de bactérias e fungos não seguem distribuição normal, mesmo quando transformados logaritmicamente. Os histogramas dos dados são distorcidos à direita, o que indica a ocorrência de altos valores com baixa frequência. Os pontos dos gráficos quartil-quartil resultante não assentam numa linha com declive 1, o que significa que os dados não seguem a forma atribuída à distribuição de dados normais.

Foi feita a mesma análise para outras transformações, usando os testes de normalidade de *Kolmogorov-Smirnov*, *Shapiro-Wilk* e *D'Agostino & Pearson* (dados não

exibidos), e em qualquer uma das situações obteve-se sempre o mesmo resultado, os dados seguem uma distribuição não-normal.

Todos estes indícios enunciam a necessidade de realizar todos os testes estatísticos desta análise considerando métodos não-paramétricos, e a importância da normalização dos dados para análise diferencial.

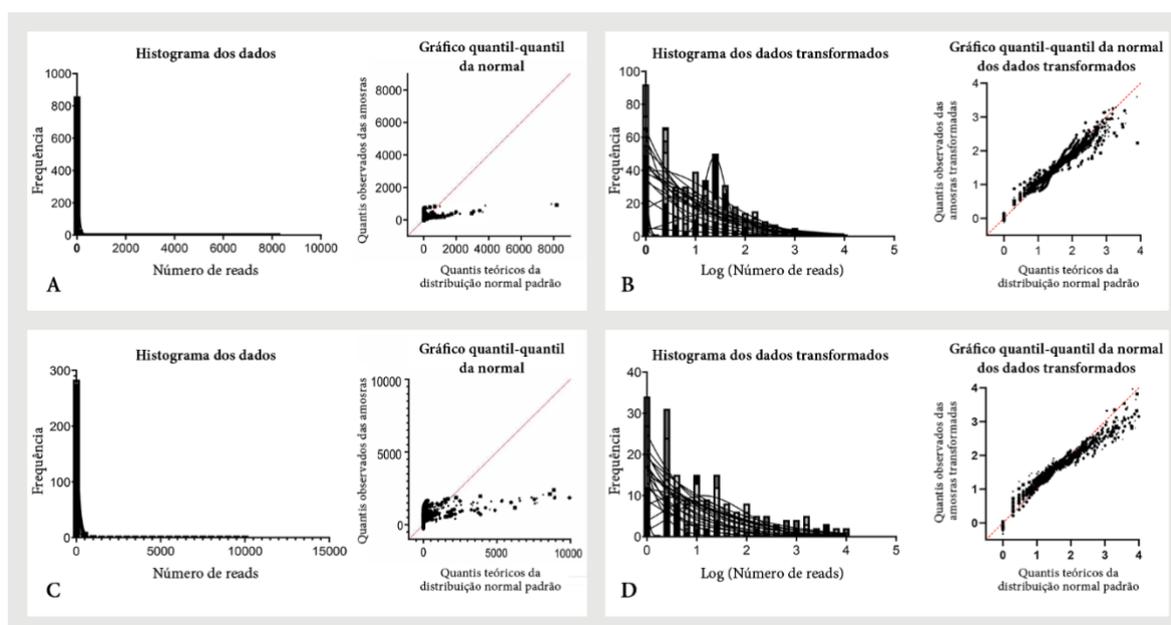


Figura 22. Histogramas e gráficos quantil-quantil dos dados de bactérias (A, B) e fungos (C, D), antes (A, C) e após transformação logarítmica (B, D) obtidos através do *software GraphPad Prism 8*.

4.7. Descrição do microbioma

As 21 amostras para a comunidade bacteriana formam uma biblioteca de 860 OTUs, composta por 208703 *reads* (figura 23A). Em média, têm 9938 *reads*, sendo que a amostra mais pequena contém 3228 *reads* e a maior 15052 *reads*.

Por sua vez, os fungos constituem uma biblioteca menor em número de OTUs, com apenas 281, mas formada por um número muito mais elevado de *reads*, 302108 (figura 23B). Em média, as amostras têm 14386 *reads*, a mais pequena com 2045 *reads* e a maior com 28725 *reads*, valores com uma distribuição mais alargada que os anteriores.

Há alguma heterogeneidade entre as amostras em relação ao número de *reads*, mas nenhuma delas se revelou *outlier* através do cálculo da variação interquartilica dos dados.

4. Resultados e Discussão

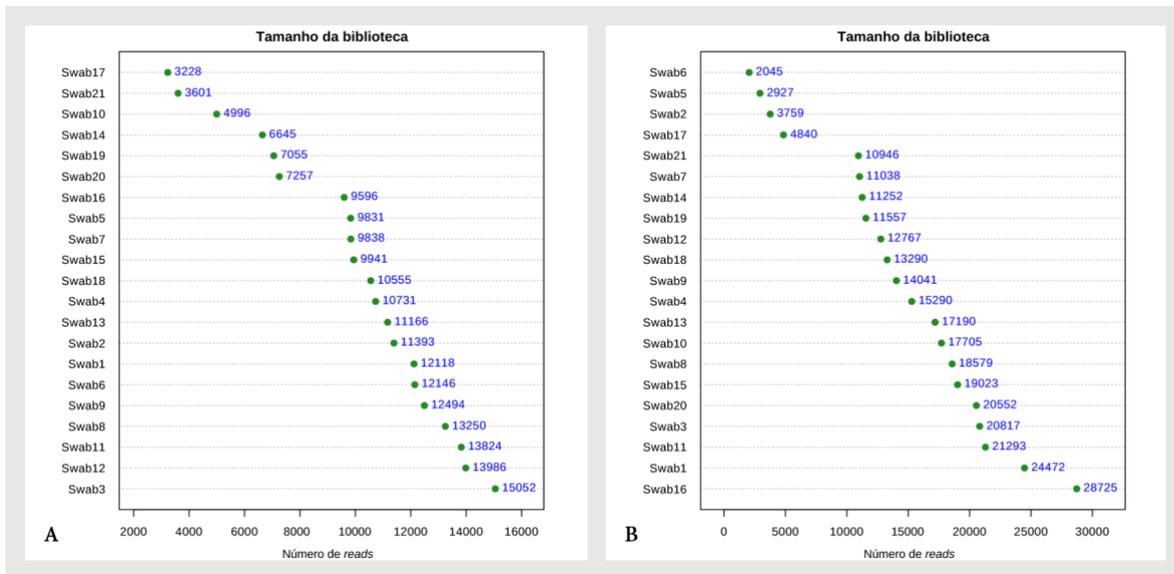


Figura 23. Bibliotecas de bactérias (A) e fungos (B) do microbioma das lesões na casaca no sobreiro, obtidas no *MicrobiomeAnalyst*.

4.7.1. Alfa diversidade

Para medir a diversidade dentro das amostras (tabela 9 e 10) recorreremos aos estimadores de diversidade mais comuns, os índices de *Chao1* (Chao, 1984), de *Simpson* (Simpson, 1949) e de *Shannon* (Shannon e Weaver, 1949).

O primeiro calcula o número de OTUs esperadas, a partir dos dados de abundância das amostras, estimando possíveis OTUs raras que não foram encontradas por subamostragem. O índice de *Simpson* permite calcular a dominância de OTUs dentro das amostras – valores próximos de 100% indicam que uma OTU domina a amostra. Por fim, o índice de *Shannon* é usado para determinar a equitatividade das amostras, isto é, a distribuição das abundâncias das OTUs – um valor próximo de 100% pressupõe OTUs em igual proporção.

Tal como apontado anteriormente, a comunidade de bactérias tem tendencialmente mais OTUs do que a de fungos. Em ambas as comunidades o número de OTUs observado, praticamente nunca está dentro do intervalo estimado para o número de OTUs esperado, isto porque existem espécies não descobertas em praticamente todos os levantamentos taxonômicos ou inventários de espécies (Chao e Chiu, 2016). Por sua vez, os valores díspares de dominância (entre 1,28% e 73,28%) e de equitatividade (entre 10,28% e 70,82%) das amostras apontam, uma vez mais, para microbiomas heterogêneos.

Tabela 9. Métricas de alfa diversidade aplicadas aos dados de sequenciação de bactérias dos esfregaços recolhidos, calculadas através dos estimadores de riqueza e ou variância *Chao1*, *Simpson* e *Shannon*.

	OTUs observadas	OTUs esperadas	Dominância	Equitatividade
Esfregaço 1	137	227,46 ± 36,79	7,31%	45,29%
Esfregaço 2	235	267,23 ± 12,23	3,32%	60,01%
Esfregaço 3	41	50,00 ± 8,04	34,35%	22,88%
Esfregaço 4	198	215,11 ± 8,86	14,66%	44,33%
Esfregaço 5	291	329,68 ± 13,82	2,61%	64,35%
Esfregaço 6	284	355,72 ± 23,76	2,60%	62,78%
Esfregaço 7	239	318,44 ± 26,32	13,17%	46,65%
Esfregaço 8	260	379,45 ± 39,13	4,10%	59,20%
Esfregaço 9	279	344,81 ± 22,90	3,63%	60,04%
Esfregaço 10	221	271,65 ± 21,04	1,28%	70,82%
Esfregaço 11	311	402,00 ± 25,28	5,54%	54,95%
Esfregaço 12	307	396,38 ± 25,96	3,37%	59,64%
Esfregaço 13	347	411,77 ± 19,54	3,45%	64,08%
Esfregaço 14	280	336,00 ± 18,43	3,30%	63,31%
Esfregaço 15	42	94,50 ± 33,44	26,04%	23,55%
Esfregaço 16	74	144,00 ± 33,37	73,28%	10,28%
Esfregaço 17	114	147,07 ± 15,56	12,75%	42,01%
Esfregaço 18	357	428,19 ± 22,96	2,26%	68,82%
Esfregaço 19	131	144,13 ± 9,01	3,69%	58,03%
Esfregaço 20	189	248,91 ± 22,03	5,64%	54,15%
Esfregaço 21	195	276,43 ± 24,88	8,87%	46,76%

Tabela 10. Métricas de alfa diversidade aplicadas aos dados de sequenciação de fungos dos esfregaços recolhidos, calculadas através dos estimadores de riqueza e ou variância *Chao1*, *Simpson* e *Shannon*.

	OTUs observadas	OTUs esperadas	Dominância	Equitatividade
Esfregaço 1	50	61,00 ± 8,47	16,26%	30,45%
Esfregaço 2	50	60,00 ± 6,66	36,62%	23,62%
Esfregaço 3	28	41,20 ± 10,21	31,46%	19,66%
Esfregaço 4	97	97,46 ± 0,88	14,02%	39,70%
Esfregaço 5	86	103,11 ± 8,86	19,13%	35,44%
Esfregaço 6	62	83,08 ± 11,65	15,76%	34,60%
Esfregaço 7	128	152,39 ± 11,01	35,46%	27,83%
Esfregaço 8	112	126,44 ± 8,15	12,62%	37,59%
Esfregaço 9	88	111,00 ± 12,46	18,30%	30,51%
Esfregaço 10	96	111,11 ± 9,73	34,66%	26,35%
Esfregaço 11	96	119,10 ± 13,14	24,35%	27,81%
Esfregaço 12	70	75,00 ± 4,13	43,74%	19,05%
Esfregaço 13	132	159,19 ± 13,13	22,43%	34,46%
Esfregaço 14	140	160,65 ± 10,48	5,20%	50,28%
Esfregaço 15	115	181,00 ± 32,96	26,43%	29,02%
Esfregaço 16	60	77,10 ± 10,41	21,56%	28,03%
Esfregaço 17	65	66,20 ± 1,84	21,71%	35,68%
Esfregaço 18	110	133,33 ± 13,64	9,28%	42,10%
Esfregaço 19	75	113,50 ± 22,94	16,07%	36,38%
Esfregaço 20	112	131,25 ± 10,87	26,28%	30,40%
Esfregaço 21	70	87,50 ± 10,11	24,62%	28,58%

4.7.2. Diversidade taxonômica

A comparação quantitativa direta das abundâncias relativas de cada amostra a diferentes níveis taxonômicos (figuras 24-27), permite-nos um vislumbre do panorama geral da composição das comunidades bacterianas e fúngicas das lesões na casca do sobreiro.

O microbioma bacteriano dos sangramentos é composto por uma diversidade de microrganismos evidente a nível do filo (figura 24), mas sobretudo a nível de famílias (figura 25). De uma forma geral, as amostras são dominadas por Proteobactérias, estando também representados, com alguma abundância, os filos Actinobacteria, Firmicutes e Bacteroidetes (figura 24).

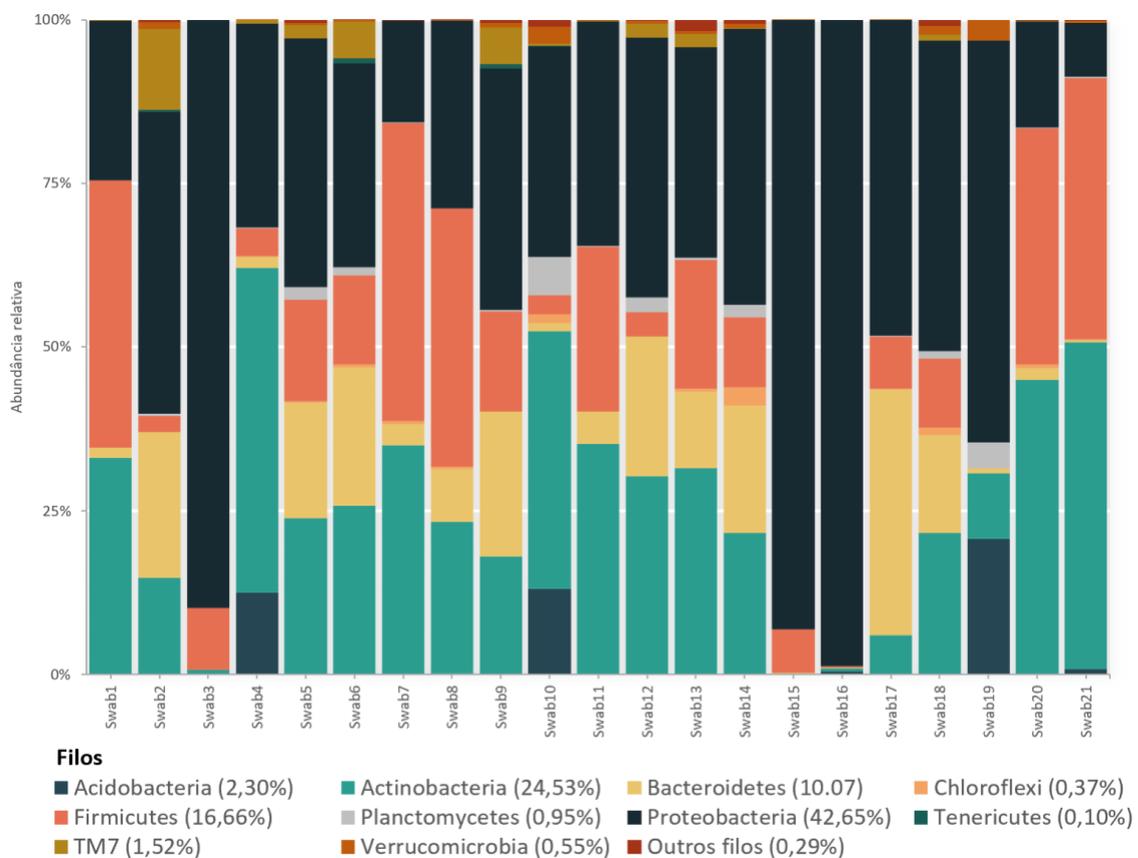
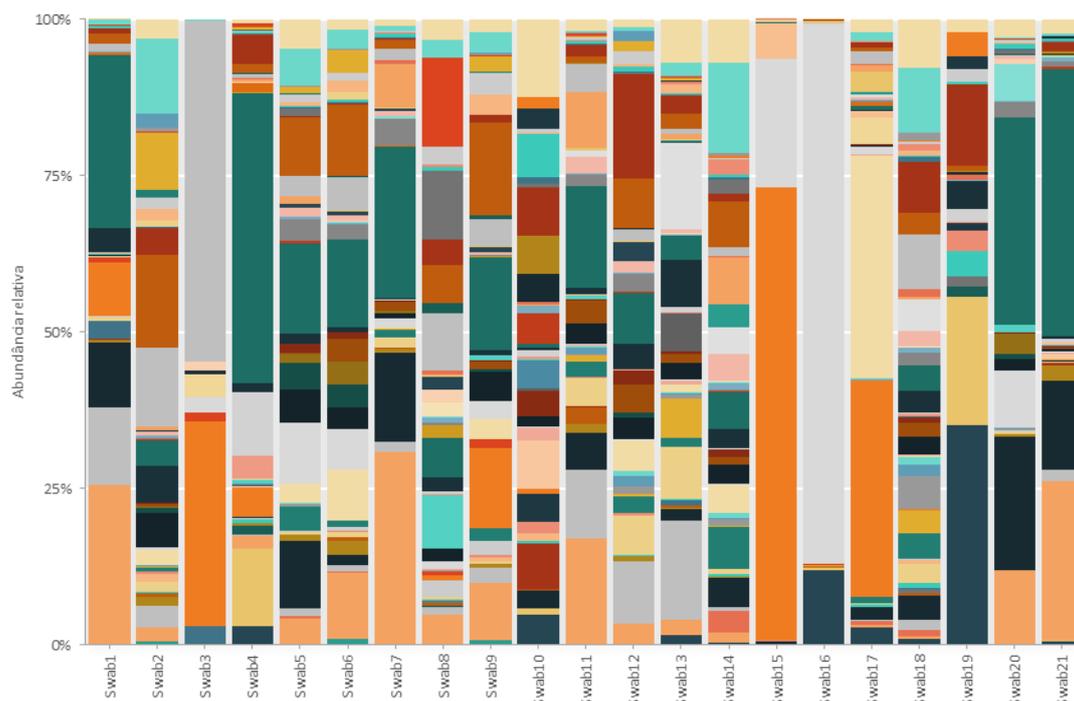


