The association between overlooked microbial eukaryotes and plant holobiont: functionality and diversity

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Alfredo Mari

aus Pescara, Italien

Köln, 2018

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung für Pflanze-Mikroben Interaktionen (Direktor: Prof. Dr. Paul Schulze-Lefert) angefertigt.

MAX-PLANCK-GESELLSCHAFT

Max-Planck-Institut für Pflanzenzüchtungsforschung

Berichterstatter:	Prof. Dr. Gunther Döhlemann
	Prof. Dr. Eric Kemen
Schriftführer:	Dr. Korbinian Schneeberger
Prüfungsvorsitzender:	Prof. Dr. Achim Tresch
Tag der Disputation:	23.07.2018

"I got power, poison, pain and joy inside my DNA,

I got hustle though, ambition flow inside my DNA"

Kendrick Lamar, Pulitzer prize 2017. DNA

VI

Table of Contents

II

Publications:

Agler, M.T., **Mari, A.,** Almario J., Dombrovski, N., Hacquard, S. & Kemen, E.M. (2018) Accurate systems biology-based analyses in multi-kingdom microbiome studies by overhauling amplicon sequencing and data analysis; Methods in Ecology and Evolution. *In revision*

Mari,A., Agler.M.T., Roux,F., Alonso-Blanco,C., Ågren,J., Durán-Ballesteros,P., Almario,J. & Kemen, E.M. (2018) Microalgae shape bacterial and fungal diversity of plant holobiont. *In preparation*

List of tables

List of Figures

VI

List of Abbreviations and Acronyms

-Brac	Network inferred from samples not containing *Bracteacoccus sp.*
%	Percent
+Brac	Network inferred from samples containing *Bracteacoccus sp.*
3N-BBM+V	Triple nitrogen-bold basal medium plus vitamin b1-12
ANOVA	Analysis of variance
bp	base pairs
BV3	Bacterial 16S, V3 region
BV5	Bacterial 16S, V5 region
CDC	Ciruelos de Coca, Spain (sampling site)
CH	Chemin de Garaut, France (sampling site)
cm	Centimeter
Col	Columbia
CoNET	Co-occurrence Network Inference
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
e.g	Exempli gratia
EDTA	Etylenediaminotetraacetic acid
EPS	Extracellular polymeric substances
ERG	Ergenzingen, Germany (sampling site)
EY	Eyach, Germany (sampling site)
FITS2	Fungal ITS2 region
Ftrad	Fungal ITS1 region
Fwd	Forward
gDNA	genomic DNA
GPS	Global positioning system
h	Hours
IAA	Indole-3-acetic acid
ITS	Internal transcribed spacer
JUG	Jugendherberge, (Tübingen), Germany (sampling site)
K6	Kirchenstellinsfurter Str., Germany (sampling site)

VIII

KFA	Kramfors A, Sweden (sampling site)
KFB	Kramfors B, Sweden (sampling site)
KFC	Kramfors C, Sweden, (sampling site)
Ksk	Keskwick
LA	Lanta, France (sampling site)
LB	Lysogeny broth
M	Molar
MAR	Marjaliza, Spain (sampling site)
masl	Meters above the sea level
mg	Milligram
min	Minutes
mL	Milliliters
mm	Millimeter
mM	Millimolar
MPIPZ	Max Planck Institute for Plant Breeding Research
ng	Nanogram
nm	Nanometers
NMAS	Neff modified amoebae salinae
ºC	Celsius degree
OITS2	Oomycetal ITS2 region
Otrad	Oomycetal ITS1 region
OTU	Operational Taxonomic Unit
PCR	Polymerase chain reaction
PDA	Potato dextrose agar medium
PERMANOVA	Permutated multivariate ANOVA
PFN	Pfrondorf, Germany (sampling site)
pH	Negative decimal logarithm of H^+ concentration
PV4	Eukaryotic 18S, V4-5 region
PV9	Eukaryotic 18S, V9 region
qPCR	Quantitative polymerase chain reaction
rDNA	Ribosomal DNA
Rev	Reverse

RNA	Ribonucleic acid
rpm	round per minute
SACf	*Saccharomyces boulardii* qPCR primer forward
SACr	*Saccharomyces boulardii* qPCR primer reverse
sec	Seconds
Sf	San feliu
SLY	San Leonardo de Yagüe, Spain (sampling site)
sp.	Species
Taq	*Thermophilus aquaticus*
UPP	Uppsala, Sweden (sampling site)
UV	Ultra-violet
WH	Wendelsheim, Germany (sampling site)
Ws	Wassilewskjia
µg	Microgram
µl	Microliters
µM	Micromolar

Abstract:

The plant holobiont is a relatively recent term in biology, introduced to address the assembly composed by the plant itself and its associated microbes. In order to address evolutionary selection within this frame, it becomes necessary to study plant-associated microbes, or microbiome, as a whole.

A number of studies underpin the hypothesis that the picture of the plant microbiome is not yet complete in a number of aspects: from the role of abiotic factors in shaping microbial communities, to the whole microbial composition itself. The full comprehension of the microbiome scenario is, however, crucial to address the plant holobiont and therefore plant health.

A still overlooked component of the plant microbiome, which I called overlooked microbial eukaryotes, includes a vast range of microbes, spanning from photoautotrophs to parasites, and has been proven to have prominent roles in other contexts such as soil or freshwaters. I applied ecological approaches as well as field experiments in order to address this overlooked component as a key of the still incomplete picture of the plant holobiont.

Amplicon sequencing on natural *A.thaliana* populations from 15 sites in Germany, France, Spain, and Sweden, revealed that overlooked microbial eukaryotes are not occasional partners of the plant holobiont, both epiphytical and endophytical. On the contrary, they are able to widely interact with key hubs of the plant leaf microbiome such as *Sphingomonas sp* and the family of Caulobacteraceae, independently of the surrounding abiotic factors.

Among others, microalgae have proven to be major shapers of microbial diversity for bacteria, fungi, and oomycetes. A newly established network analysis tool revealed that the presence of selected microalgae like *Bracteacoccus sp.* are linked to fluctuations of the pathogen *Pseudomonas viridiflava,* opening potential new fields in plant immunity research.

The role of microalgae within the plant holobiont is also likely expressed by the symbiosis with lichenising fungi. A tight association between lichenising fungi and potential algal partners was found in shaping the leaf microbial diversity, prominently the epiphytic compartment.

Amplicon sequencing on a time course-common garden experiment revealed that, beside microalgae, consumers like amoeboid organisms of the groups Lobosa, Conosa, Ciliophora also shape microbial diversity and follow specific succession patterns over time.

Furthermore, the key role of overlooked microbial eukaryotes in the plant holobiont seems to be stable even at low concentrations. In fact, pools of low abundant overlooked microbial eukaryotes shape microbial diversity, to an extent never assessed before.

In all these experiments, a key finding was the marginal effect of plant ecotype, as well as the marginal effect of latitudinal-climatic factors (with the exception of oomycetes), compared to the impact of overlooked microbial eukaryotes.

My work gives novel insights into the ecology and the successions of overlooked microbial eukaryotes, revealing scenarios in which primary producers shape microbial diversity through the presence of single taxon such as *Bracteacoccus sp,* or together with lichenising fungi. They also influence microbial population through consumers and predators like Ciliophora and Lobosa, which undergo successions over time. My work further supports the autonomy of the assembly of microbial community from the plant host genetic background and from province related features, underpinning a core and multilevel role of overlooked microbial eukaryotes in the plant holobiont.

Zusammenfassung:

Der Pflanzenholobiont ist ein neuartiger Begriff in der Biologie, der eingeführt wurde, um die Gesamtheit der Pflanze selbst und ihrer assoziierten Mikroben zu beschreiben. Um evolutionäre Selektion in diesem Rahmen zu verstehen, ist es notwendig pflanzenassoziierte Mikroben oder *Mikrobiome* in ihrer gesamten Komplexität zu untersuchen.

Eine Reihe von Studien untermauert die Hypothese, dass das Bild des Pflanzenmikrobioms in viele Aspekten noch nicht vollständig ist. Diese Aspekte beinhalten die Rolle der abiotischer Faktoren bei der Zusammensetztung mikrobieller Gemeinschaften bis hin zur gesamten mikrobiellen Zusammensetzung selbst. Das vollständige Verständnis des Pflanzenholobionts ist jedoch entscheidend, da die Zusammensetzung des Mikrobiomes Ausschlag gebend für die pflanzliche Gesundheit ist.

Eine noch immer verborgene Komponente des Pflanzenmikrobiomes, die im Verlauf als verdeckte mikrobielle Eukaryoten bezeichnete werden, umfasst eine große Bandbreite von Mikroben, die von photoautotrophen Organismen bis zu Parasiten reichen. In anderen Habitaten wie Böden und Süßwasser wurde bereits gezeigt, dass diese versteckten mikrobiellen Eukaryoten eine herausragende Rolle spielen. Ich habe sowohl ökologische Ansätze als auch Feldexperimente angewendet, um diese verborgene Komponente als Schlüssel für das noch unvollständige Bild des Pflanzenholobionts zu untersuchen. Amplikon-Sequenzierung von natürlichen *A.thaliana*-Populationen an 15 Standorten in Deutschland, Frankreich, Spanien und Schweden zeigte, dass versteckte mikrobielle Eukaryoten keine gelegentlichen Partner von Pflanzenholobionten sind und sowohl epiphytisch als auch endophytisch vorkommen. Des Weieteren sind sie in der Lage, unabhängig von den umgebenden abiotischen Faktoren mit Schlüsselorganismen des Pflanzenblatt-Mikrobioms, wie *Sphingomonas sp.* und der Familie der Caulobacteraceae, in Wechselwirkung zu treten. Unter anderem haben sich Mikroalgen als Hauptverursacher der mikrobiellen Vielfalt für Bakterien, Pilze und Oomyceten erwiesen. Das neu etablierte Netzwerkanalyse-Tool zeigte, dass ausgewählte Mikroalgen wie *Bracteacoccus sp.* zu einer geringeren Konnektivität des Erregers *Pseudomonas viridiflava* führten und eröffnet neue Möglichkeiten in der Pflanzenimmunitätsforschung. Die Rolle von Mikroalgen innerhalb von Pflanzenholobionten wird wahrscheinlich auch durch die Symbiose mit lichenisierenden Pilzen beeinflusst. Ein Einfluss von

lichenisierenden Pilzen und deren potentiellen Algenpartnern auf der mikrobiellen Diversität von Blättern, insbesondere in dem epiphytischen Kompartiment, wurde bereits nachgewiesen.

Die Amplicon-Sequenzierung bei einem Zeit-abhängigen Feld-Experiment ergab, dass Konsumenten wie amöboide Organismen der Gruppen Lobosa, Conosa, Ciliophora neben Mikroalgen auch die mikrobielle Diversität beeinflussen und im Laufe der Zeit bestimmte Sukzessionsmuster aufweisen. Darüber hinaus scheint die Schlüsselrolle versteckter mikrobieller Eukaryoten in Pflanzenholobionten auch bei niedrigen Konzentrationen stabil zu sein. In der Tat, prägen Gruppen von wenig abundanten versteckten mikrobiellen Eukaryoten die mikrobielle Diversität in einem Ausmaß, das zuvor noch nie bestimmt wurde.

In allen diesen Experimenten war ein Hauptergebnis der marginale Effekt des Pflanzenökotyps sowie der marginale Effekt von Breitengrad und Klimafaktoren (mit Ausnahme von Oomyceten) im Vergleich zu den Auswirkungen verborgener mikrobieller Eukaryoten.

Meine Arbeit gibt wichtige Einblicke in die Ökologie und das zeitliche Auftreten von versteckten mikrobiellen Eukaryoten. Darüber hinaus enthüllen meine Ergebnisse Szenarien, in denen Primärproduzenten die mikrobielle Diversität durch die Anwesenheit eines einzelnen Taxons wie *Bracteacoccus sp* oder zusammen mit lichenisierenden Pilzen formen. Sie beeinflussen auch die mikrobielle Gemeinschaften durch Konsumenten und Prädatoren wie Ciliophora und Lobosa, welche in ihrem Auftreten einem zeitlichen Muster folgen. Meine Arbeit unterstützt auch eine Unabhängigkeit der Zusammensetzung der mikrobiellen Gemeinschaft von dem genetischen Hintergrund der Pflanze und von Breitgrad-bezogenen Merkmalen, was eine Kern- und Multilevel-Rolle verborgener mikrobieller Eukaryoten in Pflanzenholobionten untermaue

1.   Introduction

1.1 Plant microbiome: ecological importance and implications for plant health

In their native environment, plants are surrounded by thousands of microbes, which form numerous communities and influence the plant health and productivity (Berendsen et al., 2012; Buée et al., 2009; Lindow and Brandl, 2003; Vorholt, 2012). Some of these associations can be beneficial (mutualistic), neutral (commensalistic), or deleterious (pathogenic) (Thrall et al., 2007).

Associations between plant and microbes, both pathogenic and mutualistic have been studied for centuries. One of the first being described was mycorrhizal symbiosis (Bary, 1863) and nodulation symbiosis of *Rhizobium sp.* (Leonard, 1943).

Recently, more than one hundred years after the first approach to mutualistic plant-microbe interactions, the focus is enlarging, trying to encompass a much higher number of plant
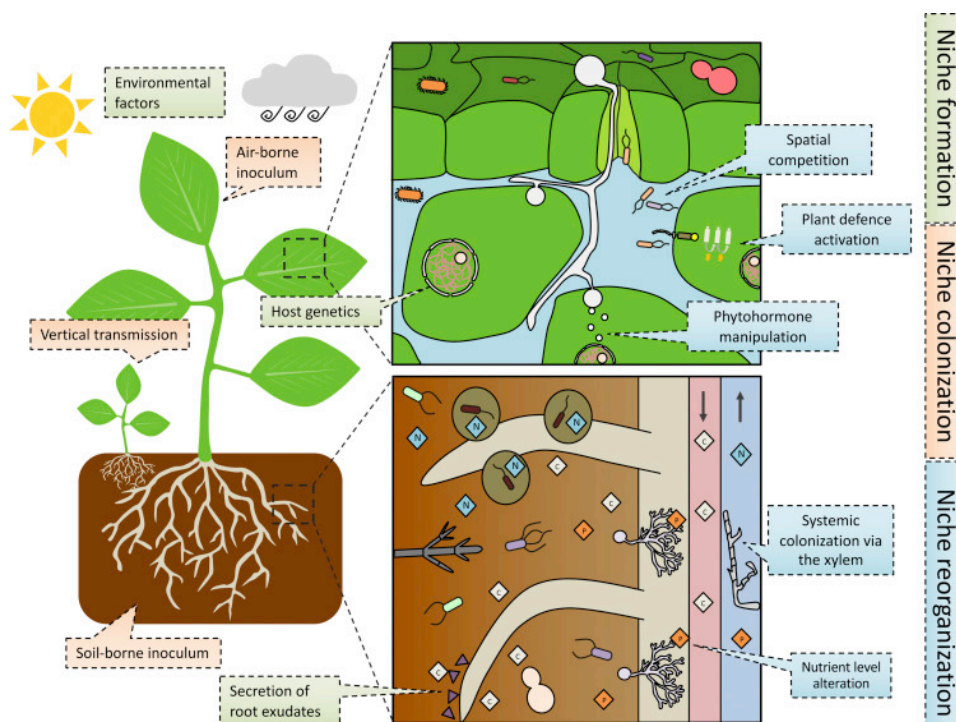
Figure 1.1: The plant holobiont, adapted from Kroll et al., 2018

interactors, including those microbes that are not cultivable in the laboratory environment. This generated the so-called culture-independent approaches (Hugenholtz et al., 1998).

The technological advance of high throughput culture-independent methods was crucial to disentangle complex and multivariate microbial interactions. Innovative approaches such

as metagenome or amplicon sequencing (Caporaso et al., 2012) (discussed later), allowed an integrated and complex formulation of a new biological concept: the so-called "holobiont" (Rosenberg and Zilber-Rosenberg, 2016). It is an entity comprehending a multicellular host and its associated microbiota, in which evolutionary selection takes place between host and microbes and between microbe and microbe (Hassani et al., 2018).

In this frame, it becomes clear that plant evolutionary pathway shall never be separated from the microbial consortia included in the holobiont. In a study of 2015 (Hardoim et al., 2015), it was reported that bacterial cells colonizing each gram of root, not only outnumber the cells of the entire host plant, but also represent more microbes than people existing on Earth.

The importance of plant-associated microbiome has been widely assessed in the past years, from being associated with the evolution of multicelluarity itself (McFall-Ngai et al., 2013), to the remarkable role in interfering with pathogens, suppressing them in the soil (Cordovez et al., 2015; Weller et al., 2002), or in some cases enhancing them (Partida-Martinez and Hertweck, 2005).

Ultimately, in certain plants like mosses, microbiome is so important for plant health that its absence implies the impossibility for the moss to germinate (Hornschuh et al., 2006), or to successfully reproduce (Peñuelas et al., 2014).

Studies on bacteria prove that microbial colonization of the plant host does not happen randomly (Cardinale et al., 2015), and across distantly related plant species, different microbial assemblies can be found (Schlaeppi et al., 2014). Quantitative differences are also present between the root (Berendsen et al., 2012; Philippot et al., 2013), and the phyllosphere (Vorholt, 2012).

In the light of the holobiont concept, these different microbiomes across species were analyzed both concerning host-microbe interactions and concerning microbe-microbe interactions(Hassani et al., 2018).

If on one hand there is strong evidence that microbial assemblies are linked to the host phylogeny (Lundberg et al., 2012) and to the conditions of the immune system (Lebeis et al., 2015), on the other hand, these assemblies seem to diverge only quantitatively from species to species (Schlaeppi et al., 2014), and in some cases they largely overlap (Hacquard et al., 2015). This means that not only host-derived factors are relevant in shaping microbiome assembly, but also environmental factors must be taken in consideration.

Examples of these environmental factors are biogeography –the distribution of species and ecosystem through time- (Coleman-Derr Devin et al., 2015) or soil type (Bulgarelli et al., 2012).

Among these other environmental variables there is also the presence of other microbes. For instance, the already mentioned disease suppressive soils have a positive effect on plant health. This is mostly due to the presence of a definite microbial assembly which actively counteracts the colonization of a pathogen directly, or indirectly (van der Heijden and Hartmann, 2016; Zamioudis et al., 2015). This final effect can be due to competition on the nutritional aspects (Mercado-Blanco and Bakker, 2007) or to the active secretion of antimicrobial compounds (Van Acker et al., 2014).

However, not only competitive interactions take place in phyllosphere/rhizosphere microbiome. It is now consolidated knowledge that many bacteria and fungi form together what is called a *biofilm*. The *biofilm* is a matrix-like structure formed by the Extracellular PolySaccharides (EPS) (Bogino et al., 2013; Guennoc et al., 2017). EPS are harnessed by bacteria and fungi as protection from pathogens or for facilitated inter-species chemical signaling (Steidle et al., 2001).

There has been proof that chemical signaling is a key to understand microbial dynamics, in particular *quorum sensing* is a chemical signaling process which takes place across microbes, typically gram-negative bacteria (Miller and Bassler, 2001).

Originally believed to be only a chemical-based way of monitoring population growth, *quorum sensing* has been discovered to play a major role within plant holobiont (von Bodman et al., 2003). In fact, many intra and inter-kingdom interactions rely on *quorum sensing* signals to be developed (Jarosz et al., 2011). It has been reported that microbes in the phyllosphere, as well as in the rhizosphere, actively use it or interfere between other species *quorum sensing* in order to compete or cooperate with the present microbes (Hartmann and Schikora, 2012). Finally, there has been evidence that the plant itself is able to interfere in microbial quorum sensing, thereby exerting a top-level role in shaping microbial diversity (Bauer and Mathesius, 2004).

All these studies mentioned have been of invaluable importance in opening the field of plant microbiome research and in further addressing its dynamics. However, there are a number of issues which still remain unsolved:

What role does biogeography play in plant microbiome structure and dynamics? Is it enough to explain what before was accounted just as stochasticity? Are there other

dynamics which may be of great relevance that are still not yet tackled in plant microbiome studies?

As already mentioned, biogeography seems to be of primary importance for microbe distribution. However, this has been demonstrated mainly for fungi (Coleman-Derr Devin et al., 2015; Tedersoo et al., 2014). In recent studies, it was reported that seasonality and location explained up to circa 50% of bacteria variation (Agler et al., 2016a), leaving another 50% of variation unexplained. In the same article, it was reported that additional variation portions can be explained by the presence of other microbes, like *Albugo sp.* and *Dioszegia sp.*, but the variation accounted by these two microbes still leave a large part of bacterial variability unexplained.

On the other hand, there has been evidence of other ecological processes taking place on the leaf, for instance predation due to certain protists is reported to shape bacterial diversity on the leaf and in the roots, especially for the taxonomic group Rhizaria-Cercozoa, hereafter "Cercozoa"(Flues et al., 2018).

It becomes logical to think that the picture of microbiome that had been analyzed so far may be incomplete, and therefore deserves full attention.

1.2   The incomplete picture: ecological insights from overlooked microbial eukaryotes can reveal the missing piece of the puzzle

The aforementioned *protists* (par 1.1) do not officially constitute a kingdom nor a taxonomically defined group (Adl et al., 2005; Parfrey et al., 2010). Their definition has been reformulated many times, including sometimes unicellular, sometimes multicellular organisms (Parfrey et al., 2011). One of the latest definitions encompasses an extremely high number of species, very heterogeneous in terms of phylogeny and occupied niche, on many occasions including or excluding fungal microbes as well (Parfrey et al., 2011). For the purposes of this study, I propose to focus only on some of these groups with an alternative definition. As *overlooked microbial eukaryotes*, I am going to address only those eukaryotic organisms unicellular or multicellular, which are classified neither within the fungal kingdom nor within oomycetal subphylum and that are found to be associated with the plant holobiont. (fig 1.2, 1.3).

4
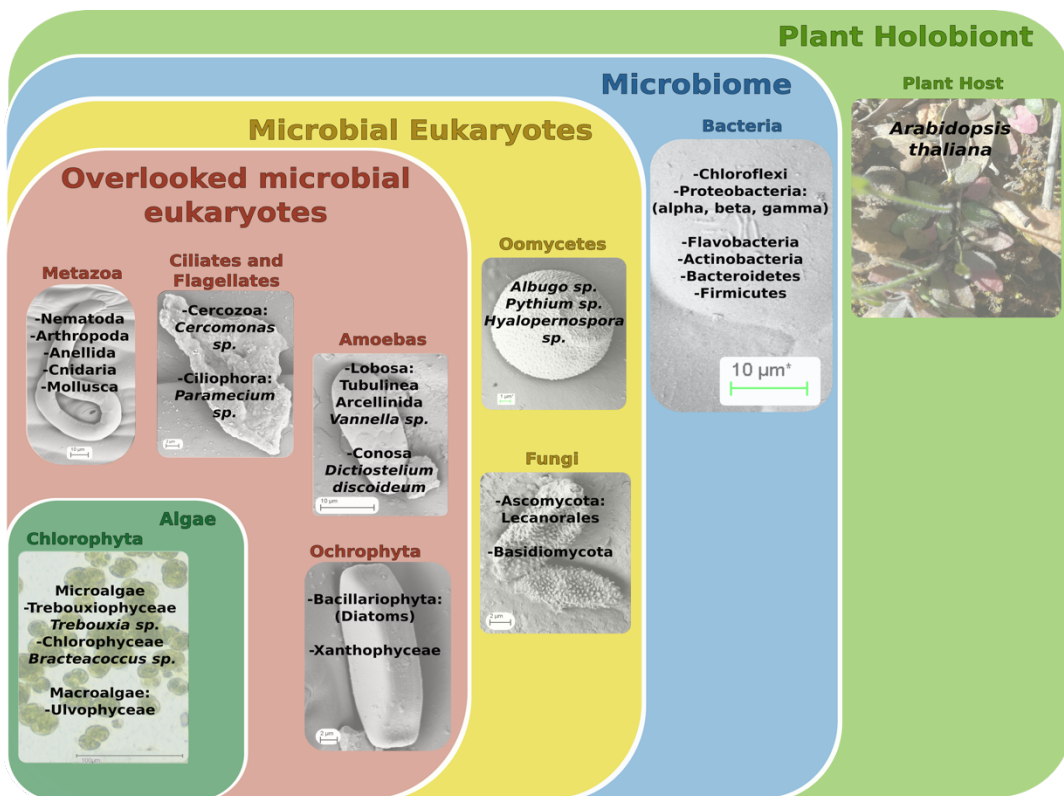
Figure 1.2 Overlooked microbial eukaryotes in the plant holobiont frame

The ecological relevance of *overlooked microbial eukaryotes* in microbial consortia has already been assessed for unrelated ecosystems such as freshwaters (Simon et al., 2016) or
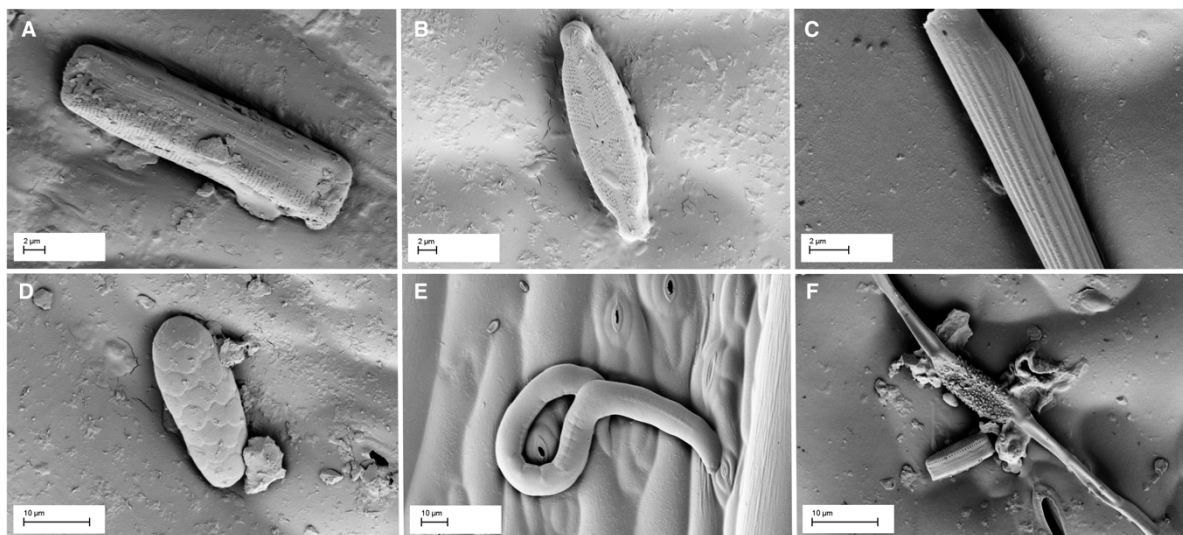
Figure 1.3 A glimpse into overlooked microbial eukaryotes communities on the leaf surface: A-B: Diathomes (Bacillariophyta), B, likely belonging to the genus Navicula sp, C: Unknown, D: Testate amoeba, (Tecofilosea), E: Nematode (Nematoda), F:bottom-left, Bacillariophyceae (Diatom), middle-elongated: germinated fungal spore, lkely belonging to the genus Cladosporium, bottom-middle-right: unknown amoeba

oceans (Moreira and López-García, 2002). This is the reason why only investigating *overlooked microbial eukaryotes* in an ecological framework would allow to unravel at

least part of the missing tile of plant holobiont, as it has been done already for the rest of the known plant associated microbiome (Vandenkoornhuyse Philippe et al., 2015). Considering plant-associated overlooked microbial eukaryotes in an ecological framework does not mean just considering the amount and the identity of species that are plant-associated. On the contrary, I propose a framework featuring the evaluation of biogeography and dispersal potential, the role of primary producers and consumers/predators, as well as the assessment of the role of low abundant taxa in microbial consortia.

1.2.1 Biogeography of overlooked microbial eukaryotes: which are the ecological drivers of microbial populations?

In recent studies on marine plankton, (Lima-Mendez et al., 2015), it has been reported how a large part of overlooked microbial eukaryotes interactions is conserved among several locations across the oceans, although over a third of the reported interactions remain on a local scale.

This issue of local versus global scale of microbial dispersal and activity is of great ecological importance and often a target of microbial biogeography studies.

Studies on biogeography of organisms, including microbes, often use the notions of province and habitat: the first one defined as a region, the biotic composition of which, reflects the legacies of historical events (de Candolle, 1820; Martiny et al., 2006). In other words, province is the latitudinal/geographical location linked to that specific community. The habitat is described as an environment defined by the combination of its abiotic and biotic characteristics (de Candolle, 1820). Of course, a province can contain different habitats and vice versa. A typical example is the Mediterranean coastal habitat, which can be individuated in extremely distant provinces like California, South Africa, and Australia. Concerning microbial ecology, four different hypotheses explaining the recovered distribution of microbes have been formulated (Martiny et al., 2006)(fig 1.4):
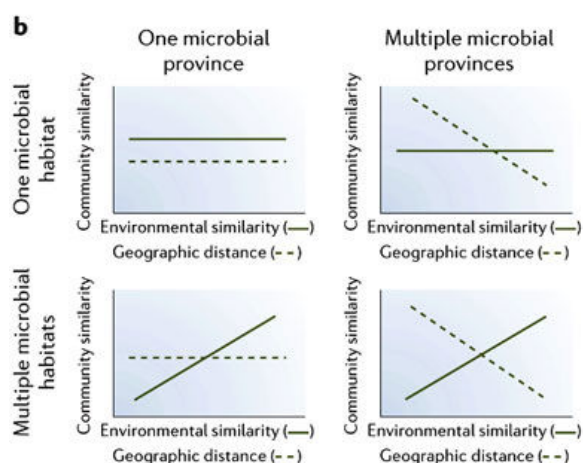
Figure 1.4: Province vs habitat: the four hypotheses of (microbial) ecology, the scheme represents the different formulations of province and habitat impact on microbial community. Adapted from Martiny et al., 2006

1. Biogeography does not exist (Cosmopolitism hypothesis): implying that the entire microbial diversity is reachable everywhere. The entire world is constituted by one province and one habitat

2. Biogeography reflects the environmental variation within the single province (Baas-Becking hypothesis (Baas-Becking, 1934)): One province, several habitats. Also known as *"everything is everywhere, environment selects"*.

3. Biogeography is in essence shaped by historical events that might influence present-day assemblages include dispersal limitation and past environmental conditions, both of which can lead to genetic divergence of microbial consortia. In short multiple provinces but only one habitat.

4. Biogeography is driven by both historical events and contemporary environmental conditions. In other words, the diversity can be shaped by multiple habitats and multiple provinces.

It has to be clarified that none of the cited hypotheses has to be considered true or false *per se*. In a recent study on Finnish freshwaters, (Heino et al., 2010) diatoms (brown algae) have been found to follow rather geographical gradients than a habitat-specific patch, thereby matching more the third mentioned hypothesis. On the other hand, a study on other fresh water green alga (*Synura petersenii*) (Boo et al., 2010) supports the hypothesis that both habitat conditions and latitudinal dispersal are equally influential for the biogeography of the studied microbe, thereby matching the fourth hypothesis.

This demonstrates how all four hypotheses can be applicable to virtually all microbes, including species which share a relatively close phylogeny.

Concerning plant microbiome, in the light of the holobiont concept, not only microbe biogeography has to be considered, but also the host biogeography.

Most of the macro-organisms, and therefore plants as well, follow the fourth hypothesis in terms of biogeography. This is why it is not surprising that, given the knowledge achieved so far on plant microbiome, both habitat and province are hypothesized to play the most

7

significant role in shaping plant microbiome composition (Redford et al., 2010). However, to the best of my knowledge, no study so far has tackled the distribution of plant-associated overlooked microbial eukaryotes. The first step to understand overlooked microbial eukaryotes impact in plant holobiont would be to investigate how much their distribution is linked to the host, or to which extent that it is independent.

1.2.2 Plant holobiont encompasses several primary producers: autotrophic organisms associated with plants

In ecology, all organisms able to produce their own food source through photosynthesis (autotrophs), are called primary producers (Lindeman Raymond L., 1942). This definition includes all macro and micro-organisms which possess chloroplasts and are able to perform photosynthesis. This encompasses not only plants, but also macroalgae, microalgae, as well as some strains of photosynthetic bacteria such as *Chloroflexi* and *Cyanobacteria*. The distinction between macroalgae and microalgae is mainly morphology-based, since phylogenetically both are composed of the same taxonomical groups. There is also an ecological divergence, since macroalgae are eminently aquatic, while microalgae are also terrestrial (Bonanno and Orlando-Bonaca, 2018). Microalgae are a polyphyletic group including several taxonomic groups, sometimes distantly related (*Chlorophyta* and *Ochrophyta*, the main ones) (Metting, 1996).

Biogeography of microalgae however does not seem to be connected with their phylogeny. For example, *Klebsomorbidium sp.*, a free living terrestrial alga which a recent study (Ryšánek et al., 2015) found to be genetically homogeneous on a global scale, whereas heterogeneous on a local scale.

In the past years, many studies have tackled the distribution and the phylogeny of microalgae, however, very little knowledge is available on their possible interactions with higher plants, and therefore their impact on the plant holobiont (Škaloud and Rindi, 2013). A more explored field regarding microalgae and plant holobiont is the one belonging to lichens.

Lichens are stratified, often leafy structures called *thalli* which includes a fungus, typically belonging to the phylum of Ascomycota, and a microalga (De Bary, 1879). Entire families of fungi are recovered in nature only within lichens, and consolidated studies indicate that lichenising fungi evolved earlier than believed (Lutzoni et al., 2001). This indicates that most of the non-lichenising fungi have actually evolved from lichenising ancestors through

independent gene losses. This finding underpins the ecological and phylogenetical relevance of lichenising fungi and triggers the question of what could be their interaction within the plant holobiont (Lutzoni et al., 2001).

Concerning the algal symbiont, otherwise known as *phycobiont*, the most widespread one belongs to the family of *Trebouxiophyceae*, in most of the cases, to the genus *Trebouxia sp.* (Ahmadjian, 1960)

Lichens have been widely studied and characterized across ecosystems and most of them show behavior of pioneer colonizers (Lawrey, 1991).
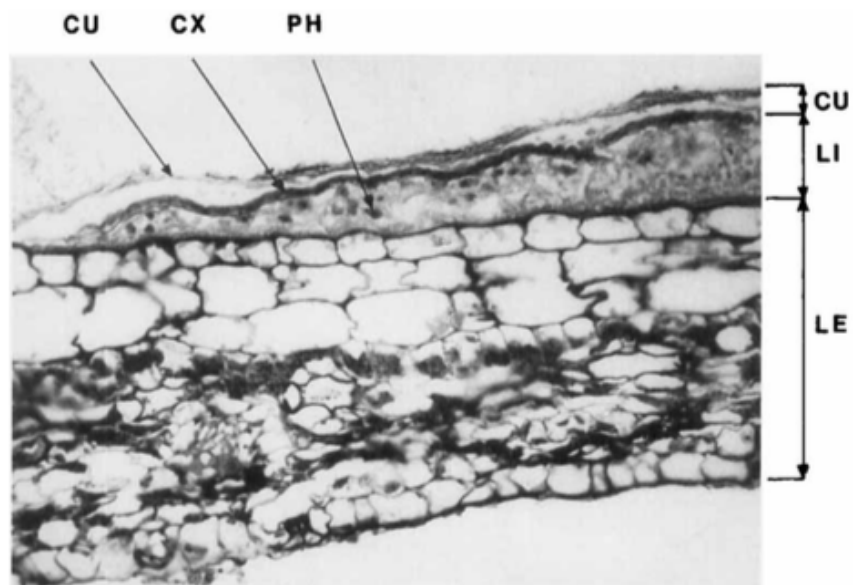
Figure 1.5: Structure of foliicoulous lichen; Vertical section through Strigula smaragdula. CU: leaf cuticle; CX: lichen cortex; PH: lichen photobiont (algae belonging to the genus Cephnleuros); LI: lichen; LE: leaf. Adapted from Serousiaux et al., 1989

Certain kind of lichens, known as "foliicoulous lichens" are reported to live eminently on tree leaves (Lücking, 1999) on top or under the cuticle (fig 1.5), however, this seems to be a phenomenon limited to a certain climatic province -the tropical one (Lücking, 1999)- with a few exceptions in temperate areas like the Balcans and a few forests in France (Serusiaux, 1989).

Most of the studies conducted on these lichens support microhabitat factors as key factor for foliicoulous lichen establishment (Lücking, 1999). The main ones appear to be the floating temperature of the shady understory, and the so called light gaps (Favero-Longo and Piervittori, 2010; Rogers et al., 1994) -physical openings among the forest canopy, large enough to let UV light to reach the understory.

Interestingly, little or no evidence has been shown on the role of the plant host on which the lichen is living (Favero-Longo and Piervittori, 2010; Lücking, 1999; Serusiaux, 1989). In fact, the lack of difference in lichen community between primary and secondary forest suggest that environmental factors and microhabitat factors are the most prominent variables determining lichen community.

At present, little is known about foliicoulous lichen interaction with bacteria. However, many studies on other lichens (fructicose, foliose) typical of other substrates, indicate a signature of a lichen microbiome presence (Bates et al., 2011; Grube and Berg, 2009; Grube et al., 2009; Mushegian et al., 2011; Rogers, 1988).

The bacterial species recovered (mainly α–proteobacteria) seem more bound to the phycobiont rather than to the fungal partner (Grube et al., 2009). The ability of those bacteria to stimulate algal growth by secreting auxin highly underpins the link between bacteria and phycobiont (Grube and Berg, 2009). However, the bacterial species found to be related to lichens are also involved in phosphorus and nitrogen mobilization (Bates et al., 2011), which happens also to benefit the fungal partner. The genera found to be prominent in lichen microbiome are typically the ones known to occur in the endophytic compartment of the leaf (Grube and Berg, 2009). Examples are *Burkholderia*, *Stenotrophomonas*, *Pseudomonas* (Grube and Berg, 2009), but also *Methilobacterium* and *Sphingomonas* (Grube et al., 2009; Mushegian et al., 2011). Other studies indicate a major role of the Rhizobiales order in general (Bates et al., 2011).

