



HAL
open science

Genetic architecture of the interactions between English oak (*Quercus robur* L.) and the microbial community of its phyllosphere

Boris Jakuschkin

► **To cite this version:**

Boris Jakuschkin. Genetic architecture of the interactions between English oak (*Quercus robur* L.) and the microbial community of its phyllosphere. Symbiosis. Université de Bordeaux, 2015. English. NNT : 2015BORD0170 . tel-01665488

HAL Id: tel-01665488

<https://theses.hal.science/tel-01665488>

Submitted on 16 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

**DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE SCIENCES ET ENVIRONNEMENTS
SPÉCIALITÉ: ECOLOGIE ÉVOLUTIVE, FONCTIONNELLE ET DES
COMMUNAUTÉS

Par Boris JAKUSCHKIN

**Genetic architecture of the interactions between English oak (*Quercus robur* L.)
and the microbial community of its phyllosphere**

Sous la direction de : Cécile Robin
(co-encadrants: Virgil Fievet, Corinne Vacher)

Soutenue le 15 december 2015

Membres du jury :

M. LAVERGNE Sébastien	Chargé de recherche HDR CNRS Univ Grenoble	Rapporteur
M. ROUX Fabrice	Chargé de recherche HDR INRA Toulouse	Rapporteur
M. MARÇAIS Benoit	Directeur de recherche INRA Nancy	Rapporteur
M. CANDRESSE Thierry	Directeur de recherche INRA Univ Bordeaux	Président
M. MAIGNIEN Lois	Maitre de conférences CNRS Univ Brest	Examineur
M. PLOMION Christophe	Directeur de recherche INRA Univ Bordeaux	Examineur
M. FIEVET Virgil	Maitre de conférences INRA Univ Bordeaux	Examineur
Mme. ROBIN Cécile	Directrice de recherche INRA Univ Bordeaux	Directrice de thèse

Résumé

De nombreux et divers micro-organismes vivent dans les tissus interne et externe des feuilles des plantes, la phyllosphère. Ils influencent de nombreux traits, les interactions biotiques, le flux d'énergie, la tolérance au stress de leur hôte et en fin de compte la valeur sélective de leurs hôtes. Il a été montré que plusieurs traits quantitatifs de plantes structurent la communauté microbienne de la phyllosphère. Ainsi des Loci de ces traits quantitatifs (Quantitative Trait Loci QTL) liés à la structure de cette communauté étaient attendus. L'objectif principal de ce travail était de rechercher des régions génomiques chez le chêne (*Quercus robur* L.), dont l'effet se prolonge jusqu'au niveau de la communauté, influençant ainsi le microbiote de la phyllosphère. Tout d'abord, nous avons étudié la composition et le réseau d'interactions du microbiote de la phyllosphère partant un intérêt particulier à *Erysiphe alphitoides*, un agent pathogène majeur pour les chênes. Nous avons montré que l'infection par *E. alphitoides* est accompagnée par des changements dans la composition de la communauté fongique foliaire, mais pas dans la composition de la communauté bactérienne. Nous avons souligné certains partenaires d'interaction d'*E. alphitoides* et nous avons montré que le réseau d'interactions microbiennes, contrairement aux résultats précédents, été dominé par des interactions positives. Ensuite nous avons effectué une analyse QTL de descripteurs de la communauté microbienne dans une population de pleins frères. Nous avons trouvé 8 QTL correspondant à des traits de la communauté microbienne: composition fongique et bactérienne, diversité fongique, pourcentage de *Erysiphe alphitoides*. Trois d'entre eux sont colocalisés avec un QTL de sensibilité à *E. alphitoides*, suggèrent un fort déterminisme génétique de la résistance à l'oïdium chez *Q. robur*. Enfin, nous présentons les résultats préliminaires d'une étude d'association génétique et discutons nos résultats avec une perspective évolutive.

Mots clés : microbiote de la phyllosphère, séquençage de dernière génération, *Erysiphe alphitoides*, *Quercus robur*, génétique des communautés, réseaux d'interactions microbiennes

Abstract

Numerous and various microorganisms inhabit inner and outer tissues of plant leaves, the phyllosphere. They influence many plant traits, biotic interactions, energy flux, host stress tolerance and ultimately the fitness of their hosts. Many plant quantitative traits were shown to structure the phyllosphere microbial community. Hence quantitative trait loci (QTLs) linked to the structure of this community were expected. The main objective of this work was to search for genomic regions in oak (*Quercus robur* L.), whose effect extends to the community level, influencing the phyllosphere microbiota. First, we studied the composition and the interaction network of oak phyllosphere microbiota with specific focus on *Erysiphe alphitoides*, a major biotrophic pathogen of oak. We showed that infection by *E. alphitoides* is accompanied by changes in the foliar fungal community composition but not in the bacterial community composition. We highlighted likely interaction partners of *E. alphitoides* and we showed that the complex microbial interaction network, in contrast to previous findings, was dominated by positive interactions. Next we performed QTL analysis of microbial community descriptors in a full-sib mapping population of oak. We found 8 QTLs for microbial community traits: fungal and bacterial composition, fungal diversity, and percentage of *Erysiphe alphitoides* reads. Three of these QTLs colocalized with a QTL for powdery mildew sensibility, suggesting for strong genetic determinism of powdery mildew resistance in *Q. robur*. Finally we present preliminary results of a genetic association study and discuss our findings within an evolutionary perspective.

Keywords: phyllosphere microbiota, next-generation sequencing, *Erysiphe alphitoides*, *Quercus robur*, community genetics, microbial interaction networks

Acknowledgements:

Many thanks to my supervisors Cécile Robin, Corinne Vacher and Virgil Fievet for their mentoring and encouragement. This thesis would not be possible without the work of Catherine Bodénès, Marie-Laure Desprez-Loustau, Antoine Kremer and Christophe Plomion. Many thanks for crash course in QTL analysis to Jérôme Bartholomé and Jean-Marc Gion. Thanks to the current and former staff of the sequencing facilities: Thibaut Decourcelle, Adline Delcamp, Jennifer Dupiot-Chiron, Christophe Hubert, Sarah Monllor, Frank Salin. Many thanks to the technical staff for their help and advise in the lab and out in the field: Xavier Capdevielle, Olivier Fabreguettes, Gilles Saint-Jean, Martine Martin. Thanks to the BioGeCo staff. Thanks to the doctoral school and EVOLTREE travel grants, which allowed me to participate to events in France and abroad, where I have met many inspiring researchers and learned about many interesting projects. I also would like to thank members of two PhD committees: Antoine Kremer, Marc Buée, Stéphane Uroz, Jennifer Rowntree, and Stéphane Compant for their valuable advise and fruitful discussions. Many thanks to all the software developers of great powerful tools under GPL: Bio-Linux, LibreOffice, L^AT_EX, L^AX, QIIME, R, RStudio and many others. I owe a great debt of gratitude to my family for their constant support and encouragement. Also many thanks to my friends, with whom we shared good times and who stood by me when it was tough: Jean-Paul M, Cécile and Paul F, Clémence F, Celine S, Daria K, Anna Z, Gianluca M, Andjela D, Petra U, Victor S, Gabriel G, Nikita and Kira F, Giorgi B, Levan K, Georgiy and Natasha T, Laure V, Paul F and Sacha B. Special thanks for their support to Virgil F, Sergey S and Guillaume D. And I want to express my gratitude to Olga for her understanding and support.

Contents

Contents	V
1 Introduction	1
1.1 Community genetics	1
1.2 The phyllosphere	11
Definition of the term	11
Topography and microclimatic conditions	11
Microbial inhabitants of the phyllosphere	14
Microbiota functions in the phyllosphere	17
1.3 Plants strategy to cope with the diversity of biotic interactions might select for specific networks of interacting species in the associated communities	18
1.4 Powdery mildew	19
1.5 Metabarcoding for studying microbial community	21
Fungi	22
Bacteria	22
1.6 Objectives of this work	23
1.7 Bibliography	24
2 Deciphering the pathobiome	41
2.1 Introduction	42
2.2 Material and methods	43
Study site and sampling design	43
Leaf DNA extraction	44
Amplification and sequencing of the fungal assemblages	44
Amplification and sequencing of the bacterial assemblages	45
Processing of the fungal 454 pyrosequencing data	46
Processing of the bacterial Illumina data	47
Statistical analyses	48
2.3 Results	49
Taxonomic composition of oak foliar microbial communities	49
Relationship between the abundance of powdery mildew symptoms and the number of reads assigned to <i>E. alphitoides</i>	52

	Relationship between the level of infection by <i>E. alphitoides</i> and the composition of foliar microbial communities	52
	Inferred interactions between <i>E. alphitoides</i> and other members of the foliar community	55
2.4	Discussion	57
2.5	Acknowledgements	60
2.6	References	61
2.7	Supplementary methods S1 - Preliminary experiment aimed at assessing the effect of oak leaf storage on fungal community composition.	68
	Study site and sampling design	68
	DNA extraction and fungal communities sequencing	68
	Statistical analyses	69
	Results	69
	Discussion	70
	References	70
3	QTLs for microbial community descriptors	71
3.1	Introduction	72
3.2	Material and Methods	73
	Experimental site and leaf-sampling design	73
	DNA extraction	74
	Amplification and sequencing of the fungal assemblages	74
	Amplification and sequencing of the bacterial assemblages	74
	Processing of the fungal 454 pyrosequencing data	75
	Processing of the bacterial Illumina data	76
	Construction of foliar microbial networks	76
	Choice of traits for QTL analysis	77
	Genotyping and map construction	77
	Number and type of markers	79
	QTL analysis	80
3.3	Results	81
	Sequencing data	81
	Fungal data	81
	Bacterial data	81
	Correlation among traits	82
	Chromosomal loci for the microbial community descriptors (QTLs)	82
3.4	Discussion	85
3.5	Bibliography	88
4	Association mapping study	96
4.1	Introduction	96
4.2	Material and Methods	97
	Population structure and Sampling	97

	DNA extraction	98
	Amplification and sequencing of the fungal assemblages	98
	Amplification and sequencing the bacterial assemblages	99
	Processing of the fungal Ion Torrent sequencing data	99
	Processing of the bacterial Illumina sequencing data	100
	Statistical analysis and choice of traits	100
4.3	Results	102
	Sequencing data	102
	Fungal data	102
	Bacterial data	103
	Differences between maternal and offspring trees	104
4.4	Discussion	104
4.5	Bibliography	107
5	General discussion	111
5.1	Recapitulation of the major findings	111
5.2	Possible evolutionary mechanisms	113
5.3	Bibliography	117
6	Appendix	122
6.1	Appendix A - An evolutionary ecology perspective to address forest pathology challenges of today and tomorrow	122
6.2	Appendix B - Learning Ecological Networks from Next-Generation Sequencing Data	146
6.3	Appendix C - Résumé en français	186

Chapter 1

Introduction

Today it is commonly accepted that microbe-free plants probably do not exist (Partida-Martínez & Heil, 2011). The entirety of Bacteria, Archea and Fungi inhabiting the plant outer and inner tissues is called plant **microbiota**¹. It impacts plant hormones, nutrient cycling, plant resistance to biotic and abiotic stress, in essence plant growth and survival, two key components of fitness. Plant fitness therefore depends on the plant per se and its microbiota, which collectively form a **holobiont**² (Vandenkoornhuysen *et al.*, 2015). The holobiont concept requires a holistic perception for the understanding of plant ecology and evolution. A broad overview of plant fitness thus needs: 1) an exhaustive description of the microbiota, 2) the determination of the core microbiota, 4) an understanding of the community assemblage rules shaping those communities, 5) an understanding of the functions of the microbiota, 6) an understanding of the plant-microbiota interactions, and 7) of the interactions occurring between the members of the microbiota. The community genetics approach provides tools to better understand how the plant influences its microbiota.

1.1 Community genetics (Plant genetics structures associated communities)

Community genetics³ unifies the diverse theories, concepts and techniques of community ecology, evolutionary biology, genetics and genomics to understand the ecology and evolution of interactions among species within communities (Antonovics, 1992; Agrawal, 2003; Whitham *et al.*, 2006; Rowntree *et al.*, 2011; Hersch-Green *et al.*, 2011). The idea was born out of the recognition that most ecological systems

¹**Microbiota:** the sum of all microorganisms (Bacteria, Archea, Fungi) that inhabit a specific ecological niche

²**Holobiont:** the animal or plant with all of its associated microorganisms

³**Community genetics:** the study of the role of genetic variation in influencing species interactions and determining community structure

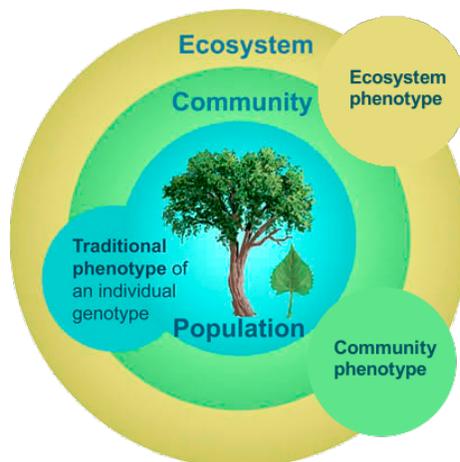


Figure 1.1: Extended phenotype (modified after Whitham *et al.* (2006))

do not involve simple pairwise interactions among species, but multiple complex interactions within and among trophic levels, and that traditional coevolutionary models were too limited to explain species interactions in this context (Neuhauser *et al.*, 2003; Rowntree *et al.*, 2011).

An important milestone in the development of community genetics as an autonomous research area was the 2003 Ecology special feature edited by A.A. Agrawal (2003). In this feature, Whitham *et al.* (2003) reviewed the evidence for the presence of an **extended phenotype**⁴ sensu Dawkins (1982), i.e., the effects of genes at levels higher than the population (Figure 1.1), in multiple ecological systems and highlighted the role of genetic variation in important traits of **dominant**⁵ and of **keystone species**⁶ in structuring natural communities. Neuhauser *et al.* (2003) demonstrated in turn that neither ecological nor evolutionary models alone could explain the response of communities under strong selection and highlighted areas of applied research where a community genetic approach could answer tangible questions with a large socio-economic impact. For example, a community genetics perspective helps to understand how natural enemies (predators, parasitoids, pathogens) may affect the rate of resistance evolution of herbivores to plant defenses because of the role that natural enemies play in biasing the mortality of resistant vs. susceptible herbivore genotypes (Neuhauser *et al.*, 2003). In the case of strong resistance (e.g., plants expressing high doses of *Bt* toxin), the density of feeding-stage herbivores on *Bt* plants will initially be orders of magnitude lower than the density of feeding-stage herbivores on non-*Bt* plants. However, resistant individuals will form a much higher fraction of the herbivores feeding on the *Bt* plants than on the non-*Bt* plants. A pos-

⁴**Extended phenotype:** the effects of genes at levels higher than the population.

⁵**Dominant species:** a species, which garners a disproportionate share of resources and contributes most to productivity.

⁶**Keystone species:** a species whose effect is large, and disproportionately large relative to its abundance

itively density-dependent attack rate implies lower risk at low densities, which would allow resistant genotypes to 'escape' from attacks by natural enemies in the *Bt* crop. The result would be accelerated resistance evolution. An inversely density-dependent attack rate, on the other hand, would imply increased risk at lower herbivore densities and a delay in the evolution of resistance (Neuhauser *et al.*, 2003). Community genetics has largely focused on the role of genetic variation in plant populations for associated microbial, arthropod and plant communities (Whitham *et al.*, 2006; Johnson & Stinchcombe, 2007; Hughes *et al.*, 2008; Bailey *et al.*, 2009; Whitham *et al.*, 2012). It supports the hypothesis that intraspecific genetic variation has cascading effects on communities and ecosystems. For instance, genetic differences among individual plants influence abundance, composition and diversity of herbivorous and predaceous arthropod species (Table 1.1 on the following page), the performance and the coexistence of competing plant species (Table 1.2 on page 6), the microbial communities (Table 1.3 on page 7), and the flow of energy and nutrients through ecosystems (Table 1.4 on page 9).

Every biotic interaction of an individual is considered as a part of its **community phenotype**⁷, whether the partners interact directly, e.g. they are involved in a predator-prey or a host-pathogen interaction, or indirectly through indirect trait-mediated interactions. The community phenotype was shown to be partially inheritable, and high levels of heritability were reported (broad sense) (Johnson & Agrawal, 2005; Bailey *et al.*, 2006; Shuster *et al.*, 2006; Schweitzer *et al.*, 2008a; Whitham *et al.*, 2008). Host plant genotype explained for example up to 41% of the variation in the diversity of insect communities associated with the evening primrose (Johnson & Agrawal, 2005) and 50–63% of the community composition on a single cottonwood species (Shuster *et al.*, 2006). In these cases, some traits and thus some genes must drive the community phenotype more than others and Quantitative Traits Loci (QTL) for community descriptors are thus expected (Whitham *et al.*, 2003). A highly accessible study system to apply community genetics concept is the plant and its **phyllosphere**⁸ associated communities, as many of the plant traits exert direct effects on the phyllosphere habitat with the potential of extended effects at the community and ecosystem levels (Whipps *et al.*, 2008; Bodenhausen *et al.*, 2014). These traits vary between individual plants and are ultimately controlled by plant genotype (Hunter *et al.*, 2010; Reisberg *et al.*, 2013).

⁷**Community phenotype:** effects of genes at community level

⁸**Phyllosphere:** foliar tissues and foliar surfaces (see Section 1.2 on page 11)

Table 1.1: Examples of plant genetic variability effects on arthropod communities.

Organism	Effects of plant genetic variability	References
Late goldenrod (<i>Solidago altissima</i>)	Goldenrod clones differed in resistance to 15 of 16 herbivorous insect species. Authors found heritable variation in resistance for 10 of these insects.	Maddox & Root, 1987
Arroyo willow (<i>Salix lasiolepis</i>)	The total community structure of four sawfly species varied among willow clones. Between 9 and 50% of variation in the density of these species was due to willow clone genotype.	Fritz & Price, 1988
Seaside tansy (<i>Borrchia sp.</i>)	Plant clone and environment were found to have significant effect on the success of the gall midge <i>Asphondylia borrichiae</i> , and on the parasitoid community composition. The effect of environment was usually stronger than that of genotype.	Stiling & Rossi, 1996
Cottonwoods interspecific hybrids (<i>Populus angustifolia</i> x <i>P. fremontii</i>)	Plant cross type structured the arthropod community (207 species) of individual trees (common garden and wild stands). Plant genetic diversity accounted for nearly 60% of the variation in arthropod diversity.	Wimp <i>et al.</i> , 2004
Common evening-primrose (<i>Oenothera biennis</i>)	Genotypic differences among plants (14 genotypes, 926 plants) accounted for as much as 41% of the variation in arthropod diversity. There were significant plant genotype-by-habitat interactions (five natural habitats).	Johnson & Agrawal, 2005
Late goldenrod (<i>Solidago altissima</i>) continued...	Increasing population genotypic diversity in <i>S. altissima</i> determined arthropod diversity and community structure.	Crutsinger <i>et al.</i> , 2006

Organism	Effects of plant genetic variability	References
Common evening-primrose (<i>Oenothera biennis</i>)	Genetic variation in heritable plant traits (% of leaf water content, trichome density, and % of leaf N) explained 49% of the variation in aphid population growth rate.	Johnson, 2008
Quaking aspen (<i>Populus tremuloides</i>)	Varying plant genotypic diversity and watering treatment altered multivariate arthropod community composition and structure.	Kanaga <i>et al.</i> , 2009
Maya nut (<i>Brosimum alicastrum</i>)	More genetically similar trees host more similar communities of epiphytic plants, leaf litter invertebrates and trunk dwelling invertebrates.	Zytyńska <i>et al.</i> , 2011
Quaking aspen (<i>Populus tremuloides</i>)	Arthropod community similarity was better explained by geographic (56%) and environmental (32%) distance than genetic distance in clonal aspen stands. There was a significant relationship between genetic and arthropod diversity in hybrid stands.	Meneses <i>et al.</i> , 2012
Populus fremontii, <i>P. angustifolia</i> and their F ₁ hybrids	Plant hybridization and genotype impact litter-dwelling spiders.	Wojtowicz <i>et al.</i> , 2014
Sessile oak (<i>Quercus petraea</i>) Pedunculate oak (<i>Q. robur</i>)	Contrary to previous findings authors found no significant influence of oak genotype on arthropod communities.	Gossner <i>et al.</i> , 2015

Table 1.2: Examples of plant genetic variability effects on associated plant communities.

Organism	Effects of plant genetic variability	References
Scots pine (<i>Pinus sylvestris</i>)	Genetically determined chemical diversity of monoterpenes of individual trees was significantly positively associated with the species richness of the ground vegetation beneath the tree.	Iason <i>et al.</i> , 2005
Smooth cordgrass (<i>Spartina alterniflora</i>)	Different clone genotypes influenced recruitment and growth of other plant species.	Proffitt <i>et al.</i> , 2005
Scots pine (<i>Pinus sylvestris</i>)	The monoterpene phenotype of <i>P. sylvestris</i> had an effect on the patterning of the understory vegetation.	Pakeman <i>et al.</i> , 2006
Black mustard (<i>Brassica nigra</i>)	Genetic variation in the concentration of an allelopathic secondary compound in <i>B. nigra</i> is necessary for the coexistence of <i>B. nigra</i> and its competitor species.	Lankau & Strauss, 2007
Coyote brush (<i>Baccharis pilularis</i>)	Genetically based architectural dimorphism within a dominant native shrub, <i>B. pilularis</i> , influenced richness and cover of colonizing plant species (both native and exotic species).	Crutsinger <i>et al.</i> , 2010
Cottonwoods interspecific hybrids (<i>Populus angustifolia</i> x <i>P. fremontii</i>)	Understory biomass and cover increase along the genetic gradient from <i>P. angustifolia</i> to <i>P. fremontii</i> . Along the same hybridization gradient, species richness decreases and species composition shifts.	Lamit <i>et al.</i> , 2011
Narrowleaf cottonwood (<i>Populus angustifolia</i>)	Tree genotype affected lichen species richness, total lichen cover, lichen species composition, and rough bark cover, whereas environment only influenced composition and there were no genotype by environment interactions.	Lamit <i>et al.</i> , 2015

Table 1.3: Examples of plant genetic variability effects on microbial communities.

Organism	Effects of plant genetic variability	References
Mountain birch (<i>Betula pubescens</i>)	Mountain birch trees showed heritable variation in their foliar endophytic fungi.	Elamo <i>et al.</i> , 1999
F ₁ hybrids of <i>Populus fremontii</i> x <i>P. angustifolia</i> , backcross <i>P. angustifolia</i> , and <i>P. fremontii</i>	Introgression of Fremont's cottonwood markers was strongly and negatively related to the concentration of condensed tannins and positively related to endophytes, indicating that these compounds may be a chemical mechanism that inhibits fungal colonization in cottonwoods.	Bailey <i>et al.</i> , 2005
F ₁ hybrids of <i>Populus fremontii</i> x <i>P. angustifolia</i> , backcross <i>P. angustifolia</i>	Microbial community composition above- and below-ground was correlated to the concentration of condensed tannin in the foliage of the associated tree genotype.	Schweitzer <i>et al.</i> , 2006
European beech (<i>Fagus sylvatica</i>)	Dissimilarity between fungal assemblages was correlated with the genetic distance rather than the geographic distance between trees.	Cordier <i>et al.</i> , 2012a
Garlic mustard (<i>Alliaria petiolata</i>)	Variation in allelochemical production among <i>A. petiolata</i> individuals altered the plant's impact on soil communities, especially arbuscular mycorrhizal fungi. These altered soil communities in turn lead to a gradient of impacts on native tree seedlings.	Lankau, 2011
Balsam poplar (<i>Populus balsamifera</i>)	Host genotypes had a structuring effect on the composition of foliar fungal communities.	Bálint <i>et al.</i> , 2013

continued...

Organism	Effects of plant genetic variability	References
Norway spruce (<i>Picea abies</i>)	Authors found a statistically significant clone-specific effect on ectomycorrhizal fungal diversity and dominating fungal species.	Velmalá <i>et al.</i> , 2013
Thale cress (<i>Arabidopsis thaliana</i>)	Different <i>Arabidopsis</i> accessions exhibited different bacterial communities, indicating that plant host genetic factors shape the associated microbiota.	Bodenhausen <i>et al.</i> , 2014
Wine grape (<i>Vitis vinifera</i>)	Varietal patterns in grape-surface microbiota suggest a genetic component to host-microbial interactions on the grape surface.	Bokulich <i>et al.</i> , 2014
Narrowleaf cottonwood (<i>Populus angustifolia</i>)	Endophytic fungal community structure, including composition and abundance, varied among genotypes.	Lamit <i>et al.</i> , 2014
Balsam poplar (<i>Populus balsamifera</i>)	The phyllosphere fungal communities are strongly influenced by the interaction of environment and host genetic properties.	Bálint <i>et al.</i> , 2015
Lettuce accessions (<i>Lactuca spp.</i>)	Different lettuce genotypes supported significantly different phyllosphere fungal communities.	Hunter <i>et al.</i> , 2015
Cereals (<i>Triticum aestivum</i> , <i>Hordeum vulgare</i> , <i>Avena sativa</i> , <i>Secale cereale</i> , <i>Triticum x Secale</i>)	Host genotype at both the species and cultivar level is important in shaping the phyllosphere fungal communities, whereas fungicide treatment and location had minor effects.	Sapkota <i>et al.</i> , 2015

Table 1.4: Examples of plant genetic variability effects on ecosystems.

Organism	Effects of plant genetic variability	References
<i>Populus fremontii</i> , <i>P. angustifolia</i> and their F ₁ hybrids	Authors found that the concentration of condensed tannins is genetically based and is the best predictor of ecosystem-level processes. Condensed tannin inputs from foliage explained 55–65% of the variation in soil net nitrogen mineralization.	Schweitzer <i>et al.</i> , 2004
Trembling aspen (<i>Populus tremuloides</i>)	Aspen genotype identity explains a large portion of variation in leaf litter decomposition rates, and therefore influences nutrient cycling.	Madritch <i>et al.</i> , 2006
Populus model system	Condensed tannins in plants consistently slow rates of leaf litter decomposition and reduce nutrient availability in terrestrial ecosystems.	Schweitzer <i>et al.</i> , 2008b
<i>Populus fremontii</i> , <i>P. angustifolia</i> and their F ₁ and backcross hybrids	Tree genetic variation correlated with soil microbial community composition and nutrient availability (inorganic Nitrogen, Ammonium, Nitrate).	Schweitzer <i>et al.</i> , 2010
Various tree species	Tree genotypes differ in their compatibility with particular ectomycorrhizal fungal species, and therefore could have cascading effects on soil C sequestration.	Hoeksema & Classen, 2012
Various terrestrial plant species	Because different genotypes of terrestrial plant species support different soil biota and soil processes, shifts in genetics are likely to have cascading effects on ecosystems.	Fischer <i>et al.</i> , 2014

continued...

Organism	Effects of plant genetic variability	References
<i>Populus fremontii</i> , <i>P. angustifolia</i>	Closely related tree species differentially modulate elemental (C, N) fluxes in streamwater during decomposition, with initial litter characteristics likely driving nutrient cycling during decomposition.	Pastor <i>et al.</i> , 2014

1.2 The phyllosphere, a link between the plant and the world

Definition of the term

Two definitions of the term “phyllosphere” (Greek “phyllo” - leaf) are commonly found in the literature. One refers to all aboveground parts of a plant (e.g. Lindow & Brandl, 2003; Vorholt, 2012; Bulgarelli *et al.*, 2013), the other refers solely to the leaf surface and the leaf tissues (e.g. Andrews, 1992; Morris, 2001; Rastogi *et al.*, 2013; Peñuelas & Terradas, 2014). Further in this work, I will use the latter, unless indicated otherwise:

“The phyllosphere is the microenvironment extending from the leaf surface outward to the outer edge of the boundary layer surrounding the leaf and inward into the leaf tissues” (Morris, 2001)

The phyllosphere has been studied as a microbial ecosystem since the 1950s (Kerling, 1958; Morris, 2001). Since the dawn of the next-generation molecular technologies (Handelsman, 2004; Liu *et al.*, 2007; Bartram *et al.*, 2011; Di Bella *et al.*, 2013), their application in the area of phyllosphere microbiology is rapidly increasing and so is the number of new studies providing further insights into community variation, drivers, functions, and interactions with biotic and abiotic components (Whipps *et al.*, 2008; Jumpponen & Jones, 2010; Redford *et al.*, 2010; Finkel *et al.*, 2011; Cordier *et al.*, 2012b; Bálint *et al.*, 2013; Rastogi *et al.*, 2013; Kembel *et al.*, 2014).

Topography and microclimatic conditions

Millions of years of plant evolution have selected for adaptations to allow plants to survive in nearly all possible habitats on earth. These adaptations are reflected in the astonishing variability of microscopic surface structures on plant leaves (Figure 1.2 on the next page). The epidermal cells create a sculpturing of the surfaces in the micro-scale, but on a smaller scale, these surface cells are structured as well (Koch *et al.*, 2009). Epidermal cells can differ in the shape of the outer (periclinal) wall and the shape of the boundaries to the adjacent cells (anticlinal walls). If the anticlinal walls are sunken, they provide favorable conditions for microbial colonization, possibly due to better capability to retain surface water (Morris, 2001). Periclinal walls are covered with a continuous extracellular hydrophobic membrane, the cuticle. One of the most important properties of the cuticle is the transpiration barrier function. This property is based on the integrated (intracuticular) and superimposed (epicuticular) waxes, which can differ in composition, thickness and microtopography (Andrews, 1992; Yeats & Rose, 2013). Epicuticular wax structures and composition are important in determining fungal pathogen development and, thus, pathogenicity (Yeats & Rose, 2013). In addition to protection against water loss, the cuticle prevents leaching of ions from the inside of the cells. These plant adaptations to the

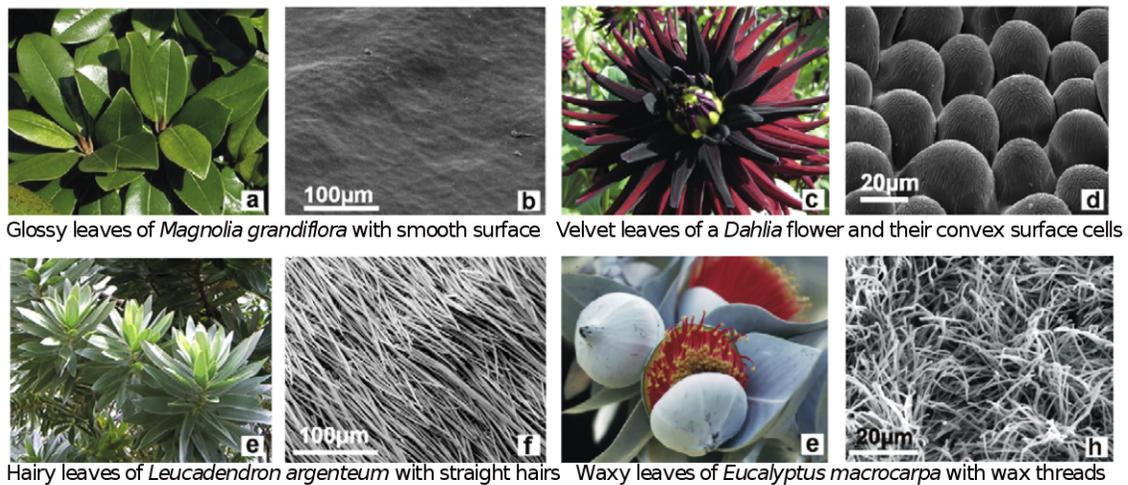
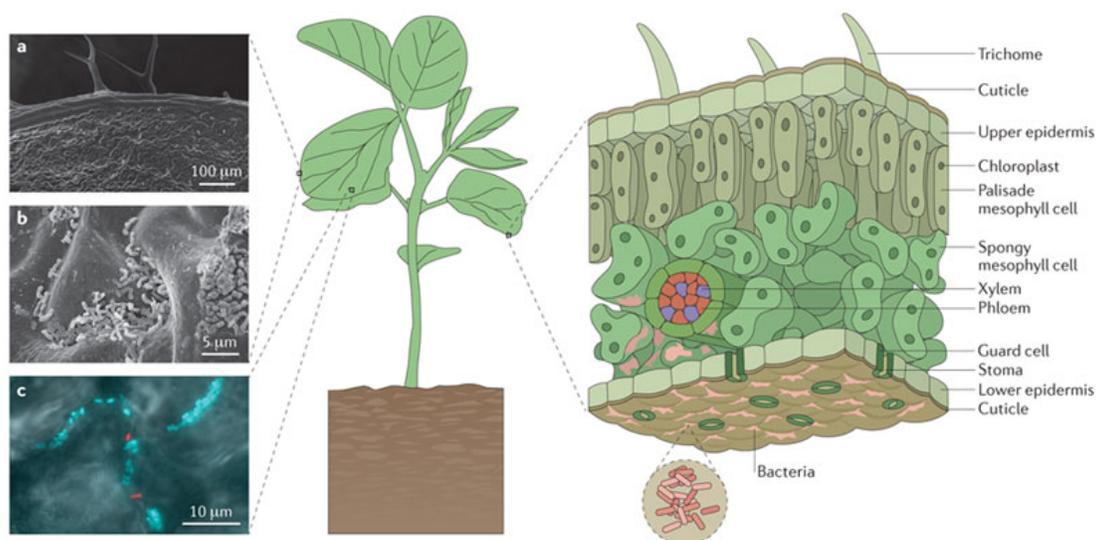


Figure 1.2: Macroscopically visible optical appearance of plant surfaces and their surface micro-structures, shown in the SEM micrographs. In (a) leaves appear glossy because of the flat surface structure of the surface, shown in (b). In (c) the flower leaves appears velvet, because of the micro-structure of the epidermis cells shown in (d). In (e) the white appearance of the leaves is caused by a dense layer of hairs, shown in (f). (g) Shows a white or bluish leaf surface, which is densely covered with three-dimensional waxes (h) (Koch *et al.*, 2009)

terrestrial life, create unfavorable conditions for microbial survival on the leaf surfaces – an oligotrophic environment with limited water availability (Derridj, 1996). Mercier and Lindow (2000) showed for various plant species that the population size of epiphytic bacteria is limited by the abundance of carbon sources on the leaf surface. They hypothesized that this is due to limitations in wettability and/or diffusion of nutrients across the leaf surface. Additionally the available nutrients are highly heterogeneously distributed on the leaf surface (Leveau & Lindow, 2001; Lindow & Brandl, 2003), which leads to aggregated distribution of micro-organisms on the leaves and their non-random associations with surface structures likely providing nutrients and water, e.g. along veins, anticlinal walls and trichomes (Monier & Lindow, 2004, 2005; Yadav *et al.*, 2005). To overcome the water limiting conditions, some bacteria developed the ability to increase water permeability of leaf cuticle, which in turn increased the availability of water and dissolved compounds in the phyllosphere (Schreiber *et al.*, 2005; Van Der Wal & Leveau, 2011). From the perspective of a microbial cell the topography of the leaf is one of hills, peaks, valleys, and craters (Figure 1.3 on the following page). Additionally to the aforementioned features common to most plants, some structural elements might be present or absent depending on plant taxonomic identity. These include trichomes and emergences, stomata, and hydathodes (Andrews, 1992). Leaf surface is a dynamic structure and its topography changes during leaf development and maturation and with it the conditions of



Nature Reviews | Microbiology

Figure 1.3: The phyllosphere. The middle panel shows a soybean plant, and the right panel shows a magnified schematic cross-section of a leaf. Various microorganisms, and in particular bacteria (shown) and fungi (not shown), can colonize the surface of the leaf and the apoplast. a) Scanning electron micrograph (SEM) showing the edge of an *Arabidopsis thaliana* leaf with trichomes visible on the upper side. b) Sem of epiphytic *Sphingomonas* sp. Fr 1 on the surface of a leaf. The bacterial cells are located in the grooves formed by the junctions of the epidermal plant cells. c) Epifluorescence image of *Pantoea agglomerans* str. 299R constitutively expressing cfp, and *Pseudomonas syringae* pv. tomato str. DC3000 producing mCherry, on an *A. thaliana* leaf. Parts a, b images courtesy of G. Innerebner, ETH Zurich; part c image courtesy of M. Remus-Emsermann, ETH Zurich (Vorholt, 2012)

leaf as microbial habitat. For instance young holm oak (*Quercus ilex*) leaves possess trichomes on **adaxial**⁹ and **abaxial**¹⁰ surfaces and shed the adaxial trichomes during the maturation, which decreases the wettability of the adaxial part of the leaf (Fernández *et al.*, 2014). Whereas for the birch trees it was shown that the number of trichomes seems to be established during the early leaf development, as their density decreases with leaf expansion (Valkama *et al.*, 2004). The adaxial surface of young cranberry (*Vaccinium macrocarpon*) leaves is characterized by regular pattern of broad expanses on the cell surfaces, which are damaged during the leaf maturation, creating irregular abrasions, and altered microclimatic conditions on the leaf surface (Mechaber *et al.*, 1996).

The phyllosphere environment encompasses a variety of microclimatic conditions

⁹**adaxial:** upper surface of a leaf

¹⁰**abaxial:** lower surface of a leaf

with cyclic and noncyclic environmental variables, including temperature, relative humidity, dew, rain, wind and radiation (Andrews, 1992; Morris, 2001; Peñuelas *et al.*, 2012). These conditions are of varying importance for the phyllosphere microorganisms: depending on whether they reside as epiphytes in exposed sites on the leaf surface or as endophytes, avoiding the stresses on the exterior by inhabiting the interior of the leaf. Leaves are surrounded by a boundary layer, a zone within a few millimeters of each leaf in which the diffusion resistance and heat transfer rates of the air are different from those of the ambient air, and which is influenced by the microtopography of the leaf (Koch *et al.*, 2009). Relative humidity within the boundary layer is generally greater than that of ambient air. Much of the water vapor in the boundary layer originates from stomata as leaves transpire (Morris, 2001). The temperature at the leaf surface may be as much as 5–7°C greater than that of the air and it can vary by as much as about 4°C across a single leaf, the leaf edges generally being cooler than the central part of the leaf (Morris, 2001; Pincebourde & Woods, 2012). Some fungal and bacterial pathogen infections can also lead to thermal changes on a leaf by either impeding (Oerke *et al.*, 2006) or increasing (Melotto *et al.*, 2006) its transpiration. The microbes on leaves within a plant canopy are exposed to a combination of both direct and diffuse UV light depending on the position of the leaf within the canopy and adaxial or abaxial position of the microbes on the leaf (Lindow, 2006). Tolerance to UV irradiation is reflected in the relatively high frequency of occurrence of pigmented bacteria and fungi on plants (Sundin, 2002; Lindow & Brandl, 2003; Jacobs *et al.*, 2005; Andrews, 1992; Atamna-Ismaeel *et al.*, 2012; Mizuno *et al.*, 2012).

As shown above many leaf morphological properties influence the microclimatic conditions of the phyllosphere and exert direct or indirect effects on its microbial inhabitants. Additionally leaf chemical properties can effect leaf microbiota on different levels. Leaf emit volatile organic compounds (VOC), which may constitute a carbon source for microbes (e.g. methane) or may protect the plant by inhibiting microbial growth (e.g. monoterpenes) (Junker & Tholl, 2013). Furthermore the chemical properties of leaves can have far reaching effects by influencing microbial communities on litter and thereby also the decomposition rate and the nutrient cycling (Pastor *et al.*, 2014). As it was shown for aspen, soil microbial communities adapt to rapidly process specific genotypes of aspen leaf litter, accelerating nutrient cycling. Hence the extended phenotype of aspen feeds back to influence the fitness of causal genotypes through soil processes (Madritch & Lindroth, 2011). It can not be ruled out, that leaf litter and soil microbiota influence the microbial community assembly of a tree phyllosphere as it was shown for crop plants (Copeland *et al.*, 2015).

Microbial inhabitants of the phyllosphere

The majority of the $1.5 \cdot 10^8 \text{km}^2$ of terrestrial surface of the Earth is covered with plants (Figure 1.4 on the next page). The total leaf surface area of these plants,

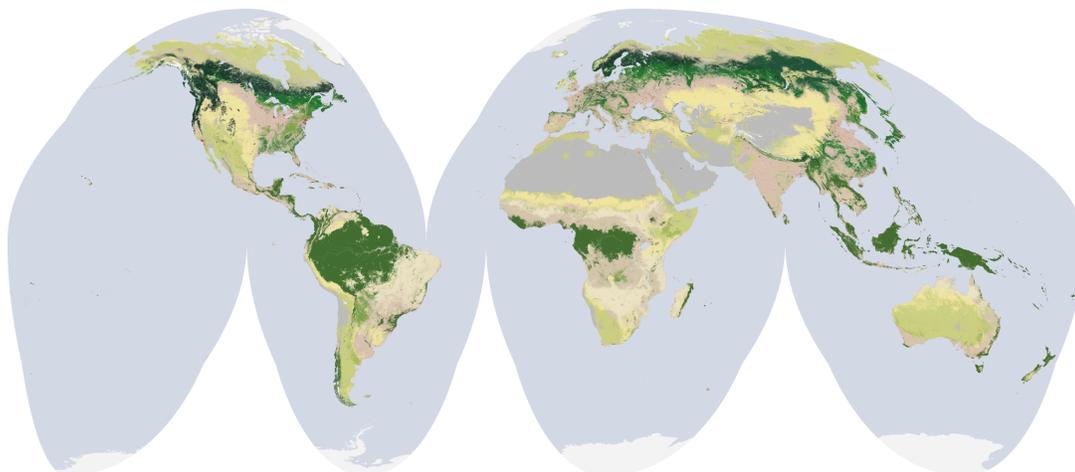


Figure 1.4: Global vegetation cover. Map by Robert Simmon, based on data from the MODIS Land Cover Group, Boston University.

that might be colonized by microbes, has been estimated to be about $2 - 6 \cdot 10^8 \text{ km}^2$, or approximately twice as great as the land surface (Morris, 2001; Vorholt, 2012). Assuming that there are on average $10^6 - 10^7$ bacteria per square cm of leaf surface, the planetary phyllosphere bacterial population could comprise up to 10^{26} cells (Morris & Kinkel, 2002; Vorholt, 2012). The total size of the fungal population of the planetary phyllosphere has not yet been estimated but is expected to be very high, yet lower than that of bacteria (Vorholt, 2012).

A wide range of microorganisms survives and proliferates in the phyllosphere. Natural openings in the leaf tissue, such as stomata and hydathodes, allow ready passage of microorganisms between the external and internal domains of the phyllosphere (Morris, 2001). But some fungal parasites developed the ability to breach through the cuticle and plant cell wall to reach plant cytoplasm (s. Section 1.4 on page 19). Bacterial and fungal populations on leaves are hyperdiverse (Arnold *et al.*, 2000; Yang *et al.*, 2001; Lambais *et al.*, 2006; Jumpponen & Jones, 2009; Zimmerman & Vitousek, 2012; Lau *et al.*, 2013), but some phyla and groups are frequently found on many plant species. Among Bacteria Proteobacteria (and within this phylum, the classes Alphaproteobacteria and Gammaproteobacteria in particular), predominate the phyllosphere of distinct plant species (Whipps *et al.*, 2008; Vorholt, 2012; Kembel *et al.*, 2014; Delmotte *et al.*, 2009; Redford & Fierer, 2009). Ascomycota dominate phyllosphere fungal communities of many plant species with the families Taphrinaceae and Mycosphaerellaceae seeming to be commonplace (Jumpponen & Jones, 2009; Cordier *et al.*, 2012b; Lau *et al.*, 2013; Douanla-Meli *et al.*, 2013; Meiser *et al.*, 2013).

Bacterial cells on leaves are commonly found as aggregates which are generally embedded within extracellular polymeric substances and can include multiple bacterial and fungal species (Lindow & Brandl, 2003; Monier & Lindow, 2005; Whipps

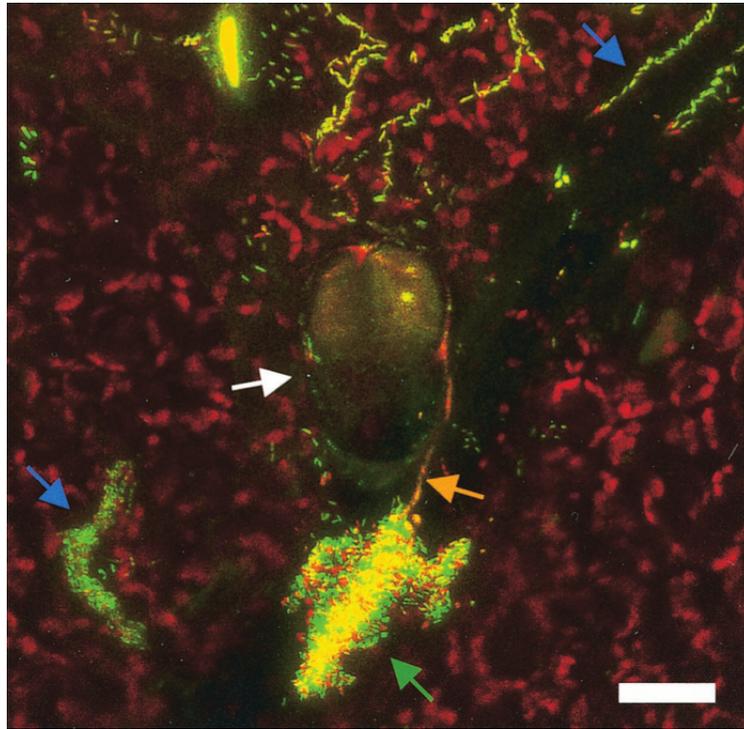


Figure 1.5: Fluorescence micrograph of the natural microbial flora colonizing the bean phyllosphere. A large mixed bacterial aggregate (green arrow), which also includes a fungal hypha (orange arrow), has formed at the base of a glandular trichome (white arrow). Bacteria are present also at plant cell junctions and on veins (blue arrows). The red background originates from the autofluorescent chloroplasts within the leaf epidermal cells. Gram-negative bacteria are green and gram-positive bacteria are red. Bar: 20 μm . (from Lindow & Brandl 2003)

et al., 2008; Vorholt, 2012; Danhorn & Fuqua, 2007). These aggregates are particularly common at the depressions formed at the junctions of epidermal cells, along the veins and at the bases of trichomes (Figure 1.5) (Lindow & Brandl, 2003; Monier & Lindow, 2004). The formation of aggregates provides protection from environmental stresses and increases epiphytic survival. However about 70% of the overall space on the leaf-surface, and up to 85% in the interveinal regions, remains uncolonized (Andrews, 2006).

Potential sources of colonists for the phyllosphere are the soil, seed, insect vectors, local vegetation and plant debris, air, precipitation, or buds and overwintering shoots (Lindemann *et al.*, 1982; Andrews, 1992; Kinkel, 1997; Morris *et al.*, 2008; Bodenhausen *et al.*, 2013; Bulgarelli *et al.*, 2013; Kembel *et al.*, 2014; Zarraonaindia *et al.*, 2015). Microbial population dynamics on leaves is a function of four population processes: immigration, emigration, growth, and death (Lindow & Andersen, 1996; Upper *et al.*, 2003; Woody *et al.*, 2007; Whipps *et al.*, 2008; Maignien *et al.*, 2014). The importance of these processes varies during the growing season of the

plant and is strongly influenced by the environmental conditions - wind, rainfall, solar radiation, temperature, and moisture levels (Kinkel, 1997; Pietrarelli & Balestra, 2006; Zimmerman & Vitousek, 2012; Lau *et al.*, 2013). For instance intense rains can lead to large and rapid decreases in leaf surface bacterial populations by washing off the bacterial cells and depositing them on the ground (Hirano & Upper, 2000). The bacterial and fungal communities vary throughout the growing season (Lindow & Andersen, 1996; Redford & Fierer, 2009; Jumpponen & Jones, 2010; Peñuelas *et al.*, 2012; Rastogi *et al.*, 2012). For the crop plants, the bacterial communities at the beginning of the growing season more strongly resemble soil communities, but as the season progresses, they differentiate and become more leaf-specific (Copeland *et al.*, 2015). Whereas in a greenhouse experiment the bacterial communities of *Arabidopsis thaliana* phyllosphere initially resembled airborne communities before converging to leaf-specific communities (Maignien *et al.*, 2014). The findings of Woody *et al.* (2007) for resident and immigrant populations of the yeast-like fungus *Aureobasidium pullulans* (*Ap*) suggest strong competition from resident *Ap* populations against immigrating *Ap*, whereas immigrants have little impact on residents. In addition they suggest that the *Ap* dynamics are dominated by density-dependent processes and resource limitation. While the immigration of *Ap* during the early stages of leaf development is essential, it likely plays a minor role during the later stages of leaf development (Woody *et al.*, 2007). Another study showed that resident bacterial aggregates of *Pantoea agglomerans* facilitated survival of immigrant bacterial cells of *P. agglomerans* itself and of *Pseudomonas fluorescens* (Monier & Lindow, 2005). As can be seen the microbial population already resident in the phyllosphere is likely to influence the ability of incoming organisms to successfully colonize the leaf surface (Hunter *et al.*, 2010; Venturi & da Silva, 2012). Antagonistic interactions between phyllosphere microorganisms are of special interest for biological control of pathogens (Kiss, 2003; Sylla *et al.*, 2013).

Microbiota functions in the phyllosphere

The relationship between phyllosphere microbiota and their host ranges from mutualistic to parasitic (Friesen *et al.*, 2011; Mitter *et al.*, 2013; Vandenkoornhuyse *et al.*, 2015). Phyllosphere microbiota mediates plant functional traits either by altering the existing plant pathways or by providing novel biochemical capabilities (Friesen *et al.*, 2011). It impacts nutrient and energy fluxes, e.g. many bacterial taxa found in the phyllosphere are able to fix atmospheric nitrogen (Pati & Chandra, 1993; Müller *et al.*, 2003; Peñuelas *et al.*, 2012; Kembel *et al.*, 2014). Other common bacteria in the phyllosphere - methylotrophs (Mizuno *et al.*, 2012) - can utilize methane and/or methanol as sole carbon and energy sources, methanol being a byproduct of cell wall-associated pectin metabolism in growing plant cells (Galbally & Kirstine, 2002; Delmotte *et al.*, 2009; Iguchi *et al.*, 2015). Furthermore phyllosphere microbiota can improve stress tolerance of their host to biotic and abiotic stresses. For example bacterial and fungal endophytes procure increased drought tolerance (Hamilton &

Bauerle, 2012; Naveed *et al.*, 2014). Other microorganisms increase plant pathogen resistance (Arnold *et al.*, 2003; Kiss, 2003; Podolich *et al.*, 2015) through competition (Innerebner *et al.*, 2011), antibiosis (Stockwell *et al.*, 2002), mycoparasitism (Ownley *et al.*, 2010), or through modulation of plant immunity, which induces systemic resistance primed for accelerated activation (Bakker *et al.*, 2007; Vandenkoornhuys *et al.*, 2015). Furthermore microbiota mediates the outcome of plant-insect interactions (Omacini *et al.*, 2001) by killing (Lacey *et al.*, 2009) or deterring herbivores (Crawford *et al.*, 2010), attracting herbivores (Ponzio *et al.*, 2013), impacting their life-cycle (Laine, 2004), or attracting predators and parasitoids (van Nouhuys & Laine, 2008; Ponzio *et al.*, 2013).

1.3 Plants strategy to cope with the diversity of biotic interactions might select for specific networks of interacting species in the associated communities

Plants have evolved sophisticated mechanisms to deal with biotic interactions and to respond with defensive or conducive compounds either to combat pathogens and herbivorous insects or to accommodate beneficial organisms (Kessler & Baldwin, 2002; Dicke & Hilker, 2003; Chisholm *et al.*, 2006; Jones & Dangl, 2006; Pierik *et al.*, 2014). The success of any response depends on its fitness consequences and needs to be viewed within a community context (Whitham *et al.*, 2008; Dicke, 2009).

Multi-species interactions form complex networks of interactions. These interactions are of different types: direct or indirect, ranging from antagonistic to beneficial. Leaf-associated species interact directly through predator-prey, host-parasite, mutualistic, facilitative or competitive interactions. The nature of the interaction between a given pair of species is dynamic, depending both on the biotic and the abiotic environment. For instance, changes in biotic and abiotic conditions can tip the balance away from mutualistic exchange towards exploitative outcomes (Kiers & Van Der Heijden, 2006). In the case of interacting species associated with plant leaves, the abiotic environment corresponds to the physical and chemical properties of plant leaves and we therefore expect direct interactions to be controlled by plant traits, which are themselves controlled by plant genotype. Species associated with plant leaves also interact indirectly, the interaction being mediated by the plant (Stout *et al.*, 2006). For instance, plants respond to herbivore damage by phenotypic changes in allelochemistry, physiology, morphology, and phenology (Ohgushi, 2005; Ponzio *et al.*, 2013) and these changes may impact other species living on plants (Dicke & Hilker, 2003; Kniskern *et al.*, 2007). These changes therefore represent the mechanistic basis of trait-mediated indirect interactions among species (Utsumi & Ohgushi, 2008).



Figure 1.6: *Quercus* leaves expressing symptoms of *Erysiphe alphitoides* infection.
© Malcolm Storey, 2011

Plants traits are heritable, also are interactions among species if they depend on these traits. We are thus expecting a genetic control of the network of interactions occurring on the plant foliage. The idea has been suggested recently for insects by Utsumi *et al.* (2010), and extended to microbes by Miki & Jacquet (2010).

1.4 Powdery mildew, a putative agent of phyllosphere community change

Powdery mildews (Ascomycota, Erysiphales) are some of the world's most frequently encountered plant pathogenic fungi. They are often conspicuous owing to the profuse production of conidia (Figure 1.6 and Figure 1.7 on the next page) that give them their common name (Glawe, 2008). Powdery mildew is a very common disease of pedunculate (*Quercus robur*) and sessile (*Quercus petraea*) oaks in Western Europe. This disease can be very severe, particularly in seedlings growing in naturally regenerating forest, decreasing the growth of the plant and causing its death in some cases (Feau *et al.*, 2012). Molecular data for ITS sequences has suggested that at least four genetic lineages cause oak powdery mildew disease in European forests: *Erysiphe alphitoides*, *E. hypophylla*, *E. quercicola*, and *Phyllactinia guttata* (Mougou *et al.*, 2008; Desprez-Loustau *et al.*, 2010).

An infection is initiated when an ascospore or conidium lands on a susceptible host, germinates, and forms a germ tube that elongates to form a hypha with appressoria, penetration pegs, and haustoria. Appressoria are short, lateral hyphal outgrowths or swellings that produce penetration pegs to infect host cells (Figure 1.7,

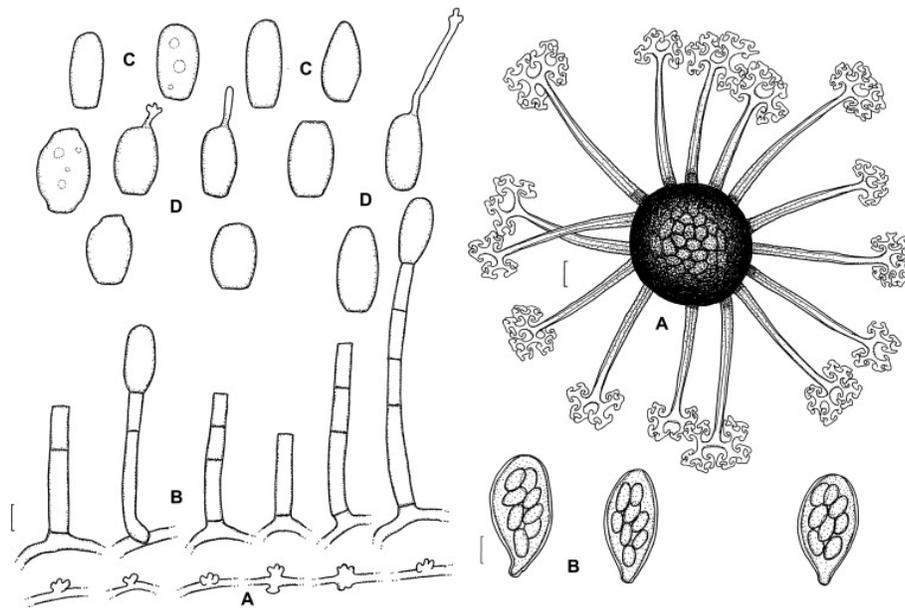


Figure 1.7: Left panel - *Erysiphe alphitoides*, anamorph. (A) Appressoria. (B) Conidiophores. (C) Primary conidia. (D) Secondary conidia (partly with germ tubes). Bar = 10 μm . Right panel - *E. alphitoides* teleomorph. (A) Chasmothecium. (B) Asci. Bars = 25 μm (chasmothecium), 10 μm (asci). U. Braun (Takamatsu *et al.*, 2007)

(A), left panel). Penetration pegs are narrow protrusions produced from appressoria that penetrate the walls of host cells by means of turgor pressure and enzymatic activity (Glawe, 2008). If the fungus is successful in breaching the host wall, the penetration peg extends into the host cell, invaginating the cytoplasm, and enlarging to form a feeding structure, the haustorium. The haustorium is surrounded by extrahaustorial matrix, a noncytoplasmic, gel-like material that may help protect the fungus from host responses. The haustorium, extrahaustorial matrix, and host cell are the sites of molecular signaling events. Host responses and nutrient transfer determine whether a successful parasitic relationship develops and is maintained (Glawe, 2008).

E. alphitoides can be found on a wide range of oak species, including *Quercus dentata*, *Q. crispula*, *Q. petraea*, *Q. robur*, *Q. serrata*, Fagaceae (occasionally attacking *Aesculus hippocastanum*, Hippocastanaceae; *Mangifera indica*, Anacardiaceae; *Paeonia lutea*, Paeoniaceae). Its distribution stretches from Asia (Armenia, Azerbaijan, China, Georgia, India, Iran, Iraq, Israel, Japan, Kazakhstan, Korea, Lebanon, Nepal, Russia, Sri Lanka, Turkey, Turkmenistan), over North Africa (Ethiopia, Morocco) to all Europe, and it was introduced in South Africa, Australia, New Zealand, North and South America (Takamatsu *et al.*, 2007). Oak powdery mildew fungi are widely thought to survive from one growing season to the next by generating chasmothecia, the sexual fruiting bodies, or by mycelia infecting shoot primordia under bud scales

(Feau *et al.*, 2012).

In general phytopathogens may have strong impact on the dynamics of the multitrophic plant-based food web. They impact herbivore performance, preference and population dynamics (Laine, 2004; Stout *et al.*, 2006), they modify the interaction between the herbivore and its carnivorous natural enemies and consecutively the insect population structure (Biere *et al.*, 2002; Cardoza *et al.*, 2003; van Nouhuys & Laine, 2008). Plant pathogens drive plant density, diversity and composition (Tack & Dicke, 2013; Bagchi *et al.*, 2014). *E. alphitoides* infection in particular negatively influences the diversity and number of gall-forming wasps (Zargaran *et al.*, 2012). Furthermore it modifies the insect community structure as different herbivores respond differently to the pathogen presence on oak (Tack *et al.*, 2012). In turn the insect defoliation increases the risk of subsequent severe powdery-mildew infection (Marçais & Desprez-Loustau, 2012), which in synergy can lead to death of the host plant. The prevalence of powdery-mildew infection on oaks, its impact on the oak-associated community, high variability in susceptibility to it among the individual trees, and the known genetic variability for resistance to it in oak (Robin *et al.*, 2010; Mougou-Hamdane, 2009) make *E. alphitoides* a suitable species to link oak genetics with the oak-associated community in the sense of the community genetics framework.

1.5 Metabarcoding for studying microbial community

The study of microbial communities is boosting with the technological advances of Next-Generation Sequencing (NGS) techniques (Di Bella *et al.*, 2013; Knief, 2014), which provide lowered cost experimental tools without the cloning process inherent in conventional capillary-based methods. Up until ten years ago, the vast majority of microbial communities' descriptions were realized with culture-dependent methods, which largely under-estimate species richness as many species cannot be grown in culture (Christen, 2008). The recent revolution in DNA sequencing brought about by NGS techniques allows a more thorough description of microbial communities and we are now, in principle, able to detect any microbial organism in nature, even those that cannot be isolated or grown in the lab (Horton & Bruns, 2001; Anderson & Cairney, 2004; Hamelin, 2006; Cardenas & Tiedje, 2008; Peay *et al.*, 2008). In the case of the phyllosphere, culture-independent approaches have shown that although assumptions regarding the dominant inhabitants were largely correct, the diversity of phyllosphere communities is far greater than previously recognized (Whipps *et al.*, 2008).

A very powerful technique for species identification in microbiology is the screening for DNA barcodes with NGS techniques. A DNA barcode, in its simplest definition, is one or few relatively short gene sequences taken from a standardized portion of the genome and used to identify species. (Kress *et al.*, 2015; Schlaeppi & Bulgar-

elli, 2015). Different barcodes are necessary to provide adequate species identification across lineages.

Fungi

The most popular DNA barcode for mycological studies is the internal transcribed spacer (ITS) region of the nuclear ribosomal repeat unit (Schoch *et al.*, 2012). Rapidly evolving ITS1 and ITS2 subloci are thereby widely used for fungal species identification (Figure 1.8). ITS1 and ITS2 yield similar results when used as DNA metabarcodes for fungi (Blaalid *et al.*, 2013). However, the variability of ITS1, at least on average, exceeds that of ITS2 (Nilsson *et al.*, 2008).

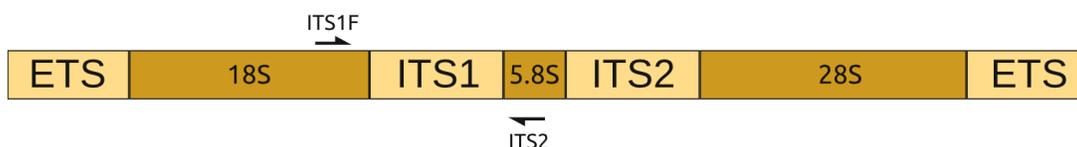


Figure 1.8: Overview over the fungal ITS region. ETS - external transcribed spacer; ITS - internal transcribed spacer.

Bacteria

The 16S small ribosomal subunit gene (16S rRNA) has been widely used to study and characterize bacterial community compositions (Hamady & Knight, 2009). The 16S rRNA gene contains conserved regions as well as nine hypervariable regions (V1-V9) (Figure 1.9). Sequencing of one or more of these hypervariable regions and comparing them to large bacterial databases can be effectively used to distinguish between taxa (Youssef *et al.*, 2009).

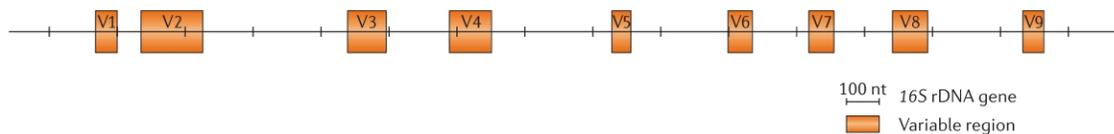


Figure 1.9: Overview over the bacterial 16S rDNA gene. V1, V2, ... - hypervariable regions (after Kuczynski *et al.*, 2011)

1.6 Objectives of this work

The first objective of the dissertation at hand is the description of the bacterial and fungal communities of the phyllosphere microbiota of English oak (*Quercus robur* L.), and the study of species association network with a specific focus on *Erysiphe alphitoides*, the causing agent of powdery-mildew disease. The phyllosphere environment is described as a stressful environment. Will this kind of environment promote positive or negative interactions among microbial members of the phyllosphere? And how will a pathogen effect the interaction network? Will it team up with the residents to successfully infect the leaf or will it compete with the residents for the resources? In Chapter 2 we tackle the first objective and present our findings in the manuscript, which is currently in revision for resubmission to Microbial Ecology.

The second objective is the determination of the genetic basis explaining the structural features of the oak phyllosphere microbiota, as well as its association network properties. Are there genes, whose effects extend to the community level and influence the structure and interactions of the phyllosphere microbiota? In Chapter 3 we deal with this objective and present the results of the QTL search for traits describing the properties of the phyllosphere microbiota and its association network.

The third objective is not only to confirm but also to find alternative loci in oak genome explaining the phyllosphere microbiota structure and its interaction network properties. To achieve this objective we perform an association study. In Chapter 4 we present our preliminary analysis of the association study.

Finally we develop an evolutionary perspective on our results in Chapter 5.

1.7 Bibliography

- Agrawal, A.A. (2003). Community genetics: new insights into community ecology by integrating population genetics. *Ecology*, 84, 543–544.
- Anderson, I.C. & Cairney, J.W.G. (2004). Diversity and ecology of soil fungal communities: Increased understanding through the application of molecular techniques. *Environmental Microbiology*, 6, 769–779.
- Andrews, J.H. (1992). Biological control in the phyllosphere. *Annual Review of Phytopathology*, 30, 603–35.
- Andrews, J.H. (2006). Population Growth and the Landscape Ecology of Microbes on Leaf Surfaces. In: *Microbial ecology of aerial plant surfaces*. pp. 239–250.
- Antonovics, J. (1992). Toward community genetics. In: *Plant Resistance to Herbivores and Pathogens: Ecology, Evolution, and Genetics* (eds. Fritz, R. & Simms, E.). University of Chicago Press, Chicago, IL., pp. 426–449.
- Arnold, A.E., Gilbert, G.S., Coley, P.D. & Thomas, A. (2000). Are tropical fungal endophytes hyperdiverse? *Ecology letters*, 3, 267–274.
- Arnold, A.E., Mejía, L.C., Kylo, D., Rojas, E.I., Maynard, Z., Robbins, N. & Herre, E.A. (2003). Fungal endophytes limit pathogen damage in a tropical tree. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 15649–15654.
- Atamna-Ismaeel, N., Finkel, O.M., Glaser, F., Sharon, I., Schneider, R., Post, A.F., Spudich, J.L., von Mering, C., Vorholt, J.a., Iluz, D., Béjà, O. & Belkin, S. (2012). Microbial rhodopsins on leaf surfaces of terrestrial plants. *Environmental microbiology*, 14, 140–6.
- Bagchi, R., Gallery, R.E., Gripenberg, S., Gurr, S.J., Narayan, L., Addis, C.E., Freckleton, R.P. & Lewis, O.T. (2014). Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature*, 506, 85–8.
- Bailey, J.K., Deckert, R., Schweitzer, J.A., Rehill, B.J., Lindroth, R.L., Gehring, C. & Whitham, T.G. (2005). Host plant genetics affect hidden ecological players: links among *Populus*, condensed tannins, and fungal endophyte infection. *Canadian Journal of Botany*, 83, 356–361.
- Bailey, J.K., Schweitzer, J.a., Ubeda, F., Koricheva, J., LeRoy, C.J., Madritch, M.D., Rehill, B.J., Bangert, R.K., Fischer, D.G., Allan, G.J. & Whitham, T.G. (2009). From genes to ecosystems: a synthesis of the effects of plant genetic factors across levels of organization. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364, 1607–1616.

