ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE BIOTECNOLOGICHE, BIOCOMPUTAZIONALI, FARMACEUTICHE E FARMACOLOGICHE

Ciclo 37

Settore Concorsuale: 03/D1 - CHIMICA E TECNOLOGIE FARMACEUTICHE, TOSSICOLOGICHE E NUTRACEUTICO-ALIMENTARI

Settore Scientifico Disciplinare: CHIM/11 - CHIMICA E BIOTECNOLOGIA DELLE FERMENTAZIONI

MICROBIOME-BASED APPLICATIONS FOR SUSTAINABLE FOOD PRODUCTION AND WASTE-STREAMS VALORIZATION

Presentata da: Enrico Nanetti

Supervisore: Marco Candela

Coordinatore Dottorato

Maria Laura Bolognesi

Co-supervisore

Nicola Secchi

Esame finale anno 2025

Borsa di dottorato del Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005), risorse FSE REACT-EU, Azione IV.4 “Dottorati e contratti di ricerca su tematiche dell’innovazione” e Azione IV.5 “Dottorati su tematiche Green."

Codice CUP: J35F21003320006

"But you know, happiness can be found even in the darkest of times, if one only remembers to turn on the light."

Albus Dumbledore

TABLE OF CONTENTS

ABSTRACT

The transition to circular, green economy calls for new ways to minimize the environmental impact of production processes while implementing new bio-based pathways to circularize the disposal of recalcitrant waste-streams. Microbiomes, defined as ultracomplex microbial communities – with their genes and genomes – in a given ecosystem, represent an untapped source of functionalities to be used as eco-friendly biofertilizers/pesticides, as well as for the biotransformation of biomasses into platform chemicals (e.g., Volatile Fatty Acids, VFAs), with the concomitant production of several bioactive compounds (e.g., antibacterials, antivirals, antifungals, anticancers, flavoring agents, sweeteners, additives, etc.). In this view, my PhD thesis will be focused on natural microbiomes in relation to two main research topics.

First, we used Next Generation Sequencing (16S/ITS marker gene sequencing and shotgun metagenomics) to characterize the microbiome of the grapevine (*Vitis vinifera* L.) roots – and surrounding soil – across two worldwide famous Italian viticultural sites, i.e., the "Consorzio Tutela Lambrusco DOC" protected designation of origin area (PDO, Emilia Romagna, Italy) and the "Consorzio del Vino Nobile di Montepulciano DOCG" protected designation of origin area (PDO, Tuscany, Italy). In both cases, we focused on microbial communities at the soil-root interface as a determinant of the wine terroir, with potential effects on wine final organoleptic properties, and with particular emphasis on microbial traits responsible for plant biofertilization/stimulation, named Plant Growth-Promoting (PGP) functions. Indeed, decades of research have demonstrated the role of microbial communities in providing a plethora of beneficial functions for plant nutrition, growth, and stress tolerance, making them a strategic player for the transition to a more sustainable agriculture and viticulture. PGP functions include drought resistance, biological nitrogen fixation, phosphorus solubilization, exudation of bacterial siderophores, production of antimicrobials, phytohormones, and competition with pathogens and pests, among others. Our results provide glimpses of the presence of multiple PGP microorganisms in the investigated wine-producing sites, with some degree of regional specificities at the local scale, highlighting the role of PGP microbes for grapevine growth and paving the way to design new microbiome-based inoculants for wine production, aiming to increase product quality and sustainability.

The second topic analyzed in this thesis is related to the biotransformation of lignocellulose (LC) waste-streams into platform chemicals and other useful molecules by natural microbiomes. Lignocellulose is the most abundant polymer on Earth and is composed of carbohydrates (mainly cellulose, hemicellulose, and pectin) and aromatic compounds (lignin), holding the potential to be deconstructed for several biotechnological applications, but recalcitrant to hydrolysis. In this context, we explored the gut microbiome (GM) of the wild Alpine ibex (*Capra ibex* L.), a poorly studied species with a strict plant-based diet, and thus a good target for the biodiscovery of microbial species, hubs, and pathways involved in the biotransformation of LC biomass. Fecal samples were collected at Stelvio National Park from the Alpine ibex in spring, summer, and fall 2020, and the GM was investigated by means of multi-omic techniques from a compositional and functional perspective. By obtaining species-level genome bins, we provided glimpses of *C. ibex* gut microbiome bacterial strains as a possible microbiome-based solution for the bioconversion of lignocellulose to high-value compounds, such as volatile fatty acids and alcohols. Besides, a preliminary investigation of the Alpine ibex GM biosynthetic gene clusters (BGCs) emphasizes how wild ruminants may be regarded as a rich reservoir of genetic functionalities of industrial and pharmaceutical interest.

Chapter 1 – Soil and plant-associated microbiomes: a focus on Italian grapevines

1.1 General introduction

1.1.1 The plant holobiont

Plants, as all macro-organisms, live in close association with complex and dynamic microbial consortia, and thus can be regarded as '*holobionts*'. The holobiont is the entity encompassing the host (e.g., animals, plants, and fungi), as well as the whole set of microbial cells intimately interacting with it[1,2,3]. Microbial communities associated with multicellular hosts make up their *microbiota* (or *microbiome*, if one considers also the entire 'theatre of activity' of the holobiont, meaning its chemo-physical properties, as well as the spectra of molecules, metabolites, and toxins produced and released by the microbiota) and comprise a plethora of prokaryotic and eukaryotic taxa (e.g., Bacteria, Archaea, Fungi, Algae, Viruses, etc.) with a pivotal role in the regulation of the host's physiology, ranging from growth and nutrition up to immune modulation, homeostasis, and stress tolerance[4]. This makes holobionts, and microbiomes in particular, an increasingly studied topic, with the aim to unravel the many biological processes which are still poorly – or not fully – understood, from both the ecological and the biotechnological perspective[5,6].

The plant holobiont hosts variegated and multifaceted microbiotas in all of its organs (i.e., roots, stem, leaves, flowers, fruits, and seeds), with some microbial taxa shared amongst the different compartments and others showing a certain degree of specificity in relation to the organ chemo-physical/biological features[7,8,9,10]. As a matter of fact, each plant compartment can be regarded as a different ecological niche, encompassing both core and unique microbes acquired throughout the plant life, mainly derived from the surrounding environment, and, to a minor extent, vertically transmitted through the seeds following sexual reproduction[9,10,11]. Specifically, microorganisms have the ability to colonize both the internal plant niches (endophytes) as well as the external surfaces (epiphytes) of the entire plant body[12,13]. Plant microbial colonization generally begins from the bulk soil, which is an extremely diverse and rich microbial ecosystem functioning as a microbial reservoir for plants[7,8,14]. Indeed, recent observations highlight how microbial communities colonizing above-ground aerial portions of the plant, e.g., leaves and flowers, share more taxa with the surrounding soil rather than with each other in grapevines from different continents[8,15]. Furthermore, several studies targeting different plants show a decreasing trend of bacterial species richness and biodiversity from the bulk soil up to the aerial plant organs[7,16,17]. Plant microbial colonization, of both endophytes and epiphytes, emerges as a result of the biochemical crosstalk established at the soil-plant interface by means of plant root exudation of small organic molecules synthesized by the host[18]. Microorganisms are able to colonize all plant niches, starting from the roots up to the aerial portions, through several entry points, including inter-cellular junctions and ruptures in the main or lateral root epidermis, as well as wounds in the rhizoplane (i.e., the root surface including its associated soil particles). Following entry, microbes which are able to 'elude' the plant immune response, including beneficial taxa which have finely co-evolved with their host, may migrate through the root cortex until they reach the vascular cylinder, and finally to the other organs following the vascular route (xylem and phloem)[19,20,21]. As for leaves, flowers, and fruits, besides the soil-root-stele axis, additional minor ways of entry have been described. These include stomata (i.e., microscopic openings in plant tissues, generally more numerous on the underside of leaves and across thin

stems, that allow gas exchange with the atmosphere), trichomes (i.e., tiny glandular or non-glandular outgrowths of the plant epidermis), superficial wounds caused, for instance, by herbivores, and secreting hydathodes[22,23]. Further, environmental microorganisms may reach plants by means of wind, rain, watercourses, and animals (including humans)[24,25,26,27]. For example, terrestrial arthropods, including Hymenoptera, Lepidoptera, Coleoptera, etc., have been proved to have a direct influence on the plant holobiont, being capable of injecting microbes into plants via complex buccal structures and being actively involved in pollination processes[28,29]. In this direction, a previous work by McFrederick et al. (2017) shed light on trophic interactions occurring between insects of the *Megachilidae* family and the flowers of several plant species, pointing out the role of these insects in shaping the plant microbiota[30].

Once colonization has successfully occurred, microorganisms will start building up a complex network of symbiotic relationships with the host, contributing to the biological homeostasis of the holobiont. For the host, these relationships range from neutral (commensalism) to beneficial (mutualism) and, in some cases, may also be detrimental (parasitism)[4]. In the last decades, breakthrough technological advances, especially in the fields of molecular biology and bioinformatics, enabled researchers to look more deeply into plant-microbiota biochemical pathways and ecological interactions, paving the way for the onset of a plethora of studies targeting holobionts, from both the ecological and biotechnological point of view. Next Generation Sequencing (NGS) techniques, for instance, despite their limitations related to the intrinsic complexity of microbiome research, have revolutionized our understanding of the plant holobiont over the past decade, by means of lab assays, genome analysis of individual strains or consortia, as well as meta-omics, including metabarcoding, metagenomics, metatranscriptomics, and metabolomics, that can be carried out alongside 'classical' microbiology protocols (cultivation, isolation, and characterization of microbes) in order to shed light on holobiont biology[31].

Given the role of microorganisms for the plant fitness and metabolic plasticity, as a consequence of the intimate metabolic networks established at the holobiont level, beneficial microorganisms colonizing plant tissues have gained increasing attention in the scientific scenario[18,32]. Specifically, mutualistic plant-associated taxa can be grouped into two main macro-categories, collectively referred to as *Plant Growth-Promoting Microorganisms* (PGPM), which are *Plant Growth-Promoting Bacteria* (PGPB) and *Arbuscular Mycorrhizal Fungi* (AMF)[33,34]. These beneficial microbes are able to perform a myriad of metabolic processes actively supporting plant health, including: solubilization, mobilization, and uptake of both macro- and micro-nutrients (bio-fertilization), production of phytohormones involved in plant growth and stress tolerance (bio-stimulation), niche competition against microbial pathogens, production of small antimicrobials (antibiotics, bacteriocins), and production of bioactive lytic enzymes (bio-protection)[18,35] (**Fig. 1.1**).

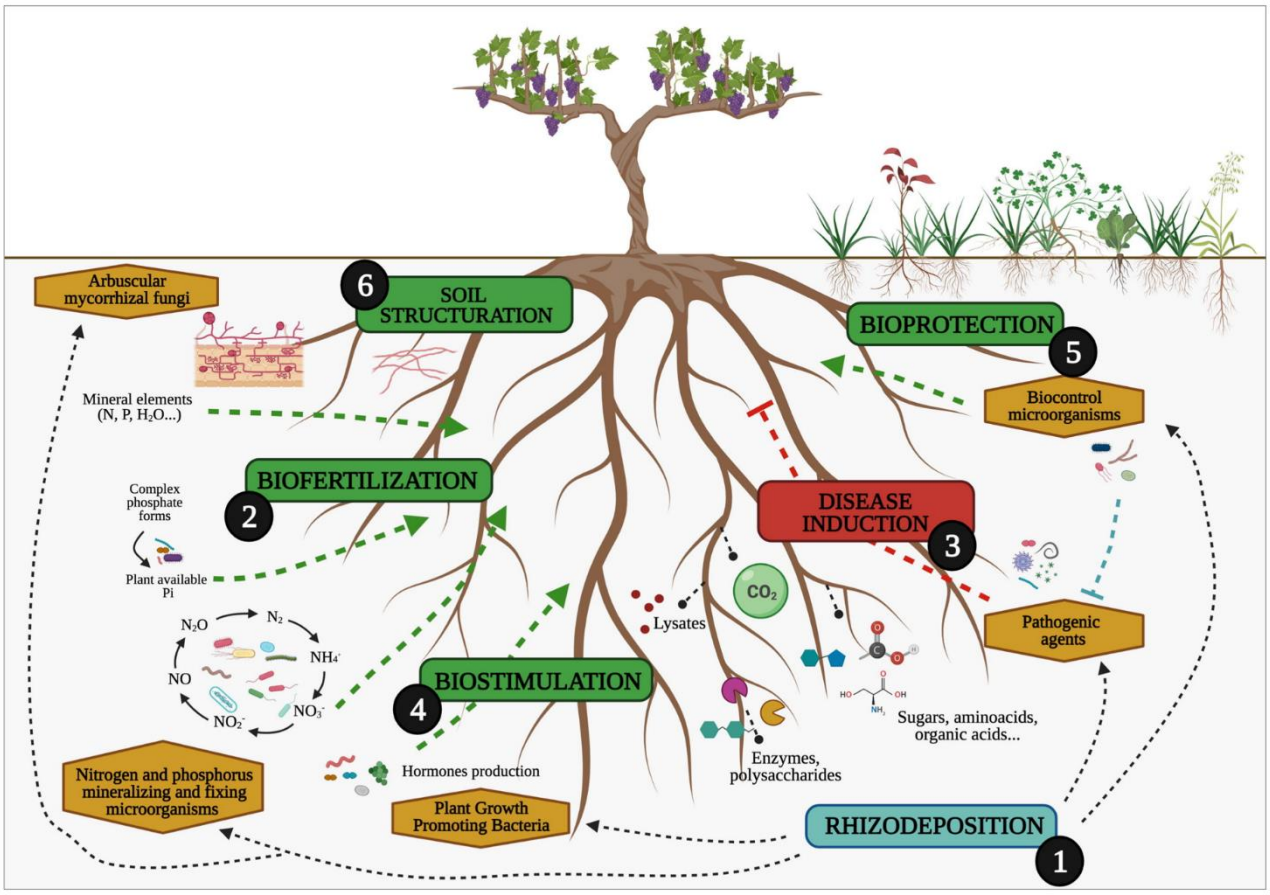


Fig. 1.1 – Schematic illustration of the complex biochemical crosstalk occurring at the soil-plant interface, which represents the main colonization route of plants by the microbiota. 1) Plant root exudation (rhizodeposition) modulates and attracts microorganisms, including PGPB and AMF. 2) Bacteria and fungi involved in nutrient cycling (e.g., nitrogen, phosphorous, and minerals) possess the genetic repertoire needed for plant bio-fertilization. 3) Pathogens able to elude plant immune systems, including bacteria, fungi, viruses, protozoans, and animals, may replicate and induce infection in the plant. 4) Bacteria and fungi equipped for phytohormone biosynthesis or induction are actively involved in plant bio-stimulation pathways. 5) Ecological niche competition and production of assorted antimicrobial metabolites/enzymes are amongst the main mechanisms of pathogen suppression (bio-protection) by the plant microbiota. 6) Soil structure and edaphic factors, such as pH, temperature, nutrients, and texture, exert a strong influence on soil and plant microbial communities, which, in turn, may modify soil properties themselves (modified from[18]).

Nitrogen (N), phosphorus (P), and iron (Fe) are some of the most important mineral nutrients for the growth and development of plants, and are elements that, by their nature, occur in soil mainly in forms that cannot be directly assimilated by plants, thus microorganisms are the main source of supply and mobilization of these nutrients for the plant holobiont[36]. Diazotrophic nitrogen-fixing bacteria (e.g., those belonging to the genera *Rhizobium*, *Bradyrhizobium*, *Frankia*, *Nostoc*, and *Azotobacter*) are naturally capable of converting atmospheric nitrogen (N_2, which is chemically inert due to the triple covalent bond) into bio-available forms such as ammonium (NH_4^+) and nitrate (NO_3^-)[37,38], which facilitates nitrogen uptake and incorporation into cell macromolecules and molecular

machineries (e.g., proteins, nucleic acids, amino sugars, and so on). Indeed, studies have confirmed that diazotrophic microbes may provide their plant host with up to 90% of the bio-available nitrogen needed for growth[39]. Mineralization and solubilization of P from organic decaying matter and from soil particles, respectively, is carried out by several PGPM, including *Pseudomonas*, *Enterobacter*, *Arthrobacter*, *Burkholderia*, and *Bacillus*, as well as AMF, i.e., widespread members of the plant microbiota secreting a wide range of organic acids (e.g., lactic acid, citric acid, and malic acid), as well as lytic enzymes of the phosphatase family, with a pivotal role in enriching soil with orthophosphate, mainly in the forms of HPO_4^{2-} and $H_2PO_4^{1-}$, which can be directly absorbed and exploited by plants as part of nucleic acids and phospholipids[37,40,41]. As for iron, despite its high concentrations in the soil, the Fe^{3+} form is prevalent, and it is poorly soluble and therefore difficult to assimilate by plants. Consequently, the ferric ion must be chelated by a siderophore, and, once the Fe-siderophore complex has entered the root cells, it is usually reduced to Fe^{2+} and used at the cellular level. Siderophores are low molecular weight Fe chelating agents with an extremely heterogeneous chemical structure, which can be produced both by the microbiota (mainly by PGPB and AMF) and by the plant itself (in this case, they are referred to as phyto-siderophores). The presence of these molecules helps the plant reach optimal levels of iron in its organs, which is a key component of some proteins (e.g., cytochromes and iron-sulfur proteins), especially when the plant suffers from iron deficiency[37,42]. Besides, previous research has proved that bacterial siderophores may hamper phyto-pathogen spread by reducing bio-available iron in the environment, as an example of trophic competition for the ecological niche[42,43]. Several soil microorganisms are also able to produce and release a vast array of molecules that fall within the definition of phytohormones (e.g., auxins, abscisic acid, cytokinins, ethylene, gibberellins, etc.). These molecules regulate various plant physiological processes involved in growth, development, and resistance to multiple biotic and abiotic stresses[41]. Among those, auxins are undoubtedly the best-characterized phytohormones, especially the indole-3-acetic acid (IAA), involved in numerous metabolic processes, including cell division and differentiation, root growth, and, together with abscisic acid (ABA), resistance to drought and salt stress[37,41]. Examples of PGPM directly involved in plant bio-stimulation by synthesizing phytohormones are *Pseudomonas*, *Bacillus*, and *Azospirillum*[37,41,44], and genetic engineering of these and other taxa represents a promising avenue towards more sustainable eco-friendly agricultural approaches. For instance, a recent proof-of-concept work by Pham and colleagues demonstrated that a metabolically-engineered strain of the obligate methanotrophic bacterium *Methylotuvimicrobium alcaliphilum 20Z* is able to increase germination, as well as shoot and root elongation, in wheat (*Triticum aestivum* L.) thanks to the overproduction of L-tryptophan and the phytohormone derived from it (i.e., IAA). In fact, wheat seeds treated with the engineered strain showed an increase in the above-mentioned parameters of over 100% compared to the untreated control group[45]. Moreover, microbes of the genera *Pseudomonas*, *Bacillus*, and *Azospirillum*, among others, are also able to regulate the levels of ethylene (a gaseous phytohormone that, when excessively produced in stress conditions, has adverse effects on the plant) in the plant holobiont through the biosynthesis of the enzyme 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase, which converts 1-aminocyclopropane-1-carboxylic acid (ACC), a precursor of ethylene, into alpha-ketobutyric acid and ammonia, thus decreasing the ethylene peak that occurs in stress conditions, such as metal-derived toxicity, hyper salinity, drought, attack by pathogens and/or parasites, and extreme temperatures[46,47]. Additional bioactive phytohormones produced by many players of the plant microbiota are cytokinins and

gibberellins, which, together with other hormones and secondary metabolites, take part in the ultra-complex signaling network at the base of plant growth and development[48]. PGPM can also limit or prevent damage caused by phytopathogenic soil taxa through various mechanisms, the main one being the production of antimicrobial compounds. As a matter of fact, several bacteria of the genera *Pseudomonas*, *Rhizobium*, *Streptomyces*, *Bacillus*, and *Stenotrophomonas*, among others, are able to produce a wide range of metabolites with antibacterial, antifungal, and antiviral activities, such as 2,4-diacetylphloroglucinol (DAPG) and hydrocyanic acid[49,50]. Besides, bacteriocins are potent and structurally heterogenous ribosomally-synthesized antimicrobial peptides produced by bacteria and potentially involved in plant defense against pathogens by means of several, and often unknown, mechanisms, including membrane integrity disruption[51]. The compound named thuricin-17, for instance, which is produced and excreted in the extracellular medium by *Bacillus thuringiensis*, has been shown to promote the growth of canola plants (*Brassica napus* L.) at high temperatures, displaying a potential role for bacteriocins in supporting plant growth under climate change conditions[52]. Another study by Kim at al., (2019) showed that *Streptomyces* derived lanthipeptides were able to hinder infection by saprophytic *Fusarium oxysporum* in strawberries[23]. PGPM suppress the activity of phytopathogens also by producing hydrolytic enzymes like chitinases, glucanases, proteases, and lipases, which weaken and break down components of the plasma membrane and cell wall, such as cellulose, hemicellulose, chitin, peptidoglycan, lipids, and membrane proteins[37,53]. For example, β-1,3-glucanases synthesized by species of *Streptomyces*, *Paenibacillus*, and *Bacillus*, among others, are able to deconstruct the coating of many phytopathogenic fungi found in soil (e.g., *Fusarium* spp.)[37]. Finally, other relevant compounds useful for the plant holobiont and produced by the microbiota are the so-called Volatile Organic Compounds (VOCs). Generally, PGPM give off VOCs ranked as alkenes, esters, alcohols, ketones, and terpenes, which act as signal molecules involved in the hormonal crosstalk between the plant and the microbiota, increasing the holobiont resistance to biotic and abiotic stresses and influencing the absorption of mineral nutrients[54,55,56].

1.1.2 The soil-root interface and its associated Plant Growth-Promoting Microorganisms

As already mentioned, soil is a rich and dynamic microbial ecosystem, and is in direct contact with plant roots, in which it is possible to distinguish two highly diversified microbiomes: the endophytic community, which populates the endosphere (i.e., plant internal tissues), and the epiphytic community, which thrives in the rhizosphere. The rhizosphere is defined as the small portion of soil (about 1-3 mm thick) that surrounds the roots and that is directly influenced by root biological activity. Specifically, the rhizosphere is the region of the soil being persistently influenced by rhizodeposition of exudates, adhesives, and sloughed cells, and can be regarded as the soil-plant interface colonized by bacterial and fungal species that exert growth-promoting and adaptive benefits[32] (**Fig. 2.1**).

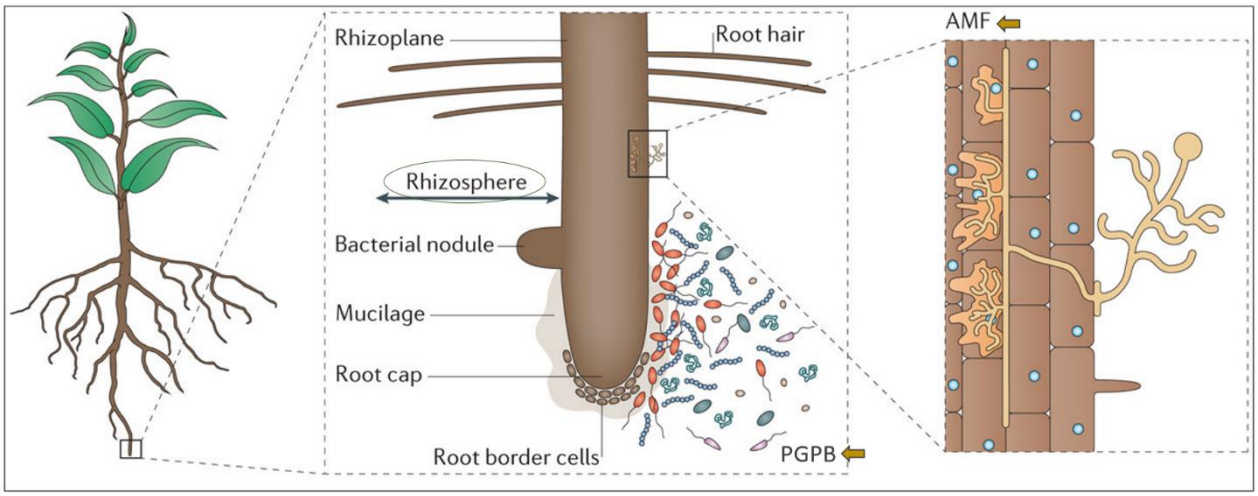


Fig. 2.1 – The rhizosphere (green circle) is the thin portion of soil, usually just few millimeters thick, surrounding plant roots and directly influenced by root exudates (mainly sugars, amino acids, organic acids, and vitamins). A simplified structure of the root terminal portion (tip), where microbial activity is concentrated due to the high release of organic exudates and to the disruption of root border cells, is reported. PGPB and AMF (brown arrows) are worth of note (modified from[57]).

The structure of the rhizosphere microbiota is influenced by several factors, both abiotic, such as the type of soil and the different pedoclimatic conditions, and biotic, such as the nature of root exudates, the plant species, the plant genotype, its development stage, as well as the presence or absence of pathologies[35,57,58]. The '*rhizosphere effect*', in fact, is the phenomenon causing the microbial community of the rhizosphere to be different from that of the bulk soil, implying that plant roots are able to actively recruit and accumulate specific microorganisms in their surroundings, selecting from the rich pool that populates the bulk soil[32,59]. Previous studies have shown that this recruitment occurs in a species-specific and genotype-specific manner, with the health status and developmental stage of the host plant being also relevant. Indeed, all terrestrial plants release low molecular weight compounds into the soil, such as sugars, organic acids, amino acids, vitamins, and phenols, which act both as signal molecules and energy substates for the associated microorganisms[37,41]. As a matter of fact, it was estimated that from 10 to 40% of photosynthetically fixed carbon and up to 15% of total plant nitrogen[41,60,61] are released in the rhizosphere via complex transmembrane transport systems or following the disruption of root border cells, producing a nutrient-rich mucigel which is promptly colonized by microbes[62]. In cucumbers (*Cucumis sativus* L.), for instance, citric acid secretion is involved in the recruitment to the rhizosphere of the well-known PGP bacterium *Bacillus amyloliquefaciens SQR9*[63]. Similar studies have been performed on many plant species, including the model species *Arabidopsis thaliana*, which is able to attract PGPM (e.g., *B. subtilis*) through the release of amino acids into the environment[64]. Furthermore, Yuan and colleagues observed how, in *A. thaliana*, infection by phytopathogenic *Pseudomonas syringae* is effectively repelled following the recruitment at root level of *Bacillus subtilis FB17*, which is attracted by malic acid secretion and is able to activate the plant's immune response at the systemic level[65]. *Bacillus* represents one of the best studied and most widespread groups of PGPB in soil and plants, together with *Pseudomonas*, given the plethora of PGP traits that these taxa can provide. Several Bacilli, such as *B. subtilis*, *B. amyloliquefaciens*, *B. cereus*, *B.*

licheniformis, and *B. megaterium*, in fact, are able to mobilize soil nutrients while producing phytohormones, siderophores, enzymes (e.g., chitinases), assorted antibiotics/bacteriocins, and also low molecular weight bio-stimulating molecules (e.g., VOCs). Besides, the ability of Bacilli to form resistant endospores that can persist for long underground and are stable in the dehydrated form, makes *Bacillus* a suitable taxon for applications as rhizobacterial inoculants (i.e., rhizobacterial-based technologies such as biofertilizers and biocontrol agents) commercialized as an alternative to synthetic fertilizers in order to foster sustainable agricultural practices and crop production[66]. Additional factors known to affect the rhizosphere microbiota are the plant health status and its developmental stage. In the first case, Berendsen and colleagues, for instance, highlighted that *Arabidopsis* root microbiome is subjected to severe oscillations following *Peronospora* infection[58]. As for the growth phase, a study on the same plant model species showed that four rhizosphere-associated bacterial phyla (i.e., Acidobacteria, Actinobacteria, Bacteroidetes, and Cyanobacteria), as well as specific genera within those phyla, followed distinct abundance patterns associated with plant development (seedling, vegetative, bolting, and flowering stage) and root exudation. These results suggest that the plant is able to select a subset of microbes at different stages of development, presumably to perform specific functions, via root exudates that are differentially produced throughout its life cycle, to help orchestrate rhizosphere microbiome assemblage and expand the holobiont metabolic repertoire[10,67]. In particular, the rhizosphere microbiota is generally enriched with a multitude of PGPM directly involved in the biogeochemical cycles of nutrients and in plant bio-stimulation/bio-protection[37,40]. In this context, one of the best-known examples of symbiosis that takes place at the root level, starting from the rhizosphere and then extending to the internal root tissues, is represented by the symbiotic relationship between the *Fabaceae* family (legumes) and soil Proteobacteria mainly belonging to the genera *Rhizobium*, *Azorhizobium*, *Sinorhizobium*, *Mesorhizobium*, and *Bradyrhizobium*. After complex and finely regulated biochemical crosstalk, these microbes induce the formation of root nodules where atmospheric nitrogen can be reduced by means of the nitrogenase enzymatic complex, making it bio-available for the plant holobiont[68]. Another interesting aspect related to the rhizospheric microbiota concerns the induction of the so-called Induced Systemic Resistance (ISR) in plants. ISR is a complex immune pathway, aimed at blocking the action of phytopathogens and triggered by microbial molecular elicitors, such as chitin, flagellin, and lipo-polysaccharides, which are perceived by the plant receptors and activate hormonal/immune responses needed to counteract the pathogen action[69,70]. For example, a study conducted by Van Peer and colleagues emphasized how the presence of *Pseudomonas fluorescens* in the rhizosphere of carnation (*Dianthus caryophyllus* L.) significantly increases the plant's resistance to the pathogenic fungus *Fusarium oxysporum*[71]. Besides, it was also noted that the PGP consortium including *Bacillus pumilus*, *Bacillus subtilis* and *Curtobacterium flaccumfaciens* generates a synergistic effect that increases resistance against several pathogens in cucumber plants[72]. Overall, *Bacillus* taxa are good antagonists against both fungal (e.g., *Alternaria alternata*, *Aspergillus flavus*, *Botrytis cinerea*, *Colletotrichum acutatum*, *Fusarium* spp., *Verticillium dahlia*, etc.) and bacterial (e.g., *Xanthomonas compestris*) plant pests[66]. Plus, in a recent study by Mendes and collaborators, it was seen that the rhizospheric microbiota of a bean cultivar (*Phaseolus vulgaris* L.), mainly composed of bacteria belonging to the families *Oxalobacteraceae*, *Burkholderiaceae*, and *Sphingobacteriaceae*, was significantly enriched in several genes related to antifungal activity[73]. The antimicrobial properties of rhizospheric soils against native soil pathogens are linked to the

phenomenon referred to as '*disease-suppressive soils*', as these properties can be transferred between different soils by direct human action[74]. Interestingly, these studies underline how rhizosphere microbiome research has been carried out for decades but the number of related works has peaked ever since thanks to the advent of NGS methodologies[31]. Overall, despite the differences related to the above-mentioned factors, the dominant prokaryotic taxa found most often associated with the plant rhizosphere across the entire plant kingdom, whether beneficial, commensal, or parasitic, belong to the phyla Proteobacteria (e.g., *Sphingomonas*, *Rhizobium*, *Mesorhizobium*, *Bradyrhizobium*, *Pseudomonas*, *Acidobacter*, *Enterobacter*, *Burkholderia*), Bacteroidetes (e.g., *Niastella*, *Pedobacter*), Acidobacteria (e.g., *Acidobacteriales*, *Solibacterales*), Actinobacteria (e.g., *Mycobacterium*, *Nocardioides*, *Streptomyces*, *Glycomices*), and Firmicutes (e.g., *Bacillus*, *Paenibacillus*, *Clostridium*)[75]. Together with prokaryotes, the rhizospheric microbiota encompasses numerous eukaryotic microorganisms, and, among those, we find mainly fungi belonging to the phyla Ascomycota, Basidiomycota, Zygomycota, and Glomeromycota, which are able to colonize and establish close mutualistic relationships with plant roots[76]. In this view, the most studied and best-characterized are arbuscular mycorrhizal fungi (phylum Glomeromycota, e.g., *Glomus* spp.), due to the peculiar symbiotic relationship that they establish with the root systems of almost 90% of all terrestrial plants[77]. In fact, mycorrhizal symbiosis represents a key aspect of plant biology, due to the strong influence it has on plant productivity. This highlights the roles of mycorrhizae in promoting plant mineral nutrition (e.g., nitrogen and phosphorus acquisition) in exchange for photosynthetically-fixed carbohydrates, in conferring resistance to biotic and abiotic stresses, and in connecting the root systems of different individuals through a complex network of hyphae that allows the movement and exchange of nutrients between all the organisms involved in the symbiosis[77]. The presence of microorganisms, such as PGPB and AMF, in the microbial communities of the rhizosphere is therefore an essential element for the health of all plant holobionts, indicating that shifts in the specific richness and abundance of plant microbiome members can lead to a dysbiotic condition where the probiotic action of mutualistic microorganisms is lacking, thus increasing the probability of nutritional deficiencies and pathologies[18].

A restricted subset of the rhizosphere microbiota, which is selected by the holobiont, is also capable of crossing the rhizosphere and entering the roots, travelling past the root epidermis, through the root cortex, and up to the central vascular cylinder. This is the microbial community of the plant endosphere, referred to as the plant endophytic microbiota, which first inhabits root inter- and intra-cellular spaces, and then reaches all plant compartments[8,31]. Notably, previous experiments proved that less than 30% of the rhizosphere microorganisms may find their way inside the roots, and, among these, less than 5% may enter xylem and phloem and travel to aerial plant organs[8]. Most plant endophytes derive from the soil (horizontal acquisition, that does not exclude the vertical seed-mediated route) and the root represents the initial contact point with the plant, by means of root exudates, signal molecules, and the respective bacterial receptors[31,78]. For example, plant-secreted flavonoids activate specific bacterial genes, resulting in the production of the *nod* factors in Rhizobia, that are necessary for successful interaction between plant and bacterial cells, which are hosted intracellularly within root nodules where they find the proper conditions for reducing atmospheric nitrogen into ammonia[79]. Similarly, rice root exudates induce the expression of genes involved in adherence and signal transduction in the endophyte *Azoarcus* spp., while flagella synthesis is downregulated, indicating that the bacterium is primed for the switch from the rhizosphere to the plant endosphere[80]. Furthermore,

Compant and collaborators reported how microbial species of the genus *Burkholderia* first colonize the root endosphere, and then spread through the xylem to the stomatal chambers, without mixing up with the epiphytic community of the leaf[81]. In spite of this finely-tuned selection, the plant endosphere represents a highly diverse and dynamic environment for microbes, providing multiple niches in different plant compartments, and allowing them to perform a vast array of functions, ranging from mutualism (PGPM) to pathogenicity. Finally, in line with what was discussed for the rhizosphere, the endosphere usually sees the massive presence of probiotic AMF, called endomycorrhizae[77]. Endophytic microbes have lately received high attention, because of the increasing awareness of the importance of host-associated microbiota for the functioning and performance of the plant holobiont. Nevertheless, we are only beginning to elucidate the molecular details of the interplay between plants and endophytes, weakening our understanding of the selective forces in the plant environment, as well as the mechanisms of adaptation of endophytes to this environment[31].

References

[1]Hardoim, Pablo R., Leo S. van Overbeek, and Jan Dirk van Elsas. "Properties of bacterial endophytes and their proposed role in plant growth." *Trends in microbiology* 16.10 (2008): 463-471.

[2]Ryan, Robert P., et al. "Bacterial endophytes: recent developments and applications." *FEMS microbiology letters* 278.1 (2008): 1-9.

[3]Rosenberg, Eugene, and Ilana Zilber-Rosenberg. "Microbes drive evolution of animals and plants: the hologenome concept." *MBio* 7.2 (2016): 10-1128.

[4]Berg, Gabriele, et al. "Microbiome definition re-visited: old concepts and new challenges." *Microbiome* 8 (2020): 1-22.

[5]Rosenberg, Eugene, and Ilana Zilber-Rosenberg. "Symbiosis and development: the hologenome concept." *Birth Defects Research Part C: Embryo Today: Reviews* 93.1 (2011): 56-66.

[6]Gordon, Jeffrey, et al. "Superorganisms and holobionts." *Microbe* 8.4 (2013): 152-153.

[7]Martins, Guilherme, et al. "Characterization of epiphytic bacterial communities from grapes, leaves, bark and soil of grapevine plants grown, and their relations." *Plos one* 8.8 (2013): e73013.

[8]Zarraonaindia, Iratxe, et al. "The soil microbiome influences grapevine-associated microbiota." *MBio* 6.2 (2015): 10-1128.

[9]Coleman-Derr, Devin, et al. "Plant compartment and biogeography affect microbiome composition in cultivated and native Agave species." *New Phytologist* 209.2 (2016): 798-811.

[10]Griggs, Reid G., et al. "Sources and assembly of microbial communities in vineyards as a functional component of winegrowing." *Frontiers in Microbiology* 12 (2021): 673810.

[11]Cregger, M. A., et al. "The Populus holobiont: dissecting the effects of plant niches and genotype on the microbiome." *Microbiome* 6 (2018): 1-14.

[12]Pinto, Cátia, and Ana Catarina Gomes. "Vitis vinifera microbiome: from basic research to technological development." *BioControl* 61.3 (2016): 243-256.

[13]Frank, Anna Carolin, Jessica Paola Saldierna Guzmán, and Jackie E. Shay. "Transmission of bacterial endophytes." *Microorganisms* 5.4 (2017): 70.

[14]Wagner, Maggie R., et al. "Host genotype and age shape the leaf and root microbiomes of a wild perennial plant." *Nature communications* 7.1 (2016): 12151.

[15]Mezzasalma, Valerio, et al. "Geographical and cultivar features differentiate grape microbiota in Northern Italy and Spain vineyards." *Frontiers in microbiology* 9 (2018): 946.

[16]Andreolli, Marco, Silvia Lampis, and Giovanni Vallini. "Diversity, distribution and functional role of bacterial endophytes in Vitis vinifera." *Endophytes: Biology and Biotechnology: Volume 1* (2017): 233-266.

[17]Dong, Chun-Juan, et al. "Bacterial communities in the rhizosphere, phyllosphere and endosphere of tomato plants." *PloS one* 14.11 (2019): e0223847.

[18]Bettenfeld, Pauline, et al. "The microbiota of the grapevine holobiont: a key component of plant health." *Journal of advanced research* 40 (2022): 1-15.

[19]Baldan, Enrico, et al. "Identification of culturable bacterial endophyte community isolated from tissues of Vitis vinifera "Glera"." *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology* 148.3 (2014): 508-516.

[20]Yousaf, Sohail, et al. "Pyrosequencing detects human and animal pathogenic taxa in the grapevine endosphere." *Frontiers in Microbiology* 5 (2014): 327.

[21]Hacquard, Stéphane, et al. "Microbiota and host nutrition across plant and animal kingdoms." *Cell host & microbe* 17.5 (2015): 603-616.

[22]Mercado-Blanco, Jesús. "Life of microbes inside the plant." *Principles of Plant-Microbe Interactions: Microbes for Sustainable Agriculture*. Cham: Springer International Publishing, 2014. 25-32.

[23]Kim, Da-Ran, et al. "Function and distribution of a lantipeptide in strawberry Fusarium wilt disease–suppressive soils." *Molecular Plant-Microbe Interactions* 32.3 (2019): 306-312.

[24]Bozoudi, Despina, and Dimitris Tsaltas. "Grape microbiome: potential and opportunities as a source of starter cultures." *Grape and Wine Biotechnology* (2016).

[25]Fort, Thomas, et al. "Foliar fungal communities strongly differ between habitat patches in a landscape mosaic." *PeerJ* 4 (2016): e2656.

[26]Abdelfattah, Ahmed, et al. "Revealing cues for fungal interplay in the plant–air interface in vineyards." *Frontiers in plant science* 10 (2019): 922.

[27]Del Frari, Giovanni, et al. "Characterization of the wood mycobiome of Vitis vinifera in a vineyard affected by esca. Spatial distribution of fungal communities and their putative relation with leaf symptoms." *Frontiers in plant science* 10 (2019): 910.

[28]Bucci, Enrico M. "Xylella fastidiosa, a new plant pathogen that threatens global farming: Ecology, molecular biology, search for remedies." *Biochemical and biophysical research communications* 502.2 (2018): 173-182.

[29]Chandi, Ravinder Singh. "Integrated management of insect vectors of plant pathogens." *Agricultural Reviews* 42.1 (2021): 87-92.

[30]McFrederick, Quinn S., et al. "Flowers and wild megachilid bees share microbes." *Microbial ecology* 73 (2017): 188-200.

[31]Compant, Stéphane, et al. "The plant endosphere world–bacterial life within plants." *Environmental Microbiology* 23.4 (2021): 1812-1829.

[32]Sasse, Joelle, Enrico Martinoia, and Trent Northen. "Feed your friends: do plant exudates shape the root microbiome?." *Trends in plant science* 23.1 (2018): 25-41.

[33]Mercado-Blanco, Jesús, et al. "Belowground microbiota and the health of tree crops." *Frontiers in Microbiology* 9 (2018): 1006.

[34]Noceto, Pierre-Antoine, et al. "Arbuscular mycorrhizal fungi, a key symbiosis in the development of quality traits in crop production, alone or combined with plant growth-promoting bacteria." *Mycorrhiza* 31.6 (2021): 655-669.

[35]Trivedi, Pankaj, et al. "Plant–microbiome interactions: from community assembly to plant health." *Nature reviews microbiology* 18.11 (2020): 607-621.

[36]Van Der Heijden, Marcel GA, Richard D. Bardgett, and Nico M. Van Straalen. "The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems." *Ecology letters* 11.3 (2008): 296-310.

[37]Singh, Monika, et al. "Plant growth promoting rhizobacteria: application in biofertilizers and biocontrol of phytopathogens." *PGPR amelioration in sustainable agriculture*. Woodhead Publishing, 2019. 41-66.

[38]Pathania, Priyanka, et al. "Role of plant growth-promoting bacteria in sustainable agriculture." *Biocatalysis and Agricultural Biotechnology* 30 (2020): 101842.

[39]Richardson, Alan E., et al. "Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms." (2009): 305-339.

[40]Rolli, Eleonora, et al. "Root-associated bacteria promote grapevine growth: from the laboratory to the field." *Plant and Soil* 410 (2017): 369-382.

[41]Qu, Qian, et al. "Rhizosphere microbiome assembly and its impact on plant growth." *Journal of agricultural and food chemistry* 68.18 (2020): 5024-5038.

[42]Ghosh, Swapan Kr, Tanmay Bera, and Ananda M. Chakrabarty. "Microbial siderophore–A boon to agricultural sciences." *Biological Control* 144 (2020): 104214.

[43]Sasirekha, Bakthavatchalu, and Shivakumar Srividya. "Siderophore production by Pseudomonas aeruginosa FP6, a biocontrol strain for Rhizoctonia solani and Colletotrichum gloeosporioides causing diseases in chilli." *Agriculture and Natural Resources* 50.4 (2016): 250-256.

[44]Cassán, Fabricio, et al. "Everything you must know about Azospirillum and its impact on agriculture and beyond." *Biology and Fertility of Soils* 56 (2020): 461-479.

[45]Pham, Diep Ngoc, et al. "Development of an engineered methanotroph-based microbial platform for biocatalytic conversion of methane to phytohormone for sustainable agriculture." *Chemical Engineering Journal* 429 (2022): 132522.

[46]Gamalero, Elisa, and Bernard R. Glick. "Bacterial modulation of plant ethylene levels." *Plant physiology* 169.1 (2015): 13-22.

[47]del Carmen Orozco-Mosqueda, Ma, Bernard R. Glick, and Gustavo Santoyo. "ACC deaminase in plant growth-promoting bacteria (PGPB): An efficient mechanism to counter salt stress in crops." *Microbiological Research* 235 (2020): 126439.

[48]Khade, Omkar, and K. Sruthi. "The rhizosphere microbiome: A key modulator of plant health and their role in secondary metabolites production." *Biotechnology of Emerging Microbes*. Academic Press, 2024. 327-349.

[49]Olanrewaju, Oluwaseyi Samuel, Bernard R. Glick, and Olubukola Oluranti Babalola. "Mechanisms of action of plant growth promoting bacteria." *World Journal of Microbiology and Biotechnology* 33 (2017): 1-16.

[50]Mishra, Jitendra, and Naveen Kumar Arora. "Secondary metabolites of fluorescent pseudomonads in biocontrol of phytopathogens for sustainable agriculture." *Applied Soil Ecology* 125 (2018): 35-45.

[51]Sugrue, Ivan, R. Paul Ross, and Colin Hill. "Bacteriocin diversity, function, discovery and application as antimicrobials." *Nature Reviews Microbiology* (2024): 1-16.

[52]Nazari, Mahtab, and Donald L. Smith. "Mitigation of drought or a combination of heat and drought stress effects on canola by Thuricin 17, a PGPR-produced compound." *Frontiers in Sustainable Food Systems* 7 (2023): 1237206.

[53]Veliz, Esteban A., Pilar Martínez-Hidalgo, and Ann M. Hirsch. "Chitinase-producing bacteria and their role in biocontrol." *AIMS microbiology* 3.3 (2017): 689.

[54]Cellini, Antonio, et al. "Bacterial volatile compound-based tools for crop management and quality." *Trends in Plant Science* 26.9 (2021): 968-983.

[55]Weisskopf, Laure, Stefan Schulz, and Paolina Garbeva. "Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions." *Nature Reviews Microbiology* 19.6 (2021): 391-404.

[56]Ravelo-Ortega, Gustavo, Javier Raya-González, and José López-Bucio. "Compounds from rhizosphere microbes that promote plant growth." *Current Opinion in Plant Biology* 73 (2023): 102336.

[57]Philippot, Laurent, et al. "Going back to the roots: the microbial ecology of the rhizosphere." *Nature reviews microbiology* 11.11 (2013): 789-799.

[58]Berendsen, Roeland L., et al. "Disease-induced assemblage of a plant-beneficial bacterial consortium." *The ISME journal* 12.6 (2018): 1496-1507.

[59]Edwards, Joseph, et al. "Structure, variation, and assembly of the root-associated microbiomes of rice." *Proceedings of the National Academy of Sciences* 112.8 (2015): E911-E920.

[60]Jones, Davey L., Christopher Nguyen, and Roger D. Finlay. "Carbon flow in the rhizosphere: carbon trading at the soil–root interface." (2009): 5-33.

[61]Lu, Tao, et al. "Investigation of rhizospheric microbial communities in wheat, barley, and two rice varieties at the seedling stage." *Journal of agricultural and food chemistry* 66.11 (2018): 2645-2653.

[62]Canarini, Alberto, et al. "Root exudation of primary metabolites: mechanisms and their roles in plant responses to environmental stimuli." *Frontiers in plant science* 10 (2019): 157.

[63]Zhang, Nan, et al. "Effects of different plant root exudates and their organic acid components on chemotaxis, biofilm formation and colonization by beneficial rhizosphere-associated bacterial strains." *Plant and soil* 374 (2014): 689-700.

[64]Massalha, Hassan, et al. "Live imaging of root–bacteria interactions in a microfluidics setup." *Proceedings of the National Academy of Sciences* 114.17 (2017): 4549-4554.

[65]Yuan, Jun, et al. "Root exudates drive the soil-borne legacy of aboveground pathogen infection." *Microbiome* 6 (2018): 1-12.

[66]Aloo, Becky N., B. A. Makumba, and Ernest R. Mbega. "The potential of Bacilli rhizobacteria for sustainable crop production and environmental sustainability." *Microbiological Research* 219 (2019): 26-39.

[67]Chaparro, Jacqueline M., Dayakar V. Badri, and Jorge M. Vivanco. "Rhizosphere microbiome assemblage is affected by plant development." *The ISME journal* 8.4 (2014): 790-803.

[68]Fierer, Noah. "Embracing the unknown: disentangling the complexities of the soil microbiome." *Nature Reviews Microbiology* 15.10 (2017): 579-590.

[69]Jankiewicz, U., and M. Kołtonowicz. "The involvement of Pseudomonas bacteria in induced systemic resistance in plants." *Applied Biochemistry and Microbiology* 48 (2012): 244-249.

[70]Villena, Julio, et al. "Receptors and signaling pathways for recognition of bacteria in livestock and crops: prospects for beneficial microbes in healthy growth strategies." *Frontiers in immunology* 9 (2018): 2223.

[71]Van Peer, R., G. J. Niemann, and B. Schippers. "Induced resistance and phytoalexin accumulation in biological control of Fusarium wilt of carnation by Pseudomonas sp. strain WCS 417 r." *Phytopathology* 81.7 (1991): 728-734.

[72]Raupach, Georg S., and Joseph W. Kloepper. "Mixtures of plant growth-promoting rhizobacteria enhance biological control of multiple cucumber pathogens." *Phytopathology* 88.11 (1998): 1158-1164.

[73]Mendes, Lucas William, et al. "Influence of resistance breeding in common bean on rhizosphere microbiome composition and function." *The ISME journal* 12.1 (2018): 212-224.

[74]Raaijmakers, Jos M., and Mark Mazzola. "Soil immune responses." *Science* 352.6292 (2016): 1392-1393.

[75]Ling, Ning, Tingting Wang, and Yakov Kuzyakov. "Rhizosphere bacteriome structure and functions." *Nature communications* 13.1 (2022): 836.

[76]Gramaje, David, et al. "Exploring the temporal dynamics of the fungal microbiome in rootstocks, the lesser-known half of the grapevine crop." *Journal of Fungi* 8.5 (2022): 421.

[77]van Der Heijden, Marcel GA, et al. "Mycorrhizal ecology and evolution: the past, the present, and the future." *New phytologist* 205.4 (2015): 1406-1423.

[78]Bergna, Alessandro, et al. "Tomato seeds preferably transmit plant beneficial endophytes." *Phytobiomes Journal* 2.4 (2018): 183-193.

[79]Geurts, René, and Ton Bisseling. "Rhizobium Nod factor perception and signalling." *The plant cell* 14.suppl_1 (2002): S239-S249.

[80]Shidore, Teja, et al. "Transcriptomic analysis of responses to exudates reveal genes required for rhizosphere competence of the endophyte Azoarcus sp. strain BH72." *Environmental Microbiology* 14.10 (2012): 2775-2787.

[81]Compant, Stéphane, et al. "Endophytic colonization of Vitis vinifera L. by plant growth-promoting bacterium Burkholderia sp. strain PsJN." *Applied and environmental microbiology* 71.4 (2005): 1685-1693.

1.2 *Vitis vinifera* L. – microbiome and microbial terroir

1.2.1 *Vitis vinifera*-associated microbiomes

The European grapevine (*Vitis vinifera* L.) is a woody perennial plant, commonly grown in the mediterranean basin, of paramount global economic importance for the production of wine, with millions of hectares planted each year and millions of tons of grapes harvested. Grapevine is a holobiont whose growth is influenced by a rich microbiota, with microorganisms (beneficial, neutral, or harmful) interacting with each other and with the host to regulate plant functions and expand the holobiont metabolic and genetic asset. As a matter of fact, the concept of the microbiome as a ‘‘second plant genome”, bringing a supplementary source of genes and functions, is perfectly depicted in the grapevine holobiont. Therefore, symbiotic grapevine microbes are paramount for the fitness of the holobiont within its environment and may provide the grapevine with phenotypic plasticity under environmental stress scenarios[1]. Microorganisms colonize each organ of *V. vinifera*, from the roots to the aerial parts, and can be regarded as endophytes (found in the internal tissues) or epiphytes (found in contact with the environment) in the rhizosphere (root surface), lignosphere (wood), caulosphere (trunk), phyllosphere (leaf surface), anthosphere (flowers), carposphere (fruits), and spermosphere (seeds)[2,3]. Despite having been extensively investigated and reviewed[3,4,5,6,7,8,9], the majority of the currently published studies focused on the bacterial and fungal fractions of the grapevine microbiota, whereas archaea, micro-eukaryotes, and viruses were less investigated[1].

Soil surrounding roots represents the main microbial reservoir for the grapevine, and plant-microbiome interactions usually start taking place in this rich ecosystem. Roots exude several compounds (mostly small hydrophilic molecules, but also lipids and cellular content) and this rhizodeposition actively selects microbes from the soil, resulting in a specific subset of soil taxa colonizing the rhizosphere and then the endosphere, with a decreasing trend in biodiversity from bulk soil to aerial organs. Indeed, it has been postulated that plant exudates transit from the root cells cytoplasm into the rhizosphere via complex transport proteins, and then they start to shape microbial communities and modulate plant-microbial interactions acting both as signals and nourishment[10], paving the way for microbial colonization. After being chemically-attracted to the grapevine by rhizodeposition, the most common gateways exploited by microbes to enter the roots are inter-cellular junctions in the root epidermis, as well as wounds on the rhizoplane, and, following successful colonization, microbial cells which overcome the vine innate immunity are free to penetrate the root cortex and be spread throughout the host via vascular structures[1]. The soil-as-reservoir assumption was demonstrated by many studies that found microbial species shared between the vineyard soil and the different vine organs. In this direction, Zarraonaindia et al. (2015) carried out a comprehensive metagenomic analysis on the spatial and temporal dynamics of the bacterial communities associated with grapevine organs (leaves, flowers, grapes, and roots), and surrounding bulk soil, in order to shed light on microbiome assembly features in relation to different parameters. Interestingly, they found that, despite marked significant differences between below and aboveground bacterial communities, the majority of vine-associated taxa, also in aerial organs, originated in the soil, and that their distribution reflected the influence of biogeographic/edaphic factors and vineyard management practices as well. These results imply that differences in the soil bacterial communities of different vineyards are reflected not only in the roots, but also aboveground, pointing at the bulk soil as a paramount “microbial seed bank” that may also reach the grapes, thereby exerting a more direct impact on wine characteristics[11]. In line with these

findings, Zhang and collaborators showed both similarities and differences in bacterial and fungal microbiome structure comparing grape berries, leaves, and vineyard soils by means of metabarcoding (16S rRNA and ITS region, respectively) analysis. In particular, despite the presence of some compartment-specific taxa, all the investigated sites shared the main phyla and genera, e.g., the well-known PGPB *Pseudomonas* and *Bacillus*, but with different abundances. The relevant degree of overlap in the dominant microbes between soil, leaves, and grapes, led the authors to consider the vineyard soil as the primary source of microorganisms for the entire holobiont[12]. This assumption was then corroborated by further studies highlighting strong associations between soil bacterial community composition and the grapevine microbiota, including microbes associated with grapes and wine[13]. Besides the soil-root interface, additional entry modes of microbes in aerial vine structures have also been described, and these include stomata, trichomes, wounds, hydathodes, atmospheric agents, insects, and other animals (horizontal transmission), as well as seeds (vertical transmission)[14,15,16,17].

Table 1.1 summarizes decades of grapevine-microbiome research, as reviewed by Bettenfeld and colleagues, and lists the most common bacterial and fungal taxa, at the phylum and genus level, which are generally found in different grapevine compartments after colonization. As shown in the table, in spite of the microbial selection leading to taxonomic differences between compartments, it is reasonable to delineate a continuum of microbial taxa across all vine regions, leading to the presence of a *V. vinifera* core microbiota encompassing microorganisms shared between the different organs of the same individual, as well as between different individuals living in different areas and/or with different characteristics. The dominant bacterial phyla of the grapevine core microbiota are: Actinobacteria, Proteobacteria, Firmicutes, Bacteroidetes, Gemmatimonadetes, Acidobacteria, Nitrospirae, Chloroflexi, Verrucomicrobia, and Planctomycetes. As for fungi, the most widespread groups are Ascomycota and Basidiomycota, whereas a greater number of phyla is detected only in belowground regions (i.e., soil, rhizosphere, and root). At the genus level, the grapevine core microbiota is more easily delineated for the bacterial fraction, encompassing taxa like *Bacillus*, *Pseudomonas*, and *Sphingomonas*. Notably, previous studies are in agreement in assuming the presence of a solid core microbiota for *V. vinifera*, which is not influenced by external factors of any kind[4,11,18,19].

Table 1.1 – Summary of unique and shared taxa between below and aboveground grapevine regions as illustrated in[1]

	Bacterial Phyla	**Bacterial Genera**	**Fungal Phyla**	**Fungal Genera**
Phyllosphere	Acidobacteria Actinobacteria Bacteroidetes Firmicutes Gemmatimonadetes Proteobacteria	*Arthrobacter* *Bacillus* *Blastococcus* *Curtobacterium* *Enterococcus* *Flavobacterium* *Methylobacterium* *Pantoea* *Pseudomonas* *Sphingomonas* *Streptococcus*	Ascomycota Basidiomycota Zygomycota	*Alternaria* *Aureobasidium* *Cladosporium* *Guehomyces* *Epicoccum* *Mucor* *Pandora* *Rhizopus* *Sporormiella*
Reproductive organs	Acidobacteria Actinobacteria Bacteroidetes Firmicutes Proteobacteria	*Bacillus* *Blastococcus* *Enterobacter* *Erwinia* *Gaella* *Massilia* *Methylobacterium* *Micrococcus* *Pseudomonas* *Sphingomonas*	Ascomycota Basidiomycota	*Alternaria* *Aureobasidium* *Botrytis* *Cladosporium* *Cryptococcus* *Davidiella* *Guehomyces* *Penicillium* *Sporobolomyces* *Rhodotorula*
Wood	Acidobacteria Actinobacteria Bacteroidetes Chloroflexi Proteobacteria Verrucomicrobia	*Achromobacter* *Bacillus* *Bradyrhizobium* *Cellulomonas* *Curtobacterium* *Pseudomonas* *Sphingomonas* *Xanthomonas*	Ascomycota Basidiomycota	*Cladosporium* *Alternaria* *Chaetomium* *Aureobasidium*
Soil, roots, rhizosphere	Acidobacteria Actinobacteria Bacteroidetes Chloroflexi Firmicutes Planctomycetes Proteobacteria Verrucomicrobia	*Bacillus* *Blastococcus* *Clostridium* *Flavobacterium* *Gaella* *Methylobacterium* *Micrococcus* *Nitrososphaera* *Pseudomonas* *Rhizobium* *Steroidobacter* *Sphingomonas*	Ascomycota Basidiomycota Chytridiomycota Glomeromycota Mortierellomycota Mucoromycota Zygomycota	*Alternaria* *Archeospora* *Aspergillus* *Dactylonectria* *Fusarium* *Glomus* *Mortierella* *Mucor* *Paraglomus* *Penicillium* *Peziza* *Phaeoacremonium* *Sclerocystis* *Trichoderma*

Nevertheless, several factors are known to affect the microbiota of healthy grapevines, and these are both endogenous (i.e., plant genotype, age, and phenological stage) and exogenous (i.e., plant location, agronomical practices, climate, and soil characteristics) (**Fig. 3.1**).

Plant genetic diversity, which plays a key role in plant immune response and exudation pathways, has been seen to influence vine microbiota both at the scion and rootstock level[1,20]. Different rootstocks secrete different exudates, driving microbial selection and structuring starting from the rhizosphere, and confer resistance to different

pathogens[21]. A recently published work clearly demonstrated the influence of rootstock genetics in shaping *Vitis* microbiota by considering four different rootstock varieties of the same age, grown on the same soil, and managed identically, in order to minimize the variability from all external factors but the rootstock genotype. Bacterial amplicon sequencing revealed sharp segregation and clustering of rhizospheric microbiotas in relation to the rootstock variety, despite the presence of some core microbiota components and similar phylogenetic richness, emphasizing the role played by the rootstock in shaping different microbial communities starting from the rhizosphere[22]. These findings were corroborated by Marasco and colleagues, who investigated five different rootstocks of Barbera plants and showed that they strongly influence the selection and recruitment of soil bacteria, that may also colonize the aboveground plant compartments, with a potential cascade effect on the quality of the fruit. Notably, although richness, diversity, and bacterial community networking were affected by the rootstock, cultivation-based assays revealed that rootstock-specific microbiotas encoded similar PGP traits, carried out by different bacteria, potentially to provide the grapevine with conserved ecological services (e.g., mineral absorption, hormone homeostasis regulation, and production of secondary metabolites)[5]. D'amico et al. (2018) also highlighted the influence of different rootstocks on *V. vinifera* mineral nutrition and consequent yield, as some specifically rootstock-selected microbes may be responsible for regulating potassium absorption in grapevines[23]. Besides, previous studies have demonstrated that vine genetics is somehow related to the microbiota of the aerial organs as well (scion), but these associations may be rather complex to elucidate. In this sense, Singh and others saw that niche differentiation between grapevine phyllosphere and carposphere was a greater driver than host genetics in modulating the microbiota of plants grown in Montpellier. By means of high throughput 16S and ITS profiling, the authors reported that, while some taxa (e.g., *Vagococcus*) may be specifically associated with the grapevine genotype, taxonomy, alpha, and beta diversity of the microbiota were more severely affected by the plant organ[24]. Age and phenology are also linked to the grapevine microbiota, as clearly demonstrated by previous research concerning both bacteria and fungi[25,26,27], although they may not be the most influencing factors. Indeed, a study undertaken across long-term cultivated Italian vineyards reported how fungal (e.g., Ascomycota and Basidiomycota) and bacterial (e.g., Proteobacteria and Firmicutes) rhizosphere taxa were more influenced by biogeography (site-dependent variation of the microbiome) rather than plant age[28]. Furthermore, while some works seem to rule out the effect of the phenological stage on the grapevine microbiota, other studies reveal high microbial variations across the vine growing season. In the first case, Novello et al. (2017) noted that the microbiota associated with *V. vinifera* cultivar Pinot Noir differed from that of the bulk soil (despite several shared features), and these variations were independent of the phenological stage. In this work, the authors considered two sharply distinct stages of vine development (i.e., flowering and early fruiting time), and their NGS outputs suggested that the composition of the vineyard microbiota outcompetes the plant phenological stage as a driver modelling the holobiont characteristics[29]. On the contrary, different publications reported the opposite trend, showing a plant phenology-microbiota association. For example, spatial and temporal dynamics of fungal communities associated with grapes, flowers, leaves, soil, and roots in two Australian vineyards were demonstrated to vary in relation to both the grapevine habitat and developmental stage, showing seasonal community succession patterns over the vine annual growth cycle[30].