Figura 24. Gráfico de barras empilhadas da abundância relativa de bactérias a nível do filo, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em "Outros filis". Média das abundâncias relativas indicada junto ao respetivo *taxon*.

O mesmo pode ser constatado ao nível taxonômico das famílias de bactérias (figura 25), onde a elevada heterogeneidade das amostras, e entre amostras, está claramente rubricada pelas diferenças entre as suas abundâncias relativas. A este nível, considerando todo o microbioma, destacam-se as Micrococcaceae, as Enterobacteriaceae, as Aerococcaceae, as Halomonadaceae e as Rhodobacteraceae (figura 25).

O microbioma de lesões na casca no declínio do sobreiro



Famílias

■ Acetobacteraceae (2,96%)	■ Achleplasmataceae (0,10%)	■ Acidobacteriaceae (1,67%)
■ Aerococcaceae (7,21%)	■ Alteromonadaceae (0,27%)	■ Aurantimonadaceae (3,09%)
■ Bacillaceae (4,49%)	■ Beijerinckiaceae (0,15%)	■ Beutenbergiaceae (0,58%)
■ Bogoriellaceae (0,26%)	■ Bradyrhizobiaceae (0,37%)	■ Burkholderiaceae (0,12%)
■ Carnobacteriaceae (0,34%)	■ Caulobacteraceae (0,39%)	■ Cellulomonadaceae (1,39%)
■ Chitinophagaceae (0,18%)	■ Chthoniobacteraceae (0,33%)	■ Clostridiaceae (0,28%)
■ Conexibacteraceae (0,29%)	■ Cyclobacteriaceae (1,30%)	■ Cytophagaceae (0,61%)
■ Enterobacteriaceae (7,98%)	■ Enterococcaceae (0,21%)	■ Erythrobacteraceae (0,45%)
■ Euzebyaceae (0,25%)	■ Flammeovirgaceae (0,16%)	■ Flavobacteriaceae (3,02%)
■ Gaiellaceae (0,42%)	■ Geodermatophilaceae (0,18%)	■ Halomonadaceae (6,77%)
■ Hyphomicrobiaceae (1,99%)	■ Hyphomonadaceae (0,50%)	■ Idiomarinaceae (0,45%)
■ Intrasporangiaceae (0,97%)	■ Isosphaeraceae (0,59%)	■ Kineosporiaceae (0,33%)
■ Koribacteraceae (0,24%)	■ Lachnospiraceae (0,59%)	■ Lactobacillaceae (0,39%)
■ Leuconostocaceae (0,35%)	■ Methylobacteriaceae (0,24%)	■ Methylocystaceae (0,74%)
■ Microbacteriaceae (2,03%)	■ Micrococcaceae (12,74%)	■ Mogibacteriaceae (0,13%)
■ Moraxellaceae (0,12%)	■ Mycobacteriaceae (0,29%)	■ Nitriliruptoraceae (1,01%)
■ Nocardiodiaceae (0,32%)	■ Nocardiospiceae (0,34%)	■ Oxalobacteraceae (0,12%)
■ Paenibacillaceae (0,24%)	■ Phyllobacteriaceae (0,81%)	■ Planococcaceae (1,18%)
■ Porphyromonadaceae (0,37%)	■ Promicromonosporaceae (0,22%)	■ Propionibacteriaceae (0,24%)
■ Pseudomonadaceae (1,35%)	■ Pseudonocardiceae (0,21%)	■ Rhodobacteraceae (5,35%)
■ Rhodospirillaceae (0,26%)	■ Ruminococcaceae (0,13%)	■ Solibacteraceae (0,31%)
■ Sphingobacteriaceae (4,05%)	■ Sphingomonadaceae (3,31%)	■ Streptomycetaceae (0,78%)
■ Unclassified Acidimicrobiales (0,16%)	■ Unclassified Actinomycetales (0,62%)	■ Unclassified Alphaproteobacteria (0,11%)
■ Unclassified EW055 (0,52%)	■ Unclassified JG30 KF CM45 (0,23%)	■ Unclassified Rhizobiales (0,80%)
■ Unclassified Solirubrobacterales (0,27%)	■ Unclassified TM7 1 (0,09%)	■ Unclassified TM7 3 (0,91%)
■ Unclassified WD2101 (0,30%)	■ Veillonellaceae (0,77%)	■ Verrucomicrobiaceae (0,13%)
■ Weeksellaceae (0,19%)	■ Xanthomonadaceae (2,81%)	■ Outras famílias (3,00%)

Figura 25. Gráfico de barras empilhadas da abundância relativa de bactérias a nível da família, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em “Outras famílias”. Média das abundâncias relativas indicada junto ao respetivo *taxon*.

As *Micrococcaceae*, do filo *Actinobacteria*, encontradas nas amostras são, sobretudo, do género *Nesterenkonia*. Espécies de *Nesterenkonia*, bactérias relativamente patogénica em humanos, têm sido isoladas a partir de nichos ecológicos muito distintos, incluindo ambientes extremos, especialmente em habitats hipersalinos (Chander et al.,

4. Resultados e Discussão

2017), como por exemplos em salgueiros (*Populus euphratica*) e amostras de solo de desertos da China (Li, 2004; Liu et al., 2015).

Atente-se para a presença das Proteobacterias da família Enterobacteriaceae, à qual pertencem praticamente todas as bactérias que têm vindo a ser associadas ao AOD (identificadas na secção 1.3.2.). A maioria das Enterobacteriaceae encontradas não foram classificadas até ao género, mas foi identificado o género *Dickeya* com relativa abundância. Bactérias deste género estão associadas a numerosas doenças de culturas e de plantas, afetando uma ampla gama de hospedeiros em todo o mundo (Toth et al., 2011). De salientar que este género já tem sido identificado em *Quercus* e é bastante semelhante ao género *Brenneria* (Brady et al., 2012) que tem sido associado ao AOD.

Estranhamente, as bactérias Firmicutes detetadas, da família Aerococcaceae, são na sua maioria de *Marinilactibacillus psychrotolerans*. Tal como as *Nesterenkonia*, trata-se um microrganismo halofílico, um extremófilo que se pode desenvolver em ambientes com altas concentrações de sal (Ishikawa, 2003). Estes são também alcalifílicos, pois encontram condições ótimas de crescimento em habitats extremamente alcalinos (Ishikawa, 2003).

Relativamente às Halomonadaceae encontradas, também elas Proteobacteria, são sobretudo do género *Candidatus Portiera*. Neste género, incluem-se bactérias endossimbióticas primárias obrigatórias da mosca *Bemisia tabaci*, uma das pragas agrícolas mais prejudiciais a nível mundial (Jiang et al., 2013).

Finalmente, as Proteobacteria da família Rhodobacteraceae, cuja maioria das OTUs não conseguiu ser classificada a nível do género, são membros-chave da comunidade bacteriana responsável pela colonização primária do biofilme formado em águas costeiras do Mediterrâneo Oriental (Elifantz et al., 2013).

Relativamente a fungos, a nível do filo (figura 26) há uma clara dominância de Ascomycota, mas esta não se traduz a nível das famílias (figura 27), nível taxonómico onde volta a imperar a diferença.

Neste microbioma, destacam-se os fungos das famílias Davidiellaceae, Dothioraceae, Saccharomycetaceae e Trichocomaceae, das ordens Pleosporales e Capnodiales, e da classe Sordariomycetes (figura 27), todos eles naturalmente pertencentes ao filo Ascomycota. Sobre estes, apenas é possível tentar enquadrar a presença nestas amostras para os fungos com família identificada. Os restantes enquadram-se em ordens e classes muito abrangentes, com elevada multiplicidade de características, que lhes permite a colonização de variadíssimos habitats.

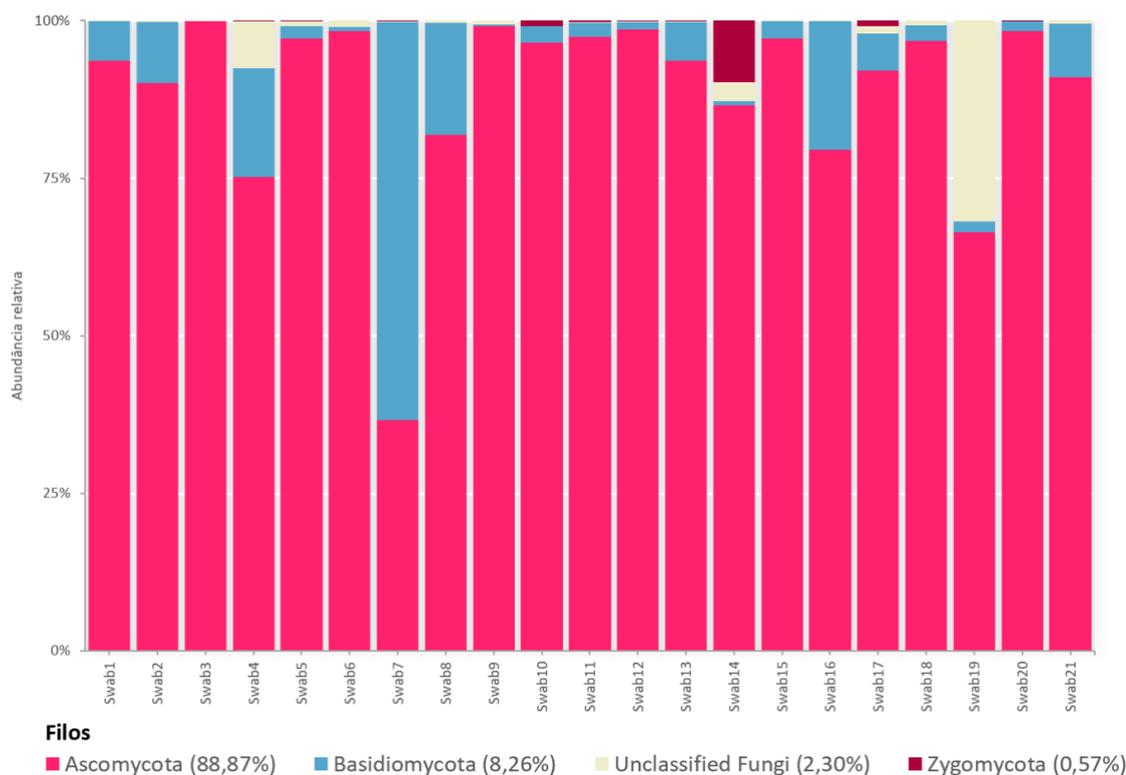


Figura 26. Gráfico de barras empilhadas da abundância relativa de fungos a nível do filo, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em “Outros filis”. Média das abundâncias relativas indicada junto ao respetivo *taxon*.

Os fungos da família *Davidiellaceae* identificados nestas amostras, correspondem na sua totalidade ao género *Cladosporium*. Estes microrganismos cosmopolitas, são frequentemente encontrados em todos os tipos de plantas, são isolados a partir do solo, do ar, alimentos e outras matérias orgânicas, ou colonizam, como invasores secundários, lesões foliares provocadas por outros patógenos (Bensch et al., 2012).

Os *Dothioraceae* mais encontrados nas amostras, correspondem a uma espécie com importância biotecnológica, o *Aureobasidium pullulans*. Estes produtores do exopolissacarídeo pululano (utilizado para diversos fins nas indústrias alimentar e farmacêutica, por exemplo), podem ser encontrados em ambientes normalmente ligados a flutuações de água, como a filosfera das plantas (Zalar et al., 2008). Têm também capacidade de explorar habitats ricos em stressantes osmóticos, normalmente hipersalinos (Zalar et al., 2008).

Os *Saccharomycetaceae* encontrados compõem uma diversidade de géneros, onde se destacam os *Lachancea*. Estes, normalmente, correspondem a leveduras utilizadas em processos de fermentação de vinhos (Benito et al., 2015; Gobbi et al., 2013) e outros alimentos. Vivem como saprófitas, muitas vezes em associação com plantas, animais e as suas interfaces (Mestre et al., 2010), tendo sido inclusive isolados a partir da

4. Resultados e Discussão

casca de carvalhos (Freel et al., 2015), mas também em plantas de ambientes salinos marinhos (Fell et al., 2004).

Por fim, na última família de fungos encontrada, nos Trichocomaceae, foram identificados microrganismos do gênero *Penicillium*, sobretudo *P. spinulosum*. Estes fungos com potencial inflamatório (Jussila et al., 2002), fazem parte da flora fúngica de edifícios, normalmente danificados pela humidade (Jussila et al., 2002), mas também são dos fungos mais dominantes em serrim de madeiras (Samson, 2000).