- Bailey, J.K., Wooley, S.C., Lindroth, R.L. & Whitham, T.G. (2006). Importance of species interactions to community heritability: a genetic basis to trophic-level interactions. *Ecology letters*, 9, 78–85.
- Bakker, P.a.H.M., Pieterse, C.M.J. & van Loon, L.C. (2007). Induced Systemic Resistance by Fluorescent *Pseudomonas* spp. *Phytopathology*, 97, 239–243.
- Bálint, M., Bartha, L., O’Hara, R.B., Olson, M.S., Otte, J., Pfenninger, M., Robertson, A.L., Tiffin, P. & Schmitt, I. (2015). Relocation, high-latitude warming and host genetic identity shape the foliar fungal microbiome of poplars. *Molecular Ecology*, 24, 235–248.
- Bálint, M., Tiffin, P., Hallström, B., O’Hara, R.B., Olson, M.S., Fankhauser, J.D., Piepenbring, M. & Schmitt, I. (2013). Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). *PLoS one*, 8, e53987.
- Bartram, A.K., Lynch, M.D.J., Stearns, J.C., Moreno-Hagelsieb, G. & Neufeld, J.D. (2011). Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Applied and environmental microbiology*, 77, 3846–52.
- Biere, A., Elzinga, J.a., Honders, S.C. & Harvey, J.a. (2002). A plant pathogen reduces the enemy-free space of an insect herbivore on a shared host plant. *Proceedings. Biological sciences / The Royal Society*, 269, 2197–2204.
- Blaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M. & Kauserud, H. (2013). ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular ecology resources*, 13, 218–24.
- Bodenhause, N., Bortfeld-Miller, M., Ackermann, M. & Vorholt, J.a. (2014). A Synthetic Community Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota. *PLoS Genetics*, 10, e1004283.
- Bodenhause, N., Horton, M.W. & Bergelson, J. (2013). Bacterial Communities Associated with the Leaves and the Roots of *Arabidopsis thaliana*. *PLoS ONE*, 8, e56329.
- Bokulich, N.a., Thorngate, J.H., Richardson, P.M. & Mills, D.a. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proceedings of the National Academy of Sciences of the United States of America*, 111, E139–48.
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., Ver Loren van Themaat, E. & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology*, 64, 807–38.

- Cardenas, E. & Tiedje, J.M. (2008). New tools for discovering and characterizing microbial diversity. *Current Opinion in Biotechnology*, 19, 544–549.
- Cardoza, Y.J., Teal, P.E.a. & Tumlinson, J.H. (2003). Effect of Peanut Plant Fungal Infection on Oviposition Preference by *Spodoptera exigua* and on Host-Searching Behavior by *Cotesia marginiventris*. *Environmental Entomology*, 32, 970–976.
- Chisholm, S.T., Coaker, G., Day, B. & Staskawicz, B.J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124, 803–14.
- Christen, R. (2008). Global Sequencing: A Review of Current Molecular Data and New Methods Available to Assess Microbial Diversity. *Microbes and Environments*, 23, 253–268.
- Copeland, J.K., Yuan, L., Layeghifard, M., Wang, P.W. & Guttman, D.S. (2015). Seasonal Community Succession of the Phyllosphere Microbiome. *Molecular Plant-Microbe Interactions*, 28, 274–285.
- Cordier, T., Robin, C., Capdevielle, X., Desprez-Loustau, M.L. & Vacher, C. (2012a). Spatial variability of phyllosphere fungal assemblages: genetic distance predominates over geographic distance in a European beech stand (*Fagus sylvatica*). *Fungal Ecology*, 5, 509–520.
- Cordier, T., Robin, C., Capdevielle, X., Fabreguettes, O., Desprez-Loustau, M.L. & Vacher, C. (2012b). The composition of phyllosphere fungal assemblages of European beech (*Fagus sylvatica*) varies significantly along an elevation gradient. *New Phytologist*, 196, 510–9.
- Crawford, K.M., Land, J.M. & Rudgers, J.a. (2010). Fungal endophytes of native grasses decrease insect herbivore preference and performance. *Oecologia*, 164, 431–44.
- Crutsinger, G.M., Collins, M.D., Fordyce, J.a., Gompert, Z., Nice, C.C. & Sanders, N.J. (2006). Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science (New York, N.Y.)*, 313, 966–968.
- Crutsinger, G.M., Strauss, S.Y. & Rudgers, J.E.R.a. (2010). Genetic variation within a dominant shrub species determines plant species colonization in a coastal dune ecosystem. *Ecology*, 91, 1237–1243.
- Danhorn, T. & Fuqua, C. (2007). Biofilm formation by plant-associated bacteria. *Annual review of microbiology*, 61, 401–22.
- Dawkins, R. (1982). *The Extended Phenotype*. Oxford University Press, Oxford.

- Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C. & Vorholt, J.a. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 16428–16433.
- Derridj, S. (1996). Nutrients on the leaf surface. In: *Aerial Plant Surface Microbiology*. pp. 25–42.
- Desprez-Loustau, M.L., Feau, N., Mougou-Hamdane, A. & Dutech, C. (2010). Interspecific and intraspecific diversity in oak powdery mildews in Europe: coevolution history and adaptation to their hosts. *Mycoscience*, 52, 165–173.
- Di Bella, J.M., Bao, Y., Gloor, G.B., Burton, J.P. & Reid, G. (2013). High throughput sequencing methods and analysis for microbiome research. *Journal of microbiological methods*, 95, 401–414.
- Dicke, M. (2009). Behavioural and community ecology of plants that cry for help. *Plant, Cell and Environment*, 32, 654–665.
- Dicke, M. & Hilker, M. (2003). Induced plant defences: from molecular biology to evolutionary ecology. *Basic and Applied Ecology*, 14, 3–14.
- Douanla-Meli, C., Langer, E. & Talontsi Mouafo, F. (2013). Fungal endophyte diversity and community patterns in healthy and yellowing leaves of Citrus limon. *Fungal Ecology*, 6, 212–222.
- Elamo, P., Helander, M.L., Saloniemi, I. & Neuvonen, S. (1999). Birch family and environmental conditions affect endophytic fungi in leaves. *Oecologia*, 118, 151–156.
- Feau, N., Lauron-Moreau, A., Piou, D., Marçais, B., Dutech, C. & Desprez-Loustau, M.L. (2012). Niche partitioning of the genetic lineages of the oak powdery mildew complex. *Fungal Ecology*, 5, 154–162.
- Fernández, V., Sancho-Knapik, D., Guzmán, P., Peguero-Pina, J.J., Gil, L., Karabourniotis, G., Khayet, M., Fasseas, C., Heredia-Guerrero, J.A., Heredia, A. & Gil-Pelegri, E. (2014). Wettability, polarity and water absorption of Quercus ilex leaves: effect of leaf side and age. *Plant physiology*, 166, 168–180.
- Finkel, O.M., Burch, A.Y., Lindow, S.E., Post, A.F. & Belkin, S. (2011). Geographical location determines the population structure in phyllosphere microbial communities of a salt-excreting desert tree. *Applied and environmental microbiology*, 77, 7647–7655.
- Fischer, D.G., Chapman, S.K., Classen, a.T., Gehring, C.a., Grady, K.C., Schweitzer, J.a. & Whitham, T.G. (2014). Plant genetic effects on soils under climate change. *Plant and Soil*, 379, 1–19.

- Friesen, M.L., Porter, S.S., Stark, S.C., von Wettberg, E.J., Sachs, J.L. & Martinez-Romero, E. (2011). Microbially Mediated Plant Functional Traits. *Annual Review of Ecology, Evolution, and Systematics*, 42, 23–46.
- Fritz, R.S. & Price, P.W. (1988). Genetic Variation Among Plants and Insect Community Structure: Willows and Sawflies. *Ecology*, 69, 845.
- Galbally, I.E. & Kirstine, W. (2002). The production of methanol by flowering plants and the global cycle of methanol. *Journal of Atmospheric Chemistry*, 43, 195–229.
- Glawe, D.A. (2008). The powdery mildews: a review of the world’s most familiar (yet poorly known) plant pathogens. *Annual Review of Phytopathology*, 46, 27–51.
- Gossner, M.M., Brändle, M., Brandl, R., Bail, J., Müller, J. & Opgenoorth, L. (2015). Where Is the Extended Phenotype in the Wild? The Community Composition of Arthropods on Mature Oak Trees Does Not Depend on the Oak Genotype. *PLOS ONE*, 10, e0115733.
- Hamady, M. & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome research*, 19, 1141–52.
- Hamelin, R.C. (2006). Molecular epidemiology of forest pathogens: from genes to landscape. *Canadian Journal of Plant Pathology*, 28, 167–181.
- Hamilton, C.E. & Bauerle, T.L. (2012). A new currency for mutualism? Fungal endophytes alter antioxidant activity in hosts responding to drought. *Fungal Diversity*, 54, 39–49.
- Handelsman, J. (2004). Metagenomics : Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews*, 68, 669–685.
- Hersch-Green, E.I., Turley, N.E. & Johnson, M.T.J. (2011). Community genetics: what have we accomplished and where should we be going? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366, 1453–1460.
- Hirano, S.S. & Upper, C.D. (2000). Bacteria in the Leaf Ecosystem with Emphasis on *Pseudomonas syringae* - a Pathogen, Ice Nucleus, and Epiphyte. *Microbiology and Molecular Biology Reviews*, 64, 624–653.
- Hoeksema, J.D. & Classen, A.T. (2012). Is plant genetic control of ectomycorrhizal fungal communities an untapped source of stable soil carbon in managed forests? *Plant and Soil*, 359, 197–204.
- Horton, T.R. & Bruns, T.D. (2001). The molecular revolution in ectomycorrhizal ecology: peeking into the black-box. *Molecular ecology*, 10, 1855–1871.
- Hughes, R.A., Inouye, B.D., Johnson, M.T.J., Underwood, N. & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, 11, 609–623.

- Hunter, P.J., Hand, P., Pink, D., Whipps, J.M. & Bending, G.D. (2010). Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca* species) phyllosphere. *Applied and Environmental Microbiology*, 76, 8117–8125.
- Hunter, P.J., a.C. Pink, D. & Bending, G.D. (2015). Cultivar-level genotype differences influence diversity and composition of lettuce (*Lactuca* sp.) phyllosphere fungal communities. *Fungal Ecology*, pp. 8–11.
- Iason, G.R., Lennon, J.J., Pakeman, R.J., Thoss, V., Beaton, J.K., Sim, D.a. & Elston, D.a. (2005). Does chemical composition of individual Scots pine trees determine the biodiversity of their associated ground vegetation? *Ecology Letters*, 8, 364–369.
- Iguchi, H., Yurimoto, H. & Sakai, Y. (2015). Interactions of Methyloprophs with Plants and Other Heterotrophic Bacteria. *Microorganisms*, 3, 137–151.
- Innerebner, G., Knief, C. & Vorholt, J.a. (2011). Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. *Applied and environmental microbiology*, 77, 3202–10.
- Jacobs, J.L., Carroll, T.L. & Sundin, G.W. (2005). The role of pigmentation, ultraviolet radiation tolerance, and leaf colonization strategies in the epiphytic survival of phyllosphere bacteria. *Microbial Ecology*, 49, 104–113.
- Johnson, M.T.J. (2008). Bottom-up effects of plant genotype on aphids, ants, and predators. *Ecology*, 89, 145–154.
- Johnson, M.T.J. & Agrawal, A.A. (2005). Plant genotype and environment interact to shape a diverse arthropod community on Evening Primrose (*Oenothera biennis*). *Ecology*, 86, 874–885.
- Johnson, M.T.J. & Stinchcombe, J.R. (2007). An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution*, 22, 250–257.
- Jones, J.D.G. & Dangl, J.L. (2006). The plant immune system. *Nature*, 444, 323–329.
- Jumpponen, A. & Jones, K.L. (2009). Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist*, 184, 438–48.
- Jumpponen, A. & Jones, K.L. (2010). Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. *New Phytologist*, 186, 496–513.
- Junker, R.R. & Tholl, D. (2013). Volatile organic compound mediated interactions at the plant-microbe interface. *Journal of chemical ecology*, 39, 810–25.

- Kanaga, M.K., Latta, L.C., Mock, K.E., Ryel, R.J., Lindroth, R.L. & Pfrender, M.E. (2009). Plant genotypic diversity and environmental stress interact to negatively affect arthropod community diversity. *Arthropod-Plant Interactions*, 3, 249–258.
- Kembel, S.W., O’Connor, T.K., Arnold, H.K., Hubbell, S.P., Wright, S.J. & Green, J.L. (2014). Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences*, 111, 13715–13720.
- Kerling, L. (1958). De microflora of het blad van *Beta vulgaris*. *European Journal of Plant Pathology*, 64, 402–410.
- Kessler, A. & Baldwin, I.T. (2002). Plant responses to insect herbivory: the emerging molecular analysis. *Annual review of plant biology*, 53, 299–328.
- Kiers, E.T. & Van Der Heijden, M.G.A. (2006). Mutualistic stability in the arbuscular mycorrhizal symbiosis: Exploring hypotheses of evolutionary cooperation. *Ecology*, 87, 1627–1636.
- Kinkel, L.L. (1997). Microbial population dynamics on leaves. *Annual review of phytopathology*, 35, 327–47.
- Kiss, L. (2003). A review of fungal antagonists of powdery mildews and their potential as biocontrol agents. *Pest Management Science*, 59, 475–483.
- Knief, C. (2014). Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Frontiers in plant science*, 5, 216.
- Kniskern, J.M., Traw, M.B. & Bergelson, J. (2007). Salicylic acid and jasmonic acid signaling defense pathways reduce natural bacterial diversity on *Arabidopsis thaliana*. *Molecular plant-microbe interactions : MPMI*, 20, 1512–1522.
- Koch, K., Bhushan, B. & Barthlott, W. (2009). Multifunctional surface structures of plants: An inspiration for biomimetics. *Progress in Materials Science*, 54, 137–178.
- Kress, W.J., Garcia-Robledo, C., Uriarte, M. & Erickson, D.L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution*, 30, 25–35.
- Kuczynski, J., Lauber, C.L., Walters, W.a., Parfrey, L.W., Clemente, J.C., Gevers, D. & Knight, R. (2011). Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics*, 13, 47–58.
- Lacey, L.A., Horton, D.R., Jones, D.C., Headrick, H.L. & Neven, L.G. (2009). Efficacy of the biofumigant fungus *Muscodor albus* (Ascomycota: Xylariales) for control of codling moth (Lepidoptera: Tortricidae) in simulated storage conditions. *Journal of economic entomology*, 102, 43–49.

- Laine, A.I. (2004). A powdery mildew infection on a shared host plant affects the dynamics of the Glanville fritillary butterfly populations. *Oikos*, 2, 329–337.
- Lambais, M.R., Crowley, D.E., Cury, J.C., Büll, R.C. & Rodrigues, R.R. (2006). Bacterial diversity in tree canopies of the Atlantic forest. *Science (New York, N.Y.)*, 312, 1917.
- Lamit, L.J., Lau, M.K., Næsborg, R.R., Wojtowicz, T., Whitham, T.G. & Gehring, C.A. (2015). Genotype variation in bark texture drives lichen community assembly across multiple environments. *Ecology*, 96, 960–971.
- Lamit, L.J., Lau, M.K., Sthultz, C.M., Wooley, S.C., Whitham, T.G. & Gehring, C.a. (2014). Tree genotype and genetically based growth traits structure twig endophyte communities. *American journal of botany*, 101, 467–78.
- Lamit, L.J., Wojtowicz, T., Kovacs, Z., Wooley, S.C., Zinkgraf, M., Whitham, T.G., Lindroth, R.L. & Gehring, C.a. (2011). Hybridization among foundation tree species influences the structure of associated understory plant communities. *Botany*, 89, 165–174.
- Lankau, R.a. (2011). Intraspecific variation in allelochemistry determines an invasive species' impact on soil microbial communities. *Oecologia*, 165, 453–463.
- Lankau, R.a. & Strauss, S.Y. (2007). Mutual feedbacks maintain both genetic and species diversity in a plant community. *Science (New York, N.Y.)*, 317, 1561–1563.
- Lau, M.K., Arnold, a.E. & Johnson, N.C. (2013). Factors influencing communities of foliar fungal endophytes in riparian woody plants. *Fungal Ecology*, 6, 365–378.
- Leveau, J.H. & Lindow, S.E. (2001). Appetite of an epiphyte: quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 3446–53.
- Lindemann, J., Constantinidou, H.a., Barchet, W.R. & Upper, C.D. (1982). Plants as sources of airborne bacteria, including ice nucleation-active bacteria. *Applied and Environmental Microbiology*, 44, 1059–1063.
- Lindow, S. (2006). Phyllosphere Microbiology: A Perspective. In: *Microbial ecology of aerial plant surfaces* (eds. Bailey, M., Lilley, A., Timms-Wilson, T. & Spencer-Phillips, P.). CABI; Wallingford; UK, pp. 1–20.
- Lindow, S.E. & Andersen, G.L. (1996). Influence of immigration on epiphytic bacterial populations on navel orange leaves. *Applied and Environmental Microbiology*, 62, 2978–2987.
- Lindow, S.E. & Brandl, M.T. (2003). Microbiology of the Phyllosphere. *Applied and Environmental Microbiology*, 69, 1875–1883.

- Liu, Z., Lozupone, C., Hamady, M., Bushman, F.D. & Knight, R. (2007). Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic acids research*, 35, e120.
- Maddox, G.D. & Root, R.B. (1987). Resistance to 16 diverse species of herbivorous insects within a population of goldenrod, *Solidago altissima*: genetic variation and heritability. *Oecologia*, 72, 8–14.
- Madritch, M., Donaldson, J.R. & Lindroth, R.L. (2006). Genetic identity of *Populus tremuloides* litter influences decomposition and nutrient release in a mixed forest stand. *Ecosystems*, 9, 528–537.
- Madritch, M.D. & Lindroth, R.L. (2011). Soil microbial communities adapt to genetic variation in leaf litter inputs. *Oikos*, 120, 1696–1704.
- Maignien, L., DeForce, E.A., Chafee, M.E., Eren, A.M. & Simmons, S.L. (2014). Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *mBio*, 5, 10.
- Marçais, B. & Desprez-Loustau, M.L. (2012). European oak powdery mildew: impact on trees, effects of environmental factors, and potential effects of climate change. *Annals of Forest Science*, 71, 633–642.
- Mechaber, W.L., Marshall, D.B., Mechaber, R.a., Jobe, R.T. & Chew, F.S. (1996). Mapping leaf surface landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 4600–4603.
- Meiser, A., Bálint, M. & Schmitt, I. (2013). Meta-analysis of deep-sequenced fungal communities indicates limited taxon sharing between studies and the presence of biogeographic patterns. *New Phytologist*, pp. 623–635.
- Melotto, M., Underwood, W., Koczan, J., Nomura, K. & He, S.Y. (2006). Plant Stomata Function in Innate Immunity against Bacterial Invasion. *Cell*, 126, 969–980.
- Meneses, N., Bailey, J.K., Allan, G.J., Bangert, R.K., Bowker, M.a., Rehill, B.J., Wimp, G.M., Lindroth, R.L. & Whitham, T.G. (2012). Arthropod Community Similarity in Clonal Stands of Aspen: A Test of the Genetic Similarity Rule. *Ecoscience*, 19, 48–58.
- Mercier, J. & Lindow, S.E. (2000). Role of Leaf Surface Sugars in Colonization of Plants by Bacterial Epiphytes. *Applied and Environmental Microbiology*, 66, 369–374.
- Miki, T. & Jacquet, S. (2010). Indirect interactions in the microbial world: Specificities and similarities to plant-insect systems. *Population Ecology*, 52, 475–483.

- Mitter, B., Brader, G., Afzal, M., Compant, S., Naveed, M., Trognitz, F. & Sessitsch, A. (2013). Advances in Elucidating Beneficial Interactions Between Plants, Soil, and Bacteria. *Advances in Agronomy*, 121, 381–445.
- Mizuno, M., Yurimoto, H., Yoshida, N., Iguchi, H. & Sakai, Y. (2012). Distribution of Pink-Pigmented Facultative Methylophilic Bacteria on Leaves of Vegetables. *Bioscience, Biotechnology, and Biochemistry*, 76, 578–580.
- Monier, J.M. & Lindow, S.E. (2004). Frequency, Size, and Localization of Bacterial Aggregates on Bean Leaf Surfaces. *Applied and Environmental Microbiology*, 70, 346–355.
- Monier, J.M. & Lindow, S.E. (2005). Aggregates of resident bacteria facilitate survival of immigrant bacteria on leaf surfaces. *Microbial ecology*, 49, 343–352.
- Morris, C.E. (2001). Phyllosphere. In: *eLS*. pp. 1–8.
- Morris, C.E. & Kinkel, L.L. (2002). Fifty years of phyllosphere microbiology: significant contributions to research in related fields. In: *Phyllosphere microbiology* (eds. Lindow, S.E., Hecht-Poinar, E.I. & Elliot, V.J.). American Phytopathological Society (APS Press), St. Paul, pp. 365–375.
- Morris, C.E., Sands, D.C., Vinatzer, B.a., Glaux, C., Guilbaud, C., Buffière, A., Yan, S., Dominguez, H. & Thompson, B.M. (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *The ISME journal*, 2, 321–334.
- Mougou, A., Dutech, C. & Desprez-Loustau, M.L. (2008). New insights into the identity and origin of the causal agent of oak powdery mildew in Europe. *Forest Pathology*, 38, 275–287.
- Mougou-Hamdane, A. (2009). *Interaction Chene-oidium : caracterisation moleculaire et adaptation locale du parasite, resistance genetique de l'hote*. Ph.D. thesis, Bordeaux 1.
- Müller, T., Müller, M., Behrendt, U. & Stadler, B. (2003). Diversity of culturable phyllosphere bacteria on beech and oak: the effects of lepidopterous larvae. *Microbiological Research*, 158, 291–297.
- Naveed, M., Mitter, B., Reichenauer, T.G., Wieczorek, K. & Sessitsch, A. (2014). Increased drought stress resilience of maize through endophytic colonization by *Burkholderia phytofirmans* PsJN and *Enterobacter* sp. FD17. *Environmental and Experimental Botany*, 97, 30–39.
- Neuhauser, C., Andow, D.A., Heimpel, G.E., May, G., Shaw, R.G. & Wagenius, S. (2003). COMMUNITY GENETICS: EXPANDING THE SYNTHESIS OF ECOLOGY AND GENETICS. *Ecology*, 84, 545–558.

- Nilsson, R.H., Kristiansson, E., Ryberg, M. & Hallenberg, N. (2008). Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evolutionary Bioinformatics*, 4, 193–201.
- van Nouhuys, S. & Laine, A.L. (2008). Population dynamics and sex ratio of a parasitoid altered by fungal-infected diet of host butterfly. *Proceedings. Biological sciences / The Royal Society*, 275, 787–795.
- Oerke, E.C., Steiner, U., Dehne, H.W. & Lindenthal, M. (2006). Thermal imaging of cucumber leaves affected by downy mildew and environmental conditions. *Journal of Experimental Botany*, 57, 2121–2132.
- Ohgushi, T. (2005). Indirect Interaction Webs: Herbivore-Induced Effects Through Trait Change in Plants. *Annual Review of Ecology, Evolution, and Systematics*, 36, 81–105.
- Omacini, M., Chaneton, E.J., Ghersa, C.M. & Müller, C.B. (2001). Symbiotic fungal endophytes control insect host-parasite interaction webs. *Nature*, 409, 78–81.
- Ownley, B.H., Gwinn, K.D. & Vega, F.E. (2010). Endophytic fungal entomopathogens with activity against plant pathogens: Ecology and evolution. *Bio-Control*, 55, 113–128.
- Pakeman, R.J., Beaton, J.K., Thoss, V., Lennon, J.J., Campbell, C.D., White, D. & Iason, G.R. (2006). The extended phenotype of Scots pine *Pinus sylvestris* structures the understorey assemblage. *Ecography*, 29, 451–457.
- Partida-Martínez, L.P. & Heil, M. (2011). The microbe-free plant: fact or artifact? *Frontiers in plant science*, 2, 100.
- Pastor, A., Compson, Z.G., Dijkstra, P., Riera, J.L., Martí, E., Sabater, F., Hungate, B.a. & Marks, J.C. (2014). Stream carbon and nitrogen supplements during leaf litter decomposition: contrasting patterns for two foundation species. *Oecologia*, 176, 1111–1121.
- Pati, B. & Chandra, A. (1993). Diazotrophic bacterial population and other associated organisms on the phyllosphere of some crop plants. *Zentralblatt fur Mikrobiologie*, 148, 392–402.
- Peay, K.G., Kennedy, P.G. & Bruns, T.D. (2008). Fungal Community Ecology: A Hybrid Beast with a Molecular Master. *BioScience*, 58, 799.
- Peñuelas, J., Rico, L., Ogaya, R., Jump, a.S. & Terradas, J. (2012). Summer season and long-term drought increase the richness of bacteria and fungi in the foliar phyllosphere of *Quercus ilex* in a mixed Mediterranean forest. *Plant biology (Stuttgart, Germany)*, 14, 565–75.

- Peñuelas, J. & Terradas, J. (2014). The foliar microbiome. *Trends in Plant Science*, 19, 278–280.
- Pierik, R., Ballaré, C.L. & Dicke, M. (2014). Ecology of plant volatiles: Taking a plant community perspective. *Plant, Cell and Environment*, 37, 1845–1853.
- Pietrarelli, L. & Balestra, G.M. (2006). Effects of simulated rain on *Pseudomonas syringae* pv. tomato populations on tomato plants. *Journal of Plant Pathology*, 88, 245–251.
- Pincebourde, S. & Woods, H.A. (2012). Climate uncertainty on leaf surfaces: The biophysics of leaf microclimates and their consequences for leaf-dwelling organisms. *Functional Ecology*, 26, 844–853.
- Podolich, O., Ardanov, P., Zaets, I., Pirttilä, A.M. & Kozyrovska, N. (2015). Reviving of the endophytic bacterial community as a putative mechanism of plant resistance. *Plant Soil*, pp. 367–377.
- Ponzio, C., Gols, R., Pieterse, C.M.J. & Dicke, M. (2013). Ecological and phytohormonal aspects of plant volatile emission in response to single and dual infestations with herbivores and phytopathogens. *Functional Ecology*, 27, 587–598.
- Proffitt, C.E., Chiasson, R.L., Owens, A.B., Edwards, K.R. & Travis, S.E. (2005). *Spartina alterniflora* genotype influences facilitation and suppression of high marsh species colonizing an early successional salt marsh. *Journal of Ecology*, 93, 404–416.
- Rastogi, G., Coaker, G.L. & Leveau, J.H.J. (2013). New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS microbiology letters*.
- Rastogi, G., Sbodio, A., Tech, J.J., Suslow, T.V., Coaker, G.L. & Leveau, J.H.J. (2012). Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *The ISME journal*, 6, 1812–22.
- Redford, A.J., Bowers, R.M., Knight, R., Linhart, Y. & Fierer, N. (2010). The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental Microbiology*, 12, 2885–2893.
- Redford, A.J. & Fierer, N. (2009). Bacterial succession on the leaf surface: a novel system for studying successional dynamics. *Microbial ecology*, 58, 189–98.
- Reisberg, E.E., Hildebrandt, U., Riederer, M. & Hentschel, U. (2013). Distinct Phyllosphere Bacterial Communities on Arabidopsis Wax Mutant Leaves. *PLoS ONE*, 8, e78613.
- Robin, C., Mougou-Hamdane, A., Gion, J.M., Kremer, A. & Desprez-Loustau, M.L. (2010). Quantitative trait loci for resistance to pathogens in pedunculate oak (*Quercus robur* L.). In: *Forest ecosystem genomics and adaptation*. Madrid.

- Rowntree, J.K., Shuker, D.M. & Preziosi, R.F. (2011). Forward from the crossroads of ecology and evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366, 1322–1328.
- Sapkota, R., Knorr, K., Jørgensen, L.N., Hanlon, K.a.O. & Nicolaisen, M. (2015). Host genotype is an important determinant of the cereal phyllosphere mycobiome. *New Phytologist*.
- Schlaeppli, K. & Bulgarelli, D. (2015). The Plant Microbiome at Work. *MPMI*, 28, 212–217.
- Schoch, C.L., Seifert, K.a., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.a., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, a.N., Wingfield, M.J., Aime, M.C., An, K.D., Bai, F.Y., Barreto, R.W., Begerow, D., Bergeron, M.J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, a.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, a., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.a., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.a., Ge, Z.W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.B., Hosaka, K., Houbraken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., Kovacs, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, a.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, a.R., Methven, a.S., Meyer, W., Moncalvo, J.M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.a., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, a., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.O., Sung, G.H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.a., Urbina, H., Vagvolgyi, C., Vialle, a., Vu, T.D., Walther, G., Wang, Q.M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, a., Zamora, J.C., Zhang, N., Zhuang, W.Y. & Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109, 6241–6246.
- Schreiber, L., Krimm, U., Knoll, D., Sayed, M., Auling, G. & Kroppenstedt, R.M. (2005). Plant-microbe interactions: identification of epiphytic bacteria and their ability to alter leaf surface permeability. *New Phytologist*, 166, 589–594.

- Schweitzer, J.A., Bailey, J.K., Bangert, R.K., Hart, S.C. & Whitham, T.G. (2006). The role of plant genetics in determining above- and below-ground microbial communities. In: *Microbial ecology of aerial plant surfaces* (eds. Bailey, M.J., Lilley, A.K., Timms-Wilson, T.M. & Spencer-Phillips, P.T.N.). CABI; Wallingford; UK, pp. 107–119.
- Schweitzer, J.A., Bailey, J.K., Fischer, D.G., LeRoy, C.J., Lonsdorf, E.V., Whitham, T.G. & Hart, S.C. (2008a). Plant-soil-microorganism interactions: heritable relationship between plant genotype and associated soil microorganisms. *Ecology*, 89, 773–781.
- Schweitzer, J.a., Bailey, J.K., Rehill, B.J., Martinsen, G.D., Hart, S.C., Lindroth, R.L., Keim, P. & Whitham, T.G. (2004). Genetically based trait in a dominant tree affects ecosystem processes. *Ecology Letters*, 7, 127–134.
- Schweitzer, J.a., Fischer, D.G., Rehill, B.J., Wooley, S.C., Woolbright, S.a., Lindroth, R.L., Whitham, T.G., Zak, D.R. & Hart, S.C. (2010). Forest gene diversity is correlated with the composition and function of soil microbial communities. *Population Ecology*, 53, 35–46.
- Schweitzer, J.a., Madritch, M.D., Bailey, J.K., LeRoy, C.J., Fischer, D.G., Rehill, B.J., Lindroth, R.L., Hagerman, A.E., Wooley, S.C., Hart, S.C. & Whitham, T.G. (2008b). From Genes to Ecosystems: The Genetic Basis of Condensed Tannins and Their Role in Nutrient Regulation in a *Populus* Model System. *Ecosystems*, 11, 1005–1020.
- Shuster, S.M., Lonsdorf, E.V., Wimp, G.M., Bailey, J.K. & Whitham, T.G. (2006). Community heritability measures the evolutionary consequences of indirect genetic effects on community structure. *Evolution*, 60, 991.
- Stiling, P. & Rossi, A.M. (1996). Complex Effects of Genotype and Environment on Insect Herbivores and Their Enemies. *Ecology*, 77, 2212.
- Stockwell, V.O., Johnson, K.B., Sugar, D. & Loper, J.E. (2002). Antibiosis Contributes to Biological Control of Fire Blight by *Pantoea agglomerans* Strain Eh252 in Orchards. *Phytopathology*, 92, 1202–1209.
- Stout, M.J., Thaler, J.S. & Thomma, B.P.H.J. (2006). Plant-mediated interactions between pathogenic microorganisms and herbivorous arthropods. *Annual Review of Entomology*, 51, 663–89.
- Sundin, G. (2002). Ultraviolet radiation on leaves: Its influence on microbial communities and their adaptations. In: *Phyllosphere microbiology* (eds. Lindow, S., Hecht-Poinar, E. & Elliott, V.). American Phytopathological Society (APS Press), pp. 27–41.

- Sylla, J., Alsanus, B.W., Krüger, E., Becker, D. & Wohanka, W. (2013). In vitro compatibility of microbial agents for simultaneous application to control strawberry powdery mildew (*Podosphaera aphanis*). *Crop Protection*, 51, 40–47.
- Tack, A.J.M. & Dicke, M. (2013). Plant pathogens structure arthropod communities across multiple spatial and temporal scales. *Functional Ecology*, 27, 633–645.
- Tack, A.J.M., Gripenberg, S. & Roslin, T. (2012). Cross-kingdom interactions matter: fungal-mediated interactions structure an insect community on oak. *Ecology Letters*, 15, 177–185.
- Takamatsu, S., Braun, U., Limkaisang, S., Kom-Un, S., Sato, Y. & Cunnington, J.H. (2007). Phylogeny and taxonomy of the oak powdery mildew *Erysiphe alphitoides sensu lato*. *Mycological Research*, 111, 809–826.
- Upper, C.D., Hirano, S.S., Dodd, K.K. & Clayton, M.K. (2003). Factors that Affect Spread of *Pseudomonas syringae* in the Phyllosphere. *Phytopathology*, 93, 1082–1092.
- Utsumi, S., Ando, Y. & Miki, T. (2010). Linkages among trait-mediated indirect effects: A new framework for the indirect interaction web. *Population Ecology*, 52, 485–497.
- Utsumi, S. & Ohgushi, T. (2008). Host plant variation in plant-mediated indirect effects: moth boring-induced susceptibility of willows to a specialist leaf beetle. *Ecological Entomology*, 33, 250–260.
- Valkama, E., Salminen, J.P., Koricheva, J. & Pihlaja, K. (2004). Changes in leaf trichomes and epicuticular flavonoids during leaf development in three birch taxa. *Annals of Botany*, 94, 233–242.
- Van Der Wal, A. & Leveau, J.H.J. (2011). Modelling sugar diffusion across plant leaf cuticles: The effect of free water on substrate availability to phyllosphere bacteria. *Environmental Microbiology*, 13, 792–797.
- Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. (2015). Tansley review The importance of the microbiome of the plant holobiont. *New Phytologist*.
- Velmala, S.M., Rajala, T., Haapanen, M., Taylor, a.F.S. & Pennanen, T. (2013). Genetic host-tree effects on the ectomycorrhizal community and root characteristics of Norway spruce. *Mycorrhiza*, 23, 21–33.
- Venturi, V. & da Silva, D.P. (2012). Incoming pathogens team up with harmless 'resident' bacteria. *Trends in Microbiology*, 20, 160–164.
- Vorholt, J.A. (2012). Microbial life in the phyllosphere. *Nature Reviews Microbiology*, 10, 828–840.

- Whipps, J.M., Hand, P., Pink, D. & Bending, G.D. (2008). Phyllosphere microbiology with special reference to diversity and plant genotype. *Journal of applied microbiology*, 105, 1744–1755.
- Whitham, T.G., Bailey, J.K., Schweitzer, J.a., Shuster, S.M., Bangert, R.K., LeRoy, C.J., Lonsdorf, E.V., Allan, G.J., DiFazio, S.P., Potts, B.M., Fischer, D.G., Gehring, C.a., Lindroth, R.L., Marks, J.C., Hart, S.C., Wimp, G.M. & Wooley, S.C. (2006). A framework for community and ecosystem genetics: from genes to ecosystems. *Nature reviews. Genetics*, 7, 510–523.
- Whitham, T.G., Difazio, S.P., Schweitzer, J.a., Shuster, S.M., Allan, G.J., Bailey, J.K. & Woolbright, S.a. (2008). Extending genomics to natural communities and ecosystems. *Science (New York, N.Y.)*, 320, 492–495.
- Whitham, T.G., Gehring, C.a., Lamit, L.J., Wojtowicz, T., Evans, L.M., Keith, A.R. & Smith, D.S. (2012). Community specificity: Life and afterlife effects of genes. *Trends in Plant Science*, 17, 271–281.
- Whitham, T.G., Young, W.P., Martinsen, G.D., Gehring, C.a., Schweitzer, J.a., Shuster, S.M., Wimp, G.M., Fischer, D.G., Bailey, J.K., Lindroth, R.L., Woolbright, S. & Kuske, C.R. (2003). Community and ecosystem genetics: A consequence of the extended phenotype. *Ecology*, 84, 559–573.
- Wimp, G.M., Young, W.P., Woolbright, S.a., Martinsen, G.D., Keim, P. & Whitham, T.G. (2004). Conserving plant genetic diversity for dependent animal communities. *Ecology Letters*, 7, 776–780.
- Wojtowicz, T., Compson, Z.G., Lamit, L.J., Whitham, T.G. & Gehring, C.A. (2014). Plant genetic identity of foundation tree species and their hybrids affects a litter-dwelling generalist predator. *Oecologia*, 176, 799–810.
- Woody, S.T., Ives, A.R., Nordheim, E.V. & Andrews, J.H. (2007). Dispersal, density dependence, and population dynamics of a fungal microbe on leaf surfaces. *Ecology*, 88, 1513–1524.
- Yadav, R.K.P., Karamanoli, K. & Vokou, D. (2005). Bacterial Colonization of the Phyllosphere of Mediterranean Perennial Species as Influenced by Leaf Structural and Chemical Features. *Microbial Ecology*, 50, 185–196.
- Yang, C.H., Crowley, D.E., Borneman, J. & Keen, N.T. (2001). Microbial phyllosphere populations are more complex than previously realized. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 3889–3894.
- Yeats, T.H. & Rose, J.K.C. (2013). The formation and function of plant cuticles. *Plant physiology*, 163, 5–20.

- Youssef, N., Sheik, C.S., Krumholz, L.R., Najar, F.Z., Roe, B.a. & Elshahed, M.S. (2009). Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Applied and environmental microbiology*, 75, 5227–36.
- Zargaran, M.R., Erbilgin, N. & Ghosta, Y. (2012). Changes in Oak Gall Wasps Species Diversity (Hymenoptera: Cynipidae) in Relation to the Presence of Oak Powdery Mildew (*Erysiphe alphitoides*). *Zoological Studies*, 51, 175–184.
- Zarraonaindia, I., Owens, S.M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax, S., Bokulich, N.a., Mills, D.a., Martin, G., Taghavi, S., van der Lelie, D. & Gilbert, J.A. (2015). The soil microbiome influences grapevine-associated microbiota. *mBio*, 6, 1–10.
- Zimmerman, N.B. & Vitousek, P.M. (2012). Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 13022–7.
- Zytynska, S.E., Fay, M.F., Penney, D. & Preziosi, R.F. (2011). Genetic variation in a tropical tree species influences the associated epiphytic plant and invertebrate communities in a complex forest ecosystem. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366, 1329–1336.

Chapter 2

Deciphering the pathobiome: intra- and inter-kingdom interactions involving the pathogen *Erysiphe alphitoides*.

RUNNING TITLE: Deciphering the pathobiome

Boris Jakuschkin^{1,2}, Virgil Fievet^{1,2}, Loïc Schwaller^{3,4}, Sarah Ouadah^{3,4}, Stéphane Robin^{3,4}, Cécile Robin^{1,2}, Corinne Vacher^{1,2,*}

1. INRA, UMR1202 BIOGECO, F-33610 Cestas, France
2. Univ. Bordeaux, BIOGECO, UMR 1202, F-33615 Pessac, France
3. AgroParisTech, UMR 518 MIA, F-75005 Paris, France
4. INRA, UMR 518 MIA, F-75005 Paris, France

* Author for correspondence: Corinne Vacher, INRA, UMR1202 BIOGECO, Domaine de l'Hermitage, 69 route d'Arcachon, F-33612 Cestas Cedex, France; corinne.vacher@pierroton.inra.fr; Phone: +33 (0)5 57 12 27 24; Fax: +33 (0)5 57 12 28 81

Abstract

Plant-inhabiting microorganisms interact directly with each other, forming complex microbial interaction networks. These networks may be disrupted by the arrival of a new species, such as a pathogen infecting the plant. Here we decipher intra-kingdom and cross-kingdom interactions involving a pathogen species, by using a Bayesian method of network inference. We use this method to highlight the most likely interactions between *Erysiphe alphitoides*, the

causal agent of oak powdery mildew, and other foliar microorganisms of pedunculate oak (*Quercus robur* L.). Our results showed that infection by *E. alphitoides* is accompanied by drastic changes in the foliar fungal community composition but no significant change in the bacterial community composition. Our analyses highlighted 13 fungal Operational Taxonomic Units (OTUs) and 13 bacterial OTUs that are likely to interact directly with *E. alphitoides*. Four of the fungal OTUs were phyllosphere yeasts. The fungal endophytes *Mycosphaerella punctiformis* and *Monochaetia kansensis* were highlighted as potential antagonists of *E. alphitoides*. These potential interactions will have to be validated experimentally. The study of the temporal dynamics of microbial networks during the course of infection also appears to be a promising research avenue, that will undoubtedly give deeper insights into the ecology of this disease. Overall, we showed that combining metagenomics and network ecology may improve biological control of plant diseases, by highlighting potential antagonists of pathogen species.

Keywords: community ecology, microbial diversity, ecological network, network inference, high-throughput sequencing, pathobiome, powdery mildew, *Quercus*.

2.1 Introduction

Almost all plant tissues host a microbial community (Turner *et al.*, 2013). These microbes may interact with each other directly, through pairwise ecological interactions such as predation, parasitism or mutualism (Faust & Raes, 2012). These interactions may involve micro-organisms belonging to the same kingdom, or different kingdoms (Frey-Klett *et al.*, 2011). These microbes may also interact indirectly through shared resources or shared environment (Frey-Klett *et al.*, 2011; Little *et al.*, 2008; Kemen, 2014). Some of them modulate the physicochemical conditions within plant tissues and impact other plant-associated organisms. Such interactions, called plant-mediated indirect interactions (Stout *et al.*, 2006), can involve fungal pathogens. The latter have been shown to activate or deregulate plant immune responses, thereby influencing other plant-associated organisms, including herbivorous arthropods (Stout *et al.*, 2006; Tack *et al.*, 2012; Tack & Dicke, 2013) and the resident plant microbiota (Rovenich *et al.*, 2014). Our aim here is to decipher the complex relationships between a fungal pathogen and the resident microbial community, by combining high-throughput sequencing techniques (Di Bella *et al.*, 2013) and network inference methods (Faust & Raes, 2012).

The notion that the pathogens live and interact with other microorganisms has led to the development of the “pathobiome” concept, which represents the pathogenic agent integrated within its biotic environment (Vayssier-Taussat *et al.*, 2014). Elucidating the components of the pathobiome is one of the prerequisites for understanding the pathogenesis, persistence, transmission and evolution of pathogenic agents (Vayssier-Taussat *et al.*, 2014). Pathobiome studies will promote renewed

approaches of plant protection paralleling those currently developed to study human diseases (Thiele *et al.*, 2013). The role of commensal flora in animal health is indeed well recognized and its manipulation to treat human diseases is a highly promising and competitive field of current medical research (Clemente *et al.*, 2012; Ursell *et al.*, 2013). Similar approaches should be considered to improve plant health (Berlec, 2012; Newton *et al.*, 2010).

Here we study the pathobiome of the fungus *Erysiphe alphitoides*, the causal agent of oak powdery mildew (Mougou *et al.*, 2008; Mougou-Hamdane *et al.*, 2010). This fungus, which covers oak leaves with spots or patches of white mycelium and spores, is nowadays causing one of the most common diseases in European forests (Mougou *et al.*, 2008). Like most powdery mildews, *E. alphitoides* grows superficially on leaves, sometimes covering up to 80% of the upper leaf surface, whilst deriving resources from leaf cells with haustoria (Glawe, 2008). *E. alphitoides* interacts directly with certain foliar microorganisms, such as the mycoparasites belonging to the *Ampelomyces* genus (Kiss, 2003), and indirectly with other microorganisms by altering the foliar traits. Oak infection by *E. alphitoides* indeed changes leaf chemistry and physiology (Hajji *et al.*, 2009; Copolovici *et al.*, 2014). It reduces net CO₂ assimilation, alters stomatal conductance, increases leaf nitrogen content and alters the composition of volatile emissions. Moreover, infection is suspected to increase leaf tannin content (Zargaran *et al.*, 2012). In the case of severe infection, leaf life-span is drastically reduced (Hajji *et al.*, 2009).

To better understand the relationships between *E. alphitoides* and other foliar micro-organisms, we thoroughly analyzed the foliar fungal and bacterial communities of three oak trees (*Quercus robur* L.) with different levels of susceptibility to the pathogen. We assessed the composition of the entire foliar microbial community, including both epiphytes and endophytes. We then addressed the following questions: (1) Does the number of sequences assigned to *E. alphitoides* in high-throughput sequencing datasets reflect the abundance of symptoms on the leaves? (2) Does infection by *E. alphitoides* is accompanied by changes in the composition of the foliar microbial community? (3) Which microbial species are the most likely to interact directly with *E. alphitoides*?

2.2 Material and methods

Study site and sampling design

We sampled three trees which belonged to a full-sib family of *Quercus robur* L. planted for a field experiment established in 2000 at INRA-Bourran, located in the South-West of France (44°19'43" N, 0°25'26" W) (Scotti-Saintagne *et al.*, 2004). These trees were previously characterized as highly susceptible, intermediately resistant or strongly resistant to powdery mildew based on five years of visual observations (Desprez-Loustau M-L, personal communication). They were located within

5 m from one another and were thus expected to experience similar environmental conditions. On the 29th of July 2011, we collected 10 leaves each from four branches at approximate height of 2 m, resulting in 40 leaves per tree. The leaves taken were evenly spaced between the trunk and the tip of the branch. We collected leaves strictly from the first flush. We measured the powdery mildew infection level as the proportion of the upper leaf area displaying the symptoms. We also measured the position of each leaf within the canopy: its distance to the ground, to the base of the branch, to the tree trunk and the orientation of the branch (South-West versus North-East). Each leaf was then cut at the base of the petiole and placed into a sterile bag containing silica gel. The bags, sealed and made airtight, were then stored in dark at +18°C in the laboratory. A preliminary experiment had been performed to assess the effect of oak leaf storage in silica gel on the fungal community composition inferred by 454 pyrosequencing. We did not find any significant difference between the communities of fresh leaves and leaves stored in silicagel (Methods S1 2.7 on page 68).

Leaf DNA extraction

Leaf DNA extraction was performed three months after the sampling date. We prepared the leaf material for the DNA extraction in a laminar flow hood. In order to avoid contaminating the samples, the hood and all tools were disinfected with sodium hypochlorite followed by 70% ethanol and 40 min of UV exposure. Plastic ware and tungsten carbide beads were autoclaved for 2 h prior to utilization. Similarly, the tools used for leaf handling were sterilized prior to each use, by dipping them into 70% alcohol and flaming over an electric bunsen burner. Using a metal corer we excised two discs of 0.5 cm 2 each on both sides of the middle leaf vein. The four discs were evenly distributed on the leaf, with their position not being chosen according to *E. alphitoides* symptoms. The four leaf discs were placed in a collection microtube (QIAGEN DNeasy[®] 96 Plant Kit) containing one tungsten carbide bead. We used Geno/Grinder[®] (SPEX SamplePrep, Metuchen, NJ, USA) to homogenize the leaf material at 1 600 strokes per minute for 90 s. Homogenized samples were stored at -80 °C overnight. We used QIAGEN DNeasy 96 Plant Kit to extract the DNA (following the protocol for frozen plant tissue). We determined the DNA concentration with NanoDrop 8000 and diluted all DNA working solutions to 1 ng μl^{-1} .

Amplification and sequencing of the fungal assemblages

We chose the rapidly evolving internal transcribed spacer 1 (ITS1) as the DNA metabarcoding marker for fungi (Schoch *et al.*, 2012). We used titanium fusion primers for the PCR amplification as follows. The forward primers were composed of the 454 A-adaptor, a sample-specific 6-bp tag, and the ITS1F primer (Gardes & Bruns, 1993). The reverse primer was composed of the B-adaptor and the ITS2 primer (White *et al.*, 1990) (Tab. S1). The PCR reactions (50 μl) contained 2.5

Titanium fusion adaptors:

adaptor A:	CGTATCGCCTCCCTCGCGCCATCAG
adaptor B:	CTATGCGCCTTGCCAGCCCGCTCAG

ITS1 primers:

ITS2 (forward in PCR):	GCTGCGTTCTTCATCGATGC
ITS1F (reverse in PCR):	CTTGGTCATTTAGAGGAAGTAA

Multiplex tags:

ACTCAG	ACAGTA	TGCTAG	ATGCAT
AGCTCT	CGAGAG	CGACTA	CACATC
TAGTCT	GTCTGT	CTGCGC	GATGTA
TGTGCA	ATACGA	AGTACG	TACTGC
GCTACT	ATCGCA	ACGAGA	GACACA
TCATCA	TGCGTC	TATCGT	TCTGTG
TGCAGT	CTCAGA	GCAGAC	GCGTGC
TCACAT	CACTCG	ATGTCG	GCACTG
GATCAC	TAGATA	CTGATG	CGATGT
CGTAGC	GTAGCG	CAGTAC	TACGAT

Concatemer scheme:

5'- 454 adaptor A + multiplex tag + forward ITS2 -3'
5'- 454 adaptor B + reverse ITS1F -3'

Table S1: 454 pyrosequencing primer sequences (forward) used for the amplification of fungal rITS1 and concatemer scheme

units Silverstar Taq polymerase (Eurogentec), 5 μ l 10x PCR buffer supplied by the manufacturer, 2 μ l MgCl₂ (50 mM), 2 μ l dNTPs (5 mM), 1 μ l of each primer (10 μ M), 0.5 μ l of 100 mg ml⁻¹ BSA, and 5 ng DNA template. The PCR reaction was carried out as follows: an initial phase at 95 °C (5 m), followed by 35 cycles at 95 °C (30 s), 54 °C (60 s) and 72 °C (90 s), and a final step at 72 °C (10 m). Samples were amplified in triplicate and pooled before purification with AMPure magnetic beads (Beckman Coulter Genomics, Danvers, MA, USA). After the DNA concentration measurement was performed using a NanoDrop 8000, the samples were pooled in equimolar ratios to obtain the amplicon libraries. We used 40 tags in total, hence we obtained one amplicon library per tree. The libraries were sequenced in separate 1/8 th runs on 454 GS Junior sequencer (454 Life Sciences, Branford, CT, USA) at the *Centre de Génomique Fonctionnelle de Bordeaux* (CGFB, Bordeaux, France). Sequencing was unidirectional and started with the A adaptor. The 454 sff files are available from the European Nucleotide Archive (<http://www.ebi.ac.uk/ena/data/view/PRJEB7319>).

Amplification and sequencing of the bacterial assemblages

We applied the microbiome profiling method developed by Gloor *et al.* (2010), which uses combinatorial sequence tags attached to PCR primers amplifying the

HVR6 primers:				
left primer (forward primer):	CAACGCGARGAACCTTACC			
right primer (reverse primer):	ACAACACGAGCTGACGAC			
<hr/>				
Multiplex tags:				
left tags:	CATGCG	GACTGT	ACGTA	AGTA
	GCAGT	CGTCGA	CACTAC	ATGA
	TAGCT	GTCGC	TGAC	TGCA
right tags:	ACTGC	GCGAC	TACT	CGA
	AGCTA	TACGT	TCAT	TAC
	ACAGTC	GTAGTG	TGCA	
	TCGACG	GTCA	AGT	
<hr/>				
Concatemer scheme:				
	5'- left tag + left primer -3'			
	5'- right tag + right primer -3'			

Table S2: Illumina primer sequences used for the amplification of bacterial HVR6 and concatemer scheme

hypervariable V6 region of the 16S rRNA gene (Huse *et al.*, 2008). This type of tagging combined with the Illumina paired-end protocol enabled for the examination of hundreds of samples with far fewer primers (Tab. S2) than is required when tags are incorporated at only a single end. The PCR mix (50 μ l) contained 2.5 units Silverstar Taq polymerase (Eurogentec), 5 μ l 10x PCR buffer supplied by the manufacturer, 2 μ l MgCl₂ (50 mM), 2 μ l dNTPs (5 mM), 1 μ l of 10 μ M forward primer, 1 μ l of 10 μ M reverse primer, 0.5 μ l of 100 mg ml⁻¹ BSA, and 5 ng DNA template. The PCR was carried out as follows: an initial phase at 95 °C (2 m), followed by 30 cycles at 95 °C (30 s), 56 °C (30 s) and 72 °C (30 s), and a final extension at 72 °C (5 m). Samples were amplified in triplicate and pooled before purification with AMPure magnetic beads (Beckman Coulter Genomics, Danvers, MA, USA). We quantified the PCR products using the Nanodrop 8000 and pooled them at equimolar ratios. The final DNA library was sequenced on the Illumina platform GAIIx at the Centre de Génomique Fonctionnelle de Bordeaux (CGFB, Bordeaux, France). The fastq files are available from the European Nucleotide Archive (<http://www.ebi.ac.uk/ena/data/view/PRJEB7319>).

Processing of the fungal 454 pyrosequencing data

We used QIIME v.1.7.0 (Caporaso *et al.*, 2010) for quality filtering, demultiplexing, and denoising. After the removal of tags and primers, all sequences shorter than 100 nucleotides were discarded. Quality filtering was done with a sliding window test of quality scores having a window size of 50 nucleotides and a minimum quality score of 25. The sequences were truncated at the beginning of the poor quality window and tested for the minimum sequence length of 100 nucleotides. Be-

cause the highly conserved ribosomal genes flanking the ITS1 marker may distort sequence clustering and similarity searches, we removed these from the dataset using Fungal ITS Extractor 1.1 (Nilsson *et al.*, 2010). ITS1 sequences shorter than 100 nucleotides were discarded and sequences longer than 280 nucleotides were trimmed at the 3'-end. Before the clustering step we combined the dataset with a larger, unpublished 454 fungal dataset acquired by sampling leaves on 203 oak trees planted in the same field test. In doing so we aimed to reduce the number of singletons. We discarded singleton reads before clustering to minimize spurious Operational Taxonomic Units (OTUs). OTU clustering was done with UPARSE-OTU (Edgar, 2013), an algorithm implemented within Usearch v.7 (Edgar, 2010) that performs chimera filtering and OTU clustering simultaneously. The algorithm identifies a set of OTU representative sequences at the 97% sequence similarity threshold, that corresponds roughly to the differences at the species level (Nilsson *et al.*, 2008). Unfortunately, other *Erysiphe* species, such as *E. quercicola* and *E. hypophylla*, have an ITS sequence which has less than a 3% difference from the *E. alphitoides* ITS sequence (Mougou *et al.*, 2008). ITS sequences of these species therefore cluster together at 97% similarity threshold. To circumvent this issue we performed a BLAST search of all sequences against a custom taxonomic database of the *Erysiphe* genus. We extracted all sequences with successful hits to *E. alphitoides* into a separate file, and assigned the corresponding read counts. Taxonomy assignments were done by BLAST searches against Fungal ITS Database (Nilsson *et al.*, 2009) (as of March 15, 2012). We modified the taxonomic file associated with this database to equalize the length of taxonomic ranks, by using the NCBI Taxonomy Browser. Additionally we corrected a few false taxonomic assignments in the case of *Erysiphe* species and removed a few non-fungal sequences. We used this database to assign the taxonomy to the most abundant sequence of each OTU, using QIIME (v.1.7.0) with standard settings (BLAST searches, e-value: 0.001, minimum percent identity: 90). To exclude possible *Q. robur* sequences, we also compared the most abundant sequence of each OTU with the sequences deposited in GenBank, using the BLASTN algorithm (Altschul, 1990). We then constructed the OTU table in QIIME and transformed it from BIOME into tab delimited text format for downstream analyses with R (R Core Team, 2013).

Processing of the bacterial Illumina data

We assembled the paired-end reads using PANDAseq (Masella *et al.*, 2012). Reads were then demultiplexed following the method described in Gloor *et al.* (2010) using a custom Bash shell script. We removed oak chloroplast sequence prior to the clustering step. OTU clustering and OTU table construction were done by following the same procedure than for fungi. Taxonomic assignments were performed with the RDP Classifier program (v.2.5) (Wang *et al.*, 2007) as it is implemented into QIIME (v.1.7.0), with standard settings and Greengenes 13_5 as the reference sequences. The origin of OTU representative sequences (archaeal, bacterial, nuclear eukaryote,

mitochondrial or chloroplast origin) was identified using Metaxa (Bengtsson *et al.*, 2011) and we retained only the bacterial OTUs.

Statistical analyses

All statistical analyses were done in R (R Core Team, 2013). We first investigated whether the number of reads assigned to *E. alphitoides* reflected the level of infection estimated by visual inspection of foliar symptoms. The effect of tree on the percentage of symptomatic upper leaf area (arc-sine transformed) was tested, using a generalized linear model (GLM) with a Gaussian distribution. A second model was built to test differences in number of *E. alphitoides* reads amongst trees. To control for overdispersion, we used negative binomial generalized linear model (NBGLM) with the total number of reads per sample (log-transformed) as an offset. In the case of the susceptible tree, we also examined the relationship between the percentage of symptomatic upper leaf area and the number of *E. alphitoides* reads, by using NBGLM.

We then investigated the relationship between infection by *E. alphitoides* and the composition of foliar microbial communities, by using data collected for the susceptible tree. We used principle component analysis (PCA) on rarefied datasets to visualize variations in microbial community composition between leaves. The rarefaction threshold was 1226 reads for fungal data and 114 reads for bacterial data. The reads assigned to *E. alphitoides* were not included in the fungal dataset because they were used as an explanatory variable for variations in fungal community composition. To assess the robustness of our results to rarefaction, we performed 100 rarefactions, followed by 100 PCAs. We calculated the mean and the standard error of the mean of the coordinates of the samples on two first axes of the PCA. The mean coordinates (named comp1 and comp2) were used as proxies for microbial community composition. The effect of the percentage of *E. alphitoides* reads on these two variables was tested with a GLM with a Gaussian distribution. Moreover, we examined the associations between microbial OTUs and leaf type (highly infected by *E. alphitoides* versus "uninfected") by using the abundance-based version of the Pearson's phi coefficient of association. A leaf was considered as highly infected if the percentage of reads assigned to *E. alphitoides* was higher than 5%. Associations were calculated for all OTUs representing more than 0.5% of the sequences for the tree susceptible to *E. alphitoides*.

We finally highlighted the most likely interactions between *E. alphitoides* and other foliar microorganisms, by using a Bayesian method for network inference (Schwaller *et al.*, 2015) implemented in the R package saturnin (Schwaller, 2015). Network inference aimed at disentangling the processes underlying the patterns of co-association and co-exclusion between all pairs of OTUs. These processes may be (i) direct ecological interactions between OTUs (predation, parasitism, mutualism, commensalism or amensalism; see Faust & Raes 2012), (ii) shared environmental requirements (e.g. two OTUs are associated because they are both adapted to shade

leaves) and (iii) shared interacting partners (e.g. two OTUs are associated because they are both parasites of a third OTU). Here we constructed microbial networks having only (i) as edges. We constructed one microbial network per tree. Each network had OTUs as nodes and putative direct interactions between OTUs as edges. Network construction was based on a subset of the most abundant OTUs, composed of 48 bacterial OTUs and 66 fungal OTUs (including *Erysiphe alphitoides*). These OTUs were the same for the three trees. Each of them represented at least 0.5% of the sequences for at least one of the three trees. The model developed by Schwaller *et al.* (2015) aims at separating (i) from associations due to (iii). The model indeed considers a class of acyclic undirected networks (called spanning trees) as possible structures of dependency between the nodes. These structures only include direct dependencies. Environmental variables cannot be included in the model as covariates. Therefore, to remove associations due to (ii), we fitted the number of reads of each OTU using NBGLMs with environmental variables as predictors and the total number of reads per sample (log-transformed) as an offset. GLMs with a Poisson distribution were used for a few OTUs. The Pearson residuals obtained from the fitted models were then used to infer the direct interactions between the OTUs, making use of gaussian copulas to model the bivariate distributions between the residuals associated to any pair of OTUs. The position of a leaf within the canopy was taken as a proxy for its abiotic environment (e.g. temperature, UV exposure). The four variables describing the position of the sampled leaves - their distance to the ground, to the base of the branch, to the tree trunk and the orientation of the branch - were included as predictors in the GLMs, as single factors. The inference yielded, for any possible edge, the probability of its belonging to the true network. The prior probability of interactions was fixed to 0.5. Only edges with posterior appearance probability higher than 0.5 were kept in the interaction network. The sign of the interactions was retrieved by looking at the sign of the maximum a posteriori of the correlation in the associated bivariate gaussian copula. The inferred microbial networks consequently had positive and negative edges, that were visualized with the Gephi (v. 0.8.2 beta) software (Bastian & Heymann, 2009). Based on the theoretical outcome of each interaction type for each interacting partner (positive (+), negative (-) or null (0); see (Faust & Raes, 2012)), positive edges may indicate mutualism (+/+) or commensalism (+/0). Negative edges may indicate competition (-/-) or amensalism (-/0). The correspondence between network edges and parasitism or predation (+/-) is less straightforward to predict.

2.3 Results

Taxonomic composition of oak foliar microbial communities

Sequencing the fungal ITS1 amplicon libraries yielded 297888 sequences of which 259075 (86.97%) passed the quality filtering procedures, and 251039 (84.27%) se-

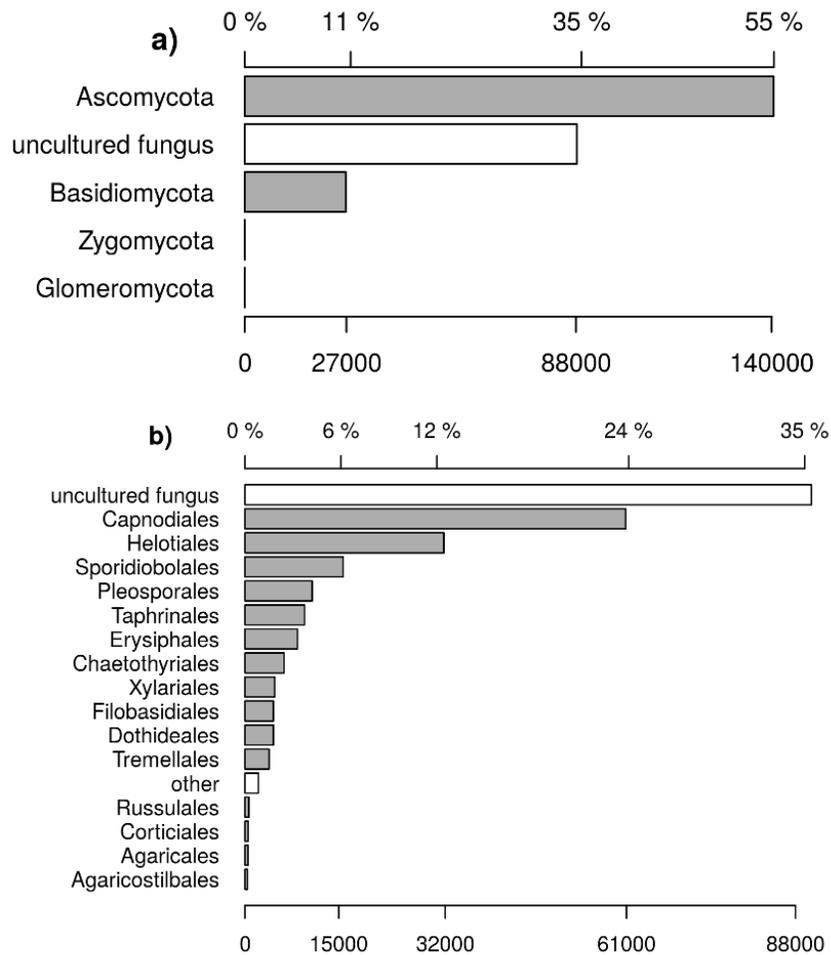


Figure S1: Taxonomical distribution of the fungal OTUs (a) at the phylum level and (b) at the order level

quences remained after discarding the singleton sequences and after chimera filtering. The number of sequences per sample ranged from 1202 to 3725 with a mean of 2166 and a median of 2146. These sequences clustered into 1074 OTUs. Two OTUs assigned to *Q. robur* were excluded from the dataset. Of the remaining OTUs a putative taxonomy could be assigned to 747 OTUs. In agreement with previous findings on foliar fungal communities (Jumpponen & Jones, 2009; Cordier *et al.*, 2012; Douanla-Meli *et al.*, 2013), the taxonomically assigned OTUs were dominated by Ascomycota (492 OTUs, 55.1% of the high quality sequences), followed by Basidiomycota (252 OTUs, 10.5%). One OTU was assigned to Glomeromycota and two to Zygomycota. Taxonomic composition at the phylum and order level is available in the supplementary materials (Figure S1).

Sequencing of the bacterial 16S amplicon pool yielded $30.5 \cdot 10^6$ sequences of which $28.1 \cdot 10^6$ were oak chloroplast sequences that were filtered out. OTU clustering

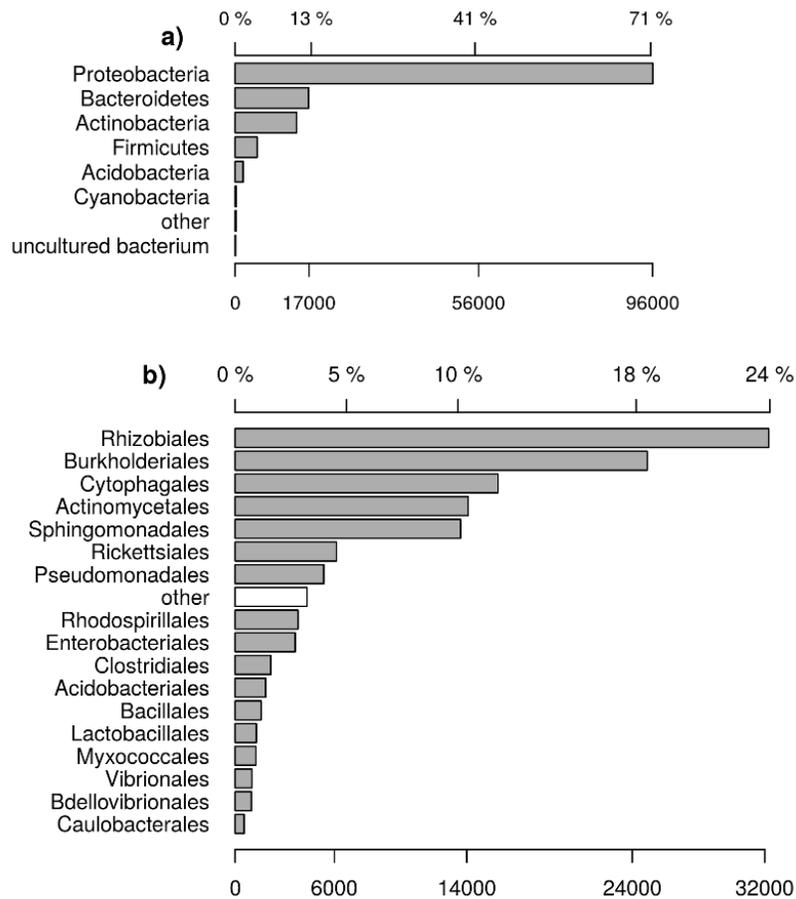


Figure S2: Taxonomical distribution of the bacterial OTUs (a) at the phylum level and (b) at the order level

yielded 1507 OTUs, of which 820 were of bacterial origin. The number of bacterial sequences per sample ranged from 73 to 11706 with a mean of 1160 and a median of 668. Within the bacterial OTUs, 816 were taxonomically assigned. The bacterial phyla associated with the oak leaves were dominated by Proteobacteria (406 OTUs, 49% of the high quality bacterial sequences), which comprised Alphaproteobacteria (205 OTUs, 25%), Betaproteobacteria (96 OTUs, 12%), Gammaproteobacteria (76 OTUs, 9%), and Deltaproteobacteria (29 OTUs, 4%). Other phyla were Firmicutes (175 OTUs, 21%), Actinobacteria (114 OTUs, 14%), Bacteroidetes (83 OTUs, 10%), Acidobacteria (11 OTUs, 1%), and others (31 OTUs, 4%). Detailed taxonomic composition at the phylum and order level is available in the supplementary materials (Figure S2).