It is still debated whether lichen microbiome assembly is correlated with the lichen species (Bates et al., 2011), or not (Mushegian et al., 2011). Instead, what is not debated is that the major correlator for lichen microbiomes is actually the substrate nature, and the microhabitat conditions that the resemblance with plant leaf endophytes underlines. The discovery of an impact of lichens on the plant holobiont in plants other than tree or mosses and in temperate ecosystems would open new fields in plant ecology and microbiology research.

1.2.3 Micropredators: heterotrophic microbes shaping the overall diversity of plant associated microbes

I have already mentioned that heterotrophic organisms are certainly part of the microbial consortia associated with plants, especially roots (par 1.1). Among the most most well-known interactors is the taxonomic group of Rhizaria-Cercozoa (Flues et al., 2018).

However, not just Rhizaria-Cercozoa are part of plant-associated microbes. It has been recently reviewed that other organisms belonging to the groups of Lobosa, Conosa and Ciliophora, populate plant holobiont (Smith and Wilkinson, 2007). From here onwards I will refer to them as "amoeboid organisms". In that study, it has been pointed out that certain testate amoebas, belonging to Lobosa taxonomic group can be found in 59% of the cases on mosses, and on trees in 16% of the cases. Remaining percentages are assigned to algae (2%), forbs (6%), and grasses (9%). Such a remarkable trace of presence on mosses, suggests that this occurrence may be other than stochastic. This conclusion is supported by further studies (Fiz-Palacios et al., 2013) which found other member of Arcellinidae (a Lobosa sub-group, including testate amoebas) to live more prominently on mosses than on soil.

The dynamics behind these co-occurrences remain to be clarified. However, the fact that distribution of amoeboid organisms would be rather influenced by the substrate abundance/quality (Heger et al., 2013; Schipper et al., 2001) seems not to be debated. The dispersal potential seems as well to have a role in amoeboid organisms distribution, which of course is different between taxonomic groups (Heger et al., 2013). For example Ciliophora appear to have a more intense cosmopolitan behavior (Borror, 1980; Foissner, 1999) than other taxonomic groups like Lobosa or Conosa. Any link with plant genotype is yet to be clarified.

Concerning substrate, happens that for most of the amoeboid organisms, this is constituted of bacteria and other microbial eukaryotes. It is not surprising then, that micro-heterotrophs diversity is majorly connected with bacterial diversity (Ploch et al., 2016).

I have already mentioned the peculiarity of bacterial communities associated with plants (see par 1.1). If there is a connection between bacterial communities and amoeboid organisms, it would be logical that terrestrial amoeboid organisms were then associated with higher plants indirectly through a bacterial-feeding connection.

Such a hypothesis is underpinned by a study on the main taxonomic groups among amoeboid organisms (Fiz-Palacios et al., 2013), which showed a coincidence of genetic differentiation of amoebas into Lobosa and Conosa group with the corresponding differentiation of higher plants, especially of mosses.

These studies together show that the dispersal, and therefore the colonization of different environments by amoeboid organisms, is mainly driven by circumstantial factors such as substrate nature, prey distribution, and environmental factors (Heger et al., 2013).

This importance of environmental factors and nutrient source, justifies their large diffusion of amoeboid organisms. Studies on meat processing plants and on drinking water (Poitelon et al., 2009; Vaerewijck et al., 2008) have found eukaryotic diversity strikingly similar to the one found on mosses or on the soil, of course with quantitative differences, indicating an extremely high dispersal potential and adaptability.

The large dispersal of amoeboid organisms, however, does not imply minor roles in microbial diversity shaping and in general, in ecology. On the contrary: a number of studies underlines that their relationship with bacteria is not only just related to feeding (Delafont et al., 2015; Nowack et al., 2016; Okubo et al., 2018) but also to mutualistic interactions and gene transfer. In fact, endocytobiosis is a phenomenon which involves amoeboid organisms and bacteria, and is fundamental for bacteria pathogenicity, especially towards mammals (Scheid, 2014). It has been proven that amoebas often host bacterial populations inside their cytoplasm and lysosomes (Cosson and Soldati, 2008). This often determines an increased pathogenicity of the endosymbiont bacteria, as it has been proven for *Legionella sp.*(Cosson and Soldati, 2008; Okubo et al., 2018). In the same study, it is proposed that this "amoebic passage" is actually crucial for the bacterium to also develop opportunistic pathogenicity against mammals. Concerning plants, no evidence has been found so far of such effect on plant pathogens. However, a study from 2007 (Hilbi et al., 2007), reviews that amoebas like *Dictiostelium discoideum* are actually used to determine which strains of *Pseudomonas aeruginosa*, (relatively close to the plant pathogens *P.syringae* and *P.viridiflava*) is virulent or not. It is remarkable to notice that the virulence of such bacteria, is tightly connected to the presence of LasR genes, involved in quorum sensing signaling. Amoeboid organisms are well known to be able to interfere with bacterial quorum sensing and therefore be of major importance for the development of pathogenicity, especially of bacteria (Cosson and Soldati, 2008; Hilbi et al., 2007; Scheid, 2014). These findings would underpin the hypothesis that also plant pathogens could acquire or strengthen their pathogenicity through an "amoebic passage" as well. However, apart from the consolidated knowledge on Cercozoa already demonstrated in previous studies (Flues et al., 2018; Ploch et al., 2016; Sapp Melanie et al., 2018), the role of all these other taxonomic groups (Lobosa, Conosa, Ciliophora, Metazoa) in association with higher plant hosts like crops or just *A.thaliana* has hardly been addressed so far. A proceeding in this sense will be of great advance in understanding community dynamics and successions, including prey-predator interactions.

1.2.4 Keystone species are not necessarily abundant: the role of rare taxa in microbial assembly

Within microbial consortia, especially in oceans environments, the so-called rare taxa have been studied for a long time (Lynch and Neufeld, 2015). Rare microbial communities show a remarkable diversity, however, their role remains to be fully clarified (Fuhrman, 2009; Sogin et al., 2006). There have been different theories proposed to explain the presence of rare microbes, including the transient hypothesis **(Pedrós-Alió, 2011)**. It proposes that the reason why they display low abundance is just due to the fact that they are close to extinction. Another theory, actually more grounded to several observation in several ecosystems, hypothesize the active role of rare taxa, which not necessarily are on the way to be extinct, triggering effects which are disproportioned compared to their size or population abundance (Campbell et al., 2011; Jones and Lennon, 2010; Shade et al., 2012). An interesting theory developed from these observations hypothesize the role of rare taxa as a reservoir of organisms that are waiting for more favorable conditions to bloom (Epstein, 2009; Lennon and Jones, 2011). In some cases, this "stand by" to wait for favorable conditions, has been interpreted as a crucial mechanism to foster resilience of microbial networks after a major perturbation (Ainsworth et al., 2015; Fuhrman, 2009).

Recently an interesting study has tackled this hypothesis thoroughly by addressing the ecology of the so called CRT (Conditional Rare Taxa) (Shade et al., 2014). In this study, CRT are identified through a ratio between skewness and kurtosis of their abundance curves. The CRT calculated in this manner, in datasets sequenced with an adequate depth, seem to be responsive (bloom) explicitly after an induced mechanical perturbation (artificial lake water mixing). This finding confirms the aforementioned hypothesis of rare taxa being crucial for microbial communities' resilience.

Despite the insights into rare microbe ecology, it is still unclear whether or not the biogeography of rare taxa follows the biogeography of abundant taxa (Galand et al., 2009; Newton et al., 2013; Reveillaud et al., 2014; Youssef et al., 2010).

Another issue which still remains debated is what could be set as a threshold to define a taxon as rare or not rare. Some studies (Reveillaud et al., 2014) set this threshold to the 0.001% of the overall abundance. However, this arbitrary solution poses the problem that such low concentrations of microorganisms can become difficult to detect and distinguish from the noise (Fuhrman, 2009). Furthermore, in geographical sampling, it would become

important to define a threshold per site, in order not to consider rare, what may be abundant in the next site (Fuhrman, 2009).

This is the reason why several studies tackling the biological question of the importance of rare taxa, often choose different thresholds, sometimes different by several order of magnitude (Campbell et al., 2011; Debroas et al., 2015; Galand et al., 2009).

Most of the previous mentioned studies have been focused on ocean, or human microbiome. Little is known so far about the role of terrestrial rare taxa. A hint comes from a recent study that identifies soil rare bacterial community as responsible for a decreased plant fresh weight, and therefore triggering a negative effect on plant production (Hol et al., 2010).

In conclusion, there has been evidence that rare bacterial biosphere is of crucial importance for microbial community dynamics. However, little or no knowledge has been collected on rare overlooked microbial eukaryotes. Few studies have tackled the biogeography of them (Schiaffino et al., 2016), however, no conclusion has been drawn on their possible impact on the rest of the microbial community. I hypothesize that low abundant overlooked microbial eukaryotes can have an effect comparable to low abundant bacteria, on microbial community.

1.3 Methods to study microbiome: high throughput amplicon sequencing

Amplicon sequencing has been widely used in a number of studies aiming to characterize the complexity of microbiomes, on plants and on other systems (Hacquard et al., 2015). Amplicon sequencing aims to target a so-called barcode sequence within the target organism genome, amplify it, and determine its sequence (Caporaso et al., 2012). The obtained reads are then paired, clustered in Operational Taxonomical Units (OTUs), and flagged with the extracted representative sequence (Caporaso et al., 2010). A taxonomy is then assigned to the representative sequence, based on available sequence references. The relative abundance of a certain species is meant as the abundance of the OTU represented by the representative sequence whose taxonomy assignment match that certain specie (Caporaso et al., 2010). The whole approach therefore relies entirely on the amplification of the barcode sequence, which ideally should be able to discriminate organisms at species or at least genus level.

The traditional strategy was to design very specific primers, able to target an entire class, order or kingdom, and, if they were specific enough, automatically exclude any other

possible misamplification. However, this approach still creates a number of other abundance-related biases (Schirmer et al., 2015). Of course, the more specific the primers are, the more successful is the strategy. In case of usage of universal primers, the risk of misamplification grows steeply (Agler et al., 2016a, 2016b).

In the case of the present study on plant microbiome, however, the choice of universal primers to amplify overlooked microbial eukaryotes is mandatory. In fact, as already mentioned previously, overlooked microbial eukaryotes encompass a vast range of phylogenetically heterogeneous organisms. This means that the barcode sequence of overlooked microbial eukaryotes and the host will have similar barcode primer sites. This feature makes it necessary to develop a new solution to allow a universal primer approach targeting specifically overlooked microbial eukaryotes and avoid the misamplification of the host DNA.

1.3.1 Network analysis majorly strengthen ecological analysis

Amplicon sequencing data have been recently classified in the category of the so-called "big data". big data are data whose dimension and complexity is so massive, that usual
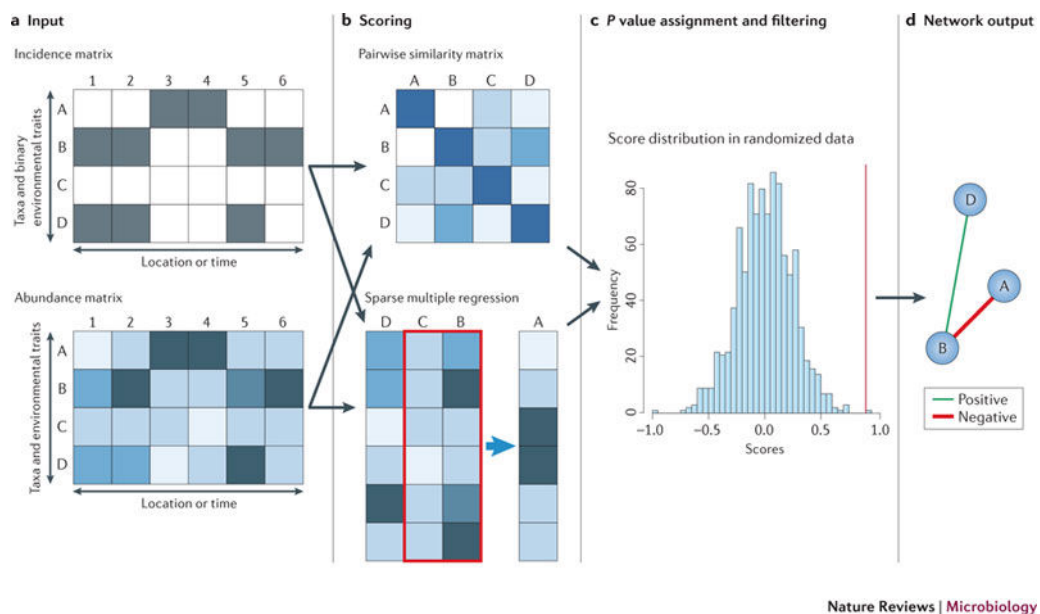
Figure 1.6: Inferring a network, from the input data (a), to the scoring via similarity matrices(b), filtering out the high p-values (c) to the final output (d). Adapted from Faust and Raes (2012).

software facilities are not fast or efficient enough to process and mine them. While this classification is still under debate, it is actually consolidated knowledge that microbiome sequencing data should be approached with tools similar to the ones commonly used for big data analysis. These includes data subsetting, clustering, filtering, and network analysis.

Network analysis has already been used many times to disentangle microbiome data (Barberán et al., 2012; Faust et al., 2015; Lima-Mendez et al., 2015). In summary, it uses the principles of graph theory in order to represent microorganisms as nodes and their positive or negative co-occurrence as edges (Barabasi, 2009; Faust and Raes, 2012). Network analysis can be deployed in numerous contexts and for different subsets of data. In the past, it allowed to draw holistic conclusions on the dynamics of certain species, without losing the complete picture of surrounding microbial interactions (Lima-Mendez et al., 2015). Since network analysis was deployed for the first time on biological data, there has been a growing demand for tools able to infer networks in a more solid and reliable way (Faust and Raes, 2016) and as well as accessory tools for network comparison. For the first need, several new techniques were developed and proposed to the scientific community, in some cases developing extremely precise and powerful instruments, able to combine the best characteristics of already present tools in a multi-layer fashion (Pilosof et al., 2017). Concerning network comparison, the tools available so far, mainly allow a qualitative comparison (Goenawan et al., 2016; Landeghem et al., 2016), which only in few cases can restitute a detailed picture of the discrepancies within the two networks, and in no case provide supporting statistics.

Therefore, it is necessary to improve network comparison by the deployment of a new tool able to compare networks quantitatively and provide supporting statistics.

1.4 Aim of the thesis and experimental setup

With the term *overlooked microbial eukaryotes* I aim to encompass all eukaryotic organisms assigned neither to the fungal nor to oomycetal group associated to the plant holobiont. Those organisms and microorganisms, also colloquially known as *protists*, are extremely heterogeneous both in terms of ecological niche and in terms of phylogeny (Adl et al., 2005; Parfrey et al., 2010, 2011), including microalgae (Chlorophyta), parasitic plasmodia (Apicomplexa), as well as predator amoebas (Lobosa, Conosa).

The plant microbiome has been dissected and deeply investigated so far (Hassani et al., 2018), mainly for roots, and first piece of knowledge about certain groups of overlooked microbial eukaryotes start now to be achieved (Flues et al., 2018; Sapp Melanie et al., 2018).

However, many taxonomical groups remain neglected, from predators to photoautotrophs. A number of studies underpin the hypothesis that the picture of the plant microbiome is not

yet complete (Agler et al., 2016a; Hassani et al., 2018). Until this picture becomes complete, it will be hard to fully understand plant microbiomes and plant holobiont dynamics and its direct impact on plant health (Hassani et al., 2018).

In this study, I aim to contribute in completing this picture, by specifically addressing overlooked microbial eukaryote classes through the following biological questions:

- What are the factors determining the dispersal of overlooked microbial eukaryotes associated with plants? Are they more linked to biogeographical factors or to microhabitat features?

- Are they important within microbial consortia? What is their connectivity with keystone species?

- Narrowing down the focus: which ones are the most important? What is the role of microalgae in microbial consortia? And what the role of amoeboid organisms?

- Is there any microbial succession taking place? And what is the role of the host genotype?

- Is the relevance of overlooked microbial eukaryotes dependent on their abundance? What is the role of low abundant overlooked microbial eukaryotes?

In order to tackle these questions, I designed two experimental setups:

The first one consisted of a latitudinal sampling of natural, stable populations of *A.thaliana*. I chose *A.thaliana* in order to allow easier lab experimental follow-ups, given the extensive
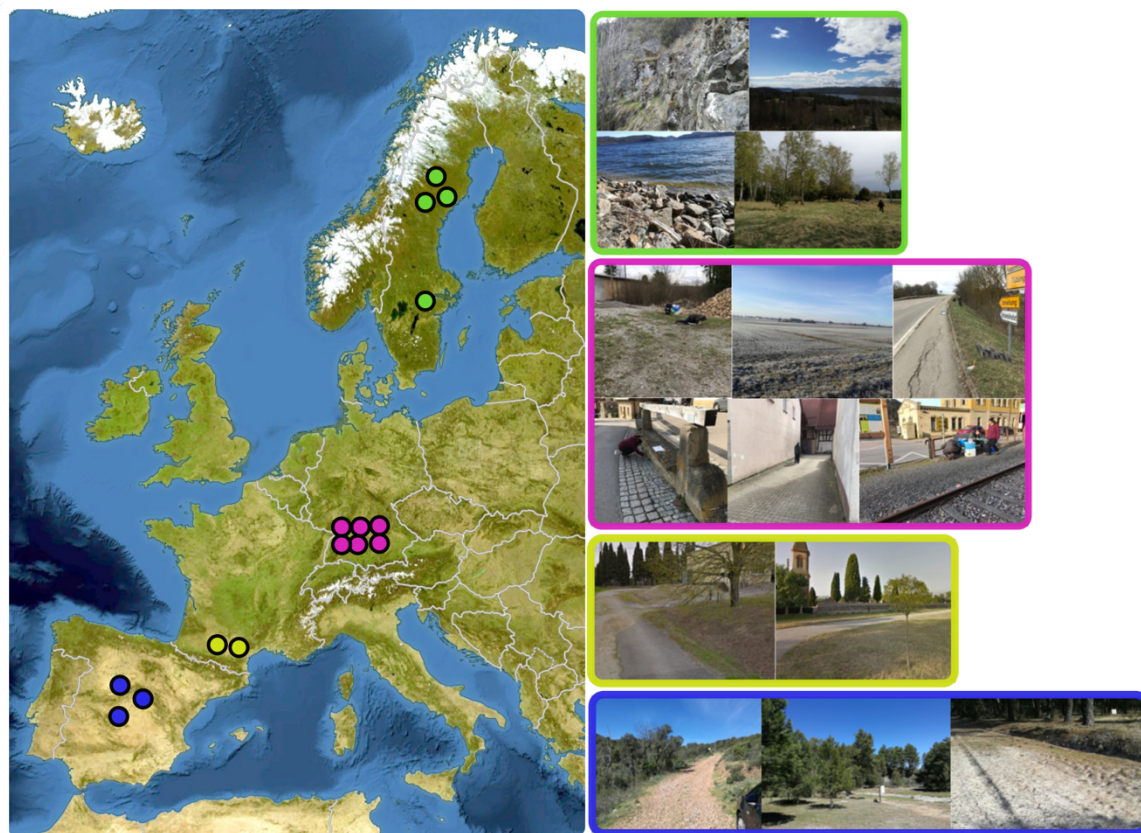
Figure 1.7 The biomes and provinces of the investigated sites

knowledge and resources available concerning the genetics of this species. In order to collect as much data as possible on the latitudinal distribution, I chose 15 sites across 4 countries: Germany, Spain, France and Sweden. Each site had a different history in terms of soil, neighboring plants and anthropogenic impact. I proposed the following classification of the sites based on these variables in: natural, agricultural, urban, suburban. I called these classifiers *biomes* which were equally distributed across the four countries, apart few exceptions (see M&M).

The second setup consisted of a common garden experiment in Cologne field soil, in which four ecotypes of *A.thaliana* (Ksk, Ws-0, Col-0, Sf-2) were planted and sampled monthly starting from November and ending in March. Both setups were set running for three consecutive years: 2015, 2016, 2017.

I then performed a microbiological survey using amplicon sequencing, through Illumina technology. I chose to amplify bacterial, fungal and oomycetal markers, respectively 16S (V3 and V5 regions) for bacteria, internal spacer sequences ITS1 and ITS2 for fungi and

oomycetes, as described in (Agler et al., 2016a). Concerning overlooked microbial eukaryotes I chose as reference markers eukaryotic 18S (V4-5 and V8-9 regions), as suggested in (Hadziavdic et al., 2014).

With collected data, I aim to draw ecological conclusions based on a holistic approach, therefore narrowing down the focus on single genera or specie as little as possible. In order to do so, I chose to use tools like network analysis, which allow a wide and reliable representation of microbial interactions. I also used supporting statistics like ANOVA and Mantel test to disentangle the network analysis results. A proposal of methodological improvement to compare differential network analysis will be also deployed in order to validate the data and if possible to draw further conclusions

2.   Results

2.1 Ecological relevance of microbial eukaryotes in the plant holobiont

2.1.1 Sequencing taxonomical markers from microbial eukaryotes: blocking oligos substantially increase sequencing resolution

The amplification of universal eukaryote taxonomic markers within a plant system is highly susceptible to off-target amplification from the plant DNA itself. These amplicons typically reach almost 90% of the total amplicons (Hanshew et al., 2013). Another issue to be considered is that the target of this amplification are typically taxa which are less abundant than other
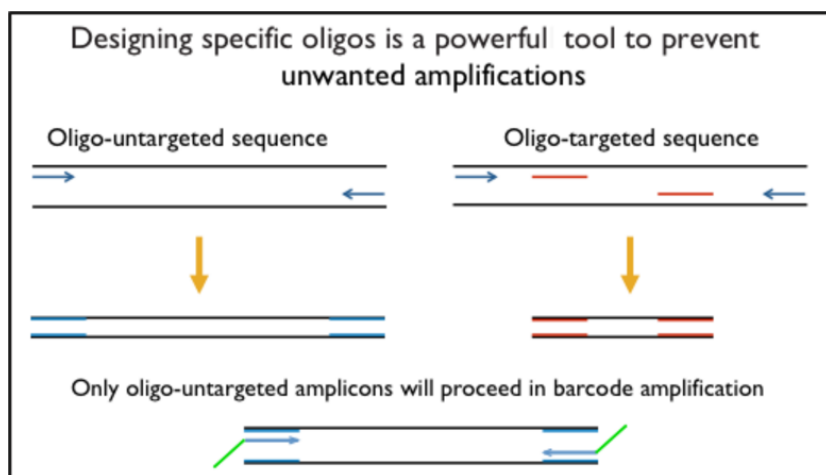
Figure 2. 2; Blocking oligos target specifically the host DNA determining the formation of a shorter amplicon lacking the barcoded primer binding site. The off target amplicon is therefore not further amplified in the second step of the PCR. With the blue arrow are depicted the universal primers, with the red segments are depicted the blocking oligos
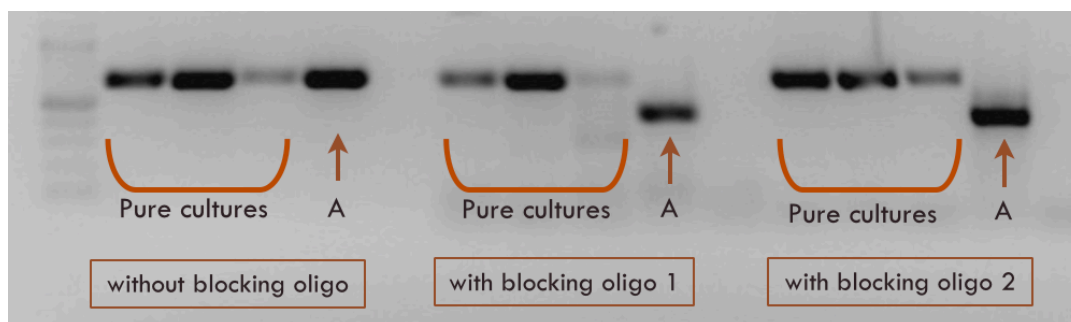
Figure 2. 1; The presence of oligos determines the formation of the shorter amplicon only for A.thaliana DNA, but not for overlooked microbial eukaryotes DNA. PCR with the adding of blocking oligos, on the left the control. Letter A indicates the well in which is ran Arabidopsis thaliana DNA, among the pure cultures used were Vanella sp., and Cercomonas sp., see materials and methods for further details.

communities like bacteria(Martiny et al., 2006). The approach I designed is aiming to prevent the amplification of the host DNA by selectively and physically blocking its complete amplification. This has the side effect of favouring microbial eukaryote rDNA

over host rDNA amplification, by reducing the competition for amplification among the fragments, rendering *de facto* microbial eukaryote rDNA more "detectable" for the *taq* polymerase, and therefore more amplifiable. The approach followed to achieve this goal included the design of the so-called host DNA blocking oligos: DNA sequences matching specifically the 18S of *Arabidopsis thaliana*, able to align within the fragment flanked by the universal primer sites with much higher specificity than primers themselves. This interference leads to amplification of a smaller fragment that lacks universal primer binding sites. In the second step of the PCR, the smaller, off target fragment is therefore lost and not further barcoded (fig 2.1).

This process allows to lower significantly the amount of barcoded plant amplicon in the two step PCR process.

In order to design the oligos, I started from *A.thaliana* 18S. I excised in silico random oligos 35 to 40 bp long. I then proceeded to select the most effective oligos based on careful evaluation of melting temperature, GC content, but most importantly specificity of the oligos to *A.thaliana* 18S itself. I then tested the candidate oligos by adding them to the PCR reagents (see M&M) and by testing them on *A.thaliana*, as well as on DNA from pure isolates of overlooked microbial eukaryotes (see M&M). The expected outcome of this test is to see bands of equal length for microbial eukaryotes tester with or without the addition
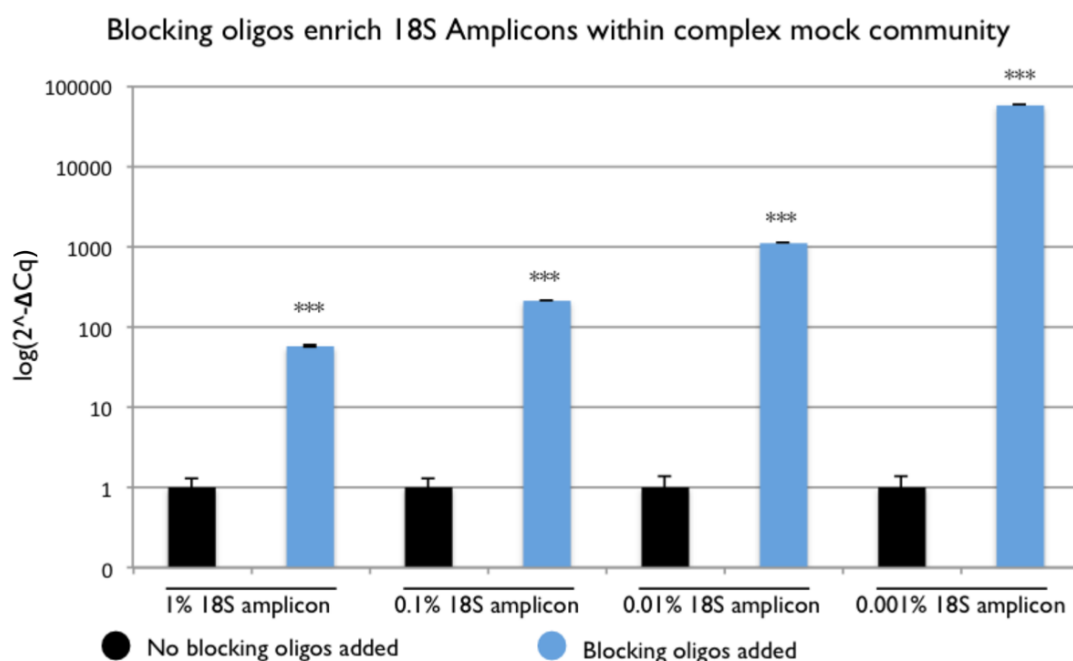
Figure 2. 3; The presence of the oligos in the amplified mock community determines a higher amount of target microbial eukaryote 18S amplicon. Results obtained through qPCR on the amplicons of the 2step PCR. Percentages indicate the initial concentration of tester 18S amplicon (S.boulardii) before the 2step PCR. Colors indicate the adding of oligos (blue), and the non-adding (black). Each treatment and mock has been repeated in three technical replicates. stars indicate significant p-value according to Wilcoxon-Mann-Whitney test, in this case, all the treatments show a pvalue < 0.001

of blocking oligos. Instead, for *A.thaliana* full length amplicon is expected (with universal primer binding sites) only in case oligos are not added. With the addition of the oligos, we should see only the shorter amplicon depicted in fig 2.1 and no full-length amplicon (with universal primer binding site).

The results in fig 2.2 confirm the ability of these oligos to effectively block the amplification of full-length *A.thaliana* amplicon.

However, the efficiency of blocking oligos in a realistic context (like DNA populations extracted from the field) remains an open question. In order to test the efficiency of the candidate blocking oligos, I prepared different DNA mock communities, including bacteria DNA, plant DNA, and different concentrations of *Saccharomyces boulardii* DNA (fig 2.3, x axis -18S amplicon-) as a microbial eukaryote probe (see M&M). After conducting a normal two step PCR for library preparation (see M&M), I quantified *S.boulardii* amplicons through qPCR. Figure 2.3 shows how the usage of blocking oligos increases 18S amplicon (*S.boulardii*) up to 100.000 fold. These results show that the designed oligos are particularly suitable for low starting concentrations of microbial eukaryote DNA to significantly decrease off target DNA amplification. For these reasons this method can be effectively used for heterogeneous taxonomic groups such as microbial eukaryotes, as well as rare taxa analysis (see chapter 2.4).

2.1.2 Shedding light on leaf-associated overlooked microbial eukaryotes: more than 14 subdivisions compose the mosaic

This experimental setup deployed on the geographical sampling allowed to discover a completely unexplored diversity within the leaf habitat. After quality filtering, there were clustered 3725 OTUs (Operational Taxonomic Units) whose taxonomy was assigned to more than 300 species. According to current taxonomy, more than 14 macro taxonomical groups (subdivisions) of overlooked microbial eukaryotes are represented. The recovered diversity is summarized in fig 2.4.

For the endophytic compartment, unclassified reads represent 39% of the recovered reads, another 48% are represented by Metazoa, leaving 13% of the reads being represented by all the other 13 taxonomical groups. The epiphytic compartment is composed by less than 15% of unclassified OTUs, while more than 40% is represented by Metazoa, leaving circa 40% to the known taxa. These 40% are largely dominated by Chlorophyta and Cercozoa (see supplementary figure 1)

These results are the very first indication of the presence of taxonomical groups never
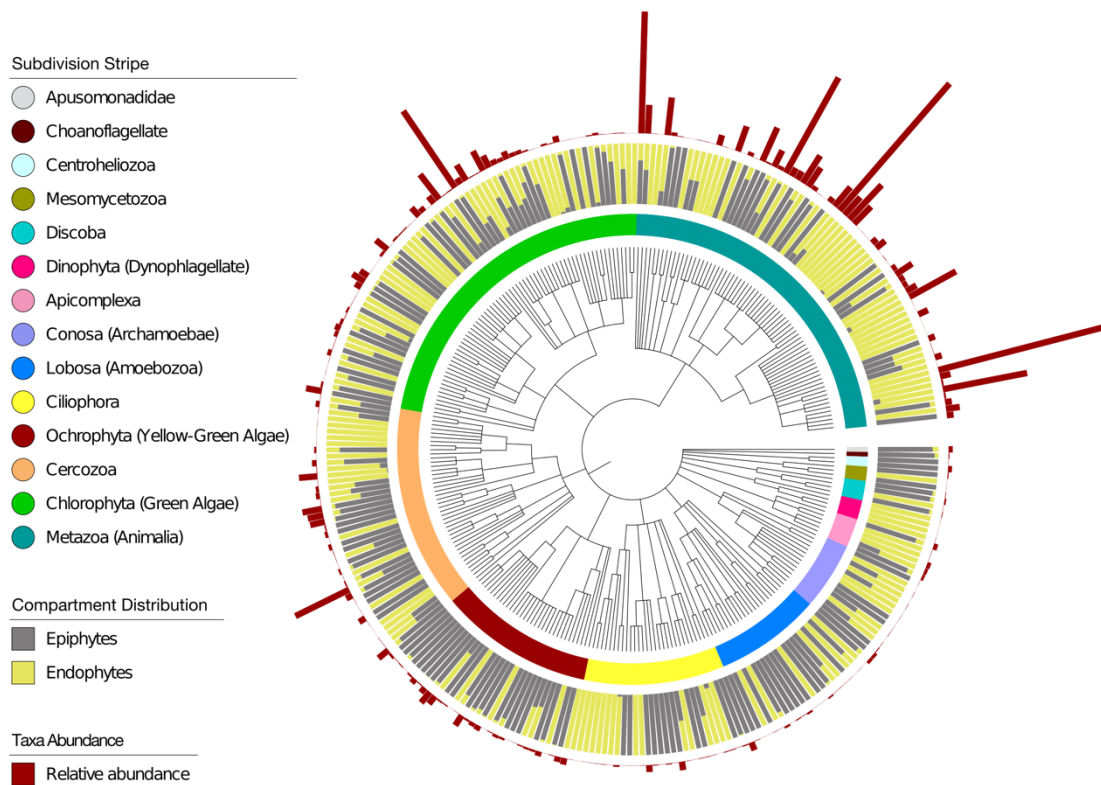
Figure 2. 4; Taxa based phylogenetic tree displaying the recovered overlooked microbial eukaryotes diversity. Each leaf represents a species, the Subdivision stripe represents the major taxonomic groups to which each leaf belongs, the compartment distribution stripe represents the compartment distribution, calculated as percentage of occurrence in endophytic or epiphytic samples: endophyte (yellow) or epiphyte (grey), and finally a relative abundance stripe, in which is depicted the relative reads abundance overall.

associated before with the plant holobiont. The presence of parasitc groups like Apicomplexa, and of predators like Ciliophora or Lobosa opens new chapters in plant-associated microbiology and underpins a potential predator-prey relationship taking place on the leaf.

On the other hand, the presence of multiple photoautotrophic groups (Chlorophyta and Ochrophyta) also inside the mesophyll appears surprising and triggers the biological question of what could be the role of photosynthetic organisms in a host that already performs photosynthesis.

2.1.2.1 Connected microbes: overlooked microbial eukaryotes as a fundamental part of microbial food webs

In order to obtain a picture of leaf-associated microbiome as a whole, I also had to include a focus on bacteria, fungi and oomycetes in my setup. This allowed me to deeply and widely analyse biotic interactions including overlooked microbial eukaryotes, which have remained neglected for years.

The goal was to have a detailed dataset of microbial composition of the compartment, across sites and location, in order to draw meaningful and solid conclusion on biotic interactions on the leaf.

To achieve this I added the following components to my experimental setup: 16S rDNA (V3 and V5) for bacteria, and the rDNA intergenic sequence ITS (1 and 2) for fungi and oomycetes. After having assessed an almost complete overlap between the couples of loci in terms of recovered diversity, I decided to restrict the analysis on the loci that were providing more resolved and rich diversity (Agler et al., 2016a). This means that the results displayed from this paragraph onwards are coming from bacteria V5, fungal ITS2, oomycete ITS1 (see M&M).

The recovered bacteria diversity is composed of 11224 OTUs assigned to 24 phyla. It appears dominated by Proteobacteria and Actinobacteria, followed by Firmicutes and Bacteroidetes as shown in fig 2.5. Among these four phyla, Proteobacteria are by far the most abundant one, counting 5644 OTUs alone, assigned to 326 species. With the exception of Firmicutes, mainly found in endophytic compartment, Proteobacteria, Actinobacteria and Bacteroidetes are equally distributed across epiphytic and endophytic compartment. The fungal scenario appears less diverse, with 8503 OTUs assigned to 6 phyla only. Ascomycetes are the most prominent, showing 5195 OTUs clustered in 760 taxa, Basidiomycetes follow with 1019 OTUs clustered in 290 taxa.

Finally, oomycetes with 2776 OTUs show the lowest diversity, with only 55 taxa represented.

The results shown in this paragraph are in line with what has already been described and reviewed in previous studies (Agler et al., 2016a; Hassani et al., 2018; Vorholt, 2012).

Adding a broad and accurate amplicon sequencing of microbial eukaryotes makes it possible to investigate previously non-targeted inter-kingdom interactions, as well as classical ecological questions such as the role of primary producers, consumers and decomposers on a microbial scale.

