Recapping coffee production in Mozambique: assessment of plant and microbial genetic diversity

INOCÊNCIA DA PIEDADE ERNESTO TAPAÇA

SCIENTIFIC ADVISORS

Doctor Ana Isabel Faria Ribeiro

Doctor Isabel Cristina dos Santos Lourenço Marques

THESIS PRESENTED TO OBTAIN THE DOCTOR DEGREE (PhD) IN BIOLOGY

2025

Recapping coffee production in Mozambique: assessment of plant and microbial genetic diversity

INOCÊNCIA DA PIEDADE ERNESTO TAPAÇA

SCIENTIFIC ADVISORS

Doctor Ana Isabel Faria Ribeiro

Doctor Isabel Cristina dos Santos Lourenço Marques

THESIS PRESENTED TO OBTAIN THE DOCTOR DEGREE (PhD) IN BIOLOGY

President

Doctor Ricardo Manuel de Seixas Boavida Ferreira, Professor Catedrático do Instituto Superior de Agronomia da Universidade de Lisboa.

Members

Doctor Fábio Luíz Partelli, Professor Titular da Universidade Federal do Espírito Santo, Brasil;
Doctor Ivete Sandra Alberto Maquia, Investigadora Assistente da Universidade Eduardo Mondlane, Moçambique;
Doctor Maria Paula Abranches Alvarinhas Fareleira, Investigadora Auxiliar do Instituto Nacional de Investigação Agrária e Veterinária;
Doctor José Domingos Cochicho Ramalho, Investigador Auxiliar com Agregação do Instituto Superior de Agronomia da Universidade de Lisboa;
Doctor Ana Isabel Faria Ribeiro, Investigadora Auxiliar com Agregação do Instituto Superior de Agronomia da Universidade de Lisboa.

Funding: Camões, Instituto da Cooperação e da Língua (Projeto TRICAFÉ), and FCT – Fundação para a Ciência e Tecnologia, I.P. by project reference UIDB/00239/2020 of the Forest Research Centre and DOI identifier 10.54499/UIDB/00239/2020

Hinc
Patriam
Sustinet

INSTITUTO
SUPERIOR DE
AGRONOMIA
Universidade de Lisboa

U LISBOA

UNIVERSIDADE
DE LISBOA

Acknowledgments

First of all, I thank God for the gift of life, I know that without him in my life nothing would be possible.

I express my deep and sincere gratitude to my supervisors Professor Ana Isabel Ribeiro and Isabel Marques for their guidance, commitment, and tireless teaching of various techniques and scientific methodologies for the completion of this thesis. I would also like to thank them for their total availability in resolving the problems that arose throughout the work. I feel very proud to mention them as my mentors in professional training. I am grateful for their trust, friendship, advice, patience, and, above all for the opportunity to work with this team.

To Professor Ana Isabel Ribeiro, for crossing my path in 2017: The opportunity you gave me changed everything in my life and today opens another stage of achievements, my doctorate. Words fail to describe the wonderful person you are, besides being a supervisor you have been a true mother to me. God placed you in my life at the right moment, your teachings will remain and you will continue to be the mother, supervisor, and friend you have always been. May God bless you always.

To Dr. Isabel Marques, I express my deepest gratitude for all the support provided throughout this journey, which is now coming to an end. Your constant patience and availability were essential, providing me with scientific guidance and critical review from the beginning to the conclusion of this work. I am truly grateful for all your dedication and continuous encouragement throughout this journey. Thank you, thank you, and thank you.

To the Instituto Superior de Agronomia (ISA) for welcoming me as a student, especially to the Plant-Environment Interactions & Biodiversity lab and to all the professors in the Ferreira Lapa building, particularly on the third and top floors.

I would also like to express my gratitude to the Centro de Estudos Florestais (CEF) in Lisbon for their warm welcome and support throughout my doctorate. To all the professors, researchers, and colleagues, my sincere thanks for the invaluable guidance and sharing of knowledge.

I also want to extend my gratitude to Dr. Inês Graça for her trust, kindness, and valuable contribution to some laboratory work, as well as to Riccardo Corti and Samuele Pedrazzani. In a short time, we managed to get to know each other and turn our work team into a beautiful friendship. Thank you for crossing my path, for all your support, and for your contribution that allowed this work to become a reality.

Paula Batista Santos, you crossed my path in 2017 and stayed forever. You never stopped encouraging me, which strengthened our beautiful friendship, Thank you.

To all my colleagues from the lab meetings, especially Dr. Cochicho Ramalho, for being part of this work and contributing to this thesis. To Dr Teresa Vaz, for her availability encouragement and friendship.

To Dr. Maria Romeiras, for accepting my participation as a student in one of her classes and for her encouragement and help during the course.

To my father, may he rest in peace I wish so much that he were alive at this moment, but God wanted it this way. I dedicate this title to him, for I know he would be happy to receive it. His teachings were invaluable. A heartfelt thank you to my mother for being part of my story.

To my husband Dane António Martinho and my dear children, who felt my absence so much; today I affirm that it was worth it. I thank you from the bottom of my heart for every word of comfort. In every day of distress and uncertainty, you never stopped encouraging me. My loves, I love you so much.

To all my siblings, especially Katia Tapaça and Noémio Tapaça, who were involved in my struggle, directly or indirectly, thank you very much.

A special thanks to my friends from Mozambique: Julieta Americano, Janete Americano, Orneia Uaqueio, Lena Manhique, Ivete Maquia, Fátima Ismael, Salva Somueque, Crimildo Cassamo, Alberto Charrua, Alberto Guilengue, Yara Chiluva, Obete Madacussengua, Sérgio Ernesto, and Camilo Bruno, for the companionship, joy, and moments of relaxation and friendship created over the years. They never failed to offer words of encouragement and constant support in times of greater uncertainty.

Finally, I would like to thank the Portuguese Cooperation (Camões, Instituto da Cooperação e da Língua IP, and the Fundação para a Ciência e Tecnologia) for funding the scholarship for the project Sustainable Coffee Production in the Gorongosa National Park in an Integrated Agroforestry System in the context of deforestation, climate change, and food security, especially to Professor Ana Isabel Ribeiro-Barros as the project coordinator for the opportunity and funding to pursue a doctorate in Biology through the research grant. I also thank the Gorongosa National Park (PNG) for their constant availability in providing materials and transport for collecting coffee samples in the Gorongosa Mountain in Sofala, Mozambique.

Dedicatory

I dedicate this work to my father Marcos Lenadi Tapaça (in memory).

To my husband Dane António Martinho and my children: Tamalelane Dane Martinho, Malakhay Inocência Martinho, and Shadday Dane Martinho.

Resumo

O café está entre as culturas tropicais mais importantes, tendo a sua cadeia de valor um impacto global. O cultivo e o comércio de café são dominados por duas espécies principais: *Coffea arabica* (café Arábica) e *Coffea canephora* (café Robusta). Moçambique é considerado um hotspot de diversidade de espécies silvestres de *Coffea*. Este estudo teve como objetivos avaliar a diversidade genética de *Coffea* spp. em Moçambique e analisar a flora microbiana associada ao cafeeiro Arábica (*C. arabica*) cultivado em sistema agroflorestal no Parque Nacional da Gorongosa. A análise molecular identificou apenas a ocorrência de *Coffea racemosa*, sugerindo erosão genética causada pela pressão antropogénica. Apesar disso, *C. racemosa* apresentou uma diversidade genética significativa. Nas plantações da Gorongosa, o *pool* genético de *C. arabica* revelou-se reduzido, reforçando a necessidade urgente de diversificação dos cultivares. Foram também estudados os microbiomas do solo, folhas e frutos ao longo de um gradiente de altitude e irradiância. Os consórcios microbianos mostraram um padrão complexo influenciado apenas pela elevação. Entre os microrganismos, destacam-se os promotores do crescimento vegetal, produtores de antibióticos e biorremediadores, bem como alguns patógenos vegetais. Este trabalho reforça a importância da conservação de espécies silvestres de *Coffea* e da diversificação genética em práticas de cultivo adaptadas às mudanças climáticas, evidenciando a relevância do microbioma na saúde do solo e na qualidade dos frutos. Os resultados fornecem subsídios importantes para estratégias de sustentabilidade e maneio de culturas em regiões tropicais.

Palavras-chave: cafeeiro, diversidade genética, Gorongosa, microbioma, sustentabilidade.

Abstract

Coffee is among the most important tropical crops, with its value chain having a global impact. Coffee cultivation and trade are dominated by two main species: *Coffea arabica* (Arabica coffee) and *Coffea canephora* (Robusta coffee). Mozambique is considered a hotspot of diversity for wild *Coffea* species. This study aimed to assess the genetic diversity of *Coffea* spp. in Mozambique and analyze the microbial flora associated with *C. arabica* cultivated under agroforestry in Gorongosa National Park. Molecular analysis identified only the occurrence of the wild species *Coffea racemosa*, suggesting genetic erosion caused by anthropogenic pressure. Despite this, *C. racemosa* presented significant genetic diversity. In the Gorongosa plantations, the genetic pool of *C. arabica* was limited, reinforcing the urgent need for cultivar diversification. The study also examined the microbiomes of coffee soils, leaves, and fruits along a gradient of altitude and irradiance. The microbial consortia showed a complex pattern influenced only by elevation. Among the microorganisms, plant growth promoters, antibiotic producers, bioremediators, and some plant pathogens stood out. This work emphasizes the importance of conserving wild *Coffea* species and promoting genetic diversification in cultivation practices adapted to climate change, highlighting the relevance of the microbiome for soil health and bean quality. The results provide important insights for sustainable strategies and coffee crop management in tropical regions.

Keywords: coffee, genetic diversity, Gorongosa, microbiome, sustainability.

Resumo Alargado

O género *Coffea* pertence à família botânica Rubiaceae engloba e 130 espécies. Entre estas, apenas duas possuem relevância comercial: *Coffea arabica* L. (café Arábica) e *Coffea canephora* Pierre ex A. Froehner (café Robusta), correspondendo a aproximadamente 57% e 43% da produção anual, respetivamente. O café figura entre as *commodities* mais importantes a nível global, sendo crucial para a subsistência de cerca de 25 milhões de pequenos agricultores e envolvendo mais de 100 milhões de pessoas ao longo de toda a cadeia de valor, principalmente em países tropicais.

As mudanças climáticas representam o maior desafio para o setor do café. Sendo o aumento da temperatura, os períodos prolongados de seca e os episódios de chuva intensa os fatores de maior risco. Projeções de vários modelos climáticos sugerem que *C. arabica* enfrentará uma drástica redução nas atuais regiões produtoras de café até 2050, com perdas que podem chegar até aos 50% da área atual de cultivo. Embora mais resiliente, *C. canephora* também poderá ser afetada, principalmente na África Ocidental. Considerando o cenário ambiental e as pressões antropogénicas globais, bem como a reduzida diversidade genética dos cultivares comerciais de *C. arabica*, várias abordagens estão a ser implementadas para garantir a sustentabilidade desta importante cultura. Entre as estratégias promissoras, destaca-se a introdução de espécies silvestres na cadeia de valor. *Coffea racemosa* Lour., *Coffea zanguebariae* Lour. e *Coffea liberica* W.Bull ex Hiern foram identificadas como alternativas apropriadas para os programas de melhoramento genético. Estas espécies apresentam relevante tolerância a stresses bióticos e abióticos, juntamente com períodos de maturação mais curtos e atributos únicos de sabor e aroma, representando um excelente reservatório genético para adaptação e resiliência ao ambiente, bem como um importante recurso para a criação de novos *blends* de café.

Os sistemas agroflorestais (SAF) constituem outra abordagem indispensável (e mais imediata) para promover a sustentabilidade do cultivo de café. Nestes sistemas, as árvores de sombra desempenham um papel fundamental ao proteger as plantas de café da luz solar direta, além de melhorar o microclima (temperatura e humidade do ar), com a consequente redução da evapotranspiração. Quando implantados em ecossistemas florestais degradados, os SAF também têm o potencial de contribuir

para a restauração da biodiversidade e promover o sequestro de carbono, tornando-se uma estratégia eficaz tanto para a produção agrícola, quanto para a regeneração ambiental.

Outro fator importante na evolução e adaptação do café é o seu microbioma, constituído por uma complexa comunidade de microrganismos que habitam a superfície da planta (epifíticos), os tecidos internos dos diferentes órgãos (endofíticos) e o solo ao redor das raízes. Esta comunidade inclui arqueias, bactérias e fungos, influenciando coletivamente o crescimento e a adaptação das plantas. As ameaças à sustentabilidade do cultivo de café motivaram investigações recentes sobre a influência e o potencial desses microrganismos na melhoria do desempenho da cultura e da qualidade dos grãos. Além disso, o uso de microrganismos que podem atuar como agentes de biocontrolo é visto como uma alternativa viável e promissora ao uso de pesticidas. Por outro lado, o uso de microrganismos fixadores de azoto, solubilizadores de fósforo e produtores de sideróforos representa uma estratégia sustentável para reduzir a dependência de fertilizantes químicos. Assim, o uso de consórcios de microrganismos promotores do crescimento, pode potencialmente reduzir os custos financeiros e ecológicos, melhorando o desenvolvimento das plantas e mitigando o impacto dos stresses ambientais. Além disso, o microbioma dos grãos de café desempenha um papel essencial no controle dos processos de fermentação natural, prevenindo a transferência de outros grupos microbianos oriundos das folhas, da superfície dos frutos, de frutos muito maduros e do solo, evitando aromas indesejados no produto final.

Moçambique possui um número considerável de espécies silvestres de café. No entanto, o café nunca foi incluído no conjunto de culturas industriais estratégicas para Moçambique, por questões que se prendem com a extensão da área apropriada ao cultivo das espécies comerciais. Nos últimos anos, vários esforços para promover o setor do café em Moçambique têm vindo a ser realizados. Como resultado, foi restaurada a plantação de *C. zanguebariae* na ilha do Ibo, na província de Cabo Delgado, e implementadas plantações de *C. arabica* em Sanga e Mahua (Niassa), Alto Molocue e Gurue (Zambézia), Tsangano e Angonia (Tete), Chimanimani (Manica) e Gorongosa (Sofala).

O Parque Nacional da Gorongosa (PNG) estabeleceu as suas primeiras plantações de café em 2013, com o objetivo de restaurar a floresta tropical da Serra da Gorongosa através da implementação de um sistema agroflorestal que integra árvores nativas e café (SAF-café). Esta iniciativa centra-se na no uso sustentável dos recursos naturais, baseada na conservação da biodiversidade, na promoção da saúde humana e no desenvolvimento comunitário. É importante notar que a taxa de desflorestação na região tem sido alarmante, com uma perda florestal de cerca de 40% em 2019. Tendo em conta os desafios globais para o setor do café, e sob a égide do Projeto Tricafé (www.tricafe.org), este estudo teve como objetivo contribuir para o Projeto de Restauração da Gorongosa, preenchendo algumas lacunas de pesquisa e complementando a ciência do café. Nesse contexto, foram formuladas três questões relacionadas com a biodiversidade: (i) qual o estado atual do património genético de *Coffea* em Moçambique e seu respetivo potencial para suportar pressões climáticas?; (ii) o *pool* genético de *C. arabica* implantado na Gorongosa é suficiente para garantir a sustentabilidade dos SAF-café?; (iii) qual a importância da altitude e da sombra na diversidade microbiana do café e funções associadas? A resposta a essas questões constitui uma linha de base para a integração de espécies nativas na cadeia de valor do café, e para maximizar o uso de espécies microbianas nativas no desenvolvimento de soluções de base biológica, como biorremediação, biofertilização ou fermentação. Neste contexto, foram definidos os seguintes objetivos:

I. Efetuar o levantamento e a caracterização da diversidade genética de *Coffea* no Sul (Maputo, Gaza e Inhambane) e no Centro (Sofala) de Moçambique por meio de marcadores de DNA (cloroplastidiais, microssatélites - SSR e polimorfismos de nucleotído único - SNP).

II. Analisar a diversidade e as funções da flora microbiana associada ao café Arábica (solo, folhas e frutos) cultivado no SAF da Gorongosa ao longo de um gradiente de altitude e sombra.

A tese está organizada em cinco capítulos. No primeiro capítulo é apresentada uma introdução geral ao café, de uma perspetiva global ao contexto local, bem como os objetivos da tese. O Capítulo II centra-se na caracterização genética e genómica de *C. arabica* e de *C. racemosa*, integrando estudos de taxonomia, diversidade,

estrutura populacional e atributos genéticos funcionais. Neste trabalho, avaliou-se a ocorrência de espécies de *Coffea* em Moçambique, por comparação com os dados bibliográficos e de herbário. A análise de taxonomia molecular, baseada em marcadores cloroplastídiais, identificou apenas uma espécie silvestre, *C. racemosa*, com elevado grau de separação genética em relação a *C. arabica*. O estudo suporta a hipótese de erosão genética, causada pela enorme pressão antropogénica e ambiental na região. Apesar desta pressão, a diversidade genética intra e inter-polucional em *C. racemosa* é considerável e a estrutura geográfica bem delineada. Por outro lado, o *pool* genético de *C. arabica* nas plantações da Gorongosa revelou-se reduzido, salientando a necessidade urgente de diversificação de cultivares no SAF-café. Adicionalmente, a análise das vias funcionais com base em SNPs sugere que as vias de sinalização de stresses são mais robustas em *C. racemosa*. Esta abordagem inovadora reforça a importância da implementação de programas de conservação de espécies silvestres de *Coffea*, da introdução destas espécies na cadeia de valor do café, bem como do aumento da diversidade genómica em práticas adaptadas às mudanças climáticas.

Os Capítulos III e IV são dedicados ao estudo do microbioma do solo e de folhas e frutos, respetivamente, ao longo de um gradiente de irradiância (pleno sol, sombra moderada, sombra intensa) e altitude (600, 800 e 900 m). Relativamente ao solo (Capítulo III), foi efetuada uma análise integrada, incluindo as propriedades físico-químicas, atividades enzimáticas e diversidade taxonómica. O fósforo disponível, o carbono orgânico e os nitratos foram significativamente mais elevados a 800 m, provavelmente devido às maiores atividades enzimáticas de β-glucosidase e fosfatases. A biomassa microbiana (carbono e nitrogénio) e a humidade foram significativamente maiores a 600 e 900 m, o que pode ser atribuído à abundância e riqueza de microrganismos do solo. De facto, a análise do metagenoma revelou um padrão complexo de consórcios microbianos (bactérias, arqueias, fungos), influenciado pela altitude (e sem efeito significativo da irradiância), com uma grande diversidade de bactérias, cuja análise funcional revelou adaptações metabólicas distintas nas diferentes altitudes, a saber: (i) estímulos ambientais e interações microbianas (900m); (ii) degradação da matéria orgânica, fermentação e biorremediação (800 m); (iii) decomposição de compostos derivados de plantas (600 m). No geral, os resultados indicam uma maior eficácia do consórcio microbiano a 800 m. Relativamente aos frutos e folhas (Capítulo IV), a dinâmica das comunidades

microbianas (bactérias e fungos) foi igualmente influenciada apenas pela altitude. Para ambos os grupos microbianos, observou-se um grau considerável de variação taxonómica entre frutos e folhas e ao longo do gradiente de altitude, mais proeminente para bactérias. Os táxons bacterianos dominantes incluíram: (i) promotores do crescimento de plantas, muitos dos quais produzem hormonas de crescimento (como auxinas e giberelinas), solubilizam fósforo, fixam azoto e melhoram a absorção de nutrientes; (ii) produtores de antibióticos, sideróforos ou enzimas que inibem o crescimento de agentes patogénicos das plantas; (iii) biorremediadores, com capacidade de degradação de compostos orgânicos nocivos ou poluentes; e (iv) patógenos vegetais. O conjunto de bactérias está associado a 74 vias metabólicas metacíclicas diferencialmente abundantes ao longo do gradiente de altitude. Relativamente aos fungos, a comunidade é dominada por saprotóficos. Estes fungos podem desempenhar um papel significativo na decomposição da matéria orgânica e a sua presença pode ter efeitos benéficos (*e.g.* fermentação) ou prejudiciais (patogenecidade). O crescimento excessivo destes organismos em grãos de café pode ainda levar à deterioração do fruto, com acumulação de sabores indesejados, e em última análise, contaminação por micotoxinas.

No último capítulo, Capítulo V, são apresentadas as conclusões gerais e considerações finais deste estudo. Tendo em conta o cenário global, a implementação de programas de conservação e gestão de espécies nativas de *Coffea*, bem como a sua inclusão na cadeia de valor do café (melhoramento genético ou aproveitamento do grão para novos *blends*) é imprescíndivel. Por outro lado, a diversidade microbiana do solo e da endosfera de *C. arabica*, revela-se bastante rica e diversificada, dentro do conceito de *terroir*, que inclui altitude, solo, clima e práticas agrícolas. Em conjunto, a análise do microbioma sugere que a variação dos táxons microbianos ao longo do gradiente de altitude é essencial para o desempenho das plantas, incluindo as características físico-químicas dos grãos de café. Qualquer uma das comunidades microbianas, possui potencial para o desenvolvimento de soluções de base biológica, como biorremediação, biofertilização ou fermentação. A presença de algumas espécies patogénicas, sem manifestação aparente de sintomas nas plantas, poderá dever-se à tolerância do cultivar em estudo e/ou ao efeito antagonista de outros microrganismos. Estes resultados ressaltam a importância de estudar e preservar a interação entre plantas e microrganismos, tanto para a sustentabilidade agrícola quanto para o aproveitamento de recursos biológicos na produção de café.

Abbreviations

ITS - Internal Transcribed Spacer

LISC - Herbarium of the Tropical Research Institute, University of Lisbon

PNG - Gorongosa National Park

UEM - Eduardo Mondlane University

UFES - Federal University of Espírito Santo

NCBI - National Center for Biotechnology Information

BLAST - Basic Local Alignment Search Tool

MatK - Maturase K

RbcL - Large subunit of the gene encoding RuBisCO

RuBisCO - Ribulose 1,5-bisphosphate carboxylase/oxygenase

ITS - Internal Transcribed Spacer

16S rRNA- Ribosomal RNA molecule

SNP - Single Nucleotide Polymorphisms

SSR - Simple Sequence Repeats

DNA - Deoxyribonucleic Acid

PCR - Polymerase Chain Reaction

dNTPs - Deoxynucleotide Triphosphates

$MgCl_2$ - Magnesium Chloride

Taq - Thermus aquaticus

List of Figures

List of Tables

Supplementary Information

List of Contents

Chapter 1

General Introduction

1.1. Origin, distribution and socioeconomic importance of *Coffea*

The genus *Coffea* belongs to the botanical family Rubiaceae, tribe *Coffea*, and has two recognized subgenera: *Coffea* and *Baracoffea* (Davis *et al*., 2006; Guerreiro-Filho *et al*., 2008). To date, 130 *Coffea* species have been identified (Davis & Rakotonasolo, 2021). Comparative studies based on molecular analyzes have demonstrated a strong correspondence between their phylogenetic origin and geographic distribution, clustering *Coffea* spp. in four large phylogenetic groups, West Africa, Central Africa, East Africa, and Madagascar (Figure 1.1) (Lashermes *et al*., 1997; Souza *et al*., 2004; Razafinarivo *et al*., 2013; Davis & Rakotonasolo, 2021).These are distributed in three sections: *Mascarocoffea*, predominantly from Madagascar and the Mascarene Islands; *Mozambicoffea*, native to East Africa; and *Eucoffea* originating from Central and West Africa (Davis *et al*., 2011).

Two coffee species support the value chain: *Coffea arabica* L. (Arabica coffee) and *Coffea canephora* Pierre ex A Froehner (Robusta coffee), which account respectively for *ca.* 57% and 43% of the annual production (ICO, 2023). The first is endemic to Ethiopia, Sudan, and Kenya, and the second to Zaire, Uganda, Sudan, Tanzania, and the Congo basin (Davis *et al*., 2006; Milard, 2017). *C. arabica*, is a tetraploid species (2n=4x=44) resulting from the natural hybridization between *C. eugenioides* Moore and *C. canephora* (Charrier & Berthaud, 1985; Lashermes *et al*., 1999; Bawin *et al*., 2021). This species reproduces mainly by self-fertilization, with 90% of its flowers fertilized by pollen from the same flower (Sakiyama *et al*., 1999). In contrast, *C. canephora* is diploid (2n=2x=22) and a predominantly allogamous species (Charrier *et al*., 1985; Fonseca *et al*., 2007; Teixeira *et al*., 2011).

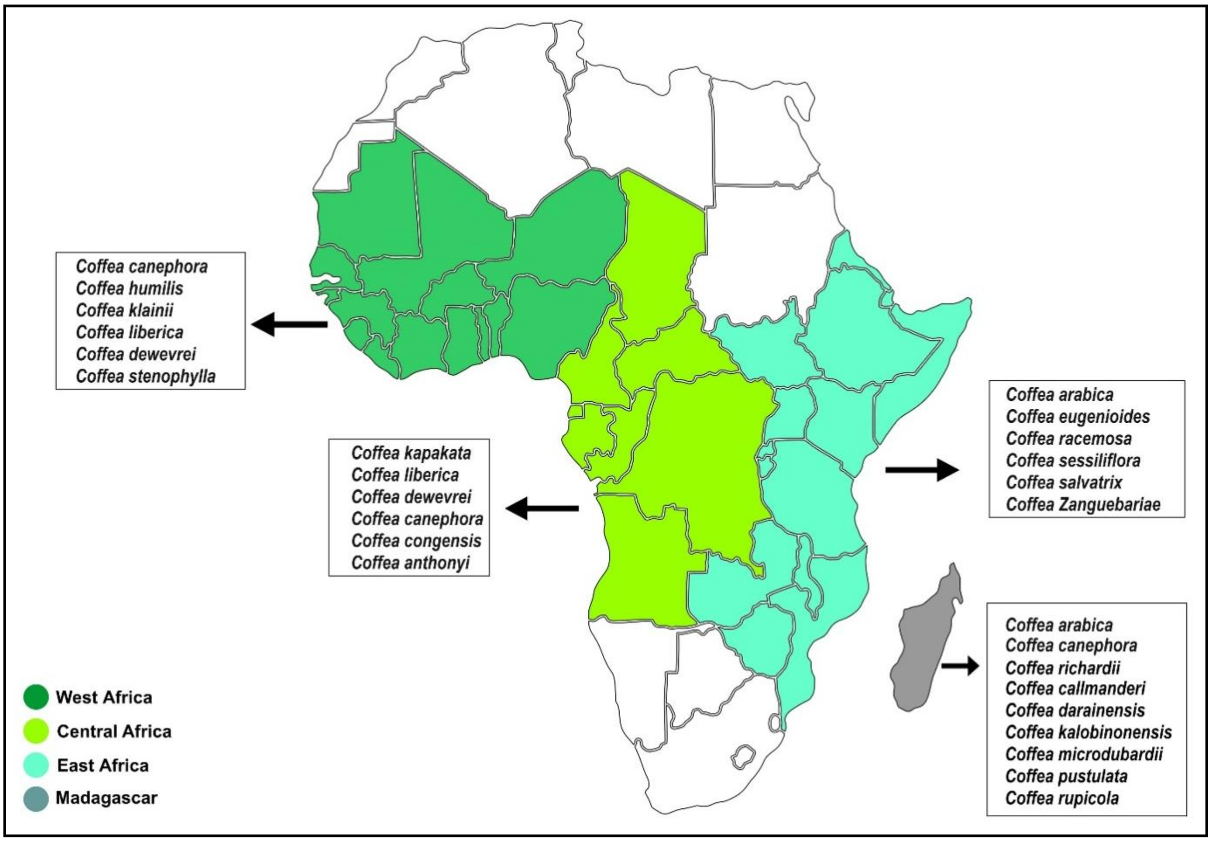


Figure 1.1 Natural distribution of some of the most recognized *Coffea* species in Africa.

Globally, coffee production is crucial to the livelihoods of *ca.* 25 million smallholder farmers, involving approximately 100-125 million people across the value chain, mostly in tropical countries (Voora *et al*., 2019; Figure 1.2). It is among the most traded commodities in the world, in a matrix that includes *e.g.* oil, natural gas, gold, cocoa, and sugar. In 2022/2023, the top five coffee producers (Brazil, Vietnam, Indonesia, Colombia, and Ethiopia) contributed to *ca.* 74 % of the total production. India, Honduras, Uganda, Mexico, and Peru completed the list of the top 10 coffee-producing countries, all together accounting for *ca.* 90% of the global production (ICO, 2023). In terms of regions, South America is the largest exporter (43%), followed by Asia and Oceania (34%), Central America and Mexico (12%), and Africa (11%). Despite the climate pressure, the global coffee production is projected to increase from 168 (2022/23) to 178 million 60-kg bags (2003/2024), nearly accompanying the consumption needs (ICO, 2023; Figure 1.3).

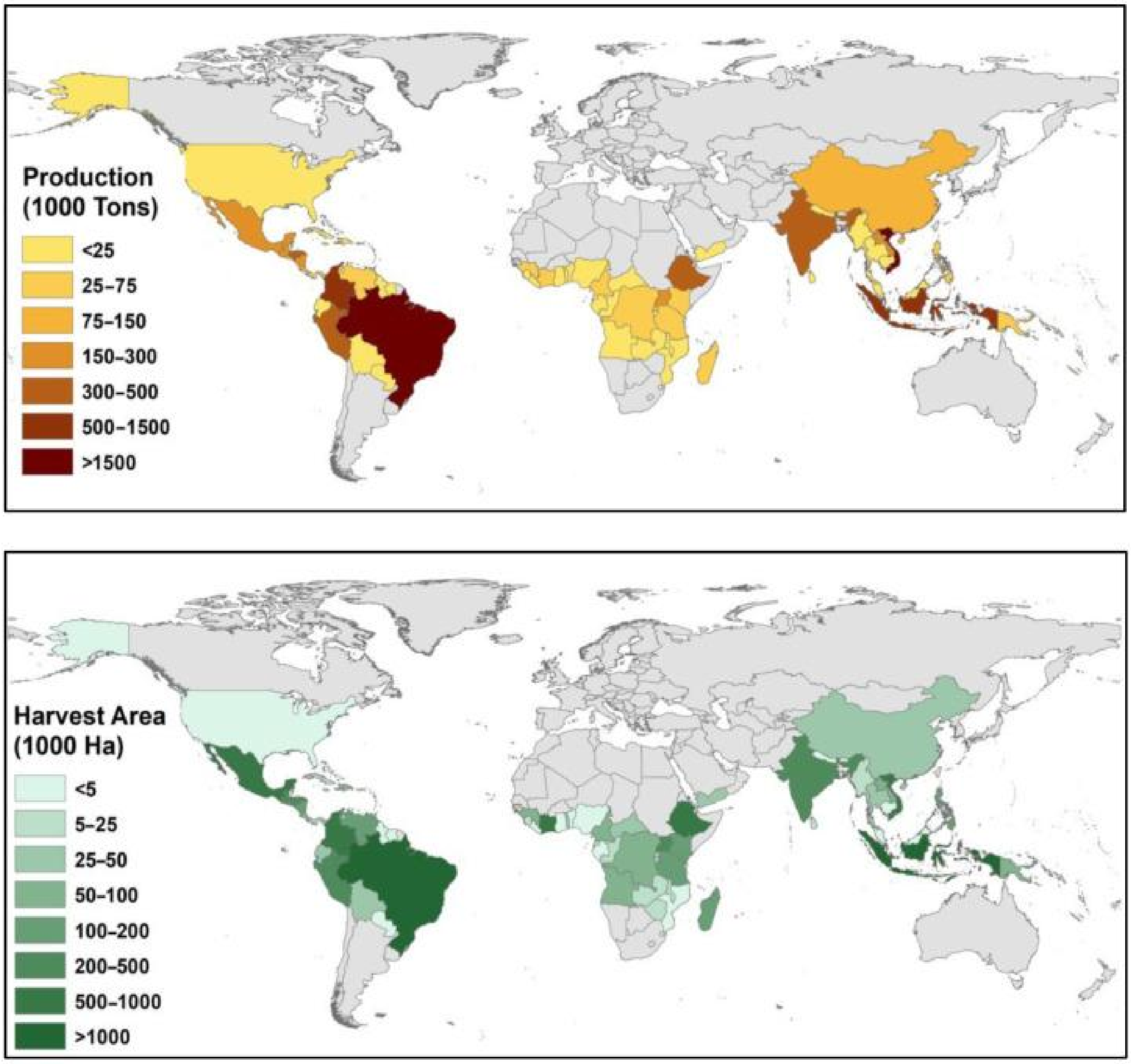


Figure 1.2 World coffee production (top) and harvested area (bottom) in 2019 (Bilen *et al.*, 2023).

However, extreme weather scenarios will impose particularly negative impacts in some countries, likely Indonesia, Côte d'Ivoire, and Uganda (ICO, 2023). In fact, the present operational environment is hampered by several factors including, climate change events, pests and diseases, labor shortages, rising prices of inputs, slow pace of breeding programs, institutional framework constraints, price volatility, and fair-trade concerns (Ngure & Watanabe, 2024).

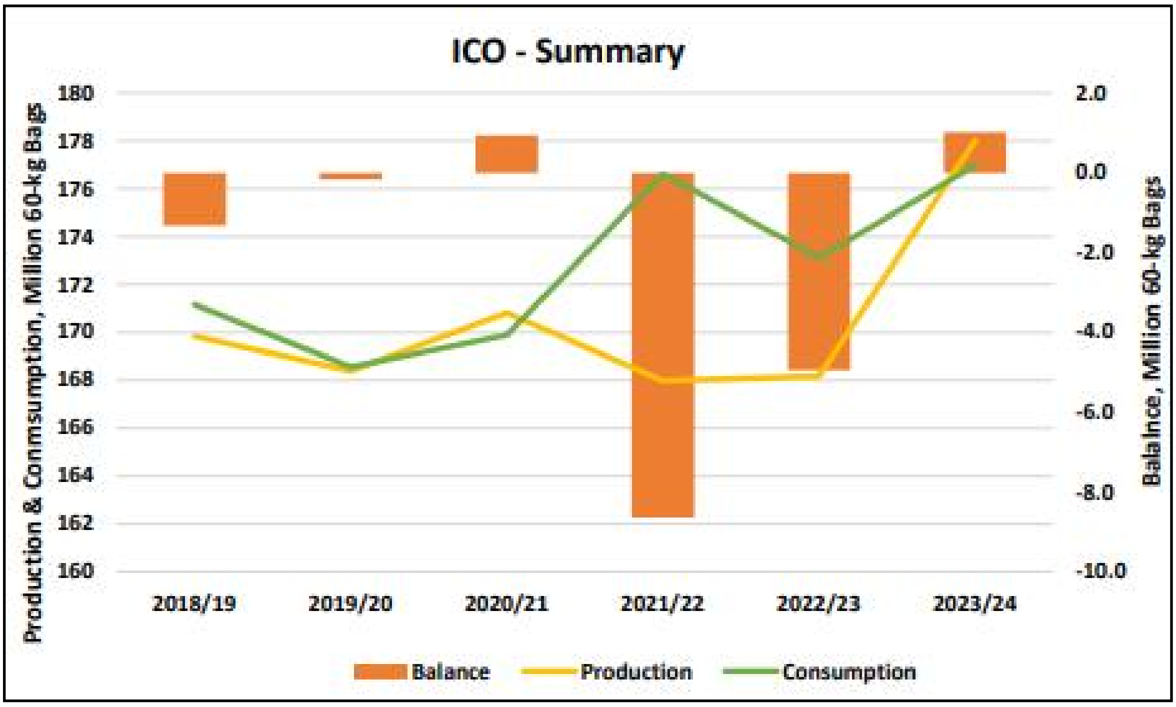


Figure 1.3 Summary of the World Coffee Market – Million 60-Kg Bags (ICO, 2023).

1.2. Sustainability of coffee production

Climate change is among the most critical issues for the coffee sector (Grüter *et al.*, 2022). High temperatures, long dry periods, and intense rainfall events may contribute to higher severity of pests and diseases incidence, as they favor the development and spread of pathogenic agents, contributing to a significant decrease in productivity and to a shift of production areas (Bunn *et al.*, 2015; Magrach & Ghazoul, 2015). This is likely due to the fact that the time required for species migration and adaptation lags behind the rate of change (Bunn *et al.*, 2015; Ovalle-Rivera *et al.*, 2015; Davis *et al.*, 2019; Moat *et al.*, 2019; de Aquino *et al.*, 2022). Projections from several climate models show that *C. arabica* will be drastically less suitable for cultivation in current coffee-producing regions by 2050, with reductions up to 50% of the actual cropping area (Magrach & Ghazoul, 2015; Gruter *et al.*, 2022; Cassamo *et al.*, 2023). The more climate-resilient *C. canephora* will also lose some areas suitable for cultivation, mainly in West Africa, but it is estimated that this may double globally, mainly in Brazil and Southeast Asia (Magrach & Ghazoul, 2015). However, it should be noted that recent studies underline higher sensitivity of *C. canephora* to high temperature than previously reported (Kath *et al.*, 2020). On the other hand, recent findings have shown that elite cultivars of both *C. arabica* and *C. canephora* might have a greater intrinsic ability to endure environmental stresses

(Rodrigues *et al*., 2016; Dubberstein *et al*., 2020; Rodrigues *et al*., 2024). Furthermore, elevated air [CO_2] can improve carbon assimilation and plant vigor (Ramalho *et al*., 2013; DaMatta *et al*., 2019), and carbon investment in reproductive structures (Rakocevic *et al*., 2021), overall associated with the upregulation of several photosynthetic-related genes (Marques *et al*., 2020). In fact, elevated air [CO_2] may even increase yield under adequate water supply (Verhage *et al*., 2017; Rahn *et al*., 2018). Additionally, increased air [CO_2] mitigates heat (Rodrigues *et al*., 2016; Scotti-Campos *et al*., 2019), and drought (Avila *et al.*, 2020; Semedo *et al*., 2021) impacts on carbon assimilation and plant growth. Still, the predicted climate conditions along this century are expected to pose fundamental challenges (and threats) to coffee producers (Koutouleas *et al*., 2022).

Considering the global environmental and anthropic scenario, as well as the generally reduced genetic diversity of commercial *C. arabica* cultivars, several approaches are being implemented to ensure the sustainability of this important crop. Among them, the introduction of *Coffea* wild-relatives in the value chain is a promising strategy to increase the adaptive and resilient capacity of the current systems (Läderach *et al*., 2011; Jaramillo *et al*., 2011; Magrach & Ghazoul, 2015; Davis *et al*., 2019). Among them, *C. racemosa* Lour., *C. zanguebariae* Lour., or *C. liberica* W.Bull ex Hiern have been identified as appropriate alternatives (Davis *et al*., 2021).These species present relevant tolerance to biotic and abiotic stresses (Halle & Faria, 1973; Guerreiro Filho, 1992; Davis *et al*., 2021), along with shorter maturation periods and unique flavor attributes, representing a rich gene pool, as well as an important resource for new coffee blends (Davis *et al*., 2021; Tapaça *et al*., 2023). Nevertheless, in the wild, *Coffea* spp. face high risk of extinction, likely with strong implications for the sustainability of the coffee sector (Davis *et al*., 2019).

Agroforestry systems (AFS), that is, the intercropping of coffee with shade trees, is another unavoidable (and more immediate) approach to promote the sustainability of the coffee crop (Koutouleas *et al*., 2022; da Silva *et al*., 2022; Cassamo *et al*., 2023). Shade trees provide protection from direct sunlight, and improve microclimate (natural cooling and higher air humidity) with consequent reduction of evapotranspiration (Partelli *et al*., 2014; Oliosi *et al*., 2016; Gomes *et al*., 2020; Gidey *et al*., 2020; Silva *et al*., 2022; Cassamo *et al*., 2022; 2023). When implanted in degraded forest ecosystems, AFS have also the potential to contribute to biodiversity restoration and carbon sequestration (Kessler *et al*., 2012; De Beenhouwer *et al.,* 2016; Ghimire *et al*., 2020; Gomes *et al*., 2020; IPCC, 2022).

1.3. Biodiversity Studies in *Coffea* as a strategy to ensure the sustainability of the coffee sector

The survival and continuity of a species in their natural habitats depend on the variety of genes present in their populations (Booy *et al*., 2000). As for any other species, the genetic diversity of coffee plays a key role in evolution and adaptation (Hughes *et al.*, 2008). Therefore, understanding and protecting different coffee species, particularly neglected wild-relatives, is extremely important to ensure the sustainability of the coffee crop. This is even more crucial considering the challenges imposed by climate change and population growth (Tapaça *et al*., 2023; Scalabrin *et al*., 2020; Salojärvi *et al*., 2024).

As referred above, *C. arabica*, the most economically important species, has a relatively low genetic diversity (Aggarwal *et al*., 2007; Lashermes *et al*., 2011; Tapaça *et al*., 2023).This is attributed to its autogamous nature (Carvalho & Krug, 1949) and to a unique polyploidization event at the origin of the tetraploid genome of this species, likely further narrowed in some cultivars (Scalabrin *et al.,* 2020). The low levels of genetic variation found in most *C. arabica* varieties are a major concern regarding the long-term sustainability of the sector, as there may not be enough genomic resilience to cope with climate change (Silvestrin *et al*., 2007; Scalabrin *et al.,* 2020). On the contrary, the allogamous wild-species *C. racemosa* maintains significantly higher levels of genetic diversity, an attribute valued for its remarkable adaptation and tolerance to climate change, especially when contrasted with *C. arabica* (Filho *et al.,* 1977; Tapaça *et al*., 2023). Indeed, despite the fact that wild species have a wealth of genetic variations and stress responsive genes, their potential contribution to the coffee value chain is far from being fully explored (Davis *et al*., 2007; Jaramillo *et al*., 2011; Davis *et al*., 2019; Tapaça *et al.,* 2023). Thus, genetic studies are essential to produce information on key genes involved in response and tolerance to environmental stressors, as well as on other valuable and desired characteristics related to grain quality, essential for breeding. This basic information will be crucial for the conservation and introduction of these native species into breeding programs to secure valuable genetic resources and to address emerging challenges to coffee production (Tapaça *et al*., 2023).

Another important driver of coffee evolution and adaptation is the associated microbiome (Choi *et al.,* 2021). This component is often neglected and detached from the research component (FAO *et al.,* 2021). The plant microbiome refers to the diverse community of microorganisms that inhabit the surface (epiphytes), and the internal tissues (endophytes) of plant organs, as well as soil around the roots (rhizosphere) (Mendes *et al*., 2013; Bulgarelli *et al*., 2013; Gomes *et al*., 2016). These microorganisms include archaea, bacteria, and fungi (Azevedo *et al*., 2000; Gnanamanickam & Immanuel, 2006; Mahatmanto *et al.,* 2023), collectively influencing the growth and fitness of plants (Duong *et al*., 2020; Afridi *et al*., 2022). Plant growth-promoting (PGP) microorganisms are an important component of the coffee microbiome (Duong *et al*., 2020), and are often involved in plant protection, nutrition, development and growth (Gomes *et al*., 2016; Egamberdieva *et al*., 2017; Duong *et al*., 2020). Ecologically, the rhizosphere microbiome has a crucial role in nutrient cycling and assimilation, as well as soil fertility and remediation (Castiglioni *et al*., 2008; Zhang *et al.,* 2008; Gomes *et al.*, 2016).

Recent concerns about the sustainability of coffee farming have motivated investigations about the coffee microbiome as a tool to understand the functional influence of microorganisms and to improve crop health and bean quality (Vaughan *et al*., 2015; Vale *et al*., 2021). Also, the use of microorganisms that can act as biocontrol agents is seen as a viable and promising alternative to the use of pesticides (Felber *et al*., 2016). Since nitrogen, phosphorus and iron are among the most limiting nutrients for plants, the use of nitrogen-fixing microorganisms, phosphorus solubilizers, and siderophore producers represent a sustainable strategy to reduce the dependence on chemical fertilizers (Aeron *et al.,* 2020). Thus, the use of PGP microbial consortia can potentially reduce ecological costs, improving plant development, and mitigating the impact of environmental stresses (Arif *et al*., 2020). In addition, the microbiome of coffee beans is essential to control natural fermentation processes and prevent the transfer of other microbial groups from leaves, fruit surface, over-ripe fruits, and soil, avoiding unwanted aromas in the final product (Vale *et al*., 2021).

1.4. Resumption of coffee production in Mozambique: the successful case of Gorongosa Mountain

Mozambique is one of the countries where a high number of wild coffee species has been described (Medina, 1955; Halle & Faria, 1973; Charrier & Berthaud, 1985; Davis *et al*., 2006). According to Medina (1955), the viability of coffee production in the country is restricted to six specific microecological niches, the low coastal lands of Inhambane, and the highlands of Gorongosa, Manica, Espungabera-Amatongas, Niassa, and Zambezia, all characterized by red fertile soils, moderate temperatures, and mean annual precipitation between 1000 and 2000 mm. The same author refers that, due to their peculiar characteristics, three *Mozambicoffea* species (*C. racemosa*, *C. zanguebariae*, and *C. eugenioides*) deserve special interest for the expansion of coffee producing regions throughout the world. However, coffee was never included among the set of strategic industrial crops for Mozambique.

In recent years several efforts to promote the coffee sector in Mozambique have been made. These included, the restoration of *C. zanguebariae* plantations in the Ibo island, province of Cabo Delgado (Navarini *et al*., 2024), and the introduction of *C. arabica* in Sanga and Mahua (Niassa), Alto Molocue and Gurue (Zambezia), Tsangano and Angonia (Tete), Chimanimani (Manica), and Gorongosa (Sofala) (Fenagri, 2022).

The Gorongosa National Park (GNP) instaled the first coffee plantations in 2013, as part of a pilot project (2013/2017) aiming at restoring the rain forest of the Gorongosa mountain through the implemenation of an agroforestry system integrating native trees and coffee (Stalmans, 2017). This initiative was based on an integrated relation between the sustainable use of natural resources, biodiversity conservation, human health, and community development. It should be noted that the rate of deforestation in the region has been alarming, with *ca*. 40% forest loss in 2019 (Stalmans, 2020).

Based on the socioeconomic indicators of the experimental phase, the Coffee Agroforestry System (CAFS) from GNP was gradually expanded from 50 to 250 ha in 2023, through a transdisciplinary approach integrating community education and development (including education and food security), best agricultural practices, coffee trading, research and advanced training (Tricafé, 2023).The last pillar was settled under the trilateral partnership between Mozambique (GNP), Brasil (Universidade Federal do Espírito Santo), and Portugal (Universidade de Lisboa, School of Agriculture), funded by the Portuguese (Camões, I.P.) and the Brazilian Cooperation (ABC) Agencies (2017/2023). Until the end of 2023, more than 800,000 coffee plants and 140,000 native trees have been planted, with the direct involvement of the entire local community (*ca*. 1000 families), and a successful trading experience in national and international markets (Tricafé, 2023). The scientific component was settled on three pillars: (i) Coffee performance: from seed to cup; (ii) Biodiversity indicators; and (iii) Carbon dynamics. Considering the projected climate scenarios, and taking into account that Mozambique is among the most vulnerable regions, Cassamo *et al*. (2022) identified the Gorongosa Mountain as one of the most promising agro-ecological areas for coffee production. However, high productivity and grain quality are strongly tight to altitude (> 900m) and moderate shade under AFS (Cassamo *et al*., 2023). The implementation of the CAFS was accompanied by the recovery of several biodiversity indicators, including birds (Ngovene, 2020; Lupaka, 2022), and insects (Alfinete, 2022; Caetano, 2022). In addition, above and belowground carbon sequestration tend to approach the levels stored in the native forest (Cossa, 2022; Real, 2022).

1.5. Objectives of the study and outline of the thesis

GNP is one of the greatest cases of success of wildlife restoration in Africa (Huntley, 2023). Before the independence of Mozambique (1975), it was considered a paradise of wildlife and landscapes. As a result of the civil war that lasted from 1975 to 1992, nearly 85% of the wildlife population was lost, and several species were extirpated. In 2004, the American philanthropist, Greg Carr, and the Mozambican president, Joaquim Chissano, established an ambitious plan to restore the Park, the *Gorongosa Restoration Project* (GRP). Since then, the Park has become the nexus between piece, human development, eco-tourism, nature conservation, and science. While restoring and preserving wild ecosystems (namely by avoiding slash-and-burn itinerant agriculture), GRP safeguards the needs of local communities, which are heavily dependent on the resources provided by the forest (*e.g.*, wood, firewood, primary health, food).

Taking into account the global challenges for the coffee sector, and under the umbrella of the Tricafé Project (www.tricafe.org), this study aimed at contributing to the GRP, bridging some research gaps and complementing emerging coffee science in Mozambique (Cassamo *et al*., 2022; 2023; Alberto *et al*., 2023; Navarini *et al*., 2024). In this context, three research questions related to biodiversity were formulated: (i) what is the actual status of the genetic heritage of *Coffea* in Mozambique and their respective potential to endure climate pressure?; (ii) is the genetic pool of *C. arabica* implanted in Gorongosa enough to ensure the sustainability of the CAFS?; (iii) how do elevation and shade affect coffee microbial diversity and associated functions?. The answer to these questions constitutes a baseline for the integration of native coffee species in the coffee value chain, and to maximize the use of native microbial species for the development of bio-based solutions, such as bioremediation, biofertilization, or fruit fermentation. Based on that, the following objectives were defined:

> I. To survey and characterize the genetic diversity of *Coffea* in southern (Maputo, Gaza, and Inhambane) and central (Sofala) Mozambique through DNA markers (Cloroplast-based, Microsatellites and Single Nucleotide Polymorphisms*).*
>
> II. To analyze the diversity and putative functions of the microbial flora associated with Arabica coffee (soil, leaves, and fruits) grown in Gorongosa along the elevational and shade gradient.

The thesis is organized in 5 chapters. In this chapter (**Chapter I**), a general introduction to coffee from a global to local perspective and the objectives of the thesis are presented. **Chapter II** reports on the genomic evaluation of *C. arabica* and *C. racemosa*, centered on taxonomy, diversity and structure, and functional attributes. **Chapters III and IV** are dedicated to high throughput Next Generation Sequencing analysis of the soil and endosphere microbial communities (bacteria, archaea, and fungi), respectively, along a gradient of irradiance and altitude. In the last chapter, (**Chapter V**), general conclusions and considerations are presented.

1.6. References

Aeron, A., Khare, E., Jha, C.K., *et al*. Revisiting the plant growth-promoting rhizobacteria: Lessons from the past and objectives for the future. *Arch Microbiol*. 2020, 202, 665-676, https://doi.org/10.1007/s00203-019-01779-w.