Relationship between the abundance of powdery mildew symptoms and the number of reads assigned to *E. alphitoides*

As expected, the three trees differed significantly in both the percentage of symptomatic upper leaf area (GLM; $F=30.72$, $df_1=2$, $df_2=113$, $p<0.001$) and the *E. alphitoides* read number (NBGLM; $\text{Chisq}=91.62$, $df=2$, $p<0.001$). Percentages of *E. alphitoides* symptoms and number of *E. alphitoides* reads were significantly higher for the tree previously characterized as highly susceptible to powdery mildew (Figure 1 on the next page).

Percentages of symptomatic upper leaf area equaled 12.9%, 1.3% and 0.4% on average for the highly susceptible, intermediately resistant or strongly resistant trees, respectively. Percentages of *E. alphitoides* reads equaled 5.4%, 0.17% and 0.01% on average, respectively. Analysis of leaves within the susceptible tree showed that the number of reads assigned to *E. alphitoides* increased significantly with the percentage of symptomatic upper leaf area (NBGLM; $\text{Chisq}=12.14$, $df=1$, $p<0.001$; Figure 2 on page 54).

Relationship between the level of infection by *E. alphitoides* and the composition of foliar microbial communities

Analysis of leaves within the susceptible tree also showed that the level of infection by *E. alphitoides* was significantly related to fungal community composition. The percentage of *E. alphitoides* reads had a significant effect on the variable comp1 (GLM; $F=13.56$, $df_1=1$, $df_2=37$, $p<0.001$) but not comp2 (GLM; $F=0.79$, $df_1=1$, $df_2=37$, $p=0.39$). Variations in the percentage of *E. alphitoides* reads thus explained compositional variations observed along the first axis of the PCA (Figure 3 on page 55). Highly infected leaves harbored a higher abundance of OTU 3 (taxonomically unassigned; contributing to 23.6% of the compositional variations along the first PCA axis) and OTU 26 (assigned to *Taphrina carpini*; 1.4%). Leaves with low abundance of *E. alphitoides* were divided into two groups differing in their fungal community composition. The first group had a high abundance of OTU 1 (assigned to *Naevala minutissima*; 4.6%), OTU 10 (unassigned; 2.4%), and OTU 19 (unassigned; 6.3%). The second group had a high abundance of OTU 1278 (assigned to *Mycosphaerella punctiformis*). This latter species was the main driver of variations in fungal community composition. Its contributions to the first and second axis of the PCA equaled 58.5% and 34.7%, respectively. Pearson correlations between OTUs relative abundances showed that it is significantly, negatively associated to the pathogen *E. alphitoides* ($r=-0.40$; $p\text{-val}=0.01$). Analyses of associations between fungal OTUs and leaf type (highly infected by *E. alphitoides* versus "uninfected") confirmed these results. OTU 3 and OTU 26 were significantly associated to highly infected leaves ($\text{phi}=0.38$, $p\text{-val}=0.02$ and $\text{phi}=0.33$, $p\text{-val}=0.03$, respectively), while OTU 1278 and OTU 19 were significantly associated to uninfected leaves ($\text{phi}=0.39$, $p\text{-val}=0.01$ and $\text{phi}=0.39$, $p\text{-val}=0.007$, respectively). OTU 28 (taxonomically as-

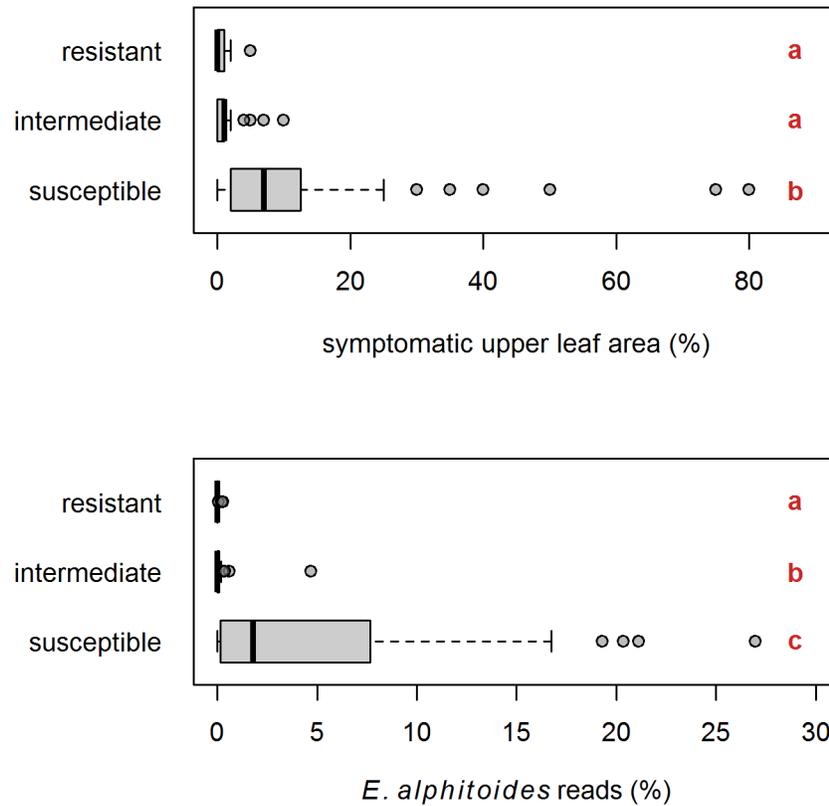


Figure 1: Percentage of (a) symptomatic upper leaf area and (b) reads assigned to the fungal pathogen *Erysiphe alphitoides* for three oak trees (*Quercus robur* L.) that were previously characterized as highly susceptible, intermediately resistant or strongly resistant to powdery mildew based on five years of visual observations. Box and whisker plots show the median percentage (thick black line), the quartiles (boxes) and the quartiles plus (or minus) one and a half times the interquartile range (whiskers). Outliers are shown as circles. Letters indicate the results of pairwise comparisons performed using Tukey's tests. Different letters indicate significant differences between trees.

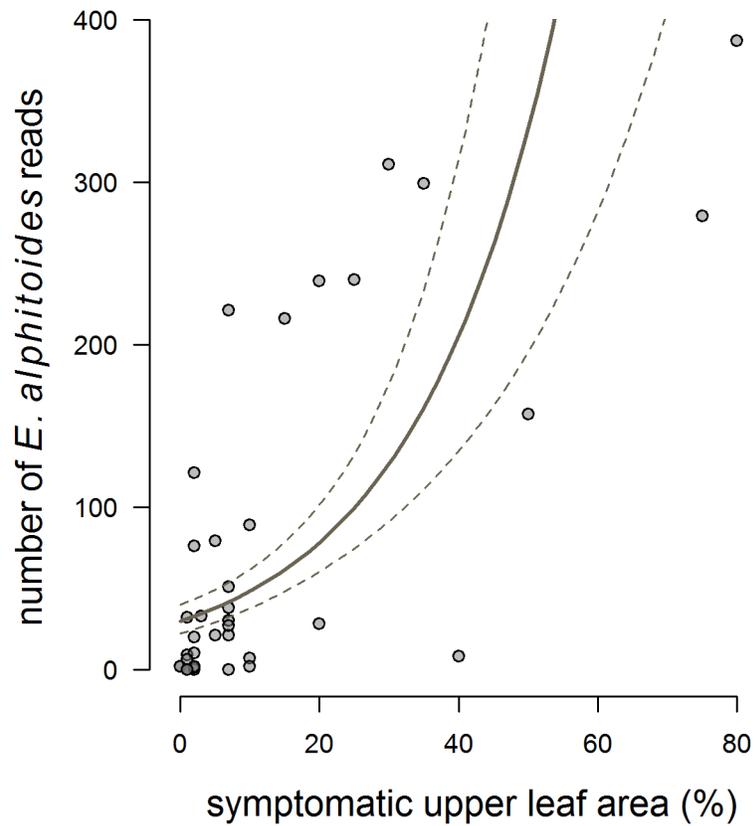


Figure 2: Relationship between the percentage of symptomatic upper leaf area and the number of reads assigned to *Erysiphe alphitoides*, across leaves sampled in an oak tree (*Quercus robur* L.) highly susceptible to *E. alphitoides*. The number of reads was obtained from a rarefied dataset having 1466 sequences per sample. The solid grey line represent predictions of a negative binomial generalized linear model. Dotted lines represent standard errors of the predictions.

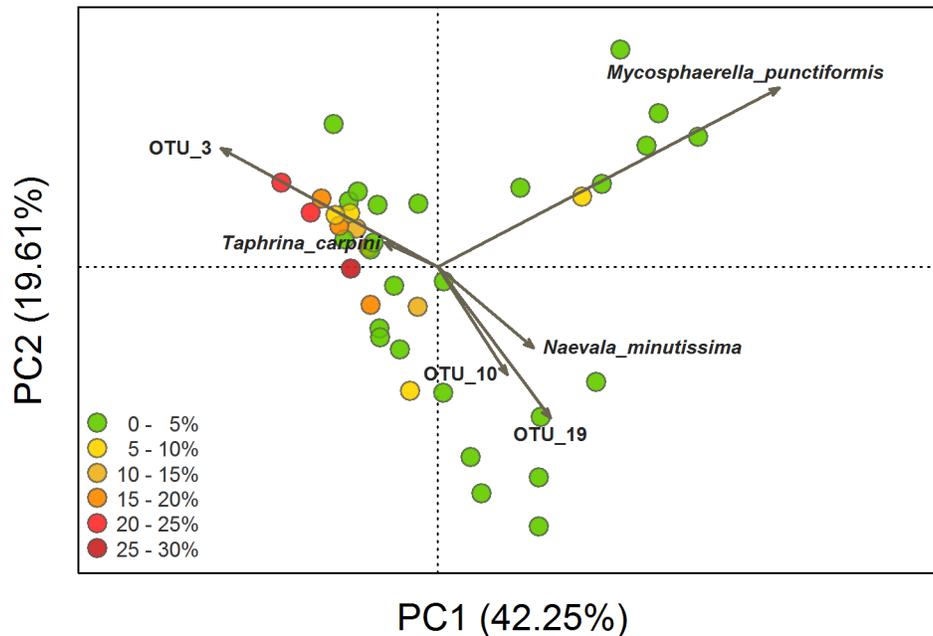


Figure 3: Principal coordinate analysis showing dissimilarities in fungal community (663 OTUs) composition among leaves of an oak tree (*Quercus robur* L.) highly susceptible to *Erysiphe alphitoides*. The arrows indicate the OTUs driving variations in fungal community composition and which contribute at least 1% to either of the principal components. The ordination is based on 100 rarefied datasets and is very robust to rarefaction. *E. alphitoides* was considered as being external to the fungal community. The colors indicate the percentage of reads assigned to *E. alphitoides* for each leaf.

signed to *Sporobolomyces gracilis*) was also strongly associated to uninfected leaves ($\phi=0.47$, $p\text{-val}=0.007$). In contrast to the fungal community, the bacterial community composition was unrelated to the level of infection by *E. alphitoides* (GLM; $F=1.45$, $df_1=1$, $df_2=37$, $p=0.23$ for comp1; $F=0.03$, $df_1=1$, $df_2=37$, $p=0.87$ for comp2). This result may nevertheless be due to the low rarefaction threshold that was used for analyzing the bacterial dataset (114 sequences per sample).

Inferred interactions between *E. alphitoides* and other members of the foliar community

We used the microbial network for the susceptible tree to highlight the most likely interactions between *E. alphitoides* and other foliar microorganisms, because *E. alphitoides* had a very low abundance in the two other trees. In that network, *E. alphitoides* had 26 interactions in total (Figure 4 on the following page), of which 13 were interactions with bacterial OTUs (6 positive, 7 negative) and 13 were interac-

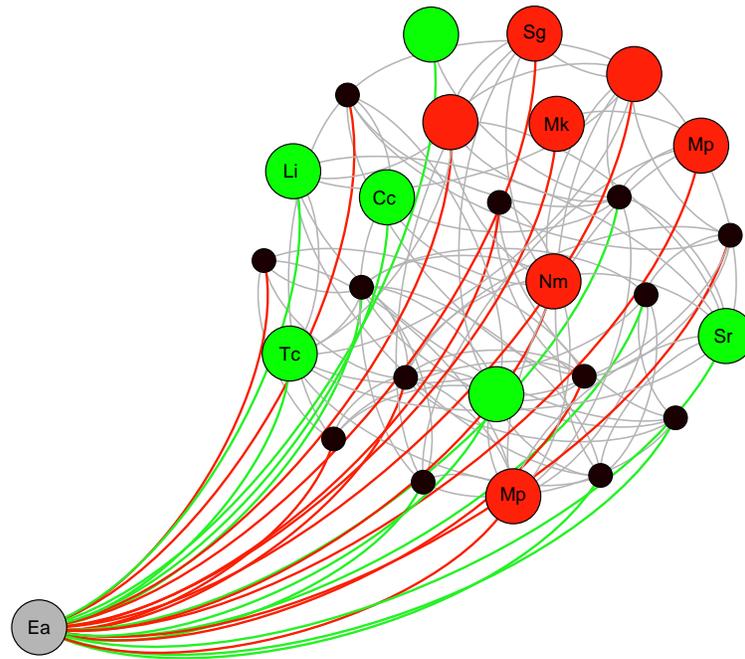


Figure 4: Model of the pathobiome of *Erysiphe alphitoides* on oak leaves (*Quercus robur* L.). Network nodes correspond to microbial OTUs. OTUs are linked if they are likely to interact together through direct ecological interactions, according to the results of a Bayesian model of network inference. The big grey node is *E. alphitoides* (*Ea*). Small black nodes correspond to bacterial OTUs. Big green nodes are fungal OTUs that are positively associated with *Ea*. Big red nodes are fungal OTUs that are negatively associated with *Ea*. Interactions between *Ea* and other OTUs are represented in green when the OTUs are positively associated, and in red when they tend to exclude each other. Interactions between the interacting partners of *Ea* are represented in grey. The names of the fungal OTUs that could be assigned to the species level is indicated. Cc: *Cladosporium cladosporioides*; Li: *Lalaria inositophila*; Nm: *Naevula minutissima*; Mp: *Mycosphaerella punctiformis*; Mk: *Monochaetia kansensis*; Tc: *Taphrina carpini*; Sg: *Sporobolomyces gracilis*; Sr: *Sporobolomyces roseus*.

tions with fungal OTUs (6 positive, 7 negative). The list of these OTUs with their taxonomic assignment is given in Table 1 on the next page. All the fungal species highlighted in ordination and association analyses, except OTU 3, were likely to interact directly with *E. alphitoides*. OTU 1 (assigned to *Naevala minutissima*), OTU 10, OTU 19 (both taxonomically unassigned), OTU 28 (assigned to *Sporobolomyces gracilis*) and OTU 1278 (assigned to *Mycosphaerella punctiformis*) were linked to *E. alphitoides* through negative edges. OTU 26 (assigned to *Taphrina carpini*) was linked to *E. alphitoides* through a positive edge. The network inference highlighted seven other fungal OTUs that were likely to interact directly to *E. alphitoides*. Both of them, OTU 2 (also assigned to *Mycosphaerella punctiformis*) and OTU 15 (assigned to *Monochaetia kansensis*), were linked negatively to *E. alphitoides*. These interactions were embedded within a larger network dominated by positive edges. The whole network for the susceptible tree had 1099 edges, including 65% of positive edges. Within-kingdom interactions were predominantly positive (71% for bacteria-bacteria interactions and 63% for fungi-fungi interactions), as well as cross-kingdom interactions (60%). The two other networks were also dominated by positive edges. The one for the intermediately resistant tree had 1112 edges, including 66% of positive edges. The one for the resistant tree had 1082 edges, including 65% of positive edges. A total of 95 positive interactions and 5 negative interactions were common to the three networks.

2.4 Discussion

The pathobiome concept has been defined in the context of microbial community interactions (Vayssier-Taussat *et al.*, 2014). It can be viewed as a subset of the microbiome of a host, centered on a pathogen species. The microbiome corresponds to the totality of microbes (bacteria, archaea and fungi) in a given environment and the totality of functions encoded by this microbiota (Schlaeppli & Bulgarelli, 2015). The pathobiome can thus be defined as the totality of microbes interacting with a given pathogen species and their influence on pathogenesis. In this study we showed that the pathobiome can be studied by combining metagenomics and network inference. We used these tools to decipher the pathobiome of *Erysiphe alphitoides*, the causal agent of oak powdery mildew (Mougou *et al.*, 2008; Mougou-Hamdane *et al.*, 2010).

A prerequisite for studying the pathobiome of a given pathogen using high-throughput sequencing is a significant relationship between the level of infection by the pathogen and the number of sequences assigned to it. Our results showed that this is the case for oak powdery mildew, both at the leaf level and at the tree level. These results parallel those obtained recently by Sapkota *et al.* (2015) for two foliar diseases of cereals. Another prerequisite is a good preservation of the microbial community composition during tissue storage. Our results suggested that the storage of oak leaves in silica-gel for a period of two weeks does not bias the composition of the

Phylum	OTU number	Interaction with <i>E. alphi-tooides</i>	Closest match in GenBank	Coverage (%)	Similarity (%)	Putative species
Bacteria	625	-0,243	KR265624.1	97	97	<i>Acidovorax</i> sp.
	27	-0,221	JN545036.1	100	98	<i>Ralstonia</i> sp.
	49	-0,204	LN829899.2	100	99	<i>Arsenophonus</i> sp.
	63	-0,195	JX993414.1	97	98	<i>Methylobacterium</i> sp.
	90	-0,187	KM253240.1	97	94	<i>Burkholderia</i> sp.
	112	-0,175	KM253240.1	97	98	<i>Burkholderia</i> sp.
	87	-0,165	HQ179145.1	97	100	<i>Staphylococcus</i> sp.
	31	0,144	JX448502.1	96	99	<i>Sphingomonas</i> sp.
	444	0,151	KR265692.1	97	98	<i>Methylobacterium</i> sp.
	1191	0,186	KM203869.1	94	97	<i>Hymenobacter</i> sp.
	42	0,19	KF681060.1	97	98	<i>Methylobacterium</i> sp.
	33	0,2	KR181798.1	97	99	<i>Fronidihabitans</i> sp.
	20	0,21	AB698681.1	97	98	<i>Methylobacterium</i> sp.
	Fungi	19	-0,244	-		
10		-0,194	-			
1		-0,179	AY853228.1	100	100	<i>Naevala minutissima</i>
1278		-0,179	KC339234.1	100	97	<i>Mycosphaerella punctiformis</i>
28		-0,161	KJ706995.1	100	99	<i>Sporobolomyces gracilis</i>
15		-0,145	KC345691.1	100	100	<i>Monochaetia kansensis</i>
2		-0,139	KC339234.1	100	100	<i>Mycosphaerella punctiformis</i>
1567		0,124	-			
9		0,146	KP836326.1	100	100	<i>Cladosporium cladosporioides</i>
25		0,178	-			
23		0,181	HG008766.1	100	99	<i>Sporobolomyces roseus</i>
20		0,241	AY239214.1	100	100	<i>Lalaria inositophila</i>
26	0,265	AF492085.1	100	100	<i>Taphrina carpini</i>	

Table 1: Taxonomic assignment of the 26 operational taxonomic units (OTUs) potentially interacting with *Erysiphe alphioides* according to the results of a Bayesian model of network inference. The interaction indicates the maximum a posteriori of the correlation in the associated bivariate Gaussian copula. The sign of the interaction indicates whether the OTUs are positively or negatively associated to *E. alphioides*. Taxonomic assignments are based on BLAST analysis of OTU seed sequences against GenBank. Coverage is the percentage of the query length covered by the alignment. Similarity is the percentage identity over the alignment.

foliar fungal community composition inferred by 454-pyrosequencing. These results nevertheless contrast with those obtained by U'Ren *et al.* (2014) for moss and lichen species. The last but not least requirement is to be able to predict intra-kingdom and cross-kingdom interactions involving a pathogen species, based on metagenomic data. Here we used the Bayesian method of network inference developed by Schwaller *et al.* (2015) to highlight the most likely microbial interactions.

To decipher the pathobiome of *E. alphitoides*, we characterized the fungal and the bacterial community composition of oak leaves (*Quercus robur* L.) differing in their level of infection by the pathogen. Our results showed that the fungal community composition changes markedly with the level of infection by *E. alphitoides*. The community composition of phyllosphere yeasts, in particular, varied significantly with infection. The yeast species *Taphrina carpini* was for instance associated to highly infected leaves, while *Sporobolomyces gracilis* was preferentially found on leaves with a low abundance of *E. alphitoides*. Another fungal species whose abundance differed significantly between infected and uninfected leaves was *Mycosphaerella punctiformis*. This species asymptotically colonizes living oak leaves as an endophyte, and then sporulates on senescent leaves (Verkley *et al.*, 2004). Its relative abundance was significantly, negatively correlated to that of *E. alphitoides*. This finding suggests that this fungal endophyte may play a protective role against oak powdery mildew. According to the microbial interaction network, there might be a direct ecological interaction between both species. Controlled inoculations will have to be performed to validate the antagonistic relationship between both species, and to understand the mechanism of interaction. According to our results, another foliar endophyte that might compete with *E. alphitoides* is *Monochaetia kansensis*. This fungus has already been observed on various *Quercus* species, including pedunculate oak (Genaro *et al.*, 2003), and produces a variety of bioactive compounds (Yogeswari *et al.*, 2012). In contrast, the ubiquitous foliar endophyte *Cladosporium cladosporioides* was positively linked to *E. alphitoides* in the microbial interaction network. This endophyte is considered as a weakness parasite that becomes virulent when the host plant is weakened by stress. Its abundance was shown to be higher in declining oak trees than in healthy trees (Ragazzi *et al.*, 2001). Our results suggest that leaf infection by *E. alphitoides* might be beneficial to this species.

Contrary to our expectations, our results suggested that mutualism, facilitation and commensalism dominate microbial interactions in the habitat formed by oak leaves. More than half of the edges in the microbial interaction networks were indeed positive. These results contrast with other findings, suggesting that competition is by far the most dominant outcome in microbial species interactions (Hibbing *et al.*, 2010; Foster & Bell, 2012). This high proportion of positive microbial interactions may be accounted for by the fact that our analyses were based on a snapshot of the microbial communities, taken about three months after leaf unfolding. Many competitive interactions may have occurred before this snapshot. The weaker competitors may have already been excluded, tipping the scales towards positive interactions. The positive interactions between *E. alphitoides* and some bacterial OTUs

may be accounted for by endosymbiosis (Hoffman & Arnold, 2010), commensalism or facilitation. Bacteria may use the dead hyphae as a nutrient source. They may also facilitate the infection (Venturi & da Silva, 2012) because they benefit from the changes induced by the fungal pathogen in the plant metabolism. These changes may indeed enlarge the quantity and/or the spectrum of substrates available to bacteria. For instance, the transpiration rate of oak leaves increases with powdery mildew infection as a result of transpiration via mycelium (Hewitt & Ayres, 1975). This change may be beneficial to bacteria, as water availability on the leaf surface was previously shown to be one of the major factors limiting bacterial growth (Lindow & Brandl, 2003).

In conclusion, our results showed that complex networks of microbe-microbe interactions occur on oak leaves. Network inference highlighted 13 bacterial OTUs and 13 fungal OTUs that are likely to interact directly with the *E. alphitoides*, the causal agent of oak powdery mildew. Among the fungal OTUs, nine could be assigned at the species level. Two foliar endophytic species - *Mycosphaerella punctiformis* and *Monochaetia kansensis* - were highlighted as potential antagonists of *E. alphitoides*. Several phyllosphere yeasts were also likely to have direct interactions with *E. alphitoides*. Controlled inoculations will have to be performed to validate these potential interactions. Additionally, the study of the temporal dynamics of microbial networks during the course of infection appears to be a promising research avenue, that will give deeper insights into the ecology of this disease and may improve its biological control. Endophyte-based biocontrol of forest diseases is still rare nowadays (Witzell *et al.*, 2014). Here we showed that combining metagenomics and network ecology may foster this method of forest protection, by highlighting potential interactions between pathogens and endophytes. The study of the effect of tree genotype on the structure and functioning of microbial networks is another promising research avenue. Our results suggest that there are differences in foliar microbial networks between trees, but replicates are needed to test the effect of tree genotype. Understanding how tree genetics shapes microbial interaction networks will strengthen the field of community genetics (Lamit *et al.*, 2015) and may have applied perspectives for tree breeding (Hacquard & Schadt, 2014; Pautasso *et al.*, 2014; Desprez-Loustau *et al.*, 2015).

2.5 Acknowledgements

We thank Xavier Capdevielle, Olivier Fabreguettes, Laure Villate and Martine Martin-Clotté (INRA, BioGeCo) for their technical help and advice during preliminary experiments and during the course of the study. We also thank Franck Salin, Thibaut Decourcelle, Adline Delcamp and Christophe Hubert (CGFB, Bordeaux) for the sequencing of the samples. Sampling and sequencing expenses were covered by the AIP Bioressource METAPHORE . Computing facilities were provided by the MCIA (*Mésocentre de Calcul Intensif Aquitain*) of the *Université de Pau et des Pays de*

l'Adour. BJ was funded by a grant from the French Ministry of Research and Education (MENRT no. 2011/AF/57). We thank Cindy E. Morris for helpful discussions about the phyllosphere. We thank Marie-Laure Desprez-Loustau, Arndt Hampe, Samantha Yeo and David Bohan and four anonymous reviewers for their very helpful comments. We also thank Julie Sappa from Alex Edelman & Associates for English language revision.

2.6 References

- Altschul, S. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215, 403–410.
- Bastian, M. & Heymann, S. (2009). Gephi : An Open Source Software for Exploring and Manipulating Networks. *Association for the Advancement of Artificial Intelligence*.
- Bengtsson, J., Eriksson, K.M., Hartmann, M., Wang, Z., Shenoy, B.D., Grelet, G.A., Abarenkov, K., Petri, A., Rosenblad, M.A. & Nilsson, R.H. (2011). Metaxa: a software tool for automated detection and discrimination among ribosomal small subunit (12S/16S/18S) sequences of archaea, bacteria, eukaryotes, mitochondria, and chloroplasts in metagenomes and environmental sequencing datasets. *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology*, 100, 471–5.
- Berlec, A. (2012). Novel techniques and findings in the study of plant microbiota: search for plant probiotics. *Plant science : an international journal of experimental plant biology*, 193-194, 96–102.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335–336.
- Clemente, J.C., Ursell, L.K., Parfrey, L.W. & Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell*, 148, 1258–70.
- Copolovici, L., Vaartnou, F., Estrada, M.P. & Niinemets, U. (2014). Oak powdery mildew (*Erysiphe alphitoides*)-induced volatile emissions scale with the degree of infection in *Quercus robur*. *Tree Physiology*, 34, 1399–1410.

- Cordier, T., Robin, C., Capdevielle, X., Desprez-Loustau, M.L. & Vacher, C. (2012). Spatial variability of phyllosphere fungal assemblages: genetic distance predominates over geographic distance in a European beech stand (*Fagus sylvatica*). *Fungal Ecology*, 5, 509–520.
- Desprez-Loustau, M.L., Aguayo, J., Dutech, C., Hayden, K.J., Husson, C., Jakushkin, B., Marçais, B., Piou, D., Robin, C. & Vacher, C. (2015). An evolutionary ecology perspective to address forest pathology challenges of today and tomorrow. *Annals of Forest Science*.
- Di Bella, J.M., Bao, Y., Gloor, G.B., Burton, J.P. & Reid, G. (2013). High throughput sequencing methods and analysis for microbiome research. *Journal of microbiological methods*, 95, 401–414.
- Douanla-Meli, C., Langer, E. & Talontsi Mouafo, F. (2013). Fungal endophyte diversity and community patterns in healthy and yellowing leaves of *Citrus limon*. *Fungal Ecology*, 6, 212–222.
- Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26, 2460–2461.
- Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10, 996–998.
- Faust, K. & Raes, J. (2012). Microbial interactions: from networks to models. *Nature Reviews Microbiology*, 10, 538–550.
- Foster, K.R. & Bell, T. (2012). Competition, not cooperation, dominates interactions among culturable microbial species. *Current Biology*, 22, 1845–1850.
- Frey-Klett, P., Burlinson, P., Deveau, A., Barret, M., Tarkka, M. & Sarniguet, A. (2011). Bacterial-Fungal Interactions: Hyphens between Agricultural, Clinical, Environmental, and Food Microbiologists. *Microbiology and Molecular Biology Reviews*, 75, 583–609.
- Gardes, M. & Bruns, T.D. (1993). ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology*, 2, 113–118.
- Gennaro, M., Gonthier, P. & Nicolotti, G. (2003). Fungal Endophytic Communities in Healthy and Declining *Quercus robur* L. and *Q. cerris* L. Trees in Northern Italy. *Journal of Phytopathology*, 151, 529–534.
- Glawe, D.A. (2008). The powdery mildews: a review of the world's most familiar (yet poorly known) plant pathogens. *Annual Review of Phytopathology*, 46, 27–51.

- Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R. & Reid, G. (2010). Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS ONE*, 5, e15406.
- Hacquard, S. & Schadt, C.W. (2014). Towards a holistic understanding of the beneficial interactions across the Populus microbiome. *New Phytologist*, pp. n/a–n/a.
- Hajji, M., Dreyer, E. & Marçais, B. (2009). Impact of Erysiphe alphitoides on transpiration and photosynthesis in Quercus robur leaves. *European Journal of Plant Pathology*, 125, 63–72.
- Hewitt, H. & Ayres, P. (1975). Changes in CO₂ and water vapour exchange rates in leaves of Quercus robur infected by Microsphaera alphitoides (powdery mildew). *Physiological Plant Pathology*, 7, 127–137.
- Hibbing, M.E., Fuqua, C., Parsek, M.R. & Peterson, S.B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, 8, 15–25.
- Hoffman, M.T. & Arnold, A.E. (2010). Diverse bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. *Applied and Environmental Microbiology*, 76, 4063–4075.
- Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. & Sogin, M.L. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, 4, e1000255.
- Jumpponen, A. & Jones, K.L. (2009). Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate Quercus macrocarpa phyllosphere. *New Phytologist*, 184, 438–48.
- Kemen, E. (2014). Microbe-microbe interactions determine oomycete and fungal host colonization. *Current Opinion in Plant Biology*, 20, 75–81.
- Kiss, L. (2003). A review of fungal antagonists of powdery mildews and their potential as biocontrol agents. *Pest Management Science*, 59, 475–483.
- Lamit, L.J., Lau, M.K., Næsberg, R.R., Wojtowicz, T., Whitham, T.G. & Gehring, C.A. (2015). Genotype variation in bark texture drives lichen community assembly across multiple environments. *Ecology*, 96, 960–971.
- Lindow, S.E. & Brandl, M.T. (2003). Microbiology of the Phyllosphere. *Applied and Environmental Microbiology*, 69, 1875–1883.
- Little, A.E.F., Robinson, C.J., Peterson, S.B., Raffa, K.F. & Handelsman, J. (2008). Rules of engagement: interspecies interactions that regulate microbial communities. *Annual Review of Microbiology*, 62, 375–401.

- Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. & Neufeld, J.D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, 13, 31.
- Mougou, A., Dutech, C. & Desprez-Loustau, M.L. (2008). New insights into the identity and origin of the causal agent of oak powdery mildew in Europe. *Forest Pathology*, 38, 275–287.
- Mougou-Hamdane, A., Giresse, X., Dutech, C. & Desprez-Loustau, M.L. (2010). Spatial distribution of lineages of oak powdery mildew fungi in France, using quick molecular detection methods. *Annals of Forest Science*, 67, 212.
- Newton, A., Gravouil, C. & Fountaine, J. (2010). Managing the ecology of foliar pathogens: ecological tolerance in crops. *Annals of Applied Biology*, 157, 343–359.
- Nilsson, R.H., Bok, G., Ryberg, M., Kristiansson, E. & Hallenberg, N. (2009). A software pipeline for processing and identification of fungal ITS sequences. *Source Code for Biology and Medicine*, 4, 1.
- Nilsson, R.H., Kristiansson, E., Ryberg, M. & Hallenberg, N. (2008). Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evolutionary Bioinformatics*, 4, 193–201.
- Nilsson, R.H., Veldre, V., Hartmann, M., Unterseher, M., Amend, A., Bergsten, J., Kristiansson, E., Ryberg, M., Jumpponen, A. & Abarenkov, K. (2010). An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. *Fungal Ecology*, 3, 284–287.
- Pautasso, M., Schlegel, M. & Holdenrieder, O. (2014). Forest Health in a Changing World. *Microbial Ecology*.
- R Core Team (2013). R: A Language and Environment for Statistical Computing.
- Ragazzi, A., Moricca, S., Capretti, P., Dellavalle, I., Mancini, F. & Turco, E. (2001). Endophytic fungi in *Quercus cerris*: Isolation frequency in relation to phenological phase, tree health and the organ affected. *Phytopathologia Mediterranea*, 40, 165–171.
- Rovenich, H., Boshoven, J.C. & Thomma, B.P. (2014). Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current Opinion in Plant Biology*, 20, 96–103.
- Sapkota, R., Knorr, K., Jørgensen, L.N., Hanlon, K.a.O. & Nicolaisen, M. (2015). Host genotype is an important determinant of the cereal phyllosphere mycobiome. *New Phytologist*.

- Schlaeppli, K. & Bulgarelli, D. (2015). The Plant Microbiome at Work. *MPMI*, 28, 212–217.
- Schoch, C.L., Seifert, K.a., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.a., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, a.N., Wingfield, M.J., Aime, M.C., An, K.D., Bai, F.Y., Barreto, R.W., Begerow, D., Bergeron, M.J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, a.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, a., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.a., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.a., Ge, Z.W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.B., Hosaka, K., Houbroken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., Kovacs, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, a.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, a.R., Methven, a.S., Meyer, W., Moncalvo, J.M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.a., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, a., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.O., Sung, G.H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.a., Urbina, H., Vagvolgyi, C., Vialle, a., Vu, T.D., Walther, G., Wang, Q.M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, a., Zamora, J.C., Zhang, N., Zhuang, W.Y. & Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109, 6241–6246.
- Schwaller, L. (2015). Saturnin: Spanning Trees Used for Network Inference. R Package. Version 1.0.
- Schwaller, L., Robin, S. & Stumpf, M. (2015). Bayesian Inference of Graphical Model Structures Using Trees.
- Scotti-Saintagne, C., Bodénès, C., Barreneche, T., Bertocchi, E., Plomion, C. & Kremer, A. (2004). Detection of quantitative trait loci controlling bud burst and height growth in *Quercus robur* L. *Theoretical and Applied Genetics*, 109, 1648–1659.

- Stout, M.J., Thaler, J.S. & Thomma, B.P.H.J. (2006). Plant-mediated interactions between pathogenic microorganisms and herbivorous arthropods. *Annual Review of Entomology*, 51, 663–89.
- Tack, A.J.M. & Dicke, M. (2013). Plant pathogens structure arthropod communities across multiple spatial and temporal scales. *Functional Ecology*, 27, 633–645.
- Tack, A.J.M., Gripenberg, S. & Roslin, T. (2012). Cross-kingdom interactions matter: fungal-mediated interactions structure an insect community on oak. *Ecology Letters*, 15, 177–185.
- Thiele, I., Heinken, A. & Fleming, R.M.T. (2013). A systems biology approach to studying the role of microbes in human health. *Current opinion in biotechnology*, 24, 4–12.
- Turner, T.R., James, E.K. & Poole, P.S. (2013). The plant microbiome. *Genome biology*, 14, 209.
- U'Ren, J.M., Riddle, J.M., Monacell, J.T., Carbone, I., Miadlikowska, J. & Arnold, a.E. (2014). Tissue storage and primer selection influence pyrosequencing-based inferences of diversity and community composition of endolichenic and endophytic fungi. *Molecular Ecology Resources*, 14, 1032–1048.
- Ursell, L.K., Van Treuren, W., Metcalf, J.L., Pirrung, M., Gewirtz, A. & Knight, R. (2013). Replenishing our defensive microbes. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 35, 810–7.
- Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.F., Jacques, M.A., Lebrun, M.H., Le Loir, Y., Ogliastro, M., Petit, M.A., Roumagnac, P. & Candresse, T. (2014). Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Frontiers in cellular and infection microbiology*, 4, 29.
- Venturi, V. & da Silva, D.P. (2012). Incoming pathogens team up with harmless 'resident' bacteria. *Trends in Microbiology*, 20, 160–164.
- Verkley, G.J., Crous, P.W., Groenewald, J., Braun, U. & Aptroot, A. (2004). *Mycosphaerella punctiformis* revisited: morphology, phylogeny, and epitypification of the type species of the genus *Mycosphaerella* (Dothideales, Ascomycota). *Mycological Research*, 108, 1271–1282.
- Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73, 5261–5267.
- White, T., Bruns, T., Lee, S. & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications*. (eds. Innis, M., Gelfand, D., Sninsky, J. & White, T.). Academic Press, London, pp. 315–322.

- Witzell, J., Martín, J.A. & Blumenstein, K. (2014). Ecological Aspects of Endophyte-Based Biocontrol of Forest Diseases. In: *Advances in Endophytic Research* (eds. Verma, V.C. & Gange, A.C.). Springer India, New Delhi, pp. 321–333.
- Yogeswari, S., Ramalakshmi, S., Neelavathy, R. & Muthumary, J. (2012). Identification and comparative studies of different volatile fractions from *Monochaetia kansensis* by GCMS. *Global Journal of Pharmacology*, 6, 65–71.
- Zargaran, M.R., Erbilgin, N. & Ghosta, Y. (2012). Changes in Oak Gall Wasps Species Diversity (Hymenoptera: Cynipidae) in Relation to the Presence of Oak Powdery Mildew (*Erysiphe alphitoides*). *Zoological Studies*, 51, 175–184.

2.7 Supplementary methods S1 - Preliminary experiment aimed at assessing the effect of oak leaf storage on fungal community composition.

Study site and sampling design

On the 5th of July 2011 at 2pm, 30 leaves were collected from the tip of three branches of a *Q. robur* tree growing in front of the lab at the INRA Pierroton forestry research station (Aquitaine, France; 44°44'N, 00°46'W). The branches were at 5m above ground and faced different orientations (North, South-East, South-West). Ten leaves of each branch, chosen randomly, were placed into sterile plastic bags containing silica gel, as performed in the present article. Bags had been filled with silicagel under the sterile hood. All materials required for filling the bags (including silica gel) had been previously placed for 40min under UV light. The bags, sealed and made airtight, were then stored in dark at ambient temperature in the laboratory. Ten other leaves were placed into commercial zipper locking plastic bags containing silica gel. Bags had been filled with silicagel under the sterile hood. The bags were also stored in dark at ambient temperature in the laboratory. The ten remaining leaves of each branch were placed in sterile plastic boxes and brought back immediately to the lab. Foliar disks were excised under the sterile hood and placed in microtubes containing one tungsten carbide bead, as described in the present article. Then the samples were immediately frozen under liquid nitrogen and ground by using Geno/Grinder® (SPEX SamplePrep, Metuchen, NJ, USA) at 1 300 strokes per minute for 180s. The samples were placed at -80°C at 4pm (2 hour after leaf collection). These three treatments - leaves dried in silica gel in sterile bags, leaves dried in silica gel in commercial zipper locking bags and fresh leaves immediately frozen after sampling - are hereafter referred to as SS, SC and FF respectively.

DNA extraction and fungal communities sequencing

On the 20th of July 2011, foliar disks were excised from the dried leaves as described in the present article. Then they were ground by using Geno/Grinder® (SPEX SamplePrep, Metuchen, NJ, USA) at 1 300 strokes per minute for 180 s. DNA was then extracted from all samples using the QIAGEN DNeasy® 96 Plant Kit. Several frozen samples (belonging to the same branch) were lost at this stage. Amplification of the fungal ITS1 region and purification of the PCR products was then performed as described in Counce *et al.* (2014). The purified PCR products were quantified (NanoDrop) and pooled in equal proportion at the branch level, resulting into 8 samples (2 FF, 3 SS and 3 SC). 454 pyro-sequencing of these samples and 26 others was then performed at the *Centre de Génomique Fonctionnelle de Bordeaux* (CGFB, Bordeaux, France) on one run of a GS-Junior sequencer (454 Life Sciences, Brandford, CT, USA). Bioinformatic analysis was performed with QIIME 1.6.0 (Caporaso

et al., 2010) following the same procedure as Cordier *et al.* (2012), with an additional step of sequence denoising performed by using the function *denoise.wrapper*.

Statistical analyses

We calculated the compositional dissimilarity matrix between samples, based on the Pearson index for 100 rarefied data sets. The rarefaction threshold was of 2692 sequences per sample. We then calculated the average dissimilarity matrix from these 100 matrices. Average dissimilarities in community composition between samples were represented by non-metric multidimensional scaling (NMDS). The effect of branch and treatment were assessed separately by performing permutational multivariate analyses of variance (PERMANOVAs) using the R *vegan* package (Oksanen *et al.*, 2014).

Results

The composition of foliar fungal communities differed significantly between branches (PERMANOVA; $F=163.9$, $df_1=2$, $df_2=5$, $p\text{-val}=0.07$), but not between treatments ($F=0.06$, $df_1=2$, $df_2=5$, $p\text{-val}=0.96$).

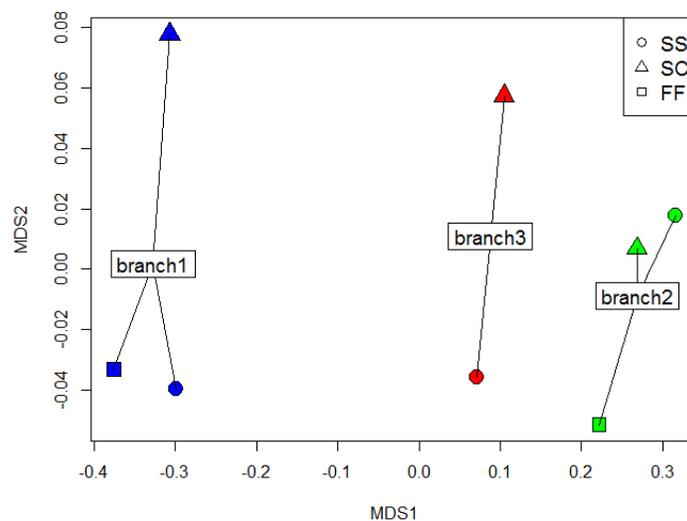


Figure 1: NMDS plot illustrating significant within-canopy differences in foliar fungal community composition in *Quercus robur*. Effect of tissue treatment - leaves dried in silica gel in sterile bags (SS), leaves dried in silica gel in commercial zipper locking bags (SC) and fresh leaves immediately frozen after sampling (FF) - on the community composition inferred by pyrosequencing was not significant.

Discussion

We showed that branches of an oak tree (*Q. robur* L.) harbor significantly different foliar fungal communities. These results parallel those found by Cordier *et al.* (2012) for foliar fungal communities of European beech (*Fagus sylvatica*) and by Leff *et al.* (2015) for foliar bacterial communities of *Ginkgo biloba*. Leaf treatment-tissues dried in silicagel versus flash-freezing of fresh tissues - did not have any significant effect on the composition of the foliar fungal community inferred by pyrosequencing, in contrast to the results obtained by U'Ren *et al.* (2014) for two species of moss and lichen. Our results thus suggest that leaf storage in silica gel is an adequate method for studying oak foliar fungal communities.

References

- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335–336.
- Coince, A., Cordier, T., Lengellé, J., Defosse, E., Vacher, C., Robin, C., Buée, M. & Marçais, B. (2014). Leaf and root-associated fungal assemblages do not follow similar elevational diversity patterns. *PLoS ONE*, 9.
- Cordier, T., Robin, C., Capdevielle, X., Desprez-Loustau, M.L. & Vacher, C. (2012). Spatial variability of phyllosphere fungal assemblages: genetic distance predominates over geographic distance in a European beech stand (*Fagus sylvatica*). *Fungal Ecology*, 5, 509–520.
- Leff, J.W., Del Tredici, P., Friedman, W.E. & Fierer, N. (2015). Spatial structuring of bacterial communities within individual *Ginkgo biloba* trees. *Environmental Microbiology*, 17, 2352–2361.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H. & Wagner, H. (2014). *vegan: Community Ecology Package*. R package version 2.2-0.
- U'Ren, J.M., Riddle, J.M., Monacell, J.T., Carbone, I., Miadlikowska, J. & Arnold, a.E. (2014). Tissue storage and primer selection influence pyrosequencing-based inferences of diversity and community composition of endolichenic and endophytic fungi. *Molecular Ecology Resources*, 14, 1032–1048.

Chapter 3

Phyllosphere microbiota structure of *Quercus robur* is genetically correlated with susceptibility to *Erysiphe alphitoides* infection

Boris Jakuschkin^{1,2}, Corinne Vacher^{1,2}, Jérôme Bartholome^{1,2}, Catherine Bodénès^{1,2}, Marie-Laure Desprez-Loustau^{1,2}, Cécile Robin^{1,2}, Virgil Fievet^{1,2}

1. INRA, UMR1202 BIOGECO, F-33610 Cestas, France
2. Univ. Bordeaux, BIOGECO, UMR 1202, F-33615 Pessac, France

Abstract

Phyllosphere microbiota influences host plants interactions with the biotic and abiotic environment. The structure of the phyllosphere microbial community is influenced by plant functional traits, which in turn are encoded in the plant genome. Therefore we hypothesized the existence of genomic regions, whose effects extend to the community level and influence the phyllosphere microbiota structure. We surveyed the phyllosphere bacterial and fungal communities of English Oak (*Quercus robur* L.) full-sib (F₁) mapping population using next-generation sequencing methods. We performed quantitative trait loci (QTL) analysis of microbial community descriptors. We found 8 QTLs for microbial community traits: fungal and bacterial composition, fungal diversity, and percentage of *Erysiphe alphitoides* reads. Three of these QTLs colocalized with a QTL for powdery mildew sensibility, suggesting for strong genetic determinism of powdery mildew resistance in *Q. robur*. We discuss our findings in the light of putative functional traits shaping microbial community of the phyllosphere.

Keywords: community genetics, phyllosphere microbiota, next-generation sequencing, Quantitative Trait Loci, *Erysiphe alphitoides*, powdery mildew, *Quercus robur*

3.1 Introduction

Microorganisms associated with plants influence the fitness of their hosts (Bolton, 2009; Hamilton *et al.*, 2010; Junker & Tholl, 2013). Plant phenotype is hence the product of the concerted and highly co-regulated expression of both plant and microbial genes (Partida-Martínez & Heil, 2011). Myriads of microorganisms inhabit inner and outer plant tissues above and below-ground and are collectively called the plant microbiota. They augment the plant genome by microbial genes, hence they extend plant genome functionality by providing novel biochemical capabilities or by modifying existing plant pathways (Friesen *et al.*, 2011). They influence carbon (C) and nitrogen (N) cycling (Delmotte *et al.*, 2009; Peñuelas *et al.*, 2012; Kembel *et al.*, 2014), plant growth (Compant *et al.*, 2010), plant defense-responses (Reinhold-Hurek & Hurek, 2011), plant biotic (Raghavendra & Newcombe, 2013) and abiotic stress tolerance (Hamilton & Bauerle, 2012), and other plant functional traits, plant biotic interactions, and ultimately ecosystem services (van Loon *et al.*, 1998; Friesen *et al.*, 2011; Bulgarelli *et al.*, 2013). Plant microbiota can help the plant to adapt to its environment and might help it to respond to rapid climate change (Compant *et al.*, 2010). Hence, the enrollment of cooperative microbial partners enables the plant to improve its fitness and to promote the evolution of cooperation processes (Vandenkoornhuyse *et al.*, 2015).

Plant genetic identity was shown to be important in structuring the above- and below-ground microbial communities (Schweitzer *et al.*, 2006, 2008; Ulrich *et al.*, 2008; Bailey *et al.*, 2009; Bálint *et al.*, 2013; Fischer *et al.*, 2014; Bodenhausen *et al.*, 2014; but see Mason *et al.*, 2015), whereby the effects of plant diversity were stronger on the above-ground microbiota (Bailey *et al.*, 2009). The effects of genetic factors are mediated by plant functional trait variation (Schweitzer *et al.*, 2010), like productivity traits (leaf water and phosphorous content, leaf mass per area, leaf thickness, thickness of the leaf epidermis, wood density, growth/mortality rate) or defense-related traits (total phenolics content, trichome density, ethylene signaling and cuticle composition) (Yadav *et al.*, 2005; Kembel *et al.*, 2014; Bodenhausen *et al.*, 2014). The quantitative nature of these traits suggests for existence of QTLs (Balint-Kurti *et al.*, 2010), whose effects extend to the community level influencing the leaf microbiota, consistent with the community genetics framework (Whitham *et al.*, 2006). Thus within this framework it was hypothesized that specific QTLs in plant host would be linked to properties of the plant associated community (Whitham *et al.*, 2003).

Forest trees offer a powerful system for testing community-level effects of genetic variation, as they are the basal resource in most food webs. They often are dom-

inant and keystone species in ecosystems, encompassing a huge number of biotic interactions, and thus having cascading effects throughout communities (Johnson & Stinchcombe, 2007). English oak (*Quercus robur*) grows under a wide range of climatic and edaphic conditions and is a dominant tree species in many forests throughout Europe (Gailing *et al.*, 2009). There are genetic maps available for oak and its genome data was recently released in public sequence repositories (Plomion *et al.*, 2015). Several QTLs were mapped for traits potentially involved in shaping oak associated microbial community through trait-mediated effects. These include leaf morphology (Saintagne *et al.*, 2004), bud burst (Scotti-Saintagne *et al.*, 2004), water use efficiency (Brendel *et al.*, 2007), and stomatal density (Gailing *et al.*, 2008). Here we hypothesized the existence of QTLs in oak, which are linked to the properties of oak phyllosphere microbiota. Those QTLs would be linked to oak traits with major effects on associated community.

3.2 Material and Methods

Experimental site and leaf-sampling design

The oaks belonged to a full-sib oak family obtained by a controlled cross realized during 1992 that provided acorns to construct a clonal bank composed of 278 clones. The female parent (accession 3P) is from the Forestry Research Station of Pierrotton (latitude 44°44'18" N, longitude 0°46'38" W) and the male parent (accession A4) originates from Arcachon (latitude 44°39'10" N, longitude 1°10'5" W). The 278 seedlings were raised in a seedbed in the nursery until age 3, and they were then transplanted in a clonal bank as stool beds. Rooted cuttings were produced from the clonal bank and planted in their second year of growth in an open-air experimental site in the South-west of France (Experimental Unit of Bourran 44°20'30" N, 0°23'35" E). The plantation comprising 207 full-sibs was planted in 2000 in a randomized incomplete block design (183 blocks with 12 cuttings each, in total 2196 individuals). This mapping population was used in numerous experiments to detect QTLs for leaf morphology traits (Saintagne *et al.*, 2004), phenology (Scotti-Saintagne *et al.*, 2004; Derory *et al.*, 2010), growth (Scotti-Saintagne *et al.*, 2004), water efficiency (Brendel *et al.*, 2007), and powdery-mildew resistance (Robin *et al.*, 2010; Mougou-Hamdane, 2009).

On the 29th of July 2011, we collected randomly three leaves each from North-East and South-West side of each tree at approximate height of 2 m, resulting in six leaves per tree and 180 trees. We collected leaves strictly from the first flush. Each leaf was cut at the base of the petiole and placed into a sterile bag containing silica gel. The bags, sealed and made airtight, were then stored in dark at +18°C in the laboratory. It was shown by Jakuschkin *et al.* (in prep., see Methods S1 on page 68) that the storage of *Q. robur* leaves in silica gel prior to the DNA-extraction (when compared to flash-freezing) did not have any significant effect on the composition of

the foliar fungal community inferred by pyrosequencing.

DNA extraction

We performed the DNA extraction under a laminar flow hood. We disinfected the hood and all tools with sodium hypochlorite followed by 70% ethanol and 40 minutes of UV exposure. All plastic ware and tungsten carbide beads were autoclaved for 2 hours prior to utilization. Using a metal corer we excised two discs on each side of the leaf middle vein – four discs of 2 cm² in total. Each set of four leaf discs was placed in a collection microtube (QIAGEN DNeasy® 96 Plant Kit) containing one tungsten carbide bead. The tools were sterilized prior to each leaf handling by dipping into 70% alcohol and flaming over an electric Bunsen burner. We used Geno/Grinder® (SPEX SamplePrep, Metuchen, NJ, USA) to grind the leaf material at 26 Hz for 90 seconds into homogenous powder. The samples were stored at -80°C overnight. We used the QIAGEN DNeasy 96 Plant Kit to extract the DNA (following the protocol for frozen plant tissue); we measured the DNA concentration with NanoDrop 8000 and diluted DNA working solutions to 1ng DNA μl^{-1} . Total DNA extracts from leaves were pooled by tree prior to the PCR step.

Amplification and sequencing of the fungal assemblages

To characterize the fungal assemblages on oak leaves we targeted the internal transcribed spacer 1 (ITS1) sublocus of the ITS region of the nuclear ribosomal repeat unit (Schoch *et al.*, 2012). The PCR reactions were pipetted under a laminar flow hood using the same disinfection protocol as for the DNA extraction. The PCR reactions (50 μl) contained 2.5 units Silverstar Taq polymerase (Eurogentec), 5 μl 10x PCR buffer supplied by the manufacturer, 2 μl MgCl₂ (50mM), 2 μl dNTPs (5mM), 1 μl of 10 μM forward primer, 1 μl of 10 μM reverse primer, 0.5 μl of 100mg/ml BSA, and 5ng DNA template. The PCR reaction was carried out as follows: an initial phase at 95°C (5 min), followed by 35 cycles at 95°C (30 sec), 54°C (60 sec) and 72°C (90 sec), and a final step at 72°C (10 min). Samples (leaf as a sample unit) were amplified in triplicate and pooled before purification with AMPure magnetic beads (Beckman Coulter Genomics, Danvers, MA, USA). The DNA concentration in each sample was quantified using Nanodrop8000 (Thermo Fischer Scientific, Waltham, MA, USA) and pooled to an equimolar concentration to obtain the amplicon libraries, five in total. The final DNA libraries were sequenced (unidirectional, starting with A adaptor) on 454 GS Junior sequencer(454 Life Sciences, Branford, CT, USA) at the Centre de Génomique Fonctionnelle de Bordeaux (CGFB, Bordeaux, France).

Amplification and sequencing of the bacterial assemblages

To characterize the bacterial assemblages on oak leaves we targeted the hypervariable region 6 (V6) of the 16S small ribosomal subunit gene (16S rRNA). We applied

the microbiome profiling method developed by Gloor *et al.* (2010), which uses combinatorial sequence tags attached to PCR primers amplifying the hypervariable V6 region of the 16S rRNA gene (Huse *et al.*, 2008). This type of tagging combined with the Illumina paired-end protocol enabled for the examination of hundreds of samples with far fewer primers than is required when tags are incorporated at only a single end. The PCR set up (50 μ l) contained 2.5 units Silverstar Taq polymerase (Eurogentec), 5 μ l 10x PCR buffer supplied by the manufacturer, 2 μ l MgCl₂ (50mM), 2 μ l dNTPs (5mM), 1 μ l of 10 μ M forward primer, 1 μ l of 10 μ M reverse primer, 0.5 μ l of 100mg/ml BSA, and 5ng DNA template. The PCR was carried out as follows: an initial phase at 95 $^{\circ}$ C (2 min), followed by 30 cycles at 95 $^{\circ}$ C (30 sec), 56 $^{\circ}$ C (30 sec) and 72 $^{\circ}$ C (30 sec), and a final extension at 72 $^{\circ}$ C (5 min). Samples were amplified in triplicate and pooled before purification with AMPure magnetic beads (Beckman Coulter Genomics, Danvers, MA, USA). Amplicon samples were then individually quantified using the Nanodrop8000 and pooled to an equimolar concentration. The final DNA library was sequenced on the Illumina platform GAIIx (Illumina, Inc, San Diego, CA, USA) at the *Centre de Génomique Fonctionnelle de Bordeaux* (CGFB, Bordeaux, France). Paired-end sequencing allowed for full coverage of the V6 region with reads overlapping at full length.

Processing of the fungal 454 pyrosequencing data

We used QIIME v.1.7.0 (Caporaso *et al.*, 2010) for quality filtering, demultiplexing, and denoising. After the removal of tags and primers, all sequences shorter than 100 nucleotides were discarded. Quality filtering was done with a sliding window test of quality scores having a window size of 50 nucleotides and a minimum quality score of 25. The sequences were truncated at the beginning of the poor quality window and tested for the minimum sequence length of 100 nucleotides. We removed the highly conserved ribosomal genes flanking the ITS1 marker region from the dataset using Fungal ITS Extractor 1.1 (Nilsson *et al.*, 2010). ITS1 sequences shorter than 100 nucleotides were discarded and sequences longer than 280 nucleotides were trimmed at the 3'-end. Before the clustering step we combined the dataset with the fungal dataset acquired from 118 leaves from three oak trees planted in the same field test (see Section 2 on page 41). In doing so we aimed to reduce the number of singletons. We discarded singleton reads before clustering to minimize spurious Operational Taxonomic Units (OTUs). OTU clustering was done with UPARSE-OTU (Edgar, 2013), an algorithm implemented within Usearch v.7 (Edgar, 2010) that performs chimera filtering and OTU clustering simultaneously. The algorithm identifies a set of OTU representative sequences at the 97% sequence similarity threshold, that corresponds roughly to the differences at the species level (Nilsson *et al.*, 2008). Unfortunately, other *Erysiphe* species, such as *E. quercicola* and *E. hypophylla*, have an ITS sequence which has less than a 3% difference from the *E. alphitoides* ITS sequence (Mougou *et al.*, 2008). ITS sequences of these species therefore cluster together at 97% similarity threshold. To circumvent this issue we performed a BLAST

search of all sequences against a custom taxonomic database of the *Erysiphe* genus. We extracted all sequences with successful hits to *E. alphitoides* into a separate file, and assigned the corresponding read counts. Taxonomy assignments were done by BLAST searches against Fungal ITS Database (Nilsson *et al.*, 2009) (as of March 15, 2012). We modified the taxonomic file associated with this database to equalize the length of taxonomic ranks, by using the NCBI Taxonomy Browser. Additionally we corrected a few false taxonomic assignments in the case of *Erysiphe* species and removed a few non-fungal sequences. We performed BLAST searches using QIIME (v.1.7.0) with standard settings (e-value: 0.001, minimum percent identity: 90) to assign the taxonomy to the most abundant sequence of each OTU. We also performed BLAST searches against the Nucleotide collection database (nr/nt database) to exclude possible *Q. robur* sequences. We then constructed the OTU table in QIIME and transformed it from BIOME into tab delimited text format for downstream analyses with R (R Core Team, 2013).

Processing of the bacterial Illumina data

We assembled the paired-end reads using PANDAseq (Masella *et al.*, 2012). Reads were then demultiplexed following the method described in Gloor *et al.* (2010) using a custom Bash shell script. We removed oak chloroplast sequence prior to the clustering step. OTU clustering and OTU table construction were done by following the same procedure than for fungi. Taxonomic assignments were performed with the RDP Classifier program (v.2.5) (Wang *et al.*, 2007) as it is implemented into QIIME (v.1.7.0), with standard settings and Greengenes 13_5 as the reference sequences. The origin of OTU representative sequences (archaeal, bacterial, nuclear eukaryote, mitochondrial or chloroplast origin) was identified using Metaxa (Bengtsson *et al.*, 2011) and we retained only the bacterial OTUs.

Construction of foliar microbial networks

We used the SparCC software (Friedman & Alm, 2012) to construct the microbial association network. SparCC approximates linear Pearson correlation coefficients between log-transformed OTU abundances. It does not require any normalization or downsampling of the data (Friedman & Alm, 2012). Given the very high total number of OTUs, we only included the most abundant fungal OTUs (representing in total 90% of the fungal sequences in the dataset) and the most abundant bacterial OTUs (representing in total 90% of the bacterial sequences in the dataset) to reduce computation time. We used 20 iterations to compute the correlation coefficients between OTU abundances. Their significance was assessed by generating 100 bootstrapped datasets (see Friedman & Alm (2012) for the bootstrap procedure). The pseudo p-values were defined as the proportion of random datasets for which a correlation value at least as extreme as the one computed for the original data was obtained. Only significant correlations (with p-value < 0.01) were retained as

network links. Each node of the network corresponds to an OTU and each link between two OTUs corresponds to a significant correlation between their abundances across all samples. The weight of each link (hereafter referred to as the association strength) is the Pearson coefficient correlation value.

Choice of traits for QTL analysis

All data handling, trait calculations and QTL analysis were performed in the R Program, version 3.1.3 (R Core Team, 2013). We measured bacterial and fungal diversity using the inverse Simpson index. The value of this index starts with 1 as the lowest possible value, with larger values indicating greater diversity. And the maximum possible value equals to the number of OTUs in the sample (richness). Inverse Simpson index is defined as $\frac{1}{D}$, where $D = \sum p_i^2$ and p_i is the proportional abundance of species i . This index is comparatively insensitive to numerically minor constituents (Bent & Forney, 2008) and it better differentiates between highly diverse samples, which is the case in our data. We chose the percentage of *Erysiphe alphitoides* reads in fungal samples as a proxy for the powdery mildew infection as the reads were previously shown to correlate with the disease symptoms (Figure 2 on page 54). Additionally we included measurements of powdery-mildew inoculation experiment performed in 2006 (Mougou-Hamdane, 2009). In this experiment excised leaves were inoculated with one strain of *Erysiphe alphitoides* and the lesion surface was measured every day. The “pm sympt” represents the average growth of lesion surface in $\mu\text{m}^2 \cdot \text{day}^{-1}$. As the measurement of the community composition we chose the axis projections of principal coordinate analysis (PCoA). We have used the Bray-Curtis dissimilarity index calculated on the bacterial and fungal datasets. Bray-Curtis index was previously shown to perform best in recovering environmental gradients (Austin, 2013). For each tree we also computed traits describing properties of the microbial association network (Table 3.1 on the next page). Changes in microbial community association networks will have implications for microbiome dynamics and functioning (Bakker *et al.*, 2014). Trait abbreviations and explanations are summarized in Table 3.2 on page 79. The pattern of phenotypic variation differed by trait (Figure 3.1 on page 81). Summary statistics of each trait are given in Table 3.3 on page 80.

Genotyping and map construction

A map was constructed for each parent of the cross following the so called double-pseudo-testcross mapping strategy (Grattapaglia & Sederoff, 1994) as previously described in Saintagne *et al.* (2004). In the pseudo-testcross configuration markers are present in one parent and absent in the other or vice versa, and are expected to segregate 1:1 in the F_1 generation (Grattapaglia & Sederoff, 1994). From the high-density map (Bodenes *et al.* in prep.), a skeletal map with uniformly distributed

Network trait	Formula
Average frequency of associations	$\frac{\sum_{i=1}^{T-1} \sum_{j=i+1}^T S_{ij} \cdot A_i \cdot A_j}{\sum_{i=1}^{T-1} \sum_{j=i+1}^T A_i \cdot A_j}$
Average frequency of positive associations	$\frac{\sum_{i=1}^{T-1} \sum_{j=i+1}^T P_{ij} \cdot A_i \cdot A_j}{\sum_{i=1}^{T-1} \sum_{j=i+1}^T A_i \cdot A_j}$
Average frequency of negative associations	$\frac{\sum_{i=1}^{T-1} \sum_{j=i+1}^T N_{ij} \cdot A_i \cdot A_j}{\sum_{i=1}^{T-1} \sum_{j=i+1}^T A_i \cdot A_j}$
Average strength of associations	$\frac{\sum_{i=1}^{T-1} \sum_{j=i+1}^T R_{ij} \cdot A_i \cdot A_j}{\sum_{i=1}^{T-1} \sum_{j=i+1}^T A_i \cdot A_j}$
Average strength of positive associations	$\frac{\sum_{i=1}^{T-1} \sum_{j=i+1}^T P_{ij} \cdot R_{ij} \cdot A_i \cdot A_j}{\sum_{i=1}^{T-1} \sum_{j=i+1}^T P_{ij} \cdot A_i \cdot A_j}$
Average strength of negative associations	$\frac{\sum_{i=1}^{T-1} \sum_{j=i+1}^T N_{ij} \cdot R_{ij} \cdot A_i \cdot A_j}{\sum_{i=1}^{T-1} \sum_{j=i+1}^T N_{ij} \cdot A_i \cdot A_j}$

R_{ij} is the association strength between OTU i and OTU j in the microbial network inferred with SparCC method, $0 \leq R_{ij} \leq 1$. S_{ij} is the associations sign between OTU i and OTU j , $S_{ij} \in \{-1, 0, 1\}$. P_{ij} is the presence of a positive association between OTU i and OTU j . $P_{ij} = 1$ if $S_{ij} = 1$ otherwise $P_{ij} = 0$. N_{ij} is the presence of a negative association between OTU i and OTU j . $N_{ij} = 1$ if $S_{ij} = -1$ otherwise $N_{ij} = 0$. T is the total number of OTUs. A_i and A_j represent the relative abundances of OTU i and OTU j in the sample, respectively.

Table 3.1: Traits describing properties of the microbial association network inferred with SparCC method.