Moreover, it is now possible to tackle other biologically meaningful questions: how are
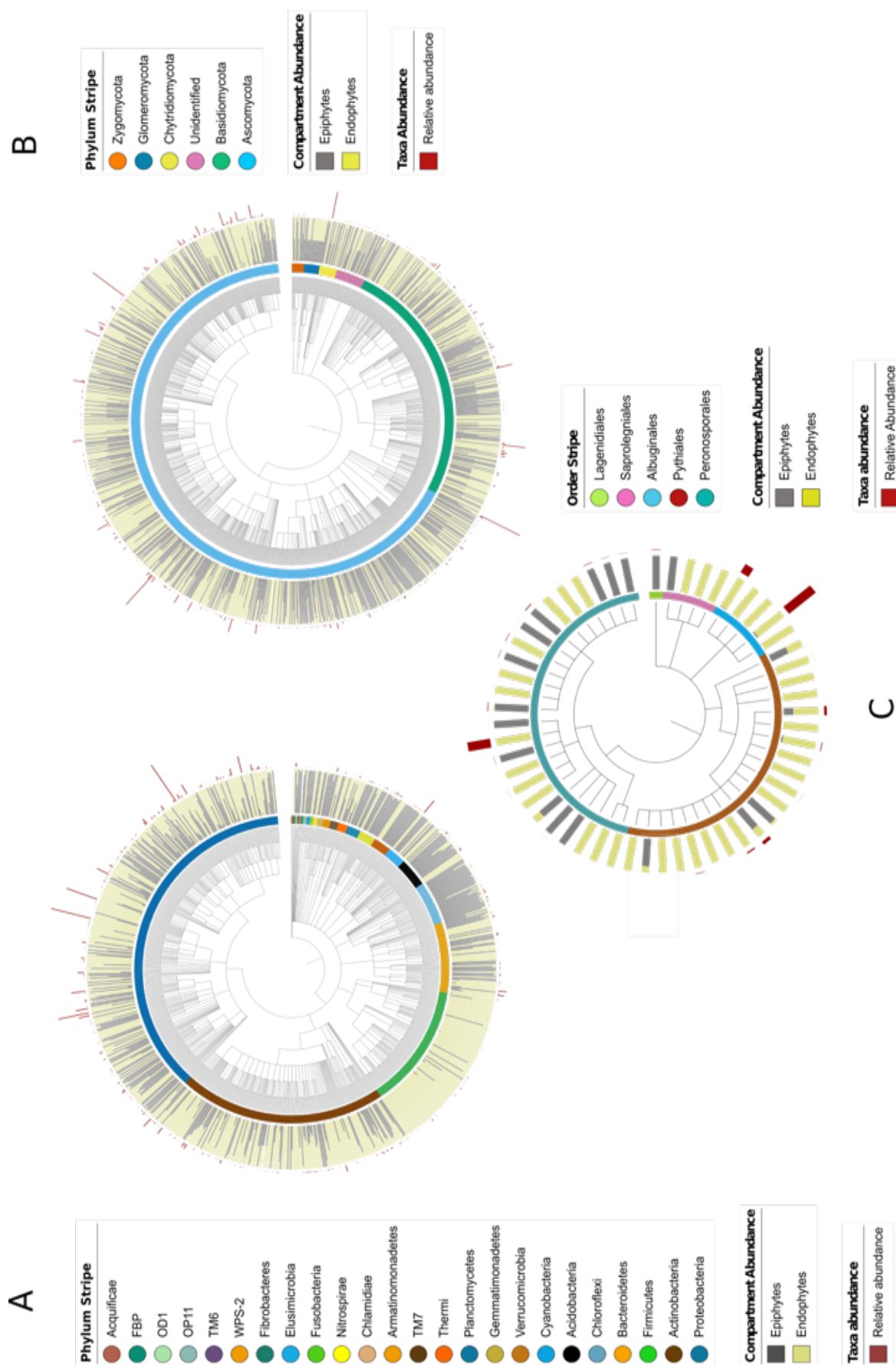
Figure 2. 5 Taxa based phylogenetic tree of the three target taxonomic groups of this study, Bacteria (A), Fungi (B), Oomycetes (C). Each leaf represents a species, the kingdoms stripe represents the major taxonomic groups to which each leaf belongs, the compartment distribution stripe represents the compartment distribution, calculated as percentage of occurrence in endophytic or epiphytic samples: endophyte (yellow) or epiphyte (grey), and finally a relative abundance stripe, in which is depicted the relative reads abundance overall.

microbes distributed across the sites and locations? In other words, is the biogeography of organisms influenced by habitat or province factors? How is the genotype of the host involved in shaping the diversity of the hosted microorganisms? And how important is the presence of other microbes in shaping any microbial population

2.1.3 Biogeography of leaf microbiome: habitat and province shape microbial diversity depending on the taxonomic group

I decided to tackle the aforementioned questions by investigating the role of geographical
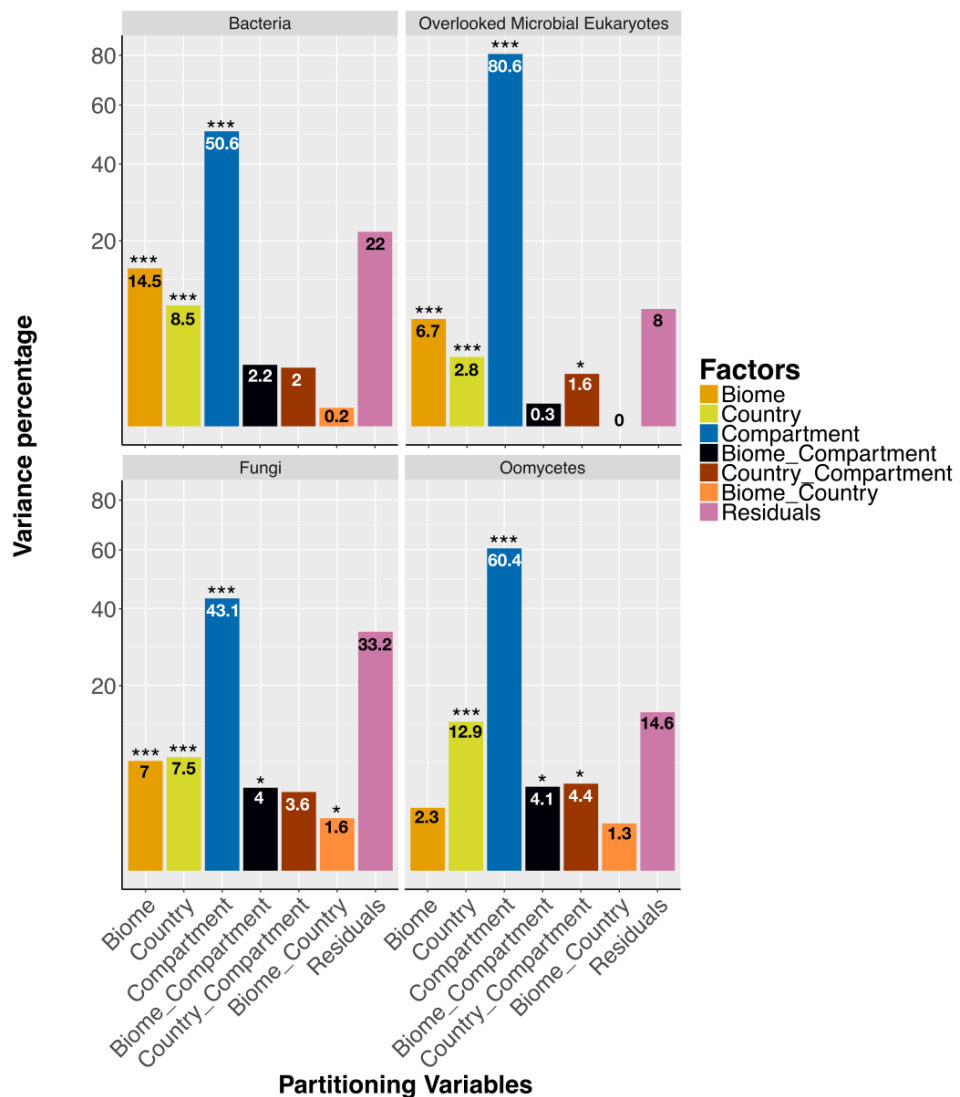
Figure 2.6 The compartment microhabitat is the main shaper of microbial diversity. Biogeography impacts microbes differently according to their taxonomy. Biome and Province (Country) have differential effects over microbial diversity. Obtained by ANOVA test on alpha diversity of bacteria, fungi, oomycetes and microbial eukaryotes.

province versus microhabitat on the community assembly. As a first step, I classified each site as natural, suburban, urban, or agricultural. This classification took into consideration the history of the soil, but mostly the proximity to anthropogenic impact. (see materials and

methods for a detailed description per site). Contextually I considered the Country as a variable able to group sites connected by the shortest possible distance. Moreover, the variable Country was usually encompassing groups of similar latitudes, with the only exception of Sweden (see introduction).

Following this classification, it is possible to draw the first picture of possible main biogeographical drivers of microbial assembly. Country mainly represents geographical distance and is therefore linked to microbial dispersal ability, and the biome represents microhabitat features such as habitat-related biotic interactions or microclimatic variables. I quantified the impact of country and biome on microbial assembly by performing ANOVA on alpha diversity tables of bacteria, fungi, oomycetes and overlooked microbial eukaryotes respectively. Results of this analysis are shown in figure 2.6. The microhabitat represented by the mesophyll is by far the main correlator with microbial alpha diversity for all the microbial groups considered: bacteria, fungi, oomycetes and overlooked microbial eukaryotes, which is in line with previous observations(Cardinale et al., 2015; Vorholt, 2012). These results find confirmation in the Principal Coordinate analyses, which finds a clear separation between the compartments but unclear clustering due to biome or geographical location (Supplementary_video1-4, https://github.com/AlfredoMari/chetproject0001/tree/master/Results). However, OTUs were defined epiphytic or endophytic based on their unique presence in one or the other compartment (see M&M). Therefore, this finding may be considered with caution. It is interesting however to notice that overlooked microbial eukaryotes exhibit higher percentage of variation accounted for the compartment compared to other taxonomic groups.

It seems that the biogeography can have a variable impact on microbiome from group to group. Bacteria and overlooked microbial eukaryotes are more affected by biome than by latitudinal distance. Fungi are equally affected, whereas for oomycetes there is a clear latitudinal trend prevailing on biome related factors.

These findings suggest that, with the exception of oomycetes, geographical distance plays a marginal role in shaping microbial diversity. This diversity appears to be more affected by biome-related variables such as substrate and soil features, as well as microhabitat features like leaf compartment. This major relevance of local factors disconnected from latitudinal, large scale features leads to the question of what other local factors may play a role in shaping microbial diversity. A plausible hypothesis is that microhabitat, as well as

biome, can be drastically influenced by biotic factors, such as microbial interactions (Wisz et al., 2013).

This triggers a couple of working questions: how are biotic interactions shaped after including overlooked microbial eukaryotes? How are typical latitude-related features (temperature, rain) involved in shaping microbiome? In order to answer these questions, I inferred a microbial network starting from single OTUs, including features like year of sampling and environmental information collected on the sites (see M&M).

2.1.4 Fitting overlooked microbial eukaryotes and environmental features in microbial networks provides insights on biotic interactions on the phyllosphere

The geographical sampling involved several locations at different provinces, highly different regarding climatic variables and proximity to anthropogenic pollution. This makes each site likely to have a unique composition and evolution of the phyllosphere microbiome.

In order to investigate the interplay between environmental variables and OTUs I inferred a scale free network, using the CoNet platform (See M&M) with stringent cutoffs chosen after recent studies (Faust and Raes, 2016) in order to minimize background noise and consider only solid interactions after permutations. I inferred the network using OTUs from all the analysed markers (18S, Bacteria, Fungi, Oomycetes). Only consensus edges holding a p-values lower than 0.05 after 1000 permutations and 100 bootstrap iterations were considered.

2.1.4.1 Chlorophyta and Metazoa are key nodes in microbial networks

The resulting network consists of 786 nodes and 25934 edges, (fig 2.7). It appears sparse, with a density of 9.2% of actual connections over potential connections. Negative interactions prevail among inter-kingdom connections, whereas positive interactions are mostly intra-kingdom. By evaluating the betweenness centrality score *g*, two nodes can be identified as the most influential hubs: one belonging to the family of Nocardiaceae and the other to the family of Pseudonocardiaceae. In terms of degree (number of connections) the highest node belongs to an unknown fungal OTU (741 connections), followed by Saprospiraceae (715 connections). Nodes such as Sphingomonadales, and *Albugo sp.* (~250 connections) appear among the top 20 organisms with highest degree (supplementary table 2, https://github.com/AlfredoMari/chetproject0001/tree/master/Results).

Overlooked microbial eukaryotes show far lower degree if compared to the nodes mentioned above, with on average 60 edges each. Overlooked microbial eukaryote nodes are assigned to Metazoa, and Chlorophyta, two of the most abundant subdivisions among

Ranking (Betweenness Centr.based)	Lineage	Degree	Betweenness Centrality
1	OTU-FITS2denovo77527	741	0.2297512
2	OTU-BV5denovo18472	715	0.18311814
3	Saprospirales	523	0.05423583
4	Saprospiraceae	523	0.05423583
5	Saprospirae	523	0.05423583
6	OTU-BV5denovo14812	455	0.04564049
7	Ellin517 (Verrucomicrobia)	320	0.00863405
8	Pedosphaerales	320	0.00863405
9	Verrucomicrobia	320	0.00863405
10	OTU-BV5denovo34137	320	0.00863405
11	Pedosphaerae	320	0.00863405
12	*Kaistobacter sp.*	319	0.01216833
13	*Albugo sp.*	275	0.00582929
14	OTU-BV5denovo572867	250	0.00473542
15	OTU-BV5denovo300126	231	0.0018659
16	*Albugo sp.*	228	0.00297411
17	*Sphingomonas sp.*	228	0.00168686
18	OTU-BV5denovo556005	221	0.00147127
19	*Sphingomonas sp.*	218	0.0016739
20	OTU-BV5denovo439380	215	0.00126771
323	*Bracteacoccus sp.*	61	0.06

Table 2.1 Bracteacoccus is not in the top 20 nodes classified by degree, however, has high betwennesss centrality. Extract from Network Metadata. Full table in Supplementary table

18S pool after unclassified. Two nodes are assigned to *Bracteacoccus sp.,* (Chlorophyta).

As shown in figure 2.7 all of microbial eukaryotes show only negative inter-kingdom correlations, with no connection between each other. These connections are limited to some bacteria (Sphingomonadales, Methylobacterium), and several oomycetal nodes. Correlation between Sphingomondales and Chlorophyta is also confirmed by a linear regression analysis which displays linear correlation between Chlorophyta and Sphingomonadales abundance (supplementary figure 2). Interestingly, this correlation takes place only endophytically.

Figure 2.7; Overlooked Microbial eukaryotes connect with the major hubs of microbial network, environemtnal variables do not connect with singular nodes. The network has been inferred from all OTU tables, odd microbial eukaryotes, Bacteria, Fungi, Oomycetes. each circle represents a node, the dimension of the node is proportional to the node degree. Nodes have the following color code: BLUE - bacteria, ORANGE - fungi, MAGENTA - oomycetes, GREEN – Overlooked microbial eukaryotes, LIGHT GREEN - unclassified, GREY – environmental variables. Connections between the nodes represent the edges, GREEN – positive, RED – negative. In white are labelled the nodes connected to overlooked microbial eukaryotes The network has been inferred using four different metrics: Kullback-Leibler, Bray-Curtis, Pearson, Spearman. The thickness of the edges is proportional to the number of metrics determining the edges.

2.1.4.2 *Bracteacoccus sp.* hampers connections towards potential pathogens such as *Pseudomonas viridiflava*

In the previous paragraph, I briefly mentioned the presence of *Bracteacoccus sp.* as a node in the inferred microbial network. However, the node assigned to *Bracteacoccus sp.,* is not in the top 20 nodes with highest closeness centrality or high degree (see supplementary table 2 https://github.com/AlfredoMari/chetproject0001/tree/master/Results).

At a first glance this would indicate a not crucial role of *Bracteacoccus sp* within microbial consortia. I argued otherwise, since interestingly, despite its low degree, *Bracteacoccus sp* shows high levels of betweenness centrality, index of centrality of a node in the network.

Inferring network on the overall OTUs for all the samples gives us an accurate picture of microbial interactions. However, this does not provide enough information on how much of the network potential rewiring is due to the presence of a single node, in this case
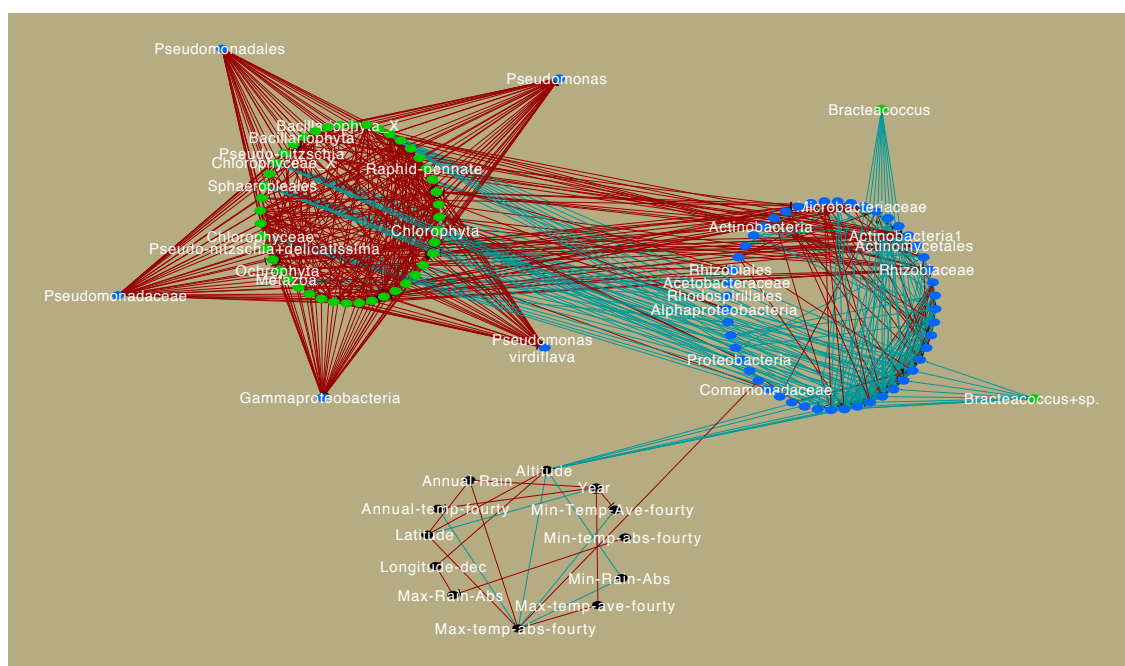
Figure 2.8 The qualitative comparison between +Brac and –Brac network reveals that pathogens like P. viridiflava are absent in networks including Bracteacoccus sp. Directed differential network based on the comparison between -Brac and +Brac using the Cytoscape app Diffany. Incoming red edges indicate decreased connections in +Brac network compared to –Brac. Incoming green edges indicate viceversa increasing connections in +Brac network compared to –Bra.

Bracteacoccus sp. In order to tackle this, I inferred two networks, respectively considering only samples without (-Brac) or with (+Brac) *Bracteacoccus sp.* using the already mentioned CoNET platform. The standard measurements of the two networks provided important details such as the number of nodes and edges: 114 nodes/ 1261 edges in –Brac

network and 51 nodes/319 edges in +Brac. The +Brac network appears to be generally denser, with higher clustering coefficient compared to –Brac (0.67 vs 0.53) and more heterogeneous (heterogeneity coefficient: 0.8 vs 0.7).

22% of the nodes overlap between the two networks. Figure 2.8 depicts a differential network after the comparison between –Brac and +Brac taking –Brac as a reference. It is already possible to notice the major differences in terms of nodes and edges, mainly nodes assigned to the order of Pseudomonadales. However, what the picture depicts, does not provide quantitative insights into such comparison.

Currently available tools infact allow to visualize differences across two different networks (DyNet, Goenawan et al., 2016) or to qualitatively visualize differential networks such as the one depicted in figure 2.8 (Diffany, Van Landeghem et al., 2016). In order to provide that quantitative insight, I designed a statistical comparison platform, able to perform an edge by edge comparison which has not been available so far.

2.1.4.2.1 A new pipeline for solid and quantitative network comparison unravels the lack of correlation between Brac+ and Brac- network

In order to quantitatively compare two networks, I designed a platform which had to be customizable, modular, and deployable in reasonable computational times. The strategy was based on converting the two networks into two distance matrices, which can be filled with any quantifiable property of the network edge. Other studies have used directly downstream network-generated distance matrices (Williams et al., 2014) to compare the networks, however this entails using only one edge score (typically the edge weight). Instead my pipeline, starting from the already calculated network, can be fed with virtually any possible edge score. It can be applied to any combination of networks inferred with any pipeline.
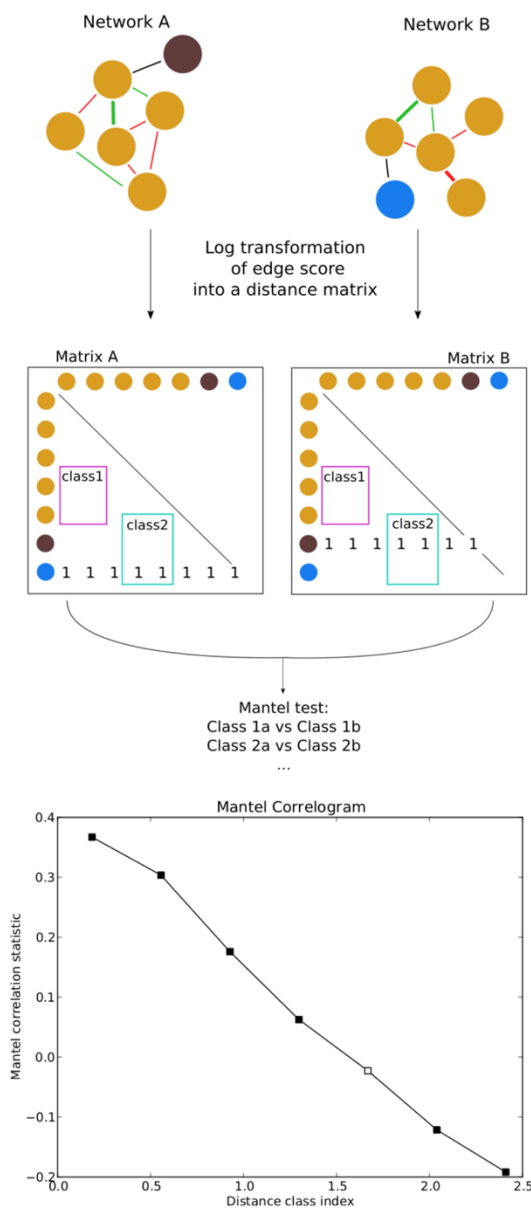
The matrices are then compared using Mantel correlogram test giving a correlation on the comparison of the matrices divided into classes (Legendre and Legendre, 1998), as summarized in figure 2.9. Mantel correlogram is performed through the qiime platform (Caporaso et al., 2010) and allows 500 permutations to calculate a pvalue (Bonferroni correction). Since the comparison is based on distance matrices produced from the edge scores of the networks, the correlation between distance classes reflects the correlation between edge score classes.

Figure 2.9 Workflow scheme of the network comparison pipeline. Nodes composing both networks are kept in both distance matrices, scores are assigned based on which edge score the user want to feed in. Edge scores between differentially present nodes are by default assumed to be not existing, and therefore set to 1. Mantel correlogram compares the distance matrices divided into classes following the Sturge's rule.
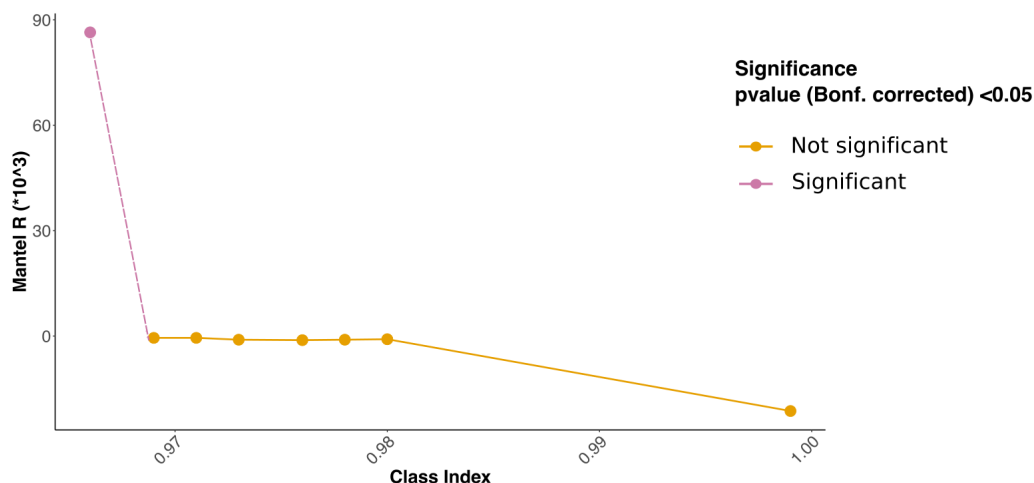
Figure 2.10; Brac- and Brac+ do not correlate in almost any edge weight class. Pipeline deployed +Bracteacoccus and −Bracteacoccus (feeding the edge weight as parameter), on the x axis are displayed the class index, calculated via Sturge rule. The values are transfromed logarithmically. The Mantel R (y axis) spans from -1 (no correlation) to 1, maximal correlation. Each correlation between the classes displays the p-value after 500 permutations, most of them are not significant, indicating no correlation between the two networks.

The platform also provides a tracking table listing which edges compose which edge class. This is a simple way to track which interactions are conserved or not across the networks, but most importantly, how much they are / they are not correlated across the networks. (see M&M).

The platform is coded in PERL language and allows PYTHON codes and R subscripts to run in parallel and in background, using the LSF management platform (Songnian et al., 1993). This way, the computational time needed to perform the comparison is substantially reduced (See M&M). Currently, a comparison between two medium networks (~300 nodes each), depending on network connectivity and number of nodes, harnesses averagely from 10 to 100 sec of CPU time.

I applied this pipeline on the aforementioned *Bracteacoccus sp+/-* comparison, basing the comparison on the edge weight variable. The results are depicted in figure 2.10, showing that most of the edge classes are not significantly correlated. This indicates that the two networks exhibit major divergence and almost do not share a core of conserved interactions, as the previous analysis (fig 2.8) suggested. The tracking table (supplementary table 3 https://github.com/AlfredoMari/chetproject0001/tree/master/Results) shows as main components of the non-correlated classes, edges including Pseudomonadales, Pseudomonadaceae, Pseudomonas and *Pseudomonas viridiflava*, thereby confirming the qualitative analysis performed with Diffany, which was individuating the same genera to be not conserved across the two networks (figure 2.8). Most of them show a corrected p-

value higher than 0.05. The only ones showing a significant pvalue are Gammaproteobacteria and Pseudomonadales. However, they display a Mantel R score still close to 0, indicating a lack of conservation of these interactions across the two networks. The results obtained with this new method largely confirm analysis displayed in figure 2.8 (Diffany-based). Therefore, it is possible to conclude then that the presence of *Bracteacoccus sp* particularly affects the connectivity of pathogens orders like Pseudomonadales, or more specifically pathogens like *P. viridiflava.* With the displayed pipeline, however, it was possible to quantify how solid this hampering of pathogens was. Moreover, the substantial consistency of the presented pipeline with previously known qualitative methods demonstrates its reliability. This method can be therefore deployed in multiple different contexts, on any pairwise network comparison, irrespective of the inferring methods.

2.2 Investigating the role of producers in terrestrial ecosystems

2.2.1 Overlooked microbial eukaryotes shaping microbial diversity:

Bracteacoccus sp., is a keyplayer, together with Metazoa, Lobosa and Ciliophora
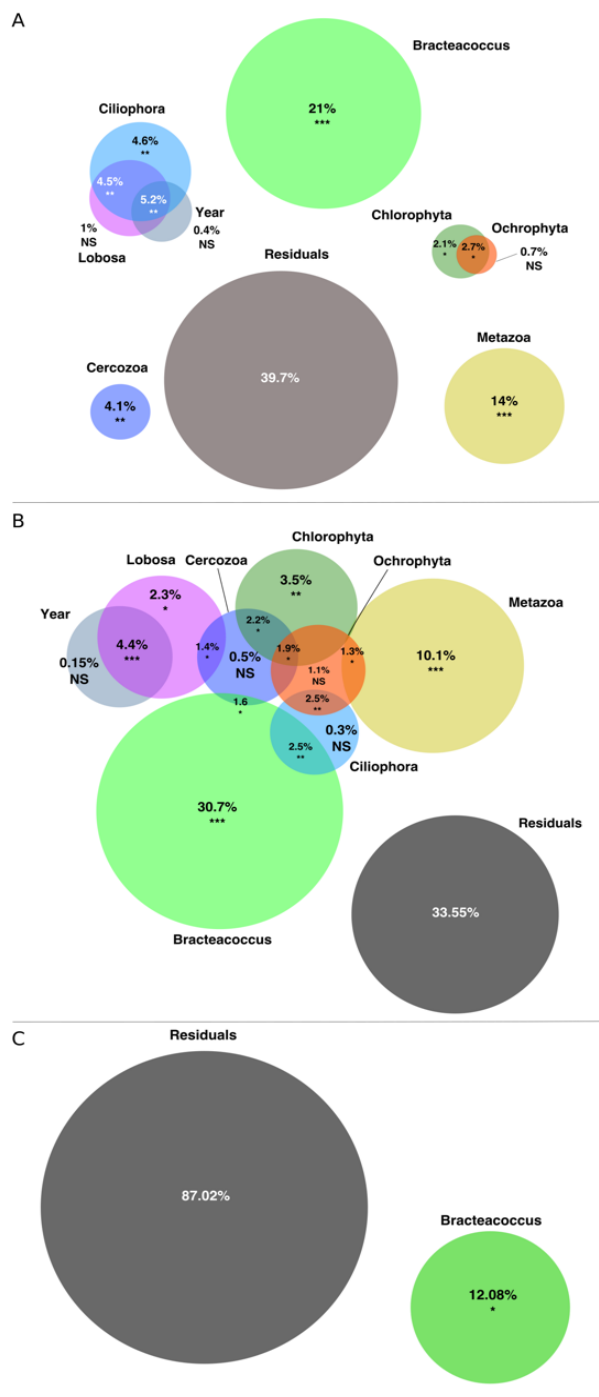
Network inferring made it possible to dissect meaningful ecological interactions between single OTUs. However, this does not allow to delineate conclusions on the overall impact of overlooked microbial eukaryote communities. I hypothesize that, beside the singular interactions displayed in network analysis, overlooked microbial eukaryotes can have remarkable effects on entire microbial communities. In order to disentangle the impact of all overlooked microbial eukaryote

*Figure 2.11a Overlooked microbial eukaryotes have a remarkable impact on alpha diversity of Bacteria, Fungi and Oomycetes. Bracteacoccus is prominent shaping factor in all cases. Bacteria are depicted in **panel A**, followed Fungi (**panel B**), and Oomycetes (**panel C**). **figure** displays ANOVA on alpha diversity of each panel, showing the variation explained (η^2 score) by the presence of each overlooked microbial eukaryotes. Only significant pvalues are displayed. Notation of pvalues follows , *= pvalue <0.05, **= pvalue<0.01, ***pvalue<0.001.*

taxonomic groups, I focused on the alpha diversity of the other three target groups analysed:
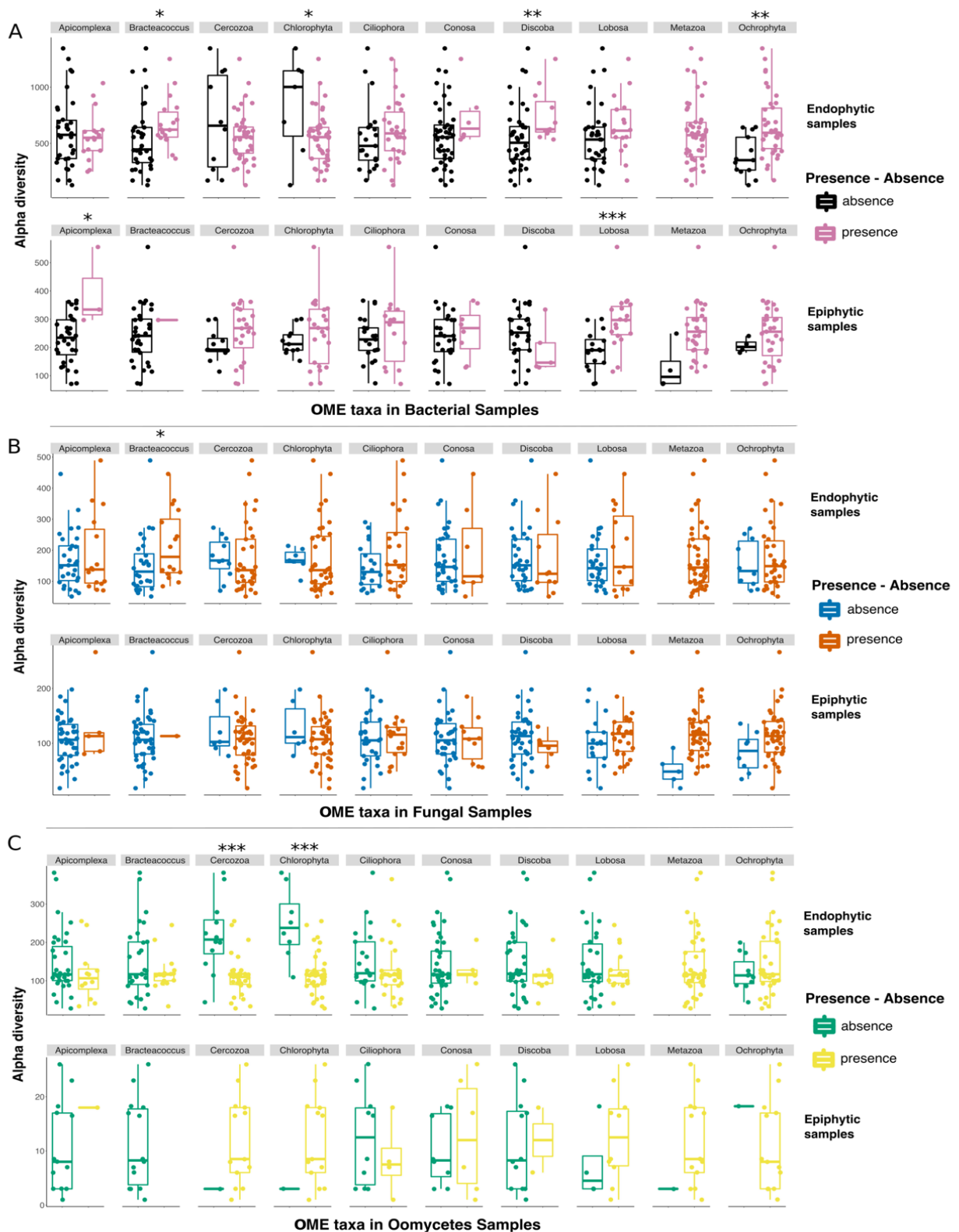
*Figure 2.11b Overlooked microbial eukaryotes (OME) have a remarkable impact on alpha diversity of Bacteria, Fungi and Oomycetes. Bracteacoccus is prominent shaping factor in all cases. Alpha diversity is often increased by overlooked microbial eukaryotes. Bacteria are depicted in **panel A**, followed Fungi (**panel B**), and Oomycetes **(panel C)**. figure depicts alpha diversity panel diversity in Endophytic samples(Up) or Epiphytes(Down) for every overlooked microbial eukaryotes group displayed on top of the boxplot.. P- values are calculated after Wilcoxon-Mann-Whitney test. Notation of pvalues follows *= pvalue <0.05, **= pvalue<0.01, ***pvalue<0.001*

bacteria, fungi and oomycetes.

For each sample, I first established whether or not it contained reads from each of *overlooked microbial eukaryotes* groups that I aim to tackle (Chlorophyta, Lobosa etc), including the candidate *Bracteacoccus sp.*, found to be a key player in network analysis. Based on this criterion, I then labelled all the samples as positive or negative. Using this labelling as a partitioning variable, I was then able to use ANOVA in order to unravel how the targeted diversity variation was distributed across biotic factors, and what is the impact of overlooked microbial eukaryotes in such scenario.

What can be observed for all the target groups is a prominent part of variation explained by *Bracteacoccus sp.,* (in the endophytic compartment) which, in the case of Fungi is responsible alone of more than 30% of variance (fig 2.11).

The major overlooked microbial eukaryotes groups involved in driving endophytic bacteria variation are Ochrophyta and Chlorophyta, together with the microalga *Bracteacoccus sp*, underpinning a major role held by photosynthetic organisms on endophytic bacteria. Notably, also heterotrophic subdivisions such as Apicomplexa and Lobosa seem to harbour differential bacterial diversity (fig 2.11 A2), especially in the epiphytic compartment (fig 2.11 A1). However, no significant variation percentage explained by Apicomplexa was found. Lobosa effect on epiphytes appears remarkable especially since I previously showed Lobosa reads as mainly endophytical. Concerning heterotrophs, Ciliophora and Lobosa, together, explain a large portion of variance. However, this joint effect is largely connected to year variation as well. Conversely Cercozoa hold 4% of variation independently from other taxa or variables.

Concerning fungal diversity, apart from the nestedness of variation accounts, patterns are similar to those of the bacteria scenario, with Lobosa having an impact only in defined years, and *Bracteacoccus* holding the major percentage. For oomycetes, only *Bracteacoccus* steers a significant variation portion.

What appears clear from my analysis on amoeboid organisms is a variable but ubiquitous portion of variation accounted for Ciliophora, Lobosa, Metazoa and Chlorophyta. Overlooked microbial eukaryotes seem in general to determine the increasing of bacterial diversity as well, with the exception of Chlorophyta. It would be possible then to connect the presence of overlooked microbial eukaryotes to a higher stability and resilience of the community due to the enhancement of alpha diversity (Elmqvist Thomas et al., 2003; Lozupone et al., 2012).