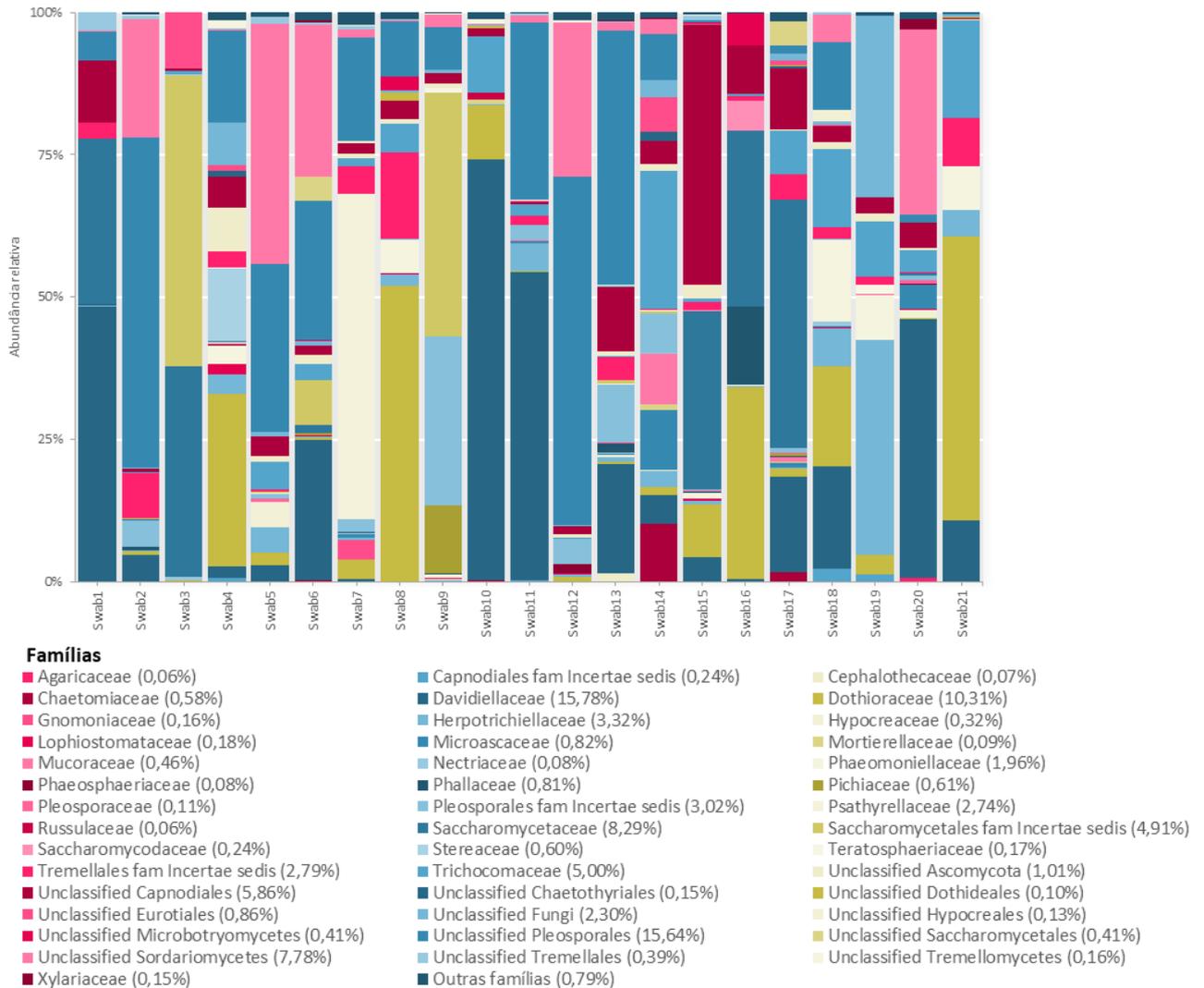


Figura 27. Gráfico de barras empilhadas da abundância relativa de fungos a nível da família, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em “Outras famílias”. Média das abundâncias relativas indicada junto ao respetivo *taxon*.

Parece existir um claro agrupamento das amostras, tanto em relação a bactérias, como em relação a fungos. São observados padrões de diversidade semelhantes entre algumas amostras, em ambos os níveis taxonómicos. Torna-se necessário procurar correspondência entre este agrupamento dos dados e os fatores ambientais recolhidos.

Para além disso, não basta olhar apenas para os microrganismos mais abundantes,

é necessário considerar espécies raras, aqueles que marcam presença em vários esfregaços e os que são o principal fator distinguidor entre grupos de amostras.

4.7.3. Microbioma nuclear

Na tentativa de estabelecer um microbioma nuclear destas lesões, isto é, de encontrar um conjunto de taxa que permaneça inalterado na composição das populações, procuraram-se microrganismos com elevada prevalência e considerável abundância relativa. Foram encontradas algumas classes de bactérias e fungos com elevada prevalência (tabela 11).

Tabela 11. Classes de bactérias e fungos detetadas nas amostras com prevalência superior a 50%.

Classes de bactérias	Prevalência	Classes de fungos	Prevalência
Gammaproteobacteria	100%	Eurotiomycetes	100%
Actinobacteria	100%	Dothideomycetes	100%
Bacilli	95%	Tremellomycetes	90%
Alphaproteobacteria	95%	Agaricomycetes	81%
Sphingobacteriia	86%	Sordariomycetes	71%
Nitriliruptoria	67%	Unclassified Fungi	67%
Planctomycetia	62%	Unclassified Ascomycota	67%
Cytophagia	62%	Saccharomycetes	62%
Clostridia	62%		
Betaproteobacteria	57%		
TM7 3	52%		

No entanto, apenas parte destes organismos cumprem os limites mínimos de deteção, normalmente utilizado em estudos do núcleo de microbiomas (Sudarshan, 2017; Wirth, 2018). A definição do microbioma nuclear é, de facto, sensível aos limites utilizados (Salonen et al. 2012). Na figura 28 estão todas as classes de microrganismos que atingem pelo menos 95% de prevalência e 0,10% de abundância relativa nas amostras, valores extremamente conservadores usados como limite de deteção (Surdashan, 2017).

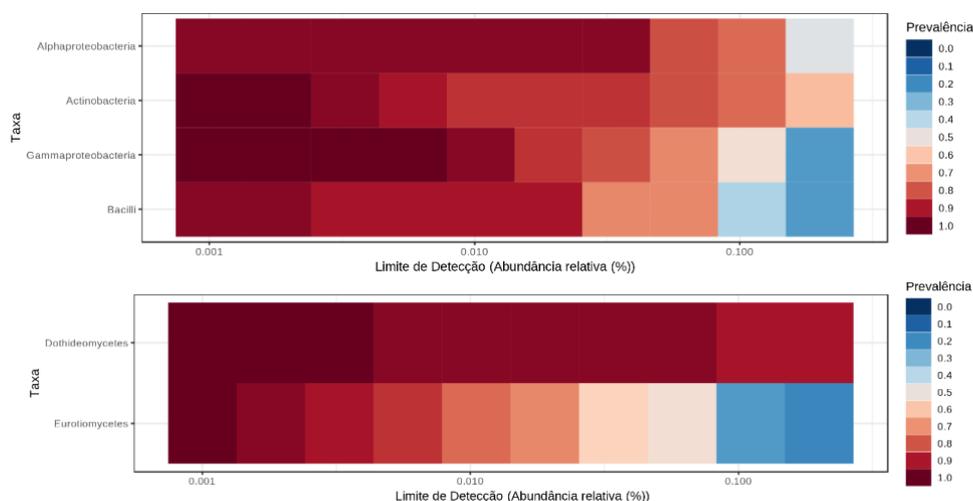


Figura 28. Classes de bactérias e fungos detetadas nas amostras com pelo menos 95% de prevalência e 0,10% de abundância relativa.

4. Resultados e Discussão

De entre as classes de microrganismos destacadas, apenas as bactérias das classes Alphaproteobacteria e Actinobacteria e os fungos Dothideomycetes, possuem prevalências elevadas com abundâncias relativas superiores a 0,10%.

Das famílias enunciadas anteriormente, as bactérias Rhodobacteraceae pertencem às Alphaproteobacteria e as Micrococcaceae fazem parte da classe Actinobacteria, e os fungos das famílias Davidiellaceae e Dothioraceae são Dothideomycetes. Assim, para além da sua elevada abundância no microbioma, estes dados indicam uma elevada prevalência deste tipo de microrganismos nas amostras.

Não sendo possível encontrar *taxa* a níveis taxonómicos inferiores, que cumprissem os limites mínimos de deteção, torna-se complicado estabelecer um grupo de bactérias e/ou fungos que constitua o núcleo do microbioma destas lesões. Podemos inferir, uma vez mais, que estes microbiomas são bastante heterogêneos.

4.7.4. Agrupamento das amostras

De facto, produzindo um *heatmap* do agrupamento hierárquico das amostras em função do algoritmo de *clustering* de Ward (Ward, 1963) a nível das classes (figuras 29, 30), podemos observar que a diversidade da composição dentro das próprias amostras é o principal fator de agrupamento entre elas. Mesmo em níveis taxonómicos elevados, não se observa a típica distribuição das amostras em função dos fatores ambientais recolhidos, característica deste tipo de gráficos.

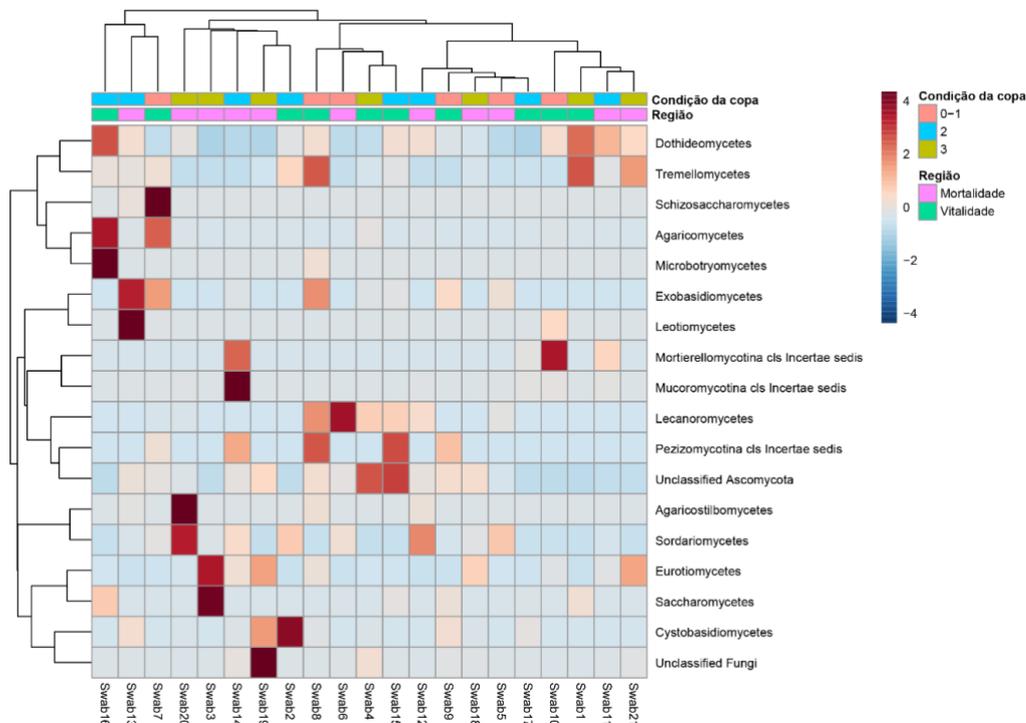


Figura 29. Heatmap do *clustering* hierárquico das amostras de fungos a nível taxonómico das classes, com base na medida de distância euclidiana e no algoritmo de *clustering* de Ward.

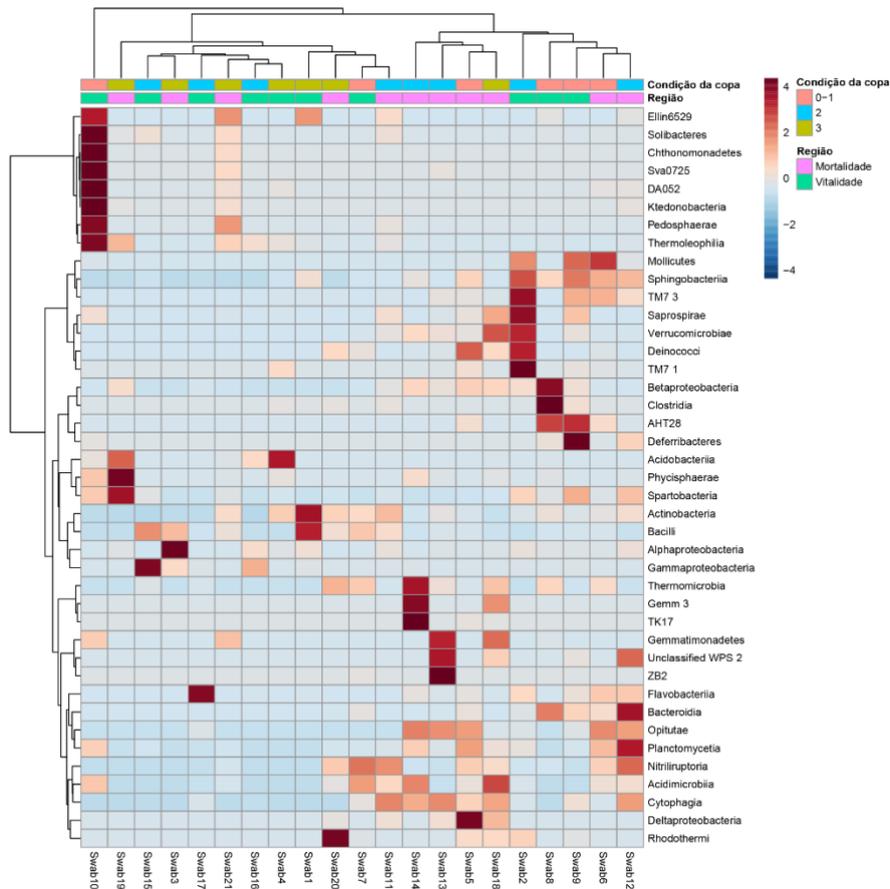


Figura 30. Heatmap do *clustering* hierárquico das amostras de bactérias a nível das classes, com base na medida de distância Euclideana e no algoritmo de *clustering* de *Ward*.

Geralmente, estes gráficos permitem a visualização de padrões de *taxa* de alta abundância, contra um fundo de *taxa* que é na sua maioria de baixa abundância ou ausente. Nestes, desvios inferiores em relação à média da população surgem a azul, enquanto desvios superiores à média aparecem a vermelho. Paralelamente, o dendrograma superior mostra os parâmetros ambientais que provocam as respostas mais análogas, e o dendrograma lateral permite perceber quais *taxa* se comportam de modo mais semelhante.

A aparente ausência de padrões de abundância e o arranjo do dendrograma superior de forma não concordante com os fatores ambientais, são então provocados pela heterogeneidade dentro da composição das amostras, o que não permite tirar ilações sobre a influência das condições ambientais recolhidas na abundância dos *taxa* a partir dos *heatmaps*.

De facto, recorrendo apenas aos dendrograma do agrupamento hierárquico das amostras (figura 31), desta vez agrupando-as em função do índice de *Bray-Curtis* (Bray e Curtis, 1957), é possível visualizar os esfregaços agrupados em *clusters*, que traduzem a sua proximidade numa matriz de distância, leia-se, em função da semelhança das suas comunidades.