Afridi, M.S., Ali, S., Salam, A., *et al*. Plant microbiome engineering: Hopes or hypes. *Biol.* 2022*,* 11, 1782, https://doi.org/10.3390/biology11121782.

Arif, I., Batool, M., Schenk, P.M., *et al*. Plant microbiome engineering: Expected benefits for improved crop growth and resilience. *Trends Biotech*. 2020, 38, 1385-1396, https://doi.org/10.1016/j.tibtech.2020.04.015.

Aggarwal, R.K., Hendre, P.S., Varshney, R.K., *et al*. Identification characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor Appl Genet*. 2007, 114, 359, https://doi.org/10.1007/s00122-006-0440-x.

Alfinete, A. Variação nas comunidades de aranhas nos diferentes estágios da sucessão ecológica no sistema agroflorestal na serra da Gorongosa**.** Master thesis, Gorongosa Bioeducation Consortium (UniZambeze, UniLurio, ISPM, ULisboa, Gorongosa National Park). 2022, 40 pp.

Avila, R.T., Almeida, W.L., Costa, L.C., *et al*. Elevated air [CO_2] improves photosynthetic performance and alters biomass accumulation and partitioning in drought-stressed coffee plants. *Env Exp Bot*. 2020, 177, 104137, https://doi.org/10.1016/j.envexpbot.2020.104137.

Azevedo, J.L., Maccheroni, W., Pereira, J.O., *et al*. Endophytic microorganisms: A review on insect control and recent advances on tropical plants. *Electron J Biotech*. 2000, 3, 40-65, http://www.ejb.org/content/vol3/issue1/full/4.

Alberto, N.J., Ramalho, J.C., Ribeiro-Barros, A.I., *et al*. Diversity in *Coffea arabica* cultivars in the mountains of Gorongosa National Park, Mozambique, regarding bean and leaf nutrient accumulation and physical fruit traits. *Agronomy*. 2023, 13, 13041162. https://doi.org/10.3390/agronomy13041162.

Bilen, C., El Chami, D., Mereu, V., *et al.* Systematic review on the impacts of climate change on coffee agrosystems. *Plants.* 2023, 12, 102, https://doi.org/10.3390/plants12010102.

Bawin, Y., Ruttink, T., Staelens, A., *et al*. Phylogenomic analysis clarifies the evolutionary origin of Coffea arabica L. J Syst Evol. 2021, 59, 953-963, https://doi.org/10.1111/jse.12694.

Bunn, C., Läderach, P., Ovalle Rivera, O., *et al.* A bitter cup: Climate change profile of global production of *arabica* and *robusta* coffee. *Clim Change*. 2015,129, 89–101, https://doi.org/10.1007/s10584-014-1306-x.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., *et al*. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol*. 2013, 64, 807-838, https://doi.org/10.1146/annurev-arplant-050312-120106.

Booy, G., Hendriks, R.J., Smulders, M.J.M., *et al*. Genetic diversity and the survival of populations. *Plant Biol*. 2000, 2, 375-395**,** https://doi.org/10.1055/s-2000-5958.

Choi, K., Khan, R., Lee, S.W., *et al*. Dissection of plant microbiota and plant-microbiome interactions. *J Microbiol*. 2021, 59, 281-291, https://doi.org/10.1007/s12275-021-0619-5.

Caetano, E. Borboletas como bioindicadoras da degradação e restauração de habitats na Serra da Gorongosa. Master Thesis, Gorongosa Bioeducation Consortium (UniZambeze, UniLurio, ISPM, ULisboa, Gorongosa NationalPark). 2022, 44 pp.

Cassamo, C.T., Mangueze, A.V.J., Leitão, A.E., *et al*. Shade and altitude implications on the physical and chemical attributes of green coffee beans from Gorongosa Mountain, Mozambique. *Agronomy*. 2022, 12, 2540, https://doi.org/10.3390/agronomy12102540.

Cassamo, C.T., Draper, D., Romeiras, M.M., *et al*. Impact of climate changes in the suitable areas for *Coffea arabica* L. production in Mozambique: Agroforestry as an alternative management system to strengthen crop sustainability. *Agric Ecosyst Environ*. 2023, 346, 108341, https://doi.org/10.1016/j.agee.2022.108341.

Carvalho, A., Krug, C.A. Agentes de polinização da flor do cafeeiro (*Coffea arabica* L.) *Bragantia*. 1949, 9, 11-24, https://doi.org/10.1590/S0006-87051949000100002

Castiglioni, P., Warner, D., Bensen, R.J., *et al*. Bacterial RNA chaperones confer abiotic stress tolerance in plants and improved grain yield in maize under water-limited conditions. *Plant Physiol*. 2008,147, 446-55, https://doi:org/10.1104/pp.108.118828.

Charrier, A., Berthaud, J. Botanical classification of coffee. In: Clifford, M.N., Willson, K.C. (eds). *Springe*r, Boston, MA. 1985, 13-47, https://doi.org/10.1007/978-1-4615-6657-1_2.

Cossa, M. Dinâmica do carbono do solo nos sistemas agroflorestais de café na Serra da Gorongosa. Master Thesis, Gorongosa Bioeducation Consortium (UniZambeze, UniLurio, ISPM, ULisboa, Gorongosa NationalPark). 2022, 54 pp.

DaMatta, F.M., Rahn, E., Läderach, P., *et al*. Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? *Clim Change.* 2019,152, 167–178, https://doi.org/10.1007/s10584-018-2346-4.

Davis, A.P., Govaerts, R., Bridson, D.M., *et al*. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Bot J Linn Soc*. 2006, 152, 465–512, https://doi.org/10.1111/j.1095-8339.2006.00584.x.

Davis, A.P., Chester, M., Maurin, O., *et al*. Searching for the relatives of *Coffea* (Rubiaceae, Ixoroideae): The circumscription and phylogeny of *Coffeae* based on plastid sequence data and morphology. *Am J Bot*. 2007, 94, 313–329, https://doi.org/10.3732/ajb.94.3.313.

Davis, A.P., Tosh, J., Ruch, N., *et al*. Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Bot J Linn Soc*. 2011, 167, 357–377, https://doi.org/10.1111/j.1095-8339.2011.01177.x.

Davis, A.P., Chadburn, H., Moat, J., *et al*. High extinction risk for wild coffee species and implications for coffee sector sustainability. *Sci Adv*. 2019, 5, eaav3473, https:/doi.org/10.1126/sciadv.aav3473.

Davis, A.P., Gargiulo, R., Almeida, I.N.d.M., *et al*. Hot coffee: The Identity, climate profiles, agronomy, and beverage characteristics of *Coffea racemosa* and *Coffea zanguebariae*. *Front Sustain Food Syst*. 2021, 5, 740137, https://doi.org/10.3389/fsufs.2021.740137.

Davis, A.P., Rakotonasolo, F. Six new species of coffee (*Coffea*) from northern Madagascar. *Kew Bull*. 2021, 76, 497–511, https://doi.org/10.1007/s12225-021-09952-5.

De Beenhouwer, M., Geeraert, L., Mertens, J., *et al*. Biodiversity and carbon storage co-benefits of coffee agroforestry across a gradient of increasing management intensity in the SW Ethiopian highlands. *Agric Ecosyst Environ*. 2016, 222, 193–199, https://doi.org/10.1016/j.agee.2016.02.017.

Dubberstein, D., Lidon, F.C., Rodrigues, A.P., *et al*. Resilient and sensitive key points of the photosynthetic machinery of *Coffea* spp. To the single and superimposed exposure to severe drought and heat stresses. *Front Plant Sci*. 2020, 11, 01049, https://doi.org/10.3389/fpls.2020.01049.

Duong, B., Marraccini, P., Maeght, J.L., *et al*. Coffee microbiota and its potential use in sustainable crop management. A review. *Front Sustain Food Syst*. 2020, 4, 607935, https://doi.org/10.3389/fsufs.2020.607935.

de Aquino, S.O., Kiwuka, C., Tournebize, R., *et al*. Adaptive potential of *Coffea canephora* from Uganda in response to climate change. *Mol Ecol* 2022, 31, 1800-1819, https://doi.org/10.1111/mec.16360.

da Silva, J.M.V. de O., Souza, M.N., Rangel, O.J.P., *et al*. Sistemas agroflorestais e consórcios na cultura do café. In: Souza, M.N. (ed). Méricadas, Canoas, RS. 2022, 4, 171-200, https://doi.org/10.4322/mp.978-65-84548-10-7.c6.

Egamberdieva, D., Wirth, S.J., Alqarawi, A.A., *et al*. Phytohormones and beneficial microbes: essential components for plants to balance stress and fitness. *Front Microbiol*. 2017, 8, 02104, https://doi.org/10.3389/fmicb.2017.02104.

Gnanamanickam, S.S., Immanuel, J.E. Epiphytic bacteria, their ecology and functions. In: Gnanamanickam, S.S. (eds) Plant-associated bacteria. *Springer, Dordrecht*. 2006, 131-153, https://doi.org/10.1007/978-1-4020-4538-7_4

FAO. Crop statistics. 2021. [(Accessed on February 2024)]. Availableonline: http://www.fao.org/faostat/en/#home.

Felber, A.C., Orlandelli, R.C., Rhoden, S., *et al*. Bioprospecting foliar endophytic fungi of *Vitis labrusc*a Linneaus, Bordô and Concord cv. *Ann Microbiol*. 2016, 66, 765–775, https://doi.org/10.1007/s13213-015-1162-6.

Fenagri - Frederação Nacional de Associações Agrárias de Moçambique. Café: Uma cultura que promete prosperidade ao sector familiar, 2022, 4, https://fenagri.co.mz/wp-content/uploads/2022/10/FENAGRI_Newsletter_004.pdf.

Filho, H.P.M., Carvalho, A., Medina, D.M., *et al*. Germoplasma de *Coffea racemosa* e seu potencial de melhoramento do cafeeiro. Bragantia. 1977, 36, 43–46, https://doi.org/10.1590/S0006-87051977000100040.

Fonseca, A.F.A., Ferrão, R.G., Ferrão, M.A.G., *et al*. Jardins clonais, produção de sementes e mudas. In: Ferrão, R.G., Fonseca, A.F.A., Bragança, S.M., *et al.* (eds). *Incaper*, Vitória, ES. 2007, 227-255, http://biblioteca.incaper.es.gov.br/digital/handle/item/700.

Ghimire, P., Bolakhe, S. Agroforestry systems: Biodiversity, carbon stocks and contribution to rural livelihood. *J Agric For Univ*. 2020, 4, 197–205, https://doi.org/10.3126/jafu.v4i1.47071.

Gidey, T., Oliveira, T.S., Crous-Duran, J., *et al*. Using the yield-safe model to assess the impacts of climate change on yield of coffee (*Coffea arabica* L.) under agroforestry and monoculture systems. *Agrofor Syst*. 2020, 94, 57–70, https://doi.org/10.1007/s10457-019-00369-5.

Gomes, E.A., Silva, U.de C., Paiva, C.A. de O., *et al*. Microrganismos promotores do crescimento de plantas. Embrapa milho e sorgo, Sete Lagoas. 2016, 51, http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1063799.

Gomes, L.C., Bianchi, F.J.J.A., Cardoso, I.M., *et al*. Agroforestry systems can mitigate the impacts of climate change on coffee production: A spatially explicit assessment in Brazil. *Agric Ecosyst Environ*. 2020, 294, https://doi.org/10.1016/j.agee.2020.106858.

Grüter, R., Trachsel, T., Laube, P., *et al*. Expected global suitability of coffee, cashew and avocado due to climate change. *PLoS One*. 2022, 17, e0261976, https://doi.org/10.1371/journal.pone.0261976.

Guerreiro-Filho, O. *Coffea racemosa* Lour. *Une revue. Café, Cacao, Thé*. 1992, 36, 171–186, https://agritrop.cirad.fr/408807/.

Guerreiro-Filho, O., Mendes, A.N.G., Carvalho, G.R., *et al.* Origem e classificação botânica do cafeeiro. In: de Carvalho, C.H.S. (ed). *Embrapa café*, Brasília, DF. 2008, 27-32.

Halle, F., Faria, M.T. Le *Coffea racemosa* Lour. *Agron Moçamb.* 1973, *7*, 243–250.

Hughes, A.R., Inouye, B.D., Johnson, M.T.J., *et al*. Ecological consequences of genetic diversity. *Ecol Lett*. 2008, 11, 609-623, https://doi.org/10.1111/j.1461-0248.2008.01179.x.

ICO – International Coffee Organization. Coffee market report 2023. 2023. Available online: https://icocoffee.org/specialized-reports/ (accessed on June 2024).

IPCC – Intergovernmental panel on climate change. Mitigation pathways compatible with 1.5°C in the context of sustainable development. 2022, https://doi.org/10.1017/9781009157940.

Jaramillo, J., Muchugu, E., Vega, F.E., *et al.* Some like it hot: The influence and implications of climate change on coffee berry borer (*Hypothenemus hampei*) and coffee production in East Africa. *PLoS One*. 2011, 6, https://doi:org/10.1371/journal.pone.0024528.

Kath, J., Byrareddy, V.M., Craparo, A., *et al*. Not so robust: Robusta coffee production is highly sensitive to temperature. *Glob Chang Biol*. 2020, 26, 3677-3688, https://doi.org/10.1111/gcb.15097.

Kessler, M., Hertel, D., Jungkunst, H.F., *et al.* Can joint carbon and biodiversity management in tropical agroforestry landscapes be optimized? *PLoS One*. 2012, 7, e47192, https://doi.org/10.1371/journal.pone.0047192.

Koutouleas, A., Sarzynski, T., Bordeaux, M., *et al*. Shaded-coffee: Anature-based strategy for coffee production under climate change? A review. *Front Sustain Food Syst.* 2022a, 6, 877476, https://doi.org/10.3389/fsufs.2022.877476.

Laderach, P., Lundy, M., Jarvis, A., *et al*. Expected impact of climate change on coffee production chains. In: Filho WL (ed). *Springer*, Berlin, Heidelberg. 2011, 703–23, https://doi.org/10.1007/978-3-642-14776-0_42.

Lashermes, P., Combes, M.C., Trouslot, P., *et al*. Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. *Theor Appl Genet.* 1997, 94, 947-955, https://doi.org/10.1007/s001220050500.

Lashermes, P., Combes, M.C., Robert, J., *et al*. Molecular characterisation and origin of the *Coffea arabica* L. genome. *Mol Gen Genet*. 1999, 261, 259-66, https://doi.org/10.1007/s004380050965.

Lashermes, P., Combes, M.C., Ansaldi, C., *et al*. Analysis of alien introgression in coffee tree (*Coffea arabica* L). *Mol Breed*. 2011, 27, 223-232, https://doi.org/10.1007/s11032-010-9424-2.

Lupaka, C. Efeito da predação de artrópodes por aves e morcegos na plantação de café de sombra da Serra da Gorongosa. Master Thesis, Gorongosa Bioeducation Consortium (UniZambeze, UniLurio, ISPM, ULisboa, Gorongosa National Park). 2022, 34pp.

Magrach, A., Ghazoul, J. Climate and pest-driven geographic shifts in global coffee production: Implications for forest cover, biodiversity and carbon storage. *PloS One*. 2015, 10, e0133071, https://doi.org/10.1371/journal.pone.0133071.

Mahatmanto, T., Sunarharum, W.B., Fanny A Putri, F., *et al*, The microbiology of arabica and robusta coffee cherries: a comparative study of indigenous bacteria with presumptive impact on coffee quality. *FEMS Microbiology Letters*, 2023, 370, fnad024, https://doi.org/10.1093/femsle/fnad024

Marques, I., Fernandes, I., David, P.H.C., *et al*. Transcriptomic leaf profiling reveals differential responses of the two most traded coffee species to elevated [CO_2]. *Int J Mol Sci*. 2020, 21, 9211, https://doi.org/10.3390/ijms21239211.

Medina, A.R. Possibilidades de produção de café em Moçambique. Revista do Café Português. 1955, II, 16-33.

Mendes, R., Garbeva, P., Raaijmakers, J.M., *et al*. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. FEMS *Microbiol Rev*. 2013, 37, 634-63, https://doi.org/10.1111/1574-6976.12028.

Moat, J., Gole, T. W., Davis, A. P., *et al*. Least concern to endangered: Applying climate change projections profoundly influences the extinction risk assessment for wild Arabica coffee. *Glob Chang Biol.* 2019, 25, 390–403, https://doi.org/10.1111/gcb.14341.

Millard, C.J., Watson, P.J., Fairall, L., *et al*. Targeting class i histone deacetylases in a "complex" environment. *Trends Pharmacol Sci*. 2017, 38, 363-377, https://doi.org/10.1016/j.tips.2016.12.006.

Navarini, L., Scaglione, D., Del Terra, L. *et al*. Mozambican *Coffea* accessions from Ibo and Quirimba Islands: Identification and geographical distribution. *AoB Plants*. 2024, 16, https://doi.org/10.1093/aobpla/plae004.

Ngovene, A. Pode o café de sombra ajudar a restaurar a biodiversidade na floresta da Serra da Gorongosa. Master Thesis, Gorongosa Bioeducation Consortium (UniZambeze, UniLurio, ISPM, ULisboa, Gorongosa National Park). 2020, 50 pp.

Ngure, G.M., Watanabe, K.N. Coffee sustainability: Leveraging collaborative breeding for variety improvement. *Front Sustain Food Syst*. 2024, 8, 1431849, https://doi.org/10.3389/fsufs.2024.1431849.

Oliosi, G., Giles, J.A.D., Rodrigues, W.P., *et al*. Microclimate and development of *Coffea canephora* cv. *Conilon* under different shading levels promoted by Australian cedar (*Toona ciliata* M. Roem. var. Australis). *Austr J Crop Sci*. 2016, 10, 528–538, https://doi.org/10.21475/ajcs.2016.10.04.p7295x.

Ovalle-Rivera, O., Läderach, P., Bunn, C., *et al*. Projected shifts in *Coffea arabica* suitability among major global producing regions due to climate change. *PLoS One*. 2015, 10, e0124155, https://doi.org/10.1371/journal.pone.0124155.

Partelli, F.L., Araújo, A.V., Vieira, H.D., *et al*. Microclimate and development of ’conilon’ coffee intercropped with rubber trees. *Pesq Agropec Bras*. 2014, 49, 872–881, https://doi.org/10.1590/s0100-204x2014001100006.

Rahn, E., Vaast, P., Läderach, P., *et al*. Exploring adaptation strategies of coffee production to climate change using a process based model. *Ecol. Model*. 2018, 371, 76–89, https://doi.org/10.1016/j.ecolmodel.2018.01.009.

Rakocevic, M., Batista, E.R., Pazianotto, R.A.A., *et al*. Leaf gas exchange and bean quality fluctuations over the whole canopy vertical profile of Arabic coffee cultivated under elevated CO_2. *Funct Plant Biol*. 2021, 48, 469–482, https://doi.org/10.1071/FP20298.

Ramalho, J.C., Rodrigues, A.P., Semedo, J.N., *et al*. Sustained photosynthetic performance of *Coffea* spp. under long-term enhanced [CO_2]. *PLoS One*. 2013, 8, e82712, https://doi.org/10.1371/journal.pone.0082712.

Razafinarivo, N.J., Guyot, R., Davis, A.P., *et al*. Genetic structure and diversity of coffee (*Coffea*) across Africa and the Indian Ocean islands revealed using microsatellites. *Annal sof Botany*. 2013, 111, 229-248, https://doi.org/10.1093/aob/mcs283.

Real, A. Carbono armazenado na vegetação lenhosa num sistema agroflorestal de cafeeiro e numa floresta natural na Serra da Gorongosa, Moçambique. Master Thesis, Gorongosa Bioeducation Consortium (UniZambeze, UniLurio, ISPM, ULisboa, Gorongosa NationalPark). 2022, 66 pp.

Rodrigues, A.P., Pais, I.P., Leitão, A.E., *et al*. Uncovering the wide protective responses in *Coffea* spp. leaves to single and superimposed exposure of warming and severe water deficit. *Front Plant Sci*. 2024, 14, 1320552, https://doi.org/10.3389/fpls.2023.1320552.

Rodrigues, W.P., Martins, M.Q., Fortunato, A.S., *et al*. Long-term elevated air [CO_2] strengthens photosynthetic functioning and mitigates the impact of supra-optimal temperatures in tropical *Coffea arabica* and *C. canephora* species. *Glob Change Biol*. 2016, 22, 415–431, https://doi.org/10.1111/gcb.13088.

Sakiyama, N.S., Pereira, A.A., Zambolim, L., *et al*. Melhoramento de *café arábica* em melhoramento de espécies cultivadas. In: Borém, A. (ed). UFV: Viçosa, MG. 1999, 189–204.

Scalabrin, S., Toniutti, L., Di Gaspero, G., *et al*. A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Sci Rep*. 2020, 10, 4642, https://doi.org/10.1038/s41598-020-61216-7.

Scotti-Campos, P., Pais, I.P., Ribeiro-Barros, A.I., *et al*. Lipid profile adjustments may contribute to warming acclimation and to heat impact mitigation by elevated [CO_2] in *Coffea* spp. *Environ Exp Bot*. 2019, 167, 103856, https://doi.org/10.1016/j.envexpbot.2019.103856.

Semedo, J.N., Rodrigues, A.P., Lidon, F.C., *et al.* Intrinsic non-stomatal resilience to drought of the photosynthetic apparatus in *Coffea* spp. can be strengthened by elevated air CO_2. *Tree Physiol*. 2021, 41, 708–727, https://doi.org/10.1093/treephys/tpaa158.

Silvestrini, M., Junqueira, M.G., Favarin, A.C., *et al*. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genet Resour Crop Evol*. 2007, 54, 1367–1379, https://doi.org/10.1007/s10722-006-9122-4.

Souza, A.L. de., Garcia, R., Bernardino, F.S., *et al*. Coffee hulls in the diet of sheep: intake and apparent digestibility. *Ver Bras Zootec*. 2004, 33, 2170-2176, https://www.feedipedia.org/node/2809.

Stalmans, M. Serra da Gorongosa - Feedback visit and some thoughts on the way forward. *Tech Rep*. 2017, 19.

Stalmans, M., Victor, M. Forest cover on Gorongosa Mountain - assessment of satellite imagery. *Tech Rep*. 2020, 14 pp.

Salojärvi, J., Rambani, A., Yu, Z., *et al*. The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars. *Nat Genet*. 2024, 56, 721–73, https://doi.org/10.1038/s41588-024-01695-w.

Tapaça, I.d.P.E., Mavuque, L., Corti, R., *et al*. Genomic evaluation of *Coffea arabica* and its wild relative *Coffea racemosa* in Mozambique: Settling resilience keys for the Coffee crop in the context of climate change. *Plants*. 2023, 12, 2024,https://doi.org/10.3390/plants12102044.

Teixeira, A.L., Rocha, R.B., Ramalho, A.R., *et al.* Melhoramento genético, registro e proteção de cultivares de *Coffea canephora* para o estado de rondônia / Porto Velho, RO: *Embrapa Rondônia*. 2011, 23, 143, https://ainfo.cnptia.embrapa.br/digital/bitstream/item/57882/1/doc143-cafe.pdf.

Tricafé. Produção sustentável de café no Parque Nacional da Gorongosa em sistema agroflorestal integrado, no contexto da desflorestação, alterações climáticas e segurança alimentar. Relatório final de execução técnica. 2023, 19 pp.

Vale, F.X.R., Zambolim, L., Jesus Jr, W.C., *et al*. Effect of climatic factors on the occurrence and development of coffee rust.*1st Brazilian Coffee Research Symposium*. 2021, 171-174.

Vaughan, M.J., Mitchell, T., McSpadden Gardener, B.B., *et al*. What's inside that seed we brew? A new approach to mining the coffee microbiome. *Appl Environ Microbiol*. 2015, 81, 6518-27, https://doi.org/10.1128/AEM.01933-15.

Verhage, F.Y.F., Antem, N.P.R., Sentelhas, P.C., *et al*. Carbon dioxide fertilization offsets negative impacts of climate change on arabica coffee yield in Brazil. *Clim Change*. 2017, 144, 671–685, https://doi.org/10.1007/s10584-017-2068-z.

Voora, V., Bermúdez, S., Larrea, C., *et al*. Global market report: Coffee. IISD 2019, 6, https://www.iisd.org/system/files/publications/ssi-global-market-report-coffee.pdf. (Accessed on 30 July 2019).

Zhang, H., Kim, M.S., Sun, Y., *et al*. Soil bacteria confer plant salt tolerance by tissue-specific regulation of the sodium transporter HKT1. *Mol Plant Microbe Interact*. 2008, 21, 737-44, https://doi.org/10.1094/MPMI-21-6-0737.

Chapter 2

Genomic evaluation of *Coffea arabica* and its wild relative *Coffea racemosa* in Mozambique: settling resilience keys for the coffee crop in the context of climate changes

Modified after Tapaça, I.d.P.E.; Mavuque, L., Corti, R., Pedrazzani, S., Maquia, I.S.A., Tongai, C., Partelli, F.L., Ramalho, J.C., Marques, I., Ribeiro-Barros, A.I. Genomic evaluation of *Coffea arabica* and its wild relative *Coffea racemosa* in Mozambique: settling resilience keys for the coffee crop in the context of climate change. Plants. 2023, 12, 2044, https://doi.org/10.3390/plants12102044.

2.1. Introduction

Coffee (*Coffea arabica* L.) plays a prominent agricultural, social, and commercial role, standing as one of the world's largest agricultural supply chains. The livelihoods of almost 25 million people are directly dependent on coffee cultivation, and more than 100 million people in over 80 countries are involved across the entire value chain (ICO, 2022). The coffee market is also growing due to increasing consumption in emerging economies and a stronger interest in specialty coffee (ICO, 2022). Despite the expansion of the coffee sector, the market relies mostly on genotypes from two species: *Coffea arabica* L. (Arabica coffee), which dominates the world market, and one of its ancestors, *Coffea canephora* Pierre ex A. Froehner (Robusta coffee) (DaMatta & Ramalho 2006; Davis *et al*., 2011).

However, the two-mainstream species are highly sensitive to changing temperatures and water availability (DaMatta & Ramalho 2006; Dubberstein *et al*., 2020; Fernandes *et al*., 2021), with visible impacts across the coffee value chain (Venancio *et al*., 2020; ICO, 2022). Traditionally, Arabica cultivars have an optimal annual mean temperature ranging between 18–21 °C, with temperatures above 23 °C accelerating fruit ripening, which can cause bean quality loss (DaMatta *et al*., 2019). Seasonal high temperatures above 33 °C and dryer seasons can also reduce floral initiation and increase the production of abnormal reproductive structures and flower abortion (de Oliveira *et al*., 2020). Drought decreases coffee yield and quality, especially in *C. canephora* (Byrareddy *et al*., 2021). The effects of drought are also aggravated in Arabica plantations under full sunlight exposure (Cassamo *et al*., 2023). According to the International Coffee Organization, 2020 and 2021 were already marked by a global reduction of Arabica and Robusta stocks associated with the vulnerability of these species to extreme weather events, with notable uneven increases in market prices (ICO, 2022). The future seems not to be better as modeling analyses predict that the supply chain will be severely affected by climate change across this century, with significant effects on coffee yield and quality (Lãderach *et al*., 2017; Chemura *et al*., 2021; Lara-Estrada *et al*., 2021; Cassamo *et al*., 2023).

The quite low levels of genetic variation found in most commercial coffee cultivars (Scalabrin *et al.,* 2020) constitute a major concern regarding the long-term sustainability of the sector since there might not be enough genomic resilience to keep pace with climatic change (Silvestrini *et al*., 2007). Looking back at coffee, wild relatives offer the potential to increase the adaptive capacity of agricultural systems to diseases and climatic pressures, representing a large pool of new, untapped, genetic variation (Davis *et al*., 2007; Jaramillo *et al*., 2011; Davis *et al*., 2019). Therefore, considering the global environmental and anthropogenic scenario, as well as the narrow genetic plasticity of commercial coffee cultivars, several approaches are being implemented to ensure the sustainability of this important crop. The introduction of wild relatives and underutilized species, such as *C. racemosa* Lour., *C. zanguebariae* Lour., or *C. liberica* W.Bull ex Hiern in the value chain, has been pointed out as one of the most promising approaches (Davis *et al*., 2021). These species are considered highly resilient to environmental pressures, particularly high temperatures, and extreme precipitation events (Halle *et al*., 1973; Guerreiro, 1992; Davis *et al*., 2021). Additionally, *C. racemosa* (and probably *C. zanguebariae*, which is often misclassified as *C. racemosa*) shows relevant resistance to several pests and diseases (Guerreiro, 1992; Davis *et al*., 2021; Davis *et al*., 2021). Such tolerance to abiotic and biotic stresses, together with the short ripening period and unique flavor attributes, make these species an outstanding gene pool, as well as an important resource for new coffee blends (Davis *et al*., 2021).

Another recommended approach to mitigate climate change impacts on the coffee crop is the shift from intensive production under the full sun (and monocrop systems) to agroforestry systems (AFS) using native or other economic-important trees for shade, constituting refuges for biodiversity and providing multiple ecosystem services (e.g., timber, food, carbon sequestration, or nutrient cycling) (Gomes *et al*., 2020; Gidey *et al*., 2020; Cassamo *et al*., 2022; 2023). The effectiveness of this strategy, which provides a better micro-environment for coffee plants, is however dependent on several factors, such as the agroecological conditions, shade density, crop management, and the cultivars used (Koutouleas *et al*., 2022a; Koutouleas *et al*., 2022b). Successful coffee AFS systems have been implemented in several countries in the Americas and Africa countries (Nesper *et al*., 2017; Gomes *et al*., 2020; De Leijster *et al*., 2021; Cassamo *et al*., 2022; 2023). Among them, the coffee AFS system established in the Gorongosa Mountain, which is part of the Gorongosa National Park (GNP) in Mozambique, has been pointed out as one of the most emblematic cases, not only in terms of the coffee crop sustainability but also regarding the positive socio-economic benefits to local communities, with a direct impact in the reversion of the accelerated loss and degradation of the tropical rainforest (Cassamo *et al*., 2022; 2023).

Located in southern Africa, Mozambique might contribute significantly to the coffee value chain, although it is currently not included in the list of coffee-producing countries: (i) it is a promising source of coffee wild-relatives (Davis *et al*., 2021; Halle *et al*., 1973; Davis *et al*., 2021), and (ii) it has a remarkable abundance of native forests (Sitoe *et al*., 2012). A recent molecular study solved ambiguities between *C. racemosa* and *C. zanguebariae* and elucidated their distribution in north eastern Mozambique (Cabo Delgado Province) *(*Davis *et al*., 2021*)*. The same authors pinpointed the knowledge gap regarding the current distribution of these species in the rest of the country. Indeed, cultivation of *C. racemosa* in central Mozambique (Inhambane Province) was first reported in 1876 (Hiern *et al*., 1876), and the last full description dates from 1973 (Halle *et al*., 1973). More recently, a new study mapped the distribution and suitability of *C. arabica* plantations across the country (Cassamo *et al*., 2023). Four regions (Manica, Sofala, Zambezia, and Nampula) were identified as suitable for Arabica, particularly under AFS management.

In this study, we have assessed, for the first time, the occurrence of *Coffea* species in southern and central Mozambique (Figure 2.1A) and the impact of genetic diversity on the long-term sustainability of the AFS implemented in Gorongosa Mountain. Specifically, we first aimed at understanding the degree of genetic relationships between *Coffea* species in Mozambique. For that, we used plastid markers to understand the phylogenetic relationship between these two species. Then, we used nuclear simple sequence repeat (SSR) polymorphisms to understand the patterns of genetic diversity, differentiation, and genetic structure. For that, we compared the cultivated *C. arabica* plants (Figure 2.1B) with the ones farmed in northern Mozambique (Niassa) together with the wild relative, *C. racemosa*, that was found during our field surveys (Figure 2.1C). To complement the SSR study, we investigated coffee genome-wide diversity using single nucleotide polymorphisms (SNP) generated by Genotype-by-Sequencing (GBS) on a reduced sampling set, allowing us to further detect the patterns of diversity and functional pathways involved, as well as to test possible differences with SSR markers. This is the first study that reveals the underlying genomic mechanisms explaining the different adaptation abilities of the cultivated *C. arabica* and the wild relative, *C. racemosa*.

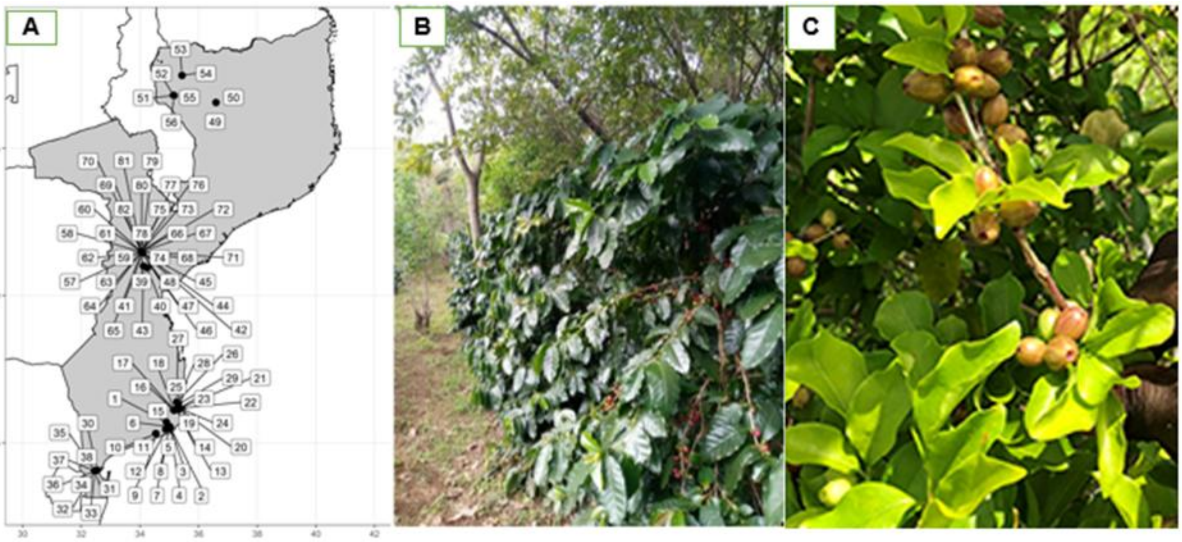


Figure 2.1 (A) Sampling of *Coffea* species in Mozambique. See Table S2.1 for the ID of samples. (B) Cultivation of *Coffea arabica* in the Gorongosa Mountain under agroforestry systems. (C) Wild plants of *Coffea racemosa*.

2.2. Materials and Methods

2.2.1. Plant Sampling and DNA Isolation

Thirty-five samples of *C. arabica* (27 from the Gorongosa Mountain, Sofala Province; and eight from Niassa Province) and 48 wild relatives collected in three provinces from central and southern Mozambique (Maputo, Inhambane, and Sofala) (Figure 2.2) were included in the analysis. This region gathered most of the historical collections recorded for *Coffea* in Mozambique. The study also included three additional commercial genotypes from Coffee Rust Research Center (CIFC), totaling 86 samples. The main variety of *C. arabica* cultivated in the Gorongosa Mountain is a commercial variety imported from Zimbabwe due to similar agroecological conditions in both countries. The cultivar is labeled Costa Rica (CR) and it is claimed to be tolerant to coffee leaf rust and coffee berry disease. The CIFC cultivars are certified hybrids of the Catimor group (CR-95). *C. arabica* is a tetraploid species with 2n = 44 chromosomes that usually behave genetically as diploid (Cubry *et al*., 2008) while *C. racemosa* is a diploid species with 2n = 22 chromosomes (Silva, 1956). In each site, 6 to 10 individuals were randomly collected with a minimum sampling distance of 10 m. Samples, locations, and geographic coordinates are shown in Table S2.1.

Fresh leaves were collected for each sample, dried on silica gel, and stored at −80 °C until DNA was extracted. Total genomic DNA was extracted from 100 mg of ground leaves using the InnuPrep Plant DNA kit (Analytik Jena Innuscreen GmbH, Jena, Germany) according to the manufacturer's protocol. Mean yield and purity were evaluated spectrophotometrically by readings of OD230, OD260, and OD280 (Nanodrop 2000, Thermo Fisher Scientific, Waltham, MA, USA) and visualized by 1% agarose gel electrophoresis under UV light.

2.2.2. Plastid Barcode Sequencing and Phylogenetic Analysis

Two barcode organelle regions (*rbcL* and *matK*) previously used in *Coffea* (Table S2.3) were first amplified to detect the degree of haplotype variation using the 86 samples. Polymerase chain reactions (PCR) were performed in 20 µL reactions using Biotaq DNA polymerase (Bioline, London, UK), 2X reaction buffer (Bioline, London, UK), 1 µM forward and reverse primers, 2 mM $MgCl_2$, and dNTPs 0.8 mM (Promega, Maddison, WI, USA), 0.2 U Taq Meridian Bioscience (MI, Italy), 0.28 mg/mL BSA, and 40 ng μL^{-1} of genomic DNA. Cycle sequencing reactions were carried out using the Bio-Rad PCR System MyCycler™ thermocycler.

The PCR program for *rbcL* consisted of 4 min at 94 °C followed by 35 cycles of 30 s at 94 °C, 1 min at 55 °C, and 1 min at 72 °C, with a final extension of 10 min at 72 °C. For *matK*, amplifications consisted of 5 min at 95 °C followed by 40 cycles of 30 sec at 95 °C, 30 s at 52 °C, and 1 min at 72 °C, with a final extension of 5 min at 72 °C. Amplified products were purified using QIAquick purification columns (QIAgen, Madrid, Spain), as described in the manufacturer's protocol, and sent for sequencing (Macrogen, Madrid, Spain). Consensus alignments for each gene were created in Geneious v.11.1.5 (Biomatters, Ltd., Auckland, New Zealand) using the MAFFT alignment algorithm v.7.450 (Katoh, 2013) and manually checked.

A phylogeny based on a maximum likelihood (ML) analysis was performed using the two plastid genes concatenated into a single matrix. Additionally, data from other representatives of the same species studied here, as well as representative outgroup taxa, were extracted from the NCBI database (Table S2.4). The best-fitting nucleotide substitution model was estimated using jModelTest2 v. 2.1.6 (Darriba *et al.*, 2012) (GTR) and used as input for RAxML v.8.2.12 with 1.000 bootstrap iterations (Stamatakis, 2014). In addition, genealogical haplotype relationships of the collected samples were inferred using the median-joining method in Popart v1.7 (Leigh *et al.*, 2015).

2.2.3. Single-Sequence Polymorphic Repeats

The 86 samples were genotyped at 14 nuclear single-sequence polymorphic repeats (SSRs) previously developed for *Coffea* (Table S2.5). Based on the initial research, we selected these 14 SSRs markers as they produced robust and highly polymorphic amplified bands across all collections of the samples under study. Amplifications were performed in 20 µL reaction volume containing 1 µM forward and reverse primers, 2X Buffer Meridian Bioscience (MI, Italy), 0.5 U of TAQ Meridian Bioscience (MI, Italy), and 40 ng µL−1 of genomic DNA on a Bio-Rad PCR System MyCycler™ thermocycler. Allele sizes were determined using Peak Scanner version 1.0 (Life Technologies, Carlsbad, CA, USA) and revised manually.

2.2.4. GBS Library Preparation, Sequencing, and SNP Calling

Genomic DNA (0.3~0.6 µg) of a subset of *Coffea* samples (29 total; Table S2.6) was double-digested using 10 µL of the restriction enzymes EcoRI and Mse I for 5 h at 37 °C, then 20 min at 65 °C, and final incubation at 12 °C. The resulting digested fragments were cleaned and subsequently quantified using agarose gel electrophoresis and the Qubit®2.0 fluorometer. Digested fragments were ligated to EcoR I and Mse I adapters containing sample-specific barcodes with T4 DNA ligase (NEB) for 4 h at 16 °C, then 20 min at 65 °C, and final incubation at 12 °C. Individually barcoded samples were cleaned and size-selected (350–500 bp) using agarose gel. After dilution to 1 ng µL−1, the Agilent®2100 bioanalyzer was used to assess insert size. Each library was then PCR-amplified to the desired concentration and paired-end sequenced on an Illumina®HiSeq PE150.

FastQC (Andrews, 2010) was used to remove adapters, and low-quality reads, e.g., uncertain nucleotides > 10% and base quality < 5 in more than 50% of either read, consistent with an error rate < 0.1%. Assembled reads were mapped against the reference genome of *C. arabica* downloaded from the NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCF_003713225.1, accessed on 4 April 2021) using BWA version 0.7.16 (Li *et al*., 2009) with the default parameters. The resulting individual files were converted into BAM files with SAMtools version 1.16.1 (Li *et al*., 2009), removing duplicate reads. Sequencing yielded a total of 170,720,052 raw reads, which were reduced to 51,115,669 after quality filtering (Table S2.6). Overall, an average of 79% of cleaned reads were mapped to the reference genome. Calling of variants (SNPs) was performed for the 28 sequenced *Coffea* samples using GATK 4.2.6.1 (Mckenna *et al*., 2010) with base quality score recalibration, indel realignment, duplicate removal, and performed SNP and INDEL discovery. Genotyping across samples was performed simultaneously using standard hard filtering parameters or variant quality score recalibration according to GATK Best Practices recommendations (Depristo *et al*., 2011). Filtering of SNPs included those with a sequencing depth of 3 to 50 for each sample and an average quality > 20. To exclude SNP calling errors caused by incorrect mapping or indels, two adjacent SNPs separated by <5 bp were not called. A total of 3,058,824 SNPs were found, including 1,461,205 intergenic SNPs, 185,956 intronic SNPs, 115,488 exonic SNPs, 7908 SNPs in splice site, 772,163 upstream, 449,139 downstream, 39,278 in UTR3, and 27,687 in UTR5. The location and annotation of SNPs were based on the data retrieved from the reference genome of *C. arabica* as mentioned above. Associated genes were mapped to the KEGG (Kaneisa *et al*., 2000) pathway and were examined if they were enriched in particular pathways based on the hypergeometric distribution test. Fisher's exact test was used to identify pathways significantly enriched ($p < 0.05$) with *Coffea* genes.

2.2.5. Genetic Diversity, Structure, and Differentiation

Since genetic data of the two species were diploidized (e.g., only a maximum of two alleles were found), we used the Bayesian program STRUCTURE v.2.3.4 (Pritchard, 2000) to test whether any discrete genetic structure existed among samples and species. The analysis was performed assuming K = 1 to K = 10 genetic clusters (K), with 10 repetitions per K. Models were run assuming ancestral admixture and correlated allele frequencies using run lengths of 200,000 interactions for each K after 50,000 burn-in steps. The optimum K value was determined using STRUCTURE HARVESTER (Earl *et al*., 2012), which identifies the optimal K based on both the posterior probability of the data for a given K and the ΔK (Evanno *et al*., 2000). The results of the replicates at the best-fit K were then post-processed using CLUMPAK (Kopelman *et al*., 2015). To visualize the genetic structure, a Principal Components Analysis (PCA) and a Neighbor-Joining (NJ) tree were constructed with 10,000 bootstraps in the adegenet R package (Jombart *et al*., 2011).

Differentiation between sites was analyzed by conducting an analysis of molecular variance (AMOVA) using Arlequin 3.5.2.2 (Excoffier *et al*., 2010). This approach is derived from the analysis of the variance framework based on Wright's fixation indices defined by (Cockerham, 1969). Pairwise differentiation between species and sites was also computed based on Nei's Genetic Distance and the coefficient of differentiation (Fst). Genetic diversity was assessed by calculating the number of alleles (Na), observed heterozygosity (Ho), expected heterozygosity (He), and fixation index (F), using diveRsity (Keenan *et al*., 2013) and PopPr R packages (Kamvar *et al*., 2014).

We also calculated linkage disequilibrium (LD) pruning the SNPs using Plink v1.9 (Purcell *et al*., 2007) with a window of 50 SNPs and a step size of five makers. PLINK was used to measure pairwise LD between multi-SNPs based on the allele frequency correlations. The LD decay plot was drawn using R (http://www.R-project.org/, accessed on 22 June 2021). Functional annotation of the SNPs was defined using the Blast2GO V5.0 tool (Gõtz *et al*., 2008) (E-value ≥ $1 \times 10{-}5$) implemented in the KEGG database (Kanehisa *et al*., 2000).

2.3. Results

2.3.1. Plastid Relationships among *Coffea* Samples

Two main maternal lineages were retrieved in phylogenetic analyses: one cluster containing all the samples of *C. arabica*, and the other all samples of *C. racemosa* (Figure 2.2). The cluster containing all *C. arabica* cultivars was phylogenetically apart from all *C. racemosa.* In contrast, three subclusters were retrieved within the *C. racemosa* lineage (Figure 2.2).

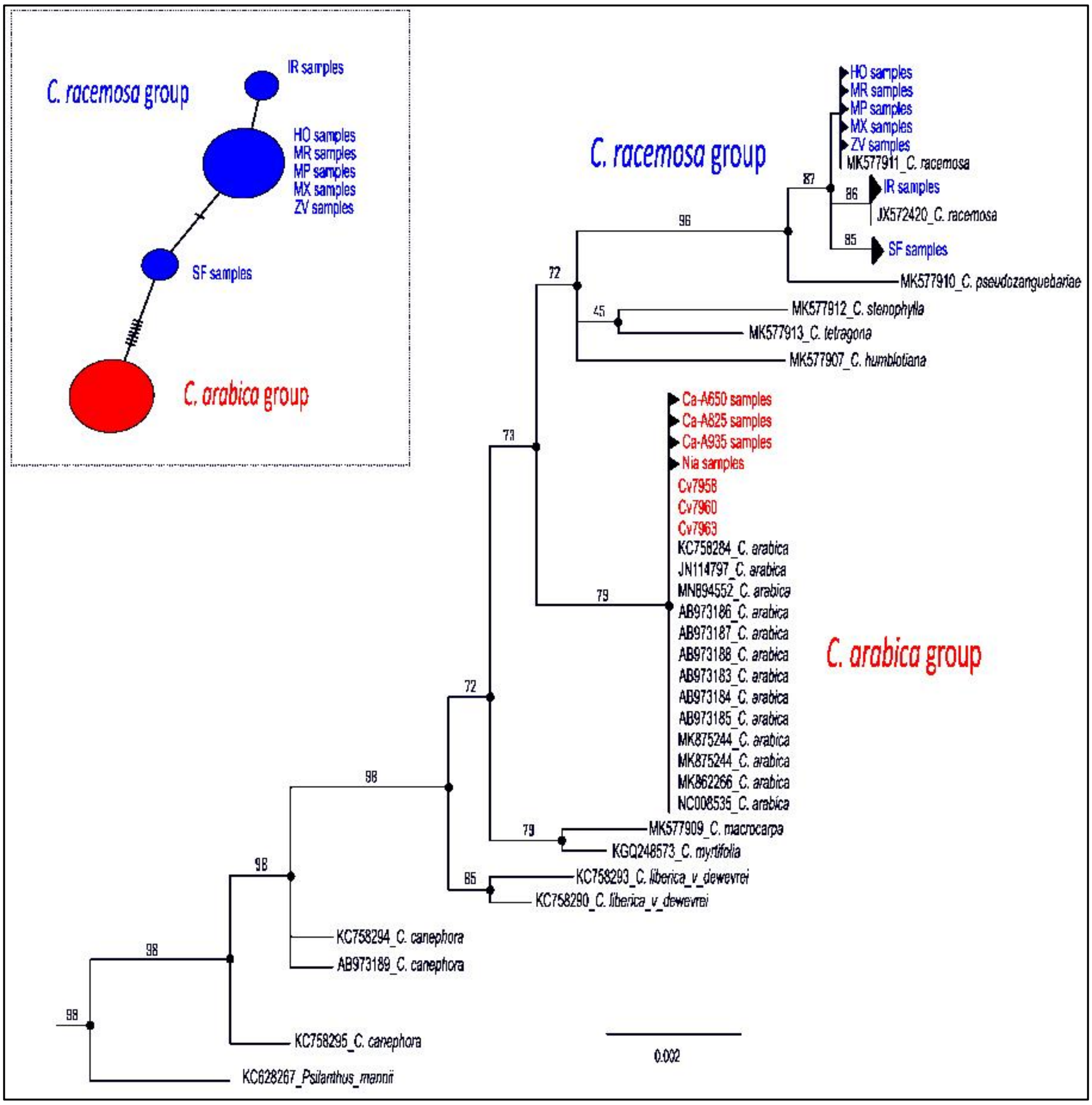


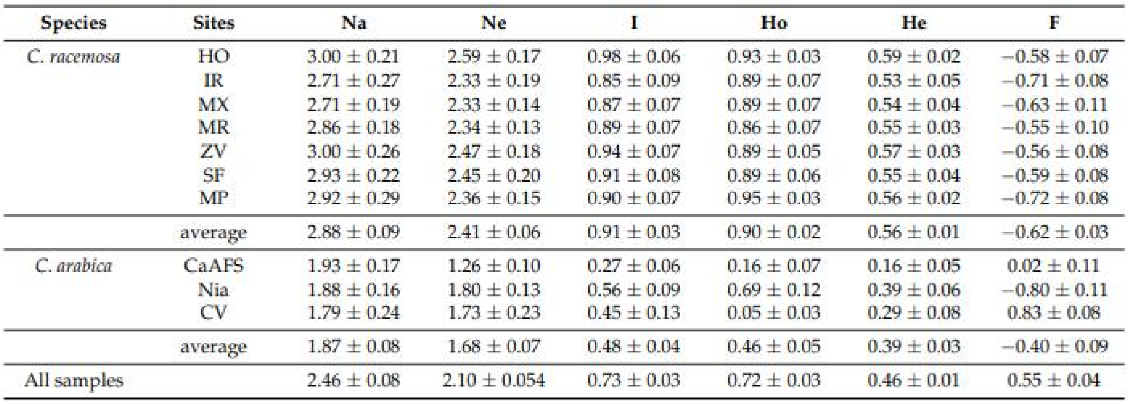
Figure 2.2 Plastid RaxML tree obtained for *Coffea*. Color codes indicate the two main groups of species: *C. arabica* (red) and *C. racemosa* (blue) included in this study. Bootstrap values higher than 50% are indicated above branches. Black codes indicate NCBI numbers and species retrieved from GenBank. Insert on the top right indicates the haplotype network using a median-joining method including only *Coffea* sampled for this study.