2.2.2 Chloroplastic DNA confirms presence of algal reads in the

endophytic compartment

The previously shown findings of the major role of algae, made it imperative
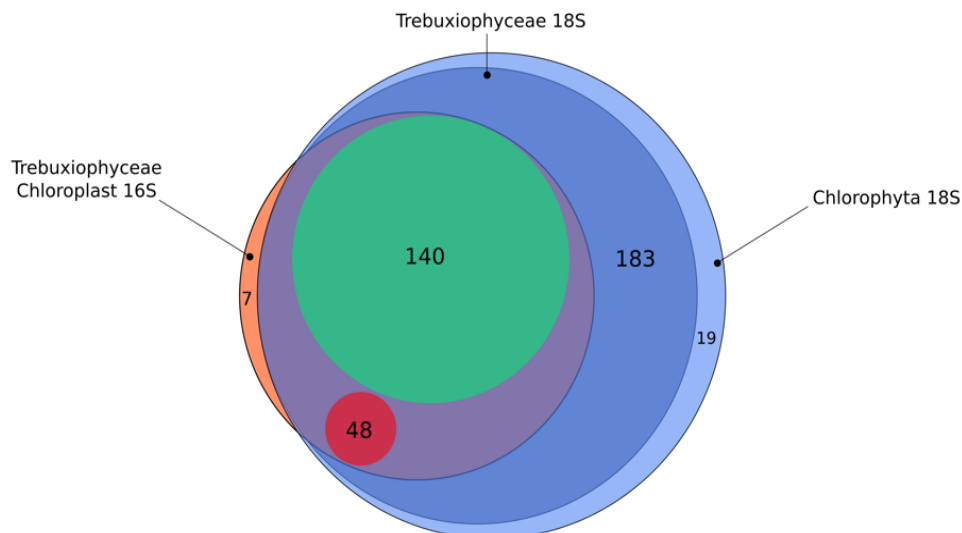
Figure 2. 12a A large part of samples contain algal 18S but not chloroplast 16S, introducing the hypothesis of active and inactive algae. The figure displays the number of samples which overlap between Trebouxiophyceae 18S and algal chloroplastic 16S, numbers inside the venn diagrams represents the number of samples. The so-called "Active Trebouxiophyceae" samples (GREEN) contain both chloroplastic 16S and Treb 18S, the "Inactive Trebouxiophyceae" (in BLUE) contain 18S but not 16S, finally in RED are depicted the absent, the samples in which no Trebouxiophyceae 18S nor chloroplastic 16S was found

to better understand algal distribution across the leaf compartments. In order to do so, I checked for the presence of algal chloroplast DNA within samples containing Chlorophyta 18S.

Within Chlorophyta, more than 60% of the reads belong to the phylum of Trebouxiophyceae. I therefore decided to focus on this phylum. In total, 11 OTUs of Trebouxiophyceae chloroplastic DNA, and only 2 OTUs accountable for other chlorophyte genera were recovered. The presence of algal plastid 16S DNA overlaps with 18S in 96,4 % of the 16S samples, supporting the hypothesis of a photosynthetic activity of such organisms.

However, the contrary is not always true the contrary: samples containing algal 18S overlap with chloroplastic 16S in roughly 50% of the total algal 18S samples (fig 2.12a), meaning that another 50% contains Trebouxiophyceae 18S but no plastidial 16S (fig 2.12a). These "algae without chloroplast" are mainly concentrated in the endophytic leaf samples (fig 2.12b). Interestingly, no clear trend is found across the collection sites or across other variables underpinning a minor role of environmental variables and location factors as
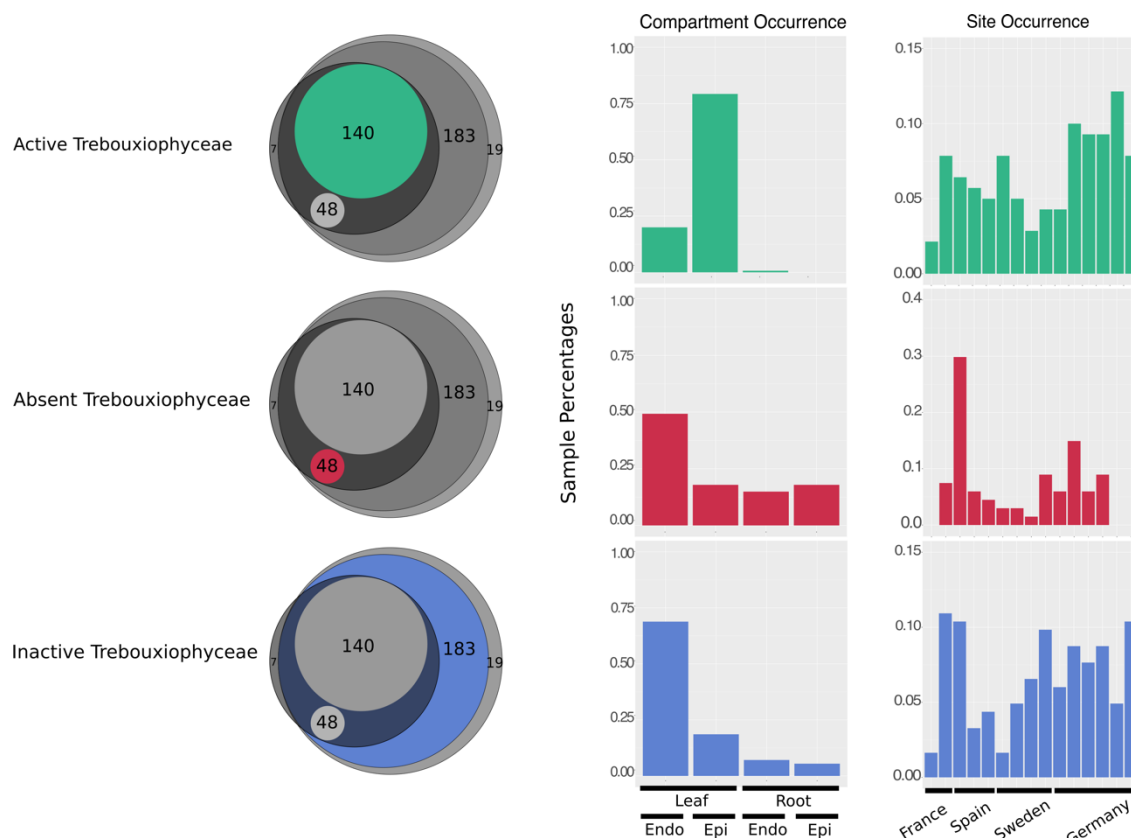
Figure 2. 12b Active Trebouxiophyceae are mainly epiphytic, Inactive distribute mainly endophytically. No clear distribution over site is found. Overlap between Trebouxiophyceae 18S and algal chloroplastic 16S, like in figure 2.12a. Plots on the right side display the distribution of Active, Inactive, Absent Trebouxiophyceae, respectively, over compartment, and over site. Plots display sample occurrence in the indicated condition

shown in fig 2.7. These findings would support the hypothesis of competitive algal photosynthetic activity on the leaf surface and can open contextually new fields of investigation on potential algal interaction inside the mesophyll.

2.2.3 Endophytic lichens belong to Lecanorales order

In the context of discovering potential roles played by the algal community on the leaf, I wondered whether this occurrence of algae in the endophytic compartment could be correlated with the presence of specific orders of fungi, which normally maintain lichen symbiosis with microalgae and cyanobacteria. As previous studies already clarified

(Hawksworth, 1988), fungi of the classes of Leotiomycetes, Dothideomycetes, Eurotiomycetes, and of the orders Helotiales and Lecanorales normally engage lichen symbiosis with certain species of algae of the family of Trebouxiohyceae. I investigated all the classes mentioned above and what is possible to observe is a co-occurrence especially of Lecanorales with the 18S belonging to the algal class Trebouxiophyceae, with very high overlap (up to 98 %).

However, only the just mentioned Lecanorales show a substantial overlap also with algal chloroplastic 16S, of more than 70% of the samples, as shown in fig 2.13. Whereas no sample negative either for Trebouxiophyceae 18S or Chloroplastic 16S, contains Lecanorales reads.

Concerning site distribution, I have observed that these Lecanorales-Trebouxiophyceae samples are preferably distributed in urban sites, such as ERG or EY, located in Germany.
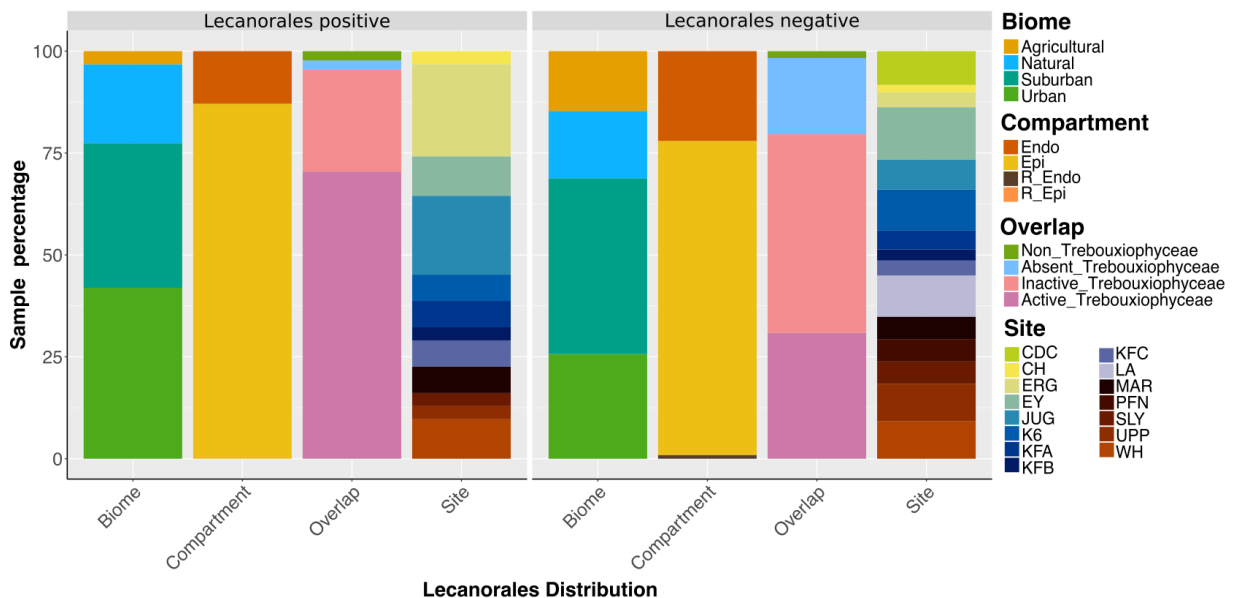
Figure 2.13; Lecanorales distribute mainly across active Trebouxiophyceae. They are distribute mainly across German urban sites, and mainly epiphytically. Distribution of the samples clustered by the presence or absence of the fungal family of Lecanorales.

Concerning compartment distribution, more than 70% belong to epiphytic compartment, while 12% is accounted for endophytic samples (see fig 2.13). Especially in the sites UPP, EY and KFC, (Germany and Sweden), I found reads belonging to the lichen family of Parmeliaceae. All these samples contained *Trebouxia sp* reads as well. Among them, few OTUs appear belonging to the lichen genus *Parmelia.*

This suggests not just a co-presence of algae within the analysed ecosystems, but also a developed network of mutual interactions in which algae can play a major role.

2.2.3.1 Lichenising fungi impact on bacteria diversity only when co-occurring with Algae

In order to further dissect the putative lichen presence on the leaf, I then analysed the impact of Chlorophyta and lichenising fungi on bacteria alpha diversity, via ANOVA. I especially focused on the already mentioned Lecanorales. Notably, Lecanorales show a significant effect only in the samples in which is shared the presence of algal chloroplast. This finding highly underpins the hypothesis of lichen presence and active role in shaping bacterial diversity of the leaf. These findings (fig 2.14) show for the first time that not only can lichens be found on the surface of and inside the mesophyll of a weed like *A.thaliana*, but they can also be a major shaper of leaf microbiome.

Figure 2. 14; Lecanorales have impact on bacterial diversity only when co-occurring with algal chloroplast. First evidence of lichens presence on the leaf. Obtained by ANOVA on the geographical sampling bacterial alpha diversity. The percentages display the η^2 score considering as partitioning variables the presence or absence of the fungal order of Lecanorales, as well as the presence/absence of Chlorophyta as a whole. * indicates p-value <0.05,** a pvalue <0.01, *** a p-value <0.001

2.2.3.2 Dissecting the role of lichen on the leaf. Time course over different ecotypes reveals fluctuations across time, but not across genotype

In order to dissect further aspects of this algal-fungal co-occurrence, I analysed data from a separate experiment, a common garden experiment in which four different ecotypes of *Arabidopsis thaliana* (KSK, WS0, Sf2, Col-0) were planted in the same field over three years, harvested and sequenced monthly (See Materials and Methods). By plotting the

relative abundance of Chlorophyta across different ecotypes we can observe no clear trend,
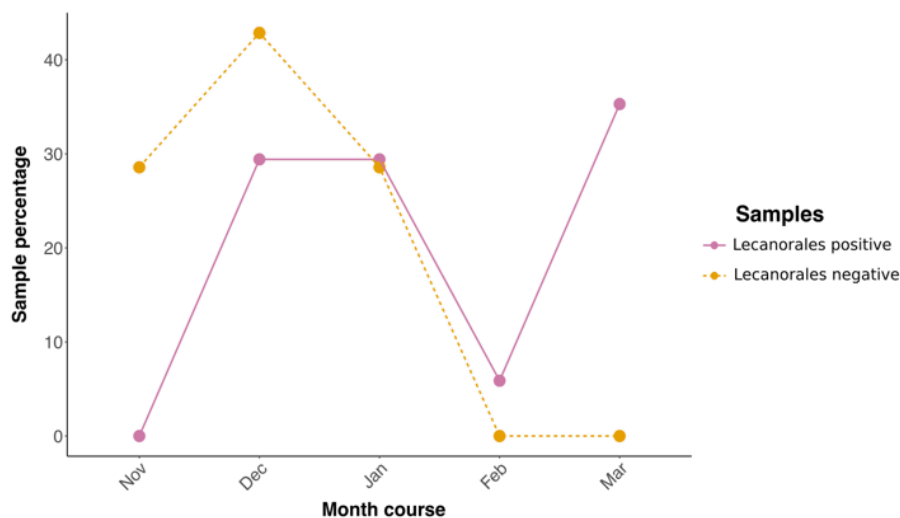
Figure 2. 15; Lecanorales follow a month distribution, with the highest peak in March Obtained through Lecanorales occurrence across time in field experiment - time course.

having all genotypes hosting all the displayed families (supplementary figure 3). Concerning Lecanorales, conversely, we observe an occurrence peak in December, decreasing in the following months (fig 2.15). Given this scenario, I hypothesized that the
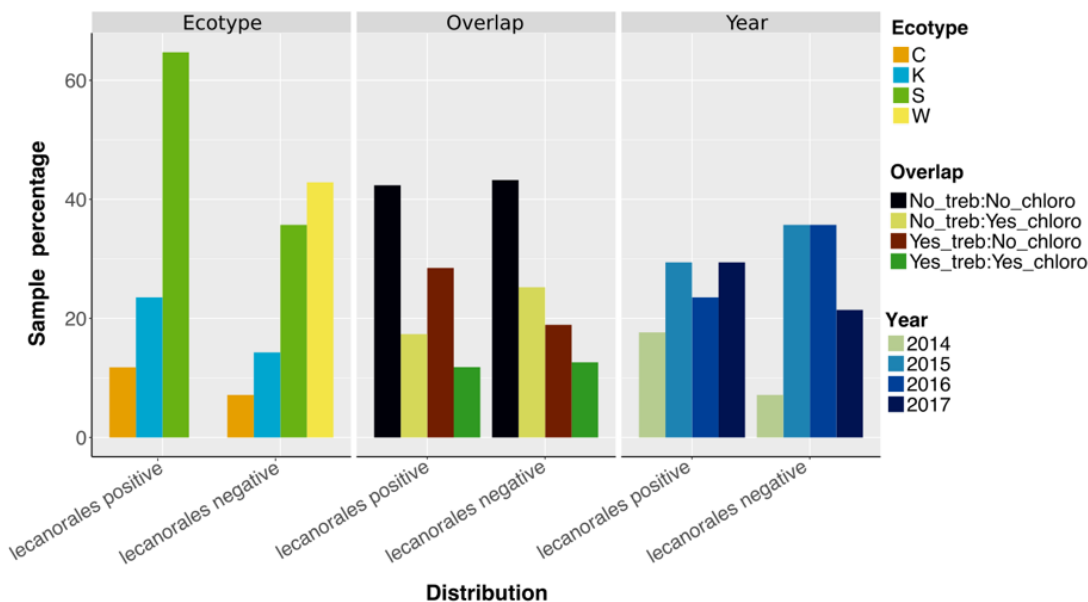
Figure 2. 16 Within the time course, Lecanorales are not distributed consistently compared to the geographical sampling. Lecanorales distribution across the remaining variables, ecotype, overlap (of algal chloroplastic 16S and algal 18S), and year variation. Notably, the ecotype Ws-0 does not show any presence of Lecanorales across all three years.

reason of this shift could be found in the presence of other microbes, therefore we investigated whether or not Lecanorales or Trebouxiophyceae had a similar or stronger

impact on bacteria diversity, as seen previously in the latitudinal sampling. I therefore performed ANOVA on bacteria diversity in the same way as mentioned in the previous paragraph. I also included the presence-absence of algal chloroplastic DNA as partitioning
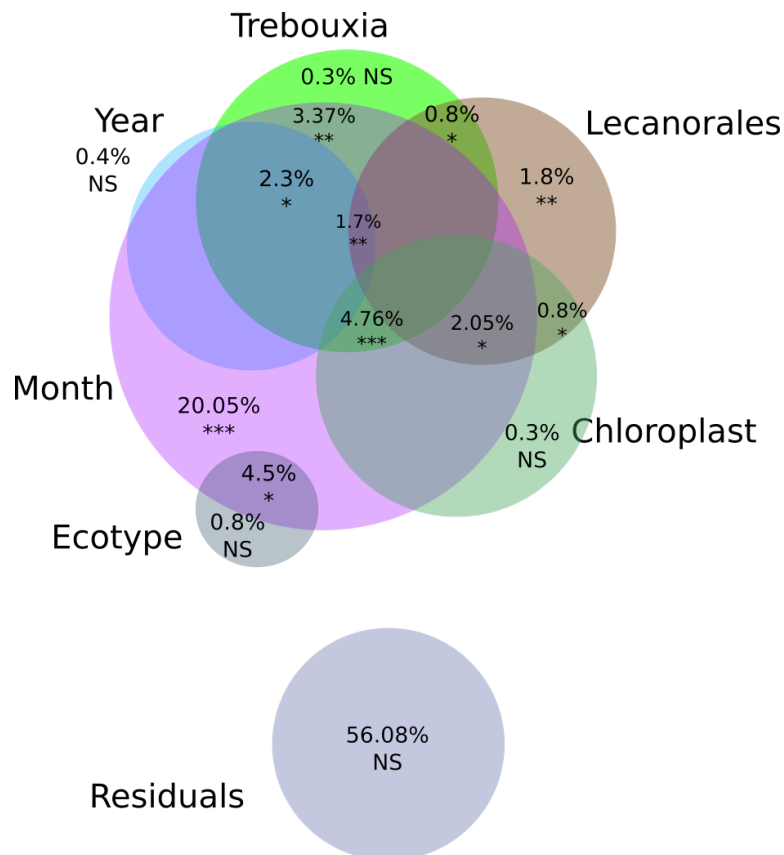
Figure 2.17; Lecanorales and Trebouxia are confirmed to be connected variables in shaping bacterial diversity, irrespectively from ecotype. Evidence of lichens role connected to time successions. Obtained through ANOVA on common garden field experiment bacterial alpha diversity. Percentages display the η^2 score. P-value code: *=p-value <0.05, **=p-value <0.01, ***=p-value <0.001, non- significant p-values are labeled with NS.

variable. The results of ANOVA (fig 2.1) show significant percentages of variation accounted for Trebouxia (more than 13% cumulatively) and Lecanorales (total 7.2%). Remarkably, these results are consistent with the ones found in the geographical sampling (fig 2.14). However, differently from geographical sampling, in the common garden experiment, Lecanorales account alone for a small but significant portion of variation. Another feature worth to mention is that most of effect size of Lecanorales and *Trebouxia sp.* are linked to month variation, underpinning a link to microbial successions within microbiome. These results, together with the latitudinal sampling suggest that the lichen life on the plant and inside the plant might be not stochastic and a potentially important aspect of plant holobiont.

2.2.4 Mechanisms underlying algae role in leaf ecosystem do not involve differential sugar productions

All these results suggest a mutual interaction between algae belonging to the family of Trebouxiophyceae and lichenising fungal partner. However, I still could not exclude that algal presence could have had an impact on the microbiome by actively providing nutrients, mainly sugar.

In order to investigate this issue, I first assumed that high abundant microalgae could influence more nutrient composition (Jones and Mayfield, 2012). I therefore considered the most abundant phylum in the ecological sampling (Trebouxiophyceae – supplementary figure 4), and within it, one of the most abundant and free living organism, *Microthamnion kuetzingianum*.

I decided to investigate the nutrient composition of the leaf surface with a leaf wash experiment (see M&M) in order to reveal whether the application of microalgae on the leaf surface led to a different composition of epiphytic sugar content. It is known that algae are able to form their own biofilm by the secretion of the so-called EPS (extracellular polymeric substances) which typically are composed by sugar (Ramanan et al., 2015a). I hypothesized that *M.kuetzingianum* presence on the leaf could lead to a change in EPS composition of the leaf surface. The results in figure 2.18 show the absolute equality between the control and the two algal replicates tested, indicating that the tested algae do not modify the sugar composition of the leaf surface. However, in replicate MK2 a slightly

higher quantity of glucose is visible. It is interesting to notice though the high basal
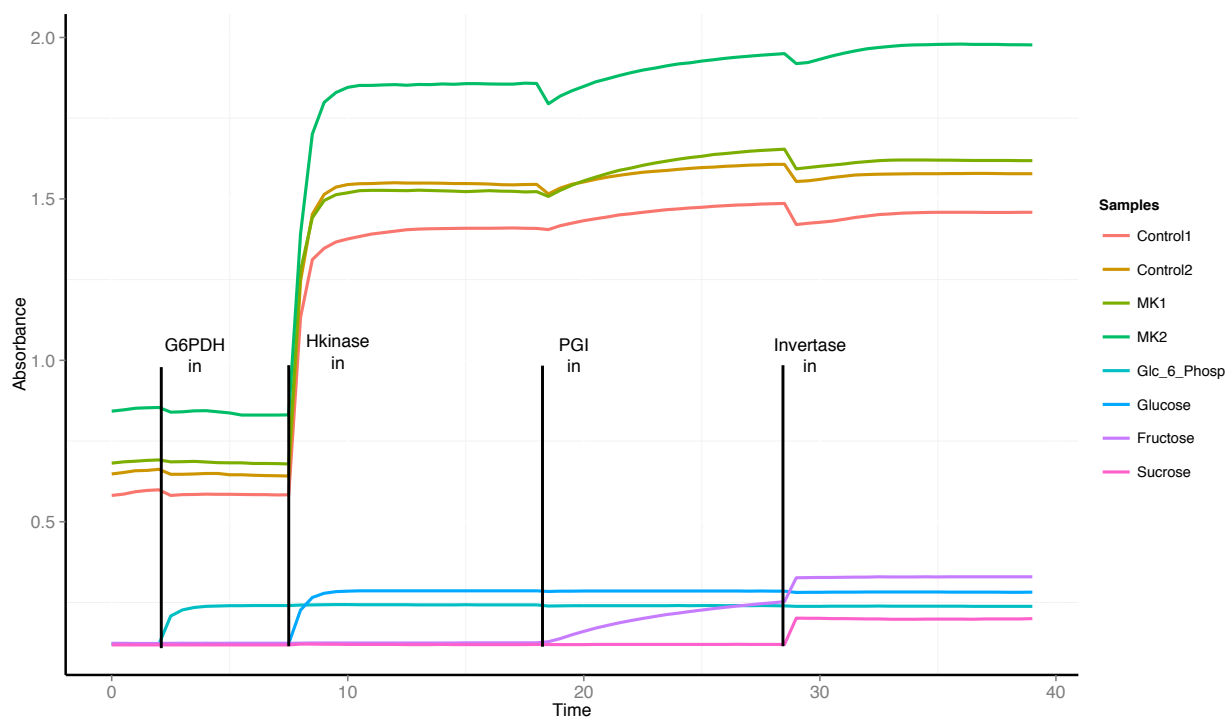
Figure 2. 18 Sugar composition of the leaf surface does not change after the application of M. kuetzingianum. Obtained through enzimatic test on leaf washes for plants treated with Microthamnion kuetzingianum (MK) in two biological replicates (MK1, MK2) and the control, solution of MgCl2 in two independent biological replicates (Control 1 and 2). Separately, technical controls were added, containing sterile solutions of the displayed pure sugars: Glc-6P, Glucose, Fructose, and Sucrose. The increasing of the absorbance of the technical controls after the adding of the correct enzyme (G6DPH for Glc-6-Phosph, Hexokinase for glucose, PGI for Fructose and Invertase for Sucrose –see complete experimental design in M&M--) indicate clearly that the assay is technically reliable.

concentration of free glucose apparently present on the leaf surface, which potentially could be of microbial origin. These results do not exclude the presence of algal derived EPS on the leaf, which is still likely, given the high amount of sugar in general recovered. They instead underpin a marginal role, if any, of *M. kuetzingianum* as possible major constructor of biofilm.

2.3 Dissecting the role of heterotrophs and amoeboid organisms: predators on the leaf surface

As previously shown in paragraph 2.1.2, several subdivisions revealed by 18S sequencing have heterotrophic lifestyle. I also showed in paragraph 2.2.1 that a relatively large part of variation for bacteria, fungi and oomycetes is explained by the simple presence of overlooked microbial eukaryotes groups such as Lobosa, Ciliophora and Metazoa. In order to better tackle this question, I used the time course in order to observe which was the underlying pattern.

2.3.1 Successions of consumers: different time for different predators

In order to dissect time course data, I performed ANOVA in the same way as shown in paragraph 2.2.1, including the presence of presumed producer families as partitioning variables (figure 2.19). Apart from the month and the year, which are the main drivers of diversity, respectively 25% and 10% of variance highly nested with each other, what follows is Lobosa and Ciliophora (fig 2.18).
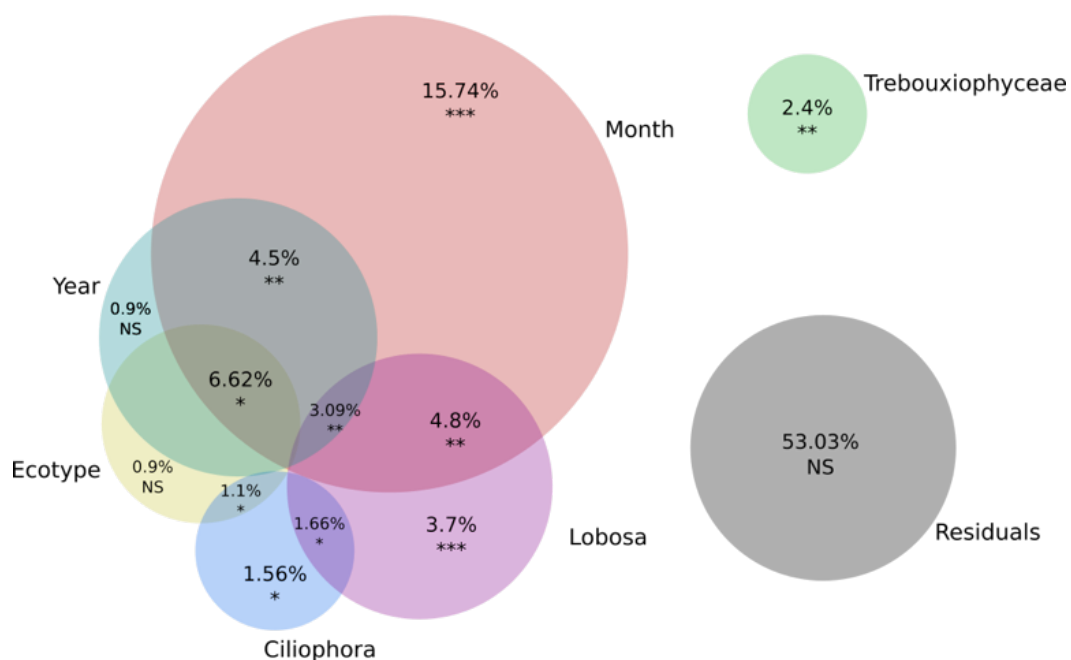
Figure 2.19; Month variation is the key shaper of microbial diversity , its nestedness underpins microbial successions. Obtained through ANOVA on bacteria alpha diversity on the experimental common garden field sampling, percentages display η^2. pvalue code: *=pvalue <0.05, **=pvalue <0.01, ***=pvalue <0.001, non significant pvalues are labled with NS

What is really surprising is that a portion of variation is also explained by the joint variable Lobosa or Ciliophora and ecotype, a variable, which alone has no significant impact on bacteria diversity.

Given the relevant impact of the month variable, I then dissected the composition of Lobosa, Ciliophora and Metazoa over the sampled months. As shown in figure 2.20 for both Lobosa and Ciliophora, in the month of November there is a peak of overall abundance
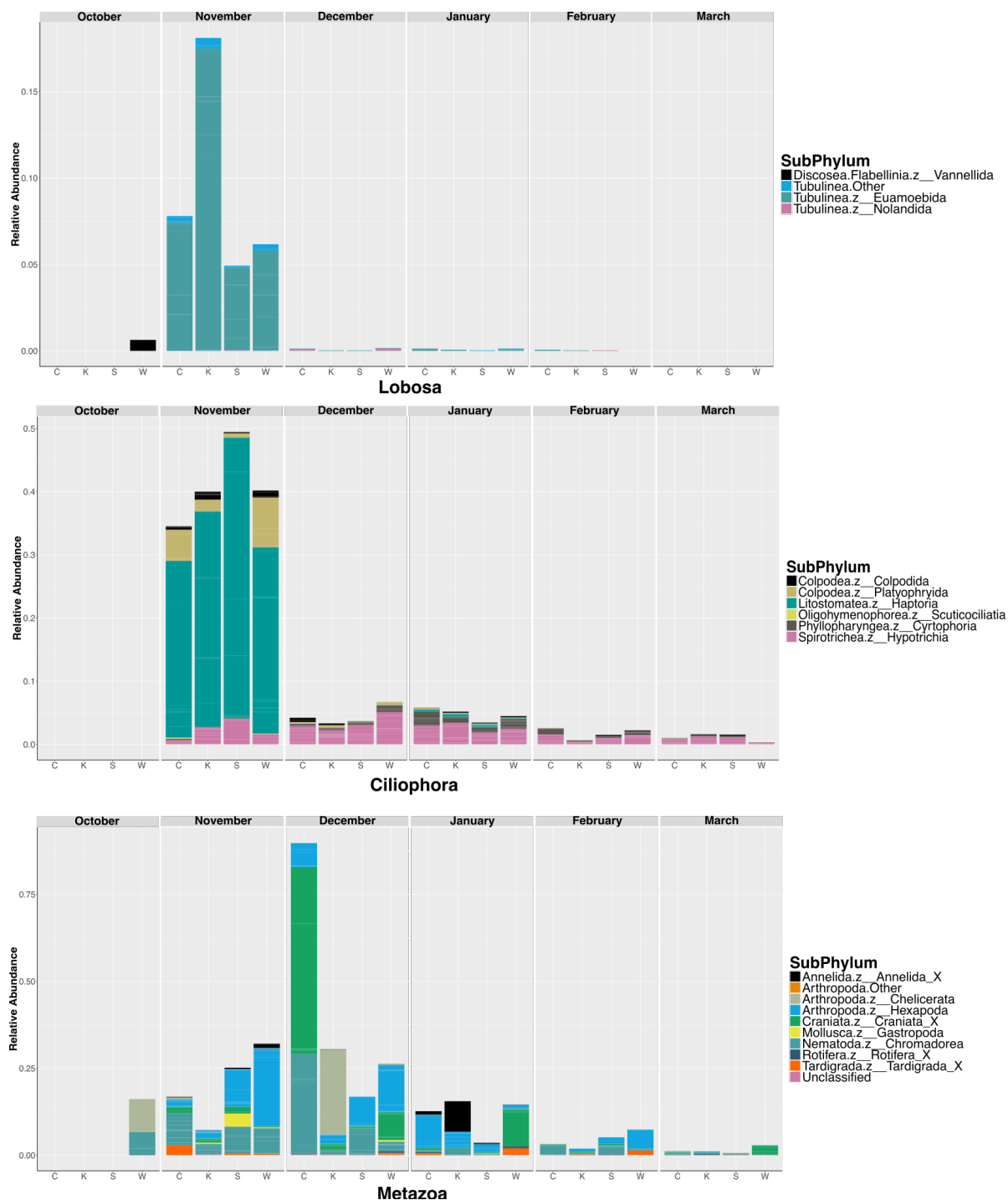
Figure 2.20; Successions of heterotrophic groups underpin a role of Metazoa in common garden experiment. On x axis are clustered the different months of sampling, from October until March, divided in all the ecotypes (Col-O, KsK, Sf-2, Ws-O), each color represent a subphylum, as indicated in the legend.

and diversity, followed by a steep decrease of both in the following months. Whereas, in the same context, Metazoa show a different trend, displaying a peak in December and a following decrease of both abundance and diversity. This major shift between November

and December for Ciliophora and Lobosa, but not for Metazoa, suggests a radical change in biotic interactions rather than a change in environmental conditions, which remained stable across those two months (See Mapfile_Common_garden.xlsx at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M). These findings indicate as well that ecotype has a role only in defining pioneer organisms such as Lobosa and Ciliophora, whereas major predators take over the leaf only in a second moment.

2.4 United they stand: rare taxa play an important role in shaping the diversity, but only if taken together

The results shown so far highlight the impact of overlooked microbial eukaryotes on other kingdoms. It was already mentioned the highly variable abundance across overlooked microbial eukaryotes, with entire groups being low abundant (Lobosa) and others conversely being highly abundant (Metazoa).
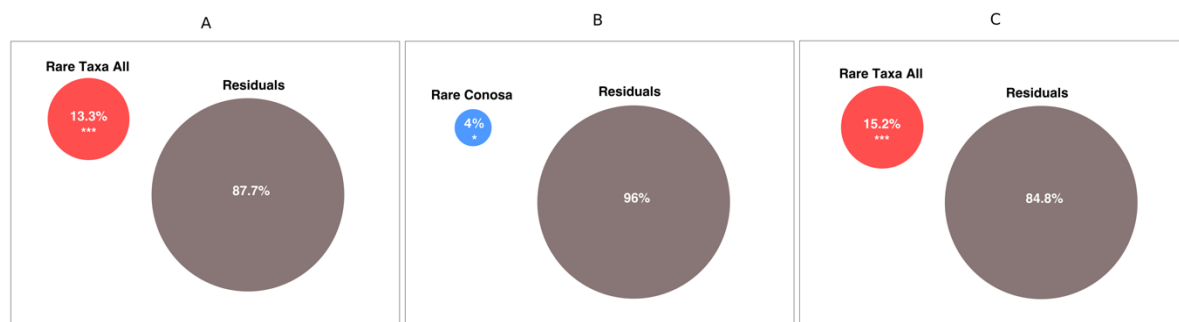
*Figure 2.21a; Rare overlooked microbial eukaryotes have a remarkable impact on alpha diversity of Bacteria, Fungi and Oomycetes. However, only if taken as a pool. Bacterial alpha diversity is increased by rare overlooked microbial eukaryotes. Bacteria are depicted in **panel A**, followed by Fungi (**panel B**), and Oomycetes (**panel C**). picture displays ANOVA on panel alpha diversity, percentages display η^2 score, showing the variation explained by the presence of each overlooked microbial eukaryotes. Only significant pvalues are displayed. Notation of pvalues follows *= pvalue <0.05, **= pvalue<0.01, ***pvalue<0.001.*

I consider this scenario of mixed high abundant and very low abundant taxonomic groups as an ideal situation to study the impact of rare taxa in an ecological perspective (Shade et al., 2014). Network analysis is usually abundance-limited, implying that can hardly deal with low abundance or even rare taxa. On the other hand, there is evidence that rare taxa can hold a key role in shaping microbial diversity and fostering community stability and resilience (Lozupone et al., 2012; Shade et al., 2011). I chose an approach similar to the one used for evaluating the impact of single taxa, with the difference that I considered as rare taxa only the ones present in the 10th percentile of the overall overlooked microbial eukaryotes abundance. I then tested these taxa both grouped together and by subdivision,

in order to see any possible synergistic effect. These taxa presence/absence (listed in
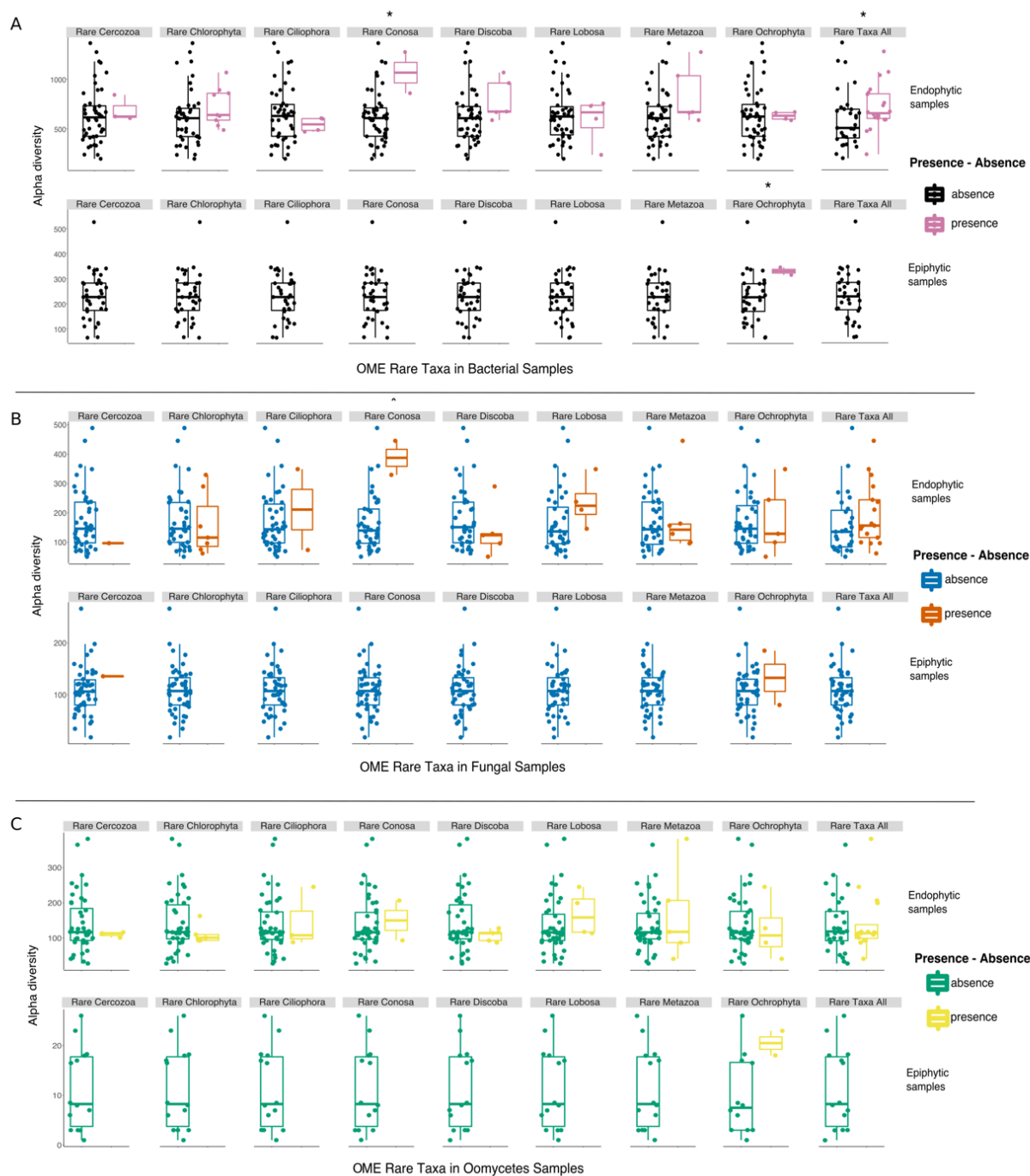
*Figure 2.21b; Rare overlooked microbial eukaryotes (Rare OME) have a remarkable impact on alpha diversity of Bacteria, Fungi and Oomycetes. However, only if taken as a pool. Bacterial alpha diversity is increased by rare overlooked microbial eukaryotes. Bacteria are depicted in **panel A**, followed by Fungi (**panel B**), and Oomycetes **(panel C)**. Boxplots depicts panel diversity in Endophytic samples(Up) or Epiphytes(Down) for every overlooked microbial eukaryotes group displayed on top of the boxplot. P- values are calculated after Wilcoxon-Mann-Whitney test. Notation of pvalues follows *= pvalue <0.05, **= pvalue<0.01, ***pvalue<0.001.*

Supplementary_table_3.xlsx

https://github.com/AlfredoMari/chetproject0001/tree/master/Results),   were   then

considered as partitioning variable for ANOVA test on bacteria, fungal, oomycetal alpha diversity, as a whole or divided by subdivision.