4. Resultados e Discussão

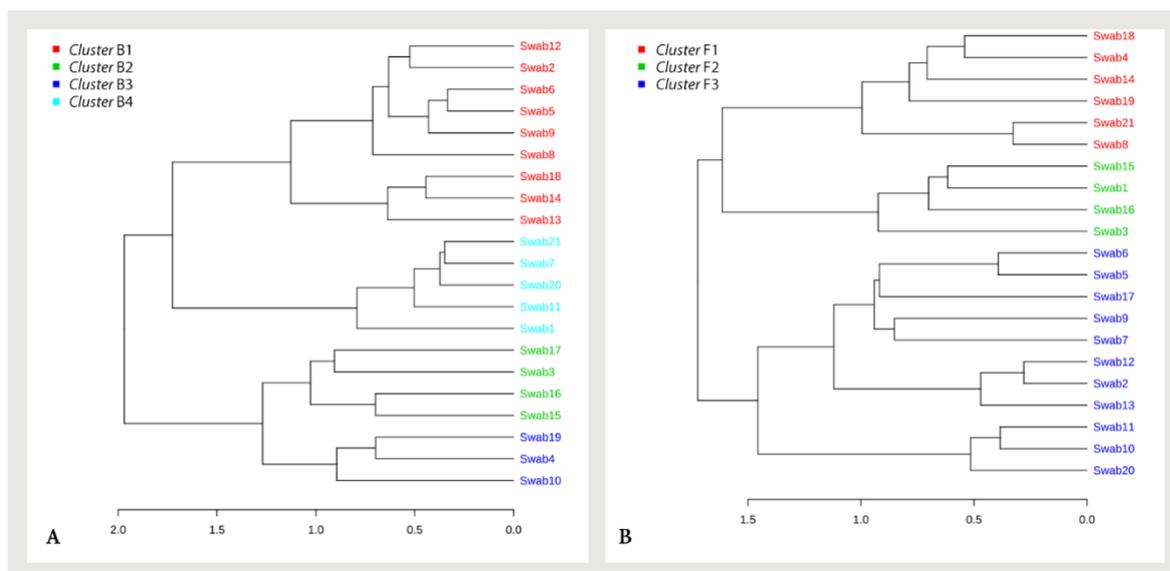


Figura 31. Dendrogramas das amostras de bactérias (A) e de fungos (B), gerado no *MicrobiomeAnalyst* a nível taxonômico da família, utilizando o índice de *Bray-Curtis* como medida de distância e o método de *Ward* como algoritmo de *clustering*.

Tal como apontado anteriormente, é possível agrupar as amostras em função da composição das suas comunidades. As comunidades de famílias de bactérias dos diferentes esfregaços, faz com seja possível identificar quatros grupos distintos (figura 31A). A nível da diversidade fúngica, são identificados três grupos de amostras (figura 31B). No entanto não existe sobreposição entre estes, e os fatores ambientais (tabela 12).

Tabela 12. Comparação entre os *clusters* de bactérias (B1-B4) e de fungos (F1-F3), e os dados ambientais recolhidos.

Cluster	ID	Região	Condição da Copa	Cluster	ID	Região	Condição da Copa
B1	Esfregaço 2	Vitalidade	2	F1	Esfregaço 4	Vitalidade	3
	Esfregaço 5	Mortalidade	0-1		Esfregaço 8	Vitalidade	0-1
	Esfregaço 6	Mortalidade	0-1		Esfregaço 14	Mortalidade	2
	Esfregaço 8	Vitalidade	0-1		Esfregaço 18	Mortalidade	3
	Esfregaço 9	Vitalidade	0-1		Esfregaço 19	Mortalidade	3
	Esfregaço 12	Mortalidade	2		Esfregaço 21	Mortalidade	3
	Esfregaço 13	Mortalidade	2		F2	Esfregaço 1	Vitalidade
	Esfregaço 14	Mortalidade	2	Esfregaço 3		Mortalidade	3
	Esfregaço 18	Mortalidade	3	Esfregaço 15		Vitalidade	2
				Esfregaço 16		Vitalidade	2
B2	Esfregaço 3	Mortalidade	3	F3	Esfregaço 2	Vitalidade	2
	Esfregaço 15	Vitalidade	2		Esfregaço 5	Mortalidade	0-1
	Esfregaço 16	Vitalidade	2		Esfregaço 6	Mortalidade	0-1
Esfregaço 17	Vitalidade	2	Esfregaço 7		Vitalidade	0-1	
B3	Esfregaço 4	Vitalidade	3		Esfregaço 9	Vitalidade	0-1
	Esfregaço 10	Vitalidade	0-1		Esfregaço 10	Vitalidade	0-1
	Esfregaço 19	Mortalidade	3		Esfregaço 11	Mortalidade	2
B4	Esfregaço 1	Vitalidade	3		Esfregaço 12	Mortalidade	2
	Esfregaço 7	Vitalidade	0-1		Esfregaço 13	Mortalidade	2
	Esfregaço 11	Mortalidade	2		Esfregaço 17	Vitalidade	2
	Esfregaço 20	Mortalidade	3	Esfregaço 20	Mortalidade	3	
	Esfregaço 21	Mortalidade	3				

4.7.5. Beta diversidade

Para avaliar as diferenças entre as comunidades microbianas das amostras, e dos grupos de amostras, comparando-as explicita e visualmente, enquanto lhe é atribuída em simultâneo a significância estatística correspondente, socorremo-nos dos métodos de ordenação. O nMDS foi realizado juntamente com o teste estatístico de análise de semelhanças (ANOSIM) (dados não exibidos) e o PCoA com os testes estatísticos PERMANOVA e análise de dispersões permutacionais (PERMDISP) (figuras 32, 33).

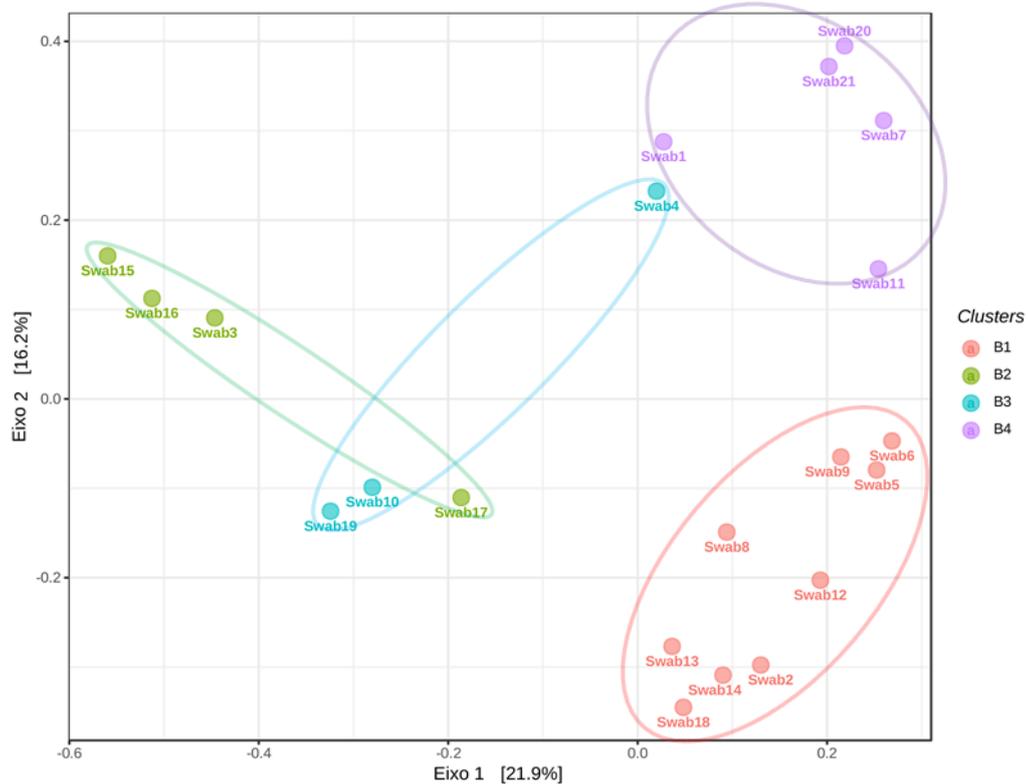


Figura 32. Análise de beta diversidade de bactérias, através do método de ordenação PCoA gerado no *MicrobiomeAnalyst* a nível taxonómico da família, utilizando o índice de *Bray-Curtis* como medida de distância. Agrupadas as amostras hierarquicamente mais semelhantes segundo o dendrograma da figura 31A. Soma dos três primeiros eixos = 49,9%.

PERMANOVA: $F\text{-value}=2,284$; $R^2=0,287$; $p\text{-value}<0,001$, PERMDISP: $F\text{-value}=2,570$; $p\text{-value}<0,088$;

B1:B2 PERMANOVA: $F\text{-value}=2,613$; $R^2=0,192$; $p\text{-value}<0,003$;

B1:B3 PERMANOVA: $F\text{-value}=2,201$; $R^2=0,180$; $p\text{-value}<0,013$;

B1:B4 PERMANOVA: $F\text{-value}=2,228$; $R^2=0,157$; $p\text{-value}<0,001$;

B2:B3 PERMANOVA: $F\text{-value}=1,336$; $R^2=0,211$; $p\text{-value}<0,048$;

B2:B4 PERMANOVA: $F\text{-value}=2,875$; $R^2=0,291$; $p\text{-value}<0,013$;

B3:B4 PERMANOVA: $F\text{-value}=2,299$; $R^2=0,277$; $p\text{-value}<0,02$.

O PCoA das comunidades bacterianas a nível taxonómico da família (figura 32) mostra o agrupamento dos *clusters* definidos anteriormente. Neste, as amostras são organizadas num referencial, onde a sua semelhança em termos de composição se reflete numa maior proximidade entre os pontos. Vemos que é possível separar os grupos em função dos *clusters*, havendo apenas sobreposição dos *clusters* B2 e B3 no PCoA a duas dimensões, situação que não se verifica no PCoA 3D (dados não exibidos). Cada

4. Resultados e Discussão

eixo tem uma porcentagem associada que significa a variação dos dados que este captura, sendo que a soma dos três primeiros eixos é igual a 49,9%, o que significa que este modelo captura esse exato valor da variação dos dados.

Os *p-values* das PERMANOVA são todos inferiores a 0,05, apontando para o significado estatístico dos dados, o que indica que todos os grupos são significativamente diferentes entre si. Por sua vez, os valores do PERMDISP indicam que as variâncias dos grupos são, também, significativamente diferentes.

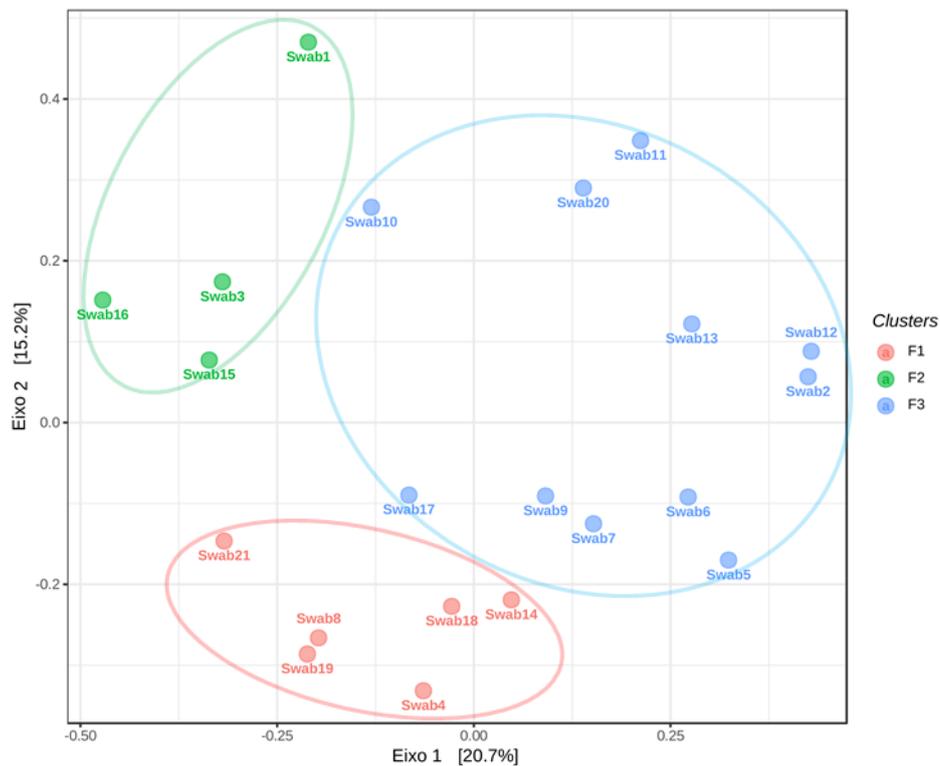


Figura 33. Análise de beta diversidade de fungos, através do método de ordenação PCoA gerado no *MicrobiomeAnalyst* a nível taxonômico da família, utilizando o índice de *Bray-Curtis* como medida de distância. Agrupadas as amostras hierarquicamente mais semelhantes segundo o dendrograma da figura 31A. Soma dos três primeiros eixos = 49,6%.

PERMANOVA: $F\text{-value}=2,560$; $R^2=0,221$ $p\text{-value}<0,001$. PERMDISP: $F\text{-value}=0,166$; $p\text{-value}<0,849$;

F1:F2 PERMANOVA: $F\text{-value}=2,284$; $R^2=0,222$; $p\text{-value}<0,004$;

F1:F3 PERMANOVA: $F\text{-value}=2,975$; $R^2=0,166$; $p\text{-value}<0,001$;

F2:F3 PERMANOVA: $F\text{-value}=2,292$; $R^2=0,150$; $p\text{-value}<0,011$.

Os PCoA das famílias de fungos presentes nos diferentes *clusters* (figura 33), também definidos anteriormente, apontam para conclusões idênticas. Neste modelo, que explica 49,6% da variação dos dados, os *clusters* estão bem definidos a duas dimensões, todos os grupos são significativamente diferentes entre si (PERMANOVA significativa), mas não as suas variâncias (PERMDISP não significativo).

4.7.6. Alfa diversidade dos grupos

Refazendo a análise da alfa diversidade, desta vez agrupando as amostras por *clusters*, podemos explorar estas diferenças da composição das comunidades apontadas pelas análises anteriores. Os *clusters* foram comparados em relação ao número de famílias observadas, ao número de famílias esperado (estimado pelo índice de *Chao1*), à dominância de famílias (calculado a partir do índice de *Simpson*) e à equitatividade entre famílias (obtido com o índice de *Shannon*), quer em termos de bactérias (figura 34), quer de fungos (figura 35). A comparação destas métricas foi acompanhada do teste estatístico de *Kruskal-Wallis*, um equivalente não-paramétrico da ANOVA, que permite avaliar se as diferenças sobre as médias de populações são ou não significativas.