2.3.2. Genetic Diversity in *C. arabica* and the Wild Relative *C. racemosa*

Based on nuclear microsatellites, a total of 101 alleles were found among all samples: 56 in the set of *C. racemosa* samples and 41 in *C. arabica* samples or 62 when including also the three cultivars of *C. arabica* from the group Catimor. The average number of alleles and the levels of observed and expected heterozygosity were always lower in *C. arabica* than in *C. racemosa* (Table 2.1). The mean Shannon Information Index (I) varied from 0.459 among *C. arabica* to 0.905 in *C.*

racemosa and was particularly low in the *C. arabica* cultivars of the Gorongosa Agroforestry System (CaAFS) (0.268). The fixation index was negative in all *C. racemosa* accessions, as well as in the *C. arabica* cultivars from Niassa, while the cultivars from the CaAFS and CIFC collection showed positive values of fixation (Table 2.1). Estimates of genetic diversity based on SNPs revealed extremely low genetic diversity in *C. arabica* (Ho = 1.1± 0.02; He = 2.6 ± 0.9) when compared with *C. racemosa* samples (Ho = 2.9 ± 0.4; He = 3.1 ± 0.12).

Table 2.1 Estimates of genetic diversity based on SSRs for *C. arabica* and *C. racemosa*. HO: Homoine, IR: Inharrime, MX: Maxixe; MR: Morrumbene; ZV: Zavala (all from Inhambane Province); MP: Maputo Province; SF: Sofala Province; CaAFS: Gorongosa (Sofala Province); Nia: Niassa Province; CV: CIFC cultivars; Na: number of alleles; Ne: Number of effective alleles; I: Shannon's Information Index; Ho: observed heterozygosity; He: expected heterozygosity; F = fixation index.



Species	Sites	Na	Ne	I	Ho	He	F
C. racemosa	HO	3.00 ± 0.21	2.59 ± 0.17	0.98 ± 0.06	0.93 ± 0.03	0.59 ± 0.02	−0.58 ± 0.07
	IR	2.71 ± 0.27	2.33 ± 0.19	0.85 ± 0.09	0.89 ± 0.07	0.53 ± 0.05	−0.71 ± 0.08
	MX	2.71 ± 0.19	2.33 ± 0.14	0.87 ± 0.07	0.89 ± 0.07	0.54 ± 0.04	−0.63 ± 0.11
	MR	2.86 ± 0.18	2.34 ± 0.13	0.89 ± 0.07	0.86 ± 0.07	0.55 ± 0.03	−0.55 ± 0.10
	ZV	3.00 ± 0.26	2.47 ± 0.18	0.94 ± 0.07	0.89 ± 0.05	0.57 ± 0.03	−0.56 ± 0.08
	SF	2.93 ± 0.22	2.45 ± 0.20	0.91 ± 0.08	0.89 ± 0.06	0.55 ± 0.04	−0.59 ± 0.08
	MP	2.92 ± 0.29	2.36 ± 0.15	0.90 ± 0.07	0.95 ± 0.03	0.56 ± 0.02	−0.72 ± 0.08
	average	2.88 ± 0.09	2.41 ± 0.06	0.91 ± 0.03	0.90 ± 0.02	0.56 ± 0.01	−0.62 ± 0.03
C. arabica	CaAFS	1.93 ± 0.17	1.26 ± 0.10	0.27 ± 0.06	0.16 ± 0.07	0.16 ± 0.05	0.02 ± 0.11
	Nia	1.88 ± 0.16	1.80 ± 0.13	0.56 ± 0.09	0.69 ± 0.12	0.39 ± 0.06	−0.80 ± 0.11
	CV	1.79 ± 0.24	1.73 ± 0.23	0.45 ± 0.13	0.05 ± 0.03	0.29 ± 0.08	0.83 ± 0.08
	average	1.87 ± 0.08	1.68 ± 0.07	0.48 ± 0.04	0.46 ± 0.05	0.39 ± 0.03	−0.40 ± 0.09
All samples		2.46 ± 0.08	2.10 ± 0.054	0.73 ± 0.03	0.72 ± 0.03	0.46 ± 0.01	0.55 ± 0.04

2.3.3. Genetic Structure

Microsatellite data based on SSRs revealed a total of five genetic clusters among all samples, based on the highest LnP (D) and ΔK values obtained in STRUCTURE HARVESTER (Figure S2.1). The different genetic membership retrieved divided *C. racemosa* samples from Sofala and Maputo vs. Inhambane (HO, IR, MX, MR, and ZV) provinces (Figure 2.3A). In *C. arabica*, genetic memberships divided samples from the Gorongosa CaAFS, Niassa (Nia), and the three cultivars from CIFC included in this study (CV) (Figure 3A). However, GBS data retrieved one single genetic membership per species, which segregated *C. racemosa* from *C. arabica* samples (Figure 2.3B).

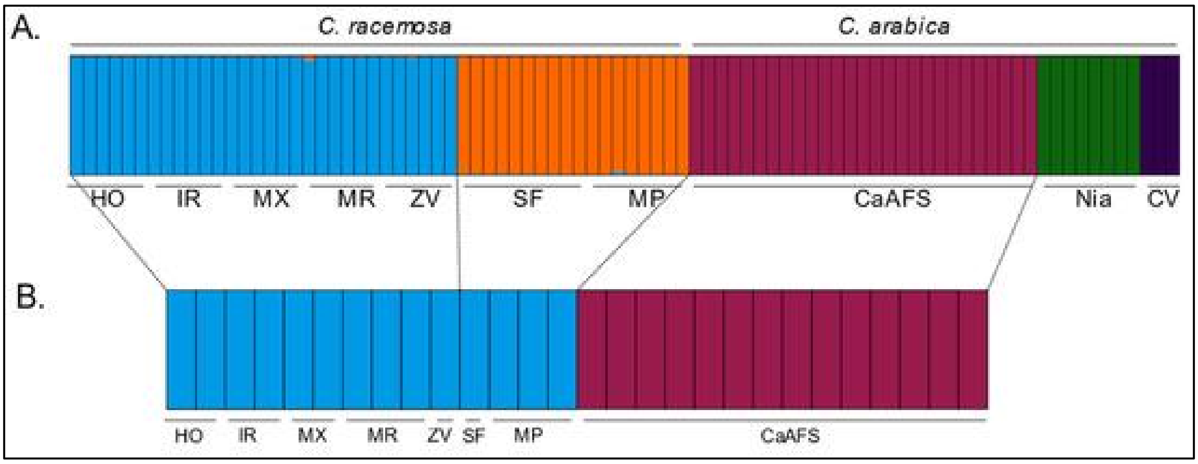


Figure 2.3 Genetic structure of *C. racemosa* and *C. arabica*. Genetic clusters are based on the best assignment group found for SSRs (**A**; K = 5) and GBS markers (**B**; K = 2). Colors indicate an assignment probability, according to different genetic clusters. Each sample is represented by a vertical bar. HO: Homoine, IR: Inharrime, MX: Maxixe; MR: Morrumbene; ZV: Zavala (all from Inhambane Province); MP: Maputo Province; SF: Sofala Province; CaAFS: Gorongosa (Sofala Province); Nia: Niassa Province; CV: CIFC cultivars.

No genetic admixture was detected between samples, either using SSRs or GBS data. Linkage disequilibrium (LD) was overall low in *C. racemosa* but significantly higher in *C. arabica* (Figure S2.2). In both species, LD values did not change significantly ($p > 0.05$) with the increasing physical distance of SNPs.

Results were generally compatible with the topology of NJ trees and the PCA patterns, which also isolated *C. racemosa* from *C. arabica* samples (Figure 2.4). It is worth highlighting that Sofala and Maputo are segregated from Inhambane in STRUCTURE, which can also be observed in the NJ tree and PCA from SSR data (Figure 2.4A,C) but are not well discriminated using the GBS data (Figure 2.4B,D).

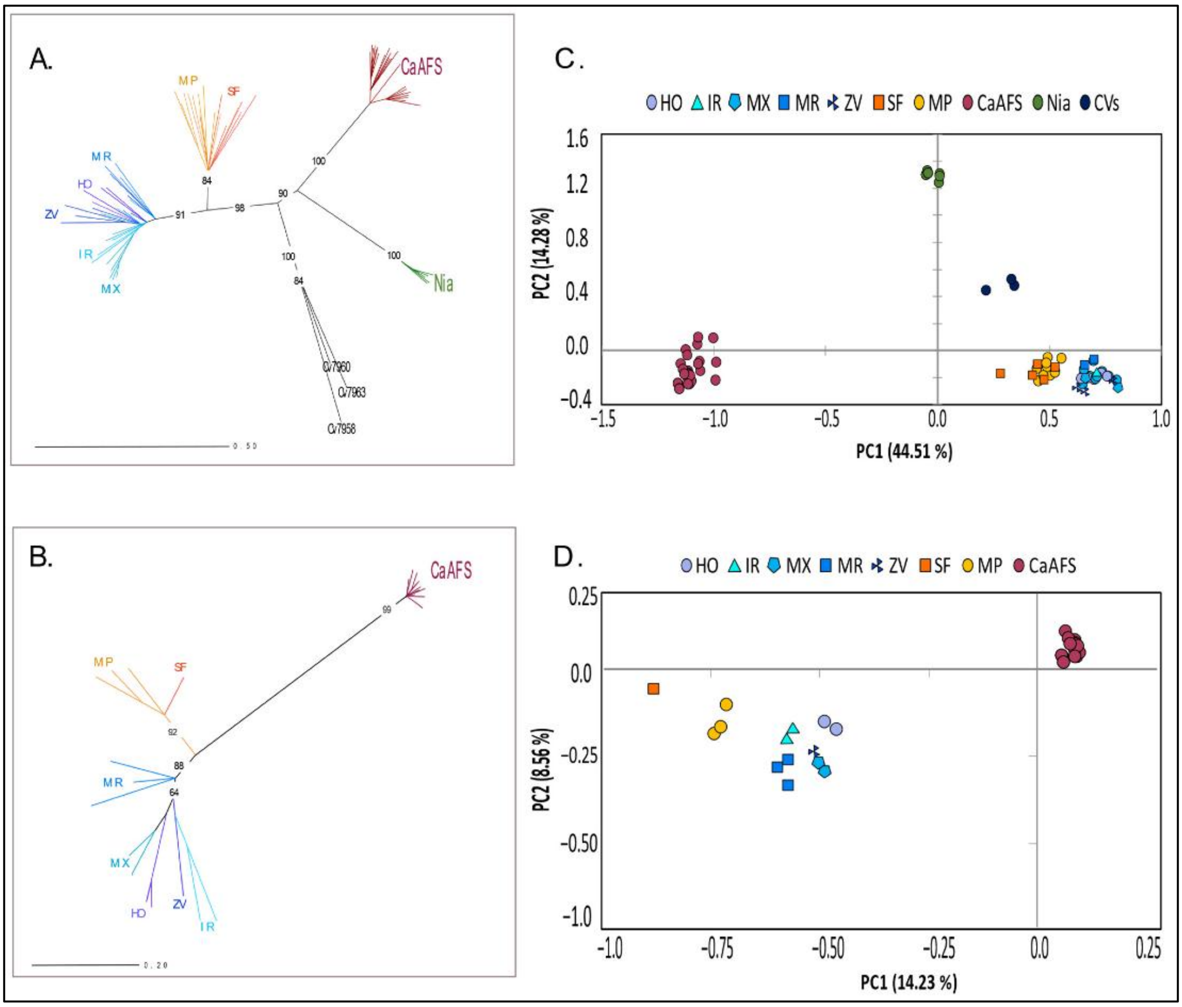


Figure 2.4 Genetic relationships among *C. racemosa* and *C. arabica* samples based on Nei's Genetic Distance. Unrooted Neighbor-Joining (NJ) tree showing relationships among the sampled individuals using the scored nSSRs markers (A) and GBS data (B). Numbers associated with branches indicate bootstrap values >50 based on 1000 replications. Principal Coordinate Analysis (PCA) scatterplots using the scored nSSRs markers (C) and GBS data (D). The percentage of explained variance of each axis is given in parentheses. HO: Homoine, IR: Inharrime, MX: Maxixe; MR: Morrumbene; ZV: Zavala (all from Inhambane Province); MP: Maputo Province; SF: Sofala Province; CaAFS: Gorongosa (Sofala Province); Nia: Niassa Province; CV: CIFC cultivars.

2.3.4. Genetic Differentiation between Species and Sites

Overall, genetic differentiation was significantly high (AMOVA FST = 0.5044, PHI = 0.673, $p < 0.001$). The variance was equally attributed among the $K = 5$ groups found by STRUCTURE (50.45% and 67.32%), and within sites (49.55%; 32.67%) based on SSR and GBS data, respectively.

A large genetic divergence was found between *C. arabica* and *C. racemosa* either using pairwise genetic differences of SSRs based on Nei's Genetic Distance or using

Fst values from GBS (Figure 2.5). It is worth mentioning the large range of divergence found between *C. arabica* cultivars from Gorongosa and the ones collected in Niassa, and even with the CIFC cultivars, where the highest level of divergence was found. Genetic distances were lower between the cultivars of *C. racemosa* sampled in Sofala and Maputo than the ones from Inhambane (HO, IR, MX, MR, and ZV), supporting the previously reported results of genetic structure.

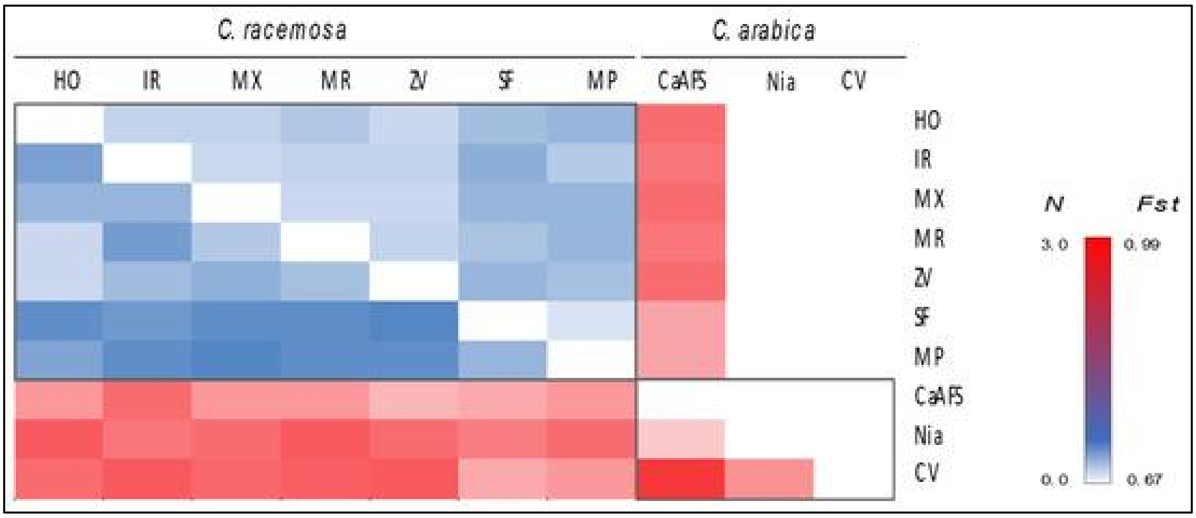


Figure 2.5 Pairwise differentiation between sites based on Nei's Genetic Distance using SSRs (below diagonal) and sites differentiation coefficient (Fst) values from GBS (above diagonal) in *C. racemosa* and *C. arabica*. HO: Homoine, IR: Inharrime, MX: Maxixe; MR: Morrumbene; ZV: Zavala (all from Inhambane Province); MP: Maputo Province; SF: Sofala Province; CaAFS: Gorongosa (Sofala Province); Nia: Niassa Province; CV: CIFC cultivars.

2.3.5. Annotation and Functional Pathways of SNPs

Sequencing yielded a total of 170,720,052 raw reads, which were reduced to 51,115,669 after quality filtering (Table S2.6). Overall, an average of 79% of cleaned reads were mapped to the reference genome. A total of 3,058,824 SNPs were found, including 1,461,205 intergenic SNPs, 185,956 intronic SNPs, 115,488 exonic SNPs, 7908 SNPs in splice site, 772,163 upstream, 449,139 downstream, 39,278 in UTR3, and 27,687 in UTR5 (Figure S2.2). In both species, SNPs were involved in 191 KEGG pathways (Table S2.2) being top-regulated: 'Plant–pathogen interaction', 'Protein processing in endoplasmic reticulum', and 'Phenylpropanoid biosynthesis' (Figure 2.6). Only three pathways showed significant differences between the two species: the 'Plant–pathogen interaction' ($F_{2,1} = 25.892$, $p < 0.05$) and the 'Plant hormone signal transduction' ($F_{2,21} = 26.034$, $p < 0.05$) had more SNPs involved in *C.*

racemosa than in *C. arabica* while 'Amino sugar and nucleotide sugar metabolism' was higher in *C. arabica* than in *C. racemosa* ($F_{2,23}$ = 22.056, $p < 0.05$; Figure 2.6). Interestingly, SNPs linked with '*Caffeine* metabolism' (Table 2.S2) showed no significant differences between the two cultivars ($F_{2,23}$ = 1.741, $p > 0.05$).

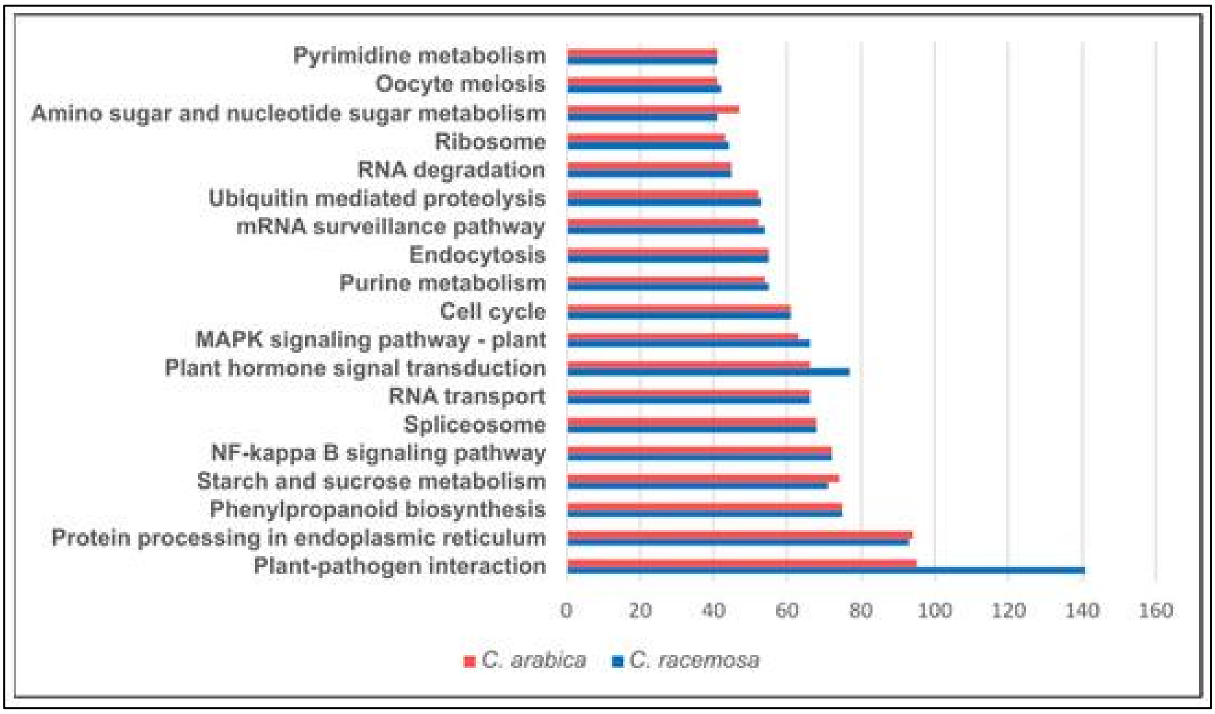


Figure 2.6 Top KEGG pathways involved among *C. arabica* and *C. racemosa*.

2.4. Discussion

2.4.1. Assessment of *Coffea* Species in Southern and Central Mozambique Using Plastid Markers

Despite the fact that other wild relatives are supposed to occur in Mozambique such as *C. zanguebariae*, for which herbarium data suggested a sparse distribution in these regions (Bridson,1998; Bridson, 2003; Davis *et al*., 2006; Davis *et al*., 2009), our field expeditions (Figure 2.1) found only one wild *Coffea* species, *C. racemosa* (known also as Inhambane coffee). Molecular analysis based on plastid markers congruently found two main maternal lineages, splitting this species from *C. arabica* (Figure 2.2). The cluster grouping *C. arabica* cultivars was phylogenetically apart from *C. racemosa*, supporting a single maternal origin scenario for each species (Wicke *et al*., 2011). In contrast, three subclusters were retrieved within the *C. racemosa* lineage, suggesting different origins for this species in Mozambique (Figure 2.2).

One explanation for the absence of other wild relatives in these areas could be the genetic drift of *C. zanguebariae* from southern and central Mozambique due to environmental and anthropogenic pressure. Indeed, during our expeditions, we could not validate many historical herbaria locations (personal observations).

This was not unexpected and might be interconnected with the fact that (i) Mozambique is among the most disaster-prone countries on a global scale (https://drmkc.jrc.ec.europa.eu/inform-index (Accessed on 10 March 2023), and has gone through a series of natural shocks over the last decades, e.g., the flooding of 2000 and 2017, the Earthquake of 2006, the cyclone Favio in 2007, Idai and Kenneth in 2019, the storm Dando in 2012, or the current tropical storm Freddy that hazards the country as we write this article (https://sicnoticias.pt/arquivo/ciclone-idai/2019-03-22-Os-desastres-naturais-em-Mocambique-desde-as-cheias-de-2000 (accessed on 10 March 2023; https://reliefweb.int/country/moz (accessed on 10 March 2023); (ii) the related resettlements of local communities; and (iii) the dynamics of land use and land cover (Cloves *et al*., 2020).

The second possibility is that the species' identity has been mistaken in the past. *C. racemosa* has been reported as endemic to southern and central Mozambique, distributed across coastal and riverine forests as well as deciduous woodlands and bushlands (0 to 500 m above sea level—a.s.l.), while *C. zanguebariae* was considered endemic to northern Mozambique, distributed across dry deciduous forests and riverine and coastal thickets (0 to 350 m a.s.l.) (Bridson, 1998; Bridson, 2003).

On the other hand, despite the taxonomic advances to discriminate the two species, *C. racemosa* and *C. zanguebariae* are in fact so similar that they have often been believed to be the same, and only recently, DNA markers allowed accurate species discrimination (Davis *et al*., 2021). It is, thus, possible that many records have misidentified these species although further field expeditions should be done in the north of Mozambique.

2.4.2. Low Genetic Diversity in *C. arabica* in Comparison with the Wild Relative *C. racemosa*

C. arabica presented very low levels of genetic diversity in comparison with *C. racemosa* (Table 2.1). In our study, the mean number of alleles (Na) and effective alleles (Ne) was consistently below two in the cultivars from the three provenances, Gorongosa, Niassa, and CIFC. This value is much lower than those reported in other studies based on SSR markers, which ranged from ca. 3 to 6 (Anthony *et al*., 2002; Moncada *et al*., 2004; Maluf *et al*., 2005; Dida *et al*., 2021). However, the observed heterozygosity (Ho) and expected heterozygosity (He) values were within the expected range considering the low genetic values usually reported in microsatellite studies of *C. arabica* (Combes *et al*., 2000; Dida *et al*., 2021).

In addition, our estimates of genetic diversity based on SNPs also revealed extremely low values in *C. arabica* when compared with *C. racemosa*. In fact, the overall congruent results found between SSR and SNP data suggest that these independent markers can detect similar patterns of genetic diversity. However, between the two markers, SSRs remain the most cost-effective and rapid marker being widely used in most genetic population studies.

The low levels of heterozygosity are likely due to the autogamous nature of *C. arabica* (Aggarwal *et al*., 2007; Aerts *et al*., 2013., Gadissa *et al*., 2018; Dida *et al*., 2021) and the single polyploidization event at the origin of the tetraploid genome of this species, which was probably narrowed further in some cultivars of this species (Scalabrini *et al*., 2020). Indeed, the Shannon diversity index (I) of all *C. arabica* cultivars used in this study was very low (0.48 on average), supporting the genetic bottleneck hypothesis in commercial Arabica varieties (Moncada *et al*., 2004). In contrast, the consistently high diversity levels observed in *C. racemosa* are likely due to the allogamy of the species (Filho *et al*., 1977). These genetic diversity results were comparable to those reported in other tropical trees, such as *Warburgia salutaris* from southern Mozambique (Senkoro *et al*., 2020) or *Ocotea rotundata* from the northern Andean forests (Marques *et al*., 2022), suggesting that *C. racemosa* retains high levels of genetic diversity, especially when compared with *C. arabica*.

The absence of gene flow between the two species would explain the finding of different genetic clusters (and the absence of genetic admixture) that segregated all *C. racemosa* from the *C. arabica* sample, either when considering STRUCTURE results (Figure 2.3), the topology of NJ trees, or PCA patterns (Figure 2.4). The large genetic divergence found between *C. arabica* and *C. racemosa* is also supported by the pairwise genetic differences of SSRs based on Nei's Genetic Distance and the

Fst values from GBS (Figure 2.5). It is also worth mentioning the large range of divergence found between *C. arabica* cultivars from Gorongosa and the ones implemented in Niassa, and even with the CIFC cultivars, where the highest level of divergence was found. Gene flow that usually results from pollen and seed migration plays a significant role in preventing genetic differentiation among populations while contributing to the conservation of genetic diversity (Li *et al*., 2018).

The autogamous nature of the cultivated *C. arabica* contributes to such differentiation and is a concern in light of environmental changes. In contrast, pollination by birds or insects and the dispersion of seeds are likely to occur in *C. racemosa*, contributing to the patterns of genetic diversity and structure found in this study. Additionally, although hybrids between the diploid wild *C. racemosa* and the tetraploid cultivated *C. arabica* would be possible, hybrid triploid plants are expected to be infertile (Husband, 2004).

As a wild relative of coffee and despite the important role that *C. racemosa* might have to implement sustainable changes in the coffee sector, fundamental basic data, such as the type of breeding system, and the type of pollinators and dispersers involved in this species are unknown. This is particularly significant as we found differences in the functional pathways of these two species that could reflect differences in the tolerance to environmental stresses (Figure 2.6). For instance, the '*Plant–pathogen interaction*' and '*Plant hormone signal transduction*' were significantly more represented in *C. racemosa*, suggesting that the signaling pathways related to stress tolerance are more robust in this species (Botero *et al*., 2018).

On the other hand, *C. arabica* was more enriched in SNPs involved in amino sugar and nucleotide sugar metabolism than *C. racemosa*. Some enzyme proteins in these pathways are also involved in stress response in plants and thus, a greater number of genes from these pathways may be redundant, as they are important to maintain pivotal functions, including cell wall synthesis and cell repair processes (e.g., associated with pectin synthesis) (Li *et al*., 2018; Zhu *et al*., 2020). Interestingly, SNPs linked with the '*Caffeine metabolism'* showed no significant differences between the two cultivars (Table S2.2), even though *C. racemosa* is sought as a "naturally decaffeinated" bean due to its low levels of caffeine (Hamon *et al*., 2017; Davis *et al*., 2021). Based on local records, this species is thought to produce an aromatic drink with low caffeine levels (Lains & Silva, 1954). This highlights the need for more studies on the functional traits of *C. racemosa*, namely the quality of its beans.

2.4.3. Implications for the Management of the Gorongosa Agroforestry System

The use of AFS in Gorongosa Nacional Park is seen as a promising and compatible approach to help adapt to climate change while reconciling biodiversity conservation and local development (Jezeer *et al*., 2017; Cassamo *et al*., 2022; 2023). The fact that coffee originates from high-altitude forest regions and can develop in shady areas (Davis *et al*., 2006), together with the historical context of coffee in Mozambique, were the main reasons for the implantation of this system in the Gorongosa Mountains. However, our results showed very low genetic diversity values in *C. arabica* plants used in the AFS, which could affect the long-term sustainability of this system.

In this context, the introduction of new Arabica cultivars in the Gorongosa AFS would be an asset, given the low genetic diversity of the implanted cultivar. More efforts involving coffee producers should be developed to create awareness of the importance of conserving *C. racemosa*. Additionally, attempts to introduce *C. racemosa* and other crop-wild relatives into the value chain should be placed on the agenda. Wild coffee species are already being farmed in Kwa Zulu Natal in South Africa (Hluhluwe) (Davis *et al*., 2021), and although yields are low when compared with the widely used *C. arabica* and *C. canephora*, their specific attributes regarding environmental stress tolerance and flavor could be useful for new blends (Davis *et al*., 2021). Recognizing the global socioeconomic importance of coffee, particularly in many developing countries that largely depend on this commodity, it is vital to innovate the coffee value chain, introducing more resilient species, increasing genomic diversity, and adopting climate-smart practices.

2.5. Conclusion

Here, we show for the first time how genetic diversity is needed to assure sustainable agriculture practices. Even though the implementation of AFSs is of interest since they can offset deforestation in tropical environments, while increasing biodiversity, productivity, social profitability, and environmental stewardship, guaranteeing the genetic diversity of the species/cultivars is an essential condition to ensure the long-term sustainability of AFSs. In this sense, the introduction of crop wild relatives in coffee AFSs provides an opportunity to increase the productivity and resilience of agricultural systems as they contain useful genetic diversity, which as reported here is not present in cultivated Arabica coffee.

2.6. References

Aerts, R., Berecha, G., Gijbels, P., *et al.* Genetic variation and risks of introgression in the wild *Coffea arabica* gene pool in south-western Ethiopian montane rainforests. *Evol Appl*. 2013, 6, 243–252, https://doi.org/10.1111/j.1752-4571.2012.00285.x.

Aggarwal, R.K., Hendre, P.S., Varshney, R.K., *et al*. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor Appl Genet.* 2007, *114*, 359–372, https://doi.org/10.1007/s00122-006-0440-x.

Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: http://www.bioinformatics.babraham.ac.uk/projects/fastqc (accessed on 15 May 2023).

Anthony, F., Combes, M.C., Astorga, C., *et al*. The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theor Appl Genet.* 2002, 104, 894–900,https://doi.org/10.1007/s00122-001-0798-8.

Botero, D., Alvarado, C., Bernal, A., *et al*. Network analyses in plant pathogens. *Front Microbiol.* 2018, 9, 35, https://doi.org/10.3389/fmicb.2018.00035.

Bridson, D.M. Coffea. In: Polhill, R.M., Bridson, D.M., Verdcourt, B. (eds). *Brookfield Balkema*: Rotterdam. 1998, 703–723.

Bridson, D.M. *Coffea*. In: Pope, G.V. (ed). *Royal Botanic Gardens*, Kew, London. 2003, 452–463.

Byrareddy, V., Kouadio, L., Mushtaq, S., *et al*. Coping with drought: Lessons learned from robusta coffee growers in Vietnam. *Clim Serv.* 2021, 22, 100229, https://doi.org/10.1016/j.cliser.2021.100229.

Cassamo, C.T., Draper, D., Romeiras, M.M., *et al*. Impact of climate changes in the suitable areas for *Coffea arabica* L. production in Mozambique: Agroforestry as an alternative management system to strengthen crop sustainability. *Agric Ecosyst Environ.* 2023, 346, 108341, https://doi.org/10.1016/j.agee.2022.108341.

Cassamo, C.T., Mangueze, A.V.J., Leitão, A.E., *et al*. Shade and altitude implications on the physical and chemical attributes of green coffee beans from Gorongosa Mountain, Mozambique. *Agronomy.* 2022, 12, 2540, https://doi.org/10.3390/agronomy12102540.

Chemura, A., Mudereri, B.T., Yalew, A.W., *et al*. Climate change and specialty coffee potential in Ethiopia. *Sci Rep.* 2021, 11, 8097, https://doi.org/10.1038/s41598-021-87647-4.

Clovis, G., Telina, R., Frederique, M., *et al*. Land-use and land-cover change 2000-2016 in Mozambique. *Zenodo*. 2020, available at https://doi.org/10.5281/zenodo.4242928 (accessed on 15 May 2023).

Cockerham, C.C. Variance of gene frequencies. *Evol.* 1969, 23, 72–84.

Combes, M.C., Andrzejewski, S., Anthony, F., *et al*. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. *Mol Ecol.* 2000, 9,1178–1180, https://doi.org./10.1046/j.1365-294x.2000.00954-5.x.

Cubry, P., Pascal, M., Hyacinthe, L., *et al*. Diversity in coffee assessed with SSR markers: Structure of the genus *Coffea* and perspectives for breeding. *Genome.* 2008, 51, 50–63, https://doi.org./10.1139/g07-096.

DaMatta, F.M., Ramalho, J.D.C. Impacts of drought and temperature stress on coffee physiology and production: A review. *Braz J Plant Physiol.* 2006, 18, 55–81, https://doi.org/10.1590/S1677-04202006000100006.

DaMatta, F.M., Rahn, E., Läderach, P., *et al*. Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? *Clim Chang.* 2019, 152, 167–178, https://doi.org/10.1007/s10584-018-2346-4.

Darriba, D., Taboada, G.L., Doallo, R., *et al*. J Model Test 2: More models, new heuristics and parallel computing. *Nat Methods*. 2012, 9, 772, https://doi.org/10.1038/nmeth.2109.

Davis, A.P., Chadburn, H., Moat, J., *et al*. High extinction risk for wild coffee species and implications for coffee sector sustainability. *Sci Adv.* 2019, 5, 3473–3489, https://doi.org/10.1126/sciadv.aav3473.

Davis, A.P., Chester, M., Maurin, O., *et al.* Searching for the relatives of *Coffea* (Rubiaceae, Ixoroideae): The circumscription and phylogeny of *Coffeae* based on plastid sequence data and morphology. *Am J Bot.* 2007, 94, 313–329, https://doi.org/10.3732/ajb.94.3.313.

Davis, A.P., Gargiulo, R., Almeida, I.N.d.M., *et al*. Hot Coffee: The identity, climate profiles, agronomy, and beverage characteristics of *Coffea racemosa* and *Coffea zanguebariae*. *Front Sustain Food Syst.* 2021, 5, 740137, https://doi.org/10.3389/fsufs.2021.740137.

Davis, A.P., Gole, T.W., Baena, S., *et al*. The impact of climate change on indigenous Arabica Coffee (*Coffea arabica*): Predicting future trends and identifying priorities. *PLoS One.* 2012, *7*, e47981, https://doi.org/10.1371/journal.pone.0047981.

Davis, A.P., Govaerts, R., Bridson, D.M., *et al*. A Global assessment of distribution, diversity, endemism, and taxonomic effort in the rubiaceae. *Ann Missouri Bot Gard.* 2009, 96, 68–78, https://www.jstor.org/stable/40390048.

Davis, A.P., Govaerts, R., Bridson, D.M., *et al*. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Bot J Linn Soc.* 2006, 152, 465–512, http://dx.doi.org/10.1111/j.1095-8339.2006.00584.x.

Davis, A.P., Tosh, J., Ruch, N., *et al*. Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Bot J Linn Soc.* 2011, 167, 357–377, https://doi.org/10.1111/j.1095-8339.2011.01177.x.

De Leijster, V., Santos, M.J., Wassen, M.W., *et al*. Ecosystem services trajectories in coffee agroforestry in Colombia over 40 years. *Ecosyst Serv.* 2021, 48, 101246, https://doi.org/10.1016/j.ecoser.2021.101246.

de Oliveira, R.R., Ribeiro, T.H.C., Cardon, C.H., *et al.* Elevated temperatures impose transcriptional constraints and elicit intraspecific differences between coffee genotypes. *Front Plant Sci.* 2020, 11, 1113, https://doi.org/10.3389/fpls.2020.01113.

DePristo, M.A., Banks, E., Poplin, R., *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011, 43, 491-498, https://doi.org/0.1038/ng.806.

Dida, G., Bantte, K., Disasa, T. *et al*. Molecular characterization of Arabica Coffee (*Coffea arabica* L.) germplasms and their contribution to biodiversity in Ethiopia. *Plant Biotechnol Rep.* 2021, 15, 791–804, https://doi.org/10.1007/s11816-021-00721-1.

Dubberstein, D., Lidon, F.C., Rodrigues, A.P., *et al*. Resilient and sensitive key points of the photosynthetic machinery of *Coffea* spp. to the single and superimposed exposure to severe drought and heat stresses. *Front. Plant Sci.* 2020, 11, 1049, https://doi.org/10.3389/fpls.2020.01049.

Earl, D.A., von Holdt, B.M. Structure Harvester: A website and program for visualizing structure output and implementing the evanno method. *Conserv Genet Resour.* 2012, 4, 359–361, https://doi.org/10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., Goudet, J., *et al*. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol.* 2005, 14, 2611–2620, https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Excoffier, L., Lischer, H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010, 10, 564–567, https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Fernandes, I., Marques, I., Paulo, O.S., *et al*. Understanding the impact of drought in *Coffea* genotypes: Transcriptomic analysis supports a common high resilience to moderate water deficit but a genotype dependent sensitivity to severe water deficit. *Agronomy.* 2021, 11, 2255, https://doi.org/10.3390/agronomy11112255.

Filho, H.P.M., Carvalho, A., Medina, D.M., *et al*. Germoplasma de *Coffea racemosa* e seu potencial de melhoramento do cafeeiro. *Bragantia.* 1977, 36, 43–46, https://doi.org/10.1590/S0006-87051977000100040.

Gadissa, F., Tesfaye, K., Dagne, K., *et al*. Genetic diversity and population structure analyses of *Plectranthus edulis* (Vatke) Agnew collections from diverse agro-ecologies in Ethiopia using newly developed EST-SSRs marker system. *BMC Genet.* 2018, 19, 92, https://doi.org/10.1186/s12863-018-0682-z.

Gidey, T., Oliveira, T.S., Crous-Duran, J., *et al*. Using the yield-SAFE model to assess the impacts of climate change on yield of coffee (*Coffea arabica* L) under agroforestry and monoculture systems. *Agrofor Syst.* 2020, 94, 57–70, https://doi.org/10.1007/s10457-019-00369-5.

Gomes, L.C., Bianchi, F.J.J.A., Cardoso, I.M., *et al*. Agroforestry systems can mitigate the impacts of climate change on coffee production: A spatially explicit assessment in Brazil. *Agric Ecosyst Environ.* 2020, 294, 106858, https://doi.org/10.1016/j.agee.2020.106858.

Götz, S., García-Gómez, J.M., Terol, J., *et al*. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 2008, 36, 3420–3435, https://doi.org/10.1093/nar/gkn176.

Guerreiro-Filho, O. *Coffea racemosa* Lour. *Une revue. Café, Cacao, Thé*. 1992, 36, 171–186, https://agritrop.cirad.fr/408807/.

Davis, A.P., Mieulet, D., Moat, J., *et al*. Arabica-like flavour in a heat-tolerant wild coffee species. *Nat Plants.* 2021, *7*, 413–418.

Halle, F., Faria, M.T. Le *Coffea racemosa* Lour. *Agron Moçamb.* 1973, *7*, 243–250.

Hamon, P., Grover, C.E., Davis, A.P., *et al*. Genotyping-by-sequencing provides the first well-resolved phylogeny for coffee (*Coffea*) and insights into the evolution of caffeine content in its species: GBS coffee phylogeny and the evolution of caffeine content. *Mol Phylogenet Evol*. 2017, 109, 351-361, https://doi:.org/10.1016/j.ympev.2017.02.009.

Husband, B.C. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biol J Linn Soc.* 2004, 82, 537–546. https://doi.org/10.1111/j.1095-8312.2004.00339.x.

ICO - International Coffee Organization. What’s New. 2022 Available online: https://www.ico.org/ (accessed on 22 November 2022).

Jaramillo, J., Muchugu, E., Vega, F.E., *et al*. Some like it hot: The influence and implications of climate change on coffee berry borer (*Hypothenemus hampei*) and coffee production in East Africa. *PloS One.* 2011, 6, e24528, https://doi.org/10.1371/journal.pone.0024528.

Jezeer, R.E., Verweij, P.A., Santos, M.J., *et al*. Shaded coffee and cocoa—double dividend for biodiversity and small-scale farmers. *Ecol Econ.* 2017, 140, 136–145, https://doi.org/10.1016/j.ecolecon.2017.04.019.

Jombart, T., Ahmed, I. Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*. 2011, 27, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521.

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., *et al*. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *Peer J*. 2014, e281, https://doi.org/10.7717/peerj.281.

Kanehisa, M., Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000, 28, 27–30, https://doi.org/10.1093/nar/28.1.27.

Katoh, K., Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol.* 2013, 30, 772–780, https://doi.org/10.1093/molbev/mst010.

Keenan, K., Mcginnity, P., Cross, T.F., *et al*. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol.* 2013, 4, 782–788, https://doi.org/10.1111/2041-210X.12067.

Kopelman, N.M., Mayzel, J., Jakobsson, M., *et al*. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour.* 2015, 15, 1179–1191, https://doi.org/10.1111/1755-0998.12387.

Koutouleas, A., Sarzynski, T., Bertrand, B., *et al*. Shade effects on yield across different *Coffea arabica* cultivars – how much is too much? A meta-analysis. *Agron Sustain Dev*. 2022b, 42, 55, https://doi.org/10.1007/s13593-022-00788-2.

Koutouleas, A., Sarzynski, T., Bordeaux, M., *et al*. Shaded-coffee: A nature-based strategy for coffee production under climate change? A review. *Front Sustain Food Syst*. 2022a, 6, 158, https://doi.org/10.3389/fsufs.2022.877476.

Läderach, P., Ramirez–Villegas, J., Navarro-Racines, C., *et al*. Climate change adaptation of coffee production in space and time. *Clim Chang.* 2017, 141, 47–62, https://doi.org/10.1007/s10584-016-1788-9.

Laíns e Silva, H.J. Subsídios para o fomento da cultura do Café em Moçambique. *Agência Geral do Ultramar*. 1954, 10792-L22.

Lara-Estrada, L., Rasche, L., Schneider, U.A., *et al*. Land in Central America will be comeless suitable for coffee cultivation under climate change. *Reg Environ Chang.* 2021, 21, 88, https://doi.org/10.1007/s10113-021-01803-0.

Leigh, J.W., Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol Evol.* 2015, 6, 1110–1116, https://doi.org/10.1111/2041-210X.12410.

Li, H., Durbin, R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics.* 2009, 15, 25, 1754-1760, https://doi.org/10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., *et al*. The sequence alignment/Map format and SAMtools. *Bioinformatics*. 2009, 25, 2078–2079, https://doi.org/10.1093/bioinformatics/btp352.

Li, S., Gan, X., Han, H., *et al*. Low within-population genetic diversity and high genetic differentiation among populations of the endangered plant *Tetracen tronsinense* Oliver revealed by inter-simple sequence repeat analysis. *Ann For Sci.* 2018, 75, 74, https://doi.org/10.1007/s13595-018-0752-4.

Maluf, M.P., Silvestrini, M., de Campos Ruggiero, L.M., *et al*. Genetic diversity of cultivated *Coffea arabica* inbred lines assessed by RAPD, AFLP and SSR marker systems genetics and plant breeding. *Sci Agric*. 2005, 62, 366–373, https://doi.org/10.1590/S0103-90162005000400010.

Marques, I., Draper, D., Riofrío, L., *et al*. Early signs of the effects of forest fragmentation on the genetic diversity and structure of the threatened Ecuadorian Tree *Ocotea rotundata* (Lauraceae). *Forests.* 2022, 13, 1940, https://doi.org/10.3390/f13111940.

McKenna, A., Hanna, M., Banks, E., *et al*. The genome analysis toolkit: a Map reduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010, 20, 1297–1303. https://doi.org/10.1101/gr.107524.110.

Moncada, P., McCouch, S. Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. *Genome.* 2004, 47, 501–509, https://doi.org/10.1139/g03-129.

Nesper, M., Kueffer, C., Krishnan, S., *et al*. Shade tree diversity enhances coffee production and quality in agroforestry systems in the Western Ghats. *Agric Ecosyst Environ.* 2017, 247,172–181, https://doi.org/10.1016/j.agee.2017.06.024.

Pritchard, J.K., Stephens, M., Rosenberg, N.A., *et al*. Association mapping in structured populations. *Am J Hum Genet.* 2000, 67, 170–181, https://doi.org/10.1086/302959.

Purcell, S., Neale, B., Todd-Brown, K., *et al*. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007, *81*, 559–575, https://doi.org/10.1086/519795.

Scalabrin, S., Toniutti, L., Di Gaspero, G., *et al*. A single polyploidization event at the origin of the tetraploid genome of *Coffea arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Sci Rep.* 2020, 10, 4642, https://doi.org/10.1038/s41598-020-61216-7.

Senkoro, A.M., Talhinhas, P., Simões, F., *et al*. The genetic legacy of fragmentation and overexploitation in the threatened medicinal African pepper-bark tree, *Warburgia salutaris*. *Sci Rep.* 2020, 10, 19725, https://doi.org/10.1038/s41598-020-76654-6.

Silva, H.L. Número de cromossomos em *Coffea racemosa* Lour. *Bragantia.* 1956, 15, XVII–XVIII, https://doi.org/10.1590/S0006-87051956000100034.

Silvestrini, M., Junqueira, M.G., Favarin, A.C., *et al*. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genet Resour Crop Evol.* 2007, 54, 1367–1379, https://doi.org/10.1007/s10722-007-9295-5.

Sitoe, A., Salomão, A., Wertz-Kanounnikff, S., *et al*. The context of REDD+ in Mozambique: Drivers, agents and institutions. *CIFOR*, Nairobi. 2012.

Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014, 30, 1312–1313 https://doi.org/10.1093/bioinformatics/btu033.

Venancio, L.P., Filgueiras, R., Mantovani, E.C., *et al*. Impact of drought associated with high temperatures on *Coffea canephora* plantations: A case study in Espírito Santo State, Brazil. *Sci Rep.* 2020, 10, 19719, https://doi.org/10.1038/s41598-020-76713-y.

Wicke, S., Schneeweiss, G.M., de Pamphilis, C.W., *et al*. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol Biol.* 2011, 76, 273–297, https://doi.org/10.1007/s11103-011-9762-4.

Zhu, Y., Jia, X., Wu, Y., *et al*. Quantitative proteomic analysis of *Malus halliana* exposed to salt-alkali mixed stress reveals alterations in energy metabolism and stress regulation. *Plant Growth Regul.* 2020, 90, 205–222,https://doi.org/10.1007/s10725-019-00563-6

Chapter 3

Irradiance level and elevation shape the soil microbiome communities of *Coffea arabica* L. under shaded or full sun management

Modified after Tapaça, I.d.P.E., Obieze, C.C., Pereira, G.V.d.M. Fangueiro, D., Coutinho, J., Fraga, I., Partelli, F.L., Ramalho, J.C., Marques, I., Ribeiro-Barros, A.I. Irradiance level and elevation shape the soil microbiome communities of *Coffea arabica* L. *Env Microbiome*. 19, 75, (2024). https://doi.org/10.1186/s40793-024-00619-9

3.1. Introduction

The *Coffea* genus comprises 130 species, from which only *Coffea arabica* L. (Arabica coffee) and *Coffea canephora* Pierre ex A. Froehner (Robusta coffee) are commercially significant (Davis & Rakotonasolo, 2021). These account for approximately 57% and 43% of the global coffee production, respectively (ICO, 2023). Coffee is a leading commodity worldwide, involving around 100 million workers throughout its value chain (ICO, 2022). However, the sustainability of the coffee sector is increasingly at risk due to climate change impacts, notably global warming and drought (DaMatta *et al*., 2018; Cassamo *et al*., 2023). To counteract this impact, the implementation of climate-smart strategies, such as agroforestry, is undeniably the most straight forward approach (Gomes *et al*., 2020; Cassamo *et al*., 2022; 2023).

Theoretically, effective shaded coffee agroforestry systems (AFS) mitigate high temperatures and minimize evapotranspiration (DaMatta, 2004; Gomes *et al*., 2020; de Carvalho *et al*., 2021). However, the success depends on the effective embedding of crop management practices (including soil, water, as well as compatible shade trees and coffee cultivars) within specific agro-ecological contexts (Koutouleas *et al*., 2022a,b).

For example, the incidence of natural enemies of coffee biotic stressors may be either boosted or reduced depending on the vegetation structure and composition (Perfecto *et al*., 2014; Avelino *et al*., 2018; Avelino *et al.,* 2020; Gonzalez *et al*., 2021; Newson *et al*., 2021). Another relevant aspect is the impact of shade which may have a negative (e.g. Chen *et al*., 2017), positive (Bote & Struik, 2011) or neutral (Cassamo *et al*., 2022) impact, depending on the interaction between genotype and environment, as well as the intensity of shade or irradiation (Koutouleas *et al.,* 2022ab). Altitude is also a crucial factor in coffee cultivation, driving the physical and chemical characteristics of coffee beans (Mintesnot & Dechassa, 2018; Guimarães *et al*., 2019; Paudel *et al*., 2021; Cassamo *et al*., 2022). Finally, the coffee microbiome emerges as core component of the system (Caldwell *et al*., 2015; Duong *et al.,* 2020; Ge *et al*., 2023; de Sousa *et al*., 2022a; Veloso *et al*., 2023), due to the its preponderant role in plant evolution, health and productivity (Hassani *et al*., 2018; Choi *et al*., 2021).

Soils are among the most important reservoirs of biodiversity, hosting ca. 1/4 of living organisms in terrestrial ecosystems (Decaëns *et al*., 2006; Bardgett & van der Putten, 2014). These include bacteria, archaea, fungi, protists, and many other eukaryotes (e.g. nematodes, mites, ants, beetles, earthworms), which provide a set of supporting services (soil formation, and nutrient cycling), provision services (food, freshwater, fuel, fiber, biochemicals, genetic resources), and regulating services (climate regulation, pest and disease regulation, water regulation, remediation, and pollination) (FAO *et al.,* 2020; Sokol *et al*., 2022).