The striking result observable in bacteria and oomycetes (fig 2.21a) is that the rare taxa taken in pool accounts for significant portions of variation whereas no rare taxa taken by subdivision lead to a comparable effect. (fig 2.21a) Fungi constitute an exception, in fact, rare taxa as a whole do not have any accountable influence, except for Conosa, which account less than 5% of variation.

These findings underpin the hypothesis of a relevant role held by overlooked microbial eukaryotes rare taxa in shaping microbial diversity, however mostly when present together in the same samples. Among all the components of the rare taxa pool, Conosa seem to harbor the major effects. However, the low number of samples considered (2) does not allow to draw conclusion from this last result. A remarkable conclusion that is possible to delineate is that overlooked microbial eukaryotes rare taxa as a pool (but not singularly) enhance bacterial alpha diversity which is known to lead to a more stable and resilient community (Elmqvist Thomas et al., 2003; Lozupone et al., 2012)

3. Discussion

Community composition and interactions between terrestrial organisms and microorganisms on plant hosts has been investigated since centuries. However many features like biogeography or niche definition have remained debated (Agler et al., 2016a; Coleman-Derr Devin et al., 2015). This was mainly due to the substantial lack of sufficiently large scaled studies, able to provide statistical support to formulated working hypotheses (Prosser et al., 2007).

It is consolidated knowledge that the plant host harbours communities of micro-organisms and constitutes a micro-ecosystem itself (Hardoim et al., 2015). It is also well known that the plant is able to recruit beneficial microbes (Berendsen et al., 2012). The presence of these beneficial communities often coincides with enhanced plant fitness and enhanced resistance to disease. For crops, this can be translated into more stable and prosperous production (Berendsen et al., 2012).

The studies focusing on plant microbiomes have usually covered bacteria, fungi, and oomycetes (Hacquard et al., 2016; Vorholt, 2012). However, especially concerning the leaf, most of them have found high portions of microbial variation which still remained unexplained (Agler et al., 2016a). Therefore, underpinning the hypothesis of the presence of other interactors, possibly with top-level role. I hypothesize that this missing part of the mosaic can be represented by unidentified microbial eukaryotes.

The knowledge gathered so far on the plant holobiont regarding microbial eukaryotes often suffers from a too narrow focus, on one hand able to give precise insights on singular species, but on the other hand hardly able to reflect the bigger picture of ecological interactions (Bonkowski, 2004; Ploch et al., 2016; Sapp Melanie et al., 2018).

What I aim at, in this study, is to investigate interactions among plant microbiota focussing on unidentified microbial eukaryotes, within the frame of the plant leaf.

These overlooked microbial eukaryotes, also colloquially known as "protists", are extremely diverse (Parfrey et al., 2010), both in terms of biology and of phylogeny (Adl et al., 2005). Many of them are still unclassified and their cultivability is not yet assessed (Parfrey et al., 2011). That is the reason why I opted for a culture-independent approach (geographical sampling on natural *A.thaliana* populations) using Illumina amplicon sequencing technology deployed with universal primer approach. This choice allows to encompass the broadest possible phylogenetic picture of microbes (bacteria, fungi,

oomycetes and remaining microbial eukaryotes) without losing taxonomic resolution. In order to follow this approach, it was required in first place to develop the strategy mentioned in paragraph 3.1.

Following this path, I aimed to answer the following biological questions:

1. What are the abiotic and biotic factors influencing the diversity of phyllosphere microbes? Is their distribution driven by geographical province, or by micro-habitat variables?

2. What is the role of producers among microbial consortia? Are they also involved in mutualistic relationships? If this is the case, which kind of interactions do they engage in and how is the host involved?

3. What is the role of consumers in the phyllosphere? Do they actively shape microbial diversity? Are there microbial successions due to predation? Is the plant host actively modulating consumer behaviour?

4. Does the impact of microbes only depend on their abundance? What is the role of the rare biosphere in the plant holobiont? Do they have remarkable impact on the microbial assembly?

In order to answer these questions, I conducted two experiments: geographic sampling of natural populations of *A.thaliana* across Spain, Sweden, France and Germany (15 sites in total) and a common garden experiment in which four different ecotypes were sampled monthly between November and March. Both experiments were conducted over three years.

What follows is the discussion of the results of these experiments in the light of the aforementioned biological questions.

3.1 Methodological advances in amplicon sequencing and downstream data analysis

3.1.1 Blocking oligonucleotides enhance sequencing resolution and improve low abundant taxa's detectability

The objective of the experimental setup was to delineate the main features of microbial eukaryotes associated with *A.thaliana*. As mentioned above, I accomplished that by choosing Illumina amplicon sequencing, following an approach established recently (Agler et al., 2016a). I expanded this approach to this experimental design for addressing unidentified microbial eukaryotes. I chose to do so by designing barcoded primers on V4 and V9 regions of eukaryotic 18S, in order to obtain the broadest possible picture of microbial eukaryotes landscape (par 2.1.1).

However, targeting eukaryote 18S with universal primer approach had the disadvantage that high abundant off-target DNA fragments are often strongly amplified, mainly host DNA, sacrificing read depth and masking diversity (Hanshew et al., 2013). My design of blocking oligos, which bind to the template between the binding sites of the primers (paragraph 2.1.1) efficiently solves this problem by increasing the detectability of microbial eukaryotes more than 100000-fold.

Similar approaches have been attempted before: particularly a previous method described peptide nucleic acid that are highly specific to non-target templates and which physically block their amplification (Lundberg et al., 2013). These oligos work efficiently even in single-step amplifications, however their production is expensive, limiting rapid development of multiple oligos for new loci or for blocking several non-targets. Other approaches, such as using oligos modified with a C3 spacer (Vestheim and Jarman, 2008) are also more costly and worked best when they block the universal primer binding site, instead of the sequence between the binding sites.

The approach I propose, does not interfere with the primer binding site, and allows multiple blocking oligos to be deployed together. It is easy to design and not expensive.

By using this tool, I demonstrate that amplicon sequencing with a universal primer approach can be applied to microbial eukaryotes without necessarily losing resolution of target sequences. On the contrary, it improves detectability especially of low abundant taxa. Moreover, the versatility of the oligos and the low costs of the strategy indicate this approach as the most suitable in large scale amplicon sequencing such as this study.

3.1.2 A new network comparison pipeline provides customizable, modular and solid statistics for ecological studies

For the last years, microbiome analysis, has involved large deployment of graph theory in order to infer and model ecological interactions between OTUs or taxa in the shape of edges between nodes (Barabasi, 2009; Barberán et al., 2011). This analysis strategy is statistically very solid, and is perhaps one of the best tool to handle complex data such as microbiome data (Faust and Raes, 2012). Given the complexity and heterogeneity of microbiome data, it often happens that the inferring of the network may involve only part of the dataset. This creates the necessity of further analysis platform able to compare networks inferred from different subsets of the main dataset. Typically, in previous studies, such comparison has been achieved by comparing general network properties such as clustering coefficient, density, radius etc.(Faust et al., 2015). However, to the best of my knowledge, no available platform provides further calculation of the edge conservation between two networks, in a pairwise manner (not as a general parameter), supported by a solid statistical validation. In paragraph 2.1.4.2.1 I introduced a new bioinformatics platform able to extract further information from a comparison between two networks. The case displayed regarded two networks inferred under the same conditions, on two data sets respectively including and excluding the algae *Bracteacoccus sp*.

Current platforms like the Cytoscape app VennDiagrams mainly compare node and edge overlap in a qualitative manner (Saito et al., 2012). The app Dynet, instead, compares provided networks by calculating a core of shared nodes and then drawing a resulting network based on a comparison with the shared core (Goenawan et al., 2016). This approach is more insightful than the one provided by VennDiagrams, however, it ignores nodes and edges which are lost from one network to another, revealing a bias for highly connected nodes and shared nodes (Goenawan et al., 2016). Moreover, it calculates the divergence between the two networks by calculating a rewiring score (D score) which is prominently based on properties of nodes, rather than edges. This feature, together with the shared core bias, mines the possibility to compare very different networks, possibly lacking a sufficiently large core.

A more successful approach is achieved by the Cytoscape app Diffany, which offers a calculation of differentials and consensus network based on a customizable reference network (usually one of the networks being compared) and it is based on edge scores such

as edge weight (Landeghem et al., 2016). Differential and consensus network are then displayed in a directed resulting network which allows a visual comparison of the negatively or positively correlated edges across the two networks compared.

While Diffany provides a more accurate edge comparison compared to Dynet, it does not discriminate between a lost edge and a changed interaction type of the edge itself (Landeghem et al., 2016). Both platforms (Dynet and Diffany) include the edge weight as unique edge value in the comparison and do not allow any further parameter to be tested. Moreover, Diffany and Dynet do not provide statistical validation of their comparison, since they do not perform any statistical test on permutated edges.

The platform that I present here goes beyond the classical comparison of general network measurements, and beyond the calculation of network rewiring. It aims at achieving a detailed comparison of the two networks based on virtually any inferred edge score.

After dividing the edges in classes based on the provided score, it provides information on the composition of the classes. This allows the user to visualize which edge starting from which node is positively/negatively conserved between the two networks. Lost edges are by default assigned a lack of correlation score (1), and specifically tracked down. This way they are clearly identifiable as lost edges.

As the most important and last step, it succeeds in calculating solid and reliable statistics (through Mantel test and Mantel correlogram after 500 permutations) on edge classes. Thereby the user can quantify how strong the edge (lack of) correlation is.

Other approaches have also tackled network comparison based on statistics, by using distance metrics such as Jaccard diversity index, instead of Mantel correlogram (Widder et al., 2014), or by using PERMANOVA (Williams et al., 2014). Both approaches open glimpses on general correlation of the two networks based on edge score. However, in the last example, PERMANOVA was calculated directly on the incidence matrix. This calculation provides a general correlation score, similar to the one provided by regular Mantel test. Instead, my approach based on a distance matrix inferred from edge features compares the two networks edge by edge (grouped in classes). Thereby it offers a more detailed picture, since every component of the correlation can be tracked down and analysed further. Moreover, the implementation in perl guarantees a fast and efficient deployment, with an ordered and trackable output, since it makes large use of LSF system (Songnian et al., 1993) to parallelize and run subscripts in background.

In order to finally validate the platform with real biological data, I compared two networks inferred from samples containing or not containing the species *Bracteacoccus sp*. both with

Diffany and with the presented pipeline, customized to evaluate edge weight, and thereby comparable to Diffany. The working hypothesis was the matching of the qualitative divergences recovered with Diffany with the quantitative results achieved with my pipeline. While the results largely agree in terms of which edges differentially connect in the two networks, my pipeline gives further insights into possible correlation between high weight and conservation of the edges across the two networks.

This pipeline can therefore be considered as an important methodological step forward. Its modularity and customizability allow it to be applied in virtually any kind of pairwise network comparison on virtually any property of the inferred edge. In the ecology field, it can allow more solid and reliable statistical analysis on network inferring and comparison, especially for large scale microbiome study.

3.2 Overlooked microbial eukaryotes biogeography appears to be shaped by microhabitat factors rather than by geographical province

At the beginning of this chapter I mentioned a few biological questions as object of this study. The first one concerned the biogeography of the microbes associated with plant holobiont. From a recent study (Agler et al., 2016a), it emerged that a large portion of variability, especially concerning bacteria, remains still ambiguously assigned to sampling location and seasonality, or just unexplained.

I aimed to disentangle the biological meaning of this location. The formulated working hypothesis was that both overlooked microbial eukaryotes and other microbes distribute differently across site, triggering the question: is this differential distribution due to latitudinal driven dispersal, or to microhabitat variables?

In order to test this hypothesis, I analysed the microbial composition of natural populations of *A. thaliana* leaves 15 sites divided into 4 biomes (Natural, Suburban, Agricultural, Urban) equally distributed across 4 countries (France, Spain, Sweden, Germany).

Displayed results indicate that biogeography of microbes differs across different taxonomic groups: for bacteria and overlooked microbial eukaryotes biome is leading, for oomycetes – province, for fungi - both.

This last result agrees with previous studies which reveal a latitudinal pattern over fungal diversity distribution (Tedersoo et al., 2014).

For all the displayed groups, the compartment seems to harbour the most of the variability. However, this finding may be treated with caution, since the epiphytic OTUs were obtained by filtering the OTUs from whole leaf samples against endophytic OTUs (M&M).

The results I found on bacteria community match what has previously been discovered in a study on bacterial biogeography on populations living on Arctic and Antarctica ice shelves respectively (Varin et al., 2012). In that case, there have been detected quantitative discrepancies between the two poles communities, however the two populations were qualitatively similar, being dominated by Cyanobacteria and Proteobacteria. According to my results, Proteobacteria are also the main phylum composing bacteria leaf microbiome. Some studies interpreted the cause of this differential abundance between the two poles to differential "colonisation potential" (Cavicchioli, 2015). The colonisation potential is defined as the efficiency of a taxonomic group to establish a new community in a previously not colonized niche (Martiny et al., 2006). It is higher in case the taxonomic group has not restricted feeding habits and if the group has large population densities and broad range of dispersal possibilities, like in the case of bacteria.

In this case, this means that bacteria can potentially colonise every latitude The divergent populations are then determined by microhabitat conditions, like the biome, following what is known as Baas-Becking hypothesis (Baas-Becking, 1934).

In conclusion, the Baas-Becking hypothesis fits the displayed results on bacteria, and may mean that the role of the plant host is rather marginal within microbial assembly, as it has been previously shown for roots (Schlaeppi et al., 2014).

Of course, different taxonomic groups have different colonisation potential. In the case of oomycetes, whose dimensions, dispersal efficiency and biology are radically different compared to bacteria, I expect them to have lower colonisation potential.

In fact, as reported in paragraph 2.1.3, oomycetes are significantly affected by latitudinal factors rather than by microhabitat factors. This would mean that for oomycetes, the dispersal is more crucial than the adaptation to a different microhabitat.

In fact, by checking the differential composition of oomycetal taxa across the sites, I found that the whole genus *Albugo sp.* is completely absent in Sweden. Possibly because the fact of being an obligate biotroph could have harmed the dispersal over a geographical barrier such as Baltic sea, or, more likely, because of less evident anthropogenic contribution in *Albugo sp.* dispersal (Santini et al., 2013).

Regarding overlooked microbial eukaryotes, most of the biogeography studies were based on freshwater, ocean or soil samples (Martiny et al., 2006). However, these studies found

similar results compared to what is reported in paragraph 2.1.3. In fact, I found that microbial eukaryotes are affected by both province and biome (figure 2.6), even though biome is leading over province by 4%. This is consistent with what was found in a recent study (Schiaffino et al., 2016), in which microbial eukaryotes of fresh waters of Patagonia and Antarctica lakes, were found to be affected by province and by microhabitat conditions such as water pH, conductivity, or presence of other microbes. In other contexts, this similar importance of habitat and province has been questioned, underpinning a prevalent role of latitudinal distribution (Filker et al., 2016). However, in that case, the geographical distance was so high (Austrian Alps vs Chile and Ethiopia) and connected to so many different biomes itself, that discrimination between habitat and province was hard if not impossible. It is interesting to notice the great intra-kingdom variability in terms of colonisation potential, but also of feeding habits of overlooked microbial eukaryotes. Their heterogeneous composition, including pathogens, saprotrophs, symbionts, and photoautotrophs, delineates a scenario in which each class and possibly each genus can be differentially impacted from province or microhabitat conditions (Heger et al., 2013).

The aforementioned study (Schiaffino et al., 2016) reported a decrease of latitudinal effect together with the decrease of body size. I did not observe such correlation in my data, although the reported findings on bacteria (biome driven) and overlooked microbial eukaryotes (province-biome driven), would underpin this hypothesis.

It has to be noticed that the mentioned studies on biogeography of microbes in free-living conditions (oceans/ice/freshwater) are substantially in line with the results presented in this study on plant holobiont. This may indicate that the role of the plant host, at least concerning the phyllosphere, shall be rather limited.

These findings trigger one biological question: if biomes and province are almost equally important for overlooked microbial eukaryotes, what are the characteristics of the biome and of the different provinces? In other words: if biome and country impact microbial diversity in different ways, how can we dissect them in order to extract the real factors shaping microbial diversity?

Concerning provinces, I considered as relevant features the environmental factors listed in M&M such as temperature and precipitation rate. This choice is due to the fact that such variables are often latitude dependent. Moreover, for some locations, it was possible to retrieve climatic data only from one station for two or three sites.

Concerning biome, I classified them based on the soil history and based on proximity to anthropogenic areas. This means that each biome is likely to constitute an ensemble of

(micro)habitat factors, which include of course specific or even unique biotic interactions taking place. In order to tackle biotic interactions, I addressed the co-occurrence/co-absence of the recovered taxa, which have been reported to be e phaenomenon typical of microhabitats and specific biomes (Martiny et al., 2006).

3.2.1 Overlooked microbial eukaryotes and microhabitat: biotic interactions are not influenced by province related variables, underpinning a bond with the biome variable

In order to disentangle the biome effect and the province effect, a comprehensive analysis of occurring biotic interaction across my samples was needed. I then chose to infer a microbial network starting from the original clustered OTUs. This is probably the most powerful instrument to date to deal with highly complex systems such as ecological interactions (Barberán et al., 2011; Faust and Raes, 2012). This is the reason why I deployed it on my data, obtaining a picture of ecological interactions over three years of sampling, including environmental features as nodes. Notably, overlooked microbial eukaryotes connect with key hubs like *Sphingomonas sp.* and *Methilobacterium sp.*, underpinning a possible association, but excluding an association with the environmental nodes, which remained disconnected from almost every node. The fact that environmental factors such as temperature and latitude do not connect with any of the nodes is in line with a theory underpinning that biotic interactions are rather linked to microhabitat variables and do not form any large-scale pattern, for instance province-based (Wisz et al., 2013). As a proof, they do connect in case of reduced scale network inference, for instance by building networks from single biomes or sites or in case the abundance thresholds are lowered (fig 2.8).

It is interesting to notice the striking abundance of negative over positive interactions, the latter being distributed solely intra-kingdom and not inter-kingdom.

There have been studies which considered the quality of microbial interactions as the main feature to build reference models on microbial consortia stability and dynamics (Wisz et al., 2013). Specifically, it has been reported that a high network stability is correlated to the prevalence of competitive (negative) interactions (Coyte et al., 2015). In this experiment context, this would imply a high stability of the inferred interactions, confirmed by the fact that no nodes connect to the "Year" variable.

As a second confirmation, alpha diversity analysis (fig 2.11) shows no significance of the year on the variation of microbial assembly in all investigated groups. However, what do

positive and negative interaction mean in an ecological frame? Several studies up to now reported that during favourable abiotic conditions, biotic interactions tend to be negative and competition rate is much higher (Callaway et al., 2002; Mod et al., 2016). Whereas, when the microhabitat conditions become harsh, for instance after a perturbation, mutualistic interactions become the majority of interactions (Mod et al., 2016). The absolute prevalence of negative interactions would then suggest a relatively stable habitat in which no major perturbations are likely to have happened, at least recently before sampling.

These results, together with the findings reported in chapter 2.1.3, show that both geographical and habitat factors are almost equal shapers of microbial eukaryote diversity. Concerning province as shaper of overlooked microbial eukaryote diversity, it is not possible to draw major conclusions, apart from acknowledging that geography may play a role in terms of dispersal potential rather than climatic variables. Along these lines, since analysed overlooked microbial eukaryotes are phylogenetically very heterogeneous, an analysis on geographical dispersal coupled with phylogenetic analysis on every class or even genus singularly would be needed. However, a candidate-based approach would open new interesting perspectives on latitudinal biogeography of overlooked microbial eukaryotes.

The key outcome of this study concerning biogeography is assessing the importance of biome, microhabitat and biotic factors correlating with the diversity of overlooked microbial eukaryotes.

The primacy of habitat, biotic factor and local scale communities in the leaf environment would seem sound, if we consider that the plant holobiont, especially within the leaf, offers at least two microhabitats (endophytic and epiphytic compartment). Those two environments, are differently affected by abiotic (rain, UV rays, temperature) and biotic factors (the presence of apoplast in endophytic compartment, presence of other microbes) and therefore definitely able to shape microbial communities. It remains to be tackled what could be the role of the host in such a consortium development. I will address this question later in the chapter.

Through network analysis I highlighted the major interactions between OTUs. However, this leaves the impact of overlooked microbial eukaryotes groups, and other biotic interactors, such as bacteria, fungi and oomycetes still unclear. In other words, which overlooked microbial eukaryotes have which kind of impact on microbe-microbe interactions? Do they influence the increase or decrease of microbial diversity? In order to

tackle these questions, I chose to narrow my analysis to two ecologically important taxonomic groups: green algae (Chlorophyta), and amoeboid organisms (Lobosa, Ciliophora), together with Metazoa.

3.3 The leaf in the leaf: microalgae in both compartments reveal unexpected microbial dynamics

In all the investigated sites, and across all the three years, I found a substantial part of reads belonging to the subdivision Chlorophyta, known also as green algae. Remarkably I found their presence not just on the leaf surface but also inside the leaf mesophyll. I also detected at low abundance but repeatedly the presence of Chlorophyta inside not just leaves, but also root endophytic compartment (par 2.2.2, figure 2.11). This finding is unlikely due to contamination for a number of reasons: (1) the fact of having found such presence across all three years of sampling, in such different locations; (2) the fact of having manually harvested the samples always in a team of multiple people, and (3) in case of the few root samples, also the usage of a different DNA extraction kit. The almost total overlap of algal chloroplastic DNA with algal 18S (more than 97% of the 16S samples) (fig 2.11), supports the hypothesis of an active role held by green algae within plant microbiota.

This hypothesis is also underpinned by network analysis. Among several thousands of OTUs, after stringent cut-off filters and bootstrapping, only Chlorophyta and Metazoa were displayed in the network, engaging mainly negative interactions with bacteria and fungi. Among them, a single taxon, *Bracteacoccus sp.* was also found as relevant node. It is important to specify that network analysis implies ecological correlations between OTUs, but does not allow any further ecological conclusion on alpha diversity, which needs to be addressed by supporting statistics or specific experimental designs.

By performing the supporting variance analysis on alpha diversity distribution, I displayed not only that Chlorophyta have a significant impact on microbial consortia, but also that I could quantify how abiotic and biotic factors (like the simple presence-absence of microalgae) can condition alpha diversity of microbiologically crucial kingdoms (Bacteria, Fungi). Interestingly, Chlorophyta often significantly accounted for large portions of variance of bacterial and fungal alpha diversity. I also showed that the Chlorophyceae *Bracteacoccus sp* accounted for large portions of variance of bacteria, fungi and oomycetes. The importance of *Bracteacoccus sp.* will be further discussed in chapter 3.3.2.

These results open for discussion what has always been discussed only for water environments: what is the relationship between plants and algae? Are algae actively interacting with the plant or mainly with other microbes? Large knowledge has been acquired on the so-called phycosphere, and on its regulatory systems (Ramanan et al., 2015a). The presence of a biofilm-based phycosphere in the leaf would open new fields of microbiological analysis.

3.3.1 Phycosphere in the phyllosphere, microalgae shape microbial interactions by connecting to the key hubs, likely through biofilm

It is consolidated knowledge that microalgae form biofilm in almost every surface they colonize (Ramanan et al., 2015a). It has been reported as well that such biofilms are created by algae generally by the secretion of the so-called EPS (Extracellular Polymeric Substances). However, their creation and maintenance can benefit of the contribution of several species of bacteria.

An interesting study has demonstrated that beside biofilm formation, bacteria are also involved in algal colonisation, since algal growth seems to be enhanced in an environment already colonized by bacteria (Schnurr and Allen, 2015).

These elements fit my findings on recovered algal diversity found on leaves. In fact, the leaf surface is likely to be colonized in first place by bacteria (Hassani et al., 2018), and hypothetically, an algal colonisation could be therefore facilitated.

The habitat created within the biofilm is called often phycosphere. Numerous studies on phycosphere report the specie-specificity of the phycosphere itself. This means that each microalga is able to recruit differential bacterial or fungal species, mainly for trophic reasons (Ramanan et al., 2015a). However, certain classes of bacteria appear to be more recurrent in different phycospheres. The well-known ones belong to the class of Methylobacteriaceae, but also to Sphingomonadaceae, and Caulobacteraceae (Ramanan et al., 2015a). These last two families were represented in several nodes of the microbial network and directly engaged connections with the recovered algal nodes.

In general bacteria-algal interaction has been proven to be tight in other environments, either by quorum sensing interference (Teplitski et al., 2016) or by mutualism (Ramanan et al., 2015b). A well described mutual interaction is one that takes place between microalgae, especially Trebouxiophyceae and the bacterial family of Rhodobacteraceae. An active exchange of nutrient and micronutrients between the two interactors has been

reported, mainly via the delivery of carbon from the algal side, usually in EPS form, receiving back from the bacterium micronutrient like cobalamin (vitamin B12) and auxin (IAA). This specific kind of mutualism has been observed mainly in freshwater environment and in biofilms (Ramanan et al., 2015b).

This information is consistent with preliminary results not shown in this thesis. In fact, I found a trend of samples containing high levels of Rhodobacterales reads, being also abundant in Trebouxiophyceae 18s reads as well as algal chloroplastic 16S. Given these observations, I hypothesize that such a mutualism could occasionally occur also on terrestrial plants.

These findings underpin the hypothesis not only of an active outer algal biofilm, but also of a possible endophytic phycosphere, in which selected bacteria taxa such as *Methilobacter sp.* and *Sphingomonas sp.*, together with micro algae can have a key role in shaping microbiota diversity.

3.3.2 *Bracteacoccus sp* may be the main player in the biofilm niche

Among other microalgae, network analysis brought to my attention *Bracteacoccus sp.* which already displayed high betweenness centrality and was later found to be a key shaper of alpha diversity, of bacteria, fungi and oomycetes.

The sequences assigned to *Bracteacoccus sp.* found in our data set most likely belong to *Bracteacoccus occidentalis* or *Bracteacoccus bullatus*. Possibily, multiple *Bracteacoccus sp* strains may be involved.

Concerning the potential role of *Bracteacoccus sp.* in the microbial consortium, it is still difficult to hypothesize, since recent studies report it as very cosmopolitan microalgae (Lewis and McCourt, 2004), sometimes on the verge of being considered pathogen. In the past, it has been considered responsible for an algal blooming in the historical cave of Lascaux in France (Lefevre, 1974).

Interestingly, two other species of *Bracteacoccus sp.* were found to be present in forest floor and forest litter, where its presence seems to vary accordingly to seasonality and soil horizon, however, most of its prerogatives remain unknown (Maltsev, 2013; Maltsev et al., 2017).

It is however likely that *Bracteacoccus sp.* could play a role in biofilm formation, similarly to other already investigated Chlorophyceae (Ramanan et al., 2016). A speculative hypothesis would be to consider it as the reservoir alga, which in determined condition is

able to rapidly reproduce and become the main biofilm developer, given its already mentioned ability to bloom in certain condition.

Despite this lack of consolidated knowledge, it seems that *Bracteacoccus sp.* prefers to engage commensalistic interactions with bacteria and fungi, as the shown ANOVA underpins (fig 2.10). It has been found to engage symbiosis especially with the species *Rhizidium sp.* (Picard et al., 2013) However, no reads assigned to *Rhizidium sp.* were found in my geographical sampling.

In conclusion, I found that micro algae belonging to Chlorophyta have major impact on bacteria diversity as well as oomycetes diversity. A hypothesis on how they can act as modulators of diversity can be formed considering the formation of a biofilm. Biofilm development could possibly be supported by occasional mutualisms with the class of Rhodobacterales, or more likely with bacteria of the genus *Sphingomonas* and *Methilobacterium*, correlators of algae in the inferred network, and already known as "usual suspects" from previous studies (Vorholt, 2012).

Another remarkable finding that I showed is the prominent correlation between green algae and fungal diversity. In nature, a peculiar symbiosis between plants and fungi – lichens -, has been studied over centuries. I argue that this symbiosis could occur also on *A.thaliana,* and have impact on the rest of microbes.

3.4 The fungal connection: Lichens

3.4.1 Lichens appear in both endophytic and epiphytic compartment, and they impact bacterial diversity on the leaf

In order to support this hypothesis, I investigated whether or not lichens can be considered as key players in leaf microbiome and whether or not they have an impact on microbial diversity. In order to address the first question, I mined the geographical sampling data, looking for not only for lichenising fungi, but also for the corresponding phycobiont. As fungal class tester, I chose the order of Lecanorales, known to be composed eminently by lichenising fungi, engaging symbiosis mainly with members of Trebouxiophyceae family (Ahmadjian, 1967; Favero-Longo and Piervittori, 2010). As reported in paragraph 2.2.2, the overlap between Trebouxiophyceae 18S and Lecanorales ITS is almost in 98% of the samples. Moreover, I displayed that an overlap between Lecanorales and algal chloroplast 16S is in more than 70% of the samples, underpinning the hypothesis of an active symbiosis between lichenising fungi and algae taking place on the leaf. It is likely that most of the

lichens found in this context live on the leaf surface. In fact, more than 70% of the lichenised samples are epiphytic (par 2.2.3). However, a marginal quantity of lichens can still be found endophytically in 12% of the samples, especially belonging to the genus *Parmelia* (see paragraph 2.2.3).

In order to answer the second initial question about the importance of lichens on the rest of microbial consortia, I performed ANOVA on bacteria alpha diversity, keeping Lecanorales and algal chloroplast as partitioning variables. The reason for choosing bacteria as a target was due to the fact that it has been reported several times that bacteria are the kingdom that is majorly influenced by lichens within microbial consortia (Bates et al., 2011; Mushegian et al., 2011).

The results shown in paragraph 2.2.3.1 clearly display that Lecanorales have a small but significant impact on bacterial diversity only when nested with the algal chloroplast variable (2.2%). The fact that this percentage is also nested with the compartment variable underpins the hypothesis that Lecanorales have impact on bacterial diversity only in a specific compartment, which is likely to be the epiphytic one, since it shows the highest Lecanorales abundance. This hypothesis would be in line with the literature, claiming that bacteria associated with lichens closely resemble the bacterial community of the leaf surface (Grube and Berg, 2009). Moreover, other studies highlight the tight associations between lichens and Alphaproteobacteria (Bates et al., 2011; Grube and Berg, 2009), very abundant in my samples, and interestingly also with *Methilobacterium sp.* (Grube et al., 2009), which was one of the main hubs connecting with chlorophyte nodes in the network displayed in figure 2.7.

Other possible associations involve the interplay between certain classes of bacteria such as phosphorus and nitrogen mobilising bacteria, reported to be associated with lichens (Grube and Berg, 2009; Lücking, 1999), and of fundamental importance in the leaf microbiome (Vorholt, 2012).

These findings support the hypothesis of an active role of lichens within leaf microbiome of *Arabidopsis thaliana* and its relevant contribution in shaping leaf bacteria community.

3.4.2 The key role of lichens in the phyllosphere microbiome is likely dependent on microhabitat conditions and may involve different phycobionts. Host genotype is not directly involved

Algae and lichenising fungi influence bacteria diversity in plants. However, few biological questions still remain to be tackled. For instance: is the genotype of the host relevant for colonisation? How much are the microhabitat conditions relevant for lichens to establish on the leaves? And how important is the leaf age for lichens colonisation? How could possibly lichens colonise the plant and interact with bacterial population?

In order to address those, I analysed sequencing data from a common garden experiment including four different ecotypes (KsK, Col-0, Ws-0, Sf-2) sampled monthly from November till March over three years (see M&M).

Results displayed in chapter 2.2.3.2 show that abundance fluctuations of Chlorophyta and Lecanorales follow different patterns over months with Chlorophyta having the highest abundance peak in November, while Lecanorales in December. Lecanorales distribution over years and Lecanorales overlap with algal chloroplastic DNA does not seem to follow a definite pattern in this experimental setup. Concerning ecotype instead, presence of Lecanorales seems to negatively correlate with the ecotype Ws-0. In fact, no Lecanorales reads were found on this ecotype.

In order to address the factors which may co-determine Lecanorales impact on the leaf, I performed ANOVA on bacterial alpha diversity, considering among others: month variation, ecotype, and the presence of phycobiont as partitioning variables. As candidate phycobiont I chose *Trebouxia sp.* as indicated by numerous studies on the subject (Favero-Longo and Piervittori, 2010).

The results displayed in figure 2.15 and 2.16 depict the connection between Lecanorales and the genus *Trebouxia* to be weak in terms of co-occurrence itself, however significant in terms of effect size on bacteria diversity.

This finding suggest that a bond between Lecanorales and Trebouxia may occur occasionally. However, in case this happens, they are able together to exert a remarkable impact on bacterial diversity. This circumstantial bond hypothesis is supported by the fact that Lecanorales can be coupled with phycobionts other than *Trebouxia sp.* In fact, some studies (Rambold et al., 1998) found that Lecanorales engage symbiosis mainly with Trebouxiophyceae, but in few cases they can be involved also with Chlorophyceae, as well as Cyanobacteria.

68

Concerning time course, from alpha diversity analysis on common garden experiment, emerges high nestedness between Lecanorales accounted variation and month variation. This is in line with what is shown in figure 2.15 about monthly fluctuations of Lecanorales, and likely indicates that influence on bacteria community is correlated with Lecanorales abundance fluctuation over the months, (for instance the abundance peak in December).

From the same analysis, it appears that ecotype does not nest with Lecanorales in shaping bacterial diversity.

This indicates that the Lecanorales distribution over ecotype shown in figure 2.16 may be circumstantial. Overall, this result is in agreement with previous studies on foliicoulous lichens, showing that the plant host genotype has low and insignificant effect on lichens development and harbouring of beneficial bacteria (Lücking, 1999; Serusiaux, 1989).

Regarding microhabitat impact on lichens, it seems clear that at least the compartment has massive effect on determining the presence of lichenising fungi and phycobionts (par 2.2.3.1). Concerning environmental factors other than compartment, there can be noticed a remarkable adaptability (chapter 2.2.3.1) of the lichens found on the plant to different habitats and provinces. In fact, the tight link between bacterial diversity and lichenising fungi both in the geographical sampling and in the common garden experiment, suggests that lichens may be a crucial shaper for plant associated bacteria, conserved across differential provinces and biomes.

By looking especially into the common garden experiment, there can be noticed that the previous ploughing and mulching of the soil prevents almost any neighbour plant or rocks to be close to the plant. This feature suggests that neighbouring plants or rocks may be of negligible importance for lichens establishment and underpins the possibility that lichen colonisation of the plant may be airborne (Tormo et al., 2001) or soilborne (Belnap et al., 2001).

If the presence of lichens seems to be a common phenomenon on *A.thaliana*, the minute distribution of the species across the plants seems not to be driven by the genotype of the plant host, but again, by microhabitat factors. The site specificity (and compartment specificity) of some genera like *Parmelia sp.*, found only in Germany and Sweden (chapter 2.2.3), supports this hypothesis. Further support comes on one hand from the concentration of Lecanorales abundance in german sites, especially in ERG and JUG and EY and on the other hand from the extreme scarcity of Lecanorales from Spanish locations (fig 2.13)

The tight bond between habitat and lichenising fungi is further supported by numerous studies on lichens in general and foliicolous lichens as well (Bruun et al., 2006; Lücking,

1999; Rogers et al., 1994; Serusiaux, 1989). In these studies, the correlation between lichenising fungi and elevation was highlighted (Bruun et al., 2006). These findings would fit with the data shown with the already mentioned prevalence of Lecanorales in higher latitudes such as Germany and Sweden, and low abundance in Spanish locations, which are also consistently much higher in terms of elevation (see M&M, table 4.1).