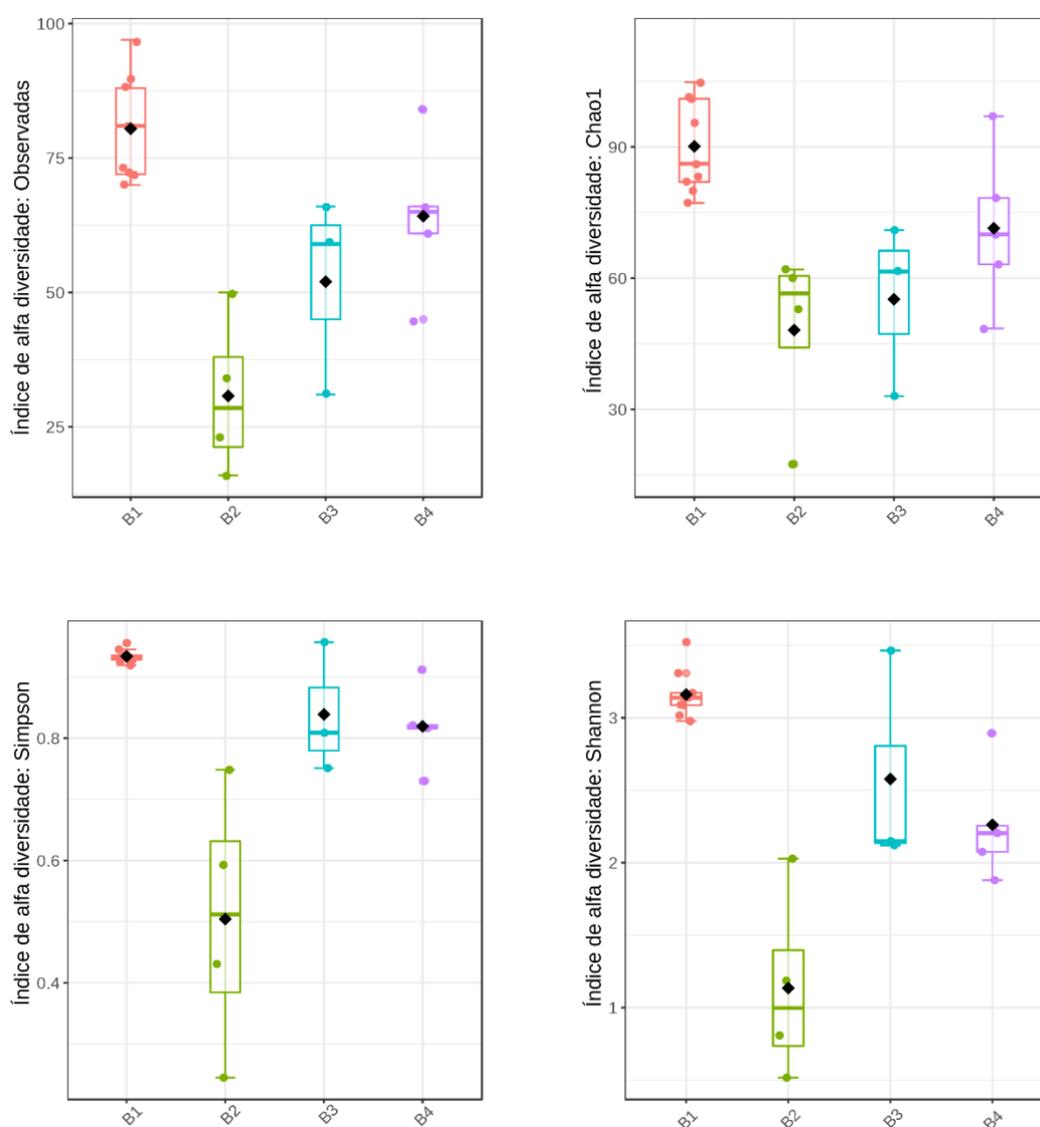


Figura 34. Boxplots das diferentes métricas de alfa diversidade entre os *clusters* de bactérias, a nível da família.

Observadas: $p\text{-value} < 0,002$; $Kruskal\text{-Wallis} = 14,376$; Chao1: $p\text{-value} < 0,004$; $Kruskal\text{-Wallis} = 13,187$;

Shanon: $p\text{-value} < 0,002$; $Kruskal\text{-Wallis} = 14,382$; Simpson: $p\text{-value} < 0,003$; $Kruskal\text{-Wallis} = 13,802$.

4. Resultados e Discussão

Apesar dos valores significativos obtidos na análise anterior, em relação ao agrupamento das amostras em *clusters*, a presença de alguns pontos fora dos *boxplots*, indica que estes são outliers em relação ao *cluster* que pertencem. Isto é, essas amostras são diferentes do restante conjuntos de amostras do *cluster* em que se incluem, em relação ao que o respetivo índice de alfa diversidade representa.

Os *p-values* do teste de *Kruskal-Wallis* apontam para valores significativos para todas as métricas, indicando que os *clusters* são significativamente diferentes entre si em termos do número de famílias de bactérias observado e esperado, e em relação à dominância (valores reduzidos do índice de *Simpson* indicam a presença de famílias dominantes) e à equitatividade (valores elevados do índice de *Shannon* indicam famílias em iguais proporções) das famílias bacterianas dentro das comunidades de cada *cluster* (figura 34).

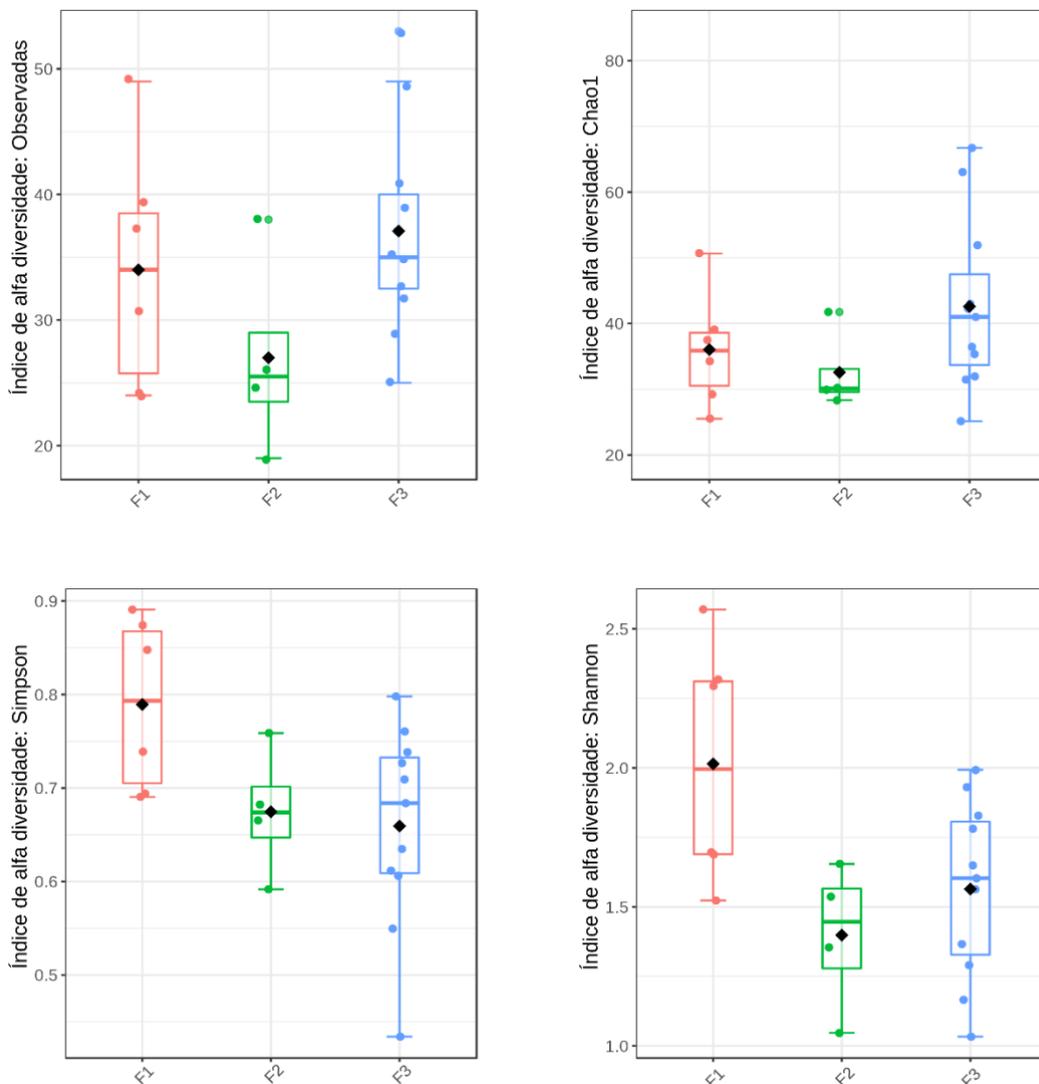


Figura 35. *Boxplots* das diferentes métricas de alfa diversidade entre os *clusters* de fungos a nível da família.

Observadas: *p-value*<0,213; *Kruskal-Wallis*=3,095; *Chao1*: *p-value*<0,225; *Kruskal-Wallis* =2,987;
Shanon: *p-value*<0,062; *Kruskal-Wallis* =5,556; *Simpson*: *p-value*<0,064; *Kruskal-Wallis* =5,490.

Em relação a fungos (figura 35), os *p-values* do teste de *Kruskal-Wallis* indicam que não existem diferenças significativas em relação ao número observado, número esperado, dominância e equitatividade das famílias de fungos dos diferentes *clusters*.

4.7.7. Diversidade taxonômicas dos grupos

Uma vez mais, a informação destas métricas pode ser traduzida à escala visual utilizando os gráficos das abundâncias relativas das famílias de bactérias (figura 36) e de fungos (figura 37), restringindo a dimensão dos dados à escala dos *clusters* identificados.

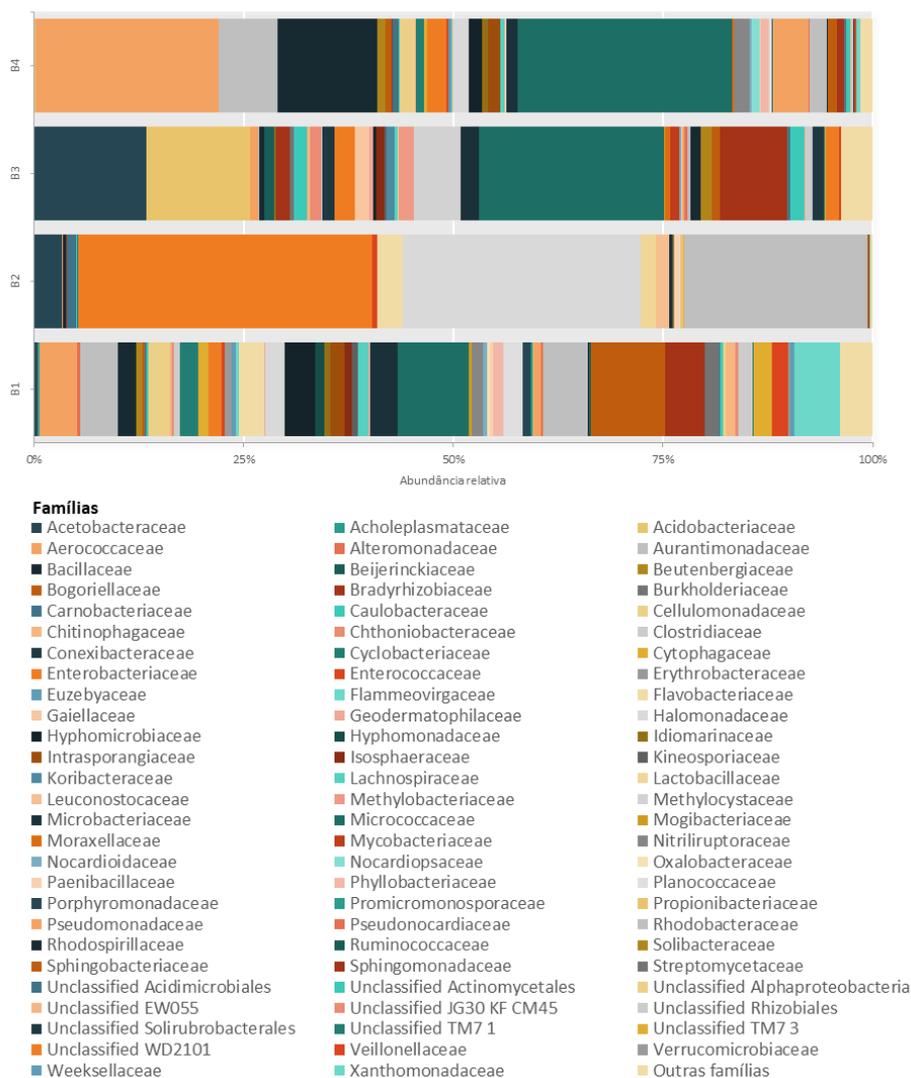


Figura 36. Gráfico de barras empilhadas da abundância relativa dos clusters de bactérias a nível das famílias, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em "Outros famílias".

Na figura 36 podemos observar a elevada equitatividade de famílias de bactérias do *cluster* B1, sendo as famílias mais abundantes as *Sphingobacteriaceae* (8,77%), as *Micrococcaceae* (8,53%), as *Rhodobacteraceae* (5,43%) e as *Xanthomonadaceae* (5,43%). Em sentido oposto, no *cluster* B2 existe uma elevada dominância de três famílias, as *Enterobacteriaceae* (35,05%), as *Halomonadaceae* (28,31%) e as *Rhodobacteraceae*

4. Resultados e Discussão

(21,86%). Os *clusters* B3 e B4, ao contrário dos anteriores, não são marcados nem por pela dominância, nem pela equitatividade das suas comunidades, estando num patamar intermédio entre estas duas características. No primeiro dominam as bactérias das famílias Micrococcaceae (21,98%), Acetobacteraceae (13,34%) e Acidobacteriaceae (12,46%), e no último, as Micrococcaceae (25,60%), as Aerococcaceae (21,94%) e as Bacillaceae (11,88%).

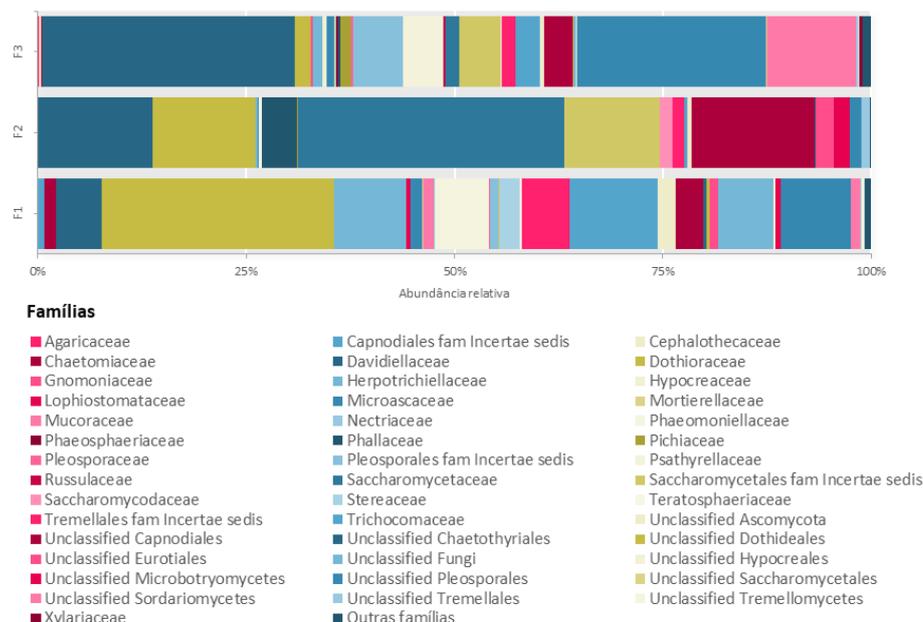


Figura 37. Gráfico de barras empilhadas da abundância relativa dos clusters de fungos a nível das famílias, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em “Outros famílias”.

Tal como apontado anteriormente, não existem diferenças significativas entre os diferentes *clusters* em termos de dominância e equitatividade das famílias de fungos (figura 37), sendo que estes apenas se distinguem pela sua composição. As famílias mais abundantes no *cluster* F1 são as Dothioraceae (27,95%) e as Trichocomaceae (10,56%), no *cluster* F2 são as Saccharomycetaceae (31,90%), Unclassified Capnodiales (14,83%), as Davidiellaceae (13,78%), as Dothioraceae (12,39%) e as Saccharomycetales fam Incertae sedis (11,44%), por fim, o *cluster* F3 é dominado por Davidiellaceae (30,35%), Unclassified Pleosporales (22,63%) e Unclassified Sordariomycetes (10,69%).

A maioria destas famílias de bactérias e de fungos foi referenciada anteriormente por ter uma abundância relativa elevada no microbioma total, as restantes revelam-se agora influentes dentro do *cluster* que representam.

Assim, no *cluster* B1, as Sphingobacteriaceae, que têm diversas aplicações industriais conhecidas, costumam ser isoladas a partir de vários tipos de solos e compostos, mas podem também usar o ser humano como hospedeiro, provocando infeções e doenças (Lambiase, 2014). Ainda no mesmo *cluster*, as bactérias da família

Xanthomonadaceae incluem bactérias associadas a plantas, muitas das quais causam doenças graves em culturas, arbustos e árvores, incluindo o sobreiro, através da oclusão dos vasos do xilema que vai levar ao stresse hídrico e murchidão do hospedeiro (Mhedbi-hajri et al., 2011).