Exogenous factors such as plant location, agronomical practices, climate, and soil characteristics are also noteworthy[1]. Among those, we know that the grapevine microbiota is extremely sensitive to chemical soil inputs, as seen in studies comparing microbial community structure under conventional, organic, or biodynamic agriculture. Overall, many articles demonstrate that organic farming is associated with greater microbiome richness and diversity in vineyards[1,31]. Hendgen and colleagues, for instance, unraveled the impact of different management practices on the soil microbiota, which is strictly tied to the vine microbiota, using an amplicon sequencing approach and targeting the Rheingau wine region in Germany. According to their findings, even though fungal species richness remained unaffected, organically-treated vineyards showed significantly higher bacterial richness[32]. Based on the work of Setati et al. (2015), however, farming practices appear highly relevant also in shaping fungal biodiversity, comparing South African Cabernet Sauvignon grapes across three differentially-managed vineyards[33]. Finally, given the assumption that the soil microbial composition has a major impact on the microbiota of the whole grapevine, it is easy to understand why soil edaphic factors and properties, such as grain size, moisture, salinity, nutrient/mineral content, altitude, temperature, and pH, may exert cascade effects on the grapevine holobiont, although ranking the importance of these factors seems to be rather difficult[1]. Taken together, these data highlight the importance and need for a more detailed and methodical characterization of the grapevine-associated microbiotas, considering different organs and different environmental/agronomical contexts, in order to fill out the still very large gap regarding the ultra-complex microbial communities (with their dynamics, functions, and potential applications for sustainable viticulture) which are paramount for the growth and fitness of *V. vinifera*.

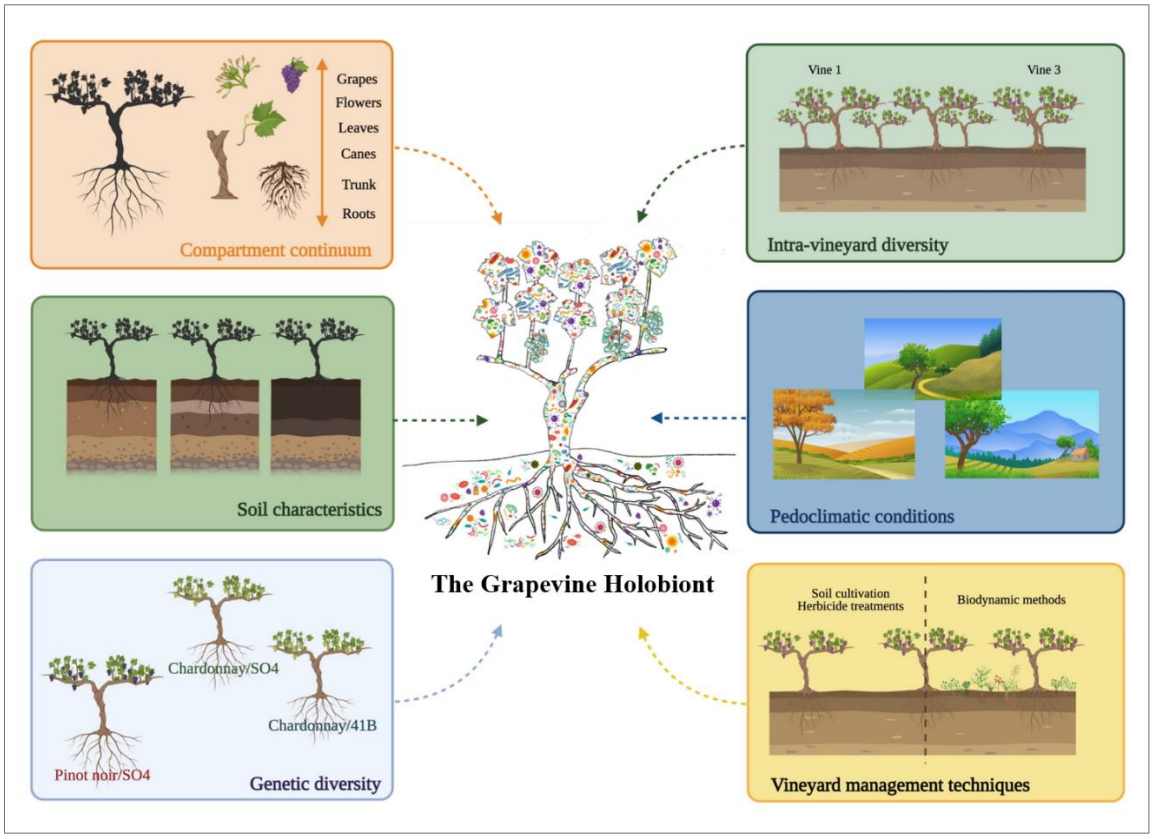


Fig. 3.1 – Schematic illustration of the grapevine holobiont, with its associated microorganisms, and the main endogenous and exogenous drivers known to shape the microbiota (modified from[1]).

1.2.2 The grapevine PGPM at the soil-root interface and the microbial terroir

Plant Growth-Promoting Microorganisms are ubiquitous and represent an intriguing research target given their potential to replace chemicals in viticultural practices, placing them amongst the multiple factors at the origin of the notion of wine ‘‘terroir”. The wine terroir encompasses all the biotic and abiotic parameters of a given viticultural region, including the localization of the vineyard and the winemaking processes used, that explain the production of a wine with unique organoleptic characteristics and sensory attributes, different from those of wines originated elsewhere[34,35]. In this context, the soil-root interface is pivotal, as bacteria and fungi thriving in this ecosystem can directly improve grapevine growth, production yield, and final product quality[1,36]. Therefore, many studies have focused on PGPM found across different vineyards and in grapevine root systems, emphasizing the ecological services that these microbes may provide, while trying to better elucidate the link between soil-rhizosphere-root microbiomes, overall plant fitness, and wine local features (i.e., the so-called *microbial terroir*)[34]. In their work, Vergani et al. characterized the endophytic microbiota and growth parameters of *V. vinifera* (var. Chardonnay) after bio-inoculation with two potential PGPB, namely *Rhizobium* sp. GR12-GFP and *Kosakonia* sp. VR04-mSc, which successfully colonized the endosphere of micro-propagated grapevines under controlled conditions. Interestingly, after three weeks of inoculation, *Rhizobium* sp. GR12-GFP strongly promoted the growth of the root system of the investigated grapevine plantlets, validating the PGP effects of this strain. Indeed, the annotated genome sequence of *Rhizobium* sp. GR12 reveals the presence of multiple PGP traits, including stress response, siderophore production, and auxin biosynthesis. The grapevine endophytic bacterial community was then characterized both via high throughput 16S rRNA gene sequencing and cultivation approaches, unraveling the ability of both inoculated strains to modulate the grapevine microbiota, via direct antagonisms as well as indirect interactions, with cascade effects on the PGP asset of the plant. However, invasion by *Kosakonia* sp. VR04-mSc, other than being associated with no significant increase in plant biomass, led to microbial dysbiosis, that likely hampered plant development. Despite its limitations due to the many differences between microbiotas in the lab and in the field, this study emphasizes the potential of PGP bacteria for reducing chemicals in viticulture, while confirming the need of preserving native microorganisms when attempting to engineer the plant microbiome, in order to ensure holobiont stability[37]. In this direction, some field trials have also been performed. Rolli and colleagues, for example, assessed the ability of 15 PGP bacterial strains to promote grapevine growth in two Italian vineyards over two years, considering both young and adult plants. The aim of their study was to evaluate how bacteria known for their PGP traits in the lab (i.e., auxin, ACC deaminase, protease, exopolysaccharides, and siderophore production, as well as P and N uptake), and isolated from the rhizosphere and endosphere of different plants, could exert their probiotic influence also in the field, where it is impossible to rule out environmental factors. Notably, the authors reported that the majority of the tested isolates consistently boosted grapevine growth in the examined vineyards, by measuring parameters such as shoot length, shoot diameter, node number, number of grape bunches, and total fruit yield. According to the authors, the PGP effects detected in the field could be attributed to beneficial interactions between the inoculated isolates and the natural microbiomes of the grapevine, indicating that lab-pre-screened PGPB may be a useful biotechnological resource for grapevine crop management, and may also be at the base of future large-scale use of eco-friendly natural products instead of chemical fertilizers[38]. These findings are in line with another field trial that recently demonstrated

how scaled-up PGPB inoculants can effectively enhance grapevine growth and fruit yield/quality over time, possibly by means of pathways such as N fixation, P and K solubilization, as well as phytohormone and siderophore production, which were previously tested in the lab[39]. Besides, it has also been proposed that probiotic bacteria found in soil and grapevine roots may ameliorate the holobiont fitness by inducing the biosynthesis of antioxidant secondary metabolites like terpenes and membrane sterols[40,41]. Similar conclusions can be drawn also in relation to arbuscular mycorrhizal fungi (AMF), whose role in grapevine development and growth has been well established[42,43,44], as well as their bio-protection properties[45,46]. Torres and collaborators, for instance, showed that mycorrhizal symbiosis modified the profile of metabolites in *V. vinifera* cultivar Tempranillo berries. Specifically, the levels of glucose and amino acids clearly increased in berries of mycorrhized Tempranillo grapevines, whereas phenolic compounds were not severely affected. The experiment was performed in pot and plants were inoculated with a commercial formulation containing a mixture of five AMF (i.e., *Septoglomus deserticola*, *Funneliformis mosseae*, *Rhizoglomus intraradices*, *Rhizoglomus clarum*, and *Glomus aggregatum*), and two PGPB belonging to the genera *Bacillus* and *Paenibacillus*. Overall, the authors proposed that grapevine mycorrhizal inoculation, by enhancing amino acid content in grapes, may represent an alternative to the application of chemical N compounds, and this may also alter the aromatic characteristics of the wine. Plus, the presence of probiotic bacterial genera in the above-mentioned bioinoculant suggests that PGPB can act synergistically with AMF to benefit host plants, and thus future viticultural practices should take further advantage of PGP strains in order to diminish environmental hazard while safeguarding crop yield, local microbial diversity, and wine quality[47].

Despite these and other studies emphasized the involvement of grapevine-associated microbiotas in vineyard productivity[48], often finding robust correlations between different microorganisms and the final quality of grapes/wine[49], the causal effect of the microbiome on the wine terroir remains hard to delineate[35,49]. A previous review on the matter explained how the microbial terroir for wine grapes may be a pillar for the wine industry. Indeed, grapevines with different traits (e.g., grape size, shape, color, flavor, yield of fruit, etc.) lead to different regional, and widely appreciated, wine features (i.e., the terroir), also at local biogeographic scale, and small shifts in climate, precipitation, soil edaphic factors, and agricultural approaches are known to be associated with shifts in these traits, together with grapevine cultivar and rootstock family. The grapevine microbiota that coexists with the plant, providing it with a plethora of probiotic functions, may also be one of the key factors affecting these traits[34]. In this sense, several works have established strong associations between biogeography and the vineyard/grapevine microbiomes (that, as we know, are intimately interconnected), though often lacking mechanistic understanding of the underlying processes leading to the presence of a well-established microbial terroir that may influence regionalized wine properties[29,50,51]. The work from Gobbi et al. (2022) is probably the most extensive survey focusing on vineyard microbial communities worldwide (both fungal and bacterial) and highlighting a connection between vineyard location and microbial biodiversity on different geographic scales (this paper will be further discussed in section 1.3.3). Briefly, the authors brilliantly delineated a global microbial terroir by meta-analyzing soil samples from 200 vineyards across 13 countries and four continents, suggesting that the microbiome should be considered as an important variable in the definition of homogeneous terroir units from which unique wines are derived. Furthermore, despite the microbial terroir appears to be dependent on several biotic and abiotic factors, some of

which being poorly understood, Gobbi and colleagues were able to exploit microbiome data to develop a random forest model which can be used to predict the geographical origin of microbial samples with reasonable precision, based on microbial taxa and patterns detected. These outstanding results provide evidence that vineyard microbiotas are undoubtedly related to spatial distance, with a possible link to typical wine features, to such an extent that microbial taxonomic composition and structure can be used to predict the location of sampled vineyards around the globe, but also at the local scale[51].

Wine terroir and PGPM have been long studied, and thousands of papers have been published[52], but more recently it has been posited that microbes establish interactions between plants, environment, and human factors that may relate wine sensory attributes to its geographic origin. Growing evidence suggests that Plant Growth-Promoting Microorganisms at the soil-root interface may be paramount for the correct build-up of these interactions, supporting vine growth through countless metabolic pathways, with cascade effects on the final products[1,53,54]. However, the underlying mechanisms connecting PGPM and wine terroir remain somewhat enigmatic, hampering the definition of unmistakable microbial terroirs worldwide. To overcome these limitations, the integration of multi-omic and environmental datasets (including microbiological and chemical analyses) will be pivotal, starting from the biodiversity of microorganisms found associated with vineyards and plants. As a matter of fact, the 'omics' era is allowing us to explore more deeply the nature and consequences of the ultra-complex interactions that support the growth of the grapevine holobiont in the context of its terroir[35].

References

[1]Bettenfeld, Pauline, et al. "The microbiota of the grapevine holobiont: a key component of plant health." *Journal of advanced research* 40 (2022): 1-15.

[2]Berg, Gabriele, et al. "Unraveling the plant microbiome: looking back and future perspectives." *Frontiers in microbiology* 5 (2014): 148.

[3]Pinto, Cátia, and Ana Catarina Gomes. "Vitis vinifera microbiome: from basic research to technological development." *BioControl* 61.3 (2016): 243-256.

[4]Andreolli, Marco, Silvia Lampis, and Giovanni Vallini. "Diversity, distribution and functional role of bacterial endophytes in Vitis vinifera." *Endophytes: Biology and Biotechnology: Volume 1* (2017): 233-266.

[5]Marasco, Ramona, et al. "Grapevine rootstocks shape underground bacterial microbiome and networking but not potential functionality." *Microbiome* 6 (2018): 1-17.

[6]Mezzasalma, Valerio, et al. "Geographical and cultivar features differentiate grape microbiota in Northern Italy and Spain vineyards." *Frontiers in microbiology* 9 (2018): 946.

[7]Darriaut, Romain, et al. "Grapevine decline is associated with difference in soil microbial composition and activity." *Oeno One* 55.3 (2021): 67-84.

[8]Aleynova, Olga A., et al. "The Endophytic Microbiome of Wild Grapevines Vitis amurensis Rupr. and Vitis coignetiae Pulliat Growing in the Russian Far East." *Plants* 12.16 (2023): 2952.

[9]Zuzolo, Daniela, et al. "The rootstock shape microbial diversity and functionality in the rhizosphere of Vitis vinifera L. cultivar Falanghina." *Frontiers in Plant Science* 14 (2023): 1205451.

[10]Sasse, Joelle, Enrico Martinoia, and Trent Northen. "Feed your friends: do plant exudates shape the root microbiome?." *Trends in plant science* 23.1 (2018): 25-41.

[11]Zarraonaindia, Iratxe, et al. "The soil microbiome influences grapevine-associated microbiota." *MBio* 6.2 (2015): 10-1128.

[12]Zhang, Shiwei, et al. "Relations among epiphytic microbial communities from soil, leaves and grapes of the grapevine." *Frontiers in Life Science* 10.1 (2017): 73-83.

[13]Zhou, Jia, et al. "Wine terroir and the soil bacteria: an amplicon sequencing–based assessment of the Barossa Valley and its sub-regions." *Frontiers in Microbiology* 11 (2021): 597944.

[14]Mercado-Blanco, Jesús. "Life of microbes inside the plant." *Principles of Plant-Microbe Interactions: Microbes for Sustainable Agriculture*. Cham: Springer International Publishing, 2014. 25-32.

[15]Frank, Anna Carolin, Jessica Paola Saldierna Guzmán, and Jackie E. Shay. "Transmission of bacterial endophytes." *Microorganisms* 5.4 (2017): 70.

[16]Kim, Ki Woo. "Plant trichomes as microbial habitats and infection sites." *European Journal of Plant Pathology* 154 (2019): 157-169.

[17]Chandi, Ravinder Singh. "Integrated management of insect vectors of plant pathogens." *Agricultural Reviews* 42.1 (2021): 87-92.

[18]Knapp, Dániel G., et al. "Above-ground parts of white grapevine Vitis vinifera cv. Furmint share core members of the fungal microbiome." *Environmental Microbiology Reports* 13.4 (2021): 509-520.

[19]Gramaje, David, et al. "Exploring the temporal dynamics of the fungal microbiome in rootstocks, the lesser-known half of the grapevine crop." *Journal of Fungi* 8.5 (2022): 421.

[20]Wagner, Maggie R., et al. "Host genotype and age shape the leaf and root microbiomes of a wild perennial plant." *Nature communications* 7.1 (2016): 12151.

[21]Wallis, Christopher M., Anna K. Wallingford, and Jianchi Chen. "Grapevine rootstock effects on scion sap phenolic levels, resistance to Xylella fastidiosa infection, and progression of Pierce's disease." *Frontiers in Plant Science* 4 (2013): 502.

[22]Dries, Leonie, et al. "Rootstocks shape their microbiome—bacterial communities in the rhizosphere of different grapevine rootstocks." *Microorganisms* 9.4 (2021): 822.

[23]D'Amico, Federica, et al. "The rootstock regulates microbiome diversity in root and rhizosphere compartments of Vitis vinifera cultivar Lambrusco." *Frontiers in microbiology* 9 (2018): 2240.

[24]Singh, Prashant, et al. "Assessing the impact of plant genetic diversity in shaping the microbial community structure of Vitis vinifera phyllosphere in the Mediterranean." *Frontiers in Life Science* 11.1 (2018): 35-46.

[25]Schreiner, R. Paul, and Keiko L. Mihara. "The diversity of arbuscular mycorrhizal fungi amplified from grapevine roots (Vitis vinifera L.) in Oregon vineyards is seasonally stable and influenced by soil and vine age." *Mycologia* 101.5 (2009): 599-611.

[26]Andreolli, Marco, et al. "Diversity of bacterial endophytes in 3 and 15 year-old grapevines of Vitis vinifera cv. Corvina and their potential for plant growth promotion and phytopathogen control." *Microbiological research* 183 (2016): 42-52.

[27]Ji, Wei, et al. "Characterization of rhizosphere bacterial community and berry quality of Hutai No. 8 (Vitis vinifera L.) with different ages, and their relations." *Journal of the Science of Food and Agriculture* 99.10 (2019): 4532-4539.

[28]Manici, L. M., et al. "Long-term grapevine cultivation and agro-environment affect rhizosphere microbiome rather than plant age." *Applied Soil Ecology* 119 (2017): 214-225.

[29]Novello, Giorgia, et al. "The rhizosphere bacterial microbiota of Vitis vinifera cv. Pinot Noir in an integrated pest management vineyard." *Frontiers in microbiology* 8 (2017): 1528.

[30]Liu, Di, and Kate Howell. "Community succession of the grapevine fungal microbiome in the annual growth cycle." *Environmental Microbiology* 23.4 (2021): 1842-1857.

[31]Hernandez, Maria M., and Cristina M. Menéndez. "Influence of seasonality and management practices on diversity and composition of fungal communities in vineyard soils." *Applied Soil Ecology* 135 (2019): 113-119.

[32]Hendgen, Maximilian, et al. "Effects of different management regimes on microbial biodiversity in vineyard soils." *Scientific reports* 8.1 (2018): 9393.

[33]Setati, Mathabatha E., Daniel Jacobson, and Florian F. Bauer. "Sequence-based analysis of the Vitis vinifera L. cv Cabernet Sauvignon grape must mycobiome in three South African vineyards employing distinct agronomic systems." *Frontiers in microbiology* 6 (2015): 1358.

[34]Gilbert, Jack A., Daniel van der Lelie, and Iratxe Zarraonaindia. "Microbial terroir for wine grapes." *Proceedings of the National Academy of Sciences* 111.1 (2014): 5-6.

[35]Fabres, Pastor Jullian, et al. "A concise review on multi-omics data integration for terroir analysis in Vitis vinifera." *Frontiers in plant science* 8 (2017): 1065.

[36]Visconti, Fernando, Roberto López, and Miguel Ángel Olego. "The Health of Vineyard Soils: Towards a Sustainable Viticulture." *Horticulturae* 10.2 (2024): 154.

[37]Vergani, Lorenzo, et al. "Deciphering the interaction of bacteria inoculants with the recipient endophytic community in grapevine micropropagated plants." *Applied and Environmental Microbiology* 90.2 (2024): e02078-23.

[38]Rolli, Eleonora, et al. "Root-associated bacteria promote grapevine growth: from the laboratory to the field." *Plant and Soil* 410 (2017): 369-382.

[39]Shi, Huimin, et al. "Effect of Microbial Inoculants Endowed with Multifarious Plant Growth-Promoting Traits on Grape Growth and Fruit Quality under Organic Fertilization Scenarios." *Agronomy* 14.3 (2024): 491.

[40]Salomon, María Victoria, et al. "Bacteria isolated from roots and rhizosphere of Vitis vinifera retard water losses, induce abscisic acid accumulation and synthesis of defense-related terpenes in in vitro cultured grapevine." *Physiologia plantarum* 151.4 (2014): 359-374.

[41]Salomon, María Victoria, et al. "Rhizosphere associated bacteria trigger accumulation of terpenes in leaves of Vitis vinifera L. cv. Malbec that protect cells against reactive oxygen species." *Plant Physiology and Biochemistry* 106 (2016): 295-304.

[42]Nogales, Amaia, et al. "Differential growth of mycorrhizal field-inoculated grapevine rootstocks in two replant soils." *American Journal of Enology and Viticulture* 60.4 (2009): 484-489.

[43]Özdemir, G., et al. "Effect of inoculation with mycorrhizal fungi on growth and nutrient uptake of grapevine genotypes (Vitis spp.)." (2010).

[44]Thangavel, Palaniswamy, et al. "Arbuscular mycorrhizae: natural modulators of plant–nutrient relation and growth in stressful environments." *Archives of Microbiology* 204.5 (2022): 264.

[45]Hao, Zhipeng, et al. "Local and systemic mycorrhiza-induced protection against the ectoparasitic nematode Xiphinema index involves priming of defence gene responses in grapevine." *Journal of Experimental Botany* 63.10 (2012): 3657-3672.

[46]Hao, Zhipeng, et al. "Arbuscular mycorrhiza affects grapevine fanleaf virus transmission by the nematode vector Xiphinema index." *Applied Soil Ecology* 129 (2018): 107-111.

[47]Torres, Nazareth, et al. "Aminoacids and flavonoids profiling in Tempranillo berries can be modulated by the arbuscular mycorrhizal fungi." *Plants* 8.10 (2019): 400.

[48]Sabir, Ali, et al. "Growth and mineral acquisition response of grapevine rootstocks (Vitis spp.) to inoculation with different strains of plant growth-promoting rhizobacteria (PGPR)." *Journal of the Science of Food and Agriculture* 92.10 (2012): 2148-2153.

[49]Antolín, M. Carmen, et al. "Dissimilar responses of ancient grapevines recovered in Navarra (Spain) to arbuscular mycorrhizal symbiosis in terms of berry quality." *Agronomy* 10.4 (2020): 473.

[50]Miura, Toshiko, et al. "Is microbial terroir related to geographic distance between vineyards?." *Environmental Microbiology Reports* 9.6 (2017): 742-749.

[51]Gobbi, Alex, et al. "A global microbiome survey of vineyard soils highlights the microbial dimension of viticultural terroirs." *Communications Biology* 5.1 (2022): 241.

[52]Dries, Leonie. *Microbial structure and diversity in the rhizosphere of grape vines (Vitis vinifera L.)*. Diss. Hochschule Geisenheim University, 2024.

[53]Holland, Taylor C., et al. "Response of grapevine rootstocks to soil inocula from different sources." *American Journal of Enology and Viticulture* 69.1 (2018): 94-100.

[54]Trouvelot, Sophie, et al. "Arbuscular mycorrhiza symbiosis in viticulture: a review." *Agronomy for sustainable development* 35 (2015): 1449-1467.

1.3 *Vitis vinifera* L. – regional terroir characteristics and PGPM in Italian vineyards

1.3.1 Aim of the research

The aim of the research illustrated in paragraphs 1.3.2 and 1.3.3 is the characterization of the microbial communities associated with *Vitis vinifera* roots – and surrounding soil – across two worldwide famous Italian viticultural sites, i.e., the "Consorzio Tutela Lambrusco DOC" protected designation of origin area (PDO, Emilia Romagna, Italy) and the "Consorzio del Vino Nobile di Montepulciano DOCG" protected designation of origin area (PDO, Tuscany, Italy). Bulk soil, rhizosphere, and root samples were collected in the field under natural conditions (**Fig. 4.1**) and analyzed in the lab by means of multi-omics and bioinformatics. The main goal of my work was to delve deeper into natural microbiome biodiversity and ecology, with a focus on soil and grapevine-associated PGPM as a determinant of the regional microbial terroir of two renowned – and still poorly investigated – Italian vineyards. Indeed, basic microbiome research is at the base of our understanding of the complex microbiome dynamics that may translate into grape/wine local attributes. In this scenario, more precise knowledge of the regional specificity and probiotic potential of the grapevine/vineyard microbiomes, which is thereafter provided, is paramount for future implementation of environmentally sustainable and effective viticultural strategies. The studies reported in the first chapter of my thesis may also pave the way to design new microbiome-based inoculants for viticultural production, aiming to increase product quality and sustainability while preserving and protecting local microbial diversity, which is essential for the holobiont homeostasis.

Fig. 4.1 – Field sampling of bulk soil and grapevine roots across Italian vineyards for microbiota analysis.

1.3.2 Study I – Composition and biodiversity of soil and root-associated microbiome in *Vitis vinifera* cultivar Lambrusco distinguish the microbial terroir of the Lambrusco DOC protected designation of origin area on a local scale

This section of the thesis was published as an Original Research Article in *Frontiers in Microbiology* (2023); https://doi.org/10.3389/fmicb.2023.1108036. Contents of the article have not been modified except for graphical purposes and bibliography format. Numeration of main and supplementary figures and tables is consistent with the original article.

Authors:

Enrico Nanetti[1], Giorgia Palladino[1,2], Daniel Scicchitano[1,2], Giulia Trapella[1,2], Nicolò Cinti[1], Marco Fabbrini[1,3], Alice Cozzi[1], Giovanni Accetta[4], Carlo Tassini[4], Luigi Iannaccone[4], Marco Candela[1,2] and Simone Rampelli[1]*

*Corresponding Author

Affiliations:

[1]Unit of Microbiome Science and Biotechnology, Department of Pharmacy and Biotechnology (FaBiT), Alma Mater Studiorum – University of Bologna, Bologna, Italy

[2]Fano Marine Center, The Inter-Institute Center for Research on Marine Biodiversity, Resources and Biotechnologies, Fano, Italy

[3]Microbiomics Unit, Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Bologna, Italy

[4]Istituto Tecnico Statale “Ignazio Calvi”, Finale Emilia, Italy

Abstract

Introduction: Wines produced from the same grape cultivars but in different locations possess distinctive qualities leading to different consumer's appreciation, preferences, and thus purchase choices. Here, we explore the possible importance of microbiomes at the soil-plant interface as a determinant of the terroir properties in grapevine production, which confer specific growth performances and wine chemo-sensory properties at the local scale.

Methods: In particular, we investigated the variation in microbial communities associated with the roots of *Vitis vinifera* cultivar Lambrusco, as well as with surrounding bulk soils, in different vineyards across the "Consorzio Tutela Lambrusco DOC" protected designation of origin area (PDO, Emilia Romagna, Italy), considering viticultural sites located both inside and outside the consortium in two different seasons (June and November 2021).

Results: According to our findings, rhizospheric and soil microbiomes show significant structural differences in relation to the sampling site, regardless of seasonality, while endophytic microbiomes seem to be completely unaffected by such variables. Furthermore, a deeper insight into the microbial terroir of PDO areas highlighted the presence of some rhizospheric microorganisms enriched inside the consortium and characterizing the PDO regardless of both sampling season and farming strategy. These include *Bacillus*, *Paenibacillus*, and *Azospirillum*, which are all well-known plant growth-promoting bacteria.

Discussion: Taken together, our results suggest a connection between soil and root microbiomes of *V. vinifera* cultivar Lambrusco and the local designation of origin, emphasizing the potential role of PDO-enriched plant growth-promoting bacteria in vine growing and final quality of the Lambrusco DOC wine.

Keywords

microbial terroir, microbiomes, *Vitis vinifera*, rhizosphere, plant growth-promoting bacteria

Introduction

Wine is a fermented product of paramount economic and cultural importance for the agri-food sector worldwide[1,2]. Therefore, vineyards are widely distributed and *Vitis vinifera* is one of the most cultivated fruit crops around the globe[3,4]. The local-scale pedoclimatic variation, also known as *terroir*, is a matter of growing interest for wine production, because considered of vital importance for the determination of the local wine quality characteristics and consequent consumer's appreciation, preferences, and purchase choices[5,6]. Indeed, wines produced from the same grape cultivars but belonging to different *terroirs* possess distinctive qualities and economic value. To legally protect such local regional products, many geographical pedigrees – such as the Protected Designation of Origin (PDO) in Europe – have been released. However, establishing which factors underlie connections between *terroir* properties and the specific wine-associated chemo-sensory properties remains difficult and is mainly ascribed to general environmental characteristics that affect grapevine growth and health[7].

Recent studies suggested that the specific microbial communities associated with *V. vinifera* may be a key element of the *terroir*, as microbiome processes essential for vine growing and wine production show spatially defined patterns linked to the vineyard location[3,8,9,10]. Vitulo et al. (2019), for instance, found that the geographic indication is a good driver of microbiome differentiation of the vine bark when comparing plants from two Italian wine-producing regions (Piedmont and Tuscany)[11]. Similar results were obtained by Mezzasalma et al. (2017), who defined a certain fraction of the grape berries microbiome that significantly varied in relation to the geographical area[12]. The same association has been highlighted when investigating the soil and the root microbiomes, with distinct microbial characteristics for different viticultural regions that probably correspond to a regional-specific contribution to the qualities of the grapes and wine[8,13,14,15,16,17].

In Italy, in pedoclimatic regions including well-defined delimitations of PDO production, the same grapes are cultivated inside and outside the PDO sites, with similar yields but different properties. This opens the question of the importance of microbiome variations at the soil-plant interface in determining the local *terroir* quality at the local scale, with the cascade implications for the PDO production. In order to provide some glimpses in this direction, we aim at investigating the presence of differences in the microbiome-dependent *terroir* features (rhizospheric, endophytic, and bulk soil microbiomes) in plant specimens of *V. vinifera* cultivar Lambrusco sampled across three different vineyards from the same pedoclimatic region but located inside and outside the "Consorzio Tutela Lambrusco DOC" PDO area, in Emilia Romagna, Italy. In particular, two vineyards were positioned immediately inside the PDO area, and another vineyard just outside the PDO area. In Emilia-Romagna, the Italian region leading Lambrusco's production globally (https://www4.istat.it/it/archivio/207188), it is of primary economic importance to safeguard the "Consorzio Tutela Lambrusco DOC" PDO. Indeed, with 42 million bottles sold in 2020 (https://magazine.wein.plus/news/three-consortia-of-lambrusco-producers-in-emilia-romagna-are-merging-new-protection-association-consorzio-tutela-lambrusco-monitors-eight-doc-areas), Lambrusco DOC is one of the best-selling Italian wines in the world. Moreover, its PDO territory overlaps with that of the Balsamic Vinegar of Modena (that is produced from the same grapes as Lambrusco DOC wine), which showed a production turnover of 370 million euros in 2021 (https://www.consorziobalsamico.it/consortium/economic-data/?lang=en). For all these reasons, we

think that a finer characterization of the microbial *terroir* on the boundaries of the PDO area can contribute to a better safeguard and enhancement of the production. Specifically, for capturing the full variation due to different agricultural practices, the two vineyards within the PDO area were subjected to different agronomic approaches, i.e., organic and conventional agriculture. Furthermore, to test the hypothesis that the composition and/or the diversity of the microbiome at the soil-root interface constitute a signature for defining and protecting a PDO area, we explored the microbiome structure at two time points (June and November) in order to get a full picture of the root-associated microbial *terroir* at different stages of plant maturation. All the plants included in this study were *V. vinifera* cultivar Lambrusco, grafted on the hybrid *Vitis berlandieri* × *Vitis riparia* KOBER 5BB. In addition to enriching our understanding of the importance of soil and root-associated microbiomes in defining the wine *terroir* and the relative PDO area, this study may provide further economic incentive for agricultural and oenological practices that safeguard regional microbial *terroir* and biodiversity.

Materials and methods

Study site

Grapevine roots and soil samples were collected from three different vineyards in Emilia Romagna (Italy) across two timepoints. In particular, from each site, i.e., Bondeno (44.953 N/11.305 E, Ferrara), Finale Emilia (44.839 N/11.285 E, Modena), and Medolla (44.816 N/11.062 E, Modena; **Figure 1**), 15 plants and two bulk soils were retrieved both in June and November 2021 (immediately after the grape harvest), for a total of 90 root and 12 soil samples. Furthermore, for each root sample, the rhizospheric and endophytic microbial communities were both analyzed. The three vineyards considered in this study were characterized by different agronomic managements and biogeographical features. Specifically, both Finale Emilia and Medolla sites are located inside a protected designation of origin (PDO) area but differ in terms of the agricultural approach employed (chemical-based vs. organic, respectively) while the Bondeno site is found outside the PDO area and a traditional chemical-based farming approach is used. A schematic summary of samples distribution across the three sites and the two timepoints (June and November) is provided by **Supplementary Table 1.**

Samples collection and pre-treatment

For the microbiome characterization, each plant root was investigated considering two different ecosystems, namely rhizospheric soil and root endophytic ecosystem, and samples of bulk soil were also analyzed, for a total of 102 samples (**Supplementary Table 1**). Grapevine thin lateral roots were collected after digging 10-20 cm under the plants. Bulk soil samples were collected near the area where plants were located, after removing the top centimeters of surface soil and the grass cover, if present. All samples were collected wearing sterile gloves, placed inside a sterile 50 ml Falcon tube and stored at −80°C until further processing.

In order to separate the two plant compartments (i.e., rhizosphere and endosphere), roots were thoroughly treated as previously described in D'Amico et al. (2018)[18]. In brief, approximately 3 cm of terminal roots portions, including

tips, were dissected using sanitized scissors and tweezers to standardize the quantity of starting material. Then, root segments were placed in 15 ml Falcon tubes filled with 2.5 ml of modified PBS buffer (130 mM NaCl, 7 mM Na_2HPO_4, 3 mM NaH_2PO_4, pH 7.0, and 0.02% Silwet L-77) and left on a shaking platform at 180 rpm for 20 min to perform washing. After removing the roots, the washing buffer was centrifuged at 1.500 × *g* for 20 min and the resulting pellet was regarded as the rhizospheric soil. Roots were then re-washed under the same conditions and transferred to another 15 ml Falcon tube containing 2.5 ml of modified PBS buffer before undergoing 10 cycles of sonication as follows: 30-s pulses at 160 W with 30-s breaks in an ultrasonic bath (Branson 1800, Branson Ultrasonic Corporation, Danbury, CT, United States). After washing and sonication, roots were grinded by means of mortar and pestle in order to reach the root inner portions. This procedure led to a total of 180 samples (90 rhizospheres + 90 roots) that were analyzed together with the 12 bulk soil samples. All samples were kept frozen at −80°C until genomic DNA extraction.

DNA extraction and sequencing

Total genomic DNA was extracted from all the 192 samples, i.e., bulk soils (0.25 g), rhizospheres (approximately 0.25 g), and smashed roots, using the DNeasy PowerSoil Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions with minor modifications: a FastPrep instrument (MP Biomedicals, Irvine, CA, United States) was used for the homogenization step with a cycle consisting of three 1-min steps at 5.5 movements per sec with 5-min incubations in ice between each run and, at the end of the protocol, DNA elution was preceded by a 5-min incubation in ice. Then, DNA was quantified by using NanoDrop ND1000 (NanoDrop Technologies, Wilmington, DE) and diluted in PCR grade water to the final concentration of 5 ng/μl before amplification. Five microliters of diluted DNA were used as template for the PCR reaction. PCR was performed in a final volume of 50 μl containing 25 ng of genomic DNA, 2X KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland) and 200 nmol/L of 341F and 785R primers carrying Illumina overhang adapter sequences for amplification of the V3-V4 hypervariable regions of the 16S rRNA gene. Specifically, the thermal cycle consisted of initial denaturation at 95°C for 3 min followed by 25 cycles of denaturation (95°C for 30 s), annealing (55°C for 30 s), and extension (72°C for 30 s), with a final extension step at 72°C for 5 min[19]. PCR amplicons were cleaned up with Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, United States). Indexed libraries were prepared by limited-cycle PCR using Nextera technology. Indexing was followed by a second clean-up step, as already described, and then libraries were quantified using Qubit 3.0 fluorimeter (Invitrogen), normalized to 4 nM and pooled. Before sequencing, the sample pool was denatured with 0.2 N NaOH and diluted to 4.5 pM with a 20% PhiX control. Sequencing was performed on Illumina MiSeq platform using a 2 × 250 bp paired end protocol, according to the manufacturer's instructions (Illumina, San Diego, CA, United States). Sequencing reads were deposited in the ENA archive with the accession code PRJEB57815.

Bioinformatics and biostatistics

Raw sequences were analyzed using a pipeline which combines PANDAseq 2.11[20] and QIIME2 2021.8.0[21] for all 192 samples. High-quality reads (min/max length = 350/550 bp) were retained thanks to the "fastq filter" function of the Usearch 11.0.667 algorithm[22] and then binned into amplicon sequence variants (ASVs) using DADA2 2021.8.0[23].

Samples with less than 1,000 high-quality reads were discarded and not used for subsequent analyses. The VSEARCH algorithm 2021.8.0.[24] and the SILVA database (December 2017 release;[25]) were employed for taxonomic classification. All unassigned and eukaryotic sequences were discarded. ASVs table was then rarefied to retain a number of 1,138 sequences per sample. The QIIME2 feature table rarefy plugin was used to perform rarefaction. Statistical analyses were carried out using the R software (R Core Team; www.r-project.org – last access: April 2022), v. 4.2.0, implemented with the packages "Made4" 1.72.0[26], "vegan" 2.6-4 (https://cran.r project.org/web/packages/vegan/index.html – last access: October 2022), "pairwiseAdonis" 0.4[27], and STAT 0.1.0 (https://cran.r-project.org/web/packages/STAT/index.html –last access: April 2019). Beta diversity based on unweighted UniFrac distances was computed, and the separation of data in the Principal Coordinates Analysis (PCoA) was assessed with a permutation test with pseudo-F ratios (function "adonis" in the vegan package and function pairwiseAdonis in the homonymous package). Kruskal-Wallis test was used to assess significant differences in alpha diversity distributions between groups. Bacterial genera with the largest contribution to the ordination space were detected by the function envfit of the R package vegan on the genus relative abundances. *p* values, when necessary, were corrected for multiple testing by means of the Benjamini-Hochberg method, with a false discovery rate (FDR) ≤ 0.05 considered to be statistically significant. Linear discriminant analysis (LDA) effect size (LEfSe,[28]), aimed at identifying discriminant rhizospheric taxa between vineyards located inside and outside the Lambrusco DOC PDO viticultural area, regardless of the agricultural practices employed, was performed on genus-level relative abundance tables, retaining only taxa with LDA score threshold of ±2 (on a log10 scale) and value of $p \leq 0.2$. The online Galaxy Version interface (https://huttenhower.sph.harvard.edu/galaxy/, last accessed in October 2022) was used to run LEfSe. All taxa identified by LEfSe, thus significantly enriched either inside or outside PDO sites, were then tested for their putative ability to support plant growth by the presence of some well-known plant growth-promoting (PGP) genes. For this purpose, starting from the QIIME2 genera level taxonomic assignment, an oligotyping procedure[29] was implemented to detect the species belonging to the genera previously identified by LEfSe through the "Minimum Entropy Decomposition" (MED) module and the global earth microbiomes (GEM) catalogue (November 2020 release;[30]). For each genus, the command line was "decompose <ASVs representative sequences fasta.file> − g-M 1-V5." The-M integer defines the minimum substantive abundance of an oligotype, and the-V integer defines the maximum variation allowed in each node. The node representative sequence of each oligotype was used for species profiling with the QIIME2 feature-classifier plug-in[21], selecting the VSEARCH algorithm 2021.8.0[24] and the GEM database[30]. Then, the aminoacidic sequence of some well-characterized PGP proteins, obtained from the reference sequence of the NCBI protein database (https://www.ncbi.nlm.nih.gov; accessed from the 1st to the 31st October 2022), was recovered and blasted against the non-redundant protein sequences NCBI database selecting as target organisms for our queries the bacterial taxa identified by the oligotyping procedure. Unclassified members of a specific taxon were considered when it was impossible to assign the ASVs at the species level, or when the number of oligotypes not assigned at the species level but assigned at higher taxonomic levels overcame the number of species-level matches. Alignments were filtered according to a query coverage of at least 40% and an alignment percentage of identity of at least 20%. The PGP functions selected for our analysis were: nitrogen fixation, phosphorous solubilization, iron chelation, production of the phytohormone indole-3-acetic acid

(IAA), and production of the enzyme 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase. For each of these functions, we selected some marker genes from previous scientific literature resulting in 15 protein sequences recovered and blasted. Specifically, the chosen marker genes were NifB, NifE, NifH, NifN, NifV, and NifU (i.e., nitrogen fixation genes) for nitrogen fixation, the alkaline phosphatase phoA and the glucose dehydrogenase GDH for phosphorous solubilization, three markers of two relevant bacterial siderophores for iron chelation (namely EntF/EntS for enterobactin and FslA for rhizoferrin), three genes directly involved in IAA synthesis (i.e., ipdC, aro10, and aldH) and the AcdS gene encoding the enzyme ACC deaminase (see **Supplementary Table 3** for genes accession and version numbers). Moreover, the presence of the same marker genes was verified across the entire rhizospheric microbiome by means of Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2 v. 2.5.0) analysis[31]. Notably, during the process of ASVs sequences matching to the KEGG database ([32], queried on January 10th, 2023), two out of 15 reference proteins (i.e., FslA and aro10) were not found in the KEGG database (**Supplementary Table 4**). Bacterial co-abundance groups (CAGs) were determined as formerly described by Schnorr et al. (2014)[33]. In brief, the Kendall correlation test was used to evaluate the associations among bacterial genera, which were visualized using hierarchical Ward clustering with a Spearman correlation distance metric and used to define CAGs at the genus level. The significant associations observed were controlled for multiple testing with the *q* value method (FDR ≤ 0.05;[34]). Permutational multivariate ANOVA (PERMANOVA;[35]) was employed to verify whether the CAGs were significantly different from one another. The Wiggum plot network analysis was carried out using cytoscape software v. 3.9.1 (http://www.cytoscape.org/, last accessed in November 2022) as previously described[36].

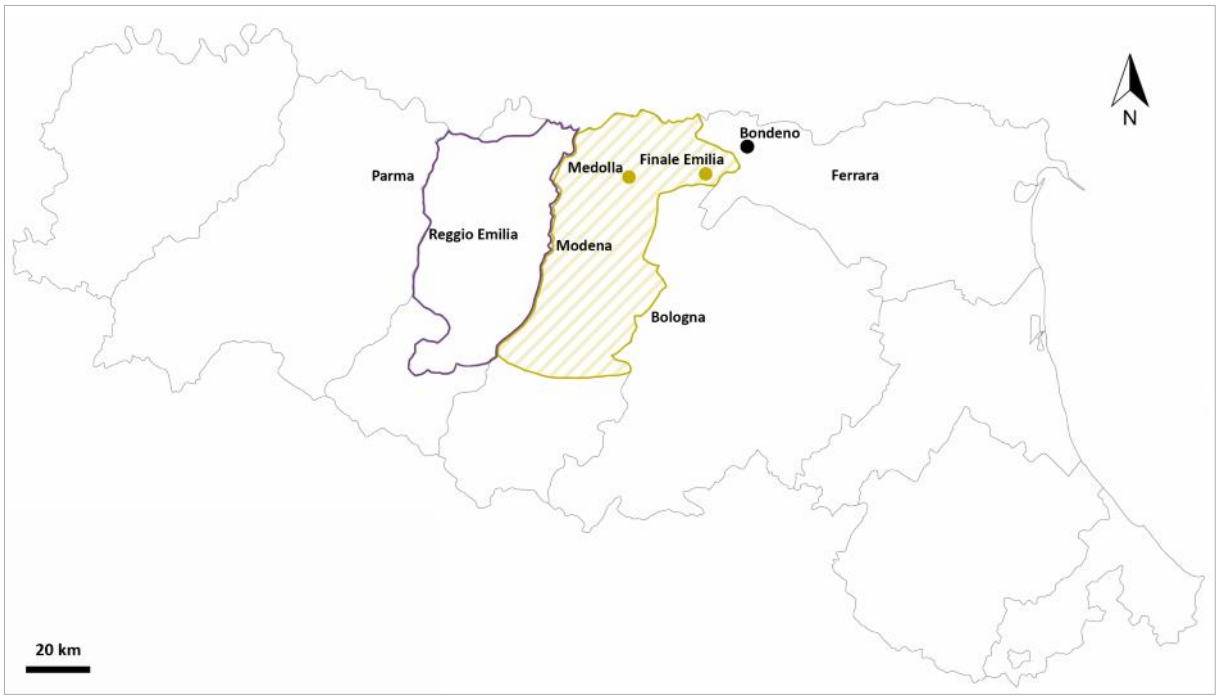


Fig. 1 – Sampling locations. Map of Emilia Romagna (Italy) showing the present study sampling sites in Bondeno (FE), Finale Emilia (MO) and Medolla (MO). Sampling locations are represented as yellow-green dots if located inside the Lambrusco DOC PDO viticultural area and as black dots if located outside the borders of such area. Borders of the entire "Consorzio Tutela Lambrusco DOC" are drawn (with a darker color for the Reggio Emilia territory and with a lighter color for the Modena territory).

Results

Microbiome composition and biodiversity in soil, rhizosphere, and root samples from viticultural farms located inside and outside the Lambrusco DOC PDO viticultural area

A total of 90 *V. vinifera* Cultivar Lambrusco roots samples and 12 bulk soil samples were taken from three different viticultural farms in June and November 2021 in Emilia Romagna, Italy (**Figure 1**). In particular, from each vineyard (located in Bondeno, Finale Emilia and Medolla) 30 roots (15 in June and 15 in November) and four bulk soils (two in June and two in November) were retrieved, resulting in 102 samples. Among those, all the 90 roots were treated as previously described in order to separate the rhizospheric from the endophytic compartment, leading to a total of 180 *V. vinifera* samples and 12 bulk soil samples. The selected farms were characterized by different designation of origin and by different agricultural practices: (i) Bondeno (non-PDO area, conventional farming), (ii) Finale Emilia (PDO area, conventional farming), and (iii) Medolla (PDO area, organic farming). For the three sites and the two timepoints, microbiome compositional structure was investigated by NGS sequencing of the 16S rRNA gene (V3-V4 hypervariable regions), resulting in ≃1,5 M high-quality reads, with an average of 9,581 ± 2,329 reads per sample (mean ± SD), which were binned in 31,264 ASVs (samples with less than 1,000 high-quality reads were not analyzed). Firstly, the bulk soil microbiome was characterized by a significantly higher degree of biodiversity with respect to both rhizospheric and endophytic compartments ($p \leq 0.05$, Kruskal-Wallis test, **Supplementary Figure 1**). When we sought for differences among farms, we only observed a gradual increase of the soil biodiversity from Bondeno, to Finale Emilia and Medolla, with a trajectory that mirrored the path from non-PDO to PDO area and from conventional to organic management. However, these differences are only appreciable at soil level, with the rhizosphere and root compartments from the different farms showing comparable levels of biodiversity (**Figure 2**). Beta-diversity analysis revealed a clear pattern toward segregation of the rhizospheric microbial communities according to the sampling location, but not to the sampling season, as shown by the unweighted UniFrac distances (permutation test with pseudo F-ratio, $p \leq 0.001$; **Figure 3A**). Interestingly, the same PCoA indicates that a similar trend can be observed also for the bulk soils, as if the differences detected into the rhizosphere compartment mirrored differences in the soil. In order to identify those bacterial genera most contributing to the separation of the rhizospheric samples in the PCoA, the relative abundances of such taxa were superimposed in the unweighted UniFrac beta diversity plot (**Figure 3B**). Our results indicate that some bacterial genera are more represented in a particular farm regardless of the season. Specifically, the genus *Pirellula*, *Micromonospora*, and *Nocardioides* are the most characteristic of the Bondeno farm, while *Pseudomonas*, *Flavobacterium*, *Acinetobacter*, *Pir4 lineage*, and

Planctomyces can be associated with Finale Emilia samples and, finally, *Skermanella*, *Gaiella*, *Solirubrobacter*, and *Rubrobacter* are the most distinguishing of the Medolla farm.

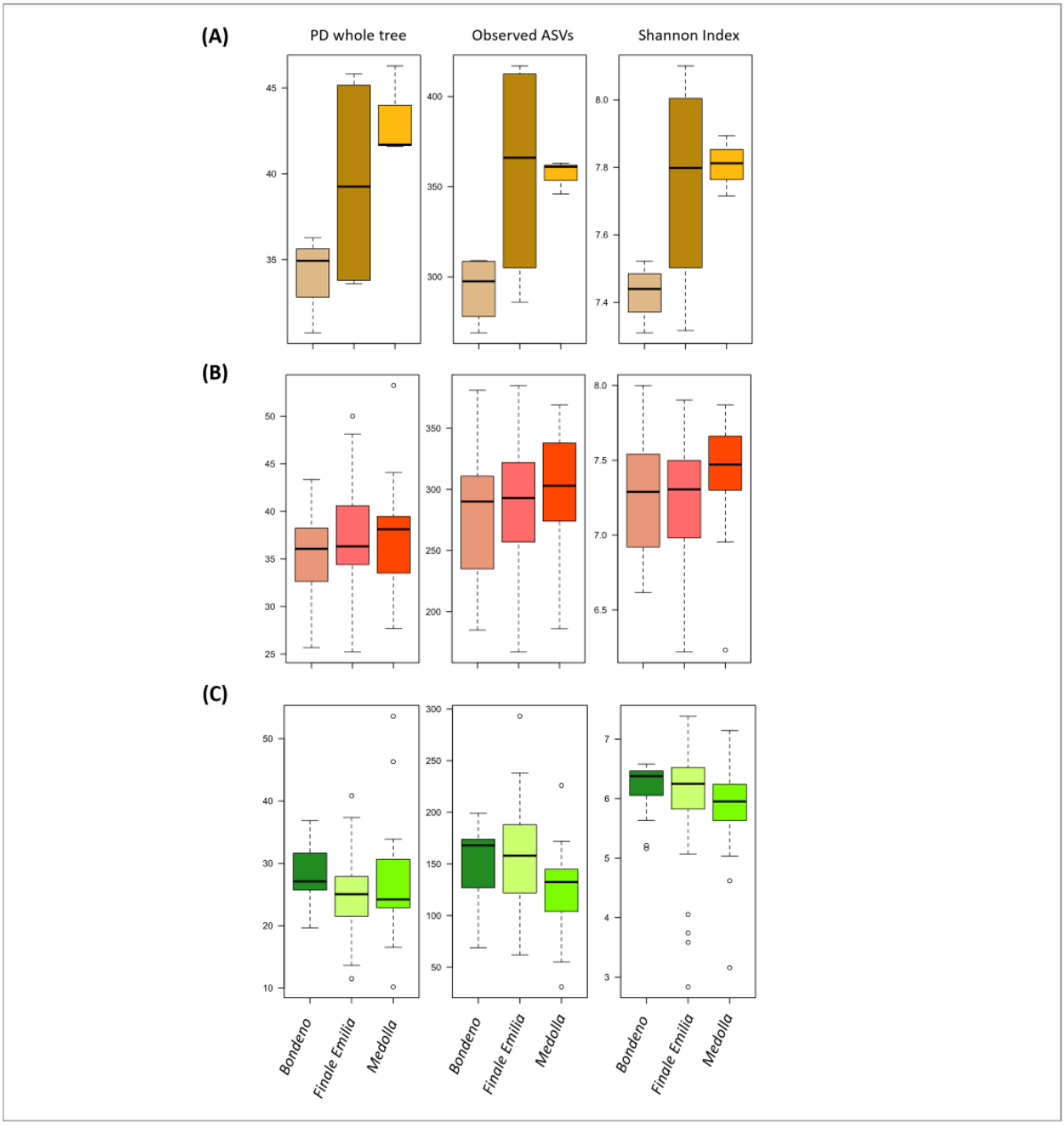


Fig. 2 – Alpha diversity of bulk soil and of *Vitis vinifera* rhizosphere and root microbiomes in the three studied farms. Box-plots showing the distributions of the Faith's Phylogenetic Diversity (PD whole tree), Observed ASVs and Shannon Index calculated for the bulk soil **(A)**, the rhizosphere **(B)**, and the root **(C)** in the three sampled vineyards (located in Bondeno, Finale Emilia, and Medolla). The only significant differences were observed for the bulk soil samples ($p \leq 0.05$, Kruskal–Wallis test).

Conversely, it is noteworthy to point out that the only ASVs detected across the entire rhizospheric cohort were assigned to uncultured members of the *Planctomycetaceae* and to uncultured members of the *Tepidisphaeraceae*. For these taxa, coefficients of variations were 0.4 (mean ± SD % rel. ab., 3.6 ± 1.5) and 0.6 (mean ± SD % rel. ab., 1.7 ± 1.0), respectively, meaning that these taxa were present in the rhizospheres at comparable levels, independently

of site and season, constituting a sort of core bacterial group for *V. vinifera* cultivar Lambrusco. Interestingly, the root samples show no significant structural differences across different sites and seasons (permutation test with pseudo *F*-ratio, $p > 0.05$) and a sharp segregation appears in the PCoA only when comparing the endophytic cluster with the entire set of the bulk soil samples (permutation test with pseudo *F*-ratio, $p \leq 0.001$; **Supplementary Figure 2**).

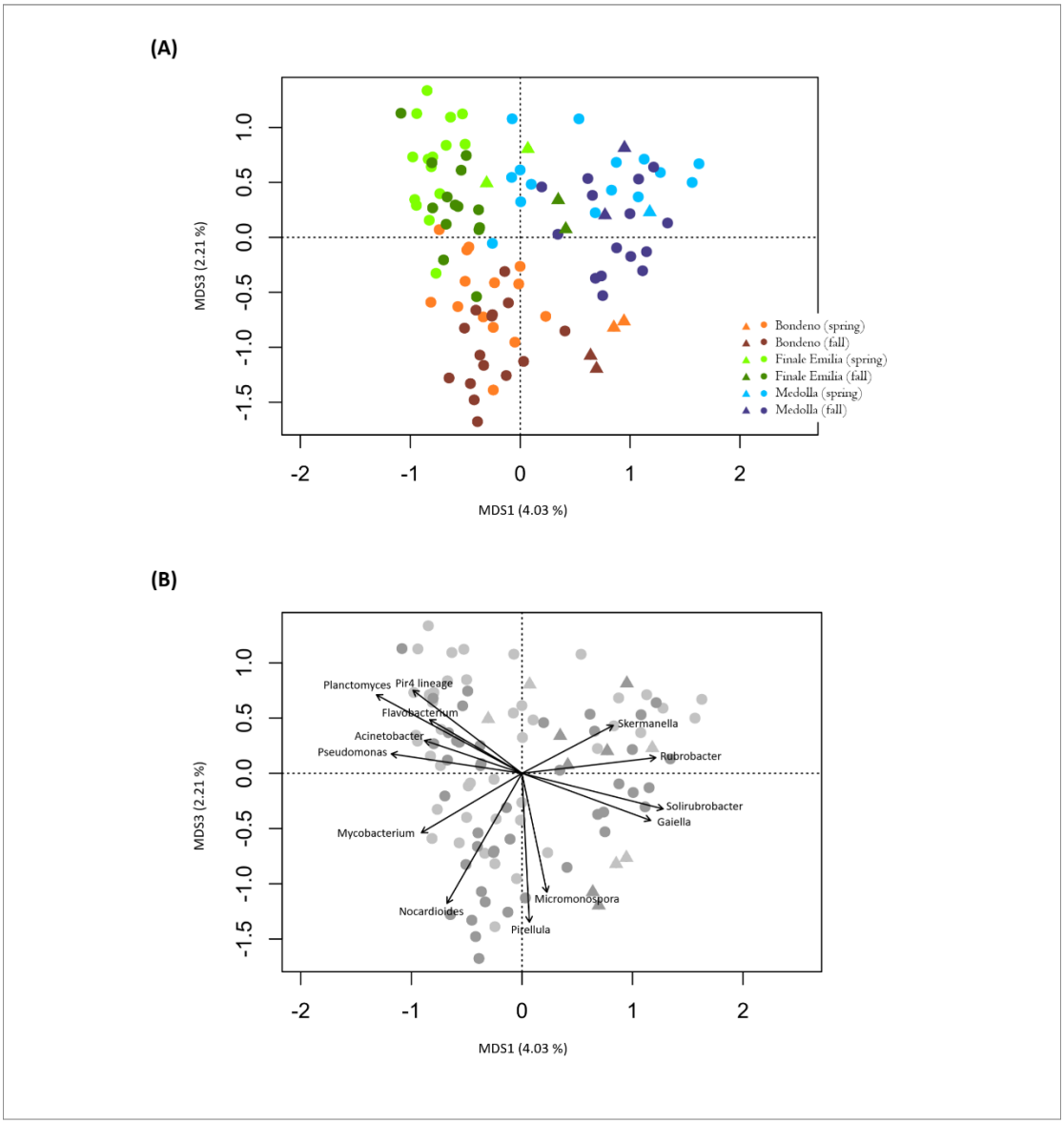


Fig. 3 – **(A)** Principal Coordinates Analysis (PCoA) based on unweighted UniFrac distances showing the variation of *Vitis vinifera* rhizosphere (dots) and bulk soil (triangles) microbiomes across sites, i.e., Bondeno (orange-red), Finale Emilia (green), and Medolla (blue) and seasons (lighter shades for spring and darker shades for fall; permutation test with pseudo *F*-ratio, $p \leq 0.001$). The first and third principal components (MDS1 and MDS3) are plotted and the percentage of variance in the dataset explained by each axis is highlighted. **(B)** The same graph as in **(A)** has been reprinted in order to visualize the bacterial genera most contributing to segregations, whose relative abundance was superimposed in the PCoA plot (function envfit of the R package vegan) considering only genera with a $p \leq 0.001$.

The LEfSe was finally used to identify rhizospheric bacterial genera that discriminated PDO-associated from non-PDO microbiomes, regardless of farming site, season and type of management (**Figure 4**). In particular, genera associated with Lambrusco DOC PDO area were *Bacillus*, *Pseudarthrobacter*, unclassified members of the order Gaiellales, *Planctomyces*, *Skermanella*, *Pir4 lineage*, *Microlunatus*, and *Paenibacillus*. On the other hand, genera less representing the PDO area were *Nocardioides*, *Micromonospora*, and *Pirellula*, unclassified members of the family *Gemmatimonadaceae* and of the order Acidimicrobiales, *Mycobacterium*, *Legionella*, and *Chthoniobacter*. Notably, such taxa were also generally more represented into the correspondent bulk soil microbiome, with the exception of *Pseudarthrobacter* and *Pir4 lineage* for what concerns the PDO area and *Nocardioides*, *Legionella*, and *Chthoniobacter* for what concerns the non-PDO area (**Supplementary Table 2**).