Another factor not to be neglected is the high sensitivity of lichens towards human pollution (Serusiaux, 1989). What one sees in the samples is a substantial absence of lichenising fungi in scenarios heavily modified by humans such as agricultural sites, however this is not true for urban sites, which conversely display high abundances of Lecanorales. Natural sites show almost equal number of samples in which Lecanorales are present or absent (fig 2.13).

In order to address the colonisation of the plant and possible mechanisms of interaction with bacteria, it has to be underlined that susceptibility to monthly variation, and especially the decreasing occurrence towards March, underpins an evolution of the lichen effect on bacteria on month basis, rather than on a year base. This is not common for lichen establishment, which usually requires years (Favero-Longo and Piervittori, 2010; Mushegian et al., 2011), and therefore underpins again the hypothesis of a circumstantial bond between lichenising fungi and phycobiont. In fact, it remains unclear whether or not the phycobiont and lichenising fungus would have the actual time to form a lichen structure, or would rather exert influence on bacteria diversity autonomously. Further conclusions on the time of colonisation and on leaf age influencing may be too preliminary to be proposed, and would need further experiments, even though, some studies propose an actual link between the two (Mushegian et al., 2011).

These findings, together with the common garden data indicate for the first time that lichenising fungi and microalgae of the genus Trebouxia can have an important contribution in modulating leaf bacterial microbiome, especially by harbouring differential bacteria community, as it was already shown in previous studies on bark or rocks (Rogers, 1988).

This does not provide the actual evidence of a lichen symbiosis taking place on the leaf, however, my results support this hypothesis for two reasons mainly: (1) because the tight association between *Active Trebouxiophyceae* and lichenising fungi in bacteria shaping is conserved across the sites and the experiments, and (2) because the main natural bacterial partners of lichens are also found in my data (fig 2.7). It remains however unclear the mechanisms with which lichenising fungi and phycobiont are shaping bacterial diversity.

If a proper lichen colonisation is taking place on the leaf I propose that this may happen through the steering of microhabitat above all other features, and without the relevant contribution of the plant host genotype. The displayed results do not underpin a major role of neighbouring plants or rocks in acting as reservoir of lichens, but rather underpin an airborne or soilborne origin of lichen coloniser. No major conclusion can be inferred on plant age effect on lichen colonisation.

In conclusion, this tight bond I showed between bacterial community and lichenising fungi can therefore provide great support in opening new fields in leaf microbiome research.

3.5 Consumers and grazing: connected through different timing, Lobosa, Conosa, Ciliophora and Metazoa exert a striking effect on microbial assembly

Phycobionts belonging to Chlorophyta, however, are just one of the 14 overlooked microbial eukaryote groups present among the assigned taxonomy of the recovered OTUs. Many others, with completely different feeding habits like heterotrophs deserve attention, since their crucial role in actively shaping microbiome and specifically root microbiome. It is in fact consolidated and reported multiple times their role in selectively grazing bacteria and other eukaryotes (Krome et al., 2010).

It is well known that grazers can shape microbial diversity by triggering microbiome fluctuations that in some cases can be considered as proper microbial successions (Bonkowski, 2004). In this paragraph, I am willing to tackle this subject, starting from the following biological questions: which overlooked microbial eukaryotes group has the major impact on microbial diversity? Do microbial successions actually take place on the leaf? And how are heterotrophic organisms involved?

I tackled these questions starting from the first: narrowing my research to the ones that seem to hold higher impact on microbial diversity, as described in chapter 2.2.1: Lobosa, Ciliophora and Metazoa, which explain significant parts of diversity of Bacteria and Fungi. A relevant role of amoeboid organisms (Lobosa, Ciliophora) in shaping bacterial diversity is not new, and actually was inferred in previous publications, even though it was discovered only in fresh waters and soil (Fiz-Palacios et al., 2013; Heger et al., 2013; Smith and Wilkinson, 2007). To the best of my knowledge, this is the first time that a relevance of Lobosa and Ciliophora within the phyllosphere microbiome is reported. The absence of

representatives of Lobosa and Ciliophora (typically low abundant) in the inferred network (Mahé et al., 2017), can be explained by the abundance bias typical of network analysis (Faust and Raes, 2012), confirmed by the presence of the much more abundant Metazoa. Network inferring and alpha diversity analysis underpinned a first level role of these three groups on the phyllosphere. However, both network analysis and ANOVA displayed respectively in par 2.1.4.1 and 2.2.1 are not able to indicate what role is held by the plant itself in the whole leaf scenario, nor they are able to give more information on the abundance dynamics over time.

I therefore investigated this biological question by performing the common garden experiment whose results are displayed in paragraph 2.3. As just mentioned for the geographical sampling, amoeboid organisms (Lobosa, Ciliophora) have a striking impact especially on bacteria community. What is remarkable from the common garden experiment (in figure 2.19) is a confirmation of what was shown for the alpha diversity analysis on geographical sampling, with Lobosa and Ciliophora having significant impact on bacterial diversity. Secondly, it is important to highlight that month variation is the major factor impacting bacteria diversity. Ecotype does not shape any part of variance, unless coupled with month. Lack of ecotype effect on plant microbial assembly was already argued in root studies, first in 2012 (Bulgarelli et al., 2012), and afterwards confirmed in 2014 (Schlaeppi et al., 2014).

Apparently, this is also the case for amoeboid organisms, which have differential patterns over months, but their distribution across genotypes is largely quantitative (fig 2.19). In fact, none of the four *A.thaliana* ecotypes tested showed differential overlooked microbial eukaryotes composition in general (supplementary figure 5).

This tight connection with month variation and lack of bond with ecotype would suggest that the lifestyle of these heterotrophs, is shaped by biotic interaction (microhabitat), rather than by province patterns. This hypothesis would be confirmed by recent studies which identified a latitudinal effect on amoebas only on very large scale such as comparisons between southern and northern hemisphere (Smith and Wilkinson, 2007). In all the other cases, biotic interactions and microhabitat factors seem to be the most important driver of amoeboid organism diversity (Heger et al., 2013). Furthermore, there have been studies on water distribution networks, as well as on meat processing plants (Vaerewijck et al., 2008), which found populations of Lobosa and Ciliophora strikingly similar to the ones depicted in this study, supporting the hypothesis of a large distribution potential, shaped afterwards by microhabitat factors, *e.g* the availability of the substrate. (Grigulis et al., 2013; Heger et

al., 2013; Schipper et al., 2001). According to the ANOVA shown in figure 2.19, these microhabitat factors should evolve on a monthly basis, contributing in developing a microbial succession.

3.5.1 Predator and prey revisited: when the prey can become a symbiont

Since month variation was found to be the major driver of bacterial diversity, I argued if also Lobosa, Ciliophora and Metazoa were following a similar pattern over the months. Results shown in paragraph 2.3.1 display a different distribution over time of Lobosa, Ciliophora and Metazoa, the first two having an occurrence peak in November and a steep descent immediately after. Metazoa instead seem to increase their occurrence just coincidentally to Lobosa and Ciliophora descent, in December. This descent is unlikely to be due to dramatic environmental changes since we did not register any of them through temperature and precipitation measurement taken on site daily (see M&M). Moreover, a major environmental change should have had similar effects across diverse kingdoms. I found otherwise. In fact, other groups of less influential heterotrophs follow different abundance patterns over time, such as Cercozoa (supplementary figure 6). It is therefore more likely, that these two events (Ciliophora and Lobosa in steep descent, Metazoa in arousal) could be linked together. The enhanced presence of Nematoda in December, decreasing progressively later (fig 2.20), could support the hypothesis of a voracious grazing on Lobosa and Ciliophora by Nematoda, known to feed prominently on ciliates and amoebas (Griffiths, 1994). This hypothesis however remains speculative, since no statistical test was performed on such co-occurrence. Alternatively, such inverse trend could be due to some major change within biotic interactions, possibly involving bacteria community.

In fact, it is consolidated knowledge that bacteria are not just food source for amoeboid organisms but occasionally they can be endosymbionts, or even parasites. One case is for example *Legionella sp.,* or *Chlamydia sp.* (Scheid, 2014), which can multiply inside the phagolysosomes of species like *Acanthamoeba castellani* or *D. discoideum* (Cosson and Soldati, 2008). It is known that amoeboid organisms are sensitive to bacteria quorum sensing (Hilbi et al., 2007), and therefore susceptible to bacterial community fluctuations. The month fluctuation of bacteria population may be correlated to the steep descent of amoeboid abundance in December, even though Bacteria fluctuations over months are largely quantitative. However, this hypothesis may be considered with caution.

Many studies in the past years have underpinned the importance of the interplay between amoeboid organisms and bacteria, specifically considering the taxonomic group of Cercozoa (Bass et al., 2005; Bonkowski, 2004; Hess and Melkonian, 2013). Aside few studies performed in controlled conditions (Flues et al., 2018), poor knowledge has been achieved concerning their ecological importance across the leaf environment.

3.5.2 Cercozoa: the big absent

Within the displayed results, I found a relevant impact held by Cercozoa only on bacterial alpha-diversity (fig 2.11). However, their contribution in alpha diversity partitioning remains unclear, especially concerning compartment allocation. Many studies have tackled the question whether or not Cercozoa have an impact on bacterial community, or in general what could be their role in microbiota consortia. A number of them (Flues et al., 2018; Krome et al., 2009; Ploch et al., 2016) claim the actual primacy of Cercozoa as principal microbial eukaryotes present in the soil and able to shape microbial diversity in the soil. This finding remains undisputed. However, on the leaves, the studies conducted so far were eminently focused on Cercozoa, by explicitly targeting ITS instead of 18S (Bass et al., 2005). This approach, which does not clear the primer bias from amplicon sequencing, gives a more detailed but very narrow scenario, that would hardly encompass enough variables or microorganisms necessary to draw major conclusions. Furthermore, other studies claiming the relevance of Cercozoa also on the leaves, had much less coverage in terms of biomes and latitude, compared to this study (Sapp Melanie et al., 2018).

As a conclusion, I showed that heterotrophic organisms belonging to Lobosa, Ciliophora and Metazoa have relevant but differential impact on established microbial communities in natural environments, mainly related to microhabitat and not to province distribution.

Having investigated groups which are typically low abundant (Ciliophora, Lobosa), opens the field to the question of how much the impact of the overall discussed microbial eukaryotes is affected by the abundance of the single organisms. There is controversial evidence that low abundance taxa may have the same or even more impact on microbial assemblies (Hol et al., 2010).

3.6 Rare taxa guarantee microbiome resilience, however, they only have effect if taken together

The ecological importance of low abundant taxa in microbiology has been documented multiple times mainly concerning bacteria (Fuhrman, 2009; Martiny et al., 2006).

Regarding plant-microbe interactions, it has been recently reported that plant roots are able to selectively recruit beneficial bacterial taxa which are low abundant in the soil, but become prosperous inside the root structure itself(Saleem et al., 2016). Such microbiome assembly is crucial for plant health maintenance(Berendsen et al., 2012). Given these findings, I wanted to further investigate the role of rare overlooked microbial eukaryotes within microbial assembly. I hypothesized a significant role of rare overlooked microbial eukaryotes in bacterial, fungal and oomycetal diversity.

I then specifically targeted overlooked microbial eukaryotes classified as 10th percentile of the overall abundance, and irrespectively of the taxonomic assignment. Following this procedure, I selected a pool of organisms belonging to the subphyla of Conosa, Lobosa, Chlorophyta, Metazoa, Dynophyta, Ochrophyta, Cercozoa, Ciliophora.

I then analyzed the overall impact of rare taxa, alone and in combination with one another, respectively on bacterial, fungal and oomycetal alpha diversity.

The results displayed in paragraph 2.4 clearly indicate that the considered rare taxa lead to an increased alpha diversity among bacterial samples, uniquely in the endophytic compartment. Taken together they account for 15% of bacterial variation, however interestingly no single taxonomic group of rare taxa is able to harbour any significant shift alone. On fungal diversity, no clear conclusion can be generalized, given the low number of samples taken into consideration. On oomycetal diversity, despite the significant accounted percentage of variation, no clear trend triggered by rare taxa is visible.

These findings show for the first time that low abundant overlooked microbial eukaryotes can be essential for the actual shaping of bacterial diversity. Moreover, this seems not to be an effect due to a specific taxon, but actually on the co-occurrence of all the screened low abundant taxa together. This feature of a pool of organisms which together influence the dynamics of an entire population, has already been reviewed within bacterial communities (Fuhrman, 2009; Lynch and Neufeld, 2015).

My findings indicate that overlooked microbial eukaryote rare taxa pool is correlated to an increased alpha diversity among bacteria. This impact may be crucial for community resilience after perturbation since it is known that a higher alpha diversity is a key factor promoting community resilience (Elmqvist Thomas et al., 2003; Lozupone et al., 2012).

In fact, already in previous studies, low abundant bacteria were linked to the promotion of community resilience (Ainsworth et al., 2015; Fuhrman, 2009; Lynch and Neufeld, 2015; Shade et al., 2014; Sogin et al., 2006).

These findings would underpin the extension of this hypothesis also to overlooked microbial eukaryotes.

Further data on time series and experiments on synthetic communities are required in order to test the provided hypothesis. Until then, my proposal has to be taken with caution.

3.7 Conclusions and further perspectives

In this thesis work I aimed to explore one side of plant-microbe interactions often neglected: the one regarding what I addressed as "overlooked" microbial eukaryotes. Colloquially also known as "protists", these organisms have never been thoroughly and broadly investigated in the terrestrial environment, specifically in association with the plant host (Hassani et al., 2018).

I hypothesized that overlooked microbial eukaryotes have an active and top-level role in shaping microbial community and thereby indirectly have remarkable effects on plant health.

Specifically I aimed to tackle:

- The role of primary producers (microalgae)
- The role of selected consumers (amoeboid organisms)
- The role of rare taxa

In order to accomplish it, I designed the experimental setup including latitudinal sampling on natural *A.thaliana* populations across 15 locations in Europe. I combined it with a common garden experiment in order to disentangle over time microbial successions on different ecotypes of *A. thaliana*. The deployment of high throughput methods such as amplicon sequencing, coupled with the development of new techniques like the one of the "blocking oligos", allowed me to recover unprecedented data to be investigated. By using and improving network analysis, as well as other supporting statistics, I confirmed my initial hypothesis that overlooked microbial eukaryotes are key factors involved in shaping other microbe diversity. Their distribution is shaped by microhabitat factors, as well as by latitudinal variables. Certainly, the host is an active shaper and recruiter of overlooked microbial eukaryote diversity, as the striking divergence between endophytic and epiphytic compartment suggests. However, this role may be better appreciable in a context of host species or genera comparison (Schlaeppi et al., 2014), while remains marginal in the context of ecotypes, as the displayed data indicate (par 2.2.3.2; 2.3.1).

This study is the first one providing evidence of an active role of microproducers (chlorophyta) in shaping microbial diversity in the phyllosphere in both compartments: epiphytic and endophytic.

I propose that this role manifest itself through the link with lichens, and as well through the constitution of biofilm. Further conclusions would need an experimental confirmation, however, one option would be to start especially from one promising candidate: *Bracteacoccus sp*.

Overlooked microbial eukaryotes manifest their effect on microbiota composition also through micropredators, with taxonomic groups such as Conosa, Lobosa and Ciliophora. Here for the first time they are revealed to be a key interactor shaping bacterial and fungal diversity, together with higher organisms from the Metazoa kingdom. (fig 3.1)

Finally, I showed that low abundant overlooked microbial eukaryotes are also key players in microbial assembly, likely to be crucial in determining the resilience of microbial community and therefore granting stability to the who
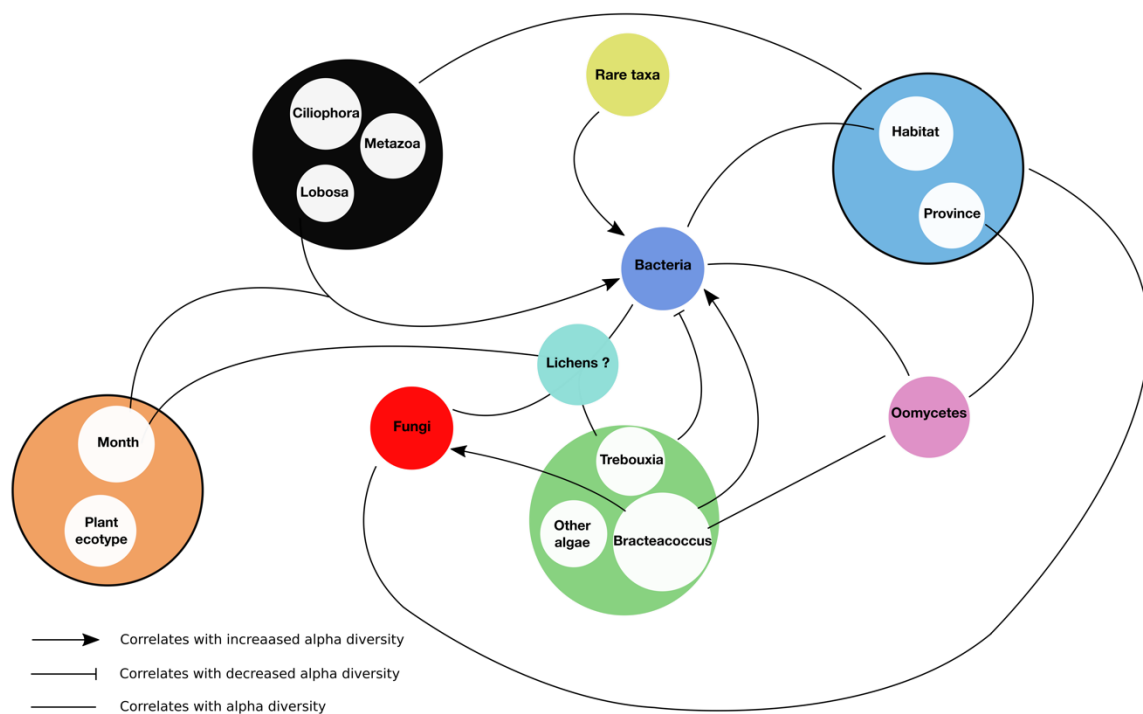
Figure 3.1 The illustrated microbe-microbe interaction after the inclusion of overlooked microbial eukaryotes. Summary of the findings, by cross comparison of alpha diversity results between the two experiments: geographical sampling: blue bubble, and common garden experiment (orange bubble).

le microbiome.

These results taken together indicate that for a full comprehension of plant-microbe interactions, overlooked microbial eukaryotes cannot be ignored any more, as they can account for variation that until now remained unexplained.

Ultimately, this study opens several new fields and perspectives on plant microbiome and ecology studies on terrestrial environments, unravelling a hidden world which is worth paying attention to

4 Materials and Methods

Site	LA	CH	CDC	MAR	SLY	KFA	KFB	KFC	UPP	ERG	WH	PFN	EY	K6	JUG
Country	FR	FR	ES	ES	ES	SE	SE	SE	SE	DE	DE	DE	DE	DE	DE
Latitude (GPS coord)	43.56	43.31	41.21	39.58	41.8	63.02	63.07	62.81	59.43	48.5	48.51	48.56	48.45	48.54	48.56
Longitude (GPS coord)	1.65	1.52	-4.55	-3.93	-3.11	18.32	18.32	18.2	17.02	8.81	8.94	9.11	8.78	9.09	9.13
Elevation (masl)	246	198	787	1024	1022	16	31	179	33	471	387	455	386	433	431
Average Temp (annual)(°C)	12.5	12.8	12.31	12.68	9.82	3.1	3.1	3.1	6.32	9.21	9.21	9.21	9.21	9.21	9.21
Average Minimum Temp (annual)(°C)	5.1	5.4	5.8	6.2	3.5	-1.27	-1.27	-1.27	1.95	4.83	4.83	4.83	4.83	4.83	4.83
Absolute Minimum Temp (month)(°C)	0.8	1.1	-0.3	-0.9	-2.3	-34.7	-34.7	-34.7	-33.4	-25	-25.6	-25.8	-25.6	-25.6	-25.6
Average Maximum Temp (°C) (annual)	19.9	20.2	18.5	19.2	16.1	7.36	7.36	7.36	10.71	13.7	13.7	13.7	13.7	13.6	13.6
Absolute Maximum Temp (month) (°C)	26.7	26.9	30.6	32.5	28.2	32	32	32	32.6	37.3	37.3	37.7	37.7	37.3	37.7
Average Annual Precip (mm)	749	719	439.7	488.5	657.6	718.39	718.39	718.39	440.65	702	703	703.2	703.5	702	704
Absolute Minimum Precip (month) (mm)	43	44	14.9	12.6	25.9	0.7	0.7	0.7	1.7	0.4	0.4	0.4	0.4	0.4	0.4
Absolute Maximum Precip (month) (mm)	78	76	51.5	62.8	78.3	323.7	323.7	323.7	110.7	197	195	199.9	200	199.6	198

Table 4.1 : GPS coordinates and climatic variables belonging to each site.

4.1 Materials

4.1.1 Sites coordinates and features

Detailed informations of the investigated sites are listed in table 4.1, including environmental variables, collected through meteorological stations in nearest proximity of the sites (C. Alonso-Blanco, J. Ågren, F. Roux, personal communication). Plants were sampled during early flowering time, respectively March in France and Germany, April in Spain, May in Sweden. Monitoring of the populations and on field assistance was provided for France, Spain, Germany, and Sweden, respectively by Fabrice Roux, Carlos Alonso-Blanco, Detlef Weigel, Jon Ågren

4.1.2 Plant materials

Plants used for the performed experiments were of the following ecotypes of *A.thaliana*:

-Keskwick-1 (KSK)

-Wassilewskija-0 (Ws-0)

-San Feliu-2 (Sf-2)

-Columbia – 0 (Col-0)

A.thaliana DNA for blocking oligos testing was extracted following the protocol described in (Agler et al., 2016a) from Ws-0 leaves

4.1.3 Microbiological material

Strain	Isolation from	Growing medium	Origin/Collector
Albugo laibachii Nc-14	*A.thaliana* Ws-0	\	Norwich-UK/Group Kemen
Albugo candida Nc-2	*A.thaliana* Ws-0	\	Norwich-UK/Group Kemen
Vannella sp.,	*A.thaliana* in EY site	NMAS	Tübingen-DE/Group Kemen
Cercomonas braziliensis,	*A.thaliana* in EY site	NMAS	Tübingen-DE/Group Kemen
Saccharomyces boulardii	Perenterol-Forte	PDA	Perenterol®
Sphingomonas sp.,	*A.thaliana* Ws-0 plants/infected by A.laibachii	LB	Cologne-MPIPZ/Group Kemen
Bacillus sp.	*A.thaliana* Ws-0 plants/infected by A.laibachii	LB	Cologne-MPIPZ/Group Kemen
Microthamnion kuetzingianum	Scotland freshwaters	3N-BBM+V	Bigelow labs US

Table 4.2 : Microbiological material used in the present studies.

The microbial strains used in this study are listed in table 4.2

4.1.4 Chemicals

NMAS medium:

Stock1

NaCl	12g
$MgSO_4$ x $7H_2O$	0.4g
$CaCl_2$ x $6H_2O$	0.6g
H_2O	500 ml

Stock2

Na_2HPO_4	14.2g
KH_2PO_4	13.6g
H_2O	500 ml

NMAS: water solution of stock1 (0.5%), and stock2 (0.5%) Autoclave.

DNA extraction buffer:

Tris pH 8.0 (2M)	0.25ml
NaCl (5M)	0.40ml
EDTA (Ethylenediaminetetraacetic acid) (0.5 M)	40µl
SDS (Sodiumdodecilsulfate) (10 %)	500µl
H_2O	up to 10ml
Proteinase K (20 mg/ml)	Add 5 µl to 1 ml buffer just before usage
Lysozyme	Add 1mg to 1ml buffer just before usage

3N-BBM+V:

first we created the following stock solutions, each one diluted in 100 ml distilled water:

$NaNO_3$	7.5g
$CaCl_2$ x 2 H_2O	0.25g
$MgSO_4$ x 7H_2O	0.75g
K_2HPO_4 x 3H_2O	0.75g
KH_2PO_4	1.75 g
NaCl	0.25g

Trace element solution: following compounds in 100ml

$FeCl_3$ x 6 H_2O	9.7mg
$MnCl_2$ x 4 H_2O	4.1mg
$ZnCl_2$	0.5mg
$CoCl_2$ x 6 H_2O	0.2mg
Na_2MoO_4 x 2 H_2O	0.4mg

Thiaminhydrochloride (vitamin B1)	0.12g
Cyanocobalamin (vitamin B12)	0.1g

3N-BBM+V: water solution 1% concentrated of stock solutions 1 to 6, 0.6%concentrated for stock 7. Afterwards the medium was autoclaved prior to add stock 8 and 9, each one with a concentration in the final medium of 0.1%.

Stocks 8 and 9 were sterilized through filter sterilization, using nitrocellulose filters with pores ø 5µm (Whatman- Merck, Darmstadt, DE).

4.2 Methods

4.2.1 Common garden experiment setup

The, experiment was performed at the MPIPZ ground (N Coord. 50° 57' 21.56" E Coord. 6° 51' 40.20"). Sowing, planting sampling, DNA extraction and library preparation was performed by Samuel Kroll. To include genetic host variability, 4 global *A. thaliana* genotypes were used (Ws-0, Col-0, Ksk-1 and Sf-2). Genotypes were chosen based on their variation in susceptibility towards *A. laibachii* and ability to survive within the field conditions (Agler et al., 2016a). Seeds were surface sterilized and stored in 0.1% Agarose for one-week prior sowing at 4°C. Seeds were then pipetted on jiffy's, which had been soaked in water prior usage and kept under short day conditions (10 h light, 14 h darkness, 23/20 °C, 60% humidity) for two weeks in the greenhouse.

4.2.2 Sampling procedure

Within the latitudinal sampling, for each sampling site, *Albugo sp.* infected and *Albugo sp.* uninfected samples were collected, and among those, whole leaf and endophytic sample types were collected as follows:

Using sterilized tweezers, the whole leaf samples were roughly cleaned from dirt, inserted into 2mL tubes, and stored immediately in dry ice, later at -80 ºC until further usage. The endophytic samples instead underwent surface sterilization.

For each of the types three to four replicates were collected. Each replicate is made of a single *Arabidopsis thaliana* rosette in case the diameter of the rosette was found to be equal or below 3 cm, in case of higher diameter, few single leaves from the rosette were taken.

For each site presenting *Albugo sp.* infection, there were collected from a minimum of 12 to a maximum of 16 samples, depending on the availability of plants. In the sites not presenting *Albugo sp.* infection, 6 to 8 uninfected samples were collected, half whole leaf and half endophytes. Across the years 2015, 2016, 2017, and across all the 15 sites, a total of 397 samples were collected.

Concerning common garden experiment, only whole leaf samples were collected.

Additional root endophytic samples and rhizoplane samples were collected by Paloma Duran and processed as following: total DNA was extracted from root episphere and root endosphere samples (four technical replicates) from three natural sites, using the FastDNA® SPIN Kit

for Soil (MP Biomedicals, Solon, USA) following the manufacturer's instructions. Samples were homogenized in the Lysis Matrix E tubes using the Precellys®24 tissue lyzer (Bertin

Technologies, Montigny-le-Bretonneux, France) at 6,200 rotations per second for 30 seconds. DNA samples were eluted in 60 μl nucleases-free water a and used for bacterial, fungal, oomycetal and overlooked microbial eukaryotes community profiling. For details of each sample, please consult the file Mapfile_geographical_sampling.xlsx and Mapfile_Common_garden.xlsx   at https://github.com/AlfredoMari/chetproject0001/tree/master/M&M

4.2.2.1 Surface sterilisation

On the site of collection endophytic samples were processed in order to remove the epiphytic fraction. Using sterilized tweezers each replicate was first inserted into 15mL tube (Corning Incorporated, USA), then washed with the following reagents:

- sterile milliQ water ~7mL for 15 sec
- 70% ethanol (Carl Roth, Karlsruhe, DE) ~7mL for 15 sec
- 2% bleach (Carl Roth, Karlsruhe, DE) ~10mL for 20 sec
- sterile milliQ water ~10mL for 25 sec (this step was repeated three times)

At the end of the process, samples are stored in dry ice, and later at -80ºC, until usage.

4.2.3 Sequencing design

In order to accomplish the sequencing, for bacteria, fungi and oomycetes, I used the barcodes primers reported in (Agler et al., 2016a). For overlooked microbial eukaryotes I first selected two pairs of primers based on the achievements of (Hadziavdic et al., 2014). The choice was justified by the ability of the primer to amplify the broadest possible range of organisms. From this publication were chosen both the 18S V4 (F566: 5'-CAGCAGCCGCGGTAATTCC-3'; R1200: 5'-CCCGTGTTGAGTCAAATTAAGC-3') and 18S V9 (F1422: 5'-ATAACAGGTCTGTGATGCCC-3'; R1797: 5'-TGATCCTTCTGCAGGTTCACCTAC-3') primer.

4.2.3.1 Barcodes design

For Bacteria, Fungi and Oomycetes, the barcodes used where the same reported in (Agler et al., 2016c). For 18S, the barcodes (forward or reverse) were formed by concatenating the following sequences:

- Illumina adapter (P5 or P7)
- Index

- Linker
- Primer (forward or reverse)

The Illumina adapter P5 was used for forward barcode design, and P7 for reverse primer design. In total 1 forward barcode and 50 reverse barcodes were produced. Each of them shows the same sequence in all parts except for the 12 bp index, this one different in each barcode (Schirmer et al., 2015). The linker was composed by a random 13bp dna sequence, and analysed for possible hairpins or self/paired untargeted amplifications. Final linker was chosen based on the lowest possible self-pairing and pair-pairing score calculated via DNAStar software (Madison, Wisconsin, USA)

Finally, the whole new barcode sequence was finalized by concatenating the primer separately, forward and reverse.

Sequencing primers were obtained by the reverse complement sequence of the primer sequence.

Further information and complete sequences of primer and barcode sequences may be found within the file named Illumina_barcodes.xlsx at the following path: https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Oligos

4.2.3.2 Blocking oligos design

In order to design blocking oligos, I first amplified the sequences 18SV4 and 18SV9 primers from *A. thaliana* and *Albugo laibachii.*, using the aforementioned primers.

After cloning in *E.coli* DH5α (PJet kit, Thermo Fisher, following the instruction of the manufacturer), sanger sequencing was performed in order to obtain the sequence of the fragment.

Generation of sequential oligos given the 18S sequence of the targets was performed as reported in(Agler et al., 2016c). Further testing and choice of the most efficient oligo is described in chapter 2.1. Sequences of the blocking oligos are reported in table 4.3

4.2.3.2.1 Mock community qPCR experiment

In order to test the efficiency of blocking oligos I built two backbone mock communities, one in which was present *Albugo laibachii*, therefore named *Infected mock*, and another without *Albugo laibachii*, therefore named *Uninfected mock*.

Composition of the two backbone mock communities is reported in table 4.3

Specie	Uninfected mock		Infected mock	
	Final concentration	Percentage	Final concentration	Percentage
A.thaliana (Ws-0)	9.6 ng/µl	97%	8.6 ng/µl	87%
Bacillus sp.	0.15 ng/µl	1.5%	0.15 ng/µl	1.5%
Sphingomonas sp.	0.15 ng/µl	1.5%	0.15 ng/µl	1.5%
Albugo laibachii	-	-	1 ng/µl	10%

Table 4.3 : Composition of the mock community.

These backbone mocks, uninfected and infected were splitted in four equal aliquots, in each of them a different percentage of *Saccharomyces boulardii* gDNA was then added, corresponding respectively:

- 1%, equal to 0.1 ng/µl
- 0.1% equal to 0.01 ng/µl
- 0.01% equal to 0.001 ng/µl
- 0.001% equal to 0.0001 ng/µl

Oligos sequences for mock community experiment	
ID	Sequence (5'-3')
G013 (SACf)	AACCTTGAGTCCTTGTG
G014 (SACr)	AATACGCCTGCTTTG
Gc002	TTGTCCCTTCGGTCGGCGATACGCTCCTGGTCTTA
Gc003	GTGCCAGCGGAGTCCTATAAGCAACATCCGCTGAT
Gc010	TGGATTTCTGATTCGAGCGTCCGGTCCGCTTCTTTAGGA
Gc012	CGGTGCTGACAAGGTCATTTAAAGTAAACGACTGCCAATC

Table 4.4 : primers and blocking oligos sequences: G013-014: qPCR primers for S.boulardi. Gc002-003: blocking oligos, Fwd and Rev respectively for A.thaliana. Gc010-012: blocking oligos, Fwd and Rev respectively, for A.laibachii

The communities were analysed through qPCR using the primers SACf and SACr and the blocking oligos Gc002-Gc003, and Gc010-Gc012, respectively targeting *A.thaliana* and *Albugo sp..* displayed in table 4.4.

Each sample, (8 in total) was divided in two and underwent endpoint PCR respectively, with (w) or without (w/o) blocking oligos. Each mix contained: 4µl of 5_x NEB buffer, 0.4 µl dNTPs (µM), 0.4 µl SACf (10µM), 0.4 µl SACr (10µM), 0.1 µl of Phusion Taq polymerase. In the mix containing oligos (w), also 0.65 µl of each blocking oligo (Gc002,Gc003,Gc010,Gc012) was added. Therefore, the quantity of nuclease free water was 11.2 µl in the mix with blocking oligos (w) and 13.8 µl in the mix without oligos (w/o). The mix, was afterwards added with the template and set into a PCR cycler (Eppendorf, Hamburg, DE) , the following protocol was used: 95 ºC for 3 min, followed by 35 cycles of 95 ºC for 30 sec, 55 ºC for 30 sec, 72 ºC for 1 min, then 72 ºC for 5 min.

The amplified samples were then aliquoted in 10µl and cleaned from free dNTPs and primers. An enzymatic cleanup was performed by adding the following mix: Antarctic Phosphatase 0.5µl, Exonuclease I 0.5 µl, Antarctic phosphatase buffer: 1.22 µl, (New England Laboratories, Ipswich, Massachussets, USA) incubated for 30 min at 37 ºC, followed by a step at 85 ºC for 15 min.

With the cleaned up samples was then performed a qPCR using the *S.boulardii* 18S internal primers SACf and SACr listed in table 4.4. qPCR was prepared in 96 well plates, using Sybr Mix as amplification kit (Bio-Rad, Hercules, California, US). The mix, per sample, included 1.05 µl of DNA template, 6 µl of primer mix (Fwd and Rev primer, nuclease free water) concentrated 240nM, 8.4 µl of Sybr Mix. Each amplification was carried in three technical replicates. As negative control I included nuclease free water, as positive control pure *S.boulardii* DNA, 50 ng/µl concentrated. S. boulardii DNA was extracted as reported in par 4.2.3.3

The protocol used for qPCR in termocycler (Bio-Rad, Hercules, California, US) was the following: 95 ºC for 3 min, followed by 35 cycles of 95 ºC for 10 sec, 55 ºC for 30 sec, 72 ºC for 30 sec, fluorescence measuring (488nm). At the end of the protocol, melting temperature of the fragments was measured by measuring fluorescence after increments of 0.5 ºC, starting from 55 ºC, ending at 95 ºC.

There was no fragment among the samples presenting a melting curve different from the positive control. The negative control showed no amplification in none of the three replicates. Results of this experiments are showed in paragraph 2.1.1

4.2.3.3 DNA extraction

Extraction of DNA took place as follows: circa 100 mg of frozen samples were first ground in 2 ml eppendorf tubes (Eppendorf, Hamburg, DE) using autoclaved plastic pestles. Afterwards, 600 µl of extraction buffer, the sample was shaken for 30sec and then incubated at 37 ºC for 45 minutes. 0.2mg of acid washed quarz beads of respectively 0.5 and 0.1 mm diameter (Carl Roth, Karlsruhe, DE) were added to the sample, and beaten in Precellys24 (Bertin technologies, France) using the protocol stated in Agler et al., 2016a. Afterwards 4 µl of RNAse A 50 µg/mL was added and the whole sample incubated for further 45 min at 37 ºC.

The sample was then washed once with 650 µl of Phenol/Chloroform/Isoamylalcohol the supernatant was then transferred and further washed with Chloroform/Isoamylalcohol 24:1. Afterwards 40 µl of Sodium Acetate 3M was added, together with 1500 µl of 100% Ethanol (Carl Roth, Karlsruhe, DE). After delicate inversion, the sample was stored overnight at -20ºC. Ultimately it was centrifuged at maximum speed for 45 min and washed two times with 70% ethanol. The pellet was finally re-suspended in 30µl of Tris HCl ph8 10mM and incubated with a water solution of Cheelex 20% (Bio-Rad, Hercules, California, US) for 30 min at room temperature.

After centrifugation, the supernatant was recovered and stored at -20ºC till usage.