Presentes no *cluster* B3, as Acetobacteraceae estão mundialmente associadas à produção de vinagre, estando amplamente distribuídas em ambientes alcoólicos e ácidos, com alguns géneros a seres isolados a partir de flores, mosquitos e outros insetos (Komagata et al., 2014). Também no *cluster* B3, as Acidobacteriaceae são extremamente abundantes e diversificadas, têm a capacidade de degradar carbonos simples bem como polissacáridos vegetais, incluindo celulose, e costumam ser encontradas numa vasta gama de ambientes acídicos (Campbell, 2014).

Finalmente, presentes no último *cluster*, as Bacillaceae são amplamente distribuídas em ambientes naturais, e os seus habitats são descritos como tão variados quanto os nichos de onde já foram recolhidas amostras, isto graças à sua característica mais marcante, a capacidade de formar endósporos de alta resistência que lhes permitem sobreviver a condições adversas por longos períodos de tempo (Mandic-Mulec et al., 2015)

4.8. Influência das variáveis ambientais no microbioma

Apesar de termos identificado que as variáveis ambientais não correspondem ao agrupamento das amostras que melhor reflete a sua proximidade, as variáveis ambientais também incutem diversidade nos microbiomas.

Fizemos diversos ensaios para testar diferentes agrupamentos de amostras e procurar a influência da região de onde as amostras foram recolhidas (percentagem de declínio do montado) e a condição da copa da árvore de onde foi feita a recolha de amostras. Não foram obtidas diferenças estatisticamente significativas entre as amostras agrupadas em função da condição da folhagem da árvore de onde foram feitos os esfregaços, o que aponta para que as diferenças observadas entre os grupos não estão dependentes desse fenótipo medido.

Já em relação ao fator ambiental região, a situação é particular. Removendo da análise as amostras recolhidas no Baixo Sorraia em 2017 (esfregaços 1, 2, 3 e 4), e assim excluindo a variabilidade introduzidas pela recolha de amostras em anos e ZIFs diferentes (Costa et al., 2010), foi possível obter resultados significativos. No entanto, não foi encontrado um registo irrevogável para que este pudesse ser considerado um fator influenciador, pois mesmo que os dados apontem para que este desempenhe algum papel no microbioma destas lesões, os testes estatísticos são *borderline*.

4. Resultados e Discussão

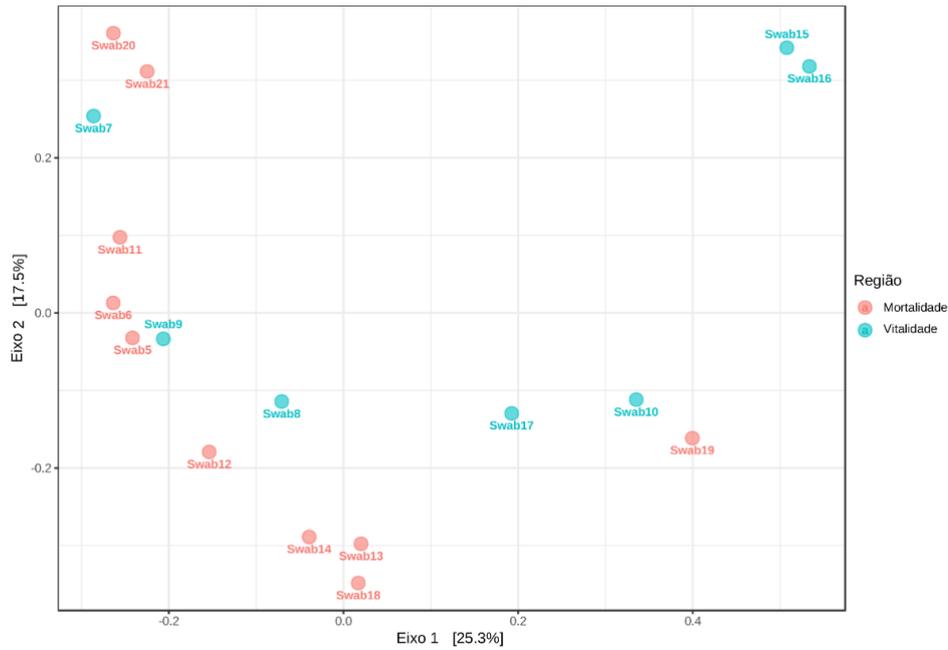


Figura 40. Análise de beta diversidade de bactérias da região do Divor, através do método de ordenação PCoA gerado no *MicrobiomeAnalyst* a nível taxonómico da família, utilizando o índice de *Bray-Curtis* como medida de distância. Soma dos três primeiros eixos = 55,2%. PERMANOVA: $F\text{-value}=1,499$; $R^2=0,091$; $p\text{-value}<0,045$.

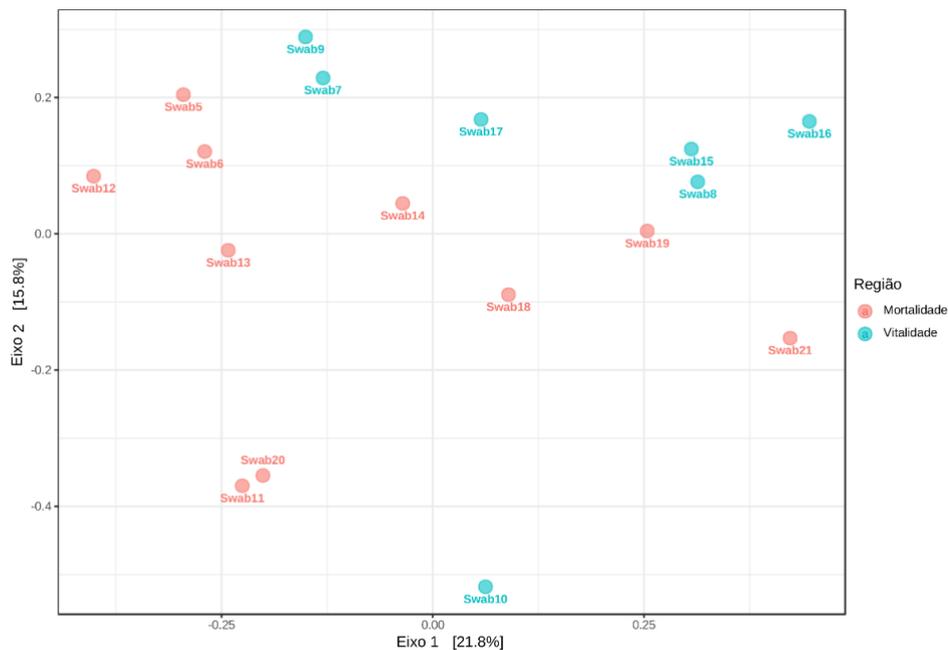


Figura 41. Análise de beta diversidade de fungos da região do Divor, através do método de ordenação PCoA gerado no *MicrobiomeAnalyst* a nível taxonómico da família, utilizando o índice de *Bray-Curtis* como medida de distância. Soma dos três primeiros eixos = 49,8%. PERMANOVA: $F\text{-value}=2,136$; $R^2=0,125$; $p\text{-value}<0,013$.

Isto reflete-se no $p\text{-value}$ da PERMANOVA obtida para o PCoA das bactérias recolhidos nas diferentes regiões da ZIF Divor a nível da família (figura 40), valor muito próximo do limite de 0,05 definido para aceitar a hipótese de que os grupos definidos são significativamente distintos. Para fungos, este valor é significativo, indicando que as

comunidades fúngicas da região de vitalidade são significativamente diferentes das da região de mortalidade (figura 41).

Isto reflete-se no *p-value* da PERMANOVA obtida para o PCoA das bactérias recolhidos nas diferentes regiões da ZIF Divor a nível da família (figura 40), valor muito próximo do limite de 0,05 definido para aceitar a hipótese de que os grupos definidos são significativamente distintos. Para fungos, este valor é significativo, indicando que as comunidades fúngicas da região de vitalidade são significativamente diferentes das da região de mortalidade (figura 41).

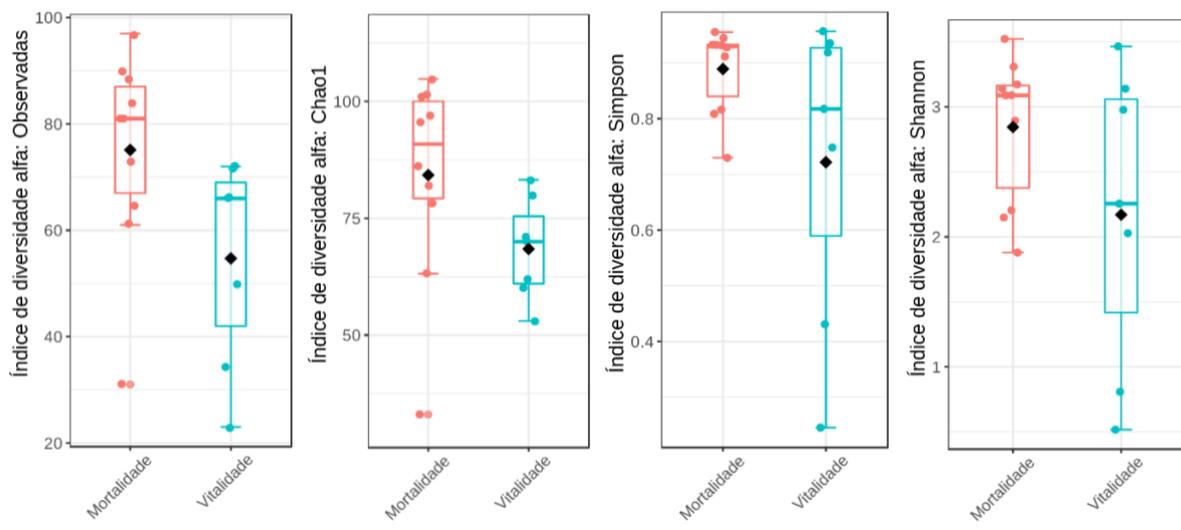


Figura 42. Métricas de alfa diversidade entre as bactérias das duas regiões, a nível da família. Observadas: *p-value*<0,045; Mann-Whitney=56; Chao1: *p-value*<0,043; Mann-Whitney=56; Shanon: *p-value*<0,315; Mann-Whitney=46; Simpson: *p-value*<0,417; Mann-Whitney=44.

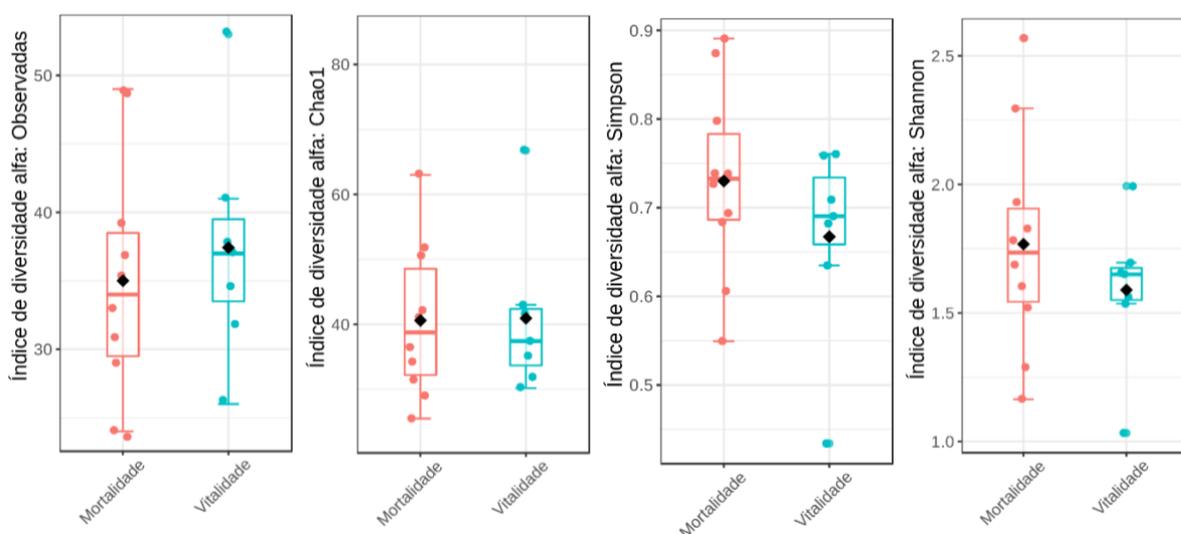


Figura 43. Métricas de alfa diversidade entre os fungos das duas regiões, a nível da família. Observadas: *p-value*<0,463; Mann-Whitney=27; Chao1: *p-value*<0,887; Mann-Whitney=33; Shanon: *p-value*<0,4747; Mann-Whitney=43; Simpson: *p-value*<0,364; Mann-Whitney=45.

4. Resultados e Discussão

Aplicando as métricas de alfa diversidade a estes dados (figuras 42, 43), concluímos que as amostras da região de mortalidade possuem um maior número de famílias de bactérias observadas e esperadas do que as da região de vitalidade, mas esta diferença é uma vez mais *borderline*, e que, em relação à dominância e equitatividade das famílias, não existem diferenças significativas entre as duas regiões (figura 42).

Por sua vez, as comunidades de fungos não apresentam diferenças significativas em relação a nenhum dos índices calculados (figura 43), mas têm tendencialmente menos famílias observadas e esperadas, comparativamente às comunidades de bactérias.



Figura 44. Gráfico de barras empilhadas da abundância relativa de bactérias agrupadas por região a nível das famílias, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em “Outros filós”.

Na região de vitalidade, as bactérias mais abundantes (figura 44) pertencem às famílias Halomonadaceae (17,38%), Enterobacteriaceae (15,89%), Micrococcaceae (8,11%) e Aerococcaceae (7,57%). Por sua vez, na região de mortalidade predominam as Micrococcaceae (12,18%), as Aerococcaceae (5,45%), as Aurantimonadaceae (5,45%), as Bacillaceae (5,25%) e as Sphingomonadaceae (5,03%).

Em relação a fungos (figura 45), a zona de vitalidade é maioritariamente constituída pelas famílias de fungos Dothioraceae (20,37%), Saccharomycetaceae (14,88%), Davidiellaceae (13,12%), Psathyrellaceae (5,54%) e pelas ordens Capnodiales (11,35%) e Saccharomycetales fam Incertae sedis (5,26%). Na zona de mortalidade os fungos mais abundantes são das famílias Davidiellaceae (23,33%), Trichocomaceae (7,31%), Dothioraceae (6,99%), Herpotrichiellaceae (5,98%), da ordem Pleosporales (21,21%) e da classe Sordariomycetes (10,76%).