Therefore, research on soil biodiversity is giving a step forward, envisioning the maximization of ecosystem goods and services, as well as the elucidation of biological, ecological and evolutionary processes (Bardgett & van der Putten, 2014). In coffee, Caldwell *et al*. (2015) reported high microbial diversity in soils from intensive, organic, and transition farms in Brazil, highlighting the potential of plant growth-promoting bacteria to improve coffee production and counteract environmental constraints. Tran (2022) published a data set of the rhizosphere microbiome of *C. canephora* in the Central Highlands region of Vietnam, the second largest coffee producer in the world after Brazil. The reported taxonomical diversity was also considerably high, and reflected in the associated functions, particularly regarding biosynthetic processes. More recently, Ge *et al.* (2023) analyzed the effect of altitude on the diversity of microbial communities in the rhizosphere of *C. arabica* in Yunnan, the most expressive center of coffee production and trading in China. In line with the previous studies, the authors observed that microbial diversity and richness was high, and essentially driven by soil pH and altitude.

In this study, we conducted a comprehensive analysis of the soil physicochemical properties and microbial communities of *C. arabica* cultivated under AFS in the evergreen rainforest of Gorongosa Mountain, part of the Gorongosa National Park (GNP), Mozambique. GNP is one of the most interesting and valuable case studies for the development of climate mitigation and/or adaptation strategies (Pringle, 2017; Matos *et al.,* 2021), owing to its exclusive biodiversity and anthropo-climate vulnerability (Stalmans & Victor, 2020; IPCC, 2022). Specifically, we aimed to unveil the influence of elevation and canopy shade in bacteria, archaea, and fungi communities.

3.2. Materials and methods

3.2.1. Experimental design

The study was conducted in the Gorongosa Mountain, belonging to the Gorongosa National Park, Sofala province, Mozambique (Lat. 18º 24’ 14’’S, Long. 34º 06’ 31.5’’E). *C. arabica* plants were implanted 1.5 m apart within a row and 3 m between rows, at a density of ca. 2222 plants ha^{-1}. The split-plot design of Cassamo *et al*., (2022) was used to assess the impact of altitude and/or light conditions on the coffee soil properties. This included three different altitudes (main plots): ca. 600 m (18º 30’ 53’’ S, 34º 03’ 05’’ E), ca. 800 m (18’ 30’’ 04’’ S, 34º 02’ 58’’ E), and ca. 900 m (18º 28’ 54’’ S, 34º 02’ 43” E) above sea level (a.s.l.); and three levels of light (sub-plots) per altitude: deep shade (DS, average diurnal PPFD of 127 ± 28 $\mu mol\ m^{-2}\ s^{-1}$), moderate shade (MS, 725 ± 48 $\mu mol\ m^{-2}\ s^{-1}$), and full Sun (FS, 1268 ± 52 $\mu mol\ m^{-2}\ s^{-1}$). The main canopy forest trees were *Khaya anthotheca* (Welw.) C. DC., *Erythrina lysistemon* Hutch, and *Albizia adianthifolia* (Schumach.) W. Wight.

3.2.2. Soil sampling

At each altitude and shade level, three coffee plants were randomly selected and soil samples were collected with an auger at a depth 0-10 cm and 10-20 cm. One sub-set was air-dried, ground, sieved (2 mm particle size) and used for physicochemical analysis. The other part was used to prepare composite samples from the two depths, stored on ice during the field collection and once in the lab kept at - 80°C until DNA extraction.

3.2.3. Chemical and enzymatic analyses

Soil samples were characterized in terms of Olsen extractable phosphorus (P-Olsen), pH (in a suspension 1:2.5 with H_2O), soil organic carbon (SOC), and microbial biomass carbon and nitrogen (MBC and MBN, respectively), electric conductivity, extractable potassium (Egner Rhiem K), mineral nitrogen (NO_3^- and NH_4^+) and moisture using standard protocols (Olsen *et al*. 1954; Egnér *et al*. 1960; Claessen, 1997). MBC and MBN were determined with the fumigation-extraction method using calibration values of K_{EC} = 0.45 for C and K_{EN} = 0.54 for N (Joergensen, 1996; Joergensen & Mueller, 1996). The quantification of C and N based on K_2SO_4 extraction was conducted using a near-infrared detector for carbon, and chemiluminescence for nitrogen. In both cases, samples underwent combustion at 950 ºC in a Formacanalyzer (Skalar, Breda, Netherlands).

The activities of β-glucosidase, acid phosphatase and alkaline phosphatase (phosphomonoesterases) were determined according to the protocol of Tabatabai (1994), using p-nitrophenyl-β-D-glucopyranoside (*p*NG) as substrate for β-glucosidase, and p-nitrophenyl phosphate (*p*NP) as substrate for both phosphomonoesterases. Urease activity was determined as described by Kandeler (1995) without buffering. For all enzymatic activities, the absorbance values of the extracts were determined in a segmented flow analyzer system with a preliminary dialysis step to remove color and microparticle interferences. β-glucosidase, acid- and alkaline phosphatase activities were expressed in µg p-nitrophenol h^{-1} g^{-1} dry soil, while urease activity was expressed in µg N-NH_4^+ 2 h^{-1} kg^{-1} dry soil.

Descriptive statistics and statistical analyses for soil properties were performed using RStudio version 4.1.1. (R Core Team, 2021). The heterogeneity of the variance was first tested, and the original data were normalized by log-transformation when necessary. A GLM analysis was used to analyze the effects of altitude, shade, and their interactions on the soil properties.

3.2.4. DNA extraction and amplicon sequencing

Microbial DNA was extracted from soil samples using the DNeasy Power Soil Pro Soil DNA Isolation Kit (Qiagen, Germany City, MD, USA) following the manufacturer instructions. DNA integrity and concentration were determined by 1% agarose gel electrophoresis and fluorometric quantification using a fluorometer (Qubit 2.0, Invitrogen, CA, USA), respectively. Amplicon libraries targeting the V4 hypervariable region of the 16S rRNA for bacteria (Bakt_341F: CCTACGGGNGGCWGCAG and Bakt_805R: GACTACHVGGGTATCTAATCC) and archaea (Bakt_341F: CCTACGGGNGGCWGCAG and Bakt_805R: GACTACHVGGGTATCTAATCC) and the ITS2 region for fungi (3F: GCATCGATGAAGAACGCAGC and 4R: TCCTCCGCTTATTGATATGC) were used following the Amplicon Metagenomic Sequencing Library Preparation guide (http://emea.support.illumina.com) (Abdelfattah *et al*. 2018).

Sequencing libraries were generated using the TruSEq DNA PCR-free sample preparation kit (Illumina, San Diego, CA, USA) following the manufacturer’s instructions. The final libraries were sequenced using the Illumina Miseq300 PE to generate 300 bp paired-end reads through Macrogen sequencing services (Macrogen, Seul, Korea).

3.2.5. Taxonomic Diversity

Paired-end sequence reads were demultiplexed using the MiSeq reporter software (Illumina Inc., CA, USA) and checked for quality using FastQC v.0.11.9 (Babraham Institute, Cambridge, UK). Paired-reads were trimmed at both 5´ and 3’ ends eliminating poor quality nucleotides, denoised, merged, and chimeric sequences using the DADA2 denoiser (Callahan *et al*., 2016) and then incorporated into QIIME 2 (Bolyen *et al*., 2019).

The resulting Amplicon Sequence Variant (ASV) count table was depleted of singletons, and representative sequences taxonomically classified using a trained classifier of the SILVA reference (Release 132) (Quast *et al.,* 2013). Alpha diversity analysis was conducted based on observed ASVs, Shannon entropy and Pielou’s evenness indices, while community dissimilarity was assessed using Bray-Curtis distance, which was visualized through non-metric multidimensional scaling (NMDS) ordination. Permutational multivariate analysis of variance (PERMANAOVA) was used to test for the significance of the microbial community dissimilarity across the different gradients of altitude and shades investigated. Canonical correspondence analysis (CCA) was also performed to determine the relationship of the microbial communities to soil physicochemical parameters along the shade and altitudinal gradients. Prior to computation of microbial diversity, the counts were normalized to 11,786 (bacteria), 2,414 (archaea) and 67,302 (fungi) ASV.

The detection of biomarkers across gradients of altitude and shade was performed using Linear discriminant analysis Effect Size (LEfSe) (Segata *et al.,* 2011). Phylotypes with an LDA score ≥ 3.0 and a False Discovery Rate (FDR)-adjusted *p*-value ≤ 1.0 were considered to be differentially abundant. Except stated otherwise, data analysis and visualization were performed using R software v.4.1.1 (R Core Team, 2021).

3.2.6. Functional prediction

Prediction of bacterial community functions was done using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States - PICRUSt2

(Douglas *et al.,* 2020). The functional prediction was done by aligning the 16s rRNA marker gene representative sequences to a reference multiple-sequence alignment and reference phylogeny utilizing HMMER (Eddy, 2011), EPA-NG (Barbera *et al.,* 2019) and GAPPA (Czech *et al*., 2020). Subsequently, gene families were predicted using a hidden state prediction tool – Castor (Louca & Doebeli, 2018), after a normalization of the 16S rRNA gene copies. The predicted gene families were thereafter collapsed into KEGG pathways using MinPath (Ye & Doak, 2009). Differentially abundant pathways across the different altitudes were afterward determined using LefSe (LDA score ≥ 2.0; *p*-value ≤ 0.05).

3.3. Results

3.3.1. Chemical characteristics and enzymatic activities of soils

The level of shade only affected available phosphorus (P-Olsen; $p = 0.014$), MBN ($p = 0.014$) and the level of moisture ($p = 0.006$) (Table S3.1). In contrast, altitude exhibited a more pronounced effect on soil characteristics, influencing nearly all parameters except for pH and urease activity, i.e. P-Olsen ($p = 0.008$), SOC ($p = 0.009$), MBC ($p = 0.0001$), MBN ($p = 0.0001$), electrical conductivity ($p = 0.002$), Egner Rhiem K ($p = 0.025$), NO_3^- ($p = 0.0001$), NH_4^+ ($p = 0.0001$), moisture ($p = 0.0001$), β-glucosidase ($p = 0.008$), acid phosphatase ($p = 0.0001$), and alkaline phosphatase ($p = 0.0001$) (Table S3.1). Overall, at 800 m there were significantly higher levels of P-Olsen, SOC, and nitrate (NO_3^-), when compared to the 600 m and 900 m. MBC and MBN, as well as moisture content, were found to be lower at the 800 m, highlighting a distinct environmental profile at this altitude (Table 3.1). Electrical conductivity and NH_4^+ increased with altitude, while the opposite was found for Egner Rhiem K (Table 3.1). Apart from urease, soil enzymatic activities also exhibited sensitivity to altitude changes (Table S3.1), with significantly higher activities of β-glucosidase and phosphatases (acid and alkaline) at 800 m (Table 3.1).

Table 3.1 Summary of the chemical properties and enzymatic activities of *C. arabica* soils. Mean values ± SD (n = 9) are indicated. Different superscripts indicate significant differences between altitudes, for each variable (ANOVA followed by Tukey-HSD, both for a 95% of confidence).P-Olsen: available phosporous; SOC: soil organic carbon; MBC: microbial biomass carbon; MBN: microbial biomass nitrogen; Egner Rhiem K: extractable potassium; NO_3^-: nitrate; NH_4^+: ammonium.

	600 m	**800 m**	**900 m**
Soil variables			
P-Olsen (mg kg^{-1})	5.40 ± 2.02 ^{a}	8.34 ± 2.37 ^{b}	5.94 ±5.43^{a}
pH (H_2O)	5.41 ± 0.33 ^{a}	5.17 ± 0.33 ^{a}	5.67 ± 0.25 ^{a}
SOC (g kg^{-1})	64.60 ± 9.47 ^{a}	74.50 ±12.56 ^{b}	66.37 ± 14.54 ^{a}
MBC (mg C kg^{-1})	146.32 ± 37.07 ^{b}	85.15 ± 21.53 ^{a}	138.28 ± 33.94 ^{b}
MBN (mg N kg^{-1})	18.65 ± 6.16 ^{b}	12.65 ± 4.21 ^{a}	21.64 ± 5.21 ^{b}
Electrical conductivity (µS cm^{-1})	84.92 ± 25.36 ^{a}	113.40 ± 61.08 ^{b}	149.17 ± 62.43 ^{c}
Egner Rhiem K (mg kg^{-1})	158.91 ± 20.78 ^{c}	140.46 ± 50.89 ^{b}	119.54 ± 47.91 ^{a}
N-NO_3^- (mg kg-1)	9.35 ± 3.54 ^{a}	15.50 ± 7.04 ^{a}	8.21 ± 3.59 ^{a}
N-NH_4^+ (mg kg^{-1})	19.42 ± 2.68 ^{a}	27.22 ± 8.56 ^{b}	30.14 ± 8.63 ^{c}
Moisture (g kg^{-1})	68.31 ± 4.30 ^{b}	57.20 ± 4.55 ^{a}	66.68 ± 6.36 ^{b}
Enzymatic activities			
Urease (µg N-NH_4^+ g^{-}1 h^{-1}) dry soil.	39.25 ± 10.67 ^{a}	32.54 ± 9.31 ^{a}	35.53 ± 7.53 ^{a}
β-glucosidase (µg p-nitrophenol g^{-1} h^{-1})	58.20 ± 12.81 ^{a}	65.73 ± 10.87 ^{b}	54.03 ± 12.85 ^{a}
Acid phosphatase (µg p-nitrophenol g^{-1} h^{-1})	310.98 ± 64.91 ^{a}	686.61 ± 149.97 ^{c}	530.51 ± 123.39 ^{b}
Alkaline phosphatase (µg p-nitrophenol g^{-1} h^{-1})	122.43 ± 42.42 ^{a}	189.96 ± 42.47 ^{b}	139.04 ± 28.15 ^{a}

3.3.2. *Coffea* alpha diversity between different altitudes and shade trees

A total of 89,590 high quality reads were obtained for bacteria, 91,977 for archaea and 93,507 for fungi with high Q values and adequate GC contents (Table S3.2). For bacteria, the observed number of Operational Taxonomic Units (OTUs), Shannon entropy and Pielou's evenness revealed high species richness in all locations, although no significant differences were detected between the different levels of canopy shade (Figure 3.1; Table S3.3).

However, altitude had a strong influence on the bacterial richness with significantly lower values observed at the intermediate elevation of 800 m (Figure 3.1; Table S3.4), while archaea richness increased with altitude (Figure 3.1; Table S3.5; Table S3.6). In contrast, neither shade (Table S3.7) nor altitude (Table S3. 8) had an impact on fungi richness, whether considering the observed number of OTUs, Shannon entropy, or Pielou's evenness (Figure 3.1).

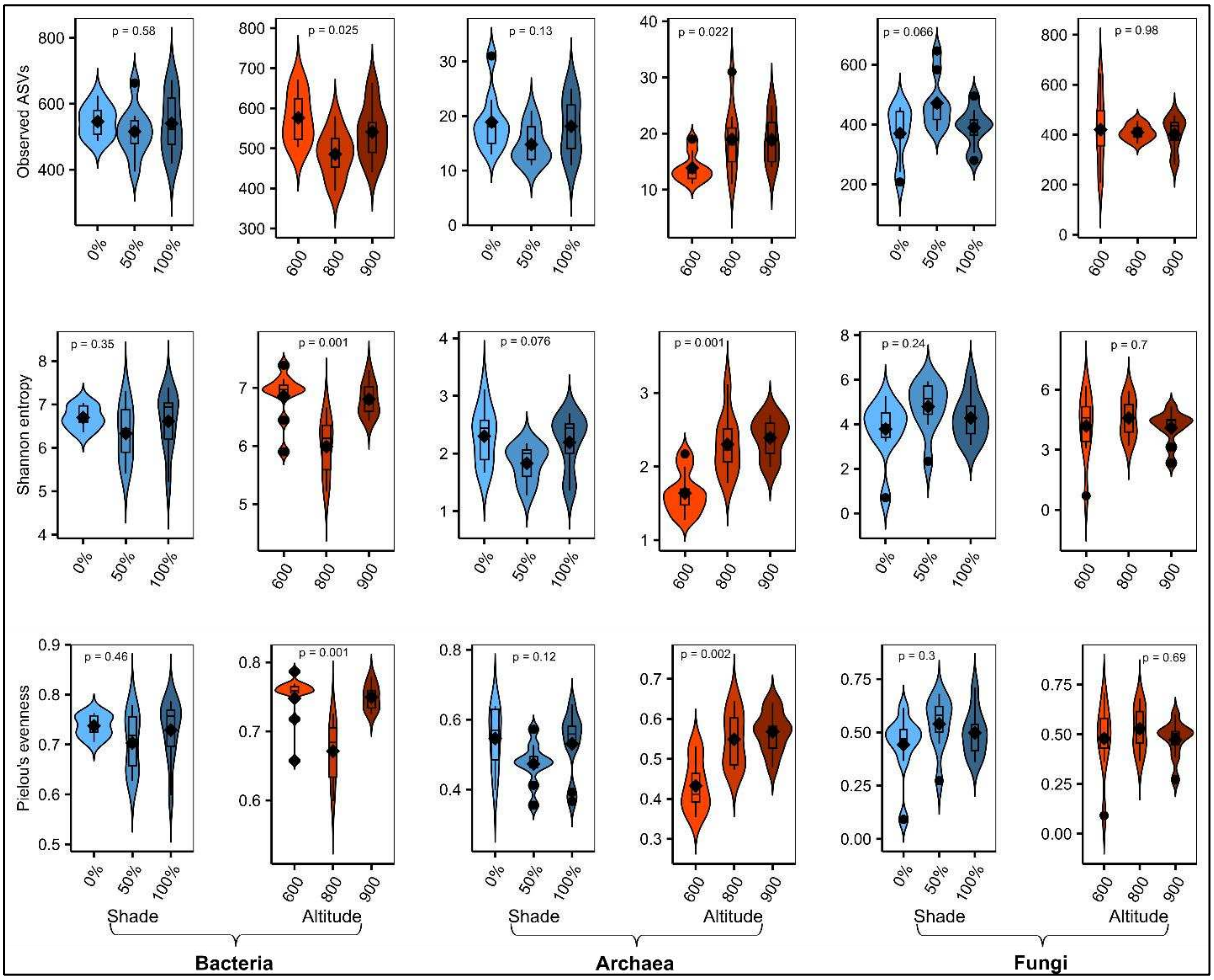


Figure 3.1 Observed ASVs, Shannon–Wiener index of diversity and Pielou's evenness in soil samples of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m) and under different levels of canopy shading (0%, 50% and 100%) considering bacteria, archaea, and fungi. Black-filled dots on the boxplots depict mean values. Comparisons between altitude and shade levels are based on Kruskal-Wallis (Bonferroni adjusted $p < 0.05$).

3.3.3. *Coffea* beta diversity between different altitudes and shade levels

Nonmetric multidimensional scaling analysis revealed a strong bacteria differentiation between altitude fields (Figure 3.2A). These differences in the multivariate space were significant when considering the effect of altitude (PERMANOVA $R^2 = 0.493$, $p = 0.001$), but not when considering the effect of shade (PERMANOVA $R^2 = 0.052$, $p = 0.797$). The same was observed in the multivariate space of archaea (Figure 3.2B) where altitude had a strong effect on the community structure (PERMANOVA $R^2 = 0.373$, $p = 0.001$), whereas no significant effect was detected when considering the level of shade (PERMANOVA $R^2 = 0.047$, $p = 0.745$). Altitude also influenced fungi beta diversity (Figure 3.2C), as significant differences were found between altitude levels (PERMANOVA $R^2 = 0.177$, $p = 0.001$), but not when considering the effect of shade (PERMANOVA $R^2 = 0.110$, $p = 0.068$).

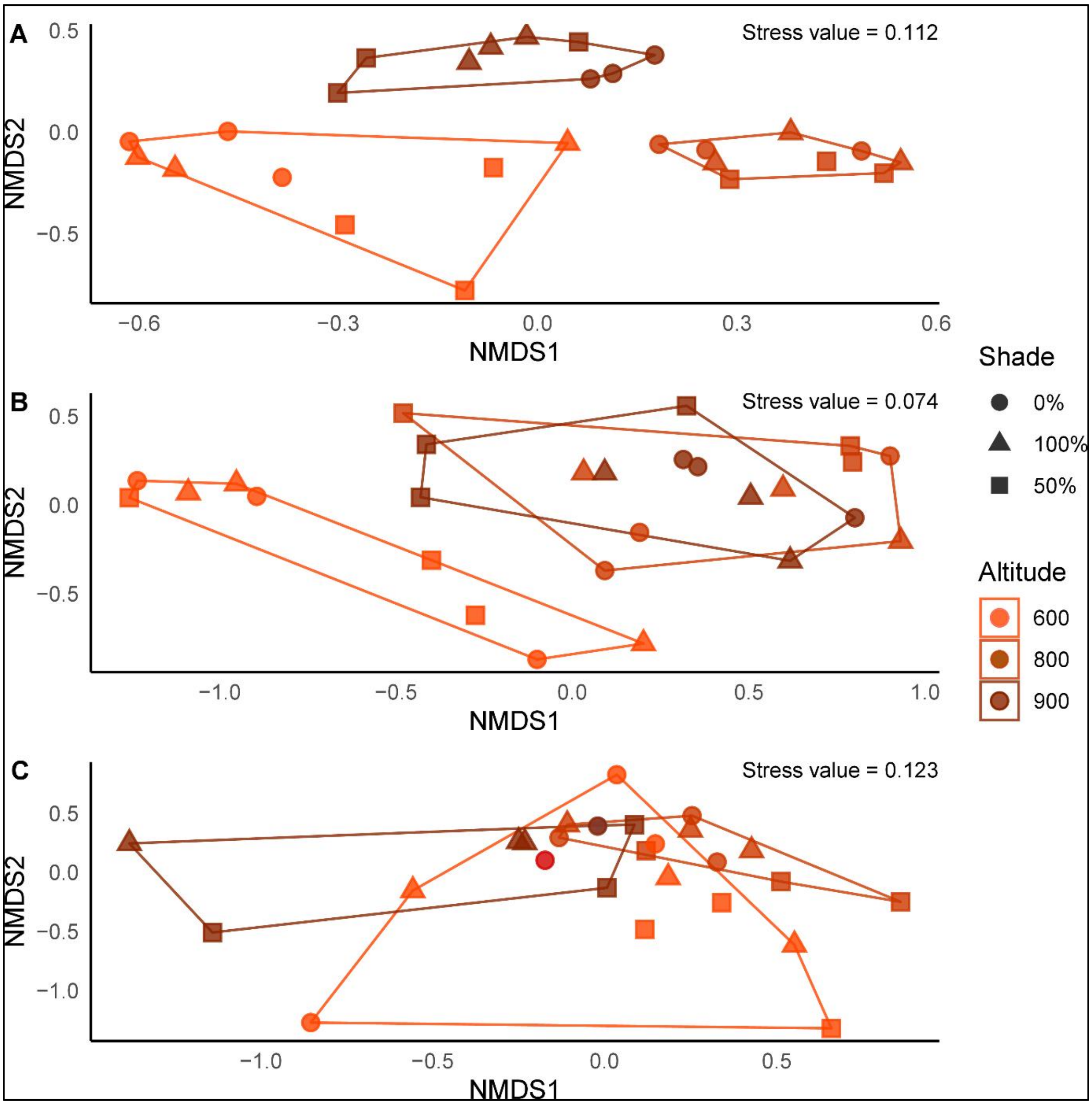


Figure 3.2 Nonmetric multidimensional scaling (NMDS) analysis of Bacteria (A), Archaea (B), and Fungi (C) community differences in soil samples of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m).

Altitude alone accounted for 46% of the bacterial community variation ($R^2 = 0.02$; $p = 0.001$). However, the effect of shade on the bacterial community was not significant ($R^2 = 0.01$; $p = 0.113$). In the case of the archaeal community, altitude played an even more prominent role, explaining 61% of its variation ($R^2 = 0.01$; $p = 0.001$), while the impact of shade remained insignificant ($R^2 = 0.01$; $p = 0.277$). Regarding the fungal community, altitudinal changes significantly affected its composition ($R^2 = 0.158$; $p = 0.001$). Additionally, canopy cover had some influence on the fungal community, though to a lesser extent ($R^2 = 0.038$; $p = 0.016$). Notably, there was still 81% of the variation in the fungal community that remained unexplained by either altitude or canopy cover.

To further determine the microbial community – soil physicochemical relationship, we performed a canonical correspondence analysis (CCA) (Figure 3.3). The results indicated that the model for bacteria explained 33% of the community variation (adjusted $R^2 = 0.33$; $p = 0.001$). Both altitude ($f = 6.87$; $p = 0.001$) and shade ($f = 1.53$; $p = 0.05$) significantly influenced the community assembly. Also, several soil physiochemical parameters, including NH_4^+, pH, EC, MBN, MBC and moisture significantly influenced the dissimilarities of the bacterial community along the altitudinal and shade gradients. Accordingly, there was high relative abundance of the bacterial families *Flavobacteriaceae* and *Ligionellaceae* at 900 m altitude, while *Streptomycetaceae* and *Shingobacteriaceae* were dominant at 600 m altitude.

The CCA model for the archaeal community explained 34% (adjusted $R^2 = 0.34$; $p = 0.001$) of the community dissimilarities along the shade and altitudinal gradients, though only altitude (shade: $f = 1.55$, $p = 0.09$; altitude: $f = 6.83$; $p = 0.001$) had significant influence on the archaeal community. The dominance of *Methanobacteriaceae* at 800 m correlated with the concentration of NO_3^-. Other parameters that influenced the archaeal community were similar to those influencing the bacterial community. The CCA analysis further revealed that a significantly smaller variation of the fungal community dissimilarities was explained by the constrained variables compared to bacteria and archaea (adjusted $R^2 = 0.16$; $p = 0.001$). However, contrary to the archaeal community, both altitude ($f = 3.04$; $p = 0.001$) and shade ($f = 1.54$; $p = 0.001$) significantly influenced the fungal community assembly, while the influence of N-NH_4^+, pH, EC, MBN, MBC and moisture were consistent, regardless of microbial domain.

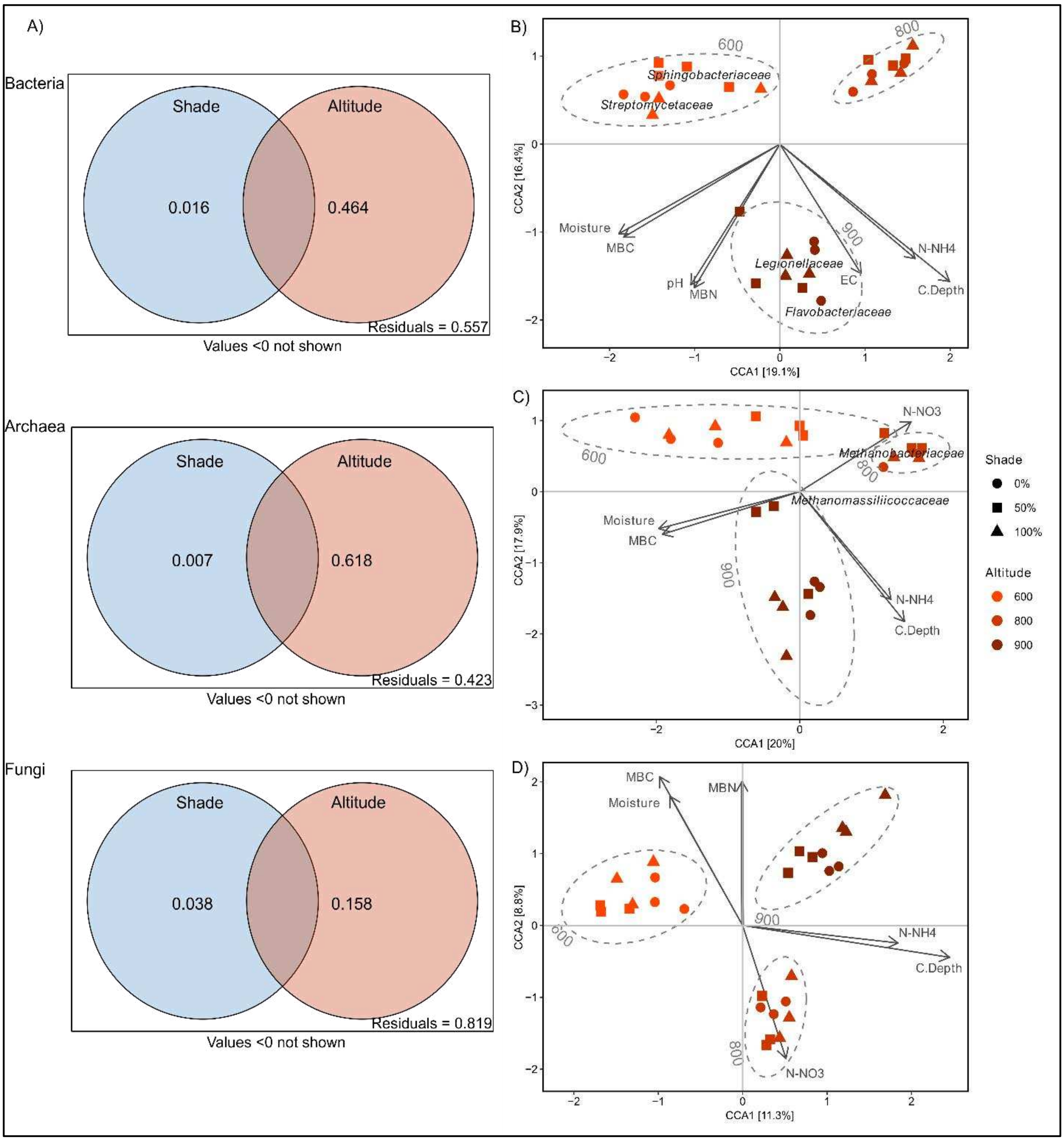


Figure 3.3 Bacterial, archaea and fungal community differentiation explained by shade and altitude based on variation partitioning analysis (A). Canonical correspondence analysis (CCA) explaining bacterial (B), archaeal (C), and fungal (D) community relationship to soil physicochemical parameters along a shade and altitudinal gradient. The ellipses represent 95% confidence interval in multivariate space according to each group's centroid.

3.3.4. *Coffea* dominant and differentially abundant phylotypes

The composition of the microbial communities was dominated by several distinct phyla (Figure 3.4). For bacterial communities, four predominant phyla were identified: *Proteobacteria* (24%), *Verrucomicrobia* (23%), *Actinobacteria* (19%), and *Acidobacteria* (15%). In contrast, the archaeal community was overwhelmingly dominated by *Thaumarchaeota,* which constituted 98% of its population. Among the fungi, *Mucoromycota* was the most prevalent (54%), followed closely by Ascomycota (40%) (Figure 3.4). At this taxonomic rank, there were no differentially abundant archaea or fungi across the different altitudes and shades, but for bacteria, some phyla had distinct patterns across altitudes. For instance, higher altitudes increased the relative abundance of *Acidobacteria* (LDA = 5.57; adjusted *p* = 0.005), *Proteobacteria* (LDA = 5.46; adjusted *p* = 0.006) and *Planctomycetes* (LDA = 4.97; adjusted *p* = 0.001). At the lowest altitude of 600 m, the abundance of *Bacteroidetes*, *Chloroflexi*, *Nitrospirae*, and *Synegistetes* were significantly higher, while the biomarkers of the mid-level altitude (800 m) included *Verrucomicrobia* and *Actinobacteria* (Figure S4.1).

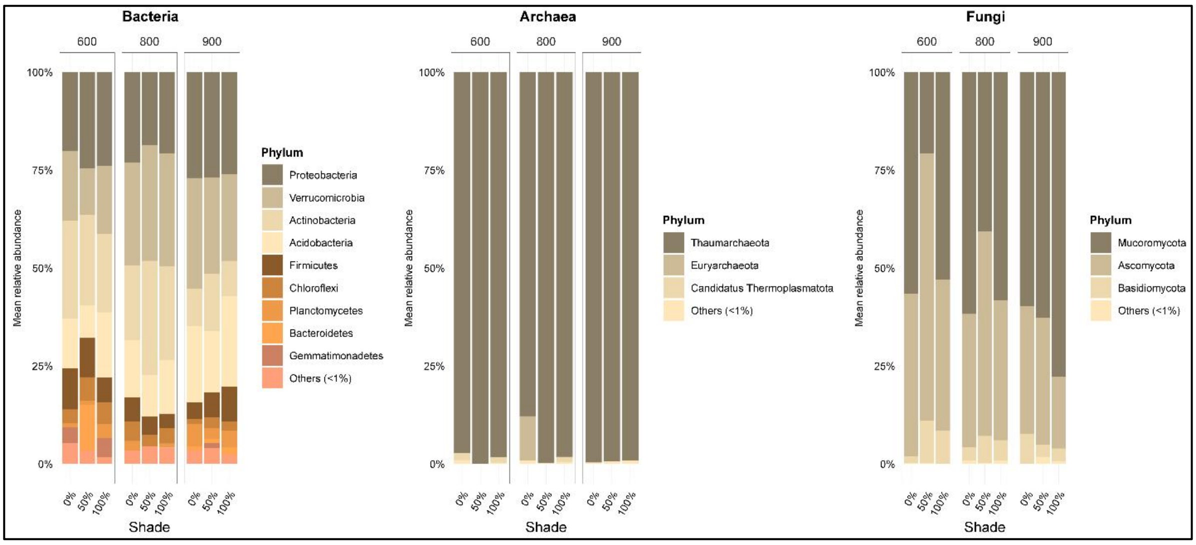


Figure 3.4 *Phylum* rank taxonomic profile of the microbial communities of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m) and under different levels of shade from native trees (0%, 50%, and 100%).

At the genus level, bacterial communities showed a predominance of *Chthoniobacter* (22%), followed by *Rhodoplanes* (8%), *Acidobacterium* (6%) and *Arthrobacter* (5%) (Figure 3.5). *Nitrososphaera* was the dominant archaea (85%), especially at an altitude of 600 m while the abundance of *Nitrosopumilus* (13%) increased at higher altitudes (Figure 3.6). In the case of fungi, *Linnemannia* (41%) was predominant in all altitudes, followed by *Mortierella* (10%), *Fusarium* (7%) and *Penicillium* (5%) (Figure 3.5).

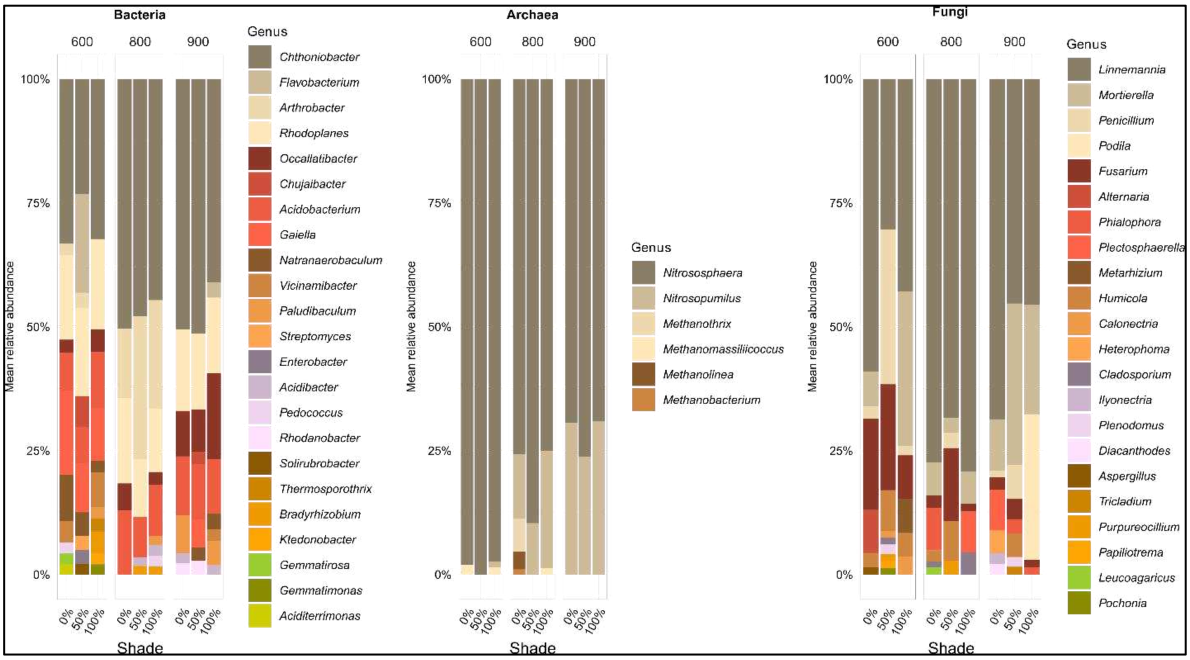


Figure 3.5 Genus level taxonomic profile of the microbial communities of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m) and under different levels of shade from native trees (0%, 50% and 100%). Only phylotypes with a relative abundance of at least 1% in any of the samples are presented.

At the genus taxonomic rank, 103 bacterial phylotypes were differentially abundant across altitudinal levels (Figure 3.6). In the case of bacteria, *Gaiella* (LDA 5.45; FDR-adjusted $p< 0.01$) and *Natranaerobaculum* (LDA 5.06; FDR-adjusted $p = 0.03$) were among the most predominant biomarker phylotypes at the lowest altitude of 600 m, *Chthoniobacter* (LDA 5.80; FDR-adjusted $p< 0.01$) and *Stella* (LDA 4.89; FDR-adjusted $p< 0.01$) were predominant biomarkers at 800 m, while *Occallatibacter* (LDA 5.30; FDR-adjusted $p = 0.01$) and *Paludibaculum* (LDA 4.89; FDR-adjusted $p< 0.02$) were predominant biomarkers at 900 m. Only four archaea phylotypes were differentially abundant at the different altitudinal levels: *Nitrososphaera* (LDA 6.11; FDR-adjusted $p< 0.01$) and *Methanomassiliicoccus* (LDA 5.12; FDR-adjusted $p = 0.60$) at 600 m and *Nitrosopumilus* (LDA 6.15; FDR-adjusted $p< 0.01$) and *Methanocaldococcus* (LDA 3.60; FDR-adjusted $p= 0.20$) at 900 m (Figure 3.6). Additionally, among the classified fungal phylotypes, 30 genera were differentially abundant across altitudinal levels. Of these, *Exophiala* (LDA 4.27; FDR-adjusted $p = 0.05$) and *Cladophialophora* (LDA 4.21; FDR-adjusted $p = 0.04$) were among the predominant biomarker phylotypes at 600 m, *Gliocladiopsis* (LDA 4.44; FDR-adjusted $p = 0.05$) and *Cystofilobasidium* (LDA 4.12; FDR-adjusted $p< 0.01$) at 800 m and *Podila* (LDA 5.58; FDR-adjusted *p* = 0.01) and *Cutaneotrichosporon* (LDA 4.05; FDR-adjusted *p*< 0.01) at 900 m (Figure 3.6).

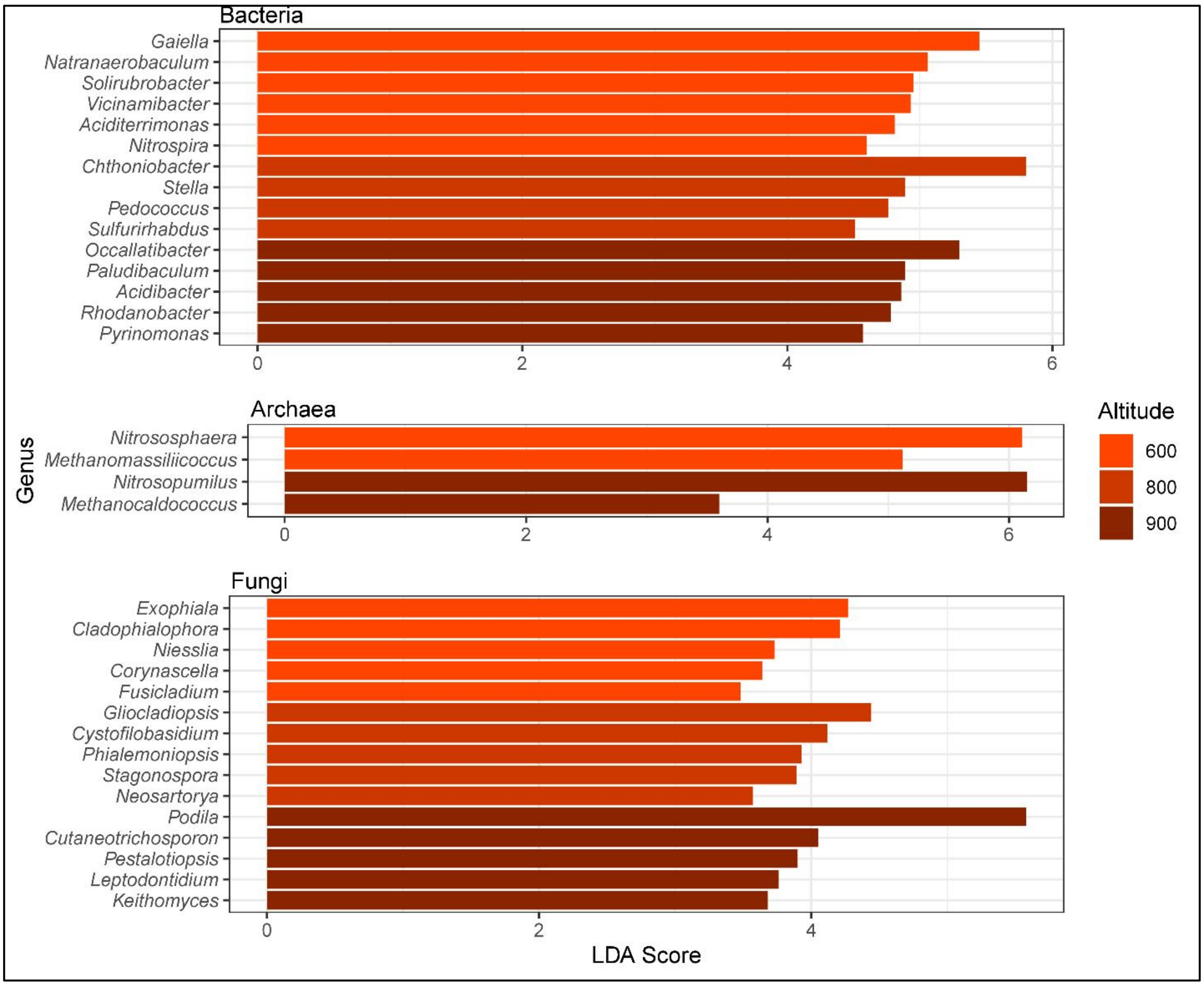


Figure 3.6 Differentially abundant Bacteria, Archaea and Fungi at the genus rank, across altitudes (600 m, 800 m, and 900 m). The differentially abundant data for bacteria and fungi (LDA ≥ 2.0; FDR-adjusted *p* ≤ 1.0) are a subset that was sorted according to adjusted p-values, LDA scores and altitudinal groupings. The complete list of differentially abundant genera is presented in Tables S3.9 and S3.10.

The prediction of bacterial community functions revealed significant variations across different altitudes, highlighting distinct metabolic adaptations that potentially support plant growth and productivity. As detailed in Figure 3.7, several ecologically relevant KEGG pathways were identified as differentially abundant at 900 m, 800 m, and 600 m, each corresponding to specific environmental stimuli and microbial interactions, organic matter decomposition and fermentation, and the breakdown of plant-derived compounds, respectively. The LDA scores and FDR-adjusted *P*-values for these pathways are presented in Table S3.10. At the highest elevation of 900 meters, bacterial communities showed a predominance of pathways associated with the response to environmental stimuli and microbial interactions. Key pathways included those related to oxidative stress response, bacterial chemotaxis, and biofilm formation.

At the mid-level elevation of 800 meters, the bacterial community was enriched in pathways involved in the decomposition of organic matter, fermentation processes, and bioremediation. Notable pathways included those related to amino acid and carbohydrate metabolism, butanoate and propanoate metabolism, and the degradation of aromatic compounds and chlorocyclohexane.

At the lowest elevation of 600 meters, the bacterial community function was characterized by pathways related to the breakdown of complex plant-derived compounds. Pathways such as cellulose and lignin degradation, starch and sucrose metabolism, and phenylpropanoid biosynthesis were significantly represented.

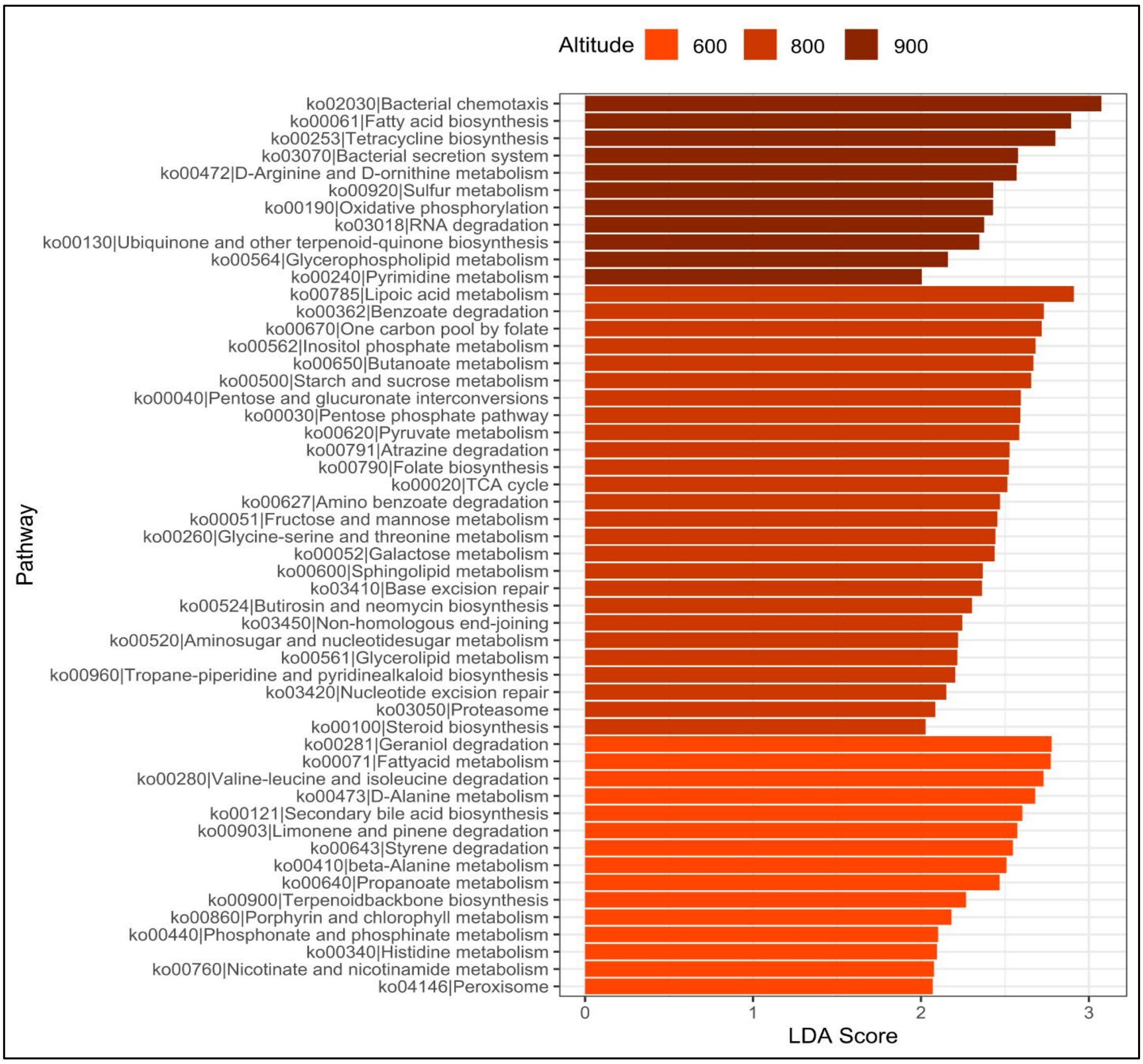


Figure 3.7 PICRUSt2 predicted differentially abundant KEGG pathways along the altitudinal gradient ($p \leq 0.01$; LDA ≥ 2.0).

3.4. Discussion

Soil microbiome encompasses diverse microorganisms residing near plant roots, playing a pivotal role in ecosystem functioning. It not only shapes the structure and composition of biodiversity but also critically influences plant health and fitness. Its interactions with plants significantly contribute to various ecosystem processes, including nutrient cycling, soil fertility, and plant disease resistance, thereby underpinning the overall health and resilience of ecosystems. Recent research has devoted special attention to the nexus plants-microbes-environment, as an essential component of climate-smart, resilient, and sustainable agriculture (Arif *et al.,* 2020; Trivedi *et al*., 2020; Ajala *et al*., 2022). In this study, a comprehensive analysis of the soils under the canopy projection of *C. arabica* was conducted to evaluate the shifts in microbial composition and associated functions along a gradient of elevation (600 m, 800 m, and 900 m) and shade (0%, 50%, 100%).

A set of soil variables (organic carbon, microbial biomass carbon and nitrogen, available phosphorus, extractable potassium, nitrate, ammonium, moisture, pH, electrical conductivity), and microbial enzyme activities (urease, β-glucosidase, acid- and alkaline phosphatases) were analysed. In addition, metabarcoding was performed to study microbial communities (bacteria, archaea, and fungi), and their putative functions.

Most soil parameters were highly variable (high standard deviation) and only influenced by altitude, with a particularly differentiated pattern at 800 m, where available phosphorus (P-Olsen), soil organic carbon (SOC) and nitrate (NO_3^-) were significantly higher, and microbial biomass carbon and nitrogen (MBC and MBN, respectively) and moisture were significantly lower when compared to 600 m and 900 m (Table 3.1). Considering that soil elementary composition along the elevation and shade gradients was quite homogeneous (fine clay, acidic, non-saline, non-calcareous, 8-14% organic matter; Table S3.11), this pattern might be driven by the vegetation composition, i.e. *C. arabica*, shade tree species and other natural vegetation (Sauvadet *et al*., 2020; Phour & Sindhu, 2023), which in turn shapes the microbial communities (Yue *et al*., 2023). Gota *et al*., (2024) reported that the influence of agroforestry species on soil chemical properties is specific to each species and is not altered by altitude changes. Accordingly, Sauvadet *et al.*, (2020) reported that shade trees were the main drivers of soil composition in cocoa agroforestry systems (AFS), improving pH, NH_4^+, NO_3^-, total C and N, biomass and P-Olsen contents. However, soil function improvement varied across AFS, being low with fruit species (*Canarium* and *Dacryodes*), moderate with tree legumes (*Albizia*), and high with timber trees (*Milicia* and *Ceiba*).