4.2.3.4 Library preparation

Library preparation had the goal to amplify the selected regions (Bacteria 16S V3 &V5, Fungal ITS1 & ITS2, Oomycetes ITS1 & ITS2 and overlooked microbial eukaryotes 18S V4 & V9). Prior to library preparation, a test endpoint PCR with bacteria BV5 primers (see Agler et al., 2016b) was performed in order to discriminate samples poorly purified or contaminated by PCR inhibitors. In that case, the sample was further cleaned with Chloroform/Isoamylalcohol 24:1, precipitated with ethanol and then cleaned with Cheelex as described in the paragraph above. The sample was afterwards diluted with Tris HCl 10mM to reach the concentration of 50 ng/µl, and with it was performed the first PCR step. In the first step PCR, blocking oligos for each primer set were inserted as described in Agler et al., 2016. Each reaction was composed by: Q5 High-GC buffer: 3.85µl, Q5 5_x buffer: 3.85µl, dNTPs(10µM): 0.45µl, Q5 Enzyme: 0.2µl, Primers(10µM) (Fwd & Rev, (BV3/BV5/FITS2/Ftrad/OITS2/Otrad/PV4/PV9): 0.16µl, Blocking oligos (10µM): 0.5µl. Since for overlooked microbial eukaryotes I inserted two pairs of oligos, one for *Albugo sp.*, and one for *A.thaliana*, the added nuclease free water was 9.83µl for bacteria, fungal

and oomycete mix, whereas for overlooked microbial eukaryotes I added 8.83µl. 1µl (~50ng) of the sample template was added to the reaction. Each reaction was carried in three technical replicates. The protocol used in the termocycler (Eppendorf, Hamburg, DE) was: 95ºC for 40 sec, followed by 10 cycles of 95ºC for 35 sec, 55ºC for 45 sec, 72ºC for 15 sec, 72ºC for 2 min.

Afterwards the three replicates were combined, mixed, and an aliquot of 10µl was taken and subjected to enzymatic clean-up with Antarctic phosphatase and Exonuclease I as described in paragraph 4.2.3.2.1. The second step PCR was carried in a single replicate. One single reaction included: Q5 High-GC buffer: 10µl, Q5 5_x buffer: 10µl, dNTPs(10µM): 1µl, Q5 Enzyme: 0.5µl, Barcoded Primers(10µM) (Fwd & Rev, (BV3/BV5/FITS2/Ftrad/OITS2/Otrad/PV4/PV9): 0.83µl, nuclease free water: 26.34µl. To each reaction was added 0.67µl of the cleaned-up 1st step PCR product.

The protocol used in the thermocycler was: 95ºC for 40 sec, followed by 25 cycles of 95ºC for 35 sec, 55ºC for 45 sec, 72ºC for 15 sec, 72ºC for 2 min.

The barcoded amplicons were then cleaned using Ampure XP beads purification. 50µl of barcoded amplicon was added with 40µl of magnetic beads solution, shaken 5 min at 700 rpm at room temperature, then washed twice with 200µl of 80% ethanol, dried, and then resuspended in 25µl of TrisHCl 10mM.

4.2.3.5 Pooling and Sequencing

Libraries were then quantified using PicoGreen (ThermoFischer, Waltham, Massachusetts, USA) following the indications of the manufacturer, concentration was calculated using a standard curve set with known concentrations of Salmon Sperm DNA(Invitrogen, Waltham, Massachusetts, USA), and evaluated reliable only when R^2 score of the curve was higher or equal to 0.98.

Libraries quantified in this manner where pooled together, aiming to pool 8 pmol for each library. Pooled libraries were cleaned using Ampure XP beads as described in paragraph 4.2.3.2.1, and re-suspended in one tenth of the original volume of Tris HCl 10mM (total 200 µl). Quality assessment and eventual presence of contamination was assessed through bioanalyzer (Agilent, Santa Clara, California, US). Bands composing the pool were compared with standard libraries chosen random from the not pooled ones. Bioanalyzer curves and bands can be observed in supplementary figures at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Libraries.

Final concentration of the pooled libraries was finally assessed through Qbit (Promega, Fitchburg, Wisconsin, US) following the indications of the manufacturer. Due to the fact that only 50 barcoded primers per locus were available, each Miseq run could contain a maximum of 50 libraries per run. In total 8 runs were performed. We used Illumina MiSeq to sequence the libraries, loading the machine as described in Agler et al., 2016.

4.2.4 Reads processing

All 8 runs were yielding reads for 8 GB per run or more. Raw reads, both forward and reverse, as well as index reads were checked for quality of the basecalling through FastQC (Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc): all runs displaying an average Phred score(Schirmer et al., 2015) per read lower than 30 or more than 2 N bases per read were discarded and re-sequenced. Raw reads and FastQC profiles for consultation are available upon request.

4.2.4.1 OTU clustering pipeline

Each run was processed in order to remove the adapter sequence, pair forward and reverse read, then demultiplexed and clustered in OTUs with 0.97 confidence as described in Agler et al., 2016. Regarding bacteria (BV3-BV5), fungi (FITS2, Ftrad), and oomycetes (OITS2,Otrad), I applied the downstream steps of the presented pipeline described in Agler et al., 2016. Regarding overlooked microbial eukaryotes, I performed the taxonomy assignment by modifying the pipeline presented in Agler et al., 2016. I created two reference databases, respectively one for the 18S V4 and one for V9 region, based on in silico PCR on the original Pr2 database (Guillou et al., 2013).

Here follows a summary of the database preparation:

- Isolate from the PR2 database the V4-5 and V8-9 regions
- Cluster OTUs from the resulting two databases
- Pick up representative sequences from the two databases
- Format the two new databases in order to avoid mismatches and mistakes in downstream analyses, this included:
- Removal of empty lines
- Removal of empty spaces
- Removal of sequences with degenerated bases

The reference databases obtained in this manner are available upon request. Through these new databases, I then performed a parallelized taxonomy assignment using the qiime platform (Caporaso et al., 2010).

4.2.4.2 OTU table filtering and rarefaction

In order to create the final OTUtable I crafted the perl script OTU_tab_gen.pl available at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts . In short, from the otu mapfile and the taxonomy assignments the script:

Created a native otu table

Discarded the otus present in less than 2 samples and abundant less than 50 reads

Divided the tables in one containing only endophytes and the other containing only whole leaf samples

Resulting tables were processed through the perl script OTU_filter_pipe.pl, available at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts , which in short was performing:

A cross filtering between the otus present in whole leaf tables but not in the endophytic tables, this way producing an epiphytic only otu table, containing only obligate epiphytes.

Merging respectively endophytes (root and leaf) and epiphytes (root and leaf) in two separate tables

Primer set	Compartment	Rarefaction threshold
BV3-(bacteria)	Endophyte	4415 reads
BV5-(bacteria)	Epiphyte	504 reads
FITS2-(fungi)	Endophyte	2052 reads
Ftrad-(fungi)	Epiphyte	1033 reads
OITS2-(oomycetes)	Endophyte	5309 reads
Otrad-(oomycetes)	Epiphyte	309 reads
PV4-(microbial eukaryotes)	Endophyte	3055 reads
PV9-(microbial eukaryotes)	Epiphyte	412 reads

Table 4.5 : rarefaction thresholds chosen for each marker/compartment

Filtering the untargeted taxa.

Concerning rarefaction, I screened the number of reads per samples through BIOM tool set, available at www.biom-format.org choosing as rarefaction thresholds the ones stated in table 4.5

After rarefaction, tables have been merged by marker and stored for further usage.

For all downstream analysis, only the following loci were chosen:

- BV5
- FITS2
- OITS2
- PV4

This choice was made after considering the substantial overlap in terms of taxa found by each couple of regions, as stated in (Agler et al., 2016a). Regarding the choice between PV4 and PV9, I found there a s well a prominent overlap, as stated in figure 4.1. In this case, the discrimination was made towards the sequence which has been reported to have major taxa resolution(Hadziavdic et al., 2014).

Figure 4.1 : Class overlap between 18S V4 region (blue) and 18S V9 region (orange)

4.2.5 Network inferring and comparing

Network analysis were obtained through inferring of microbial network using the platform CoNET (Faust and Raes, 2016).

4.2.5.1 CoNET and inferring pipeline

In order to infer the network displayed in figure 2.7 and 2.8, the Co-Net pipeline was used in the command line form, and embedded in a sequential script (Co_Net_Embedder.pl), available at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts . Most of the parameters used were set as default in the three embedded Java subscripts, as suggested by (http://psbweb05.psb.ugent.be/conet/microbialnetworks/conet_new.php) and by the developer, (Karoline Faust, personal communication).

The parameters left to the customization of the user were only the abundance filter threshold - minimal number of reads to take in consideration to represent an OTU as a node –and the guessing parameter –The number of edges to be considered in the automatic distance metrics thresholds- (Karoline Faust, 2017). For the figure 2.7, filterparameter and guessingparameter were respectively set at 20 and 1000, for the figure 2.8 at 10 and 300.

4.2.5.2 Network comparing pipeline

The network comparing pipeline is coded in the script Network_comparison.pl available at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts   The conceptual workflow of the pipeline is the following:

Extraction of the chosen edge score out of the network table (/Subscript: Extractor.R)

Formation of the distance matrices based on the edge scores of the two networks (/Subscript: matrix_maker_weight.R)

Mantel correlogram and Mantel test between being perfomed from the two distance matrices. (compare_distance_matrices.py - QIIME)

Dissection of the results of the correlogram and tracking of the nodes composing the single classes (/Subscript: zweite.R and Backtracker.R)

The pipeline offers the possibility to extract two variables from the network table: edge betweenness and edge weight. For the purpose of this thesis only edge weight variable has been used. Further variables can be implemented easily.

4.2.6 Overlooked microbial eukaryotes presence/absence assessment and rare taxa evaluation

Within the results sections, it has been made large use of analysis of samples containing or not containing a certain taxonomic group. This labelling according to presence of certain taxa is visible in the mapfiles (https://github.com/AlfredoMari/chetproject0001tree/master/M%26M). This labelling was achieved through the script map_adder.R available at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts.

In short, the script performed for each sample the sum of the reads assigned to the chosen taxonomic group, if this sum was positive or equal to 0, the sample was labelled accordingly, respectively as X_taxa_present or X_taxa_absent. Concerning rare taxa, a modification of the mentioned script was used. The modification first selected the species whose abundance sum was equal or below the 10th percentile. Presence or absence of those species was then assessed in the same manner of the script Variable_partitioner.R at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts.

4.2.7 Phylogenetic tree construction

The construction of the phylogenetic trees was taxonomy based and achieved as described in (Agler et al., 2016b)

4.2.8 Alpha diversity calculation

Alpha diversity was calculated using the script alpha_diversity.py (QIIME(Caporaso et al., 2010)), using as metric the Chao1 index.

4.2.9 Statistical Methods

Statistical tests used for this study were performed through R programming language.

4.2.9.1 ANOVA

ANOVA (was performed through the function *aov* and *anova* within the *vegan* package (Oksanen et al., 2018). A baseline ANOVA performing script can be found under ANOVA_baseline.R   at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts.

For all the different ANOVAs, this baseline script was adapted from time to time including the different required variables.

4.2.9.2 Wilcoxon/Mann-Whitney test and PCoA

Wilcoxon-Mann Whitney test was performed through the R base function *wilcox.test* . Principal Component Analysis was performed through the capscale function within vegan package. The code producing supplementary videos 1 to 4 is available in PCoA_3d_movie.R   at https://github.com/AlfredoMari/chetproject0001/tree/master/M%26M/Scripts.

4.2.9.3 Graphical interfaces

Plots and videos were achieved through the R package *ggplot2* (Wickham et al., 2016), *rgl*, (Adler et al., 2018), and for the supplementary videos, *magick* (Ooms, 2018).

4.3 Epiphytic sugar experiment

4.3.1 Plant growth and leaf wash

A. thaliana seeds were stratified on moist soil (Einheitserde type A240, Stender, Germany) for seven days at 4 °C in darkness, before transfer into growth chambers with short day conditions (10 h light, 14 h darkness, 23/20 °C, 60 % humidity). After three weeks, seedlings were singularised and grown for three weeks further. After six weeks plants were used for experiments.

M.kuetingianum was cultured in 3N-BBM+V for 7 days, centrifuged for 5 minutes at 1200 rpm, the supernatant discarded and substituted with autoclaved solution of $MgCl_2$, 10mM. Washed cells were adjusted to reach $2x10^4$ cells/ml.

Algal suspension was sprayed on the plants in the quantity of circa 1mL per plant using an airbrush gun (Conrad, Hirschau, DE). Control plants were sprayed with autoclaved $MgCl_2$ 10mM.

Both test and control were sprayed on 54 single plants. Afterwards, plants were put back in growth chambers, following the same growing conditions mentioned above, inside a transparent plastic bag in order to conserve a higher moisture.

The bags were removed 3 days after. After four more days (total of one week of post-spraying growth), plants were ready for the leaf wash.

During the leaf wash, leaves were detached from the rosette, put in 50 ml tubes (Cornig Incorporated, USA) filled with 20mL autoclaved water, inserted in the tube in groups of 10 leaves, and then shaken for circa 5-10 sec. The liquid was retained and the procedure was repeated till all leaves had been washed.

The solution was then concentrated via speed-vac (Concentrator plus, Eppendorf, Hamburg, DE) and then resuspended in 500 µl. Leaf wash obtained in this manner was then analysed via enzymatic assay.

4.3.2 Enzymatic assay

The measurement of glucose, fructose, and sucrose, was performed as described in (Velterop and Vos, 2001) using transparent 96 well plate (Corning Incorporated, USA) filled separately with leaf washes from the control and the test, as well as sugar standards of glucose, sucrose, fructose, in the concentration of 10mM. all the reactions were carried in the volume of 50 µl, reading each well with set absorbance of 340 nm.

5 References:

Adl, S.M., Simpson, A.G.B., Farmer, M. a, Andersen, R. a, Anderson, O.R., Barta, J.R., Bowser, S.S., Brugerolle, G., Fensome, R. a, Fredericq, S., et al. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J. Eukaryot. Microbiol. *52*, 399–451.

Adler, D., Murdoch, D., and others (2018). rgl: 3D Visualization Using OpenGL., URL https://CRAN.R-project.org/package=rgl

Agler, M.T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S.-T., Weigel, D., and Kemen, E.M. (2016a). Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. PLOS Biol. *14*, e1002352–e1002352.

Agler, M.T., Mari, A., Dombrowski, N., Hacquard, S., and Kemen, E.M. (2016b). New insights in host-associated microbial diversity with broad and accurate taxonomic resolution. BioRxiv 050005.

Ahmadjian, V. (1960). Some New and Interesting Species of Trebouxia, a Genus of Lichenized Algae. Am. J. Bot. *47*, 677–683.

Ahmadjian, V. (1967). A Guide to the Algae Occurring as Lichen Symbionts : Isolation , Culture , Cultural Physiology , and Identification Phycologia *6*.2-3

Ainsworth, T.D., Krause, L., Bridge, T., Torda, G., Raina, J.-B., Zakrzewski, M., Gates, R.D., Padilla-Gamiño, J.L., Spalding, H.L., Smith, C., et al. (2015). The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. ISME J. *9*, 2261–2274.

Baas-Becking, L.G.M. (1934). Geobiologie; of inleiding tot de milieukunde (WP Van Stockum & Zoon NV).

Barabasi, A.-L. (2009). Scale-Free Networks: A Decade and Beyond. Science *325*, 412–413.

Barberán, A., Bates, S.T., Casamayor, E.O., and Fierer, N. (2011). Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J. *6*, 343–351.

Barberán, A., Bates, S.T., Casamayor, E.O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J. *6*, 343–351.

Bary, A. de (1863). Developpement de quelques champignons parasites. Ann. Sci. Nat. Ser 4 Bot.

Bass, D., Moreira, D., López-García, P., Polet, S., Chao, E.E., Heyden, S. von der, Pawlowski, J., and Cavalier-Smith, T. (2005). Polyubiquitin Insertions and the Phylogeny of Cercozoa and Rhizaria. Protist *156*, 149–161.

Bates, S.T., Cropsey, G.W.G., Caporaso, J.G., Knight, R., and Fierer, N. (2011). Bacterial Communities Associated with the Lichen Symbiosis. Appl. Environ. Microbiol. *77*, 1309–1314.

Bauer, W.D., and Mathesius, U. (2004). Plant responses to bacterial quorum sensing signals. Curr. Opin. Plant Biol. *7*, 429–433.

Belnap, J., Büdel, B., and Lange, O.L. (2001). Biological Soil Crusts: Characteristics and Distribution. In Biological Soil Crusts: Structure, Function, and Management, (Springer, Berlin, Heidelberg), pp. 3–30.

Berendsen, R.L., Pieterse, C.M.J., and Bakker, P. a H.M. (2012). The rhizosphere microbiome and plant health. Trends Plant Sci. *17*, 478–486.

Bogino, P.C., Oliva, M. de las M., Sorroche, F.G., and Giordano, W. (2013). The Role of Bacterial Biofilms and Surface Components in Plant-Bacterial Associations. Int. J. Mol. Sci. *14*, 15838–15859.

Bonanno, G., and Orlando-Bonaca, M. (2018). Trace elements in Mediterranean seagrasses and macroalgae. A review. Sci. Total Environ. *618*, 1152–1159.

Bonkowski, M. (2004). Protozoa and plant growth: the microbial loop in soil revisited. New Phytol. *162*, 617–631.

Boo, S.M., Kim, H.S., Shin, W., Boo, G.H., Cho, S.M., Jo, B.Y., Kim, J.-H., Kim, J.H., Yang, E.C., Siver, P.A., et al. (2010). Complex phylogeographic patterns in the freshwater alga Synura provide new insights into ubiquity vs. endemism in microbial eukaryotes. Mol. Ecol. *19*, 4328–4338.

Borror, A.C. (1980). Spatial Distribution of Marine Ciliates: Micro-Ecologic and Biogeographic Aspects of Protozoan Ecology*†. J. Protozool. *27*, 10–13.

Bruun, H.H., Moen, J., Virtanen, R., Grytnes, J.-A., Oksanen, L., Angerbjörn, A., and Ezcurra, E. (2006). Effects of altitude and topography on species richness of vascular plants, bryophytes and lichens in alpine communities. J. Veg. Sci. *17*, 37–46.

Buée, M., De Boer, W., Martin, F., Van Overbeek, L., and Jurkevitch, E. (2009). The rhizosphere zoo: an overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and of some of their structuring factors. Plant Soil *321*, 189–212.

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., et al. (2012). Revealing

structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature *488*, 91–95.

Callaway, R.M., Brooker, R.W., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., Paolini, L., Pugnaire, F.I., Newingham, B., Aschehoug, E.T., et al. (2002). Positive interactions among alpine plants increase with stress. Nature *417*, 844–848.

Campbell, B.J., Yu, L., Heidelberg, J.F., and Kirchman, D.L. (2011). Activity of abundant and rare bacteria in a coastal ocean. Proc. Natl. Acad. Sci. *108*, 12776–12781.

de Candolle, A.P. (1820). Essai élémentaire de géographie botanique (FS Laeraule).

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., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. *6*, 1621–1624.

Cardinale, M., Grube, M., Erlacher, A., Quehenberger, J., and Berg, G. (2015). Bacterial networks and co-occurrence relationships in the lettuce root microbiota. Environ. Microbiol. *17*, 239–252.

Cavicchioli, R. (2015). Microbial ecology of Antarctic aquatic systems. Nat. Rev. Microbiol. *13*, 691–706.

Coleman-Derr Devin, Desgarennes Damaris, Fonseca-Garcia Citlali, Gross Stephen, Clingenpeel Scott, Woyke Tanja, North Gretchen, Visel Axel, Partida-Martinez Laila P., and Tringe Susannah G. (2015). Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species. New Phytol. *209*, 798–811.

Cordovez, V., Carrion, V.J., Etalo, D.W., Mumm, R., Zhu, H., Wezel, V., P, G., and Raaijmakers, J.M. (2015). Diversity and functions of volatile organic compounds produced by Streptomyces from a disease-suppressive soil. Front. Microbiol. *6*.

Cosson, P., and Soldati, T. (2008). Eat, kill or die: when amoeba meets bacteria. Curr. Opin. Microbiol. *11*, 271–276.

Coyte, K.Z., Schluter, J., and Foster, K.R. (2015). The ecology of the microbiome: Networks, competition, and stability. Science *350*, 663–666.

De Bary, A. (1879). Die erscheinung der symbiose (Verlag von Karl J. Trübner).

Debroas, D., Hugoni, M., and Domaizon, I. (2015). Evidence for an active rare biosphere within freshwater protists community. Mol. Ecol.

Delafont, V., Samba-Louaka, A., Bouchon, D., Moulin, L., and Héchard, Y. (2015). Shedding light on microbial dark matter: a TM6 bacterium as natural endosymbiont of a free-living amoeba. Environ. Microbiol. Rep. *7*, 970–978.

Elmqvist Thomas, Folke Carl, Nyström Magnus, Peterson Garry, Bengtsson Jan, Walker Brian, and Norberg Jon (2003). Response diversity, ecosystem change, and resilience. Front. Ecol. Environ. *1*, 488–494.

Epstein, S.S. (2009). Microbial awakenings.

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. Nat. Rev. Microbiol. *10*, 538–550.

Faust, K., and Raes, J. (2016). CoNet app: inference of biological association networks using Cytoscape. F1000Research *5.1519*.

Faust, K., Lima-Mendez, G., Lerat, J.-S., Sathirapongsasuti, J.F., Knight, R., Huttenhower, C., Lenaerts, T., and Raes, J. (2015). Cross-biome comparison of microbial association networks. Front. Microbiol. *6-1200*.

Favero-Longo, S.E., and Piervittori, R. (2010). Lichen-plant interactions. J. Plant Interact. *5*, 163–177.

Filker, S., Sommaruga, R., Vila, I., and Stoeck, T. (2016). Microbial eukaryote plankton communities of high-mountain lakes from three continents exhibit strong biogeographic patterns. Mol. Ecol. *25*, 2286–2301.

Fiz-Palacios, O., Romeralo, M., Ahmadzadeh, A., Weststrand, S., Ahlberg, P.E., and Baldauf, S. (2013). Did Terrestrial Diversification of Amoebas (Amoebozoa) Occur in Synchrony with Land Plants? PLOS ONE *8*, e74374.

Flues, S., Blokker, M., Dumack, K., and Bonkowski, M. (2018). Diversity of Cercomonad Species in the Phyllosphere and Rhizosphere of Different Plant Species with a Description of Neocercomonas epiphylla (Cercozoa, Rhizaria) a Leaf-Associated Protist. J. Eukaryot. Microbiol. 1-13.

Foissner, W. (1999). Protist Diversity: Estimates of the Near-Imponderable. Protist *150*, 363–368.

Fuhrman, J.A. (2009). Microbial community structure and its functional implications. 459-14-05.

Galand, P.E., Casamayor, E.O., Kirchman, D.L., and Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. Proc. Natl. Acad. Sci. *106*, 22427–22432.

Goenawan, I.H., Bryan, K., and Lynn, D.J. (2016). DyNet: Visualization and analysis of dynamic molecular interaction networks. Bioinformatics *32*, 2713–2715.

Griffiths, B.S. (1994). Microbial-feeding nematodes and protozoa in soil: Their effectson microbial activity and nitrogen mineralization in decomposition hotspots and the rhizosphere. Plant Soil *164*, 25–33.

Grigulis, K., Lavorel, S., Krainer, U., Legay, N., Baxendale, C., Dumont, M., Kastl, E., Arnoldi, C., Bardgett, R.D., Poly, F., et al. (2013). Relative contributions of plant traits and soil microbial properties to mountain grassland ecosystem services. J. Ecol. *101*, 47–57.

Grube, M., and Berg, G. (2009). Microbial consortia of bacteria and fungi with focus on the lichen symbiosis. Fungal Biol. Rev. *23*, 72–85.

Grube, M., Cardinale, M., de Castro Jr, J.V., Müller, H., and Berg, G. (2009). Species-specific structural and functional diversity of bacterial communities in lichen symbioses. ISME J. *3*, 1105–1115.

Guennoc, C.M., Rose, C., Labbe, J., and Deveau, A. (2017). Bacterial Biofilm Formation On Soil Fungi: A Widespread Ability Under Controls. BioRxiv 130740.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., et al. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. *41*, D597–D604.

Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A.C., Dangl, J.L., Knight, R., Ley, R., et al. (2015). Microbiota and Host Nutrition across Plant and Animal Kingdoms. Cell Host Microbe *17*, 603–616.

Hacquard, S., Kracher, B., Hiruma, K., Münch, P.C., Garrido-Oter, R., Thon, M.R., Weimann, A., Damm, U., Dallery, J.-F., Hainaut, M., et al. (2016). Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. Nat. Commun. *7-11362*.

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E.M., and Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. PloS One *9*, e87624–e87624.

Hanshew, A.S., Mason, C.J., Raffa, K.F., and Currie, C.R. (2013). Minimization of chloroplast contamination in 16S rRNA gene pyrosequencing of insect herbivore bacterial communities. J. Microbiol. Methods *95*, 149–155.

Hardoim, P.R., Overbeek, L.S. van, Berg, G., Pirttilä, A.M., Compant, S., Campisano, A., Döring, M., and Sessitsch, A. (2015). The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. Microbiol. Mol. Biol. Rev. *79*, 293–320.

Hartmann, A., and Schikora, A. (2012). Quorum Sensing of Bacteria and Trans-Kingdom Interactions of *N*-Acyl Homoserine Lactones with Eukaryotes. J. Chem. Ecol. *38*, 704–713.

Hassani, M.A., Durán, P., and Hacquard, S. (2018). Microbial interactions within the plant holobiont. Microbiome *6*, 58.

Hawksworth, D.L. (1988). The variety of fungal-algal symbioses, their evolutionary significance, and the nature of lichens. Bot. J. Linn. Soc. *96*, 3–20.

Heger, T.J., Mitchell, E.A.D., and Leander, B.S. (2013). Holarctic phylogeography of the testate amoeba Hyalosphenia papilio (Amoebozoa: Arcellinida) reveals extensive genetic diversity explained more by environment than dispersal limitation. Mol. Ecol. *22*, 5172–5184.

van der Heijden, M.G.A., and Hartmann, M. (2016). Networking in the Plant Microbiome. PLoS Biol. *14*, 1–9.

Heino, J., Bini, L.M., Karjalainen, S.M., Mykrä, H., Soininen, J., Vieira, L.C.G., and Diniz-Filho, J.A.F. (2010). Geographical patterns of micro-organismal community structure: are diatoms ubiquitously distributed across boreal streams? Oikos *119*, 129–137.

Hess, S., and Melkonian, M. (2013). The Mystery of Clade X: Orciraptor gen. nov. and Viridiraptor gen. nov. are Highly Specialised, Algivorous Amoeboflagellates (Glissomonadida, Cercozoa). Protist *164*, 706–747.

Hilbi, H., Weber, S.S., Ragaz, C., Nyfeler, Y., and Urwyler, S. (2007). Environmental predators as models for bacterial pathogenesis. Environ. Microbiol. *9*, 563–575.

Hol, W.H.G., Boer, W.D., Termorshuizen, A.J., Meyer, K.M., Schneider, J.H.M., Dam, N.M.V., Veen, J.A.V., and Putten, W.H.V.D. (2010). Reduction of rare soil microbes modifies plant–herbivore interactions. Ecol. Lett. *13*, 292–301.

Hornschuh, M., Grotha, R., and Kutschera, U. (2006). Moss-associated methylobacteria as phytosymbionts: an experimental study. Naturwissenschaften *93*, 480–486.

Hugenholtz, P., Goebel, B.M., and Pace, N.R. (1998). Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J. Bacteriol. *180*, 4765–4774.

Jarosz, L.M., Ovchinnikova, E.S., Meijler, M.M., and Krom, B.P. (2011). Microbial Spy Games and Host Response: Roles of a Pseudomonas aeruginosa Small Molecule in Communication with Other Species. PLOS Pathog. *7*, e1002312.

Jones, C.S., and Mayfield, S.P. (2012). Algae biofuels: versatility for the future of bioenergy. Curr. Opin. Biotechnol. *23*, 346–351.

Jones, S.E., and Lennon, J.T. (2010). Dormancy contributes to the maintenance of microbial diversity. Proc. Natl. Acad. Sci. *107*, 5881–5886.

Karoline Faust (2017). CoNet documentation. User manual available at http://psbweb05.psb.ugent.be/conet/manual.php.

Krome, K., Rosenberg, K., Dickler, C., Kreuzer, K., Ludwig-Müller, J., Ullrich-Eberius, C., Scheu, S., and Bonkowski, M. (2009). Soil bacteria and protozoa affect root branching via effects on the auxin and cytokinin balance in plants. Plant Soil *328*, 191–201.

Krome, K., Rosenberg, K., Bonkowski, M., and Scheu, S. (2010). Grazing of protozoa on rhizosphere bacteria alters growth and reproduction of Arabidopsis thaliana. Soil Biol. Biochem. *41*, 1866–1873.

Landeghem, S.V., Parys, T.V., Dubois, M., Inzé, D., and de Peer, Y.V. (2016). Diffany: an ontology-driven framework to infer, visualise and analyse differential molecular networks. BMC Bioinformatics *17*, 18.

Lawrey, J.D. (1991). Biotic Interactions in Lichen Community Development: A Review. The Lichenologist *23*, 205–214.

Lebeis, S.L., Paredes, S.H., Lundberg, D.S., Breakfield, N., Gehring, J., McDonald, M., Malfatti, S., Rio, T.G. del, Jones, C.D., Tringe, S.G., et al. (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science aaa, 10.1126-8764.

LEFÈVRE, M. (1974). La 'Maladie Verte' De Lascaux. Stud. Conserv. *19*, 126–156.

Legendre, P., and Legendre, L. (1998). Numerical ecology: second English edition. Dev. Environ. Model. *20*.

Lennon, J.T., and Jones, S.E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat. Rev. Microbiol. *9*, 119–130.

Leonard, L.T. (1943). A Simple Assembly for Use in the Testing of Cultures of Rhizobia. J. Bacteriol. *45*, 523–527.

Lewis, L.A., and McCourt, R.M. (2004). Green algae and the origin of land plants. Am. J. Bot. *91*, 1535–1556.

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S., Ignacio-Espinosa, J.C., Roux, S., Vincent, F., et al. (2015). Determinants of community structure in the global plankton interactome. Science *348*, 1262073.

Lindeman Raymond L. (1942). The Trophic-Dynamic Aspect of Ecology. Ecology *23*, 399–417.

Lindow, S.E., and Brandl, M.T. (2003). Microbiology of the Phyllosphere. *69*, 1875–1883.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. Nature *489*, 220–230.

Lücking, R. (1999). Ecology of Foliicolous Lichens at the 'Botarrama' Trail (Costa Rica), a Neotropical Rainforest. IV. Species Associations, their Salient Features and Their Dependence on Environmental Variables. The Lichenologist *31*, 269–289.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., del Rio, T.G., et al. (2012). Defining the core Arabidopsis thaliana root microbiome. Nature *488*, 86–90.

Lundberg, D.S., Yourstone, S., Mieczkowski, P., Jones, C.D., and Dangl, J.L. (2013). Practical innovations for high-throughput amplicon sequencing. Nat. Methods *10*, 999–1002.

Lutzoni, F., Pagel, M., and Reeb, V. (2001). Major fungal lineages are derived from lichen symbiotic ancestors. Nature *411*, 937–940.

Lynch, M.D.J., and Neufeld, J.D. (2015). Ecology and exploration of the rare biosphere. Nat. Rev. Microbiol. *13*, 217–229.

Mahé, F., Vargas, C. de, Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., Sernaker, S., et al. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat. Ecol. Evol. *1*, 0091.

Maltsev, Y.I. (2013). Ecological Features of Algae Communities in Forest Floor of Pine Plantations of Different Types of Landscapes in Steppe Area of Ukraine. Biolohichnyi Visnyk Melitopol Koho Derzhavnoho Pedahohichnoho Universytetu Im. Bohdana Khmel Nyt Koho Melitopol *3*, 330–339.

Maltsev, Y.I., Pakhomov, A.Y., and Maltseva, I.A. (2017). Specific features of algal communities in forest litter of forest biogeocenoses of the steppe zone. Contemp. Probl. Ecol. *10*, 71–76.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., et al. (2006). Microbial biogeography: putting microorganisms on the map. Nat. Rev. Microbiol. *4*, 102–112.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H.V., Domazet-Lošo, T., Douglas, A.E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl. Acad. Sci. *110*, 3229–3236.

Mercado-Blanco, J., and Bakker, P.A.H.M. (2007). Interactions between plants and beneficial Pseudomonas spp.: exploiting bacterial traits for crop protection. Antonie Van Leeuwenhoek *92*, 367–389.

Metting, F.B. (1996). Biodiversity and application of microalgae. J. Ind. Microbiol. *17*, 477–489.

Miller, M.B., and Bassler, and B.L. (2001). Quorum Sensing in Bacteria. Annu. Rev. Microbiol. *55*, 165–199.

Mod, H.K., Heikkinen, R.K., le Roux, P.C., Wisz, M.S., and Luoto, M. (2016). Impact of biotic interactions on biodiversity varies across a landscape. J. Biogeogr. *43*, 2412–2423.

Moreira, D., and López-García, P. (2002). The molecular ecology of microbial eukaryotes unveils a hidden world. Trends Microbiol. *10*, 31–38.

Mushegian, A.A., Peterson, C.N., Baker, C.C.M., and Pringle, A. (2011). Bacterial Diversity across Individual Lichens. Appl. Environ. Microbiol. *77*, 4249–4252.

Newton, R.J., Huse, S.M., Morrison, H.G., Peake, C.S., Sogin, M.L., and McLellan, S.L. (2013). Shifts in the Microbial Community Composition of Gulf Coast Beaches Following Beach Oiling. PLOS ONE *8*, e74265.

Nowack, E.C.M., Price, D.C., Bhattacharya, D., Singer, A., Melkonian, M., and Grossman, A.R. (2016). Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of Paulinella chromatophora. Proc. Natl. Acad. Sci. *113*, 12214–12219.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., et al. (2018). vegan: Community Ecology Package, URL https://CRAN.R-project.org/package=vegan.

Okubo, T., Matsushita, M., Nakamura, S., Matsuo, J., Nagai, H., and Yamaguchi, H. (2018). Acanthamoeba S13WT relies on its bacterial endosymbiont to backpack human pathogenic bacteria and resist Legionella infection on solid media. Environ. Microbiol. Rep. 1758-2229 *0*.

Ooms, J. (2018). magick: Advanced Graphics and Image-Processing in R., URL https://CRAN.R-project.org/package=magick

Parfrey, L.W., Grant, J., Tekle, Y.I., Lasek-Nesselquist, E., Morrison, H.G., Sogin, M.L., Patterson, D.J., and Katz, L.A. (2010). Broadly Sampled Multigene Analyses Yield a Well-Resolved Eukaryotic Tree of Life. Syst. Biol. *59*, 518–533.

Parfrey, L.W., Walters, W.A., and Knight, R. (2011). Microbial Eukaryotes in the Human Microbiome: Ecology, Evolution, and Future Directions. Front. Microbiol. *2. 59.5.518-533*.

Partida-Martinez, L.P., and Hertweck, C. (2005). Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature *437*, 884–888.

Pedrós-Alió, C. (2011). The Rare Bacterial Biosphere. Annu. Rev. Mar. Sci. *4*, 449–466.

Peñuelas, J., Asensio, D., Tholl, D., Wenke, K., Rosenkranz, M., Piechulla, B., and Schnitzler, J.P. (2014). Biogenic volatile emissions from the soil. Plant Cell Environ. *37*, 1866–1891.

Philippot, L., Raaijmakers, J.M., Lemanceau, P., and Putten, W.H. van der (2013). Going back to the roots: the microbial ecology of the rhizosphere. Nat. Rev. Microbiol. *11*, 789–799.

Picard, K.T., Letcher, P.M., and Powell, M.J. (2013). Evidence for a facultative mutualist nutritional relationship between the green coccoid alga Bracteacoccus sp. (Chlorophyceae) and the zoosporic fungus Rhizidium phycophilum (Chytridiomycota). Fungal Biol. *117*, 319–328.

Pilosof, S., Porter, M.A., Pascual, M., and Kéfi, S. (2017). The multilayer nature of ecological networks. Nat. Ecol. Evol. *1*, 0101–0101.

Ploch, S., Rose, L.E., Bass, D., and Bonkowski, M. (2016). High Diversity Revealed in Leaf-Associated Protists (Rhizaria: Cercozoa) of Brassicaceae. J. Eukaryot. Microbiol. *63*, 635–641.

Poitelon, J.-B., Joyeux, M., Welté, B., Duguet, J.-P., Peplies, J., and DuBow, M. s. (2009). Identification and phylogeny of eukaryotic 18S rDNA phylotypes detected in chlorinated finished drinking water samples from three Parisian surface water treatment plants. Lett. Appl. Microbiol. *49*, 589–595.

Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P., Green, J.L., Green, L.E., Killham, K., Lennon, J.J., et al. (2007). The role of ecological theory in microbial ecology. Nat. Rev. Microbiol. *5*, 384–392.

Ramanan, R., Kang, Z., Kim, B.-H., Cho, D.-H., Jin, L., Oh, H.-M., and Kim, H.-S. (2015a). Phycosphere bacterial diversity in green algae reveals an apparent similarity across habitats. Algal Res. *8*, 140–144.

Ramanan, R., Kim, B.-H., Cho, D.-H., Oh, H.-M., and Kim, H.-S. (2015b). Algae–bacteria interactions: evolution, ecology and emerging applications. Biotechnol. Adv. *34*, 14–29.

Ramanan, R., Kim, B.-H., Cho, D.-H., Oh, H.-M., and Kim, H.-S. (2016). Algae–bacteria interactions: Evolution, ecology and emerging applications. Biotechnol. Adv. *34*, 14–29.

Rambold, G., Friedl, T., and Beck, A. (1998). Photobionts in Lichens: Possible Indicators of Phylogenetic Relationships? The Bryologist *101*, 392–397.

Redford, A.J., Bowers, R.M., Knight, R., Linhart, Y., and 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.

Reveillaud, J., Maignien, L., Eren, A.M., Huber, J.A., Apprill, A., Sogin, M.L., and Vanreusel, A. (2014). Host-specificity among abundant and rare taxa in the sponge microbiome. ISME J. *8*, 1198–1209.

Rogers, R.W. (1988). Succession and Survival Strategies in Lichen Populations on a Palm Trunk. J. Ecol. *76*, 759–776.

Rogers, R.W., Barnes, A., and Conran, J.G. (1994). Lichen Succession Wilkiea Macrophylla Leaves. The Lichenologist *26*, 135–147.

Rosenberg, E., and Zilber-Rosenberg, I. (2016). Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. MBio *7*, e01395-15.