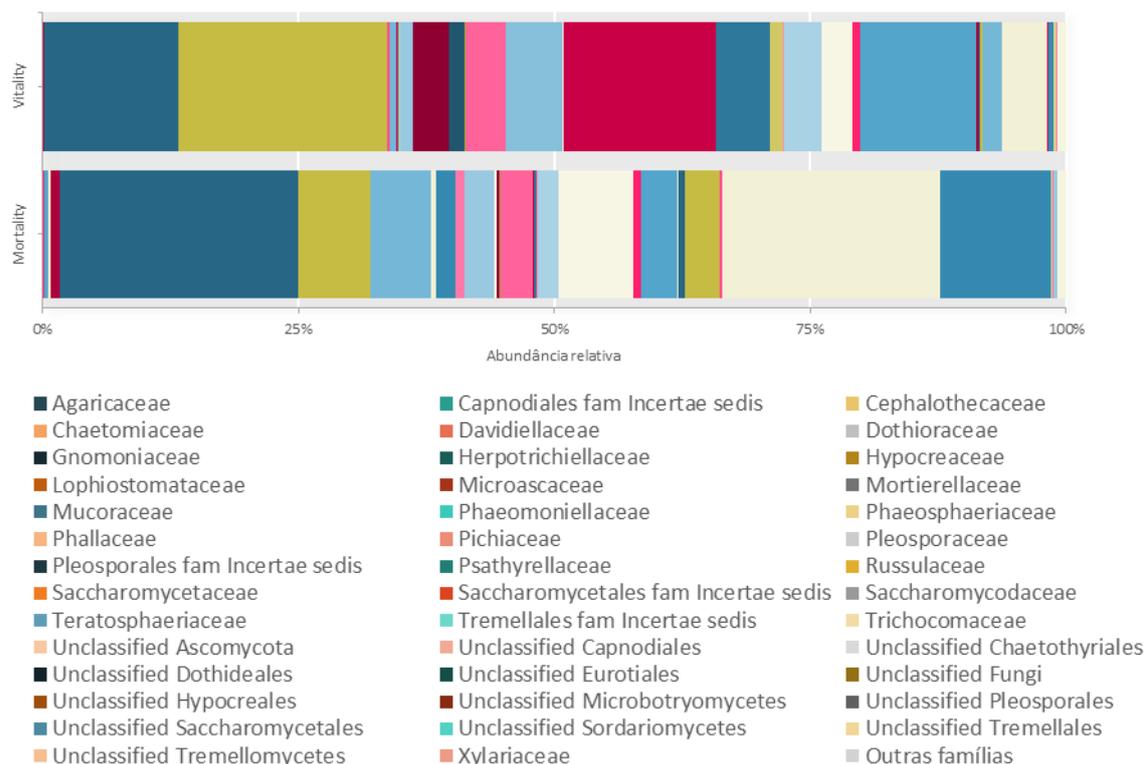


Figura 45. Gráfico de barras empilhadas da abundância relativa de fungos agrupadas por região a nível das famílias, obtido no *MicrobiomeAnalyst*. Taxa com número total de *reads* inferior a 200 são agrupados em "Outros filós".

Algumas destas famílias, de bactérias e de fungos, de maior abundância repetem-se nas duas regiões. De lembrar que também a região de vitalidade corresponde a um montado com algum declínio registado, ainda que com menor intensidade em relação à região contraposta (0-1% de árvores em declínio, contra 9-22%).

À exceção das bactérias Aurantimonadaceae e Sphingomonadaceae, e dos fungos Psathyrellaceae e Herpotrichiellaceae, os restantes foram detetados ou pela sua prevalência nas amostras, ou pela elevada abundância no microbioma.

A família Aurantimonadaceae foi recentemente descrita, sendo composta por dois géneros que foram isolados em ambientes marinhos, um deles foi isolado em corais doentes, e do segundo, não há registo da sua presença em ambientes terrestres (Jurado et al., 2006).

4. Resultados e Discussão

A maioria das Sphingomonadaceae encontradas não teve classificação atribuída para além da família, mas foi detetada a presença de uma elevada porção de *Sphingomonas*, particularmente *Sphingomonas wittichii*. Estes organismos, que demonstram capacidades biotecnológicas únicas, estão amplamente distribuídos e são responsáveis pela indução de corrosão, por provocar doenças em plantas e por principalmente degradar matéria orgânica (White et al., 1996).

Os Psathyrellaceae são praticamente representados na totalidade por fungos do género *Coprinopsis*. Estes cogumelos são normalmente usados como modelo experimental para o estudo do desenvolvimento multicelular de fungos, sendo espécies frequentes geograficamente bem distribuídos (Kilaru et al., 2006).

Por fim, fungos Herpotrichiellaceae encontrados, são na sua maioria do género *Exophiala*. Este género de leveduras negras possui uma variedade de espécies clinicamente relevantes, regularmente indicados como os agentes causadores de micoses em humanos (Hoog et al., 2003).

5. Conclusão

O estudo exploratório aqui apresentado, teve o objetivo de compreender se o microbioma detetado nas lesões encontradas na casca de sobreiros é característico do declínio agudo presente em carvalhos.

De uma forma geral, as comunidades microbianas enquadram-se com o que tem sido descrito na literatura, sendo semelhantes àquelas encontradas em outras árvores. Isto é assegurado pela presença de bactérias dos filos Actinobacteria, Firmicutes e Bacteroidetes, e dos fungos do filo Actinobacteria.

Se a presença de algumas das famílias de bactérias e sobretudo de fungos encontradas é esperada, dado serem capazes de colonizar um diverso leque de ambientes ou estarem normalmente associadas a plantas, a ocorrência de algumas destas é, no mínimo, inusitada. Com destaque para as famílias de bactérias Micrococcaceae, Aerococcaceae, Rhodobacteraceae, e os fungos da família Davidiellaceae, que são normalmente microrganismos associadas a habitats aquáticos hipersalinos.

As bactérias Xanthomonadaceae merecem também algum destaque, dado a sua relação com doenças provocadas por stresse hídrico (fator abiótico que surge muitas vezes associado ao declínio de carvalhos) e, também, as Acidobacteraceae, pela sua capacidade de degradar celulose (podem estar de algum modo relacionados com a formação das lesões). No entanto, a presença de Enterobacteriaceae é mesmo o maior indicador do declínio agudo do carvalho. Está documentado que espécies desta família são os principais patógenos responsáveis pela ocorrência destes sintomas.

Os resultados demonstram comunidades diversas entre si, não agrupadas pelos fatores ambientais identificados, pois nem o grau de desfolha das árvores associado ao declínio crónico, nem o local de amostragem em zonas de maior vitalidade ou mortalidade, se revelaram responsáveis pelo agrupamento dos dados recolhidos.

As amostras agrupam-se sim, por semelhança de comunidades, mas de acordo com fatores ainda não identificados. Especula-se que a natureza multifatorial deste tipo de condições possa resultar em comunidades bastante distintas ao longo de todo o episódio de desenvolvimento da doença, convergindo numa composição diversa e variada do microbioma, provavelmente também ele afetado por fatores bióticos e abióticos.

Ainda assim, esta investigação possibilitou a otimização de protocolos para caracterização do microbioma bacteriano e fúngico destas lesões no sobreiro, culminando na sua descrição detalhada e inédita. Futuramente, crê-se que possa motivar novos estudos para entender os fatores microbiológicos associados ao estado de saúde da floresta de sobreiro.

Bibliografia

- APCOR (s.d. a) [infográfico] <http://www.apcor.pt/media-center/estatisticas/> [25 de janeiro de 2019]
- APCOR (s.d. b) [fotografia] <https://www.apcor.pt/a-chegada-dos-tiradores-de-cortica-ii/> [25 de janeiro de 2019]
- APFC (2016a) 1.^a reunião, consulta pública e 2.^a reunião do PGF da ZIF do Baixo Sorraia [imagem] http://www2.icnf.pt/portal/florestas/gf/zif/resource/doc/zif/publicit/plano-gestao/zif_241_10_baixosorraia_pgf [4 de dezembro de 2018]
- APFC (2016b) 1.^a reunião, consulta pública e 2.^a reunião do PGF da ZIF do Divor [imagem] http://www2.icnf.pt/portal/florestas/gf/zif/resource/doc/zif/publicit/plano-gestao/zif_240_10_divor_pgf [4 de dezembro de 2018]
- Benito, Á., Calderón, F., Palomero, F., e Benito, S. (2015). Combine Use of Selected *Schizosaccharomyces pombe* and *Lachancea thermotolerans* Yeast Strains as an Alternative to the Traditional Malolactic Fermentation in Red Wine Production. *Molecules* 20, 9510–9523.
- Bensch, K., Braun, U., Groenewald, J.Z., e Crous, P.W. (2012). The genus *Cladosporium*. *Stud. Mycol.* 72, 1–401.
- Brady, C., Denman, S., Kirk, S., Venter, S., Rodríguez-Palenzuela, P., e Coutinho, T. (2010). Description of *Gibbsiella quercinecans* gen. nov., sp. nov., associated with Acute Oak Decline. *Syst. Appl. Microbiol.* 33, 444–450.
- Brady, C., Hunter, G., Kirk, S., Arnold, D., e Denman, S. (2014a). Description of *Brenneria roseae* sp. nov. and two subspecies, *Brenneria roseae* subspecies *roseae* ssp. nov. and *Brenneria roseae* subspecies *americana* ssp. nov. isolated from symptomatic oak. *Syst. Appl. Microbiol.* 37, 396–401.
- Brady, C., Hunter, G., Kirk, S., Arnold, D., e Denman, S. (2014b). *Rahnella victoriana* sp. nov., *Rahnella bruchi* sp. nov., *Rahnella woolbedingensis* sp. nov., classification of *Rahnella genomospecies* 2 and 3 as *Rahnella variigena* sp. nov. and *Rahnella inusitata* sp. nov., respectively and emended description of the genus R. *Syst. Appl. Microbiol.* 37, 545–552.
- Brady, C., Hunter, G., Kirk, S., Arnold, D., e Denman, S. (2014c). *Gibbsiella greigii* sp. nov., a novel species associated with oak decline in the USA. *Syst. Appl. Microbiol.* 37, 417–422.
- Brady, C.L., Cleenwerck, I., Denman, S., Venter, S.N., Rodríguez-Palenzuela, P., Coutinho, T.A., e De Vos, P. (2012). Proposal to reclassify *Brenneria quercina* (Hildebrand and

- Schroth 1967) Hauben et al. 1999 into a new genus, *Lonsdalea* gen. nov., as *Lonsdalea quercina* comb. nov., descriptions of *Lonsdalea quercina* subsp. *quercina* comb. nov., *Lonsdalea quercina* subsp. *ib.* Int. J. Syst. Evol. Microbiol. 62, 1592–1602.
- Bray, J.R., e Curtis, J.T. (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecol. Monogr. 27, 325–349.
- Broberg, M., Doonan, J., Mundt, F., Denman, S., e McDonald, J.E. (2018). Integrated multi-omic analysis of host-microbiota interactions in acute oak decline. Microbiome 6, 21.
- Camilo-Alves, C.S.P., Vaz, M., Da Clara, M.I.E., e Ribeiro, N.M.D.A. (2017). Chronic cork oak decline and water status: new insights. New For. 48, 753–772.
- Campbell, B.J. (2014). The family acidobacteriaceae. In The Prokaryotes: Other Major Lineages of Bacteria and The Archaea, (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 405–415.
- Caporaso, J.G., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Mcdonald, D., Muegge, B.D., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.
- Chander, A.M., Nair, R.G., Kaur, G., Kochhar, R., Dhawan, D.K., Bhadada, S.K., e Mayilraj, S. (2017). Genome Insight and Comparative Pathogenomic Analysis of *Nesterenkonia jeotgali* Strain CD08_7 Isolated from Duodenal Mucosa of Celiac Disease Patient. Front. Microbiol. 8, 1–12.
- Chao, A. (1984). Nonparametric Estimation of the Scandinavian Journal of Statistics. Scand. J. Stat. 11, 265–270.
- Chao, A., e Chiu, C.-H. (2016). Species Richness: Estimation and Comparison. Wiley StatsRef Stat. Ref. Online 1–26.
- Costa, A., Pereira, H., e Madeira, M. (2010). Analysis of spatial patterns of oak decline in cork oak woodlands in Mediterranean conditions. Ann. For. Sci. 67, 204–204.
- D’agostino, R.B., Belanger, A., e Ralph B., D.J. (2012). A Suggestion for Using Powerful and Informative Tests of Normality. Am. Stat. 44:4, 316–321.
- Denman, S., e Webber, J. (2009). Oak declines: new definitions and new episodes in Britain. Q. J. For. 103, 285–290.
- Denman, S., Kirk, S., e Webber, J. (2010). Forestry Commission Practice Note - Managing acute oak decline.
- Denman, S., Brady, C., Kirk, S., Cleenwerck, I., Venter, S., Coutinho, T., e De Vos, P. (2012). *Brenneria goodwinii* sp. nov., associated with acute oak decline in the UK. Int. J. Syst. Evol. Microbiol. 62, 2451–2456.
- Denman, S., Brown, N., Kirk, S., Jeger, M., e Webber, J. (2014). A description of the symptoms of Acute Oak Decline in Britain and a comparative review on causes of

- similar disorders on oak in Europe. *Forestry* 87, 535–551.
- Denman, S., Plummer, S., Kirk, S., Peace, A., e McDonald, J.E. (2016). Isolation studies reveal a shift in the cultivable microbiome of oak affected with Acute Oak Decline. *Syst. Appl. Microbiol.* 39, 484–490.
- Denman, S., Doonan, J., Ransom-Jones, E., Broberg, M., Plummer, S., Kirk, S., Scarlett, K., Griffiths, A.R., Kaczmarek, M., Forster, J., et al. (2018). Microbiome and infectivity studies reveal complex polyspecies tree disease in Acute Oak Decline. *ISME J.* 12, 386–399.
- Dhariwal, A. (2017). *Statistical, Visual and Functional Analysis of Microbiome Data* (Dissertação de Mestrado) McGill University, Montreal.
- Dhariwal, A., Chong, J., Habib, S., King, I.L., Agellon, L.B., e Xia, J. (2017). MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res.* 45, W180–W188.
- Dias, Carlos (2017, 7 de junho). Abate ilegal de sobreiros e azinheiras para plantar eucaliptos repete-se em Tomar [Versão electrónica]. Público. Disponível em <https://www.publico.pt/2017/06/07/local/noticia/abate-ilegal-de-sobreiros-e-azinheiras-para-plantar-eucaliptos-repetese-em-tomar-1774887>
- Dias, Carlos (2018, 5 de julho). Câmara ordena o corte de quase uma centena de sobreiros para “salvar vidas” [Versão electrónica]. Público. Disponível em <https://www.publico.pt/2018/07/05/local/noticia/camara-de-setubal-ordena-o-corte-de-quase-uma-centena-de-sobreiros-para-salvar-vidas-1836729>
- Dias, Carlos (2019, 8 de março). População declara guerra a dois sobreiros que "sujam" as campas no cemitério [Versão electrónica]. Público. Disponível em <https://www.publico.pt/2019/03/08/local/noticia/sobreiros-nao-incomodam-mortos-irritam-vivos-cortados-1864329>
- Doonan, J., Denman, S., Gertler, C., Pachebat, J.A., Golyshin, P.N., e McDonald, J.E. (2015). The intergenic transcribed spacer region 1 as a molecular marker for identification and discrimination of Enterobacteriaceae associated with acute oak decline. *J. Appl. Microbiol.* 118, 193–201.
- Doonan, J., Denman, S., Pachebat, J.A., e McDonald, J.E. (2019). Genomic analysis of bacteria in the Acute Oak Decline pathobiome. *Microb. Genomics* 5.
- Elentir (2011) Monte De La Peneda [fotografia] www.flickr.com/photos/elentir/5850309915/ [22 de julho de 2018]
- Elifantz, H., Horn, G., Ayon, M., Cohen, Y., e Minz, D. (2013). Rhodobacteraceae are the key members of the microbial community of the initial biofilm formed in Eastern Mediterranean coastal seawater. *FEMS Microbiol. Ecol.* 85, 348–357.
- Fadrosh, D.W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R.M., e Ravel, J.