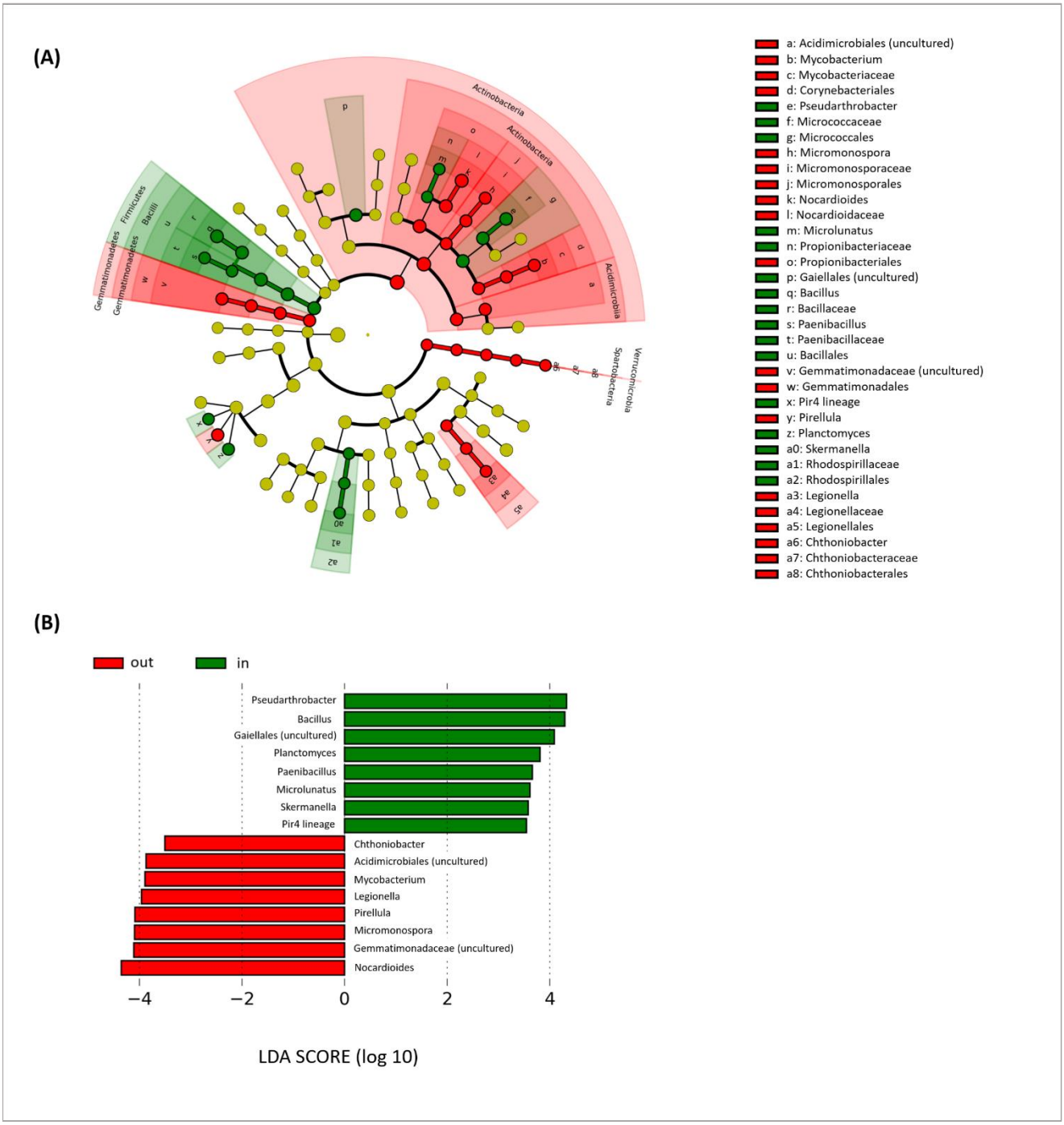


Fig. 4 – Rhizospheric microbiome signatures of PDO production sites. **(A)** Cladogram of microbial taxa differentially represented between farms located inside (Finale Emilia and Medolla, green) and outside (Bondeno, red) the PDO area at phylum to genus level. Only genera whose relative abundance was higher than 0.5% in at least 33% of the rhizospheric samples are represented. The diameter of each circle is proportional to the genera relative abundance within the entire rhizospheric cohort. **(B)** Linear discriminant analysis (LDA) scores of discriminating genera between the abovementioned groups (the logarithmic threshold for discriminative features was set to 2.0). Plots were obtained by LDA effect size (LEfSe) analysis.

Seasonal variations of the *Vitis vinifera* rhizospheric microbiome network across the Lambrusco DOC PDO region

To identify specificities of the rhizospheric microbiome structure associated with the Lambrusco DOC PDO region and seasonality, we established co-abundance associations of genera and then clustered correlated bacterial taxa into four co-abundance groups (CAGs), describing microbiome configurations across the entire dataset (**Supplementary Figure 3**). The dominant (i.e., the most abundant) genera in these CAGs were *Pirellula* (red), *Nocardioides* (blue), *Pseudomonas* (pink), and *Bacillus* (green). The network-establishing CAGs relationships are named Wiggum plots, where genera abundances are represented as a circle proportional to the genus normalized over-abundance (**Figure 5**). The microbiome variation from Bondeno to Finale Emilia and Medolla through the two different seasons was accompanied by distinctive CAGs dominance, and most relevantly by abundances of the *Pirellula* and *Nocardioides* CAGs (Bondeno), the *Pseudomonas* CAG (Finale Emilia) and the *Bacillus* CAG (Medolla). When we sought for shared network topological features among Finale Emilia and Medolla microbiome structure, distinctive of the PDO region and not included in the control site (Bondeno), we found that nodes corresponding to *Bacillus* and *Rhizobium* were over-abundant during spring, whereas *Pseudarthrobacter* and *Microlunatus* nodes were over-abundant in the fall season. When combined, such results underline a sort of seasonal dynamic, very peculiar to the Lambrusco DOC PDO region independently of the type of management. Remarkably, most of these taxa constitute a subgroup of the species previously identified by LEfSe.

Understanding the importance of PDO-related taxa for grapevine biology

Plant growth-promoting microorganisms regulate plant physiological reactions and foster plant growth with several mechanisms. Here, we sought for some of these functions within the reference genomes of the taxa revealed by LEfSe. Specifically, we first used oligotyping[29] to identify the bacterial species (or higher taxonomic levels in some cases, as explained above) nested by the ASVs sequences belonging to the genera identified by LEfSe. In particular, ASVs sequences coding for *Bacillus*, *Pseudarthrobacter*, *Planctomyces*, *Paenibacillus*, *Microlunatus*, *Skermanella*, *Pir4 lineage*, and uncultured members of Gaiellales (i.e., the PDO-related taxa identified by LEfSe) along with ASVs sequences coding for *Chthoniobacter*, *Nocardioides*, *Micromonospora*, *Pirellula*, *Legionella*, *Mycobacterium*, and uncultured members of Acidimicrobiales and *Gemmatimonadaceae* (i.e., the non-PDO taxa identified by LEfSe) were processed using the "Minimum Entropy Decomposition" (MED) module and the global earth microbiomes (GEM) catalogue[30]. We found that *Bacillus korlensis*, *Bacillus mediterraneensis*, *Bacillus tuaregi*, *Bacillus niacini*,

Bacillus jeotgali, *Bacillus lonarensis*, and *Bacillus litoralis*, unclassified species of the genus *Bacillus*, species of the genus *Pseudarthrobacter*, species of the genus *Planctomyces*, unclassified species of the order Gaiellales, *Azospirillum brasilense*, *Azospirillum thiophilum*, species of the Pirellulales order, *Microlunatus phosphovorus*, *Paenibacillus castaneae*, *Paenibacillus harenae*, *Paenibacillus ferrarius*, *Paenibacillus beijingensis*, and *Paenibacillus uliginis*, and unclassified species of the genus *Paenibacillus*, were the taxa characterizing the PDO-area.

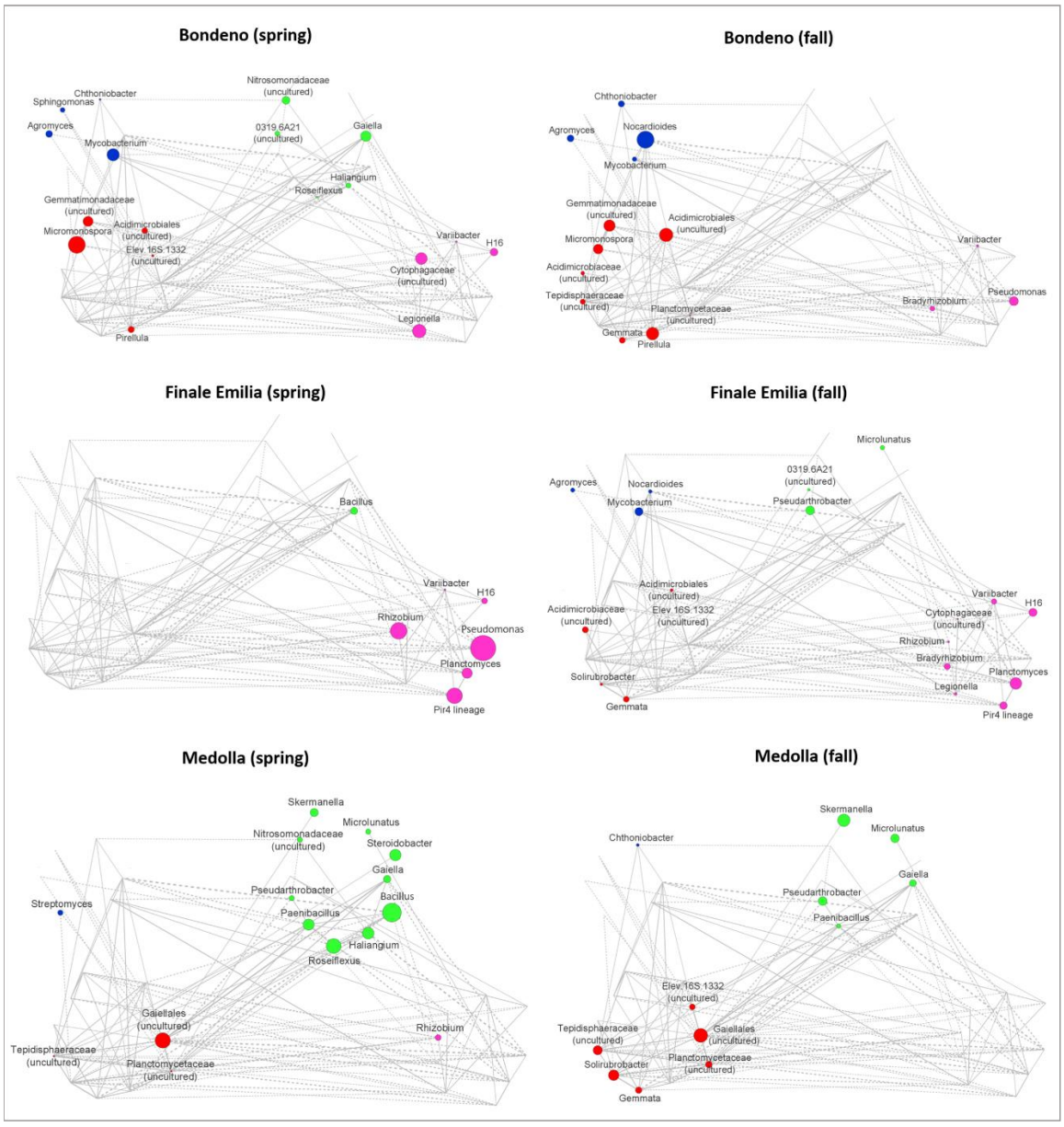


Fig. 5 – Declination of rhizospheric *Vitis vinifera* co-abundance groups according to the sampling site and to seasonality. Co-abundance groups (CAGs) are named according to the dominant bacterial genus in each group: *Pirellula* (red), *Nocardioides* (blue), *Pseudomonas* (pink), and *Bacillus* (green). Each node represents a bacterial genus and the size of the corresponding circle is proportional to its over-abundance on the average value within the population. The connections between nodes constitute positive (solid lines) and negative (dashed lines) Kendall correlations between genera (FDR ≤ 0.05). For CAGs definition see **Supplementary Figure S3**.

Notably, when applying oligotyping and GEM database, the ASVs sequences previously assigned to *Skermanella* were assigned to *A. brasilense* and *A. thiophilum*. We chose to retain both *Skermanella* and *Azospirillum* genomes for the following analysis, also because of the high level of overlapping found between the 16S rRNA sequences of these two taxa[37]. On the other hand, *Nocardioides massiliensis*, *Nocardioides allogilvus*, *Nocardioides exalbidus*, *Nocardioides halotolerans*, and *Nocardioides szechwanensis*, unclassified species of the genus *Nocardioides*, *Micromonospora cremea*, *Micromonospora marina*, *Micromonospora nigra*, *Micromonospora sediminis*, *Gemmatirosa kalamazoonesis*, *Pirellula staleyi*, *Legionella fallonii*, *Legionella saoudiensis*, *Mycolicibacterium moriokaense*, and *Mycolicibacterium sphagni*, unclassified species of the order Acidimicrobiales and *Chthoniobacter flavus* were the taxa most distinguishing the non-PDO area. Interestingly, the oligotyping procedure and the GEM database identified *Mycolicibacterium* species nested in the ASVs belonging to the genus *Mycobacterium*. In this regard, a recent comprehensive phylogenomic study by Gupta et al. (2018)[38] revealed that *Mycolicibacterium* can be actually regarded as a distinct clade previously classified as *Mycobacterium* and now forming a novel microbial genus. Then, in order to investigate the presence of potential PGP traits related to all these microorganisms, the NCBI reference genomes of all of these taxa were scanned for genes associated with nitrogen fixation (essential for plant growth), phosphorous solubilization (important for plant P uptake), siderophore production (for growth in iron-limiting conditions), indole-3-acetic acid (IAA) phytohormonal secretion (beneficial to increase water and nutrient absorption), and 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (for ethylene precursor degradation and regulation of plant stress response). Among the taxa identified by LEfSe as PDO-characterizing, species belonging to the genera *Bacillus*, *Skermanella*/*Azospirillum*, *Paenibacillus*, and unclassified species of the order Pirellulales contained most of the PGP features, whereas species from the genera *Pseudarthrobacter*, *Planctomyces*, and *Microlunatus*, together with unclassified members of the order Gaiellales, contained only one or two out of the five investigated PGP traits (**Figure 6A**). Conversely, if we look at the microbial taxa related to the non-PDO area (**Figure 6B**), species of the genus *Nocardioides* (i.e., unclassified *Nocardioides* and *N. exalbidus*) are the only ones in which more than two PGP traits out of five have been detected. Furthermore, two important PGP functions, namely nitrogen fixation and IAA production, have been scarcely observed in non-PDO related taxa (with the first only detected in species of unclassified *Nocardioides* while the latter entirely absent in non-PDO related taxa). PICRUSt2 confirmed most of the findings[31], with some exceptions, above all for what concerns siderophore production and ACC deaminase production (**Supplementary Table 5**). This can be attributed at least in part to the fact that two markers used in our analysis and necessary for predicting the functionalities are absent in the databases provided with PICRUSTt2.

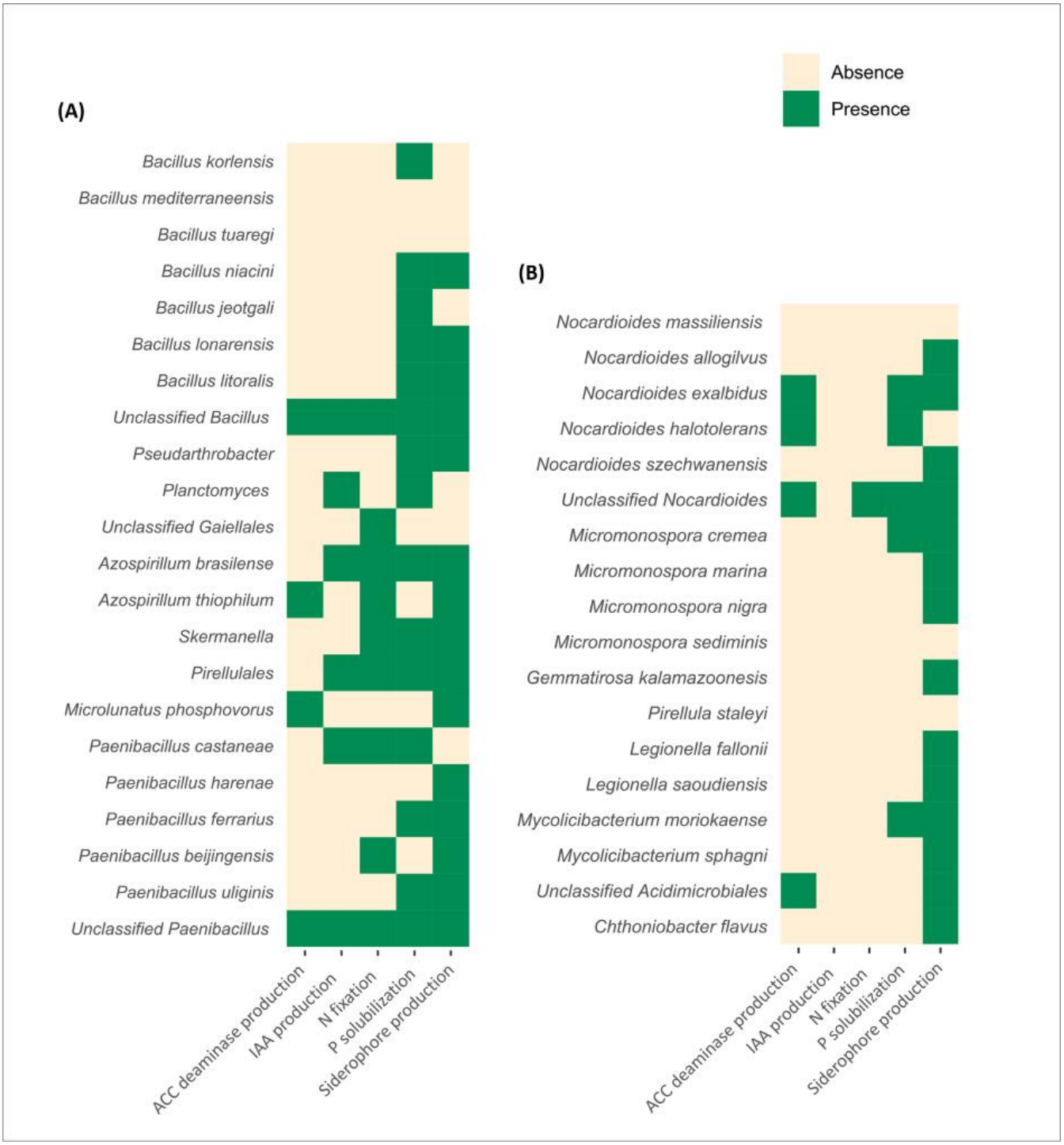


Fig. 6 – Schematic illustration showing the potential presence of PGP traits within the genomes of PDO-related **(A)** and non-PDO related **(B)** rhizospheric species. Each taxon was tested for the presence/absence of a specific set of PGP functions. The selected functions were nitrogen fixation, phosphorous solubilization, iron chelation, production of indole-3-acetic acid (IAA), and production of 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase. Green squares for putative presence of PGP activities, ivory squares for absence of PGP activity.

Discussion

This study aimed at characterizing the soil-plant microbiome dimension of the viticultural *terroir* of *V. vinifera* cultivar Lambrusco from the "Consorzio Tutela Lambrusco DOC" PDO area, in Emilia Romagna, Italy. This was made possible by comparing microbiomes from viticultural farms – of the same age – located immediately inside or outside the PDO area, thus controlling for complex variations associated with differences in pedoclimatic regions. The amplicon sequencing revealed a clear differentiation of the soil and rhizospheric microbiomes according to the sampling location, but not to the sampling season, and the bulk soil microbial diversity was higher within the PDO area rather than outside. When we applied the LEfSe analysis, we detected eight bacterial genera significantly differentially recruited by the plants grown inside the PDO consortium with respect to the non-PDO area. Specifically, the eight genera, and where possible the nested species, included *Bacillus* (*B. korlensis*, *B. mediterraneensis*, *B. tuaregi*, *B. niacini*, *B. jeotgali*, *B. lonarensis*, *B. litoralis*, and other unclassified species), *Pseudarthrobacter*, *Planctomyces*, *Paenibacillus* (*P. castaneae*, *P. harenae*, *P. ferrarius*, *P. beijingensis*, *P. uliginis*, and other unclassified species), *Microlunatus* (*M. phosphovorus*), and *Skermanella/Azospirillum* (*A. brasilense* and *A. thiophilum*), *Pir4 lineage* and uncultured members of the order Gaiellales. Stinkingly, such group of bacteria was also detected in the correspondent bulk soil samples of the same PDO areas, emphasizing the commonly accepted hypothesis that soil can function as rich microbial *reservoir* for those microorganisms that interact with the plant holobiont at the root level[39,40]. Even if not entailing significant variations in the microbiomes compositional structure, seasonality was shown to be associated with relevant changes in the rhizosphere microbiome network topology, with features characterizing the seasonal dynamics in the PDO area. Interestingly, most of the network seasonal variations related to the PDO area involved PDO-related taxa, which seem to modulate their abundance in response to seasonality. All the PDO-related species identified by LEfSe (except for *B. mediterraneensis* and *B. tuaregi*) presented at least one of the PGP traits potentially involved in the biostimulation and biofertilization of grapevine, with some species combining multiple PGP traits, such as *Bacillus*, *Skermanella*/*Azospirillum*, and *Paenibacillus*, possibly exerting a multifactorial probiotic role for the plant growth and biology. In particular, the PGP features detected in PDO-related species included the abilities to produce ACC deaminase, IAA, and siderophores, of solubilizing phosphorous from soil particles and soil organic matter and of biofertilizing soil through nitrogen fixation, which are all features that play an important role in microbiome-root crosstalk and plant growth/adaptation[41]. With such a specific microbiome configuration, PDO-related bacteria may induce modification of the root system architecture as previously demonstrated[42], and thus enhance nutrients and water uptake by the grapevines, with a resulting higher resistance to environmental stresses, better plant health and, consequently, improvement of the organoleptic properties of the Lambrusco wine, probably contributing to the regional *terroir*[13,43,44]. Conversely, non-PDO-related species show far less PGP traits. In this regard, the only widespread function identified is connected to the production of some well-known siderophores. Referring to the available literature on the PDO-related bacteria, we noticed that *Bacillus* is widely found on the root of grapevines in several different studies[18,45,46,47]. In particular, the higher abundance of *Bacillus* in the PDO area is quite interesting since *Bacillus* is a well-known plant growth-promoting rhizobacterium which can have many beneficial effects on plant growth[48]. These include, for instance: improvement of iron acquisition[49], regulation of the Na+/K+ efflux[50], and modulation of plant physiology by IAA

production[51]. Additionally, *Bacillus* is able to promote plant root length, photosynthetic pigment formation, and shoot germination through the production of the ACC deaminase enzyme, which also enhances tolerance to salinity stress[52]. Specifically, for grapevine plantlets, it has been shown that *Bacillus* can upregulate melatonin synthesis and reduce the production of malondialdehyde and reactive oxygen species in salt and drought stress conditions[53]. Further, when we sought for plant growth-promoting features of *Paenibacillus*, we found that it can be important for enhancing drought tolerance by upregulating dehydration-responsive genes, RD29A and RD29B[54], and for improving root surface area, root projection area and root fork numbers by IAA production, nitrogen fixation, and phosphorous solubilization[55,56]. Functional genes related to plant growth-promoting activity were also previously identified in *Pseudarthrobacter*, that is an aerobic auxin-producing bacterium[57], *Azospirillum*, a noteworthy diazotrophic microorganism which stimulates plant growth in different ways, e.g., by enhancing roots development and lateral root formation by IAA production[58,59] and *M. phosphovorus* that has been reported as phosphorous accumulator in wastewater treatment plants[60]. Finally, our data clearly show that all the above-mentioned PDO microbiome specificities are limited to the soil and rhizospheric microbiome ecosystems, while the corresponding root microbiomes, possibly under a strong host-driven selection pressure[18,61,62], remain constant in the three different farms, independently of the PDO or non-DOP location. Collectively, all the taxa we found characterizing the PDO area are commonly detected in grapevine rhizospheres[10,13,18,45,46,47,63]. However, here their concomitant presence at high abundance, their network structure and their characteristic seasonal dynamics may represent a key feature of the "Consorzio Tutela Lambrusco DOC" microbial *terroir*, possibly contributing to the peculiarity of the regional wine product, generally supporting the strategic importance of the soil-plant microbiome interface in defining microbiome-associated *terroir* specificities of relevance for the overall product quality. Future studies on higher number of sites within and outside the PDO area, based on shotgun metagenomics and possibly providing for a more extensive sampling, are needed to better unravel the contribution of the root-associated microbiomes, as well as of specific PGP species and/or strains, to the specific regional characteristics of grapevines and associated local products. Finally, examining in depth the link between root microbiome and grapevine may also provide helpful information for vineyard management, productivity and precision oenology, as well as elements to be safeguarded as pivotal features of the microbial *terroir* of Lambrusco grapevine, especially in the context of the current global change scenario, where we are witnessing a continuous loss of microbial diversity in several ecosystems, including soil.

Data availability statement

The data presented in the study are deposited in the ENA archive (https://www.ebi.ac.uk/ena/), accession number PRJEB57815.

Author contributions

MC and SR designed research project and acquired funds. SR, GA, CT, and LI performed sampling activity. EN, DS, GP, and GT performed laboratory experiments. MC contributed to analytical tools and laboratory equipment. EN, SR, AC, MF, and DS analyzed the data. EN and SR wrote the original draft. GP, DS, GT, NC, MF, AC, GA, CT, LI, and MC reviewed and edited the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the "Controlling Microbiomes Circulations for Better Food Systems" (CIRCLES) project, which was funded by the European Union's Horizon 2020 research and innovation program under grant agreement no. 818290.

Acknowledgments

We would like to thank the Istituto Tecnico Statale "Ignazio Calvi" (Finale Emilia, Italy) for their primary contribution to the sample collection for the present study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found in the appendix of the thesis (Study I) and online at https://www.frontiersin.org/articles/10.3389/fmicb.2023.1108036/full#supplementary-material.

References

[1]Gamalero, Elisa, et al. "Discovering the bacteriome of Vitis vinifera cv. Pinot Noir in a conventionally managed vineyard." *Scientific Reports* 10.1 (2020): 6453.

[2]Martinho, Vítor João Pereira Domingues. "Contributions from literature for understanding wine marketing." *Sustainability* 13.13 (2021): 7468.

[3]Pinto, Cátia, et al. "Unravelling the diversity of grapevine microbiome." *PLoS One* 9.1 (2014): e85622.

[4]Smyczek, Sławomir, et al. "Economic sustainability of wine tourism services and direct sales performance–emergent profiles from Italy." *British Food Journal* 122.5 (2020): 1519-1529.

[5]Savelli, Elisabetta, et al. "PDO labels and food preferences: results from a sensory analysis." *British Food Journal* 123.3 (2021): 1170-1189.

[6]Spognardi, Sara, et al. "Impact of organic and "protected designation of origin" labels in the perception of olive oil sensory quality." *British Food Journal* 123.8 (2021): 2641-2669.

[7]Van Leeuwen, Cornelis, and Gerard Seguin. "The concept of terroir in viticulture." *Journal of wine research* 17.1 (2006): 1-10.

[8]Gilbert, Jack A., Daniel van der Lelie, and Iratxe Zarraonaindia. "Microbial terroir for wine grapes." *Proceedings of the National Academy of Sciences* 111.1 (2014): 5-6.

[9]Bokulich, Nicholas A., et al. "Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics." *MBio* 7.3 (2016): 10-1128.

[10]Novello, Giorgia, et al. "The rhizosphere bacterial microbiota of Vitis vinifera cv. Pinot Noir in an integrated pest management vineyard." *Frontiers in microbiology* 8 (2017): 1528.

[11]Vitulo, Nicola, et al. "Bark and grape microbiome of Vitis vinifera: influence of geographic patterns and agronomic management on bacterial diversity." *Frontiers in Microbiology* 9 (2019): 3203.

[12]Mezzasalma, V., Sandionigi, A., Bruni, I., Bruno, A., Lovicu, G., Casiraghi, M., et al. (2017). Grape microbiome as a reliable and persistent signature of field origin and environmental conditions in Cannonau wine production. *PLoS One* 12:e0184615.

[13]Zarraonaindia, Iratxe, et al. "The soil microbiome influences grapevine-associated microbiota." *MBio* 6.2 (2015): 10-1128.

[14]Coller, Emanuela, et al. "Microbiome of vineyard soils is shaped by geography and management." *Microbiome* 7 (2019): 1-15.

[15]Oyuela Aguilar, Mónica, et al. "Influence of vintage, geographic location and cultivar on the structure of microbial communities associated with the grapevine rhizosphere in vineyards of San Juan Province, Argentina." *PLoS One* 15.12 (2020): e0243848.

[16]Biget, Marine, et al. "The drivers of vine-plant root microbiota endosphere composition include both abiotic and plant-specific factors." *Oeno One* 55.3 (2021): 299-315.

[17]Gobbi, Alex, et al. "A global microbiome survey of vineyard soils highlights the microbial dimension of viticultural terroirs." *Communications Biology* 5.1 (2022): 241.

[18]D'Amico, Federica, et al. "The rootstock regulates microbiome diversity in root and rhizosphere compartments of Vitis vinifera cultivar Lambrusco." *Frontiers in microbiology* 9 (2018): 2240.

[19]Turroni, Silvia, et al. "Fecal metabolome of the Hadza hunter-gatherers: a host-microbiome integrative view." *Scientific reports* 6.1 (2016): 32826.

[20]Masella, Andre P., et al. "PANDAseq: paired-end assembler for illumina sequences." *BMC bioinformatics* 13 (2012): 1-7.

[21]Bolyen, Evan, et al. "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." *Nature biotechnology* 37.8 (2019): 852-857.

[22]Edgar, Robert C. "Search and clustering orders of magnitude faster than BLAST." *Bioinformatics* 26.19 (2010): 2460-2461.

[23]Callahan, Benjamin J., et al. "DADA2: High-resolution sample inference from Illumina amplicon data." *Nature methods* 13.7 (2016): 581-583.

[24]Rognes, Torbjørn, et al. "VSEARCH: a versatile open source tool for metagenomics." *PeerJ* 4 (2016): e2584.

[25]Quast, Christian, et al. "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools." *Nucleic acids research* 41.D1 (2012): D590-D596.

[26]Culhane, Aedín C., et al. "MADE4: an R package for multivariate analysis of gene expression data." *Bioinformatics* 21.11 (2005): 2789-2790.

[27]Martinez Arbizu, Pedro. "pairwiseAdonis: Pairwise multilevel comparison using adonis." *R package version 0.4* 1 (2020).

[28]Segata, Nicola, et al. "Metagenomic biomarker discovery and explanation." *Genome biology* 12 (2011): 1-18.

[29]Eren, A. Murat, et al. "Oligotyping analysis of the human oral microbiome." *Proceedings of the National Academy of Sciences* 111.28 (2014): E2875-E2884.

[30]Nayfach, Stephen, et al. "A genomic catalog of Earth's microbiomes." *Nature biotechnology* 39.4 (2021): 499-509.

[31]Douglas, Gavin M., et al. "PICRUSt2 for prediction of metagenome functions." *Nature biotechnology* 38.6 (2020): 685-688.

[32]Wixon, Joanne, and Douglas Kell. "The Kyoto encyclopedia of genes and genomes--KEGG." *Yeast (Chichester, England)* 17.1 (2000): 48-55.

[33]Schnorr, Stephanie L., et al. "Gut microbiome of the Hadza hunter-gatherers." *Nature communications* 5.1 (2014): 3654.

[34]Dabney, Alan, John D. Storey, and G. R. Warnes. "qvalue: Q-value estimation for false discovery rate control." *R package version* 1.0 (2010).

[35]Anderson, Marti J. "A new method for non-parametric multivariate analysis of variance." *Austral ecology* 26.1 (2001): 32-46.

[36]Claesson, Marcus J., et al. "Gut microbiota composition correlates with diet and health in the elderly." *Nature* 488.7410 (2012): 178-184.

[37]Zhu, Wentao, et al. "Genomic analysis of Skermanella stibiiresistens type strain SB22 T." *Standards in Genomic Sciences* 9 (2014): 1211-1220.

[38]Gupta, Radhey S., Brian Lo, and Jeen Son. "Phylogenomics and comparative genomic studies robustly support division of the genus Mycobacterium into an emended genus Mycobacterium and four novel genera." *Frontiers in microbiology* 9 (2018): 67.

[39]Singh, Monika, et al. "Plant growth promoting rhizobacteria: application in biofertilizers and biocontrol of phytopathogens." *PGPR amelioration in sustainable agriculture*. Woodhead Publishing, 2019. 41-66.

[40]Ling, Ning, Tingting Wang, and Yakov Kuzyakov. "Rhizosphere bacteriome structure and functions." *Nature communications* 13.1 (2022): 836.

[41]Qu, Qian, et al. "Rhizosphere microbiome assembly and its impact on plant growth." *Journal of agricultural and food chemistry* 68.18 (2020): 5024-5038.

[42]Baldan, Enrico, et al. "Beneficial bacteria isolated from grapevine inner tissues shape Arabidopsis thaliana roots." *PLoS One* 10.10 (2015): e0140252.

[43]Rolli, Eleonora, et al. "Improved plant resistance to drought is promoted by the root-associated microbiome as a water stress-dependent trait." *Environmental microbiology* 17.2 (2015): 316-331.

[44]Rolli, Eleonora, et al. "Root-associated bacteria promote grapevine growth: from the laboratory to the field." *Plant and Soil* 410 (2017): 369-382.

[45]Longa, Claudia MO, et al. "Soil microbiota respond to green manure in organic vineyards." *Journal of applied microbiology* 123.6 (2017): 1547-1560.

[46]Marasco, Ramona, et al. "Grapevine rootstocks shape underground bacterial microbiome and networking but not potential functionality." *Microbiome* 6 (2018): 1-17.

[47]Berlanas, Carmen, et al. "The fungal and bacterial rhizosphere microbiome associated with grapevine rootstock genotypes in mature and young vineyards." *Frontiers in Microbiology* 10 (2019): 1142.

[48]Hashem, Abeer, Baby Tabassum, and Elsayed Fathi Abd_Allah. "Bacillus subtilis: A plant-growth promoting rhizobacterium that also impacts biotic stress." *Saudi journal of biological sciences* 26.6 (2019): 1291-1297.

[49]Xu, Xihui, et al. "Modeling microbial communities from atrazine contaminated soils promotes the development of biostimulation solutions." *The ISME journal* 13.2 (2019): 494-508.

[50]Tahir, Muhammad, et al. "Regulation of antioxidant production, ion uptake and productivity in potato (Solanum tuberosum L.) plant inoculated with growth promoting salt tolerant Bacillus strains." *Ecotoxicology and Environmental Safety* 178 (2019): 33-42.

[51]Ansari, Firoz Ahmad, Iqbal Ahmad, and John Pichtel. "Growth stimulation and alleviation of salinity stress to wheat by the biofilm forming Bacillus pumilus strain FAB10." *Applied Soil Ecology* 143 (2019): 45-54.

[52]Din, Bashir Ud, et al. "Mechanistic elucidation of germination potential and growth of wheat inoculated with exopolysaccharide and ACC-deaminase producing Bacillus strains under induced salinity stress." *Ecotoxicology and environmental safety* 183 (2019): 109466.

[53]Jiao, Jian, et al. "Melatonin-producing endophytic bacteria from grapevine roots promote the abiotic stress-induced production of endogenous melatonin in their hosts." *Frontiers in Plant Science* 7 (2016): 1387.

[54]Liu, Wenshan, Edward Sikora, and Sang-Wook Park. "Plant growth-promoting rhizobacterium, Paenibacillus polymyxa CR1, upregulates dehydration-responsive genes, RD29A and RD29B, during priming drought tolerance in arabidopsis." *Plant Physiology and Biochemistry* 156 (2020): 146-154.

[55]Navarro-Noya, Yendi E., et al. "Isolation and characterization of nitrogen fixing heterotrophic bacteria from the rhizosphere of pioneer plants growing on mine tailings." *Applied Soil Ecology* 62 (2012): 52-60.

[56]Li, Hui-Ping, et al. "Newly isolated Paenibacillus monticola sp. nov., a novel plant growth-promoting rhizobacteria strain from high-altitude spruce forests in the Qilian Mountains, China." *Frontiers in Microbiology* 13 (2022): 833313.

[57]Park, Min-Kyu, et al. "Complete genome sequence of a plant growth-promoting bacterium Pseudarthrobacter sp. NIBRBAC000502772, isolated from shooting range soil in the Republic of Korea." *The Microbiological Society of Korea* 56.4 (2020): 390-393.

[58]Arzanesh, Mohammad Hossein, et al. "Wheat (Triticum aestivum L.) growth enhancement by Azospirillum sp. under drought stress." *World Journal of Microbiology and Biotechnology* 27 (2011): 197-205.

[59]Fukami, Josiane, Paula Cerezini, and Mariangela Hungria. "Azospirillum: benefits that go far beyond biological nitrogen fixation." *Amb Express* 8.1 (2018): 73.

[60]Dorofeev, A. G., et al. "Role of phosphate-accumulating bacteria in biological phosphorus removal from wastewater." *Applied biochemistry and microbiology* 56 (2020): 1-14.

[61]Liu, Hongwei, et al. "Inner plant values: diversity, colonization and benefits from endophytic bacteria." *Frontiers in microbiology* 8 (2017): 2552.

[62]Afzal, Imran, et al. "Plant beneficial endophytic bacteria: Mechanisms, diversity, host range and genetic determinants." *Microbiological research* 221 (2019): 36-49.

[63]Bettenfeld, Pauline, et al. "The microbiota of the grapevine holobiont: a key component of plant health." *Journal of advanced research* 40 (2022): 1-15.

1.3.3 Study II – Zonation of the *Vitis vinifera* microbiome in Vino Nobile di Montepulciano PDO production area

This section of the thesis was published as an Original Research Article in *Communications Biology* (2024); https://doi.org/10.1038/s42003-024-07261-8. Contents of the article have not been modified except for graphical purposes and bibliography format. Numeration of main and supplementary figures and tables is consistent with the original article.

Authors:

Giorgia Palladino[1,2], Enrico Nanetti[1], Daniel Scicchitano[1,2], Nicolò Cinti[1,2], Lucia Foresto[1,2], Alice Cozzi[1], Antonio Gonzalez Vara Rodriguez[3], Nicolò Interino[4], Jessica Fiori[4,5], Silvia Turroni[1], Marco Candela[1,2] and Simone Rampelli[1,2*]

*Corresponding Author

Affiliations:

[1]Unit of Microbiome Science and Biotechnology, Department of Pharmacy and Biotechnology (FaBiT), Alma Mater Studiorum – University of Bologna, Bologna, Italy

[2]Fano Marine Center, The Inter-Institute Center for Research on Marine Biodiversity, Resources and Biotechnologies, Fano, Italy

[3]Unit of Enzymology, Department of Pharmacy and Biotechnology, University of Bologna, viale Risorgimento 4, 40136 Bologna, Italy

[4]Laboratorio di Proteomica Metabolomica e Chimica Bioanalitica, IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy

[5]Department of Chemistry "G. Ciamician", University of Bologna, Via Selmi 2, 40126 Bologna, Italy

Abstract

The microbial dimension of the terroir is crucial for wine quality, as microbiomes contribute to plant biofertilization, stress tolerance, and pathogen suppression. While microbial terroir can act as a biological signature at large scale, data for local contexts is lacking, hindering the characterization of regional microbial diversity in vineyards. Here, we define the microbial terroir of vineyards across the 12 sub-areas (Additional Geographic Units - AGUs) of the "Consorzio del Vino Nobile di Montepulciano DOCG" PDO area (Italy), a world-renowned wine-producing region. Rhizospheres of *Vitis vinifera* cultivar Sangiovese and soil samples were collected throughout the 2022 viticultural season and analyzed through an integrated metabarcoding/shotgun metagenomic approach, targeting bacteria and fungi. Wine metabolomics was also performed, projecting compositional and functional variations of the microbial terroir at the AGUs level into a corresponding variation in the product metabolic profile. Our findings reveal a unique taxonomic configuration of the Vino Nobile di Montepulciano terroir compared to other vineyards, with microbiomes being "AGU-specific" in taxonomic abundances and plant growth-promoting functions, confirming the potential relevance of characterizing and preserving the microbial terroir to safeguard high-quality traditional wines.

Highlights

- The territory of the "Consorzio del Vino Nobile di Montepulciano DOCG" has a characteristic microbial terroir, with specific declinations in the 12 different production areas (AGUs);

- 10 taxa have been identified as the core of the PDO "Consorzio del Vino Nobile di Montepulciano DOCG" microbial terroir, encoding several PGP features;

- Microbial terroir specificity at the AGU level provides specific PGP traits, such as the P solubilisation in the southern part of the production area and the ACC deaminase in the western part;

- These represent adaptive microbiome features to the local peculiarities of the territory, linked to the local variation of the product metabolic profile;

- The local diversity of the microbial terroir is a neglected part of the traditional terroir components that needs to be preserved and protected for the production of diverse and high-quality traditional wines.

Introduction

It is culturally common to associate wine with the place of production, with specific and recognizable characteristics, so much so that the place of origin is one of the main factors guiding wine purchase decisions[1]. The uniqueness of the relationship between wine and its territory of origin is defined by the concept of terroir, which includes local pedoclimatic, biotic and abiotic factors, combined with traditional agricultural practices, to explain the distinctive regional characteristics of the product[2] (International Organization of Vine and Wine, Definition of vitivinicultural "terroir" - https://www.oiv.int/public/medias/379/viti-2010-1-en.pdf). Today, the concept of wine terroir has spread throughout the world and is regulated by wine-producing countries through the legal definition of appellations of origin, such as the Protected Designation of Origin (PDO) in Europe. In Italy, wines made from identical grape cultivars but grown in different PDO areas with similar yields, are recognized as different products with different organoleptic characteristics[3,4,5]. Therefore, much is attributed to the components of wine terroir and, among them, to the vineyard microbiome communities, as possible and previously neglected new key determinants of terroir features that are associated with geographical location and are reported to be directly relevant to vine growing, grape quality, and winemaking[6,7]. Indeed, a reliable biological signature of the vineyard microbiome depending on the geographical location of the vineyard has recently been demonstrated[8], but little is known about its variations at finer local spatial scales[9], possibly matching different PDO areas, particularly in terms of the local diversity of plant growth-promoting (PGP) microorganisms as determinants of growth promotion, yield enhancement, and product quality[7]. In this context, we hypothesized that the interplay between bulk and rhizospheric soil microbiomes may represent an integral component of terroir, influencing nutrient uptake, and the overall terroir expression in defining the unique qualities of vineyards. Thus, the fine characterization of bulk and rhizospheric soil in the different PDO terroirs may provide important highlights on the relevance of local soil microbiome diversity in defining the distinct organoleptic characteristics of wines from specific regions[10]. To provide some insights in this direction, here we aimed to investigate possible differences in microbiome-dependent terroir characteristics (rhizospheric and bulk soil microbiomes) in plant samples of *Vitis vinifera* cultivar Sangiovese collected from 12 different sub-areas located within the "Consorzio del Vino Nobile di Montepulciano DOCG" PDO area, in Tuscany, Italy. In particular, Montepulciano and its territory are considered an excellence in the Italian food and wine context, with the "Vino Nobile di Montepulciano" renowned all over the world, with 7 million bottles sold and a production turnover of 65 million euros in 2022, for a total estimated value of around 1 billion euros, including the value of assets (https://www.ansa.it/canale_terraegusto/notizie/vino/2023/02/15/vino-nobile-montepulciano-distretto vale-1-mld-di-euro_14425b81-3f63-4d41-b29a-db1469fbed30.html). Montepulciano territory has recently been divided into 12 production areas (i.e., Additional Geographical Units - AGUs), called "Pievi", each of them showing different characteristics in terms of altitude, pedoclimatic characteristics, soil composition and chemistry (https://www.doctorwine.it/en/pot-pourri/miscellanea/the-nobile-revolution-pieve, last access February 2024). The possibility of subdividing the production area was also made possible by the fact that the wines exhibited different organoleptic profiles, which reflected the specific characteristics of the terroir. This paved the way for the characterization of the microbiome determinants of this territorial uniqueness. In particular, we proposed a finer

characterization of the microbial terroir within the 12 AGUs, in order to add a microbiome dimension to the terroir features, to better understand and thus safeguard the local diversity of Italian wine production. In addition to enriching our understanding of the importance of soil and root-associated microbiomes in defining wine terroir within the Vino Nobile di Montepulciano PDO area, this study may provide further economic incentives for agricultural and oenological practices that preserve regional microbial terroir and biodiversity.

Results

Microbial characteristics of viticultural terroirs of *V. vinifera* cultivar Sangiovese for the production of Vino Nobile di Montepulciano

A total of 336 root samples (rhizosphere) of *V. vinifera* cultivar Sangiovese and 56 bulk soil samples were collected from 14 different vineyards in the 12 AGUs in July, August, September, and October 2022 in Montepulciano (Tuscany), Italy (**Figure 1**).

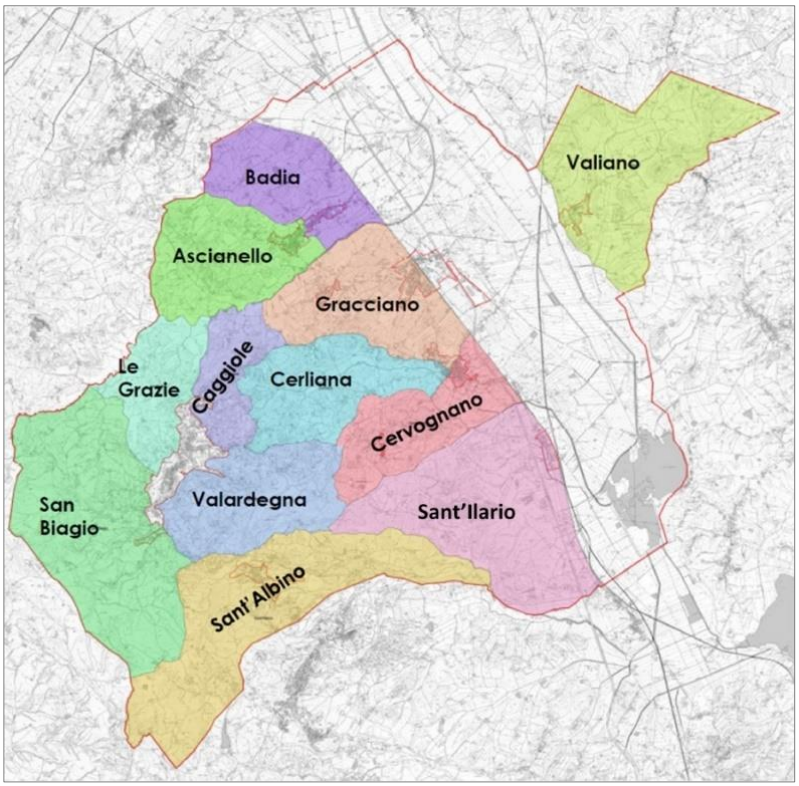


Fig. 1 – Map of the 12 production areas (i.e., Additional Geographical Units - AGUs) recognized by the "Consorzio del Vino Nobile di Montepulciano DOCG" (Tuscany, Italy). The production areas are indicated by different colors with the names in bold, Valiano, Valardegna, San Biagio, Sant'Albino, Le Grazie, Gracciano, Cervognano, Cerliana, Caggiole, Badia, Ascianello, and Sant'Ilario (map source: Consorzio del Vino Nobile di Montepulciano).

Specifically, for each vineyard, 6 rhizospheric samples and 1 bulk soil sample were retrieved at each timepoint. All selected vineyards were located within the PDO area. Information on sites and plant characteristics, rootstock families and management, as well as physical and chemical variables of the vineyards soils are provided, for each AGU, in **Supplementary Table 1** and **Supplementary Table 2**, respectively. For the 12 AGUs and the 4 time points, the composition of the soil and rhizosphere microbiomes was first investigated by next-generation sequencing of the bacterial 16S rRNA gene (V3-V4 hypervariable regions) and fungal ITS (internal transcribed spacer ITS2 region), with 332 (292 rhizosphere and 40 soil) and 64 (50 rhizosphere and 14 soil) samples successfully sequenced, respectively. This resulted in 3,654,656 high-quality reads, with an average of 11,008 ± 4,723 reads per sample (mean ± SD), for 16S rRNA gene sequencing data, and in 382,144 high-quality reads (5,971 ± 3,240) for ITS sequencing data. Reads were binned into 57,395 amplicon sequence variants (ASVs) for 16S rRNA gene sequencing and 740 ASVs for ITS sequencing.

In order to identify the soil microbiome peculiarities of the microbial terroir within the "Consorzio del Vino Nobile di Montepulciano DOCG", we compared its bacterial and fungal composition with bulk soils from vineyards from all over the world, including Chile, Argentina, USA, South Africa, Australia, Spain, France, Italy, Hungary, Portugal, Denmark, Germany, and Croatia[8] (**Figure 2**). We observed the effect of geographical distance on the composition and structure of soil microbial communities, both bacterial and fungal, with individual countries significantly segregating in the Principal Coordinates Analysis (PCoA) plots (permutation test with pseudo-F ratio, p-value ≤ 0.001) (**Figure 2A** and **B**). At the national scale, i.e., considering only bulk soil samples from Montepulciano and other Italian vineyards, we also observed a significant segregation of vineyards according to region of origin (p-value ≤ 0.001) (**Figure 2C** and **D**).

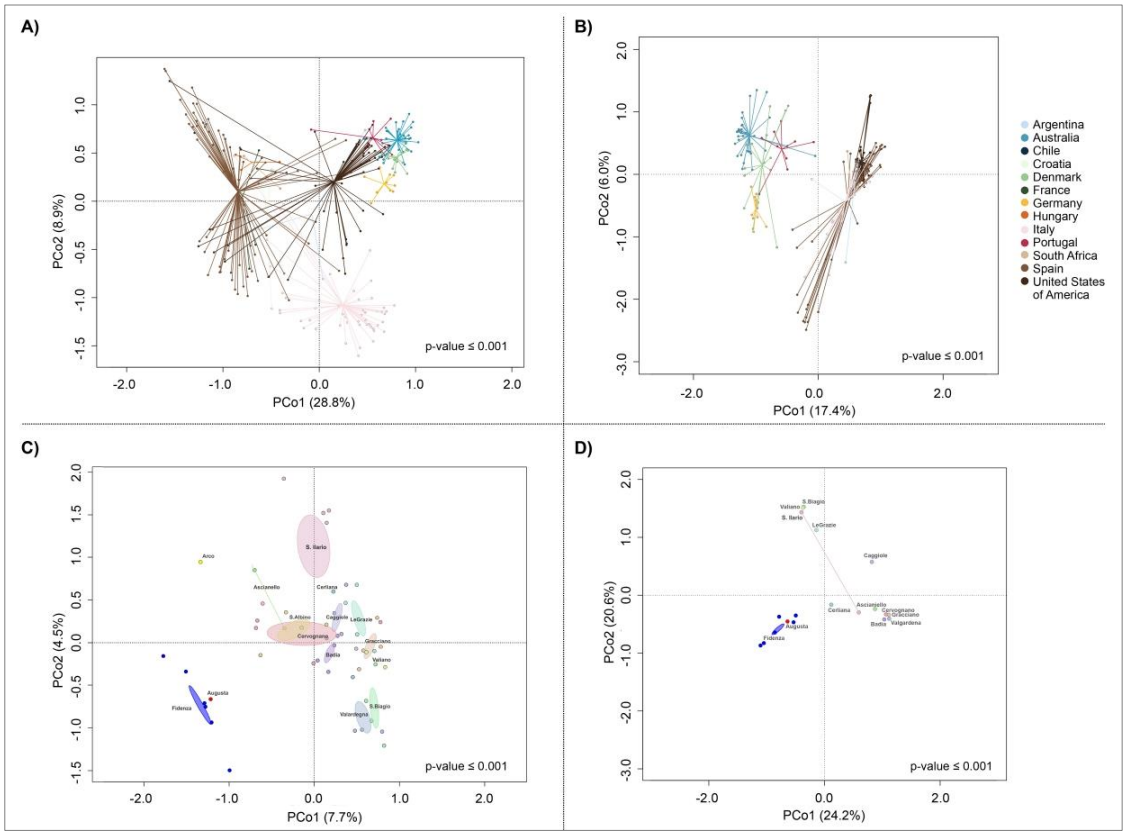


Fig. 2 – The bulk soil microbiome of Montepulciano vineyards shows a clear differentiation compared to other vineyards around the world. Comparisons were made for both 16S rRNA and ITS sequencing, using data from Gobbi et al. (2022)[8]. **(A)** Principal Coordinate Analysis (PCoA) based on unweighted UniFrac distances showing the variation of *Vitis vinifera* cultivar Sangiovese bulk soil bacterial composition at a wide geographical scale (worldwide), including Montepulciano samples in the Italian site (permutation test with pseudo-F ratio, p-value ≤ 0.001). **(B)** PCoA based on Bray-Curtis distances showing the variation of *V. vinifera* cultivar Sangiovese bulk soil fungal composition at a wide geographical scale (worldwide), including Montepulciano samples in the Italian site (p-value ≤ 0.001). **(C)** The same graph as in **(A)**, but at a finer geographical scale, including only Italian samples from Gobbi et al. (2022)[8] and Montepulciano samples (p-value ≤ 0.001). **(D)** The same graph as in **(B)**, but at a finer geographical scale, including only Italian samples from Gobbi et al. (2022)[8] and Montepulciano samples (p-value ≤ 0.001). For **(C)** and **(D)**, sample origin is indicated on each graph. For all PCoA plots, the first and second principal components are plotted and the percentage of variance in the dataset explained by each axis is shown.

When investigating the soil microbial taxa responsible for the geographical segregation, we identified 5 bacterial genera **(Figure 3A**) and 5 fungal genera (**Figure 3B**) whose variation in relative abundance was significantly different between Montepulciano and any other vineyard in Italy and worldwide (Kruskal-Wallis test controlled for multiple testing using False Discovery Rate – FDR, p-value ≤ 0.05). Specifically, for bacterial taxa, we found that the genera *Ilumatobacter*, *Microlunatus*, and *Hydrogenispora* were almost exclusively present in the Montepulciano consortium, while *Gemmata* and *Nocardioides*, widely distributed in the different soils, characterized the Montepulciano area in terms of relative abundance. As for fungal taxa, the genera *Rhizopus*, *Gongronella*, *Lipomyces*, and *Penicillium* were almost exclusively present in the Montepulciano consortium, while *Mortierella* characterized the Montepulciano soil in terms of relative abundance.

We then tried to define a core soil microbiome of the Vino Nobile di Montepulciano area, looking for taxa present in the bulk soil of all AGUs. We identified 5 microbial genera with this characteristic, namely *Nocardioides*, *Solirubrobacter*, *Gemmatimonas*, *Haliangium*, and *Pirellula*. Interestingly, *Nocardioides* was both a core taxon and a genus that distinguished the Montepulciano territory from vineyards in the rest of the world, and for this reason it could be considered the main marker characterizing the microbial terroir of Vino Nobile. Interestingly, these core genera were also present in all 12 AGUs when considering the rhizospheric soil, indicating a continuity between soil and rhizosphere in the Montepulciano territory. This continuity was further confirmed with a Procrustes correlation test using the protest function in R, comparing the beta diversity distribution of soil and rhizospehric samples and resulting in a significant correlation (p-value = $2*10^{-4}$ for bacterial community and p-value = 0.01 for fungal community).

<u>Spatial distance determines the similarity of microbial communities in vineyards at local scales across the Montepulciano territory</u>

Aware of the continuity between soil and rhizosphere microbiomes, as shown in the previous paragraph, we then aimed to identify the specificities of microbial terroir associated with the recent zonation in the 12 different AGUs of the Montepulciano territory, considering both bacterial and fungal counterparts.

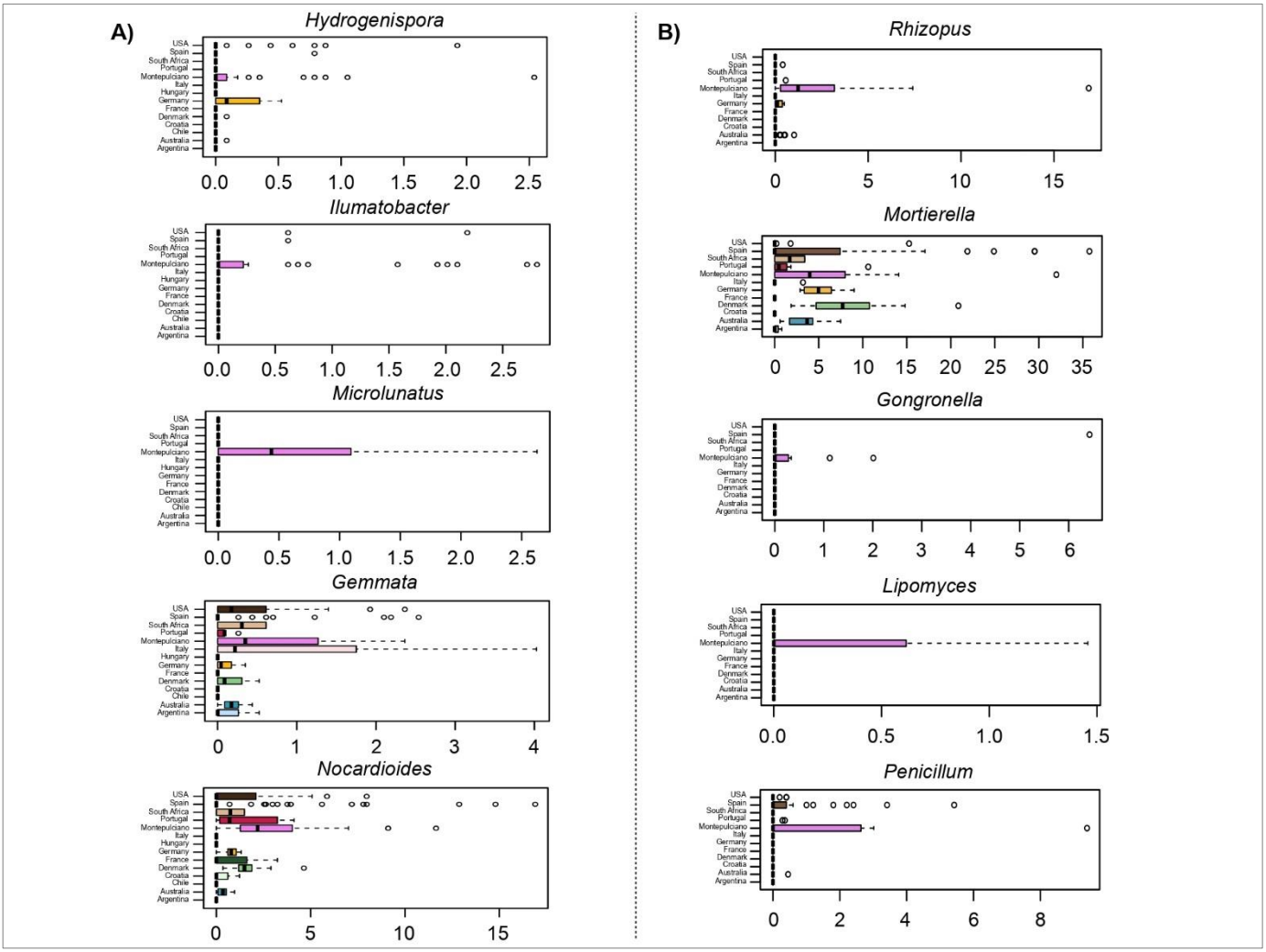


Fig. 3 – Microbial taxa distinguishing bulk soil samples of the Montepulciano territory from other vineyards worldwide. Boxplots showing the relative abundance distribution of bacterial **(A)** and fungal **(B)** genera differentially represented in bulk soil between Montepulciano and other vineyards worldwide (from Gobbi et al., 2022[8]) (Kruskal-Wallis test controlled for multiple testing using FDR with n=178 independent samples, p-value ≤ 0.05).

We found that the differences in the rhizosphere microbiome were explained by the geographical distance between the different AGUs, with the AGUs of Sant'Ilario (southeast of the territory) and San Biagio (west), located on opposite borders of the territory, having the most different bacterial and fungal configurations, and other AGUs having intermediate configurations between the two extremes. This segregation pattern was robust to seasonality, agronomical practices and management, vine clone type, rootstock family, altitude, and soil composition (permutation test with pseudo-F ratio, p-value ≤ 0.01) (**Figure 4**). Specifically, when comparing microbiomes across time points, we found the same segregation, as if the main factor driving microbiome differentiation was geographical origin at a very local scale (AGUs) rather than plant maturity and season (Procrustes test, p-value ≤ 0.01) (**Supplementary Figure 1**). In support of this evidence, we also found that rhizosphere microbiome separation in the PCoA correlated with geographical separation in terms of distance (in meters) between vineyards (p-value ≤

0.003). AGUs also showed a different alpha-diversity configuration among them (**Supplementary Figure 2**), however we did not observe a common pattern based on geographical distribution.

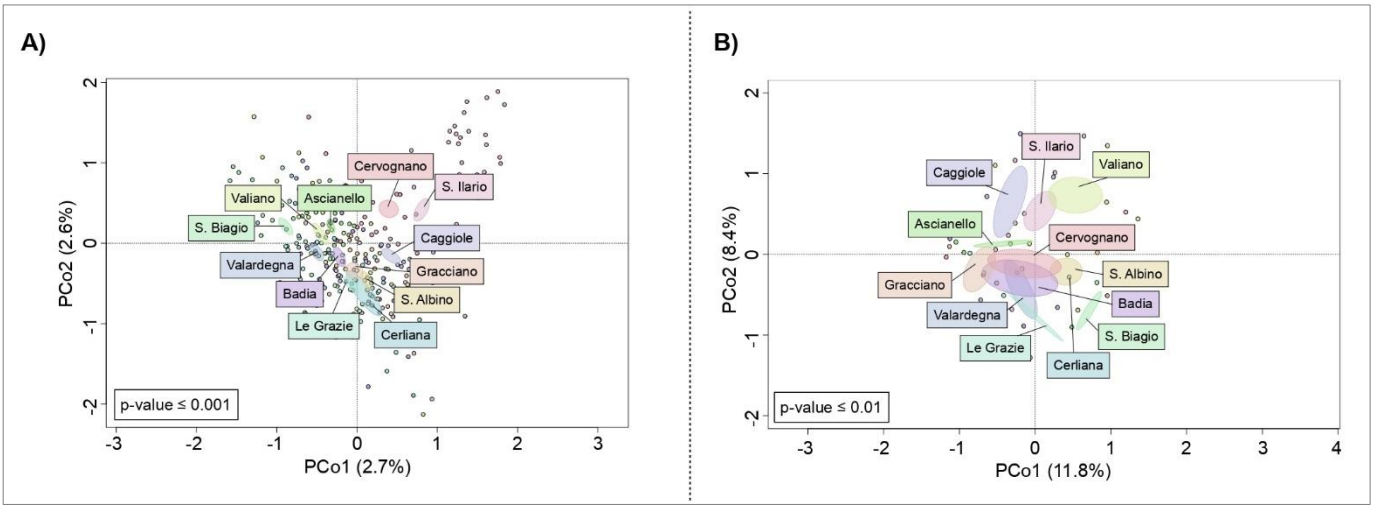


Fig. 4 – Spatial distance determines the dissimilarity of rhizospheric microbial communities in vineyards at a local scale in the Montepulciano territory. Principal coordinate analysis (PCoA) based on unweighted UniFrac distances between the bacterial **(A)** and fungal **(B)** profiles of *Vitis vinifera* cultivar Sangiovese rhizospheres in the different Additional Geographical Units (AGUs) of the Montepulciano territory. The first and second principal components are plotted and the percentage of variance in the dataset explained by each axis is shown. P-values are calculated with a permutation test with pseudo-F ratio, taking into account the contribution of seasonality, agronomical practices and management, vine clone type, rootstock family, altitude, and soil composition (p-value ≤ 0.01).