Ryšánek, D., Hrčková, K., and Škaloud, P. (2015). Global ubiquity and local endemism of free-living terrestrial protists: phylogeographic assessment of the streptophyte alga Klebsormidium. Environ. Microbiol. *17*, 689–698.

Saito, R., Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.-L., Lotia, S., Pico, A.R., Bader, G.D., and Ideker, T. (2012). A travel guide to Cytoscape plugins. Nat. Methods *9*, 1069–1076.

Saleem, M., Law, A.D., and Moe, L.A. (2016). Nicotiana Roots Recruit Rare Rhizosphere Taxa as Major Root-Inhabiting Microbes. Microb. Ecol. *71*, 469–472.

Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau, M.L., Capretti, P., Chandelier, A., Cech, T., Chira, D., Diamandis, S., Gaitniekis, T., et al. (2013). Biogeographical patterns and determinants of invasion by forest pathogens in Europe. New Phytol. *197*, 238–250.

Sapp Melanie, Ploch Sebastian, Fiore-Donno Anna M., Bonkowski Michael, and Rose Laura E. (2018). Protists are an integral part of the Arabidopsis thaliana microbiome. Environ. Microbiol. *20*, 30–43.

Scheid, P. (2014). Relevance of free-living amoebae as hosts for phylogenetically diverse microorganisms. Parasitol. Res. *113*, 2407–2414.

Schiaffino, M.R., Lara, E., Fernández, L.D., Balagué, V., Singer, D., Seppey, C.C.W., Massana, R., and Izaguirre, I. (2016). Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. Environ. Microbiol. *18*, 5249–5264.

Schipper, L.A., Degens, B.P., Sparling, G.P., and Duncan, L.C. (2001). Changes in microbial heterotrophic diversity along five plant successional sequences. Soil Biol. Biochem. *33*, 2093–2103.

Schirmer, M., Ijaz, U.Z., D'Amore, R., Hall, N., Sloan, W.T., and Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic Acids Res. *43*, e37–e37.

Schlaeppi, K., Dombrowski, N., Oter, R.G., Ver Loren van Themaat, E., and Schulze-Lefert, P. (2014). Quantitative divergence of the bacterial root microbiota in Arabidopsis thaliana relatives. Proc. Natl. Acad. Sci. U. S. A. *111*, 585–592.

Schnurr, P.J., and Allen, D.G. (2015). Factors affecting algae biofilm growth and lipid production: A review. Renew. Sustain. Energy Rev. *52*, 418–429.

Serusiaux, E. (1989). Foliicolous lichens: ecological and chorological data. Bot. J. Linn. Soc. *100*, 87–96.

Shade, A., Read, J.S., Welkie, D.G., Kratz, T.K., Wu, C.H., and McMahon, K.D. (2011). Resistance, resilience and recovery: aquatic bacterial dynamics after water column disturbance. Environ. Microbiol. *13*, 2752–2767.

Shade, A., Hogan, C.S., Klimowicz, A.K., Linske, M., McManus, P.S., and Handelsman, J. (2012). Culturing captures members of the soil rare biosphere. Environ. Microbiol. *14*, 2247–2252.

Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., and Gilbert, J.A. (2014). Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. MBio *5*, e01371-14.

Simon, M., López-García, P., Deschamps, P., Restoux, G., Bertolino, P., Moreira, D., and Jardillier, L. (2016). Resilience of Freshwater Communities of Small Microbial Eukaryotes Undergoing Severe Drought Events. Front. Microbiol. *7*.

Škaloud, P., and Rindi, F. (2013). Ecological Differentiation of Cryptic Species within an Asexual Protist Morphospecies: A Case Study of Filamentous Green Alga Klebsormidium (Streptophyta). J. Eukaryot. Microbiol. *60*, 350–362.

Smith, H.G., and Wilkinson, D.M. (2007). Not all free-living microorganisms have cosmopolitan distributions – the case of Nebela (Apodera) vas Certes (Protozoa: Amoebozoa: Arcellinida). J. Biogeogr. *34*, 1822–1831.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., and Herndl, G.J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." Proc. Natl. Acad. Sci. *103*, 12115–12120.

Songnian, Z., Xiaohu, Z., Jingwen, W., and Pierre, D. (1993). Utopia: A load sharing facility for large, heterogeneous distributed computer systems. Softw. Pract. Exp. *23*, 1305–1336.

Steidle, A., Sigl, K., Schuhegger, R., Ihring, A., Schmid, M., Gantner, S., Stoffels, M., Riedel, K., Givskov, M., Hartmann, A., et al. (2001). Visualization of N-Acylhomoserine Lactone-Mediated Cell-Cell Communication between Bacteria Colonizing the Tomato Rhizosphere. Appl. Environ. Microbiol. *67*, 5761–5770.

Susanne B. von Bodman, W. Dietz Bauer, and Coplin, D.L. (2003). Quorum Sensing in Plant-Pathogenic Bacteria. Annu. Rev. Phytopathol. *41*, 455–482.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., Ruiz, L.V., Vasco-Palacios, A.M., Thu, P.Q., Suija, A., et al. (2014). Global diversity and geography of soil fungi. Science *346*, 1256688.

Teplitski, M., Chen, H., Rajamani, S., Gao, M., Merighi, M., Sayre, R.T., Robinson, J.B., Rolfe, B.G., and Bauer, W.D. (2016). Chlamydomonas reinhardtii Secretes Compounds That Mimic Bacterial Signals and Interfere with Quorum Sensing Regulation in Bacteria 1. 137–146.

Thrall, P.H., Hochberg, M.E., Burdon, J.J., and Bever, J.D. (2007). Coevolution of symbiotic mutualists and parasites in a community context. Trends Ecol. Evol. *22*, 120–126.

Tormo, R., Recio, D., Silva, I., and Muñoz, A.F. (2001). A quantitative investigation of airborne algae and lichen soredia obtained from pollen traps in south-west Spain. Eur. J. Phycol. *36*, 385–390.

Vaerewijck, M.J.M., Sabbe, K., Baré, J., and Houf, K. (2008). Microscopic and Molecular Studies of the Diversity of Free-Living Protozoa in Meat-Cutting Plants. Appl. Environ. Microbiol. *74*, 5741–5749.

Van Acker, H., Van Dijck, P., and Coenye, T. (2014). Molecular mechanisms of antimicrobial tolerance and resistance in bacterial and fungal biofilms. Trends Microbiol. *22*, 326–333.

Vandenkoornhuyse Philippe, Quaiser Achim, Duhamel Marie, Le Van Amandine, and Dufresne Alexis (2015). The importance of the microbiome of the plant holobiont. New Phytol. *206*, 1196–1206.

Varin, T., Lovejoy, C., Jungblut, A.D., Vincent, W.F., and Corbeil, J. (2012). Metagenomic Analysis of Stress Genes in Microbial Mat Communities from Antarctica and the High Arctic. Appl. Environ. Microbiol. *78*, 549–559.

Velterop, J.S., and Vos, F. (2001). A rapid and inexpensive microplate assay for the enzymatic determination of glucose, fructose, sucrose, L-malate and citrate in tomato (Lycopersicon esculentum) extracts and in orange juice. Phytochem. Anal. *12*, 299–304.

Vestheim, H., and Jarman, S.N. (2008). Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. Front. Zool. *5*, 12.

Vorholt, J. a (2012). Microbial life in the phyllosphere. Nat. Rev. Microbiol. *10*, 828–840.

Weller, D.M., Raaijmakers, J.M., Gardener, B.B.M., and Thomashow, L.S. (2002). Microbial Populations Responsible for Specific Soil Suppressiveness to Plant Pathogens. Annu. Rev. Phytopathol. *40*, 309–348.

Wickham, H., Chang, W., and RStudio (2016). ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics.

Widder, S., Besemer, K., Singer, G.A., Ceola, S., Bertuzzo, E., Quince, C., Sloan, W.T., Rinaldo, A., and Battin, T.J. (2014). Fluvial network organization imprints on microbial co-occurrence networks. Proc. Natl. Acad. Sci. *111*, 12799–12804.

Williams, R.J., Howe, A., and Hofmockel, K.S. (2014). Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. Front. Microbiol. *5*.

Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dormann, C.F., Forchhammer, M.C., Grytnes, J.-A., Guisan, A., et al. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biol. Rev. *88*, 15–30.

Youssef, N.H., Couger, M.B., and Elshahed, M.S. (2010). Fine-Scale Bacterial Beta Diversity within a Complex Ecosystem (Zodletone Spring, OK, USA): The Role of the Rare Biosphere. PLOS ONE *5*, e12414.

Zamioudis, C., Korteland, J., Pelt, J.A.V., Hamersveld, M. van, Dombrowski, N., Bai, Y., Hanson, J., Verk, M.C.V., Ling, H.-Q., Schulze-Lefert, P., et al. (2015). Rhizobacterial volatiles and photosynthesis-related signals coordinate MYB72 expression in Arabidopsis roots during onset of induced systemic resistance and iron-deficiency responses. Plant J. *84*, 309–322.

6 Acknowledgements:

My PhD was a long journey, an experience that without doubt changed in better my personal and professional life. There is a number of persons which I have to acknowledge for this. I like to start with the professional side.

First of, I want to acknowledge my supervisor, Eric Kemen. He gave me the possibility of dealing with a great and fascinating challenge: design experimental setups and developing data analysis for a substantially unknown field in plant biology. His balance between pushing for autonomy and supporting attitude on one hand made me learn enormous amounts of skills hands-on and on the other hand developed my mindset in a more scientific direction. Thanks also to CEPLAS for funding me in all these years

I have to thank also my collaborators in Germany and abroad: Detlef Weigel, Fabrice Roux from INRA (France), Carlos Alonso-Blanco from CNB (Spain), and Jon Ågren from Uppsala University for being of crucial help and guidance during the geographical sampling and of course Stephane Hacquard for the fruitful discussions. I thank them also for the productive exchange and collaboration also after the sampling and as well for their patience with shipping companies. I thank also my TAC committee members, Stijn, and Michael Bonkowski for the nice directions and discussion on my project. I thank then Gunther Döhlemann, Achim Tresch, Korbinian Schneeberger to be part of my thesis committee, along with Eric.

Thanks also to Sam Kroll for taking care of the common garden experiment in almost all the field and lab aspects.

These experiments generated a lot of data, and there was no chance I could reliably analyze those without the help of several people:

Matthew Agler for his invaluable helps and hints which made me learn from scratch programming languages like perl and amplicon sequencing pipelines. In his shoes probably I would not have the same patience.

Juliana Almario: her almost infinite knowledge of microbial ecology and statistics made me learn those concepts that for years were not entering my brain.

Karoline Faust and Gipsi Lima Mendez from Leuven University, Katharine Coyte from MSKCC (New York) and Leo Lahti from Turku University for the productive discussions and the hints concerning network analysis and its interpretation.

Thomas Nakano for his nice hints in programming, especially its crucial role in my struggle with R, Stackoverflow is good, but Tom is better.

The main experiment that I displayed involved a lot of travels and what is a travel without a travel companion? I have to thank Paloma Durán for the very precise and detailed organisation of the trip as well as her help and hints in designing my own experiment in the ecological frame. A very big thank to the students that I had the pleasure to supervise, especially to Laurenz Gewiss, one of the most reliable students one can ever imagine.

Finally, I have to thank all (ex) the members of Kemen Lab: Ariane, Sam, Jonas, Aleksandra, Ronny, Martin, Paul for the great support and feedback they gave me within and outside the labmeeting.

A last thank to the whole plant microbe interaction department in MPIPZ for the feedback given over these year.

Now it is time for the personal acknowledgements. During these years sometimes the path was easy, sometimes not, and here I want to thank those who were there in both cases:

First of course my super girlfriend Aleksandra, who made my life better and better on a daily basis, and succeeded even in making me smile while writing the thesis. She must be the one.

My lovely friends Francesco and Giulia: like Kashkaval or Burrancino, they were always flavoring my life in Germany and before Germany.Vangelis and Roussana for the amazing support and 3h long coffee together. Chidi, definitely my running model, thank you for the Marathon. Sayan, because through him I learnt that maybe other opinions may be acceptable as well. Thanks to Lukas, literally the first person I met in Germany, it was the best welcome I could imagine.

A big thank to Paloma, with whom everything started and everything finished. Truly the best travel mate ever had. Then Maria, Tommy, Mathias, and all the Tsuda people for making lab life much more reggaeton, and then all the Hispanic community of MPI. Rafa and his awesome parties and his laughs on words like Profumiera, Miguel, Virginia and Fernando, Eva, Nacho, Annahit.

A big thank to everyone that I met during my journey, which made me smile even once. Hope I managed to do the same.

A frank thank you also to Roberto Saviano, Bonobo, John Oliver, and Kendrick Lamar and for the continuous inspiration that they provide me.

In finale un ringraziamento speciale alla mia famiglia, che pur sparsa in 4 posti diversi non mi ha mai fatto mancare il proprio supporto e la propria vicinanza. Ed infine ai miei incancellabili veterani: Nicola, Alessandro, Saulle, Claudio ed Eman: se sono un po' più simpatico rispetto a qualche anno fa, é soprattutto merito loro.
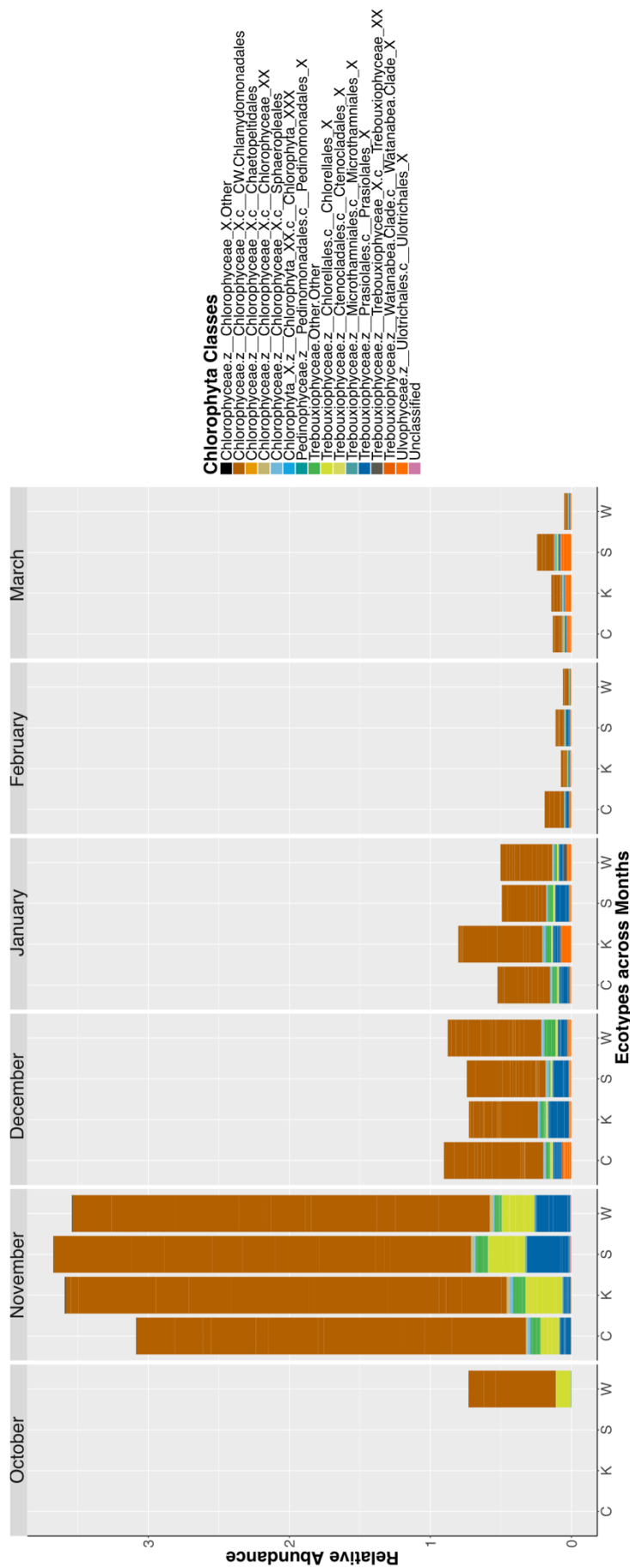
7 Supplementary figures
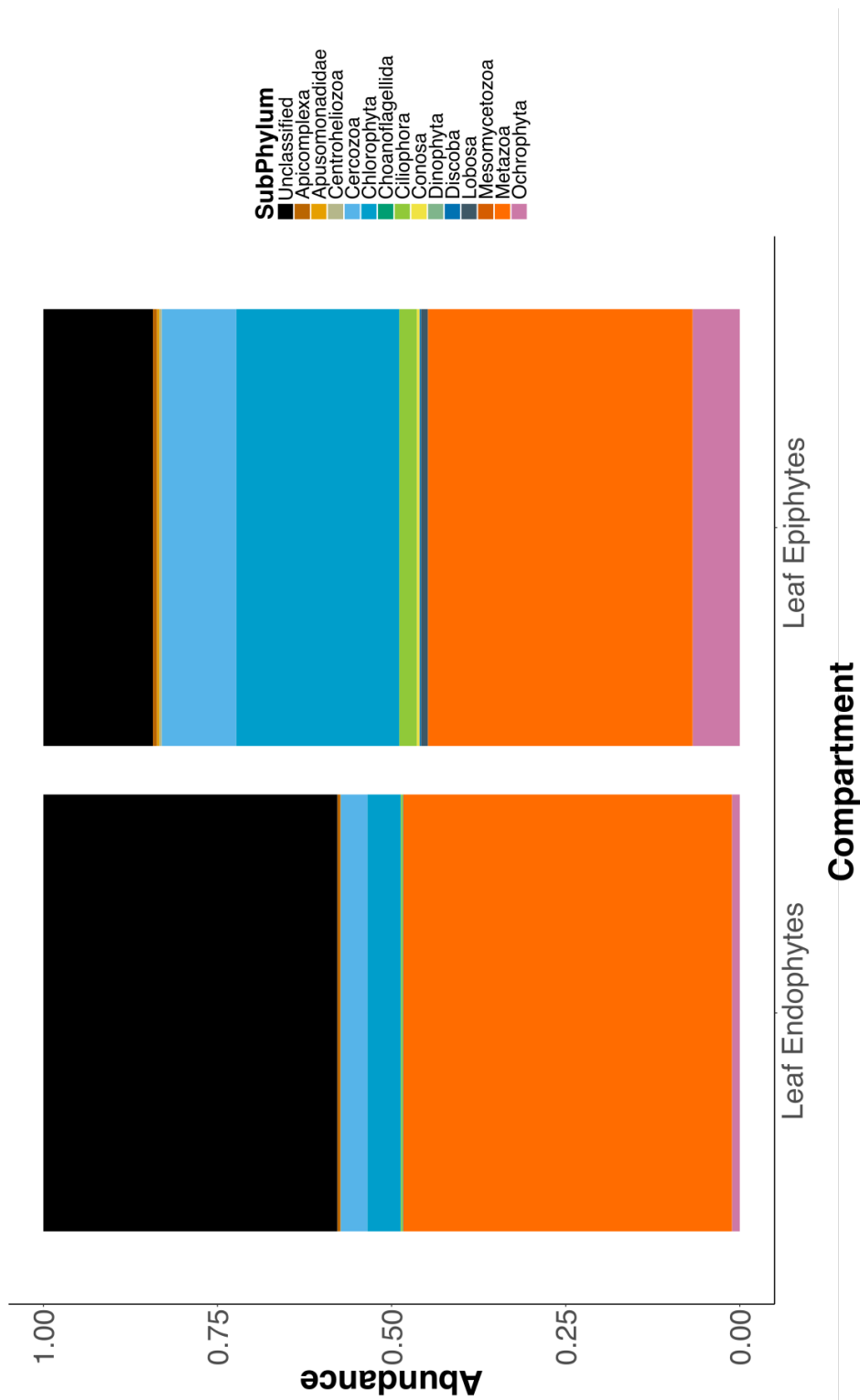
Supplementary Figure 1: Beside unclassified OTUs, Metazoa are the most abundant Overlooked Microbial eukaryotes group across compartments, Chlorophyta follows but only in epiphytes. Relative abundance calculated overall the ample over the three years
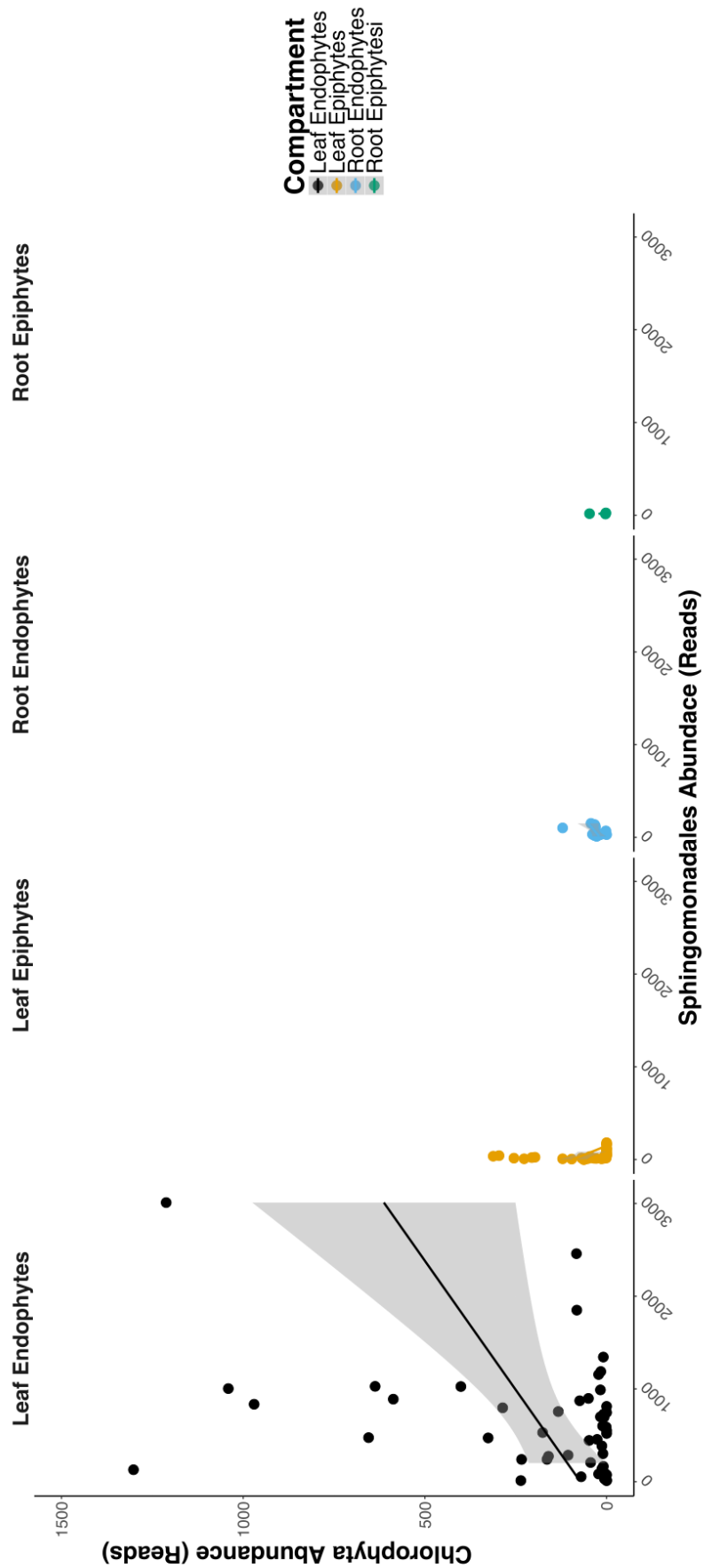
Supplementary Figure 2 Chlorophyta abundance correlates with Sphingomonadales abundance, only in the leaf endophytic compartment, linear regression on absolute reads abundance Absolute abundance calculated overall the ample over the three years

Supplementary Figure 3 Chlorophyta relative abundance in the common garden experiment reveals a peak in November, and underpins the marginal impact of A.thaliana ecotype. Relative abundance calculated overall the ample over the three years
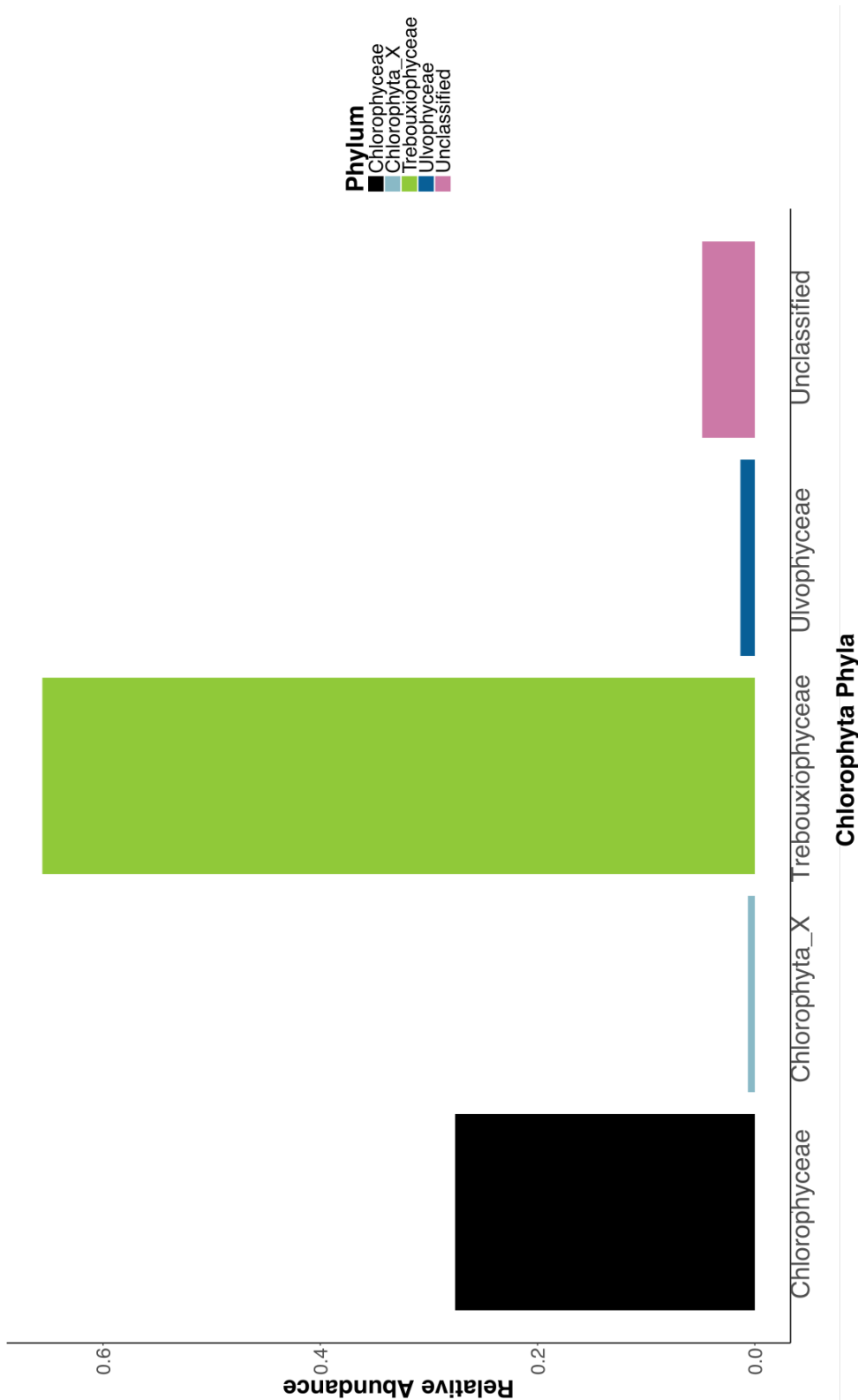
*Supplementary Figure 4 Trebouxiophyceae are the most abundant phylum across sites in geographical sampling, Relative abundance calculated overall the ample over the three years*

Overlooked Microbial Eukaryotes Phyla

Annelida
Apicomplexa_X
Apusomonadidae_Group.1
Arthropoda
Ascomycota
Bacillariophyta
Basidiomycota
Chlorophyceae
Chlorophyta_X
Chrysophyceae.Synurophyceae
Chytridiomycota
Colpodea
Craniata
Cryptomycota
Discosea.Flabellinia
Embryophyceae

Endomyxa
Endomyxa.Phytomyxea
Entomophthoromycota
Euglenozoa
Filosa.Granofilosea
Filosa.Imbricatea
Filosa.Sarcomonadea
Filosa.Thecofilosea
Heterolobosea
Ichthyosporea
Kickxellomycota
Klebsormidiophyceae
Labyrinthulea
Litostomatea
Microsporidiomycota
Mollusca

Mucoromycota
Nematoda
Oligohymenophorea
Oomycota
Pedinophyceae
Phyllopharyngea
Rotifera
Spirotrichea
Tardigrada
Trebouxiophyceae
Tubulinea
Ulvophyceae
Unclassified
Xanthophyceae
Zoopagomycota
Zygnemophyceae

Supplementary Figure 5 Overlooked microbial eukaryotes relative abundance over time and ecotype does not reveal major difference in term of phyla composition. Relative abundance calculated overall the ample over the three years
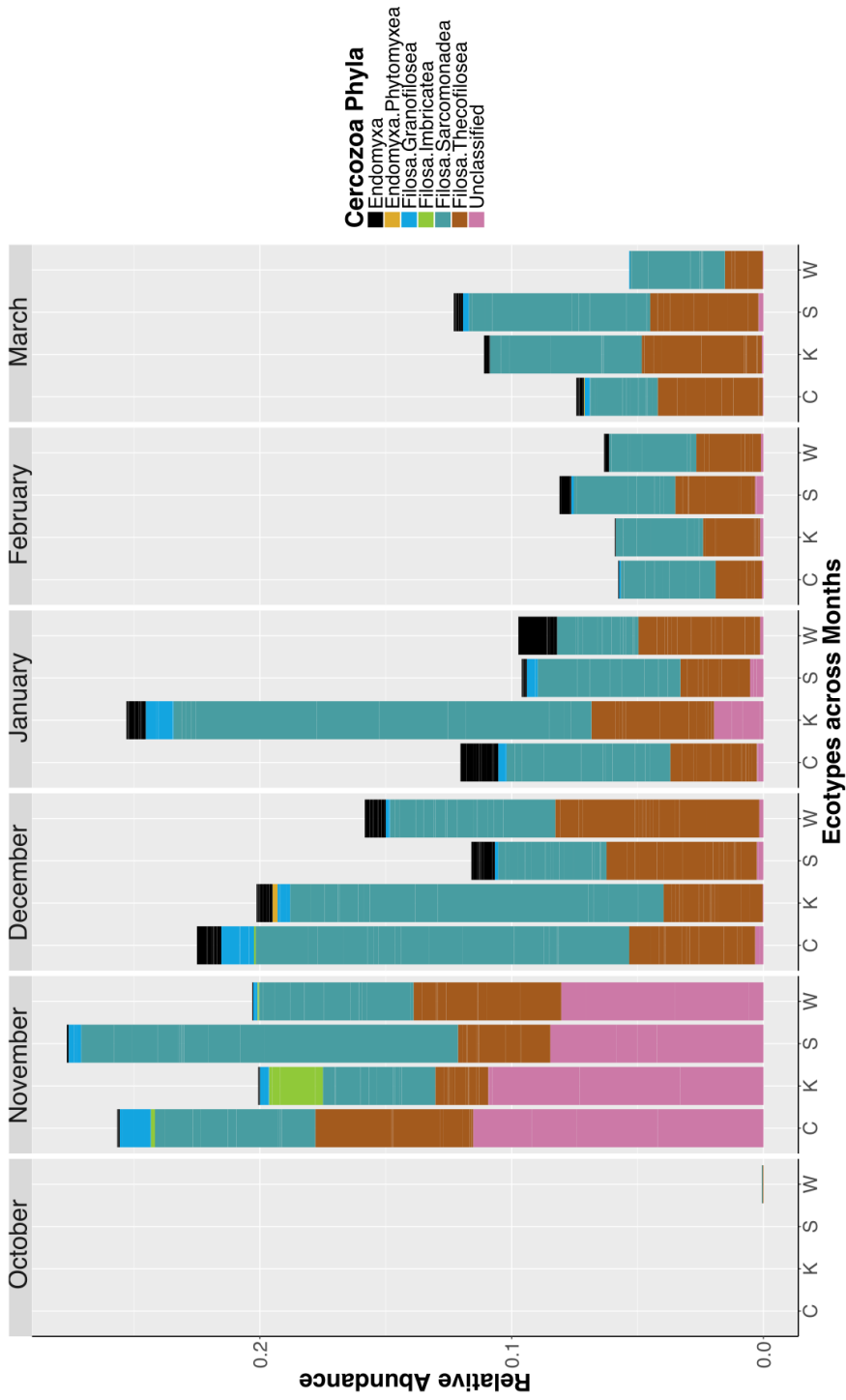
Supplementary Figure 6 Cercozoa show a maximum peak of abundance and diversity in November, irrespective of the ecotype. Relative abundance calculated over all sample over the three years

8 Curriculum Vitae

Alfredo Mari

PERSONAL DETAILS

Full Name: Alfredo Mari
Tel.: (+49) 0221 5062 347
Email: mari@mpipz.mpg.de
LinkedIn: www.linkedin.com/in/alfredomari/

SKILLS AND COMPETENCES

Rapid adaptability to new projects and methods
Leadership and human management skills
Systematic and detailed working methodology
Competence in analyzing large data sets, familiarity with current statistical methods
Languages: fluent in English and Italian, advanced in German, basic in French
IT: fluent in R, Perl, Unix, Office, Cytoscape, Adobe Illustrator, basic user: Python, Adobe Photoshop.

PROFESSIONAL CAREER

PhD Fellow at Max Planck Institute for Plant Breeding Research (Köln, Germany)

11/2014 – present

Aim to discover and reconstruct the role of microbial eukaryotes within plant leaf microbiome. A project in collaboration with university of Madrid (ES), Uppsala (SWE), Toulouse(FR), Tübingen (DE). Thesis focus: Reconstruction of microbial networks and study on the impact of stochastic variables.

ACCOMPLISHMENTS:

Coordination and management of international academic collaboration
Environmental study design and organization
Supervision of bachelor students and student helpers, both theoretical and practical
Development and implementation of new methods for Next Generation Sequencing data analysis, including application of graph theory and deployment of respective data visualization tools
Development of data processing pipelines with a reproducibility-oriented and human-error-minimizing focus

Leader of Visions in Science 2016 Organising Committee *01/2016-09/2016*:

Organized a multidisciplinary scientific conference in Berlin sponsored by Max Planck Society
ACCOMPLISHMENTS:

Result-oriented coordination of a team of eight people including scheduling, work assignment, conflict management.

Coordination of 2nd Max Planck Career Fair, including sponsorship seeking, budget handling and efficient networking.

Master Intern at University of Turin (Turin, Italy) *12/2013–10/2014*

Aimed to disentangle the tradeoff between pathogenesis and symbiosis between plants and fungi. Thesis title: "Characterization of LORE1 *Lotus japonicus* transposonic mutants in mycorrhizal phenotype, an insight into the early stage of arbuscular mycorrhizal symbiosis".

ACCOMPLISHMENTS:

Basic handling of medium size datasets (microarray data), including basic statistical analysis, mainly univariate.

Supervision of two student helpers in practical basic laboratory analysis

Intern at Cornell University – Boyce Thompson Institute (Ithaca, NY, USA) *09/2012–01/2013*

Aimed to gain experience in new fields of plant biology such as epigenetics.

ACCOMPLISHMENTS:

Development of efficient workflows and systematic processing of large sample sets, from collection until lab analysis.

Bachelor Intern at University of Turin (Turin, Italy) *01/2012–04/2012*

Aimed to clarify the cellular localization of water-channel proteins (aquaporins) within plant-fungus symbiosis, thesis title: "Cellular Localization and characterization of LjNIP, overexpressed aquaporin in *G.margarita* mycorrhized roots"

ACCOMPLISHMENTS:

Learning and deployment of multiple lab techniques in a result-oriented project, including laser microscopy and laser micro-dissection

ACHIEVEMENTS AND CERTIFICATES

Ko Shimamoto Travel Award, participation in Molecular Plant Microbe Interaction conference 2016, (Portland,OR-USA)

2nd prize best science slam, Center of Advanced European Studies and Research, Bonn 2015

Scuola Superiore Sant'Anna (Pisa, Italy) Student Scholarship – 2009/2014

PUBLICATIONS

"Early Lotus japonicus root transcriptomics responses to symbiotic and pathogenic exudates"
M. Giovannetti, <u>A. Mari</u>, M. Novero & Paola Bonfante *Frontiers in plant science* June 2015

Agler, M.T., Mari, A., **Almario J., Dombrovski, N., Hacquard, S. & Kemen, E.M. (2018)** **Accurate systems biology-based analyses in multi-kingdom microbiome studies by overhauling amplicon sequencing and data analysis; Methods in Ecology and Evolution.** *In revision*

EDUCATION

Master degree in Molecular and Industrial Biotechnology, University of Pisa, Pisa, Italy (22/10/14)

Bachelor degree in Agro-Industrial Biotechnology, University of Pisa, Pisa, Italy (12/07/2014)

Bachelor degree in Agricultural sciences and Biotechnology, Scuola Superiore Sant'Anna, Italy (25/10/2013)

Abitur, "Liceo Classico Statale Lorenzo Costa", La Spezia, Italy (14/07/2009)

9 Affidavit / Eidesstaatliche Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit − einschließlich Tabellen, Karten und Abbildungen − , die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie − abgesehen von unten angegebenen Teilpublikationen − noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Gunther Döhlemann betreut worden.

Ich versichere, dass ich alle Angaben wahrheitsgemäß nach bestem Wissen und Gewissen gemacht habe und verpflichte mich, jedmögliche, die obigen Angaben betreffenden Veränderungen, dem Dekanat unverzüglich mitzuteilen

Köln, 28.05.2018

Alfredo Mari