- (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* 2, 6.
- FAO (2010) [fotografia] <https://www.amorim.com/a-cortica/localizacao-do-montado/> [25 de janeiro de 2019]
- Fell, J.W., Statzell-Tallman, A., e Kurtzman, C.P. (2004). *Lachancea meyersii* sp. nov., an ascosporegenous yeast from mangrove regions in the Bahama Islands. *Stud. Mycol.* 50, 359–363.
- Fraiman, R., e Muniz, G. (2001). Trimmed means for functional data. *Test* 10, 419–440.
- Freel, K.C., Charron, G., Leducq, J.B., Landry, C.R., e Schacherer, J. (2015). *Lachancea quebecensis* sp. Nov., a yeast species consistently isolated from tree bark in the Canadian province of Québec. *Int. J. Syst. Evol. Microbiol.* 65, 3392–3399.
- Freire, Emília (2018, 10 de julho). Uma frente nacional em defesa do Sobreiro e da Cortiça [Versão electrónica]. *Vida Rural*. Disponível em <https://www.vidarural.pt/insights/uma-frente-nacional-em-defesa-do-sobreiro-e-da-cortica/>
- Gobbi, M., Comitini, F., Domizio, P., Romani, C., Lencioni, L., Mannazzu, I., e Ciani, M. (2013). *Lachancea thermotolerans* and *Saccharomyces cerevisiae* in simultaneous and sequential co-fermentation: A strategy to enhance acidity and improve the overall quality of wine. *Food Microbiol.* 33, 271–281.
- Goodrich, J.K., Di Rienzi, S.C., Poole, A.C., Koren, O., Walters, W.A., Caporaso, J.G., Knight, R., e Ley, R.E. (2014). Conducting a Microbiome Study. *Cell* 158, 250–262.
- Gotelli, N.J., e Chao, A. (2013). *Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data* (Elsevier Ltd.).
- Gotelli, N.J. e Colwell, R.K. (2011) *Estimating Species Richness. Biological Diversity: Frontiers in Measurement and Assessment*, Oxford University Press, United Kingdom, 39-54.
- Gottschalk, K.W., e Wargo, P.M. (1996). Oak decline around the world. *USDA Interag. Gypsy Moth Res. Forum* 3–13.
- Guiomar, N., Palheiro, P., Loureiro, C., Salgueiro, A., Tomé, J., e Fernandes, J.P.A. (2011). *Manual De Boas Práticas De Gestão Dos Espaços Florestais Na Bacia Drenante Da Albufeira De Castelo De Bode*.
- Harmon, P., e Harmon, C. (2014). Sudden Oak Death and Ramorum Blight.
- Herlemann, D.P.R., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., e Andersson, A.F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* 5, 1571–1579.
- Herrera, A., e Cockell, C.S. (2007). Exploring microbial diversity in volcanic environments: A review of methods in DNA extraction. *J. Microbiol. Methods* 70, 1–12.

- Hoog, G.S. De, Vicente, V., Caligiorne, R.B., Kantarcioglu, S., Tintelnot, K., e Haase, G. (2003). Species Diversity and Polymorphism in the *Exophiala spinifera* Clade Containing Opportunistic Black Yeast-Like Fungi. *J. Clin. Microbiol.* *41*, 4767–4778.
- ICNF (2013). Áreas dos usos do solo e das espécies florestais de Portugal continental em 1995, 2005 e 2010.
- Illumina (2017). An introduction to Next-Generation Sequencing Technology. 1–16.
- Illumina (2018). Illumina Methods Guide.
- Ishikawa, M. (2003). *Marinilactibacillus psychrotolerans* gen. nov., sp. nov., a halophilic and alkaliphilic marine lactic acid bacterium isolated from marine organisms in temperate and subtropical areas of Japan. *Int. J. Syst. Evol. Microbiol.* *53*, 711–720.
- Jiang, Z.-F., Xia, F., Johnson, K.W., Brown, C.D., Bartom, E., Tuteja, J.H., Stevens, R., Grossman, R.L., Brumin, M., White, K.P., et al. (2013). Comparison of the Genome Sequences of “*Candidatus Portiera aleyrodidarum*” Primary Endosymbionts of the Whitefly *Bemisia tabaci* B and Q Biotypes. *Appl. Environ. Microbiol.* *79*, 1757–1759.
- Jurado, V., Gonzalez, J.M., Laiz, L., e Saiz-Jimenez, C. (2006). *Aurantimonas altamirensis* sp. nov., a member of the order Rhizobiales isolated from Altamira Cave. *Int. J. Syst. Evol. Microbiol.* *56*, 2583–2585.
- Jussila, J., Komulainen, H., Kosma, V.M., Pelkonen, J., e Hirvonen, M.R. (2002). Inflammatory potential of the spores of *Penicillium spinulosum* isolated from indoor air of a moisture-damaged building in mouse lungs. *Environ. Toxicol. Pharmacol.* *12*, 137–145.
- Kaltenbach, T. (2019) Dead cork oak [fotografia] <http://www.wood-land.net/forestry/combato-of-cork-oak-die-back/> [25 de janeiro de 2019]
- Kaltenbach, T. The decline of Cork oak (*Quercus suber*) and Holm oak (*Q. ilex/rotundifolia*) in South Western Portugal (Dissertação de Mestrado). Highschool of Applied Science and Art (HAWK), Göttingen.
- Kilaru, S., Hoegger, P.J., e Kües, U. (2006). The laccase multi-gene family in *Coprinopsis cinerea* has seventeen different members that divide into two distinct subfamilies. *Curr. Genet.* *50*, 45–60.
- Kim, H.N., Jin, H.Y., Kwak, M.J., Khaine, I., You, H.N., Lee, T.Y., Ahn, T.H., e Woo, S.Y. (2017). Why does *Quercus suber* species decline in Mediterranean areas? *J. Asia-Pacific Biodivers.* *10*, 337–341.
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., e Glöckner, F.O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* *41*, e1–e1.
- Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez,

- A., Kosciolk, T., McCall, L.-I., McDonald, D., et al. (2018). Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* *16*, 410–422.
- Komagata, K., Iino, T., e Yamada, Y. (2014). The family Acetobacteraceae. In *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 3–78.
- Lambiase, A. (2014). The family sphingobacteriaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*, E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt, e F. Thompson, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 907–914.
- Levy, S.E., e Myers, R.M. (2016). Advancements in Next-Generation Sequencing. *Annu. Rev. Genomics Hum. Genet.* *17*, 95–115.
- Li, W.-J. (2004). *Nesterenkonia halotolerans* sp. nov. and *Nesterenkonia xinjiangensis* sp. nov., actinobacteria from saline soils in the west of China. *Int. J. Syst. Evol. Microbiol.* *54*, 837–841.
- Liu, J.-M., Tuo, L., Habden, X., Guo, L., Jiang, Z.-K., Liu, X.-F., Chen, L., Zhang, Y.-Q., e Sun, C.-H. (2015). *Nesterenkonia populi* sp. nov., an actinobacterium isolated from *Populus euphratica*. *Int. J. Syst. Evol. Microbiol.* *65*, 1474–1479.
- Love, M.I., Anders, S., e Huber, W. (2014). Differential analysis of count data - the DESeq2 package.
- Lusa (2018, 9 de outubro). Proprietário de terreno em Braga multado por "mutilação" de 27 sobreiros. Sapoz4. Disponível em <https://24.sapo.pt/atualidade/artigos/proprietario-de-terreno-em-braga-multado-por-mutilacao-de-27-sobreiros>
- Mandic-Mulec, I., Stefanic, P., e van Elsas, J.D. (2015). Ecology of Bacillaceae. *Microbiol. Spectr.* *3*.
- Manion, Paul D. (1981) [imagem] *Tree Disease Concepts*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Massey Jr., F. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Am. Stat. Assoc. J.* *46*, 68–78.
- Matthias, H. (2013) Cork Oaks [fotografia] www.flickr.com/photos/sill-padrino/12080710903 [22 de julho de 2018]
- McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., Desantis, T.Z., Probst, A., Andersen, G.L., Knight, R., e Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* *6*, 610–618.
- McMurdie, P.J., e Holmes, S. (2014). Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput. Biol.* *10*.

- Mendes, Rita B. (2019, 23 de abril). Bactéria perigosa detetada em jardins de Gaia. Tvi24. Disponível em <https://tvi24.iol.pt/sociedade/xylella-fastidiosa/bacteria-perigosa-detetada-em-jardins-de-gaia>
- Mestre, M.C., Ulloa, J.R., Rosa, C.A., Lachance, M.A., e Fontenla, S. (2010). *Lachancea nothofagi* sp. nov., a yeast associated with *Nothofagus* species in Patagonia, Argentina. *Int. J. Syst. Evol. Microbiol.* 60, 2247–2250.
- Mhedbi-Hajri, N., Jacques, M., e Koebnik, R. (2011). Adhesion Mechanisms of Plant-Pathogenic *Xanthomonadaceae*. pp. 71–89.
- Michel A, Seidling W, editors (2016) Forest Condition in Europe: 2016 Technical Report of ICP Forests. Report under the UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP). BFW- Documentation 23/2016. Vienna: BFW Austrian Research Centre for Forests. 206 p.
- Moricca, S., Linaldeddu, B.T., Ginetti, B., Scanu, B., Franceschini, A., e Ragazzi, A. (2016). Endemic and Emerging Pathogens Threatening Cork Oak Trees: Management Options for Conserving a Unique Forest Ecosystem. *Plant Dis.* 100, 2184–2193.
- de Muinck, E.J., Trosvik, P., Gilfillan, G.D., Hov, J.R., e Sundaram, A.Y.M. (2017). A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome* 5, 68.
- Muniz. L. F. (s.d.) [fotografia] <http://greentopia.com.br/materiais-sustentaveis-pisos/> [22 de julho de 2018]
- Naves, Filomena (2017, 14 de outubro). Sobreiros a morrer e espécies em perigo a perder terreno [Versão electrónica]. *Diário de Notícias*. Disponível em <https://www.dn.pt/sociedade/interior/sobreiros-a-morrer-e-especies-em-perigo-a-perder-terreno-8841634.html>
- Nachar, N. (2008). The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. *Tutor. Quant. Methods Psychol.* 4, 13–20.
- Nilsson, R.H., Larsson, K.H., Taylor, A.F.S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F.O., Tedersoo, L., et al. (2019). The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47, D259–D264.
- Oliveira, M.C.D.S., Regitano, L.C.D.A., Roese, A.D., Anthonisen, D.G., Patocínio, E., Parma, M.M., Scagliusi, S.M.M., Timóteo, W.H.B., e Jardim, S.N. (2007). Fundamentos teórico-práticos e protocolos de extração e de amplificação de dna por meio da técnica de reação em cadeia da polimerase.
- Paes, Â.T. (2009). Por dentro da estatística O que fazer quando a distribuição não é normal? *Einstein Educ Contin Saúde* 7, 3–4.

- Pereira-Leal, J.B., Abreu, I.A., Alabaça, C.S., Almeida, M.H., Almeida, P., Almeida, T., Amorim, M.I., Araújo, S., Azevedo, H., Badia, A., et al. (2014). A comprehensive assessment of the transcriptome of Cork oak (*Quercus suber*) through EST sequencing. *BMC Genomics* 15, 1–14.
- Pinto, Armando S. (2015, 18 de janeiro). E por que não tentarmos salvar os sobreiros? [Versão electrónica]. Público. Disponível em <https://www.publico.pt/2015/01/18/economia/opiniao/e-por-que-nao-tentarmos-salvar-os-sobreiros-1682545>
- Qiu, P., e Hawkins, D. (2003). A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions. *J. R. Stat. Soc. Ser. D (The Stat.* 52, 151–164.
- Quail, M., Smith, M.E., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., e Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics* 13, 341.
- Reuter, J.A., Spacek, D. V., e Snyder, M.P. (2015). High-Throughput Sequencing Technologies. *Mol. Cell* 58, 586–597.
- Samson, R.A. (2000). Integration of Modern Taxonomic Methods For *Penicillium* and *Aspergillus* Classification (CRC Press).
- Sá, Ana C.(2019, 9 de maio). PDR 2020. Há 20 M€ contra pragas do pinheiro, sobreiro, azinheira, castanheiro e eucalipto. *Agricultura e Mar Actual*. Disponível em <http://agriculturaemar.com/pdr-2020-ha-20-me-contra-pragas-do-pinheiro-sobreiro-azinheira-castanheiro-e-eucalipto/>
- Sapp, M., Lewis, E., Moss, S., Barrett, B., Kirk, S., Elphinstone, J., e Denman, S. (2016). Metabarcoding of Bacteria Associated with the Acute Oak Decline Syndrome in England. *Forests* 7, 95.
- Schirmer, M., Ijaz, U.Z., D'Amore, R., Hall, N., Sloan, W.T., e Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* 43, e37–e37.
- Shannon, C.E. & Weaver, W. (1949) The mathematical theory of communication. The University of Illinois Press, Urbana, 117.
- Shokralla, S., Porter, T.M., Gibson, J.F., Dobosz, R., Janzen, D.H., Hallwachs, W., Golding, G.B., and Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Sci. Rep.* 5, 9687.
- Silva, Alexandre (2017, 30 de março). Abate de sobreiros em Alcoentre já chegou à Assembleia da República. *Fundamental*. Disponível em <https://www.fundamental-diario.pt/20170330abate-sobreiros-azambuja-ja-chegou-assembleia-da-republica>
- Simpson, J. (1949). Measurements of diversity. *Nature* 163, 688.

- Smahi, H., Belhoucine-guezouli, L., Franceschini, A., e Scanu, B. (2017). *Phytophthora* species associated with cork oak decline in a Mediterranean forest in western Algeria. *IOBC-WPRS Bull.* 127, 123–129.
- Srivastava, M.S., e Hui, T.K. (1987). On assessing multivariate normality based on Shapiro-Wilk w statistic. *Stat. Probab. Lett.* 5, 15–18.
- Tedersoo, L., Bahram, M., Polme, S., Koljalg, U., Yorou, N.S., Wijesundera, R., Ruiz, L. V., Vasco-Palacios, A.M., Thu, P.Q., Suija, A., et al. (2014). Global diversity and geography of soil fungi. *Science* (80-.). 346, 1256688–1256688.
- Thorsen, J., Brejnrod, A., Mortensen, M., Rasmussen, M.A., Stokholm, J., Al-Soud, W.A., Sørensen, S., Bisgaard, H., e Waage, J. (2016). Large-scale benchmarking reveals false discoveries and count transformation sensitivity in 16S rRNA gene amplicon data analysis methods used in microbiome studies. *Microbiome* 4, 62.
- Toth, I.K., van der Wolf, J.M., Saddler, G., Lojkowska, E., Hélias, V., Pirhonen, M., Tsrör (Lahkim), L., e Elphinstone, J.G. (2011). *Dickeya* species: An emerging problem for potato production in Europe. *Plant Pathol.* 60, 385–399.
- Vandenkoornhuysse, P., Quaiser, A., Duhamel, M., Le Van, A., e Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytol.* 206, 1196–1206.
- Vargha, A., e Delaney, H.D. (1998). The Kruskal-Wallis Test and Stochastic Homogeneity. *J. Educ. Behav. Stat.* 23, 170–192.
- Wang, Brian (2017) [imagem] 16S rRNA indexed primers amplify phylogenetic markers for microbiome sequencing analysis. Integrated DNA Technologies.
- Ward, J.H. (1963). Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* 236–244.
- Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5, 1–18.
- White, D.C., Suttont, S.D., e Ringelberg, D.B. (1996). The genus *Sphingomonas*: Physiology and ecology. *Curr. Opin. Biotechnol.* 7, 301–306.
- Zalar, P., Gostinčar, C., de Hoog, G.S., Uršič, V., Sudhadham, M., e Gunde-Cimerman, N. (2008). Redefinition of *Aureobasidium pullulans* and its varieties. *Stud. Mycol.* 61, 21–38.