Random forest[11] was then used to identify rhizospheric bacterial and fungal genera that distinguished the 12 AGUs, and then combined with the Kruskal-Wallis test among relative taxon abundances in each AGU, to extract as much information as possible from our analysis. All significantly discriminating genera identified were represented as a heatmap using their relative abundance in each AGU (**Figure 5**). For the bacterial component of the rhizospheric soil, 24 genera were found to be discriminant among AGUs, 11 of which belonging to uncultured or unassigned genera. As for the fungal counterpart of the rhizospheric soil, 6 genera were identified as discriminating among AGUs. Such patterns reflected a sort of gradient describing the variation in relative abundance of these microorganisms along the Montepulciano territory, from Sant'Ilario to San Biagio and *vice versa*, crossing all other AGUs in an intermediate configuration between the two extremes. This was very clear when we superimposed the gradient of relative abundance of microorganisms on the map of the territory (**Figure 6**). Interestingly, these characteristics of the rhizosphere were confirmed at the level of soil microbiome (**Supplementary Figures 3** and **4**).

<u>Understanding the functional peculiarities of the microbial terroir in the Vino Nobile di Montepulciano PDO area</u>

We performed shotgun metagenomics on a subset of 28 samples, one bulk soil sample and one rhizosphere sample for each vineyard, representative of each Montepulciano AGU at the first time point, to obtain a more accurate picture of the pattern of variation of bacterial PGP functions across the Vino Nobile di Montepulciano PDO area. We retained ~390M high-quality reads, with an average of 14M ± 9M (mean ± SD) paired-end sequences per sample.

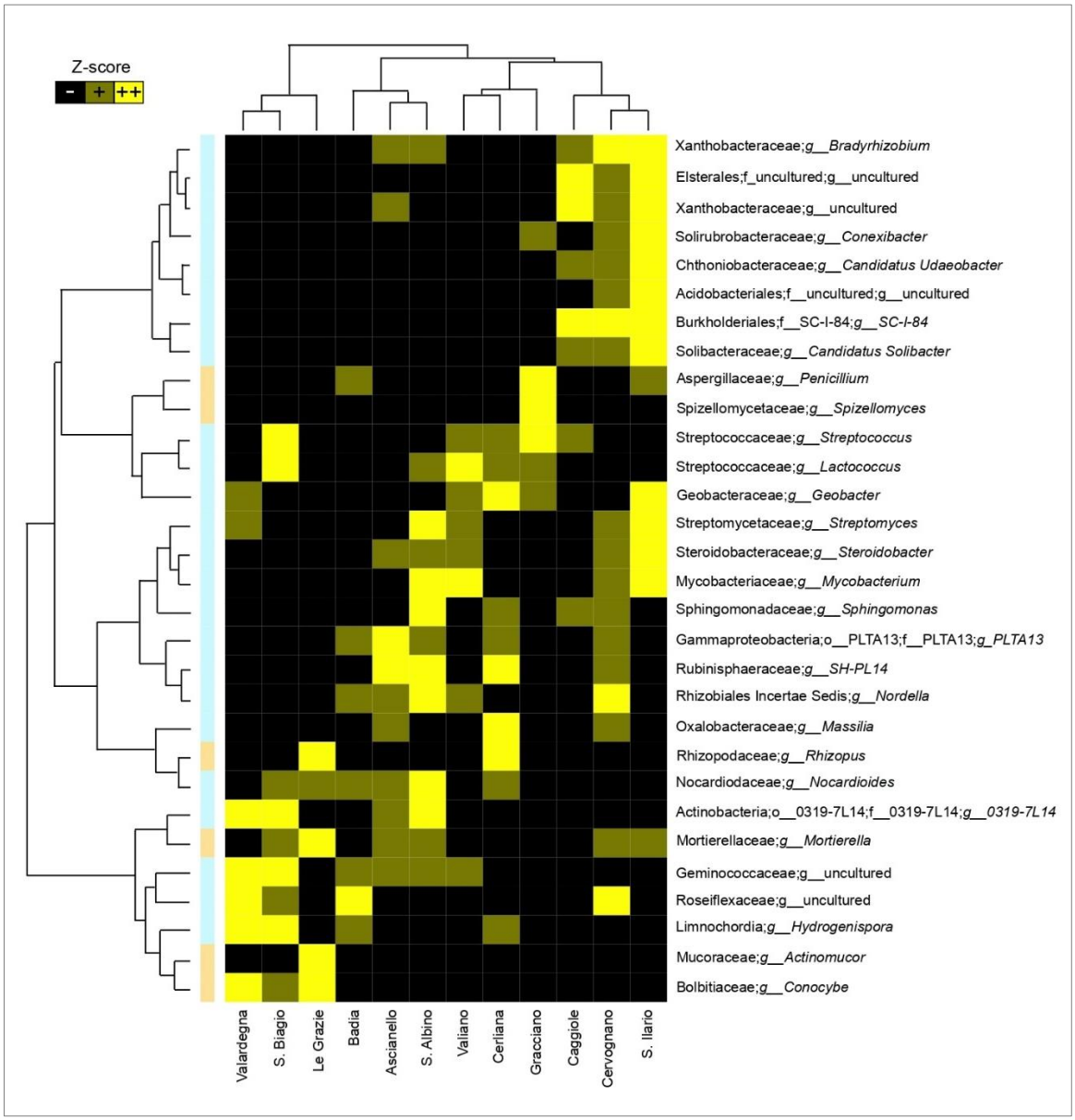


Fig. 5 – The AGU-related taxa show a relative abundance gradient in the Montepulciano area, describing the rhizospheric microbial variation across the 12 AGUs (Additional Geographical Units). Heatmap showing all significantly discriminating genera among the rhizosphere microbiomes of the 12 AGUs (Random Forest combined with the Kruskal-Wallis test among relative taxon abundances in each AGU, p-value ≤ 0.05). Relative taxon abundance is represented in the heatmap through z-score. The vertical bar is colored blue for bacteria and orange for fungi.

Reads were first aligned to known PGP genes[9] to screen for potential microbial ability to support plant growth at the soil-root interface. The microbial PGP functions selected for our analysis were: (i) nitrogen (N) fixation; (ii) phosphorous (P) solubilization; (iii) iron chelation; (iv) production of the phytohormone indole-3-acetic acid (IAA); and (v) production of the enzyme 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase.

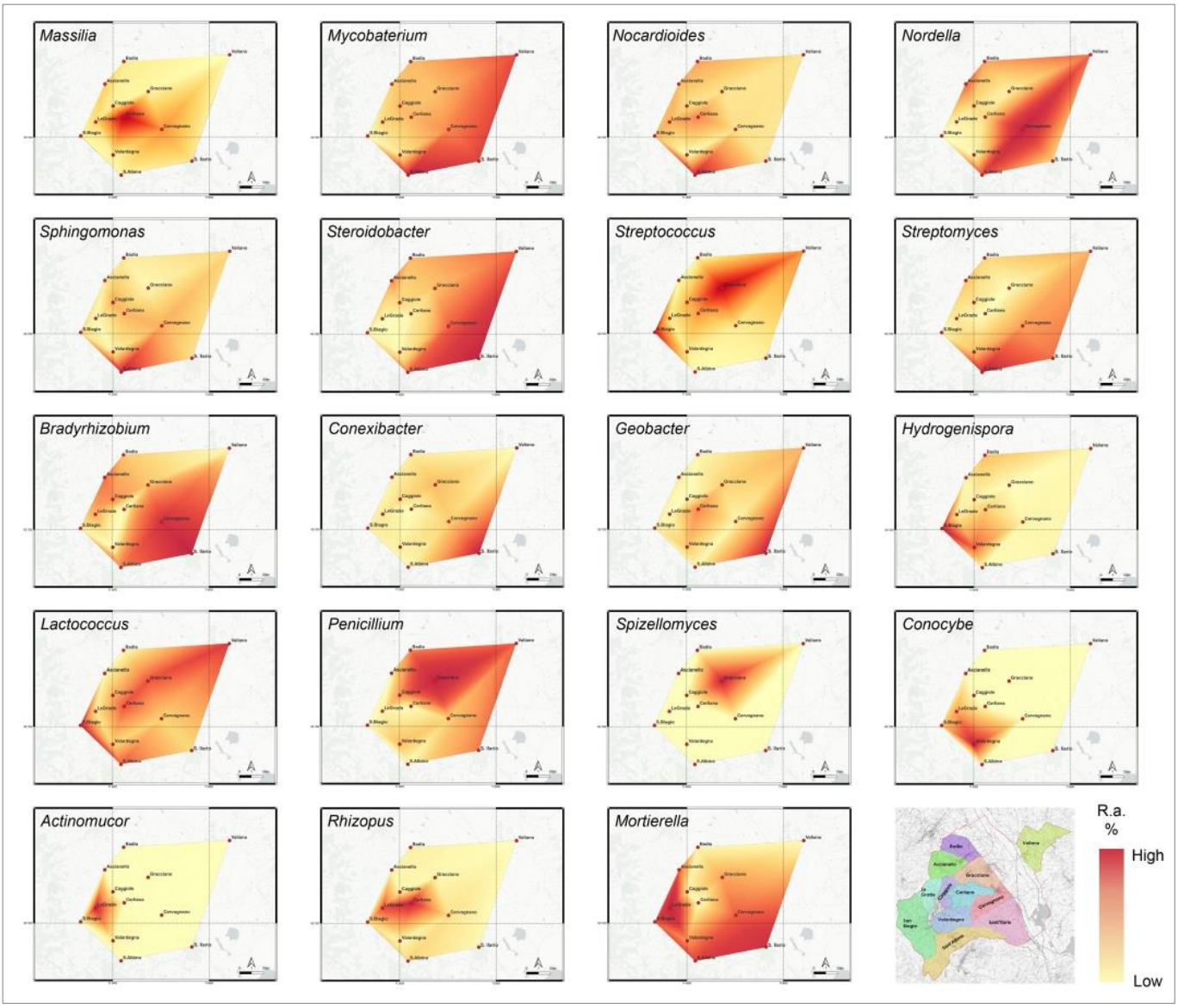


Fig. 6 – The AGU-related taxa show a pattern of relative abundance variation across the 12 AGUs (Additional Geographical Units). Maps of the relative abundance of discriminating microbial components of the rhizosphere, both bacterial and fungal. The AGU map is shown at the bottom right for reference to the AGU location, together with a color code for the relative abundance percentage (r.a.%). All maps were created using the QGIS open-source tool (https://www.qgis.org/it/site/).

We found that the rhizosphere microbiome of each AGU showed its own peculiar functional profile of microbial PGP traits, with the AGUs of the southeastern area showing an overall greater potential for P solubilization, while those of the western area showing a greater propensity for ACC deaminase production (p-value = 0.05, Wilcoxon test, **Figure 7**).

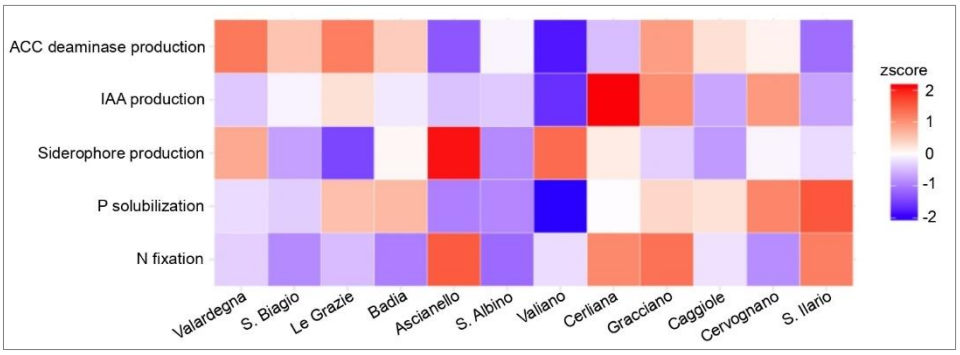


Fig. 7 – Functional plant growth-promoting (PGP) profile of Additional Geographical Units (AGU) rhizosphere microbiomes. Heatmap of PGP functions identified in rhizosphere samples from the different AGUs. PGP functions were normalized in copies per million ((reads count for an enzyme in a given sample/(gene length/1000))/(n° reads per sample/10^6) and represented in the heatmap through z-score. Abbreviations: ACC = 1-aminocyclopropane-1-carboxylic acid; IAA = indole-3-acetic acid; P = phosphorous; N = nitrogen.

It is also interesting to note that the functional potential for siderophore and IAA production showed an isolated peak in the northern and central part of the territory, respectively. Finally, the N fixation potential was more homogeneous in the production area (z-score between -1 and 1), although it was more present in the AGUs of Sant'Ilario, Gracciano, Cerliana, and Ascianello. Notably, all PGP functions were present at similar levels in the soil microbiomes of the corresponding AGUs, again supporting the continuity between the two ecosystems also from a functional point of view ($r > 0.98$, p-value < 0.0001, Pearson's correlation) (**Supplementary Table 3**). We then used the entire set of 28 metagenomes to reconstruct metagenome-assembled genomes (MAGs), with the aim of increasing the taxonomic resolution of 16S rRNA gene sequencing analysis and matching the functional potential to the corresponding taxonomy. We were able to reconstruct 37 MAGs with more than 50% completeness and less than 5% contamination, 15 of which were taxonomically assigned to the previous AGU-associated bacterial taxa and 4 to the core bacterial genera of the production area (**Supplementary Table 4**). Specifically, of the 15 MAGs assigned to AGU-associated bacterial genera, three were assigned to unclassified species of Conexibacteriaceae, one to *Massilia yuzhufengensis*, one to *Bradyrhizobium algeriense*, one to unclassified species of the genus *Mycobacterium*, one to *Nocardioides iriomotensis*, two to unclassified species of the genus *Nocardioides*, one to unclassified species of the genus *Sphingomonas*, two to *Steroidobacter denitrificans*, and three to unclassified species of the genus *Streptomyces*. The 4 MAGs assigned to core genera included the three MAGs assigned to *Nocariodes* and one to unclassified species of the genus *Solirubrobacter*. The entire set of MAGs was further processed by directly aligning them to the previous PGP gene sequences used to screen for the potential ability to support plant growth at the soil-root interface in each AGU. This analysis was integrated by applying METABOLIC[12], a software that computes the contribution of microorganisms to biogeochemical transformations and cycles of carbon, N and sulfur (**Figure 8** and **Supplementary Table 5**). Looking specifically at the MAGs assigned to the core taxa, we found that they covered a very broad range of functions capable of supporting soil fertility and plant health. In particular, the two MAGs assigned to unclassified *Nocardioides* (bin.197 and bin.92) and the one assigned to *N. iriomotensis* (bin.111) encoded for genes involved in siderophore production, N fixation, nitrite ammonification, nitrate reduction, iron reduction, organic substrate

fermentation, acetate oxidation, and organic carbon oxidation. In addition, the MAG assigned to *Solirubrobacter* (bin.178) carried the genes necessary for the oxidation of organic carbon from amino acids and complex carbohydrates, including several glycosyl hydrolases, such as GH5, GH65, GH113, GH39, and GH15, which are involved in the degradation of various polysaccharides, such as mannans and glucans.

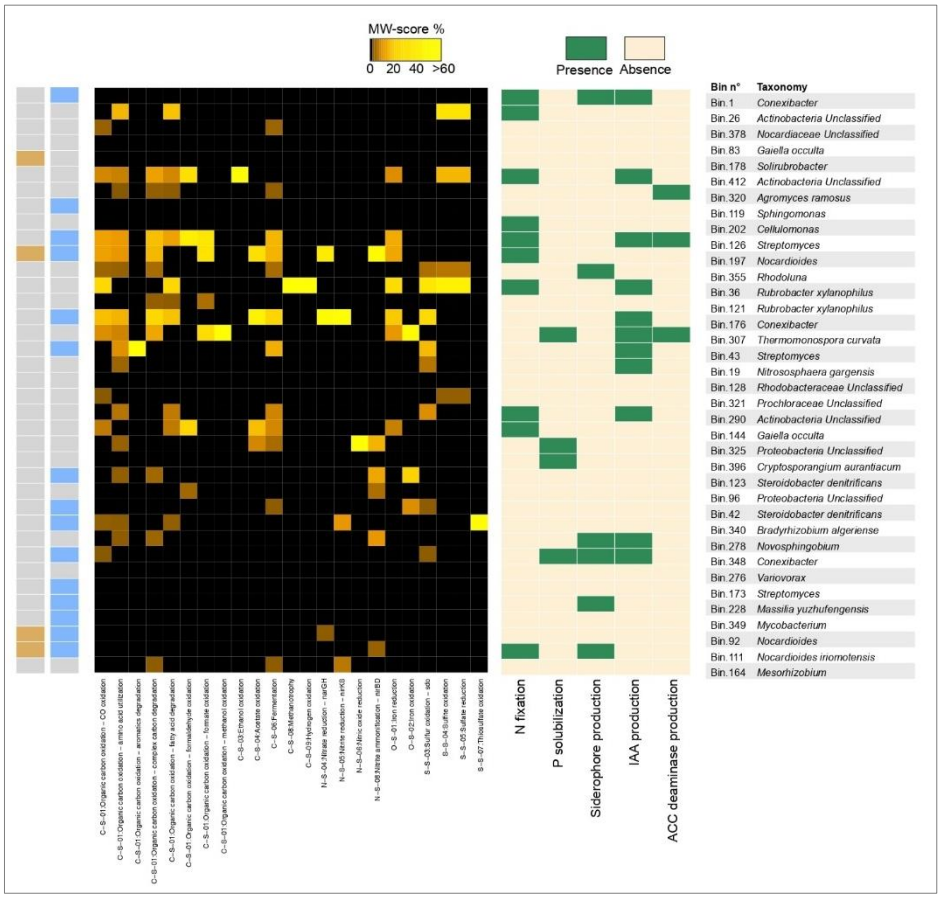


Fig. 8 – Presence of plant growth-promoting (PGP) functions in reconstructed metagenome-assembled genomes (MAGs) across the Montepulciano territory. Gradient heatmap of metabolic functions identified in the MAGs and presence/absence table of the PGP functions. From left to right: (i) core taxa identified among MAGs in beige; (ii) MAGs with taxonomy corresponding to discriminant taxa in light blue; (iii) METABOLIC functions identified in the MAGs in a gradient of MW-score percentage (black ”0”, yellow “> 60%”); (iv) PGP pathways identified in the MAGs (green = presence; light beige = absence); and (v) MAGs number and corresponding taxonomy. The MW-score represents the metabolic potential of each MAG within the Montepulciano territory, based on its coverage (how much this MAG is present in the territory), and on the presence or absence of the genes in the MAG (whether the analyzed function is present or not).

We then analyzed the MAGs assigned to AGU-discriminant taxa for PGP functions previously shown to be discriminant for the different AGUs, specifically P solubilization, which was higher in AGUs on the southeastern side of the territory, and ACC deaminase production, which was higher in AGUs on the western side. Among the MAGs characterizing the AGUs in the southeastern side of the territory, we found that bin.348 (*Conexibacter*) was

equipped with metabolic pathways for P solubilization, while bin.126 (*Streptomyces*), characterizing the AGUs in the western side, carried the genes necessary for ACC deaminase production. Looking more widely at the functional features characterizing the 15 AGU-related MAGs, we found that the two MAGs for *S. denitrificans* (bin.123 and bin.42), the three for *Conexibacter* (bin.1, bin.176, and bin.348) and the one for *B. algeriense* (bin.340), whose higher abundances were associated with the AGUs in the southeastern part of the territory, encoded the functions necessary for IAA production, sulfur oxidation, N fixation, nitrate reduction, nitrite ammonification, nitrite reduction, fermentation of organic substrates, iron oxidation and reduction, siderophore production, thiosulfate oxidation, fermentation of organic compounds, acetate oxidation, and oxidation of organic carbon from different sources including amino acids, fatty acids, and complex carbohydrates. Furthermore, the MAG assigned to *M. yuzhufengensis* (bin.228), characteristic of the AGUs in the central part of the territory, carried the genes devoted to siderophore production. Finally, the three MAGs assigned to the core taxon *Nocardioides* (bin.111, bin.197, and bin.92) were more abundant in the AGUs in the southern part of the territory, as were the three MAGs assigned to *Streptomyces* (bin.126, bin.173, and bin.43). In particular, the latter encoded the genes necessary for ACC deaminase and siderophore production, N fixation, iron reduction, fermentation of organic molecules, sulfur oxidation, and oxidation of organic carbon from different sources such as complex carbohydrates and aromatic compounds.

Wine metabolomics highlights association between product features and the variation of the microbial terroir in the different AGUs

In order to explore matches between wine characteristics and the variation of the microbial terroir in the different AGUs, a metabolomic analysis of the wines of 2022 vintage (the year of the sampling campaign) from wineries where grapes were taken exclusively from the same AGU (namely, Sant'Ilario, Caggiole, Cervognano, Le Grazie, Valardegna, and Valiano) was conducted. Interestingly, the diversity of the wines metabolic profiles matched the variation in the overall rhizosphere microbiome configuration in the corresponding AGU (p-value = 0.04, Procrustes test, **Supplementary Figure 5C**). In order to identify the wine metabolites responding to changes in the rhizosphere microbiomes in the different AGUs, the analytical components were superimposed on the PCoA plot of the Unweighted UniFrac distances between bacterial and fungal profiles of the rhizosphere microbiomes at the different AGUs. Interestingly, several wine metabolites, as L-acetylcarnitine, L-methionine, quercetin, and citicoline for the bacterial configuration, and adenine for the fungal configuration, were significantly associated with terroir specificities of the rhizosphere microbiome in the different AGUs (p-value < 0.05, "envfit" function in the vegan R package, **Supplementary Figure 5A** and **B**).

Discussion

In the present study, we characterized the soil-plant microbiome dimension of the variation of the viticultural terroir of *V. vinifera* cultivar Sangiovese from the "Consorzio del Vino Nobile di Montepulciano DOCG" PDO area in Tuscany, Italy. This was made possible by comparing the microbiomes associated with the bulk soil and rhizosphere of vineyards located in the 12 AGUs recently established by the consortium. At first, we explored specificities in the bulk soil microbiome from Montepulciano territory with respect to other vineyards in Italy and around the world. Data revealed a clear differentiation of the Montepulciano bulk soil vineyard microbiomes compared to all other vineyards, with 10 microbial genera characterizing the Montepulciano territory. Specifically, 5 bacterial and 5 fungal taxa were identified, namely the bacteria *Hydrogenispora*, *Ilumatobacter*, *Microlunatus*, *Gemmata*, and *Nocardioides*, and the fungi *Rhizopus*, *Mortierella*, *Gongronella*, *Lipomyces*, and *Penicillium*. Among these, *Nocardioides* was also part of the core microbiome of Montepulciano vineyards, together with other four taxa, specifically *Solirubrobacter*, *Gemmatimonas*, *Haliangium*, and *Pirellula*, as they were present in all AGUs. It is noteworthy that the genus *Nocardioides* has recently been reported as one of the beneficial microorganisms capable of counteracting and preventing *Fusarium oxysporum* infection in crops[13], as well as being a potential carrier of other multiple PGP traits[9]. *Gemmatimonas* and *Haliangium* have also previously been associated with plant growth benefits[14,15]. Confirming previous findings, we then highlighted the continuity between the soil and rhizosphere microbiomes in the Montepulciano territory, with evidence that the microbial specificity of the territory (soil) directly reflects the microbiome configuration at the soil-root interface, potentially determining different interactions that differentially affect plant growth and biology[5]. In order to derive microbiome specificities at AGU level, we focused on the rhizosphere microbiome and we found that samples clustered according to sampling location, but not to sampling season. In particular, we observed a west-southeast gradient in the relative abundance of some microbial genera (from Sant'Ilario to Argiano AGUs), which correlated with the geographical distances between AGUs and the pedology of the area, suggesting that previous observations of variation in microbial terroir associated with a national and regional geographical scale[8] are also valid at a local scale. The taxa identified included 24 bacterial and 6 fungal genera, some of which have been previously associated with different PGP functions[13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29]. To deeply explore these observation, we applied shotgun metagenomics to a subset of samples to investigate the presence of PGP functions that could potentially promote plant growth through soil biofertilization and grapevine bio-stimulation by enhancing nutrient and water uptake, and providing higher resistance to environmental stressors, better plant health, and also probably improved wine organoleptic characteristics, thus contributing to regional terroir[5,30,31,32,33]. We found that the potential microbial contribution to the biogeochemical cycles of N and carbon, which are critical for soil fertility and plant health, is widely diffused in the analyzed genomes, with the four MAGs assigned to the core taxa (i.e., *Nocardioides* and *Solirubrobacter* as genera present in all AGUs; notably, *Nocardioides* was both a core taxon and a genus that distinguished the Montepulciano territory from vineyards in the rest of the world), carrying the necessary genes for nitrification and denitrification pathways, as well as organic carbon oxidation and fermentation using different substrates, including complex carbohydrates, such as mannans and glucans, known components of plant cell walls[34,35]. The AGUs of the southeastern part of the production area (the side delimited by Sant'Ilario) showed a greater potential for P

solubilization, while those of the western part (delimited by S.Biagio) showed a greater propensity for ACC deaminase production. Consolidating this evidence, we also found that one MAG associated with the southeastern part of the territory, assigned to *Conexibacter*, encoded genes devoted to P solubilization, while MAGs associated with the western and southern parts, assigned to *Streptomyces*, carried the ACC deaminase gene. These differences match local peculiarities of the terroir, with the southeastern AGU-associated microbiome providing an extra means of P provision from a local soil, which resulted generally depleted of this important nutrient[36], and the western AGU-associated microbiome potentially helping plants to respond to salt and drought stress[37,38]. This set of metabolic potentials, either common to all AGUs or specific to some of them, represents a promising avenue for leveraging microbial terroir as a mediator between soil resources and plant requirements[7], with possible implications on local product quality and productivity, possibly contributing to wine differentiation depending on the AGU of origin. Indeed, by controlling for plant P availability, the root-associated microbiome can influence several sensory characteristics, including the aroma, appearance, flavor, and taste, of its associated wines[39,40]. On the other hand, by counteracting excessive drought stress, root microbiome ACC deaminases protect against delays in fruit ripening and the consequent loss of varietal character, which is crucial for flavor development[41,42]. To provide insights into possible connections between the observed compositional and functional variations in the microbial terroir at the AGU level and the corresponding organoleptic characteristics of the produced wines, metabolomic profiles of the wines were analyzed. Interestingly, several associations between wine metabolites and the terroir microbiomes were observed across the different AGUs. Specifically, varying concentrations of L-acetylcarnitine, L-methionine, quercetin, citicoline, and adenine in the wines from Sant'Ilario, Caggiole, Cervognano, Le Grazie, Valardegna, and Valiano, respectively, were linked to corresponding AGU-level specificities in the rhizosphere microbiome structure. These molecules have been previously reported as key determinant of vine characteristics. For instance, L-acetylcarnitine in grapes can influence the synthesis of esters that enhance the wine's aroma profile[43] while L-methionine has been associated with the production of volatile compounds that contribute to the aromatic complexity of wine[44]. Finally, quercetin contributes to the color, flavor, and health benefits of wine[45]. On the other hand, there is no documented evidence in the literature that the presence of different concentrations of citicolines and adenines in wine influences its organoleptic profile. Taken together, these finding suggest a possible connection between local features in the terroir microbiomes in the different AGUs and corresponding organoleptic features of the produced wines. Our findings, showing the potential relevance of the local diversity of terroir microbiome in maintaining plant health and productivity, and potentially wine product quality, became relevant when placed in the current context of global change, leading to nutrient and soil depletion and loss of microbial diversity[46,47,48,49]. In this scenario, the characterization of the root-associated microbiome encoding PGP functions could represent the first step towards new strategies to improve the sustainability and resilience of viticulture, integrating management strategies for the protection and preservation of the local microbial terroir features as a key aspect in the product quality[50]. Our results, coupled with the growing evidence of the significant influence of both soil and plant microbiomes on the sensory properties of the final product[51,52], may lay the foundations for a new perspective in which the local variation of microbiome features in terroir needs to be protected as a biodiversity treasure highly linked to the local diversity of wine production and traditions. This is particularly relevant in cases where the product is closely linked to its

geographical origin, such as within a PDO area, like DOCG in Italy, or when the vineyard location is indicated by the term "cru", which immediately links the product to a precise growing location. The local microorganisms carrying the genes for PGP functions could be those best suited and preserved to thrive in the local pedoclimatic conditions and contribute to healthy plant development. This will require the integration of current viticultural strategies with a precise and tailored microbiological approach. It will entail the combination of the isolation of PGP microorganisms with the metagenomic approach, thereby enabling a comprehensive investigation of their functions through genome sequencing and targeted functional assessments. This represents an unexplored frontier aimed at safeguarding and enhancing the properties and qualities of wine in a context of global change by exploiting the natural microbiomes of the vineyard.

Methods

Study sites, sample and metadata collection, and sample pre-treatment

Grapevine roots and soil samples were collected from the 12 production areas (AGUs) within the "Consorzio del Vino Nobile di Montepulciano DOCG" PDO area in Tuscany, Italy. All plants were *V. vinifera* cultivar Sangiovese, apparently healthy, more than 15 years old and used for the production of Vino Nobile di Montepulciano. Sampling was carried out at 4 different time points throughout the production season in 2022 (i.e., pre-harvest in July, August and September, and post-harvest in October) for a total of 336 root samples and 56 soil samples. Specifically, for each of the 14 vineyards, 6 rhizospheric samples and 1 bulk soil sample were retrieved at each timepoint. Chemical features of the vineyard soil (e.g. P and N concentrations) were measured at the time of samples collection using a multiparametric probe. In particular, for the microbial characterization of the rhizospheric soil, thin lateral roots of the grapevine were collected after digging 10-20 cm below the ground surface, whereas bulk soil was collected in the plant proximity at the same depth after removing the surface soil and grass cover, if present[9]. All samples were collected using sterile gloves, placed in a sterile 50-ml Falcon tube, transported to the laboratory on ice, and stored at −80°C until further processing. To separate the rhizospheric soil from the root surface, roots were treated as previously described[9,53]. Briefly, approximately 3 cm of terminal root portions, including tips, were dissected using sterilized scissors and tweezers. The root segments were then placed in 15-ml Falcon tubes filled with 2.5 ml of modified PBS buffer (130 mM NaCl, 7 mM Na_2HPO_4, 3 mM NaH_2PO_4, pH 7.0, and 0.02% Silwet L-77) and left on a shaking platform at 180 rpm for 20 min for washing. Roots were removed from the tubes and the washing buffer was centrifuged at 1,500 × *g* for 20 min, with the resulting pellet regarded as the rhizospheric soil. Metadata, such as agronomical practices and management, vine clone type, rootstock family, and altitude, were collected during the sampling campaigns by inspecting the vineyards and asking the winemakers directly. Soil composition was retrieved from a previous publication[54].

Microbial DNA extraction and sequencing

Prior to DNA extraction, rhizosphere and soil samples for fungal analysis were treated with 500 µl of lyticase solution (for 1 ml of solution: 978 µl Tris-EDTA, 2 µl β-mercaptoethanol, and 20 µl lyticase 10 U/ml) and incubated at 37°C

for 30 min to facilitate fungal cell wall lysis and nucleic acid recovery. Microbial DNA was extracted from the rhizospheric soil after the pre-treatment described above and from 0.25 g of bulk soil using the DNeasy PowerSoil Pro Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions with a few modifications. Briefly, the homogenization step was performed using a FastPrep instrument (MP Biomedicals, Irvine, CA, USA) with a cycle consisting of three 1-min steps at 5.5 movements per sec with 5-min incubation on ice between each step. At the end of the protocol, DNA elution was preceded by a 5-min incubation on ice. The resulting DNA was quantified using a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and diluted in PCR-grade water to a final concentration of 5 ng/µl before amplification. For characterization of the bacterial fraction of the rhizospheric and bulk soil microbiome, the V3-V4 hypervariable regions of the 16S rRNA gene were PCR-amplified in a final volume of 50 µl containing 25 ng of genomic DNA, 2X KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland) and 200 nmol/L of 341F (S-D-Bact-0341-b-S-17, 5′-CCTACGGGNGGCWGCAG-3′) and 785R (S-D-Bact-0785-a-A-21, 5′-GACTACHVGGGTATCTAATCC-3') primers[55] carrying Illumina overhang adapter sequences. The thermal cycle consisted of an initial denaturation at 95°C for 3 min, followed by 25 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s, with a final extension step at 72°C for 5 min[56]. For characterization of the fungal component, ITS2 was PCR-amplified as above using ITS3 and ITS4 primers[57] carrying Illumina overhang adapter sequences. The thermal cycle consisted of an initial denaturation at 95°C for 3 min, followed by 30 cycles at 95°C for 30 s, 56°C for 30 s, and 72°C for 1 min, with a final extension step at 72°C for 5 min. All PCR amplicons were cleaned up with Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA) and indexed libraries were prepared by limited-cycle PCR using Nextera technology. Indexing was followed by a second clean-up as described above. Libraries were then quantified using a Qubit 3.0 fluorimeter (Invitrogen, Waltham, MA, USA), normalized to 4 nM and pooled. Prior to sequencing, the sample pool was denatured with 0.2 N NaOH and diluted to 4.5 pM with a 20% PhiX control. Sequencing was performed on an Illumina MiSeq platform using a 2 × 250-bp paired-end protocol, according to the manufacturer's instructions (Illumina, San Diego, CA, United States). A subset of 28 representative samples (14 bulk soils and 14 rhizospheric soils) were further processed for shotgun sequencing. DNA libraries were prepared using the QIAseq FX DNA Library Kit (QIAGEN) according to the manufacturer's instructions. Shortly, 100 ng of each DNA sample was fragmented to 450-bp size, end-repaired, and A-tailed fragments using FX enzyme mix with the following thermal cycle: 4°C for 1 min, 32°C for 8 min, and 65°C for 30 min. Adapter ligation was performed by incubating DNA samples at 20°C for 15 min in the presence of DNA ligase and Illumina adapter barcodes. A first purification step with Agencourt AMPure XP magnetic beads (Beckman Coulter) was performed, followed by library amplification with limited-cycle PCR according to the manufacturer's instructions, and a further purification step. Samples were pooled at an equimolar concentration of 4 nM to obtain the final library. Sequencing was performed on an Illumina NextSeq platform using a 2 × 150-bp paired-end protocol, following the manufacturer's instructions (Illumina). All sequencing reads were deposited in the ENA archive under the accession number PRJEB75007.

Bioinformatic analysis of microbiome data

For marker gene analysis (16S rRNA gene and ITS), samples were processed using a pipeline combining PANDAseq[58] and QIIME 2[59]. The "fastq filter" function of the Usearch11 algorithm[60] was applied to retain high-

quality reads (min/max length = 350/550 bp). Based on the phred Q score probabilities, reads with an expected error per base E = 0.03 (i.e., three expected errors every 100 bases) were discarded. The retained reads were then binned into ASVs using DADA2[61]. Taxonomic assignment was performed using the VSEARCH algorithm[62] and the SILVA database[63] (December 2017 release) for bacteria and the UNITE database[64] (May 2021 release) for fungi. For the bacterial analysis, all sequences assigned to eukaryotes (including mitochondrial and chloroplast sequences) or unassigned were discarded, whereas for the fungal analysis, all sequences not assigned to the fungal kingdom were discarded. Normalization was performed to the lowest number of reads for all samples. Beta diversity was estimated by computing unweighted UniFrac distances and Bray-Curtis distances for bacterial and fungal communities, respectively. In order to identify the soil microbiome peculiarities of the bacterial and fungal composition within the "Consorzio del Vino Nobile di Montepulciano DOCG" with bulk soils from vineyards from all over the world, we retrieved sequencing data from the study of Gobbi et al. (2022)[8]. Although Gobbi and co-workers[8] applied a different set of primers for bacterial and fungal analysis, the same bioinformatic pipeline as described above was applied for sequence processing. In order to avoid bias of ASVs assignment, giving the different DNA regions taken into account, we compared Gobbi's and our dataset at genus levels. All statistical analyses were performed using the R software (R Core Team; www.r-project.org - last accessed March 2021), v. 4.1.2, with the packages "Made4"[65], "vegan" (https://cran.r-project.org/web/packages/vegan/index.html, v2.6-6.1) and "heatmap3"[66]. Data separation in the PCoA was tested using a permutation test with pseudo-F ratio (function "adonis" in the vegan package). The Wilcoxon rank-sum test and Kruskal-Wallis test were used to assess significant differences in relative taxon abundance between groups. P-values were corrected for multiple testing using the Benjamini–Hochberg method, with a FDR ≤ 0.05 considered statistically significant. Specifically, we used the Procrustes test to compare microbiomes across time points, taking into account seasonality and management practices (i.e., agronomical practices and management, vine clone type, rootstock family, altitude, and soil composition) as well. Random forest[11] with default parameters was used to assess discriminant taxa between AGUs. For the graphical representation of **Figures 4B** and **5**, we used the soil ITS configuration of Cerliana AGU. In order to confirm soil-rhizospheric continuity of microbial community composition, we performed a Procrustes correlation test on the beta distribution using the "protest" function in R. Variations in wine metabolites related to the microbiome configurations were estimated by correlation analysis calculated using the "envfit" function in the vegan R package. QGIS software (https://qgis.org/it/site/, last access February 2024) was used to construct maps of bacterial and fungal distribution in the Montepulciano territory based on relative taxon abundance. The geographical coordinates of longitude and latitude were used to plot the exact sampling locations in the software. The distribution of relative abundance across samples was obtained using the Triangulated Irregular Network interpolation method in QGIS (TIN interpolation).

For shotgun metagenomics sequencing, KneadData (https://github.com/biobakery/kneaddata, v0.10.0) was used with default parameters to trim and remove low-quality (q<20) reads, tandem-repeated sequences (based on fastqc output) and host reads (*V. vinifera* RefSeq id: GCF_000003745.3). High-quality reads were assembled with MegaHit[67] and the resulting contigs were processed with MetaWRAP[68] for MAG generation. Bins were evaluated for completeness and contamination with CheckM[69]. Only MAGs with >50% completeness and <5% contamination were retained for subsequent analyses. Taxonomic assignment of MAGs was performed with PhyloPhlAn 3.0[70]. All MAGs were tested

for their ability to support plant growth by searching for known PGP genes[9]. The microbial PGP functions selected for our analysis were: (i) N fixation (i.e., the genes NifB, NifE, NifH, NifN, NifV, and NifU); (ii) P solubilization (i.e., the alkaline phosphatase phoA and the glucose dehydrogenase GDH); (iii) iron chelation (i.e., the bacterial siderophores-encoding genes EntF/EntS for enterobactin and FslA for rhizoferrin); (iv) production of the phytohormone IAA (i.e., three genes directly involved in IAA synthesis, namely ipdC, aro10, and aldH); and (v) production of the enzyme ACC deaminase (i.e., AcdS gene encoding the enzyme). The amino acid sequence of selected PGP proteins, obtained from the reference sequence of the NCBI protein database (https://www.ncbi.nlm.nih.gov; last access May 2023) (**Supplementary Table 6**), was recovered and blasted against the MAG sequences using BlastP[71]. Alignments were filtered for query coverage >40% and identity percentage >20%. Reads count for each enzyme within the PGP functions was normalized in copies per million ((reads count for an enzyme in a given sample/(gene length/1000))/(n° reads per sample/10^6). Finally, the entire set of MAGs was processed using the METABOLIC software[12] with default parameters to compute their contribution to biogeochemical transformations and cycles of carbon, N, and sulfur. The software calculates the MW-score percentage, which was used for the color gradient of the heatmap in **Figure 8** to show the metabolic potential of each MAG within the Montepulciano territory, based on its coverage (how much this MAG is present in the territory) and the presence or absence of the genes in the MAG (whether the analyzed function is present or not).

Wine analytical characterization

All wine samples underwent two types of analytical characterization: untargeted headspace solid-phase microextraction gas chromatography-electron impact mass spectrometry (HS-SPME-GC-Orbitrap) and untargeted metabolomics analysis using micro liquid chromatography-high-resolution mass spectrometry (microLC-ESI-QTOF) in data-dependent acquisition (DDA) mode. For HS-SPME-GC analysis, a TriPlus RSH SMART robotic autosampler ensured consistent pre-analytical preparation. Each 20 mL glass vial contained 5 mL of wine and 1.5 g of NaCl, sealed with a magnetic cap featuring a pierceable septum. A technical blank of bi-distilled water and a surrogate matrix blank of a 90:10 v:v water-ethanol solution were also analyzed. Samples were incubated at 40°C for 5 minutes with orbital mixing at 250 rpm to achieve gas phase equilibrium. After thermal conditioning, the septum was pierced, and a DVB/Carbon WR/PDMS smart SPME fiber (80 µm thickness) was exposed 10 mm above the liquid for 30 minutes to extract analytes. Extraction conditions were optimized based on previous studies[72,73]. Following extraction, the fiber was retracted and transferred to the gas chromatograph's injection port, where it was thermally desorbed at 250°C for 3 minutes. The SPME fiber was thermally cleaned at 240°C for 5 minutes between analyses. GC-MS analysis utilized a TRACE GC 1610 Series (Thermo Fisher Scientific, Waltham, MA, USA) gas chromatograph interfaced with an Orbitrap Exploris MS (Thermo Fisher Scientific, Waltham, MA, USA) analyzer, operating in electron impact mode (70 eV). The capillary GC column was TG-5MS (30 m x 0.25 mm ID, 0.15 µm film thickness), with helium as the carrier gas at a flow rate of 1.2 mL/min. The temperature program started at 40°C (5 min hold), ramped at 7 °C/min to 250°C (5 min hold), totaling 40 minutes. The transfer line and ionization source temperatures were maintained at 270°C, and the GC operated in split mode with a 1:20 split ratio. Mass spectra were recorded in full scan mode (50-700 Dalton) to collect total ion current chromatograms. For LC-HRMSMS metabolome analysis, 1 mL aliquots of liquid wine were taken from commercial bottles after vigorous shaking.

Interpooled Quality Control (QC) samples were created by pooling equal aliquots (200 µL) from each sample and underwent the same preparation as experimental samples. Wine aliquots and QCs were vortex-mixed for 30 seconds and sonicated for 10 minutes in a sonicator bath. After sonication, samples were transferred to Spin-X Centrifuge Tube Filters (0.22 µm, cellulose acetate membrane) and centrifuged at 14,000 rpm for 10 minutes. The filtered extracts were collected in glass microvials for analysis. To minimize bias, all experimental samples were randomized before the analytical run. Additionally, a QC sample was injected repeatedly (10 times) prior to the first sample injection for system equilibration and conditioning. Sample analyses were conducted in Data Dependent Acquisition (DDA) mode. LC-MS analysis was performed using an Eksigent M5 MicroLC system (Sciex) coupled with a TripleTOF 6600+ mass spectrometer featuring an OptiFlow Turbo V Ion Source (Sciex). Analyses were conducted in positive ionization mode with the column temperature set at 35°C. A 5 µL sample volume was loaded onto a Phenomenex Luna Omega Polar C18 column (100 × 1.0 mm I.D., 1.6 µm, 100 Å). Chromatographic separation occurred over 25 minutes at a constant flow rate of 30 µL/min, following this gradient elution program: 0-2 minutes, 0.2% eluent B; 2-5 minutes, 0.2-15% eluent B; 5-15 minutes, 15-70% eluent B; 15-18 minutes, 70-98% eluent B; 18-20 minutes, 98% eluent B; 20-22 minutes, 98-0.2% eluent B; 22-25 minutes, 0.2% eluent B. Equilibration time between chromatographic runs was 3 minutes. Mobile phase A consisted of 0.1% formic acid, while mobile phase B was acetonitrile with 0.1% formic acid. IonSpray voltage (ISV) was set to 5,000 V, and the Curtain Gas supply pressure (CUR) was 30 PSI. Nebulizer and heater gas pressures were set at 30 and 40 PSI, respectively, with the ion spray probe temperature at 300°C. The declustering potential was 80 V, and analyses were performed using a collision energy of 40 eV.

Statistics and Reproducibility

Each subsection of the methods contains detailed explanations of statistical approaches used in this paper. In brief, all statistical analyses were conducted using the R software (R Core Team; www.r-project.org, accessed March 2021), version 4.1.2, with the packages "Made4"[65], "vegan" (https://cran.r-project.org/web/packages/vegan/index.html, version 2.6-6.1), and "heatmap3"[66]. A permutation test with a pseudo-F ratio (function "adonis" in the vegan package) was employed to assess the suitability of data separation in the PCoA. The Wilcoxon rank-sum test and Kruskal-Wallis test were employed to ascertain whether there were significant discrepancies in relative taxon abundance between the various groups. P-values were corrected for multiple testing using the Benjamini–Hochberg method, with a false discovery rate (FDR) of ≤ 0.05 considered statistically significant. Procrustes test was employed to compare microbiomes across time points, with consideration given to the influence of seasonality and management practices (including agronomical practices and management, vine clone type, rootstock family, altitude, and soil composition). A Procrustes correlation test on the beta distribution was also performed to assess the soil-rhizospheric continuity of microbial community composition. Variations in wine metabolites related to the microbiome configurations were estimated by correlation analysis, calculated using the "envfit" function in the vegan R package.

Data availability

Both metagenomics and metabarcoding data generated during the current study are available in the ENA archive under the accession number PRJEB75007.

Acknowledgements

This work was carried out in the context of the "Controlling Microbiomes Circulations for Better Food Systems" (CIRCLES) project, which was funded by the European Union's Horizon 2020 research and innovation program under grant agreement no. 818290. We gratefully acknowledge the Consorzio del Vino Nobile di Montepulciano for providing access to vineyards and their invaluable assistance during sampling campaigns and covariates collection. Their cooperation significantly contributed to the success of this study.

Author contributions

M.C. and S.R.: conceptualization. G.P., E.N., D.S. N.I., J.F., and S.R.: data curation, formal analysis. M.C. and S.R.: project administration, resources, supervision. G.P., E.N., D.S., N.I. and S.R.: visualization. G.P. and S.R.: writing – original draft. E.N., D.S., N.C., L.F., A.C., A.G.V.R., N.I., J.F., S.T. and M.C.: writing – review and editing.

Competing interest

The authors declare no competing interest.

Supplementary material

The Supplementary material for this article can be found in the appendix of the thesis (Study II).

References

[1]D'amico, Mario, Giuseppe Di Vita, and Luisa Monaco. "Exploring environmental consciousness and consumer preferences for organic wines without sulfites." *Journal of cleaner production* 120 (2016): 64-71.

[2]Van Leeuwen, Cornelis, and Gerard Seguin. "The concept of terroir in viticulture." *Journal of wine research* 17.1 (2006): 1-10.

[3]Bokulich, Nicholas A., et al. "Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics." *MBio* 7.3 (2016): 10-1128.

[4]D'Agata, Ian. *Italy's native wine grape terroirs*. Univ of California Press, 2019.

[5]Sessitsch, Angela, et al. "Microbiome interconnectedness throughout environments with major consequences for healthy people and a healthy planet." *Microbiology and Molecular Biology Reviews* 87.3 (2023): e00212-22.

[6]Oyuela Aguilar, Mónica, et al. "Influence of vintage, geographic location and cultivar on the structure of microbial communities associated with the grapevine rhizosphere in vineyards of San Juan Province, Argentina." *PLoS One* 15.12 (2020): e0243848.

[7]Belda, Ignacio, et al. "From vineyard soil to wine fermentation: microbiome approximations to explain the "terroir" concept." *Frontiers in Microbiology* 8 (2017): 821.

[8]Gobbi, Alex, et al. "A global microbiome survey of vineyard soils highlights the microbial dimension of viticultural terroirs." *Communications Biology* 5.1 (2022): 241.

[9]Nanetti, Enrico, et al. "Composition and biodiversity of soil and root-associated microbiome in Vitis vinifera cultivar Lambrusco distinguish the microbial terroir of the Lambrusco DOC protected designation of origin area on a local scale." *Frontiers in Microbiology* 14 (2023): 1108036.

[10]Marasco, Ramona, et al. "Grapevine rootstocks shape underground bacterial microbiome and networking but not potential functionality." *Microbiome* 6 (2018): 1-17.

[11]Liaw, A. "Classification and regression by randomForest." *R news* (2002).

[12]Zhou, Zhichao, et al. "METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks." *Microbiome* 10.1 (2022): 33.

[13]Zhao, Fengyan, et al. "Vermicompost can suppress Fusarium oxysporum f. sp. lycopersici via generation of beneficial bacteria in a long-term tomato monoculture soil." *Plant and Soil* 440 (2019): 491-505.

[14]Shang, J., and B. Liu. "Application of a microbial consortium improves the growth of Camellia sinensis and influences the indigenous rhizosphere bacterial communities." *Journal of applied microbiology* 130.6 (2021): 2029-2040.

[15]Liu, Chao, et al. "Soil bacterial communities of three types of plants from ecological restoration areas and plant-growth promotional benefits of Microbacterium invictum (strain X-18)." *Frontiers in Microbiology* 13 (2022): 926037.

[16]Antoun, Hani, et al. "Potential of Rhizobium and Bradyrhizobium species as plant growth promoting rhizobacteria on non-legumes: effect on radishes (Raphanus sativus L.)." *Plant and soil* 204 (1998): 57-67.

[17]Dell'Amico, Elena, Lucia Cavalca, and Vincenza Andreoni. "Improvement of Brassica napus growth under cadmium stress by cadmium-resistant rhizobacteria." *Soil Biology and Biochemistry* 40.1 (2008): 74-84.

[18]Luo, Hong, Heather E. Hallen-Adams, and Jonathan D. Walton. "Processing of the phalloidin proprotein by prolyl oligopeptidase from the mushroom Conocybe albipes." *Journal of Biological Chemistry* 284.27 (2009): 18070-18077.

[19]Yang, Suijuan, et al. "Growth-promoting S phingomonas paucimobilis ZJSH 1 associated with D endrobium officinale through phytohormone production and nitrogen fixation." *Microbial biotechnology* 7.6 (2014): 611-620.

[20]Evstatieva, Yana, et al. "Production of plant growth regulatory metabolites of Rhizopus arrhizus KB-2." *Bulgarian Journal of Agricultural Science* 26.3 (2020).

[21]Ozimek, Ewa, and Agnieszka Hanaka. "Mortierella species as the plant growth-promoting fungi present in the agricultural soils." *Agriculture* 11.1 (2020): 7.

[22]Masuda, Yoko, et al. "Enhancement of the nitrogen-fixing activity of paddy soils owing to iron application." *Soil Science and Plant Nutrition* 67.3 (2021): 243-247.

[23]Hanrahan-Tan, Deirdre G., et al. "The effects of nitrogen and phosphorus on colony growth and zoospore characteristics of soil Chytridiomycota." *Journal of Fungi* 8.4 (2022): 341.

[24]Kazerooni, Elham Ahmed, et al. "Actinomucor elegans and podospora bulbillosa positively improves endurance to water deficit and Salinity Stresses in tomato plants." *Journal of Fungi* 8.8 (2022): 785.

[25]Qi, Mingsheng, et al. "Identification of beneficial and detrimental bacteria impacting sorghum responses to drought using multi-scale and multi-system microbiome comparisons." *The ISME journal* 16.8 (2022): 1957-1969.

[26]Yan, Yuanyuan, et al. "Reductive soil disinfestation enhances microbial network complexity and function in intensively cropped greenhouse soil." *Horticulturae* 8.6 (2022): 476.

[27]Jaffar, Nur Sulastri, Roslina Jawan, and Khim Phin Chong. "The potential of lactic acid bacteria in mediating the control of plant diseases and plant growth stimulation in crop production-A mini review." *Frontiers in plant science* 13 (2023): 1047945.

[28]Nazari, Mateus Torres, et al. "Using Streptomyces spp. as plant growth promoters and biocontrol agents." *Rhizosphere* (2023): 100741.

[29]Zhang, Mei, et al. "Responses of fungal assembly and co-occurrence network of rhizosphere soil to Amaranthus palmeri invasion in Northern China." *Journal of Fungi* 9.5 (2023): 509.

[30]Baldan, Enrico, et al. "Beneficial bacteria isolated from grapevine inner tissues shape Arabidopsis thaliana roots." *PLoS One* 10.10 (2015): e0140252.

[31]Rolli, Eleonora, et al. "Improved plant resistance to drought is promoted by the root-associated microbiome as a water stress-dependent trait." *Environmental microbiology* 17.2 (2015): 316-331.

[32]Rolli, Eleonora, et al. "Root-associated bacteria promote grapevine growth: from the laboratory to the field." *Plant and Soil* 410 (2017): 369-382.

[33]Zarraonaindia, Iratxe, et al. "The soil microbiome influences grapevine-associated microbiota." *MBio* 6.2 (2015): 10-1128.

[34]Kirui, Alex, et al. "Carbohydrate-aromatic interface and molecular architecture of lignocellulose." *Nature communications* 13.1 (2022): 538.

[35]Grieß-Osowski, Annika, and Cătălin Voiniciuc. "Branched mannan and xyloglucan as a dynamic duo in plant cell walls." *The Cell Surface* 9 (2023).

[36]Elhaissoufi, Wissal, et al. "Phosphate bacterial solubilization: a key rhizosphere driving force enabling higher P use efficiency and crop productivity." *Journal of Advanced Research* 38 (2022): 13-28.

[37]Vurukonda, Sai Shiva Krishna Prasad, et al. "Enhancement of drought stress tolerance in crops by plant growth promoting rhizobacteria." *Microbiological research* 184 (2016): 13-24.

[38]del Carmen Orozco-Mosqueda, Ma, Bernard R. Glick, and Gustavo Santoyo. "ACC deaminase in plant growth-promoting bacteria (PGPB): An efficient mechanism to counter salt stress in crops." *Microbiological Research* 235 (2020): 126439.

[39]Schreiner, R. Paul, Jungmin Lee, and Patricia A. Skinkis. "N, P, and K supply to Pinot noir grapevines: Impact on vine nutrient status, growth, physiology, and yield." *American Journal of Enology and Viticulture* 64.1 (2013): 26-38.

[40]Skinner, Paul W., et al. "Sensory attributes of wines made from vines of differing phosphorus status: This article is published in cooperation with the 21th GIESCO International Meeting, June 23-28 2019, Thessaloniki, Greece. Guests editors: Stefanos Koundouras and Laurent Torregrosa." *Oeno One* 53.2 (2019).

[41]Gambetta, Gregory A., et al. "The physiology of drought stress in grapevine: towards an integrative definition of drought tolerance." *Journal of experimental botany* 71.16 (2020): 4658-4676.

[42]Duan, Bingbing, et al. "1-Aminocyclopropane-1-carboxylate deaminase-producing plant growth-promoting rhizobacteria improve drought stress tolerance in grapevine (Vitis vinifera L.)." *Frontiers in Plant Science* 12 (2021): 706990.

[43]Boss, Paul K., et al. "Potential grape-derived contributions to volatile ester concentrations in wine." *Molecules* 20.5 (2015): 7845-7873.

[44]Pripis-Nicolau, L., et al. "Methionine catabolism and production of volatile sulphur compounds by Oenococcus oeni." *Journal of Applied Microbiology* 96.5 (2004): 1176-1184.

[45]Careri, Maria, et al. "Direct HPLC analysis of quercetin and trans-resveratrol in red wine, grape, and winemaking byproducts." *Journal of agricultural and food chemistry* 51.18 (2003): 5226-5231.

[46]Schaller, K. "Intensive viticulture and its environmental impacts: nitrogen as a case study." *XXV International Horticultural Congress, Part 2: Mineral Nutrition and Grape and Wine Quality 512*. 1998.

[47]Thomidis, Thomas, Nikitas Karagiannidis, and Constantinos Karagiannidis. "Effect of New Environmentally Friendly Fertilizers in a Commercial Vineyard." *Journal of Agricultural Science* 8.1 (2016).

[48]Poore, Joseph, and Thomas Nemecek. "Reducing food's environmental impacts through producers and consumers." *Science* 360.6392 (2018): 987-992.

[49]Giffard, Brice, et al. "Vineyard management and its impacts on soil biodiversity, functions, and ecosystem services." *Frontiers in Ecology and Evolution* 10 (2022): 850272.

[50]Darriaut, Romain, et al. "Grapevine rootstock and soil microbiome interactions: Keys for a resilient viticulture." *Horticulture research* 9 (2022): uhac019.

[51]Knight, Sarah, et al. "Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir." *Scientific reports* 5.1 (2015): 14233.

[52]Zhou, Jia, et al. "Wine terroir and the soil bacteria: an amplicon sequencing–based assessment of the Barossa Valley and its sub-regions." *Frontiers in Microbiology* 11 (2021): 597944.

[53]D'Amico, Federica, et al. "The rootstock regulates microbiome diversity in root and rhizosphere compartments of Vitis vinifera cultivar Lambrusco." *Frontiers in microbiology* 9 (2018): 2240.

[54]Costantini, Edoardo AC, Pierluigi Bucelli, and Simone Priori. "Quaternary landscape history determines the soil functional characters of terroir." *Quaternary International* 265 (2012): 63-73.

[55]Klindworth, Anna, et al. "Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies." *Nucleic acids research* 41.1 (2013): e1-e1.

[56]Turroni, Silvia, et al. "Fecal metabolome of the Hadza hunter-gatherers: a host-microbiome integrative view." *Scientific reports* 6.1 (2016): 32826.

[57]Yu, Yongjie, et al. "ITS3/ITS4 outperforms other ITS region and 18S rRNA gene primer sets for amplicon sequencing of soil fungi." *European Journal of Soil Science* 73.6 (2022): e13329.

[58]Masella, Andre P., et al. "PANDAseq: paired-end assembler for illumina sequences." *BMC bioinformatics* 13 (2012): 1-7.

[59]Bolyen, Evan, et al. "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." *Nature biotechnology* 37.8 (2019): 852-857.

[60]Edgar, Robert C. "Search and clustering orders of magnitude faster than BLAST." *Bioinformatics* 26.19 (2010): 2460-2461.

[61]Callahan, Benjamin J., et al. "DADA2: High-resolution sample inference from Illumina amplicon data." *Nature methods* 13.7 (2016): 581-583.

[62]Rognes, Torbjørn, et al. "VSEARCH: a versatile open source tool for metagenomics." *PeerJ* 4 (2016): e2584.

[63]Quast, Christian, et al. "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools." *Nucleic acids research* 41.D1 (2012): D590-D596.

[64]Nilsson, Rolf Henrik, et al. "The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications." *Nucleic acids research* 47.D1 (2019): D259-D264.

[65]Culhane, Aedín C., et al. "MADE4: an R package for multivariate analysis of gene expression data." *Bioinformatics* 21.11 (2005): 2789-2790.

[66]Zhao, Shilin, et al. "Heatmap3: an improved heatmap package with more powerful and convenient features." *BMC bioinformatics* 15 (2014): 1-2.

[67]Li, Dinghua, et al. "MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph." *Bioinformatics* 31.10 (2015): 1674-1676.

[68]Uritskiy, Gherman V., Jocelyne DiRuggiero, and James Taylor. "MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis." *Microbiome* 6 (2018): 1-13.

[69]Parks, Donovan H., et al. "CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes." *Genome research* 25.7 (2015): 1043-1055.

[70]Asnicar, Francesco, et al. "Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0." *Nature communications* 11.1 (2020): 2500.

[71]Altschul, Stephen F., et al. "Basic local alignment search tool." *Journal of molecular biology* 215.3 (1990): 403-410.

[72]Rossi, Leucio, et al. "Optimization of HS-SPME-GC/MS analysis of wine volatiles supported by chemometrics for the aroma profiling of trebbiano d'abruzzo and pecorino white wines produced in Abruzzo (Italy)." *Molecules* 28.4 (2023): 1534.

[73]Zhang, Lin, et al. "Characterization of wine volatile compounds from different regions and varieties by HS-SPME/GC-MS coupled with chemometrics." *Current Research in Food Science* 6 (2023): 100418.

Chapter 2 – Gut microbiome (GM) of wild herbivores for waste-streams valorization

2.1 General introduction

2.1.1 The GM of wild herbivores as an untapped source of functionalities

The gut microbiome (GM) of herbivorous holobionts is a very dynamic microbial ecosystem and represents an untapped source of functional biodiversity, as it may offer unique insights into microbial adaptations to specialized plant-based diets across different environments. Overall, the gut microbiome refers to the vast community of microorganisms (bacteria, fungi, archaea, viruses, etc.,) that populate the whole gastrointestinal tract of an organism, making up a unique set of genetic functions that are paramount in the holobiont framework. Indeed, gut-dwelling microorganisms play crucial roles in digestion, metabolism, immune function, homeostasis, and health of all animals[1,2,3,4,5].