Abbreviation	Trait explanation
fung div	fungus diversity (inverse Simpson index for fungal data)
bact div	bacterial diversity (inverse Simpson index for bacterial data)
fung comp1	1st principal component of PCoA with Bray-Curtis distance on fungal community data
fung comp2	2nd principal component on fungal data
bact comp1	1st principal component of PCoA with Bray-Curtis distance on bacterial community data
bact comp2	2nd principal component on bacterial data
Erys %reads	percentage of <i>Erysiphe alphitoides</i> in the sample
pm sympt	visual assessment of powdery-mildew infection in 2006: linear growth of lesion surface after controlled inoculation of excised leaves ($\mu\text{m}\cdot\text{day}^{-1}$)
netw int freq	average frequency of network associations
netw pint freq	average frequency of positive network associations
netw nint freq	average frequency of negative network associations
netw int strngth	average strength of network associations
netw pint strngth	average strength of positive network associations
netw nint strngth	average strength of negative network associations

Table 3.2: Microbial community traits, listing their abbreviation and explanation of traits.

markers (approximately 2 cM per marker interval) was extracted and used for QTL identification.

Number and type of markers

The female framework map was composed of 345 SNP markers and covered a distance of 686.8 cM, with an average spacing between markers of 2.01 cM (max: 5.5 cM). The male framework map comprised 341 SNP markers over a total distance of 768.6 cM, with an average marker spacing of 2.3 cM (max: 8.7 cM).

Trait	Phenotyped individuals	Min	Max	Mean	SD
fung div	181	2.88	31.71	16.44	6.63
bact div	174	1.48	47.68	21.82	7.52
fung comp1	181	-0.39	0.23	0	0.16
fung comp2	181	-0.22	0.29	0	0.11
bact comp1	174	-0.26	0.38	0	0.13
bact comp2	174	-0.19	0.27	0	0.09
Erys %reads	181	0	0.17	0.01	0.02
pm sympt	143	0	897.94	321.95	327.67
netw int freq	181	0.29	0.48	0.39	0.03
netw pint freq	181	0.14	0.36	0.23	0.04
netw nint freq	181	0.08	0.24	0.16	0.03
netw int strngth	181	0.08	0.16	0.11	0.02
netw pint strngth	181	0.04	0.13	0.07	0.02
netw nint strngth	181	0.02	0.06	0.04	0.01

Table 3.3: Phenotypic values in microbial traits in the *Q. robur* mapping population.

QTL analysis

Identification of quantitative trait loci (QTL) was performed using composite interval mapping (Zeng, 1994) with three marker covariates, a window size of 50cM in R/qtl (Broman *et al.*, 2003). The LOD threshold was established separately for each trait using 1,000 permutations at a significance threshold of 0.05. The position and effect of significant QTL were refined using the Haley–Knott regression method and then assessed for additive effects and percent variation explained by fitting a model containing all QTLs identified for a given trait in R/qtl. An approximate Bayesian credible interval was calculated for each QTL.

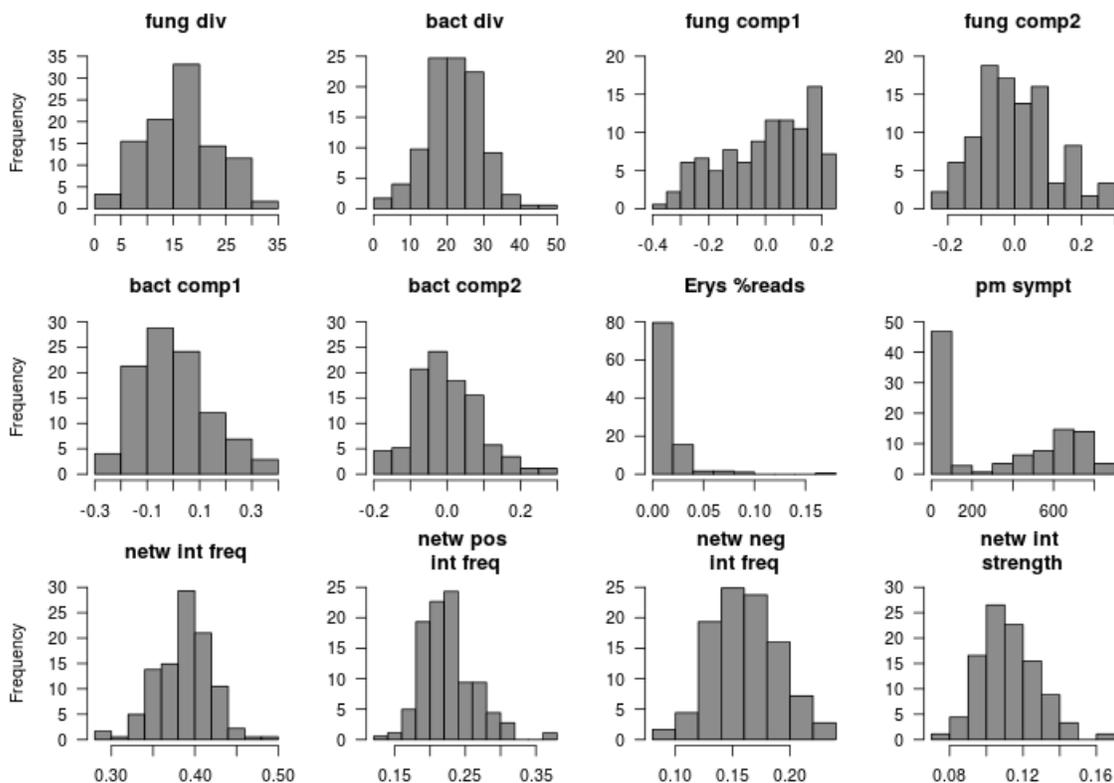


Figure 3.1: Phenotypic variation in microbial traits in the *Q. robur* mapping population.

3.3 Results

Sequencing data

Fungal data

Out of the initial 718193 raw reads, 94668 were found to be of poor read quality, too short, chimeric or singletons, leaving 623525 (87%) sequences for subsequent analyses. The number of sequences per sample ranged from 1764 to 6804 with the mean of 3445 (median: 2961). A total of 1443 OTUs were identified. Out of these OTUs a putative taxonomy could be assigned to 981 OTUs. The 462 not-assigned OTUs represented 25.5% of the reads and were confirmed as uncultured fungi.

Bacterial data

Sequencing of the bacterial amplicon pool yielded $31 \cdot 10^6$ sequences, of which $30.6 \cdot 10^6$ were chloroplast sequences, mitochondria sequences, poor quality, too short, chimeric or singletons, leaving 351000 sequences for subsequent analyses. In total 1102 bacterial OTUs were identified. The number of reads per sample ranged from

83 to 11890 with a mean of 1939 and a median of 1526. All OTUs were assigned at differing taxonomical resolution.

Correlation among traits

Significant correlations were observed between traits (Table 3.4 on the next page). Strong and significant correlations were observed between association network traits: average frequency of associations (netw int freq) and average strength of associations (netw int strngth). Three pairs of traits showing very strong correlations were average frequency of positive associations (netw pint freq) and average strength of positive associations (netw pint strngth), average frequency of negative associations (netw nint freq) and average strength of negative associations (netw nint strngth), average strength of associations (netw int strngth) and average strength of positive associations (netw pint strngth). These pairs were expected to give very similar or almost identical results in the QTL analysis, consequently we kept only one trait from each pair for the subsequent QTL analysis: netw pint freq, netw nint freq, netw int strngth. Furthermore we observed moderate significant correlations between fungal diversity (fung div) and fungal community composition (fung comp1), between fungal (fung comp1) and bacterial (bact comp1) community composition, between *Erysiphe alphitoides* read abundance (Erys %reads) and bacterial (bact comp1) community composition, between fungal community composition (fung comp1) and association network properties (netw int freq, netw nint freq), between network properties (netw int freq and netw pint freq, netw pint freq and netw nint freq). Among the weak significant correlations following correlations were to point out: between *E. alphitoides* read abundance (Erys %reads) and the visual assessment of powdery-mildew symptoms from 2006 (pm sympt), between *E. alphitoides* read abundance and fungal composition (fung comp1), and between *E. alphitoides* read abundance and network traits (netw pint freq, netw nint freq).

Chromosomal loci for the microbial community descriptors (QTLs)

We found 11 chromosomal regions for eight out of 12 traits: fung div, fung comp1, fung comp2, bact comp1, Erys %reads, pm sympt, netw nint freq and netw int strngth (Table 3.5 on page 84). QTLs were identified on linkage groups (LG) 4, 6, 8, 9 and 10 explaining between 2.98% and 55.26% of phenotypic variation. The QTL with the largest effect was for the powdery-mildew symptoms observations from inoculation experiment (pm sympt) and was identified on male (m) LG10. QTL for fung div was found on female (f) LG6. QTLs for fung comp1 and fung comp2 were found on LG9f and LG10m respectively. QTLs for bact comp1, fung comp2, Erys %reads, and pm sympt co-localized on LG10m. QTLs for netw nint freq and fung div co-localized on LG6f. QTLs for netw nint freq and fung comp1 co-localized on LG9f (Figure 3.2 on page 86).

Trait	LG	Position LOD (cM)	<i>p</i> value	Marker	Bayes credible interval			% Vari- ation ex- plained	total PEV (a)	Estimate			
					Marker	Position LOD (cM)	Marker				Position LOD (cM)		
fung div	6f	20.62	0.017	s_1A7TMS_539	c6.loc11	c6.loc40	11	2.29	40	2.33	7.88	-	3.75
fung comp1	9f	33	0.089	c9.loc33	s_1BONE7_524	v_9193_250	8.48	1.86	41.12	2.89	7.13	-	-0.08
fung comp2	10m	10.07	0.056	s_1ARAYV_169	s_1A4JSZ_947	c10.loc26	0	1.84	26	1.73	6.56	-	0.06
bact comp1	10m	4	<0.001	c10.loc4	s_1A4JSZ_947	c10.loc5	0	7.43	5	7.38	16.5	-	-0.11
Erys %reads	10m	1	<0.001	c10.loc1	s_1A4JSZ_947	c10.loc4	0	4.29	4	3.18	8.12	-	-0.01
pm sympt	8f	0	<0.001	s_1A9STZ_1334	s_1A9STZ_1334	c8.loc5	0	6.89	5	6.39	7.47	-	-179.02
pm sympt	6m	20.43	0.008	s_1B10TI_1062	c6.loc18	c6.loc34	18	2.93	34	2.65	5.49	58.95	-151.56
pm sympt	10m	3	<0.001	c10.loc3	c10.loc2	v_11565_84	2	25.79	6	25.63	55.26	58.95	-499.31
netw mint freq	6f	2.22	3.17	v_3496_117	s_1A6BXG_1262	c6.loc17	0	2.87	17	1.7	7.13	16.64	-0.02
netw mint freq	9f	9.73	4.22	s_1AKBB0_560	c9.loc5	v_10305_387	5	3.44	20.22	3.3	9.26	16.64	0.02
netw int strngth	4m	28	3.01	c4.loc28	v_402_776	s_1AKV9K_1217	21.65	1.86	63.69	1.37	2.98	-	-0.01

Table 3.5: Summary for QTL for microbial community descriptors in oak full-sib mapping population, listing trait abbreviation, chromosome, position of the QTL (cM), LOD, p-value, significant marker, markers, positions and LOD of the bayes credible interval, the % variation explained by the QTL, total % variation explained by the additive model if applicable, and the additive contribution to the phenotype. Explanations of trait abbreviations are listed in Table 3.2 on page 79

3.4 Discussion

We found a major QTL on LG 10m for powdery-mildew lesion growth to be associated with QTLs for bacterial and fungal phyllosphere community composition. This QTL overlapped with a QTL for the number of reads of *E. alphitoides*, suggesting number of reads of *E.a.* is a good proxy of leaf infection by the pathogen (Jakuschkin *et al.*, Section 2). Furthermore the stability of this QTL over two different seasons (2006 and 2011) and experimental designs (inoculation experiment vs. natural infection) suggests strong genetic determinism of powdery mildew resistance in *Q. robur*. Powdery mildew infection is associated with changes in the fungal and bacterial community composition, the latter being however more subtle. *E. alphitoides* competes with other fungal species, and resident bacteria tend to facilitate or take advantage of the infection (Jakuschkin *et al.*, Section 2). Oak resistance/sensitivity to *Erysiphe alphitoides* might therefore drive phyllosphere microbial community. Overall this result extends the knowledge of fungal-mediated interactions on community demonstrated in insects (Tack *et al.*, 2012; Zargaran *et al.*, 2012; Milanovic *et al.*, 2014), and on microbial community (Section 2) and suggests that evolutionary history of pathogen resistance in a plant can shape the entire community of interacting species. The origin of oak powdery-mildew species in Europe is still unclear, also the coevolution between this species and European *Quercus* species (Desprez-Loustau *et al.*, 2011). Meanwhile, their coevolutionary history might have left signatures within oak genome and oak associated microbial community.

The putative oak resistance locus we found would open new exciting areas in trees-pathogens molecular evolution. Plant resistance genes are crucial in plant defense against pathogens (Jones & Dangl, 2006). Resistance genes often fail to provide long-term resistance in an agricultural context, and they often persist as long-lived balanced polymorphisms in nature (Stahl *et al.*, 1999; Tian *et al.*, 2002; Bakker *et al.*, 2006). The maintenance of polymorphisms is based on the balance between costs and benefits of resistance on one side and virulence on the other in a coevolving host-pathogen pair (Anderson & May, 1982; Frank, 1992), but, like for powdery-mildew infection, many plant-pathogen interactions lack such specificity (Barrett *et al.*, 2009). Karasov *et al.* (2014) showed that selection for a resistant gene in *A. thaliana* involves multiple non-homologous effectors and multiple pathogen species. The resistant gene polymorphism in *A. thaliana* may thus not be maintained through a tightly coupled interaction involving a single coevolved R gene and effector pair. The stable polymorphism is more likely maintained through complex and diffuse community-wide interactions. Our results are in line with their finding as the genomic region associated to a major pathogen in oak is also associated with the oak microbial community phenotype.

We found QTL for fungal diversity that possibly co-localize with QTLs for leaf stomatal density (Gailing *et al.*, 2008). Oak powdery-mildew causing *Erysiphe* species do not indirectly penetrate the mesophyll via stomata and we were not expecting genes for stomatal density to be linked to powdery-mildew infection. However,

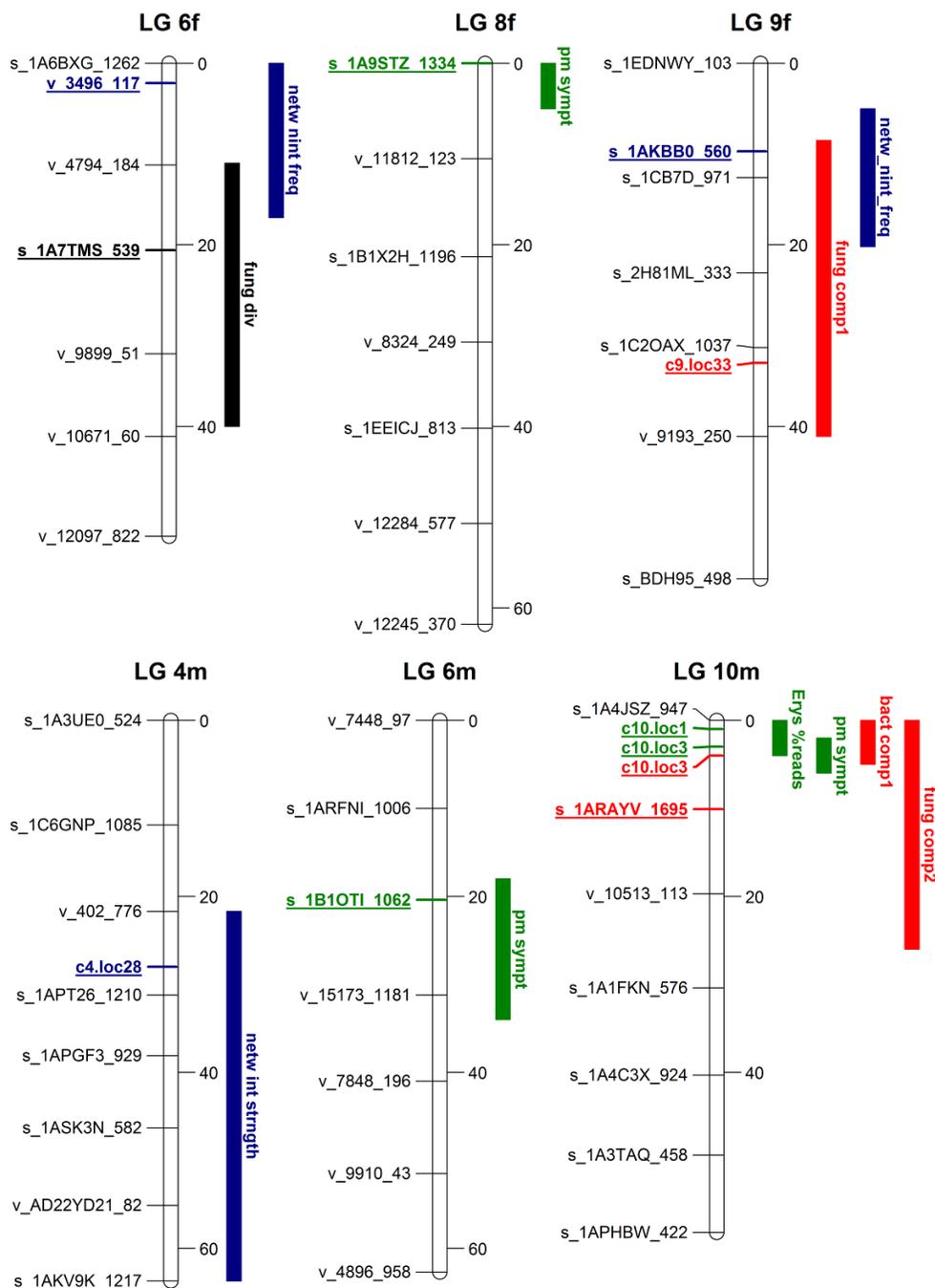


Figure 3.2: Microbial community quantitative trait loci (QTL) locations on *Q. robur* linkage groups (LG). Map distances are in centimorgans (cM). Colored bars show the confidence intervals of QTLs. Color coding: black for diversity traits, blue for network traits, red for microbial composition traits, blue for network traits. Markers with highest LOD value for each traits are marked in bold and underlined. Explanations of traits abbreviations are listed in Table 3.2 on page 79

variation for that trait might influence fungal diversity by selecting against species which need to colonize the leaves through stomatal openings. A reduced stomatal density might increase competition during the colonization step. Interval confidences of QTL for fungal diversity and QTL for Negative Interaction overlap on LG 6f (Figure 3.2). Because stomata constitute a resource-rich micro-environment compared with other leaf parts, we were expecting the genomic region associated with stomatal density to be linked to fungal and/or bacterial species in general. Further experiments are needed to get a better insight of the microbial phyllosphere community response to an increase in competition.

Finally, a genomic approach using oak genome (Plomion *et al.*, 2015) will shed light on putative plant function that matter in shaping phyllosphere community. Horton and colleagues' (2014) search for genes associated with phyllosphere here is a mark microbiota of *Arabidopsis thaliana* highlighted defense responses, signal transduction, kinase-related activities, and processes associated with the cell wall modification as putative biological functions associated with those microbial communities. Bodenhausen *et al.* (2014) established a synthetic community approach to study the effect of host genotype on bacterial composition and abundance in the phyllosphere. They showed strong influence of cuticle thickness and permeability on bacterial composition and density as it represents the initial interaction surface with microorganisms (Beattie, 2002). The anatomy and surface features of the leaves often determine plant resistance to infection and intrusion (Yadav *et al.*, 2005). Aspects of stomata, cuticle and leaf morphology can be involved in resistance (Niks & Rubiales, 2002). Trichomes can also act as physical barriers to infection (Martin & Glover, 2007) and this can delay germination by trapping the spores (Mmbaga *et al.*, 1994), because they reduce the probability of germ tubes to reach leaf surface (Wynn, 1976; Shaik, 1985). Leaf and cuticular or epidermal thicknesses have also been associated with powdery mildew resistance (Commenil, 1997). Indeed, biotrophy by powdery mildew is initially dependent on the morphology of the epidermal surface and the adhesion capacity of pathogen spores to the cuticle/cell wall (Vanacker *et al.*, 2000; Kruger *et al.*, 2002). All abovementioned traits might provide good candidate genes linked to the community phenotype QTLs we found in this study, and will be investigated in the next future with the help of a genomic approach using oak genome (Plomion *et al.*, 2015).

3.5 Bibliography

- Anderson, R.M. & May, R.M. (1982). Coevolution of hosts and parasites. *Parasitology*, 85, 411.
- Austin, M. (2013). Inconsistencies between theory and methodology: a recurrent problem in ordination studies. *Journal of Vegetation Science*, 24, 251–268.
- Bailey, J.K., Schweitzer, J.a., Ubeda, F., Koricheva, J., LeRoy, C.J., Madritch, M.D., Rehill, B.J., Bangert, R.K., Fischer, D.G., Allan, G.J. & Whitham, T.G. (2009). From genes to ecosystems: a synthesis of the effects of plant genetic factors across levels of organization. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364, 1607–1616.
- Bakker, E.G., Toomajian, C., Kreitman, M. & Bergelson, J. (2006). A genome-wide survey of R gene polymorphisms in Arabidopsis. *The Plant cell*, 18, 1803–1818.
- Bakker, M.G., Schlatter, D.C., Otto-Hanson, L. & Kinkel, L.L. (2014). Diffuse symbioses: roles of plant-plant, plant-microbe and microbe-microbe interactions in structuring the soil microbiome. *Molecular ecology*, 23, 1571–1583.
- Bálint, M., Tiffin, P., Hallström, B., O’Hara, R.B., Olson, M.S., Fankhauser, J.D., Piepenbring, M. & Schmitt, I. (2013). Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). *PloS one*, 8, e53987.
- Balint-Kurti, P., Simmons, S.J., Blum, J.E., Ballaré, C.L. & Stapleton, A.E. (2010). Maize leaf epiphytic bacteria diversity patterns are genetically correlated with resistance to fungal pathogen infection. *Molecular plant-microbe interactions*, 23, 473–84.
- Barrett, L.G., Kniskern, J.M., Bodenhausen, N., Zhang, W. & Bergelson, J. (2009). Continua of specificity and virulence in plant host-pathogen interactions: causes and consequences. *New Phytologist*, 183, 513–529.
- Beattie, G. (2002). Leaf surface waxes and the process of leaf colonization by microorganisms. In: *Phyllosphere microbiology* (eds. Lindow, S., Hecht-Poinar, E. & Elliott, V.). APS PRESS, pp. 3–26.
- Bengtsson, J., Eriksson, K.M., Hartmann, M., Wang, Z., Shenoy, B.D., Grelet, G.A., Abarenkov, K., Petri, A., Rosenblad, M.A. & Nilsson, R.H. (2011). Metaxa: a software tool for automated detection and discrimination among ribosomal small subunit (12S/16S/18S) sequences of archaea, bacteria, eukaryotes, mitochondria, and chloroplasts in metagenomes and environmental sequencing datasets. *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology*, 100, 471–5.

- Bent, S.J. & Forney, L.J. (2008). The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *The ISME journal*, 2, 689–695.
- Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M. & Vorholt, J.a. (2014). A Synthetic Community Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota. *PLoS Genetics*, 10, e1004283.
- Bolton, M.D. (2009). Primary metabolism and plant defense - fuel for the fire. *Molecular plant-microbe interactions : MPMI*, 22, 487–497.
- Brendel, O., Le Thiec, D., Scotti-Saintagne, C., Bodénès, C., Kremer, A. & Guehl, J.M. (2007). Quantitative trait loci controlling water use efficiency and related traits in *Quercus robur* L. *Tree Genetics & Genomes*, 4, 263–278.
- Broman, K.W., Wu, H., Sen, S. & Churchill, G.a. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19, 889–890.
- Bulgarelli, D., Schlaeppli, K., Spaepen, S., Ver Loren van Themaat, E. & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology*, 64, 807–38.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunencko, T., Zaneveld, J. & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335–336.
- Commenil, P. (1997). The development of the grape berry cuticle in relation to susceptibility to bunch rot disease. *Journal of Experimental Botany*, 48, 1599–1607.
- Compant, S., van der Heijden, M.G.a. & Sessitsch, A. (2010). Climate change effects on beneficial plant-microorganism interactions. *FEMS microbiology ecology*, 73, 197–214.
- Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C. & Vorholt, J.a. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 16428–16433.
- Derory, J., Scotti-Saintagne, C., Bertocchi, E., Le Dantec, L., Gaignic, N., Jauffres, A., Casasoli, M., Chancerel, E., Bodénès, C., Alberto, F. & Kremer, A. (2010). Contrasting relationships between the diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity*, 104, 438–48.

- Desprez-Loustau, M.L., Feau, N., Mougou-Hamdane, A. & Dutech, C. (2011). Inter-specific and intraspecific diversity in oak powdery mildews in Europe: coevolution history and adaptation to their hosts. *Mycoscience*, 52, 165–173.
- Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26, 2460–2461.
- Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10, 996–998.
- Fischer, D.G., Chapman, S.K., Classen, a.T., Gehring, C.a., Grady, K.C., Schweitzer, J.a. & Whitham, T.G. (2014). Plant genetic effects on soils under climate change. *Plant and Soil*, 379, 1–19.
- Frank, S.A. (1992). Models of plant-pathogen coevolution. *Trends in Genetics*, 8, 213–219.
- Friedman, J. & Alm, E.J. (2012). Inferring correlation networks from genomic survey data. *PLoS Computational Biology*, 8, e1002687.
- Friesen, M.L., Porter, S.S., Stark, S.C., von Wettberg, E.J., Sachs, J.L. & Martinez-Romero, E. (2011). Microbially Mediated Plant Functional Traits. *Annual Review of Ecology, Evolution, and Systematics*, 42, 23–46.
- Gailing, O., Langenfeld-Heyser, R., Polle, A. & Finkeldey, R. (2008). Quantitative trait loci affecting stomatal density and growth in a *Quercus robur* progeny: implications for the adaptation to changing environments. *Global Change Biology*, 14, 1934–1946.
- Gailing, O., Vornam, B., Leinemann, L. & Finkeldey, R. (2009). Genetic and genomic approaches to assess adaptive genetic variation in plants: Forest trees as a model. *Physiologia Plantarum*, 137, 509–519.
- Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R. & Reid, G. (2010). Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS ONE*, 5, e15406.
- Grattapaglia, D. & Sederoff, R. (1994). Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. *Genetics*, 137, 1121–1137.
- Hamilton, C.E. & Bauerle, T.L. (2012). A new currency for mutualism? Fungal endophytes alter antioxidant activity in hosts responding to drought. *Fungal Diversity*, 54, 39–49.
- Hamilton, C.E., Dowling, T.E. & Faeth, S.H. (2010). Hybridization in endophyte symbionts alters host response to moisture and nutrient treatments. *Microbial Ecology*, 59, 768–775.

- Horton, M.W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B.D., Subramanian, S., Vetter, M.M., Vilhjálmsson, B.J., Nordborg, M., Gordon, J.I. & Bergelson, J. (2014). Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nature communications*, 5, 5320.
- Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. & Sogin, M.L. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, 4, e1000255.
- Johnson, M.T.J. & Stinchcombe, J.R. (2007). An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution*, 22, 250–257.
- Jones, J.D.G. & Dangl, J.L. (2006). The plant immune system. *Nature*, 444, 323–329.
- Junker, R.R. & Tholl, D. (2013). Volatile organic compound mediated interactions at the plant-microbe interface. *Journal of chemical ecology*, 39, 810–25.
- Karasov, T.L., Kniskern, J.M., Gao, L., DeYoung, B.J., Ding, J., Dubiella, U., Lastra, R.O., Nallu, S., Roux, F., Innes, R.W., Barrett, L.G., Hudson, R.R. & Bergelson, J. (2014). The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature*, 512, 436–440.
- Kembel, S.W., O’Connor, T.K., Arnold, H.K., Hubbell, S.P., Wright, S.J. & Green, J.L. (2014). Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences*, 111, 13715–13720.
- Kruger, W., Carver, T. & Zeyen, R. (2002). Effects of inhibiting phenolic biosynthesis on penetration resistance of barley isolines containing seven powdery mildew resistance genes or alleles. *Physiological and Molecular Plant Pathology*, 61, 41–51.
- van Loon, L.C., Bakker, P.a. & Pieterse, C.M. (1998). Systemic resistance induced by rhizosphere bacteria. *Annual review of phytopathology*, 36, 453–483.
- Martin, C. & Glover, B.J. (2007). Functional aspects of cell patterning in aerial epidermis. *Current Opinion in Plant Biology*, 10, 70–82.
- Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. & Neufeld, J.D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, 13, 31.
- Mason, C.J., Pfammatter, J.A., Holeski, L.M. & Raffa, K.F. (2015). Foliar bacterial communities of trembling aspen in a common garden. *Canadian Journal of Microbiology*, 61, 1–7.

- Milanovic, S., Mihajlovic, L., Karadzic, D., Jankovsky, L., Aleksic, P., Jankovic-Tomanic, M. & Lazarevic, J. (2014). Effects of pedunculate oak tree vitality on gypsy moth preference and performance. *Archives of Biological Sciences*, 66, 1659–1672.
- Mmbaga, M.T., Steadman, J.R. & Roberts, J.J. (1994). Interaction of bean leaf pubescence with rust urediniospore deposition and subsequent infection density. *Annals of Applied Biology*, 125, 243–254.
- Mougou, A., Dutech, C. & Desprez-Loustau, M.L. (2008). New insights into the identity and origin of the causal agent of oak powdery mildew in Europe. *Forest Pathology*, 38, 275–287.
- Mougou-Hamdane, A. (2009). *Interaction Chene-oidium : caracterisation moleculaire et adaptation locale du parasite, resistance genetique de l'hoite*. Ph.D. thesis, Bordeaux 1.
- Niks, R.E. & Rubiales, D. (2002). Potentially durable resistance mechanisms in plants to specialised fungal pathogens. *Euphytica*, 124, 201–216.
- Nilsson, R.H., Bok, G., Ryberg, M., Kristiansson, E. & Hallenberg, N. (2009). A software pipeline for processing and identification of fungal ITS sequences. *Source Code for Biology and Medicine*, 4, 1.
- Nilsson, R.H., Kristiansson, E., Ryberg, M. & Hallenberg, N. (2008). Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evolutionary Bioinformatics*, 4, 193–201.
- Nilsson, R.H., Veldre, V., Hartmann, M., Unterseher, M., Amend, A., Bergsten, J., Kristiansson, E., Ryberg, M., Jumpponen, A. & Abarenkov, K. (2010). An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. *Fungal Ecology*, 3, 284–287.
- Partida-Martínez, L.P. & Heil, M. (2011). The microbe-free plant: fact or artifact? *Frontiers in plant science*, 2, 100.
- Peñuelas, J., Rico, L., Ogaya, R., Jump, a.S. & Terradas, J. (2012). Summer season and long-term drought increase the richness of bacteria and fungi in the foliar phyllosphere of *Quercus ilex* in a mixed Mediterranean forest. *Plant biology (Stuttgart, Germany)*, 14, 565–75.
- Plomion, C., Aury, J.M., Amselem, J., Alaeitabar, T., Barbe, V., Belser, C., Bergès, H., Bodénès, C., Boudet, N., Boury, C., Canaguier, A., Couloux, A., Da Silva, C., Duplessis, S., Ehrenmann, F., Estrada-Mairey, B., Fouteau, S., Francillonne, N.,

- Gaspin, C., Guichard, C., Klopp, C., Labadie, K., Lalanne, C., Le Clainche, I., Leplé, J.C., Le Provost, G., Leroy, T., Lesur, I., Martin, F., Mercier, J., Michotey, C., Murat, F., Salin, F., Steinbach, D., Faivre-Rampant, P., Wincker, P., Salse, J., Quesneville, H. & Kremer, A. (2015). Decoding the oak genome: public release of sequence data, assembly, annotation and publication strategies. *Molecular Ecology Resources*, pp. n/a–n/a.
- R Core Team (2013). R: A Language and Environment for Statistical Computing.
- Raghavendra, A.K.H. & Newcombe, G. (2013). The contribution of foliar endophytes to quantitative resistance to *Melampsora* rust. *New Phytologist*, 197, 909–918.
- Reinhold-Hurek, B. & Hurek, T. (2011). Living inside plants: bacterial endophytes. *Current opinion in plant biology*, 14, 435–43.
- Robin, C., Mougou-Hamdane, A., Gion, J.M., Kremer, A. & Desprez-Loustau, M.L. (2010). Quantitative trait loci for resistance to pathogens in pedunculate oak (*Quercus robur* L.). In: *Forest ecosystem genomics and adaptation*. Madrid.
- Saintagne, C., Bodénès, C., Barreneche, T., Pot, D., Plomion, C. & Kremer, a. (2004). Distribution of genomic regions differentiating oak species assessed by QTL detection. *Heredity*, 92, 20–30.
- Schoch, C.L., Seifert, K.a., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.a., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, a.N., Wingfield, M.J., Aime, M.C., An, K.D., Bai, F.Y., Barreto, R.W., Begerow, D., Bergeron, M.J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, a.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, a., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.a., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.a., Ge, Z.W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.B., Hosaka, K., Houbroken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., Kovacs, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, a.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, a.R., Methven, a.S., Meyer, W., Moncalvo, J.M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.a., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, a., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos,

- S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.O., Sung, G.H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.a., Urbina, H., Vagvolgyi, C., Vialle, a., Vu, T.D., Walther, G., Wang, Q.M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, a., Zamora, J.C., Zhang, N., Zhuang, W.Y. & Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109, 6241–6246.
- Schweitzer, J.A., Bailey, J.K., Bangert, R.K., Hart, S.C. & Whitham, T.G. (2006). The role of plant genetics in determining above- and below-ground microbial communities. In: *Microbial ecology of aerial plant surfaces* (eds. Bailey, M.J., Lilley, A.K., Timms-Wilson, T.M. & Spencer-Phillips, P.T.N.). CABI; Wallingford; UK, pp. 107–119.
- Schweitzer, J.A., Bailey, J.K., Fischer, D.G., LeRoy, C.J., Lonsdorf, E.V., Whitham, T.G. & Hart, S.C. (2008). Plant-soil-microorganism interactions: heritable relationship between plant genotype and associated soil microorganisms. *Ecology*, 89, 773–781.
- Schweitzer, J.a., Fischer, D.G., Rehill, B.J., Wooley, S.C., Woolbright, S.a., Lindroth, R.L., Whitham, T.G., Zak, D.R. & Hart, S.C. (2010). Forest gene diversity is correlated with the composition and function of soil microbial communities. *Population Ecology*, 53, 35–46.
- Scotti-Saintagne, C., Bodénès, C., Barreneche, T., Bertocchi, E., Plomion, C. & Kremer, A. (2004). Detection of quantitative trait loci controlling bud burst and height growth in *Quercus robur* L. *Theoretical and Applied Genetics*, 109, 1648–1659.
- Shaik, M. (1985). Race-Nonspecific Resistance in Bean Cultivars to Races of *Uromyces appendiculatus* var. *appendiculatus* and its Correlation with Leaf Epidermal Characteristics. *Phytopathology*, 75, 478.
- Stahl, E.a., Dwyer, G., Mauricio, R., Kreitman, M. & Bergelson, J. (1999). Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature*, 400, 667–671.
- Tack, A.J.M., Gripenberg, S. & Roslin, T. (2012). Cross-kingdom interactions matter: fungal-mediated interactions structure an insect community on oak. *Ecology Letters*, 15, 177–185.
- Tian, D., Araki, H., Stahl, E., Bergelson, J. & Kreitman, M. (2002). Signature of balancing selection in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11525–11530.

- Ulrich, K., Ulrich, A. & Ewald, D. (2008). Diversity of endophytic bacterial communities in poplar grown under field conditions. *FEMS Microbiology Ecology*, 63, 169–180.
- Vanacker, H., Carver, T.L. & Foyer, C.H. (2000). Early H₂O₂ accumulation in mesophyll cells leads to induction of glutathione during the hyper-sensitive response in the barley-powdery mildew interaction. *Plant physiology*, 123, 1289–1300.
- Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. (2015). Tansley review The importance of the microbiome of the plant holobiont. *New Phytologist*.
- Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73, 5261–5267.
- Whitham, T.G., Bailey, J.K., Schweitzer, J.a., Shuster, S.M., Bangert, R.K., LeRoy, C.J., Lonsdorf, E.V., Allan, G.J., DiFazio, S.P., Potts, B.M., Fischer, D.G., Gehring, C.a., Lindroth, R.L., Marks, J.C., Hart, S.C., Wimp, G.M. & Wooley, S.C. (2006). A framework for community and ecosystem genetics: from genes to ecosystems. *Nature reviews. Genetics*, 7, 510–523.
- Whitham, T.G., Young, W.P., Martinsen, G.D., Gehring, C.a., Schweitzer, J.a., Shuster, S.M., Wimp, G.M., Fischer, D.G., Bailey, J.K., Lindroth, R.L., Woolbright, S. & Kuske, C.R. (2003). Community and ecosystem genetics: A consequence of the extended phenotype. *Ecology*, 84, 559–573.
- Wynn, W.K. (1976). Appressorium Formation over Stomates by the Bean Rust Fungus: Response to a Surface Contact Stimulus. *Phytopathology*, 66, 136.
- Yadav, R.K.P., Karamanoli, K. & Vokou, D. (2005). Bacterial Colonization of the Phyllosphere of Mediterranean Perennial Species as Influenced by Leaf Structural and Chemical Features. *Microbial Ecology*, 50, 185–196.
- Zargaran, M.R., Erbilgin, N. & Ghosta, Y. (2012). Changes in Oak Gall Wasps Species Diversity (Hymenoptera: Cynipidae) in Relation to the Presence of Oak Powdery Mildew (*Erysiphe alphitoides*). *Zoological Studies*, 51, 175–184.
- Zeng, Z.B. (1994). Precision mapping of quantitative trait loci. *Genetics*, 136, 1457–68.

Chapter 4

Association mapping of phyllosphere microbiota properties in *Quercus robur* L.

Boris Jakuschkin^{1,2}, Corinne Vacher^{1,2}, Santiago C. González-Martínez^{1,2}, Cécile Robin^{1,2}, Virgil Fievet^{1,2}

1. INRA, UMR1202 BIOGECO, F-33610 Cestas, France

2. Univ. Bordeaux, BIOGECO, UMR 1202, F-33615 Pessac, France

4.1 Introduction

In the previous QTL experiment we have identified 11 QTLs for eight traits describing phyllosphere microbial structure (fung comp1, fung comp2, bact comp1), fungal diversity (fung div), association network properties (netw nint freq, netw int strngth), and severity of powdery-mildew infection (Erys %reads, pm sympt). The majority of these QTLs were minor QTLs explaining less than 10% of the phenotypic variation. As mentioned earlier the phyllosphere as a microbial habitat is likely influenced by a broad set of traits, which themselves are complex and are controlled by multiple genes, e.g. genetic control of phenolic contents (DeWoody *et al.*, 2013). The outcome of each trait is determined by the combined effects of variants in numerous genes and nongenetic factors. Therefore many genetic loci of minor effects are expected for each of these traits. The main limitation of a QTL experiment is that only those loci segregating in the mapping population under consideration can be identified. Because the number of recombination events per meiosis is relatively small the mapping resolution of a QTL experiment is typically low, i.e. the confidence intervals of detected QTLs are very big. Another difficulty of detecting a QTL of a polygenic trait is that the effect sizes of individual variants are likely

too small to be detected via cosegregation within pedigrees. Rapid advances in sequencing technologies allowed for discovery and scoring of vast numbers of SNPs allowing to design and perform genome wide linkage disequilibrium association mapping studies (GWASs) on unrelated individuals. Association analysis searches for correlation between genomic region and phenotype in population, whereas a linkage analysis tests for correlation in families. Association analysis can achieve much better mapping resolution than linkage analysis, because close relatives share stretches of genetic material over larger parts of chromosomes than do unrelated individuals (Hirschhorn & Gajdos, 2011). However, GWA mapping suffers from the limitation that it generates false positives due to population structure. Statistical methods to control for population structure have been developed to reduce the inflation of false positives associations. Knowledge of kinship relationship among individuals in the population provides additional important information for control of population stratification (Gauderman *et al.*, 1999). But an alternative is the complementary use of traditional QTL mapping in controlled crosses and the use of GWA mapping (Brachi *et al.*, 2010; Horton *et al.*, 2014). Genome-wide association studies (GWASs) are most often used for studying the genetic basis of human diseases and by 2013 they have identified over 2000 SNPs associated with more than 300 polygenic diseases and traits (Manolio, 2013). But this methodology also has great utility for studying the natural variation in model organisms like *Arabidopsis thaliana* (Atwell, 2010) and mouse (Flint & Eskin, 2012), in economically important plants (Ogura & Busch, 2015) and cattle (Olsen *et al.*, 2011). Here we aim to identify the genetic regions that contribute to the structure of oak phyllosphere community. The results of this study should complement our findings from the previously performed QTL experiment. I present and discuss preliminary analyses.

4.2 Material and Methods

Population structure and Sampling

The open-air experimental site was located in a clearing surrounded by pine trees (*Pinus pinaster* Ait.) and broadleaved species (predominantly *Q. robur*) in the South-West of France (Pierroton, Cestas, 44°44'23" N, 0°46'1" W). The site was fenced to prevent grazing by mammalian herbivores. This study site was designed to study the effects of powdery-mildew infection on oak seedlings (Desprez-Loustau *et al.*, 2014) and to perform genome-wide association study (GWASs) to identify the genetic basis of powdery-mildew resistance (Barres *et al.* unpublished work). The half-sib trees originate from 15 open pollinated progenies. The trees were planted in 2009 from acorns collected in 2008. They were planted in nine blocks: three replicated blocks per treatment (control, fungicide treated, inoculated with *E. alphioides*), at a 10 cm x 10 cm distance in each block, in eight rows of 38 plants. And the trees of the inoculation treatment were not treated for one year previously to our

sampling date. On 3rd of July 2012, we sampled three leaves from each of the 218 trees, which were previously successfully genotyped and used in the GWAS study by Barres *et al.*. The trees were sampled within the three replicated blocks, which previously belonged to the control treatment (no treatment). All trees were sampled on the same day. At that moment the trees were three years old and were between 30 and 80 cm tall. Additionally we sampled 12 of the 15 maternal trees (12 leaves per tree, four leaves per cardinal direction, 144 leaves), which were interspersed among other trees at the campus of Inra Pierroton research facility, but had non-overlapping canopies with their neighbors. Three maternal trees were fallen in previous years due to construction works. The leaves were sampled at the lowest accessible branches at approximate height of eight to ten meters. Each leaf was cut at the base of the petiole and placed into a sterile bag containing silica gel. The bags, sealed and made airtight, were then stored in dark at +18°C in the laboratory. This storage method was shown to be adequate for studying oak foliar fungal communities (Section 2.7 on page 68).

DNA extraction

Leaf DNA extraction was performed four months after the sampling date. We prepared the leaf material for the DNA extraction in a laminar flow hood. In order to avoid contaminating the samples, the hood and all tools were disinfected with sodium hypochlorite followed by 70% ethanol and 40 min of UV exposure. Plastic ware and tungsten carbide beads were autoclaved for 2 h prior to utilization. Similarly, the tools used for leaf handling were sterilized prior to each use, by dipping them into 70% alcohol and flaming over an electric bunsen burner. Using a metal corer we excised two discs of 0.5 cm² each on both sides of the middle leaf vein. The four discs were evenly distributed on the leaf, with their position not being chosen according to *E. alphitoides* symptoms. The four leaf discs were placed in a collection microtube (QIAGEN DNeasy® 96 Plant Kit) containing one tungsten carbide bead. We used Geno/Grinder® (SPEX SamplePrep, Metuchen, NJ, USA) to homogenize the leaf material at 1 600 strokes per minute for 90 s. Homogenized samples were stored at -80 °C overnight. We used the QIAGEN DNeasy 96 Plant Kit to extract the DNA (following the protocol for frozen plant tissue). We determined the DNA concentration with NanoDrop 8000 and diluted all DNA working solutions to 1 ng µl⁻¹. We pooled the samples (798) by tree obtaining 218 DNA samples for the progeny and 12 for the maternal trees.

Amplification and sequencing of the fungal assemblages

To characterize the fungal assemblages on oak leaves we targeted the internal transcribed spacer 1 (ITS1) sublocus of the ITS region of the nuclear ribosomal repeat unit (Schoch *et al.*, 2012). Hereby we have used the primers and the protocol described in Section 3.2 on page 74. The DNA concentration in each sample was

quantified using the Nanodrop8000 and pooled to an equimolar concentration to obtain the amplicon libraries, five in total. The final DNA libraries were sequenced on Ion PGM™ Platform using IonXpress™ Library Kit (Life Technologies, Thermo Fischer Scientific, Waltham, MA, USA) at the Centre de Génomique Fonctionnelle de Bordeaux (CGFB, Bordeaux, France).

Amplification and sequencing the bacterial assemblages

To characterize the bacterial assemblages on oak leaves we targeted the hypervariable region 6 (V6) of the 16S small ribosomal subunit gene (16S rRNA). We used the primers and the protocol described in Section 3.2 on page 74. However we designed new barcodes of uniform length of 8 nucleotides, because the company upgraded their sequencing kits allowing for reads length of 250 bp. New barcode design was used to allow for sequencing of all samples in one sequencing run and to facilitate sequence processing during the demultiplexing step. Amplicon samples were then individually quantified using the Nanodrop8000 and pooled to an equimolar concentration. The final amplicon library was sequenced on the Illumina platform at the sequencing center of Bordeaux (Plateforme génome-transcriptome Bordeaux). Paired-end sequencing allowed for full coverage of the V6 region with reads overlapping at full length. The final DNA library was sequenced on the Illumina platform GAIIX (Illumina, Inc, San Diego, CA, USA) at the Centre de Génomique Fonctionnelle de Bordeaux (CGFB, Bordeaux, France).

Processing of the fungal Ion Torrent sequencing data

Ion torrent sequencing output file was a mixture of sense and anti-sense reads. We created a reverse complement of the sequencing file. Both files were processed in the same way until the ITS extraction step, after which the output files of Fungal ITS Extractor 1.1 (Nilsson *et al.*, 2010) were joined to a single file. We used QIIME v.1.7.0 (Caporaso *et al.*, 2010) for quality filtering, demultiplexing, and denoising. After the removal of tags and primers, all sequences shorter than 100 nucleotides were discarded. Quality filtering was done with a sliding window test of quality scores having a window size of 50 nucleotides and a minimum quality score of 25. The sequences were truncated at the beginning of the poor quality window and tested for the minimum sequence length of 100 nucleotides. Highly conserved ribosomal genes flanking the ITS1 marker were removed using Fungal ITS Extractor 1.1. ITS1 sequences shorter than 100 nucleotides were discarded and sequences longer than 280 nucleotides were trimmed at the 3'-end. We discarded singleton reads before clustering to minimize spurious Operational Taxonomic Units (OTUs). OTU clustering was done with UPARSE-OTU (Edgar, 2013), an algorithm implemented within Usearch v.7 (Edgar, 2010) that performs chimera filtering and OTU clustering simultaneously. The algorithm identifies a set of OTU representative sequences at the 97% sequence similarity threshold, that corresponds roughly to the differences at the species level

(Nilsson *et al.*, 2008). We filtered out *E. alphitoides* sequences as described in Section 2.2 on page 47. Taxonomy assignments were done by BLAST searches against Fungal ITS Database (Nilsson *et al.*, 2009) (as of March 15, 2012). We modified the taxonomic file associated with this database to equalize the length of taxonomic ranks, by using the NCBI Taxonomy Browser. Additionally we corrected a few false taxonomic assignments in the case of *Erysiphe* species and removed a few non-fungal sequences. We performed BLAST searches using QIIME (v.1.7.0) with standard settings (e-value: 0.001, minimum percent identity: 90) to assign the taxonomy to the most abundant sequence of each OTU. We also performed BLAST searches against the Nucleotide collection database (nr/nt database) to exclude possible *Q. robur* sequences. We then constructed the OTU table in QIIME and transformed it from BIOME into tab delimited text format for downstream analyses with R (R Core Team, 2013).

Processing of the bacterial Illumina sequencing data

We assembled the paired-end reads using PANDAseq (Masella *et al.*, 2012). Reads were then demultiplexed following the method described in Gloor *et al.* (2010) using a custom Bash shell script. We removed oak chloroplast sequence prior to the clustering step. OTU clustering and OTU table construction were done by following the same procedure than for fungi. Taxonomic assignments were performed with the RDP Classifier program (v.2.5) (Wang *et al.*, 2007) as it is implemented into QIIME (v.1.7.0), with standard settings and Greengenes 13_5 as the reference sequences. The origin of OTU representative sequences (archaeal, bacterial, nuclear eukaryote, mitochondrial or chloroplast origin) was identified using Metaxa (Bengtsson *et al.*, 2011) and we retained only the bacterial OTUs.

Statistical analysis and choice of traits

All data handling, trait calculations and QTL analysis were performed in the R Program, version 3.1.3 (R Core Team, 2013). We measured bacterial and fungal diversity using the inverse Simpson index, which is defined as $\frac{1}{D}$, where $D = \sum p_i^2$ and p_i is the proportional abundance of species i . This index is comparatively insensitive to numerically minor constituents (Bent & Forney, 2008) and it better differentiates between highly diverse samples, which is the case in our data. We chose the percentage of *Erysiphe alphitoides* reads in fungal samples as proxy for the powdery mildew infection (Chapter 2 on page 54). As the measurement of the community composition we chose the axis projections of principal coordinate analysis (PCoA). We have used the Bray-Curtis dissimilarity index calculated on the fungal and bacterial datasets (Figure 4.1 on the next page a) and b) respectively). Bray-Curtis index was previously shown to perform best in recovering environmental gradients (Austin, 2013). Trait abbreviations and explanations are summarized in Table 4.1 on page 102. Summary statistics of each trait are given in Table 4.2 on page 102.

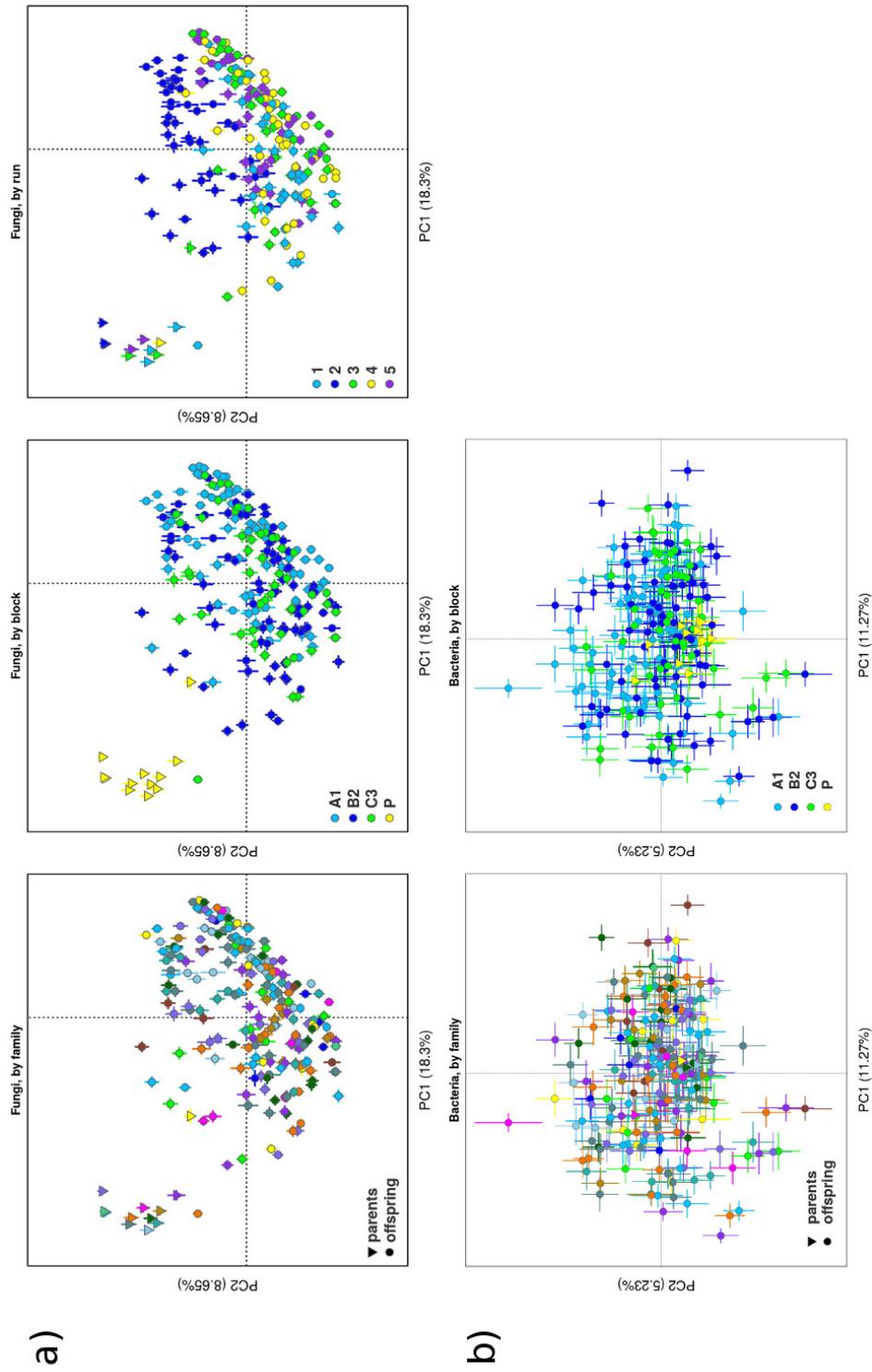


Figure 4.1: PCoA of a) fungal and b) bacterial community composition data using Bray-Curtis dissimilarity index. Dots represent mean values of 100 PCoAs on respective rarefied datasets, whiskers show the SD. a) from left to right: there is no obvious family effect, maternal samples (triangles) cluster separately from offspring samples (dots), coloration by run/ sequencing pool shows distinctive run effect.. b) left panel: there is no obvious family effect, right panel: there is no distinctive difference between maternal and offspring samples.

Abbreviation	Trait explanation
fung div	fungus diversity (inverse Simpson index for fungal data)
bact div	bacterial diversity (inverse Simpson index for bacterial data)
fung comp1	1st principal component of PCoA with Bray-Curtis distance on fungal community data
fung comp2	2nd principal component on fungal data
bact comp1	1st principal component of PCoA with Bray-Curtis distance on bacterial community data
bact comp2	2nd principal component on bacterial data
Erys %reads	percentage of <i>Erysiphe alphitoides</i> reads in the sample

Table 4.1: Microbial community traits, listing their abbreviation and explanation of traits

4.3 Results

Sequencing data

Fungal data

Out of the initial 2089540 raw reads, 1561541 were found to be of poor read quality, too short, chimeric or singletons, leaving 527999 (25%) sequences for subsequent analyses. The number of sequences per sample ranged from 754 to 3709 with the

Trait	Phenotyped offspring	Min	Max	Mean	SD
fung div	217	2.39	61.63	26.84	11.35
bact div	217	2.54	33.58	16.93	5.79
fung comp1	217	-0.5	0.30	0.03	0.16
fung comp2	217	-0.24	0.27	-0.01	0.12
bact comp1	217	-0.38	0.39	0	0.16
bact comp2	217	-0.33	0.36	0	0.10
Erys %reads	217	0	0.28	0.01	0.03

Table 4.2: Phenotypic values in microbial traits in the *Q. robur* offspring population.

Trait	Phenotyped offspring	Min	Max	Mean	SD
fung div	11	3.15	16.57	9.04	3.75
bact div	11	7.59	28.14	21.25	6.36
fung comp1	11	-0.55	-0.25	-0.48	0.08
fung comp2	11	0.14	0.37	0.26	0.07
bact comp1	11	-0.10	0.04	-0.01	0.05
bact comp2	11	-0.12	0.06	-0.05	0.05
Erys %reads	11	0	0.01	0	0

Table 4.3: Phenotypic values in microbial traits in the *Q. robur* maternal trees.

mean of 2158 (median: 2098). A total of 1124 OTUs were identified. Out of these OTUs a putative taxonomy could be assigned to 934 OTUs. The 190 not-assigned OTUs represented 2.11% of the reads and were confirmed as uncultured fungi. The samples were grouped into five amplicon libraries and were sequenced in five separate sequencing runs. Analysis of a possible sequencing run effect on fungal diversity and fungal composition revealed that the amplicon library number two was significantly different from other libraries (fung div, GLM, $F_{4,212}=17.62$, $p<0.001$; fung comp2, GLM, $F_{4,212}=40.67$, $p<0.001$). Additionally sequences from the amplicon library one were significantly different from the libraries three and five, when tested for fungal composition (fung comp1, GLM, $F_{4,212}=4.11$, $p=0.003$). The data revealed also block effect for diversity (fung div, LMM with library as random effect, $F_{2,215}=3.39$, $p=0.03$) and composition (fung comp1, LMM with library as random effect, $F_{2,215}=13.93$, $p<0.001$). Overall, there was no family effect, but on the percentage of *Erysiphe alphitoides* reads (Erys %reads, LMM with library and block as random effects, $F_{13,204}=3.18$, $p=0.0002$).

Bacterial data

Sequencing of the bacterial amplicon pool yielded $22.9 \cdot 10^6$ sequences, of which $22.5 \cdot 10^6$ were chloroplast sequences, mitochondria sequences, poor quality, too short, chimeric or singletons, leaving 391867 sequences for subsequent analyses. In total 1395 bacterial OTUs were identified. The number of reads per sample ranged from 123 to 12440 with a mean of 1633 and a median of 1052. All OTUs were assigned at differing taxonomical resolution. We found no block effect on bacterial diversity and on bact comp1, however there was block effect on bact comp2 (GLM, $F_{2,214}=6.84$, $p=0.001$). Like for the fungal data, we found no family effect on bacterial diversity and composition.

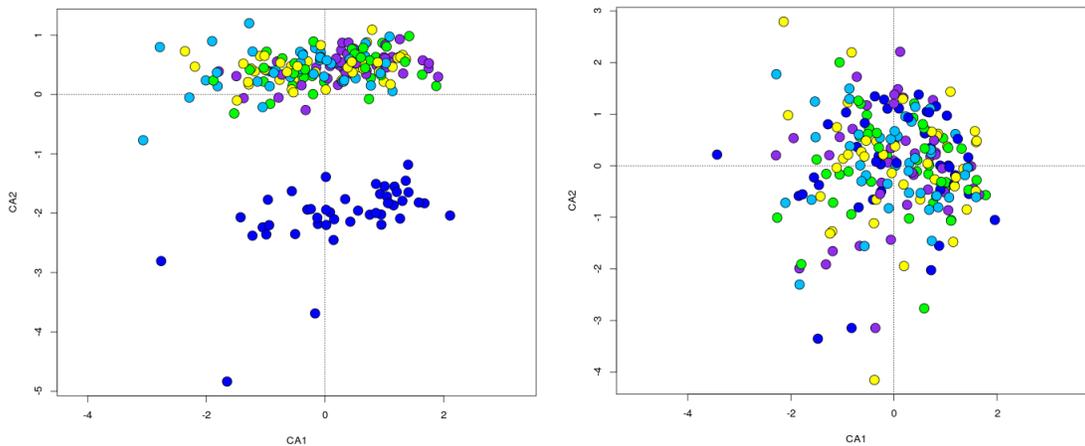


Figure 4.2: Canonical correspondence analysis (CCA) of the fungal dataset using `cca` function from `vegan` package in R (Oksanen *et al.*, 2014). Only the offspring data is shown to emphasize the run effect. Different colors correspond to different sequencing runs. Left panel shows obvious run effect. Right panel shows CCA with run being partialled out from the analysis as confounding effect.

Differences between maternal and offspring trees

It is notable, that we observed differences between maternal mature trees and young offspring trees in fungal and bacterial diversity. Samples of maternal trees had lower fungal diversity (LMM with run as random effect, $F_{1,229}=34.97$, $p<0.001$) and higher bacterial diversity (GLM, $F_{1,229}=7.72$, $p=0.005$).

4.4 Discussion

The confounding effects of sequencing run/amplicon library in fungal data and block effect in fungal and bacterial data need to be dealt with before conducting the GWAS. One possibility is to work on the residuals of the model, which takes these effects into account. Alternatively, for the microbial composition data, it is possible to use constrained ordination analysis to 'partial-out' the confounding effects, for instance canonical correspondence analysis (CCA), and then use the residuals from CCA as phenotypes in GWAS (Figure 4.2) (Horton *et al.*, 2014).

In the preliminary analysis we found no family effect on bacterial and fungal diversities and on the bacterial and fungal compositional data. Following explanations are possible: (a) it could be that the small sizes of the families and large variations in the traits within each family blur the differences between the families and do not allow to reach the statistical power to show significance due to the large influence of outliers; (b) it is possible, that our chosen traits are not sensitive enough to depict the differences between the families, for instance the first two coordinates of PCoA of the bacterial data explain 16.53% of the total compositional variability between

the samples, thus much of the variability remains unexplained, which also could explain the differences between the families; (c) the structure and composition of the microbial community are shaped by many genes of variable minor effects and the differences between the families in some of these genes could be diminished due to high number of possible gene combinations within each progeny; (d) maternal trees originate from a very narrow geographical range (within radius of about one hundred meters) and they could be too similar in the scored traits (Table 4.3 on page 103, exception: bacterial diversity) to produce family signature between the progenies due to the adaptation to the similar environmental pressures experienced by their ancestors; (e) the non-genetic component of these traits is too big and obscures the differences among genotypes: for instance, the stochastic variation during the initial colonization process of the phyllosphere may determine the developmental trajectory of the microbial community on newly emerged leaves (Maignien *et al.*, 2014); and (f) it is possible that poorly characterized/sequenced taxa, occasional sequencing artifacts, and the choice of OTU clustering threshold obscure the effects of host genetics, which probably has the strongest effect on most common/sequenced taxa within samples (Horton *et al.*, 2014). Environmental factors may conceal the genetic differences – rain, for instance, represents a strong perturbation and can wash off big parts of the phylloplane community (Hirano & Upper, 2000), changing the proportions of the microbial components and inducing recolonization processes anew. Also the trees within each block grew so close that their canopies were in contact with each other and visually formed a hedge. This could contribute to the reduced differences in microbial community traits among closely spaced trees. Nevertheless, we can expect to detect SNPs linked to genes involved in the structuring of phyllosphere community, especially because of (c) possible involvement of many loci of minor effect. Barres *et al.* (unpublished) has successfully identified 18 SNPs linked to the differences in powdery-mildew susceptibility using the GWA analysis within the same experimental setup. As we have seen previously powdery-mildew is linked to changes in microbial communities (Chapters 2 and 3).

Another notable result is the striking difference in fungal and bacterial diversity between old and young trees and the fact that these differences do not have the same sign: mature trees have lower fungal and higher bacterial diversities than young trees. This result needs to be viewed with caution because of the different locations of the trees and therefore experienced different environmental conditions. Nonetheless it inspires to speculate about possible reasons. First striking difference is the position of the sampled leaves. Leaves of the young trees were sampled about 40–60 cm above-ground, while the height of lower branches of the sampled mature trees was about 8–10 m. Differences in the leaf height will likely mean differences in inoculum sources and pressures. It is likely that grass as inoculum source will have greater impact on low positioned leaves, as well as soil. In the same time the lack of protective upper leaf stories of a big canopy results in stronger fluctuations of humidity levels, and prolonged periods of UV radiation in young oak plants. But leaf morphology of trees of different age is also different. Leaves do not only differ in size but also

in other characteristics, like leaf thickness and leaf water content. For instance, leaf water content of eight perennial species was shown to be strongly linked to the size of epiphytic bacterial population (Yadav *et al.*, 2005).

4.5 Bibliography

- Austin, M. (2013). Inconsistencies between theory and methodology: a recurrent problem in ordination studies. *Journal of Vegetation Science*, 24, 251–268.
- Bengtsson, J., Eriksson, K.M., Hartmann, M., Wang, Z., Shenoy, B.D., Grelet, G.A., Abarenkov, K., Petri, A., Rosenblad, M.A. & Nilsson, R.H. (2011). Metaxa: a software tool for automated detection and discrimination among ribosomal small subunit (12S/16S/18S) sequences of archaea, bacteria, eukaryotes, mitochondria, and chloroplasts in metagenomes and environmental sequencing datasets. *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology*, 100, 471–5.
- Bent, S.J. & Forney, L.J. (2008). The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *The ISME journal*, 2, 689–695.
- Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., Bergelson, J., Cuguen, J. & Roux, F. (2010). Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS genetics*, 6, e1000940–e1000940.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7, 335–336.
- Desprez-Loustau, M.L., Saint-Jean, G., Barrès, B., Dantec, C.F. & Dutech, C. (2014). Oak powdery mildew changes growth patterns in its host tree: host tolerance response and potential manipulation of host physiology by the parasite. *Annals of Forest Science*, 71, 563–573.
- DeWoody, J., Viger, M., Lakatos, F., Tuba, K., Taylor, G. & Smulders, M.J.M. (2013). Insight into the genetic components of community genetics: QTL mapping of insect association in a fast-growing forest tree. *PloS one*, 8, e79925.
- Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26, 2460–2461.
- Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10, 996–998.
- Flint, J. & Eskin, E. (2012). Genome-wide association studies in mice. *Nature Reviews Genetics*, 13, 807–817.

- Gauderman, W.J., Witte, J.S. & Thomas, D.C. (1999). Family-Based Association Studies. *Journal of the National Cancer Institute Monographs*, pp. 31–37.
- Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R. & Reid, G. (2010). Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS ONE*, 5, e15406.
- Hirano, S.S. & Upper, C.D. (2000). Bacteria in the Leaf Ecosystem with Emphasis on *Pseudomonas syringae* - a Pathogen, Ice Nucleus, and Epiphyte. *Microbiology and Molecular Biology Reviews*, 64, 624–653.
- Hirschhorn, J.N. & Gajdos, Z.K.Z. (2011). Genome-Wide Association Studies: Results from the First Few Years and Potential Implications for Clinical Medicine. *Annual Review of Medicine*, 62, 11–24.
- Horton, M.W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B.D., Subramanian, S., Vetter, M.M., Vilhjálmsson, B.J., Nordborg, M., Gordon, J.I. & Bergelson, J. (2014). Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nature communications*, 5, 5320.
- Maignien, L., DeForce, E.A., Chafee, M.E., Eren, A.M. & Simmons, S.L. (2014). Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *mBio*, 5, 10.
- Manolio, T.a. (2013). Bringing genome-wide association findings into clinical use. *Nature reviews. Genetics*, 14, 549–58.
- Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. & Neufeld, J.D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, 13, 31.
- Nilsson, R.H., Bok, G., Ryberg, M., Kristiansson, E. & Hallenberg, N. (2009). A software pipeline for processing and identification of fungal ITS sequences. *Source Code for Biology and Medicine*, 4, 1.
- Nilsson, R.H., Kristiansson, E., Ryberg, M. & Hallenberg, N. (2008). Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evolutionary Bioinformatics*, 4, 193–201.
- Nilsson, R.H., Veldre, V., Hartmann, M., Unterseher, M., Amend, A., Bergsten, J., Kristiansson, E., Ryberg, M., Jumpponen, A. & Abarenkov, K. (2010). An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. *Fungal Ecology*, 3, 284–287.