In the present study, the canopy shade trees in the experimental plots located at 800 m were mostly legumes (*Albizia adianthifolia* and *Erythrina lysistemon*), both included on the top list of N_2-fixing fertilizers (Sileshi *et al.*, 2014), while at 600 m and 900 m non-legume trees were also dominant. In addition, the differences observed between the three altitudes might also be associated to climate variations, particularly rainfall and temperature, which shape vegetation composition and structure, implying differences in the amount and composition of litter (Zhang *et al.*, 2014; Giweta, 2020). This hypothesis is corroborated by the enzyme activity assays, particularly β-glucosidase, and phosphatases, which presented higher activities at 800 m, i.e.65.73 ± 10.87, 686.61 ± 149.97, and 189.96 ± 42.47 µg p-nitrophenol g^{-1} h^{-1}. Enzymatic activities in soils are important for the decomposition of organic matter and mineralization of nutrients, and are useful indicators of soil biological activity and deductively of soil health (Stenberg, 1999; Ezeokoli *et al.*, 2020).

Metabarcoding analysis revealed high species richness in bacterial communities, with the lowest indexes recorded at 800 m (Figure 3.1; Table S3.4). This might explain the lowest MBC and MBN values obtained in this altitude. As for the soil parameters, altitude was the most important driver of bacterial communities, explaining 46% of the variation (R^2= 0.493, p=0.001), while shade did not produce significant changes (R^2 = 0.01; p = 0.113) (Figures 3.2A and 3.3).

The predominant phyla (*Proteobacteria*, *Verrucomicrobia*, *Actinobacteria*, and *Acidobacteria*) and the average Shannon diversity indexes were similar to those reported previously for *Coffea* spp. soils, i.e. between 6 and 7 (Caldwell *et al.*, 2015; Ge *et al.*, 2023; de Sousa *et al.*, 2022a; Veloso *et al.*, 2023). Although the abundance of various phyla varied along the elevation gradient (Figure 3.4), key plant growth-promoting (PGP) functions such as, nutrient cycling, organic matter decomposition, phosphate solubilization, soil aggregation, and biocontrol, were consistently present across all clusters, in line with the studies of Fierer *et al*. (2007, 2012). Accordingly, the predominant genera, i.e. *Chthoniobacter*, *Rhodoplanes*, *Acidobacterium*, and *Arthrobacter* (Figure 3.5), incorporated these major PGP attributes (Bill *et al*., 2021; Catania *et al*., 2022; Marian *et al*., 2022).

While the presence of *Rhodoplanes* and *Arthrobacter* in *Coffea* soils has been previously reported, the presence of *Chthoniobacter* and *Acidobacterium* seems to be exclusive of our agro-ecological system (Cabrera-Rodríguez *et al*., 2020; Duong *et al*., 2020; de Sousa *et al*., 2022b; Veloso *et al*., 2023). A set of 103 biomarker phylotypes were differentially abundant across the altitude gradient (Figure 3.6), from which the most prominent were: *Gaiella* and *Natranaerobaculum* at 600 m; *Chthoniobacter* and *Stella* at 800 m; and *Occallatibacter* and *Paludibaculum* at 900 m. All of them are included in the group of PGP bacteria with vital functions in forest ecosystems. The genus *Gaiella* is bound to microbe-microbe interactions (Sun *et al*., 2018; Wang *et al*., 2021), playing an important role in biocontrol, e.g. of *Fusarium oxyspororun* in tomato (Zhao *et al*., 2019) and strawberry (Lazcano *et al.,* 2021); some members also contribute to nitrogen cycling through the reduction of nitrate to nitrite (Albuquerque *et al*., 2011).

Natranaerobaculum is typical from hypersaline soda environments (Fazi *et al*., 2018), but its occurrence in forest soils seems to be related to health biomarkers (Kim *et al.,* 2022). *Chthoniobacter* is involved in nutrient cycling and produces secondary metabolites (Jenkins *et al*., 2009; Fu *et al*., 2022), associated with the control of bacterial wilt (Chen *et al*., 2020) and growth of beneficial bacterial communities (Fu *et al.,* 2022). Although *Stella* and *Paludibaculum* bacteria are not common soil taxa, their ecological function might be related to organic matter decomposition (Zavarzin, 1970). Additionally, *Paludibaculum* may also enhance plant chlorophyll content (Yoneda *et al*., 2021). *Occallatibacter* may play a relevant function in denitrification (Truu *et al*., 2020) as well as soil remediation (de Souza *et al*., 2022b). Despite the fact that the basic PGP functions are likely maintained along the elevational gradient, taking into account that P-Olsen, SOC and NO_3^- were significantly higher at 800 m, we hypothesize that the microbial consortium at this altitude is more effective in nutrient cycling.

Interestingly, despite the presence of legume trees and unlike what has been reported in similar studies in coffee (Duong *et al*., 2020; de Sousa *et al*., 2022a;

Veloso *et al*., 2023; Bullergahn *et al*., 2024), the presence of diazotrophs (Nitrogen-fixing bacteria) was not prominent in this study. However, collectively, a considerable set of symbiotic (*Bradyrhizobium, Mezorhizobium*, and *Rhizobium*) and non-symbiotic (e.g.*Agrobacterium*, *Azospirillum*, *Bacillus*, *Burkholderia*, *Clostridium*, *Microvirga*, *Nitrospirillum*, *Paraburkholderia*) genera is represented in the bacterial pool.

The richness and diversity of archaea was considerably lower than that of bacteria, and fungi. Similarly to the soil parameters and bacterial communities, the patterns of richness and diversity also varied with altitude. However, in this case, the lowest values of richness and diversity were recorded at the lowest altitude of 600 m. *Thaumarchaeota*, typically including ammonia-oxidizing archaea (AOA) (Stahl *et al.,* 2012), was the dominant phylum along the altitudinal gradient, similarly to the observations of Siles and Margesin (2016) for alpine forest soils, Caldwell *et al.,* (2015) for the coffee rhizosphere, and Aqell *et al*. (2024) for the Himalayas. Additionally, this is the most prominent endophytic phylum in several crops, including coffee (Oliveira *et al*., 2013). *Euryarchaeota*, which is associated to methanogenesis (Stahl *et al*., 2012), was scorable only at 800 m and under full sun (0% shade) and less represented (< 1%) in all other comparison groups. The irregular and low abundance (< 10%) of this phylum has also been reported in other soils, e.g. along an elevation gradient in alpine forest soils (Siles & Margesin, 2016), and in the rhizosphere of full-sun coffee in consociation with other food crops (Caldwell *et al*., 2015).

Both phyla are associated to a wide range of PGP functions, e.g. improvement of plant growth, tolerance to biotic and abiotic stress, and nutrient solubilization and assimilation (Stahl & de la Torre, 2012; Chow *et al*., 2022). In line with the findings of Siles and Margesin (2016) and Caldwell *et al.,* (2015), in this study, *Nitrososphaera* (phylum *Thaumarchaeota*) was the most abundant genus (70-100%) in all comparison groups. *Nitrosopumilus*, another AOA belonging to *Thaumarchaeota*, was the second most abundant genus (< 30%), particularly at 800 m and 900 m. These two genera play a crucial role in the nitrogen cycle (converting NH_3^+ into NO_2^-), i.e. bioavailability, soil fertility, and environmental regulation of nitrogen.

Fungi richness and diversity was considerably high (410.26 ± 91.32 observed ASVs and 4.29 ± 1.19 Shannon indexes, on average) (Figure 3.1) and within the range of the values reported for the soil microbiome of *C. arabica* (e.g. Bez *et al.,* 2023; Ge *et al.,* 2023). Alpha-diversity of fungi was neither affected by shade (Table 3. S7), and altitude (Table S3. 8), while beta-diversity (Figure 3.2C) was driven by altitude (R^2 = 0.177, *p* = 0.001), but unaffected by shade (R^2 = 0.110, *p* = 0.068). *Mucoromycota* and *Ascomycota* were the dominant phyla (> 90% in total) (Figure 3.5). As for the other microbial domains, both phyla include soil mycorrhizal fungi, saprotrophic decomposers and endophytes, with key functions in e.g. nutrient cycling and assimilation, or plant protection (Challacombe *et al*., 2019; Pawłowska *et al*., 2019). *Linnemannia* (*Mucoromycota*) was the predominant genus in all comparison groups. Interestingly, this genus has never been reported previously in *Coffea* (Caldwell *et al*., 2015; Duong *et al.,* 2020; de Sousa *et al*., 2021; de Sousa 2022a,b; Bez *et al*., 2023; Ge *et al.,* 2023; Ochoa-Henriquez *et al*., 2024). *Linnemannia* and *Mortierella* (the second most abundant genus in our study) have been reported as strong elicitors of plant growth in maize (Li *et al.,* 2018), wheat (Johnson *et al.,* 2019), and Arabidopsis (Ozimek *et al*., 2018; Vandepol *et al.,* 2022), likely driven by phytohormones (Vandepol *et al*., 2022). In the *Ascomycota* phylum, the most prevalent genera were *Fusarium* and *Penicillium*.

These genera are linked to a broad spectrum of ecological functions, including decomposition, nutrient cycling, biocontrol, and bioremediation (Altaf *et al.,* 2014; Abdel-Azeem *et al.,* 2019). Such functions are visible among the 30 biomarkers differentially abundant across the three altitudes, from which *Exophiala* and *Cladophialophora* (600 m), *Gliocladiopsis* and *Cystofilobasidium* (800 m), and *Podila* and *Cutaneotrichosporon* (900 m) were the predominant genera (Figure 3.6). For instance, *Exophiala* promotes plant stress tolerance and growth (Wang *et al*., 2020); *Cladophialophora*is associated to plant protection and productivity (Harsonowati*et al*., 2020); *Gliocladiopsis* enhances tolerance against biotic ad abiotic factors (Singh *et al*., 2017; Singh *et al.,* 2018); and *Cysttofilobasidium* is also a potential biocontrol agent (Vero *et al*., 2011). While the ecological roles of *Podila* and *Cutaneotrichosporon* remain less defined, the presence of *Podila* in alpine forest soils has been documented (Telagathoti *et al*., 2022), and *Cutaneotrichosporo* shows promise in biotechnology for the food and cosmetic industries (Stellner *et al*., 2023), as well as in biofuel and bioplastic production (Di Fidio *et al*., 2021).

The KEGG pathway data across different altitudes revealed distinct metabolic adaptations of microbial communities to their respective environments (Figure 3.7). Firstly, at 800 m the metabolic pathways to lipoic acid and benzoate degradation indicated a rich organic matter environment and a microbial ability to utilize diverse organic compounds (Zhou *et al.,* 2022). This altitude also shows active fermentation processes, as evidenced by butanoate metabolism and pyruvate metabolism pathways, hinting at the decomposition of organic materials. Notably, the atrazine degradation pathway suggests a capability to bioremediate certain pollutants, reflecting potential exposure to agricultural chemicals in this mid-altitude environment (Sandhya *et al.,* 2021). This observation suggests a robust microbial activity adapted towards organic matter turnover and pollutant degradation, which aligns with the higher levels of soil organic carbon and nitrate observed at this altitude, corroborating the hypothesis formulated above regarding the higher effectiveness of the microbial consortium at this altitude.

At 900 m, the high abundance of pathways such as bacterial chemotaxis and secretion systems (Figure 3.7) suggests a heightened sensitivity to environmental stimuli and an advanced capacity for microbial interactions, possibly aiding adaptation to cooler, more variable high-altitude conditions. The biosynthesis of tetracycline and ubiquinone is usually related to microbial defense and stress response, vital for survival in a potentially challenging high-altitude environment (Ahmad *et al.,* 2020). Key pathways such as oxidative phosphorylation and sulfur metabolism highlight the efficiency in energy production and nutrient cycling, crucial in nutrient-limited high-altitude conditions.

At the lowest altitude of 600 m, the dominance of pathways related to the breakdown of plant-derived compounds, such as geraniol, limonene, and pinene degradation (Ndao & Adjallé, 2023) suggests a closer interaction between plants and microbes, likely due to the more direct influence of vegetation at this altitude. Basic cellular processes are represented by pathways like D-alanine metabolism and steroid biosynthesis, suggesting a diverse microbial community engaged in various ecological functions. Moreover, phosphonate and phosphinate metabolism pathways at this altitude underline roles in phosphorus cycling, essential for both plant growth and microbial activity within soil ecosystems (Bi *et al.*, 2022). Altogether, the microbial community at this altitude seems highly adapted to decomposing plant materials, thereby facilitating nutrient release and availability for plant uptake.

Overall, the functional prediction analysis underscores the adaptive metabolic strategies employed by bacterial communities across different altitudes. These strategies are integral to supporting the growth and productivity of *C. arabica* by enhancing nutrient cycling, organic matter decomposition, and resilience to environmental stressors. The differential abundance of these KEGG pathways not only reflects the unique environmental conditions at each altitude but also highlights the potential of utilizing these microbial functions for sustainable coffee cultivation and ecosystem management.

3.5. Conclusion

In summary, this study highlights the significant impact of elevation on the microbiome of *C. arabica* soils within agroforestry systems. At different elevations, particularly at 800 m with predominantly legume trees, distinct microbial communities and soil chemistry profiles emerged, suggesting elevation-specific microbial functions likely related to nutrient cycling and plant growth. Despite the variations in microbial biomass and diversity, plant growth-promoting functions remained consistent across the altitudinal gradient. The study offers valuable insights into sustainable coffee cultivation, emphasizing the role of the soil microbiome in ecosystem health. Understanding these complex microbial interactions paves the way for developing climate-smart agricultural practices that leverage natural processes for improved crop resilience and productivity, aligning with the goals of sustainable agroecosystem management.

3.6. References

Abdel-Azeem, A.M., Abdel-Azeem, M.A., Darwish, A.G., *et al*. *Fusarium*: Biodiversity, ecological significances, and industrial applications. In: Yadav, A., Mishra, S., Singh, S., *et al*. (eds). *Springer*, Cham. 2019, 201-261, https://doi.org/10.1007/978-3-030-10480-1_6.

Abdelfattah, A., Malacrinò, A., Wisniewski, M., *et al*. Metabarcoding: A powerful tool to investigate microbial communities and shape future plant protection strategies. *Biol Control*. 2018, 120, 1-10, https://doi.org/10.1016/j.biocontrol.2017.07.009.

Ahmad, F., Zhu, D., Sun, J., *et al*. Bacterial chemotaxis: A way forward to aromatic compounds biodegradation. *Environ Sci Eur*. 2020, 32, 52. https://doi.org/10.1186/s12302-020-00329-2.

Ajala, O.A., Ajibade, F.O., Oluwadipe, O.R., *et al*. Microbial impact on climate-smart agricultural practices. In: Kumar, A., Singh, J., Ferreira, L.F.R. (eds). *Kidlington*, Duxford, Cambrige MA. 2022, 203–236, https://doi.org/10.1016/B978-0-323-90571-8.00009-2.

Albuquerque, L., França, L., Rainey, F.A., *et al. Gaiellaocculta* gen. Nov., sp. Nov., a novel representative of a deep branching phylogenetic lineage within the class Actinobacteria and proposal of Gaiellaceae fam. Nov. and Gaiellales ord. Nov. *Syst App Microbiol*. 2011, 34, 595–599, https://doi.org/10.1016/j.syapm.2011.07.001.

Altaf, R., Rauf, C.A., Naz, F., *et al*. Surveillance and morphological characterization of *Fusarium* isolates associated with lentil wilt. *Pak J Phytopathol*. 2014, 26, 85-90.

Arif, I., Batool, M., Schenk, P.M., *et al*., Plant microbiome engineering: Expected benefits for improved crop growth and resilience. *Trends Biotech*. 2020, 38, 1385-1396, https://doi.org/10.1016/j.tibtech.2020.04.015.

Avelino, J., Allinne, C., Cerda, R., *et al*. Multiple- disease system in coffee: From crop loss assessment to sustainable management. *Ann Rev Phytopathol*. 2018, 56, 611–635, https://doi.org/10.1146/annurev-phyto-080417-050117.

Avelino, J., Vilchez, S., Segura-Escobar, M.B., *et al*. Shade tree *Chloroleucon eyrycyclum* promotes coffee leaf rust by reducing uredospore wash-off by rain. *Crop Protect.* 2020, 129, 105038, https://doi.org/10.1016/j.cropro.2019.105038.

Barbera, P., Kozlov, A.M., Czech, L., *et al*. EPA-ng: Massively parallel evolutionary placement of genetic sequences. *Syst Biol*. 2019, 68, 365–369, https://doi.org/10.1093/sysbio/syy054.

Bardgett, R.D, van der Putten, W.H. Belowground biodiversity and ecosystem functioning. *Nature*. 2014, 515, 505–511, https://doi.org/10.1038/nature13855.

Bez, C., Esposito, A., Musonerimana, S., *et al*. Comparative study of the rhizosphere microbiome of *Coffea arabica* grown in different countries reveals a small set of prevalent and keystone taxa. *Rhizosphere*. 2023, 25, 100652, https://doi.org/10.1016/j.rhisph.2022.100652.

Bi, W.X., Weng, B.S., Dengua, Y., *et al*. Responses of phosphate-solubilizing microorganisms mediated phosphorus cycling to drought-flood abrupt alternation in summer maize field soil. *Front Microbiol*. 2022, 12, https://doi.org/10.3389/fmicb.2021.768921.

Bill, M., Chidamba, L., Gokul, J.K., *et al*. Bacterial community dynamics and functional profiling of soils from conventional and organic cropping systems. *App Soil Ecol*. 2021, 157, 103734, https://doi.org/10.1016/j.apsoil.2020.103734.

Bolyen, E., Rideout, J.R., Dillon, M., *et al*. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotech*. 2019, 37, 852–857, https://doi.org/10.1038/s41587-019-0209-9.

Bote, A.D., Struik, P.C. Effects of shade on growth, production and quality of coffee (*Coffea arabica*) in Ethiopia. *J Hortic For*. 2011, 3, 336–341, Available from: Microsoft Word - Bote and Struik PDF (wur.nl).

Bullergahn, V.B, Menezes, K.M.S., Veloso, T.G.R., *et al*. Diversity of potential nitrogen-fixing bacteria from rhizosphere of the *Coffea arabica* L. and *Coffea canephora* L. 3 *Biotech*. 2024, 14, 27, https://doi.org/10.1007/s13205-023-03875-7

Cabrera-Rodríguez, A., Trejo-Calzada, R., la Peña, C.G., *et al*. A metagenomic approach in the evaluation of the soil microbiome in coffee plantations under organic and conventional production in tropical agroecosystems. *Emirates J Food Agric*. 2020, 32, 263-270, https://doi.org/10.9755/ejfa.2020.v32.i4.2092.

Caldwell, A.C., Silva, L.C.F., da Silva, C.C., *et al*. Prokaryotic diversity in the rhizosphere of organic, intensive, and transitional coffee farms in Brazil. *PLoS One*. 2015, 10, e0106355, https://doi.org/10.1371/journal.pone.0106355.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., *et al*. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016, 13, 581–583, https://doi.org/10.1038/nmeth.3869.

Cassamo, C.T., Draper, D., Romeiras, M.M., *et al*. Impact of climate changes in the suitable areas for *Coffea arabica* L. production in Mozambique: Agroforestry as an alternative management system to strengthen crop sustainability. *Agric Ecosyst Environ*. 2023, 346, 103841, https://doi.org/10.1016/j.agee.2022.108341.

Cassamo, C.T., Mangueze, A.V.J., Leitão, A.E., *et al*. Shade and altitude implications on the physical and chemical attributes of green coffee beans from Gorongosa Mountain, Mozambique. *Agronomy*. 2022, 12, 2540, https:// doi.org/10.3390/agronomy12102540.

Catania, V., Bueno, R.S., Alduina, R., *et al.* Soil microbial biomass and bacterial diversity in southern European regions vulnerable to desertification*. Ecol Indic. 2022,* 145, 109725, https://doi.org/10.1016/j.ecolind.2022.109725.

Challacombe, J.F., Hesse, C.N., Bramer, L.M., *et al*. Genomes and secretomes of *Ascomycota* fungi reveal diverse functions in plant biomass decomposition and pathogenesis. *BMC Genomics*. 2019, 20, 976, https://doi.org/10.1186/s12864-019-6358-x.

Chen. C., Liua, W., Jiang, X., *et al*. Effects of rubber-based agroforestry systems on soil aggregation and associated soil organic carbon: Implications for land use. *Geoderma.* 2017, 299, 13-24, https://doi.org/10.1016/j.geoderma.2017.03.021.

Chen, S., Qi, G., Ma, G., *et al*. Biochar amendment controlled bacterial wilt through changing soil chemical properties and microbial community. *Microbiol Res*. 2020, 231,126373, https://doi.org/10.1016/j.micres.2019.126373.

Choi, K., Khan, R., Lee, S.W. *et al*. Dissection of plant microbiota and plant-microbiome interactions. *J Microbiol*. 2021, 59, 281-291, https://doi.org/10.1007/s12275-021-0619-5.

Chow, C., Padda, K.P., Puri, A., *et al*. An archaic approach to a modern issue: Endophytic archaea for sustainable agriculture. *Curr Microbiol*. 2022, 79, 322, https://doi.org/10.1007/s00284-022-03016-y.

Claessen, M.E.C. Manual de análise de solos, 2ª edição. *Embrapa Solos*, Rio de Janeiro. 1997.

Czech, L., Barbera, P., Stamatakis, A., *et al*. Genesis and Gappa: Processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics*. 2020, 36, 3263–3265, https://doi.org/10.1093/bioinformatics/btaa070.

DaMatta, F.M. Ecophysiological constraints on the production of shaded and unshaded coffee: a review. *Field Crops Res*. 2004, 86, 99–114, https://doi.org/10.1016/j.fcr.2003.09.001.

DaMatta, F.M., Avila, R.T., Cardoso, A.A., *et al*. Physiological and agronomic performance of the coffee crop in the context of climate change and global warming: a review. *J Agric Food Chem*. 2018, 66, 5264-5274, https://doi.org/10.1021/acs.jafc.7b04537.

Davis, A.P., Gargiulo, R., Almeida, I.N.D.M., *et al*. Hot coffee: The identity, climate profiles, agronomy, and beverage characteristics of *C. racemosa* and *C. zanguebariae*. *Front Sustain Food Syst.* 2021*,* 5, 740137, https://doi.org/.3389/fsufs.2021.740137.

Davis, A.P., Rakotonasolo, F. Six new species of coffee (*Coffea*) from northern Madagascar*. Kew Bull*. 2021, 76, 497-511, https://doi.org/10.1007/s12225-021-09952-5.

Davis, A.P., Tosh, J., Ruch, N., *et al*. Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea. Bot J Linn Soci*. 2011,167, 357–377, https://doi.org/10.1111/j.1095-8339.2011.01177.x.

de Carvalho, A.F., Fernandes-Filho, E.I., Daher, M., *et al.* Microclimate and soil and water loss in shaded and unshaded agroforestry coffee systems. *Agroforest Syst*. 2021, 5, 119–134, https://doi.org/10.1007/s10457-020-00567-6.

de Sousa, L.P., Guerreiro-Filho, O., Mondego, J.M.C., *et al.*The rhizosphere microbiomes of five species of coffee trees. *Microbiol Spectrum*. 2022a, 10, https://doi.org/10.1128/spectrum.00444-22.

de Souza, J.P., de Araújo Pereira, A.P., Pedrinho, A., *et al*. Land use and roles of soil bacterial community in the dissipation of atrazine. *Sci Total Env*. 2022b, 827,154239, https://doi.org/10.1016/j.scitotenv.2022.154239.

de Sousa, L.P., Guerreiro-Filho, O., Costa Mondego, J.M., *et al*. Differences between the leaf mycobiome of *Coffea arabica* and wild coffee species and their modulation by caffeine/chlorogenic acid content. *Microorg*. 2021, 9, 2296, https://doi.org/10.3390//microorganisms9112296.

Decaëns, T., Jiménez, J.J., Gioia, C., *et al*. The values of soil animals for conservation biology. *Eur J Soil Biol.* 2006*,* 42, S23–S38, https://doi.org/10.1016/j.ejsobi.2006.07.001.

Di Fidio, N., Minonne, F., Antonetti, C., *et al*. *Cutaneotrichosporon oleaginosum*: Aversatile whole-cell biocatalyst for the production of single-cell. *Catalysts.* 2021, 11, 1291, https://doi.org/10.3390/catal11111291.

Douglas, G., Maffei, V., Zaneveld, J., *et al*. PICRUSt2 for prediction of metagenome functions. *Nat Biotech*. 2020, 38, 1-5, https://doi.org/10.1038/s41587-020-0548-6.

Duong, B., Marraccini, P., Maeght, J.L., *et al*. Coffee microbiota and its potential use in sustainable crop management. A review. *Front Sustain Food Syst.* 2020*,* 4, 607935, https://doi.org/10.3389/fsufs.2020.607935.

Eddy, S.R. Accelerated Profile HMM Searches. *PloS Comput Biol*. 2011, 7, e1002195, https://doi.org/10.1371/journal.pcbi.1002195.

Egnér, H., Riehm, H., Domingo, W.R., *et al*. Untersuchungen über die chemische Bodenanalyse als Grundlage für die Beurteilung des Nährstoffzustandes der Böden. II. Chemische Extraktionsmethoden zur Phosphor- und Kaliumbestimmung. K. *Lantbr Ann*. 1960, 26, 199-215.

Ezeokoli, O.T., Bezuidenhout, C.C., Maboeta, M.S., *et al*. Structural and functional differentiation of bacterial communities in post-coal mining reclamation soils of South Africa: Bioindicators of soil ecosystem restoration. *Sci Rep*. 2020, 10, 1759, https://doi.org/10.1038/s41598-020-58576-5.

FAO, ITPS, GSBI, SCBD, EC. State of knowledge of soil biodiversity - status, challenges and potentialities, Report 2020. *FAO*, Rome. 2020.

Fazi, S., Butturini, A., Tassi, F., *et al*. Biogeochemistry and biodiversity in a network of saline–alkaline lakes: Implications of ecohydrological connectivity in the Kenyan Rift Valley. *Ecohydrol Hydrobiol.* 2018, 18, 96-106, https://doi.org/10.1016/j.ecohyd.2017.09.003.

Fierer, N., Bradford, M.A., Jackson, R.B., *et al*. Toward an ecological classification of soil bacteria. *Ecol*. 2007, 88, 1354-1364, https://doi.org/10.1073/pnas.1215210110

Fierer, N., Leff, J.W., Adams, B.J., *et al.* Effects of three regeneration methods on the growth and bacterial community diversity of *Populus* × *euramericana*. *Plos One*. 2022, 17, e0273306, https://doi.org/10.1371/journal.pone.0273306.

Fu, Q., Lai, J.L., Ji, X.H., *et al*. Alterations of the rhizosphere soil microbial community composition and metabolite profiles of *Zea mays* by polyethylene-particles of different molecular weights. *J Hazard Mater*. 2022, 423, 127062, https://doi.org/10.1016/j.jhazmat.2021.127062.

Ge, Y., Zhang, f., Xie, C., *et al*., Effects of different altitudes on *Coffea Arabica* rhizospheric soil chemical properties and soil microbiota. *Agronomy*. 2023, 13, 471, https://doi.org/10.3390/agronomy13020471.

Giweta, M. Role of litter production and its decomposition, and factors affecting the processes in a tropical forest ecosystem: A review. *J Ecol Environ*. 2020, 44, 11, https://doi.org/10.1186/s41610-020-0151-2.

Gomes, L.C, Bianchi, F.J.J.A., Cardoso, I.M., *et al*. Agroforestry systems can mitigate the impacts of climate change on coffee production: A spatially explicit assessment in Brazil. *Agric Ecosyst Environ*. 2020, 294, 106858, https://doi.org/10.1016/j.agee.2020.106858.

Gonzalez, C.G., Van Cauwelaert, E.M., Boyer, D., *et al*. High-order interactions maintain or enhance structural robustness of a coffee agro-ecosystem network. *BioRxiv*. 2021, 47, https://doi.org/10.1016/j.ecocom.2021.100951.

Gota, H.G., Madalcho, A.B., Kerse, B.L., *et al*. The impact of native trees, *Cordia africana* and *Ficus sur*, and the economically valuable *Manihot esculenta* on soil chemical properties in an agroforestry system. *Trees People*. 2024, 15, 100471, https://doi.org/10.1016/j.tfp.2023.100471.

Guerrero, R., Margulis, L., Berlanga, M., *et al*. Symbiogenesis: The holobiont as a unit of evolution. *Int J Microbiol*. 2013, 16, 133–143, https://doi.org/10.2436/20.1501.01.188.

Guimarães, R.J., Borem, F.M., Shuler, J., *et al*. Coffee growing and post-harvest processing. In: Farah A (ed). *RSC*, London. 2019, 150–178, https://doi.org/10.1039/9781782622437-00026.

Hassani, M.A., Durán, P., Hacquard, S., *et al*. Microbial interactions within the plant holobiont. *Microbiome*. 2018, 6, 58,https://doi.org/10.1186/s40168-018-0445-0.

Harsonowati, W., Marian, M., SuronoNarisawa, K., *et al*. The effectiveness of a dark septate endophytic fungus, *Cladophialophora chaetospira* sk51, to mitigate strawberry *fusarium* wilt disease and with growth promotion activities. *Front Microbiol*. 2020, 11, 00585, https://doi:10.3389/fmicb.2020.00585.

ICO - International Coffee Organization. Monthly coffee market report 2020/21. 2022. Available at: https:// www.ico.org/Market-Report-21-22-e.asp (accessed on 26 March 2024).

ICO - International Coffee Organization. Coffee Market Report 2021/2022. 2023. Available at: https://icocoffee.org/specialized-reports/ (accessed 20 September 2024).

IPCC - Intergovernmental Panel on Climate Change. Climate change 2022 - impacts, adaptation and vulnerability. Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change. *Cambridge University Press*, Cambridge, New York. 2022.

Jenkins, S.N., Waite, I.S., Blackburn, A., *et al.* Actinobacterial community dynamics in long term managed grasslands. *Antonie van Leeuwenhoek*. 2009, 5, 319–334, https://doi.org/10.1007/s10482-009-9317-8.

Joergensen, R.G. The fumigation-extraction method to estimate soil microbial biomass: Calibration of the KEC value. *Soil Biol Biochem*. 1996, 28, 25–31, https://doi.org/10.1016/0038-0717(95)00102-6.

Joergensen, R.G., Mueller, T. The fumigation-extraction method to estimate soil microbial biomass: Calibration of the KEN value. *Soil Biol Biochem*. 1996, 28, 3–37, https://doi.org/10.1016/0038-0717(95)00101-8.

Johnson, J.M., Ludwig, A., Furch, A.C.U., *et al*. The beneficial root-colonizing fungus *Mortierella hyaline* promotes the aerial growth of arabidopsis and activates calcium-dependent responses that restrict *Alternaria brassicae* induced disease development in roots. *Mol Plant Microbe Interact.* 2019*,* 32, 351–363, https://doi.org/10.1094/MPMI-05-18-0115-R.

Kandeler, E. Urease activity by colorimetric technique. In: Schinner, F., Kandeler, E., Öhlinger, R., *et al.* (eds). *Springer-Verlag*. Berlin. 1995, 171-174.

Kim, H., Park, Y-H., Yang, J.E., *et al*. Analysis of major bacteria and diversity of surface soil to discover biomarkers related to soil health. *Toxics*. 2022, 10, 117, https://doi.org/10.3390/toxics10030117.

Koutouleas, A., Sarzynski, T., Bertrand, B., *et al*. Shade effects on yield across different *Coffea arabica* cultivars – How much is too much? A meta-analysis. *Agron Sustain Dev*. 2022b, 42, 55, https://doi.org/10.1007/s13593-022-00788-2.

Koutouleas, A., Sarzynski, T., Bordeaux, M., *et al.* Shaded-coffee: A nature-based strategy for coffee production under climate change? *A review. Front Sustain Food Syst*. 2022a, 6, 877476, https://doi.org/10.3389/fsufs.2022.877476.

Lazcano, C., Boyd, E., Holmes, G., *et al*. The rhizosphere microbiome plays a role in the resistance to soil-borne pathogens and nutrient uptake of strawberry cultivars under field conditions. *Sci Rep.* 2021*,* 11, 3188, https://doi.org/10.1038/s41598-021-82768-2.

Li, F., Chen, L., Redmile-Gordon, M., *et al*. *Mortierella elongata*'s roles in organic agricultura and crop growth promotion in a mineral soil. *Land Degrad Dev*. 2018, 29, 1642–1651, https://doi.org/10.1002/ldr.2965.´

Louca, S., Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics*. 2018, 34, 1053–1055, https://doi.org/10.1093/bioinformatics/btx701.

Marian, M., Licciardello, G., Vicelli, B., *et al*. Ecology and potential function sof plant-associated microbial communities in cold environments. *FEMS Microbiol Ecol*. 2022, 98, fiab161, https://doi.org/10.1093/femsec/fiab161.

Matos, A., Barraza, L., Ruiz-Mallén, I., *et al.* Linking conservation, community knowledge, and adaptation to extreme climatic events: A case study in Gorongosa National Park, Mozambique. *Sustain*. 2021, 13, 6478, https://doi.org/10.3390/su13116478.

Mintesnot, A., Dechassa, N. Effect of altitude, shade, and processing methods on the quality and biochemical composition of green coffee beans in Ethiopia. *East Afr J Sci*, 2018, 12, 87–100, https://doi.org/10.20372/eajs.v12i2.495.

Ndao, A., Adjallé, K. Overview of the biotransformation of limonene and α-pinene from wood and citrus residues by microorganisms. *Waste*. 2023, 1, 841-859, https://doi.org/10.3390/waste1040049.

Newson, J., Vandermeer, J., Perfecto, I., *et al*. Differential effects of ants as biological control of the coffee berry borer in Puerto Rico. *Biol Control.* 2021*,*160, 104666, https://doi.org/10.1016/j.biocontrol.2021.104666.

Nikolaou, G., Neocleous, D., Christou, A., *et al*. Implementing sustainable irrigation in water-scarce regions under the impact of climate change. *Agronomy.* 2020*,* 10, 1120, https://doi.org/10.3390/agronomy10081120.

Ochoa-Henriquez, V.H., Faggioli, V., Gómez-Godínez, L.J., *et al*. Colombian coffee (*Coffea arabica* L.) plantations: a taxonomic and functional survey of soil fungi. *Front Sustain Food Syst*. 2024, 8, 1345383, https://doi.org/10.3389/fsufs.2024.1345383.

Oliveira, M.N.V., Santos, T.M.A., Vale, H.M.M., *et al*. Endophytic microbial diversity in coffee cherries of *Coffea arabica* from south eastern Brazil. *Can J Microbiol*. 2013, 59, 221-230, https://doi.org/10.1139/cjm-2012-0674.

Olsen, S.R., Cole, C.V., Sterling, R., *et al*. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. *USDA Circular*. 1954.

Ozimek, E., Jaroszuk-Ściseł, J., Bohacz, J., *et al*. Synthesis of indole acetic acid, gibberellic acid and ACC-deaminase by *Mortierella* strains promote winter wheat seedlings growth under different conditions. *Int J Mol Sci*. 2018,19, 30340353, https://doi.org/10.3390/ijms19103218.

Paudel, M., Parajuli, K., Regmi, S., *et al*. Effect of altitude and shade on production and physical attributes of coffee in Gulmi, Syangja and Palpa districts of Nepal. *J Agric Nat Resour*. 2021, 4, 222–238, https://doi.org/10.3126/janr.v4i1.33275.

Pawłowska, J., Okrasińska, A., Kisło, K., *et al.* Carbon assimilation profiles of mucorale an fungi show their metabolic versatility. *Sci Rep*. 2019, 9, 11864, https://doi.org/10.1038/s41598-019-48296-w.

Perfecto, I., Vandermeer, J., Philpott, S.M., *et al.* Complex ecological interactions in the coffee agroecosystem. *Ann Rev Ecol Evol System*. 2014, 45, 137–158, https://doi.org/10.1146/annurev-ecolsys-120213-091923.

Phour, M., Sindhu, S.S. Soil salinity and climate change: Microbiome-based strategies for mitigation of salt stress to sustainable agriculture. In: Parray, J.A. (ed) *Springer,* Cham. 2023, 191-243, https://doi.org/10.1007/978-3-031-21079-2_13.

Pringle, R. Upgrading protected areas to conserve wild biodiversity. *Nature*. 2017, 546, 91–99, https://doi.org/10.1038/nature22902.

Quast, C., Pruesse, E., Yilmaz, P., *et al*. The Silva ribosomal RNA gene database project: Iimproved data processing and web-based tools. *Nucleic Acids Res*. 2013, 41, D590-6, https://doi.org/10.1093/nar/gks1219.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Version 4.1.1. R Core *Team, Vienna*. 2021, URL https://www.R-project.org/.

Sandhya, M., Ziqiu, L., Shimei, P., *et al*. Recent advanced technologies for the characterization of xenobiotic-degrading microorganisms and microbial communities. *Front Bioeng Biotech*. 2021, 9, 63205910, https://doi.org/10.3389/fbioe.2021.632059.

Sauvadet, M., Saj, S., Freschet, G.T., *et al*. Cocoa agroforest multifunctionality and soil fertility explained by shade tree litter traits. *J Appl Ecol.* 2020*,* 57, 476–487, https://doi.org/10.1111/1365-2664.13560.

Segata, N., Izard, J., Waldron, L., *et al*. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011, 12, R60, https://doi.org/10.1186/gb-2011-12-6-r60.

Siles, J.A., Margesin, R. Abundance and diversity of bacterial, archaeal, and fungal communities along an altitudinal gradient in Alpine forest soils: what are the driving factors? *Microb Ecol*. 2016, 72, 207–220, https://doi.org/10.1007/s00248-016-0748-2.

Sileshi, G.W., Mafongoya, P.L., Akinnifesi, F.K., *et al*. Agroforestry: fertilizer trees. In: Van Alfen, N. (ed). *Encyc Agric and Food System*. 2014, 222-234.

Singh, S., Dubey, O.P., Upadhyay, R.S., *et al*. Evaluation of the potential of endophytic *Gliocladiopsis* sp. as a biological control agent against *Sclerotinia* *sclerotiorum* in *Brassica juncea*. *Biol Control.* 2018, 123, 99-106, https://doi.org/.1016/j.biocontrol.2018.05.012.

Singh, S., Upadhyay, R.S., Sharma, S., *et al*. Endophytic *Gliocladiopsi*s sp. confers drought stress tolerance in potted wheat plants (*Triticum aestivum*). *Braz J Microbiol*. 2017, 48, 489-497, https://doi.org/10.1016/j.bjm.2017.01.004.

Sokol, N.W., Slessarev, E., Marschmann, G.L., *et al*. Life and death in the soil microbiome: how ecological processes influence biogeochemistry*. Nat Rev Microbiol*. 2022, 20, 415–430, https://doi.org/10.1038/s41579-022-00695-z.

Stahl, D.A., de la Torre, J.R. Physiology and diversity of ammonia-oxidizing archaea. *Annu Rev Microbiol*. 2012, 63, 83-101, https://doi.org/ 10.1146/annurev-micro-092611-150128.

Stalmans, M., Victor, M. Forest cover on Gorongosa mountain. Assessment of satellite imagery 2019. 2020. https://gorongosa.org/wp-content/uploads/2020/10/SerraGorongosa_2019ForestCover_15May2020.pdf (accessed 19 September 2024).

Stellner, N.I., Rerop, Z.S., Mehlmer, N., *et al*. Expanding the genetic toolbox for *Cutaneo trichosporon oleaginosus* employing newly identified promoters and a novel antibiotic resistance marker. *BMC Biotechnol*. 2023, 23, 40, https://doi.org/10.1186/s12896-023-00812-7

Stenberg, B. Monitoring soil quality of arable land: Microbiological indicators. *Acta Agric Scand Sect B - Soil Plant Sci*. 1999, 49, 1–24, https://doi.org/10.1080/09064719950135669.

Sun, W., Xiao, E., Krumins, V., *et al.* Rhizosphere microbial response to multiple metal (loid)s in different contaminated arable soils indicate scrop-specific metal-microbe interactions. *Appl Environ Microbiol*. 2018, 84, e00701-18, https://doi.org/10.1128/AEM.00701-18.

Tabatabai, M.A. Soilenzymes. In: Weaver RW, Angle JS, bottomley PS (eds) methods of soil analysis, Part 2: Microbiological and biochemical properties. *Soil Sci Soc Am*, Book Ser 5, 1994, 5, 775–833.

Tapaça, I.d.P.E., Mavuque, L., Corti, R., *et al*. Genomic evaluation of *Coffea arabica* and its wild relative *Coffea racemosa* in Mozambique: Settling resilience keys for the coffee crop in the context of climate change. *Plants*. 2023, https://doi.org/10.3390/plants12102044.

Telagathoti, A., Probst, M., Mandolini, E., *et al*. *Mortierellaceae* from subalpine and alpine habitats: New species of *Entomortierella*, *Linnemannia*, *Mortierella*, *Podila* and *Tyroliella* gen nov. *Stud Mycol*. 2022, 103, 25–58, https://doi.org/10.3114/sim.2022.103.0.

Tran, D. M. Rhizosphere microbiome dataset of Robusta coffee (*Coffea canephora* L.) grown in the Central Highlands, Vietnam, based on 16S rRNA metagenomics analysis. *Data Brief*. 2022, 42, 108106, https://doi.org/10.1016/j.dib.2022.108106.

Trivedi, P., Leach, J.E., Tringe, S.G., *et al.* Plant–microbiome interactions: From community assembly to plant health. *Nat Rev Microbiol*. 2020, 18, 607–621, https://doi.org/10.1038/s41579-020-0412-1.

Truu, M., Nõlvak, H., Ostonen, I., *et al.* Soil bacterial and archaeal communities and their potential to perform N-cycling processes in soils of Boreal forests growing on well-drained peat. *Front Microbiol.* 2020*,* 11, 591358, https://doi.org/10.3389/fmicb.2020.591358.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., *et al*. The importance of the microbiome of the plant holobiont. *New Phytol*. 2015, 206, 1196–1206, https://doi.org/10.1111/nph.13312.

Vandepol, N., Liber, J., Yocca, A., *et al*. *Linnemannia elongata* (Mortierellaceae) stimulates *Arabidopsis thaliana* aerial growth and responses to auxin, ethylene, and reactive oxygen species. *Plos One*. 2022, 17, e0261908, https://doi.org/10.1371/journal.pone.0261908.

Veloso, T.G.R., da Silva, M.d.C.S., Moreira, T.R., *et al*. Microbiomes associated with *Coffea arabica* and *Coffea canephora* in four different floristic domains of Brazil. *Sci Rep*. 2023, 13, 18477, https://doi.org/10.1038/s41598-023-45465-w.

Vero, S., Garmendia, G., Garat, M.F., *et al*. *Cystofilobasidium infirmominiatum* as a biocontrol agent of postharvest diseases on apples and citrus. *Acta Hortic*. 2011, 905,169-180, https://doi.org/10.17660/ActaHortic.2011.905.

Wang, Y., Wang, H., Cheng, H., *et al*. Niche differentiation in the rhizosphere and endosphere fungal microbiome of wild *Paris polyphylla* sm. *Peer J*. 2020, 8, e8510. https://doi.org/10.7717/peerj.8510.

Ye, Y., Doak, T.G. A parsimony approach to biological pathway reconstruction / inference for genomes and metagenomes. *PLoS Comp Biol.* 2009, 5, e1000465, https://doi.org/10.1371/journal.pcbi.1000465.

Yoneda, Y., Yamamoto, K., Makino, A., *et al*. Novel plant-associated *Acidobacteria* promotes growth of common floating aquatic plants, duckweeds. *Microorg*. 2021, 9, 1133, https://doi.org/10.3390/microorganisms9061133

Yue, H., Yue, W., Jiao, S., *et al*. Plant domestication shapes rhizosphere microbiome assembly and metabolic functions. *Microbiome*. 2023, 11, 70, https://doi.org/10.1186/s40168-023-01513-1.

Zavarzin, G.A. The notion of microflora of dispersion in the carbon cycle. *J Gen Biol.* IzvAkadNauk USSR. 1970, 31, 386-393.

Zhang, H., Yuan, W., Dong, W., *et al*. Seasonal patterns of litter fall in forest ecosystem worldwide. *Ecol Complex*. 2014, 20, 240-247, https://doi.org/10.1016/j.ecocom.2014.01.003.

Zhao, F., Zhang, Y., Dong, W., *et al*. Vermicompost can suppress *Fusarium oxysporum* f. sp. lycopersici via generation of beneficial bacteria in a long-term tomato monoculture soil. *Plant Soil*. 2019, 440, 491–505, https://doi.org/10.1007/s11104-019-04104-y.

Zhou, Z., Tran, P.Q., Breister, A.M., *et al*. Metabolic: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome*, 2022, 10, 33, https://doi.org/10.1186/s40168-021-01213-8.

Chapter 4

Ecological function and diversity of endosphere microbiome in leaves and fruits of *Coffea arabica* L. across elevation and shade gradients

Modified after Obieze, C.C., Tapaça, I.d.P.E., Pereira, G.V.d.M. Partelli, F.L., Ramalho, J.C., Marques, I., Ribeiro-Barros, A.I. Ecological function and diversity of endosphere microbiome in leaves and fruits of Coffea arabica L. across elevation and shade gradients. *e-DNA, in press*

4.1. Introduction

The hologenome theory of evolution postulates that plants and animals are not independent entities, as the unit of natural selection and evolution (the holobiont) also includes their associated ("inherited") microbes (Rosenberg and Zilber-Rosenberg 2016). Indeed, the rhizospheric, endophytic, and epiphytic plant microbiome consists of diverse and structured microbial communities with functional attributes that ultimately determine plant and soil health and productivity (Berg *et al*., 2014). Additionally, with increasing food demand, climate change, and anthropogenic activities leading to the loss of arable land, the plant microbiome is emerging as a climate-smart tool for sustainable agriculture (Berg *et al*., 2014; Afridi *et al*., 2022; Negi *et al*., 2024). Therefore, research on plant-associated microorganisms is gaining momentum globally (Compant *et al*., 2019; Dastogeer *et al*., 2020; Xu *et al*., 2021; Trivedi *et al*., 2022), driven by the urgent need to explore their potential in enhancing crop resilience, improving soil health, and contributing to the sustainability of agricultural systems in the face of environmental challenges.

Coffee is one of the most traded commodities in the world, supporting the livelihoods of ca. 25 million smallholder farmers in ca. 80 tropical countries, and involving 100-125 million people across the value chain (Voora, Bermúdez, and Larrea 2019). Among the 130 *Coffea* species (Davis and Rakotonasolo 2021), only two support the coffee market: Coffea arabica L. (Arabica coffee) and Coffea canephora Pierre ex A. Froehner (Robusta coffee), contributing approximately for 57% and 43% of the annual production, respectively (ICO 2023). Despite its socioeconomic importance, the sustainability of the coffee value chain is severely constrained by climate change events, incidences of pests and diseases, labor shortages, rising prices of inputs, slow pace of breeding programs, institutional framework constraints, price volatility, and concerns over fair trade (Ngure and Watanabe 2024).

Climate change is imposing serious challenges to the coffee sector (Grüter *et al*., 2022; Koutouleas *et al*., 2022a). Recent projections foresee that the suitable land for cultivation of *C. arabica* will drastically decrease in traditionally coffee-growing regions by the middle of the century, with reductions up to 50% of the current cropping area (Grüter *et al*., 2022; Cassamo *et al*., 2023). The more climate-resilient *C. canephora* will also lose some areas suitable for cultivation (Magrach and Ghazoul 2015). Under this scenario, and considering the low plasticity and adaptive evolution of the commercial species, several approaches are being implemented to ensure the sustainability of this important crop. Among them, the introduction of wild-relatives in the value chain (Davis *et al.*, 2021; Tapaça *et al.*, 2023), the use of agroforestry systems (AFS) (Koutouleas *et al.*, 2022a,b; Cassamo *et al.*, 2023), as well as the adoption of nature-based solutions are crucial for ensuring the long-term sustainability of the coffee crop (Vogt 2019; de Resende *et al.*, 2021). The use of the plant microbiome to improve plant growth, environmental resilience, health, coffee beverage attributes, and waste management is one of the most promising strategies (Duong *et al.*, 2020; da Silva Vale *et al.*, 2021; Asad *et al.*, 2023; Castillo-González *et al.*, 2023; de Melo Pereira *et al.*, 2024).

Coffee microbiome studies have been conducted mostly in soils, rhizosphere, and the fermentation process of coffee beans. Soils are among the most important reservoirs of biodiversity, which provide important supporting, provision, and regulating services (FAO *et al.*, 2020; Sokol *et al.*, 2022). Therefore, research in coffee soil and rhizosphere is a step forward (e.g. Caldwell *et al.*, 2015; Cabrera-Rodríguez *et al.*, 2020; Duong *et al.*, 2020; de Sousa, Guerreiro-Filho, Mondego 2022; Tran 2022; Ge *et al.*, 2023; Veloso *et al.*, 2023; Bullergahn *et al.*, 2024; Tapaça *et al.*, 2024). It is consensual that microbial diversity and richness are high, driven by soil pH and altitude, and in case of AFS, by the canopy shade species. Additionally, their functional attributes are beyond plant health and productivity. In fermentation processes, yeasts, lactic acid bacteria, and other soil-related microorganisms have been identified as the main drivers of coffee quality, influencing the accumulation of aroma-enhancing compounds such as alcohols, aldehydes, organic acids, esters, ketones, and terpenoids (da Silva Vale *et al.*, 2019, 2021; de Melo Pereira *et al.*, 2024).

More recently, the analysis of the coffee microbiome has been extended to endophytes. Veloso *et al.* (2023) reported the richness of microbiome in fruits of *C. arabica* and *C. canephora* across various floristic domains of Brazil, highlighting the importance of wild yeast for potential use in fermentation processes. Castillo-González *et al.* (2023), observed that, in comparison to conventional and organic practices, agroforestry enhances functional diversity of endophytic fungal communities in coffee leaves. Regarding root endophytes, Fulthorpe *et al.* (2020) suggest the existence of a core microbiome modulated by environmental conditions and associated to specific plant functional traits. However, a robust understanding of the structure and composition of microbial endophytic communities along different microenvironments and the relationship between belowground and aboveground communities remains elusive (Dos Santos Gomes *et al*., 2024). To address this gap and expand upon the work of Tapaça *et al*. (2024) on coffee soil microbiome, we employed a next generation sequencing approach to perform an in-depth analysis of bacterial and fungal communities in the leaves and fruits of *C. arabica* cultivated under AFS along an elevation and canopy shade gradient. This study examines variations along both elevation and canopy shade gradients, providing a comprehensive insight into the microbiomes influencing coffee plant health and productivity.