Herbivores, which primarily rely on plant material for nutrition, have co-evolved with highly specialized GMs to help break down complex plant fibers (such as cellulose, hemicellulose, lignin, and pectin), that animal-encoded enzymes cannot entirely digest, detoxify plant secondary metabolites (such as alkaloids and tannins), and ferment metabolic products (mainly monosaccharides) into usable short-chain fatty acids (SCFAs)[6,7,8]. SCFAs, also known as volatile fatty acids (VFAs), are simple monocarboxylic acids with a short aliphatic chain (2-6 carbon atoms) and represent a critical component of herbivore nutrition, serving as a primary energy source derived from the fermentation of complex plant fibers in the gut[4,7,9,10]. Microbial fermentation products of complex plant carbohydrates, including acetate, propionate, and butyrate (i.e., the main endpoints in the intestine), are absorbed by gut cells and can be either utilized directly as energy source or be released into the bloodstream to reach other districts[11,12,13]. In fact, SCFAs not only fuel metabolic processes by providing energy to intestinal cells, but also play vital roles in maintaining gut health and supporting immune functions. In mammals, for instance, SCFAs regulate epithelial barrier function, as well as mucosal and systemic immunity, via evolutionary conserved processes that involve G protein-coupled receptor signaling or histone deacetylase activity[12,13]. Besides, for many herbivores, SCFAs represent a significant portion of their overall caloric intake, underscoring their essential role in sustaining herbivore nutrition and health. In ruminants, for example, SCFAs, which are primarily produced during the fermentation of dietary fiber by the extremely complex rumen microbiota, are absorbed through the rumen wall and may provide up to 70% of the ruminant's total energy requirements. Specifically, it was estimated that acetate typically accounts for 60-70% of total fatty acid production, while propionate and butyrate make up around 20-30% and 5-15%, respectively. In these animals, acetate serves as a primary fuel for the synthesis of milk fat and body fat, while propionate plays a key role in gluconeogenesis (glucose production) in the liver, and butyrate is vital for maintaining rumen epithelial health and optimal rumen function[9,14].

Wild herbivores often host a more diverse array of microorganisms than their domesticated counterparts, enabling them to better break down complex plant fibers, detoxify naturally occurring toxins, and survive in variable ecosystems subjected to strong annual vegetation shifts[15,16,17]. As a matter of fact, wild herbivores are distinct from

domesticated ones in terms of their diet, environment, and evolutionary pressures, with their microbiomes being shaped by natural dietary diversity matching ever-changing ecosystems. Peculiar microbial adaptations detected in wild herbivores are thus linked to the consumption of a wide variety of plants, including grasses, leaves, fruits, and bark, which vary greatly in terms of nutrient composition and secondary metabolites layout[7,8,16,18]. This dietary diversity places unique selective pressures on the GM of wild herbivorous holobionts, resulting in a rich and adaptable microbial community that may include microorganisms specialized for digesting rough plant compounds or for metabolizing specific molecules that would be otherwise indigestible or even toxic[10,19,20]. As such, the great complexity and richness of the GM of wild herbivores are related to factors like climate, habitat, and seasonal changes of plant availability and composition, which are capable of influencing the structure and function of microbial populations in the gut[16]. Despite we still lack a comprehensive mechanistic understanding of GM dynamics in wild herbivores populations, hampering the exploitation of their full potential, some research in this direction has been carried out, emphasizing the pivotal role of microbes in cooperating with their hosts to cope with a complex polysaccharides-based diet.

Couch and colleagues, for instance, demonstrated a clear diet-driven seasonal plasticity of the wild African buffalo (*Syncerus caffer*) GM at Kruger National Park (South Africa). Briefly, feces from these wild ruminants were collected longitudinally, over a 3-year period, selecting seasons characterized by different dietary regimes of the individuals, and a metabarcoding approach was used for microbiome analysis. Interestingly, researchers highlighted a striking shift in microbial community structure when comparing the dry season (i.e., resource-restricted diet) with periods of high availability of green vegetation, suggesting a certain degree of adaptation of the buffalo microbiome to different food intake[21]. In line with these assumptions, a recent work reported that the GM of the North American mule deer (*Odocoileus hemionus*) is partly influenced by seasonality and food availability, showing compositional shifts between winter and spring[22]. Similar results were obtained by Su and co-workers, who described the GM of two deer species (i.e., *Moschus moschiferus* and *Moschus berezovskii*) after fecal sampling both in summer and winter. Also in this case, 16S gene sequencing allowed to point out marked species-related and season-related differences in GM structure and function when looking at the two main gut microbial phyla, namely Firmicutes and Bacteroidetes[23]. Furthermore, comparable patterns have been observed for other ruminants[16,24], but also for other herbivores, such as bamboo-eating giant pandas[25,26], primates[27], and rodents[28], for instance. These studies underline the importance of environmental and seasonal factors in shaping the microbiomes of wild herbivores, with implications for understanding microbial evolutionary adaptations (providing valuable insight into animal biology), as well as for unravelling the potential of these natural microbial systems for biotechnological/industrial applications. Indeed, this vast, yet largely unexplored, microbial richness holds the potential to harbor novel enzymes and metabolic pathways that could have significant applications in biotechnology, medicine, and agriculture, making this an exciting frontier for future scientific discovery.

For example, wild herbivores often host microbes capable of neutralizing plant toxins. Indeed, plants produce a variety of secondary metabolites (e.g., alkaloids, glucosinolates, terpenes, and polyphenols, among others) in response to biotic and abiotic stresses, which are mainly exploited as deterring agents against herbivores, insects, or pathogens. Microorganisms found across several ecological niches, including animals' organs, have been discovered

capable of detoxifying these metabolites and use them as energy, while providing benefit to their host via detoxifying symbiosis[19]. In this view, Berasategui et al. reported the ability of the GM of the insect *Hylobius abietis* (Coleoptera), a species that feeds on conifers rich in terpenoid resins that are toxic to other insects, to degrade the diterpene acids of Norway spruce, thanks to the presence of several genes of a previously described diterpene degradation gene cluster[29]. Consistently with this, a full taxonomic and functional profiling of the diamondback moth *Plutella xylostella* (Diptera), obtained by metagenomic sequencing, revealed the important role of specific gut bacteria (e.g., *Enterobacter cloacae*, *Enterobacter asburiae*, and *Carnobacterium maltaromaticum*) in the breakdown of plant cell walls, detoxification of plant phenolics, and synthesis of amino acids. This evidence demonstrates the beneficial interactions between insects and their GM, that may also represent potential biotechnological targets to provide novel pest management approaches[30]. Similar findings emerge also when targeting herbivorous mammals, as shown by Kohl and colleagues. In their research, they investigated the GM of desert woodrats (*Neotoma lepida*), whose diet consists of highly toxic bushes, highlighting the crucial role of gut microbes in allowing herbivores to consume toxic plants. In particular, ingested toxins prompted a plastic response of the microbial community by increasing the abundance of genes devoted to metabolizing them, and this adaptation mechanism seemed to be disrupted by antibiotics. These results indicate that microbes can enhance the ability of hosts to eat plant secondary compounds, expanding their dietary niche and opening new intriguing biotechnological avenues[31]. Overall, the above-mentioned works emphasize the ability of herbivore-associated microbes to neutralize plant toxins, underlining that increasing knowledge of these microbial traits could be exploited for improving animal feed safety, developing probiotics for livestock to prevent poisoning, and creating bioengineering strategies to detoxify crops for human consumption[19,32].

In addition, antimicrobial compounds produced by the GM of wild herbivores could potentially be developed into new antibiotics and antimicrobial agents. In this context, Garcia-Gutierrez and collaborators provided an extensive review based on the assumption that close co-evolution of microbes in the gut has led to the development of specialized antimicrobials that may serve as novel alternatives to traditional antibiotics, possibly mitigating the global concern of antimicrobial resistance (AMR). In fact, the gastrointestinal tract represents a suitable target for finding novel antimicrobials, due to the vast array of microbes that inhabit it, and whose coexistence is functionally balanced by symbiotic or antagonistic relationships. In the gut habitat, competitive responses are in fact fueled through the production of antimicrobial agents against other organisms occupying the same environmental niche. Specifically, gut microorganisms have been seen capable of synthesizing a plethora of bioactive antimicrobials, including both non-ribosomal peptides (NRPs) and ribosomally-synthesized bacteriocins like lasso peptides, sactibiotics, lantibiotics, bacteriolysins, and so on[33]. Considering that wild herbivores GMs are amongst the most complex microbial consortia on Earth, it is more than reasonable to believe that these vibrant ecosystems may hide the presence of countless new antimicrobial agents that could be exploited to address the current growing health challenge of antibiotic resistance. Corroborating these ideas, Youngblut and others performed an unparalleled *in silico* investigation, based on large-scale metagenome assemblies combined with bioinformatic approaches for secondary metabolite detection, in order to characterize novel gut microbial diversity for bioprospecting of bioactive natural products, catalytic and carbohydrate-binding enzymes, new probiotics, as well as unknown potential pathogens and AMR pathways. Guided by the still limited knowledge of these aspects in non-model holobionts, researchers

developed an extensive metagenome assembly-based pipeline which was tested across five vertebrate classes (namely Mammalia, Aves, Reptilia, Amphibia, and Actinopterygii), with the vast majority of samples obtained from wild individuals, including herbivores. Besides suggesting that a great portion of the genetic diversity in vertebrates' gut is still unknown and regarded as 'dark matter', this multi-species dataset identified almost 2,000 biosynthetic gene clusters (BGCs) spread across 1,522 reconstructed species-level genome bins (SGBs). Strikingly, most of the detected clusters encoded for non-ribosomal peptide synthetases (NRPSs) and ribosomally-synthesized and post-translationally modified sactipeptides (RiPPs), which are both involved in antibiotic defence mechanisms, as well as for carbohydrate-active enzymes and AMR markers. Therefore, this comprehensive work substantially expands the known taxonomic and functional diversity of the vertebrate GM and may help guide future natural product discovery, bioprospecting of novel carbohydrate-active enzymes, as well as elucidate AMR transmission routes starting from wild animals and their microbiomes[34].

Finally, one of the major potential benefits of studying wild herbivore microbiomes is the discovery of novel enzymes and metabolic pathways that could be optimized for breaking down and valorizing plant material at the industrial level. Enzymes such as cellulases, hemicellulases, and lignases, for example, could be used in industrial processes for biofuel/biogas production, platform chemicals generation, waste degradation, and plant-based food processing[6,7] (further aspects related to plant biomass bioconversion will be further discussed in the following paragraphs of this thesis).

Despite all these huge potentials, recently unveiled thanks to the groundbreaking technological advances of the past decades, studying the GM of wild herbivores still poses logistical challenges, such as capturing samples in remote or difficult-to-access environments, as well as ethical concerns related to invasive sampling methods, that should be minimized in favor of non-invasive techniques[15,16]. These and other drawbacks must be overcome in order to enrich our understanding of herbivores ultra-complex microbial communities across wild natural habitats and for unlocking their full potential. In fact, future efforts aimed in this direction could not only boost animal conservation biology by understanding how microbiomes contribute to the health and survival of endangered species, but will be also pivotal to shed light on microbiome-based biotechnologies addressed towards human health and environmental sustainability[35].

2.1.2 Lignocellulose-degrading natural microbial communities: a focus on herbivores

Amongst the many functions performed by complex microbial communities found in natural habitats, including herbivores gastrointestinal tract, lignocellulose (LC) degradation and valorization certainly stand out. Lignocellulose is the complex polymeric structure that makes up plant cell walls (PCW), and has evolved in order to provide shape and rigidity to plants thriving in different ecological niches[36,37,38,39,40]. LC is a very complex molecule primarily composed of three energy-rich biopolymers, namely cellulose, hemicellulose, and lignin, which are found in different proportions according to plant species, age, and phenological stage, and are hierarchically arranged to build up a solid matrix in plant cells. Wood in angiosperm trees, for example, generally contains 42-50% cellulose, 25-30%

hemicellulose, and 20-25% lignin, with the remaining portions consisting of molecules in smaller amounts[36,37]. Cellulose is a linear polymer made of glucose units linked by β-1,4-glycosidic bonds which forms crystalline microfibrils, providing rigidity and strength to the plant cell wall, but making it resistant to microbial degradation and enzymatic deconstruction. Hemicellulose is a heterogeneous highly branched polysaccharide composed of various C5 and C6 sugars (e.g., xylose, arabinose, mannose, galactose, glucose, and glucuronic acid) with a typically amorphous structure, serving as a matrix around cellulose fibers and providing flexibility and porosity to the cell. Overall, hemicellulose has a lower degree of polymerization and is less crystalline than cellulose, making it more easily metabolized. Finally, lignin is a complex, three-dimensional polyphenolic polymer resulting from the enzymatic polymerization of phenolic monomers (referred to as 'monolignols'), which are primarily p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol (linked via C-O and C-C bonds), and whose abundance varies depending on plant species. Lignin is highly heterogeneous, cross-linked, and amorphous, acting as a glue that binds cellulose and hemicellulose chains together, providing structural integrity to the cell wall, but is also highly resistant to degradation by most organisms and severely recalcitrant. Furthermore, other compounds can be found together with cellulose, hemicellulose, and lignin in PCW, adding further structural heterogeneity, and these include, for example, pectin[6,38,39,40,41,42]. The ultra-complex molecular structure of LC presents both opportunities and challenges for its use as a renewable biomass source in a circular economy perspective[41,42]. Indeed, LC can be regarded as the most abundant, nutrient-rich, renewable substance on the planet, derived mainly from agriculture, forestry, and paper/pulp industries, that regrettably accumulates in large quantities as unused valuable waste-streams[42,43,44]. According to the Food and Agriculture Organization (FAO), for instance, of the hundred million tons of agriculturally-derived LC each year, roughly 90%, including empty fruit bunches, fibers, fronds, trunks, kernels, etc., is discarded as waste instead of being properly taken advantage of[36].

Despite this, LC represents a very promising feedstock for biofuel production and other green biotechnological applications, but, unfortunately, the industrial-scale conversion of LC still faces significant obstacles[43,45]. These include, above all, pre-treatment and enzyme production costs, the need for anaerobic environments, the formation of inhibitory compounds, and the consequent inefficiency of fermentation processes. In particular, being LC highly resistant to degradation, pre-treatment methods (e.g., mechanical, chemical, or enzymatic) are required to break open the lignin barrier and increase the accessibility of cellulose and hemicellulose. Among those, chemical pre-treatments (carried out mainly through strong acid or alkaline reagents) have been observed to be effective, but may lead to the formation of inhibitory compounds (e.g., furfural and phenolic derivatives) which can inhibit downstream microbial or enzyme activity during fermentation. Dealing with these inhibitors requires additional detoxification processes, which can add to the already expensive costs of LC deconstruction procedures. Besides, these processes generate waste molecules which can pose environmental hazard and must be properly disposed[43,44,45,46]. From the enzymatic point of view, enzyme mixtures are necessary to degrade and bio-transform cellulose, hemicellulose, and lignin, but the production of high-yield, un-inhibited, cost-effective enzymes remains a severe challenge. Overall, the enzymatic degradation of LC is a multi-step pathway that involves several enzymatic classes. Cellulose is broken down by cellulases mainly referred to as endoglucanases (which cleave internal β-1,4-glycosidic bonds), exoglucanases (which remove cellobiose units from the ends of the cellulose chains), and β-glucosidases (which hydrolyze

cellobiose into two glucose molecules). In contrast, hemicellulases, like xylanases and mannanases, cleave the hemicellulose backbone, producing simpler sugars such as xylose, arabinose, or mannose, which can be further metabolized and valorized. Finally, lignin degradation, that is more challenging due to lignin aromatic cross-linked structure, requires enzymes such as ligninases, peroxidases, and laccases, which catalyze the oxidation of aromatic compounds and break down the internal lignin bonds. Interestingly, lignin degradation produces a variety of smaller aromatic compounds (e.g., vanillin and syringaldehyde) that can be further converted into valuable chemicals or biofuels[36,37,38,45,46,47]. A very thorough and updated list of all the enzyme classes involved in carbohydrate and lignin metabolism is reported in the online version of the CAZy database (http://www.cazy.org/). These classes include glycoside hydrolases, glycosyl transferases, polysaccharide lyases, carbohydrate esterases, auxiliary activities redox proteins, as well as a plethora of carbohydrate-binding modules, and they all perform their catalytic/binding activity towards sugar/phenolic chains with different chemical structures and found widespread in the biosphere[48]. Despite the huge number of proteins currently characterized, scaling up their catalytic potential from the lab to the industry is posing many problems, also related to different biomass types, since variations in LC composition (e.g., from different plant sources) can severely affect overall process efficiency[43,44,49].

Aiming to overcome cost, time, impact, and efficiency-related issues of LC biotechnological transformation, bio-based methods involving microorganisms (and their enzymes) have been extensively studied, with very promising results (*learning from nature*). Microbial metabolism of LC consists of three main steps, i.e., hydrolysis, acidogenic fermentation, and methanogenesis (summarized in **Fig. 1.2**)[43,45,50,51].

Specifically, hydrolysis is the initial and crucial step in LC degradation, during which microorganisms break down the complex polymeric matrix into simpler monomeric sugars/alcohols and small oligosaccharides. In this phase, cellulose and hemicellulose are deconstructed into their specific building units, such as hexoses (like glucose and galactose) and pentoses (like xylose, mannose, and arabinose)[50,51]. To do so, bacteria and other microbes rely on two main kinds of hydrolytic machineries, i.e., cellulosomes and free secreted proteins. As a matter of fact, while secreted enzymes act individually to digest biomass, cellulosomes, which were first identified in Clostridia, are multienzyme complexes that work together to efficiently degrade cellulose and hemicellulose. These complexes are made up of several different enzymatic subunits and scaffold proteins, tightly anchored to the bacterial cell wall thanks to the interactions among cohesin and dockerin modules, that are coordinated within a single molecular super-structure, enhancing the hydrolysis of the targets. Synergy between freely diffusive and cell surface-tethered cellulosomal systems may result in increased cellulose and hemicellulose deconstruction compared with the single pathways alone[52,53,54]. At the same time, lignin-degrading microorganisms can hydrolyze lignin aromatic structure into smaller phenolic compounds, which may serve as carbon and energy sources for other (syntrophic) microbes. Enzymes devoted to this task encompass lignin peroxidases, manganese peroxidases, laccases, oxidases, and several other oxidative proteins that generate reactive radical species that can degrade the lignin structure[36,37,42,44].

Once simple monomers become available, they are fermented by acidogenic microorganisms (typically anaerobic bacteria or fungi) that are capable of converting them into SCFAs (mainly acetate, propionate, butyrate, and valerate) and other intermediate products like methanol, lactate, succinate, molecular hydrogen (H_2), and carbon dioxide

(CO_2), by means of several different metabolic pathways. This step is crucial for the anaerobic degradation of LC, as acidogenesis-derived shorter-chain products can be used by other microorganisms in subsequent stages[10]. Building on previous scientific literature, amongst the many microbes capable of hydrolyzing and fermenting LC, the most common bacteria include *Bacteroides*, *Parabacteroides*, *Ruminococcus*, *Ruminobacter*, *Ruminiclostridium*, *Enterococcus*, *Bacillus*, *Paenibacillus*, *Treponema*, *Prevotella*, *Eubacterium*, *Clostridium*, *Lachnoclostridium*, *Lachnospira*, *Cellvibrio*, *Luteimonas*, *Fibrobacter*, *Arthrobacter*, *Pseudomonas*, *Selenomonas*, *Butyrivibrio*, *Pseudobutyrivibrio*, *Acetivibrio*, *Cellulomonas*, *Proteiniphilum*, *Fermentimonas*, *Christensenella*, and *Coprococcus* (primarily belonging to the phyla Proteobacteria, Firmicutes, and Bacteroidetes), while the most common fungi belong to the genera *Orpinomyces*, *Neocallimastix*, *Piromyces*, *Caecomyces*, and *Anaeromyces* (Neocallimastigomycota)[10,37,43,45,47,48,49,50,55].

Methanogenesis is typically the final step of the cascade, as it converts the volatile fatty acids and gases previously synthesized into methane (CH_4), a high-energy product that can be used as a renewable biofuel, completing the anaerobic LC degradation process[56,57]. From a biochemical perspective, methanogenesis may occur in different ways, but is always confined to strictly anaerobic niches, such as the digestive tracts of herbivores (e.g., ruminants and termites), sludges, or anaerobic digesters used for biogas production. Acetoclastic methanogens, for example, utilize acetic acid to produce methane, whereas hydrogenotrophic methanogens rely solely on hydrogen gas and carbon dioxide for the same purpose. Moreover, some methanogens can use methylated compounds (e.g., methanol or methylamines) as substrates for methane production[55,56,57]. Archaea, and particularly those of the phylum Euryarchaeota, are the only known methanogens in nature, and some archaeal methanogenic taxa include *Methanosaeta*, *Methanococcus*, *Methanosarcina*, *Methanoculleus*, *Methanothermobacter*, *Methanobrevibacter*, *Methanospirillum*, *Methanothrix*, *Methanosphaera*, *Methanomicrobium*, and so on. These organisms are uniquely adapted to produce methane through specialized biochemical pathways that are absent in bacteria and eukaryotes, playing vital roles for global carbon cycling[10,43,55,56,57].

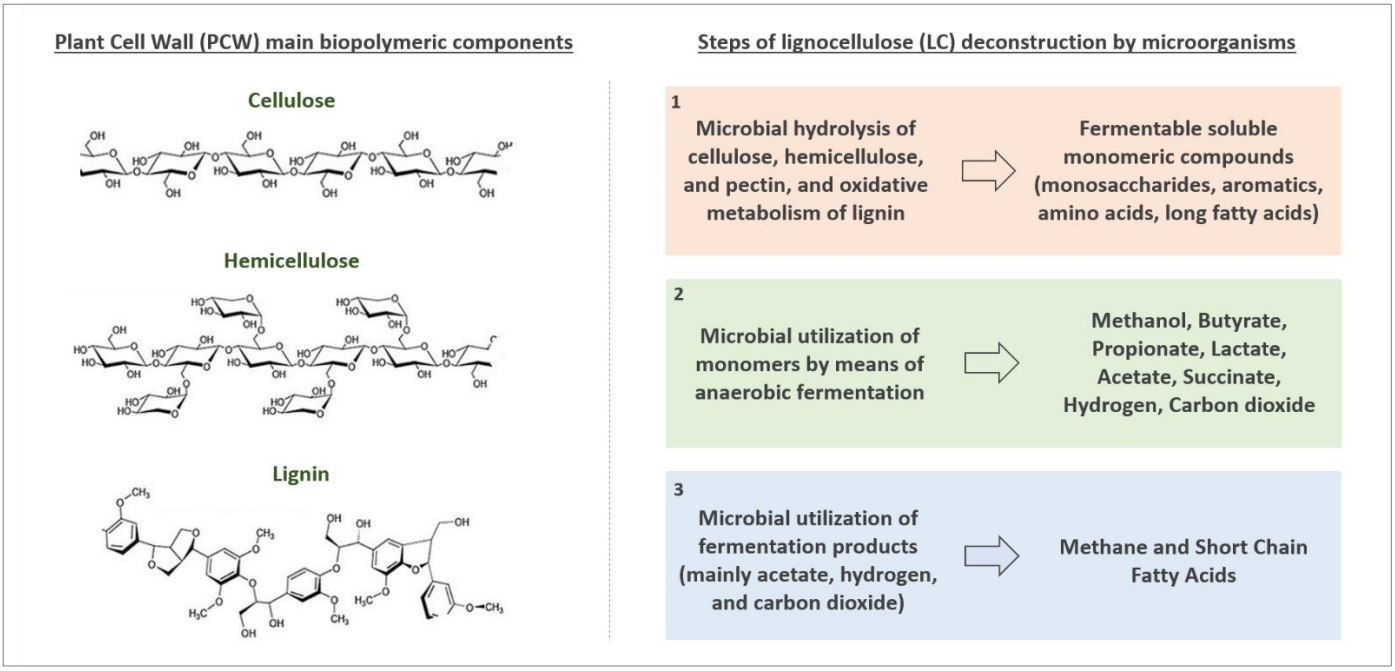


Fig. 1.2 – In the left portion of the figure, the molecular structure of the main biopolymers that make up plant cells, namely cellulose, hemicellulose, and lignin, is represented. On the right, microbial metabolism of the lignocellulose degradation cascade is summarized in boxes 1 (hydrolysis), 2 (acidogenic fermentation), and 3 (methanogenesis). The first trophic level includes the enzymatic degradation of the main polymers of the plant cell walls into small compounds by specialist cellulolytic, hemicellulolytic, and ligninolytic microorganisms (orange box). At the second level, soluble sugars and other metabolites are anaerobically fermented into short chain fatty acids (SCFAs), alcohols, organic acids, and gases (green box). Finally, microbial utilization of fermentation products leads to the production of methane and further SCFAs (blue box).

All the LC degradation steps listed above represent a multi-stage, microbial-driven process that occurs in natural environments, especially in soil/sediment/compost and animal guts, and that is fulfilled thanks to the multifaceted and dynamic ecological interactions that take place across natural microbiomes[10,43,58]. If exhaustively disentangled, these microbial traits can also be harnessed in industrial biotechnological applications for biofuel production and waste-streams valorization[43]. However, in spite of their extensive potential, many microbial strains still do not perform properly at the industrial level[59]. In this view, natural microbiomes, which have efficiently evolved to recycle organic polymers and mobilize nutrients in the biosphere, both in free-ranging and host-associated environments, represent a promising source of metabolic potential, which has not been fully elucidated and exploited yet[37,38,39,40]. In fact, several natural microorganisms, including bacteria, archaea, and fungi, have the ability to degrade LC to fermentable monomers, which can be further bio-transformed into value-added compounds, making natural microbiomes suitable for finding new and more effective strains, enzymes, and metabolic pathways devoted to this difficult task, with potential cascade effects on human economy[7,42,43,50]. In order to accomplish these goals, taxonomic analyses of the microbes involved in LC degradation are pivotal. Although a gene-centric approach, such as 16S rRNA, 18S rRNA, and ITS gene amplicon sequencing, can effectively characterize microbiomes structurally, it does not identify the microorganisms responsible for a metabolic process or their genes. In this context, whole metagenome sequencing, often coupled with assembly and binning strategies, is highly effective, especially for studying the uncultivated portion of LC-degrading microbiomes with diverse metabolic functions. Nevertheless, a fully extended characterization of microbiomes from different natural environments can only be appropriately carried out using a combination of integrated omics techniques (including meta-transcriptomics, meta-proteomics, and metabolomics), together with efficient statistical and computational tools for large-scale-generated data interpretation. For example, shotgun metagenomics coupled with metabolomics has aided in the investigation and understanding of the metabolic capabilities of microbial populations for LC biomass degradation[43]. Overall, despite some degree of error-susceptibility, mostly caused by inadequate sample size and non-standardized molecular and bioinformatic pipelines, integration of meta-omics sciences is essential to delve deeper into LC-degrading microbial communities, aiming to take advantage of their full potential[43,49,50,51,60]. Many microbial ecosystems have been explored in this sense, including manure[61,62,63], sewage sludge[64,65], municipal solid waste[66,67], food waste[68,69], sediment and soil[70,71,72], etc., and, among those, a very promising target is represented by the digestive tract of herbivorous animals, being herbivores evolutionarily-specialized plant food eaters[16,21]. Indeed, several studies have suggested the importance of biomimicking herbivorous gut systems to achieve better LC digestion performances at the industrial level[43,64,73]. Peng

and colleagues, for example, conducted an extremely comprehensive study that shed light on the ability of fungal and bacterial consortia from goat feces to breakdown LC. Remarkably, their work included parallel enrichment experiments with different LC substrates, 16S-V4 and ITS2 metabarcoding, shotgun metagenomics, and metabolomics. The authors were able to reconstruct metagenome-assembled genomes (MAGs) for both prokaryotes (Bacteria and Archaea) and anaerobic fungi (i.e., low-abundance members of the digestive tract that contain a wealth of biomass-degrading enzymes). Despite being ascribed to typical herbivore GM phyla (i.e., Firmicutes, Bacteroidetes, Euryarchaeota, and Neocallimastigomycota), more than 90% of these genomes were for the first time annotated at the species level, and encoded a plethora of carbohydrate-active enzymes (CAZymes), mostly belonging to the glycoside hydrolase (GH) and carbohydrate esterase (CE) classes. Aiming to decipher the actual metabolic capacity of microbial consortia originated from goat feces, different microbial cultures were set and their metabolite layout was evaluated, indicating that both bacteria and fungi, despite some inter-domain differences, were able to efficiently digest different LC feedstocks while producing value-added compounds such as sugars, methane, and SCFAs. Overall, this research emphasized that herbivores GMs may serve as a rich and untapped resource of strains, pathways, and enzymes that could be applied to convert plant waste into sugar substrates, biogases, and SCFAs for green biotechnology[49]. Besides, although a number of recent studies on ruminants have used metagenomics to assess the metabolic potential in the rumen (whose microorganisms are hard to cultivate), less attention has been paid to the hindgut, adding more scientific novelty to Peng's investigation. In fact, microbes from this habitat are partly derived from the rumen but may be better adapted to deal with more recalcitrant plant material that is not completely processed in the foregut, making them also more robust and resilient in culture[49]. One noteworthy rumen metagenomic survey considering both captive (*Bos Taurus* and *Ovis aries*) and wild (*Rangifer tarandus* and *Cervus elaphus*) herbivores has been provided by Glendinning et al., where the researchers reconstructed 391 MAGs from 16 microbial phyla belonging to 279 novel species, and identified several CAZymes and polysaccharide utilization loci (PUL) within these sequences[7]. From a more applied perspective, Ariaeenejad and co-workers managed to isolate and characterize a novel alkali-thermostable xylanase from the camel rumen that could be exploited industrially. Following cloning, purification, and structural/functional characterization, the isolated enzyme showed high thermal stability, good activity in a broad range of pH and temperature, and high effectivity in recalcitrant LC biomass degradation[74]. In line with these findings, promising candidate enzymes for LC degradation were discovered in other microbial symbionts of plant-feeding animals, such as the widely-studied termites[75,76,77], as well as moths[77], isopods[6], fish[78], wallaby[79], yak[80], buffalo[81], cow[50,82], sheep[83], elephants[84], and so on[85]. In particular, a recent survey showed that herbivores natural microbiomes are best suited to carry out LC digestion rather than biogas reactors. Specifically, by means of metagenomics and meta-transcriptomics, it was seen that microbiomes found in fecal/ruminal samples from elephants and cows, respectively, outperformed industrial biogas reactors communities, fed with maize silage, cow manure, and chicken manure, in terms of LC hydrolysis rates and strategies, with the latter showing lower abundance of glycoside hydrolases and carbohydrate esterases compared to the two investigated natural plant biomass-degrading systems[84]. In addition, Bredon and collaborators shed light on the wealth of CAZymes and lignin-modifying enzymes in invertebrate herbivores, via shotgun metagenome sequencing of different isopod species. In the holobiont framework, isopods (Crustacea) and termites (Isoptera) are excellent models to study LC degradation, harboring

diverse and rich microbial communities in their digestive tissues, which co-evolved with the host to cope with strict herbivory[6,75]. In this research, both freshwater and terrestrial isopods have been discovered to host microorganisms encoding all the currently known CAZymes classes, organized in PULs and cellulosomes, further supporting the idea that herbivores could be an interesting source of valuable enzymes for biotechnological industries of biomass conversion, aiming to obtain value-added compounds from waste while reducing the impact of fossil fuel-based processes[6]. To sum up, an ever-increasing number of studies showed that LC-degrading microbiomes from herbivores gut could provide potentially beneficial features for industrial applications, that could lead, in the future, to the sustainable and green utilization of LC, whose valorization to platform chemicals will be briefly discussed in the paragraph below.

2.1.3 Microbiome-based production of platform chemicals from lignocellulose biomass

Microbiome-based valorization of LC plant biomass to platform chemicals represents a new captivating frontier for biotechnological research in the light of circular green economy[10,71,86]. As a matter of fact, scaled-up implementation of bio-based naturally-occurring microbial processes to obtain commercially high-value compounds may open up new ways to reduce environmental impacts and costs derived from the use of fossil fuels as raw materials[56,64,68]. In this intricate landscape, attention has been paid to the potential of microbial biorefineries for the production of SCFAs starting from abundant zero-cost LC waste[86,87,88].

SCFAs are emerging as valuable platform chemicals due to their versatility in producing a wide range of industrially relevant compounds, and can be derived from renewable biomass sources through microbial fermentation[87,88]. In general, SCFAs serve as precursors for the synthesis of bio-based chemicals such as biofuels (e.g., ethanol, butanol), monomers for building biodegradable plastics (e.g., polyhydroxyalkanoates), specialty chemicals (e.g., butyrolactone, hexanoic acid), and pharmaceuticals (e.g., acetylsalicylic acid, fluoro-propionate, gamma-hydroxybutyrate). Furthermore, SCFAs can be converted into high-value materials, including bio-composites and bio-surfactants, making them integral for the development of sustainable industrial processes. Their broad availability, low-cost production potential, and chemical reactivity underline the efficacy of SCFAs as key intermediates for the transition to a more sustainable, circular bioeconomy[68,89,90,91,92]. A schematic overview of useful products derived from SCFAs after microbial biomass digestion and fermentation is provided in **Fig. 2.2**.

Specifically, acetate (C2), propionate (C3), and butyrate (C4) can be regarded as the main fermentation products of LC microbial metabolism[10,11,14], produced by microorganisms after hydrolysis of complex polymeric chains, and are considered potential platform chemicals. Platform chemicals are molecules that act as central building blocks in the production of a broad array of chemical products, by undergoing further chemical or biological transformations, making them key intermediates in the production of diverse industrial materials. SCFAs are considered platform chemicals due to their ability to serve as starting points for the synthesis of numerous valuable compounds across different industries[89,90,91,92]. In particular, acetate serves as a key building block in a variety of industrial applications, including biofuel, bioplastics, synthetic fibers, solvents, fragrances, pharmaceuticals, etc. Looking at the possible

industrial uses of propionate, it can be considered a precursor to produce textiles, biodegradable plastics, food additives, antifreeze agents, superabsorbent polymers, and paints, for instance. Finally, butyrate can be exploited for producing fuels, solvents, bioplastics, pharmaceuticals, specialty chemicals, bio-based surfactants, detergents, and so on. For all these reasons, microbiome-derived SCFAs are key players for offering renewable, environmentally friendly alternatives to petroleum-based substances, enhancing the environmental sustainability of industrial chemical production[89,93,94,95].

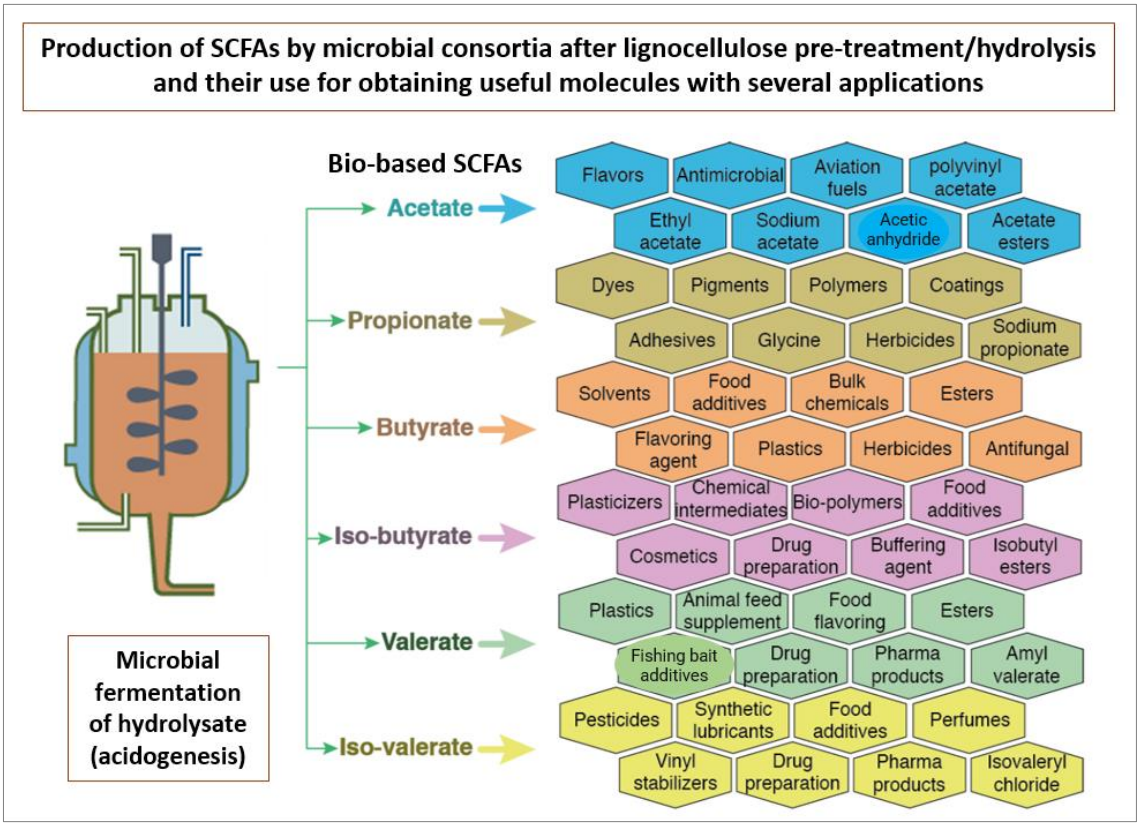


Fig. 2.2 – Graphical summary of the main possible uses of bio-based SCFAs (i.e., the microbial acidogenic fermentation endpoints) as high-value-added intermediates for several human purposes (e.g., lipids, bioplastics, proteins, pharmaceuticals, ethanol, and antimicrobials production). A simplified integrated biorefinery for feasible and convenient next-generation processes providing sustainable solutions for global waste-streams valorization is shown. In brief, the process will begin with the introduction of waste/hydrolysate into the acidogenesis reactor, where SCFAs are produced by tailored microbial mono- or mixed-cultures. Then, optimization of SCFAs extraction will be needed for ensuring further valorization (in a mixture or after separation of the different acids produced by microorganisms, and in either batch or fed-batch fermentation processes). This circular cascade encompasses biosynthesis, recovery, separation (optional), and conversion of SCFAs starting from lignocellulose or other waste biomass (modified from[89]).

Thus, microbial production of short-chain fatty acids from lignocellulosic biomass represents a promising alternative for the generation of bio-based chemicals[92,94,95]. In this view, advances in metabolic engineering and synthetic

(micro)biology are paramount for the optimization of microbial strains/consortia able to enhance the yield and selectivity of SCFAs production during fermentation[2,89,90]. Shahab et al. addressed the issue starting from the assumption that microbial consortia are a more suited alternative to monocultures for complex bio-transformations of LC to SCFAs, thanks to the division of labor that occurs in microbiomes and to their less susceptibility to contamination than pure cultures. Attempting to mimic natural mechanisms and their biodiversity, the authors engineered simplified synthetic microbial consortia composed of a cellulolytic fungus (*Trichoderma reesei*), facultative anaerobic lactic acid bacteria that funneled carbohydrates to lactate generation, and three anaerobic lactate-fermenting SCFA-producing strains (i.e., *Clostridium tyrobutyricum*, *Veillonella criceti*, and *Megasphaera elsdenii*). Sequential inoculation schemes were set starting from different types of LC biomass and exploiting syntrophic relationships between these heterogeneous microbes. During the experiments, after LC was digested and anaerobic conditions were reached in the reactor, SCFA-producing strains could efficiently synthesize a wide range of valuable molecules, including acetate, propionate, butyrate, valerate, hexanoate, ethanol, propanol, butanol, 1,2-propanediol, polyhydroxybutyrate, and lipids. These intriguing results expand the available toolbox for successfully engineering stable and controllable synthetic microbial communities capable of bio-transforming LC to value-added compounds[96]. Consistently with this, Nguyen and colleagues compared the ability of two different microbial consortia to produce biogas and SCFAs starting from four different types of LC biomass. In this survey, both rumen fluid and anaerobic sludge microbiomes were considered, with the first found to be far more efficient, indicating the possible application of rumen microorganisms for SCFAs generation from plant biomass. In fact, the high yield of SCFAs in the rumen fluid reactor was probably related to the high abundance of specific hydrolytic and acidogenic bacteria like *Fibrobacter* and *Prevotella*. These findings suggest that the use of rumen fluid microorganisms, together with tailored extraction of the produced acids, may represent an alternative solution to enhance the environmental and economic benefits of LC waste-streams disposal[97]. Analogous results were seen also considering other LC-rich matrices, such as municipal waste-streams[98] and food waste[99,100], for instance, further demonstrating the microbial potential for the production of bio-based products from organic wastes as alternatives to fossil-based products. Looking at animal-associated microbiomes, a multi-omics survey targeting bamboo rats (*Rhizomys pruinosus*), i.e., herbivorous holobionts with LC-rich bamboo-based diets, confirmed that natural GMs may be a precious source of diverse CAZymes associated with LC degradation and biosynthesis of SCFAs, amino acids, and vitamins[101].

Overall, while the microbial production of SCFAs from renewable feedstocks, including LC biomass, holds significant promise as an industrial platform, there are still several challenges that need to be addressed for scaling up these processes and making them economically viable, while reducing costs, environmental risk, carbon footprint, and time need[89,90]. These obstacles include, for example, difficulties in microbial strain/enzyme development, better biomass pretreatment methods, and more efficient SCFAs purification systems[91,92]. In fact, technologies that, after optimization of SCFAs biorefineries, will enable the selective recovery of these compounds from fermentation broths (such as membrane filtration, ion-exchange chromatography, or adsorption technologies) would significantly improve the economic competitiveness of microbiome-based SCFAs production[43,89,94]. In order to overcome these several hurdles, metabolic engineering, fermentation optimization, multi-omics, and synthetic biology tools are being used, often starting by deepening our current knowledge of ultra-rich natural microbiomes. Furthermore, continued

research into integrated biorefinery models by combining the production and recovery of SCFAs with other valuable co-products may help enhance the overall economic feasibility and sustainability of the process[89,96,102]. In summary, microbial digestion of LC biomass to produce SCFAs is a promising biotechnological avenue that aligns with the goals of sustainability, renewable energy, and green chemistry, and may boost the implementation of biobased circular economies by providing valuable chemicals as well as innovative solutions for waste management.

References

[1]Biagi, Elena, et al. "Gut microbiota and extreme longevity." *Current Biology* 26.11 (2016): 1480-1485.

[2]Vrancken, Gino, et al. "Synthetic ecology of the human gut microbiota." *Nature Reviews Microbiology* 17.12 (2019): 754-763.

[3]Shanahan, Fergus, Tarini S. Ghosh, and Paul W. O'Toole. "The healthy microbiome—what is the definition of a healthy gut microbiome?." *Gastroenterology* 160.2 (2021): 483-494.

[4]Kuziel, Gavin A., and Seth Rakoff-Nahoum. "The gut microbiome." *Current Biology* 32.6 (2022): R257-R264.

[5]Glaeser, Stefanie P., et al. "A preliminary comparison on faecal microbiomes of free-ranging large baleen (Balaenoptera musculus, B. physalus, B. borealis) and toothed (Physeter macrocephalus) whales." *Microbial ecology* 83.1 (2022): 18-33.

[6]Bredon, Marius, et al. "Isopod holobionts as promising models for lignocellulose degradation." *Biotechnology for biofuels* 13 (2020): 1-14.

[7]Glendinning, Laura, et al. "Metagenomic analysis of the cow, sheep, reindeer and red deer rumen." *Scientific reports* 11.1 (2021): 1990.

[8]Alessandri, Giulia, et al. "Creating an atlas to visualize the biodiversity of the mammalian gut microbiota." *Current Opinion in Biotechnology* 73 (2022): 28-33.

[9]Russell, James B., and Jennifer L. Rychlik. "Factors that alter rumen microbial ecology." *Science* 292.5519 (2001): 1119-1122.

[10]Mizrahi, Itzhak, R. John Wallace, and Sarah Moraïs. "The rumen microbiome: balancing food security and environmental impacts." *Nature Reviews Microbiology* 19.9 (2021): 553-566.

[11]Koh, Ara, et al. "From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites." *Cell* 165.6 (2016): 1332-1345.

[12]Deleu, Sara, et al. "Short chain fatty acids and its producing organisms: An overlooked therapy for IBD?." *EBioMedicine* 66 (2021).

[13]Zhang, Dan, et al. "Short-chain fatty acids in diseases." *Cell Communication and Signaling* 21.1 (2023): 212.

[14]Aluwong, Tagang, Patricia Ishaku Kobo, and Abdullahi Abdullahi. "Volatile fatty acids production in ruminants and the role of monocarboxylate transporters: a review." *African Journal of Biotechnology* 9.38 (2010): 6229-6232.

[15]Hird, Sarah M. "Evolutionary biology needs wild microbiomes." *Frontiers in microbiology* 8 (2017): 725.

[16]Dahl, Sarah-Alica, et al. "'"Get the best out of what comes in"'–adaptation of the microbiota of chamois (Rupicapra rupicapra) to seasonal forage availability in the Bavarian Alps." *Frontiers in Microbiology* 14 (2023): 1238744.

[17]Alessandri, Giulia, et al. "The impact of human-facilitated selection on the gut microbiota of domesticated mammals." *FEMS microbiology ecology* 95.9 (2019): fiz121.

[18]Zingg, Annatina. *Seasonal variability in the diet composition of alpine ibex (Capra ibex ibex L.) in the Swiss National Park*. Diss. Verlag nicht ermittelbar, 2009.

[19]Rogowska-van der Molen, Magda A., et al. "Microbial degradation of plant toxins." *Environmental microbiology* 25.12 (2023): 2988-3010.

[20]Kline, Olivia, and Neelendra K. Joshi. "Microbial Symbiont-Based Detoxification of Different Phytotoxins and Synthetic Toxic Chemicals in Insect Pests and Pollinators." *Journal of Xenobiotics* 14.2 (2024): 753-771.

[21]Couch, Claire E., et al. "Diet and gut microbiome enterotype are associated at the population level in African buffalo." *Nature communications* 12.1 (2021): 2267.

[22]Eddington, Hyrum S., et al. "Spatiotemporal variation in the fecal microbiota of mule deer is associated with proximate and future measures of host health." *BMC Veterinary Research* 17 (2021): 1-8.

[23]Su, Rina, et al. "Comparative study of the function and structure of the gut microbiota in Siberian musk deer and Forest musk deer." *Applied Microbiology and Biotechnology* 106.19 (2022): 6799-6817.

[24]Haworth, Sarah E., et al. "Space, time and captivity: quantifying the factors influencing the fecal microbiome of an alpine ungulate." *FEMS microbiology ecology* 95.7 (2019): fiz095.

[25]Xue, Zhengsheng, et al. "The bamboo-eating giant panda harbors a carnivore-like gut microbiota, with excessive seasonal variations." *MBio* 6.3 (2015): 10-1128.

[26]Wu, Qi, et al. "Seasonal variation in nutrient utilization shapes gut microbiome structure and function in wild giant pandas." *Proceedings of the Royal Society B: Biological Sciences* 284.1862 (2017): 20170955.

[27]Orkin, Joseph D., et al. "Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest." *The ISME journal* 13.1 (2019): 183-196.

[28]Ren, Tiantian, et al. "Seasonal, spatial, and maternal effects on gut microbiome in wild red squirrels." *Microbiome* 5 (2017): 1-14.

[29]Berasategui, Aileen, et al. "Gut microbiota of the pine weevil degrades conifer diterpenes and increases insect fitness." *Molecular ecology* 26.15 (2017): 4099-4110.

[30]Xia, Xiaofeng, et al. "Metagenomic sequencing of diamondback moth gut microbiome unveils key holobiont adaptations for herbivory." *Frontiers in microbiology* 8 (2017): 663.

[31]Kohl, Kevin D., et al. "Gut microbes of mammalian herbivores facilitate intake of plant toxins." *Ecology letters* 17.10 (2014): 1238-1246.

[32]Berasategui, Aileen, et al. "Potential applications of insect symbionts in biotechnology." *Applied microbiology and biotechnology* 100 (2016): 1567-1577.

[33]Garcia-Gutierrez, Enriqueta, et al. "Gut microbiota as a source of novel antimicrobials." *Gut microbes* 10.1 (2019): 1-21.

[34]Youngblut, Nicholas D., et al. "Large-scale metagenome assembly reveals novel animal-associated microbial genomes, biosynthetic gene clusters, and other genetic diversity." *Msystems* 5.6 (2020): 10-1128.

[35]Trevelline, Brian K., et al. "Conservation biology needs a microbial renaissance: a call for the consideration of host-associated microbiota in wildlife management practices." *Proceedings of the Royal Society B* 286.1895 (2019): 20182448.

[36]Kumar, Raj, Sompal Singh, and Om V. Singh. "Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives." *Journal of industrial microbiology and biotechnology* 35.5 (2008): 377-391.

[37]Cragg, Simon M., et al. "Lignocellulose degradation mechanisms across the Tree of Life." *Current opinion in chemical biology* 29 (2015): 108-119.

[38]Wilhelm, Roland C., et al. "Bacterial contributions to delignification and lignocellulose degradation in forest soils with metagenomic and quantitative stable isotope probing." *The ISME Journal* 13.2 (2019): 413-429.

[39]Bomble, Yannick J., et al. "Lignocellulose deconstruction in the biosphere." *Current opinion in chemical biology* 41 (2017): 61-70.

[40]Ransom-Jones, Emma, et al. "Lignocellulose-degrading microbial communities in landfill sites represent a repository of unexplored biomass-degrading diversity." *Msphere* 2.4 (2017): 10-1128.

[41]Brown, Margaret E., and Michelle CY Chang. "Exploring bacterial lignin degradation." *Current opinion in chemical biology* 19 (2014): 1-7.

[42]Espro, Claudia, et al. "Sustainable production of pharmaceutical, nutraceutical and bioactive compounds from biomass and waste." *Chemical Society Reviews* 50.20 (2021): 11191-11207.

[43]Basak, Bikram, et al. "Roles of engineered lignocellulolytic microbiota in bioaugmenting lignocellulose biomethanation." *Renewable and Sustainable Energy Reviews* 207 (2025): 114913.

[44]Bugg, Timothy DH, and Rahman Rahmanpour. "Enzymatic conversion of lignin into renewable chemicals." *Current opinion in chemical biology* 29 (2015): 10-17.

[45]Llamas, Mercedes, et al. "Microbial co-cultures for biochemicals production from lignocellulosic biomass: a review." *Bioresource Technology* 386 (2023): 129499.

[46]Lu, Jiasheng, et al. "Recent progress on bio-succinic acid production from lignocellulosic biomass." *World Journal of Microbiology and Biotechnology* 37 (2021): 1-8.

[47]Thapa, Santosh, et al. "Microbial cellulolytic enzymes: diversity and biotechnology with reference to lignocellulosic biomass degradation." *Reviews in Environmental Science and Bio/Technology* 19 (2020): 621-648.

[48]Lombard, Vincent, et al. "The carbohydrate-active enzymes database (CAZy) in 2013." *Nucleic acids research* 42.D1 (2014): D490-D495.

[49]Peng, Xuefeng, et al. "Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes." *Nature microbiology* 6.4 (2021): 499-511.

[50]Liang, Jinsong, et al. "Promising biological conversion of lignocellulosic biomass to renewable energy with rumen microorganisms: A comprehensive review." *Renewable and Sustainable Energy Reviews* 134 (2020): 110335.

[51]Bhujbal, Sachin Krushna, et al. "Biotechnological potential of rumen microbiota for sustainable bioconversion of lignocellulosic waste to biofuels and value-added products." *Science of the Total Environment* 814 (2022): 152773.

[52]Bayer, Edward A., et al. "The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides." *Annu. Rev. Microbiol.* 58.1 (2004): 521-554.

[53]Schwarz, W. "The cellulosome and cellulose degradation by anaerobic bacteria." *Applied microbiology and biotechnology* 56 (2001): 634-649.

[54]Minor, Christine M., et al. "A genomic analysis reveals the diversity of cellulosome displaying bacteria." *Frontiers in Microbiology* 15 (2024): 1473396.

[55]Leis, Stefanie, et al. "Finding a robust strain for biomethanation: anaerobic fungi (Neocallimastigomycota) from the Alpine ibex (Capra ibex) and their associated methanogens." *Anaerobe* 29 (2014): 34-43.

[56]Zhou, M. I., Emma Hernandez-Sanabria, and Le Luo Guan. "Assessment of the microbial ecology of ruminal methanogens in cattle with different feed efficiencies." *Applied and environmental microbiology* 75.20 (2009): 6524-6533.

[57]Patra, Amlan Kumar, and Ryszard Puchala. "Methane mitigation in ruminants with structural analogues and other chemical compounds targeting archaeal methanogenesis pathways." *Biotechnology Advances* (2023): 108268.

[58]Cavicchioli, Ricardo, et al. "Scientists' warning to humanity: microorganisms and climate change." *Nature Reviews Microbiology* 17.9 (2019): 569-586.

[59]Basak, Bikram, et al. "Lignocellulolytic microbiomes for augmenting lignocellulose degradation in anaerobic digestion." *Trends in Microbiology* 30.1 (2022): 6-9.

[60]Griffith, Gareth W., et al. "Anaerobic fungi: neocallimastigomycota." *IMA fungus* 1 (2010): 181-185.

[61]Chen, Hong, et al. "Biohythane production and microbial characteristics of two alternating mesophilic and thermophilic two-stage anaerobic co-digesters fed with rice straw and pig manure." *Bioresource Technology* 320 (2021): 124303.

[62]Li, Yu, Jing Zhao, and Zhenhua Zhang. "Implementing metatranscriptomics to unveil the mechanism of bioaugmentation adopted in a continuous anaerobic process treating cow manure." *Bioresource Technology* 330 (2021): 124962.

[63]Ozbayram, Emine Gozde, et al. "Comparison of rumen and manure microbiomes and implications for the inoculation of anaerobic digesters." *Microorganisms* 6.1 (2018): 15.

[64]Bize, Ariane, et al. "Shotgun metaproteomic profiling of biomimetic anaerobic digestion processes treating sewage sludge." *Proteomics* 15.20 (2015): 3532-3543.

[65]Beale, D. J., et al. "An 'omics' approach towards the characterisation of laboratory scale anaerobic digesters treating municipal sewage sludge." *Water research* 88 (2016): 346-357.

[66]Acedos, Miguel G., et al. "New efficient meta-fermentation process for lactic acid production from municipal solid waste." *Microbial Cell Factories* 21.1 (2022): 233.

[67]Carrillo-Barragán, Priscilla, et al. "The stability of ethanol production from organic waste by a mixed culture depends on inoculum transfer time." *Biochemical Engineering Journal* 166 (2021): 107875.

[68]Greses, Silvia, et al. "Genome-centric metagenomics revealed the effect of pH on the microbiome involved in short-chain fatty acids and ethanol production." *Bioresource Technology* 377 (2023): 128920.

[69]Bühlmann, Christopher H., et al. "Lactic acid from mixed food wastes at a commercial biogas facility: Effect of feedstock and process conditions." *Journal of Cleaner Production* 284 (2021): 125243.

[70]Sinha, Debasree, et al. "Enhanced biogas production from Lantana camara via bioaugmentation of cellulolytic bacteria." *Bioresource Technology* 340 (2021): 125652.

[71]Singh, Rajeev, et al. "Prospects of soil microbiome application for lignocellulosic biomass degradation: An overview." *Science of the Total Environment* 838 (2022): 155966.

[72]Fernández-Bayo, J. D., et al. "The initial soil microbiota impacts the potential for lignocellulose degradation during soil solarization." *Journal of applied microbiology* 126.6 (2019): 1729-1741.

[73]Bayané, Ali, and Serge R. Guiot. "Animal digestive strategies versus anaerobic digestion bioprocesses for biogas production from lignocellulosic biomass." *Reviews in Environmental Science and Bio/Technology* 10 (2011): 43-62.

[74]Ariaeenejad, Shohreh, et al. "Mining of camel rumen metagenome to identify novel alkali-thermostable xylanase capable of enhancing the recalcitrant lignocellulosic biomass conversion." *Bioresource technology* 281 (2019): 343-350.

[75]Brune, Andreas. "Symbiotic digestion of lignocellulose in termite guts." *Nature Reviews Microbiology* 12.3 (2014): 168-180.

[76]Tsukagoshi, Hikaru, et al. "Structural and biochemical analyses of glycoside hydrolase family 26 β-mannanase from a symbiotic protist of the termite Reticulitermes speratus." *Journal of Biological Chemistry* 289.15 (2014): 10843-10852.

[77]Brennan, YaLi, et al. "Unusual microbial xylanases from insect guts." *Applied and Environmental Microbiology* 70.6 (2004): 3609-3617.

[78]McDonald, Ryan, Harold J. Schreier, and Joy EM Watts. "Phylogenetic analysis of microbial communities in different regions of the gastrointestinal tract in Panaque nigrolineatus, a wood-eating fish." *PLoS One* 7.10 (2012): e48018.

[79]Pope, P. B., et al. "Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores." *Proceedings of the National Academy of Sciences* 107.33 (2010): 14793-14798.

[80]Bao, Lei, et al. "Cloning and characterization of two β-glucosidase/xylosidase enzymes from yak rumen metagenome." *Applied biochemistry and biotechnology* 166 (2012): 72-86.

[81]Zhong, Huimin, et al. "Whole-genome sequencing reveals lignin-degrading capacity of a ligninolytic bacterium (Bacillus cereus) from Buffalo (Bubalus bubalis) rumen." *Genes* 13.5 (2022): 842.

[82]Del Pozo, Mercedes V., et al. "Microbial β-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail." *Biotechnology for biofuels* 5 (2012): 1-13.

[83]Al-Masaudi, Saad, et al. "A metagenomics investigation of carbohydrate-active enzymes along the gastrointestinal tract of Saudi sheep." *Frontiers in Microbiology* 8 (2017): 666.

[84]Güllert, Simon, et al. "Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies." *Biotechnology for biofuels* 9 (2016): 1-20.

[85]Ozbayram, Emine Gozde, Sabine Kleinsteuber, and Marcell Nikolausz. "Biotechnological utilization of animal gut microbiota for valorization of lignocellulosic biomass." *Applied Microbiology and Biotechnology* 104.2 (2020): 489-508.

[86]Lee, Wee Shen, et al. "A review of the production and applications of waste-derived volatile fatty acids." *Chemical Engineering Journal* 235 (2014): 83-99.

[87]Shafiq, Muhammad, et al. "Integrated biorefinery approach for bioconversion of fish manure to volatile fatty acids and its valorization by Schizochytrium sp for docosahexaenoic acid, squalene, and carotenoids production." *Journal of Water Process Engineering* 58 (2024): 104749.

[88]Kumar, Bikash, et al. "Current perspective on pretreatment technologies using lignocellulosic biomass: An emerging biorefinery concept." *Fuel processing technology* 199 (2020): 106244.

[89]Koubaa, Mohamed. "Integrated Biorefinery for a Next-Generation Methanization Process Focusing on Volatile Fatty Acid Valorization: A Critical Review." *Molecules* 29.11 (2024): 2477.

[90]Varghese, Vijay K., et al. "A comprehensive review on current status and future perspectives of microbial volatile fatty acids production as platform chemicals." *Science of The Total Environment* 815 (2022): 152500.

[91]Ramos-Suarez, Maria, Yue Zhang, and Victoria Outram. "Current perspectives on acidogenic fermentation to produce volatile fatty acids from waste." *Reviews in Environmental Science and Bio/Technology* 20.2 (2021): 439-478.

[92]Bruni, Cecilia, et al. "Targeted bio-based volatile fatty acid production from waste streams through anaerobic fermentation: link between process parameters and operating scale." *ACS Sustainable Chemistry & Engineering* 9.30 (2021): 9970-9987.

[93]Aboudi, Kaoutar, et al. "Biogas, biohydrogen, and polyhydroxyalkanoates production from organic waste in the circular economy context." *Sustainable Biofuels*. Academic Press, 2021. 305-343.

[94]Bhatia, Shashi Kant, and Yung-Hun Yang. "Microbial production of volatile fatty acids: current status and future perspectives." *Reviews in Environmental Science and Bio/Technology* 16 (2017): 327-345.

[95]Chen, Zheng, et al. "Anaerobic fermentation of hydrothermal liquefaction wastewater of dewatered sewage sludge for volatile fatty acids production with focuses on the degradation of organic components and microbial community compositions." *Science of the Total Environment* 777 (2021): 146077.

[96]Shahab, Robert L., et al. "A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose." *Science* 369.6507 (2020): eabb1214.

[97]Nguyen, Luong N., et al. "Application of rumen and anaerobic sludge microbes for bio harvesting from lignocellulosic biomass." *Chemosphere* 228 (2019): 702-708.

[98]Owusu-Agyeman, Isaac, et al. "Volatile fatty acids production from municipal waste streams and use as a carbon source for denitrification: the journey towards full-scale application and revealing key microbial players." *Renewable and Sustainable Energy Reviews* 175 (2023): 113163.

[99]Wu, Yang, et al. "Enhanced volatile fatty acid production from food waste fermentation via enzymatic pretreatment: new insights into the depolymerization and microbial traits." *Acs Es&T Engineering* 3.1 (2022): 26-35.

[100]Dahiya, Shikha, et al. "Acidogenic fermentation of food waste for volatile fatty acid production with co-generation of biohydrogen." *Bioresource technology* 182 (2015): 103-113.

[101]Xiao, Kangpeng, et al. "Adaptation of gut microbiome and host metabolic systems to lignocellulosic degradation in bamboo rats." *The ISME Journal* 16.8 (2022): 1980-1992.