- Ogura, T. & Busch, W. (2015). From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. *Current Opinion in Plant Biology*, 23, 98–108.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H. & Wagner, H. (2014). *vegan*: Community Ecology Package. R package version 2.2-0.
- Olsen, H.G., Hayes, B.J., Kent, M.P., Nome, T., Svendsen, M., Larsgard, a.G. & Lien, S. (2011). Genome-wide association mapping in Norwegian Red cattle identifies quantitative trait loci for fertility and milk production on BTA12. *Animal Genetics*, 42, 466–474.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*.
- Schoch, C.L., Seifert, K.a., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.a., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, a.N., Wingfield, M.J., Aime, M.C., An, K.D., Bai, F.Y., Barreto, R.W., Begerow, D., Bergeron, M.J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, a.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, a., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.a., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.a., Ge, Z.W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.B., Hosaka, K., Houbraken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., Kovacs, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, a.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, a.R., Methven, a.S., Meyer, W., Moncalvo, J.M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.a., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, a., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.O., Sung, G.H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.a., Urbina, H., Vagvolgyi, C., Vialle, a., Vu, T.D., Walther, G., Wang, Q.M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, a., Zamora, J.C., Zhang, N., Zhuang, W.Y. & Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109, 6241–6246.

- Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73, 5261–5267.
- Yadav, R.K.P., Karamanoli, K. & Vokou, D. (2005). Bacterial Colonization of the Phyllosphere of Mediterranean Perennial Species as Influenced by Leaf Structural and Chemical Features. *Microbial Ecology*, 50, 185–196.

Chapter 5

General discussion

5.1 Recapitulation of the major findings

Forest trees are most likely to be foundation species, structuring communities and ecosystems by influencing a multitude of processes at local and global scales (Ellison *et al.*, 2005): they influence energy flux and water cycling; they influence local microclimate; and they provide food source and dwelling for below- and above-ground organisms of diverse taxonomic origin, ranging from unicellular bacteria over nematodes to arthropods, birds, and mammals. These organisms, whether they interact directly or indirectly with each other, form complex interaction networks with the tree as a hub - the crucial network element connecting all the network components (Simard, 2009). Consequently hub properties will have great effect on the network composition and structure. For instance, leaf quality structures insect herbivore community and consequently influences higher trophic levels - parasitoids and predators of herbivores (Tack & Dicke, 2013). However, the herbivores might be influenced either by direct effects of tree traits or indirect effects - either of microbes, mediated by the tree, or of the tree, mediated by the microbes (Biere & Bennett, 2013). In either case microbes are pivotal to the tree associated community and they influence almost all plant functional traits (Friesen *et al.*, 2011). Phytopathogens, like powdery-mildew, can thereby exert strong and measurable effects on the plant host (Hajji *et al.*, 2009) and its associated communities, as it was shown for the insect (Tack *et al.*, 2012), bacterial (Suda *et al.*, 2009), and fungal communities (Jakuschkin *et al.*, Chapter 2 on page 41). But how are the pathogens integrated into the interlaced microbial community of a tree? In Chapter 2 we investigated the pathobiome of *Erysiphe alphitoides*, the causal agent of oak powdery-mildew. We showed that powdery mildew infection is associated with compositional changes in fungal community (Figure 5.1 on the next page). Using network inference we detected 13 bacterial and 13 fungal OTUs, which are likely to interact directly with *E. alphitoides* within a larger microbial network dominated by positive interactions. With 816 bacterial and 1073 fungal OTUs detected on three trees we showed that oak phyllosphere harbors very rich microbial community.

As it was previously discussed, besides the environmental and stochastic variables, plant functional traits play important role in phyllosphere microbial community assembly (Section 3.1 on page 72). These plant traits vary between individual plants and are ultimately controlled by plant genes (Hunter *et al.*, 2010; Reisberg *et al.*, 2013). Thus we expected to find QTLs in oak genome, which would affect properties of phyllosphere microbial community. In Chapter 3 we presented QTLs for several descriptors of oak phyllosphere community: we found QTLs for fungal diversity, fungal and bacterial composition, microbial network properties, *E. alphitoides* reads, and powdery-mildew infection symptoms (Figure 5.1). Especially notable result was the confirmation of the QTL for *E. alphitoides* on the LG 10: the QTL for *E. alphitoides* inoculation experiment from 2006 co-localized with the QTL for *E. alphitoides* read counts from 2011. Furthermore QTLs for bacterial and fungal composition were also located in this region of LG 10. Hence we showed that intraspecific genetic variation can have cascading effects beyond population, effecting oak phyllosphere microbial community. If this genetic variability is non-neutral with respect to fitness consequences for the tree host, it provides a mechanism upon which natural selection can act, affecting the evolution of oak phyllosphere microbial community.

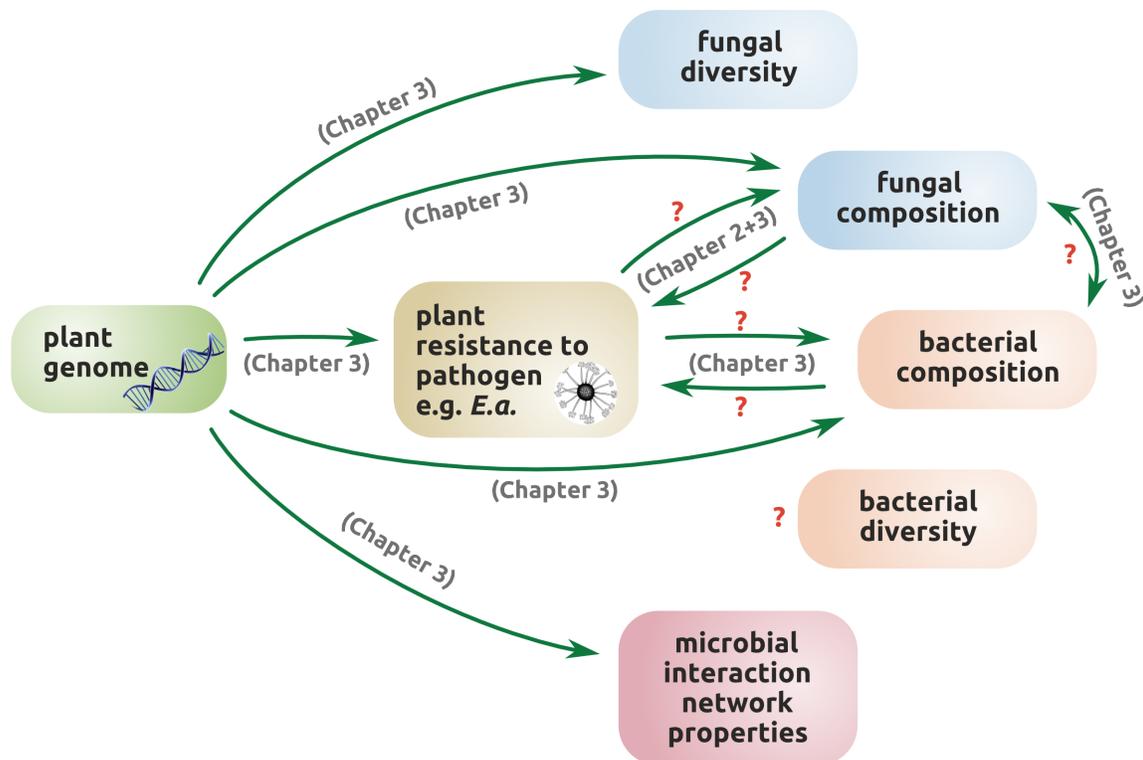


Figure 5.1: Summary of major findings of this thesis with references to the relevant chapters. Question marks indicate the detected correlations, however the causations remain to be determined.

5.2 Possible evolutionary mechanisms

The evolution of plants is tightly linked to microbial life. Eukaryotic ancestors of plantae acquired cyanobacteria, which became chloroplasts (Mereschkowsky, 1905; Bhattacharya *et al.*, 2004). Equipped with chloroplasts, early plants were ready to colonize the land and transform Earth into the green planet (McFadden, 2001). Another key role in the colonization of land by plants was played by mycorrhizal interactions, which evolved about 400 to 450 million years ago (Buckling *et al.*, 2012), and which are present in approximately 80% of land plants surveyed so far (Wang & Qiu, 2006). The tight evolutionary link can also be traced back in more recent evolutionary history of angiosperms and gymnosperms, and their endophytes. The endophyte communities in angiosperms are frequently dominated by species of Diaporthales and those in gymnosperms by species of Helotiales. Divergence of angiosperms and gymnosperms coincides with the divergence of the Diaporthales and the Helotiales about 300 million years (Ma) ago, indicating that the Diaporthalean and Helotialean ancestors of tree endophytes had been associated, respectively, with angiosperms and gymnosperms since ≥ 300 Ma (Sieber, 2007).

While the generation time for a tree, e.g. *Quercus robur*, can take as long as 30-45 years (Ducouso *et al.*, 1993), such simple unicellular plants like green algae, e.g. *Chlamydomonas reinhardtii*, can have several generations per day (Ness *et al.*, 2012). Consequently evolution of plant-microbe interactions will be different for the two cases as the generation time for most bacteria is between 20 minutes and 20 hours, and is typically less than 1 hour (Black, 2012). Therefore unicellular plants and microbes are more likely to evolve in the Red Queen-race manner. Along with multicellularity and increased body size plants have evolved longer generation times. In the years, that a typical tree needs to reach maturity, microbes will have passed many thousands of generations. Some authors argue that mutations take place at constant rates (Ochman & Wilson, 1987) and the difference in generation times between trees and microbes would, therefore, be negligible. This is not always true, as, for instance, stressful conditions lead to the emergence of hypermutator clones (Foster, 2007; Bertrand *et al.*, 2015). Another microbial “weapon” in the evolution race is lateral gene transfer, which can accelerate the spread of advantageous genes in a population. Furthermore microbes have large effective population size, which in combination with short generation times provides copious material for natural selection to act. This stands in contrast with relatively small effective population size of trees. So how come that trees did not lose the evolutionary race to microbial pathogens? “The ace up the trees’ sleeve” is their broad phenotypic plasticity, which allows them to adapt to various environmental conditions and changing environmental pressures during their life-time. It is likely that beneficial microbial partners contribute substantially to this plasticity (Desprez-Loustau *et al.*, 2015) and some of these partners may act as antagonists to pathogens (Arnold *et al.*, 2003). Therefore trees, as units of evolution, should, probably, be viewed as holobionts (Zilber-Rosenberg & Rosenberg, 2008), consisting of two components: slow evolving host genome and fast evolving

microbiome.

Plant host exert a strong selective force on the microbial colonizers as it was shown for the crop plants, whose bacterial communities differentiated from soil-like at the beginning of the growing season to leaf-specific as the season progressed (Copeland *et al.*, 2015), and for *Arabidopsis thaliana*, whose phyllosphere microbes differentiated from air-like to leaf-specific (Maignien *et al.*, 2014). Thereby various environmental variables (e.g. precipitation, UV-radiation, Section 1.2 on page 11) along with host functional traits (e.g. resource availability, leaf topography) could be acting as an ecological filter that selects for certain taxa or functional groups and influences community assembly on leaves (Kembel & Mueller, 2014). The question arises whether reciprocal case is true and the phyllosphere microbes can exert selective pressure on their tree hosts? *Erysiphe alphitoides*, for instance, can exert strong selective pressure in oak seedlings, negatively affecting their growth and even causing death in some cases (Feau *et al.*, 2012). Furthermore it was shown to influence host phenology significantly speeding up the second flush in the growing season (Desprez-Loustau *et al.*, 2014), which might influence the phyllosphere microbiota, as the available microbial inoculum pool varies throughout the year (Jumpponen & Jones, 2010; Kaneko & Kaneko, 2004). *E. alphitoides* also induced growth polycyclism changing plant architecture (Desprez-Loustau *et al.*, 2014), which in turn may affect the microclimatic conditions in the phyllosphere and influence the phyllosphere microbial community. In Chapter 2 we also have shown that *E. alphitoides* infection is associated with changes in microbial communities. It is therefore possible that some microbes, e.g. phytopathogens like *E. alphitoides*, exert (strong) selective pressure on their host and the microbial community follows this evolutionary trajectory. On the other hand it is possible that the infection is the consequence of the microbial community structure, which represents the host extended phenotype. Both options are not mutually exclusive and could act in concert to influence the evolution of plant host. Additionally, the phyllosphere microbial community might influence the evolution of the host by influencing the decomposer community (Osono, 2006; Voriskova & Baldrian, 2013), soil microbiota, and hence the nutrient cycles (Kembel & Mueller, 2014) in which the host is involved, which in turn can contribute to host fitness and evolution (Schweitzer *et al.*, 2008) (Figure 5.2 on the following page).

But which attributes of the microbial community are best suited to study its influence on host and associated microbiota evolution? Application of NGS techniques in combination with DNA metabarcoding provides great taxonomical resolution and reveals the diversity of phyllosphere microbiota (Huse *et al.*, 2008; Schoch *et al.*, 2012), bypassing the cloning bias of Sanger method and outperforming other culture based methods (Rappe & Giovannoni, 2003; Cardenas & Tiedje, 2008). Using this approach we have observed large taxonomical variability within the same host – only few OTUs were shared among all leaves, e.g. within the susceptible tree 3 bacterial and 15 fungal OTUs were shared among all 39 leaves, and 72 bacterial and 78 fungal OTUs were shared among at least 20 leaves (51.28%), while total bacterial and fungal richness for the leaves of this tree was 820 and 1071, respectively

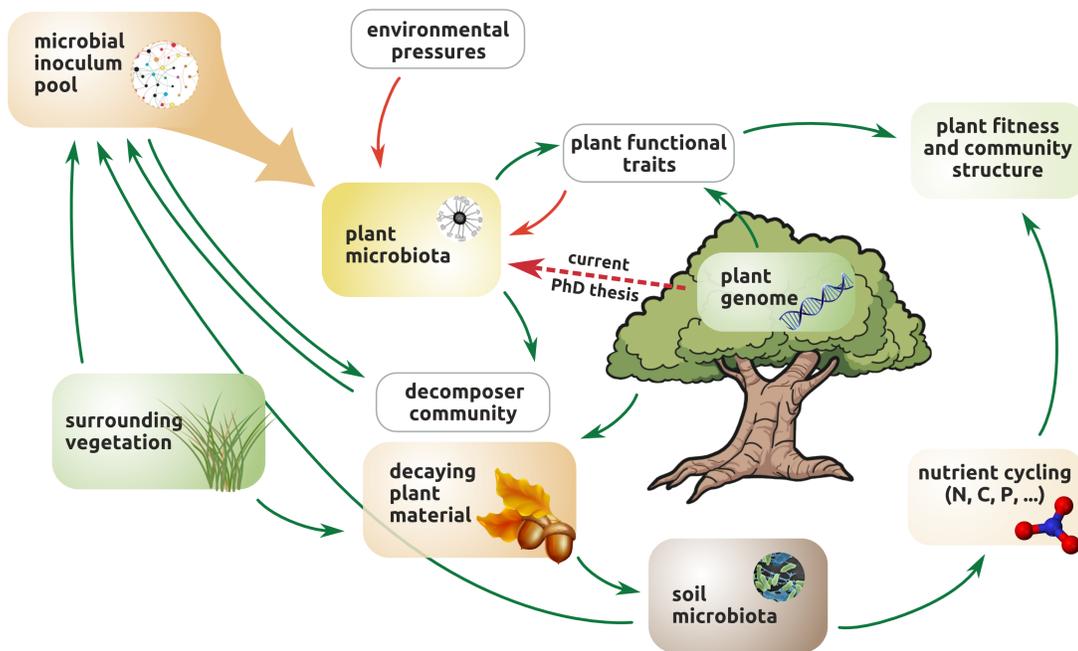


Figure 5.2: Microbiota dynamics and interactions - representation of the linkage between a plant host and its microbiota within their ecological context.

(Chapter 2). Microenvironmental differences between the leaves, e.g. UV exposure differences due to the plant architecture, along with the stochastic “lottery” probably play major role in the observed variability (Maignien *et al.*, 2014). However it is possible that the microbial assemblies on the leaves are more alike than can be revealed with the taxonomical approach. Why could it be so? Firstly, the metagenomic approach has inherent methodological weaknesses: the usage of metabarcodes and fixed similarity thresholds blurs true taxonomical composition by combining different species within same OTUs and by splitting same species into different OTUs (*Erysiphe* example, Section 2.2 on page 47; also Table 2 in Nilsson *et al.*, 2008; Blaaid *et al.*, 2013); PCR step introduces errors due to polymerase errors, primer biases, biases due to the chosen metabarcode (Soergel *et al.*, 2012; Blaaid *et al.*, 2013). Secondly, detection of a DNA barcode in the sample does not necessarily mean that the detected microorganism is metabolically active, it could be in a spore state and not contributing to the ecological interactions at this particular moment. Next, the biological species concept has its difficulties when applied to microorganisms, which do not form sexual populations Mayr (1996). Even on the strain level there can be remarkable genetic diversity as it was shown by comparison of three *E. coli* genomes, which had only about 2998 (39.2%) genes in common of total 7638 present in at least one of the three strains. The matter is additionally complicated by the possibility of lateral gene transfer, which can weaken or uncouple the relationship between ecological similarity and evolutionary relatedness (Horner-Devine & Bohannan, 2006; Doolittle & Zhaxybayeva, 2009). Hence we can expect to find microbial

species of different evolutionary origin being adapted to similar ecological niche due to convergent evolution or lateral gene transfer (Boon *et al.*, 2014). Therefore, as traits ultimately dictate the interactions among organisms and between organisms and the environment, trait-based approaches might provide more accurate community analyses (Violle *et al.*, 2007). Metagenomic, -transcriptomic, -proteomic, and -bolomic studies will become more and more attractive due to the lowering costs and technological advances and will provide tools for trait-based plant microbiota studies (Knief, 2014), which should overcome the limits of current taxonomy based approaches.

5.3 Bibliography

- Arnold, A.E., Mejía, L.C., Kylo, D., Rojas, E.I., Maynard, Z., Robbins, N. & Herre, E.A. (2003). Fungal endophytes limit pathogen damage in a tropical tree. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 15649–15654.
- Bertrand, J.C., Caumette, P., Lebaron, P., Matheron, R., Normand, P. & Sime-
Ngando, T. (2015). *Environmental Microbiology: Fundamentals and Applications*. Springer Netherlands, Dordrecht.
- Bhattacharya, D., Yoon, H.S. & Hackett, J.D. (2004). Photosynthetic eukaryotes unite: Endosymbiosis connects the dots. *BioEssays*, 26, 50–60.
- Biere, A. & Bennett, A.E. (2013). Three-way interactions between plants, microbes and insects. *Functional Ecology*, 27, 567–573.
- Blaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M. & Kauserud, H. (2013). ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular ecology resources*, 13, 218–24.
- Black, J. (2012). *Microbiology Principles and Explorations*.
- Boon, E., Meehan, C.J., Whidden, C., Wong, D.H.J., Langille, M.G.I. & Beiko, R.G. (2014). Interactions in the microbiome: communities of organisms and communities of genes. *FEMS microbiology reviews*, 38, 90–118.
- Bucking, H., Liepold, E. & Ambilwade, P. (2012). The Role of the Mycorrhizal Symbiosis in Nutrient Uptake of Plants and the Regulatory Mechanisms Underlying These Transport Processes. In: *Plant Science*. InTech.
- Cardenas, E. & Tiedje, J.M. (2008). New tools for discovering and characterizing microbial diversity. *Current Opinion in Biotechnology*, 19, 544–549.
- Copeland, J.K., Yuan, L., Layeghifard, M., Wang, P.W. & Guttman, D.S. (2015). Seasonal Community Succession of the Phyllosphere Microbiome. *Molecular Plant-Microbe Interactions*, 28, 274–285.
- Desprez-Loustau, M.L., Aguayo, J., Dutech, C., Hayden, K.J., Husson, C., Jakushkin, B., Marçais, B., Piou, D., Robin, C. & Vacher, C. (2015). An evolutionary ecology perspective to address forest pathology challenges of today and tomorrow. *Annals of Forest Science*.
- Desprez-Loustau, M.L., Saint-Jean, G., Barrès, B., Dantec, C.F. & Dutech, C. (2014). Oak powdery mildew changes growth patterns in its host tree: host tolerance response and potential manipulation of host physiology by the parasite. *Annals of Forest Science*, 71, 563–573.

- Doolittle, W.F. & Zhaxybayeva, O. (2009). On the origin of prokaryotic species. *Genome research*, 19, 744–56.
- Ducouso, A., Michaud, H. & Lumaret, R. (1993). Reproduction and gene flow in the genus *Quercus* L. *Annales Des Sciences Forestieres*, 50, 91–106.
- Ellison, A.M., Bank, M.S., Clinton, B.D., Colburn, E.a., Elliott, K., Ford, C.R., Foster, D.R., Kloeppel, B.D., Knoepp, J.D., Lovett, G.M., Mohan, J., Orwig, D.a., Rodenhouse, N.L., Sobczak, W.V., Stinson, K.a., Stone, J.K., Swan, C.M., Thompson, J., Von Holle, B., Webster, J.R., Holle, B.V. & Webster, J.R. (2005). Loss of foundation species : consequences for the structure and dynamics of forested ecosystems. *Frontiers in Ecology and the Environment*, 3, 479–486.
- Feau, N., Lauron-Moreau, A., Piou, D., Marçais, B., Dutech, C. & Desprez-Loustau, M.L. (2012). Niche partitioning of the genetic lineages of the oak powdery mildew complex. *Fungal Ecology*, 5, 154–162.
- Foster, P.L. (2007). Stress-Induced Mutagenesis in Bacteria. *Critical Reviews in Biochemistry and Molecular Biology*, 42, 373–397.
- Friesen, M.L., Porter, S.S., Stark, S.C., von Wettberg, E.J., Sachs, J.L. & Martinez-Romero, E. (2011). Microbially Mediated Plant Functional Traits. *Annual Review of Ecology, Evolution, and Systematics*, 42, 23–46.
- Hajji, M., Dreyer, E. & Marçais, B. (2009). Impact of *Erysiphe alphitoides* on transpiration and photosynthesis in *Quercus robur* leaves. *European Journal of Plant Pathology*, 125, 63–72.
- Horner-Devine, M.C. & Bohannan, B.J.M. (2006). Phylogenetic clustering and overdispersion in bacterial communities. *Ecology*, 87, S100–8.
- Hunter, P.J., Hand, P., Pink, D., Whipps, J.M. & Bending, G.D. (2010). Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca* species) phyllosphere. *Applied and Environmental Microbiology*, 76, 8117–8125.
- Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A. & Sogin, M.L. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, 4, e1000255.
- Jumpponen, A. & Jones, K.L. (2010). Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. *New Phytologist*, 186, 496–513.
- Kaneko, R. & Kaneko, S. (2004). The effect of bagging branches on levels of endophytic fungal infection in Japanese beech leaves. *Forest Pathology*, 34, 65–78.

- Kembel, S.W. & Mueller, R.C. (2014). Plant traits and taxonomy drive host associations in tropical phyllosphere fungal communities. *Botany*, 92, 303–311.
- Knief, C. (2014). Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Frontiers in plant science*, 5, 216.
- Maignien, L., DeForce, E.A., Chafee, M.E., Eren, A.M. & Simmons, S.L. (2014). Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *mBio*, 5, 10.
- Mayr, E. (1996). What is a species, and what is not? *Philosophy of Science*, 63, 262–277.
- McFadden, G.I. (2001). Chloroplast origin and integration. *Plant physiology*, 125, 50–53.
- Mereschkowsky, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biologisches Centralblatt*, 25, 593–604.
- Ness, R.W., Morgan, A.D., Colegrave, N. & Keightley, P.D. (2012). Estimate of the spontaneous mutation rate in *Chlamydomonas reinhardtii*. *Genetics*, 192, 1447–1454.
- Nilsson, R.H., Kristiansson, E., Ryberg, M. & Hallenberg, N. (2008). Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evolutionary Bioinformatics*, 4, 193–201.
- Ochman, H. & Wilson, A.C. (1987). Evolution in bacteria: Evidence for a universal substitution rate in cellular genomes. *Journal of Molecular Evolution*, 26, 74–86.
- Osono, T. (2006). Role of phyllosphere fungi of forest trees in the development of decomposer fungal communities and decomposition processes of leaf litter. *Canadian Journal of Microbiology*, 52, 701–716.
- Rappe, M.S. & Giovannoni, S.J. (2003). The Uncultured Microbial Majority. *Annual Review of Microbiology*, 57, 369–394.
- Reisberg, E.E., Hildebrandt, U., Riederer, M. & Hentschel, U. (2013). Distinct Phyllosphere Bacterial Communities on *Arabidopsis* Wax Mutant Leaves. *PLoS ONE*, 8, e78613.
- Schoch, C.L., Seifert, K.a., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.a., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, a.N., Wingfield, M.J., Aime, M.C., An, K.D., Bai, F.Y., Barreto, R.W., Begerow, D., Bergeron, M.J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, a.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, a., Cubas, P.,

- Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.a., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.a., Ge, Z.W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.B., Hosaka, K., Houbraeken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., Kovacs, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, a.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, a.R., Methven, a.S., Meyer, W., Moncalvo, J.M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.a., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, a., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.O., Sung, G.H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.a., Urbina, H., Vagvolgyi, C., Vialle, a., Vu, T.D., Walther, G., Wang, Q.M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, a., Zamora, J.C., Zhang, N., Zhuang, W.Y. & Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109, 6241–6246.
- Schweitzer, J.A., Bailey, J.K., Fischer, D.G., LeRoy, C.J., Lonsdorf, E.V., Whitham, T.G. & Hart, S.C. (2008). Plant-soil-microorganism interactions: heritable relationship between plant genotype and associated soil microorganisms. *Ecology*, 89, 773–781.
- Sieber, T.N. (2007). Endophytic fungi in forest trees: are they mutualists? *Fungal Biology Reviews*, 21, 75–89.
- Simard, S.W. (2009). The foundational role of mycorrhizal networks in self-organization of interior Douglas-fir forests. *Forest Ecology and Management*, 258, S95–S107.
- Soergel, D.a.W., Dey, N., Knight, R. & Brenner, S.E. (2012). Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *The ISME journal*, 6, 1440–4.
- Suda, W., Nagasaki, A. & Shishido, M. (2009). Powdery Mildew-Infection Changes Bacterial Community Composition in the Phyllosphere. *Microbes and Environments*, 24, 217–223.

- Tack, A.J.M. & Dicke, M. (2013). Plant pathogens structure arthropod communities across multiple spatial and temporal scales. *Functional Ecology*, 27, 633–645.
- Tack, A.J.M., Gripenberg, S. & Roslin, T. (2012). Cross-kingdom interactions matter: fungal-mediated interactions structure an insect community on oak. *Ecology Letters*, 15, 177–185.
- Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, 116, 882–892.
- Voriskova, J. & Baldrian, P. (2013). Fungal community on decomposing leaf litter undergoes rapid successional changes. *The ISME journal*, 7, 477–86.
- Wang, B. & Qiu, Y.L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*, 16, 299–363.
- Zilber-Rosenberg, I. & Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS microbiology reviews*, 32, 723–35.

Chapter 6

Appendix

- 6.1 Appendix A - An evolutionary ecology perspective to address forest pathology challenges of today and tomorrow

An evolutionary ecology perspective to address forest pathology challenges of today and tomorrow

Marie-Laure Desprez-Loustau^{1,2} · Jaime Aguayo³ · Cyril Dutech^{1,2} · Katherine J. Hayden^{4,5} · Claude Husson^{4,5} · Boris Jakushkin^{1,2} · Benoît Marçais^{4,5} · Dominique Piou^{1,6} · Cécile Robin^{1,2} · Corinne Vacher^{1,2}

Received: 26 December 2014 / Accepted: 21 April 2015
© The Authors 2015. This article is published with open access at Springerlink.com

Abstract

• **Key message** Increasing human impacts on forests, including unintentional movement of pathogens, climate change, and large-scale intensive plantations, are associated with an unprecedented rate of new diseases. An evolutionary ecology perspective can help address these challenges and provide direction for sustainable forest management.

Handling Editor: Jean-Michel Leban

Contribution of the co-authors J. Aguayo, C. Husson, and B. Marçais wrote the first draft of the part dealing with fungal diversity, C. Dutech of the part dealing with pathogen evolution, M.-L. Desprez-Loustau and K. Hayden of the part on disease resistance, C. Robin of the part on hyperparasitism, C. Vacher and B. Jakushkin of the part on the tree microbiote, M. L. Desprez-Loustau of the introduction and conclusion. D. Piou, M. L. Desprez-Loustau and B. Marçais conceived Fig. 1. All co-authors contributed to the final writing and revision of the manuscript, with a special contribution of K. Hayden. Marie-Laure Desprez-Loustau initiated and coordinated the review.

• **Context** Forest pathology has historically relied on an ecological approach to understand and address the practical management of forest diseases. A widening of this perspective to include evolutionary considerations has been increasingly developed in response to the rising rates of genetic change in both pathogen populations and tree populations due to human activities.

• **Aims** Here, five topics for which the evolutionary perspective is especially relevant are highlighted.

• **Results** The first relates to the evolutionary diversity of fungi and fungal-like organisms, with issues linked to the identification of species and their ecological niches. The second theme deals with the evolutionary processes that allow forest pathogens to adapt to new hosts after introductions or to become more virulent in homogeneous plantations. The third theme presents issues linked to disease resistance in tree breeding programs (e.g., growth-defense trade-offs) and proposes new criteria and methods for more durable resistance.

✉ Marie-Laure Desprez-Loustau
loustau@bordeaux.inra.fr

Jaime Aguayo
jaime.aguayo@anses.fr

Cyril Dutech
cdutech@bordeaux.inra.fr

Katherine J. Hayden
katherine.hayden@nancy.inra.fr

Claude Husson
chusson@nancy.inra.fr

Boris Jakushkin
boris.jakuschkin@gmail.com

Benoît Marçais
marçais@nancy.inra.fr

Dominique Piou
piou@pierroton.inra.fr

Cécile Robin
robin@bordeaux.inra.fr

Corinne Vacher
vacher@bordeaux.inra.fr

¹ INRA, UMR1202 BIOGECO, 33610 Cestas, France

² Univ. Bordeaux, BIOGECO, UMR 1202, 33600 Pessac, France

³ ANSES, Laboratoire de la Santé des Végétaux LSV, Unité de Mycologie, Domaine de Pixérécourt, 54220 Malzéville, France

⁴ INRA, UMR1136 Interactions Arbres/Microorganismes, 54280 Champenoux, France

⁵ Université De Lorraine, UMR1136 Interactions Arbres/Microorganismes, 54506 Vandœuvre-lès-Nancy, France

⁶ Département de la Santé des Forêts, Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt, 69 route d'Arcachon, 33612 Cestas Cedex, France

The last two themes are dedicated to the biotic environment of the tree–pathogen system, namely, hyperparasites and tree microbiota, as possible solutions for health management.

• **Conclusion** We conclude by highlighting three major conceptual advances brought by evolutionary biology, i.e., that (i) “not everything is everywhere”, (ii) evolution of pathogen populations can occur on short time scales, and (iii) the tree is a multitrophic community. We further translate these into a framework for immediate policy recommendations and future directions for research.

Keywords Emerging disease · Invasive pathogen · Microbiota · Mycoparasite · Fungal diversity · Coevolution · Tree breeding

1 Introduction: a brief history of forest pathology and its primary concepts

Forest pathology deals with the diseases of forest trees, which are mainly caused by fungal and oomycete pathogens, in both their fundamental and applied aspects. As such, the history of forest pathology has been shaped not only by the progress of ideas and concepts in the more general disciplines of microbiology and health sciences but also by changes in the practice of forestry. Forest pathology has obvious common roots with plant pathology, but is not simply one of its specialities. Long time scales, diversity, and heterogeneity are hallmarks of forests, in both their biological and socio-economic dimensions. Forest trees are still mostly nondomesticated species, highly diverse, growing in richly biotic and heterogeneous environments. Until recently, forest products have been harvested for the greatest part from natural or seminatural forests, with rotation periods generally exceeding 50 years (FAO 2010). As a consequence, forest pathology, as a speciality of forestry, has had a tradition of systemic and long-term thinking and, probably more so and earlier than plant pathology, an ecological approach. Integrating an evolutionary perspective is a natural extension, especially in the context of global changes.

Robert Hartig, the author of two forest pathology textbooks in 1874 and 1882, is generally recognized as the father of forest pathology, following in Anton de Bary’s footsteps, who is considered to be the father of plant pathology (Kutschera and Hossfeld 2012). In support to the germ theory of disease, de Bary provided experimental evidence of the role of *Phytophthora infestans* in potato late blight epidemics, refuting the old theory of spontaneous generation, while Hartig proposed a new concept of tree decay whereby fungi were the cause and not the consequence of the degradation (Merrill and Shigo 1979). The identification and biological characterization of the causal agents of the diseases affecting trees were still the primary objectives of forest pathologists at the end of the nineteenth and the beginning of the twentieth centuries.

From the very beginning, forest pathologists were concerned with both productivity (fungi affecting wood production and preservation) and conservation issues (pathogens threatening natural resources) (Peterson and Griffith 1999). As the potato late blight epidemics did for plant pathology, severe outbreaks caused by the introduction of exotic pathogens have marked the history of forest pathology, such as chestnut blight and white pine blister rust in North America and Europe (Anagnostakis 1987; Delatour et al. 1985). Forest health has also been recurrently marked by regionally important declines affecting a variety of species, e.g., oaks in Europe or maple in North America (Delatour 1983; Tainter and Baker 1996). These phenomena, leading to progressive debilitation of trees and ultimately mortality, highlight the potential long-term effects of weakening factors, especially soil and climate, interacting with the genetic background of trees. As early as 1963, George (Hepting 1963) wrote that “field crops [mostly annual plants]...reflect weather changes; whereas trees [...] will reflect, in addition, climate changes.” The “tree decline concept,” with biotic and abiotic factors acting and interacting in sequence as predisposing, inciting, or contributing factors was later formalized by Manion (1981). More generally, forest pathologists have long been interested in and have studied the relationships between environment, e.g., local site factors, and disease. In contrast to more intensively managed agricultural systems, forest environments are heterogeneous, and their natural constraints (fertility, water availability) are not usually overcome by external, human inputs. Moreover, direct control of diseases through fungicide applications has usually not been considered to be necessary or appropriate, for reasons such as cost, practical feasibility, and environmental concerns. Disease management in forests has thus traditionally been viewed as necessarily based on “good” silvicultural practices and preventive actions deriving from the knowledge of risk factors. Hazard-rating systems, which identify sites where pathogens are likely to have a high impact and which should thus be avoided for the planting of susceptible species, are examples of successful applications of forest epidemiology (Tainter and Baker 1996).

The issues faced by forest pathologists today would not be so different than those addressed by their predecessors, but for the dramatic acceleration and intensity of changes. The apparition of new diseases caused by alien invasive pathogens is rapidly increasing with the intensification of international trade and travels (Brasier 2008; Desprez-Loustau et al. 2010; Pyšek et al. 2010; Santini et al. 2013; Roy et al. 2014), presenting a major challenge to forest pathology today. Meanwhile, climate change is expected to affect the probability of establishment of introduced pathogens and more generally to affect the distribution and severity of diseases (e.g., Shaw and Osborne 2011; Sturrock et al. 2011). Indeed, forest pathologists pioneered the use of risk mapping for plant disease under climate change by combining epidemiological models and

climate change scenarios (Brasier 1996; Booth et al. 2000; Bergot et al. 2004). The rapid intensification of forestry is likewise influencing forest pathology. The increasing demand for wood, biomass, fuel, and other forest-derived products has led to a large expansion of planted forests, by around 5 million hectares per year on average in the last decade, with the aim to increase productivity (FAO 2010). In some countries, especially in the southern hemisphere and Southeast Asia, where forest industries have implemented large-scale, management-intensive monocultures of *Eucalyptus*, *Acacia*, and *Pinus*, plantations currently account for most of the wood supply (Paquette and Messier 2009). However, new disease emergences threaten productivity in these new forest ecosystems, as has been reported for stem cankers and leaf and shoot blight of *Eucalyptus* or root rots of *Acacia* spp. (Su-See 1999; Mohan and Manokaran 2013; Liebhold and Wingfield 2014). The strong impact of pathogen introductions and forest management practices on forest diseases can be illustrated by trends observed in the database of the French Forest Health service (Fig. 1).

Long before disease regulation was formally recognized as an ecosystem service (Millennium Ecosystem Assessment 2005), forest pathology took an empirical approach to the study of the ecological processes that regulate forest health. The overall low level of disease in natural and seminatural forests, with a few exceptions of devastating native diseases (Hansen and Goheen 2000), is consistent with the general observation that native plant populations rarely suffer from devastating epidemics, contrary to what is observed in crops without fungicide applications (Jousimo et al. 2014). A major exception to the effectiveness of natural disease regulation, at least in the short term, is constituted by diseases caused by some exotic pathogens. The awareness of the increasing impacts of human activities on ecosystems, including forests, has emphasized the need for global approaches and has changed our views on evolution of species. It has been increasingly realized that evolutionary changes can be observed at short time-scales, especially in ecosystems under strong anthropogenic influence (Palumbi 2001; Burdon and Thrall 2008; Jousimo et al. 2014). Thus, the traditional view considering evolutionary biology and ecology as two independent fields has changed dramatically (Koch et al. 2014). The speed of adaptation of organisms to keep pace with the human alterations of the environment has become a matter of concern (Carroll et al. 2014). In particular, the ever-faster pace of disease emergences and the severity of the damage they cause has emphasized the need to better understand and take into consideration the coevolutionary processes between pathogens and their host tree populations (Parker and Gilbert 2004; Schoettle and Snieszko 2007; Hendry et al. 2011; Ennos 2015). At the same time, the revolution brought by DNA technologies and their reducing costs has opened new prospects for population genetic and environmental metagenomic studies (Neale and Ingvarsson 2008; Di Bella et al. 2013),

which can shed new light on the ecology and evolution of forest diseases and open new avenues for disease control.

In the current context of human-induced accelerated changes, applying evolutionary principles to disease and pest management has been advocated by several authors as the way forward to slow the pathogen arms race in human, animal, or plant diseases, (Combes 2001; Palumbi 2001; Varki 2012; Vander Wal et al. 2014). Anton de Bary himself referred to Darwin's principle of natural selection when defining plant–parasite interactions (Kutschera and Hossfeld 2012). Indeed, evolutionary biology provides the concepts and methods to understand the diversity and adaptive characteristics of organisms, thus allowing the identification of pathogens through phylogenetics, and the prediction and even manipulation of selection and adaptation processes involved in immune or pathogenic functions (Carroll et al. 2014). The idea that “nothing in forest pathology makes sense except in the light of evolution,” paraphrasing Dobzhansky, could now seem obvious and not deserving of further development (Varki 2012). Many recent articles dealing with forest pathology indeed have an evolutionary background (e.g., Hansen and Goheen 2000; Pinon and Frey 2005; Kinloch et al. 2008; Garbelotto et al. 2010; Hayden et al. 2011; Cruikshank and Jaquish 2014; Franceschini et al. 2014). However, we contend that an evolutionary perspective is still not sufficiently emphasized and applied to the management of forest diseases. For example, evolutionary considerations, especially relating to interactions with pathogens, are still ignored in a recently proposed “novel tree breeding strategy” advocating domestication of forest trees in a changing world (Harfouche et al. 2012). In line with other recent initiatives in forest pathology (e.g., Cavers 2015; Cavers and Cottrell 2015) and taking benefit of insights from other fields (plant pathology, invasion ecology, ecology of parasitism, etc.), this article aims to show how an evolutionary ecology perspective can improve and even change our understanding of forest diseases and help address the challenges of today and tomorrow. Rather than a comprehensive review of recent developments in forest pathology, we highlight and focus on some key topics, for which the evolutionary perspective is especially relevant and could provide new directions for forest research or disease management: (i) fungal evolutionary diversity (species diversity of forest pathogens and their ecological niches); (ii) pathogen evolution (how forest pathogens become adapted to their hosts); (iii) forest resistance to disease, especially in relation to tree breeding (trade-offs, tolerance, emerging properties in populations); and (iv) the role of hyperparasites and tree microbiota in the regulation of pathogen populations and disease. Finally, we show how knowledge acquired in an evolutionary framework can translate into recommendations for forest

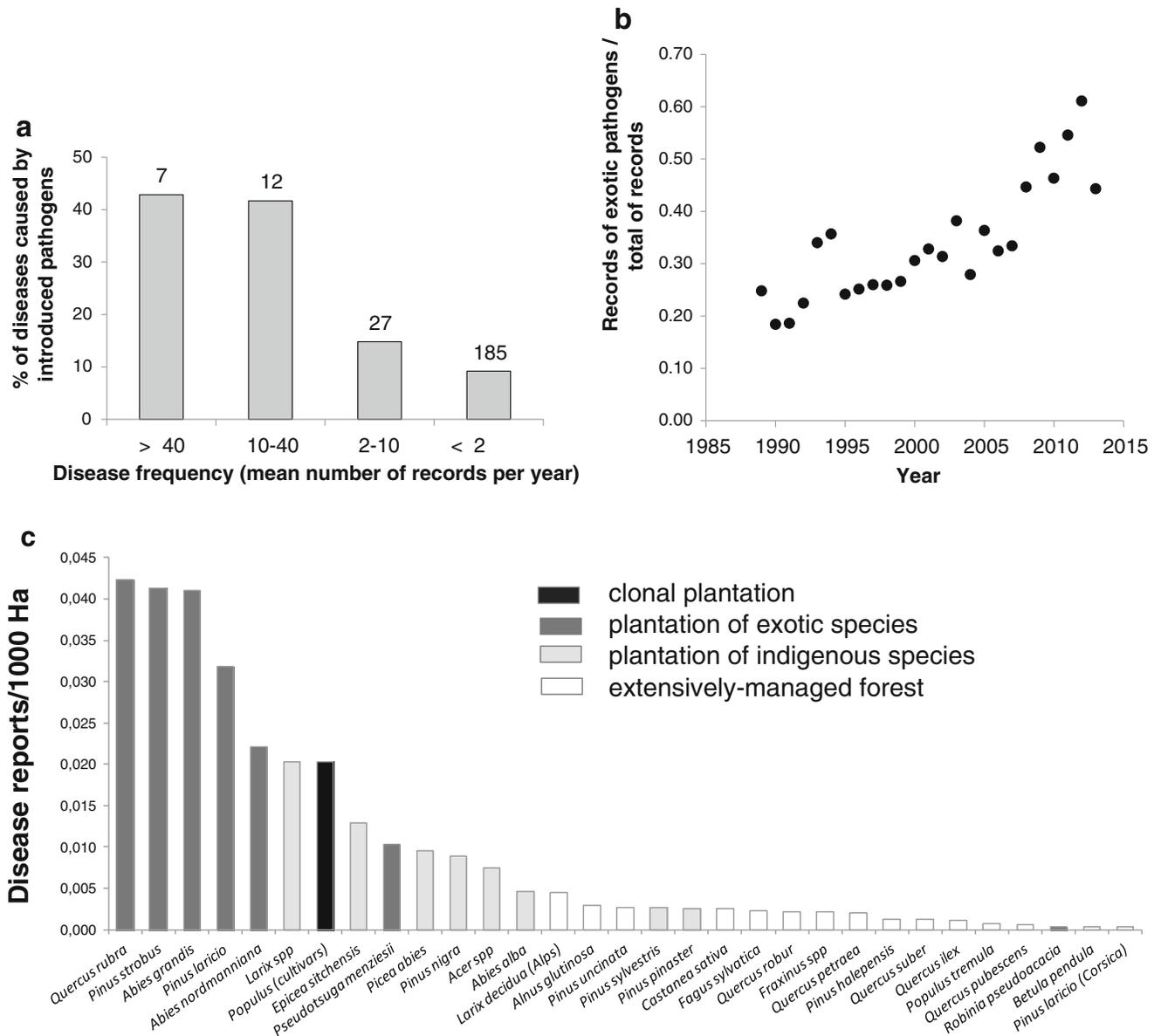


Fig. 1 The database of the French Health Service includes approximately 60,000 reports, from 1989 to 2014, which provides a general view of all pathogens (and other causes) affecting forest health. A total of 24,375 records have been identified at the pathogen species level, for a total of 231 species, among which 29 introduced species for 8634 records and 202 nonintroduced species for 15741 records. **a** Introduced pathogens are overrepresented among the most frequent forest diseases; Fisher's exact test ; $**p=0.0014$. The value above each column indicates the number of

pathogen species. The 19 more frequent pathogens represent 81% of all disease records. **b** The part of reports associated with introduced pathogens has much increased in the recent period especially since the detection of *Hymenoscyphus fraxineus* in 2008 in eastern France. **c** A much higher rate of reports is observed in forest plantations compared to less intensively managed forests (1989–2006 data). This probably includes both higher pathogen load but also higher attention paid to growth losses in these systems (lower tolerance to damage by the forester)

pathology in the future, both from an academic and an applied perspective. Although we focus here on forest pathology as the art and science of forest diseases and their control, it should be reminded that forest pathologists more generally contribute to the ecology of parasitism. In this ecosystem perspective, fungal pathogens are no longer “enemies” but key actors of the evolution and ecology of local communities, and more generally of the ecosystem health (e.g., Burdon 1987; Gilbert 2002; Hudson et al. 2006; Ostry and Laflamme 2009).

2 Towards a better knowledge of the diversity, biogeography, and ecology of forest pathogenic fungi and *Phytophthora* spp.

The emergence of new diseases has underlined the poor current knowledge of microbial evolutionary diversity. Here, we refer not only to the diversity of organisms as taxonomic units but as evolutionary lineages that have diverged and thus present specific distributions and ecological niches (Taylor et al.

2006). In many cases, new diseases of forest trees and wild animals have been caused by introduced organisms, including fungi and oomycetes, which were not even formally described before their invasive behavior (Kupferschmidt 2012). Well-documented examples are forest epidemics caused by *Cryphonectria parasitica* on American chestnut, the Dutch elm disease caused by *Ophiostoma ulmi* and *Ophiostoma novo-ulmi*, the beech bark disease caused by *Neonectria faginata*, the Port-Orford cedar rot by *P. lateralis* and sudden oak death by *Phytophthora ramorum* (Garbelotto 2008; Loo 2009). Indeed, the kingdom Fungi alone has been estimated to include between one to five million species among which only approximately 5% are identified (Blackwell 2011). Similarly, diversity in taxonomic groups such as *Phytophthora* spp. is probably still largely undescribed (Brasier 2009; Hansen et al. 2011; Hayden et al. 2013; Hüberli et al. 2013; Oh et al. 2013). Understudied areas of Asia, Central America, or Africa are thought to host a huge diversity of undescribed fungi and fungal-like organisms that probably go unnoticed on their native hosts and could potentially be destructive pathogens if introduced outside their native area. (Hyde 2001; Brasier 2009; Schmit and Mueller 2007; Blackwell 2011). The lack of baseline data on fungal diversity and biogeography makes it difficult to say whether an emerging disease is caused by an alien species or a pre-existing pathogen, which has evolved increased virulence or been favored by changing environmental conditions (Jeger and Pautasso 2008; Shaw et al. 2008; Linzer et al. 2009; Shaw and Osborne 2011). Identifying the origin of fungal introductions solely from historical and geographical arguments is rarely feasible because temporal or spatial data retracing the chronology of invasion are often missing, and species descriptions are inaccurate, making it difficult to separate different putative population sources (see for example Kim et al. 2010; Goss et al. 2014). Once the causal organism has been isolated, only a rigorous phylogenetic study associated with epidemiological studies can decipher the causes of emergence among introduction of an exotic pathogen by human-mediated long-distance transport, changes in climatic conditions, and human disturbance to forests ecosystems, in particular, changes in host density, which often interact (Anderson et al. 2004; Desprez-Loustau et al. 2007; Stukenbrock and McDonald 2008; Stenlid et al. 2011; Santini et al. 2013).

The full recognition of fungal and fungal-like diversity has long been hampered by a lack of diagnostic criteria, but the availability of improved molecular techniques now allows the recognition of fungal species consistent with the Evolutionary Species Concept (Taylor et al. 2000). For example, detailed phylogenetical studies have demonstrated that the fungal lineage causing ash dieback was different from the indigenous European species *Hymenoscyphus albidus*, which is not pathogenic, and that it was conspecific to the Asian species formerly reported as *Lambertella albida* (Queloz et al. 2011;

Husson et al. 2011; Zhao et al. 2012). In many cases, species defined from morphological traits were shown to be pools of genetically isolated lineages that differed significantly in biologically relevant features such as geographic distribution and pathogenicity (Taylor et al. 2000, 2006). Many plant diseases have thus been shown to be caused by a complex of cryptic species (de Wet et al. 2003; Fitt et al. 2006; Crous et al. 2004; Cai et al. 2011; Pérez et al. 2012). In particular, recent studies using DNA sequences have clarified the taxonomic confusion that existed for some important forest pathogens, leading to revisions such as the split of the red band needle blight of *Pinus* into two distinct species: *Dothistroma septosporum* and *Dothistroma pini* (Barnes et al. 2004), the redefinition of the genus *Ceratocystis* (de Beer et al. 2014), and the description and characterization of the *Gibberella fujikuroi* complex (Kvas et al. 2009). Another example is oak powdery mildew in Europe, which was shown to be associated with four different cryptic species often coexisting in the same stands: *Erysiphe alphitoides*, *Erysiphe quercicola*, *Erysiphe hypophylla*, and *Phyllactinia roboris* (Mougou et al. 2008; Mougou-Hamdane et al. 2010). The geographic origin of these species is still hypothetical (Desprez-Loustau et al. 2011), as is the possibility that some of them evolved in sympatry in their region of origin, as a result of evolutionary branching. Such evolutionary branching could explain their temporal niche differentiation, in which certain species predominate at different points in the season (Feau et al. 2012; Hamelin et al. 2011).

The resolution of species complexes and the reappraisal of fungal phylogeny has not only produced taxonomic clarification but also changed our views on the biology of the fungi and their ecological niche. For example, Fabre et al. (2011) showed that the two different *Dothistroma* species causing similar foliar symptoms in pines, *D. septosporum* and *D. pini*, have a different climatic niche, the latter being present in warmer climate. The use of species-specific tools proved that the disease outbreak in the 1990s in France could not be explained by a recent introduction of *D. pini*, as it was already present in the country in old herbarium samples, but rather by increased temperature conditions in recent years. The long-held belief that powdery mildews, like many biotrophic pathogens, are highly specialized was put into question when the same species were found on distantly related hosts, as were *E. alphitoides* and *E. quercicola* on various tropical hosts in plant families very distant to Fagaceae (Takamatsu et al. 2007). Similarly, *Fusarium circinatum*, part of the large *G. fujikuroi* complex (Kvas et al. 2009), which had been thought to be specialized on the *Pinus* genus, was recently shown to be able to behave as an endophyte of different grasses (Swett and Gordon 2012) and to infect maize (Swett and Gordon 2015). This finding raises the question whether grass species can act as reservoir for *F. circinatum*.

Pathogens that cause minor damage are often unseen—especially in the region of origin of diseases, where pathogens and hosts have coevolved (Parker and Gilbert 2004; Loo

2009; Jousimo et al. 2014). Tracking the origin of invasive fungi and *Phytophthora* spp. has thus improved our knowledge of these organisms in their natural environments. A few well-documented epidemics caused by introduced pathogens have allowed the reconstruction of the invasion pathways or the exclusion of potential sources by testing different scenarios of introduction, using recent advances in molecular and statistical Bayesian analysis (Grünwald and Goss 2011; Barrès et al. 2012; Dutech et al. 2012; Graça et al. 2013; Tsui et al. 2014). For instance, molecular studies completed since the European invasion of *Hymenoscyphus fraxineus* strongly suggested that the fungus originates from East Asia, (Gross et al. 2014), where no epidemics have been reported on the local ash species. The number of described *Phytophthora* species has dramatically increased in the last years, and it is now clear that forest soils host numerous and diverse resident communities of *Phytophthora* species (Hansen 2008). The bulk of the increase is attributable to better survey and discrimination of *Phytophthora* spp. in temperate forests and wild land ecosystems (e.g., Burgess et al. 2009; Oh et al. 2013). Whereas these surveys originally focused on highly pathogenic and mostly introduced *Phytophthora* species, they allowed the discovery of several species unassociated with disease symptoms (Kroon et al. 2011; Hansen et al. 2011). One example is *Phytophthora uniformis* which does not cause decline of its host, *Alnus incana*, in its native environment in Alaska (Adams et al. 2009), but is one of the parental species of *P. ×alni*, which has decimated riparian populations of black alder in Europe since the early 1990s (Brasier et al. 1995, see below for the hybridization process).

3 Pathogen evolution: understanding and predicting the success of invasive forest pathogens and the adaptation of pathogens in intensively managed plantations

Invasions by forest pathogens most often lead to new species interactions, without any previous coevolutionary history (Parker and Gilbert 2004). This differs from what is observed in widely planted crop plants, which have a worldwide distribution and are mainly affected by pathogens with which the plant species coevolved in the center of origin. New interactions in these cases occur at the intraspecific level between cultivars and strains of the pathogens, after a period of interruption in the natural coevolutionary process, as in the famous example of potato late blight in Europe (Desprez-Loustau et al. 2007). How a species introduced into a new environment can thrive and develop biotic interactions in the invaded community is a central question in invasion ecology (Elton 1958; Parker and Gilbert 2004; Facon et al. 2006; Keller and Taylor 2008; Vacher et al. 2010). In particular, the outcomes of evolutionary novelty in plant–pathogen interactions, i.e., new

encounters between pathogens and naive plants, have led to contradictory predictions. Some have argued for biotic resistance: that is, the novelty of a new environment will be too difficult to overcome, thus only a fraction of introduced pathogens will cause disease (Newcombe and Dugan 2010). Indeed, it has been shown for plants that most biological invasions fail (Williamson and Fitter 1996). In contrast, evolutionary novelty has been suggested to generally benefit plant enemies, since there will have been no selective pressure on resident plants to develop or maintain resistance to the new antagonist (Parker et al. 2006). Verhoeven et al. (2009) proposed that the outcomes of novel interactions between plants and pathogens or insects can be predicted by the mechanisms underlying the interaction. In particular, novelty is expected to benefit pathogens in the case of recognition-based defense, as occurs in the elicitor–receptor system (Jones and Dangl 2006), since, in this case, the plants have not evolved receptors able to recognize the specific elicitors of the new pathogen and ultimately to trigger resistance responses.

The success of pathogen invasions relies on various processes, from preadaptation or plasticity to postinvasion adaptation (Agosta and Klemens 2008). Deciphering the relative contributions of these processes requires genetic and phenotypic comparisons between source and introduced populations (Keller and Taylor 2008). These analyses have been only rarely performed for forest pathogens, notably in a study demonstrating key differences in transmission between native and introduced species of *Heterobasidion* (Garbelotto et al. 2010).

In many cases, successful invasive forest pathogens have jumped to new hosts with close phylogenetic relationship, often in the same genus, e.g. *C. parasitica* from Asian to American or European *Castanea* spp. (Anagnostakis 1987), *Cronartium ribicola* from Asian to American or European *Pinus* spp. (Kim et al. 2010), *O. novo-ulmi* from Asian to European and American elms (Brasier and Mehrotra 1995), *Phytophthora lateralis* from Asian to American *Chamaecyparis* (Brasier et al. 2010), and *H. fraxineus* from Asian to European ash (Zhao et al. 2012; Goss et al. 2014). The close relationship between original and new hosts fits with the experimental demonstration that the likelihood that a pathogen can infect two plant species decreases continuously with phylogenetic distance (Gilbert and Webb 2007; De Vienne et al. 2009). This relationship between ease of host jumps and phylogenetic distance could be explained by “ecological fitting” or exaptation, i.e., that infectivity traits evolved in a host interaction in its place of origin would allow the pathogen to be able to infect a closely related host species in its area of introduction (Agosta and Klemens 2008; Agosta et al. 2010). A second, not mutually exclusive hypothesis is that fewer evolutionary changes in the pathogen are required to adapt to a new host more closely related to its original host (Altizer et al. 2003; Schulze-Lefert and Panstruga 2011). Only a few mutations events in pathogen effectors and their targeted

host proteins were shown to explain successful host jumps among sister species in *Phytophthora* (Dong et al. 2014). The phylogenetic signal in plant pathogen–host range could help identify potential risks for forest ecosystems and cautions against the introduction of seedlings or trees of exotic species to an area with wild relatives among the indigenous trees (Vacher et al. 2010; Gilbert et al. 2012). However, the generality of these predictions is questioned by the observation of several invasions with apparent host jumps to more distantly related hosts, for example, in the case of *P. ramorum* (Hansen et al. 2005; Brasier and Webber 2010), oak powdery mildew (Takamatsu et al. 2007; Desprez-Loustau et al. 2011), or *Neofusicoccum parvum* (Sakalidis et al. 2013). In some instances, ecological proximity plays a larger role than phylogenetic proximity in host jumps (Roy 2001; Refregier et al. 2008). Agosta et al. (2010) proposed a scenario to explain hosts jumps by specialized pathogens with a two-step process. In a first phase, host expansion is favored by ecological perturbations, which put into contact pathogens and new hosts, with a high level of stochasticity. The success of this first phase, i.e., the initiation of a new interaction, is enabled by mechanisms such as ecological fitting or exaptation, which do not require evolutionary innovation. This phase is followed by a phase of host specialization based on evolutionary changes.

This raises the question of the evolutionary potential for introduced pathogens. Indeed, species introductions are often accompanied by severe genetic bottlenecks (Lee 2002). The resulting reduced genetic diversity could happen to be insufficient to allow adaptation to the new environment (i.e., the invasion paradox, Frankham 2005). Multiple introductions from genetically divergent populations have often been invoked to explain the successful invasion of plants or animals despite this apparent paradox (e.g., Kolbe et al. 2004; Lavergne and Molofsky 2007). Multiple introductions not only increase the propagule pressure but also increase overall population genetic variance, hence evolutionary potential. In particular, recombination can occur between previously allopatric, divergent genetic lineages, what is called genetic admixture (Rius and Darling 2014). Such intraspecific hybridization generates novel allelic combinations that can be beneficial for the colonization of new hosts and environments (Lavergne and Molofsky 2007; Rius and Darling 2014). Multiple introductions have been documented for some forest pathogens, such as *C. parasitica* in Europe (Dutech et al. 2010), *F. circinatum* in Spain (Bergal et al. 2013), or *Diplodia pinea* in South Africa (Burgess et al. 2001). Genetic admixture among genetically divergent populations has been identified in some examples, such as *C. parasitica* (Dutech et al. 2012), but the demonstration of a benefit for invasion has not yet been rigorously established (Dlugosch and Parker 2008). The respective roles of admixture and adaptation within populations thus

represent an important and as-yet unanswered question in understanding the causes of invasive success in forest pathogens.

Interspecific recombination events involving nuclear and mitochondrial genomes have been increasingly recognized as an important factor in the evolution of plant pathogens and their adaptation to new environmental conditions, for example enabling them to perform host jumps between phylogenetically divergent species (Newcombe et al. 2000; Scharld and Craven 2003; Parker and Gilbert 2004; Giraud et al. 2008a,b; Stukenbrock and McDonald 2008; Stukenbrock 2013). Partial sexual compatibility between phylogenetically close species (Le Gac and Giraud 2008) and regular non-Mendelian genetic transmission (Roper et al. 2011) are commonly observed in fungi. These genetic processes increase the probability of genetic introgression among species and, therefore, the likelihood of acquisition of new genes associated with pathogenic functions (Dhillon et al. 2015). Furthermore, the dramatic increase in worldwide movement of pathogens through plant trade (Roy et al. 2014) has favored the encounter in the same environment of phylogenetic taxa that did not develop interfertility barriers during allopatric speciation (Taylor et al. 2006; Le Gac and Giraud 2008). Recent studies on *Phytophthora alni*, the causal agent of alder decline in Europe, clearly illustrate the consequences of hybridization on the emergence of a new disease. A hybrid origin of the *P. alni* complex was first concluded from the study of morphological and physiological characters (Brasier et al. 1999). The investigation of ploidy and mitochondrial and nuclear genes in the *P. alni* complex have confirmed that the epidemic taxon *P. xalni* was the result of an interspecific hybridization between two close phylogenetic species, *P. uniformis* and *Phytophthora × multiformis* (Ioos et al. 2006; Husson et al. 2015). In addition, genetic studies using polymorphic microsatellite markers strongly suggested that the parental species *P. uniformis* was alien to Europe (Aguayo et al. 2013) and that its introduction subsequently led to its hybridization with *P. × multiformis* and to the emergence of the aggressive hybrid species *P. × alni*, which is responsible for the epidemics.

Once the species barrier has been overcome and very high mortality rates have occurred in a new host, it has long been supposed that pathogen populations should evolve towards lesser virulence, in the meaning used by evolutionary biologists, i.e., the amount of damage a parasite causes to its host, encompassing both infectivity (the ability to colonize a host) and severity of the disease (Frank 1996; Parker and Gilbert 2004). Such a decline in virulence is expected as populations move toward equilibrium so long as there are trade-offs between pathogen transmission and virulence—that is, pathogens that cause their hosts to die more quickly have fewer opportunities to spread (May and Anderson 1983; Frank 1993; Alizon et al. 2009). In the case of oak powdery mildew, for example, the severity of epidemics in Europe (associated

with mortalities) rapidly declined after the first outbreaks. However, whether this decreased disease severity is attributable to a decreased virulence of the pathogen or to other factors (e.g., changes in forest practices, such as abandonment of pollarding and coppicing, or regulation by hyperparasites) remains unclear (Mougou et al. 2008). In other, long-time invasions, no decrease in virulence has yet been observed (e.g., *O. novo-ulmi* in Europe or *C. parasitica* in North America; Jarosz and Davelos 1995). In addition to the classic virulence/transmission trade-off and subsequent intermediate optimum, several factors such as cost of virulence/resistance, trait correlations, and geographic distributions can determine the outcomes for virulence evolution (Bull 1994; Frank 1996; Gandon 2002; May and Anderson 1983). Moreover, the transmission-virulence trade-off may be released in some situations, especially when high transmission rates are favored, allowing virulence to be maintained or increased (Gandon and Michalakis 2002). The strength of the virulence-transmission link varies according to pathogen life-history traits. For example, transmission is expected to remain high even for virulent strains for pathogens that can persist on their dead host as saprobes, as is often observed for necrotroph forest pathogens (e.g., Prospero et al. 2006 for *C. parasitica*). An increase in virulence is also expected to occur as a consequence of improved pathogen transmission in artificialized ecosystems, such as agroecosystems or planted forests (Stukenbrock and McDonald 2008). Large host plants surfaces with low genetic diversity maintain large pathogen populations, with high evolutionary potential and great opportunities for transmission, thus allowing rapid adaptation to the hosts (Stukenbrock and Bataillon 2012). Increased incidence of qualitative virulence types in response to forest host density has been clearly demonstrated, as in the examples of the breakdown of rust resistance types in plantation poplars (Pinon and Frey 2005). Similarly, *C. ribicola* populations have been demonstrated to increase in the proportion of strains carrying virulence types in the presence of five-needle pines with corresponding qualitative resistance (Kinloch et al. 2004, 2008). Increased virulence in the quantitative sense (often referred to aggressiveness by plant pathologists) has been reported in some agroecosystems (Milus et al. 2009; Delmotte et al. 2014). The risk of evolution of increased virulence in response to quantitative resistance has been acknowledged for forest diseases (Ennos 2001), but there is still no clear empirical demonstration. Increasing damage caused by some endemic pathogen species following the extension of host plantations on large surfaces have been reported in the last decades (Woods et al. 2005; Lieberei 2007; Fabre et al. 2011; Xhaard et al. 2012). However, climatic changes (Fabre et al. 2011; Woods et al. 2005) or simply the increase in host density and extent (Lieberei 2007; Xhaard et al. 2012) have been invoked as causes of increasing damage in these studies. Because planted forests are expected to increase rapidly in the next future, special attention toward

predicting pathogen evolution in interaction with future forest management practices is surely required.

4 Disease resistance: revisiting the ideotype concept for breeding trees

The forces that allow disease to persist at an overall moderate level in otherwise undisturbed ecosystems have long been a subject of study and speculation (May and Anderson 1983). Wide genetic variation in host resistance and pathogen virulence have been evidenced in such ecosystems, where they are sustained by and result from coevolutionary dynamics between hosts and pathogens (Burdon and Thrall 2000; Keesing et al. 2006; Burdon et al. 2013; Ennos 2015; Jousimo et al. 2014). In contrast, the association between low diversity and high disease incidence has been called the “monoculture effect” in reference to the agricultural context in which frequent, severe epidemics require the use of pesticides to maintain yield loss below acceptable levels (King and Lively 2012). This model of intensification with a strong dependence on external, fossil fuel inputs has been increasingly put into question, and new paradigms for intensification have been widely adopted (Tittonell 2014). However, recent forest intensification has largely followed the old model, with a great expansion of monospecific, and sometimes monoclonal, plantations (Paquette and Messier 2009). The philosophy of some tree breeding strategies still strongly relies on selecting or designing an ideal tree, i.e., “superior genotypes” for mass propagation using biotechnology and genome engineering (Campbell et al. 2003, Harfouche et al. 2012). The questions raised by this approach—what should be an ideal tree? how to select the ideal tree?—have been addressed by many studies and reviews (e.g., Dickmann et al. 1994; Martin et al. 2001; Dubouzet et al. 2013; Franklin et al. 2014). Here, we would like to highlight three topics that have been comparatively little dealt with or applied in the field of tree breeding: trade-offs, tolerance, and group selection.