4.2. Material and Methods

4.2.1. Experimental design

The study was conducted in the Gorongosa Mountain, Gorongosa National Park, Sofala province, Mozambique (Lat. 18º 24’ 14’’S, Long. 34º 06’ 31.5’’E), using the split-plot design of Cassamo *et al*. (2022): *Coffea arabica* implanted under AFS with native trees - *Khaya anthotheca* (Welw.) C. DC., *Erythrina lysistemon* Hutch., and *Albizia adianthifolia* (Schumach.) W. Wight; along an elevation gradient (main plots) - ca. 600 m (18º 30’ 53’’ S, 34º 03’ 05’’ E), ca. 800 m (18’ 30’’ 04’’ S, 34º 02’ 58’’ E), and ca. 900 m (18º 28’ 54’’ S, 34º 02’ 43” E) above sea level (a.s.l.); and three levels of light (sub-plots) per altitude: deep shade (DS, average diurnal PPFD of 127 ± 28 μmol m^{-2} s^{-1}), moderate shade (MS, 725 ± 48 μmol m^{-2} s^{-1}), and full Sun (FS, 1268 ± 52 μmol m^{-2} s^{-1}).

4.2.2. Sampling

At each altitude and shade level, three coffee plants were randomly selected from which 6-10 leaves and fruits were detached. The samples were then disinfected using the protocol of Fernández-González *et al*. (2019) with the following washing sequence: 70% EtOH for 5 min; 3.7 % NaClO containing 0.01% Tween 20 for 3 min; and 3 rinses with sterile distilled water. For each sample, a 100 μl aliquot of the last wash water was taken to test for the presence of microbial contaminants. Aliquots were placed in Petri dishes containing growth medium for bacteria (LB) and fungi (PDA), followed by incubation at 28°C for 7 days. The absence of microbial growth in

LB and PDA medium was scored to confirm the success of surface decontamination. Leaves and fruits were kept at - 80°C until DNA extraction.

4.2.3. DNA extraction and amplicon sequencing

DNA from leaves and fruits was extracted using the DNeasy Power Soil Pro Soil DNA Isolation Kit (Qiagen, Germany City, MD, USA) following the manufacturer's instructions. DNA integrity and concentration were determined by 1% agarose gel electrophoresis and fluorometric quantification using a fluorometer (Qubit 2.0, Invitrogen, CA, USA), respectively. Amplicon libraries targeting the V4 hypervariable region of the 16S rRNA for bacteria (Bakt_341F: CCTACGGGNGGCWGCAG and Bakt_805R: GACTACHVGGGTATCTAATCC), and the ITS2 region for fungi (3F: GCATCGATGAAGAACGCAGC and 4R: TCCTCCGCTTATTGATATGC) were used following the Amplicon Metagenomic Sequencing Library Preparation guide (http://emea.support.illumina.com). Sequencing libraries were generated using the TruSEq DNA PCR-free sample preparation kit (Illumina, San Diego, CA, USA) following the manufacturer’s instructions. The final libraries were sequenced using the Illumina Miseq300 PE to generate 300 bp paired-end reads through Macrogen sequencing services (Macrogen, Seul, Korea).

4.2.4. Sequence processing, diversity and statistical analyses

The raw sequence reads were first quality checked to filter low quality reads at the 5´ and 3´ ends using trimommatic. The resulting sequences were subsequently merged, denoised into amplicon sequence variants (ASVs), and chimera checked using DADA2 (Callahan *et al*., 2016) as implemented in QIIME2 v2023.7 (Bolyen *et al*., 2019). Classification of the ASVs was based on VSEARCH consensus taxonomic classifier using the SILVA v138 reference database (Quast *et al*., 2013) and the UNITE database (Abarenkov *et al*., 2022), respectively for the 16S rRNA gene region and the ITS region. From the16S rRNA reads, 350 ASVs were classified as bacteria at the domain level taxonomic rank, while for ITS, 625 ASVs were classified as fungi. Accordingly, the final ASV table used for diversity estimation and statistical analyses contained only ASVs classified as bacteria (without chloroplast and mitochondria ASVs) or fungi. Alpha diversity estimation was based on observed ASVs, Shannon-entropy and Pielou’s evenness index as implemented in QIIME2.

The microbial community dissimilarity was determined based on Bray-Curtis distance and visualized using principal coordinates analysis (PCoA). The significance of the community differences was determined based on Permutational multivariate analysis of variance (PERMANOVA) (adonis function, vegan R package). Differentially abundant ASVs along the altitudinal and shade gradients were investigated using linear discriminant analysis (LDA) effect size (LEfSe) (Segata *et al.,* 2011). Functional characterization of the bacterial communities along the altitudinal and shade gradient was inferred using the default parameters of PICRUSt2 (Douglas *et al.,* 2020). Similarly, functional guilds were also assigned to the fungal community using FUN Guild (Nguyen *et al*., 2016). Out of 625 fungal ASVs, ecological guilds were assigned to 486, representing 77.8% of the dataset. Heat maps were used to visualize the differences in fungal ecological guilds across the shade and altitudinal gradient, while differentially abundant metacyc pathways were determined by LEfSe.

4.3. Results

4.3.1. Diversity along the altitude and shade gradient

A total of 2,531,667 high-quality reads were obtained for the 16S rRNA sequences and 2,624,960 for the ITS sequences after denoising and removal of chimeric sequences (Table S4.1). The observed Amplicon Sequence Variants (ASVs) index was significantly higher in the leaves than fruits for both bacteria ($p = 0.03$) and fungi ($p < 0.01$). A similar observation was made on investigation of Shannon-entropy for both bacteria ($p < 0.01$) and fungi ($p < 0.01$).

For the investigated gradients, the observed ASVs index was highly variable across both shade and altitudinal gradient, regardless of the microbial community compartment (leaves of fruits), with the averages ranging from less than 10 to more than 80 ASVs (Figure 4.1). Similarly, Shannon entropy and Pielou’s evenness indexes revealed low to moderate diversity for both bacteria (Table S4.2A) and fungi (Table S4.2B) in all sampled sites (Figure 4.1). Shade only had marginal influence on bacterial and fungal diversity, with exceptions at some altitudinal ranges as shown in Figure 4.1. Conversely, the observed ASVs ($p = 0.04$) and Pielou’s evenness ($p = 0.01$) indexes of fungal communities in the fruits were significantly influenced by altitude, while all three alpha diversity indexes were significantly influenced by altitude for bacterial communities in the leaves (Table S4.2B).

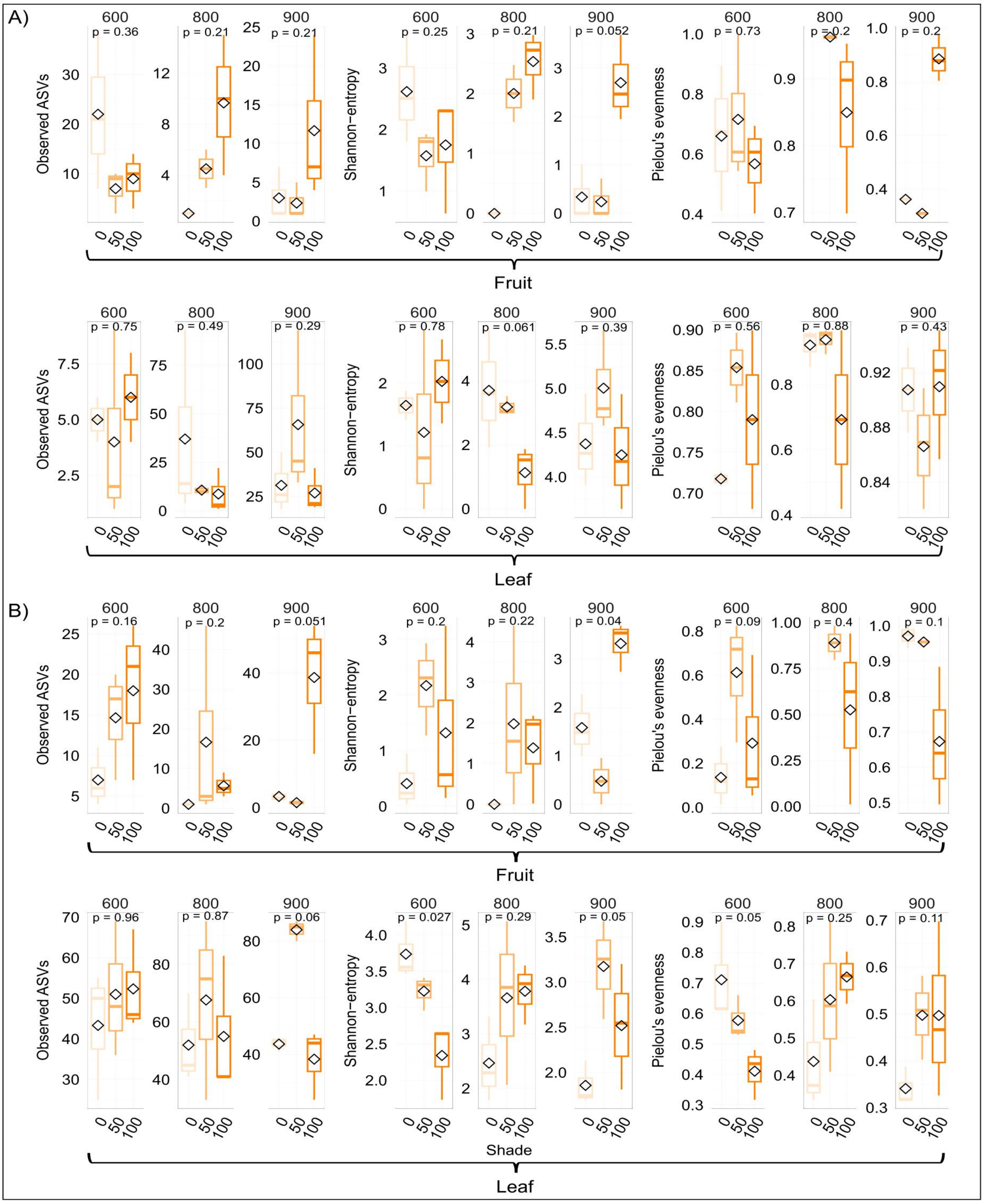


Figure 4.1 Observed ASVs, Shannon–entropy and Pielou's evenness alpha diversity indexes in fruits and leaves of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m) and under different levels of canopy shading (0%, 50% and 100%) considering bacteria (A), and fungi (B). The diamond-shapedpoints on the boxplots depict mean values. Comparisons between altitude and shade levels are based on the Kruskal-Wallis test (Bonferroni adjusted *p* < 0.05).

Principal coordinate analysis (Figure 4.2) revealed that for bacteria, 13.4% of variation was explained by PCo1 (Figure 4.2A) and 9.3% by PCo2 (Figure 4.2B). Similar results were obtained for fungi, i.e., 13.9% and 10.9%, respectively, by PCo1 and PCo2 (Figure 4.2B). Samples were widely dispersed in the PCoA space, and no specific groups could be identified. For both bacteria and fungi, the communities in the fruits and leaves partially clustered separately within the ordination space. However, based on PERMANOVA, these communities were significantly different for both bacteria (p = 0.001; R^2 = 0.06) and fungi (p = 0.001; R^2 = 0.09), but with a higher proportion of the community variation explained for fungi. Further investigation based on canopy cover and altitude revealed that the altitudinal gradient significantly influenced the bacterial and fungal community dissimilarities. Specifically, regardless the plant host compartment, the canopy-cover alone did not influence the microbial community composition, though the interaction between canopy cover and altitude significantly influenced the bacterial community in the Fruits (p = 0.002; R^2 = 0.25). In contrast, altitude was the most important variable that determined microbial community composition in the fruits and leaves, and for both bacteria and fungi. For the bacterial community in the leaves, altitude alone explained 20% of the variation (p = 0.001; R^2 = 0.20), and this observation was similar for the fungal communities both in the fruits (p = 0.005; R^2 = 0.14) and leaves (p = 0.002; R^2 = 0.18).

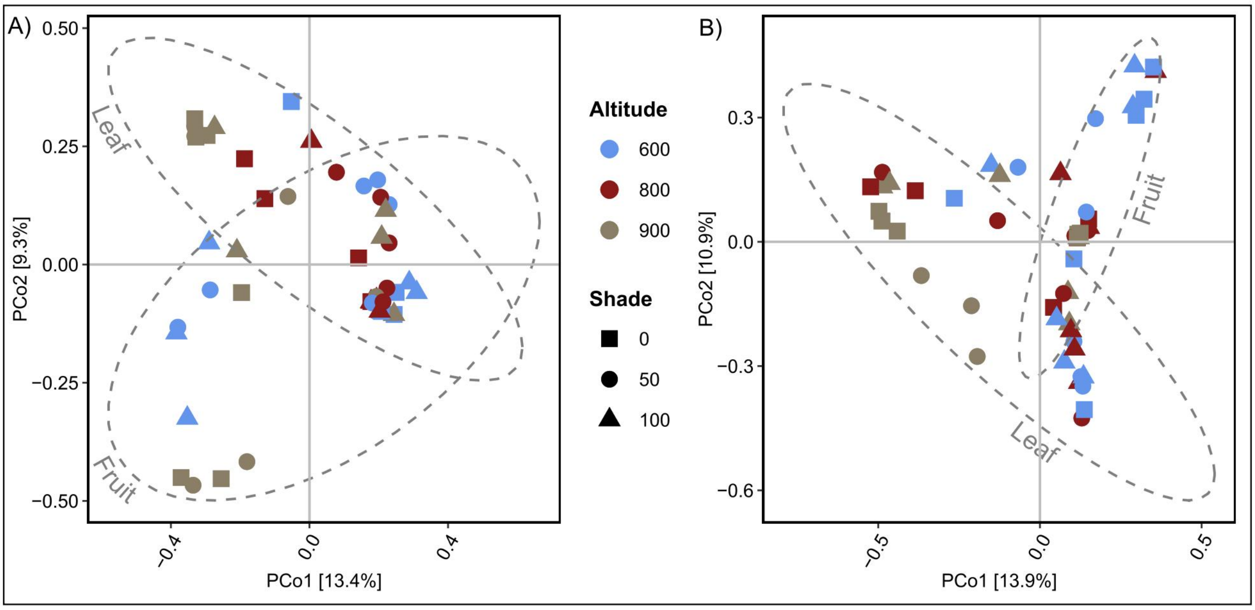


Figure 4.2 Principal Coordinate Analysis based on Bray-Curtis distance. (A) Bacteria; (B) Fungi.

4.3.2. Dominant bacterial taxa

As Taxonomic analysis revealed significant variation in bacterial taxa between fruits and leaves, as well as across the shade and altitudinal gradients (Figures 3 and 4). As depicted in Figure S1A, at 600 m, the predominant families in fruits were: (i) Enterobacteriaceae and Erwiniaceae (full sun or 0% shade); (ii) Erwiniaceae and Pseudonocardiaceae (50% shade); and (iii) Pseudonocardiaceae (100% shade). At 800 m, Oxalobacteriaceae was the only family found under full sun, while Erwiniaceae, and Nocardiaceae were predominant under moderate (50%) and dense (100%) shade, respectively. At 900 m, the predominant families were Pseudonocardiaceae (0% shade), Enterobacteriaceae (50% shade), and Erwiniaceae (100% shade). In leaves, the predominant bacterial families (Figure S1B) were: (i) at 600 m: Micrococcaceae and Beijerinckiaceae (0% and 50% shade), and Beijerinckiaceae (100% shade); (ii) at 800 m: Micrococcaceae and Nocardiaceae (full sun), Pseudonocardiaceae and Beijerinckiaceae (50% shade), and Erwinaceae (100% shade); and (iii) at 900 m: Micrococcaceae and Nocardiaceae (0% and 50% shade), and Nocardiaceae and Beijerinckiaceae (100% shade). These findings indicated that both shade levels and altitude play crucial roles in shaping the composition of bacterial communities in fruits and leaves, with distinct bacterial families dominating under varying environmental conditions.

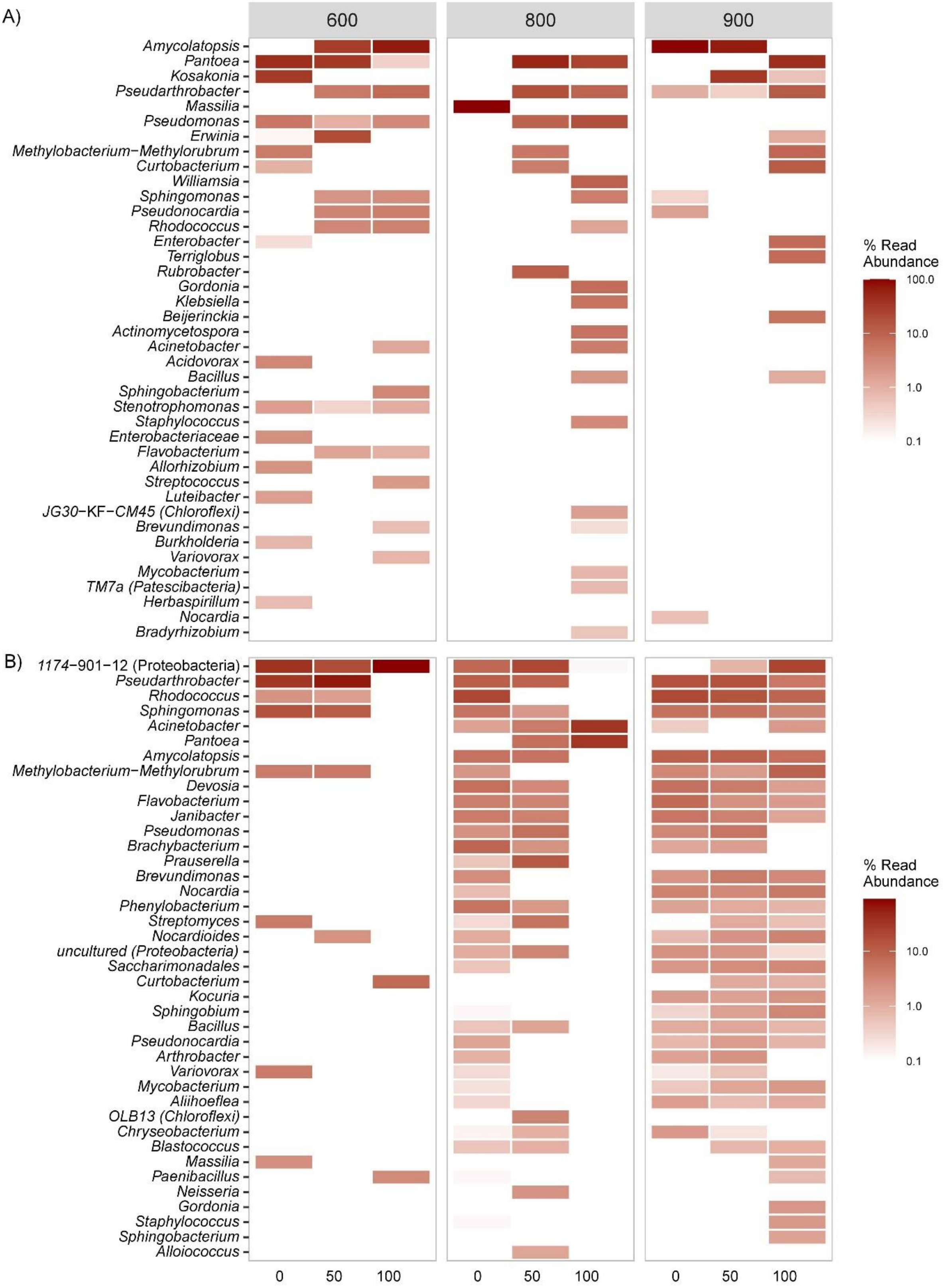


Figure 4.3 Bacterial genera-level taxonomic profile of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m) and under different levels of shade from native trees (0%, 50%, and 100%). (A) Fruits; (B) Leaves.

The variation of the taxonomic profiles was also observed at the genus level (Figure 4.3). In fruits, the most diverse and abundant genera were observed at 600 m and 0% shade: *Kosakonia, Pantoea, Pseudomonas, Acidovorax, Allorhizobium, Methylorubrum, Stenotrophomonas, Curtobacterium, Erwinia, Burkholderia, Herbaspirillum, Luteibacter, Novosphingobium, Sphingomonas, Sodalis, Aureimonas,* and *Bosea*. Some of these genera were also present, but less abundant, in fruits from plants grown under moderate and/or full shade, where *Amycolatopsis, Pseudarthrobacter, Pseudonocardia, Rhodococcus,* and *Flavobacterium* were among the most abundant. Also, under 100% shade, *Sphingobacterium, Acinetobacter, Streptococcus, Brevundimonas*, and *Variovorax* were visibly dominant compared to other levels of shade cover. At 800 m, *Pantoea, Pseudomonas*, and *Pseudarthrobacter* were among dominant genera observed in fruits from plants under 50 and 100% shade. Others included, *Methylorubrum* and *Curtobacterium* (50% shade), and *Williamsia, Gordonia, Actinomycetospora, Rhodococcus, Pseudonocardia, Sphingomonas, Nocardia, Acinetobacter, Bacillus*, *Klebsiella, Mycobacterium, Staphylococcus,* and *Brevundimonas* (100% shade). At 900 m the most diverse and abundant genera were observed in fruits from plants grown under full shade: *Kosakonia, Pantoea, Pseudomonas, Enterobacter, Methylorubrum, Curtobacterium, Erwinia, Burkholderia, Pseudarthrobacter, Lactococcus,* and *Bacillus*. In leaves, the most diverse profiles were observed at 800 m (full sun), and 900 m (all shade levels), with most of them overlapping with those identified in fruits. *Pantoea* was the most abundant genus at 800 m and 100% shade.

4.3.3. Dominant fungal taxa

The variation in fungal taxa between fruits and leaves, as well as across the shade and altitudinal gradient, was also considerable (Figures 4 and S2). As depicted in Figure S2A, the most abundant families in fruits were, (i) Eremotheciaceae (600 m and all shade levels; 800 m and 100% shade); (ii) Debaryomycetaceae (800m; 0% shade); Cladosporiaceae and Phyllostictaceae (800, 50% shade); (iii) Aspergillaceae, Saccotheciaceae, Pleosporaceae (900 and 0% shade); (iv) Cladosporiaceae, Phaeosphaeriaceae, and Amphisphaeriaceae (900 m, 50% shade); and (v) Eremotheciaceae and Aspergillaceae (900 m and 100% shade). In leaves, the predominant fungal families (Figure S2B) were: (i) Cladosporiaceae and Mycosphaerellaceae (600 m and all shade levels); (ii) Diaporthaceae (600 m, 100%

shade); (iii) Debaryomycetaceae and Nectriaceae (800m, 0% shade); (iv) Mycosphaerellaceae (800m, 50% shade); (v) Cladosporiaceae, Mycosphaerellaceae, and Trichosphaeriaceae (800 m,100% shade); and (vi) Debaryomycetaceae (900m, 100% shade).

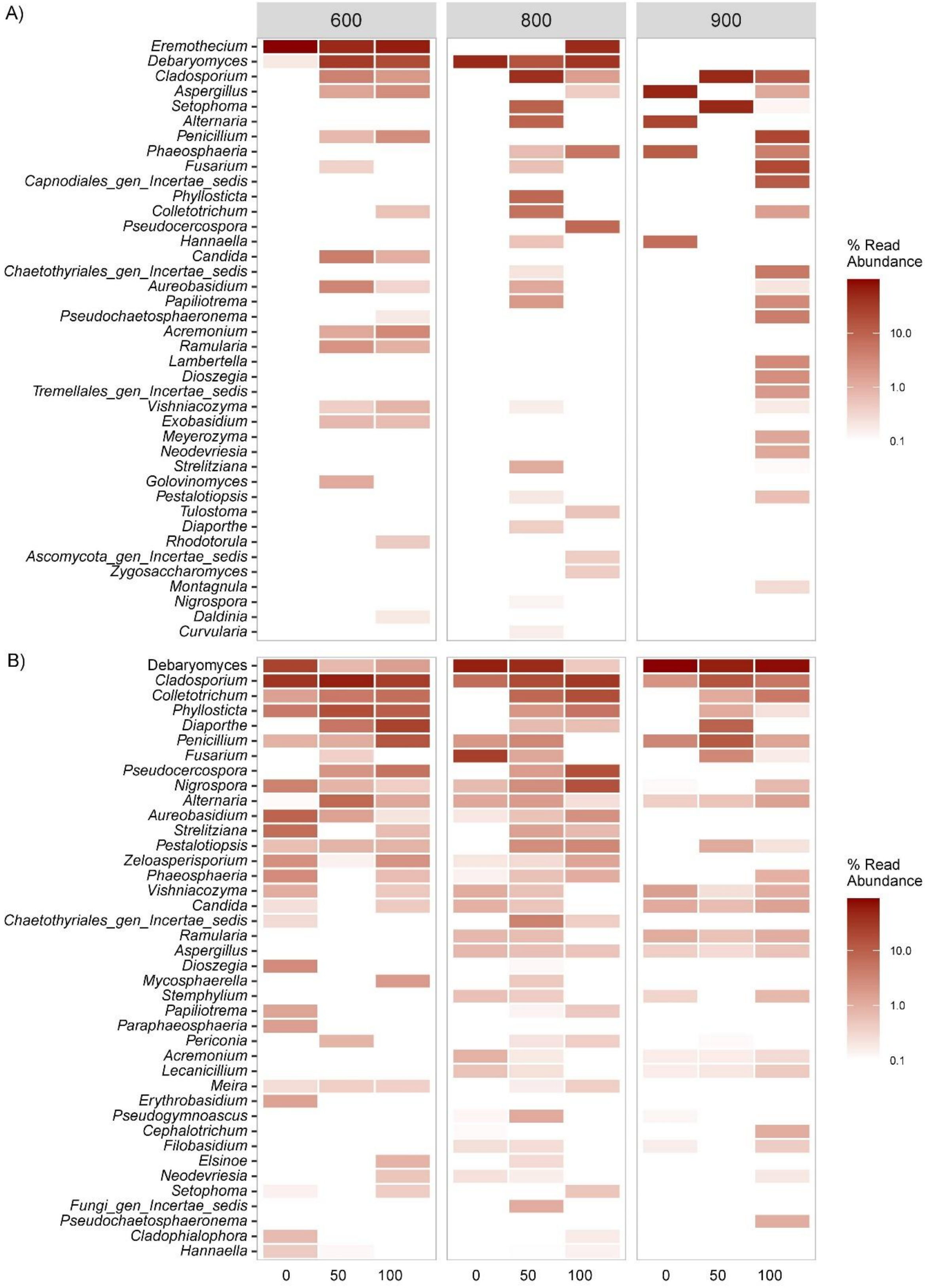


Figure 4.4 Fungal genus level taxonomic profile of *C. arabica* grown under different altitudes (600 m, 800 m, and 900 m) and under different levels of shade from native trees (0%, 50% and 100%). (A) Fruits; (B) Leaves.

In fruits, *Eremothecium* was the most prominent genus at 600 and across all shade levels, as well as at 800 under full shade (100%) (Figure S2A). The higher genera diversity and abundance was scored at 900 m and 100% shade, where more than 20 genera were identified, including *Penicillium, Fusarium, Cladosporium, Papiliotrema, Colletotrichum, Lambertella, Phaeosphaeria, Meyerozyma, Aureobasidium, Pestalotiopsis, Strelitziana, Vishniacozyma, Rhodotorula, Montagnula,* and *Setophoma*. In leaves, the diversity and abundance profiles were more uniform along the elevation and shade gradient (Figure S2B). The most abundant genera were *Debaryomyces*, *Cladosporium*, *Penicillium* and, to some extent, *Fusarium* and *Diaporthe* (600m and 100% shade), Phyllosticta (600m and 0 and 50% shade), and *Debaryomyces* (800 m and 0 and 50% shade; 900 m and all shade levels).

At the genus taxonomic rank, several bacterial and fungal genera were differentially abundant across altitudinal levels (Figure 5). In the case of bacteria, *Rhodococcus* and *Prauserella* were the most predominant biomarker phylotypes at 900 and 800 m, respectively, while no differentially phylotypes were identified at 600 m. *Debaryomyces*, *Aspergillus*, and *Aureobasidium* were the most differentially abundant fungi at 900, 800, and 600 m, respectively in the leaves, whereas *Eremothecium*, *Debaryomyces* and *Dioszegia* were the top biomarkers that differentiated the fruits fungal communities along the altitudinal gradient.

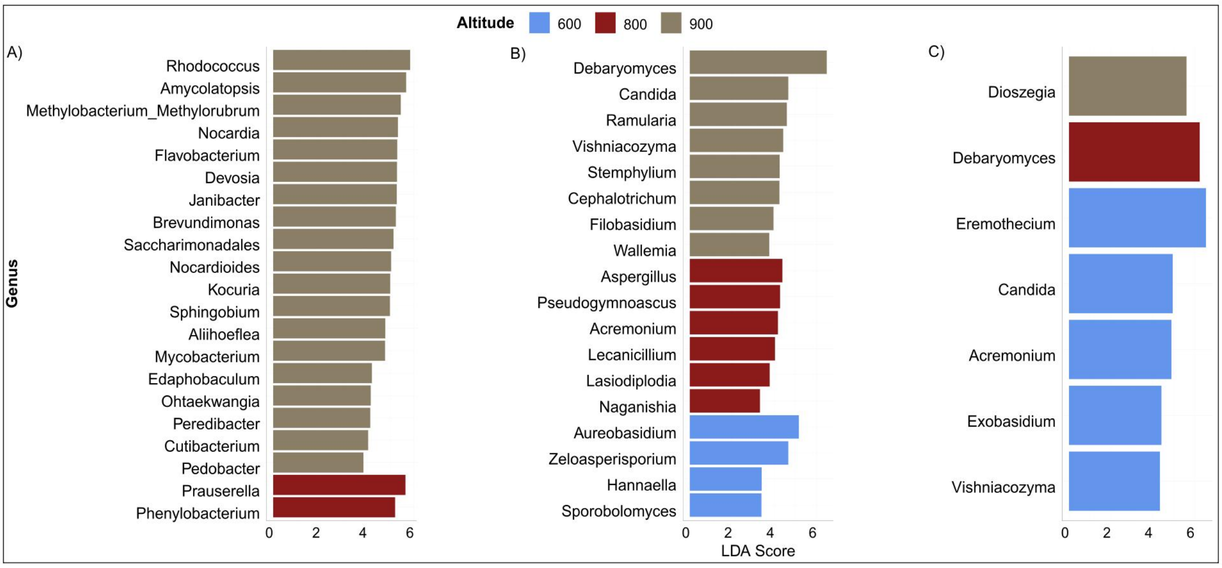


Figure 4.5 Differentially abundant bacteria in fruits (A), and fungi in leaves (B), and fruits (C) at the genus rank, across altitudes (600 m, 800 m, and 900 m). The differentially abundant data for bacteria and fungi (LDA ≥ 2.0; FDR-adjusted *p* ≤ 1.0) are a subset that was sorted according to adjusted *p*-

values, LDA scores and altitudinal groupings. The complete list of differentially abundant genera is presented in Tables S4.6 and S4.7.

4.3.4. Functional attributes

The PICRUSt2-based prediction of the bacterial community functions across shades and altitude levels identified 74 metacyc metabolic pathways that were differentially abundant along the altitudinal gradient, while no significant difference were observed in relation to canopy shades (Fig. 6). The most responsive pathways to the altitudinal change in the fruits were glycogen degradation at 600 m; superpathway of phospholipid biosynthesis at 800 m, and superpathway of heme biosynthesis from uroporphyrinogen at 900 m (Fig. 6A). In the leaves, the most responsive pathways were pyruvate fermentation, L−isoleucine biosynthesis, D−fructuronate degradation at 600 m; pyrimidine deoxyribonucleotides de novo biosynthesis, ectoine biosynthesis, sulfate reduction I (assimilatory), biotin biosynthesis at 800 m and octane oxidation, aminobutanoate degradation and enterobactin biosynthesis at 900 m (Fig. 6B).

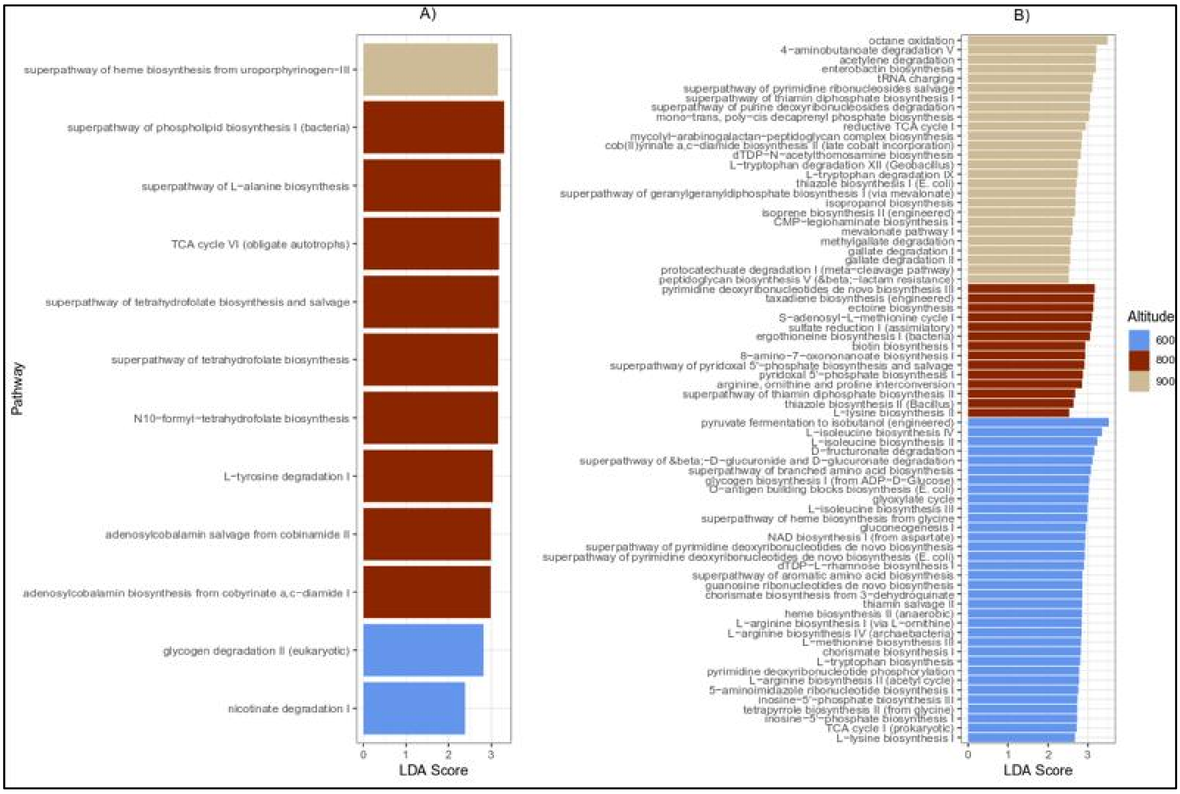


Figure 4.6 Differentially abundant metacyc pathways (LDA ≥ 2.0; FDR-adjusted $p \leq 0.05$) of bacterial communities associated with the fruits (A), and leaves (B) of *C. arabica* across altitudes (600 m, 800 m, and 900 m).

Fungal ecological guilds were also identified, with several guilds showing distinct responses to both the shade and altitudinal gradients. The ecological guilds included saprotrophs, plant pathogens, lichen parasites, and fungal parasites (Fig. 7). There were more undefined saprotrophs in the fruits at 0% canopy shade, and fungal parasites appeared to increase at 50% - 100% canopy shade. Altitudinally, undefined saproptrophs and fungal parasites increased in abundance at 900 m in the leaves and fruits, respectively.

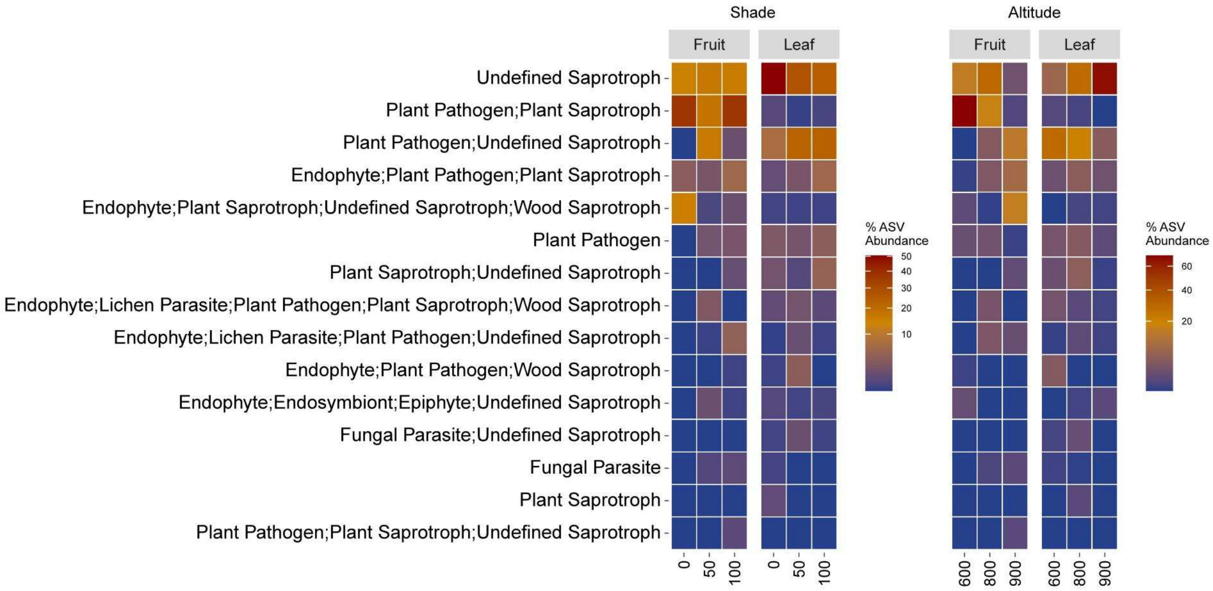


Figure 4.7 Ecological guilds of fungal communities associated with *C. arabica* in response to canopy shade and altitudinal differences.

4.4. Discusion

The study of endophytes associated with coffee plants has long been focus in coffee research, especially due to their role in major diseases such as coffee leaf rust (CLR) and coffee berry disease (CBD). Both are caused by pathogenic fungi, *Hemileia vastatrix* (CLR) and *Colletotrichum kahawe* (CBD), implying considerable losses in coffee productivity and beverage quality (McCook, 2006; 2019; Adugna, 2024). *Hemileia vastatrix* is widely distributed in all coffee producing countries, affecting commercial and wild Coffea spp., while the incidence of Colletotrichum kahawe is confined to Africa. Recently, studies on the coffee-associated microbiome, particularly those related to soil and fermentation processes, have gained prominence. The present study contributes to understanding the diversity, functions, and dynamics of coffee microbial endophytic communities across shade and elevation gradients.

While shade did not influence microbial diversity, altitude significantly shaped the composition of fungal and bacterial communities in fruits and leaves, accounting for 20% of the variation, despite some variability within the samples (Fig. 1 and 2). Analysis of the soil microbiome of the same *C. arabica* plants revealed that altitude is indeed the primary driver of microbial communities in the Gorongosa Mountain (Tapaça *et al*., 2024). Similar findings were reported by Siles *et al*., (2016) and Aqeel *et al*., (2024) for soils in alpine forests and the foothills of the Hymalayas, further supporting the significant influence of altitude on microbial diversity and community composition. One possible explanation for the greater influence of altitude over shade is that altitude affects a broader range of environmental factors that are crucial for microbial life. These include temperature, atmospheric pressure, and oxygen levels, all of which vary significantly with altitude and can exert strong selective pressures on microbial communities (Yang *et al*., 2022). In contrast, shade primarily influences light availability, which may have a more indirect effect on microbial communities, particularly those within endophytic environments (Kumar *et al*., 2017). These combined factors likely contribute to the greater impact of altitude compared to shade in determining the structure of microbial communities associated with coffee plants.

For both microbial groups, substantial taxonomic variation was observed between fruits and leaves, as well as across the shade and altitudinal gradients, with the differences being more pronounced for bacterial communities. Despite the differences, in each level of shade and altitude, the dominant bacterial taxa included: (i) plant growth-promoting bacteria, many of which produce growth hormones such as like auxins and gibberellins, solubilize phosphorus, fix nitrogen, and enhance nutrient uptake (Jia *et al*., 2022, Liu *et al*., 2022); (ii) biocontrol microbes that produce antibiotics, siderophores, or enzymes inhibiting plant pathogens (Raaijmakers *et al*., 2002; Zhu *et al*., 2023; Deb and Tatung 2024); (iii) bioremediators that degrade harmful organic compounds or pollutants, contributing to soil health and plant resilience (Antizar-Ladsilao 2010; Patil *et al*., 2024); and (iv) plant pathogens (Setubal, Moreira, and da Silva 2005). For example, the presence of *Kokasonia*, *Pseudomonas* or *Methylorubrum* in fruits and leaves might be associated to several beneficial functions. The presence of Kokasonia in plant roots or the rhizosphere has been associated to nitrogen fixation and phosphate solubilization, e.g. in peanuts (Narayanan *et al*., 2022) and rice (Ly *et al*., 2022). Many species of *Pseudomonas* are beneficial rhizobacteria that enhance plant growth by producing phytohormones (e.g., indole-3-acetic acid, IAA) and siderophores, and by inhibiting pathogens via antibiotic production (Nerek *et a*l., 2022). *Methylorubrum* can metabolize methane and other one-carbon compounds, playing an important ecological role in reducing greenhouse gases (Lindstrom 2006). They also promote plant growth by producing auxins and facilitating nitrogen fixation (Ivanova, Doronina, and Trotsenko *et al*., 2001). Pantoea exhibit diverse interactions with plants, ranging from beneficial effects, such as nitrogen fixation. phosphate solubilization, and biocontrol, to detrimental effects like causing diseases such as blights (Doni *et al*., 2021). Similarly, some species of *Stenotrophomonas* (fruits), *Burkholderia* (fruits), and *Curtobacterium* (fruits and leaves) may provide beneficial effects, e.g. nutrient assimilation, bioremediation and biological control of plant pathogens, while others may act as plant pathogens (Berg, Eberl and Hartmann 2005; Rosenblueth and Marínez-Romero 2006; Compant *et al*., 2008; Osdaghi *et al*., 2018; Horuz *et al*., 2024). Also present in fruits and leaves, Erwinia is primarily known as plant pathogens and may cause economically important diseases like fire blight (Vanneste 2000), soft rot (Toth *et al*., 2003), and bacterial wilt (Kearns and Holmes, 2015). The results illustrate the wide range of functional roles, including plant growth promotion, plant health, soil fertility and remediation, and in some cases, opportunistic pathogenesis. An interesting finding was the identification of 19 bacterial phylotypes in fruits from plants grown at 900 m (vs. two at 800 m, none at 600, and none in leaves) (Fig. 5A). These bacteria are associated with plant growth promotion and a variety of functions, ranging from bioremediation and biodegradation to nutrient cycling. The collective activity of such differential microbes may contribute to the special set of biochemical attributes identified in the corresponding green beans, like higher levels of trigonelline and monocaffeoylquinic, and low levels of feruloylquinic acids and dicaffeoylquinic acids (Cassamo *et al.,* 2022).

In this regard, PICRUSt2 analysis revealed 74 differentially abundant metacyc metabolic pathways along the altitudinal gradient (Fig. 6). At 600 m, pathways such as glycogen degradation and pyruvate fermentation were prevalent. These pathways are typically associated with energy production and resource storage, which might be essential at lower altitudes where temperature and humidity create environments that require energy-efficient microbial metabolism. Microbes may prioritize these pathways to manage fluctuations in resource availability (Pacciani-Mori *et al.*, 2020). At 800 m, the dominance of pathways such as phospholipid biosynthesis and pyrimidine deoxyribonucleotides de novo biosynthesis suggests an emphasis on cellular membrane construction and nucleic acid metabolism. The environmental conditions at this altitude, including moderate temperatures and oxygen levels, likely promote microbial growth and division, requiring efficient membrane and genetic material synthesis (Moon *et al.*, 2023). At 900 m, pathways like heme biosynthesis, octane oxidation, and enterobactin biosynthesis indicate adaptations to oxidative stress and nutrient limitation, as higher altitudes are typically characterized by lower oxygen levels and reduced nutrient availability. The heme biosynthesis pathway, for instance, is crucial for oxygen transport and utilization, making it vital in high-altitude environments where oxygen is limited (Swenson *et al.*, 2020).

In leaves, biotin biosynthesis and sulfate reduction at 800 m and aminobutanoate degradation at 900 m reflect metabolic adaptations to specific nutrient cycles and environmental stressors found at different altitudes. Enterobactin biosynthesis at 900 m further points to microbial strategies to scavenge iron, a nutrient that becomes less available at higher elevations (Li, Chen, and Bruner 2016). Thus, these altitudinal differences suggest that microbial communities adapt their metabolic functions to the specific challenges presented by their environment, ensuring survival and functional efficiency across different ecological niches.

A clear differential pattern in fungal communities is observed in fruits along the altitudinal gradient, with further variation across shade levels at 800 m and 900 m. *Eremothecium*, the most abundant genus at 600m (all shade levels) and 800 m (full shade), is a yeast well-known for fermentation capacity and for producing riboflavin (vitamin B2) (Zhang *et al.*, 2022). Some species are also known as plant pathogens, causing yest-spot diseases in crops such as soybean (Choi *et al.*, 2023) and azuka bean (Kimura *et al.*, 2009). *Penicillium*, found under full shade at 900 m, and to some extent at 600 m, is one of the most famous fungal genera for producing penicillin and some species have biocontrol potential due to their ability to inhibit plant pathogens (Toghueo *et al.*, 2020). Some species are also related to biofertilization

(Radhakrishnan *et al*., 2014) and biosynthesis of primary and defense metabolites (Radhakrishnan and Lee 2015), but also to food spoilage (Pitt and Hocking 2022) and fermentation processes (Wang *et al*., 2022; Ye *et al*., 2024). Another group particularly prominent at 900 m (full shade) was Meyerozyma associated to bioremediation (Amorim *et al*., 2018), biocontrol (Fu *et al*., 2022; Ayón-Macias *et al*., 2023) and fermentation processes (Kaur, Guleria, and Yadav 2023). The presence of *Fusarium* in fruits at 900 m should be seen with caution, as many species are known as major plant pathogens, causing diseases like fusarium wilt in a wide variety of crops (Stępień 2023). Many species produce harmful mycotoxins that contaminate agricultural products and may devastating effects on crop yield and food safety (Perincherry et al. 2019). However, it should be highlighted that the mycotoxins were not detected in green coffee beans (our lab, unpublished data). *Cladosporium* and *Colletotrichum* are other common plant pathogens found in fruits from plants grown under full shade at 900m and at moderate shade at 800 m. The first has been associated to plant growth and protection against biotic and abiotic stresses (Rǎut *et al.,* 2021), plant health (Islam 2022) and plant diseases (Temperini *et al*., 2018; Aydoğdu *et al*., 2023; Ragukula and Makandar 2024). The second is the causal agent of anthracnose in a variety of crops, including coffee, leading to severe crop losses, especially under humid conditions (Chen *et al*., 2003; Mouen *et al*., 2007). In leaves, the number of fungal taxa and the distribution along the shade and altitude gradients was more equilibrate, and as expected, the number of taxa related to fermentation was not as considerable as in fruits. While the fungal phylotypes of leaves and fruits (Fig. 5B and C) cover a wide range of roles, from plant pathogens and biocontrol agents to decomposers, producers of medically important compounds and extremophiles, saprotrophic ecological guilds seem to prevail. Saprotrophic fungi play a significant role in the decomposition of organic matter, and their presence on coffee beans can lead to both beneficial and detrimental effects. In fruits, they might help break down the residual organic matter around the coffee beans (like the cherry pulp) during natural drying processes, aiding the fermentation of the beans and potentially contributing to flavor development (de Melo Pereira *et al*., 2018). However, excessive growth on coffee beans during improper drying or storage, may lead to fruit spoilage, off-flavors, and ultimately health risks hazards due to mycotoxin contamination (Mislivec et al. 1973; Bucheli and Taniwaki 2002). In leaves, the presence of saprotrophic fungi generally plays a positive role by breaking down dead plant material, enhancing soil health, and supporting nutrient cycling (Figueiredo *et al*., 1966; Meija *et al*., 2008).

Taking together the bacterial and fungi microbiome, as well as the observations of Cassamo *et al*. (2022) and Tapaça *et al*. (2024) regarding grain quality and soil microbiome, respectively, the results suggest that the variation of microbial taxa along the altitudinal and shade gradients is essential for plant performance, including the physicochemical characteristics of the green beans.

4.5. Conclusion

The study highlights the complexity and importance of the interactions between the microbial communities associated with coffee fruits and leaves. Altitude emerged as the main factor influencing the composition of bacterial and fungal communities, while shade did not produce a significant impact on microbial diversity. The taxonomic variation observed along the altitudinal gradient points to the relevance of microbial diversity in plant performance and coffee bean quality. Beneficial endophytes were associated to several attributes related to plant growth and protection. The presence of pathogenic microorganisms underscores the need for careful crop management. The findings also highlight the importance of management strategies that preserve and utilize native microbial communities, enhancing not only productivity but also the resilience of coffee plantations against diseases and climate change.

4.6. References

Adugna, G. Coffee berry disease: A century-old anthracnose of green berries of Arabica coffee (*Coffea arabica* L.) in Africa. *J Plant Dis Prot*. 2024, 131, 315–328, https://doi.org/10.1007/s41348-023-00838-1.

Aryuman, P., Lertsiri, S., Visessanguan, W., *et al*. Glutaminase-producing *Meyerozyma* (*Pichia*) guilliermondii isolated from Thai soy sauce fermentation. *Int J Food Microbiol*. 2015, 192, 7-12, https://doi.org/10.1016/j.ijfoodmicro.2014.09.019.