[102]Patel, Alok, et al. "Valorization of volatile fatty acids derived from low-cost organic waste for lipogenesis in oleaginous microorganisms-A review." *Bioresource Technology* 321 (2021): 124457.

2.2 *Capra ibex* L. – gut microbiome dynamics, ecology, and biotechnology

2.2.1 Aim of the research

The goal of the research illustrated in paragraph 2.2.2 is the characterization of the gut microbial consortia hosted by wild populations of the Alpine ibex (*Capra ibex*) across three different seasons at Stelvio National Park (Lombardia, Italy), aiming to unravel the microbial diversity and biotechnological potential of three different seasonal GMs from a previously neglected wild herbivorous ruminant. Fecal samples were collected in the wild in spring, summer, and autumn 2020 (**Fig. 3.2**), and the ibex GM was thoroughly analyzed with a set of multi-omics and bioinformatics methods. Being wild herbivores extremely dense microbial reservoirs, we wanted to explore the Alpine ibex GM biodiversity, in terms of prokaryotic taxa, genes, and metabolites, also focusing on microbial traits that could be potentially exploited to valorize plant LC waste to industrially-relevant platform chemicals. To do so, basic knowledge of GM ecology and dynamics in wild herbivores must be heightened. In fact, this knowledge will be paramount for disentangling all the intricate microbial interconnections of herbivorous holobionts, which may be replicated in the laboratory, and, eventually, scaled up as green, eco-friendly, circular industries able to produce high-value molecules from renewable waste-streams. Therefore, the work illustrated in this part of my thesis, despite having a proof-of-concept nature, may be a starting point for the development of microbiome-based integrated biorefineries for the bioconversion of lignocellulose to high-value platform chemicals. Furthermore, in paragraph 2.2.3, I provided a synthetic glimpse of the potential of the *C. ibex* seasonal gut communities to produce specific secondary metabolites, encoded by microbial biosynthetic gene clusters, that may be suited for the pharmaceutical market.

Fig. 3.2 – Field sampling of Alpine ibex feces in the wild at Stelvio National Park (Lombardia, Italy).

2.2.2 Study III – The Alpine ibex (*Capra ibex*) gut microbiome, seasonal dynamics, and potential application in lignocellulose bioconversion

This section of the thesis was published as an Original Research Article in *iScience* (2024); https://doi.org/10.1016/j.isci.2024.110194. Contents of the article have not been modified except for graphical purposes and bibliography format. Numeration of main and supplementary figures and tables is consistent with the original article.

Authors:

Enrico Nanetti[1,10], Daniel Scicchitano[1,2,10], Giorgia Palladino[1,2], Nicolò Interino[3], Luca Corlatti[4,5], Luca Pedrotti[4], Federica Zanetti[6], Elena Pagani[6], Erika Esposito[3], Alice Brambilla[7,8], Stefano Grignolio[9], Ilaria Marotti[6], Silvia Turroni[1], Jessica Fiori[3], Simone Rampelli[1,2], and Marco Candela[1,2,11*]

*Corresponding Author

Affiliations:

[1]Unit of Microbiome Science and Biotechnology, Department of Pharmacy and Biotechnology (FaBiT), Alma Mater Studiorum - University of Bologna, Via Belmeloro 6, 40126 Bologna, Italy

[2]Fano Marine Center, The Inter-Institute Center for Research on Marine Biodiversity, Resources and Biotechnologies, 61032 Fano, Italy

[3]Department of Chemistry ‘‘G. Ciamician’’, University of Bologna, Via Selmi 2, 40126 Bologna, Italy

[4]Stelvio National Park, 23032 Bormio, Italy

[5]University of Freiburg, 79098 Freiburg, Germany

[6]Department of Agricultural and Food Sciences, University of Bologna, Viale G. Fanin 44, 40127 Bologna, Italy

[7]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich (CH), Switzerland

[8]Centro Studi Fauna Alpina, Parco Nazionale Gran Paradiso, Loc. Degioz 11, 11010 Valsavarenche, Aosta, Italy

[9]University of Ferrara, Department of Life Science and Biotechnology, via Borsari 46, I-44121 Ferrara, Italy

[10]These authors contributed equally

[11]Lead contact

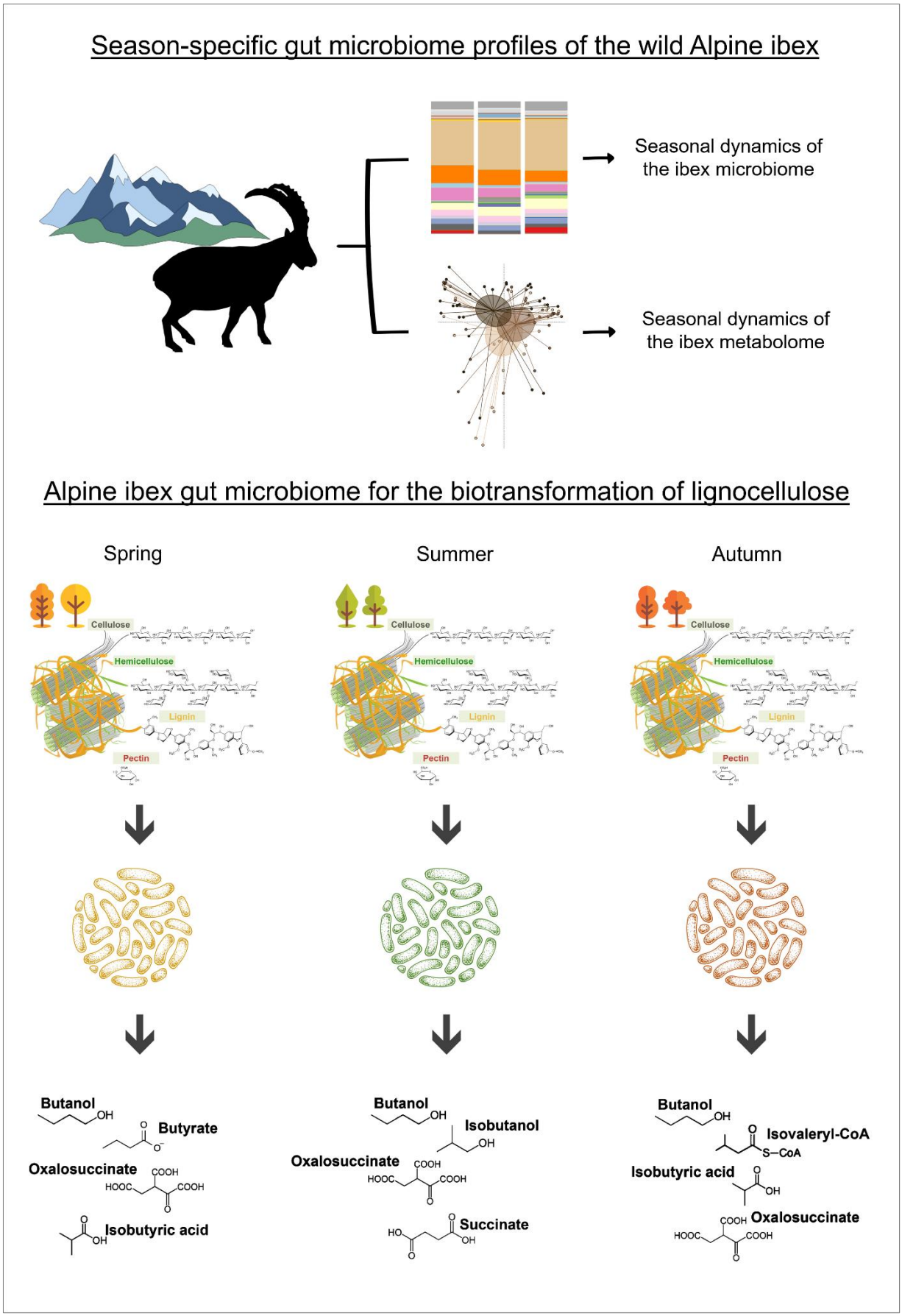
Summary

Aiming to shed light on the biology of wild ruminants, we investigated the gut microbiome seasonal dynamics of the Alpine ibex (*Capra ibex*) from the Central Italian Alps. Feces were collected in spring, summer, and autumn during non-invasive sampling campaigns. Samples were analyzed by 16S rRNA amplicon sequencing, shotgun metagenomics, as well as targeted and untargeted metabolomics. Our findings revealed season-specific compositional and functional profiles of the ibex gut microbiome that may allow the host to adapt to seasonal changes in available forage, by fine-tuning the holobiont catabolic layout to fully exploit the available food. Besides confirming the importance of the host-associated microbiome in providing the phenotypic plasticity needed to buffer dietary changes, we obtained species-level genome bins and identified minimal gut microbiome community modules of 11-14 interacting strains as a possible microbiome-based solution for the bioconversion of lignocellulose to high value compounds, such as volatile fatty acids.

Highlights

- Season-specific profiles of Alpine ibex (*Capra ibex*) gut microbiome were identified
- Gut microbiome provides Ibex with phenotypic plasticity to deal with seasonality
- Thirty-eight species-level genome bins from the Ibex gut microbiome were identified
- 3 hubs for bioconversion of lignocellulosic biomasses to fatty acids were detected

Graphical Abstract



Introduction

The gut microbiome is utterly recognized as a key element for host physiology, being involved in vital processes such as digestion, immunity, and protection[1,2,3]. Herbivorous mammals harbor complex and dynamic microbial communities in their guts, and ruminants are among the most studied animals for gut microbiome structure, dynamics, and functions[4,5,6]. Indeed, ruminants rely on gut symbionts (mainly bacteria, fungi, and protists) to break down the complex biopolymers of plant cell walls and extract energy from these dietary sources, making them an excellent model for investigating diet-host-microbiome relationships and dependencies[5,7,8,9]. However, while the cattle microbiome has been extensively studied[10,11,12] - also due to their high economic importance and role in the current global change scenario - less is known about the structure and function of microbiomes associated with wild ruminants, weakening our understanding of the full diversity and complexity of the ruminant gut microbiome and its importance in animal biology[7,9,13]. Also, wild ruminants consume a more complex and diverse diet than their domesticated relatives and are also more tolerant to roughage and lignin[14,15]. This makes their gut microbiome a possible, yet untapped, source of a diverse and undiscovered array of enzymes and taxa as a valuable natural reservoir of functionalities for the efficient digestion and valorization of lignocellulosic (LC) substrates[16,17].

Recent research has suggested the importance of wild ruminants' gut microbiomes in cooperating with their hosts to cope with their complex and diverse wild plant-based diet[8,9,15,18]. However, studies were based on 16S rRNA amplicon sequencing, making it difficult to mechanistically understand the functional role of the gut microbiome in the adaptation to seasonal dietary shifts.

To shed light on this and to explore the potential of the ibex gut microbiome for industrial LC bioconversion processes, we collected fresh fecal samples from 86 wild individuals of Alpine ibex (*Capra ibex*). Fecal samples were chosen since they are recognized as a good (and non-invasive) proxy for microbial diversity across the ruminant digestive tract[19], being possibly enriched in functionalities for degrading the most recalcitrant and indigestible portion of the plant food, as it is not completely processed in the foregut. Among large ruminant mammals, the Alpine ibex is the species using areas located at the highest elevations[20] of the Alps, where seasonality is pronounced, with a long winter season and low availability of forage that alternates with a late spring-summer season with high availability of good-quality forage. As a result, spatial behavior and habitat selection clearly vary among the seasons, as well as diet[21,22,23]. This makes the Alpine ibex an excellent case study for investigating the responses of the microbiome to the seasonality of trophic resources, including roughage or available trees and bushes, which would require the selection of specialized gut microbiome components for the digestion of more recalcitrant plant foods. Sampling was performed across 3 different seasons, namely spring (June 2020), summer (August 2020), and fall (October 2020), at the Stelvio National Park, Lombardia (Italy). Samples were collected from animals of approximately the same age, sampled non-invasively by actively searching for animals over 2-3 days and waiting for fecal deposition. Samples were then analyzed by multiomics (i.e., 16S rRNA amplicon sequencing, shotgun metagenomics, and metabolomics), in an attempt to map the compositional and functional shifts of the ibex gut microbiome in response to seasonal vegetation changes from spring to fall. Our results showed well-defined seasonal dynamics of the Alpine ibex gut microbiome, with community modules, taxa, functions, and metabolites characterized by clear seasonal patterns.

While providing some glimpses on the importance of the gut microbiome for the ibex biology, we discovered new microorganisms and community modules as potential candidates for biotechnological processes of LC bioconversion and valorization.

Results

Seasonal variation in the compositional structure of the Alpine ibex gut microbiome

A total of 86 Alpine ibex feces, 17 soil, and 8 grass samples were collected from two different sites (i.e., "Passo del Gavia" and "Valle del Braulio") at Stelvio National Park, Lombardia (Italy), in June, August, and October 2020 (see **Figure S1** for sampling coordinates and normalized difference vegetation index of the sampling area for each sampling season). In particular, we collected approximately the same number of ibex samples from each site and season to obtain a comparable subset of animal data for the three selected seasons (see **Table S1** for sampling details). The vegetation map of sampling sites highlighted that the majority of species grouped into the Magnoliophyta division, whereas the remaining species belonged to the Pinophyta division. Based on Raunkiaer's classification, plants occurred mostly as hemicryptophytes (48%), followed by chamaephytes (7%), phanerophytes (5%), geophytes (5%), and therophytes (2%). Snow on the ground (cm) from 1992 to 2020 is provided in **Figure S2**, and 2019-2020 showed one of the highest records in the period of observation.

For microbiome analysis, we performed bacterial metabarcoding (i.e., 16S rRNA gene sequencing of the V3-V4 hypervariable regions), shotgun metagenomics, and metabolomics. First, a Bray-Curtis-based principal coordinates analysis (PCoA) of the 16S rRNA gene-based taxonomic composition of the entire set of 111 samples (ibex, soil, and grass) was performed. Data revealed three distinct clusters matching the three ecosystem types (permutation test with pseudo-F ratio, $p = 0.001$), with the soil microbiome showing the highest alpha diversity, followed by the Alpine ibex gut microbiome and then by the grass microbiome (Kruskal-Wallis test, $p \leq 0.001$) (**Figure S3**). As no separation was observed between the ibex gut microbiome profiles of the two sampling areas (permutation test with pseudo-F ratio, $p = 0.66$), the two Alpine ibex subpopulations were considered as one for subsequent analyses. Indeed, the Valle del Braulio and Passo del Gavia ibex colonies share the same genetic origin, and, since the exchange of individuals between the colonies cannot be excluded, although limited in number, the two colonies can be regarded as a single meta-population[24,25].

At the phylum level, the Alpine ibex gut microbiome was characterized by two dominant phyla, namely Firmicutes (mean relative abundance ± SD, 62.8% ± 13.2%) and Bacteroidetes (19.6% ± 8.5%). Actinobacteria (6.2% ± 8.0%), Saccharibacteria (3.8% ± 4.2%), Verrucomicrobia (3.0% ± 2.9%), and Proteobacteria (2.6% ± 5.9%) were less abundant phyla. At the family level, the dominant taxa were *Ruminococcaceae* (35.8% ± 14.1%), *Lachnospiraceae* (11.0% ± 4.3%), and *Christensenellaceae* (7.6% ± 3.4%). Subdominant families were *Rikenellaceae* (6.7% ± 3.8%), *Bacteroidaceae* (4.2% ± 3.0%), *Prevotellaceae* (4.0% ± 3.5%), and *Coriobacteriaceae* (3.2% ± 3.8%). For the phylum- and family-level bacterial composition of the Alpine ibex gut microbiome across seasons, see **Figure S4**. Notably, the ibex gut microbiomes clearly segregated by season in the Bray-Curtis-based PCoA (permutation test

with pseudo-F ratio, $p = 0.001$) (**Figure 1A**). Conversely, no significant differences were found when comparing alpha diversity across seasons (Kruskal-Wallis test, $p > 0.05$) (**Figure 1B**). According to Linear discriminant analysis Effect Size (LEfSe) analysis (**Figure 1C**), genera associated with spring were *Christensenellaceae R7* group, *Ruminococcaceae NK4A214* group, *FamilyXIII AD3011* group, *Lachnospiraceae NK3A20* group, *Ruminococcus 2*, *Eubacterium hallii* group, *Streptococcus*, *Eubacterium nodatum* group, *Acetitomaculum*, and *Chthoniobacter*. Summer samples were enriched in *Solibacillus* and *Prevotella 7*, while autumn samples in *Arthrobacter*, *Ruminococcaceae UCG 010*, *Bacillus*, *Paenibacillus*, *Pseudomonas*, *Paludibacter*, *Ruminiclostridium 1*, *Odoribacter*, and *Staphylococcus*. The only component of the core Alpine ibex gut microbiome identified in our dataset, defined as the only genus present with a relative abundance ≥ 3% in at least 75% of the samples in each season[26], was *Christensenellaceae R-7* group.

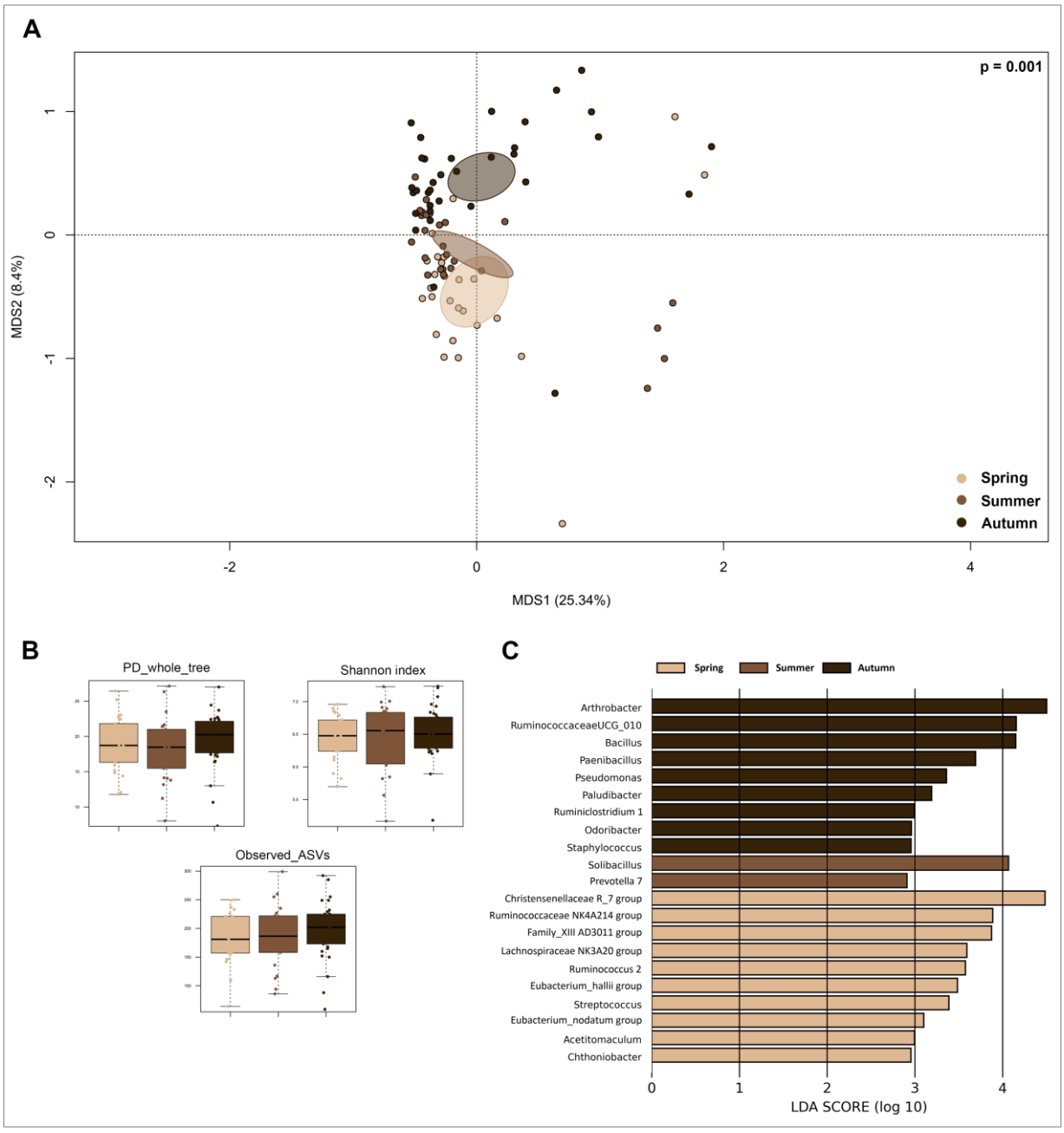


Fig. 1 – Seasonal variation of the compositional profile of the Alpine ibex gut microbiome. **(A)** Principal Coordinates Analysis (PCoA) based on Bray-Curtis distances between the Alpine ibex gut microbiome profiles across seasons, i.e., spring (June, light brown), summer (August, brown) and autumn (October, dark brown) (permutation test with pseudo-F ratio, $p = 0.001$). The first and second principal components (MDS1 and MDS2) are plotted and the percentage of variance in the dataset explained by each axis is shown. Ellipses include the 95% confidence area based on the standard error of the weighted average of samples coordinates. **(B)** Boxplots showing the alpha-diversity distributions of the Alpine ibex gut microbiome in spring, summer and autumn, based on the Faith's Phylogenetic Diversity (PD whole tree), the Shannon index and the number of observed ASVs. No significant differences were found for any of the metrics (Kruskal-Wallis test, $p > 0.05$). **(C)** Linear discriminant analysis (LDA) scores of discriminating Alpine ibex gut microbiome genera between spring, summer and autumn (the logarithmic threshold for discriminating features was set to 2.0 with $p < 0.05$). Plots were obtained by LDA effect size (LEfSe) analysis.

Seasonal changes in the Alpine ibex gut microbiome functional repertoire and metabolome

On a selected and representative subset of 12 Alpine ibex gut microbiome samples, 4 for each season, shotgun metagenomics was carried out, obtaining an average of 5.2 ± 1.0 million high-quality reads per sample. According to the PCoA based on Bray-Curtis distances between the abundance patterns of knockout (KO) genes, there was a trend toward a sample segregation by season (permutation test with pseudo-F ratio, $p = 0.088$) (**Figure 2A**). Conversely, no changes in functional diversity were observed, with alpha-diversity scores remaining constant across seasons (**Figure 2B**).

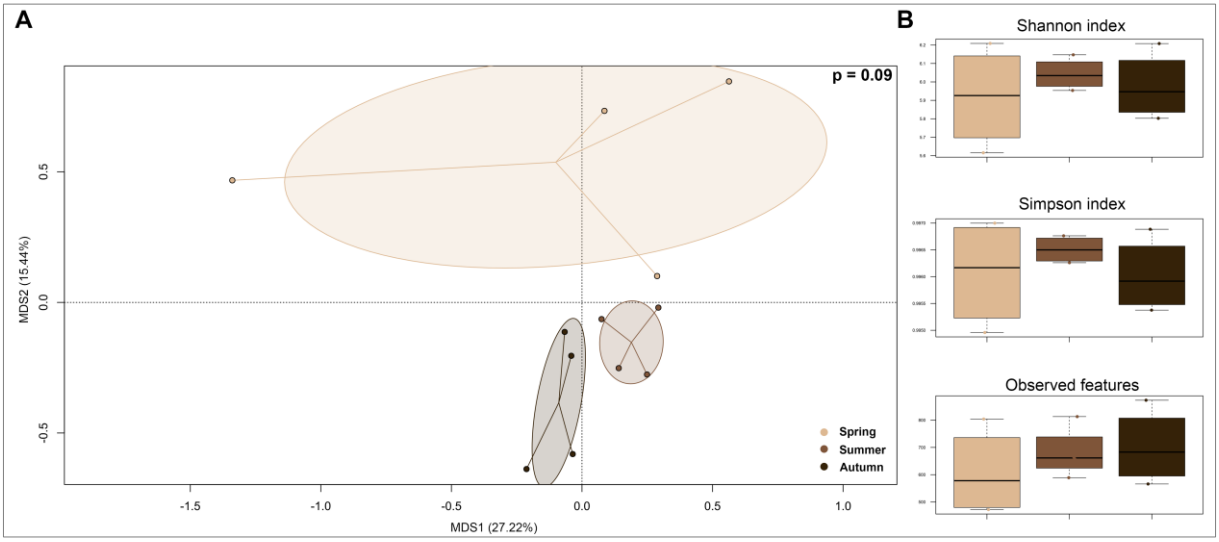


Fig. 2 – Seasonal variation of the functional profile of the Alpine ibex gut microbiome. **(A)** Principal Coordinates Analysis (PCoA) based on Bray-Curtis distances between Alpine ibex gut microbiome functional profiles of KO genes across seasons, i.e., spring (June, light brown), summer (August, brown) and autumn (October, dark brown) (permutation test with pseudo-F ratio, $p = 0.09$). The first and second principal components (MDS1 and MDS2) are plotted and the percentage of variance in the dataset explained by each axis is shown. Ellipses include the 95% confidence area based on the standard error of the weighted average of samples coordinates. **(B)** Boxplots showing the alpha-diversity distributions of the Alpine ibex gut microbiome functional profiles in spring, summer and autumn, based on the Shannon index, the Simpson index and the number of observed features. No significant differences were found for any of the metrics (Kruskal-Wallis test, $p > 0.05$).

Next, we focused on the Alpine ibex gut glycobiome, namely the set of Carbohydrate-Active enZYmes (CAZymes) encoded by the gut microbiome. Specifically, we identified 151 CAZymes in the ibex gut microbiome, representing all five classes listed in the CAZy database (http://www.cazy.org/Glycoside-Hydrolases.html), i.e., glycoside hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), and auxiliary activities (AAs). Notably, the ibex gut microbiome showed seasonal variations in the pattern of CAZymes involved in the catabolism of plant cell wall polysaccharides, including cellulases, xylanases, mannases, pectinases, β-glucosidases, and AAs enzymes (**Figure 3**). In particular, the autumn and summer samples were characterized by a higher load of cellulases, CEs, and β-glucosidases, while being depleted in auxiliary functions for lignin degradation, compared to the spring samples.

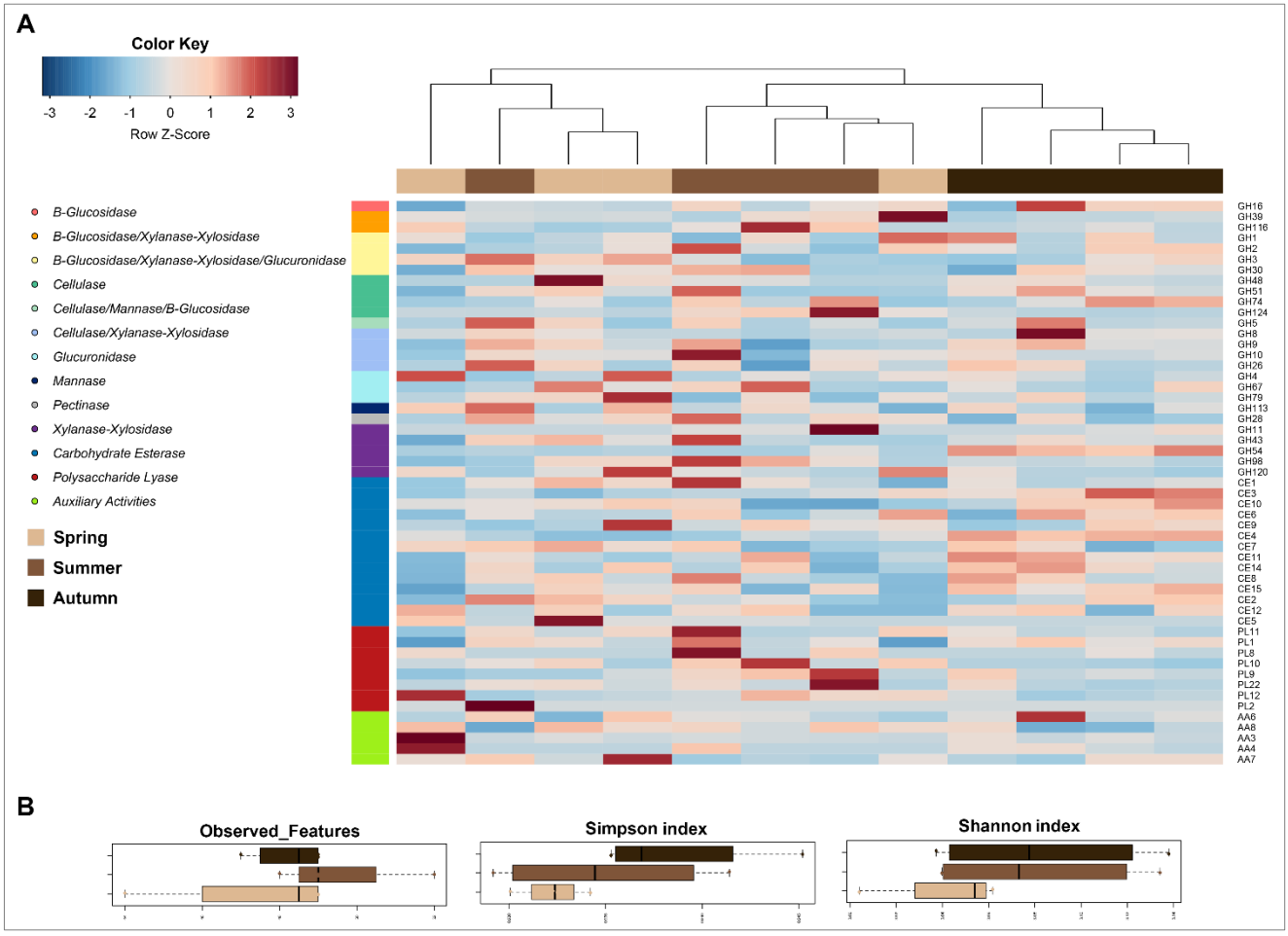


Fig. 3 – Seasonal variation of the glycobiome layout of the Alpine ibex gut microbiome. **(A)** Hierarchical Ward-linkage clustering based on the Spearman correlation coefficients of the RPKM abundances of the main plant cell wall-hydrolyzing CAZymes families of the Alpine ibex gut microbiome across seasons, i.e., spring (June, light brown), summer (August, brown) and autumn (October, dark brown). The relative Z-score is reported. Rows represent all CAZymes families grouped by the corresponding functional class. **(B)** Boxplots showing the alpha-diversity distributions of CAZymes families of the Alpine ibex gut microbiome in spring, summer and autumn, based on the number of observed features, the Simpson index and the Shannon index. No significant differences were found for any of the metrics (Kruskal-Wallis test, $p > 0.05$).

Finally, the entire set of Alpine ibex fecal samples (n = 86) was subjected to both targeted and untargeted metabolomics. According to our findings, the Alpine ibex gut metabolome, as assessed by untargeted metabolomics, showed a strong seasonal segregation (p = 0.04) (**Figure 4A**), which associated with the previously reported gut microbiome seasonal changes, as assessed by the procrustean randomized test ("protest," p value = 0.001 and correlation in a symmetrical rotation = 0.50). Similarly, we found relevant seasonal changes in the abundance profiles of the main short-chain fatty acids (SCFAs)/branched-chain fatty acids (BCFAs) in the ibex samples (**Figure 4B**). Specifically, a significant increase in acetic and isovaleric acid was observed in autumn compared to spring (Wilcoxon rank-sum test, $p \leq 0.05$), while a similar, but much smaller, trend was observed for valeric acid (p = 0.06), whereas propionic and butyric acid were significantly more abundant in spring and summer compared to autumn (Wilcoxon rank-sum test, $p < 0.01$).

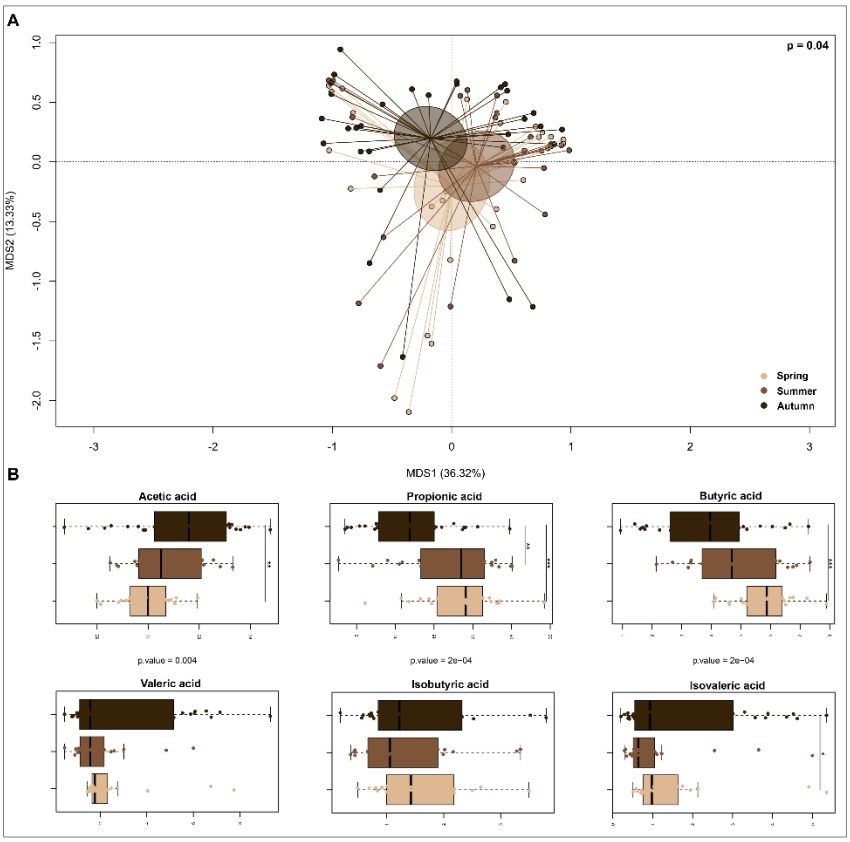


Fig. 4 – Seasonal variation of the Alpine ibex fecal metabolome. **(A)** Principal Coordinates Analysis (PCoA) based on Bray-Curtis distances between Alpine ibex fecal metabolomic profiles across seasons, i.e., spring (June, light brown), summer (August, brown) and autumn (October, dark brown) (permutation test with pseudo-F ratio, p = 0.04). The first and second principal components (MDS1 and MDS2) are plotted and the percentage of variance in the dataset explained by each axis is shown. Ellipses include the 95% confidence area based on the standard error of the weighted average of samples coordinates. **(B)** Boxplots showing the relative abundance distributions of short-chain fatty acids and branched-chain fatty acids in the Alpine ibex feces in spring, summer and autumn. Kruskal-Wallis test and Wilcoxon rank-sum test controlled for multiple testing using FDR; *p-value ≤ 0.05; **p-value ≤ 0.01; ***p-value ≤ 0.001.

Identification and characterization of SGB community modules from the Alpine ibex gut microbiome for the degradation of plant-derived biopolymers

Forty-nine high-quality metagenome-assembled genomes (MAGs) were obtained from the Alpine ibex gut microbiome and were successfully dereplicated into 38 species-level genome bins (SGBs). Only one of these SGBs (assigned to the *Acutalibacteraceae* family) showed a genetic distance < 10% compared to already available genomes from ruminant gut microbiomes, suggesting that the others could be unreported genomes. Notably, these 37 SGBs showed different distributions by season, as visualized in the PCoAs of the corresponding compositional profiles across the three different seasons, confirming the seasonal dynamics of the Alpine ibex gut microbiome also at the SGB level (**Figure S5**). Considering the SGB communities of the Alpine ibex gut microbiome characterizing each season, we next obtained the corresponding genome-scale metabolic models (GSMMs) for the degradation of the main plant components (namely cellulose, hemicellulose, lignin, and pectin) (**Figure 5**). Interestingly, for each season, we obtained a specific module of 11-14 SGBs synergistically interacting for the degradation of plant cell wall biopolymers (**Table S2**). Only three SGBs, belonging to *Akkermansia*, *Bacteroidaceae bacterium UBA4372*, and *Alistipes*, remained constant across all time points. Furthermore, based on the generated models, the primary metabolic endpoints from each SGB plant-degrading module were generally constant, with butanol and oxalosuccinate produced in all seasons, and isobutyric acid produced in spring and autumn. In contrast, butyrate, isobutanol, succinate, and isovaleryl-coenzyme A (CoA) were season-specific metabolites (**Figure 5**).

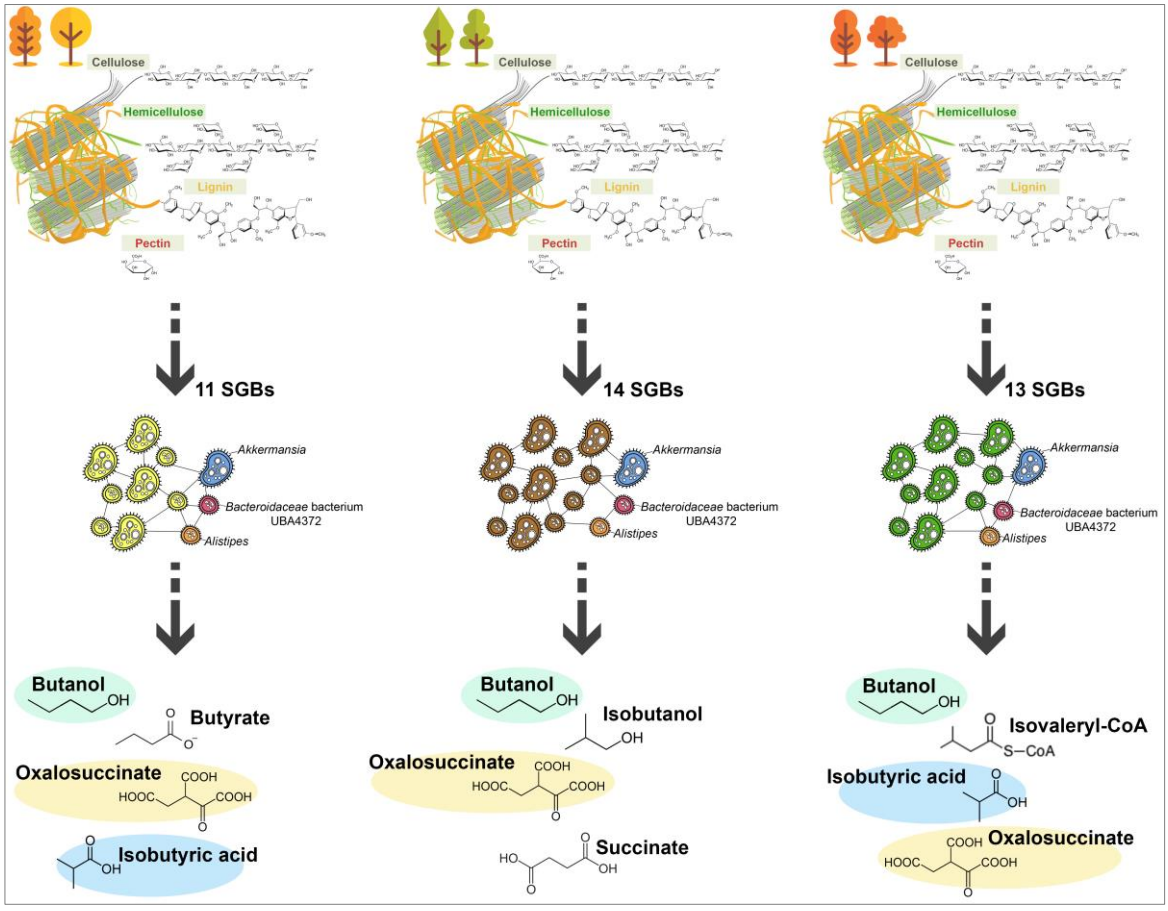


Fig. 5 – Seasonal genome-scale metabolic models of the Alpine ibex gut microbiome. Schematic representation of the main plant-derived biopolymers (top), the predicted SGB communities digesting such polymers (middle) and the resulting metabolic endpoints (bottom) in spring, summer and autumn. Despite being largely characterized by different bacterial taxa, the ibex SGB gut communities appear to be able to ferment plant fibers and produce both common and unique endpoints depending on the season.

Discussion

According to our findings, the main phyla of the Alpine ibex gut microbiome were Firmicutes and Bacteroidetes, followed by Actinobacteria, Verrucomicrobia, and Saccharibacteria, while *Ruminococcaceae*, *Lachnospiraceae*, and *Christensenellaceae* were the dominant families. To the best of our knowledge, this is the first report on the bacterial fraction of the Alpine ibex gut microbiome, which to date has only been investigated considering fungal-methanogens associations[27] or targeting specific microbial pathogens[28]. Our results are consistent with those of the few available studies conducted on mountain ungulates, such as the long-tailed goral *Naemorhedus caudatus*[29], mountain goats *Oreamnos americanus*[7], takins *Budorcas taxicolor*[30], and chamois *Rupicapra* spp.,[14,15] as well as on other wild and captive ruminants[31,32,33,34,35,36,37], suggesting the presence of a phylogenetically widespread ruminant core gut microbiome at the family level. The fine multiomic assessment of the seasonal dynamics of the Alpine ibex gut microbiome allowed us to identify sharp seasonal patterns in terms of compositional, functional, and metabolic layouts. Notably, *Christensenellaceae R-7* group was detected as the only core microbiome genus (relative abundance ≥ 3% in at least 75% of the samples in each season), confirming the impressive seasonal changes in the Alpine ibex gut microbiome at low taxonomic ranks. On the other hand, as previously reported for other alpine ruminants[35], no significant seasonal shifts were observed in alpha diversity, suggesting that a high level of microbial diversity is maintained throughout the year, possibly to ensure high redundancy in microbiome functionalities for digestion of available plants.

When we focused on the microbiome layout of CAZymes for the degradation of plant biopolymers to fermentable monosaccharides, we observed two distinct clusters according to the sampling season. In particular, one cluster was composed of summer and autumn samples enriched in CAZymes associated with a vast array of functionalities, such as cellulases, xylanases, PLs, CEs, and β-glucosidases, while the other cluster included spring samples, showing an overall lower diversity of CAZymes families, but a higher load of auxiliary enzymes dedicated to lignin oxidation and degradation. These results provide some insight into the mechanistic understanding of the functional importance of the Alpine ibex gut microbiome in the animal's adaptation to dietary shifts. More specifically, the studied year (2019-2020 period) was characterized by peculiar weather, which saw a prolonged persistence of snow on the ground until the beginning of June, with a direct effect on the available forage. Indeed, looking at the majority of the available plant species for the ibex (e.g., hemicryptophytes and geophytes), the snow sill on the ground in early June 2020 would have prevented their growth, and the ibex would have fed only on available trees and bushes, belonging to chamaephytes, phanerophytes, and therophytes, which are characterized by high lignin contents. In this condition, the Alpine ibex gut microbiome would respond adaptively, enriching for lignin-modifying functions, thus providing

the host with the necessary degree of phenotypic plasticity to exploit this available plant food. On the contrary, the higher availability of grasses and herbs during the 2020 summer-autumn period would result in a diet enriched in cellulose and hemicellulose, with a concomitant decrease in total ingested lignin. Under these conditions, the Alpine ibex gut microbiome would respond by increasing the diversity of CAZymes for cellulose and hemicellulose degradation, allowing the full exploitation of the available dietary sources. This vision is also supported at the compositional level, as, in the summer period, the ibex gut microbiome was enriched in taxa belonging to *Prevotella 7*, whose members have been suggested to be among the most important protein, hemicellulose, and pectin degraders in ruminants[38,39]. Furthermore, autumn-enriched taxa included *Ruminococcaceae UCG 010*, *Bacillus*, *Paenibacillus*, *Pseudomonas*, and *Ruminiclostridium*, which have enhanced cellulose and hemicellulose digestion capabilities, via either secreted free enzymes or extracellular multi-enzyme structures called cellulosomes[33,40,41,42,43,44]. Interestingly, the seasonal changes in the Alpine ibex gut microbiome taxonomy would explain the corresponding shifts in the overall gut metabolome layout, as well as in the measured profiles of SCFAs. In particular, the higher levels of acetic acid in autumn may be due to the prevalence, in this season, of some well-known acetate producers such as *Ruminococcaceae UCG-010*[33], *Ruminiclostridium*[40], *Bacillus*[45], and *Paenibacillus*[46]. Conversely, taxa such as *Eubacterium*, *Christensenellaceae*, *Prevotella*, and *Ruminococcus* may be correlated with increased proportions of propionic and butyric acid in spring and summer[37,47,48,49]. As the main endpoints of microbiome metabolism of plant biopolymers in the gut, SCFAs represent key molecules that support nutrition and regulate different aspects of animal physiology, including immune and metabolic homeostasis and protection against pathogenic microorganisms[31,38,50]. Although SCFAs are produced throughout the year, the Alpine ibex gut microbiome response to seasonal changes in available forage would also result in significant variation in their production profiles, raising concerns about the possible physiological importance of these changes in the holobiont metabolome. Finally, in our study, SGBs and the related metabolic models were created, allowing the identification of season-specific Alpine ibex gut microbiome community modules for the degradation of plant biopolymers (i.e., cellulose, hemicellulose, lignin, and pectin) to alcohols and organic acids, including volatile fatty acids such as butyrate, isobutyrate, and isovaleryl-CoA. These minimal modules of 11-14 interacting strains may represent new candidate microbial consortia to be exploited in circular processes for the valorization of LC biomasses, enabling their bioconversion into value-added platform chemicals[51]. In addition, given the importance of transitioning to more sustainable and secure food systems, these community modules may foster innovative applications as next-generation probiotics in cattle, allowing for improved roughage tolerance in livestock for the transition to more sustainable farming strategies, with less reliance on green grasses, which require consistent amounts of water and are likely to be negatively affected by climate change[52]. Overall, our findings support the importance of the Alpine ibex gut microbiome as a strategic evolutionary partner in the holobiont framework, providing the animal host with the necessary phenotypic plasticity to buffer seasonal changes in the available forage. This microbiome-host cooperation would be crucial for fine-tuning holobiont catabolism to fully exploit the available plant food. Besides confirming the relevance of the host-associated microbiome in the adaptation to dietary changes[53,54], we provided some insight into the possible exploitation of the Alpine ibex gut microbiome for the development of innovative biotechnological solutions, in terms of circular LC

bioconversion and valorization processes, and also as next-generation probiotics for the transition to more sustainable and secure food systems.

Limitations of the study

The main limit of our study is the lack of individual records of animal behavior and diet during the period of the study, which can be obtained by using GPS collars for animal tracking. Further, a second limitation is the lack of a second year of sampling, allowing to control for a possible annual variation. Finally, putative lignocellulose-degrading strains and hubs have been only identified as metagenomic assembled genomes, without having microbial isolates.

Method details

Study site and sampling procedure

A total of 86 fecal samples were collected at Stelvio National Park (Lombardia, Italy) from an equal number of Alpine ibex specimens, which were followed and observed until defecation. When possible, surface soil and grass samples were also collected in the proximity, for a total of 111 samples (i.e., 86 ibex feces, 17 soil and 8 grass samples). Samples were collected at two different nearby sites, namely "Passo del Gavia" (46°20'04.1"N/10°29'15.4"E) and "Valle del Braulio" (46°31'03.4"N/10°24'42.8"E), across three different seasons, namely spring (16th, 17th and 18th June 2020), summer (3rd, 4th and 5th August 2020) and fall (1st and 2nd October 2020) (**Figure S1**). A schematic summary of the sample distribution across the two sites and the three timepoints is provided in **Table S1**. All samples were collected using sterile gloves, placed in sterile plastic tubes, and kept frozen at -20°C until microbial DNA extraction. The mean values of snow on the ground (cm) for the sampling year were retrieved from the meteorological station Valdisotto Oga S. Colombano (SO, ARPA Lombardia), which is located at an altitude of 2,300 m and collects detailed measurements almost every 30 min. Data from July to June of the subsequent year (from 1992 to 2022) are reported in **Figure S2**. Coordinates of sampling sites were uploaded in the Italian Geoportale Nazionale (http://www.pcn.minambiente.it/viewer/index.php?services=progetto_natura), managed by the Ministry of Environment and providing different kinds of spatial data. In particular, a map of plant alliances of the sampling sites was retrieved form the portal. Plant species characterizing the identified phytosociological synthaxa were then inferred according to Prodromo della Vegetazione Italiana (https://www.prodromo-vegetazione-italia.org). Each species was assigned to a Raunkiaer's life form (chamaephytes, geophytes, hemicryptophytes, phanerophytes and therophytes) using Pignatti et al[70].

Microbial DNA extraction, 16S rRNA amplification and sequencing

Total microbial DNA was extracted from approximately 0.25 g of each of the 111 samples, i.e., ibex feces, soil, and grass. DNA extraction from fecal samples was performed using the DNeasy Blood & Tissue Kit (QIAGEN, Hilden, DEU) with a modified protocol[71]. In brief, fecal material was added with four 3-mm glass beads and 0.5 g of 0.1-mm zirconia beads (BioSpec Products, Bartlesville, OK, USA), and the homogenization step was performed three times

using a FastPrep instrument (MP Biomedicals, Irvine, CA, USA) at 5.5 m/s for 1 min. Samples were then heated at 95°C for 15 mins. DNA from soil and grass was extracted using the DNeasy PowerSoil Kit (QIAGEN) following the manufacturer's instructions with a minor modification: a FastPrep instrument (MP Biomedicals) was used for the homogenization step as described above. DNA was quantified using a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DEU). PCR was performed in a final volume of 50 µL containing genomic DNA (25 ng), 2X KAPA HiFi HotStart ReadyMix (Roche, Basel, CHE) and 200 nmol/L of 341F and 785R primers carrying Illumina overhang adapter sequences for amplification of the V3-V4 hypervariable regions of the 16S rRNA gene. The PCR thermal cycle consisted of an initial denaturation (95°C for 3 mins), followed by 25 cycles of denaturation (95°C for 30 s), primer annealing (55°C for 30 s) and DNA extension (72°C for 30 s); the entire reaction was completed with a final extension step (72°C for 5 mins)[72]. PCR amplicons were then cleaned up using Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA). Indexed libraries were prepared by limited-cycle PCR using Nextera technology and purified as above. Finally, the libraries were quantified using a Qubit 3.0 fluorimeter (Invitrogen, Waltham, Massachusetts, USA), normalized to a concentration of 4 nM and pooled in a single Eppendorf tube. The pool was denatured with 0.2 N NaOH and diluted to a final concentration of 4.5 pM with a 20% PhiX control. Sequencing was performed on an Illumina MiSeq platform using a 2 x 250 bp paired-end protocol, according to the manufacturer's instructions (Illumina, San Diego, CA, USA).

Metabarcoding bioinformatics and biostatistics

Raw sequences were analyzed using a pipeline combining PANDAseq[73] and QIIME 2[74]. High-quality reads (min/max length = 350/550 bp) were retained using the "fastq filter" function of the Usearch11 algorithm[75] and then binned into amplicon sequence variants (ASVs) using DADA2[76]. The VSEARCH algorithm[77] and the SILVA database (December 2017 release)[78] were used for taxonomic classification. All unassigned and eukaryotic sequences were discarded. Overall, an average sequencing depth of 10,725 ± 2,871 (mean ± SD) high-quality reads per sample was obtained, resulting in a total of 20,592 ASVs. Alpha-diversity was assessed using three different metrics, namely Faith's Phylogenetic Diversity (PD whole tree), the number of observed ASVs and the Shannon index. Beta-diversity was assessed using Bray-Curtis distances.

Statistical analyses were performed using R software (https://www.r-project.org/), v. 4.2.0, implemented with the packages "Made4"[79], "vegan"[69] (https://cran.r-project.org/web/packages/vegan/index.html), "pairwiseAdonis"[80] and "gplots"[63] (https://cran.r-project.org/web/packages/gplots/index.html). Data separation in the PCoAs was assessed using a permutation test with pseudo-F ratio (functions "adonis" in the vegan package and function "pairwiseAdonis" in the homonymous package). A procrustean randomized test (function "protest" in the vegan package) was performed to highlight significant relationship between microbiome and metabolomic distance matrices. The Kruskal-Wallis test among groups was used to assess significant differences in alpha-diversity (calculated on taxonomical annotation). P-values were corrected for multiple testing, when necessary, using the Benjamini-Hochberg method, with a false discovery rate (FDR) ≤ 0.05 considered statistically significant. Linear discriminant analysis (LDA) effect size[81] (LEfSe) was used to identify discriminant genera across the three timepoints ($p \leq 0.05$).

The online Galaxy Version interface (https://huttenhower.sph.harvard.edu/galaxy/, last accessed September 2023) was used.

Shotgun metagenomics sequencing

A subset of representative 12 Alpine ibex fecal samples (six per site, including two per season) was selected for shotgun metagenomic sequencing. The QIAseq FX DNA library kit (QIAGEN) was used for DNA library preparation according to the manufacturer's instructions. Briefly, 450-bp size, end-repaired and A-tailed fragments were generated by fragmenting 100 ng of each DNA sample using FX enzyme mix with the following thermal cycle: 4°C for 1 min, 32°C for 8 mins and 65°C for 30 mins. DNA samples were then incubated at 20°C for 15 mins to perform adapter ligation in the presence of DNA ligase and Illumina adapter barcodes. Agencourt AMPure XP magnetic beads (Beckman Coulter) were used for purification, followed by library amplification with a 10-cycle PCR and a further purification step. Samples were then pooled at an equimolar concentration (4 nM) to obtain the final library. Sequencing was performed on an Illumina NextSeq platform using a 2 × 150 bp paired-end protocol, following the manufacturer's instructions (Illumina).

Metagenomics bioinformatics and biostatistics

Raw reads were filtered for eukaryotic host DNA using bmtagger software and *Capra ibex* (NCBI GenBank accession: GCA_006410555.1) as a reference. After this filtering step, reads were processed with trimBWAstyle (https://github.com/genome/genome/blob/master/lib/perl/Genome/Site/TGI/Hmp/HmpSraProcess/trimBWAstyle.usingBam.pl) for quality trimming (quality score above 20) and length drop with default parameters. Duplicates were estimated and removed using the Picard tool EstimatedLibraryComplexity (v. 1.71). A total of 61 million high-quality microbial paired-end reads were retained, with an average of 5.2 ± 1.0 (mean ± SD) million reads per sample. The resulting reads were used to obtain a general functional annotation for each sample, using HUMAnN v. 3.0.1[82]. The output tables were then normalized using humann_renorm_table with the following parameter "—units cpm". The resulting tables were merged and then processed by removing the UNMAPPED ID and converting the UniRef90 classification into the KEGG Orthology (KO) classification. This final table was used to compute alpha-diversity indices (Shannon, Simpson, and observed features) and beta-diversity based on Bray-Curtis distances. Data separation in the Bray-Curtis-based PCoA was assessed in R using a permutation test with pseudo-F ratio (function "adonis" in the vegan package and function "pairwiseAdonis" in the homonymous package). The Kruskal-Wallis test among groups was used to assess significant differences in alpha-diversity, with P-values corrected for multiple testing as previously described. In parallel, high-quality reads were assembled using metaspades.py (v. 3.15.3) with default parameters. Each assembly was annotated using prokka[67] (v. 1.14.6) with default parameters and "—addgenes" to retrieve all classes of Carbohydrate-Active enZYmes (CAZymes), according to the latest version of the online CAZy database, namely glycoside hydrolases (GHs, EC 3.2.1.-), glycosyl transferases (GTs, EC 2.4.x.y), polysaccharide lyases (PLs, EC 4.2.2.-), carbohydrate esterases (CEs) and auxiliary activities (AAs) enzymes. Using prokka output files, open reading frames (ORFs) for each CAZyme were retrieved and used to build a reference database, dereplicated at 90% similarity and used to assess the abundance of each CAZyme in our samples. Alignment was performed using Bowtie2 v. 2.3.4.3[59] with the parameter "--end-to-end --very-sensitive"; the number

of aligned reads for each sample was then retrieved using Samtools v. 1.16[68]. Reads per kilobase of gene per million reads mapped (RPKMs) in each sample and for each gene were calculated by summing the number of reads of all mapped ORFs and processed as follows: (total reads mapped to a gene/(total reads x mean gene length)) x 10^9.

The abundance table, in terms of RPKMs, of each CAZyme family identified in our dataset was used to plot a heatmap of the CAZymes families involved in the catabolism of plant polysaccharides, assigned to the corresponding functional classes, using the heatmap.2 function in R. The Spearman distance and the ward.D2 method were used to cluster the different samples according to the obtained CAZymes abundances. The heatmap represents the Z-score of the identified CAZymes families, with clustering performed for samples. The RPKMs abundance table of the CAZymes families assigned to a specific functional class was also used to calculate alpha-diversity, using the number of observed features, the Simpson index, and the Shannon index.

Metagenome-Assembled Genomes (MAGs) reconstruction

Assemblies from each sample were used to construct Metagenome-Assembled Genomes (MAGs) using the metawrap binning module (metawrap version 1.3.2[66]). Only MAGs with completeness > 50% and contamination < 5%, as assessed through the checkm lineage_wf workflow[61], were retained. All retrieved high-quality MAGs were then dereplicated into species-level genome bins (SGBs) using the dRep dereplicate command (dRep v. 3.2.2[62]) and the following parameters: "--ignoreGenomeQuality -pa 0.9 -sa 0.95 -nc 0.30 -cm larger -centW 0". The taxonomic classification of SGBs was performed using the gtdbtk classify_wf workflow with default parameters[64], while the abundance of each SGB in each sample was obtained using the metawrap quant_bins module (metawrap v. 1.3.2). The SGBs abundance table was used to construct a presence/absence table of each SGB across samples. A phylogenetic tree including all SGBs was then built by using phylophlan[83] with the parameters "--diversity low --fast --min_num_markers 79", and used to measure UniFrac distances between samples, which were plotted in a Principal Coordinates Analysis (PCoA) graph. Finally, the SGBs were compared, using MinHash sketches implemented in the mash tool (v. 2.3), with 8,217 genomes from three of the largest ruminant gut metagenomic datasets[10,57,58], with the Genomic Encyclopedia of Bacteria and Archaea (GEBA) collection[56] and with 4,930 SGBs previously identified in a study describing the gut microbiome of different human individuals across age, geography and lifestyle[55].

Genome-scale metabolic models for the degradation of plant food substrates

Microbiome-scale metabolic models for the identification of key SGBs involved in the degradation of plant food substrates, such as cellulose, hemicellulose, pectin, and lignin, were obtained using CarveMe[60] and Metage2Metabo[65]. Specifically, CarveMe was applied to each SGB, grouped by timepoint, using the default options, to build the specific genome-scale metabolic model (GSMM) for each SGB. Metage2Metabo was then used with the parameter "metacom" to build a single metabolic network combining all the GSMMs by timepoint and retrieving the list of the minimal communities of SGBs essential for the degradation of plant components. The pipeline was repeated using as input the set of GSMMs divided by timepoint and considering the 4 main plant biopolymers.

Metabolomics

All fecal samples underwent two kinds of analytical characterization: SCFAs (Short Chain Fatty Acids) and BCFAs (Branched Chain Fatty Acids) quantitation through head space-solid phase microextraction (HS-SPME-GC-MS) and untargeted metabolomic analysis with liquid chromatography-high resolution mass spectrometry (LC-HRMS) analysis.

Reagents, materials, and solutions

All standards (purity > 99%) for acetic, propionic, butyric, isobutyric, valeric, isovaleric and d8-butyric acids (d8-BA) were provided by Sigma-Aldrich (Milan, ITA). Perchloric acid ($HClO_4$) 70 was also provided by Sigma-Aldrich. MILLEX GP syringe filter, 0.22 μm in pore size, with Polyethesulfone (PES) membrane were provided by Millipore corp. (Bedford, MA). UHPLC-MS grade acetonitrile, UHPLC-MS grade methanol and water were provided by VWR Chemicals (Radnor, PA, USA). LC-MS grade formic acid was purchased from Carlo Erba Reagents S.r.l. (Milan, ITA). The manual holder and the commercially available SPME fibers 75 μm Carboxen™/polydimethylsiloxane (CAR/PDMS) were purchased from Supelco (Bellefonte, PA, USA). Prior to first use, the SPME fiber was conditioned for 60 minutes at 300°C as per manufacturer's instructions. Individual acid stock solutions were prepared at a concentration of 1,000 ppm (μg/mL) by diluting 20 μL of acid with milliQ water in a 20-mL volumetric flask. Individual standard solutions were prepared by diluting the stock solution to final concentrations of 5, 10, 25, 50 and 100 ppm. A stock solution for internal standard (d8-BA) at a concentration of 10 mg/mL was obtained by diluting 95.2 μL of acid with H_2O in a 10-mL volumetric flask. From this stock solution, a working solution with a final concentration of 0.5 mg/mL was obtained by successive dilution with milliQ water. LC-MS analysis was performed on an Eksigent M5 MicroLC system (Sciex, Concord, Ontario, Canada) coupled to a TripleTOF 6600+ mass spectrometer with OptiFlow Turbo V Ion Source (Sciex).