The idea of an “ideal tree” derives from the ideotype concept, first coined by Donald (1968) for crop plants. In general, targeted phenotypes are mostly defined from growth-related traits (Martin et al. 2001; Harfouche et al. 2012). However, selecting fast-growing trees may be at the expense of defense traits, due to trade-offs between growth and defense (Franklin et al. 2014). A predicted lower resistance to pathogens and herbivores in faster growing plants due to the metabolic cost of resistance has gained empirical support (Lind et al. 2013), and has been observed in trees (Loehle and Namkoong 1987; Fine et al. 2006; Oliva et al. 2012, 2014). In a recent study with *Populus trichocarpa*, McKown et al. (2014) demonstrated the occurrence of trade-offs in the evolution of stomatal traits that were positively related to growth (photosynthesis and shoot elongation) and negatively to defense (foliar tannins and

Melampsora infection). Such negative relationships between growth and disease resistance could have important implications for breeding programs. A recent meta-analysis demonstrated that the selection of poultry lines for rapid growth unintentionally resulted in a strongly reduced immune function (Van der Most et al. 2011). The authors however concluded that, by explicitly taking the relative costs of growth and immune functions into consideration during selection, it may be possible to breed animals for increased growth without loss of immune function. It has also been shown for plants that the growth-defense trade-offs are complex and may vary according to the resistance traits involved, e.g., constitutive vs. induced resistance, or resistance *sensu stricto* vs. tolerance (Kempel et al. 2011). Moreover, environmental constraints also interact with the growth-defense trade-off to determine fitness. In a recent study, Franklin et al. (2014) proposed a mechanistic model of tree fitness integrating the most relevant traits involved in growth and survival (including defense allocation), in addition to limiting factors, such as light environment and drought, disturbance frequency (including pests and pathogens), and competition. This modeling framework can be used to explore the potential and limitations of artificial selection to increase productivity in a managed forest environment, by identifying optimal combinations of traits for a given goal and explicitly relating expected productivity gains to changes in tree vulnerability to biotic and abiotic risks in various environmental conditions. Using a similar framework, selection for qualitative disease resistance in limber pine has been shown to alter trees' drought and temperature tolerances, potentially changing their environmental range (Vogan and Schoettle 2015).

When dealing with plant (tree) defense against pathogens, much emphasis has been put on resistance in its restricted meaning, i.e., mechanisms that prevent or limit the development of the pathogen (see review in Telford et al. 2015). These mechanisms therefore exert a selection pressure on pathogen populations, the strength of which depends, on the plant side, on the level of resistance (from partial to complete) and the spatial scale of the deployment of resistant genotypes (McDonald and Linde 2002; REX Consortium 2013; Ennos 2015). The evolutionary constraints affecting the long-term efficacy of resistance strategies are well illustrated by past failures in the breeding programs of rust resistant poplar cultivars in Europe. The complete resistance conferred by race-specific genes widely deployed in monoclonal plantations was rapidly overcome by virulent races of the pathogen, in several boom-and-bust cycles (Pinon and Frey 2005). These failures have pushed breeders to consider other types of resistance, such as polygenic partial resistance, with the goal to slow down the evolution of virulent pathogen variants, thus increase the durability of resistance (Dowkiw et al. 2012). The increasing knowledge of plant-pathogen molecular interactions and technological advances are enabling the discovery

and functional profiling of new resistance genes, the breakdown of which would incur high fitness costs in corresponding virulent pathogens (Vleeshouwers et al. 2008; Michelmore et al. 2013). Tolerance in its restricted meaning (i.e., mechanisms that reduce the host fitness costs of infection, Roy and Kirchner 2000; Simms 2000) has been the subject of a growing interest as a potentially more stable defense component than true resistance since it was considered to place little or no selective pressure on the pathogen (Schafer 1971; Ney et al. 2013; Ennos 2015). Tolerance, however, is in nearly all cases expected to impose positive feedbacks leading to more prevalent and/or more virulent pathogens (Little et al. 2010; Miller, White, and Boots 2006; Roy and Kirchner 2000). Therefore, introducing tolerant hosts to a population, through breeding for example, has the potential to increase pathogen pressure overall (Restif and Koella 2004). More pathogens in the landscape can lead to negative effects on nontarget host species such as wild relatives (Power and Mitchell 2004), or even to host jumps to new species, since the probability of eventual emergence on new hosts is increased with a larger pathogen reservoir (Antia et al. 2003). While it has been observed that asymptomatic infections can be important drivers in the epidemiology of forest diseases (Elegbede et al. 2010; Denman et al. 2008), the potential for host tolerance to increase disease prevalence in adjacent populations or nontolerant host species is rarely acknowledged in a management context (Hayden et al. 2013; Little et al. 2010). Breeding strategies combining both resistance and tolerance mechanisms therefore appear to be the most desirable (Walters et al. 2012). This requires an accurate assessment of resistance and tolerance traits and their interactions (Simms and Triplett 1994), as investigated in the Douglas fir-*Armillaria ostoyae* pathosystem (Cruikshank and Jaquish 2014).

Just as disease tolerance has the potential to increase disease overall, a major challenge for breeders lies in the fact that characters of interest are relevant at stand level, while selection is based on individual performance. This issue was central to Colin Donald's ideotype concept, which balanced individual performance against competitive ability and population performance, but it was largely overlooked (Zhang et al. 1999). As an agronomist, Donald focused on yield, but his arguments are particularly relevant for disease resistance. Disease resistance can be visualized as an emerging property of populations, in that it is strongly affected by density- and frequency-dependent ecological and evolutionary processes (Burdon et al. 2014; Ennos 2015; Telford et al. 2015; Papaix et al. 2015). Indeed, durable resistance to pathogens is a population attribute that can only be assigned retrospectively (Johnson 1984). Rather than relying on "durable resistance genes," durable resistance should be considered in terms of a strategy to minimize the pace of evolution of virulence in the pathogen population (REX Consortium 2013; Michelmore et al. 2013). This can be achieved by a variety of approaches

based on an adequate deployment of multiple resistance genes either by pyramiding and/or by heterogeneous deployment in space and time, so as to diversify the selection pressure on the pathogen (McDonald and Linde 2002; Michelmore et al. 2013; REX Consortium 2013). Growing clonal mixtures of 5–20 genotypes for short-rotation coppice of willow has proven to be feasible and effective for both disease reduction and yield increase over a long term (McCracken et al. 2011). Nevertheless, such experiments with trees are relatively few, and are centered on genotypes that were selected on their individual performance. Going still further, breeding for genetically diverse populations has received increasing attention in recent years, especially in the context of a “Darwinian agriculture” to address future challenges of changing climate and sustainable management (e.g., Denison et al. 2003; Dawson and Goldringer 2011). In particular, human-mediated (artificial) group selection has been suggested to offer a wider scope for improvement than selection on individual fitness traits (Zhang et al. 1999; Denison et al. 2003). Natural selection is thought to operate mostly among individuals and more weakly at a group level, and is thus expected to have favored traits linked to individual performance even if costly to group productivity. As a consequence, individual traits that improve group performance could offer greater opportunities for artificial selection (Carroll et al. 2014). An application of this concept for resistance traits was recently proposed by Anche et al. (2014), who investigated the use of the basic reproduction ratio, R_0 , which is a key epidemiological parameter defined at the population level, in breeding programs using individual assessments. The theoretical basis of breeding strategies targeting population traits, i.e., an association of “cooperative phenotypes,” has been developed and applied with success in animals and cereals (Ellen et al. 2014; Carroll et al. 2014). Recently initiated tree species diversity experiments have allowed for the testing of associational effects (those linked to interactions between species) on traits including disease resistance and susceptibility (Castagneyrol et al. 2014; Hantsch et al. 2013, 2014). The same type of experiments could be used to explore within-species associational effects, which may be even higher in magnitude than between-species effects (Crawford and Rudgers 2013).

5 Hyperparasitism of fungal pathogens: a poorly understood but promising means of bio-control

The long-stable coexistence of plants and their pathogens is generally explained by coevolutionary processes (Newton et al. 2010; Jousimo et al. 2014). However, pathogens and plants are included in larger networks where interacting species may also contribute to the natural regulation of disease (Lafferty et al. 2006; Selakovic et al. 2014). In particular, trophic cascades, where negative effects of herbivores on plant

biomass are reversed by the effects of herbivore predators, have been widely studied and used in agricultural and forest entomology systems, including predatory insects, parasitoids, or disease-causing pathogens (Wainhouse 2005). Developing these methods for pathogens presents a challenge to forest pathologists and managers, in order to face increasing forest disease risks in the context of global change and new plantations. A necessary prerequisite for achieving this goal is improved understanding of mycoparasite diversity, ecology, and evolution (Roderick et al. 2012).

Because of their deleterious effect on host fitness (i.e., virulence), hyperparasites play a key role in the ecology and the evolution of parasite populations (Tollenaere et al. 2014). Koskella (2013, 2014) has nicely demonstrated that bacterial populations on horse-chestnut trees are engaged in coevolutionary dynamics with their virus phage parasites and moreover exhibit reciprocal adaptation across time and space. However, very few tri-trophic interactions involving a mycoparasite, a fungal or oomycete pathogen, and a tree host have been studied so far. This can be explained by the difficulty of detecting the hyperparasites (the biodiversity of which is even less known than that of parasites) and the complexity of these interactions due to the life cycles of both microbial protagonists. The best studied examples of hyperparasites of fungi are hypovirulent viruses, which decrease the virulence of their fungal hosts [e.g., *Cryphonectria Hypovirus 1* (CHV1) infecting the chestnut blight fungus and Dutch elm fungi *O. ulmi* and *O. novo-ulmi*]. Their role in the regulation of diseases has been established (Davelos and Jarosz 2004; Springer et al. 2013; Swinton and Gilligan 1999). Insights gained in CHV1 ecology (phenotypic and genetic diversity, transmission and evolution), which is used for biocontrol in chestnut orchards (Grente 1981) could help to improve biocontrol strategies. However, fungi can also be parasitized by other fungi (Barnett 1963; Kiss 2003). This fungal lifestyle has attracted an increased interest in recent years, with studies of the ecological impact of mycoparasites on fungal host populations (Tollenaere et al. 2014) and investigations about their mode of action at the molecular level (Mukherjee et al. 2013; Siozios et al. 2015). Metagenomic studies applied to mycoviruses (Pearson et al. 2009; Ghabrial and Suzuki 2009; Roossinck et al. 2010; Xie and Jiang 2014) and fungi (see next section) point to new hyperparasite or fungal antagonist candidates. Some mycoviruses have followed the route of invasion of their invasive fungal hosts and have been first detected in the introduced range of their hosts (Bryner et al. 2012; Feau et al. 2014; Voth et al. 2006; Schoebel et al. 2014; Martinez-Alvarez et al. 2014). However, according to the enemy release hypothesis (Mitchell and Power 2003), hyperparasites are expected to be more frequent in the native range of their hosts than elsewhere, and thus, systematic searches should be concentrated in these areas (Roderick and Navajas 2003; Hale et al. 2014; Ridout and Newcombe 2015). Such a

search for hyperparasites requires a known centre of origin of the parasite, which is far from the rule, as seen before.

Releases of biocontrol agents outside their native ranges have not always been successful (Roderick and Navajas 2003). For example, in spite of several introductions of CHV1 in Northern America, the hyperparasite never established there (Milgroom and Cortesi 2004). As it has been shown for invasive species, hyperparasite establishment success depends on demographic, genetic, and environmental factors (Hufbauer and Roderick 2005; Roderick et al. 2012; Fauvergue et al. 2012). The identification and selection of transmission and virulence traits is a necessary step for short-term success of hyperparasites strains for biological control (Robin et al. 2010; Swinton and Gilligan 1999). For different reasons, transmission and virulence are not easy to quantify for hyperparasites. Most mycoviruses have no extracellular life stages and are transmitted both vertically and horizontally (Pearson et al. 2009). The horizontal transmission of mycoviruses is controlled by the fungal vegetative incompatibility system, which, until recently, was thought to impede virus transmission in populations. However, some mycoviruses have a global distribution [e.g., CHV1 (Bryner et al. 2012), *Gremmenellia* virus (Botella et al. 2014)]. Recent results showed that in situ transmission rates of CHV1 are higher than in vitro studies had suggested (Carbone et al. 2004; Brusini and Robin 2013). Moreover, different strategies, aiming at decreasing the vegetative incompatibility reaction, finding a universal mycovirus donor, or creating vectors for mycoviruses, have been proposed to increase transmission efficiency of mycovirus within and among fungal species (Xie and Jiang 2014). Mycovirus virulence results from interactions of both parasite and hyperparasite genomes, which give it the attribute of an extended phenotype (Lambrechts et al. 2006). In conditions where transmission between hosts is not restricted, “mild” hyperparasite strains, characterized by low virulence and high transmission rates, are more apt to regulate disease than severe strains because they more readily establish in a host population than severe strains (Morozov et al. 2007). When between-host transmission is restricted by fungal vegetative incompatibility (transmission rates varying from 0 to 100%), mycoviruses exert a frequency-dependent selection on their host, which increases with increasing virulence of the mycovirus, resulting in an increase in fungal population diversity (Brusini et al. 2011). However, highest levels of polymorphism in the host population result in unfavorable conditions for mycovirus transmission. Thus, by this feedback mechanism, less virulent mycoviruses are likely favored in diverse fungal host populations. This theoretical result supports regulation efforts to impede an increase in the diversity of European *C. parasitica* populations in Europe where biological control of chestnut blight is still effective and used.

Establishment success of hyperparasites may also be strongly dependent on the biotic and abiotic environment.

Milgroom and Cortesi (2004) have argued that failures of biological control against chestnut blight disease in eastern North America could be linked to forest ecosystem characteristics, quite different from the European orchard agroecosystems where virus-infected strains are successfully released. This suggests that biological control may have greater success in forest plantations with higher tree density and lower heterogeneity than in seminatural forests, due to the facilitated development of the fungal parasite (Xie and Jiang 2014). Moreover, virus×fungus×environment interactions also affect the outcome of hyperparasite×parasite interactions (Bryner and Rigling 2011), which suggests that adaptation of mycovirus lineages to new environments can occur.

The hyperparasite’s evolutionary history can also provide useful insights to develop a successful biocontrol strategy. Genetic analyses of CHV1 populations in western and southern Europe have demonstrated that establishment and spread of the mycovirus are associated with a few recombinant virus lineages (Feau et al. 2014). These results, together with the known effect of hybridization on vigor of biological control agents (Szucs et al. 2012), the high rate of beneficial recombinations in viruses (Vuillaume et al. 2011), and the demonstrated high cost of bacterial resistance in coevolution with multiple phage viruses (Koskella et al. 2012) suggest to use several hyperparasite lineages in biological control strategies, especially in areas where releases of hyperparasites have failed until now. Coinfection strategies should be thoroughly studied in experimental systems prior to deployment, since the coexistence of different strains of a mycovirus within one host may not only lead to homologous recombinations but also result in the selection of more virulent strains, with lower transmission efficiency (Frank 1996; Mosquera and Adler 1998).

6 The tree microbiota: a rapidly evolving trait that will allow trees to face anthropogenic environmental change?

Due to their short generation time, pathogens are expected to quickly adapt to human-driven environmental change through changes in distribution and physiology, with or without host switches. By comparison, trees are expected to adapt more slowly and to experience a phenotype–environment mismatch (Carroll et al. 2014). There is, however, one facet of the tree phenotype that may evolve as quickly as the pathogens: the tree microbiota, which belongs to the extended phenotype of a tree (Dawkins 1999; Whitham et al. 2006; Partida-Martínez and Heil 2011). Indeed, virtually, all tissues of trees are colonized by microorganisms (Turner et al. 2013). These microbial communities, which are the outcome of millions of years of coevolution (Partida-Martínez and Heil 2011), modulate the tree’s resistance against multiple stresses, including pathogens

(Arnold et al. 2003; Hacquard and Schadt 2015). Within the current context of rapid environmental change, it is thus crucial to tackle the following issues: (i) Which properties of the resident microbiota protect the tree against biotic and abiotic stresses? (ii) How will these properties respond to environmental change? (iii) Can we select tree genotypes with beneficial microbiota to facilitate their adaptation to environmental change? Theoretical ecology provides a powerful framework for approaching the first issue. For instance, the biodiversity–ecosystem function (BEF) theory posits that biodiversity increases the intensity and stability of ecosystem functions (Cardinale et al. 2012). Future research aimed at improving tree performance could examine whether this relationship holds at the individual level. For instance, it could examine whether the diversity of tree microbiota improves the various components of tree fitness, including the resistance to pathogens. Such microbiota diversity–individual fitness (MIF) relationship can now be studied by using new sequencing technologies (Di Bella et al. 2013; Vayssier-Taussat et al. 2014), which allow the various properties of tree microbiota (community richness; taxonomic, phylogenetic, or functional diversity; and composition) to be readily described (Buée et al. 2009; Jumpponen and Jones 2009). By combining these novel technologies with methods for network inference (Faust and Raes 2012), the networks of microbial interactions can also be deciphered. Such an approach could enable the identification of potential pathogen antagonists, and thus facilitate protection strategies based on microbial epiphytes and endophytes (Witzell et al. 2014). This approach may also reveal network properties driving the stability of the residential microbiota, in particular its invasibility by pathogens (Kemen 2014).

The second issue is to assess how fast and in which direction the beneficial properties of the resident microbiota will respond to environmental change. There is already a large body of knowledge showing that the composition and structure of the tree microbiota strongly respond to environmental factors such as temperature, drought, atmospheric CO₂, or urbanization (Jumpponen and Jones 2010; Compant et al. 2010; Peñuelas et al. 2012; Cordier et al. 2012a; Coince et al. 2014; Rico et al. 2014). There is also increasing evidence showing that the tree microbiota is influenced by the genetic variability of trees, both at intraspecific (Schweitzer et al. 2008; Cordier et al. 2012b; Bálint et al. 2013) and interspecific (Knief et al. 2010; Redford et al. 2010; Kembel and Mueller 2014; Kembel et al. 2014) levels. Tree microbiota properties can thus be considered to be phenotypic traits that are influenced by both the environment and the tree genotype (Bálint et al. 2015). Phenotypic plasticity may induce a rapid change in tree microbiota properties, in response to anthropogenic environmental change. This change may then be followed by an evolutionary change, caused by the genetic adaptation of tree populations (Donnelly et al. 2012). Both mechanisms, plasticity and genetic adaptation of the tree microbiota, may

allow tree populations to cope with novel, emerging pathogens; thus, a current challenge for forest pathologists is to assess their relative influence.

Finally, the last challenge will be to integrate the beneficial properties of the tree microbiota into the selection process for tree genotypes (Newton et al. 2010; Chakraborty and Newton 2011). After the characterization of the beneficial properties through observations and manipulative experiments, their genetic architecture will have to be deciphered and their heritability to be assessed (Peiffer et al. 2013; Horton et al. 2014). Evolutionary models simulating the complex network of interactions formed by the novel tree genotypes, their pathogens, and microbiota (Kemen 2014) could then be used to define tree breeding strategies that will sustain forest health. However, such simulations are far from being trivial. Numerous evolutionary models have been developed in the case of simple, pairwise plant–pathogen interactions (Gilligan 2008; REX Consortium 2013). Extending these models to a greater number of species remains a challenge for evolutionary biologists (Brännström et al. 2012; Fortuna et al. 2013).

7 Conclusions and recommendations for the future

At the turn of the twenty-first century, Paul Manion (2003) reviewed some recently evolved concepts in forest pathology, which had modified the way of thinking about diseases and their practical management. Shortly more than 10 years after, the ecological view advocated in this prospective article has gained increasing recognition, and has also been the subject of substantial new evolution. It is remarkable to note that neither of the terms “global” nor “evolution” (except for that of concepts) were used by Manion. The recognition of global change in all its dimensions, including not only climate change but also accelerating global trade, habitat destruction, and other human-caused effects on the environment, has profoundly affected all areas of science and society in the last 15 years. In particular, two major changes affecting forest pathology—the world movement of species with trade, and the rise of plantation forestry to meet growing needs of an increasing human population—have led to an increasing number of emerging diseases, mobilizing the efforts of forest pathologists, and these trends are expected to continue (Essl et al. 2011; FAO 2010). In this context of global changes, characterized by strong new selective pressures put on entire ecosystems by human activities, the coupling of ecology and evolutionary biology has become essential for the understanding of forest diseases and their management. Three important advances, which have important practical implications and which also open new questions for future research, can be highlighted (see Table 1).

Table 1 Three important evolutionary concepts that have gained increased supporting evidence and have direct policy and practical implications and open directions for future research. A few key references are given; more are included throughout the main text

Supporting evidence	Practical and policy recommendations	Directions for future research
Not everything is everywhere		
<ul style="list-style-type: none"> Data on the biogeography of fungi and <i>Phytophthora</i> spp (Taylor et al. 2006) Emerging diseases in naive tree populations susceptible to new pathogens (Santini et al 2013) Emergence of aggressive pathogen species by hybridization of allopatric species with low interfertility barriers (Brasier 2001) “Lucky monoculture plantations” of exotic tree species that benefit from enemy escape (Pautasso et al. 2005) 	<ul style="list-style-type: none"> Appropriate trade regulations to decrease the risks associated with exotic pathogens (pathway approach) (Goss et al. 2009; Liebhold et al. 2012; Roy et al. 2014): <ul style="list-style-type: none"> – traceability (passports) for plants-for-planting (Parke and Grünwald 2012) – inclusion of the phylogenetic signal (Gilbert et al. 2012) and network theory applied to trade exchanges (Dehnen-Schmutz et al. 2010) in phytosanitary risk analysis Improvement of phytosanitary inspections (Liebhold et al. 2012; Eschen et al. 2015; Monitoring and surveillance <ul style="list-style-type: none"> – improved and non-specific detection: generic detection (e.g. of <i>Phytophthoras</i>) and detection without a priori identification (Fears et al. 2014); sentinel trees in areas of origin of imported plants and use of arboreta – improvement of disease management in nurseries (Parke et al. 2014) 	<ul style="list-style-type: none"> Further investigations of fungal diversity, especially in poorly studied and taxonomically rich ecosystems (Tedersoo et al. 2014) Theoretical developments for general predictions about the intensity of infection in various host-pathogen interactions; in-depth analyses of some communities, including life-history traits of host and pathogens, and genomic analyses of resistance genes in order to test predictions (Garcia-Guzman and Heil 2014; Gilbert and Webb 2007, Gilbert et al. 2012) Diversification, specialization and speciation in fungi and <i>Phytophthora</i> spp.: theoretical models and empirical approaches (Hamelin et al. 2011; Giraud et al. 2010)
Pathogen evolution can occur at short time scales		
<ul style="list-style-type: none"> Adaptation of pathogens after introduction (Gladieux et al. 2015) Breakdown of resistance developed in breeding programs and deployed on large scales (McDonald and Linde 2002) 	<ul style="list-style-type: none"> Avoid multiple introductions: regulations aimed not only at the species but also intra-specific level, to avoid admixture and increase efficiency of biological control Evolutionary breeding and forestry (Lefevre et al 2014; Cavers and Cottrell 2015) <ul style="list-style-type: none"> – use natural disease regulation mechanisms associated with tree diversity (Pautasso et al. 2005) - consideration of defense trade-offs in tree breeding (Vogan and Schoettle 2015; Franklin et al. 2014) - combination of resistance and tolerance mechanisms - study and prediction of associational traits: target an "ideapop" or "ideacomm" instead of an ideotype (Anche et al. 2014) 	<ul style="list-style-type: none"> Evolutionary potential of pathogens <ul style="list-style-type: none"> – experimental approach to the relative contributions of preadaptation and adaptation processes in invasive pathogens (Dlugosch and Parker 2008) – species barrier effect: genes explaining the determinants of host range in fungal species and their potential for evolution (Schulze-Lefert and Panstruga 2011) – trade-offs between virulence and other adaptive traits in pathogens: transmission, phenology (adaptation to climate) (Lively et al. 2014) – effect of hybridization, admixture and multiple infections on the evolution of virulence (Susi et al. 2014) – empirical studies of the impact of host tolerance and qualitative and quantitative resistance to disease on the evolution of virulence in forest systems Study of general/cross resistance mechanisms to disease in trees (Yanchuk and Allard 2009) <ul style="list-style-type: none"> – nonhost resistance (Schulze-Lefert and Panstruga 2011), recessive resistance (van Schie and Takken 2014) Development of demo-genetic models of host–parasite interactions (e.g., Bazin et al. 2014) Ecological and evolutionary mechanisms regulating disease in wild plant pathosystems (Jousimo et al. 2014) Ecoimmunology (Schulenburg et al 2009)
The tree is a multitrophic community		
<ul style="list-style-type: none"> Tree pathogens have their own enemies (Kiss 2003; Pearson et al 2009) Microbial communities around and inside trees are highly diverse (Buée et al. 2009; Peay et al. 2013) 	<ul style="list-style-type: none"> More systematic search of antagonists and hyperparasites in the native ranges of forest pathogens (Hale et al. 2014; Ridout and Newcombe 2015) Development of platforms for the coupling of metagenomic and phenotypic characterization of microbial communities (Lebeis 2014) 	<ul style="list-style-type: none"> Mycoparasitism by fungi, bacteria, virus <ul style="list-style-type: none"> – significance in natural ecosystems (Tollenaere et al. 2014) – role of hybridization, admixture and multiple infections on the efficiency and evolution of the biocontrol agents (Feau et al. 2014) Tree endophytic fungi: continuum or shift from opportunistic, weak pathogenicity, and their interaction with true pathogens (Pautasso et al. 2014).

Table 1 (continued)

Supporting evidence	Practical and policy recommendations	Directions for future research
		<ul style="list-style-type: none"> • Mechanisms through which plants control their microbiota and through which the microbiome controls plant health (Berendsen et al. 2012; Pautasso et al. 2014). – identifying the microbial species and community properties sustaining tree health (Hacquard and Schadt 2015) – understanding the genetic vs. environmental determinism of beneficial microbiota (Lively et al. 2014; Pautasso et al. 2014); how does plant immune system regulate plant microbiote? (Lebeis 2014; Kliebenstein 2014) – integration of the microbiota into epidemiological and evolutionary models of tree disease (Bálint et al. 2015)

1. Not everything is everywhere (especially forest pathogens)

The view on the processes underlying geographic variation in microbial distributions has changed from purely ecological explanations (“everything is everywhere but the environment selects,” according to Beijerinck’s and Becking’s famous claim; see in O’Malley 2008) to a combination of ecological and evolutionary (diversification) explanations supported by evidence of limited (natural) dispersal in many microbial taxa (Taylor et al. 2006; Hanson et al. 2012; Tedersoo et al. 2014). Human-mediated transfers putting into contact non-coevolved species therefore result in ecological and evolutionary dynamics that challenge processes operating in natural communities (Burdon et al. 2013). These new introductions may overwhelm the ability of the community, or even of artificial selection, to respond (Cobb et al. 2012). For instance, European and American chestnuts have been successively impacted by three exotic pests and pathogens in the last century, *Phytophthora cinnamomi*, *C. parasitica*, and *Dryocosmus kiruphilus*, showing the increasing difficulty for breeding programs to keep pace with biotic threats in woody perennials (Anagnostakis 2012).

2. Pathogen evolution can occur at short time scales

The recognition that evolution can take place at short time scales is not very new (e.g., Cook and Askew 1970), but it has been increasingly supported by evidence linked to anthropogenic changes, such as breakdowns of human-selected resistance in a medical or agricultural context (McDonald and Linde 2002; REX 2013) and adaptation of invasive organisms (Gladieux et al. 2015). Forests have already paid heavy tribute to invasions by exotic pathogens, but still, limited domestication of forest trees has up to now restricted impacts on forest pathogen virulence as those seen in agriculture. Phytosanitary problems in monospecific (monoclonal) planted forests are likely to increase in the future, since experience demonstrates that

high productivity achieved in the first rotations, often associated with enemy escape, is likely to fade away with time due to the progressive arrival of natural enemies, i.e. tree pests and diseases (Harwood and Nambiar 2014; Liebhold and Wingfield 2014). New prospects for an ecological and evolutionary intensification could benefit from rapid changes in technology and wood usages as material or fuel (e.g., engineered wood products, composite materials, wood pellets), which will likely remove some of the obstacles to the use of diverse plantations.

3. The tree is a multitrophic community

Community fingerprinting methods, and, more recently, high-throughput sequencing techniques, have revealed that the epiphytic and endophytic microbial communities are much more complex than previously realized (e.g., Yang et al. 2001; Jumpponen and Jones 2009). Pathogens themselves host a diversity of parasites, especially mycoviruses (Witzell et al. 2014). Some experimental manipulations at small spatial scales have highlighted the role of microbial communities in disease control (Arnold et al. 2003; Matos et al. 2005), while observations at larger scales have revealed environmental factors driving these communities (Cordier et al. 2012a; Peñuelas et al. 2012). Both approaches are required to predict, and ultimately manage, microbial interactions and functions in the context of global change (Kinkel et al. 2011; Berlec 2012; Borer et al. 2013; Lebeis 2014; Vayssier-Taussat et al. 2014; Hacquard and Schadt 2015). Future studies working toward this aim will benefit from the development of functional metagenomics (Knief et al. 2012), microbial interactions modeling (Faust and Raes 2012), as well as evolutionary platforms predicting the outcomes of multispecies interactions (Fortuna et al. 2013).

Uncertainties, especially regarding the emergence of new diseases, are likely a key component of the future of forest pathology (Shaw and Osborne 2011). An obvious policy recommendation is therefore to maintain the

expertise in the taxonomy, biology, epidemiology, and ecology of forest fungal pathogens needed to respond quickly to disease outbreaks in more economically developed countries (BSPP 2012). Special efforts should also be developed in regions, such as tropical areas, where high biological diversity is present but has received limited attention so far due to a lack of funding and trained researchers (Kamgan et al. 2008). Spectacular advances and reducing costs in DNA technologies have recently enabled considerable progress in the knowledge of phylogenetic diversity (inter- and intraspecies) of forest pathogens. However, an ecological and evolutionary forest pathology cannot neglect the study of phenotypes and functional traits. Investigations of a greater number of forest pathosystems (not only focused on recent introductions) and the development of high-throughput phenotyping methods are certainly deserving more efforts.

Forest pathologists have increasingly tried to disseminate their research outcomes to contribute to policy recommendations (e.g. the Montesclaros declaration, <http://www.iufro.org/science/divisions/division-7/70000/publications/montesclaros-declaration/>; Cavers 2015). Increasing links between biologists, sociologists, and economists, for example to investigate the perception of risks or the economic impacts of pathogens, or to assess the ecological service of disease regulation, could help identify the best levers to tackle forest health issues in a globalised and changing environment (Stenlid et al. 2011; Fears et al. 2014). Particular effort should also be dedicated to education, not only to raise awareness in the public and stakeholders of forest health issues and their dependence on human activities but also to attract young scientists, from diverse scientific horizons, to the field of forest pathology.

Acknowledgments The authors are very grateful to Catherine Bastien, Isabelle Goldringer, and François Lefèvre for stimulating discussions, and to Sylvain Delzon for helpful comments.

Funding BJ was funded by a grant from the French Ministry of Research and Education (MENRT no. 2011/AF/57).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Adams GC, Catal M, Trummer LM (2009) Distribution and severity of Alder *Phytophthora* in Alaska. In: Proceedings of the sudden oak death fourth science symposium (ed. by Gen. Tech. Rep. PSW-GTR-229. Department of Agriculture FS), pp. 29–49
- Agosta SJ, Klemens JA (2008) Ecological fitting by phenotypically flexible genotypes: implications for species associations, community assembly, and evolution. *Ecol Lett* 11:1123–1134
- Agosta SJ, Janz N, Brooks DB (2010) How specialists can be generalists: resolving the “parasite paradox” and implications for emerging infectious disease. *Zoologia* 27:151–162
- Aguayo J, Adams GC, Halkett F, Catal M, Husson C, Nagy ZA, Hansen EM, Marçais B, Frey P (2013) Strong genetic differentiation between North American and European populations of *Phytophthora alni* subsp. *uniformis*. *Phytopathology* 10:190–199
- Alizon S, Hurford A, Mideo N, Van Baalen M (2009) Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J Evol Biol* 22:245–259
- Altizer S, Harvell CD, Friedle E (2003) Rapid evolutionary dynamics and disease threats to biodiversity. *Trends Ecol Evol* 18:589–596
- Anagnostakis SL (1987) Chestnut blight: the classical problem of an introduced pathogen. *Mycologia* 79:23–37
- Anagnostakis SL (2012) Chestnut breeding in the United States for disease and insect resistance. *Plant Dis* 96:1392–1403
- Anche MT, de Jong MCM, Bijma P (2014) On the definition and utilization of heritable variation among hosts in reproduction ratio R0 for infectious diseases. *Heredity* 113:364–374
- Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol Evol* 19:535–544
- Antia R, Regoes RR, Koella JC, Bergstrom CT (2003) The role of evolution in the emergence of infectious diseases. *Nature* 426:658–661
- Arnold AE, Mejía LC, Kylo D, Rojas EI, Maynard Z, Robbins N, Herre EA (2003) Fungal endophytes limit pathogen damage in a tropical tree. *Proc Natl Acad Sci U S A* 100:15649–15654
- Bálint M, Tiffin P, Hallström B, O'Hara RB, Olson MS, Fankhauser JD, Piepenbring M, Schmitt I (2013) Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). *PLoS One* 8, e53987. doi:10.1371/journal.pone.0053987
- Bálint M, Barta L, Hara RBO, Olson MS, Otte J, Pfenninger M, Robertson AL, Tiffin P, Schmitt I (2015) Relocation, high-latitude warming and host genetic identity shape the foliar fungal microbiome of poplars. *Mol Ecol*. doi:10.1111/mec.13018
- Barnes I, Crous PW, Wingfield BD, Wingfield MJ (2004) Multigene phylogenies reveal that red band needle blight of *Pinus* is caused by two distinct species of *Dothistroma*, *D-septosporum* and *D-pini*. *Stud Mycol* 50:551–565
- Barnett HL (1963) The nature of mycoparasitism by fungi. *Annu Rev Microbiol* 17:1–14
- Barrès B, Carlier J, Seguin M, Fenouillet C, Cilas C, Ravigne V (2012) Understanding the recent colonization history of a plant pathogenic fungus using population genetic tools and Approximate Bayesian Computation. *Heredity (Edinb)* 109:269–279
- Bazin E, Mathe-Hubert H, Facon B, Carlier J, Ravigne V (2014) The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. *Biol Invasions* 16:875–886
- Berbegal M, Perez-Sierra A, Armengol J, Grünwald NJ (2013) Evidence for multiple introductions and clonality in Spanish populations of *Fusarium circinatum*. *Phytopathology* 103:851–861
- Berendsen RL, Pieterse CMJ, Bakker P (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci* 17:478–486
- Bergot M, Cloppet E, Perarnaud V, Déqué M, Marçais B, Desprez-Loustau ML (2004) Simulation of potential range expansion of oak disease caused by *Phytophthora cinnamomi* under climate change. *Glob Chang Biol* 10:1–14
- Berlec A (2012) Novel techniques and findings in the study of plant microbiota: search for plant probiotics. *Plant Sci Int J Exp Plant Biol* 193–194:96–102

- Blackwell M (2011) The Fungi: 1, 2, 3 ... 5.1 million species? *Am J Bot* 98:426–438
- Booth TH, Jovanovic T, Old KM, Dudzinski MJ (2000) Climatic mapping to identify high-risk areas for *Cylindrocladium quinqueseptatum* leaf blight on eucalypts in mainland South East Asia and around the world. *Environ Pollut* 108:365–372
- Borer ET, Kinkel LL, May G, Seabloom EW (2013) The world within: quantifying the determinants and outcomes of a host's microbiome. *Basic Appl Ecol* 14:533–539
- Botella L, Tuomivirta TT, Vervuurt S, Diez JJ, Hantula J (2014) Occurrence of two different species of mitoviruses in the European race of *Gremmeniella abietina* var. *abietina*, both hosted by the genetically unique Spanish population. *Fungal Biol* 116:872–882
- Brännström Å, Johansson J, Loeuille N, Kristensen N, Troost T, Hille Ris Lambers R, Dieckmann U (2012) Modelling the ecology and evolution of communities: a review of past achievements, current efforts, and future promises. *Evol Ecol Res* 14:601–625
- Brasier CM (1996) *Phytophthora cinnamomi* and oak decline in southern Europe. Environmental constraints including climate change. *Ann Sci For* 53:347–358
- Brasier CM (2001) Rapid evolution of introduced plant pathogens via interspecific hybridization. *Bioscience* 51:123–133
- Brasier CM (2008) The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathol* 57:792–808
- Brasier CM (2009) *Phytophthora* biodiversity: how many *Phytophthora* species are there? General Technical Report—Pacific Southwest Research Station, USDA Forest Service
- Brasier CM, Mehrotra MD (1995) *Ophiostoma himal-ulmi* sp. nov., a new species of Dutch elm disease fungus endemic to the Himalayas. *Mycol Res* 99:205–215
- Brasier CM, Webber J (2010) Plant pathology. Sudden larch death. *Nature* 466:824–825
- Brasier CM, Rose J, Gibbs JN (1995) An unusual *Phytophthora* associated with widespread alder mortality in Britain. *Plant Pathol* 44:999–1007
- Brasier CM, Cooke DEL, Duncan JM (1999) Origin of a new *Phytophthora* pathogen through interspecific hybridization. *Proc Natl Acad Sci U S A* 96:5878–5883
- Brasier CM, Vettraino AM, Chang TT, Vannini A (2010) *Phytophthora lateralis* discovered in an old growth *Chamaecyparis* forest in Taiwan. *Plant Pathol* 59:595–603
- Brusini J, Robin C (2013) Mycovirus transmission revisited by in situ pairings of vegetatively incompatible isolates of *Cryphonectria parasitica*. *J Virol Methods* 187:435–442
- Brusini J, Robin C, Franc A (2011) Parasitism and maintenance of diversity in fungal vegetative incompatibility system: the role of selection by deleterious cytoplasmic elements. *Ecol Lett* 14:444–452
- Bryner SF, Rigling D (2011) Temperature-dependent genotype-by-phenotype interaction between a pathogenic fungus and its hyperparasitic virus. *Am Nat* 177:65–74
- Bryner SF, Rigling D, Brunner PC (2012) Invasion history and demographic pattern of *Cryphonectria hypovirus* 1 across European populations of the chestnut blight fungus. *Ecol Evol* 2:3227–3241
- BSPP (2012) Plant pathology education and training in the UK: an audit. <http://www.bspp.org.uk>
- Buée M, Reich M, Murat C, Morin E, Nilsson RH, Uroz S, Martin F (2009) 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytol* 184:449–456
- Bull J (1994) Perspective: virulence. *Evolution* 48:1423–1437
- Burdon JJ (1987) Diseases and plant population biology. Cambridge University Press, Cambridge
- Burdon JJ, Thrall PH (2008) Pathogen evolution across the agro-ecological interface: implications for disease management. *Evol Appl* 1:57–65
- Burdon JJ, Thrall PH (2000) Coevolution at multiple spatial scales - from population to continent and beyond. *Evol Ecol* 14:261–281
- Burdon JJ, Thrall PH, Ericson L (2013) Genes, Communities and Invasive Species: Understanding the Ecological and Evolutionary Dynamics of Host–Pathogen Interactions. *Curr Opin Plant Biol* 16: 1–6
- Burdon JJ, Barrett LG, Rebetzke G, Thrall PH (2014) Guiding deployment of resistance in cereals using evolutionary principles. *Evol Appl* 7:609–624
- Burgess T, Wingfield BD, Wingfield MJ (2001) Comparison of genotypic diversity in native and introduced populations of *Sphaeropsis sapinea* isolated from *Pinus radiata*. *Mycol Res* 105:1331–1339
- Burgess TI, Webster JL, Ciampini JA, White D, Hardy GES, Stukely MJC (2009) Re-evaluation of *Phytophthora* species isolated during 30 years of vegetation health surveys in Western Australia using molecular techniques. *Plant Dis* 93:215–223
- Cai L, Graud T, Zhang N, Begerow D, Cai G, Shivas R (2011) The evolution of species concepts and species recognition criteria in plant pathogenic fungi. *Fungal Divers* 50:121–133
- Campbell MM, Brunner AM, Jones HM, Strauss SH (2003) Forestry's fertile crescent: the application of biotechnology to forest trees. *Plant Biotechnol J* 1:141–154
- Carbone I, Liu YC, Hillman BI, Milgroom MG (2004) Recombination and migration of *Cryphonectria hypovirus* 1 as inferred from gene genealogies and the coalescent. *Genetics* 166:1611–1629
- Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA, Kinzig AP, Daily GC, Loreau M, Grace JB, Larigauderie A, Srivastava D, Naem S (2012) Biodiversity loss and its impact on humanity. *Nature* 486:59–67
- Carroll SP, Jørgensen PS, Kinnison MT, Bergstrom CT, Denison RF, Gluckman P, Smith TB, Strauss SY, Tabashnik BE (2014) Applying evolutionary biology to address global challenges. *Science* 80:1–16
- Castagnyrol B, Jactel H, Vacher C, Brockerhoff EG, Koricheva J (2014) Effects of plant phylogenetic diversity on herbivory depend on herbivore specialization. *J Appl Ecol* 51:134–141
- Cavers S (2015) Evolution, ecology and tree health: finding ways to prepare Britain's forests for future threats. *Forestry* 88:1–2
- Cavers S, Cottrell JE (2015) The basis of resilience in forest tree species and its use in adaptive forest management in Britain. *Forestry* 88: 13–26
- Chakraborty S, Newton AC (2011) Climate change, plant diseases and food security: an overview. *Plant Pathol* 60:2–14
- Cobb RC, Filipe JAN, Meentemeyer RK, Gilligan CA, Rizzo DM (2012) Ecosystem transformation by emerging infectious disease: loss of large tanoak from California forests. *J Ecol* 100:712–722
- Coince A, Cordier T, Lengellé J, Defossez E, Vacher C, Robin C, Buée M, Marçais B (2014) Leaf and root-associated fungal assemblages do not follow similar elevational diversity patterns. *PLoS One* 9, e100668
- Combes C (2001) Parasitism: ecology and evolution of intimate interactions. Press, Chicago University, **552 pp**
- Compant S, van der Heijden MG, Sessitsch A (2010) Climate change effects on beneficial plant-microorganism interactions. *FEMS Microbiol Ecol* 73:197–214
- Cook LM, Askew RR (1970) Increasing frequency of the typical form of the peppered moth in Manchester. *Nature* 227:1155
- Cordier T, Robin C, Capdevielle X, Fabreguettes O, Desprez-Loustau ML, Vacher C (2012a) The composition of phyllosphere fungal assemblages of European beech (*Fagus sylvatica*) varies significantly along an elevation gradient. *New Phytol* 196:510–519
- Cordier T, Robin C, Capdevielle X, Desprez-Loustau M-L, Vacher C (2012b) Spatial variability of phyllosphere fungal assemblages: genetic distance predominates over geographic distance in a European beech stand (*Fagus sylvatica*). *Fungal Ecol* 5:509–520

- Crawford KM, Rudgers JA (2013) Genetic diversity within a dominant plant outweighs plant species diversity in structuring an arthropod community. *Ecology* 94:1025–1035
- Crous PW, Groenewald JZ, Pongpanich K, Himaman W, Arzanlou M, Wingfield MJ (2004) Cryptic speciation and host specificity among *Mycosphaerella* spp. occurring on Australian *Acacia* species grown as exotics in the tropics. *Stud Mycol* 50:457–469
- Cruikshank MG, Jaquish B (2014) Resistance and tolerance in juvenile interior Douglas-fir trees *Pseudotsuga menziesii* var. *glauca* artificially inoculated with *Armillaria ostoyae*. *For Pathol* 44:362–371
- Davelos AL, Jarosz AM (2004) Demography of American chestnut populations: effects of a pathogen and a hyperparasite. *J Ecol* 92:675–685
- Dawkins R (1999) *The extended phenotype*. Oxford University Press, Oxford
- Dawson JC, Goldringer I (2011) Breeding for genetically diverse populations: variety mixtures and evolutionary populations. In: Lammerts van Bueren ET, Myers JR (eds) *Organic crop breeding*. Wiley-Blackwell, Hoboken, pp 77–98
- de Beer ZW, Duong TA, Barnes I, Wingfield BD, Wingfield MJ (2014) Redefining *Ceratocystis* and allied genera. *Stud Mycol* 79:187–219
- de Vienne DM, Hood ME, Giraud T (2009) Phylogenetic determinants of potential host shifts in fungal pathogens. *J Evol Biol* 22:2532–2541
- de Wet J, Burgess T, Slippers B, Preisig O, Wingfield BD, Wingfield MJ (2003) Multiple gene genealogies and microsatellite markers reflect relationships between morphotypes of *Sphaeropsis sapinea* and distinguish a new species of *Diplodia*. *Mycol Res* 107:557–566
- Dehnen-Schmutz K, Holdenrieder O, Jeger MJ, Pautasso M (2010) Structural change in the international horticultural industry: some implications for plant health. *Sci Hortic* 125:1–15
- Delamour C (1983) Les dépérissements de Chênes en Europe. *Rev For France* 35:262–282
- Delamour C, Pinon J, Morelet M (1985) Histoire et avenir de la pathologie forestière en France. *Rev For France* 37:65–82
- Delmotte F, Mestre P, Schneider C, Kassemeyer HH, Kozma P, Richart-Cervera S, Rouxel M, Delière L (2014) Rapid and multiregional adaptation to host partial resistance in a plant pathogenic oomycete: evidence from European populations of *Plasmopara viticola*, the causal agent of grapevine downy mildew. *Infect Genet Evol* 27:500–508
- Denison RF, Kiers ET, West SA (2003) Darwinian agriculture: when can humans find solutions beyond the reach of natural selection? *Q Rev Biol* 78:145–168
- Denman S, Mralejo E, Kirk S, et al (2008) Sporulation of *Phytophthora ramorum* and *P. kernoviae* on asymptomatic foliage and fruit. *Proc Sudd Oak Death Third Sci Symp General Technical Report PSW-GTR-214*: 201–207
- Desprez-Loustau ML, Robin C, Buee M, Courtecuisse R, Garbaye J, Suffert F, Sache I, Rizzo DM (2007) The fungal dimension of biological invasions. *Trends Ecol Evol* 22:472–480
- Desprez-Loustau ML, Courtecuisse R, Robin C, Husson C, Moreau PA, Blancard D, Selosse MA, Lung-Escarmant B, Piou D, Sache I (2010) Species diversity and drivers of spread of alien fungi (sensu lato) in Europe with a particular focus on France. *Biol Invasions* 12:157–172
- Desprez-Loustau M-L, Feau N, Mougou-Hamdane A, Dutech C (2011) Interspecific and intraspecific diversity in oak powdery mildews in Europe: coevolution history and adaptation to their hosts. *Mycoscience* 52:163–175
- Dhillon B, Feau N, Aerts AL, Beauseigle S, Bernier L, Copeland A, Foster A, Gill N, Henriessat B, Herath P, LaButti KM, Levasseur A, Lindquist EA, Majoro E, Ohm RA, Pangilinan JL, Pribowo A, Sandler JN, Sakalidis ML, de Vries RP, Grigoriev IV, Goodwin SB, Tanguay P, Hamelin RC (2015) Horizontal gene transfer and gene dosage drives adaptation to wood colonization in a tree pathogen. *Proc Natl Acad Sci U S A* 112:3451–3456
- Di Bella JM, Bao Y, Gloor GB, Burton JP, Reid G (2013) High throughput sequencing methods and analysis for microbiome research. *J Microbiol Methods* 95:401–414
- Dickmann DI, Gold MA, Flore JA (1994) The ideotype concept and the genetic improvement of tree crops. *Plant Breed Rev* 12: 163–193
- Dlugosch K, Parker I (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* 17:431–449
- Donald CM (1968) The breeding of crop ideotype. *Euphytica* 17:385–403
- Dong S, Stam R, Cano LM, Song J, Sklenar J, Yoshida K, Bozkurt TO, Oliva R, Liu Z, Tian M, Win J, Banfield MJ, Jones AM, van der Hoorn RA, Kamoun S (2014) Effector specialization in a lineage of the Irish potato famine pathogen. *Science* 343:552–555
- Donnelly A, Caffarra A, Kelleher CT, O'Neill BF, Diskin E, Pletsers A, Proctor H, Stimmemann R, O'Halloran J, Peñuelas J, Hodkinson TR, Sparks TH (2012) Surviving in a warmer world: environmental and genetic responses. *Clim Res* 53:245–262
- Dowkiw A, Jorge V, Villar M, Voisin E, Guérin V, Faivre-Rampant P, Bresson A, Bitton F, Duplessis S, Frey P, Petre B, Guinet C, Xhaard C, Fabre B, Halkett F, Plomion C, Lalanne C, Bastien C (2012) Breeding poplars with durable resistance to *Melampsora larici-populina* leaf rust: a multidisciplinary approach to understand and delay pathogen adaptation. In: Snieszko RA et al (eds) *Proceedings of the fourth international workshop on the genetics of host-parasite interactions in forestry: disease and insect resistance in forest trees*, Gen. Tech. Rep. PSW-GTR-240. Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Albany, pp 31–38
- Dubouzet JG, Strabala TJ, Wagner A (2013) Potential transgenic routes to increase tree biomass. *Plant Sci* 212:72–101
- Dutech C, Fabreguettes O, Capdevielle X, Robin C (2010) Multiple introductions of divergent genetic lineages in an invasive fungal pathogen, *Cryphonectria parasitica*, in France. *Heredity* 105:220–228
- Dutech C, Barres B, Bridier J, Robin C, Milgroom MG, Ravigne V (2012) The chestnut blight fungus world tour: successive introduction events from diverse origins in an invasive plant fungal pathogen. *Mol Ecol* 21:3931–3946
- Elegbede CF, Pierrat J-C, Aguayo J, Husson C, Halkett F, Marçais B (2010) A statistical model to detect asymptomatic infectious individuals with an application in the *Phytophthora alni*-induced alder decline. *Phytopathology* 100:1262–1269
- Ellen ED, Rodenburg TB, Albers GA, Bolhuis JE, Camerlink I, Duijvesteijn N, Knol EF, Muir WM, Peeters K, Reimert I, Sell-Kubiak E, van Arendonk JA, Visscher J, Bijm P (2014) The prospects of selection for social genetic effects to improve welfare and productivity in livestock. *Front Genet* 5:377
- Elton CS (1958) *The ecology of invasions by animals and plants*, University of Chicago Press edth edn. University of Chicago Press, Chicago
- Ennos RA (2001) The introduction of lodgepole pine as a major forest crop in Sweden: implications for host–pathogen evolution. *For Ecol Manag* 141:85–96
- Ennos RA (2015) Resilience of forests to pathogens: an evolutionary ecology perspective. *Forestry* 88:41–52
- Eschen R, Rigaux L, Sukovata L, Vettriano AM, Marzano M, Gregoire JC (2015) Phytosanitary inspection of woody plants for planting at European Union entry points: a practical enquiry. *Biol Invasions*. doi:10.1007/s10530-015-0883-6
- Essl F, Dullinger S, Rabitsch W, Hulme PE, Hülber K, Jarosik V, Kleinbauer I, Krausmann F, Kühn I, Nentwig W, Vilà M, Genovesi P, Gherardi F, Desprez-Loustau ML, Roques A, Pysek P (2011) Socioeconomic legacy yields an invasion debt. *Proc Natl Acad Sci U S A* 108:203–207
- Fabre B, Ios R, Piou D, Marçais B (2011) Is the emergence of *Dothistroma* needle blight of pine in France caused by the cryptic species *Dothistroma pini*? *Phytopathology* 102:47–54

- Facon B, Genton B, Shykoff J, Jarne P, Estoup A, David P (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends Ecol Evol* 21:130–135
- FAO (2010) Global forest resources assessment, FAO Forestry Paper 163, Rome
- Faust K, Raes J (2012) Microbial interactions: from networks to models. *Nat Rev Microbiol* 10:538–550
- Fauvergue X, Vercken E, Malausa T, Hufbauer RA (2012) The biology of small, introduced populations, with special reference to biological control. *Evol Appl* 5:424–443
- Fears R, Aroemail EM, Pais MS, ter Meulen V (2014) How should we tackle the global risks to plant health? *Trends Plant Sci* 19:206–208
- Feau N, Lauron-Moreau A, Piou D, Marçais B, Dutech C, Desprez-Loustau ML (2012) Niche partitioning of genetic lineages involved in the oak powdery mildew complex. *Fungal Ecol* 5:154–162
- Feau N, Dutech C, Brusini J, Rigling D, Robin C (2014) Multiple introductions and recombination in *Cryphonectria hypovirus 1*: perspective for a sustainable biological control of chestnut blight. *Evol Appl* 7:580–596
- Fine PVA, Miller ZJ, Mesones I, Irazuzta S, Appel HM, Stevens MHH, Saaksjarvi I, Schultz LC, Coley PD (2006) The growth defense trade-off and habitat specialization by plants in Amazonian forests. *Ecology* 87:S150–S162
- Fitt BDL, Huang YJ, van den Bosch F, West JS (2006) Coexistence of related pathogen species on arable crops in space and time. *Annu Rev Phytopathol* 44:163–182
- Fortuna MA, Zaman L, Wagner AP, Ofria C (2013) Evolving digital ecological networks. *PLoS Comput Biol* 9, e1002928. doi:10.1371/journal.pcbi.1002928
- Franceschini S, Webber JF, Sancisi-Frey S, Brasier CM (2014) Gene × environment tests discriminate the new EU2 evolutionary lineage of *Phytophthora ramorum* and indicate that it is adaptively different. *For Pathol* 44:219–232
- Frank SA (1993) Evolution of host–parasite diversity. *Evolution* 47:1721–1732
- Frank S (1996) Models of parasite virulence. *Q Rev Biol* 71:37–78
- Frankham R (2005) Invasion biology—resolving the genetic paradox in invasive species. *Heredity* 94:385
- Franklin O, Palmroth S, Nasholm T (2014) How eco-evolutionary principles can guide tree breeding and tree biotechnology for enhanced productivity. *Tree Physiol* 34:1149–1166
- Gandon S (2002) Local adaptation and the geometry of host–parasite coevolution. *Ecol Lett* 5:246–256
- Gandon S, Michalakis Y (2002) Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *J Evol Biol* 15:451–462
- Garbelotto M (2008) Molecular analysis to study invasions by forest pathogens: examples from Mediterranean ecosystems. *Phytopathol Mediterr* 47:183–203
- Garbelotto M, Linzer R, Nicolotti G, Gonthier P (2010) Comparing the influences of ecological and evolutionary factors on the successful invasion of a fungal forest pathogen. *Biol Invasions* 12:943–957
- García-Guzmán G, Heil M (2014) Life histories of hosts and pathogens predict patterns in tropical fungal plant diseases. *New Phytol* 201:1106–1120
- Ghabrial SA, Suzuki N (2009) Viruses of plant pathogenic fungi. *Annu Rev Phytopathol* 47:353–384
- Gilbert G (2002) Evolutionary ecology of plant diseases in natural ecosystems. *Annu Rev Phytopathol* 40:13–43
- Gilbert GS, Webb CO (2007) Phylogenetic signal in plant pathogen–host range. *Proc Natl Acad Sci U S A* 104:4979–4983
- Gilbert GS, Magarey R, Suiter K, Webb CO (2012) Evolutionary tools for phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant pests and pathogens. *Evol Appl* 5:869–878
- Gilligan CA (2008) Sustainable agriculture and plant diseases: an epidemiological perspective. *Philos Trans R Soc Lond B Biol Sci* 363:741–759
- Giraud T, Enjalbert J, Fournier E, Delmotte F, Dutech C (2008a) Population genetics of fungal diseases of plants. *Parasite* 15:449–454
- Giraud T, Refrégier G, Le Gac M, De Vienne DM, Hood ME (2008b) Speciation in fungi. *Fungal Genet Biol* 45:791–802
- Giraud T, Gladieux P, Gavrillets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends Ecol Evol* 25:387–395
- Gladieux P, Feurtey A, Hood ME, Snirc A, Clavel J, Dutech C, Roy M, Giraud T (2015) The population biology of fungal invasions. *Mol Ecol*. doi:10.1111/mec.13028
- Goss EM, Larsen M, Chastagner GA, Givens DR, Grünwald NJ (2009) Population genetic analysis infers migration pathways of *Phytophthora ramorum* in US nurseries. *PLoS Pathog* 5, e1000583
- Goss EM, Tabima JF, Cooke DEL, Restrepo S, Fry WE, Forbes GA, Fieland VJ, Cardenas M, Grünwald NJ (2014) The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. *Proc Natl Acad Sci U S A* 111:8791–8796
- Graça RN, Ross-Davis AL, Klopfenstein NB, Kim M-S, Peever TL, Cannon PG, Aun CP, Mizubuti ESG, Alfenas AC (2013) Rust disease of eucalypts, caused by *Puccinia psidii*, did not originate via host jump from guava in Brazil. *Mol Ecol* 22:6033–6047
- Grente J (1981) Les variants hypovirulents de l'*Endothia parasitica* et la lutte biologique contre le chancre du châtaignier. Université de Bretagne Occidentale, Brest, France, Dissertation
- Gross A, Hosoya T, Queloz V (2014) Population structure of the invasive forest pathogen *Hymenoscyphus pseudoalbidus*. *Mol Ecol* 23:2943–2960
- Grünwald NJ, Goss EM (2011) Evolution and population genetics of exotic and re-emerging pathogens: novel tools and approaches. *Annu Rev Phytopathol* 49:249–267
- Hacquard S, Schadt CW (2015) Towards a holistic understanding of the beneficial interactions across the *Populus* microbiome. *New Phytol* 205:1425–1430
- Hale IL, Broders K, Iriarte G (2014) A Vavilovian approach to discovering crop-associated microbes with potential to enhance plant immunity. *Front Plant Sci* 5:492
- Hamelin F, Castel M, Poggi S, Andrivon D, Mailleret L (2011) Seasonality and the evolutionary divergence of plant parasites. *Ecology* 92:2159–2166
- Hansen EM (2008) Alien forest pathogens: *Phytophthora* species are changing world forests. *Boreal Environ Res* 13:33–41
- Hansen E, Goheen E (2000) *Phellinus weirii* and other native root pathogens as determinants of forest structure and process in western North America. *Annu Rev Phytopathol* 38:515–539
- Hansen EM, Parke JL, Sutton W (2005) Susceptibility of Oregon forest trees and shrubs to *Phytophthora ramorum*: a comparison of artificial inoculation and natural infection. *Plant Dis* 89:63–70
- Hansen EM, Reeser PW, Sutton W (2011) *Phytophthora* beyond agriculture. *Annu Rev Phytopathol* 50:359–378
- Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* 10:497–506
- Hantsch L, Braun U, Scherer-Lorenzen M, Bruehlheide H (2013) Species richness and species identity effects on occurrence of foliar fungal pathogens in a tree diversity experiment. *Ecosphere* 4:81
- Hantsch L, Braun U, Haase J, Purschke O, Scherer-Lorenzen M, Bruehlheide H (2014) No plant functional diversity effects on foliar fungal pathogens in experimental tree communities. *Fungal Divers* 66:139–151

- Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, ScarasciaMugnozza G (2012) Accelerating the domestication of forest trees in a changing world. *Trends Plant Sci* 17:64–72
- Harwood CE, Nambiar EKS (2014) Productivity of acacia and eucalypt plantations in Southeast Asia. 2. Trends and variations. *Int For Rev* 16:249–260
- Hayden K, Nettel A, Dodd RS, Garbelotto M (2011) Will all the trees fall? Variable resistance to an introduced forest disease in a highly susceptible host. *For Ecol Manag* 261:1781–1791
- Hayden K, Hardy G, Garbelotto M (2013) Oomycete diseases. In: Gonthier P, Nicolotti G (eds) *Infectious forest diseases*. CABI, Boston, pp 519–546
- Hendry AP, Kinnison MT, Heino M, Day T, Smith TB, Fitt G, Bergstrom CT, Oakeshott J, Jørgensen PS, Zalucki MP, Gilchrist G, Southerton S, Sih A, Strauss S, Denison RF, Carroll SP (2011) Evolutionary principles and their practical application. *Evol Appl* 4:159–183
- Hepting GH (1963) Climate and forest diseases. *Annu Rev Phytopathol* 1:31–50
- Horton MW, Bodenhausen N, Beilsmith K, Meng D, Muegge BD, Subramanian S, Vetter MM, Vilhjálmsson BJ, Nordborg M, Gordon JI, Bergelson J (2014) Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nat Commun* 10: 5320
- Hüberli D, Hardy GESJ, White D, Williams N, Burgess TI (2013) Fishing for *Phytophthora* from Western Australia's waterways: a distribution and diversity survey. *Australas Plant Pathol* 42:251–260
- Hudson PJ, Dobson AP, Lafferty KD (2006) Is a healthy ecosystem one that is rich in parasites? *Trends Ecol Evol* 21:381–385
- Hufbauer RA, Roderick GK (2005) Microevolution in biological control: mechanisms, patterns, and processes. *Biol Control* 35:227–239
- Husson C, Scala B, Caël O, Frey P, Feau N, Ioos R, Marçais B (2011) *Chalara fraxinea* is an invasive pathogen in France. *Eur J Plant Pathol* 130:311–324
- Husson C, Aguayo J, Revellin C, Frey P, Ioos R, Marçais B (2015) Evidence for homoploid speciation in *Phytophthora alni* supports taxonomic reclassification in this species complex. *Fungal Genet Biol* 77:12–21
- Hyde KD (2001) Where are the missing fungi? Does Hong Kong have any answers? *Mycol Res* 105:1514–1518
- Ioos R, Andrieux A, Marçais B, Frey P (2006) Genetic characterization of the natural hybrid species *Phytophthora alni* as inferred from nuclear and mitochondrial DNA analyses. *Fungal Genet Biol* 43:511–529
- Jarosz AM, Davelos AL (1995) Effects of disease in wild plant-populations and the evolution of pathogen aggressiveness. *New Phytol* 129:371–387
- Jeger MJ, Pautasso M (2008) Plant disease and global change—the importance of long-term data sets. *New Phytol* 177:8–11
- Johnson R (1984) A critical analysis of durable resistance. *Annu Rev Phytopathol* 22:309–330
- Jones JDG, Dangl JL (2006) The plant immune system. *Nature* 444:323–329
- Jousimo J, Tack AJM, Ovaskainen O, Mononen T, Susi H, Tollenaere C, Laine A-L (2014) Ecological and evolutionary effects of fragmentation on infectious disease dynamics. *Science* 344:1289–1293
- Jumpponen A, Jones KL (2009) Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytol* 184:438–448
- Jumpponen A, Jones KL (2010) Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. *New Phytol* 186:496–513
- Kamgan NG, Jacobs K, de Beer ZW, Wingfield MJ, Roux J (2008) *Ceratocystis* and *Ophiostoma* species, including three new taxa, associated with wounds on native South African trees. *Fungal Divers* 29:37–59
- Keesing F, Holt RD, Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecol Lett* 9:485–498
- Keller SR, Taylor DR (2008) History, chance, and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecol Lett* 11:852–866
- Kembel SW, Mueller RC (2014) Plant traits and taxonomy drive host associations in tropical phyllosphere fungal communities. *Botany* 92:303–311
- Kembel SW, O'Connor TK, Arnold HK, Hubbell SP, Wright SJ, Green JL (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proc Natl Acad Sci* 111:13715–13720
- Kemen E (2014) Microbe-microbe interactions determine oomycete and fungal host colonization. *Curr Opin Plant Biol* 20:75–81
- Kempel A, Schadler M, Chrobock T, Fischer M, van Kleunen M (2011) Tradeoffs associated with constitutive and induced plant resistance against herbivory. *Proc Natl Acad Sci U S A* 108:5685–5689
- Kim MS, Klopfenstein NB, Ota Y, Lee SK, Woo KS, Kaneko S (2010) White pine blister rust in Korea, Japan and other Asian regions: comparisons and implications for North America. *For Pathol* 40: 382–401
- King KC, Lively CM (2012) Does genetic diversity limit disease spread in natural host populations? *Heredity* 109:199–203
- Kinkel LL, Bakker MG, Schlatter DC (2011) A coevolutionary framework for managing disease-suppressive soils. *Annu Rev Phytopathol* 49:47–67
- Kinloch B, Snieszko R, Dupper G (2004) Virulence gene distribution and dynamics of the white pine blister rust pathogen in western North America. *Phytopathology* 94:751–758
- Kinloch B, Davis D, Burton D (2008) Resistance and virulence interactions between two white pine species and blister rust in a 30-year field trial. *Tree Genet Genomes* 4:65–74
- Kiss L (2003) A review of fungal antagonists of powdery mildews and their potential as biocontrol agents. *Pest Manag Sci* 59:475–483
- Kliebenstein DJ (2014) Orchestration of plant defense systems: genes to populations Daniel J. *Trends Plant Sci* 19:250–255
- Knief C, Ramette A, Frances L, Alonso-Blanco C, Vorholt JA (2010) Site and plant species are important determinants of the Methylobacterium community composition in the plant phyllosphere. *ISME J* 4:719–728
- Knief C, Delmotte N, Chaffron S, Stark M, Innerebner G, Wassmann R, von Mering C, Vorholt JA (2012) Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J* 6: 1378–1390
- Knops JMH, Tilman D, Haddad NM, Naem S, Mitchell CE, Haarstad J, Ritchie ME, Howe KM, Reich PB, Siemann E, Groth J (1999) Effects of plant species richness on invasion dynamics, disease outbreaks, insect abundances and diversity. *Ecol Lett* 2:286–293
- Koch H, Frickel J, Valiadi M, Becks L (2014) Why rapid, adaptive evolution matters for community dynamics. *Front Ecol Evol* 2:17
- Kolbe JJ, Glor RE, Rodríguez Schettino L, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* 431:177–181
- Koskella B (2013) Phage-mediated selection on microbiota of a long-lived host. *Curr Biol* 23:1256–1260
- Koskella B (2014) Bacteria-phage interactions across time and space: merging local adaptation and time-shift experiments to understand phage evolution. *Am Nat* 184:S9–S21
- Koskella B, Lin DM, Buckling A, Thompson JN (2012) The costs of evolving resistance in heterogeneous parasite environments. *Proc R Soc B* 279:1896–1903
- Kroon LPNM, Brouwer H, de Cock AWAM, Govers F (2011) The genus *Phytophthora* Anno 2012. *Phytopathology* 102:348–364
- Kupferschmidt K (2012) Attack of the clones. *Science* 337:636–638
- Kutschera U, Hossfeld U (2012) Physiological phytopathology: origin and evolution of a scientific discipline. *J Appl Bot Food Qual* 85: 1–5