Abarenkov, K., Nilsson, R.H., Larsson, K.H., *et al*. The unite database for molecular identification and taxonomic communication of fungi and other eukaryotes:

Sequences, taxa and classifications reconsidered. *Nucleic Acids Res*. 2024, 52, D791-D797, https://doi.org/10.1093/nar/gkad1039.

Afridi, M.S., Ali, S., Salam, A., *et al*. Plant microbiome engineering: Hopes or hypes. *Biol.* 2022*,* 11, 1782, https://doi.org/10.3390/biology11121782.

Amorim, S.S., Ruas, F.A.D., Barboza, N.R., *et al*. Manganese (Mn^{2+}) tolerance and biosorption by *Meyerozyma guilliermondii* and *Meyerozyma caribbica strains*. *J Environ Chem Eng*. 2018, 6, 4538-4545, https://doi.org/10.1016/j.jece.2018.06.061.

Antizar-Ladislao, B. Bioremediation: Working with bacteria. *Elem*. 2010, 6, 389-394, https://doi.org/10.2113/gselements.6.6.389.

Aqeel, M., Khalid, N., Noman, A., *et al*. Interplay between edaphic and climatic factors unravels plant and microbial diversity along an altitudinal gradient. *Env Res*. 2024, 242,117711, https://doi.org/10.1016/j.envres.2023.117711.

Asad, S., Priyashantha, A.K.H., Tibpromma, S., *et al*. Coffee-associated endophytes: Plant growth promotion and crop protection. *Biol*. 2023, 12, 911, https://doi.org/10.3390/biology12070911.

Aydoğdu, M., Kurbetli, İ., Sülü, G., *et al.* First record of *Cladosporium* species causing leaf spots on globe artichoke (*Cynara scolymus*). *Australas Plant Pathol*. 2023, 52, 529–537, https://doi.org/10.1007/s13313-023-00939-3.

Ayón-Macias, K.D., Ragazzo-Sánchez, J.A., Narváez-Zapata, J.A., *et al*. *Meyerozyma strains* as biocontrol agents against postharvest phytopathogens of jackfruit (*Artocarpus heterophyllus* Lam). *Arch Phytopath Plant Prot.* 2023, 56, 1180–1204, https://doi.org/10.1080/03235408.2023.2276098.

Berg, G., Grube, M., Schloter, M., *et al*. Unraveling the plant microbiome: Looking back and future perspectives. *Front Microbiol.* 2014, 5, 148, https://doi.org/10.3389/fmicb.2014.00148.

Berg, G., Eberl, L., Hartmann, A., *et al*. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environ Microbiol*. 2005, 7, 1673-85, https://doi.org/ 10.1111/j.1462-2920.2005.00891.x.

Bolyen, E., Rideout, J.R., Dillon, M., *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019, 37, 852–857, https://doi.org/10.1038/s41587-019-0209-9.

Bucheli, P., Taniwaki, M.H. Research on the origin, impact, and control of mycotoxins in coffee. *Food Addit Contam*. 2002, 19, 655-665.

Bullergahn, V.B., Menezes, K.M.S., Veloso, T.G.R., *et al*. Diversity of potential nitrogen-fixing bacteria from rhizosphere of the *Coffea arabica* L. and *Coffea canephora* L. 3 *Biotech*. 2024, 14, 27, https://doi.org/10.1007/s13205-023-03875-7.

Cabrera-Rodríguez, A., Trejo-Calzada, R., la Peña, C.G., *et al*. A metagenomic approach in the evaluation of the soil microbiome in coffee plantations under organic and conventional production in tropical agroecosystems. *Emir J Food Agric.* 2020, 32, 263-270, https://doi.org/10.9755/ejfa.2020.v32.i4.2092.

Caldwell, A.C., Silva, L.C.F., da Silva, C.C., *et al.* Prokaryotic diversity in the rhizosphere of organic, intensive, and transitional coffee farms in Brazil. *PLoS One*. 2015, 10, e0106355, https://doi.org/10.1371/journal.pone.0106355.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016, 13, 581–583, https://doi.org/10.1038/nmeth.3869.

Cassamo, C.T., Draper, D., Romeiras, M.M., *et al*. Impact of climate changes in the suitable areas for *Coffea arabica* L. production in Mozambique: Agroforestry as an alternative management system to strengthen crop sustainability. *Agric Ecosyst Environ*. 2023, 346, 108341, https://doi.org/10.1016/j.agee.2022.108341.

Cassamo, C.T., Mangueze, A.V.J., Leitão, A.E., *et al*. Shade and altitude implications on the physical and chemical attributes of green coffee beans from Gorongosa

Mountain, Mozambique. *Agronomy*. 2022, 12, 2540, https://doi.org/10.3390/agronomy12102540.

Castillo-González, H., Bloomberg, J., Alvarado-Picado, E., *et al*. Agricultural practices influence foliar endophytic communities in coffee plants of different varieties. *Agrosyst Geo Sci Environ.* 2024, 7, e20476, https://doi.org/10.1002/agg2.20476.

Chen, J., Shen, Y., Chen, C., *et al*. Inhibition of key citrus postharvest fungal strains by plant extracts *in vitro* and *in vivo*: A review. *Plants*. 2019, 8, 26, https://doi.org/10.3390/plants8020026.

Chen, Z., Liang, J., Rodrigues Jr., C., *et al*. *Colletotrichum gloeosporioides* can overgrow *Colletotrichum kahawae* on green coffee berries first inoculated with *C. kahawae*. *Biotech Lett*. 2005, 27, 679–682, https://doi.org/10.1007/s10529-005-4684-3.

Choi, O., Noh, H., Kang, Y., *et al*. First report of yeast spot caused by *Eremothecium coryli*on soybean in South Korea. *Can J Plant Pathol*. 2023, 46, 106–113, https://doi.org/10.1080/07060661.2023.2287164.

Compant, S., Nowak, J., Coenye, T., *et al*. Diversity and occurrence of *Burkholderia* spp. in the natural environment. FEMS *Microbiol Rev*. 2008, 32, 607-26, https://doi.org/ 10.1111/j.1574-6976.2008.00113.x.

Compant, S., Samad, A., Faist, H., *et al*. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *J Adv Res.* 2019, 19, 29–37, https://doi.org/10.1016/j.jare.2019.03.004.

Cristóbal-Martínez, A.L., de Jesús Yáñez-Morales, M., Solano-Vidal, R., *et al*. Diversity of *Colletotrichum* species in coffee (*Coffea arabica*) plantations in Mexico. *Eur J PlantPathol.* 2017*,* 147, 605–614, https://doi.org/10.1007/s10658-016-1029-0.

da Silva Vale, A., de Melo Pereira, G.V., de Carvalho Neto, D.P., *et al*. Effect of co-inoculation with *Pichia fermentans* and *Pediococcus acidilactici* on metabolite produced during fermentation and volatile composition of coffee beans. *Ferment*. 2019, *5*, 67, https://doi.org/10.3390/fermentation5030067.

da Silva Vale, A., de Melo Pereira, G.V., de Carvalho Neto, D.P., *et al*. Facility-specific ‘house’ microbiome ensures the maintenance of functional microbial communities into coffee beans fermentation: Implications for source tracking. *Environ Microbiol Rep*. 2021, 13, 470-481, https://doi.org/10.1111/1758-2229.12921.

Doni, F., Suhaimi, N.S.M., Irawan, B., *et al*. Associations of *Pantoea* with rice plants: As friends or foes*?* *Agric*. 2021, 11, 1278, https://doi.org/10.3390/agriculture11121278.

Dastogeer, K.M., Tumpa, F.H., Sultana, A., *et al.* Plant microbiome – an account of the factors that shape community composition and diversity. *Curr Plant Biol.* 2020, 23, 100161, https://doi.org/10.1016/j.cpb.2020.100161.

Davis, A.P., Gargiulo, R., Almeida, I.N.d.M., *et al*. Hot coffee: The identity, climate profiles, agronomy, and beverage characteristics of *Coffea racemosa* and *Coffea zanguebariae*. *Front Sustain Food Syst*. 2021, 5, 740137, https://doi.org/10.3389/fsufs.2021.740137.

Davis, A.P., Rakotonasolo, F. Six new species of coffee (*Coffea)* from northern Madagascar. *Kew Bull*. 2021, 76, 497–511, https://doi.org/10.1007/s12225-021-09952-5.

De Melo Pereira, G.V., Magalhães Júnior, A.I., de Abreu Ribeiro, F.A., *et al*. Fermentation process for coffee beans: Current status and future perspectives. *Bio Med Res Int*. 2018, 4839743.

De Melo Pereira, G.V.M., de Mello Sampaio, V., Wiele, N., *et al*. How yeast has transformed the coffee market by creating new flavors and aromas through modern post-harvest fermentation systems. *Trends Food Sci Technol*. 2024, 104641, https://doi.org/10.1016/j.tifs.2024.104641.

de Resende, M.L.V., Pozza, E.A., Reichel, T. *et al*. Strategies for coffee leaf rust management in organic crop systems. *Agronomy.* 2021*,* 11, 1865, https://doi.org/10.3390/agronomy11091865.

de Sousa, L.P., Guerreiro-Filho, O., Mondego, J.M.C., *et al*. The rhizosphere microbiomes of five species of coffee trees. *Microbiol Spectr*. 2022, 10, e00444-22, https://doi.org/10.1128/spectrum.00444-22.

Deb, C.R., Tatung, M. Siderophore producing bacteria as biocontrol agent against phytopathogens for a better environment: A review. *South Afr J Bot.* 2024, 165, 153-162, https://doi.org/10.1016/j.sajb.2023.12.031.

di Menna, M., Lauren, D., Hardacre, A., *et al*. Fusaria and *Fusarium* toxins in New Zealand maize plants. *Mycol pathol.* 1997, 139, 165–173, https://doi.org/10.1023/A:1006863908275.

Douglas, G.M., Maffei, V.J., Zaneveld, J.R., *et al*. PICRUSt2 for prediction of metagenome functions. *Nature Biotech*. 2020, 38, 685-688, https://doi.org./10.1038/s41587-020-0548-6.

Duong, B., Marraccini, P., Maeght, J.L., *et al.* Coffee microbiota and its potential use in sustainable crop management. A review. *Front Sustain Food Syst*. 2020, 4, 31, 607935, https://doi.org/10.3389/fsufs.2020.607935.

Dos Santos Gomes, W., Pereira, L.L., Rodrigues da Luz, J.M., *et al*. Exploring the microbiome of coffee plants: Implications for coffee quality and production. *Food Res Int*. 2024, 179, 113972, https://doi.org/10.1016/j.foodres.2024.113972.

FAO, ITPS, GSBI, SCBD, EC. State of knowledge of soil biodiversity - status, challenges and potentialities. 2020. *FAO*, Rome. 2020.

Fernández-González, A.J., Villadas, P.J., Gómez-Lama Cabanás, C., *et al*. Definition of the root endosphere and rhizosphere microbiomes from the world olive germplasm collection. *Sci Rep*. 2019, 9, 20423, https://doi.org/10.1038/s41598-019-56977-9.

Figueiredo, M.B., Hurst, P., Mendes, E.P., *et al*. The role of fungi in the degradation of coffee leaves. *Mycol*. 1966, 58, 569-576.

Frank, A.C, Saldierna Guzmán, J.P., Shay, J.E., *et al*. Transmission of bacterial endophytes. *Microorg*. 2017, 5, 70, https://doi.org/ 10.3390/microorganisms5040070.

Fu, Y., Ding, C., Fan, J., *et al*. Effects of three regeneration methods on the growth and bacterial community diversity of *Populus* × *euramericana*. *PLoS One*. 2022, 17, https://doi.org/10.1371/journal.pone.0273306.

Fulthorpe, R., Martin, A.R., Isaac, M.E., *et al*. Root endophytes of coffee (*Coffea arabica*): Variation across climatic gradients and relationships with functional traits. *Phytobiomes J*. 2020, 4, 27–39, https://doi.org/10.1094/PBIOMES-04-19-0021-R.

Ge, Y., Zhang, f., Xie, C., *et al*. Effects of different altitudes on *Coffea arabica* rhizospheric soil chemical properties and soil microbiota. *Agronomy*. 2023, 13, 471, https://doi.org/10.3390/agronomy13020471.

Grüter, R., Trachsel, T., Laube, P., *et al*. Expected global suitability of coffee, cashew and avocado due to climate change. *PLoS One,* 2022, 17, e0261976, https://doi.org/10.1371/journal.pone.0261976.

Horuz, S., Serin, M. Occurrence and pathogenicity of *Stenotrophomonas* spp. and *Paenibacillus* spp. on tomato plants in Turkey. *J Plant Pathol*. 2024, 106, 191–201, https://doi.org/10.1007/s42161-023-01540-9.

ICO – International Coffee Organization. Coffee market report. 2023. Available online: https://icocoffee.org/specialized-reports/ (accessed on June 2024).

Islam, M.T. Current status and future prospects of *Cladosporium* sp. a biocontrol agent for sustainable plant protection. *Biocontrol Sci*. 2022, 27, 185-191, https://doi.org/10.4265/bio.27.185.

Jia, H., Xi, Z., Ma, J., *et al*. Endophytic bacteria from the leaves of two types of albino tea plants, indicating the plant growth promoting properties. *Plant Growth Regul*. 2022, 96, 331–343, https://doi.org/10.1007/s10725-021-00779-5.

Kimura, S., Tokumaru, S., Kuge, K., *et al*. *Eremothecium coryli* and *E. ashbyi* cause yeast spot of azuki bean. *J Gen Plant Pathol*. 2009, 75, 322–324, https://doi.org/10.1007/s10327-009-0170-4.

Koutouleas, A., Sarzynski, T., Bertrand, B., *et al*. Shade effects on yield across different *Coffea arabica* cultivars – How much is too much? A meta-analysis. *Agron Sustain Dev*. 2022b, 42, 55, https://doi.org/10.1007/s13593-022-00788-2.

Koutouleas, A., Sarzynski, T., Bordeaux, M., *et al*. Shaded-coffee: A nature-based strategy for coffee production under climate change? A review. *Front Sustain Food Syst*. 2022a, 6, 877476, https://doi.org/10.3389/fsufs.2022.877476.

Lidstrom, M.E. Aerobic methylotrophic prokaryotes. The prokaryotes. *Springer*, New York, NY. 2006, 618–634, https://doi.org/10.1007/0-387-30742-7_20.

Liu, J.Q., Chen, S.M., Zhang, C.M., *et al*. Abundant and diverse endophytic bacteria associated with medicinal plant *Arctium lappa* L. and their potential for host plant growth promoting. *Antonie van Leeuwenhoek*. 2022, 115, 1405–1420, https://doi.org/10.1007/s10482-022-01785-x.

Lv, L., Luo, J., Ahmed, T., *et al*. Beneficial effect and potential risk of *Pantoea* on rice production. *Plants*. 2022, 11, 2608, https://doi.org/10.3390/plants11192608.

Magrach, A., Ghazoul, J. Climate and pest-driven geographic shifts in global coffee production: Implications for forest cover, biodiversity and carbon storage. *PLoS One*. 2015, 10, e0133071, https://doi.org/10.1371/journal.pone.0133071.

Manuel, L., Talhinhas, P., Várzea, V., *et al*. Characterization of *Colletotrichum kahawae* isolates causing coffee berry disease in Angola. *J Phytopathol.* 2010, 158, 310-313, https://doi.org/10.1111/j.1439-0434.2009.01613.x.

McCook, S. Global rust belt: *Hemileia vastatrix* and the ecological integration of world coffee production since 1850. *J Glob Hist*. 2006, 1, 177–195, https://doi:10.1017/S174002280600012X.

McCook, S. The devourer of dreams in: Coffee is not forever: A global history of the coffee leaf rust. 1st edition. *Ohio University Press*, Ohio, Athens. 2019, 1–16.

Mejia, L.C., Rojas, E.I., Maynard, Z., *et al*. Endophytic fungi as biocontrol agents of *Theobroma cacao* pathogens. *Biol Control.* 2008, 46, 4-14, https://doi.org/10.1016/j.biocontrol.2008.01.012.

Mislivec, P.B., Tuite, J., Krall, P.M., *et al*. Ochratoxin a production by *Aspergillus ochraceus* and *Penicillium viridicatum* on green coffee beans. *J Food Sci.* 1973, 38, 780-782.

Narayanan, M., Pugazhendhi, A., David, S., *et al*. Influence of *Kosakonia* sp. on the growth of *Arachis hypogaea* L. on arid soil. *Agronomy*. 2022, 12, 1801, https://doi.org/10.3390/agronomy12081801.

Negi, R., Sharma, B., Kumar, S. *et al*. Plant endophytes: unveiling hidden applications toward agro-environment sustainability. *Folia Microbiol*. 2024, 69, 181–206, https://doi.org/10.1007/s12223-023-01092-6.

Nerek, E., Sokołowska B. *Pseudomonas* spp. in biological plant protection and growth promotion. *AIMS Environ Sci*. 2022, 9, 493-504, https://doi.org/10.3934/environsci.2022029.

Ngure, G., Wasanabe, K. Coffee sustainability: leveraging on collaborative breeding for variety improvement. *Front Sustain Food Syst*. 2024, 8, 143849, https://doi.org/10.3389/fsufs.2024.1431849.

Nguyen, Nhu. H., Song, Z., Bates, S.T., *et al*. Fun guild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol*. 2016, 20, 241–248, https://doi.org/10.1016/j.funeco.2015.06.006.

Osdaghi, E., Taghavi, S.M., Hamzehzarghani, H., *et al*. Bacterial phytopathogens associated with dry bean seeds: Implications for seed health. *Plant Pathol*. 2018, 67, 8-24.

Patil, A., Chakraborty, S., Yadav, Y., *et al*. Bioremediation strategies and mechanisms of bacteria for resistance against heavy metals: A review. *Bioremediat J*. 2024, 1–33, https://doi.org/10.1080/10889868.2024.2375204.

Perincherry, L., Lalak-Kańczugowska, J., Stępień, Ł., *et al*. *Fusarium*-produced mycotoxins in plant-pathogen interactions. *Toxins*. 2019, 11, 664, https://doi.org/10.3390/toxins11110664.

Quast, C., Pruesse, E., Yilmaz, P., *et al*. The Silva ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013, 41, D590-596, https://doi.org/10.1093/nar/gks1219.

Raaijmakers, J.M., Vlami, M., de Souza, J.T., *et al*. Antibiotic production by bacterial biocontrol agents. *Antonie Van Leeuwenhoek*. 2002, 81, 537–547, https://doi.org/10.1023/A:1020501420831.

Radhakrishnan, R., Kang, S.M., Baek, I.Y., *et al*., Characterization of plant growth-promoting traits of *Penicillium* species against the effects of high soil salinity and root disease. *J Plant Interact*. 2014, 9, 754–762, https://doi.org/10.1080/17429145.2014.930524.

Radhakrishnan, R., Lee, I.J. *Penicillium*–sesame interactions: a remedy for mitigating high salinity stress effects on primary and defense metabolites in plants. *Environ Exp Bot.* 2015, 116, 47-60, https://doi.org/10.1016/j.envexpbot.2015.03.008.

Ramakrishna, W., Yadav, R., Li, K. *et al*. Plant growth promoting bacteria in agriculture: Two sides of a coin. *Appl Soil Ecol*. 2019, 138, 10-18, https://doi.org/10.1016/j.apsoil.2019.02.019.

Răut, I., Călin, M., Capră, L., *et al*. *Cladosporium* sp. isolate as fungal plant growth promoting agent. *Agronomy*. 2021, 11, 392, https://doi.org/10.3390/agronomy11020392.

Ragukula, K., Makandar, R. Isolation and identification of *Cladosporium tenuissimum* causing leaf blight on garden pea (*Pisum sativum* L.) in Telangana from southern India. 2024, 185, 106891, https://doi.org/10.1016/j.cropro.2024.106891.

Rojas, E.S., Gleason, M.L., Batzer, J.C., *et al*. Feasibility of delaying removal of row covers to suppress bacterial wilt of muskmelon (*Cucumis melo*). *Plant Dis*. 2011, 95, 729-734, https://doi.org/10.1094/PDIS-11-10-0788.

Rosenberg, E., Zilber-Rosenberg, I. Microbes drive evolution of animals and plants: The hologenome concept. *mBio*. 2016, 7, e01395, https://doi.org/10.1128/mBio.01395-15.

Rosenblueth, M., Martínez-Romero, E. Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact*. 2006, 19, 827-837, https://doi.org/10.1094/MPMI-19-0827.

Setubal, J.C., Moreira, L.M., da Silva, A.C.R., *et al*. Bacterial phytopathogens and genome science. *Curr Opin Microbiol*. 2005, 8, 595-600, https://doi.org/10.1016/j.mib.2005.08.015.

Siles, J.A., Margesin, R. Abundance and diversity of bacterial, archaeal, and fungal communities along an altitudinal gradient in alpine forest soils: What are the driving factors? *Microb Ecol*. 2016, 72, 207–220, https://doi.org/10.1007/s00248-016-0748-2.

Segata, N., Izard, J., Waldron, L., *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011, 12, R60, https://doi.org/10.1186/gb-2011-12-6-r60.

Sokol, N.W., Slessarev, E., Marschmann, G.L., *et al*. Life and death in the soil microbiome: How ecological processes influence biogeochemistry. *Nat Rev Microbiol*. 2022, 20, 415–430, https://doi.org/10.1038/s41579-022-00695-z.

Stępień, Ł. Plant-pathogenic *Fusarium* species. *J Fungi.* 2023, 9, 13, https://doi.org/10.3390/jof9010013.

Trotsenko, Y.A., Ivanova, E.G., Doronina, N.V., *et al*. Aerobic methylotroph bacteria as phytosymbionts. *Microbiol*. 2001, 70, 725-736, https://doi.org/10.1023/A:1013167612105.

Tapaça, I.d.P.E., Mavuque, L., Corti, R., *et al*. Genomic evaluation of *Coffea arabica* and its wild relative *Coffea racemosa* in Mozambique: Settling resilience keys for the coffee crop in the context of climate change. *Plants*. 2023, 12, 2044. https://doi.org/10.3390/plants12102044.

Tapaça, I.d.P.E., Obieze, C.C., Pereira, G.V.d.M., *et al*. Irradiance level and elevation shape the soil microbiome communities of *Coffea arabica* L. *Environ Microbiome*. 2024, 19, 75, https://doi.org/10.1186/s40793-024-00619-9.

Temperini, C.V., Pardo, A.G., Pose, G.N., *et al*. Diversity of airborne *Cladosporium* species isolated from agricultural environments of northern Argentinean Patagonia: Molecular characterization and plant pathogenicity. *Aerobiologia*. 2018, 34, 227–239, https://doi.org/10.1007/s10453-018-9509-7.

Toth, I.K., Bell, K.S., Holeva, M.C., *et al*. Soft rot *Erwiniae*: from genes to genomes. *Mol Plant Pathol*. 2003, 4, 17–30, https://doi.org/10.1046/j.1364-3703.2003.00149.x.

Tran, D.M. Rhizosphere microbiome dataset of robusta coffee (*Coffea canephora* L.) grown in the central highlands, Vietnam, based on 16S rRNA metagenomics analysis. *Data in Brief*. 2022, 42,108106, https://doi.org/10.1016/j.dib.2022.108106.

Trivedi, P., Batista, B.D., Bazany, K.E., *et al.* Plant–microbiome interactions under a changing world: Responses, consequences, and perspectives. *New Phytol.* 2022, 234, 1951–1959, https://doi.org/10.1111/nph.18016.

Vanneste, J. L. Fire blight: The disease and its causative agent, *Erwinia amylovora*. *HortScience.* 2000, 36, 1-6, https://www.cabidigitallibrary.org/doi/book/10.1079/9780851992945.0000.

Veloso, T.G.R., da Silva, M.d.C.S., Moreira, T.R., *et al.* Microbiomes associated with *Coffea arabica* and *Coffea canephora* in four different floristic domains of Brazil. *Sci Rep*. 2023, 13, 18477, https://doi.org/10.1038/s41598-023-45465-w.

Vogt, M. Variance in approach toward a ‘sustainable’ coffee industry in Costa Rica: perspectives from within; lessons and insights. *Ubiquity Press*. 2019, https://doi.org/10.5334/bce.

Voora, V., Bermúdez, S., Larrea, C., *et al.* Global market report: Coffee. *IISD*. 2019, 6, https://www.iisd.org/system/files/publications/ssi-global-market-report-coffee.pdf (accessed on 10 September 2024.

Wang, M., Zhou, W., Yang, Y., *et al*. *Penicillium* sp. Cis16 improves soluble dietary fiber content in citrus dregs fermentation. *Food Biotechnol.* 2022, 191–208, https://doi.org/10.1080/08905436.2022.2091592.

Xu, L., Pierroz, G., Wipf, H. M.L., *et al*. Holo-omics for deciphering plant-microbiome interactions. *Microbiome.* 2021*,* 9, 69, https://doi.org/10.1186/s40168-021-01014-z.

Ye, L., Hu, H., Wang, Y., *et al*. *In vitro* digestion and colonic fermentation characteristics of media-milled purple sweet potato particle-stabilized pickering emulsions. *J Sci Food Agric*. 2024, 104, 5064-5076, https://doi.org/ 10.1002/jsfa.13340.

Zhang, J.R., Ge, Y.Y., Liu, P.H., *et al*. Biotechnological strategies of riboflavin biosynthesis in microbes. *Engineering*. 2022, 12, 115-127, https://doi.org/10.1016/j.eng.2021.03.018.

Zhu, X., Chen, W.J., Bhatt, K., *et al*. Innovative microbial disease biocontrol strategies mediated by quorum quenching and their multifaceted applications: A review. *Front Plant Sci*. 2023, 13, 1063393, https://doi.org/ 10.3389/fpls.2022.1063393.

Chapter 5

General Conclusion

This study emphasizes the critical role of genetic diversity in ensuring sustainable agricultural practices. It demonstrates, for the first time, that while Agroforestry Systems (AFS) provide significant benefits, such as mitigating deforestation in tropical environments, increasing biodiversity, improving productivity, enhancing social value, and promoting environmental management, maintaining genetic diversity within species and cultivars is crucial for the long-term sustainability of these systems. The research also highlights that the incorporation of wild relatives in coffee AFS offers a valuable opportunity to increase productivity and resilience. Wild relatives of crops often possess genetic diversity that is absent in conventional cultivars, such as Arabica coffee. This diversity can be harnessed to improve agricultural systems, enhancing their adaptability and performance.

The evaluation of the occurrence of *Coffea* species in Mozambique, compared with bibliographic and herbarium data, and the molecular taxonomy analysis based on chloroplast markers, identified only one wild species in southern and central Mozambique, *C. racemosa* (Chapter II). The study supports the hypothesis of genetic erosion caused by significant anthropogenic and environmental pressure in the region. Despite this pressure, intra- and inter-population genetic diversity in *C. racemosa* is considerable, with well-defined geographic structure. On the other hand, the genetic pool of *C. arabica* in Gorongosa plantations proved to be reduced, highlighting the urgent need for cultivar diversification in coffee AFS. Additionally, functional pathway analysis based on SNPs suggests that stress signaling pathways are more robust in *C. racemosa*. Considering the global context, the implementation of conservation and management programs for native *Coffea* species, as well as their inclusion in the coffee value chain (for genetic improvement or utilization of the beans for new blends), is essential.

Furthermore, the microbial diversity of the soil (Chapter III) and the endosphere (Chapter IV) of *C. arabica* revealed itself to be quite rich and diverse, within the concept of terroir, which includes altitude, soil, climate, and agricultural practices. Altitude emerged as the main determinant in the composition of bacterial and fungal communities in fruits and leaves, while shade had no significant impact on microbial diversity. Together, the microbiome analysis suggests that the variation of microbial taxa along the altitudinal gradient is essential for plant performance, including the physicochemical characteristics of coffee beans. Each of these microbial communities holds potential for the development of biological-based solutions, such as bioremediation to improve soil quality, biofertilization to enhance plant nutrition, and fermentation, which can be explored to create unique flavor profiles in coffee. The presence of some pathogenic species, without apparent symptoms in the plants, could be due to the cultivar's tolerance to biotic factors (vertical or horizontal resistance). This tolerance suggests potential for the development of more resilient cultivars that can withstand pathogenic pressures, thus ensuring plantation health and continued production.

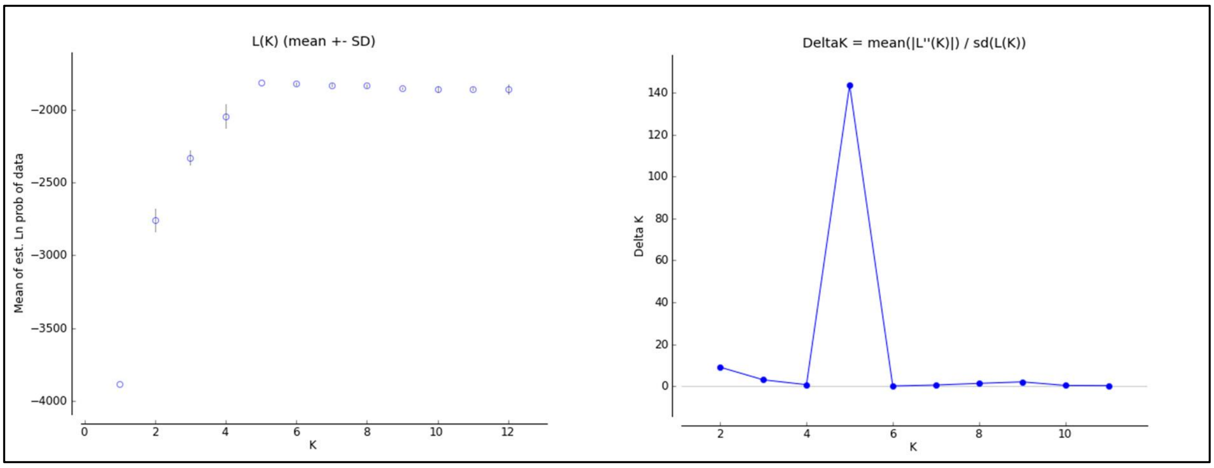
These results emphasize the importance of studying and preserving the interaction between plants and microorganisms, both for agricultural sustainability and for harnessing biological resources in coffee production. The integration of microbiome knowledge into coffee management could potentially transform agricultural practices, fostering more sustainable and higher-quality production. Therefore, further research into coffee-associated endophytes could open new avenues for the development of innovative agricultural techniques that value biodiversity and support sustainability.

6. Appendices

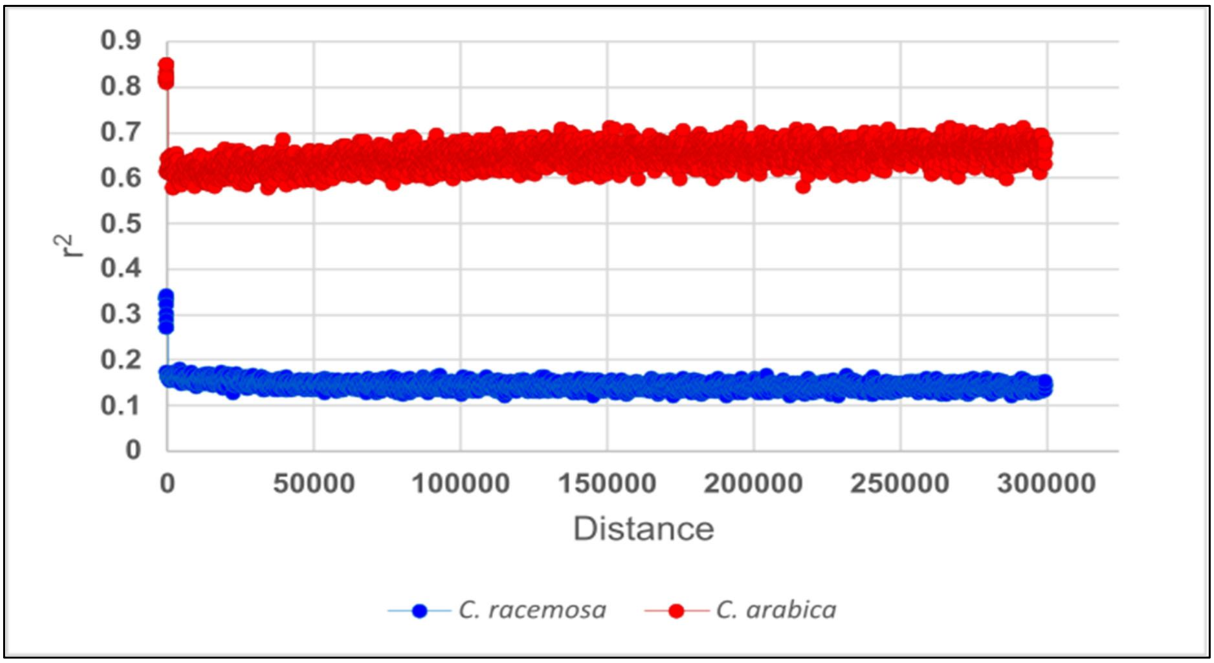
6.1. Supporting Information Files

Chapter II

Supplementary Information 1. **Figure S2.1**. Mean Ln probability of data and Delta K based on Evanno's ad hoc statistic obtained by Structure Harvester.



Supplementary Information 2. **Figure S2.2**. Scatter plot of Linkage disequilibrium decay (r^2) against the genetic distance for pairs of linked SNPs considering *C. racemosa* and *C. arabica*.



Supplementary Information 3. **Table S2.1.** Sampling information of *C. arabica* and *C. racemosa* sorted by geographical area including the individuals used for cpDNA (rbcLand matk), SSRs and SNPs analyses. Genbank numbers indicate the new sequences obtained in this study

ID	Species	Province	Population	Designation	Latitude	Longitude	RbcL	matK	SSRs	SNPs
IR01	*C. racemosa*	Inhambane	Inharrime	Wild	-24.30259	34.90368	OP207780	OP320952	x	X
IR02	*C. racemosa*	Inhambane	Inharrime	Wild	-24.47284	35.03484	OP207781	OP320953	x	X
IR03	*C. racemosa*	Inhambane	Inharrime	Wild	-24.47282	35.03481	OP207782	OP320954	x	
IR04	*C. racemosa*	Inhambane	Inharrime	Wild	-24.47329	35.02921	OP207783	OP320955	x	
IR05	*C. racemosa*	Inhambane	Inharrime	Wild	-24.47431	35.02438	OP207784	OP320956	x	
IR06	*C. racemosa*	Inhambane	Inharrime	Wild	-24.45185	35.01889	OP207785	OP320957	x	
ZV01	*C. racemosa*	Inhambane	Zavala	Wild	-24.50669	34.99925	OP207786	OP320958	x	X
ZV02	*C. racemosa*	Inhambane	Zavala	Wild	-24.50669	34.99922	OP207787	OP320959	x	
ZV03	*C. racemosa*	Inhambane	Zavala	Wild	-24.54388	34.88284	OP207788	OP320960	x	
ZV05	*C. racemosa*	Inhambane	Zavala	Wild	-24.68345	34.54291	OP207789	OP320961	x	
ZV06	*C. racemosa*	Inhambane	Zavala	Wild	-24.68415	34.54329	OP207790	OP320962	x	
ZV07	*C. racemosa*	Inhambane	Zavala	Wild	-24.68736	34.54589	OP207791	OP320963	x	
Ho01	*C. racemosa*	Inhambane	Homoine	Wild	-23.87354	35.24534	OP207792	OP320964	x	
HO03	*C. racemosa*	Inhambane	Homoine	Wild	-23.87354	35.24826	OP207793	OP320965	x	x
HO04	*C. racemosa*	Inhambane	Homoine	Wild	-23.89639	35.15764	OP207794	OP320966	x	x
HO05	*C. racemosa*	Inhambane	Homoine	Wild	-23.89639	35.15766	OP207795	OP320967	x	
HO06	*C. racemosa*	Inhambane	Homoine	Wild	-23.88346	35.15156	OP207796	OP320968	x	

HO07	*C. racemosa*	Inhambane	Homoine	Wild	-23.88332	35.15153	OP207797	OP320969	x	
MX01	*C. racemosa*	Inhambane	Maxixe	Wild	-23.86021	35.33956	OP207798	OP320970	x	
MX02	*C. racemosa*	Inhambane	Maxixe	Wild	-23.83998	35.34383	OP207799	OP320971	x	
MX03	*C. racemosa*	Inhambane	Maxixe	Wild	-23.82475	35.34215	OP207800	OP320972	x	x
MX04	*C. racemosa*	Inhambane	Maxixe	Wild	-23.82475	35.34218	OP207801	OP320973	x	
MX05	*C. racemosa*	Inhambane	Maxixe	Wild	-23.82457	35.34183	OP207802	OP320974	x	x
MX08	*C. racemosa*	Inhambane	Maxixe	Wild	-23.82476	35.34279	OP207803	OP320975	x	
MR01	*C. racemosa*	Inhambane	Morrumbene	Wild	-23.66451	35.34482	OP207804	OP320976	x	x
MR02	*C. racemosa*	Inhambane	Morrumbene	Wild	-23.61939	35.24733	OP207805	OP320977	x	x
MR03	*C. racemosa*	Inhambane	Morrumbene	Wild	-23.61931	35.24752	OP207806	OP320978	x	x
MR04	*C. racemosa*	Inhambane	Morrumbene	Wild	-23.61926	35.24751	OP207807	OP320979	x	
MR05	*C. racemosa*	Inhambane	Morrumbene	Wild	-23.61922	35.24752	OP207808	OP320980	x	
MR06	*C. racemosa*	Inhambane	Morrumbene	Wild	-23.69177	35.31211	OP207809	OP320981	x	
MP02	*C. racemosa*	Maputo	Matola	Wild	-25.92764	32.54849	OP207810	OP320982	x	
MP03	*C. racemosa*	Maputo	Matola	Wild	-25.96622	32.46106	OP207811	OP320983	x	
MP04	*C. racemosa*	Maputo	Matola	Wild	-25.95636	32.45265	OP207812	OP320984	x	x
MP05	*C. racemosa*	Maputo	Matola	Wild	-25.95733	32.45744	OP207813	OP320985	x	x
MP06	*C. racemosa*	Maputo	Matola	Wild	-25.95636	32.45255	OP207814	OP320986	x	
MP07	*C. racemosa*	Maputo	Matola	Wild	-25.95623	32.45265	OP207815	OP320987	x	
MP08	*C. racemosa*	Maputo	Matola	Wild	-25.95636	32.45232	OP207816	OP320988	x	x
MP09	*C. racemosa*	Maputo	Matola	Wild	-25.95625	32.45245	OP207817	OP320989	x	
SF02	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56341	34.08541	OP207818	OP320990	x	
SF05	*C. racemosa*	Sofala	Gorongosa	Wild	-19.06242	34.25271	OP207819	OP320991	x	X

SF06	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56271	39.08541	OP207820	OP320992	x	
SF07	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56341	34.09212	OP207821	OP320993	x	
SF08	*C. racemosa*	Sofala	Gorongosa	Wild	-19.00011	34.12044	OP207822	OP320994	x	
SF17	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56234	34.09189	OP207823	OP320995	x	
SF21	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56224	34.09196	OP207824	OP320996	x	
SF25	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56235	34.08568	OP207825	OP320997	x	
SF29	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56199	34.08572	OP207826	OP320998	x	
SF31	*C. racemosa*	Sofala	Gorongosa	Wild	-18.56167	34.08555	OP207827	OP320999	x	
NIA01	*C. arabica*	Niassa	Maua	Cultivated	-13.44471	36.59212	OP207828	OP321000	x	
NIA02	*C. arabica*	Niassa	Maua	Cultivated	-13.44471	36.59229	OP207829	OP321001	x	
NIA03	*C. arabica*	Niassa	Maua	Cultivated	-13.19508	35.14472	OP207830	OP321002	x	
NIA04	*C. arabica*	Niassa	Maua	Cultivated	-13.19508	35.14485	OP207831	OP321003	x	
NIA05	*C. arabica*	Niassa	Maua	Cultivated	-12.52288	35.43506	OP207832	OP321004	x	
NIA06	*C. arabica*	Niassa	Maua	Cultivated	-12.52288	35.43506	OP207833	OP321005	x	
NIA07	*C. arabica*	Niassa	Maua	Cultivated	-13.20231	35.16513	OP207834	OP321006	x	
NIA08	*C. arabica*	Niassa	Maua	Cultivated	-13.20232	35.16511	OP207835	OP321007	x	
Ca-A650-P1	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51435	34.05135	OP207836	OP321008	x	
Ca-A650-P2	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51431	34.05133	OP207837	OP321009	x	
Ca-A650-P3	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51439	34.05121	P207838	OP321010	x	X
Ca-A650-P6	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51428	34.05132	OP207839	OP321011	x	X
Ca-A650-P7	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51427	34.05128	OP207840	OP321012	x	X
Ca-A650-P8	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51436	34.05124	OP207841	OP321013	x	X
Ca-A650-P11	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51448	34.05127	OP207842	OP321014	x	

Ca-A650-P12	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51466	34.05122	OP207843	OP321015	x	
Ca-A650-P13	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.51469	34.05125	OP207844	OP321016	x	
Ca-A825-P16	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50061	34.04907	OP207845	OP321017	x	X
Ca-A825-P17	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50062	34.04911	OP207846	OP321018	x	
Ca-A825-P18	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50073	34.04921	OP207847	OP321019	x	X
Ca-A825-P21	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50059	34.04905	OP207848	OP321020	x	
Ca-A825-P22	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50062	34.04904	OP207849	OP321021	x	
Ca-A825-P23	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50062	34.04901	OP207850	OP321022	x	
Ca-A825-P26	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50043	34.04907	OP207851	OP321023	x	X
Ca-A825-P27	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50043	34.04906	OP207852	OP321024	x	
Ca-A825-P28	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.50081	34.04949	OP207853	OP321025	x	
Ca-A935-P31	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48163	34.04477	OP207854	OP321026	x	
Ca-A935-P32	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48162	34.04478	OP207855	OP321027	x	
Ca-A935-P33	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48163	34.04475	OP207856	OP321028	x	X
Ca-A935-P37	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48158	34.04469	OP207857	OP321029	x	X
Ca-A935-P38	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48158	34.04469	OP207858	OP321030	x	X
Ca-A935-P41	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48151	34.04467	OP207859	OP321031	x	
Ca-A935-P42	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48151	34.04466	OP207860	OP321032	x	
Ca-A935-P43	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48187	34.04402	OP207861	OP321033	x	X
Ca-A935-P44	*C. arabica* cv. Costa Rica	Sofala	PNG	Cultivated	-18.48187	34.044	OP207862	OP321034	x	
Cv7960	*C. arabica* (CIFC)		Cultivar	Cultivated			OP207863	OP321035	x	
Cv7963	*C. arabica* (CIFC)		Cultivar	Cultivated			OP207864	OP321036	x	
Cv7958	*C. arabica* (CIFC)		Cultivar	Cultivated			OP207865	OP321037	x	

Supplementary Information 4. **Table S2.2.** KEGG pathways found in *C. racemosa* and *C. arabica*.

Pathways	*C. racemosa*	*C. arabica*
Plant-pathogen interaction	141	95
Protein processing in endoplasmic reticulum	93	94
Phenylpropanoid biosynthesis	75	75
Starch and sucrose metabolism	71	74
NF-kappa B signaling pathway	72	72
Spliceosome	68	68
RNA transport	66	66
Plant hormone signal transduction	77	66
MAPK signaling pathway - plant	66	63
Cell cycle	61	61
Purine metabolism	55	54
Endocytosis	55	55
mRNA surveillance pathway	54	52
Ubiquitin mediated proteolysis	53	52
RNA degradation	45	45
Ribosome	44	43
Amino sugar and nucleotide sugar metabolism	41	42
Oocyte meiosis	42	41
Pyrimidine metabolism	41	41
Phosphatidylinositol signaling system	41	41
Autophagy - animal	40	40
Inositol phosphate metabolism	39	39
Glycolysis / Gluconeogenesis	37	37
Ribosome biogenesis in eukaryotes	36	36
Glycine, serine and threonine metabolism	35	35
ABC transporters	35	35
Pentose and glucuronate interconversions	35	34
Tyrosine metabolism	32	32
Aminoacyl-tRNA biosynthesis	32	32

Galactose metabolism	30	30
Glycerophospholipid metabolism	30	30
Carbon fixation in photosynthetic organisms	30	30
Phenylalanine metabolism	29	29
Cyanoamino acid metabolism	29	29
Regulation of actin cytoskeleton	29	29
Alanine, aspartate and glutamate metabolism	28	28
cAMP signaling pathway	28	28
Phospholipase D signaling pathway	28	28
Glutathione metabolism	27	27
Isoquinoline alkaloid biosynthesis	27	27
Nucleotide excision repair	27	27
Fanconi anemia pathway	27	27
Sphingolipid signaling pathway	27	27
Carotenoid biosynthesis	26	26
Quorum sensing	26	26
Cellular senescence	26	26
Cysteine and methionine metabolism	25	25
Oxidative phosphorylation	24	24
Glyoxylate and dicarboxylate metabolism	24	24
One carbon pool by folate	24	24
Stilbenoid, diarylheptanoid and gingerol biosynthesis	24	24
Ras signaling pathway	24	24
AMPK signaling pathway	24	24
Glycerolipid metabolism	23	23
Ascorbate and aldarate metabolism	22	22
Arginine biosynthesis	22	22
Pyruvate metabolism	22	22
Sesquiterpenoid and triterpenoid biosynthesis	22	22
Phagosome	22	22
Tight junction	22	22
Fructose and mannose metabolism	21	21
Phenylalanine, tyrosine and tryptophan biosynthesis	21	21
Nitrogen metabolism	21	21

DNA replication	21	21
MAPK signaling pathway	21	21
mTOR signaling pathway	21	21
PI3K-Akt signaling pathway	21	21
Necroptosis	21	21
TGF-beta signaling pathway	21	21
Ubiquinone and other terpenoid-quinone biosynthesis	20	20
Methane metabolism	20	20
Wntsignaling pathway	20	20
Circadian rhythm - plant	20	20
Pentose phosphate pathway	19	19
Thiamine metabolism	19	19
Proteasome	19	19
Mismatch repair	19	19
Autophagy - other	19	19
Peroxisome	19	19
Ether lipid metabolism	18	18
Flavonoid biosynthesis	18	18
Lysosome	18	18
Apoptosis	18	18
beta-Alanine metabolism	17	17
RNA polymerase	17	17
Homologous recombination	17	17
FoxOsignaling pathway	17	17
Fatty acid degradation	16	16
Terpenoid backbone biosynthesis	16	16
cGMP-PKG signaling pathway	16	16
Lysine degradation	15	15
Diterpenoid biosynthesis	15	15
Tropane, piperidine and pyridine alkaloid biosynthesis	15	15
Arginine and proline metabolism	14	14
Tryptophan metabolism	14	14
Folate biosynthesis	14	14
Hedgehog signaling pathway	14	14

Linoleic acid metabolism	13	13
alpha-Linolenic acid metabolism	13	13
Protein export	13	13
Citrate cycle (TCA cycle)	12	12
Porphyrin and chlorophyll metabolism	12	12
HIF-1 signaling pathway	12	12
Apelin signaling pathway	12	12
Focal adhesion	12	12
Carbon fixation pathways in prokaryotes	11	11
Sulfur metabolism	11	11
Metabolism of xenobiotics by cytochrome P450	11	11
Mitophagy - animal	11	11
Ferroptosis	11	11
Brassinosteroid biosynthesis	10	10
Calcium signaling pathway	10	10
Gap junction	10	10
Fatty acid biosynthesis	9	9
Steroid biosynthesis	9	9
Valine, leucine and isoleucine degradation	9	9
Hippo signaling pathway	9	9
Other glycan degradation	8	8
Propanoate metabolism	8	8
Zeatin biosynthesis	8	8
Basal transcription factors	8	8
Base excision repair	8	8
Aminobenzoate degradation	7	7
Styrene degradation	7	7
Pantothenate and CoA biosynthesis	7	7
Rap1 signaling pathway	7	7
Signaling pathways regulating pluripotency of stem cells	7	7
Cutin, suberine and wax biosynthesis	6	6
Other types of O-glycan biosynthesis	6	6
Biofilm formation - Escherichia coli	6	6
ErbBsignaling pathway	6	6

VEGF signaling pathway	6	6
Adherens junction	6	6
TNF signaling pathway	6	6
Photosynthesis	5	5
Caffeine metabolism	5	5
Taurine and hypotaurine metabolism	5	5
Lipopolysaccharide biosynthesis	5	5
Sphingolipid metabolism	5	5
Nicotinate and nicotinamide metabolism	5	5
Retinol metabolism	5	5
Glucosinolate biosynthesis	5	5
Valine, leucine and isoleucine biosynthesis	4	4
Histidine metabolism	4	4
N-Glycan biosynthesis	4	4
Various types of N-glycan biosynthesis	4	4
Glycosaminoglycan degradation	4	4
Chloroalkane and chloroalkene degradation	4	4
Naphthalene degradation	4	4
Butanoate metabolism	4	4
Vitamin B6 metabolism	4	4
Lipoic acid metabolism	4	4
Monoterpenoid biosynthesis	4	4
Notch signaling pathway	4	4
Biofilm formation - Vibrio cholerae	4	4
Steroid hormone biosynthesis	3	3
Aflatoxin biosynthesis	3	3
Mannose type O-glycan biosynthesis	3	3
Streptomycin biosynthesis	3	3
Neomycin, kanamycin and gentamicinbiosynthesis	3	3
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	3	3
Arachidonic acid metabolism	3	3
Biosynthesis of unsaturated fatty acids	3	3
Bacterial secretion system	3	3
Non-homologous end-joining	3	3

Sulfur relay system	3	3
Fatty acid elongation	2	2
Synthesis and degradation of ketone bodies	2	2
Geraniol degradation	2	2
Chlorocyclohexane and chlorobenzene degradation	2	2
Fluorobenzoate degradation	2	2
Phenazine biosynthesis	2	2
Selenocompound metabolism	2	2
D-Glutamine and D-glutamate metabolism	2	2
Polyketide sugar unit biosynthesis	2	2
Toluene degradation	2	2
Riboflavin metabolism	2	2
Biosynthesis of ansamycins	2	2
Biofilm formation - Pseudomonas aeruginosa	2	2
Jak-STAT signaling pathway	2	2
Photosynthesis - antenna proteins	1	1
Monobactam biosynthesis	1	1
Lysine biosynthesis	1	1
C5-Branched dibasic acid metabolism	1	1
Biotin metabolism	1	1
Atrazine degradation	1	1
Indole alkaloid biosynthesis	1	1
Two-component system	1	1
Neuroactive ligand-receptor interaction	1	1
SNARE interactions in vesicular transport	1	1

Supplementary Information 5. **Table S2.3**. Plastid primers used to amplify the *Coffea* samples.