HS-SPME GC-MS analysis for SCFAs and BCFAs

Solid-liquid extraction was performed as a preliminary clean-up. A perchloric acid solution (10% v/v in water) was added to frozen aliquots of fecal samples to a final concentration of 250 mg/mL. The resulting solutions were centrifuged at 15,000 rpm for 10 minutes at 4°C. After centrifugation, the supernatant was collected in a 1.5-mL glass vial and stored at -20°C. For HS-SPME analysis, 50 μL of fecal sample solution were added to 450 μL of H_2O and 10 μL of IS solution in a 4-mL glass vial, which was then capped with a pierceable septum cap. Prior to extraction, the vials containing the samples were heated at 70°C for 10 minutes under continuous stirring at 270 rpm using a poly(tetrafluoroethylene)-coated magnetic stir bar. After thermal conditioning, the septum of the vial was pierced with the needle of the SPME device and the fiber was exposed approximately 10 mm above the solid sample, allowing extraction of the analytes for 30 minutes. The optimized temperatures and times were slightly modified as pointed out by Fiori and colleagues in a previously published article[84]. After extraction, the fiber was retracted into the protective sheath, removed from the headspace glass vial and transferred without delay into the injection port of the gas chromatograph/mass spectrometer. The fiber was thermally desorbed in the injection port at 250°C for 2 minutes and the GC/MS run was started. To thermally clean the SPME fiber, it was left in the injection port for an additional 8 minutes after complete desorption of the analytes. GC-MS analysis was carried out on a TRACE GC 2000 Series (ThermoQuest CE Instruments, Austin, TX, USA) gas chromatograph, interfaced with Trace ITQ MS (ThermoQuest

CE) mass detector with 3D ion trap analyzer, operating in EI mode (70 eV). The capillary GC column was a Phenomenex ZB-WAX (30 m x 0.25 mm ID, 0.15 μm film thickness). Helium (He) was used as carrier gas at a flow rate of 1.0 mL/min. A temperature program was adopted: initial temperature was 40°C (hold time: 5 mins), then temperature ramped by 10 °C/min to 220°C (hold time: 5 mins). The temperatures of the transfer line and ionization source were maintained at 250°C and 200°C, respectively. The GC was operated in splitless mode. Mass spectra were recorded in full scan mode (34-200 amu) to collect total ion current chromatograms. Quantitation was carried out using the extracted ion chromatograms by selecting qualifier and quantifier fragment ions of the studied analytes: 43 and 60 amu for AA, 55 and 73 amu for PA, 55 and 77 amu for iBA, 60 and 87 amu for iVA, 60 and 73 amu for BA and VA, 63 and 77 amu for d8-BA.

LC-HRMSMS untargeted metabolomics

For metabolome analysis, approximately 200 mg of homogenized fecal samples were extracted by the addition of three equivalents (weight/Vv) of methanol, followed by vortex-mixing for 3 seconds and sonication for 10 minutes. The samples were then centrifuged at 14,000 rpm for 10 minutes at 4°C, and the supernatant was collected and filtered through 0.22-μm PES membranes. The obtained extracts were stored at -80°C until further analysis. In order to avoid bias, all experimental samples were randomized before sample preparation and before analytical run. Interpooled Quality Control samples (QCs) were prepared by pooling together equal aliquots (10 μL) from each sample before extraction and underwent the same treatment as experimental samples. Before injection, each methanol extract was diluted in a 1:10 ratio with Milli-Q water.

LC-MS analysis was performed on an Eksigent M5 MicroLC system (Sciex) coupled to a TripleTOF 6600+ mass spectrometer with OptiFlow Turbo V Ion Source (Sciex). Analyses were carried out in both positive and negative ionization, with the column temperature set at 35°C. In brief, 5 μL from each sample were loaded onto a Phenomenex Luna Omega Polar C18 100 × 1.0 mm I.D. 1.6 μm 100 Å. Before the first sample injection, the same QC sample was injected repeatedly, for a total of 10 times, to allow for system equilibration and conditioning. Chromatographic separation occurred in 25 minutes at a constant flow rate of 30 μL/min. The gradient elution program was as follows: 0-2 minutes, 0.2% eluent B; 2-5 minutes, 0.2-15% eluent B; 5-15 minutes, 15-70% eluent B; 15-18 minutes, 70-98% eluent B; 18-20 minutes, 98% eluent B; 20-22 minutes, 98-0.2% eluent B; 22-25 minutes, 0.2% eluent B. Equilibration time between chromatographic runs was 3 minutes. Mobile phase A consisted of 0.1% formic acid and mobile phase B was acetonitrile/0.1% formic acid. IonSpray voltage (ISV) was 5,000 V and Curtain Gas supply pressure (CUR) was 30 PSI; nebulizer and heater gas pressures were set at 30 and 40 PSI, respectively. The ion spray probe temperature was 300°C. Declustering potential was 80 V. Analyses were carried out using a collision energy of 40 eV. Sample analyses were performed in Data Independent Acquisition mode (SWATH-MS: Sequential Window Acquisition of All Theoretical Mass Spectra). The variable SWATH windows used for acquisition were obtained through the SWATH Variable Window Calculator app (Sciex). The software employs the m/z density histogram constructed from the TOF MS analysis to equalize the density of the precursors in each window across the m/z range. The overlap between windows was 1 Da. PepCal Mix (Sciex) was used to ensure steady MS and MSMS calibrations during the whole analysis timeframe.

Data analysis

SWATH raw data files were viewed using PeakView 2.2 (AB sciex). Peak picking (minimum spectral peak width of 10 ppm, minimum peak width of five scans), alignment, filtering (intensity threshold of 10,000 cps, removal of features detected in less than 50% of samples) and annotation were performed using SCIEX OS. Untargeted metabolomic analysis was based on all ion features in the SWATH-MS/MS data after peak finding, alignment and filtering. Metabolites eluted close to the solvent front (< 1 min) were excluded. Fatty acids abundances were represented by boxplots. The Kruskal-Wallis test among groups followed by post-hoc Wilcoxon rank-sum test between pairs of groups were used to assess significant differences in fatty acids abundances, with P-values corrected for multiple testing as previously described. The untargeted metabolomic data were normalized according to the Total Peak Area method, i.e., each peak area was normalized to the sum of the areas of all detected peaks in each sample[85]. The normalized table, based on negative ionization, was then used to calculate the relative abundance of each metabolite in each sample. The resulting relative abundance table was used in R as input to compute the PCoA, based on the Bray-Curtis distances between samples, using the "vegdist" function from the vegan package.

Data and code availability

High-quality reads from the samples sequenced in this study were deposited in the European Nucleotide Archive under the project accession number ENA: PRJEB70425.

Acknowledgments

This research was supported by the PON “Ricerca e Innovazione” 2014–2020 program. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author contributions

L.C. and M.C.: conceptualization. E.N., D.S., N.I., L.C., F.Z., I.M., and J.F.: data curation, formal analysis. L.C., L.P., and M.C.: project administration, resources, supervision. E.N. and D.S.: visualization. E.N., D.S., and M.C.: writing – original draft. G.P., N.I., L.C., L.P., F.Z., E.P., E.E., A.B., S.G., I.M., S.T., J.F., and S.R.: writing – review and editing.

Declaration of interests

The authors declare no competing interests.

Supplementary material

The Supplementary material for this article can be found in the appendix of the thesis (Study III) and online at https://www.cell.com/cms/10.1016/j.isci.2024.110194/attachment/a7ebc2f1-e61d-455e-a3419c6d1fd02091/mmc1.pdf.

References

[1]Hooper, Lora V., Dan R. Littman, and Andrew J. Macpherson. "Interactions between the microbiota and the immune system." *science* 336.6086 (2012): 1268-1273.

[2]McFall-Ngai, Margaret, et al. "Animals in a bacterial world, a new imperative for the life sciences." *Proceedings of the National Academy of Sciences* 110.9 (2013): 3229-3236.

[3]Lindsay, Elle C., Neil B. Metcalfe, and Martin S. Llewellyn. "The potential role of the gut microbiota in shaping host energetics and metabolic rate." *Journal of Animal Ecology* 89.11 (2020): 2415-2426.

[4]Huws, Sharon A., et al. "Addressing global ruminant agricultural challenges through understanding the rumen microbiome: past, present, and future." *Frontiers in microbiology* 9 (2018): 2161.

[5]Mizrahi, Itzhak, R. John Wallace, and Sarah Moraïs. "The rumen microbiome: balancing food security and environmental impacts." *Nature Reviews Microbiology* 19.9 (2021): 553-566.

[6]Alessandri, Giulia, et al. "Creating an atlas to visualize the biodiversity of the mammalian gut microbiota." *Current Opinion in Biotechnology* 73 (2022): 28-33.

[7]Haworth, Sarah E., et al. "Space, time and captivity: quantifying the factors influencing the fecal microbiome of an alpine ungulate." *FEMS microbiology ecology* 95.7 (2019): fiz095.

[8]Eddington, Hyrum S., et al. "Spatiotemporal variation in the fecal microbiota of mule deer is associated with proximate and future measures of host health." *BMC Veterinary Research* 17 (2021): 1-8.

[9]Su, Rina, et al. "Comparative study of the function and structure of the gut microbiota in Siberian musk deer and Forest musk deer." *Applied Microbiology and Biotechnology* 106.19 (2022): 6799-6817.

[10]Seshadri, Rekha, et al. "Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection." *Nature biotechnology* 36.4 (2018): 359-367.

[11]Wallace, R. John, et al. "A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions." *Science advances* 5.7 (2019): eaav8391.

[12]Furman, Ori, et al. "Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome assembly dynamics." *Nature communications* 11.1 (2020): 1904.

[13]Jiang, Feng, et al. "Comparative analysis of gut microbial composition and potential functions in captive forest and alpine musk deer." *Applied Microbiology and Biotechnology* 106.3 (2022): 1325-1339.

[14]Smoglica, Camilla, et al. "Microbial community and antimicrobial resistance in fecal samples from wild and domestic ruminants in Maiella National Park, Italy." *One Health* 15 (2022): 100403.

[15]Dahl, Sarah-Alica, et al. "“Get the best out of what comes in”–adaptation of the microbiota of chamois (Rupicapra rupicapra) to seasonal forage availability in the Bavarian Alps." *Frontiers in Microbiology* 14 (2023): 1238744.

[16]Glendinning, Laura, et al. "Metagenomic analysis of the cow, sheep, reindeer and red deer rumen." *Scientific reports* 11.1 (2021): 1990.

[17]Zhong, Huimin, et al. "Whole-genome sequencing reveals lignin-degrading capacity of a ligninolytic bacterium (Bacillus cereus) from Buffalo (Bubalus bubalis) rumen." *Genes* 13.5 (2022): 842.

[18]Couch, Claire E., et al. "Diet and gut microbiome enterotype are associated at the population level in African buffalo." *Nature communications* 12.1 (2021): 2267.

[19]Peng, Xuefeng, et al. "Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes." *Nature microbiology* 6.4 (2021): 499-511.

[20]Parrini, Francesca, et al. "Spatial behaviour of adult male Alpine ibex Capra ibex ibex in the Gran Paradiso National Park, Italy." *Acta Theriologica* 48 (2003): 411-423.

[21]Grignolio, Stefano, et al. "Habitat selection in adult male of Alpine ibex, Capra ibex ibex." *Folia Zoologica* 52 (2003): 113-120.

[22]Grignolio, Stefano, et al. "Seasonal variations of spatial behaviour in female Alpine ibex (Capra ibex ibex) in relation to climatic conditions and age." *Ethology Ecology & Evolution* 16.3 (2004): 255-264.

[23]Brambilla, Alice, et al. "Alpine Ibex Capra ibex Linnaeus, 1758." *Handbook of the Mammals of Europe*. Cham: Springer International Publishing, 2022. 1-27.

[24]Carro, Monica, and Luca Pedrotti, eds. *Atlante Parco nazionale dello Stelvio*. Parco nazionale dello Stelvio, 2010.

[25]Tosi, Guido. *Lo stambecco in Lombardia e sull'arco alpino*. Oikos, 2012.

[26]Neu, Alexander T., Eric E. Allen, and Kaustuv Roy. "Defining and quantifying the core microbiome: challenges and prospects." *Proceedings of the National Academy of Sciences* 118.51 (2021): e2104429118.

[27]Leis, Stefanie, et al. "Finding a robust strain for biomethanation: anaerobic fungi (Neocallimastigomycota) from the Alpine ibex (Capra ibex) and their associated methanogens." *Anaerobe* 29 (2014): 34-43.

[28]Parrini, Francesca, James W. Cain III, and Paul R. Krausman. "Capra ibex (artiodactyla: Bovidae)." *Mammalian Species* 830 (2009): 1-12.

[29]Park, Chang Eon, et al. "Geographical relationships between long-tailed goral (Naemorhedus caudatus) populations based on gut microbiome analysis." *Microorganisms* 9.9 (2021): 2002.

[30]Chen, Jun, et al. "Characterization of the gut microbiota in the golden takin (Budorcas taxicolor bedfordi)." *Amb Express* 7 (2017): 1-10.

[31]Sundset, Monica A., et al. "Novel rumen bacterial diversity in two geographically separated sub-species of reindeer." *Microbial ecology* 54 (2007): 424-438.

[32]Cunha, Isabel S., et al. "Bacteria and Archaea community structure in the rumen microbiome of goats (Capra hircus) from the semiarid region of Brazil." *Anaerobe* 17.3 (2011): 118-124.

[33]Oikonomou, Georgios, et al. "Fecal microbial diversity in pre-weaned dairy calves as described by pyrosequencing of metagenomic 16S rDNA. Associations of Faecalibacterium species with health and growth." *PloS one* 8.4 (2013): e63157.

[34]Liu, Chang, et al. "Dynamic alterations in yak rumen bacteria community and metabolome characteristics in response to feed type." *Frontiers in microbiology* 10 (2019): 1116.

[35]Yang, Chuntao, et al. "Shrub coverage alters the rumen bacterial community of yaks (Bos grunniens) grazing in alpine meadows." *Journal of Animal Science and Technology* 62.4 (2020): 504.

[36]Fan, Qingshan, et al. "Rumen microbiota of Tibetan sheep (Ovis aries) adaptation to extremely cold season on the Qinghai-Tibetan Plateau." *Frontiers in veterinary science* 8 (2021): 673822.

[37]Li, Bin, et al. "Rumen microbiota of indigenous and introduced ruminants and their adaptation to the Qinghai–Tibetan plateau." *Frontiers in Microbiology* 13 (2022): 1027138.

[38]Dao, Trong-Khoa, et al. "Understanding the role of Prevotella genus in the digestion of lignocellulose and other substrates in Vietnamese native goats' rumen by metagenomic deep sequencing." *Animals* 11.11 (2021): 3257.

[39]Betancur-Murillo, Claudia Lorena, Sandra Bibiana Aguilar-Marín, and Juan Jovel. "Prevotella: A key player in ruminal metabolism." *Microorganisms* 11.1 (2022): 1.

[40]Wang, C-M., et al. "Characterization of a novel thermophilic, cellulose-degrading bacterium Paenibacillus sp. strain B39." *Letters in applied microbiology* 47.1 (2008): 46-53.

[41]Ravachol, Julie, et al. "Combining free and aggregated cellulolytic systems in the cellulosome-producing bacterium Ruminiclostridium cellulolyticum." *Biotechnology for Biofuels* 8 (2015): 1-14.

[42]Guo, Baodi, et al. "Comparative characterization of bacterial communities in geese consuming of different proportions of ryegrass." *PloS one* 14.10 (2019): e0223445.

[43]Qian, Wenxi, et al. "Bacterial colonisation of reeds and cottonseed hulls in the rumen of Tarim red deer (Cervus elaphus yarkandensis)." *Antonie Van Leeuwenhoek* 112.9 (2019): 1283-1296.

[44]Sun, Shanshan, et al. "Insight into biodegradation of cellulose by psychrotrophic bacterium Pseudomonas sp. LKR-1 from the cold region of China: optimization of cold-active cellulase production and the associated degradation pathways." *Cellulose* 27 (2020): 315-333.

[45]Speck, E. L., and E. Freese. "Control of metabolite secretion in Bacillus subtilis." *Microbiology* 78.2 (1973): 261-275.

[46]Marwoto, B., et al. "Metabolic analysis of acetate accumulation during xylose consumption by Paenibacillus polymyxa." *Applied microbiology and biotechnology* 64 (2004): 112-119.

[47]Pascal, Victoria, et al. "A microbial signature for Crohn's disease." *Gut* 66.5 (2017): 813-822.

[48]Mukherjee, Arghya, et al. "Gut microbes from the phylogenetically diverse genus Eubacterium and their various contributions to gut health." *Gut microbes* 12.1 (2020): 1802866.

[49]Sasaki, Mari, et al. "The abundance of Ruminococcus bromii is associated with faecal butyrate levels and atopic dermatitis in infancy." *Allergy* 77.12 (2022): 3629-3640.

[50]Aluwong, Tagang, Patricia Ishaku Kobo, and Abdullahi Abdullahi. "Volatile fatty acids production in ruminants and the role of monocarboxylate transporters: a review." *African Journal of Biotechnology* 9.38 (2010): 6229-6232.

[51]Liang, Jinsong, et al. "Promising biological conversion of lignocellulosic biomass to renewable energy with rumen microorganisms: A comprehensive review." *Renewable and Sustainable Energy Reviews* 134 (2020): 110335.

[52]Cheng, Muxi, Bruce McCarl, and Chengcheng Fei. "Climate change and livestock production: a literature review." *Atmosphere* 13.1 (2022): 140.

[53]Rampelli, Simone, et al. "Metagenome sequencing of the Hadza hunter-gatherer gut microbiota." *Current Biology* 25.13 (2015): 1682-1693.

[54]Rampelli, Simone, et al. "The gut microbiome buffers dietary adaptation in Bronze Age domesticated dogs." *IScience* 24.8 (2021).

[55]Pasolli, Edoardo, et al. "Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle." *Cell* 176.3 (2019): 649-662.

[56]Mukherjee, Supratim, et al. "1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life." *Nature biotechnology* 35.7 (2017): 676-683.

[57]Stewart, Robert D., et al. "Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen." *Nature communications* 9.1 (2018): 870.

[58]Stewart, Robert D., et al. "Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery." *Nature biotechnology* 37.8 (2019): 953-961.

[59]Langmead, Ben, and Steven L. Salzberg. "Fast gapped-read alignment with Bowtie 2." *Nature methods* 9.4 (2012): 357-359.

[60]Machado, Daniel, et al. "Fast automated reconstruction of genome-scale metabolic models for microbial species and communities." *Nucleic acids research* 46.15 (2018): 7542-7553.

[61]Parks, Donovan H., et al. "CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes." *Genome research* 25.7 (2015): 1043-1055.

[62]Olm, Matthew R., et al. "dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication." *The ISME journal* 11.12 (2017): 2864-2868.

[63]Warnes, Maintainer Gregory R., et al. "Package 'gplots'." *Various R programming tools for plotting data* (2016): 112-119.

[64]Chaumeil, Pierre-Alain, et al. "GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database." (2020): 1925-1927.

[65]Belcour, Arnaud, et al. "Metage2Metabo, microbiota-scale metabolic complementarity for the identification of key species." *Elife* 9 (2020): e61968.

[66]Uritskiy, Gherman V., Jocelyne DiRuggiero, and James Taylor. "MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis." *Microbiome* 6 (2018): 1-13.

[67]Seemann, Torsten. "Prokka: rapid prokaryotic genome annotation." *Bioinformatics* 30.14 (2014): 2068-2069.

[68]Danecek, Petr, et al. "Twelve years of SAMtools and BCFtools." *Gigascience* 10.2 (2021): giab008.

[69]Oksanen, Jari. "Vegan: community ecology package." *http://vegan. r-forge. r-project. org/* (2010).

[70]Pignatti, S., R. Guarino, and M. La Rosa. "2019." *Flora d'Italia* 1 (2017): 4.

[71]Yu, Zhongtang, and Mark Morrison. "Improved extraction of PCR-quality community DNA from digesta and fecal samples." *Biotechniques* 36.5 (2004): 808-812.

[72]Turroni, Silvia, et al. "Fecal metabolome of the Hadza hunter-gatherers: a host-microbiome integrative view." *Scientific reports* 6.1 (2016): 32826.

[73]Masella, Andre P., et al. "PANDAseq: paired-end assembler for illumina sequences." *BMC bioinformatics* 13 (2012): 1-7.

[74]Bolyen, Evan, et al. "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." *Nature biotechnology* 37.8 (2019): 852-857.

[75]Edgar, Robert C. "Search and clustering orders of magnitude faster than BLAST." *Bioinformatics* 26.19 (2010): 2460-2461.

[76]Callahan, Benjamin J., et al. "DADA2: High-resolution sample inference from Illumina amplicon data." *Nature methods* 13.7 (2016): 581-583.

[77]Rognes, Torbjørn, et al. "VSEARCH: a versatile open source tool for metagenomics." *PeerJ* 4 (2016): e2584.

[78]Quast, Christian, et al. "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools." *Nucleic acids research* 41.D1 (2012): D590-D596.

[79]Culhane, Aedín C., et al. "MADE4: an R package for multivariate analysis of gene expression data." *Bioinformatics* 21.11 (2005): 2789-2790.

[80]Martinez Arbizu, Pedro. "pairwiseAdonis: Pairwise multilevel comparison using adonis." *R package version 0.4* 1 (2020).

[81]Segata, Nicola, et al. "Metagenomic biomarker discovery and explanation." *Genome biology* 12 (2011): 1-18.

[82]Beghini, Francesco, et al. "Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3." *elife* 10 (2021): e65088.

[83]Asnicar, Francesco, et al. "Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0." *Nature communications* 11.1 (2020): 2500.

[84]Fiori, Jessica, et al. "Simultaneous HS-SPME GC-MS determination of short chain fatty acids, trimethylamine and trimethylamine N-oxide for gut microbiota metabolic profile." *Talanta* 189 (2018): 573-578.

[85]Nam, Seo Lin, et al. "Towards standardization of data normalization strategies to improve urinary metabolomics studies by GC× GC-TOFMS." *Metabolites* 10.9 (2020): 376.

2.2.3 Secondary metabolites gene clusters encoded by the ibex SGB communities

Aiming to shed light on the biosynthetic potential of our previously identified Alpine ibex gut microbiome SGB communities, we applied antibiotics and secondary metabolite analysis shell, antiSMASH 7.0[1], to our SGB community modules, grouped by season, in order to detect the biosynthetic gene clusters (BGCs) nested within these reconstructed genomes, with the following parameters: (antismash -c 20 --asf --cc-mibig --cb-general --cb-subclusters --cb-knownclusters --smcog-trees --genefinding-tool prodigal--output-dir antismash_7/${i} --output-basename ${i} ${i}). Indeed, antiSMASH is currently the most widely used tool for detecting and characterizing biosynthetic gene clusters in archaea, bacteria, and fungi[1]. Further details concerning the detected BGCs were obtained by using the Biosynthetic Gene Similarity Clustering and Prospecting Engine, BiG-SCAPE[2], version 1.1.5, with default settings.

Preliminary results of this analysis are shown in **table 1.2**, considering all the BGCs, assigned to a specific functional class, that were detected in our seasonal SGB community modules. This genome mining approach led to the discovery of a total of 67 BGCs, belonging to four different functional classes (i.e., terpene, RiPPs, NRPS, and 'others'). Specifically, the spring community encompassed 20 BGCs (4 identified as terpene, 10 as RiPPs, 3 as NRPS, and 3 as others); 27 BGCs were detected in the summer community (3 identified as terpene, 13 as RiPPs, 5 as NRPS, and 6 as others); finally, the autumn SGB community contained 20 BGCs (8 identified as terpene, 4 as RiPPs, 1 as NRPS, and 7 as others). As summarized in the table, the number of BGCs was higher in the summer microbial consortium, compared to spring and autumn, and this may also be related with the overall major alpha-diversity of CAZymes in this season, considering that secondary metabolites can be regarded as a cascade result of microbial nutrient acquisition. As a matter of fact, while secondary metabolites are not directly involved in primary cell survival, they often emerge as a response to nutrient availability, stress, and environmental conditions, given that the processes of nutrient acquisition and primary metabolism can direct the flow of carbon, nitrogen, and other essential resources into secondary metabolic pathways[3,4,5].

Table 1.2 – Summary of all the Biosynthetic Gene Clusters found across the three seasonal ibex SGB communities

BGC Classes	Spring SGBs	Summer SGBs	Autumn SGBs
Terpene	4	3	8
RiPPs	10	13	4
NRPS	3	5	1
Others	3	6	7
Total Number of BGCs	**20**	**27**	**20**

Amongst the different BGCs that were found in the investigated ibex SGB communities, two classes particularly stand out, and these are RiPPs (i.e., ribosomally-synthesized and post-translationally modified peptides) and non-ribosomal peptide synthetases (NRPS).

Ribosomally-synthesized and post-translationally modified peptides (RiPPs) are a large group of fascinating and biologically diverse molecules. RiPPs are initially synthesized by ribosomes as precursor peptides, which are then

modified post-translationally by specific enzymes. These modifications can involve the addition of various chemical groups or the formation of unusual bonds, which often enhance the biological activity of the final peptide. In particular, after ribosomal synthesis, RiPPs typically come as linear precursor peptides containing a leader peptide (often serving as a signal for processing and modification, and then removed) and a core peptide that becomes bioactive after modification. Examples of post-translational modifications in RiPPs are cyclization, methylation, amination, oxidation, glycosylation, prenylation, proteolysis, etc. These modifications can significantly alter the peptide's stability, activity, and specificity[6,7,8,9,10]. From a therapeutic perspective, many RiPPs are considered to be valuable candidates for drug development, particularly in the context of emerging antibiotic resistance. In fact, some RiPPs have shown antimicrobial, anticancer, and immune-modulatory activities[7,8,10]. Nisin, for instance, a widely-studied member of class I bacteriocins, known as lantibiotics, and mainly produced by the lactic acid bacterium *Lactococcus lactis,* is characterized by a wide spectrum of antibacterial activity[11]. It works by disrupting and permeabilizing the bacterial cell coating, particularly in Gram-positive bacteria, ultimately leading to cell death[11,12]. These properties have led to the vast use of nisin for the treatment of multi-drug resistant (MDR) infections, particularly those caused by *Listeria monocytogenes*, *Staphylococcus aureus*, *Clostridium difficile*, and *Enterococcus faecalis*[11]. Nisin is also widely exploited as a food preservative[13] and may represent a potential anti-cancer drug, as it has shown the ability to induce cell death in certain cancer cell lines, such as, for instance, breast cancer[14,15]. This is just an example of the plethora of therapeutic properties already well-characterized for RiPPs, which have been extensively studied over the past decades, especially those of microbial origin. Indeed, scientists believe that many new microbiome-derived RiPPs may be yet to be discovered and exploited, increasing their potential for further therapeutic use[6,7,8].

Non-ribosomal peptide synthetases (NRPSs) are large ultra-complex multienzyme machineries capable of synthesizing several peptide molecules with great structural and functional biodiversity, referred to as non-ribosomal peptides (NRPs)[16,17,18,19,20]. Similarly to what has been reported for RiPPs, the usefulness and potential role of NRPs as drugs is now evident, strengthening the therapeutic interest towards these molecules. As a matter of fact, NRPs already include a lot of marketed drugs, such as antibacterials (e.g., penicillin, vancomycin), antitumor compounds (e.g., bleomycin), immunosuppressants (e.g., cyclosporine), molecules that can be used in obstetrics (e.g., ergometrine), antiparasitic compounds (e.g., emodepside), as well as pain killers (e.g., ergotamine), among others[18,19]. Considering that the producers of RiPPs and NRPs are mostly bacteria and fungi, research towards microbiome-derived useful secondary metabolites belonging to these classes is opening up new very intriguing scientific scenarios that will require further extensive investigation. Increasing knowledge of these topics will be possible also thanks to the recent huge technological advancements of modern genome mining informatic tools[6,20].

Finally, also terpenes, being a very large and structurally diverse group of natural secondary metabolites produced by both macro- and microorganisms, are finding applications in numerous sectors and markets, including the pharmaceutical, nutraceutical, synthetic chemistry, flavor-fragrance, and biofuel industries, for instance[21,22,23,24,25].

Notably, since gene clusters encoding for these interesting molecules were bio-prospected by genome mining the reconstructed genomes of gut bacterial symbionts of a wild herbivore (i.e., the Alpine ibex), this preliminary in silico

overview further highlights the rich biotechnological potential of natural, dynamic, and still little-explored animal-associated microbial communities, opening up new intriguing scientific perspectives that go far beyond the scope of this thesis.

References

[1]Blin, Kai, et al. "antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation." *Nucleic acids research* 51.W1 (2023): W46-W50.

[2]Navarro-Muñoz, Jorge C., et al. "A computational framework to explore large-scale biosynthetic diversity." *Nature chemical biology* 16.1 (2020): 60-68.

[3]Pandey, K. D., et al. "Secondary metabolites from bacteria and viruses." *Natural bioactive compounds* (2021): 19-40.

[4]Thirumurugan, Durairaj, et al. "An introductory chapter: secondary metabolites." *Secondary metabolites-sources and applications* 1 (2018): 13.

[5]Seibel, Elena, et al. "Bacteria from the Amycolatopsis genus associated with a toxic bird secrete protective secondary metabolites." *Nature Communications* 15.1 (2024): 8524.

[6]Montalbán-López, Manuel, et al. "New developments in RiPP discovery, enzymology and engineering." *Natural product reports* 38.1 (2021): 130-239.

[7]Hudson, Graham A., and Douglas A. Mitchell. "RiPP antibiotics: biosynthesis and engineering potential." *Current opinion in microbiology* 45 (2018): 61-69.

[8]Vogt, Eva, and Markus Künzler. "Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics." *Applied microbiology and biotechnology* 103 (2019): 5567-5581.

[9]Le, Tung, and Wilfred A. van der Donk. "Mechanisms and evolution of diversity-generating RiPP biosynthesis." *Trends in Chemistry* 3.4 (2021): 266-278.

[10]Pfeiffer, Isabel P-M., Maria-Paula Schröder, and Silja Mordhorst. "Opportunities and challenges of RiPP-based therapeutics." *Natural Product Reports* (2024).

[11]Shin, Jae M., et al. "Biomedical applications of nisin." *Journal of applied microbiology* 120.6 (2016): 1449-1465.

[12]Khan, Fazlurrahman, et al. "Multiple potential strategies for the application of nisin and derivatives." *Critical Reviews in Microbiology* 49.5 (2023): 628-657.

[13]Gharsallaoui, Adem, et al. "Nisin as a food preservative: part 1: physicochemical properties, antimicrobial activity, and main uses." *Critical reviews in food science and nutrition* 56.8 (2016): 1262-1274.

[14]Avand, Abasaleh, Vajihe Akbari, and Shahin Shafizadegan. "In vitro cytotoxic activity of a Lactococcus lactis antimicrobial peptide against breast cancer cells." *Iranian journal of biotechnology* 16.3 (2018).

[15]Mohammadi, Ali, et al. "Glycated nisin enhances nisin's cytotoxic effects on breast cancer cells." *Scientific Reports* 14.1 (2024): 17808.

[16]Sieber, Stephan A., and Mohamed A. Marahiel. "Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics." *Chemical reviews* 105.2 (2005): 715-738.

[17]Strieker, Matthias, Alan Tanović, and Mohamed A. Marahiel. "Nonribosomal peptide synthetases: structures and dynamics." *Current opinion in structural biology* 20.2 (2010): 234-240.

[18]Süssmuth, Roderich D., and Andi Mainz. "Nonribosomal peptide synthesis—principles and prospects." *Angewandte Chemie International Edition* 56.14 (2017): 3770-3821.

[19]Patel, Ketan D., et al. "Structural advances toward understanding the catalytic activity and conformational dynamics of modular nonribosomal peptide synthetases." *Natural Product Reports* 40.9 (2023): 1550-1582.

[20]Bozhüyük, Kenan AJ, et al. "Evolution-inspired engineering of nonribosomal peptide synthetases." *Science* 383.6689 (2024): eadg4320.

[21]Guimarães, Adriana G., Mairim R. Serafini, and Lucindo J. Quintans-Júnior. "Terpenes and derivatives as a new perspective for pain treatment: a patent review." *Expert opinion on therapeutic patents* 24.3 (2014): 243-265.

[22]Silvestre, Armando JD, and Alessandro Gandini. "Terpenes: major sources, properties and applications." *Monomers, polymers and composites from renewable resources*. Elsevier, 2008. 17-38.

[23]Tetali, Sarada D. "Terpenes and isoprenoids: a wealth of compounds for global use." *Planta* 249 (2019): 1-8.

[24]Fan, Mei, et al. "Application of terpenoid compounds in food and pharmaceutical products." *Fermentation* 9.2 (2023): 119.

[25]Noriega, Paco. "Terpenes in essential oils: Bioactivity and applications." *Terpenes and terpenoids—Recent advances* (2020).

Chapter 3 – Concluding remarks and future perspectives

The above-illustrated PhD thesis deals with the microbial dimension of some agricultural and natural ecosystems that have proved to be worth investigating due to their vast potential in an environmental sustainability and circular green economy framework.

In the first chapter of this work, I reported the results of two Original Research Articles, that were recently published, whose direct focus were plant-associated growth-promoting microorganisms, particularly bacteria and fungi, living in symbiosis with the grapevine holobiont, as well as thriving in the bulk soil of some well-known Italian vineyards found across two leading Italian wine-producing regions (i.e., Emilia Romagna and Tuscany).

While Study I is based solely on 16S metabarcoding amplicon sequencing for bacterial community structure profiling, Study II is more comprehensive, relying on multiple multi-omics techniques, such as ITS metabarcoding for analyzing the grapevine mycobiome, shotgun metagenomics, and untargeted metabolomics. Nevertheless, I reckon that both studies were useful for confining and defining, to some extent, the microbial terroir, namely the unique microbial configurations found in a specific vineyard and influenced by a plethora of environmental factors, of the investigated viti-vinicultural regions.

With this research, we wished to shed more light on i) the presence of specific microbial terroirs for vineyards and grapevines also at very narrow geographic scales ii) the functional potential of these microorganisms and their putative probiotic roles in supporting grapevine growth and fitness under natural conditions iii) the importance of gaining insight into grapevine microbiology in the context of global change, as native microbial communities may be the leverage to exploit for the creation of new effective bio-fertilizers/bio-pesticides/bio-stimulants. As a matter of fact, in order to counteract alarming environmental declines, the next decades must see substantial changes of the current food production paradigms, with a reduction of chemical agricultural inputs and a sharp increase of organic microbiome-based alternatives.

In this view, the two experimental surveys illustrated in this thesis work, which were carried out during my PhD, may be suited for providing valuable baseline microbiome data, which can be regarded as the pivotal starting point for a more sustainable and tailored wine production.

Future research efforts in this direction must be implemented with the goal of taking full advantage of microbial communities in viticulture. Specifically, putative plant growth-promoting microbes previously identified, whether bacteria, fungi, or others, should be isolated and cultivated in the lab, where it is more feasible to screen for plant growth-promoting traits. Whole-genome sequencing of candidate isolated microorganisms will be essential for this kind of investigation. After that, the best microbial candidates for grapevine bio-stimulation should be inoculated in pot, using both model plant species (e.g., *Arabidopsis thaliana*), as well as *V. vinifera* itself, where controlled conditions will facilitate the understanding of microbiome dynamics at the soil-root interface and their effects on plant growth.

Finally, in order to translate both bench and in silico results to the actual viticultural sector, the ultimate and most challenging research trials must be directed towards testing the effects of plant probiotics in the field, by means of vineyard inoculation of microbial strains and/or mixed consortia, that, if effective, could actually reduce the load of chemicals that are continuously spread in the soil.

The second chapter of this thesis illustrates a possible gut microbiome-based solution to bio-transform plant-derived lignocellulose into high-value-added platform chemicals, which may be exploited industrially for bio-material production, reducing the need for environmentally dangerous fossil fuels in chemical synthetic processes. These results were also recently published in a peer-reviewed scientific journal.

Specifically, three different simplified microbial consortia were identified starting from fecal samples of a wild (and neglected) herbivorous holobiont, namely the wild Alpine ibex, a ruminant belonging to the *Bovidae* family and found at Stelvio National Park. We chose to target this animal since wild herbivores thriving in highly seasonal habitats are expected to host incredibly plastic and rich GMs, since well adapted to cope with rough heterogeneous plant food, matching the changing vegetation patterns of these ecosystems throughout the year.

Indeed, our multi-omics investigation (by means of bacterial metabarcoding, shotgun metagenomics, and targeted/untargeted metabolomics) led to the biodiscovery of three microbial hubs potentially capable of digesting LC while producing platform chemicals and relevant secondary metabolites. Interestingly, these hubs were substantially different according to the sampling season, comparing spring, summer, and autumn.

Our findings clearly confirmed the potential of natural microbiomes, in this case associated with a wild herbivorous ruminant, to be exploited in biotechnological/industrial markets, starting from raw materials with high value and zero cost (i.e., LC waste-streams) that represent a suitable substrate for microbial growth.

Despite being very promising, our results should be implemented in the future by performing downstream microbiological/culturomics assays and trials, in order to highlight the actual ability of Alpine ibex-associated microbes to digest LC and produce useful compounds.

At the lab scale, these experiments should start by fecal isolation of promising candidate microbial taxa (mainly bacteria and fungi) grown in LC-rich culture media. These microorganisms should then be subjected to molecular characterization (e.g., via Sanger marker gene sequencing or whole genome sequencing) and bio-banked under appropriate storage conditions.

Batch fermentation of these strains (either in mono- or co-cultures) may provide initial glimpses on their LC bioconversion capability, via monitoring microbial growth and metabolite production in the reactor. Nevertheless, hurdles related to the several difficulties of mimicking ultra-complex natural ecosystems at the lab scale may arise, and therefore hamper the overall yield/quality of the process. These obstacles still represent one of the major challenges for the full exploitation of herbivores gut microbiomes as bioreactor-based circular biorefineries. As a matter of fact, the design and subsequent buildup of efficient synthetic microbial communities based on natural

ecosystems has not been optimized yet. Furthermore, enzyme purification from isolated microbial strains may also lead to innovative and low-impact LC valorization solutions.

In order to achieve these goals, increased knowledge of ecological rules and dynamics underlying all the intricate host-microbe and microbe-microbe interactions taking place in the herbivore gut must be obtained, and some was provided also in this thesis. In this scenario, detailed meta-omics (including meta-transcriptomics and meta-proteomics), combined with in-depth microbial network analyses, will be crucial to acquire the baseline expertise needed to assemble synthetic simplified microbiomes capable of valorizing lignocellulose to useful molecules.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to all the members of the groups and institutions, listed below, whose valuable work has been pivotal for the realization of this PhD thesis.

- Microbiome Science and Biotechnology Unit (University of Bologna), brilliantly led by my supervisor Professor Marco Candela, who hosted me for the past three years as a PhD candidate, giving me the opportunity to learn more than I could have ever imagined possible.
- Human Microbiomics Unit (University of Bologna), led by my former supervisor Professor Patrizia Brigidi.
- MicroEcoEvo Lab (University of Lisbon), led by Professor Rodrigo Costa, where I spent six wonderful months during my PhD experience abroad, consolidating my skills and knowledge related to the 'microbiome science'.
- Department of Pharmacy and Biotechnology (University of Bologna) and the coordinator of my PhD board Professor Maria Laura Bolognesi.
- Department of Chemistry "Giacomo Ciamician" (University of Bologna), and specifically Professor Jessica Fiori for the contributions given to our research.
- Department of Biological, Geological, and Environmental Sciences, where I studied Biology for 5 wonderful years.
- PON ''Ricerca e Innovazione'' 2014-2020 program.
- "Controlling Microbiomes Circulations for Better Food Systems" (CIRCLES) project (European Union's Horizon 2020 research and innovation program; grant agreement no. 818290).
- Eurovix S.p.A. – Biotecnologie per la vita – Entratico (BG), i.e., the partner company of my PhD, that has taken active part in the work illustrated in this thesis and with which I directly collaborated for six months during the course of an agro-biotech field trial under the supervision of PhD Nicola Secchi (Research & Development).
- Consorzio Vino Nobile di Montepulciano – Montepulciano (SI).
- Istituto Tecnico Statale "Ignazio Calvi" – Finale Emilia (MO).
- Parco Nazionale dello Stelvio – Bormio (SO).

Finally, my deepest and biggest gratitude goes to my fantastic **friends** and **family**, without whom none of this would have been possible, and, of course, a special thanks to my **MOM**, the pillar of everything.

PUBLICATIONS FROM THE AUTHOR

Original Research Articles

1. **Nanetti**, Enrico, et al. "Composition and biodiversity of soil and root-associated microbiome in Vitis vinifera cultivar Lambrusco distinguish the microbial terroir of the Lambrusco DOC protected designation of origin area on a local scale." *Frontiers in Microbiology* 14 (2023): 1108036. https://dx.doi.org/10.3389/fmicb.2023.1108036

2. **Nanetti**, Enrico, et al. "The Alpine ibex (Capra ibex) gut microbiome, seasonal dynamics and potential application in lignocellulose bioconversion." *iScience* (2024). https://doi.org/10.1016/j.isci.2024.110194

3. Palladino, Giorgia, **Nanetti**, Enrico, et al. "Zonation of the Vitis vinifera microbiome in Vino Nobile di Montepulciano PDO production area." *Communications Biology* 7.1 (2024): 1-14. https://doi.org/10.1038/s42003-024-07261-8

4. Radaelli, Elena, et al. "Meta-analysis of the Cetacea gut microbiome: Diversity, co-evolution, and interaction with the anthropogenic pathobiome." *Science of The Total Environment* (2024): 172943. https://doi.org/10.1016/j.scitotenv.2024.172943

5. Ruiz, Alberto, et al. "Microbiome study of a coupled aquaponic system: unveiling the independency of bacterial communities and their beneficial influences among different compartments." *Scientific Reports* 13.1 (2023): 19704. https://dx.doi.org/10.1038/s41598-023-47081-0

6. Palladino, Giorgia, et al. "Seasonal dynamics of the microbiome-host response to pharmaceuticals and pesticides in Mytilus galloprovincialis farmed in the Northwestern Adriatic Sea." *Science of The Total Environment* 887 (2023): 163948. https://dx.doi.org/10.1016/j.scitotenv.2023.163948

7. Marchi, Arianna, et al. "Towards a free wild-caught fishmeal, fish oil and soy protein in European sea bass diet using by-products from fishery and aquaculture." *Aquaculture* 573 (2023): 739571. https://dx.doi.org/10.1016/j.aquaculture.2023.739571

8. Scicchitano, Daniel, et al. "Microbiome network in the pelagic and benthic offshore systems of the northern Adriatic Sea (Mediterranean Sea)." *Scientific Reports* 12.1 (2022): 16670. https://dx.doi.org/10.1038/s41598-022-21182-8

9. Palladino, Giorgia, et al. "Plasticity of the Anemonia viridis microbiota in response to different levels of combined anthropogenic and environmental stresses." *Frontiers in Marine Science* 9 (2022): 956899. https://dx.doi.org/10.3389/fmars.2022.956899

10. Meschini, Marta, et al. "Reliability of data collected by volunteers: A nine-year citizen science study in the Red Sea." *Frontiers in Ecology and Evolution* 9 (2021): 694258. https://dx.doi.org/10.3389/fevo.2021.694258

Posters and Abstracts

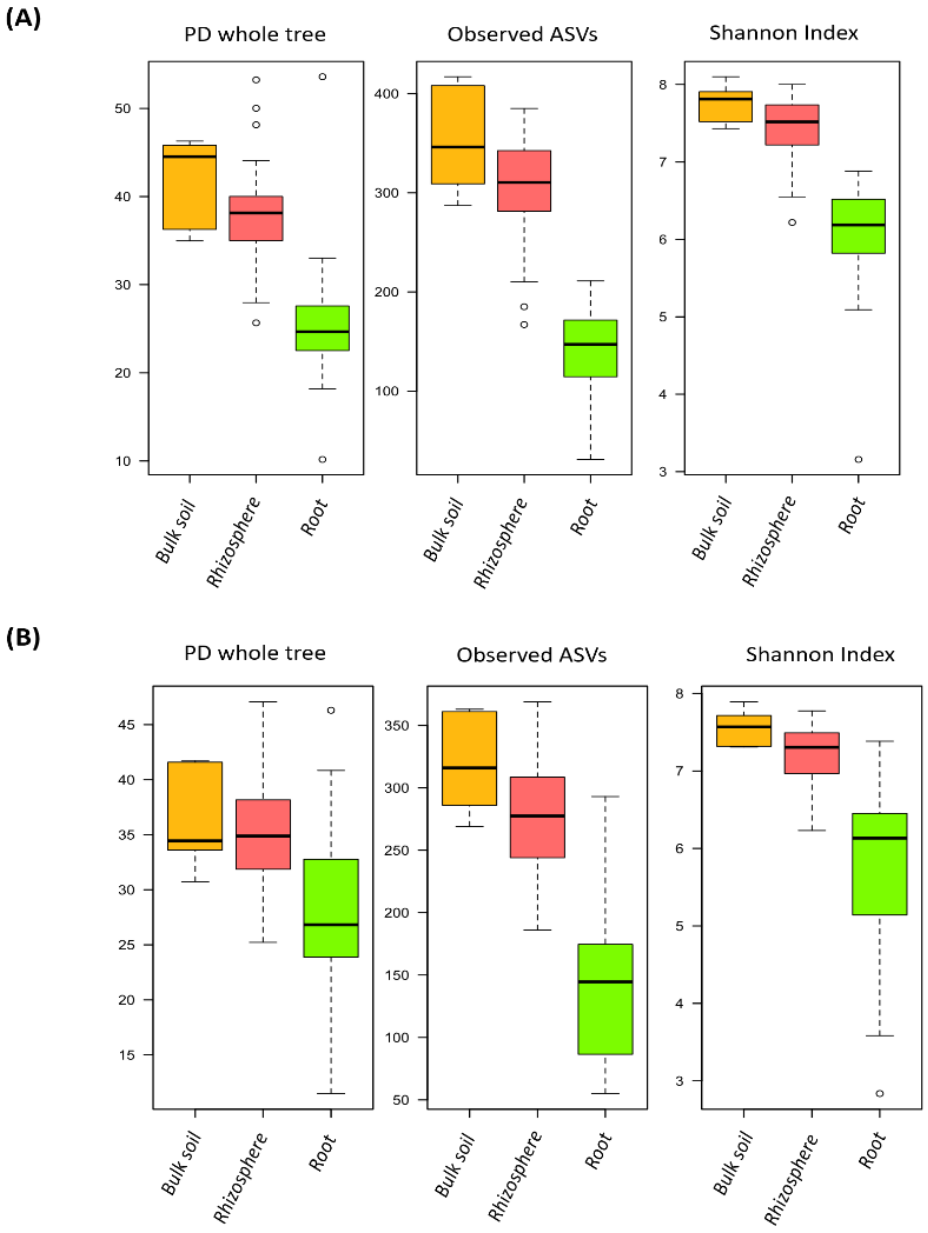
1. **Nanetti**, Enrico, et al. "Adaptive seasonal dynamics of the gut microbiome of the wild ibex from the Stelvio National Park, Italy." *FEMS Conference on Microbiology*. 2022.

2. Meschini, M., et al. "Reliability of data collected by volunteers, a nine-year citizen science study in three Red Sea touristic facilities." *5th Annual International Conference on participatory research, citizen science, crowd-innovation and fab labs for peace and development, United Nations Organization, Geneva (Switzerland), 8-10 December, 2020*. 2020.

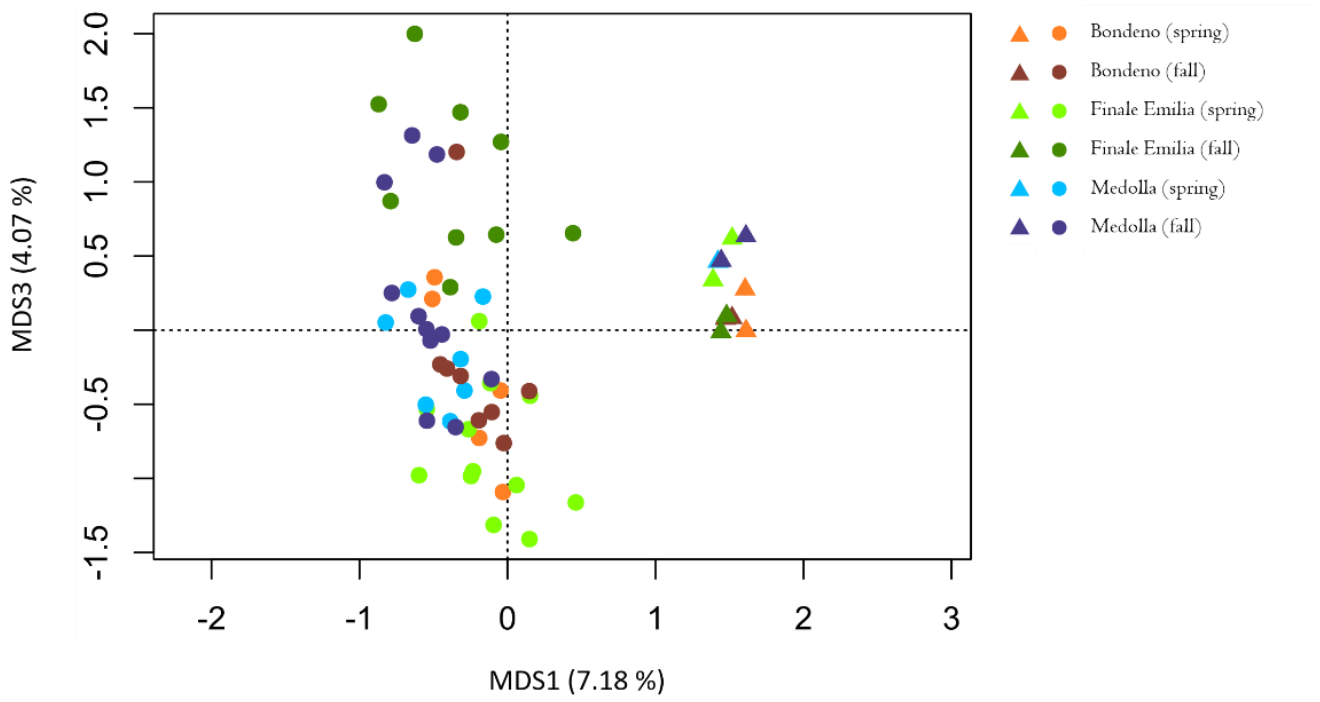
3. Radaelli, Elena, et al. "Impact of anthropogenic pathobiome on the gut microbiome of Mediterranean whales." *Impact of anthropogenic pathobiome on the gut microbiome of Mediterranean whales*. 2024.

APPENDIX

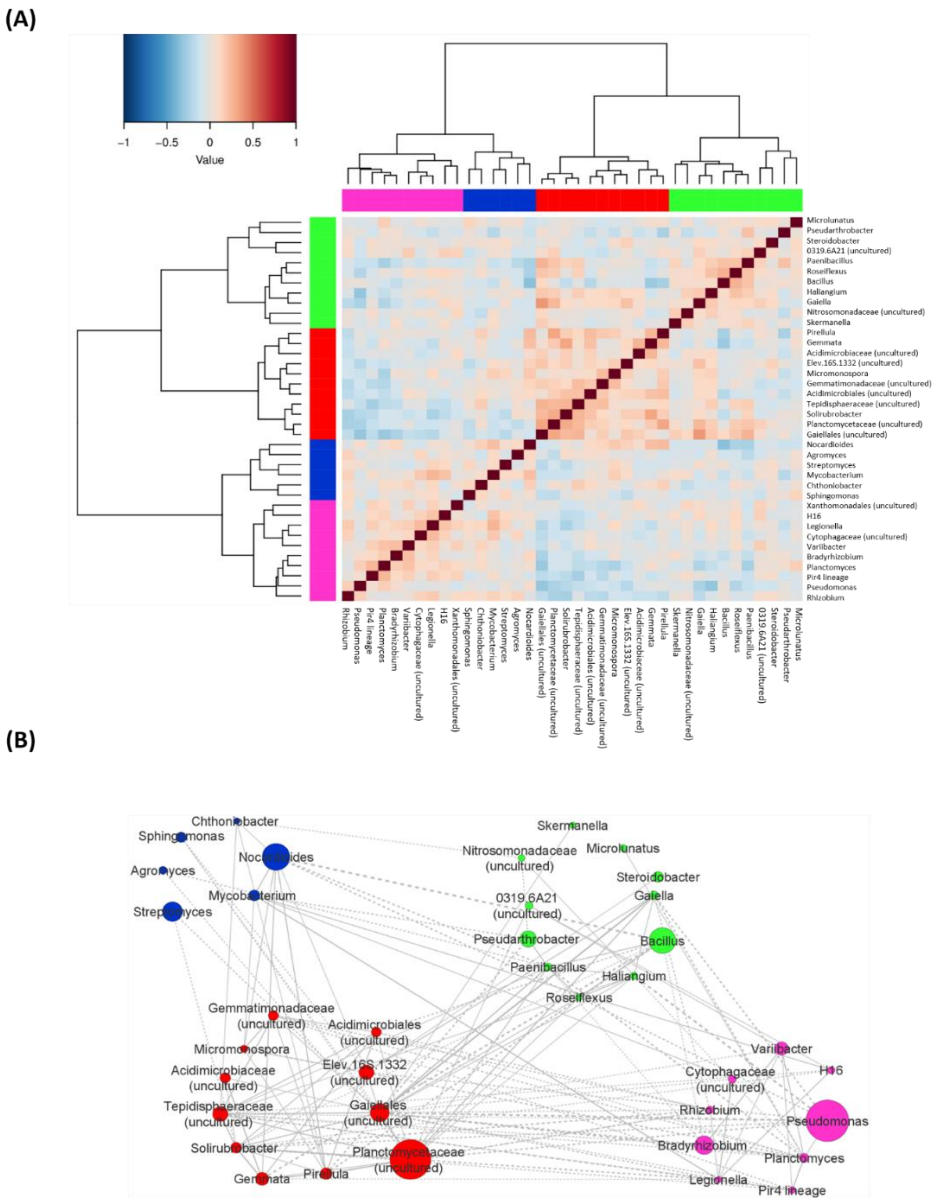
Study I



Supplementary Figure. S1 - Alpha diversity of bulk soil and of *V. vinifera* rhizosphere and root microbiomes in spring and fall. Box-plots showing the distributions of the Faith's Phylogenetic Diversity (PD whole tree), Observed ASVs and Shannon Index calculated for all samples of bulk soil and of *V. vinifera* (rhizosphere and root) in June **(A)** and November **(B)** 2021. According to all metrics, alpha diversity is higher in the bulk soil and in the rhizosphere and shows a steep decrease in the endophytic microbial communities ($P \leq 0.05$, Kruskal-Wallis test).



Supplementary Figure. S2 - Principal Coordinates Analysis (PCoA) based on unweighted UniFrac distances showing the variation of *V. vinifera* root (dots) and bulk soil (triangles) microbiomes across sites, i.e., Bondeno (orange-red), Finale Emilia (green) and Medolla (blue) and seasons (lighter shades for spring and darker shades for fall). The only significant differences emerge when comparing the two groups including all root samples and all bulk soil samples (permutation test with pseudo F-ratio, $P \leq 0.001$). The first and third principal components (MDS1 and MDS3) are plotted and the percentage of variance in the dataset explained by each axis is highlighted.



Supplementary Figure. S3 - Co-abundance associations between rhizospheric *V. vinifera* bacterial genera. **(A)** The assignment of co-abundance groups (CAGs) is based on a heat plot representing Kendall correlations between genera clustered by using the Spearman correlation coefficient and the Ward linkage hierarchical clustering method. Only genera whose relative abundance was higher than 0.5% in at least 33% of the samples are represented. Different colors indicate the four identified CAGs. **(B)** Wiggum plot correlations between the four identified CAGs. The size of the nodes is proportional to the mean genus abundance within the cohort and the connections between nodes represent positive (solid lines) and negative (dashed lines) significant Kendall correlations between genera (controlled for multiple testing using FDR, $P \leq 0.05$).