- Kvas M, Marasas WFO, Wingfield BD, Wingfield MJ, Steenkamp ET (2009) Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. *Fungal Divers* 34:1–21
- Lafferty KD, Dobson AP, Kuris AM (2006) Parasites dominate food web links. *Proc Natl Acad Sci U S A* 103:11211–11216
- Lambrechts L, Fellous S, Koella JC (2006) Coevolutionary interactions between host and parasite genotypes. *Trends Parasitol* 22:12–16
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc Natl Acad Sci U S A* 104:3883–3888
- Le Gac M, Giraud T (2008) Existence of a pattern of reproductive character displacement in Homobasidiomycota but not in Ascomycota. *J Evol Biol* 21:761–772
- Lebeis SL (2014) The potential for give and take in plant–microbiome relationships *Front. Plant Sci* 5:287
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends Ecol Evol* 17:386–391
- Lefevre F, Boivin T, Bontemps A, Courbet F, Davi H, Durand-Gillmann M, Fady B, Gauzere J, Gidoin C, Karam MJ, Lalagüe H, Oddou-Muratorio S, Pichot C (2014) Considering evolutionary processes in adaptive forestry. *Ann For Sci* 71:723–739
- Lieberei R (2007) South American leaf blight of the rubber tree (*Hevea spp.*): New steps in plant domestication using physiological features and molecular markers. *Ann Bot* 100:1125–1142
- Liebold AM, Wingfield MJ (2014) Globalization and its implications to forest health. In: Nikolakis W, Innes J (eds) *Forests and globalization: challenges and opportunities for sustainable development*. Routledge, pp 36–47
- Liebold AM, Brockerhoff EG, Garrett LJ, Parke JL, Britton KO (2012) Live plant imports: the major pathway for forest insect and pathogen invasions of the U.S. *Front Ecol Environ* 10:135–143
- Lind EM, Borer E, Seabloom E, Adler P, Bakker JD, Blumenthal DM, Crawley M, Davies K, Firn J, Gruner DS, Harpole WS, Hautier Y, Hillebrand H, Knops J, Melbourne B, Mortensen B, Risch AC, Schuetz M, Stevens C, Wragg PD (2013) Life-history constraints in grassland plant species: a growth-defence trade-off is the norm. *Ecol Lett* 16:513–521
- Linzer R, Rizzo D, Cacciola S, Garbelotto M (2009) AFLPs detect low genetic diversity for *Phytophthora nemorosa* and *P. pseudosyringae* in the US and Europe. *Mycol Res* 113:298–307
- Little TJ, Shuker DM, Colegrave N, Day T, Graham AL (2010) The coevolution of virulence: tolerance in perspective. *PLoS Pathog* 6, e1001006. doi:10.1371/journal.ppat.1001006
- Lively CM, de Roode JC, Duffy MA, Graham AL, Koskella B (2014) Interesting open questions in disease ecology and evolution. *Am Nat* 184:S1–S8
- Loehle C, Namkoong G (1987) Constraints on tree breeding: growth tradeoffs, growth strategies, and defensive investments. *For Sci* 33: 1089–1097
- Loo JA (2009) Ecological impacts of non-indigenous invasive fungi as forest pathogens. *Biol Invasions* 11:81–96
- Manion PD (1981) *Tree disease concepts*. Prentice-Hall, Englewood Cliffs, 399 pp
- Manion PD (2003) Evolution of concepts in forest pathology. *Phytopathology* 93:1052–1055
- Martin TA, Johnsen KH, White TL (2001) Ideotype Development in Southern Pines: Rationale and Strategies for Overcoming Scale-Related Obstacles. *For Sci* 47: 21–28
- Martinez-Alvarez P, Vainio EJ, Botella L, Hantula J, Diez JJ (2014) Three mitovirus strains infecting a single isolate of *Fusarium circinatum* are the first putative members of the family *Narnaviridae* detected in a fungus of the genus *Fusarium*. *Arch Virol* 159:2153–2155
- Matos A, Kerkhof L, Garland JL (2005) Effects of microbial community diversity on the survival of *Pseudomonas aeruginosa* in the wheat rhizosphere. *Microb Ecol* 49:257–264
- May RM, Anderson RM (1983) Epidemiology and genetics in the coevolution of parasites and hosts. *Proc R Soc B* 219:281–313
- McCracken AR, Walsh LRE, Moore JP, Lynch M, Cowan P, Dawson MD, Watson S (2011) Yield of willow (*Salix* spp.) grown in short rotation coppice mixtures in long-term trials. *Ann Appl Biol* 159: 229–243
- McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol* 40: 349–379
- McKown AD, Guy RD, Quamme L, Klápště J, La Mantia J, Constabel CP, El-Kassaby YA, Hamelin RC, Zifkin M, Azam MS (2014) Association genetics, geography and ecophysiology link stomatal patterning in *Populus trichocarpa* with carbon gain and disease resistance trade-offs. *Mol Ecol* 23: 5771–90
- Merrill W, Shigo AL (1979) An expanded concept of tree decay. *Phytopathology* 69:1158–1160
- Michelmore RW, Christopoulou M, Caldwell KS (2013) Impacts of resistance gene genetics, function, and evolution on a durable future. *Annu Rev Phytopathol* 51:291–319
- Milgroom MG, Cortesi P (2004) Biological control of chestnut blight with hypovirulence: a critical analysis. *Annu Rev Phytopathol* 42: 311–338
- Millennium Ecosystem Assessment (2005) *Ecosystems and human well-being: synthesis*. Island Press, Washington, DC
- Miller MR, White A, Boots M (2006) The evolution of parasites in response to tolerance in their hosts: the good, the bad, and apparent commensalism. *Evol Int J Org Evol* 60:945–956
- Milus EA, Kristensen K, Hovmoller MS (2009) Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f. sp. *tritici* causing stripe rust of wheat. *Phytopathology* 99:89–94
- Mitchell CE, Power AG (2003) Release of invasive plants from fungal and viral pathogens. *Nature* 421:625–627
- Mohan V, Manokaran P (2013) Assessment of disease problems in different clonal plantations of *Eucalyptus* spp. in South India. *J Acad Ind Res* 1:514–524
- Morozov A, Robin C, Franc A (2007) A simple model for the dynamics of a host–pathogen–hyperparasite interaction. *J Theor Biol* 249: 246–253
- Mosquera J, Adler FR (1998) Evolution of virulence: a unified framework for coinfection and superinfection. *J Theor Biol* 195:293–313
- Mougou A, Dutech C, Desprez-Loustau M-L (2008) New insights into the identity and origin of the causal agent of oak powdery mildew in Europe. *For Pathol* 38:275–287
- Mougou-Hamdane A, Giresse X, Dutech C, Desprez-Loustau ML (2010) Spatial distribution of lineages of oak powdery mildew fungi in France, using quick molecular detection methods. *Ann For Sci* 67: 212
- Mukherjee PK, Horwitz BA, Herrera-Estrella A, Schmoll M, Kenerley CM (2013) *Trichoderma* research in the genome era. *Annu Rev Phytopathol* 51:105–129
- Neale D, Ingvarsson P (2008) Population, quantitative and comparative genomics of adaptation in forest trees. *Curr Opin Plant Biol* 11:149–155
- Newcombe G, Dugan FM (2010) Fungal pathogens of plants in the Homogocene. In: Gherbawy Y, Voigt K (eds) *Molecular identification of fungi*. Springer-Verlag, Berlin, pp 3–35
- Newcombe G, Stirling B, McDonald S, Bradshaw HD Jr (2000) *Melampsora × columbiana*, a natural hybrid of *M. medusae* and *M. occidentalis*. *Mycol Res* 104:261–274
- Newton AC, Gravouil C, Fountaine JM (2010) Managing the ecology of foliar pathogens: ecological tolerance in crops. *Ann Appl Biol* 157: 343–359
- Ney B, Bancal MO, Bancal P, Bingham IJ, Foulkes J, Gouache D, Paveley N, Smith J (2013) Crop architecture and crop tolerance to fungal diseases and insect herbivory. Mechanisms to limit crop losses. *Eur J Plant Pathol* 135:561–580

- Oh E, Gryzenhout M, Wingfield BD, Wingfield MJ, Burgess TI (2013) Surveys of soil and water reveal a goldmine of *Phytophthora* diversity in South African natural ecosystems. *IMA Fungus* 4:123–131
- Oliva J, Camarero JJ, Stenlid J (2012) Understanding the role of sapwood loss and reaction zone formation on radial growth of Norway spruce (*Picea abies*) trees decayed by *Heterobasidion annosum* s.l. *For Ecol Manag* 274:201–209
- Oliva J, Stenlid J, Martinez-Vilalta J (2014) The effect of fungal pathogens on the water and carbon economy of trees: implications for drought-induced mortality. *New Phytol* 203:1028–1035
- O'Malley M (2008) 'Everything is everywhere, but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. *Stud Hist Phil Biol Biomed Sci* 39:314–325
- Ostry ME, Laflamme G (2009) Fungi and diseases—natural components of healthy forests. *Botany* 87:22–25
- Palumbi SR (2001) Humans as the world's greatest evolutionary force. *Science* 293:1786–1790
- Papaix J, Burdon JJ, Zhan J, Thrall PH (2015) Crop pathogen emergence and evolution in agro-ecological landscapes. *Evol Appl*. doi:10.1111/eva.12251
- Paquette A, Messier C (2009) The role of plantations in managing the world's forests in the Anthropocene. *Front Ecol Environ* 8:27–34
- Parke JL, Grünwald NJ (2012) A systems approach for management of pests and pathogens of nursery crops. *Plant Dis* 96:1236–1244
- Parke JL, Knaus BJ, Fieland VJ, Lewis C, Grünwald NJ (2014) *Phytophthora* community structure analyses in Oregon nurseries inform systems approaches to disease management. *Phytopathology* 104:1052–1062
- Parker IM, Gilbert GS (2004) The evolutionary ecology of novel plant–pathogen interactions. *Annu Rev Ecol Evol Systematics* 35:675–700
- Parker JD, Burkepille DE, Hay ME (2006) Opposing effects of native and exotic herbivores on plant invasions. *Science* 311:1459–1461
- Partida-Martínez LP, Heil M (2011) The microbe-free plant: fact or artifact? *Front Plant Sci* 2:100
- Pautasso M, Holdenrieder O, Stenlid J (2005) Susceptibility to fungal pathogens of forests differing in tree diversity. In: Scherer-Lorenzen M, Körner C, Schulze E-D (Eds.) *Forest diversity and function: temperate and boreal systems*. Springer, pp 263–289
- Pautasso M, Schlegel M, Holdenrieder O (2014) Forest health in a changing world. *Microb Ecol*. doi:10.1007/s00248-014-0545-8
- Pearson MN, Beever RE, Boine B, Arthur K (2009) Mycoviruses of filamentous fungi and their relevance to plant pathology. *Mol Plant Pathol* 10:115–128
- Peay K, Baraloto C, Fine PVA (2013) Strong coupling of plant and fungal structure across Amazonian rainforests. *ISME J* 7:1851–1861
- Peiffer J, Spor A, Jin Z, Koren O, Tringe SG, Dangl JL, Buckler ES, Ley RE (2013) Diversity and heritability of the Maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci USA* 5: 570–573
- Peñuelas J, Rico L, Ogaya R, Jump AS, Terradas J (2012) Summer season and long-term drought increase the richness of bacteria and fungi in the foliar phyllosphere of *Quercus ilex* in a mixed Mediterranean forest. *Plant Biol (Stuttg)* 14:565–575
- Pérez G, Slippers B, Wingfield MJ, Wingfield BD, Carnegie AJ, Burgess TI (2012) Cryptic species, native populations and biological invasions by a eucalypt forest pathogen. *Mol Ecol* 21:4452–4471
- Peterson PD, Griffith CS (1999) Hermann von Schrenk and the beginning of forest pathology in the US. *Forest History Today*, 29–34
- Pinon J, Frey P (2005) Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. In: Pei MH, McCracken AR (eds) *Rust diseases of willow and poplar*. CAB International, Wallingford, pp 139–154
- Power AG, Mitchell CE (2004) Pathogen spillover in disease epidemics. *Am Nat* 164:S79–S89
- Prospero S, Conedera M, Heiniger U, Rigling D (2006) Saprophytic activity and sporulation of *Cryphonectria parasitica* on dead chestnut wood in forests with naturally established hypovirulence. *Phytopathology* 96:1337–1344
- Pyšek P, Jarošík V, Hulme PE, Kühn I, Wilda J, Arianoutsou M, Bacher S, Chiron F, Didžiulis D, Essl F, Genovesi P, Gherardi F, Hejda M, Kark S, Lambdon PW, Desprez-Loustau M-L, Nentwig W, Pergl J, Poblajša K, Rabitsch W, Roques A, Roy DB, Shirley S, Solarz W, Vilà M, Winter M (2010) Disentangling the role of environmental and human pressures on biological invasions. *Proc Natl Acad Sci U S A* 107:12157–12162
- Queloz V, Grünig CR, Berndt R, Kowalski T, Sieber TN, Holdenrieder O (2011) Cryptic speciation in *Hymenoscyphus albidus*. *For Pathol* 41: 133–142
- Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N (2010) The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ Microbiol* 12:2885–2893
- Refregier G, Le Gac M, Jabbour F, Widmer A, Shykoff JA, Yockteng R, Giraud T (2008) Co-phylogeny of the anther smut fungi and their Caryophyllaceae hosts: prevalence of host shifts and importance of delimiting parasite species for inferring co-speciation. *BMC Evol Biol* 8:10
- Restif O, Koella JC (2004) Concurrent evolution of resistance and tolerance to pathogens. *Am Nat* 164:E90–E102
- REX Consortium (2013) Heterogeneity of selection and the evolution of resistance. *Trends Ecol Evol* 28:110–118
- Rico L, Ogaya R, Terradas J, Peñuelas J (2014) Community structures of N₂-fixing bacteria associated with the phyllosphere of a Holm oak forest and their response to drought. *Plant Biol* 16:586–593
- Ridout N, Newcombe G (2015) The frequency of modification of *Dothistroma* pine needle blight severity by fungi within the native range. *For Ecol Manag* 337:153–160
- Rius M, Darling JA (2014) How important is intraspecific genetic admixture to the success of colonising populations? *Trends Ecol Evol* 29: 233–242
- Robin C, Lanz S, Soutrenon A, Rigling D (2010) Dominance of natural over released biological control agents of the chestnut blight fungus *Cryphonectria parasitica* in south-eastern France is associated with fitness-related traits. *Biol Control* 53:55–61
- Roderick GK, Navajas M (2003) Genes in new environments: genetics and evolution in biological control. *Nat Genet* 4:889–899
- Roderick GK, Hufbauer R, Navajas M (2012) Evolution and biological control. *Evol Appl* 5:419–423
- Roossinck MJ, Saha P, Wiley GB, Quan J, White JD, Lai H, Chavarria F, Shen GA, Roe BA (2010) Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Mol Ecol* 19:81–88
- Roper M, Ellison C, Taylor JW, Glass NL (2011) Nuclear and genome dynamics in multinucleate ascomycete fungi. *Curr Biol* 21:R786–R793
- Roy BA (2001) Patterns of association between crucifers and their flower-mimic pathogens: host-jumps are more common than coevolution or cospeciation. *Evolution* 55:41–53
- Roy BA, Kirchner JW (2000) Evolutionary dynamics of pathogen resistance and tolerance. *Evolution* 54:51–63
- Roy BA, Alexander HM, Davidson J, Campbell FT, Burdon JJ, Snieszko R, Brasier C (2014) Increasing forest loss worldwide from invasive pests requires new trade regulations. *Front Ecol Environ* 12:457–465
- Sakalidis ML, Slippers B, Wingfield BD, Hardy GESJ, Burgess TI (2013) The challenge of understanding the origin, pathways and extent of fungal invasions: global populations of the *Neofusicoccum parvum*–*N. ribis* species complex. *Divers Distrib* 19:873–883
- Santini A, Ghelardini L, De Pace C, Desprez-Loustau ML, Capretti P, Chandelier A, Cech T, Chira D, Diamandis S, Gaitniekis T, Hantula J, Holdenrieder O, Jankovsky L, Jung T, Jurc D, Kirisits T, Kunca A,

- Lygis V, Malecka M, Marçais B, Schmitz S, Schumacher J, Solheim H, Solla A, Szabó I, Tsopelas P, Vannini A, Vettraino AM, Webber J, Woodward S, Stenlid J (2013) Biogeographical patterns and determinants of invasion by forest pathogens in Europe. *New Phytol* 197: 238–250
- Schafer JF (1971) Tolerance to plant disease. *Ann Rev Phytopathol* 9: 235–252
- Schardl CL, Craven KD (2003) Interspecific hybridization in plant-associated fungi and oomycetes: a review. *Mol Ecol* 12:2861–2873
- Schmit J, Mueller G (2007) An estimate of the lower limit of global fungal diversity. *Biodivers Conserv* 16:99–111
- Schoebel CN, Zoller S, Rigling D (2014) Detection and genetic characterisation of a novel mycovirus in *Hymenoscyphus fraxineus*, the causal agent of ash dieback. *Infect Genet Evol* 28:78–86
- Schoettle AW, Sniezko RA (2007) Proactive intervention to sustain high-elevation pine ecosystems threatened by white pine blister rust. *J For Res* 12:327–336
- Schulenburg H, Kurtz J, Moret Y, Siva-Jothy MT (2009) Introduction. *Ecological immunology*. *Phil Trans R Soc B* 364:3–14
- Schulze-Lefert P, Panstruga R (2011) A molecular concept connecting nonhost-resistance and pathogen speciation. *Trends Plant Sci* 16: 117–125
- Schweitzer JA, Bailey JK, Fischer DG, LeRoy CJ, Lonsdorf EV, Whitham TG, Hart SC (2008) Plant–soil–microorganism interactions: heritable relationship between plant genotype and associated soil microorganisms. *Ecology* 89:773–781
- Selakovic S, de Ruiter PC, Heesterbeek H (2014) Infectious disease agents mediate interaction in food webs and ecosystems. *Proc R Soc B* 281:20132709. doi:10.1098/rspb.2013.2709
- Shaw MW, Osborne TM (2011) Geographic distribution of plant pathogens in response to climate change. *Plant Pathol* 60:31–43
- Shaw MW, Bearchell SJ, Fitt BDL, Fraaije BA (2008) Long-term relationships between environment and abundance in wheat of *Phaeosphaeria nodorum* and *Mycosphaerella graminicola*. *New Phytol* 177:229–238
- Simms E (2000) Defining tolerance as a norm of reaction. *Evol Ecol* 14: 563–570
- Simms E, Triplett J (1994) Costs and benefits of plant responses to disease - resistance and tolerance. *Evolution* 48:1973–1985
- Siozios S, Tosi L, Ferrarini A, Ferrari A, Taroni P, Bellin D, Maurhofer M, Gessler C, Delledonne M, Pertot I (2015) Transcriptional reprogramming of the mycoparasitic fungus *Ampelomyces quisqualis* during the powdery mildew host-induced germination. *Phytopathology* 105:199–209
- Springer JC, Baines ALD, Fulbright DW, Chansler MT, Jarosz AM (2013) Hyperparasites influence population structure of the chestnut blight pathogen, *Cryphonectria parasitica*. *Phytopathology* 103: 1280–1286
- Stenlid J, Oliva J, Boberg JB, Hopkins AJM (2011) Emerging diseases in European forest ecosystems and responses in society. *Forests* 2:486–504
- Stukenbrock EH (2013) Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. *New Phytol* 199: 895–907
- Stukenbrock EH, Bataillon T (2012) A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. *PLoS Pathog* 8, e1002893
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annu Rev Phytopathol* 46:75–100
- Sturrock RN, Frankel SJ, Brown AV, Hennon PE, Kliejunas JT, Lewis KJ, Woods AJ (2011) Climate change and forest diseases. *Plant Pathol* 60:133–149
- Su-See L (1999) Forest health in plantation forests in South-East Asia. *Australas Plant Pathol* 28:283–291
- Susi H, Barres B, Pedro F, Vale PF, Laine AL (2014) Co-infection alters population dynamics of infectious disease. *Nat Commun*. doi:10.1038/ncomms6975
- Swett CL, Gordon TR (2012) First report of grass species (Poaceae) as naturally occurring hosts of the pine pathogen *Gibberella circinata*. *Plant Dis* 96:908
- Swett CL, Gordon TR (2015) Endophytic association of the pine pathogen *Fusarium circinatum* with corn (*Zea mays*). *Fungal Ecol* 13: 120–129
- Swinton J, Gilligan CA (1999) Selecting hyperparasites for biocontrol of Dutch elm disease. *Proc R Soc B* 266:437–445
- Szucs M, Eigenbrode SD, Schwarzlander M, Schaffner U (2012) Hybrid vigor in the biological control agent, *Longitarsus jacobaeae*. *Evol Appl* 5:489–497
- Tainter FH, Baker FA (1996) Principles of forest pathology. John Wiley and Sons, Inc, New York, 803 pp
- Takamatsu S, Braun U, Limkaisang S, Kom-un S, Sato Y, Cunnington JH (2007) Phylogeny and taxonomy of the oak powdery mildew *Erysiphe alphitoides* sensu lato. *Mycol Res* 111:809–826
- Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, Fisher MC (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genet Biol* 31:21–32
- Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D (2006) Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. *Philos T Roy Soc B* 361:1947–1963
- Tedersoo L, Bahram M, Pölme S, Yorou NS, Wijesundera R, Villarreal Ruiz L, Vasco-Palacios AM, Thu PQ, Suija A, Smith ME, Sharp C, Saluveer E, Saitta A, Rosas M, Riit T, Ratkowsky D, Pritsch K, Pöldmaa K, Piepenbring M, Phosri C, Peterson M, Parts K, Pärtel K, Otsing E, Nouhra E, Njouonkou AL, Nilsson RH, Morgado LN, Mayor J, May TW, Majuakim L, Lodge DJ, Lee SS, Larsson KH, Kohout P, Hosaka K, Hiiesalu I, Henkel TW, Harend H, Guo LD, Greslebin A, Grelet G, Geml J, Gates G, Dunstan W, Dunk C, Drenkhan R, Dearnaley J, De Kesel A, Dang T, Chen X, Buegger F, Brearley FQ, Bonito G, Anslan S, Abell S, Abarenkov K (2014) Global diversity and geography of soil fungi. *Science* 346:1256688
- Telford A, Cavers S, Ennos RA, Cottrell JE (2015) Can we protect forests by harnessing variation in resistance to pests and pathogens? *Forestry* 88:3–12
- Tittonell P (2014) Ecological intensification—sustainable by nature. *Curr Opin Environ Sustain* 8:53–61
- Tollenaere C, Pernechele B, Mäkinen HS, Parratt SR, Németh MZ, Kovács GM, Kiss L, Tack AJ, Laine AL (2014) A hyperparasite affects the population dynamics of a wild plant pathogen. *Mol Ecol* 23:5877–5887
- Tsui CK-M, Farfan L, Roe AD, Rice AV, Cooke JEK, El-Kassaby YA, Hamelin RC (2014) Population structure of mountain pine beetle symbiont *Leptographium longiclavatum* and the implication on the multipartite beetle–fungi relationships. *PLoS One* 9, e105455
- Turner TR, James EK, Poole PS (2013) The plant microbiome. *Genome Biol* 14:209
- Vacher C, Daudin J-J, Piou D, Desprez-Loustau M-L (2010) Ecological integration of alien species into a tree–fungus network. *Biol Invasions* 12:3249–3259
- Van der Most P, de Jong B, Parmentier HK, Verhulst S (2011) Trade-off between growth and immune function: a meta-analysis of selection experiments. *Funct Ecol* 25:74–80
- van Schie CC, Takken FL (2014) Susceptibility genes: how to be a good host. *Annu Rev Phytopathol* 52:551–581
- Vander Wal E, Garant D, Calmé S, Chapman C, Festa-Bianchet M, Millien V, Rioux-Paquette S, Pelletier F (2014) Applying evolutionary concepts to wildlife disease ecology and management. *Evol Appl* 7:856–868
- Varki A (2012) Nothing in medicine makes sense, except in the light of evolution. *J Mol Med (Berl)* 90:481–494

- Vayssier-Taussat M, Albina E, Citti C, Cosson J-F, Jacques M-A, Lebrun M-H, Le Loir Y, Ogliaastro M, Petit MA, Roumagnac P, Candresse T (2014) Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Front Cell Infect Microbiol* 5:4–29
- Verhoeven KJF, Biere A, Harvey JA, Putten WH (2009) Plant invaders and their novel natural enemies: who is naive? *Ecol Lett* 12:107–117
- Vleeshouwers VG, Rietman H, Krenk P, Champouret N, Young C, Oh SK, Wang M, Bouwmeester K, Vosman B, Visser RG, Jacobsen E, Govers F, Kamoun S, Van der Vossen EA (2008) Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. *PLoS One* 3, e2875
- Vogan PJ, Schoettle AW (2015) Selection for resistance to white pine blister rust affects the abiotic stress tolerances of limber pine. *For Ecol Manage* 344:110–119
- Voth PD, Mairura L, Lockhart BE, May G (2006) Phylogeography of *Ustilago maydis* virus H1 in the USA and Mexico. *J Gen Virol* 87: 3433–3441
- Vuillaume F, Thebaud G, Urbino C, Forfert N, Granier M, Froissart R, Blanc S, Peterschmitt M (2011) Distribution of the phenotypic effects of random homologous recombination between two virus species. *Plos Pathogens* 7, e1002028. doi:10.1371/journal.ppat.1002028
- Wainhouse D (2005) Biological control. In: Wainhouse D (ed) *Ecological methods in forest pest management*. Oxford University Press, Oxford, pp 109–126
- Walters DR, Avrova A, Bingham IJ, Burnett FJ, Fountaine J, Havis ND, Hoad SP, Hughes G, Loosely M, Oxley SJP, Renwick A, Topp CFE, Newton AC (2012) Control of foliar diseases in barley: towards an integrated approach. *Eur J Plant Pathol* 133:33–73
- Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, LeRoy CJ, Lonsdorf E, Allan GJ, DiFazio SP, Potts BM, Fischer DG, Gehring CA, Lindroth RL, Marks J, Hart SC, Wimp GM, Wooley SC (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nat Rev Genet* 7:510–523
- Williamson M, Fitter A (1996) The varying success of invaders. *Ecology* 77:1661–1665
- Witzell J, Martín JA, Blumenstein K (2014) Ecological aspects of endophyte-based biocontrol of forest diseases. In: Verma VC, Gange AC (eds) *Advances in endophytic research*. Springer India, New Delhi, pp 321–333
- Woods A, Coates KD, Hamann A (2005) Is an unprecedented *Dothistroma* needle blight epidemic related to climate change? *Bioscience* 55:761–769
- Xhaard C, Barrès B, Andrieux A, Bousset L, Halkett F, Frey P (2012) Disentangling the genetic origins of a plant pathogen during disease spread using an original molecular epidemiology approach. *Mol Ecol* 21:2383–2398
- Xie J, Jiang D (2014) New Insights into mycoviruses and exploration for the biological control of crop fungal diseases. *Annu Rev Phytopathol* 52:45–68
- Yanchuk A, Allard G (2009) Tree improvement programmes for forest health—can they keep pace with climate changes? *Unasylva* 231: 50–56
- Yang C, Crowley DE, Borneman J, Keen NT (2001) Microbial phyllosphere populations are more complex than previously realized. *Proc Natl Acad Sci U S A* 98:3889–3894
- Zhang DY, Sun GJ, Jiang XH (1999) Donald's ideotype and growth redundancy: a game theoretical analysis. *Field Crops Res* 61:179–187
- Zhao YJ, Hosoya T, Baral HO, Hosoka K, Kakishima M (2012) *Hymenoscyphus pseudoalbidus*, the correct name for *Lambertella albida* reported from Japan. *Mycotaxon* 122:25–41

6.2 Appendix B - Learning Ecological Networks from Next-Generation Sequencing Data

Learning Ecological Networks from Next-Generation Sequencing Data

Corinne Vacher^{*,†,1}, Alireza Tamaddoni-Nezhad[‡],
 Stefaniya Kamenova^{§,¶}, Nathalie Peyrard^{||}, Yann Moalic[#],
 Régis Sabbadin^{||}, Loïc Schwaller^{**,††}, Julien Chiquet^{‡‡}, M. Alex Smith[¶],
 Jessica Vallance^{§§,¶¶}, Virgil Fievet^{*,†}, Boris Jakuschkin^{*,†},
 David A. Bohan^{|||}

*INRA, UMR1202, BIOGECO, Cestas, France

†University of Bordeaux, UMR1202, BIOGECO, Pessac, France

‡Department of Computing, Imperial College London, London, United Kingdom

§INRA, UMR1349 IGEPP, Le Rheu, France

¶Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

||INRA, Toulouse, France

#Laboratoire de Microbiologie des Environnements Extrêmes (LM2E), Institut Universitaire Européen de la Mer (IUEM), Université de Bretagne Occidentale, Plouzané, France

**INRA, UMR518MIA, Paris, France

††AgroParisTech, Paris, France

‡‡LaMME, UMR CNRS 8071, Université d'Evry-Val d'Essonne, Evry, France

§§Université de Bordeaux, ISVV, UMR1065 Santé et Agroécologie du Vignoble (SAVE), Bordeaux Sciences Agro, Villenave d'Ornon, France

¶¶INRA, ISVV, UMR1065 SAVE, Villenave d'Ornon, France

|||UMR1347 Agroécologie, AgroSup/UB/INRA, Pôle Ecologie des Communautés et Durabilité des Systèmes Agricoles, Dijon Cedex, France

¹Corresponding author: e-mail address: corinne.vacher@pierroton.inra.fr

Contents

1. Introduction	2
1.1 Ecological Interactions Are Drivers of Ecosystem Functioning	2
1.2 Ecological Interactions Are Altered by Anthropogenic Activity	4
1.3 Next-Generation Sequencing Can Be Used for Monitoring Ecological Interactions	5
2. Why Learning Ecological Networks from NGS Data?	7
2.1 Limitations of Classical Methods for Resolving Ecological Interactions	7
2.2 Advantages of NGS for Identifying Species and Their Interactions	8
3. Examples of NGS-Based Ecological Networks and Their Applications	10
3.1 Deciphering Pathobiomes Using NGS-Based Microbial Networks for Improving Biological Control	10
3.2 Studying the Hologenome Theory of Evolution Using NGS-Based Microbial Networks	12

3.3	Testing the Niche Partitioning Theory with NGS-Based Trophic Networks	14
3.4	Challenges to Be Addressed to Get Predictive Insights from NGS-Based Networks	17
4.	Theoretical Methods for Deciphering Ecological Networks from NGS Data	18
4.1	The Input Data	18
4.2	Inferring Ecological Interactions Using Statistical Models	21
4.3	Learning Ecological Interactions Using Logic-Based Machine-Learning Algorithms	23
5.	Conclusion	26
	Acknowledgements	27
	Glossary	27
	References	29

Abstract

Species diversity, and the various interactions that occur between species, supports ecosystems functioning and benefit human societies. Monitoring the response of species interactions to human alterations of the environment is thus crucial for preserving ecosystems. Ecological networks are now the standard method for representing and simultaneously analyzing all the interactions between species. However, deciphering such networks requires considerable time and resources to observe and sample the organisms, to identify them at the species level and to characterize their interactions. Next-generation sequencing (NGS) techniques, combined with network learning and modelling, can help alleviate these constraints. They are essential for observing cryptic interactions involving microbial species, as well as short-term interactions such as those between predator and prey. Here, we present three case studies, in which species associations or interactions have been revealed with NGS. We then review several currently available statistical and machine-learning approaches that could be used for reconstructing networks of direct interactions between species, based on the NGS co-occurrence data. Future developments of these methods may allow us to discover and monitor species interactions cost-effectively, under various environmental conditions and within a replicated experimental design framework.



1. INTRODUCTION

1.1 Ecological Interactions Are Drivers of Ecosystem Functioning

For a considerable part of the history of ecology, ecologists have tried to observe and explain the relationships between biodiversity and ecosystem functioning (BEF relationship) and how these change with environmental conditions and human-derived stressors (see [Bohan et al., 2013](#); [Cardinale](#)

et al., 2012; chapter ‘Towards an integration of biodiversity–ecosystem functioning and food–web theory to evaluate relationships between multiple ecosystem services’ by Hines et al.; chapter ‘Linking biodiversity, ecosystem functioning and services, and ecological resilience: towards an integrative framework for improved management’ by Truchy et al.; Raffaelli et al., 2014). There is now a broad consensus that species diversity supports many ecosystem functions (Bohan et al., 2013; Cardinale et al., 2012). Ecological explanations of BEF have largely invoked two effects. First, a ‘sampling effect’ where highly productive species are more likely to be present in a diverse species pool and consequently function increases with diversity. Second, a ‘complementarity effect’ where the link between diversity and function is explained as a difference in species resource requirements (Loreau and Hector, 2001). Given that some species exploit resources more efficiently than others, then across increasingly diverse pools of species, functionality will tend to increase.

Species diversity not only supports ecosystem functions, but also influences the stability of ecosystems (Haddad et al., 2011; Loreau and de Mazancourt, 2013; Tilman et al., 2006). For system-level properties, such as total biomass/productivity and biological control, resilience of ecosystem function can arise with both the sampling and complementarity effects. As interacting species react in different ways to external stress, some species will benefit while others will not (see May, 1973). Across a portfolio of species within a diverse ecosystem, the stability of the entire portfolio will be greater than the fluctuations in the function of each species, due to an effect of averaging (Tilman et al., 1998). In high-diversity systems, stability across the portfolio of species is therefore preserved in the face of variation (insurance hypothesis, see Loreau et al., 2003; Yachi and Loreau, 1999).

These effects of biodiversity on ecosystem functioning are mediated by the ecological interactions between species (Duffy et al., 2007; Thébault and Loreau, 2006). For studying diverse systems of many interacting species, ecological network approaches have become the standard method (Ings et al., 2009; Lewinsohn et al., 2006; Poccock et al., 2012). These consider species as nodes, and the interactions between species (e.g. predator–prey interactions, host–parasite interactions, plant–pollinator interactions) as a series of links. The links may be weighted (quantitative ecological network) or not. Link weight is often defined the observed frequency of the interaction or the effect of the interaction on the performance of the interacting partners (Bascompte et al., 2006; Berlow et al., 2004; Laliberté and

Tylianakis, 2010). Building upon the BEF knowledge already acquired, the aim of network approaches is to understand how the network structure leads to the ‘emergence’ of ecosystem functions (see chapter ‘Towards an integration of biodiversity–ecosystem functioning and food–web theory to evaluate relationships between multiple ecosystem services’ by Hines et al.; Thompson et al., 2012). It has long been known that the provision of a specific ecosystem function may be maximized through management of the abundance of functionally important species (Gaston, 2010), such as pollinator honeybees (Calderone, 2012; Hagen et al., 2012). But what has become clear from network approaches is that the resilience of ecosystems and the stability of their functions rely directly on species diversity and the interactions between all species within the network (Naeem et al., 2009).

One interesting development in recent years has been a growing appreciation that ecological interactions and evolutionary dynamics can interact (Hairston et al., 2005). The structure of ecological networks and evolution can feedback, one upon the other. For instance, in plant–pollinator and plant–seed disperser mutualistic networks, network structure depends on co–evolution with and between partner species (Bascompte et al., 2006; Nuismer et al., 2013). In trophic networks (food webs), link connectance and the number of trophic levels can be affected through the evolution of body–size relationships (Loeuille and Loreau, 2006) and adaptive foraging (Beckerman et al., 2006). The evolutionary history can in turn affect network modularity and nestedness (Robinson et al., 2015; Vacher et al., 2008). These changes in ecological network structure, caused by the evolution of life–history traits, affect the functioning of ecosystems (Abrams and Matsuda, 2005; Loeuille et al., 2002), as well as their long–term stability and resilience (Kondoh, 2003; Loeuille, 2010a,b; Thébault and Fontaine, 2010).

1.2 Ecological Interactions Are Altered by Anthropogenic Activity

Anthropogenically accelerated biodiversity loss and larger global–scale environmental impacts have provoked much public concern and a drive to improve our management of the planet (Cardinale et al., 2012; Raffaelli and White, 2013; Rockström et al., 2009). Starting at the Rio ‘Earth Summit’ in 1992, this need to protect and monitor ecosystems and their functions have more recently coalesced around the concept of ecosystem services, as the ecosystem functions that directly benefits humanity (Millennium Ecosystem Assessment, 2005; see also chapter ‘Detrital

dynamics and cascading effects on supporting ecosystem services' by Mulder et al.). There is, however, little agreement on, or clear understanding of, the best practices for preventing losses in species diversity, and of how networks of interacting species may be restored (but see [Pocock et al., 2012](#)). What we do have are expectations regarding how networks may change with environmental stressors, such as agricultural intensification ([Albrecht et al., 2007](#); [Laliberté and Tylianakis, 2010](#); [Tylianakis et al., 2007](#)) or biological invasions ([Aizen et al., 2008](#); [Albrecht et al., 2014](#); [Heleno et al., 2009](#); [Lopezaraiza-Mikel et al., 2007](#); [Vacher et al., 2010](#)). However, predicting *a priori* exactly how a particular ecological network metric (e.g. species richness, number of links, interaction strength, connectance, modularity, nestedness) will change in response to a particular stressor remains difficult ([Heleno et al., 2012](#)). Defining ecologically relevant network metrics ([Blüthgen et al., 2008](#); [Fortuna et al., 2010](#); [Joppa et al., 2010](#); [Leger et al., 2015](#)) and predicting how they affect network structural stability and the persistence of the species within it ([Bascompte et al., 2003](#); [Montoya et al., 2006](#); [Rohr et al., 2014](#); [Saavedra et al., 2011](#); [Thébault and Fontaine, 2010](#)) is also very challenging.

Simply reducing or removing anthropogenic stressors may not be enough to restore the structure of ecological networks and the functioning of ecosystems. For example, in a recent grassland experiment, plant diversity in plots that received high rates of nitrogen for 10 years had not recovered to control levels 20 years after the nitrogen inputs had stopped ([Isbell et al., 2013](#)). This suggests that 'turning back the clock' to a more benign set of management practices, even where this is feasible, might not achieve restoration goals. There is, indeed, increasing evidence that network structure and dynamics modulates the trajectory and rate of change of the response to stressors, with time lags and the presence of multiple alternative stable states, due to ecological inertia in the food web. Alternative stable states have been suggested as the reasons for the slow or non-existent biological recovery of commercial marine fisheries following reductions in fishing effort, and of freshwaters following acidification ([Layer et al., 2010, 2011](#)) and eutrophication ([Scheffer et al., 2001](#)).

1.3 Next-Generation Sequencing Can Be Used for Monitoring Ecological Interactions

While ecological networks play a central role in the development of basic and applied ecological science, most empirical networks remain at best

incomplete due to persistent methodological shortcomings that hinder our capacity to resolve species (i.e. network nodes) and their interactions (i.e. network links) (Ings et al., 2009). Resolving ecological networks require considerable (and often prohibitive) investments of time and resources to observe and sample the organisms present within an ecosystem, identify them to the species level and then characterize the possible links that exist between those species. Quantifying the importance of these links, in terms of frequency and effect on the performance of the interacting partners, is even more difficult. This means that relatively few well-resolved ecological networks have been constructed, there is little replication amongst those networks (see Pocock et al., 2012), and ultimately our understanding of network-based BEF and ecosystem change is limited.

Here, we speculate that reconstructing ecological networks directly from environmental DNA (eDNA), by combining next-generation sequencing (NGS; high-throughput sequencing) (Di Bella et al., 2013) and theoretical approaches, such as statistical modelling (Faust and Raes, 2012) and machine learning (Bohan et al., 2011), will alleviate these methodological and financial constraints. Our arguments are the following: (1) NGS platforms generate several millions of DNA sequences for a few hundred dollars (Liu et al., 2012; Quail et al., 2012); (2) they permit the characterization of DNA diversity in complex environmental samples (e.g. soil, water, plant tissues, faeces, pellet, gut content, etc.) containing hundreds of microbial species and the imprint of many macro-organisms; (3) many of these organisms can be identified at the species level by using GenBank or reference taxonomic databases such as Greengenes, BOLD, SILVA and UNITE (Abarenkov et al., 2010; DeSantis et al., 2006; Kõljalg et al., 2005; Quast et al., 2013; Ratnasingham and Hebert, 2007) and (4) methods are being developed to predict species interactions from their abundance patterns and additional information such as their functional traits or the features of the environmental samples (Bohan et al., 2011; Deng et al., 2012; Faust and Raes, 2012; Kurtz et al., 2015). With these advancements, the potential of NGS techniques for resolving complex ecological networks is enormous. Below we review the types of ecological interactions for which NGS data are the most relevant, present some examples of NGS-based ecological networks and give an insight into the theoretical approaches that may be applied to NGS data to highlight ecological interactions.



2. WHY LEARNING ECOLOGICAL NETWORKS FROM NGS DATA?

2.1 Limitations of Classical Methods for Resolving Ecological Interactions

Network ecology primarily emerged based on the observation of interactions between macro-organisms (e.g. plant–pollinator, plant–herbivore, plant–seed disperser, anemone–fish; see IWDB, <https://www.nceas.ucsb.edu/interactionweb/resources.html>). These observations require considerable time and resources and have limitations. Smaller species, such as microbes and parasites, have often been overlooked (Lafferty et al., 2008) despite their importance for ecosystem functioning (Ducklow, 2008; Gilbert and Neufeld, 2014; Hudson et al., 2006). As a consequence, integration of macro-organisms and micro-organisms resolved at the species level into the same networks is far from the norm (but see Vacher et al., 2008, 2010). Short-term interactions such as those between predator and prey are also difficult to observe. This lack of completeness and integration of ecological networks critically limits our understanding of BEF and should be resolved (Fontaine et al., 2011).

Micro-organisms have not been fully integrated in network ecology largely due to the enormous difficulty associated with identifying the interacting microbes. Microbial interactions have long been studied using laboratory culture-dependent methods. However, testing the hundreds of combinations of culture conditions necessary to find the conditions required to grow successfully a single microbial species is often prohibitive (Lok, 2015). As a consequence, an estimated 85–99% of bacteria and archaea cannot yet be grown in the lab, drastically limiting our knowledge of microbial life. Culturing is not only a (inadvertently) selective environment, but also a labour intensive and tedious task that cannot be extended to the whole microbial community in a given environment. Co-culture experiments are traditionally used to detect antagonistic or mutualistic interactions between microbial taxa but the extrapolation of results to natural conditions is risky (Sher et al., 2011). High-throughput cultivation platforms, such as microfluidic chips and iChips, are currently revolutionizing this field of research. They are based on complex cultivation media, closer to those of the natural environment and may support multiple microbial species (Lok, 2015). Culturing and co-culturing microbes

nevertheless remain a complicated task (Haruta et al., 2009). Many scientists have, therefore, chosen to bypass it entirely and move to sequencing directly microbial DNA.

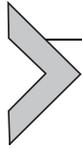
Many scientists studying trophic networks have also chosen to use DNA sequencing for characterizing predator–prey or host–parasitoid interactions. The visual and microscopic examination of guts, faeces, or pellets (Hengeveld, 1980; Pisanu et al., 2011) and the rearing of parasitoids from hosts (Eveleigh et al., 2007; Müller et al., 1999) are indeed time consuming, labour intensive and highly dependent on the observer’s experience. They tend to overlook cryptic species and are very difficult to compare between systems or researchers (reviewed by Symondson, 2002). Direct methods have expanded recently to include the analysis of ratios of naturally occurring stable isotopes (mostly carbon and nitrogen), permitting the integration of trophic information over an extended time-scale (several days to several months) (Boecklen et al., 2011). Nevertheless, the technique can only be applied to a small number of prey species (Moore and Semmens, 2008) and often does not permit obtaining accurate trophic data at individual level (Vanderklift and Ponsard, 2003). Other biomarkers, such as fatty acids (reviewed by Traugott et al., 2013) have similar constraints.

2.2 Advantages of NGS for Identifying Species and Their Interactions

The emergence of NGS techniques and associated bioinformatic pipelines has boosted the characterization of species and their interactions over the last 10 years. In particular, it has boosted the characterization of microbial communities (Di Bella et al., 2013; Hibbett et al., 2009) and facilitated the quantification of trophic links (Smith et al., 2011; Staudacher et al., 2015). NGS indeed overcomes the limitations of culture-dependent methods and provides a more thorough description of microbial communities. We are now, in principle, able to detect any microbial organism in nature, even those that cannot be isolated or grown in the lab. The first step is to collect environmental samples (e.g. samples of soil or water, tissues of plant, animal, etc.) and extract total DNA from these samples. The taxonomic description of microbial communities with NGS techniques then relies on the amplification and sequencing of DNA barcodes (Chakraborty et al., 2014). Similarly, the sequencing of faeces, pellet, or gut content gives an insight into the diet of an organism with little or no need for *a priori* information about the

target prey species (e.g. Boyer et al., 2013; Brown et al., 2012; Quéméré et al., 2013; Shehzad et al., 2012). Moreover, the use of tags (i.e. unique identifiers) to recover data from each sample after sequencing (Clarke et al., 2014; Taberlet et al., 2012) facilitates the mass screening of several hundreds of individual samples (e.g. Kartzinel et al., 2015; Mollot et al., 2014).

A DNA barcode, in its simplest definition, is one or small number of short genetic sequences taken from a standardized portion of the genome that are used to identify species (Schlaeppli and Bulgarelli, 2015). Different barcode regions are used for different taxonomic groups. The most popular DNA barcode for identifying fungal species is the internal transcribed spacer (ITS) region of the nuclear ribosomal repeat unit (Schoch et al., 2012), except for Glomeromycota (Öpik et al., 2014; Stockinger et al., 2010). The 16S small ribosomal subunit gene (16S rRNA) is the ‘gold standard’ for characterizing bacterial community composition (Sun et al., 2013), although other barcode regions have been proposed (Chakraborty et al., 2014). The 16S rRNA gene contains conserved regions as well as nine hypervariable regions (V1–V9). One or more of these hypervariable regions are usually sequenced (Gloor et al., 2010; Sun et al., 2013). Universal barcode genes are also available for algae and protozoa, but, to our knowledge, not for viruses (Chakraborty et al., 2014). The standardized barcode region for eukaryotic animals is part of the cytochrome *c* oxidase 1 (CO1) mitochondrial gene and a growing library of identified specimens exists (see BOLD, <http://www.boldsystems.org/>; Ratnasingham and Hebert, 2007). In the case of trophic ecology, the best barcode depends upon the model organism, its likely preys, and the question being addressed (Pompanon et al., 2012). Generally, as dietary samples are complex mixtures of highly degraded DNA, the best barcode should target very short DNA fragments with the same amplification efficiency across very distantly related taxa. These constraints sometimes preclude species-level taxonomic assignment of preys as the probability to encounter variable sites decreases with the sequence length. Therefore, classical DNA-based methods such diagnostic PCR followed by Sanger sequencing remains very useful for the quantitative assessment of predator–prey (Clare et al., 2009; Rougerie et al., 2011) and host–parasitoid interactions (Condon et al., 2014; Kaartinen et al., 2010; Smith et al., 2008). These molecular-based techniques have been used successfully to reveal previously unseen trophic interactions (Derocles et al., 2015; Wirta et al., 2014).



3. EXAMPLES OF NGS-BASED ECOLOGICAL NETWORKS AND THEIR APPLICATIONS

3.1 Deciphering Pathobiomes Using NGS-Based Microbial Networks for Improving Biological Control

Population growth and changes in dietary habits are likely to double human demand for food in the coming decade. However, in addition to these issues, there are inherent inefficiencies in modern agriculture where plant pathogens can claim 10–16% of the global food harvest. Reducing this percentage has become, therefore, a priority for achieving food security (Chakraborty and Newton, 2011). Biodiverse ecosystems can support natural antagonists of pathogens and they have been argued to limit or prevent disease development (Berendsen et al., 2012; Penton et al., 2014). The notion that the pathogens live and interact with other organisms in their environment has led to the development of the pathobiome concept (i.e. the pathogen species integrated within its biotic environment; Vayssier-Taussat et al., 2014). Elucidating the components of the pathobiome is one of the prerequisites for understanding the persistence, transmission and evolution of pathogen species, and for improving methods of biological control with naturally occurring antagonists.

Microbial networks give insight into the relationships between a pathogen species and the other micro-organisms interacting with the host species. Their study may enable the identification of potential pathogen antagonists and may reveal network topological properties driving the stability of the residential microbial community, including its invasibility by pathogens (Desprez-Loustau et al., 2015; Kemen, 2014). Microbial networks typically represent the patterns of co-existence between microbial taxa, irrespective of the underlying mechanism. Nodes represent microbial taxa and links indicate spatial or temporal associations between taxa. Such networks are referred to as microbial co-occurrence networks (Aires et al., 2015; Barberán et al., 2012; Kara et al., 2013; Navarrete et al., 2015), microbial association networks (Chow et al., 2014; Faust and Raes, 2012; Fuhrman, 2009), microbial correlation networks (Duran-Pinedo et al., 2011; Friedman and Alm, 2012), or networks of co-existing microbes (Chaffron et al., 2010). *Direct ecological interactions* may account for the observed co-existence patterns. As with macroscopic organisms, microbes interact with each other directly, through pairwise ecological interactions such as competition, predation, parasitism, mutualism, commensalism, or

amenalism (Faust and Raes, 2012). These interactions can involve micro-organisms belonging to the same kingdom or different kingdoms (Frey-Klett et al., 2011). Competition, which is considered to be the dominant type of microbe–microbe interaction (Foster and Bell, 2012; Hibbing et al., 2010), is expected to produce co-exclusions between microbial taxa. Amensalism is also expected to produce co-exclusions while mutualism and commensalism are expected to yield co-associations between microbial taxa. The outcome of parasitism or predation, in terms of spatial or temporal association, is less straightforward to predict, with the consumer relying on, but potentially depleting, its prey. Associations between microbial taxa may also be explained by two other mechanisms: because they both depend on the presence of a third microbial taxa (*indirect ecological interaction*) or because they are adapted to similar environmental conditions (*shared environmental preference*) (Ovaskainen et al., 2010).

Here, we reconstructed the microbial association network (Fig. 1) of the leaves of an oak tree (*Quercus robur* L.) susceptible to the fungal pathogen *Erysiphe alphitoides*, the causal agent of oak powdery mildew (Mougou et al., 2008). Forty leaves were sampled and DNA was extracted from four 0.5 cm² discs per leaf. Fungal taxa were characterized by 454 pyrosequencing of the ITS1 region (as described in Cordier et al., 2012). Bacterial taxa were characterized by sequencing the hypervariable V6 region of the 16S rRNA gene on an Illumina platform (as described in Gloor et al., 2010). The taxonomic composition of the samples was obtained by filtering and clustering the sequences using QIIME (Caporaso et al., 2010) and Usearch (Edgar, 2013). The software SparCC (Friedman and Alm, 2012) was then used to compute the correlations between the abundances of microbial taxa across samples. Only significant positive correlations (co-associations) and significant negative associations (co-exclusions) were retained as links in the microbial network. The network was visualized using Gephi (Bastian and Heymann, 2009). Its analysis revealed that *E. alphitoides* has a particular connectivity behaviour, as it is predominantly connected to the network through strong negative links. The presence of the pathogen is thus associated to the absence of other micro-organisms. To go further into the mechanistic interpretation of this network, the microbial association network should be turned into a microbial interaction network having only *direct* ecological interactions as links. This may be done by firstly using environmental conditions to screen out associations that indicate a *shared environmental preference* and, secondly, using network inference methods to predict *direct* links between microbial taxa (Faust and Raes, 2012; Kurtz et al., 2015;

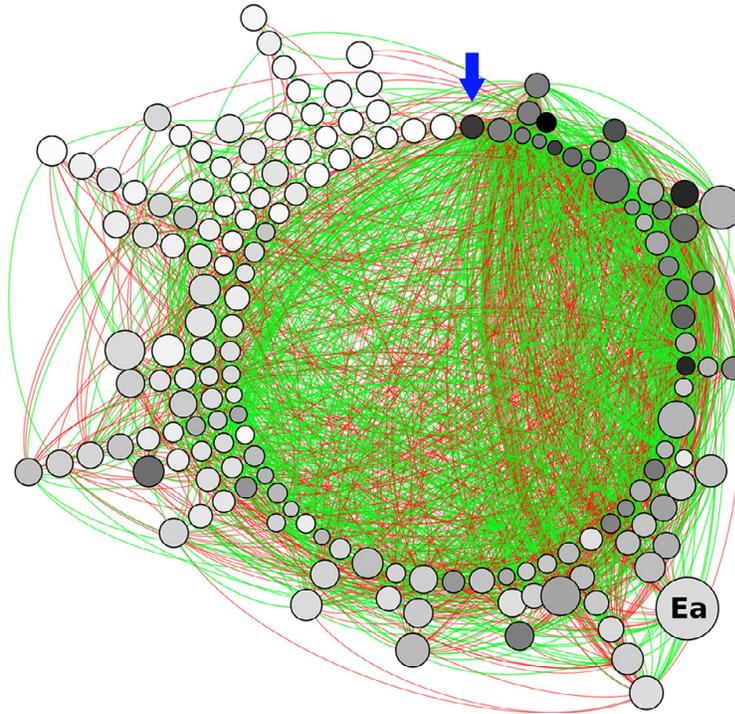


Figure 1 Microbial association network of the leaves of an oak tree (*Q. robur* L.) susceptible to the foliar fungal pathogen *E. alphitoides* (Ea). Each node represents a microbial taxon (either bacterial or fungal) and each link represents a significant correlation between their abundances. Red (grey in the print version) and green (grey in the print version) links indicate co-exclusions and co-associations, respectively. The arrow indicates the node with the highest degree (i.e. the highest total number of links). Degree decreases clockwise, with nodes stacked on the same line having the same degree. The size of the nodes is inversely proportional to the sum of the correlation coefficients: larger nodes have more number and/or stronger negative associations. Darker nodes have higher betweenness centrality (calculated on the absolute values of associations), suggesting that they are topological keystone taxa. *E. alphitoides* is predominantly connected to the network through strong negative links (co-exclusions) but is not a good candidate for topological keystone species.

[Schwaller et al., 2015](#)). Such network would allow us to identify the microbial species that directly impede or facilitate pathogen development.

3.2 Studying the Hologenome Theory of Evolution Using NGS-Based Microbial Networks

Humans are drastically and rapidly altering the environment, including climate, and many species may not be able of adapting quickly enough to these new conditions ([Carroll et al., 2014](#)). The mismatch between the phenotype of agriculturally important plants and new climate is an overarching

challenge in forestry, agriculture and conservation biology. A widespread debate concerns whether to use local versus external sources of genetic material for replanting to best anticipate responses to climate change (Carroll et al., 2014). The hologenome theory of evolution is relevant to this debate. It posits that the holobiont (the plant or animal with all of its associated micro-organisms) is a unit of selection in evolution. According to this theory, host and ‘associates’ genomes act as a consortium that copes with environmental change by promoting adaptation and evolving as a whole. During periods of rapid environmental change, a diverse microbial symbiont community can aid the holobiont in surviving, multiplying and ‘buying the time’ necessary for the host genome to evolve (Zilber-Rosenberg and Rosenberg, 2008). There is indeed empirical evidence that host speciation is under the influence of its interactions with micro-organisms (Janson et al., 2008) and that microbial communities can be one determinant shaping the adaptation and evolution of higher organisms (Dittami et al., 2015; Rosenberg et al., 2007) .

NGS techniques are useful tools to characterize the tight associations of eukaryotes with microbes. In a recent study, they were used to characterize the internal microbial communities of the algae *Caulerpa* at the scale of the Mediterranean Sea (Aires et al., 2015). These endophytic microbial communities are hypothesized to play important roles in development, defence and metabolic activities of their host algae. Three *Caulerpa* species were sampled and the bacterial lineages within each sampling unit were identified through high-throughput sequencing of the hypervariable V4 region of the 16S rRNA gene. Then, networks of co-occurring bacterial lineages were analyzed by combining percolation theory (Moalic et al., 2012; Stauffer and Aharony, 1994) and community detection algorithms (Fortunato, 2010; Lancichinetti and Fortunato, 2009; Leger et al., 2014). Here, the Bray–Curtis index was used to assess the similarity in the distribution of bacterial lineages across samples: two lineages are linked if their distribution similarity is higher than 0.62, which is the percolation threshold of the network. Second, modules were delimited by using the leading eigenvector algorithm (Fig. 2). The results revealed that a very large fraction of the bacterial community is species-specific, even in areas where distinct *Caulerpa* species occur in sympatry. These species-specific bacterial lineages account for the modular structure of the co-occurrence network. Such specificity of endophytic bacterial communities is coherent with the hologenome theory of evolution. As several *Caulerpa* taxa have extended their range through invasion in different parts of the world, including the Mediterranean Sea, future

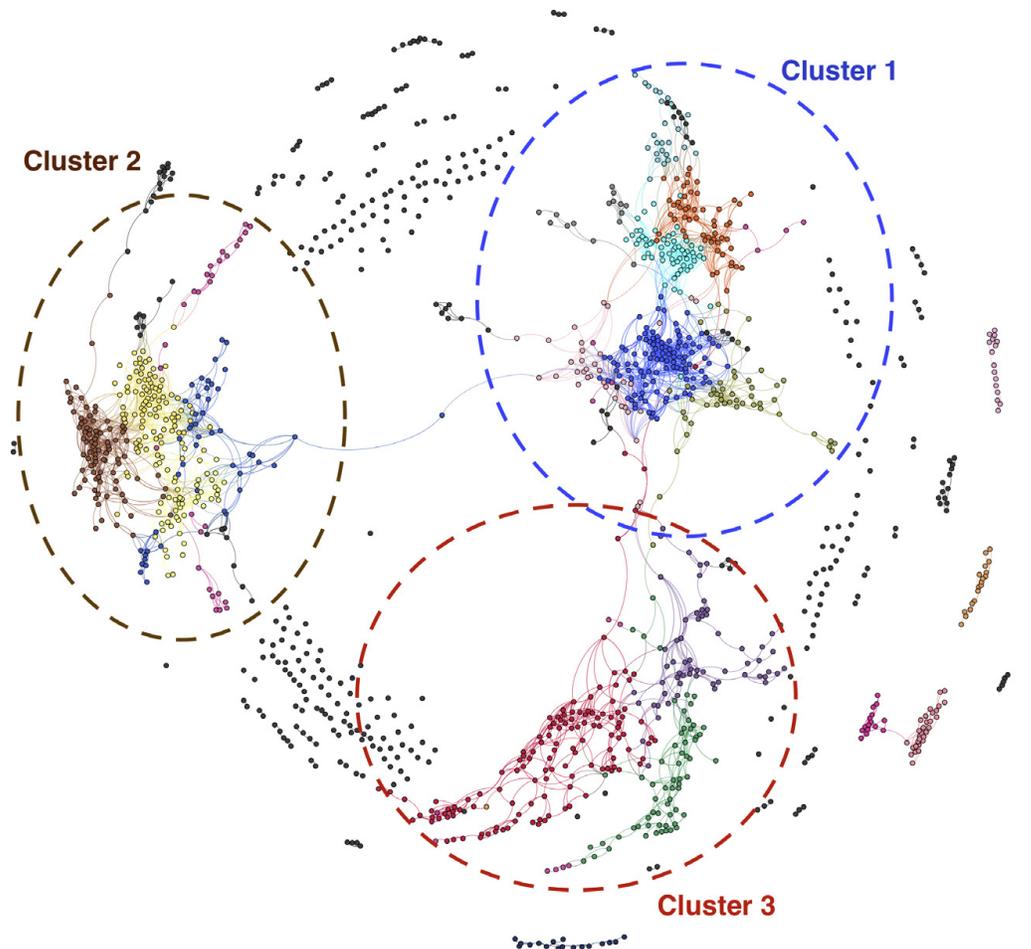


Figure 2 Associations between the endophytic bacterial lineages of the algae *Caulerpa* at the scale of the Mediterranean Sea. Each node is a bacterial lineage. Two lineages are linked if they tend to co-occur in the same samples. The Bray–Curtis index was used to assess the similarity in their distribution across samples: two lineages are linked if their distribution similarity is higher than 0.62, which is the percolation threshold of the network. Nodes are coloured according to their belonging to a module. Modules were delimited by using the leading eigenvector algorithm. The modules form three larger clusters that are, respectively, dominated by bacterial lineages specific of *C. prolifera*, *C. racemosa* var. *turbinata-uvifera*, and *C. racemosa* var. *cylindracea*.

studies should investigate the role of these microbial communities on the invasiveness of their host.

3.3 Testing the Niche Partitioning Theory with NGS-Based Trophic Networks

The niche partitioning theory is central to our understanding of biodiversity. The term niche partitioning refers to the process by which natural selection drives competing species into different patterns of resource use or different

niches (Hector and Hooper, 2002; MacArthur, 1958). This differentiation of ecological niches reduces competition and promotes co-existence between species (Chesson, 2000; Levine and HilleRisLambers, 2009). NGS techniques promise to bring significant changes in our understanding of niche partitioning because they provide information about the entire diet range of a species, while also highlighting new and unexpected trophic links. For instance, Ibanez et al. (2013) carried out a trait-based estimation of the trophic niche width of four species of grasshoppers through choice experiments and NGS study of their diet. They showed that observed trophic niche breadth in generalist herbivorous insects depends on both species-specific food preferences and habitat diversity, and that it is not an intrinsic property of the species as usually considered in theoretical studies. Kartzinel et al. (2015) also used NGS techniques, in combination with stable isotope analyses, to investigate trophic niche partitioning among sympatric large mammalian herbivores in Kenya. They observed unsuspected fine-scale resource partitioning even between species within the same trophic guild. Consequently, through this study, NGS methods illuminated mechanisms behind the large diversity of co-existing herbivorous mammal species observed in African savannas.

Here, we used NGS techniques to resolve the trophic interactions in a community of carabid beetles inhabiting European arable landscapes. These beetles significantly contribute to the biological control of pests (Kromp, 1999) but their contribution is unpredictable given their broad diet spectrum including alternative preys such as other natural pest enemies. We used NGS approach in order to investigate changes in carabid diet among 14 common arable species by analyzing the prey DNA contained in their guts. The carabid beetles were sampled in six arable fields in two different agroecosystems in Brittany, France. To cover the whole carabid diet spectrum, four barcode genes were combined, including mitochondrial 16S and COI for characterizing animal prey (Bienert et al., 2012; De Barba et al., 2014) and the chloroplast trnL for detecting consumed plant species (Taberlet et al., 2007). We obtained bipartite ecological networks showing an extensive degree of niche overlapping between carabid species (Fig. 3), confirming their generalist feeding behaviour. The use of NGS also allowed us for the first time to quantify carabid community contribution to ecosystem services (the consumption of several animal and plant pest species) and dis-services (the consumption of other service-providing organisms) (Fig. 4). These findings suggest that important ecological functions as pest regulation and intraguild predation may not be associated with the identity of particular species but are

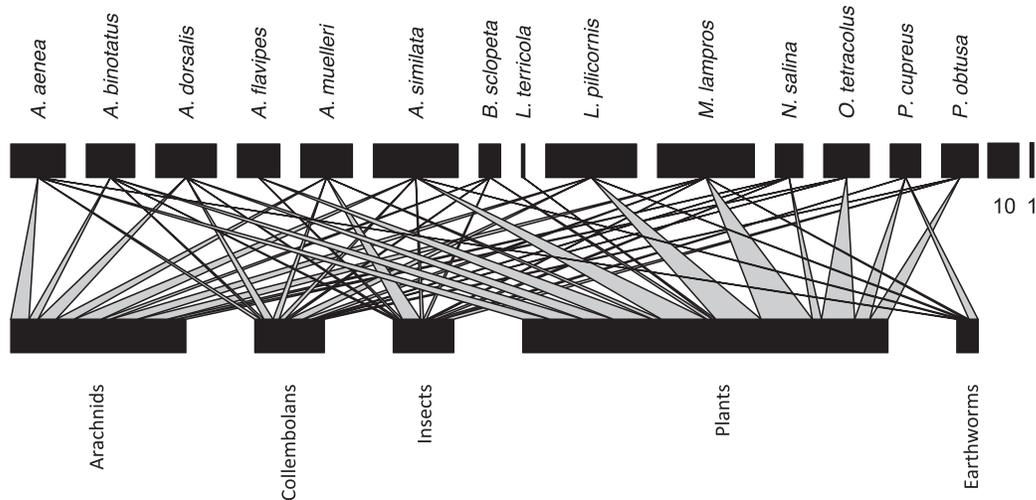


Figure 3 Quantitative predator–prey network recovered by analysing gut contents of 1414 carabid species using an NGS approach. Lower bars represent prey category (arachnids, collembolans, insects, plants, and earthworms) abundance and upper bars represent the abundance of carabid species positive for at least one prey category. Scales are indicated at right. Linkage width indicates frequency of each trophic interaction. Carabid beetles were sampled in six different fields.

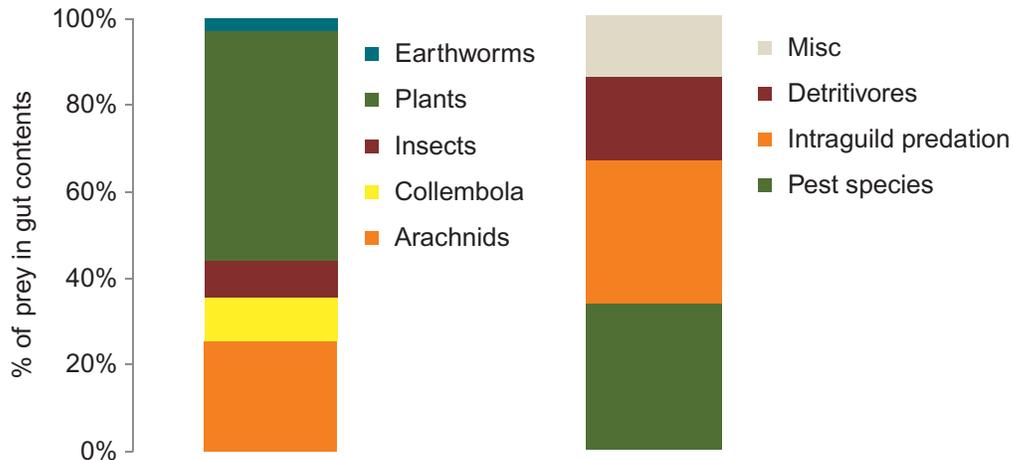


Figure 4 The relative proportions of five major prey categories (arachnids, collembolans, insects, plants, and earthworms) recovered in gut contents of 503 individuals from 14 carabid species using NGS meta-barcoding approach. Using their taxonomic identity, all animal and plant preys were assigned into three functional categories (pest species, detritivore species, and other natural predators). All preys that did not fit any functional category (e.g. tree or non-pest plant DNA) were grouped into a fourth category (miscellaneous). This approach allowed us to quantify the relative contribution of the carabid community to ecosystem services (i.e. pest control) and dis-services (i.e. control of other service-providing organisms) within agroecosystems.

more likely an emergent property of the community probably modulated by extrinsic environmental factors. Further investigations are required for see how important agronomical features such as the cropping system or management intensity impact the structure of the trophic web.

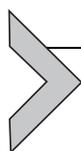
3.4 Challenges to Be Addressed to Get Predictive Insights from NGS-Based Networks

In the three-forementioned examples, the NGS-based ecological networks were static (Figs. 1–3). To predict the impact of environmental stressors on ecosystem services, future research should move from static networks to dynamic ones. In particular, it should explore whether time-varying networks have ‘early warning’ properties that can predict abrupt ecological transitions or ‘tipping points’ (Dakos and Bascompte, 2014; Faust et al., 2015). Metagenomic time series data can provide useful information on the dynamics of microbial communities (Fuhrman et al., 2015). Future research should also go beyond community composition by integrating data about functional gene expression, transcripts, proteins and metabolites. These data will yield crucial information, in addition to current co-occurrence patterns, about ‘who?’ does ‘what?’ and ‘when?’ within the community (Faust et al., 2015).