Gene Primer sequences	
MatK	F: CGTACAGTACTTTTGTGTTTACGAG R: ACCCAGTCCATCTGGAAATCTTGGTTC
RbcL	F: ATGTCACCACAAACAGAGACTAAAGC R: GAAACGGTCTCTCCAACGCAT

Supplementary Information 6. **Table S2.4**. Taxa retrieved from NCBI and used in the phylogenetic analyses

Genbank ID	Species
MK577911	*Coffea racemosa*
JX572420	*Coffea racemosa*
MK577910	*Coffea pseudozanguebariae*
MK577912	*Coffea stenophylla*
MK577913	*Coffea tetragona*
MK577907	*Coffea humblotiana*
KC758284	*Coffea arabica*
MN894552	*Coffea arabica*
AB973183	*Coffea arabica*
AB973185	*Coffea arabica*
AB973184	*Coffea arabica*
JN114797	*Coffea arabica*
MK875244	*Coffea arabica*
MK862266	*Coffea arabica*
NC008535	*Coffea arabica*
AB973187	*Coffea arabica*
AB973188	*Coffea arabica*
AB973186	*Coffea arabica*
MK577909	*Coffea macrocarpa*
GQ248573	*Coffea myrtifolia*
KC758293	*Coffea liberica*
KC758290	*Coffea liberica v. dewevrei*
KC758294	*Coffea canephora*
AB973189	*Coffea canephora*
KC758295	*Coffea canephora*
KC628267	*Psilanthus mannii*

Supplementary Information 7. **Table S2.5**. SSRs primers used to amplify the *Coffea* samples.

ID Primer	Sequences
CCRM02	F: AATGGTGGCAGTCCTGAAAGATC
	R: AACATCAACTTTCCTGGTCTTC
CCRM07	F: TAAAGGATGGTATATGTGGCTGGAGTA
	R: CCACAGCCTCGGCATTTACTATATAT
CCRM16	F: TCCTATAGCAGAAACACAAAATGACACAG
	R: GGTTTTTGGGTTCTTTTTAGCATATACA
CCRM17	F: TAAGCGTTGGAATTCCTCACTCTATCT
	R: ACAGCTAAAGAAACAATGAACCAGT
CC2P4	F: TAGCAGCAAACACTCTTCGC
	R: GCTTCTGACGGACTTGAGGA
CC5P3	F: CTGTTACAGCCTCGTCCACT
	R: TCCCGCTACTTTCATGGGAT
CC4P5	F: CTAGTCTTGGCATGTTGGGG
	R: CTTTGCTGTGGAGGTAAGGC
CC1P7	F: GGGGCAAAACAAGAACCACT
	R: ACCTTATCCAAAACCCATGTGC
CC3P6	F: CTTGGGATTGCCTAGCCCTA
	R: TCTTGCCGTTTTAGCCGATT
CC4P8	F: TGAGAAGGGACAAAGAAAGAGG
	R: ACAGTACAACATATGAGGCCAC
AJ308790	F: TTTTCTGGGTTTTCTGTGTTCTC
	R: TAACTCTCCATTCCCGCATT
AJ308779	F: TCCCCCATCTTTTTCTTTCC
	R: GGGAGTGTTTTTGTGTTGCTT
AJ308809	F: AGCAAGTGGAGCAGAAGAAG
	R: CGGTGAATAAGTCGCAGTC
SSR03	F: GGACAAAACACCGCCCAAAATA
	R: AGCGAGACAGAGGAAGGGAATATT

Supplementary Information 8. **Table S2.6**. Summary of sequencing and mapping of reads from *C. arabica* and *C. racemosa* samples

Sample	*Species*	Raw Reads	Clean Reads	GC (%)	Q 20	% mapping
IR02	*C. racemosa*	6395090	1753058	37.58	97.09	77.67%
ZV01	*C. racemosa*	8312096	8312096	37.11	96.15	75.59%
HO04	*C. racemosa*	7140830	7140830	38.61	96.42	83.51%
MR01	*C. racemosa*	5410602	5410602	36.47	96.11	78.62%
MR02	*C. racemosa*	5865342	5865342	37.78	96.05	82.58%
MX05	*C. racemosa*	7393166	7393166	36.32	95.31	84.04%
MX03	*C. racemosa*	4802168	4802168	37.23	95.13	84.82%
SF05	*C. racemosa*	8806602	8806602	37.10	96.91	76.07%
MP05	*C. racemosa*	5111332	5111332	36.71	96.36	83.34%
MP04	*C. racemosa*	8802762	8802762	37.67	95.80	82.94%
MR03	*C. racemosa*	4277686	4277686	37.81	95.16	79.69%
MP08	*C. racemosa*	4006454	4006454	37.53	96.09	81.00%
HO03	*C. racemosa*	4604378	4604378	37.61	96.71	77.52%
IR01	*C. racemosa*	5766290	5766290	37.05	96.49	74.60%
Ca-A650-P3	*C. racemosa*	5843528	5843528	38.28	94.54	80.53%
Ca-A650-P6	*C. racemosa*	6184428	6184428	38.28	96.02	81.05%
Ca-A650-P7	*C. racemosa*	7465534	7465534	38.41	96.33	86.94%
Ca-A650-P8	*C. racemosa*	6732436	6732436	38.27	96.93	86.26%
Ca-A650-P11	*C. racemosa*	4990690	4990690	38.15	94.91	77.48%
Ca-A825-P16	*C. racemosa*	4524368	4524368	38.06	95.94	83.28%
Ca-A825-P18	*C. racemosa*	5713418	5713418	45.73	96.14	28.89%
Ca-A825-P22	*C. racemosa*	7341088	7341088	38.59	97.25	82.10%
Ca-A825-P23	*C. racemosa*	5385822	5385822	38.72	94.75	78.26%
Ca-A825-P16	*C. racemosa*	4524368	4524368	38.06	95.94	83.28%
Ca-P935-P33	*C. racemosa*	4686380	4686380	38.34	94.98	82.93%
Ca-P935-P37	*C. racemosa*	6320542	6320542	38.74	96.59	76.10%
Ca-P935-P38	*C. racemosa*	8754090	8754090	37.79	97.15	85.33%
Ca-P935-P41	*C. racemosa*	5914830	5914830	38.11	96.31	86.50%
Ca-P935-P43	*C. racemosa*	4168100	4168100	37.37	96.15	85.65%

Chapter III

Supplementary Information 9. **Table S3.1.** Results of general lineal models testing the effects of altitude, shade, level, and their interactions on properties and enzymatic activities of soil samples of *C. arabica*. P-Olsen: available phosporous; SOC: soil organic carbon; MBC: microbial biomass carbon; MBN: microbial biomass nitrogen; Egner Rhiem K: extractable potassium; NO_3^-: nitrate; NH_4^+: ammonium.

Effects	**Properties**	**SS**	**Df**	**MS**	**F**	***P***
Model	P-Olsen (mg kg^{-1})	505.891	17	29.758	3.740	0.0001
	pH (H_2O)	111814.276	17	6577.310	1.007	0.474
	SOC (g kg^{-1})	5434.493	17	319.676	3.411	0.001
	MBC (mg C kg^{-1})	68406.261	17	4023.898	6.527	0.0001
	MBN (mg N kg^{-1})	1582.161	17	93.068	5.750	0.0001
	Electrical conductivity (µS cm^{-1})	85026.219	17	5001.542	1.938	0.047
	EgnerRhiem K (mg kg^{-1})	43334.052	17	2549.062	1.504	0.149
	NO_3^- (mg kg^{-1})	1432.803	17	84.283	7.718	0.0001
	NH_4^+ (mg kg^{-1})	2455.811	17	144.459	4.053	0.0001
	Moisture (g kg^{-1})	2083.806	17	122.577	7.806	0.0001
	µg N-NH_4^+ · g^{-1} · h^{-1}	1580.438	17	92.967	1.045	0.438
	β-glucosidase (µg p-Nitrophenol g^{-1} h^{-1})	4719.768	17	277.633	2.412	0.013
	Acidphosphatase (µg p-Nitrophenol g^{-1} h^{-1})	1665466.576	17	97968.622	10.711	0.0001
	Alkaline phosphatase (µg p-Nitrophenol g^{-1} h^{-1})	83057.654	17	4885.744	4.854	0.0001
Altitude	P-Olsen (mg kg^{-1})	87.644	2	43.822	5.508	0.008
	pH (H_2O)	13365.661	2	6682.830	1.023	0.370
	SOC (g kg^{-1})	1004.466	2	502.233	5.359	0.009
	MBC (mg C kg^{-1})	39770.413	2	19885.207	32.255	0 .0001
	MBN (mg N kg^{-1})	754.810	2	377.405	23.317	0.0001
	Electrical conductivity (µS cm^{-1})	37305.390	2	18652.695	7.227	0.002
	EgnerRhiem K (mg kg^{-1})	13962.918	2	6981.459	4.119	0.025
	NO_3^- (mg kg^{-1})	552.831	2	276.416	25.311	0.0001
	NH_4^+ (mg kg^{-1})	1104.984	2	552.492	15.501	0.0001

	Moisture (g kg^{-1})	1295.632	2	647.816	41.254	0.0001
	Urease (µg N-NH_4^+ g^{-1} h^{-1})	406.624	2	203.312	2.286	0.116
	β-glucosidase (µg p-Nitrophenol g^{-1} h^{-1})	1263.886	2	631.943	5.491	0.008
	Acidphosphatase (µg p-Nitrophenol g^{-1} h^{-1})	1281916.640	2	640958.320	70.079	0.0001
	Alkaline phosphatase (µg p-Nitrophenol g^{-1} h^{-1})	44570.567	2	22285.284	22.138	0.0001
Shade	P-Olsen (mg kg^{-1})	76.766	2	38.383	4.824	0.014
	pH (H_2O)	13231.332	2	6615.666	1.013	0.373
	SOC (g kg^{-1})	405.305	2	202.653	2.162	0.130
	MBC (mg C kg^{-1})	3623.825	2	1811.912	2.939	0.066
	MBN (mg N kg^{-1})	155.852	2	77.926	4.814	0.014
	Electrical conductivity (µS cm^{-1})	10984.648	2	5492.324	2.128	0.134
	EgnerRhiem K (mg kg^{-1})	8234.693	2	4117.346	2.429	0.102
	NO_3^- (mg kg^{-1})	9.491	2	4.745	0.435	0.651
	NH_4^+ (mg kg^{-1})	24.351	2	12.176	0.342	0.713
	Moisture (g kg^{-1})	183.477	2	91.739	5.842	0.006
	Urease (µg N-NH_4^+ g^{-1} h^{-1})	441.641	2	220.820	2.483	0.098
	β-glucosidase (µg p-Nitrophenol g^{-1} h^{-1})	335.750	2	167.875	1.459	0.246
	Acidphosphatase (µg p-Nitrophenol g^{-1} h^{-1})	21127.927	2	10563.963	1.155	0.326
	Alkaline phosphatase (µg p-Nitrophenol g^{-1} h^{-1})	420.710	2	210.355	0.209	0.812
Altitude * Shade	P-Olsen (mg kg^{-1})	292.137	4	73.034	9.180	0.0001
	pH (H_2O)	26299.882	4	6574.970	1.006	0.417
	SOC (g kg^{-1})	2224.362	4	556.091	5.933	0.001
	MBC (mg C kg^{-1})	5704.639	4	1426.160	2.313	0.076
	MBN (mg N kg^{-1})	304.339	4	76.085	4.701	0.004
	Electrical conductivity (µS cm^{-1})	10834.443	4	2708.611	1.049	0.396
	EgnerRhiem K (mg kg^{-1})	12239.056	4	3059.764	1.805	0.149
	NO_3^- (mg kg^{-1})	681.613	4	170.403	15.604	0.0001
	NH_4^+ (mg kg^{-1})	1106.229	4	276.557	7.759	0.0001
	Moisture (g kg^{-1})	443.193	4	110.798	7.056	0.0001

	Urease (µg N-NH$_4^+$ g^{-1} h^{-1})	403.370	4	100.842	1.134	0.356
	β-glucosidase (µg p-Nitrophenol g^{-1} h^{-1})	2328.560	4	582.140	5.058	0.0021
	Acidphosphatase (µg p-Nitrophenol g^{-1} h^{-1})	268613.196	4	67153.299	7.342	0.0001
	Alkaline phosphatase (µg p-Nitrophenol g^{-1} h^{-1})	15635.442	4	3908.860	3.883	0.0101

Supplementary Information 10. **Table S3.2.** Total number of bases, read counts, % of GC and Q values (Q20 and Q30) in each *Coffea* library concerning the sequencing of Bacteria, Archaea and Fungi.

	Bacteria					**Archaea**					**Fungi**				
SampleName	Total Bases	ReadCount	GC (%)	Q20 (%)	Q30 (%)	Total Bases	ReadCount	GC (%)	Q20 (%)	Q30 (%)	Total Bases	ReadCount	GC (%)	Q20 (%)	Q30 (%)
1.600.S0.16S	43 359 319	96 152	57.91	99.06	96.27	24 050 509	81 147	55.78	99.81	99.03	31 940 929	88 300	47.81	99.66	98.16
1.600.S100.16S	43 581 768	97 173	57.71	99.17	96.63	26 019 128	87 852	55.72	99.81	99.06	38 974 546	106 496	48.68	99.66	98.21
1.600.S50.16S	44 578 935	99 176	56.84	99.13	96.53	24 225 221	82 832	55.84	99.82	99.09	37 582 027	100 297	48.26	99.63	98.11
1.800.S0.16S	41 446 794	92 173	57.23	99.13	96.56	26 296 802	90 622	55.89	99.82	99.1	37 715 724	95 532	45.31	99.64	98.40
1.800.S100.16S	34 146 049	76 025	57.03	98.41	93.61	25 292 138	88 449	55.80	99.83	99.16	38 696 704	95 985	45.77	99.56	98.08
1.800.S50.16S	39 404 314	87 808	57.11	99.16	96.62	27 251 913	98 047	55.88	99.86	99.25	33 866 564	85 534	45.75	99.63	98.28
1.900.S0.16S	40 176 314	89 097	56.76	99.01	96.05	26 734 514	92 234	55.84	99.82	99.11	38 273 964	97 161	46.20	99.62	98.18
1.900.S100.16S	39 794 779	88 314	57.07	99.08	96.35	28 086 440	97 338	55.78	99.83	99.16	37 073 431	91 149	44.92	99.58	98.21
1.900.S50.16S	41 396 645	91 935	56.95	99.09	96.42	28 166 911	98 519	55.71	99.83	99.14	42 707 269	112 620	48.61	99.65	98.19
2.600.S0.16S	37 664 037	83 159	57.58	99.02	96.13	27 525 209	97 933	56.11	99.83	99.15	36 968 222	85 375	42.48	99.59	98.33
2.600.S100.16S	40 247 492	88 962	57.55	99.01	96.09	23 483 900	84 889	56.45	99.85	99.22	34 382 023	89 164	47.81	99.60	98.01
2.600.S50.16S	45 263 830	101 219	57.71	99.12	96.46	23 922 546	84 265	56.24	99.84	99.20	36 628 777	98 969	48.72	99.65	98.10
2.800.S0.16S	41 501 529	92 396	57.24	99.12	96.45	27 523 261	94 490	55.81	99.83	99.14	38 427 142	100 766	46.75	99.63	98.17
2.800.S100.16S	38 466 223	85 511	57.08	99.13	96.54	23 902 311	80 075	55.81	99.81	99.04	34 776 473	87 708	45.29	99.61	98.23
2.800.S50.16S	36 610 598	81 458	57.00	99.07	96.23	26 562 098	91 508	55.87	99.83	99.14	31 909 194	86 989	49.37	99.62	98.05
2.900.S0.16S	39 882 431	88 483	56.84	99.09	96.37	26 366 580	92 238	55.89	99.83	99.11	37 799 866	95 273	45.92	99.59	98.17
2.900.S100.16S	41 870 628	92 810	56.88	99.12	96.55	28 210 800	96 649	55.92	99.82	99.11	32 333 670	77 881	43.49	99.50	97.99
2.900.S50.16S	39 157 569	86 595	57.06	99.03	96.18	26 853 732	95 913	55.96	99.84	99.20	36 214 871	89 946	45.24	99.53	97.88
3.600.S0.16S	37 307 184	82 950	57.72	99.14	96.55	29 440 079	105 106	56.08	99.85	99.24	39 194 511	99 451	47.22	99.63	98.17
3.600.S100.16S	41 365 884	91 505	57.51	99.03	96.18	22 815 077	82 608	56.50	99.84	99.15	39 177 033	93 914	44.30	99.53	98.05
3.600.S50.16S	42 230 248	93 805	57.21	99.10	96.42	27 994 060	100 401	56.45	99.85	99.21	37 975 536	106 738	54.10	99.67	98.07
3.800.S0.16S	39 326 715	87 576	57.24	99.12	96.45	24 354 124	82 161	55.76	99.82	99.15	36 208 114	88 972	44.87	99.59	98.17
3.800.S100.16S	41 192 815	91 830	56.96	99.18	96.71	26 625 981	91 773	55.87	99.83	99.15	32 291 919	84 537	47.20	99.63	98.24
3.800.S50.16S	39 536 216	88 075	57.04	99.19	96.76	26 887 542	92 902	55.86	99.83	99.13	33 509 015	88 993	47.38	99.65	98.26

3.900.S0.16S	41 314 347	91 713	56.91	99.14	96.61	25 521 200	89 357	55.84	99.84	99.18	38 331 608	97 402	46.22	99.61	98.23
3.900.S100.16S	42 270 136	93 657	56.70	99.07	96.34	29 528 125	103 562	55.78	99.84	99.18	34 056 327	83 151	44.61	99.57	98.10
3.900.S50.16S	35 892 730	79 381	56.48	99.12	96.52	28 279 926	100 511	56.15	99.85	99.21	39 988 916	96 390	44.56	99.51	97.96
	40 332 797	**89 590**	**57**	**99**	**96**	**26 367 412**	**91 977**	**56**	**100**	**99**	**36 555 718**	**93 507**	**47**	**100**	**98**

Supplementary Information 11. **Table S3.3**. Pairwise alpha Bacteria diversity statistical comparison between the levels of shade based on the Wilcoxon test.

Index	**Group1**	**Group2**	**P value**
Observed OTUs	0%	50%	0.258083
	0%	100%	0.730440
	50%	100%	0.796174
Shannon entropy	0%	50%	0.190251
	0%	100%	0.931427
	50%	100%	0.297326
Pielou's evenness	0%	50%	0.386508
	0%	100%	0.666475
	50%	100%	0.297326

Supplementary Information 12. **Table S3.4.** Pairwise alpha Bacteria diversity statistical comparison between altitudinal ranges based on the Wilcoxon test. Significant values are highlighted in bold.

Index	**Group1**	**Group2**	**P value**
Observed OTUs	600	800	0.007775
	600	900	0.258083
	800	900	0.113492
Shannon entropy	600	800	0.001851
	600	900	0.489428
	800	900	0.000494
Pielou's evenness	600	800	0.001851
	600	900	0.604813
	800	900	8.23E-05

Supplementary Information 13. **Table S3.5**. Pairwise alpha Archaea diversity statistical comparison between shade trees based on the Wilcoxon test.

Index	**Group1**	**Group2**	**P value**
Observed OTUs	0%	50%	0.790447
	0%	100%	0.056472
	50%	100%	0.166879
Shannon_entropy	0%	50%	0.863307
	0%	100%	0.050309
	50%	100%	0.062526
Pielou_evenness	0%	50%	0.796174
	0%	100%	0.077005
	50%	100%	0.093912

Supplementary Information 14. **Table S3.6**. Pairwise alpha Archaea diversity statistical comparison between altitudinal ranges based on the Wilcoxon test. Significant values are highlighted in bold.

Index	**Group1**	**group2**	**P value**
Observed OTUs	600	900	0.008657
	800	900	0.893861
Shannon_entropy	600	800	0.001234
	600	900	0.000288
	800	900	0.545701
Pielou_evenness	600	800	0.002756
	600	900	0.000494
	800	900	0.489428

Supplementary Information 15. **Table S3.7**. Pairwise alpha Fungi diversity statistical comparison between shade trees based on the Wilcoxon test.

Index	**Group1**	**group2**	**P value**
Observed OTUs	0%	50%	0.796174
	0%	100%	0.039984
	50%	100%	0.062526
Shannon entropy	0%	50%	0.666475
	0%	100%	0.093912
	50%	100%	0.340107
Pielou's evenness	0%	50%	0.73044
	0%	100%	0.135911
	50%	100%	0.340107

Supplementary Information 16. **Table S3.8**. Pairwise alpha Fungi diversity statistical comparison between altitudinal ranges based on the Wilcoxon test.

Index	Group1	Group2	P value
Observed OTUs	600	800	1
	600	900	0.863307
	800	900	0.796174
Shannon entropy	600	800	0.73044
	600	900	0.604813
	800	900	0.489428
Pielou's evenness	600	800	0.73044
	600	900	0.796174
	800	900	0.386508

Supplementary Information 17. **Table S3.9**. List of all differentially abundant bacterial genera across different altitudes.

				Altitude		
Bacteria	*P* values	FDR	600	800	900	LDAscore
Chthoniobacter	0.000363	0.003079	1562300	2835200	2458700	5.8
Arthrobacter	2.48E-05	0.000741	188930	1234400	50336	5.77
Gaiella	0.00089	0.00457	694640	126670	230670	5.45
Occallatibacter	0.002632	0.009948	237900	278360	640490	5.3
Natranaerobaculum	0.015112	0.033111	338380	110840	242100	5.06
Solirubrobacter	0.000115	0.001617	228610	161470	49066	4.95
Vicinamibacter	0.014171	0.031401	242100	72523	218060	4.93
Stella	7.53E-05	0.001469	72719	227930	109760	4.89
Paludibaculum	0.006008	0.016979	169190	233110	325770	4.89
Acidibacter	0.001664	0.006906	90507	228130	233990	4.86
Aciditerrimonas	0.000335	0.002968	264090	136050	186490	4.81
Rhodanobacter	0.005641	0.016177	97.74	195.48	121200	4.78
Pedococcus	0.003712	0.01227	110150	145630	30690	4.76
Nitrospira	0.000217	0.00223	89726	9480.8	17105	4.6
Pyrinomonas	9.74E-05	0.001617	2345.8	4887	77508	4.57
Desulfonatronum	0.000127	0.001648	71937	97.74	50923	4.56
Aliidongia	7.53E-05	0.001469	18473	31863	85816	4.53
Sulfurirhabdus	1.09E-05	0.000426	14563	684.18	65779	4.51
Limnoglobus	0.001214	0.00579	56201	52975	117000	4.51
Fimbriiglobus	0.004222	0.013723	56592	78583	119540	4.5
Roseisolibacter	0.018465	0.039138	92169	31179	39780	4.48
Gemmata	6.91E-05	0.001469	33036	49261	88846	4.45
Actinoplanes	0.000251	0.002448	57862	3029.9	12413	4.44
Pseudomonas	0.028658	0.054256	8601.1	21503	62945	4.43
Nitrospirillum	0.006632	0.018474	48479	39780	86500	4.37
Vitiosangium	0.000609	0.004098	11924	7819.2	51314	4.34
Anaeromyxobacter	0.013382	0.030342	50727	6841.8	33525	4.34
Pedomicrobium	0.01692	0.03666	98131	55419	96567	4.33
Thermogemmatispora	0.000116	0.001617	39585	29224	1075.1	4.28
Actinomarinicola	0.004953	0.015044	65975	28345	32059	4.27
Thermanaerovibrio	0.00779	0.020335	54441	23849	17691	4.26

Hydrogenispora	0.00011	0.001617	25999	41246	5473.5	4.25
Edaphobacter	0.00806	0.02068	43592	78974	60403	4.25
Rugosimonospora	0.001375	0.006264	69298	34600	47893	4.24
Pseudorhodoplanes	0.000547	0.003951	11533	15638	45156	4.23
Micromonospora	0.000183	0.002103	43983	13195	45645	4.21
Brevitalea	0.007542	0.020147	28247	4984.8	36653	4.2
Gelria	0.011101	0.025468	40464	31375	11142	4.17
Reyranella	0.004345	0.013852	43494	16518	34600	4.13
Geobacter	0.000736	0.004413	13195	6841.8	33036	4.12
Nitrosospira	0.013705	0.030717	36555	10165	13488	4.12
Neobacillus	0.026456	0.051588	53171	30984	57178	4.12
Priestia	0.008995	0.022487	37630	11827	26976	4.11
Zavarzinella	0.001453	0.006295	30299	5375.7	27563	4.1
Metabacillus	0.002818	0.010122	37923	13488	30104	4.09
Terrimicrobium	0.004475	0.013852	16127	35382	14466	4.02
Jatrophihabitans	0.009808	0.023367	48675	28638	27954	4.02
Roseiarcus	0.000602	0.004098	3714.1	13977	23849	4
Variovorax	0.000769	0.004413	3029.9	5962.2	23067	4
Thermanaeromonas	0.005015	0.015044	24044	7526	27563	4
Niastella	0.009946	0.023367	6646.3	4984.8	24044	3.98
Mesorhizobium	0.018796	0.039411	17886	16127	35284	3.98
Mycolicibacterium	0.010851	0.02519	34405	22578	15932	3.97
Mycolicibacillus	0.00045	0.003655	18473	7135	1563.8	3.93
Marinobacterium	4.37E-06	0.000426	0	16323	0	3.91
Terrimonas	0.003366	0.01172	4984.8	2541.2	18962	3.91
Kribbella	0.000768	0.004413	16323	1368.4	3518.6	3.87
Acidibrevibacterium	0.005379	0.015656	23262	15052	8405.7	3.87
Domibacillus	0.006859	0.018575	19744	4887	10654	3.87
Acidiferrimicrobium	0.000214	0.00223	2834.5	15932	3225.4	3.82
Piscinibacterium	0.001381	0.006264	2736.7	15345	2052.5	3.82
Thermoflavimicrobium	0.003431	0.011737	11045	15638	2443.5	3.82
Tuwongella	0.000848	0.004493	1954.8	2541.2	14857	3.81
Risungbinella	0.018062	0.038705	6548.6	19450	14075	3.81
Luteitalea	0.000747	0.004413	5180.2	12120	195.48	3.78
Desulfofundulus	0.009498	0.023367	10751	13977	3029.9	3.74

Aureimonas	0.001098	0.005491	2932.2	5962.2	13586	3.73
Skermanella	0.002855	0.010122	3127.7	13781	11631	3.73
Sphaerobacter	0.001865	0.007578	8014.7	12218	1661.6	3.72
Oceanibaculum	0.001665	0.006906	7037.3	6157.6	15834	3.68
Laceyella	8.99E-06	0.000426	195.48	9285.3	0	3.67
Duganella	0.021963	0.044152	0	9383.1	293.22	3.67
Methyloceanibacter	0.00016	0.001947	9871.8	684.18	9480.8	3.66
Pseudonocardia	0.019882	0.040811	10458	1368.4	1661.6	3.66
Granulicella	0.008182	0.020721	3323.2	7819.2	12218	3.65
Krasilnikovia	0.000852	0.004493	10458	1661.6	2443.5	3.64
Agrobacterium	0.019536	0.040527	11240	2736.7	6353.1	3.63
Rhodopila	0.000819	0.004493	11631	10947	3323.2	3.62
Peribacillus	0.000497	0.003873	7330.5	1172.9	1661.6	3.49
Ramlibacter	0.022289	0.04435	8796.6	3811.9	9871.8	3.48
Candidatus_Protochlamydia	4.40E-06	0.000426	0	5962.2	0	3.47
Chloroflexus	0.004475	0.013852	5962.2	0	195.48	3.47
Ilumatobacter	0.002653	0.009948	6939.6	1563.8	2150.3	3.43
Acidisphaera	2.66E-05	0.000741	488.7	5082.5	0	3.41
Acidiphilium	0.002765	0.010122	0	1759.3	4887	3.39
Thermoflexus	1.03E-05	0.000426	4593.8	0	97.74	3.36
Methylocystis	0.009758	0.023367	6353.1	4789.3	2150.3	3.32
Singulisphaera	0.000738	0.004413	97.74	977.4	4202.8	3.31
Archangium	0.009945	0.023367	4007.4	390.96	3225.4	3.26
Egibacter	0.000535	0.003951	3323.2	0	586.44	3.22
Desulfomonile	0.000301	0.002791	684.18	3029.9	0	3.18
Hyphomicrobium	0.001452	0.006295	97.74	2932.2	781.92	3.15
Caedimonas	0.026374	0.051588	195.48	2834.5	2834.5	3.12
Ralstonia	0.003551	0.011939	0	2443.5	2541.2	3.1
Sinosporangium	0.007821	0.020335	977.4	2443.5	0	3.09
Kallotenue	0.006833	0.018575	2248	390.96	0	3.05
Vulgatibacter	0.02046	0.04156	390.96	684.18	2541.2	3.03
Hazenella	0.002376	0.009266	1954.8	586.44	0	2.99
Rhodovastum	0.005282	0.015607	293.22	1954.8	0	2.99
Burkholderia	0.002204	0.008772	0	97.74	1759.3	2.94
Caldanaerobacter	0.027197	0.05251	1759.3	0	1270.6	2.94

Aquihabitans	0.028114	0.053748	1661.6	488.7	0	2.92
Aldersonia	0.001217	0.00579	1563.8	0	195.48	2.89

Supplementary Information 18. **Table S3.10**. List of all differentially abundant fungi across different altitudes.

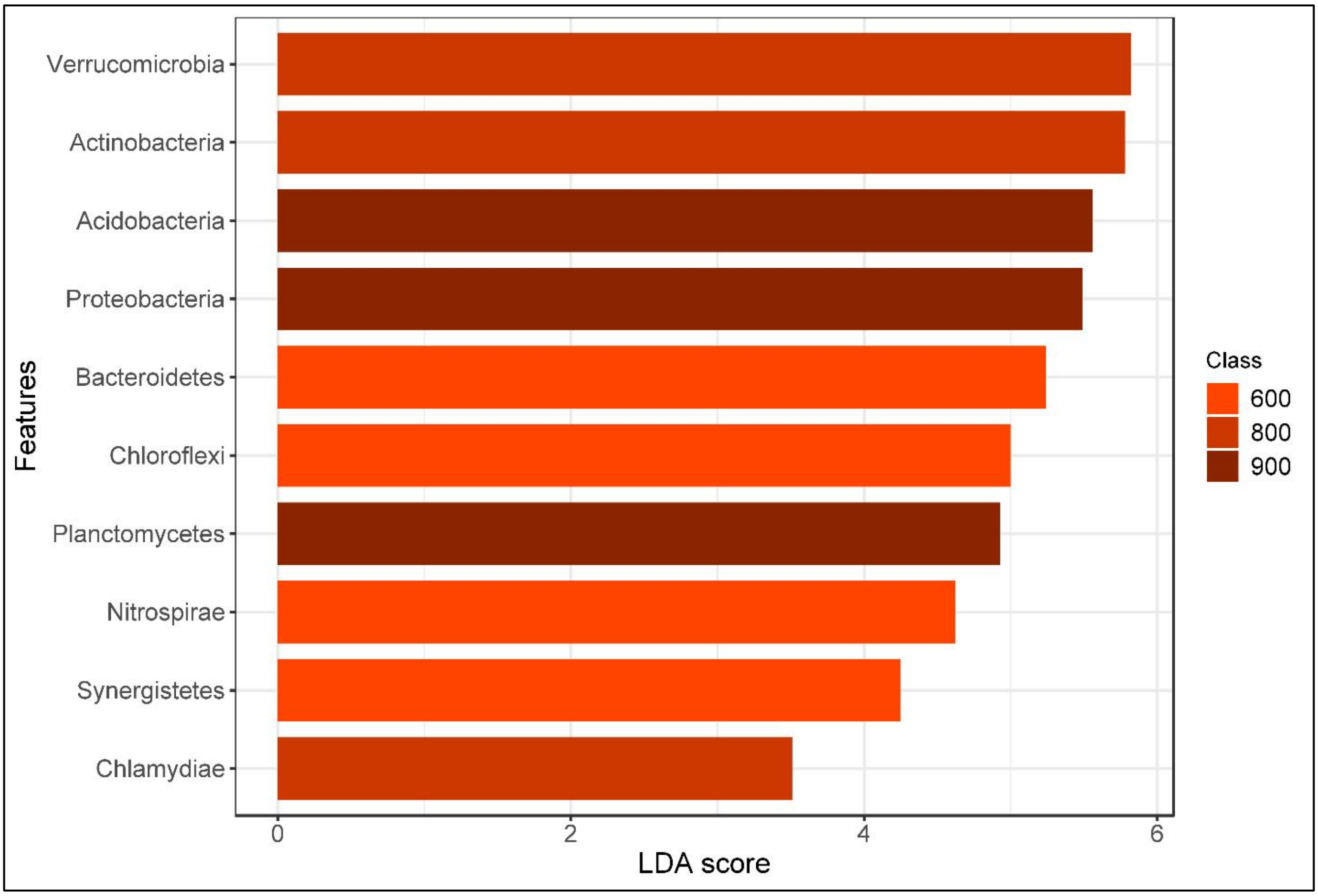
			Altitude			
Fungi	***P* values**	**FDR**	**600**	**800**	**900**	**LDAscore**
Podila	0.000147	0.005405	0	16826	751980	5.58
Calonectria	0.005185	0.053009	86739	15156	0	4.64
Ilyonectria	0.006751	0.053009	18940	17730	97105	4.6
Gliocladiopsis	0.00393	0.045713	10143	60332	5318.9	4.44
Exophiala	0.004302	0.047536	40455	4534.7	2915.2	4.27
Cladophialophora	0.003225	0.039601	36176	6052	3869.9	4.21
Cystofilobasidium	6.36E-05	0.00281	0	26220	17.048	4.12
Cutaneotrichosporon	2.71E-05	0.001995	0	0	22657	4.05
Phialemoniopsis	0.001744	0.027061	2642.4	19503	6341.8	3.93
Dactylonectria	0.006127	0.053009	8217.1	0	16349	3.91
Rhexocercosporidium	0.006828	0.053009	16213	3682.3	0	3.91
Pestalotiopsis	0.001271	0.021999	16230	596.67	16588	3.9
Stagonospora	0.002603	0.03595	1841.2	17457	3699.4	3.89
Gongronella	0.007196	0.053009	9325.2	15462	1858.2	3.83
Leptodontidium	0.001294	0.021999	0	681.91	11541	3.76
Niesslia	2.69E-05	0.001995	10723	0	0	3.73
Keithomyces	2.71E-05	0.001995	0	0	9563.8	3.68
Ganoderma	0.007008	0.053009	13434	4313.1	4279	3.66
Corynascella	3.84E-05	0.002124	8643.3	0	954.68	3.64
Neosartorya	0.001837	0.027061	1039.9	8404.6	1142.2	3.57
Fusicladium	0.00069	0.015249	6069	0	0	3.48
Neoscirrhia	0.006625	0.053009	0	4517.7	6052	3.48
Volutella	0.007139	0.053009	375.05	6052	4483.6	3.45
Xenoacremonium	0.005943	0.053009	698.96	17.048	3205	3.2
Hawksworthiomyces	0.000974	0.01956	596.67	2557.2	0	3.11
Septoria	0.00576	0.053009	1585.5	733.06	0	2.9
Preussia	0.00069	0.015249	1500.2	0	0	2.88

Glomus	0.000289	0.009138	34.096	1517.3	85.239	2.87
Mycena	0.002925	0.038026	0	0	954.68	2.68
Stegonsporium	0.00069	0.015249	0	0	698.96	2.54

Supplementary Information 19. **Table S3.11**. Results of one-way ANOVA testing the effects of altitude on soil properties. Significant variables are indicated in bold.

Properties	SS	df	MS	F	P
pH	0.06	2	0.03	0.60	0.57
Extractable P	1746.88	2	873.44	2.83	0.13
Extractable Ca	175088.88	2	87544.44	1.39	0.31
Extractable Mg	3960.66	2	1980.33	0.53	0.60
Organic Matter	1.49	2	0.74	2.35	0.17
Electrical conductivity	0.01	2	0.006	5.76	0.04
Total Limestone	*	*	*	*	*
Exchangable acidity	0.53	2	0.26	1.36	0.32
Na	2.22E-05	2	1.11E-05	0.07	2.22E-05
K	0.007	2	0.0039	2.6	0.15
Ca	4.34	2	2.17	1.39	0.31
Mg	0.27	2	0.13	0.54	0.60
CEC	6.45	2	3.23	1.02572757	0.41
Base saturation	517.17	2	258.58	1.33	0.33
*< 0.5					

Supplementary Information 20. **Figure S3.1**. Differentially abundant bacterial phylum at varying altitudes.



Chapter IV

Supplementary Information 21. **Table S4.1.** Read counts in each *Coffea* library related to the sequencing of Bacteria (16S rRNA) and Fungi (ITS). F = Fruits; L = Leaves; 1, 2, 3 = replicates; 600, 800, 900 = altitude (m); 0, 50, 100 = Shade level (0, 50, 100%).

Sample ID	**16S Rrna**	**ITS region**
	Frequency of reads	
1_600_F0__1	43269	42391
1_600_F100__1	45293	41709
1_600_F50__1	44019	41404
1_600_L0__1	53586	55577
1_600_L100__1	54707	46529
1_600_L50__1	51742	43343
1_800_F0__1	50598	49717
1_800_F100__1	45096	50845
1_800_F50__1	32193	52859
1_800_L0__1	49036	58018
1_800_L100__1	54584	51429
1_800_L50__1	42621	54515
1_900_F0__1	37789	48161
1_900_F100__1	45836	42131
1_900_F50__1	40442	38747
1_900_L0__1	49544	56521
1_900_L100__1	49981	52920
1_900_L50__1	44821	60938
2_600_F0__1	39196	57277
2_600_F100__1	43587	43179
2_600_F50__1	38612	51481
2_600_L0__1	48471	48031
2_600_L100__1	54548	51889
2_600_L50__1	50901	56779
2_800_F0__1	46825	28169
2_800_F100__1	54424	48603
2_800_F50__1	47758	48899
2_800_L0__1	46333	55632
2_800_L100__1	53886	48355
2_800_L50__1	52575	56778
2_900_F0__1	43597	38951
2_900_F100__1	48103	38402
2_900_F50__1	42421	24727
2_900_L0__1	47949	44919
2_900_L100__1	52304	44023
2_900_L50__1	42782	73380
3_600_F0__1	30820	46319
3_600_F100__1	46138	42048
3_600_F50__1	32615	52237
3_600_L0__1	45140	57214

3_600_L100__1	44553	50862
3_600_L50__1	49395	53706
3_800_F0__1	45074	33353
3_800_F100__1	50810	42583
3_800_F50__1	48000	48622
3_800_L0__1	57944	61649
3_800_L100__1	41236	56066
3_800_L50__1	53738	54468
3_900_F0__1	50710	36567
3_900_F100__1	38137	49321
3_900_F50__1	48180	31642
3_900_L0__1	52121	50740
3_900_L100__1	59425	52537
3_900_L50__1	48202	57798
Total	**2531667**	**2624960**

Supplementary Information 22. **Table S4.2A**. Observed ASVs (p = 0.04), Shannon-entropy, Pielou's evenness (p = 0.01) indexes of bacterial communities in the fruits (F) and leaves (L), along the gradient of altitude (600, 800, 900 m) and shade (0, 50, 100%).

Sample ID	Observed ASVs	Shannon-entropy	Pielou's evenness	Altitude	Shade	Host
1_600_F0__1	7	2.510996028	0.894434832	600	0	Fruit
1_600_F100__1	14	2.310900155	0.606956851	600	100	Fruit
1_600_F50__1	10	1.807117667	0.543996623	600	50	Fruit
1_600_L0__1	4	1.4172598	0.7086299	600	0	Leaf
1_600_L100__1	4	1.361926513	0.680963256	600	100	Leaf
1_600_L50__1	1	0	NA	600	50	Leaf
1_800_F100__1	4	1.905003507	0.952501753	800	100	Fruit
1_800_L0__1	4	1.922023143	0.961011571	800	0	Leaf
1_800_L100__1	3	1.531000155	0.965953551	800	100	Leaf
1_800_L50__1	10	2.967098716	0.893185714	800	50	Leaf
1_900_F0__1	1	0	NA	900	0	Fruit
1_900_F100__1	24	3.686461147	0.804032998	900	100	Fruit
1_900_F50__1	1	0	NA	900	50	Fruit
1_900_L0__1	50	4.943459942	0.875901117	900	0	Leaf
1_900_L100__1	41	4.936404825	0.921391863	900	100	Leaf
1_900_L50__1	119	5.657155663	0.820493863	900	50	Leaf
2_600_F0__1	21	1.807518502	0.411518187	600	0	Fruit
2_600_F100__1	3	0.638468567	0.402828816	600	100	Fruit
2_600_F50__1	9	1.924477917	0.607105189	600	50	Fruit

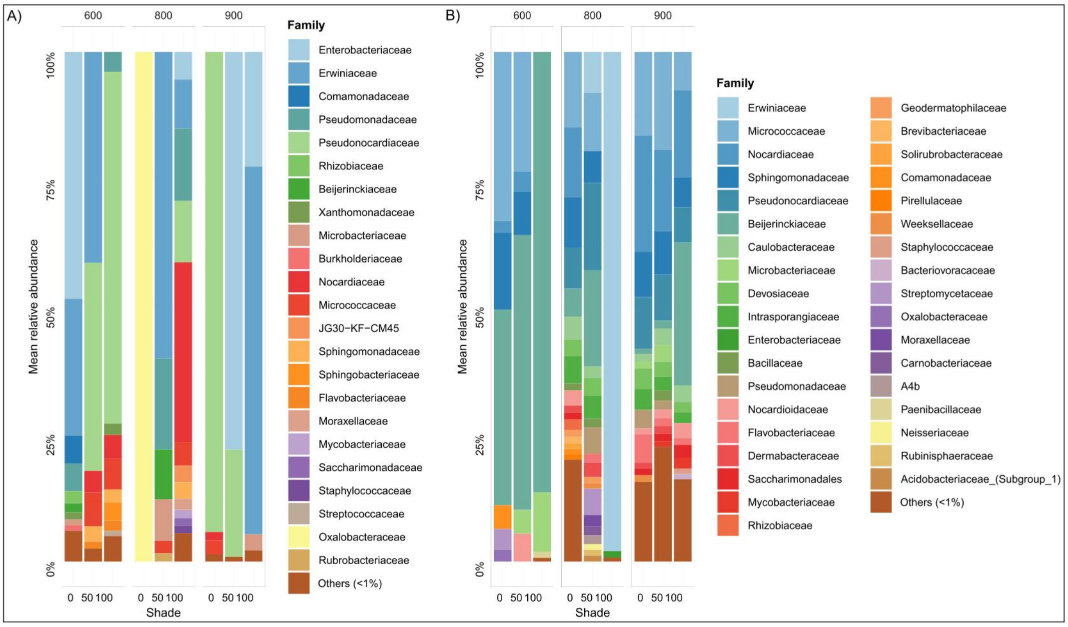
2_600_L0__1	6	1.878860874	0.726842604	600	0	Leaf
2_600_L100__1	8	2.69816932	0.899389773	600	100	Leaf
2_600_L50__1	9	2.841979487	0.896544709	600	50	Leaf
2_800_F0__1	1	0	NA	800	0	Fruit
2_800_F100__1	15	2.729400567	0.698611978	800	100	Fruit
2_800_F50__1	3	1.530493057	0.965633607	800	50	Fruit
2_800_L0__1	14	3.612031933	0.948698508	800	0	Leaf
2_800_L100__1	1	0	NA	800	100	Leaf
2_800_L50__1	9	3.060754663	0.965560593	800	50	Leaf
2_900_F0__1	1	0	NA	900	0	Fruit
2_900_F100__1	7	2.466055889	0.878426832	900	100	Fruit
2_900_F50__1	5	0.718824173	0.309580721	900	50	Fruit
2_900_L0__1	18	3.912139616	0.938179851	900	0	Leaf
2_900_L100__1	21	4.175473061	0.95063099	900	100	Leaf
2_900_L50__1	33	4.582626721	0.908459294	900	50	Leaf
3_600_F0__1	38	3.541104439	0.674762452	600	0	Fruit
3_600_F100__1	10	2.306729195	0.694394679	600	100	Fruit
3_600_F50__1	2	0.997294382	0.997294382	600	50	Fruit
3_600_L50__1	2	0.811278124	0.811278124	600	50	Leaf
3_800_F100__1	10	2.983384072	0.898088094	800	100	Fruit
3_800_F50__1	6	2.479795206	0.959315737	800	50	Fruit
3_800_L0__1	93	5.580709946	0.853429334	800	0	Leaf
3_800_L100__1	22	1.868925682	0.419095042	800	100	Leaf
3_800_L50__1	13	3.528573329	0.953555144	800	50	Leaf
3_900_F0__1	7	1.016400976	0.362049333	900	0	Fruit
3_900_F100__1	4	1.949464303	0.974732151	900	100	Fruit
3_900_F50__1	1	0	NA	900	50	Fruit
3_900_L0__1	26	4.267494494	0.907892612	900	0	Leaf
3_900_L100__1	19	3.639367203	0.856739478	900	100	Leaf
3_900_L50__1	45	4.771968974	0.868917812	900	50	Leaf

Supplementary Information 23. **Table S4.2B.** Observed ASVs (*p* = 0.04), Shannon-entropy, Pielou's evenness (*p* = 0.01) indexes of fungal communities in the fruits (F) and leaves (L), along the gradient of altitude (600, 800, 900 m) and shade (0, 50, 100%).

Sample ID	Observed ASVs	Shannon-entropy	Pielou's evenness	Altitude	Shade	Host
1_600_F0__2	6	0.037437139	0.014482662	600	0	Fruit
1_600_F100__2	26	3.252809347	0.692022352	600	100	Fruit
1_600_F50__2	17	2.936922126	0.71851959	600	50	Fruit
1_600_L0__2	50	3.467436159	0.614373585	600	0	Leaf
1_600_L100__2	46	2.652640195	0.480240869	600	100	Leaf
1_600_L50__2	36	3.412516108	0.660070718	600	50	Leaf
1_800_F0__2	1	0	NA	800	0	Fruit
1_800_F100__2	5	2.183383098	0.940331918	800	100	Fruit
1_800_F50__2	1	0	NA	800	50	Fruit
1_800_L0__2	70	2.286348506	0.373020547	800	0	Leaf
1_800_L100__2	41	3.924059555	0.732435178	800	100	Leaf
1_800_L50__2	75	5.069611845	0.813896197	800	50	Leaf
1_900_F0__2	5	2.267117876	0.976394524	900	0	Fruit
1_900_F100__2	54	3.684545688	0.640246345	900	100	Fruit
1_900_F50__2	1	0	NA	900	50	Fruit
1_900_L0__2	45	2.128614092	0.387594871	900	0	Leaf
1_900_L100__2	44	2.547088299	0.466548256	900	100	Leaf
1_900_L50__2	87	2.592665756	0.402403926	900	50	Leaf
2_600_F0__2	4	0.238451777	0.119225888	600	0	Fruit
2_600_F100__2	7	0.154091907	0.054888645	600	100	Fruit
2_600_F50__2	20	1.278236598	0.2957561	600	50	Fruit
2_600_L0__2	25	4.190988018	0.902480147	600	0	Leaf
2_600_L100__2	67	2.642909346	0.43568587	600	100	Leaf
2_600_L50__2	69	3.314174776	0.542549154	600	50	Leaf
2_800_F0__2	1	0	NA	800	0	Fruit
2_800_F100__2	3	0.018152694	0.011453075	800	100	Fruit
2_800_F50__2	3	1.556656707	0.982141033	800	50	Fruit
2_800_L0__2	45	3.320015885	0.604534722	800	0	Leaf
2_800_L100__2	41	3.174449338	0.592518623	800	100	Leaf
2_800_L50__2	95	3.855554821	0.586855336	800	50	Leaf

2_900_F0__2	2	1	1	900	0	Fruit
2_900_F100__2	46	2.737251417	0.495559105	900	100	Fruit
2_900_F50__2	2	0.954434003	0.954434003	900	50	Fruit
2_900_L0__2	41	1.697924002	0.316921609	900	0	Leaf
2_900_L100__2	24	3.200193806	0.697976004	900	100	Leaf
2_900_L50__2	85	3.256484179	0.50808013	900	50	Leaf
3_600_F0__2	11	0.956329658	0.276441266	600	0	Fruit
3_600_F100__2	21	0.570940202	0.129986098	600	100	Fruit
3_600_F50__2	7	2.314884073	0.824578344	600	50	Fruit
3_600_L0__2	55	3.558203789	0.615461408	600	0	Leaf
3_600_L100__2	44	1.732484128	0.31733782	600	100	Leaf
3_600_L50__2	48	2.961466003	0.530257097	600	50	Leaf
3_800_F100__2	9	1.977511326	0.623835367	800	100	Fruit
3_800_F50__2	46	4.402028942	0.796954751	800	50	Fruit
3_800_L0__2	41	1.79242927	0.334561245	800	0	Leaf
3_800_L100__2	83	4.257647393	0.667862127	800	100	Leaf
3_800_L50__2	33	2.066921606	0.409746256	800	50	Leaf
3_900_F0__2	3	1.485475297	0.937230563	900	0	Fruit
3_900_F100__2	16	3.535665129	0.883916282	900	100	Fruit
3_900_L0__2	45	1.744634706	0.317676871	900	0	Leaf
3_900_L100__2	47	1.811465684	0.32612057	900	100	Leaf
3_900_L50__2	80	3.674780764	0.581275318	900	50	Leaf

Supplementary Information 24. **Figure S4.1**.Family level taxonomic profile of the bacterial communities of *C. arabica* grown under different altitudes (600m, 800m, and 900m) and under different levels of canopy shades from native trees (0%, 50%, and 100%).Only phylotypes with a relative abundance of at least 1% in any of the samples are presented. (A) Fruits; (B) Leaves.



Suplementary Information 25. **Figure S4.2**. Family-level taxonomic profile of the fungal communities of *C. arabica* grown under different altitudes (600m, 800m, and 900m) and under different levels of shade from native trees (0%, 50%, and 100%). (A) Fruits; (B) Leaves.