Supplementary Table S1. Sampling details of the present study

SITE	NUMBER OF SAMPLES COLLECTED				PDO AREA	AGRICULTURAL APPROACH
	June 2021		November 2021			
	Plants	Soil	Plants	Soil		
Bondeno	15	2	15	2	No	Traditional
Finale Emilia	15	2	15	2	Yes	Traditional
Medolla	15	2	15	2	Yes	Organic

Supplementary Table S2. Relative abundance in bulk soils of bacterial genera identified by LEfSe

		MEAN RELATIVE ABUNDANCE (%) (± SD)	
	BACTERIAL GENERA	**PDO bulk soil samples**	**Non-PDO bulk soil samples**
Enriched in PDO rhizospheres	*Bacillus*	1.24 (± 0.90)	0.66 (± 0.54)
	Pseudarthrobacter	0.39 (± 0.53)	0.44 (± 0.88)
	Planctomyces	0.41 (± 0.33)	0.37 (± 0.28)
	Gaiellales (uncultured)	4.03 (± 2.53)	2.75 (± 1.99)
	Skermanella	1.09 (± 0.97)	0.64 (± 0.48)
	Pir4 lineage	0.33 (± 0.23)	0.62 (± 0.34)
	Microlunatus	0.84 (± 0.88)	0.00 (± 0.00)
	Paenibacillus	0.63 (± 0.28)	0.42 (± 0.30)
Enriched in non-PDO rhizospheres	*Nocardioides*	1.86 (± 1.17)	1.56 (± 1.06)
	Micromonospora	0.70 (± 0.55)	1.01 (± 0.59)
	Gemmatimonadaceae (uncultured)	1.71 (± 0.71)	2.26 (± 0.97)
	Pirellula	1.14 (± 0.45)	1.47 (± 0.91)
	Legionella	0.16 (± 0.09)	0.04 (± 0.05)
	Mycobacterium	0.28 (± 0.33)	0.51 (± 0.57)
	Acidimicrobiales (uncultured)	1.15 (± 1.13)	1.49 (± 0.98)
	Chthoniobacter	0.45 (± 0.38)	0.26 (± 0.19)

SD (standard deviation)

Supplementary Table S3. PGP genes accession and version numbers (NCBI protein database)

Gene	Accession	Version
NifB	WP_011024080	WP_011024080.1
NifE	WP_014404757	WP_014404757.1
NifH	WP_010870393	WP_010870393.1
NifN	WP_011241560	WP_011241560.1
NifV	WP_011241567	WP_011241567.1
NifU	NP_461477	NP_461477.1
phoA	NP_414917	NP_414917.2
GDH	NP_388275	NP_388275.1
EntF	WP_000077784	WP_000077784.1
EntS	NP_415123	NP_415123.1
FslA	WP_003037766	WP_003037766.1
ipdC	WP_035671558	WP_035671558.1
aro10	NP_010668	NP_010668.3
aldH	NP_001260290	NP_001260290.1
AcdS	XP_037178185	XP_037178185.1

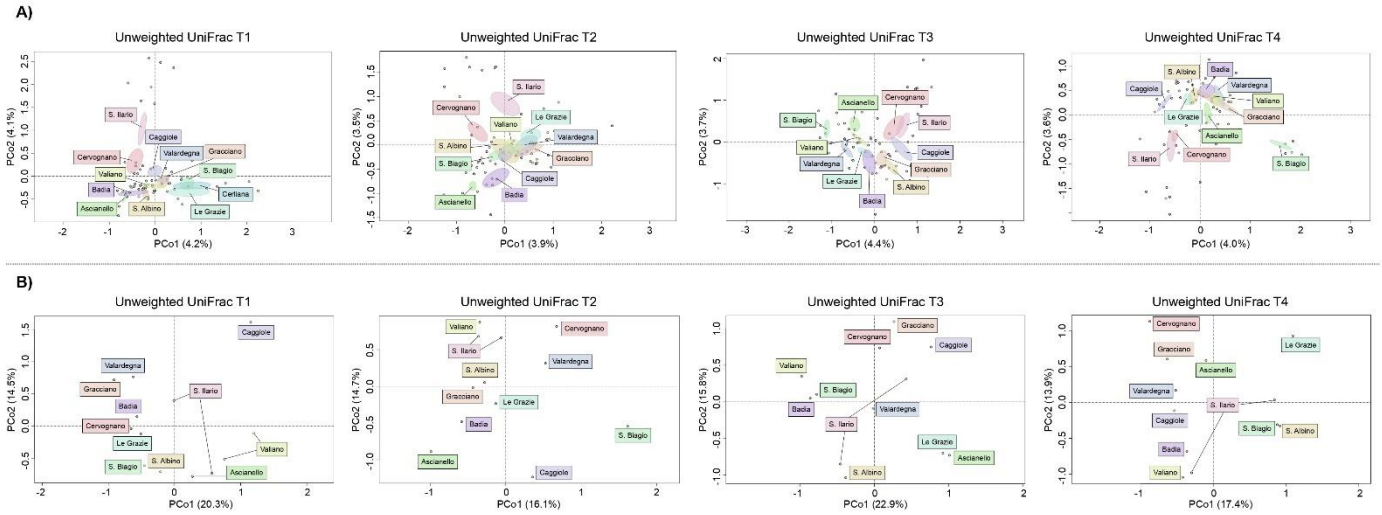
Supplementary Table S4. PGP genes K numbers (KEGG ORTHOLOGY database)

Gene	K number
NifB	K02585
NifE	K02587
NifH	K02588
NifN	K02592
NifV	K02594
NifU	K04488
phoA	K01077
GDH	K00034
EntF	K02364
EntS	K08225
FslA	-
ipdC	K04103
aro10	-
aldH	K00128
AcdS	K01505

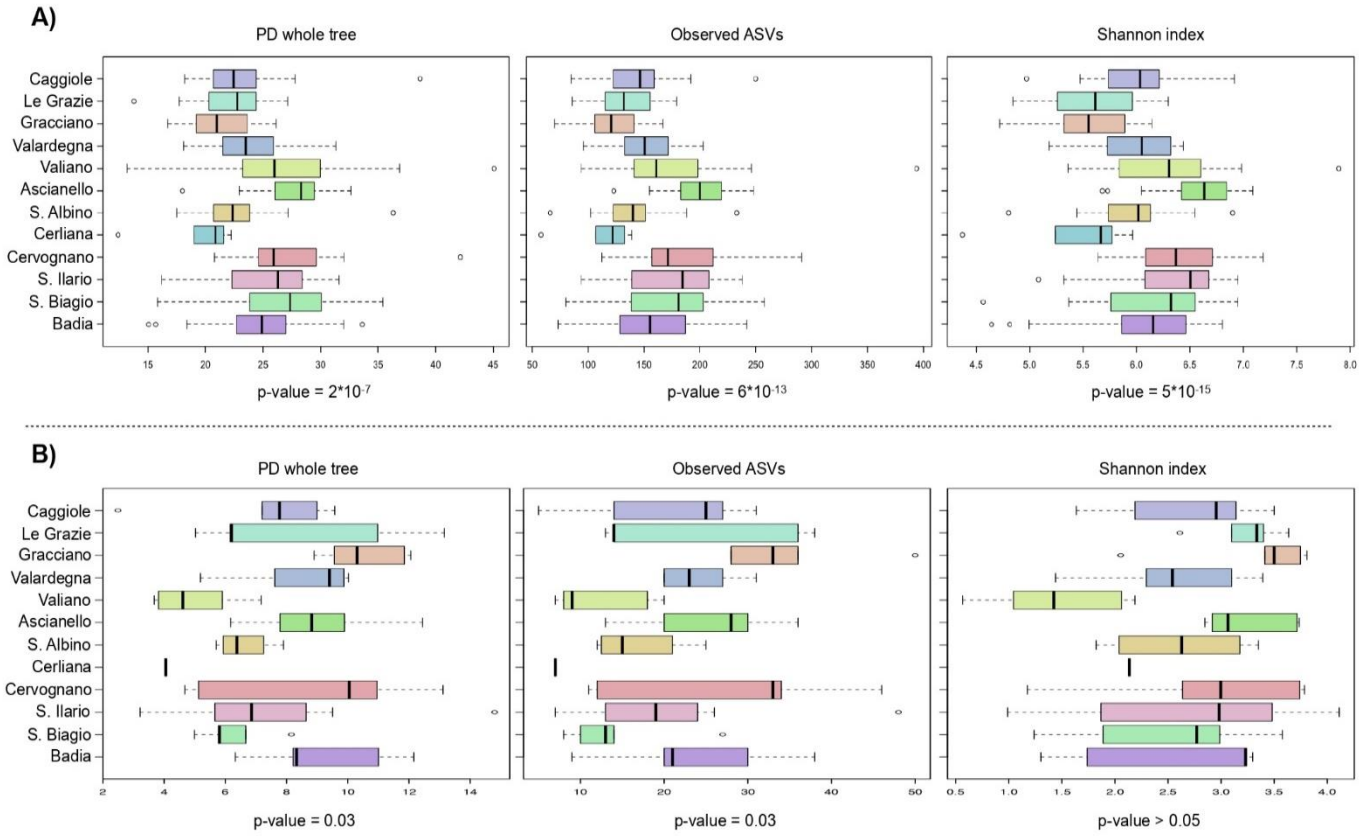
Supplementary Table S5. PGP functions identified by PICRUSt2 (KEGG ORTHOLOGY database)

PGP function	Function-encoding bacteria	Function confirmed with PICRUSt2
Nitrogen fixation	*Bacillus*	No
	Unclassified Gaiellales	Yes
	Azospirillum/Skermanella	Yes
	Pirellulales	No
	Paenibacillus	Yes
	Nocardioides	No
Phosphorous solubilization	*Bacillus*	Yes
	Pseudarthrobacter	Yes
	Planctomyces	Yes
	Azospirillum/Skermanella	Yes
	Pirellulales	Yes
	Paenibacillus	Yes
	Nocardioides	Yes
	Micromonospora	Yes
	Mycolicibacterium	Yes
Siderophore production	*Bacillus*	No
	Pseudarthrobacter	Yes
	Azospirillum/Skermanella	No
	Pirellulales	No
	Microlunatus	Yes
	Paenibacillus	No
	Nocardioides	No
	Micromonospora	No
	Gemmatirosa	Yes
	Legionella	No
	Mycolicibacterium	Yes
	Unclassified Acidimicrobiales	No
	Chthoniobacter	No
IAA production	*Bacillus*	Yes
	Planctomyces	Yes
	Azospirillum/Skermanella	No
	Pirellulales	Yes
	Paenibacillus	Yes
ACC deaminase production	*Bacillus*	No
	Azospirillum/Skermanella	No
	Microlunatus	No
	Paenibacillus	No
	Nocardioides	Yes
	Unclassified Acidimicrobiales	No

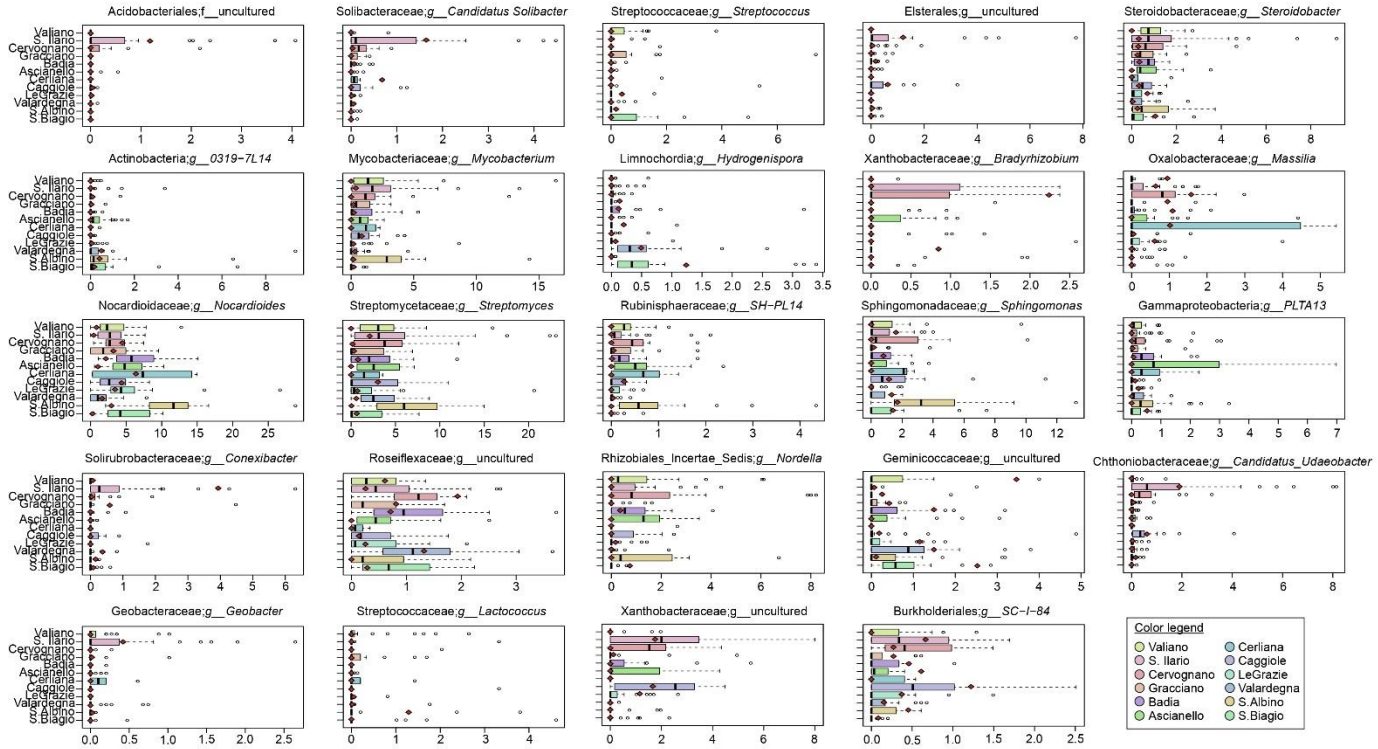
Study II



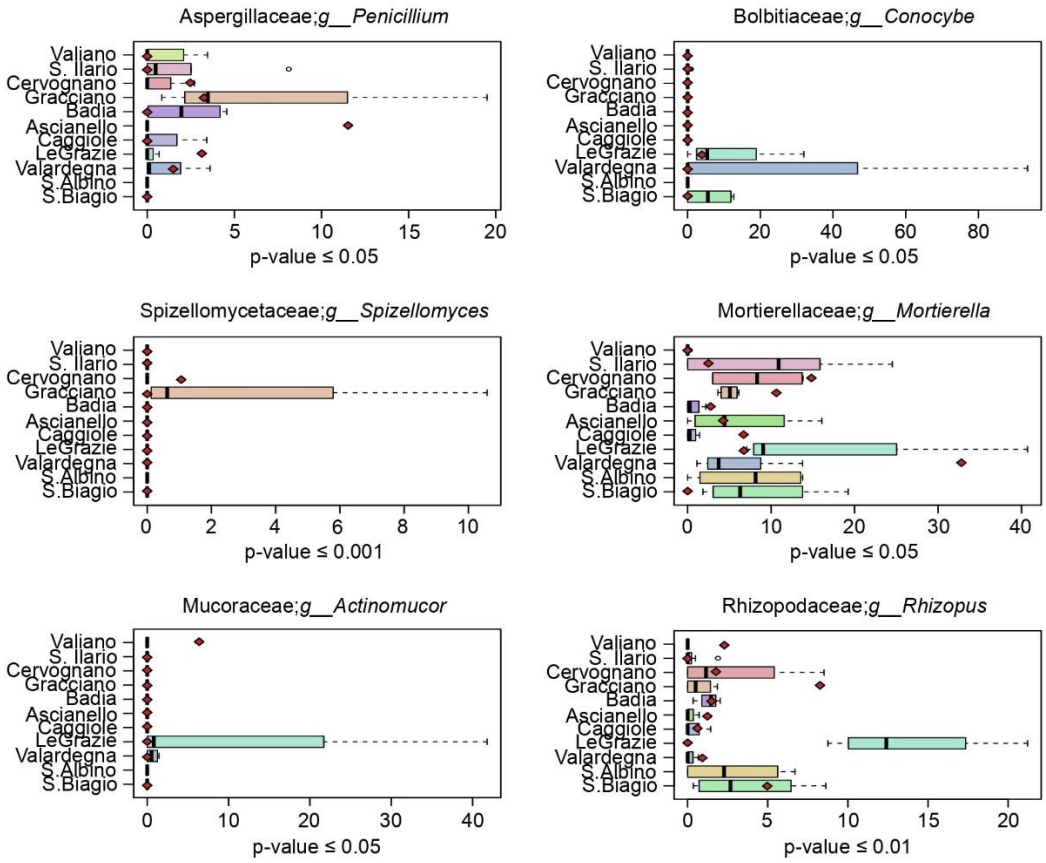
Supplementary Figure. S1 - Comparison of bacterial and fungal composition of rhizospheric soil samples of Montepulciano AGUs (Additional Geographical Units) across time points (T1 to T4). **(A)** Principal Coordinate Analysis (PCoA) based on unweighted UniFrac distances showing the variations of *Vitis vinifera* rhizospheric soil bacterial composition in the different AGUs and time points. Procrustes test between time points, p-value = 0.005. **(B)** Principal Coordinate Analysis (PCoA) based on unweighted UniFrac distances showing the variations of *Vitis vinifera* rhizospheric soil fungal composition in the different AGUs and time points. Procrustes test between time points, p-value = 0.01. The first and second principal components are plotted and the percentage of variance in the dataset explained by each axis is highlighted.



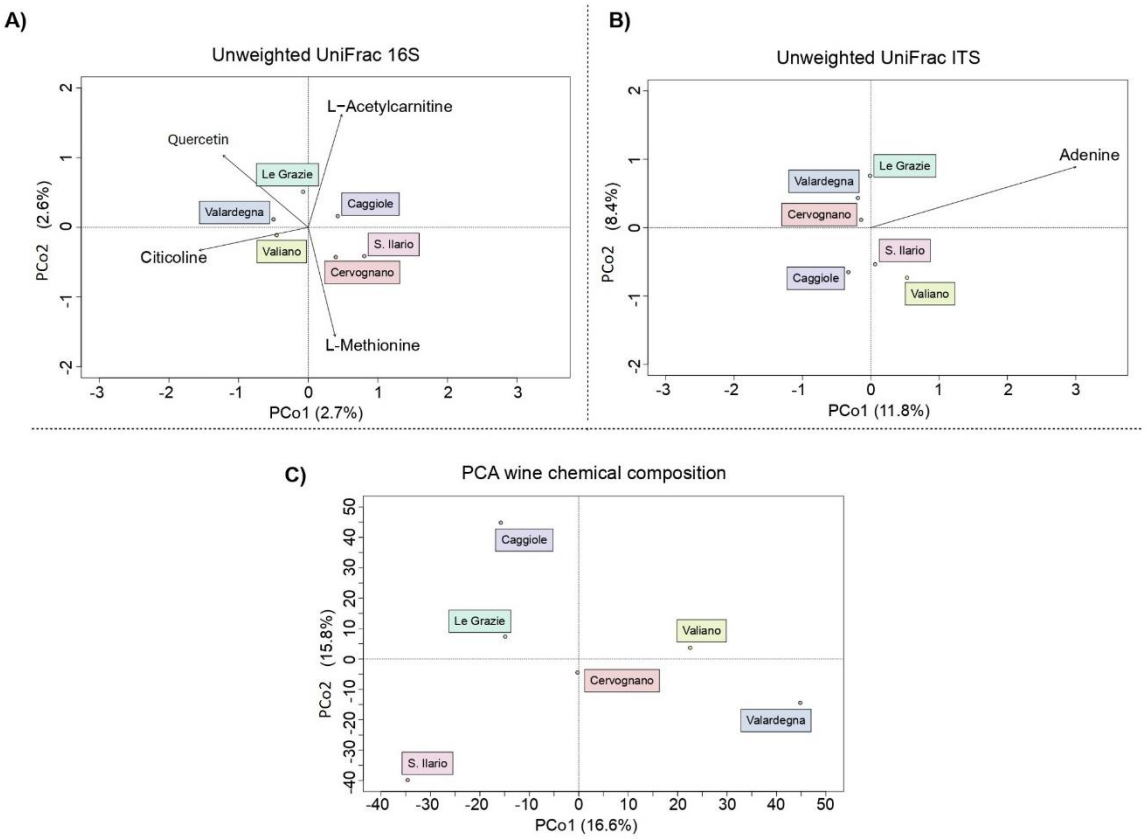
Supplementary Figure. S2 - Diversity of microbial communities across Montepulciano AGUs (Additional Geographical Units). Boxplots showing the alpha-diversity distributions of the microbiome associated with samples from different AGUs for bacterial **(A)** and fungal **(B)** communities, based on the Faith's phylogenetic diversity (PD whole tree), the number of observed amplicon sequence variants (ASVs), and the Shannon index. Significant differences between AGUs are indicated below each graph (Kruskal-Wallis test).



Supplementary Figure. S3 - Relative abundance of the 24 bacterial taxa that discriminate among Additional Geographical Units (AGUs). The relative abundance distribution of the discriminant bacterial taxa in the rhizosphere is represented in each boxplot, whereas the mean relative abundance of the same taxa in the respective soil samples is represented with a red rhombus. Significant variations in relative abundance were assessed by Kruskal-Wallis test controlled for multiple testing using False Discovery Rate (FDR) (p-value ≤ 0.001 for all taxa). AGUs are reported in the first column of boxplots and in the color legend at the bottom right.



Supplementary Figure. S4 - Relative abundance of the 6 fungal taxa that discriminate among Additional Geographic Units - AGUs. The relative abundance distribution of the discriminant fungal taxa in the rhizosphere is represented in each boxplot, whereas the mean relative abundance of the same taxa in the respective soil samples is represented with a red rhombus. Significant variations in relative abundance were assessed by Kruskal-Wallis test controlled for multiple testing using False Discovery Rate (FDR) (p-values are reported below each boxplot). AGUs are indicated next to each boxplot.



Supplementary Figure. S5 - Impact of wine chemical composition on rhizospheric microbial communities of Montepulciano AGUs (Additional Geographical Units). Principal Coordinate Analysis (PCoA) based on unweighted UniFrac distances between bacterial **(A)** and fungal **(B)** profiles of *Vitis vinifera* cultivar Sangiovese rhizospheres in the different Additional Geographical Units (AGUs) of the Montepulciano territory. Only the six AGUs selected for wine chemical composition are shown. Superimposition of wine analytical components through the envfit function are shown (p-values ≤ 0.05). **(C)** Principal Component Analysis of wine analytical components in the six selected AGUs. In all graphs, the first and second principal components are plotted and the percentage of variance in the dataset explained by each axis is highlighted.

Supplementary Table 1 - Site and plant characteristics of all vineyards in the Montepulciano consortium. Additional Geographic Units (AGUs) are given in the first column. Plant age is given in years (y), altitude in meters above sea level (masl), and soil composition in percentage of soil type (gravel, sand, silt, and clay). Soil composition was retrieved by Costantini et al. (2021). "--", no information available.

AGU	Vineyard	Plant age (y)	Rootstock	Rootstock type	Clone family	Agronomical practice	Management	Latitude	Longitude	Altitude (masl)	Soil composition (%)			
											Gravel	Sand	Silt	Clay
Valiano	Tenuta Tre Rose	> 15	420 A - Gravesac	Berlanderi x Riparia	CG5 - CV20	Plow sole	Biological	43.16022892	11.93597164	337	5	47	29	24
	Palazzo Vecchio	> 15	1103P, 110R, 775P	Berlanderi x Rupestris	VCR	Partial grassing	Conventional	43.1657283	11.90092418	320	5	47	29	24
Cervognano	Boscarelli	> 15	1103, 110R, 420A	Berlanderi x Rupestris	CH20	Partial grassing	Integrated	43.10599529	11.85076717	301	9	42	36	22
Cerliana	La Ciarliana	> 15	420A	Berlanderi x Riparia	R	Plow sole	Conventional	43.11531862	11.81196374	340	9	42	36	22
S. Albino	Fattoria del Cerro	> 15	K5BB	Berlanderi x Riparia	CH20	Plow sole	Integrated	43.07135327	11.80867656	470	9	42	36	22
Valardegna	Salcheto	> 15	1103P	Berlanderi x Rupestris	Prugnolo Grifo/Bravio	Partial grassing	Biological	43.08663725	11.80031461	422	9	42	36	22
S. Biagio	Le Bertille	> 15	--	--	--	--	--	43.10096085	11.76677398	380	9	42	36	22
S. Ilario	Podere Casanova	> 15	3309C	Riparia x Rupestris	R	Grassing	Biological	43.07415305	11.89457831	260	10	77	14	9
	Bindella	> 15	110R	Berlanderi x Rupestris	VCR	Partial grassing	Integrated	43.09568433	11.86674385	310	9	42	36	22
Le Grazie	Fattoria della Talosa	> 15	110R	Berlanderi x Rupestris	G76	Partial grassing	Biological	43.11157039	11.78225739	410	9	42	36	22
Ascianello	De' Ricci	> 15	1103P, K5BB	Berlanderi x Rupestris, Berlanderi x Riparia	CH20	Partial grassing	Biological	43.140338	11.791788	350	9	42	36	22
Gracciano	Fattoria Svetoni	> 15	--	--	--	Partial grassing	Biological	43.13461785	11.8367263	280	10	77	14	9
Caggiole	Azienda agricola Tiberini	> 15	1103P, 110R, 140R, 157.11	Berlanderi x Rupestris, Berlanderi x Riparia	SG	Partial grassing	Biological	43.12371226	11.80025501	310	9	42	36	22
Badia	Avignonesi	> 15	--	--	--	--	Biodynamic	43.157433	11.81136	280	9	42	36	22

Supplementary Table 2 - Chemical soil properties in the vineyards from different AGUs

AGU	N (mg/Kg)	P (mg/Kg)	K (mg/Kg)	pH	Hum (%)	EC	Temp °C
Ascianello	27	37	75	6.73	20	375	21.2
Badia	12	17	37	6.72	13	171	24.1
Caggiole	6	9	19	6.63	15	97	17.5
Cerliana	1	1	3	6.72	18.4	15	24.1
Cervognano	NA	NA	NA	NA	NA	NA	NA
Gracciano	12	17	36	8.25	13.2	177	24.3
Le Grazie	10	14	30	6.42	15.7	143	17.6
S.Albino	10	14	30	7.44	15.44	149	22.1
S.Biagio	42	59	119	6.67	17.4	596	29.6
S.Ilario	6	9	18	8.10	14.23	91	24.6
Valardegna	10	15	31	8.45	14.60	150	23.6
Valiano	NA	NA	NA	NA	NA	NA	NA

Supplementary Table 3 - Functional plant growth-promoting (PGP) profile of the rhizosphere and soil microbiome of Additional Geographic Units (AGUs). Abundance of PGP functions identified in the rhizosphere and bulk soil samples of different AGUs using a read-mapping approach. Data were normalized in copies per million ((reads count for an enzyme in a given sample/(gene length/1000))/(n° reads per sample/10^6).

AGU	Valardegna		S.Biagio		LeGrazie		Badia		Ascianello		S.Albino		Valiano			
Vineyard	Salcheto		Bertille		Talosa		Avignonesi		DeRicci		Cerro		PalazzoVecchio		TreRose	
Ecosystem	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil
N fixation	253.3324396	253.5608101	247.7246196	248.1121291	251.9146981	242.1440227	246.8416108	245.8961232	272.2368747	255.7700624	245.1618671	237.297965	261.4786999	258.7502105	247.0170126	289.0269023
P solubilization	519.6934706	513.8476208	517.1173927	469.8772909	539.8094403	520.8018869	541.3535649	525.0876235	502.7537045	477.5801409	504.1676451	427.0299981	462.8104732	507.0340686	490.2303127	499.6365655
Siderophore production	117.7094238	102.601028	95.15619469	86.85813397	83.7740471	89.42487643	106.68441	116.1076099	136.4253278	133.2768134	92.56504462	86.61994992	145.151287	104.0845261	107.4881701	119.5759355
IAA production	485.6309755	451.8666797	494.3931061	376.625372	504.9476595	474.7662021	492.3776903	494.6303767	484.1318267	458.8956433	485.8020601	414.2093394	461.2868008	448.963185	439.2893881	478.4947298
ACC deaminase production	23.00760103	22.72781971	21.65133262	19.59761651	22.91397761	24.6880899	21.4839648	18.71076195	18.08377848	20.64559288	20.45420007	19.14709308	20.44858956	20.73409032	13.73563757	15.15347935

AGU	Cerliana		Gracciano		Caggiole		Cervognano		S.Ilario			
Vineyard	Ciarliana		Svetoni		Tiberini		Boscarelli		Bindella		Casanova	
Ecosystem	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil	Rhizosphere	Bulk soil
N fixation	267.8728898	231.754579	270.0742233	266.9462808	254.7096695	244.9970575	248.1591998	248.1509195	278.5143788	249.2704556	259.0644038	224.6701453
P solubilization	525.4668511	421.7210059	534.4393081	530.4375281	532.1230671	502.1728681	553.0426694	488.3041181	545.1196202	468.2936365	581.5384728	622.7971817
Siderophore production	108.3057056	93.92279966	100.6326701	130.5012764	94.46120514	119.1837092	104.5912439	137.0158354	120.3663549	96.07910405	83.21318114	108.0875232
IAA production	559.1749055	414.5940388	527.1227135	481.0034529	478.4269583	436.4666336	524.0964307	462.368216	457.2386879	405.289537	497.4696901	461.1646737
ACC deaminase production	19.67825896	20.61585773	22.34349648	25.26394277	21.14718727	21.69171654	20.80845525	25.58326293	18.96274041	17.14269643	17.96467734	27.10851895

Supplementary Table 4 - Metagenome-assembled genomes (MAGs) taxonomically assigned to previous Additional Geographic Unit (AGU)-associated genera. Only MAGs with more than 50% completeness and less than 5% contamination were considered. For each MAG (column 1), the PhylophlAn assignment is reported in column 2, the respective 16S taxonomy is reported in column 3, completeness and contamination percentages are reported in column 4 and 5, and bp size is reported in column 6.

Genomic bin	PhyloPhlAn assignment	16S taxonomy	Completeness %	Contamination %	Size bp
bin.290	uSGB_61472:Other:k__Bacteria|p__Actinobacteria|c__CFGB38726|o__OFGB38726|f__FGB38726|g__GGB44089|s__GGB44089_SGB61472|t__SGB61472:0.247037	Actinobacteria Unclassified	54.47	2.91	3027411
bin.412	uSGB_61528:Other:k__Bacteria|p__Actinobacteria|c__CFGB35837|o__OFGB35837|f__FGB35837|g__GGB44133|s__GGB44133_SGB61528|t__SGB61528:0.250557	Actinobacteria Unclassified	53.09	3.02	1503685
bin.26	uSGB_55849:Other:k__Bacteria|p__Actinobacteria|c__CFGB10712|o__OFGB10712|f__FGB10712|g__GGB32443|s__GGB32443_SGB55849|t__SGB55849:0.265725	Actinobacteria Unclassified	76.83	3.51	2438761
bin.320	kSGB_35016:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Microbacteriaceae|g__Agromyces|s__Agromyces_ramosus|t__SGB35016:0.152843	Agromyces	57.4	4.10	2789788
bin.340	kSGB_31955:Species:k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Bradyrhizobiaceae|g__Bradyrhizobium|s__Bradyrhizobium_algeriense|t__SGB31955:0.26546	Bradyrhizobium	96.65	1.80	4623281
bin.202	uSGB_49283:Genus:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Cellulomonadaceae|g__Cellulomonas|s__Cellulomonas_SGB49283|t__SGB49283:0.312952	Cellulomonas	57.18	0.78	1649748
bin.1	uSGB_13269:Family:k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Solirubrobacterales|f__Conexibacteraceae|g__GGB8534|s__GGB8534_SGB13269|t__SGB13269:0.283502	Conexibacter	61.91	1.76	5199929
bin.176	uSGB_13270:Family:k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Solirubrobacterales|f__Conexibacteraceae|g__GGB8535|s__GGB8535_SGB13270|t__SGB13270:0.243761	Conexibacter	65.18	4.97	1715402
bin.348	uSGB_13270:Family:k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Solirubrobacterales|f__Conexibacteraceae|g__GGB8535|s__GGB8535_SGB13270|t__SGB13270:0.192258	Conexibacter	67.54	1.71	2165037
bin.396	kSGB_16435:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Cryptosporangiales|f__Cryptosporangiaceae|g__Cryptosporangium|s__Cryptosporangium_aurantiacum|t__SGB16435:0.276697	Cryptosporangium	52.12	0.85	1495776
bin.144	kSGB_33795:Species:k__Bacteria|p__Actinobacteria|c__Rubrobacteria|o__Gaiellales|f__Gaiellaceae|g__Gaiella|s__Gaiella_occulta|t__SGB33795:0.256381	Gaiella	73.59	4.78	1669372
bin.83	kSGB_33795:Species:k__Bacteria|p__Actinobacteria|c__Rubrobacteria|o__Gaiellales|f__Gaiellaceae|g__Gaiella|s__Gaiella_occulta|t__SGB33795:0.339586	Gaiella	69.63	3.03	2860656
bin.228	kSGB_12835:Species:k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Oxalobacteraceae|g__Massilia|s__Massilia_yuzhufengensis|t__SGB12835:0.35328	Massilia	54.81	2.81	2796737
bin.164	kSGB_32229:Species:k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhizobiales|f__Phyllobacteriaceae|g__Mesorhizobium|s__Mesorhizobium_sp_YM1C_6_2|t__SGB32229:0.212502	Mesorhizobium	59.82	3.92	5194556
bin.349	uSGB_61376:Genus:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Mycobacteriaceae|g__Mycobacterium|s__Mycobacterium_SGB61376|t__SGB61376:0.35328	Mycobacterium	98.54	0.97	2070293
bin.19	kSGB_853:Species:k__Archaea|p__Thaumarchaeota|c__Nitrososphaeria|o__Nitrososphaerales|f__Nitrososphaeraceae|g__Nitrososphaera|s__Candidatus_Nitrososphaera_gargensis|t__SGB853:0.320287	Archaea	64.29	2.76	2948997
bin.378	uSGB_55847:Family:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Corynebacteriales|f__Nocardiaceae|g__GGB38223|s__GGB38223_SGB55847|t__SGB55847:0.265526	Nocardiaceae Unclassified	56.04	3.61	1668073
bin.111	kSGB_34149:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Nocardioidaceae|g__Nocardioides|s__Nocardioides_iriomotensis|t__SGB34149:0.26546	Nocardioides	77.70	4.37	1248512
bin.197	kSGB_55752:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Nocardioidaceae|g__Nocardioides|s__Nocardioides_bacterium|t__SGB55752:0.300994	Nocardioides	51.66	2.59	1740489
bin.92	kSGB_15965:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Propionibacteriales|f__Nocardioidaceae|g__Nocardioides|s__Nocardioides_sp_Root122|t__SGB15965:0.254359	Nocardioides	57.42	4.42	2834555
bin.278	kSGB_11537:Species:k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Novosphingobium|s__Novosphingobium_sp_Rr_2_17|t__SGB11537:0.193331	Novosphingobium	57.43	4.67	2453724
bin.307	kSGB_16407:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Streptosporangiales|f__Thermomonosporaceae|g__Thermomonospora|s__Thermomonospora_curvata|t__SGB16407:0.328965	Novosphingobium/Thermomonospora	52.49	2.58	2643601
bin.321	uSGB_55819:Family:k__Bacteria|p__Cyanobacteria|c__Cyanobacteria_unclassified|o__Synechococcales|f__Prochloraceae|g__GGB41164|s__GGB41164_SGB55819|t__SGB55819:0.248764	Synechococcales Unclassified	83.49	2.91	2104713
bin.325	uSGB_61813:Other:k__Bacteria|p__Proteobacteria|c__CFGB35932|o__OFGB35932|f__FGB35932|g__GGB44394|s__GGB44394_SGB61813|t__SGB61813:0.273629	Proteobacteria Unclassified	62.1	4.73	2190116
bin.96	uSGB_61813:Other:k__Bacteria|p__Proteobacteria|c__CFGB35932|o__OFGB35932|f__FGB35932|g__GGB44394|s__GGB44394_SGB61813|t__SGB61813:0.26803	Proteobacteria Unclassified	62.22	1.03	1350259
bin.128	kSGB_67219:Species:k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodobacterales|f__Rhodobacteraceae|g__Rhodobacteraceae_unclassified|s__Rhodobacteraceae_bacterium|t__SGB67219:0.372583	Rhodobacter	52.21	2.04	1857802
bin.355	uSGB_52610:Genus:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Micrococcales|f__Microbacteriaceae|g__Rhodoluna|s__Rhodoluna_SGB52610|t__SGB52610:0.300994	Microbacteriaceae	74.22	3.53	2002286
bin.121	kSGB_17577:Species:k__Bacteria|p__Actinobacteria|c__Rubrobacteria|o__Rubrobacterales|f__Rubrobacteraceae|g__Rubrobacter|s__Rubrobacter_xylanophilus|t__SGB17577:0.300994	Rubrobacter	54.39	0.50	1851211
bin.36	kSGB_17577:Species:k__Bacteria|p__Actinobacteria|c__Rubrobacteria|o__Rubrobacterales|f__Rubrobacteraceae|g__Rubrobacter|s__Rubrobacter_xylanophilus|t__SGB17577:0.258493	Rubrobacter	60.04	3.45	3175285
bin.178	kSGB_55872:Species:k__Bacteria|p__Actinobacteria|c__Thermoleophilia|o__Solirubrobacterales|f__Solirubrobacteraceae|g__Solirubrobacter|s__Solirubrobacter_sp_CPCC_204708|t__SGB55872:0.295981	Solirubrobacter	74.11	1.94	1593893
bin.119	kSGB_24610:Species:k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae|g__Sphingomonas|s__Sphingomonas_sp_URHD0007|t__SGB24610:0.35328	Sphingomonas	66.50	3.23	2636012
bin.123	kSGB_10471:Species:k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Nevskiales|f__Sinobacteraceae|g__Steroidobacter|s__Steroidobacter_denitrificans|t__SGB10471:0.295981	Steroidobacter	78.95	3.41	2177673
bin.42	kSGB_10471:Species:k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Nevskiales|f__Sinobacteraceae|g__Steroidobacter|s__Steroidobacter_denitrificans|t__SGB10471:0.312952	Steroidobacter	74.23	4.68	3459230
bin.126	kSGB_16111:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Streptomycetales|f__Streptomycetaceae|g__Streptomyces|s__Streptomyces_sp_1114_5|t__SGB16111:0.270747	Streptomyces	68.45	3.45	3676162
bin.173	kSGB_34439:Species:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Streptomycetales|f__Streptomycetaceae|g__Streptomyces|s__Streptomyces_sp_TAA040|t__SGB34439:0.320287	Streptomyces	63.88	4.34	2875784
bin.43	uSGB_46364:Genus:k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Streptomycetales|f__Streptomycetaceae|g__Streptomyces|s__Streptomyces_SGB46364|t__SGB46364:0.35328	Streptomyces	78.38	3.39	2183585
bin.276	kSGB_12768:Species:k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Variovorax|s__Variovorax_sp_YR216|t__SGB12768:0.320287	Variovorax	66.22	4.88	2533478

Supplementary Table 5 - METABOLIC output of the reconstructed metagenome-assembled genomes (MAGs). Numerical values correspond to the gradient heatmap in Figure 8.

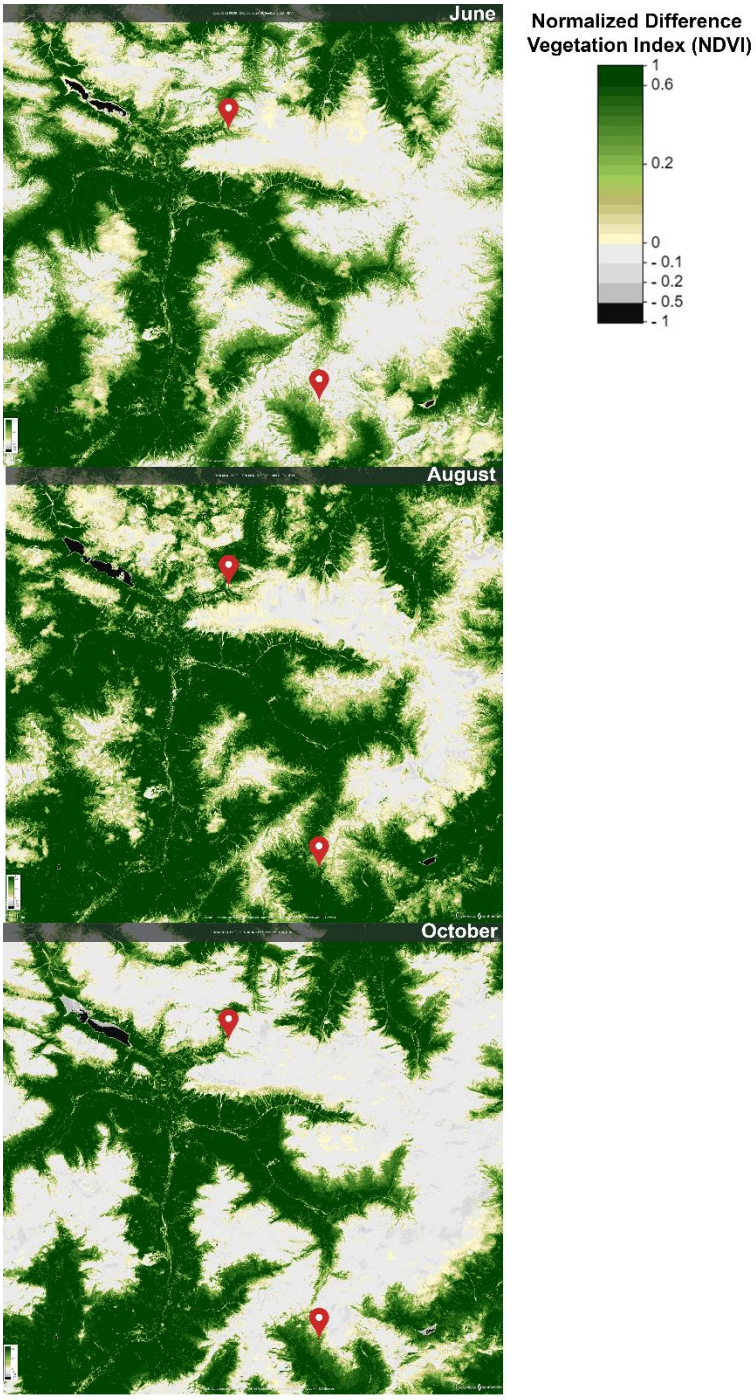
Function	bin.1	bin.111	bin.119	bin.121	bin.123	bin.126	bin.128	bin.144	bin.164	bin.173	bin.176	bin.178	bin.19	bin.197	bin.202	bin.228	bin.26	bin.276	bin.278
C-S-01:Organic carbon oxidation - CO oxidation	0	0	0	0	0	12.3	0	6.5	0	0	17.2	0	0	11.7	0.6	0	0	0	0
C-S-01:Organic carbon oxidation - amino acid utilization	0	0	0	0	2.4	10.7	0	0	0	0	14.9	0.5	3.2	10.1	0	0	15.6	0	1.9
C-S-01:Organic carbon oxidation - aromatics degradation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-01:Organic carbon oxidation - complex carbon degradation	0	0.7	0	2.1	3.7	16.9	0	0	1.7	0	23.6	0.9	0	16	0.8	0	0	0	3
C-S-01:Organic carbon oxidation - fatty acid degradation	0	0	0	1.6	0	12.5	0	6.6	0	0	17.4	0	0	0	0	0	18.2	0	0
C-S-01:Organic carbon oxidation - formaldehyde oxidation	0	0	0	0	0	43.5	0	23.1	0	0	0	0	0	0	0	0	0	0	0
C-S-01:Organic carbon oxidation - formate oxidation	0	0	0	4.4	0	35.4	0	0	0	0	0	0	0	33.6	0	0	0	0	0
C-S-01:Organic carbon oxidation - methanol oxidation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-03:Ethanol oxidation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-04:Acetate oxidation	0	0	0	0	0	0	0	16.9	0	0	44.5	0	0	30.2	0	0	0	0	0
C-S-06:Fermentation	0	0	0	0	0	15.7	0	8.4	1.6	0	22	0	0	15	0	0	0	0	0
C-S-08:Methanotrophy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-09:Hydrogen oxidation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-S-04:Nitrate reduction - narGH	0	0	0	0	0	0	0	0	0	0	58.5	0	0	39.7	0	0	0	0	0
N-S-05:Nitrite reduction - nirKS	0	0	0	0	0	0	0	0	6	0	83.7	0	0	0	0	0	0	0	0
N-S-06:Nitric oxide reduction	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N-S-08:Nitrite ammonification - nirBD	0	2.5	0	0	12.8	0	0	0	0	0	0	0	0	54.7	0	0	0	0	10.1
O-S-01:Iron reduction	0	0	0	0	0	13.8	0	7.3	0	0	19.3	0	0	13.1	0	0	0	0	0
O-S-02:Iron oxidation:	0	0	0	0	20.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S-S-03:Sulfur oxidation - sdo	0	0	0	0	0	0	0	0	0	0	25.4	0	5.4	0	0	0	0	0	0
S-S-04:Sulfite oxidation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32.9	0	0
S-S-05:Sulfate reduction	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32.9	0	0
S-S-07:Thiosulfate oxidation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Function	bin.290	bin.307	bin.320	bin.321	bin.325	bin.340	bin.348	bin.349	bin.355	bin.36	bin.378	bin.396	bin.412	bin.42	bin.43	bin.83	bin.92	bin.96
C-S-01:Organic carbon oxidation - CO oxidation	0	9.2	0	1.3	0	2.1	1.2	0	3.1	23.8	2.5	0	8.4	0	0	0	0	0
C-S-01:Organic carbon oxidation - amino acid utilization	5.5	8	1.3	0	2.8	1.8	1	0	2.6	0	0	0	7.3	0	9.9	0	0.4	0
C-S-01:Organic carbon oxidation - aromatics degradation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
C-S-01:Organic carbon oxidation - complex carbon degradation	0	12.7	2.1	0	0	0	0	0	4.2	0	0	0	11.6	0	0	0	0	0
C-S-01:Organic carbon oxidation - fatty acid degradation	6.4	0	1.5	0	0	2.2	0	0	0	24.1	0	0	8.5	0	0	0	0	1
C-S-01:Organic carbon oxidation - formaldehyde oxidation	0	0	0	0	0	0	0	0	0	0	0	0	29.8	0	0	0	0	3.6
C-S-01:Organic carbon oxidation - formate oxidation	0	26.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-01:Organic carbon oxidation - methanol oxidation	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-03:Ethanol oxidation	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
C-S-04:Acetate oxidation	0	0	0	0	8.3	0	0	0	0	0	0	0	0	0	0	0	0	0
C-S-06:Fermentation	8.1	0	2	0	4.1	0	0	0	3.9	0	3.2	0	0	1.6	14.6	0	0	0
C-S-08:Methanotrophy	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
C-S-09:Hydrogen oxidation	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
N-S-04:Nitrate reduction - narGH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.7	0
N-S-05:Nitrite reduction - nirKS	0	0	0	0	0	10.3	0	0	0	0	0	0	0	0	0	0	0	0
N-S-06:Nitric oxide reduction	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
N-S-08:Nitrite ammonification - nirBD	0	0	0	0	15.1	0	0	0	0	0	0	0	0	0	0	0	0	4.7
O-S-01:Iron reduction	0	10.4	0	0	0	0	0	0	0	26.7	0	0	9.4	0	0	0	0	0
O-S-02:Iron oxidation:	0	70.1	0	0	0	0	0	0	0	0	0	0	0	9.3	0	0	0	0
S-S-03:Sulfur oxidation - sdo	9.3	0	0	0	0	0	1.8	0	4.5	35.1	0	0	0	1.8	16.8	0	0	0
S-S-04:Sulfite oxidation	0	0	0	2.3	0	0	0	0	5.6	43.7	0	0	15.5	0	0	0	0	0
S-S-05:Sulfate reduction	0	0	0	2.3	0	0	0	0	5.6	43.7	0	0	15.5	0	0	0	0	0
S-S-07:Thiosulfate oxidation	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0

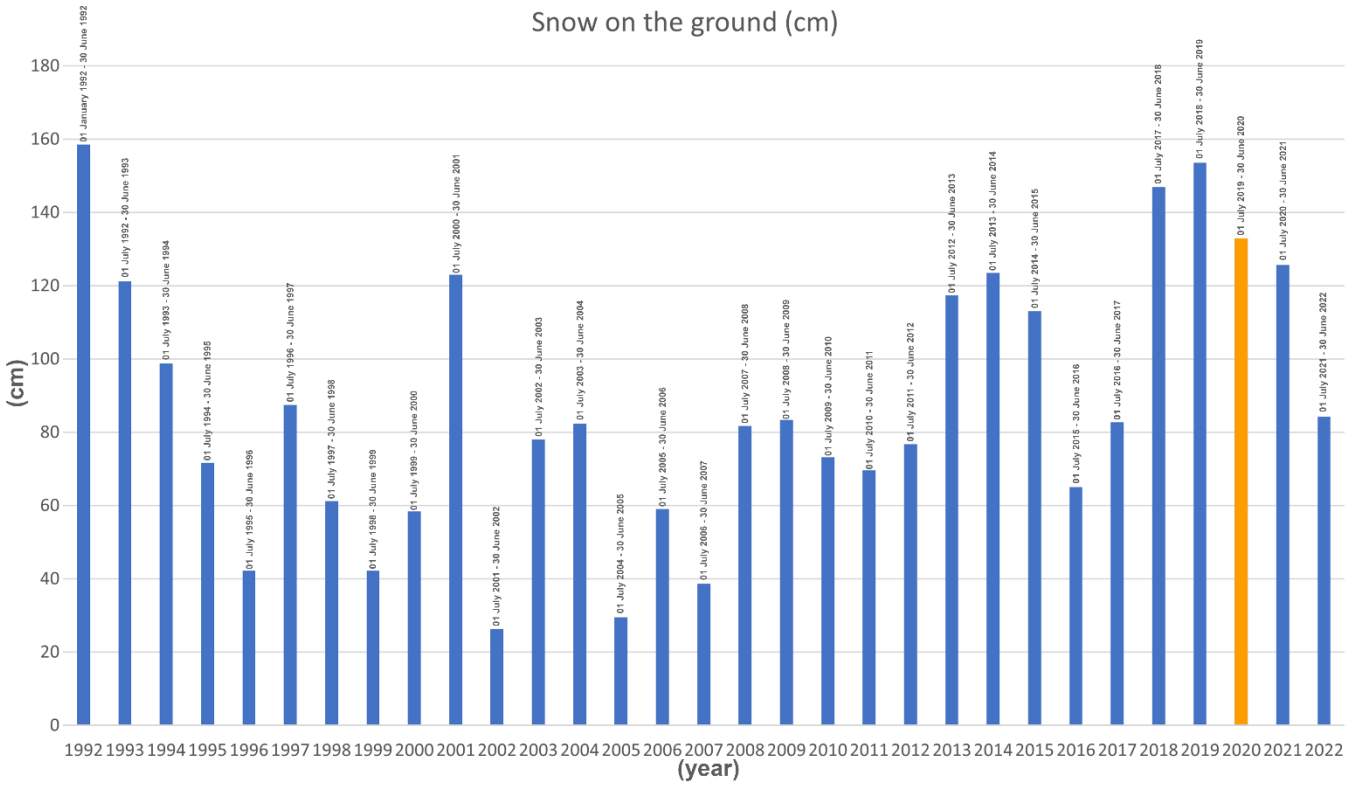
Supplementary Table 6 - Accession numbers of plant growth-promoting (PGP) proteins from the NCBI database. Protein names are given in the first column.

Protein ID	NCBI accession
NifB	WP_011024080
NifE	WP_014404757
NifH	WP_010870393
NifN	WP_011241560
NifV	WP_011241567
NifU	NP_461477
phoA	NP_414917
GDH	NP_388275
EntF	WP_000077784
EntS	NP_415123
FslA	WP_003037766
ipdC	WP_035671558
aro10	NP_010668
aldH	NP_001260290
AcdS	XP_037178185

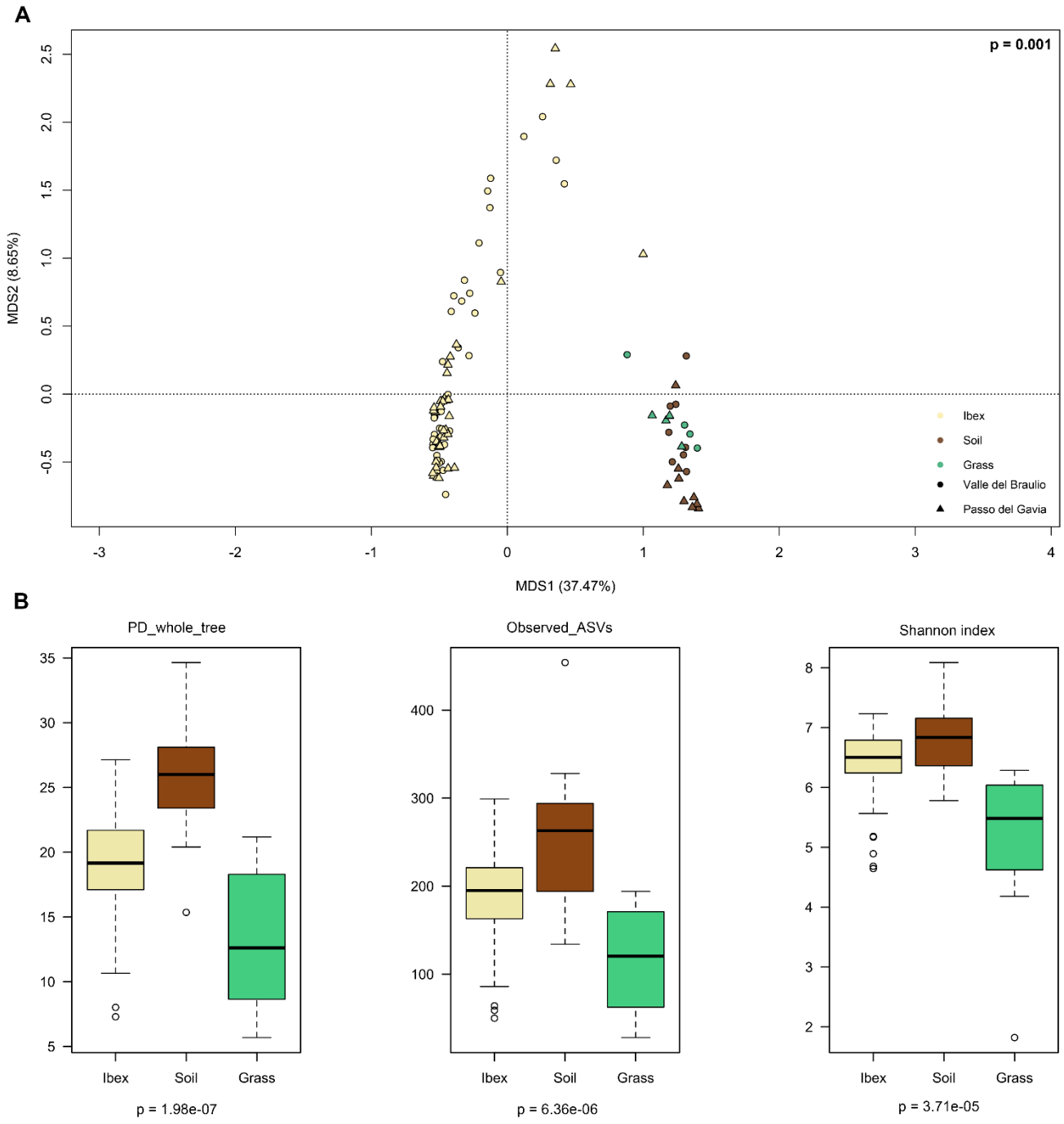
Study III



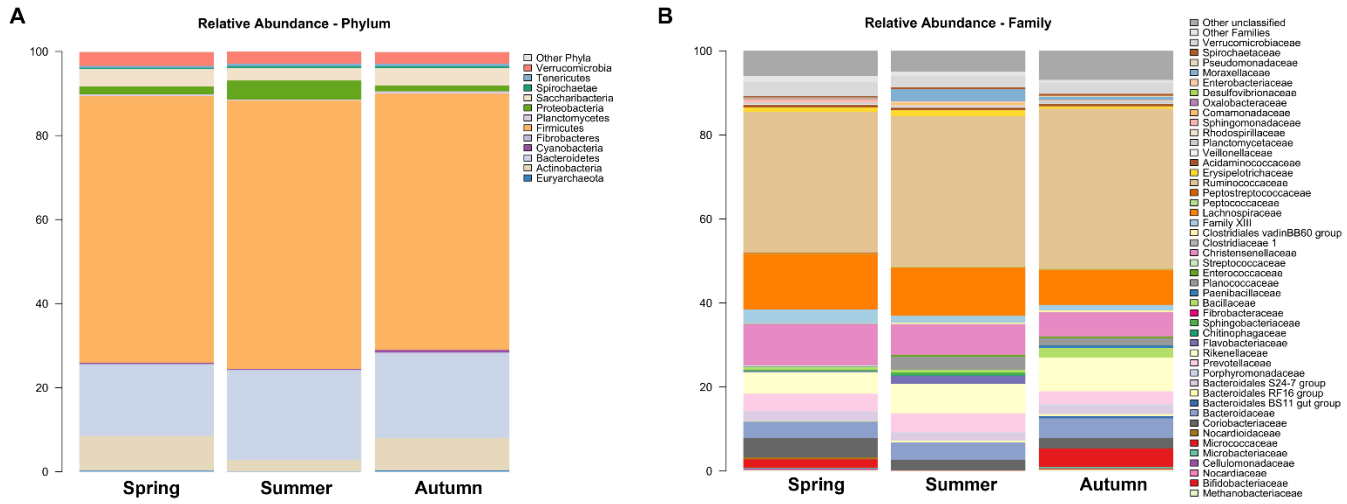
Supplementary Figure. S1 - Sampling sites. Map of the sampling sites at Stelvio National Park in Lombardia (Italy), showing the centroid of the sampling spots at "Passo del Gavia" and "Valle del Braulio". Sampling coordinates are marked in red. Satellite images from Sentinel-2 were retrieved for the three timepoints (i.e., June, August and October) (https://apps.sentinel-hub.com/eo-browser/). The days chosen for the three sampling periods were June 2nd, August 1st and October 10th, 2020, in order to have pictures with cloud cover < 20%. For each timepoint, the Normalized Difference Vegetation Index, NDVI, is shown (see color bar on the right).



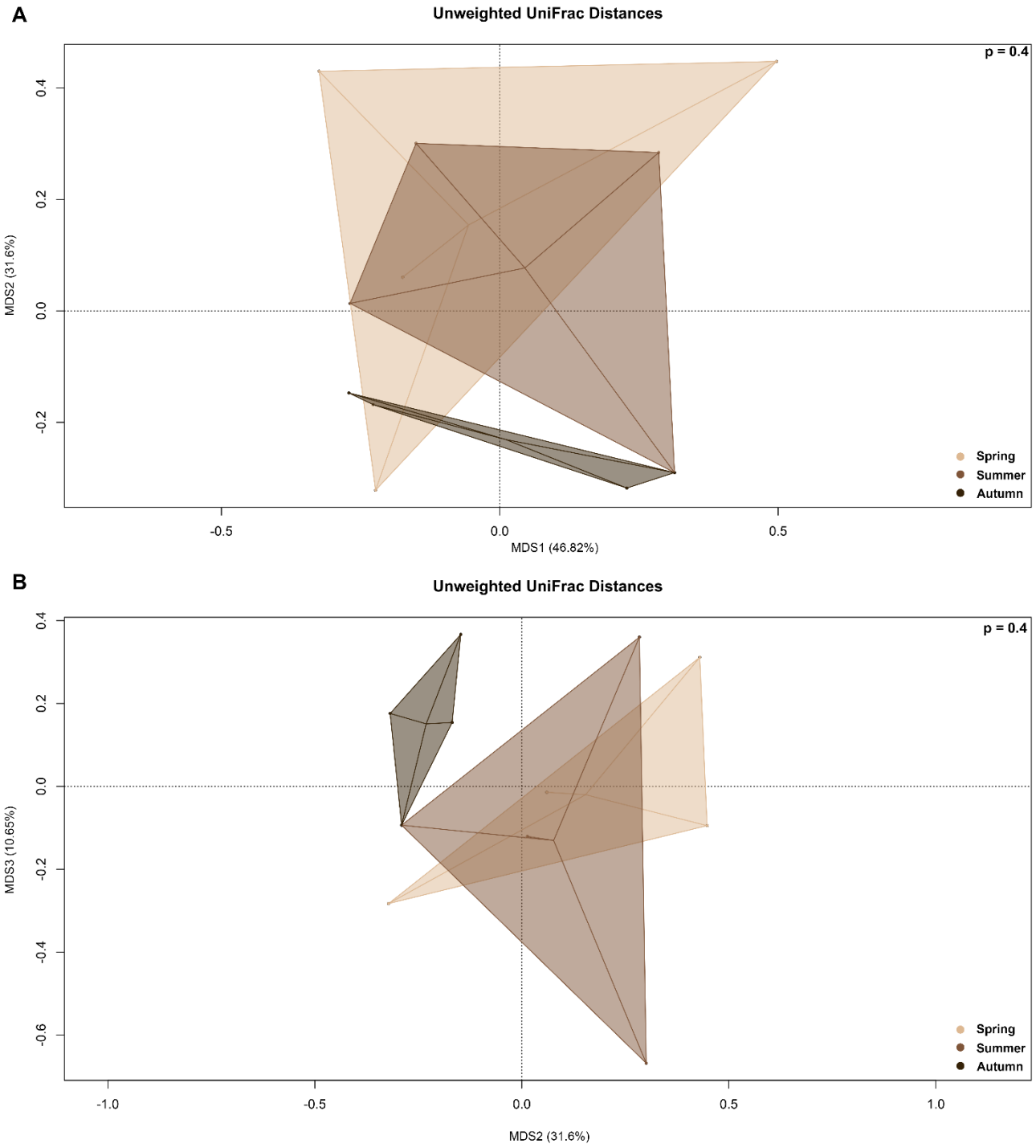
Supplementary Figure. S2 - Snow on the ground (cm) from 1992 to 2020. The mean value of snow on the ground (cm) from July 1st to June 30th of the next year (when possible) is shown. The bar regarding the period of our study is highlighted in orange.



Supplementary Figure. S3 - Comparison between the *C. ibex* gut microbiome and environmental microbiomes. **(A)** Principal Coordinates Analysis (PCoA) based on Bray-Curtis distances between the microbiome profiles of Alpine ibex gut (yellow), soil (brown) and grass (green) across the two sampling sites, i.e., "Passo del Gavia" (triangles) and "Valle del Braulio" (circles) (permutation test with pseudo-F ratio, p = 0.001). The first and second principal components (MDS1 and MDS2) are plotted and the percentage of variance in the dataset explained by each axis is shown. **(B)** Boxplots showing the alpha-diversity distributions of the *C. ibex* gut, soil and grass microbiomes, based on the Faith's Phylogenetic Diversity (PD whole tree), the number of observed ASVs and the Shannon index (Kruskal-Wallis test, $p \leq 0.001$).



Supplementary Figure. S4 - Taxonomic composition of the *C. ibex* gut microbiome. Barplots summarizing the phylum- **(A)** and family-level **(B)** bacterial composition of *C. ibex* feces in spring, summer and autumn. Only phyla and families with relative abundance > 0.5% in at least 2 samples are shown.



Supplementary Figure. S5 - Seasonal variation of SGB communities in the Alpine ibex gut microbiome. Principal Coordinates Analyses (PCoAs) based on unweighted UniFrac distances between Alpine ibex SGB gut microbial communities across seasons, i.e., spring (June, light brown), summer (August, brown) and autumn (October, dark brown) (permutation test with pseudo-F ratio, p = 0.4). The first and second principal components (MDS1 and MDS2) are plotted in **(A)** while the second and third principal components (MDS2 and MDS3) are plotted in **(B)** and the percentage of variance in the dataset explained by each axis is shown. Squared areas were drawn to enclose all samples.

Supplementary Table S1. Sampling details of the present study; number and type of samples collected

SITE	Number of samples collected								
	June 2020			August 2020			October 2020		
	Alpine ibex	Soil	Grass	Alpine ibex	Soil	Grass	Alpine ibex	Soil	Grass
Passo del Gavia	14	2	0	9	2	1	15	5	3
Valle del Braulio	12	1	0	17	2	1	19	5	3

Supplementary Table S2. SGBs within the bacterial communities in the 3 seasons interacting for the degradation of plant cell wall biopolymers. The right side highlights the bacterial genera shared between seasons

T1 Community					
MAGs	**Taxonomy**	**Shared_Genera**	Akkermansia	Alistipes	UBA4372
V04MT1_bin.1	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Oscillospiraceae;g__RGIG3566;s__				
G03MT1_bin.1	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__				
G03FT1_bin.1	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__				
G03MT1_bin.6	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Christensenellales;f__CAG-74;g__GCA-900199385;s__				
V04MT1_bin.4	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Alloprevotella;s__				
G03MT1_bin.10	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Acutalibacteraceae;g__CAG-488;s__				
G03MT1_bin.4	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__CAG-272;g__;s__				
V04MT1_bin.8	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__Alistipes sp015060115				
G03MT1_bin.8	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__				
V01MT1_bin.1	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__UBA4372;s__UBA4372 sp017622815				
V01MT1_bin.2	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__				
T2 Community					
MAGs	**Taxonomy**				
G01MT2_bin.3	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__UBA932;g__Cryptobacteroides;s__Cryptobacteroides sp017556765				
G06MT2_bin.3	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Ruminococcaceae;g__Ruminococcus;s__Ruminococcus sp017523285				
G01MT2_bin.1	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__Alistipes sp015060115				
G01MT2_bin.2	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Oscillospiraceae;g__Faecousia;s__				
G06MT2_bin.5	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__UMGS1783;g__;s__				
V16MT2_bin.2	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__HGM04593;s__HGM04593 sp017407905				
V09FT2_bin.3	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__HGM04593;s__				
G06MT2_bin.9	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Acutalibacteraceae;g__CAG-177;s__				
V16MT2_bin.5	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__				
G01MT2_bin.4	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__UBA4372;s__UBA4372 sp017622815				
G01MT2_bin.7	d__Bacteria;p__Actinobacteriota;c__Actinomycetia;o__Actinomycetales;f__Bifidobacteriaceae;g__RGIG1476;s__RGIG1476 sp017413255				
V09FT2_bin.1	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__JAAYXM01;g__SIG480;s__				
V09FT2_bin.6	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__Akkermansia sp017435365				
G06MT2_bin.2	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__				
T3 Community					
MAGs	**Taxonomy**				
V02T3_bin.6	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__Alistipes sp017621455				
V16T3_bin.1	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__HGM04593;s__HGM04593 sp017522915				
V16T3_bin.7	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__				
V02T3_bin.1	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__Alistipes sp017937765				
V16T3_bin.6	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__Akkermansia sp017477935				
V02T3_bin.5	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Ruminococcaceae;g__RGIG8367;s__				
G08FT3_bin.1	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__Akkermansia sp015061985				
G12FT3_bin.2	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Acutalibacteraceae;g__RUG420;s__				
V16T3_bin.2	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Alistipes;s__				
V16T3_bin.9	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__UBA4372;s__UBA4372 sp017622815				
G12FT3_bin.5	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__Tidjanibacter;s__				
V02T3_bin.4	d__Bacteria;p__Firmicutes_A;c__Clostridia;o__Oscillospirales;f__Oscillospiraceae;g__Faecousia;s__				
V16T3_bin.3	d__Bacteria;p__Verrucomicrobiota;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Akkermansiaceae;g__Akkermansia;s__Akkermansia sp017435365				