Future research should also aim at validating the ecological interactions learned from NGS data. In the first example (Fig. 1), co-association and co-exclusion patterns indicate which pairs of microbial taxa are the most likely to interact together. But experimental validation of these interactions is required. Reference networks, with well-known ecological interactions, are necessary to assess the relevance of the methods of network reconstruction. For instance, a machine-learning methodology to reconstruct networks from co-occurrence data and species functional traits has been demonstrated (Bohan et al., 2011). This produced hypothetical food webs in an agricultural system that bore all the hallmarks of real food webs and were validated against both the literature and direct molecular biological data for specific trophic interactions (Davey et al., 2013). Other methods have also been used successfully to reconstruct known ecological networks based on species abundance data (Aderhold et al., 2012; Faisal et al., 2010; Milns et al., 2010). The current challenge is to test all these methods on occurrence or abundance data derived from NGS datasets, by using a well-resolved empirical network as a reference.

Finally, future research must continue to improve the qualitative and quantitative reliability of NGS data, while keeping the costs down. Future

research must continue to evaluate and reduce the biases in species composition due to primers choice and amplification by PCR (Berry et al., 2011; Gonzalez et al., 2012; Lee et al., 2012; Patin et al., 2013; Pinto and Raskin, 2012), to characterize the noise profiles of the various sequencing platforms and to develop bioinformatic pipelines for taking them into account (Gaspar and Thomas, 2013; Gilles et al., 2011; Reeder and Knight, 2011). Future research must also develop new statistical methods for estimating species richness and diversity from NGS data (Haegeman et al., 2013) and improve the taxonomic assignment of the species. For instance, the combination of several barcodes targeting restricted groups of organisms could improve the taxonomic resolution of preys in dietary samples (e.g. Deagle et al., 2009). However, combining of several barcode markers could increase considerably the costs of NGS techniques. Multiplexing several barcodes within the same PCR would help improving the cost effectiveness (De Barba et al., 2014), but these still require time-consuming methodological development. Additionally, the direct comparison of the relative proportion of prey groups, by comparison of their barcode sequence counts, is highly questionable due to the step of variation in amplification among barcodes by PCR (Taberlet et al., 2012). The rapid development of direct shotgun sequencing techniques, which requires no PCR amplification step, promises to resolve these problems. In two recent studies, the shotgun sequencing of long fragments of nuclear, mitochondrial and chloroplast DNA fragments from an herbivorous mammal (Srivathsan et al., 2014) and a carnivorous coleopteran (Paula et al., 2015) provided reliable and quantitative species-level identification of their consumed preys. The shotgun sequencing simultaneously revealed detailed information about the host's genetics, as well as information about the diversity of host's parasites and bacterial symbionts. These findings open several exciting perspectives for merging multiple types of ecological interactions in the same network (Fontaine et al., 2011; Kéfi et al., 2012), or for assessing the impact of host genetic differentiation or speciation into ecological network dynamics (e.g. Robinson et al., 2015).



4. THEORETICAL METHODS FOR DECIPHERING ECOLOGICAL NETWORKS FROM NGS DATA

4.1 The Input Data

Meta-barcoding, also referred to as amplicon-based community profiling, generates a list of operational taxonomic units (OTUs) and their distribution in the environmental samples. Each OTU is a group of related sequences

(with a similarity percentage higher than 97%, in most cases) (Schlaeppli and Bulgarelli, 2015). Taxonomic assignment of OTUs is usually performed by comparing the most abundant sequence of each OTU to the sequences deposited in GenBank or in curated databases such as Greengenes, BOLD, SILVA or UNITE (Abarenkov et al., 2010; DeSantis et al., 2006; Kõljalg et al., 2005; Quast et al., 2013; Ratnasingham and Hebert, 2007). The ecological relevance of the number of sequences per OTU and per sample is a matter of debate as the quantitative link between sequence reads and species relative abundance is not always straightforward due to methodological biases. The storage of environmental samples, before DNA extraction, may cause biases in the community composition revealed by NGS technique (U'Ren et al., 2014). In the case of gut samples, differences in digestion rates between types of prey may also cause bias (Deagle et al., 2009). Finally, whatever the type of environmental sample, variations in the amplification efficiency of the barcode region among species may distort the community composition (Berry et al., 2011; Gonzalez et al., 2012; Lee et al., 2012; Patin et al., 2013; Pinto and Raskin, 2012). Studies of plant pathogen species have, however, shown that pathogen abundances measured by visual symptoms correlate with their relative abundance in NGS datasets (Sapkota et al., 2015). The number of sequences of a given OTU in a given sample therefore contains some information on the OTU abundance in that sample, despite potential biases introduced by PCR (Cotton et al., 2014). This is important, because associations or interactions between OTUs are often derived from sequence counts (Faust and Raes, 2012; Friedman and Alm, 2012; Kurtz et al., 2015).

The datasets required for reconstructing ecological networks from NGS data are the OTU table, giving the estimated abundance of each OTU in each sample, with or without some additional information on the OTUs (e.g. functional traits of the corresponding species) and on the environmental samples (e.g. host species, abiotic conditions at the time of sampling) (Table 1). Association networks can be derived from the OTU table, as shown in the two first examples (Figs. 1 and 2). In these examples, the nodes of the association networks were fungal and/or bacterial OTUs. The links were the significant correlations between their abundance across samples (Fig. 1) or the similarity in their distribution across samples (Fig. 2). In the first example, links were positive or negative depending on the correlation sign. Such correlations should not be calculated based on raw sequence counts because variations in the total number of sequences per sample (a technical bias) may yield spurious results. Correlations between normalized

Table 1 Description of the Data Entering in the Construction of NGS-Based Ecological Networks, with Their Names in the Two Main Fields of Research (Statistical Inference of Networks and Logic-Based Machine-Learning Algorithms)

Data Description	Type of Data	Name Given in Statistical Modelling	Name Given in Logic-Based Machine Learning
Measured occurrence or abundance of species/OTUs in the studied sites/samples	Input	Observed variables	Observable predicates
Additional information on the species/OTUs (e.g. functional traits) or on the sites/samples (e.g. abiotic environment)	Input	Covariates	Background knowledge
Rules of interaction (e.g. species/OTU X can eat Y if X is bigger than Y)	Input or output	Constraints on the space of possible networks to explore	Logical rules which can be part of background knowledge or can be learned from input data
Hypothetical relationships between species/OTUs that can be visualized as a network	Output	Inferred edges	Abduced links

counts may also yield spurious results and several methods have been proposed to circumvent this issue (Deng et al., 2012; Faust and Raes, 2012; Friedman and Alm, 2012; Kurtz et al., 2015).

Such networks representing *spatial or temporal associations* between microbial OTUs are not entirely satisfactory because they do not match the mechanistic framework typically used in network ecology. As mentioned above, network ecology developed mainly on the observation of *direct ecological interactions* between several sets of macro-organisms (e.g. predator–prey, plant–pollinator, plant–herbivore, plant–seed disperser, anemone–fish interactions) (Ings et al., 2009). The predator–prey network recovered by using NGS techniques to describe gut contents (Fig. 3) is similar in terms of interaction type, but not the microbial association networks of the two first examples (Figs. 1 and 2). To go beyond these association networks, the first challenge is to distinguish associations due to *direct ecological interactions* and those due to *indirect ecological interactions* (Kurtz et al., 2015; Schwaller et al., 2015). The second challenge is to take into account environmental

variables into network models (Faust and Raes, 2012) in order to remove associations due to *shared environmental preference* (Ovaskainen et al., 2010). The third challenge is to integrate ecological knowledge on OTUs, in the form of ‘rules of interaction’. This integration may improve the reliability of the network, as shown previously for macro-organisms (Bohan et al., 2011), and limits the space of possible networks that has to be explored.

Below we review the methods that can be used for addressing these challenges. These methods can be classified into two categories: statistical inference of networks and logic-based machine-learning algorithms. Both categories use similar input data and aim at constructing plausible and testable networks, which explain the observed data the best. However, they constitute different fields of research and therefore use a different vocabulary. The main terms are summarized in Table 1.

4.2 Inferring Ecological Interactions Using Statistical Models

Statistical models of graphs (Lauritzen, 1996; Whittaker, 1990) are a natural approach to depict the intricate interactions and relationships between species (or OTUs). They have proved to be a valuable framework for modelling *direct* relationships between species (or OTUs), based on conditional dependencies in occurrence and abundance data (Table 1). Here, we describe two popular approaches that might have value for reconstructing networks from NGS data.

4.2.1 Bayesian Networks and Dynamic Bayesian Networks

Bayesian Networks (BNs, Jensen and Nielsen, 2007) and their temporally explicit extension Dynamic Bayesian Networks (DBNs) framework (Dean and Kanazawa, 1989) have been used to reconstruct interactions from occurrence or abundance data. In ecology, BNs have been used to model species interactions and species–habitat relationships (Aderhold et al., 2012; Milns et al., 2010), while DBNs have been used to infer microbial networks from time series (Faust et al., 2015). The structure of BNs and DBNs can be derived from occurrence or abundance data, and possibly covariates representing environmental conditions, by using score-and-search techniques (Daly et al., 2011; Friedman et al., 1998). A score function (e.g. the Bayesian Information Criterion) is used to measure the fit between the observed data and an inferred network. Rules of iteration then move from one possible network structure to another in the space of possible network structures to evaluate fit. The aim is to find a network that maximizes the fit. Exact optimization of the score is out of reach because

of the high number of possible network structures. Various heuristics have been proposed to improve the search and its efficiency, including the ‘greedy search’ where only local improvements of the network structure are explored.

4.2.2 Gaussian Graphical Models with Sparse Regularization

Coupling Gaussian graphical models (GGM) with sparse regularization has become a popular method of inference in the last decade, because it allows us to deal with large networks involving thousands of nodes by setting many of the links to be zero. In the standard Gaussian setting, when the data are normally distributed, the approach assumes that the joint distribution of all OTUs follows a multivariate Gaussian distribution. The associated covariance matrix entirely describes the dependency structure between the OTUs. The precision matrix, which is formed as the inverse of the covariance matrix, is then a direct proxy for the expected network structure as the matrix entries are proportional to the partial covariances between the OTU abundances. Therefore, learning the graphical structure of conditional dependencies can be reduced to a problem of variable selection to optimize fit. This task can be efficiently performed by means of sparse regularization such as the Lasso (Tibshirani, 2011). Various lasso algorithms have been developed, including the ‘neighbourhood selection’ (Meinshausen and Bühlmann, 2006) and the ‘Graphical Lasso’ (Friedman et al., 2008). Some implementations can deal with millions of nodes.

To overcome the Gaussian assumption and comply with sequence count data like those obtained with NGS techniques, research is being done to broaden the applicability of GGM. Two lines of research have emerged on this topic. The first is to transform the original data into a ‘Gaussian’ setting, via simple transformations in order to use the well-understood GGM with sparse regularization framework (Liu et al., 2009). The second, which we do not discuss here, relies on the use of statistical models tailored to count data (Yang et al., 2013). The technique of data transformation has been successfully applied to ecological network inference (Kurtz et al., 2015), using a general workflow, as follows: (i) the OTU count data are preprocessed and normalized to meet the Gaussian assumption; (ii) a standard sparse GGM inference method (either ‘neighbourhood selection’ or ‘graphical-Lasso’) is used to select networks, describing direct links between OTUs; and, (iii) step (ii) is iterated several times on many random subsamples of the original data. The final network only retains the most stable edges, which appear robust because they are selected in most of the subsamples. This three-step

strategy—normalization plus inference plus stabilization—has also been successfully used on other types of genomic data (Marbach et al., 2012).

An advantage of the GGM framework is that it is well suited to theoretical analysis, which provides insights into the ‘data’ situations where the methods may either be useful or not. Most strikingly, an analysis of network selection consistency by Ravikumar et al. (2011) suggested that it should be possible to infer network structure when a network has many nodes and the sample size remains moderate, under the provision that there are no nodes in the network with very high connectance. Such statistical results are practically important because they give some confidence that networks may be inferred from NGS-based data where the number of samples is typically much smaller than the number of OTUs.

4.3 Learning Ecological Interactions Using Logic-Based Machine-Learning Algorithms

Like statistical models of graphs, the purpose of logic-based machine-learning algorithms is to construct plausible and testable networks, which best explain the observed data. The possibility of easily integrating the existing background knowledge into the learning process, e.g. in the form of ‘rules of interaction’, is a key advantage of logic-based machine-learning algorithms over statistical models (Table 1). Logic-based machine-learning algorithms have already been used successfully to automatically generate trophic networks directly from species occurrence data combined with background knowledge, including information about the species body size and functional groups (Bohan et al., 2011; Tamaddoni-Nezhad et al., 2013). They have also been applied successfully to other problems in ecology, in particular to model population dynamics (e.g. modelling phytoplankton growth for the Danish Lake Glumsø; Todorovski et al., 1998), and their relevance has also been demonstrated in many challenging domains in computational biology including predictive toxicology (e.g. King et al., 1996), pharmacophore design (e.g. Srinivasan et al., 2006) and protein structure prediction (e.g. Cootes et al., 2003).

Below, we describe two types of logic-based machine-learning algorithms that may be used to learn ecological networks from NGS data. Both types are capable of learning networks from species (or OTU) occurrence or abundance data. The first type, called inductive logic programming (ILP; Muggleton, 1991), is capable of using background knowledge, such as the existing knowledge about the species and their environment to hypothesize (learn) interactions (Table 1). In the case where background

knowledge may be incomplete or a subject of the learning itself, a new approach called meta-interpretive learning (MIL; Muggleton et al., 2014) may be used. To our knowledge, neither of these approaches have yet been used to learn ecological networks from NGS data, but the first approach has been used for classical ecological association data (co-occurrences in spatially replicated field sampling) to learn predator–prey networks (Bohan et al., 2011; Tamaddoni-Nezhad et al., 2013). Ecological interactions recovered from the high-throughput sequencing of predator gut contents (e.g. Fig. 3) or predicted by functional traits (e.g. body size or gape width), may be included as background knowledge, or may be used to validate the machine-learned network. The second approach is likely to be more appropriate to the learning of microbial networks (e.g. Figs. 1 and 2) where background knowledge on microbial OTUs is often scarce.

4.3.1 Inductive Logic Programming

ILP systems (Muggleton, 1991) use a given set of positive and negative examples $E = \{E^+ \cup E^-\}$, i.e., E is the union of positive and negative examples, background knowledge B to construct a hypothesis, H , that explains E^+ relative to B such that the extended theory is self-consistent, i.e. $B \cup H$ logically implies E^+ and $B \cup H \cup E^-$ is logically consistent. The components E , B and H are each represented as logic programmes. In the case of machine learning of trophic networks, ILP systems can be used to learn ground hypotheses H in the form of trophic relations between species (or OTUs). Background knowledge includes logical rules, such as $R \subseteq B$ (i.e. R is a subset of B), to describe the species occurrence or abundance (observable predicates) in terms of the trophic interactions (abducible predicates) (Table 1).

This approach was used to learn a trophic network from an extensive Vortis suction sampling of invertebrates in 257 arable fields across the UK (Bohan et al., 2011). These fields were part of the farm scale evaluations (FSE) of genetically modified, herbicide-tolerant (GMHT) crops. The change in invertebrates abundance data with the GMHT treatment was regarded as the primary observational data for the learning: observable predicates were represented by $abundance(X, S, up)$ (or $abundance(X, S, down)$) expressing the fact that the GMHT treatment increased or decreased the abundance of species X at site S . The aim was to learn abducible predicates $eats(X, Y)$, capturing the hypothesis that species X eats species Y . Additional information on the species were used to constrain the search for abducible predicate $eats(X, Y)$, by assuming that X should be a predator and bigger

than Y . Predicates $predator(X)$ and $bigger_than(X, Y)$ were provided as part of the background knowledge. This information was integrated within a logical rule describing the observable predicates ($abundance$) in terms of the abducible predicates ($eats$):

$abundance(X, S, Dir)$ if $predator(X)$, $bigger_than(X, Y)$, $eats(X, Y)$, $abundance(Y, S, Dir)$ where Dir can be either up or down. This rule expresses the inference that following a perturbation in the ecosystem (i.e. the introduction of GMHT crops), the increased (or decreased) abundance of species X at site S can be explained by X eating species Y , which is lower in the food chain, and by changes in the abundance of species Y . Given this model and the observable data, the Abductive ILP system Progol 5.0 was used to generate a set of ground hypotheses (i.e. hypothetical trophic links) which was visualized as a trophic network (Bohan et al., 2011). The initial study was then extended by learning the trophic network from a larger dataset (Tamaddoni-Nezhad et al., 2013). In both cases, a probabilistic ILP approach, called Hypothesis Frequency Estimation (HFE, Tamaddoni-Nezhad et al., 2012), was used for estimating probabilities of hypothetical trophic links. These probabilities were represented as the thickness of trophic links in Bohan et al. (2011) and Tamaddoni-Nezhad et al. (2013). Ecologists who examined the first machine-learned food web (Bohan et al., 2011) found that many of the learnt trophic links were corroborated by the literature. In particular, links ascribed with high probability by machine learning were shown to correspond well with those having multiple references in the literature. Novel, high probability links were also suggested, and some of these have recently been tested and confirmed by subsequent empirical studies. For example, in the hypothesized food webs, some species of spiders always appeared as prey for other predators; a result that was unexpected because spiders are obligate predators. This hypothesis was tested using molecular analysis of predator guts and it was found that in this system spiders do appear to play an important role as prey (Davey et al., 2013). This finding was reconfirmed in Section 3.3 and Fig. 3. Thus, even though some of the hypothesized links were unexpected, these were in fact confirmed later and this provided an extremely stringent test for the machine-learning approach.

4.3.2 Meta-Interpretive Learning

In the machine-learning settings described in the previous section, the search for trophic links was constrained by additional information on the species

(e.g. body size, trophic behaviour) which were provided as part of background knowledge. The logical rule stated that X may eat Y if X is a predator bigger than Y . However, for most communities and ecosystems, including microbial communities, this kind of background knowledge may not be available or it may be incomplete. MIL (Muggleton et al., 2014) is a new machine-learning approach capable of predicate invention and recursive rule learning. This new approach can be used for learning both the interactions between species (or OTUs) and the ‘rules of interaction’ directly from species occurrence or abundance data. In this case, background knowledge does not include any specific knowledge on the species but includes higher-order meta-rules, $M \subseteq B$, which are activated during the proving of examples in order to generate hypotheses, H . A recent study showed that MIL can be used to re-construct a simplified food web and learn interaction rules directly from data (Tamaddoni-Nezhad et al., 2015). We believe that this new learning setting will be useful for learning ecological networks from NGS data whenever the interaction rules are not known before hand.



5. CONCLUSION

Most interactions between species are difficult to observe, and as a consequence the ecological networks that we typically reconstruct and analyze to understand ecosystem function are incomplete. Through the examples given in this review, we have argued that NGS techniques permit the characterization of biodiversity in complex environmental samples (e.g. soil, water, plant tissues, faeces, pellet, gut content, etc.) containing hundreds of microbial OTUs and multiple macro-organism species. In the lists of co-occurring species and OTUs that we can discern with NGS techniques, there are the ‘ghosts of interactions past’ from which we could learn robust and more complete ecological networks, at any specified taxonomic resolution and for organisms from all the Kingdoms of life. In this review, we have then showed that various statistical and machine-learning approaches are available for performing such network reconstruction, but the majority have yet to be applied to NGS data. We believe that the combination of these theoretical approaches with cost-effective NGS techniques will allow us to study species interactions under all environmental conditions at high replication. However, three outstanding challenges remain to be overcome to achieve these aims. These are to:

- improve the qualitative and quantitative reliability of NGS data, while keeping the costs down;
- use well-characterized networks of ecological interactions, in order to test the validity of the various methods of network learning and reconstruction;
- develop theoretical approaches to enable the learning of temporally dynamic ecological networks.

Should these three challenges be met, then we foresee a step change in our ability to measure, understand and monitor the world's ecosystems and the functions and services they provide.

ACKNOWLEDGEMENTS

We thank the *métaprogramme EcoServ* (INRA) for funding the meeting that gave rise to this chapter. The work presented in Fig. 1 was funded by the METAPHORE project (AIP Bioressource) and a grant from the French Ministry of Research and Education (MENRT no. 2011/AF/57). We thank Cecile Robin for helpful discussions. The work presented in Fig. 2 was supported by the Portuguese Foundation for Science and Technology 609 (FCT) and FEDER with the project IBISA (PTDC/MAR/64749/2006). We thank Tania Aires, Ester Serrao and Sophie Arnaud-Haond for providing the dataset. The study presented in Figs. 3 and 4 was funded by the two French ANR projects, Landscaphid (ANR-09-STRA-05) and Peerless (ANR-12-AGRO-0006). The second author acknowledges the support of an EPSRC 'Pathways to Impact Award' during the writing of this paper.

D.B. is supported by two ANR projects, PEERLESS (ANR-12-AGRO-0006) and AGROBIOSE (ANR-13-AGRO-0001).

GLOSSARY

Barcode A short genetic sequence taken from a standardized portion of the genome that is used to identify species.

Biodiversity The variety of life, including variation among genes, species, and functional traits. It is often measured as: richness is a measure of the number of unique life forms; evenness is a measure of the equitability among life forms; and heterogeneity is the dissimilarity among life forms (Cardinale et al., 2012).

Conditional dependency The relationship between two variables conditioned on all other variables. In a Gaussian setting, conditional dependency is measured by partial correlation.

Diagnostic PCR A PCR assay which is used to test samples for the presence of DNA from a specific species or a group of organisms.

Ecosystem functions (or functioning) Ecological processes that control the fluxes of energy, nutrients, and organic matter through an environment. Primary production,

for instance, is an ecosystem function. It is the process by which plants use sunlight to convert inorganic matter into new biological tissue (Cardinale et al., 2012).

Environmental DNA (eDNA) Genetic material obtained directly from environmental samples (soil, sediment, water, etc.).

Graph A mathematical object where entities (represented by vertices) are connected with links (called edges). In graphical models, graphs are used to depict the structure of conditional dependencies between variables.

Holobiont The functional entity composed by a macro-species and its associated symbiotic microbes.

Logic-based machine learning A form of machine learning which uses logic-based inference and representation. For example, in inductive logic programming (ILP), the training examples, background knowledge, and the learned hypotheses are all represented as logic programmes.

Meta-barcoding (amplicon-based community profiling) A method of biodiversity assessment combining DNA-based identification and high-throughput sequencing. It uses universal PCR primers to mass-amplify genetic markers from mass collections of organisms or from environmental samples.

Next-generation sequencing (high-throughput sequencing) Technologies that parallelize the sequencing process, producing thousands, or millions of sequences concurrently.

Niche partitioning theory refers to evolutionary and ecological processes leading to differential resource exploitation between species in response to interspecific competition.

Operational taxonomic units (OTUs) are usually defined as clusters of similar barcode sequences (16S rDNA, ITS etc.), frequently intended to represent some degree of taxonomic relatedness.

OTU table Matrix giving the number of sequences per OTU and per sample.

Pathobiome The pathogenic agent plus the members of its biotic environment.

Sanger sequencing (chain-termination method) Method of DNA sequencing based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during *in vitro* DNA replication. Developed by Frederick Sanger and colleagues in 1977, it was the most widely used sequencing method for approximately 25 years. It can be used for fairly short strands of 100–1000 bp.

Shotgun sequencing Method used for sequencing long DNA strands that was one of the precursor technologies for full genome sequencing. DNA is fragmented into numerous small fragments that are sequenced using the chain termination method to obtain reads. Multiple overlapping reads for the target DNA are obtained and computer programmes then use the overlapping ends of different reads to assemble them into a continuous sequence.

Sparse regularization Technique from constrained optimization used to force some entries to zero in a vector of parameters. When applied to network inference, it allows to deal with thousands of nodes, by coercing many of the edges to zero.

Statistical inference The process of drawing conclusions on a population based on data.

Tag A unique DNA sequence ligated to fragments within a sequencing library for downstream sorting and identification. Tags are typically a component of adapters or PCR primers and are between 8 and 12 bp. Libraries with unique tags can be pooled together and sequenced in the same sequencing run (=multiplexing). Reads are later identified and sorted via bioinformatic pipelines.

REFERENCES

- Abarenkov, K., Nilsson, R.H., Larsson, K.H., Alexander, I.J., Eberhardt, U., Erland, S., Hoiland, K., Kjoller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A.F.S., Tedersoo, L., Ursing, B.M., Vralstad, T., Liimatainen, K., Peintner, U., Koljalg, U., 2010. The UNITE database for molecular identification of fungi—recent updates and future perspectives. *New Phytol.* 186, 281–285.
- Abrams, P.A., Matsuda, H., 2005. The effect of adaptive change in the prey on the dynamics of an exploited predator population. *Can. J. Fish. Aquat. Sci.* 62, 758–766.
- Aderhold, A., Husmeier, D., Lennon, J.J., Beale, C.M., Smith, V.A., 2012. Hierarchical Bayesian models in ecology: reconstructing species interaction networks from non-homogeneous species abundance data. *Ecol. Inform.* 11, 55–64.
- Aires, T., Moalic, Y., Serrao, E.A., Arnaud-Haond, S., 2015. Hologenome theory supported by co-occurrence networks of species-specific bacterial communities in siphonous algae (*Caulerpa*). *FEMS Microbiol. Ecol.* 91 (7): fiv067.
- Aizen, M.A., Morales, C.L., Morales, J.M., 2008. Invasive mutualists erode native pollination webs. *PLoS Biol.* 6, e31.
- Albrecht, M., Duelli, P., Schmid, B., Müller, C.B., 2007. Interaction diversity within quantified insect food webs in restored and adjacent intensively managed meadows. *J. Anim. Ecol.* 76, 1015–1025.
- Albrecht, M., Padrón, B., Bartomeus, I., Traveset, A., 2014. Consequences of plant invasions on compartmentalization and species' roles in plant–pollinator networks. *Proc. R. Soc. B* 281, 20140773.
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., Taberlet, P., 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Mol. Ecol. Resour.* 14, 306–323.
- Barberán, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6, 343–351.
- Bascompte, J., Jordano, P., Melian, C.J., Olesen, J.M., 2003. The nested assembly of plant–animal mutualistic networks. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9383–9387.
- Bascompte, J., Jordano, P., Olesen, J.M., 2006. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* 312, 431–433.
- Bastian, M., Heymann, S., 2009. Gephi: an open source software for exploring and manipulating networks. In: *International AAAI Conference on Weblogs and Social Media*.
- Beckerman, A.P., Petchey, O.L., Warren, P.H., 2006. Foraging biology predicts food web complexity. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13745–13749.
- Berry, D., Ben Mahfoudh, K., Wagner, M., Loy, A., 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Appl. Environ. Microbiol.* 77, 7846–7849.
- Berlow, E.L., Neutel, A.M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., Fox, J.W., Jansen, V.A., Jones, J.I., Kokkoris, G.D., Logofet, D.O., Mckane, A.J., Montoya, J.M., Petchey, O., 2004. Interaction strengths in food webs: issues and opportunities. *J. Anim. Ecol.* 73, 585–598.
- Cotton, T.E.A., Dumbrell, A.J., Helgason, T., 2014. What goes in must come out: testing for biases in molecular analysis of arbuscular mycorrhizal fungal communities. *PLoS One* 9, e109234.
- Derocles, S.A.P., Evans, D.M., Nichols, P.C., Evans, S.A., Lunt, D.H., 2015. Determining plant–leaf miner–parasitoid interactions: a DNA barcoding approach. *PLoS One* 10 (2), e0117872.
- Di Bella, J.M., Bao, Y., Gloor, G.B., Burton, J.P., Reid, G., 2013. High throughput sequencing methods and analysis for microbiome research. *J. Microbiol. Methods* 95, 401–414.

- Berendsen, R.L., Pieterse, C.M.J., Bakker, P.A.H.M., 2012. The rhizosphere microbiome and plant health. *Trends Plant Sci.* 17, 478–486.
- Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.-J., Taberlet, P., 2012. Tracking earthworm communities from soil DNA. *Mol. Ecol.* 21, 2017–2030.
- Blüthgen, N., Fründ, J., Vázquez, D.P., Menzel, F., Bluthgen, N., Frund, J., Vazquez, D.P., 2008. What do interaction network metrics tell us about specialization and biological traits? *Ecology* 89, 3387–3399.
- Boecklen, W.J., Yarnes, C.T., Cook, B.A., James, A.C., 2011. On the use of stable isotopes in trophic ecology. *Annu. Rev. Ecol. Evol. Syst.* 42, 411–440.
- Bohan, D.A., Caron-Lormier, G., Muggleton, S., Raybould, A., Tamaddoni-Nezhad, A., 2011. Automated discovery of food webs from ecological data using logic-based machine learning. *PLoS One* 6, e29028.
- Bohan, D.A., Raybould, A., Mulder, C., Woodward, G., Tamaddoni-nezhad, A., Blüthgen, N., Pocock, M.J.O., Muggleton, S., Evans, D.M., Astegiano, J., Massol, F., Loeuille, N., Petit, S., Macfadyen, S., 2013. Networking agroecology: integrating the diversity of agroecosystem interactions. *Adv. Ecol. Res.* 49, 1–67.
- Boyer, S., Wratten, S.D., Holyoake, A., Abdelkrim, J., Cruickshank, R.H., 2013. Using next-generation sequencing to analyse the diet of a highly endangered land snail (*Powelliphanta augusta*) feeding on endemic earthworms. *PLoS One* 8, e75962.
- Brown, D.S., Jarman, S.N., Symondson, W.O.C., 2012. Pyrosequencing of prey DNA in reptile faeces: analysis of earthworm consumption by slow worms. *Mol. Ecol. Resour.* 12, 259–266.
- Calderone, N.W., 2012. Insect pollinated crops, insect pollinators and US agriculture: trend analysis of aggregate data for the period 1992–2009. *PLoS One* 7, e37235.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunencko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336.
- Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S., Naeem, S., 2012. Biodiversity loss and its impact on humanity. *Nature* 486, 59–67.
- Carroll, S.P., Jørgensen, P.S., Michael, T., Bergstrom, C.T., Denison, R.F., Gluckman, P., Smith, T.B., Strauss, S.Y., Tabashnik, B.E., 2014. Applying evolutionary biology to address global challenges. *Science* 346, 1–16.
- Chaffron, S., Rehrauer, H., Pernthaler, J., Mering, C., 2010. A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res.* 2010, 947–959.
- Chakraborty, C., Doss, C.G.P., Patra, B.C., Bandyopadhyay, S., 2014. DNA barcoding to map the microbial communities: current advances and future directions. *Appl. Microbiol. Biotechnol.* 98, 3425–3436.
- Chakraborty, S., Newton, A.C., 2011. Climate change, plant diseases and food security: an overview. *Plant Pathol.* 60, 2–14.
- Chesson, P., 2000. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* 31, 343–366.
- Chow, C.E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., Fuhrman, J.A., 2014. Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J.* 8, 816–829.

- Clare, E.L., Fraser, E.E., Braid, H.E., Fenton, M.B., Hebert, P.D.N., 2009. Species on the menu of a generalist predator, the eastern red bat (*Lasiurus borealis*): using a molecular approach to detect arthropod prey. *Mol. Ecol.* 18, 2532–2542.
- Clarke, L.J., Czechowski, P., Soubrier, J., Stevens, M.I., Cooper, A., 2014. Modular tagging of amplicons using a single PCR for high-throughput sequencing. *Mol. Ecol. Resour.* 14, 117–121.
- Condon, M.A., Scheffer, S.J., Lewis, M.L., Wharton, R., Adams, D.C., Forbes, A.A., 2014. Lethal interactions between parasites and prey increase niche diversity in a tropical community. *Science* 343, 1240–1244.
- Cootes, A.P., Muggleton, S.H., Sternberg, M.J.E., 2003. The automatic discovery of structural principles describing protein fold space. *J. Mol. Biol.* 330, 839–850.
- Cordier, T., Robin, C., Capdevielle, X., Fabreguettes, O., Desprez-Loustau, M.-L., Vacher, C., 2012. The composition of phyllosphere fungal assemblages of European beech (*Fagus sylvatica*) varies significantly along an elevation gradient. *New Phytol.* 196, 510–519.
- Daly, R., Shen, Q., Aitken, S., 2011. Learning Bayesian networks: approaches and issues. *Knowl. Eng. Rev.* 26, 99–157.
- Dakos, V., Bascompte, J., 2014. Critical slowing down as early warning for the onset of collapse in mutualistic communities. *Proc. Natl. Acad. Sci.* 11, 17546–17551.
- Davey, J.S., Vaughan, I.P., Andrew King, R., Bell, J.R., Bohan, D.A., Bruford, M.W., Holland, J.M., Symondson, W.O.C., 2013. Intraguild predation in winter wheat: prey choice by a common epigeal carabid consuming spiders. *J. Appl. Ecol.* 50, 271–279.
- Deagle, B.E., Kirkwood, R., Jarman, S.N., 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Mol. Ecol.* 18, 2022–2038.
- Dean, T., Kanazawa, K., 1989. A model for reasoning about persistence and causation. *Comput. Intell.* 5, 142–150.
- Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., Zhou, J., 2012. Molecular ecological network analyses. *BMC Bioinformatics* 13, 113.
- DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072.
- Desprez-Loustau, M.-L., Aguayo, J., Dutech, C., Hayden, K.J., Husson, C., Jakushkin, B., Marçais, B., Piou, D., Robin, C., Vacher, C., 2015. An evolutionary ecology perspective to address forest pathology challenges of today and tomorrow. *Ann. For. Sci.* <http://dx.doi.org/10.1007/s13595-015-0487-4>.
- Dittami, S.M., Duboscq-Bidot, L., Perennou, M., Gobet, A., Corre, E., Boyen, C., Tonon, T., 2015. Host–microbe interactions as a driver of acclimation to salinity gradients in brown algal cultures. *ISME J.* <http://dx.doi.org/10.1038/ismej.2015.104>.
- Ducklow, H., 2008. Microbial services: challenges for microbial ecologists in a changing world. *Aquat. Microb. Ecol.* 53, 13–19.
- Duffy, J.E., Carinale, B.J., France, K.E., McIntyre, P.B., Thebault, E., Loreau, M., 2007. The functional role of biodiversity in ecosystems: incorporating trophic complexity. *Ecol. Lett.* 10, 522–538.
- Duran-Pinedo, A.E., Paster, B., Teles, R., Frias-Lopez, J., 2011. Correlation network analysis applied to complex biofilm communities. *PLoS One* 6, e28438.
- Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998.
- Eveleigh, E.S., McCann, K.S., McCarthy, P.C., Pollock, S.J., Lucarotti, C.J., Morin, B., McDougall, G.A., Strongman, D.B., Huber, J.T., Umbanhowar, J., Faria, L.D.B., 2007. Fluctuations in density of an outbreak species drive diversity cascades in food webs. *Proc. Natl. Acad. Sci.* 104, 16976–16981.

- Faisal, A., Dondelinger, F., Husmeier, D., Beale, C.M., 2010. Inferring species interaction networks from species abundance data: a comparative evaluation of various statistical and machine learning methods. *Ecol. Inform.* 5, 451–464.
- Faust, K., Lahti, L., Gonze, D., de Vos, W.M., Raes, J., 2015. Metagenomics meets time series analysis: unraveling microbial community dynamics. *Curr. Opin. Microbiol.* 25, 56–66.
- Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550.
- Fontaine, C., Guimaraes, P.R., Kefi, S., Loeuille, N., Memmott, J., van der Putten, W.H., van Veen, F.J.F., Thebault, E., 2011. The ecological and evolutionary implications of merging different types of networks. *Ecol. Lett.* 14, 1170–1181.
- Fortuna, M.A., Stouffer, D.B., Olesen, J.M., Jordano, P., Mouillot, D., Krasnov, B.R., Poulin, R., Bascompte, J., 2010. Nestedness versus modularity in ecological networks: two sides of the same coin? *J. Anim. Ecol.* 79, 811–817.
- Fortunato, S., 2010. Community detection in graphs. *Phys. Rep.* 486, 75–174.
- Foster, K.R., Bell, T., 2012. Competition, not cooperation, dominates interactions among culturable microbial species. *Curr. Biol.* 22, 1845–1850.
- Frey-Klett, P., Burlinson, P., Deveau, A., Barret, M., Tarkka, M., Sarniguet, A., 2011. Bacterial–fungal interactions: hyphens between agricultural, clinical, environmental, and food microbiologists. *Microbiol. Mol. Biol. Rev.* 75, 583–609.
- Friedman, J., Alm, E.J., 2012. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* 8, e1002687.
- Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441.
- Friedman, N., Murphy, K., Russell, S., 1998. Learning the structure of dynamic probabilistic networks. In: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, pp. 139–147.
- Fuhrman, J.A., 2009. Microbial community structure and its functional implications. *Nature* 459, 193–199.
- Fuhrman, J.A., Cram, J.A., Needham, D.M., 2015. Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* 13, 133–146.
- Gaston, K.J., 2010. Ecology. Valuing common species. *Science* 327, 154–155.
- Gaspar, J.M., Thomas, W.K., 2013. Assessing the consequences of denoising marker-based metagenomic data. *PLoS One* 8, e60458.
- Gilbert, J.A., Neufeld, J.D., 2014. Life in a world without microbes. *PLoS Biol.* 12, 1–3.
- Gilles, A.A., Meglec, E., Pech, N., Ferreira, S.S., Malausa, T., Martin, J.-F., 2011. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics* 12, 245.
- Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R., Reid, G., 2010. Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One* 5, e15406.
- Gonzalez, J.M., Portillo, M.C., Belda-Ferre, P., Mira, A., 2012. Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. *PLoS One* 7, e29973.
- Haddad, N.M., Crutsinger, G.M., Gross, K., Haarstad, J., Tilman, D., 2011. Plant diversity and the stability of foodwebs. *Ecol. Lett.* 14, 42–46.
- Haegeman, B., Hamelin, J., Moriarty, J., Neal, P., Dushoff, J., Weitz, J.S., 2013. Robust estimation of microbial diversity in theory and in practice. *ISME J.* 7, 1092–1101.
- Hagen, M., Kissling, W.D., Rasmussen, C., De Aguiar, M.A.M., Brown, L.E., Carstensen, D.W., Alves-Dos-Santos, I., Dupont, Y.L., Edwards, F.K., Genini, J., Guimarães, P.R., Jenkins, G.B., Jordano, P., Kaiser-Bunbury, C.N., Ledger, M.E., Maia, K.P., Marquitti, F.M.D., McLaughlin, Ó., Morellato, L.P.C., O’Gorman, E.J.,

- Trøjelsgaard, K., Tylianakis, J.M., Vidal, M.M., Woodward, G., Olesen, J.M., 2012. Biodiversity, species interactions and ecological networks in a fragmented world. *Adv. Ecol. Res.* 46, 89–120.
- H Hairston, N.G., Ellner, S.P., Geber, M.A., Yoshida, T., Fox, J.A., 2005. Rapid evolution and the convergence of ecological and evolutionary time. *Ecol. Lett.* 8, 1114–1127.
- Haruta, S., Kato, S., Yamamoto, K., Igarashi, Y., 2009. Intertwined interspecies relationships: approaches to untangle the microbial network. *Environ. Microbiol.* 11, 2963–2969.
- Hector, A., Hooper, R., 2002. Ecological experiment. *Science* 295, 639–640.
- Heleno, R., Devoto, M., Pocock, M., 2012. Connectance of species interaction networks and conservation value: is it any good to be well connected? *Ecol. Indic.* 14, 7–10.
- Heleno, R.H., Ceia, R.S., Ramos, J.A., Memmott, J., 2009. Effects of alien plants on insect abundance and biomass: a food-web approach. *Conserv. Biol.* 23, 410–419.
- Hengeveld, R., 1980. Polyphagy, oligophagy and food specialization in ground beetles (Coleoptera, Carabidae). *Neth. J. Zool.* 30, 564–584.
- Hibbett, D.S., Ohman, A., Kirk, P.M., 2009. Fungal ecology catches fire. *New Phytol.* 184, 279–282.
- Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B., 2010. Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* 8, 15–25.
- Hudson, P.J., Dobson, A.P., Lafferty, K.D., 2006. Is a healthy ecosystem one that is rich in parasites? *Trends Ecol. Evol.* 21, 381–385.
- Ibanez, S., Manneville, O., Miquel, C., Taberlet, P., Valentini, A., Aubert, S., Coissac, E., Colace, M.P., Duparc, Q., Lavorel, S., Moretti, M., 2013. Plant functional traits reveal the relative contribution of habitat and food preferences to the diet of grasshoppers. *Oecologia* 173, 1459–1470.
- Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., Veen, V., Frank, F.J., Warren, P.H., Woodward, G., 2009. Ecological networks—beyond food webs. *J. Anim. Ecol.* 78, 253–269.
- Isbell, F., Tilman, D., Polasky, S., Binder, S., Hawthorne, P., 2013. Low biodiversity state persists two decades after cessation of nutrient enrichment. *Ecol. Lett.* 16, 454–460.
- Janson, E.M., Stireman, J.O., Singer, M.S., Abbot, P., 2008. Phytophagous insect–microbe mutualisms and adaptive evolutionary diversification. *Evolution* 62, 997–1012.
- Jensen, F.V., Nielsen, T.D., 2007. *Bayesian Networks and Decision Graphs*. Springer, New York.
- Joppa, L.N., Montoya, J.M., Sole, R., Sanderson, J., Pimm, S.L., 2010. On nestedness in ecological networks. *Evol. Ecol. Res.* 12, 35–46.
- Kaartinen, R., Stone, G.N., Hearn, J., Lohse, K., Roslin, T., 2010. Revealing secret liaisons: DNA barcoding changes our understanding of food webs. *Ecol. Entomol.* 35, 623–638.
- Kara, E.L., Hanson, P.C., Hu, Y.H., Winslow, L., McMahon, K.D., 2013. A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA. *ISME J.* 7, 680–684.
- Kartzinel, T.R., Chen, P.A., Coverdale, T.C., Erickson, D.L., Kress, W.J., Kuzmina, M.L., Rubenstein, D.I., Wang, W., Pringle, R.M., 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proc. Natl. Acad. Sci.* 26, 819–824.
- Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., Boit, A., Joppa, L.N., Lafferty, K.D., Williams, R.J., Martinez, N.D., Menge, B.A., Blanchette, C.A., Iles, A.C., Brose, U., 2012. More than a meal ... integrating non-feeding interactions into food webs. *Ecol. Lett.* 15, 291–300.

- Kemen, E., 2014. Microbe–microbe interactions determine oomycete and fungal host colonization. *Curr. Opin. Plant Biol.* 20, 75–81.
- King, R.D., Muggleton, S.H., Srinivasan, A., Sternberg, M.J., 1996. Structure–activity relationships derived by machine learning: the use of atoms and their bond connectivities to predict mutagenicity by inductive logic programming. *Proc. Natl. Acad. Sci.* 93, 438–442.
- Köljalg, U., Larsson, K.-H., Abarenkov, K., Nilsson, R.H., Alexander, I.J., Eberhardt, U., Erland, S., Høiland, K., Kjoller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A.F.S., Tedersoo, L., Vrålstad, T., Ursing, B.M., 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytol.* 166, 1063–1068.
- Kondoh, M., 2003. Foraging adaptation and the relationship between food-web complexity and stability. *Science* 299, 1388–1391.
- Kromp, B., 1999. Carabid beetles in sustainable agriculture: a review on pest control efficacy, cultivation impacts and enhancement. *Agric. Ecosyst. Environ.* 74, 187–228.
- Kurtz, Z.D., Mueller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R.A., 2015. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput. Biol.* 11 (5), e1004226.
- Lafferty, K.D., Allesina, S., Arim, M., Briggs, C.J., De Leo, G., Dobson, A.P., Dunne, J.A., Johnson, P.T.J., Kuris, A.M., Marcogliese, D.J., Martinez, N.D., Memmott, J., Marquet, P.A., McLaughlin, J.P., Mordecai, E.A., Pascual, M., Poulin, R., Thieltges, D.W., 2008. Parasites in food webs: the ultimate missing links. *Ecol. Lett.* 11, 533–546.
- Laliberté, E., Tylianakis, J.M., 2010. Deforestation homogenizes tropical parasitoid–host networks. *Ecology* 91, 1740–1747.
- Lancichinetti, A., Fortunato, S., 2009. Community detection algorithms: a comparative analysis. *Phys. Rev. E* 80, 056117.
- Lauritzen, S.L., 1996. *Graphical Models*. Oxford Science Publications, New York, NY.
- Layer, K., Hildrew, A.G., Jenkins, G.B., Riede, J., Rossiter, S.J., Townsend, C.R., Woodward, G., 2011. Long-term dynamics of a well-characterised food web: four decades of acidification and recovery in the Broadstone Stream model system. *Adv. Ecol. Res.* 44, 69–117.
- Layer, K., Riede, J.O., Hildrew, A.G., Woodward, G., 2010. Food web structure and stability in 20 streams across a wide pH gradient. *Adv. Ecol. Res.* 42, 265–299.
- Lee, C.K., Herbold, C.W., Polson, S.W., Wommack, K.E., Williamson, S.J., McDonald, I.R., Cary, S.C., 2012. Groundtruthing next-gen sequencing for microbial ecology—biases and errors in community structure estimates from PCR amplicon pyrosequencing. *PLoS One* 7, e44224.
- Leger, J.-B., Vacher, C., Daudin, J., 2015. Clustering methods differ in their ability to detect patterns in species interaction networks. *Methods Ecol. Evol.* 6, 474–481.
- Leger, J.B., Vacher, C., Daudin, J.J., 2014. Detection of structurally homogeneous subsets in graphs. *Stat. Comput.* 24, 675–692.
- Levine, J.M., HilleRisLambers, J., 2009. The importance of niches for the maintenance of species diversity. *Nature* 461, 254–257.
- Lewinsohn, T.M., Prado, P.I., Jordano, P., Bascompte, J., Olesen, J.M., 2006. Structure in plant–animal interaction assemblages. *Oikos* 113, 174–184.
- Liu, H., Laferty, J., Wasserman, L., 2009. The nonparanormal: semiparametric estimation of high dimensional undirected graphs. *J. Mach. Learn. Res.* 10, 2295–2328.
- Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., Law, M., 2012. Comparison of next-generation sequencing systems. *J. Biomed. Biotechnol.* 2012, 1–11.
- Loeuille, N., 2010a. Consequences of adaptive foraging in diverse communities. *Funct. Ecol.* 24, 18–27.

- Loeuille, N., 2010b. Influence of evolution on the stability of ecological communities. *Ecol. Lett.* 13, 1536–1545.
- Loeuille, N., Loreau, M., 2006. Evolution of body size in food webs: does the energetic equivalence rule hold? *Ecol. Lett.* 9, 171–178.
- Loeuille, N., Loreau, M., Ferrière, R., 2002. Consequences of plant–herbivore coevolution on the dynamics and functioning of ecosystems. *J. Theor. Biol.* 217, 369–381.
- Lok, C., 2015. Mining the microbial dark matter. *Nature* 522, 270–273.
- Lopezaraiza-Mikel, M.E., Hayes, R.B., Whalley, M.R., Memmott, J., 2007. The impact of an alien plant on a native plant–pollinator network: an experimental approach. *Ecol. Lett.* 10, 539–550.
- Loreau, M., Hector, A., 2001. Partitioning selection and complementarity in biodiversity experiments. *Nature* 412, 72–76.
- Loreau, M., de Mazancourt, C., 2013. Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.* 16, 106–115.
- Loreau, M., Mouquet, N., Gonzalez, A., 2003. Biodiversity as spatial insurance in heterogeneous landscapes. *Proc. Natl. Acad. Sci.* 100, 12765–12770.
- MacArthur, R.H., 1958. Population ecology of some warblers of Northeastern coniferous forests. *Ecology* 39, 599–618.
- Marbach, D., Costello, J.C., Küffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., Kellis, M., Collins, J.J., Stolovitzky, G., 2012. Wisdom of crowds for robust gene network inference. *Nat. Methods* 9, 796–804.
- May, R.M., 1973. Stability and complexity in model ecosystems. *Monogr. Popul. Biol.* 6, 1–235.
- Meinshausen, N., Bühlmann, P., 2006. High-dimensional graphs and variable selection with the Lasso. *Ann. Stat.* 34, 1436–1462.
- Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well-Being: Synthesis*. Island Press, Washington, DC.
- Milns, I., Beale, C.M., Smith, V.A., 2010. Revealing ecological networks using Bayesian network inference algorithms. *Ecology* 91, 1892–1899.
- Moalic, Y., Desbruyères, D., Duarte, C.M., Rozenfeld, A.F., Bachraty, C., Arnaud-Haond, S., 2012. Biogeography revisited with network theory: retracing the history of hydrothermal vent communities. *Syst. Biol.* 61, 127–137.
- Mollet, G., Duyck, P.-F., Lefeuvre, P., Lescourret, F., Martin, J.-F., Piry, S., Canard, E., Tixier, P., 2014. Cover cropping alters the diet of arthropods in a banana plantation: a metabarcoding approach. *PLoS One* 9, e93740.
- Montoya, J.M., Pimm, S.L., Sole, R.V., 2006. Ecological networks and their fragility. *Nature* 442, 259–264.
- Moore, J.W., Semmens, B.X., 2008. Incorporating uncertainty and prior information into stable isotope mixing models. *Ecol. Lett.* 11, 470–480.
- Mougou, A., Dutech, C., Desprez-Loustau, M.-L., 2008. New insights into the identity and origin of the causal agent of oak powdery mildew in Europe. *For. Pathol.* 38, 275–287.
- Muggleton, S., 1991. Inductive logic programming. *New Generat. Comput.* 8, 295–318.
- Muggleton, S.H., Lin, D., Pahlavi, N., Tamaddoni-Nezhad, A., 2014. Meta-interpretive learning: application to grammatical inference. *Mach. Learn.* 94, 25–49.
- Müller, C.B., Adriaanse, I.C.T., Belshaw, R., Godfray, H.C.J., 1999. The structure of an aphid–parasitoid community. *J. Anim. Ecol.* 68, 346–370.
- Naeem, S., Bunker, D.E., Hector, A., Loreau, M., Perrings, C., 2009. *Biodiversity, Ecosystem Functioning, and Human Wellbeing—An Ecological and Economic Perspective*. Oxford University Press, Oxford, UK.
- Navarrete, A.A., Tsai, S.M., Mendes, L.W., Faust, K., de Hollander, M., Cassman, N.A., Raes, J., van Veen, J.A., Kuramae, E.E., 2015. Soil microbiome responses to the short-term effects of Amazonian deforestation. *Mol. Ecol.* 24 (10), 2433–2448.

- Nuismer, S.L., Jordano, P., Bascompte, J., 2013. Coevolution and the architecture of mutualistic networks. *Evolution* 67, 338–354.
- Öpik, M., Davison, J., Moora, M., Zobel, M., 2014. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. *Botany* 92, 135–147.
- Ovaskainen, O., Hottola, J., Siitonen, J., 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology* 91, 2514–2521.
- Patin, N.V., Kunin, V., Lidström, U., Ashby, M.N., 2013. Effects of OTU Clustering and PCR Artifacts on Microbial Diversity Estimates. *Microb. Ecol.* 65, 709–719.
- Paula, D.P., Linard, B., Andow, D.A., Sujii, E.R., Pires, C.S.S., Vogler, A.P., 2015. Detection and decay rates of prey and prey symbionts in the gut of a predator through metagenomics. *Mol. Ecol. Resour.* 15, 880–892.
- Penton, C.R., Gupta, V.V.S.R., Tiedje, J.M., Neate, S.M., Ophel-Keller, K., Gillings, M., Harvey, P., Pham, A., Roget, D.K., 2014. Fungal community structure in disease suppressive soils assessed by 28S LSU gene sequencing. *PLoS One* 9, e93893.
- Pinto, A.J., Raskin, L., 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS One* 7, e43093.
- Pisanu, B., Caut, S., Gutjahr, S., Vernon, P., Chapuis, J.L., 2011. Introduced black rats *Rattus rattus* on Ile de la Possession (Iles Crozet, subantarctic): diet and trophic position in food webs. *Polar Biol.* 34, 169–180.
- Pocock, M.J.O., Evans, D.M., Memmott, J., 2012. The robustness and restoration of a network of ecological networks. *Science* 335, 973–977.
- Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N., Taberlet, P., 2012. Who is eating what: diet assessment using next generation sequencing. *Mol. Ecol.* 21, 1931–1950.
- Quail, M., Smith, M.E., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., 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.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596.
- Quéméré, E., Hibert, F., Miquel, C., Lhuillier, E., Rasolondraibe, E., Champeau, J., Rabarivola, C., Nusbaumer, L., Chatelain, C., Gautier, L., Ranirison, P., Crouau-Roy, B., Taberlet, P., Chikhi, L., 2013. A DNA metabarcoding study of a primate dietary diversity and plasticity across its entire fragmented range. *PLoS One* 8, e58971.
- Raffaelli, D., White, P.C.L., 2013. Ecosystems and their services in a changing world. An ecological perspective. *Adv. Ecol. Res.* 48, 1–70.
- Raffaelli, D.G., Bullock, J.M., Cinderby, S., Durance, I., Emmett, B., Harris, J., Hicks, K., Oliver, T.H., Paterson, D., White, P.C.L., 2014. *Adv. Ecol. Res.* 51, 41–77.
- Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: the barcode of life data system. *Mol. Ecol. Notes* 7, 355–364.
- Ravikumar, P., Wainwright, M.J., Raskutti, G., Yu, B., 2011. High-dimensional covariance estimation by minimizing ℓ_1 -penalized log-determinant divergence. *Electron. J. Stat.* 5, 935–980.
- Reeder, J., Knight, R., 2011. Rapid denoising of pyrosequencing amplicon data: exploiting the rank-abundance distribution. *Nat. Methods* 7, 668–669.
- Robinson, K.M., Hauzy, C., Loeuille, N., Albrechtsen, B.R., 2015. Relative impacts of environmental variation and evolutionary history on the nestedness and modularity of tree–herbivore networks. *Ecol. Evol.* 5 (14), 2898–2915.

- Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009. A safe operating space for humanity. *Nature* 461, 472–475.
- Rohr, R.P., Saavedra, S., Bascompte, J., 2014. On the structural stability of mutualistic systems. *Science* 345, 1253–1257.
- Rosenberg, E., Koren, O., Reshef, L., Efrony, R., Zilber-Rosenberg, I., 2007. The role of microorganisms in coral health, disease and evolution. *Nat. Rev. Micro.* 5 (5), 355–362.
- Rougerie, R., Smith, M.A., Fernandez-Triana, J., Lopez-Vaamonde, C., Ratnasingham, S., Hebert, P.D.N., 2011. Molecular analysis of parasitoid linkages (MAPL): gut contents of adult parasitoid wasps reveal larval host. *Mol. Ecol.* 20, 179–186.
- Saavedra, S., Stouffer, D.B., Uzzi, B., Bascompte, J., 2011. Strong contributors to network persistence are the most vulnerable to extinction. *Nature* 478, 233–235.
- Sapkota, R., Knorr, K., Jørgensen, L.N., O’Hanlon, K.A., Nicolaisen, M., 2015. Host genotype is an important determinant of the cereal phyllosphere mycobiome. *New Phytol.* 207, 1134–1144.
- Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. *Nature* 413, 591–596.
- Schlaeppli, K., Bulgarelli, D., 2015. The plant microbiome at work. *MPMI* 28, 212–217.
- Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Bolchacova, E., Voigt, K., Crous, P.W., Miller, A.N., Wingfield, M.J., Aime, M.C., An, K.-D., Bai, F.-Y., Barreto, R.W., Begerow, D., Bergeron, M.-J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, A.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, A., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Dieguez-Uribeondo, J., Divakar, P.K., Douglas, B., Duenas, M., Duong, T.A., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., Garcia, M.A., Ge, Z.-W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.-D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.-M., Hoffmann, K., Hofstetter, V., Hognabba, F., Hollingsworth, P.M., Hong, S.-B., Hosaka, K., Houbraken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Koljalg, U., Kovacs, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, A.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, A.R., Methven, A.S., Meyer, W., Moncalvo, J.-M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.A., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramirez, J.M., Schmitt, I., Schussler, A., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.-O., Sung, G.-H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.A., Urbina, H., Vagvolgyi, C., Vialle, A., Vu, T.D., Walther, G., Wang, Q.-M., Wang, Y., Weir, B.S., Weiss, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, A., Zamora, J.-C., Zhang, N., Zhuang, W.-Y., Schindel, D., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci.* 109, 6241–6246.
- Schwaller, L., Robin, S., Stumpf, M., 2015. Bayesian Inference of Graphical Model Structures Using Trees. [arXiv:1504.02723](https://arxiv.org/abs/1504.02723).

- Shehzad, W., Riaz, T., Nawaz, M.A., Miquel, C., Poillot, C., Shah, S.A., Pompanon, F., Coissac, E., Taberlet, P., 2012. Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Mol. Ecol.* 21, 1951–1965.
- Sher, D., Thompson, J.W., Kashtan, N., Croal, L., Chisholm, S.W., 2011. Response of *Prochlorococcus* ecotypes to co-culture with diverse marine bacteria. *ISME J.* 5, 1125–1132.
- Smith, M.A., Eveleigh, E.S., McCann, K.S., Merilo, M.T., McCarthy, P.C., Van Rooyen, K.I., 2011. Barcoding a quantified food web: crypsis, concepts, ecology and hypotheses. *PLoS One* 6, e14424.
- Smith, M.A., Rodriguez, J.J., Whitfield, J.B., Deans, A.R., Janzen, D.H., Hallwachs, W., Hebert, P.D.N., 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc. Natl. Acad. Sci.* 105, 12359–12364.
- Srinivasan, A., Page, D., Camacho, R., King, R., 2006. Quantitative pharmacophore models with inductive logic programming. *Mach. Learn.* 64, 65–90.
- Srivathsan, A., Sha, J.C.M., Vogler, A.P., Meier, R., 2014. Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Mol. Ecol. Resour.* 15, 250–261.
- Staudacher, K., Jonsson, M., Traugott, M., 2015. Diagnostic PCR assays to unravel food web interactions in cereal crops with focus on biological control of aphids. *J. Pest Sci.* <http://dx.doi.org/10.1007/s10340-015-0685-8>.
- Stauffer, D., Aharony, A., 1994. *Introduction to Percolation Theory*. CRC Press, Taylor & Francis Group, London.
- Stockinger, H., Krüger, M., Schüssler, A., 2010. DNA barcoding of arbuscular mycorrhizal fungi. *New Phytol.* 187, 461–474.
- Sun, D.-L., Jiang, X., Wu, Q.L., Zhou, N.-Y., 2013. Intragenomic heterogeneity in 16S rRNA genes causes overestimation of prokaryotic diversity. *Appl. Environ. Microbiol.* 79, 5787.
- Symondson, W.O.C., 2002. Molecular identification of prey in predator diets. *Mol. Ecol.* 11, 627–641.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* 21, 2045–2050.
- Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35, e14.
- Tamaddoni-Nezhad, A., Bohan, D., Raybould, A., Muggleton, S.H., 2012. Machine learning a probabilistic network of ecological interactions. In: *Proceedings of the 21st International Conference on Inductive Logic Programming, LNAI 7207*, Springer, Berlin, pp. 332–346.
- Tamaddoni-Nezhad, A., Bohan, D.A., Raybould, A., Muggleton, S., 2015. Towards machine learning of predictive models from ecological data. In: *Proceedings of the International Conference on Inductive Logic Programming*, Springer, Berlin.
- Tamaddoni-Nezhad, A., Milani, G.A., Raybould, A., Muggleton, S., Bohan, D.A., 2013. Construction and validation of food webs using logic-based machine learning and text mining. *Adv. Ecol. Res.* 49, 225–289.
- Thébault, E., Fontaine, C., 2010. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* 329, 853–856.
- Thébault, E., Loreau, M., 2006. The relationship between biodiversity and ecosystem functioning in food webs. *Ecol. Res.* 21, 17–25.
- Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O., Hladysz, S., Kitching, R.L., Martinez, N.D., Rantala, H., Romanuk, T.N., Stouffer, D.B., Tylianakis, J.M.,

2012. Food webs: reconciling the structure and function of biodiversity. *Trends Ecol. Evol.* 27, 689–697.
- Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc. Ser. B* 73, 273–282.
- Tilman, D., Lehman, C.L., Bristow, C.E., 1998. Diversity–stability relationships: statistical inevitability or ecological consequence? *Am. Nat.* 151, 277–282.
- Tilman, D., Reich, P.B., Knops, J.M.H., 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441, 629–632.
- Todorovski, L., Džeroski, S., Kompare, B., 1998. Modelling and prediction of phytoplankton growth with equation discovery. *Ecol. Modell.* 113, 71–81.
- Traugott, M., Kamenova, S., Ruess, L., 2013. Empirically characterising trophic networks: what emerging DNA-based methods, stable isotope and fatty acid analyses can offer. *Adv. Ecol. Res.* 49, 177–224.
- Tylianakis, J.M., Tscharntke, T., Lewis, O.T., 2007. Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature* 445, 202–205.
- U’Ren, J.M., Riddle, J.M., Monacell, J.T., Carbone, I., Miadlikowska, J., Arnold, A.E., 2014. Tissue storage and primer selection influence pyrosequencing-based inferences of diversity and community composition of endolichenic and endophytic fungi. *Mol. Ecol. Resour.* 14, 1032–1048.
- Vacher, C., Daudin, J.-J., Piou, D., Desprez-Loustau, M.-L., 2010. Ecological integration of alien species into a tree–parasitic fungus network. *Biol. Invasions* 12, 3249–3259.
- Vacher, C., Piou, D., Desprez-Loustau, M.-L., 2008. Architecture of an antagonistic tree/fungus network: the asymmetric influence of past evolutionary history. *PLoS One* 3, e1740.
- Vanderklift, M.A., Ponsard, S., 2003. Sources of variation in consumer–diet? ^{15}N enrichment: a meta-analysis. *Oecologia* 136, 169–182.
- Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.-F., Jacques, M.-A., Lebrun, M.-H., Le Loir, Y., Ogliastro, M., Petit, M.-A., Roumagnac, P., Candresse, T., 2014. Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Front. Cell. Infect. Microbiol.* 4, 29.
- Whittaker, J., 1990. *Graphical Models in Applied Multivariate Statistics*. Wiley Publishing, New York.
- Wirta, H., Hebert, P.D.N., Kaartinen, R., Prosser, S., Várkonyi, G., Roslin, T., 2014. Complementary molecular information changes our perception of food web structure. *Proc. Natl. Acad. Sci.* 111, 1885–1890.
- Yachi, S., Loreau, M., 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc. Natl. Acad. Sci.* 96, 1463–1468.
- Yang, E., Ravikumar, P., Allen, G.I., Liu, Z., 2013. On Poisson graphical models. *Adv. Neural Inform. Process. Syst.* 26, 1718–1726.
- Zilber-Rosenberg, I., Rosenberg, E., 2008. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol. Rev.* 32, 723–735.

6.3 Appendix C - Résumé en français

Titre de thèse: "Architecture génétique des interactions entre le chêne pédonculé (*Quercus robur* L.) et les communautés microbiennes de sa phyllosphère."

Les micro-organismes qui habitent la surface et l'intérieur des tissus des parties aériennes des plantes constituent la phyllosphère. Ils influencent de nombreux traits des plantes, interviennent comme intermédiaire dans les interactions biotiques, les flux d'énergie, modulent les réactions aux stress de leurs hôtes et donc *in fine* influencent la fitness des plantes qu'ils colonisent. L'identité génétique de l'hôte est un facteur important qui structure les communautés microbiennes en interactions avec la plante. Cet effet est la conséquence des variations des traits fonctionnels comme la productivité (contenu en eau et en phosphore des tissus, structure et composition des feuilles, densité du bois etc...) ou des traits liés aux défenses des plantes contre les infections (contenu en composés phénoliques, densité des trichomes, voie de signalisation à l'éthylène et composition des cires de la cuticule). Ses traits quantitatifs étant sous contrôle génétique, des QTLs (Quantitative Traits Loci) des plantes dont les effets s'étendent au niveau des communautés sont à rechercher, en accord avec le cadre conceptuel de la génétique des communautés.

Le premier chapitre présente le concept de génétique des communautés suivi de la présentation de la phyllosphère comme habitat microbien et se termine par un zoom sur un de ses habitants, *Erysiphe alphitoides*, l'agent responsable du mildiou du chêne, une maladie commune du chêne pédonculé (*Quercus robur*), notre modèle d'étude. La technique de metabarcoding pour l'étude des communautés microbiennes est également présentée et détaillée dans ce chapitre.

Le chapitre suivant présente les résultats d'une expérience décrivant les communautés bactérienne et fongique du chêne pédonculé à l'aide des techniques de séquençage de deuxième génération (NGS). Ses communautés sont étudiées à l'aide de la construction de réseaux d'association avec une attention portée à l'influence du mildiou sur l'architecture de ce réseau. L'étude du pathobiome (le pathobiome intègre l'agent pathogène dans son environnement biotique) à l'aide des technologies de NGS se fonde sur la bonne représentation de l'intensité de l'infection par le nombre de séquences du pathogène détectées par ses techniques dans l'échantillon, ce que montre nos résultats. Nos résultats montrent que la composition des communautés est affectée par le niveau d'infection du pathogène, de façon très marquée chez les levures. Treize Unités Taxonomiques (OTU – Operational Taxonomic Unit) fongiques et treize Unités Taxonomiques bactériennes ont été identifiées comme étant en interaction directe avec l'agent pathogène. Les endophytes fongiques *Mycosphaerella punctiformis* et *Monochaetia kansensis* sont d'ailleurs de possible antagoniste à *E.*

alphotoides. D'une façon générale les interactions positives de type mutualiste, facilitatrices et commensales dominant au sein de la phyllosphère. Ses résultats contrastent avec les précédents résultats montrant la compétition comme le principal type d'interaction au sein de ce système.

Le troisième chapitre présente les résultats de la recherche de QTL chez le chêne liés aux variations de communautés de sa phyllosphère. Les communautés fongique et bactérienne ont ainsi été analysées à l'aide des NGS sur une collection de 180 arbres issus d'un croisement contrôlé dédié à la cartographie génétique. Huit QTLs de traits des communautés microbiennes ont ainsi pu être identifiés. Ses traits sont la composition des communautés microbiennes et fongiques, la diversité fongique, ainsi que le pourcentage de séquences attribuées à *E. alphotoides*. Trois de ses QTLs colocalisent avec des QTLs liés à la sensibilité de l'arbre à l'infection par ce dernier agent pathogène, suggérant un fort déterminisme génétique de la résistance du chêne à cet agent. Ses résultats sont discutés à la lumière des traits fonctionnels potentiellement impliqués dans la modification de structure des communautés microbiennes de la phyllosphère.

La quatrième chapitre présente l'utilisation d'une méthode de génétique d'association (GWA – Genome Wide Association study) comme méthode alternative et complémentaire à la précédente recherche de QTLs.

Le cinquième et dernier chapitre résume les résultats obtenus et les mets en perspective dans un cadre d'écologie évolutive.