UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

The impact of land use change on the soil microbiome involved in the carbon and phosphorus cycles in a tropical ecosystem from Guinea-Bissau

Inês Ruivo da Costa Lopes Cordeiro

Mestrado em Microbiologia Aplicada

Dissertação orientada por:
Professora Doutora Mónica Sebastiana
Professora Doutora Filipa Monteiro

2025

Acknowledgements

I want to start by thanking Professor Mónica Sebastiana, my internal advisor, and Professor Filipa Monteiro for their expertise, guidance, patience, kindness and constant help. Without their unwavering support, this would not have been possible. I would like to thank my lab partners Ana João Martins and Rafael Pires for being the best lab partners anyone could ever ask for. If it was not for them, especially Ana, I could not have done any of this. More than lab partners they became my good friends, which made this experience even more gratifying. I would also like to thank Professor Susana Serrazina, Professor Fernando Vaz Dias and Professor Rita Teixeira, whose help in the lab was very much appreciated. I cannot leave out my university friends, both from my bachelors and masters, for always being there and championing me all the way to the end. The same goes for my friends from Iconic, without whom I would not even be the person I am today. Lastly, my most heartfelt and deepest thank you to my parents and grandmother who have never, ever, ceased to support me in everything I have ever done.

This Dissertation was fully performed at the Faculty of Sciences of the University of Lisbon under the direct co-supervision of Professor Filipa Monteiro.

Professor Mónica Sebastiana was the internal supervisor designated in the scope of the Master in Applied Microbiology of the Faculty of Sciences of the University of Lisbon.

Abstract

Approximately 30% of the Earth's land surface is covered by forests, home to microbial soil communities that are both highly diverse and poorly studied, especially in tropical regions and particularly in Africa. It has already been established that one of the biggest threats to soil biodiversity is land management, particularly the conversion of native forests into agricultural land. In this project, the impact of land use change on the soil microbial communities involved at the carbon (C) and phosphorus (P) cycles in the tropical soils of Guinea-Bissau was studied. This was done by quantifying, through qPCR, both taxonomic (bacterial 16S rRNA, fungal ITS and archaeal 16S rRNA) and functional genes that encode enzymes involved in processes like decomposition of plant litter (e.g. *β-glu, GH6 cellulase* and *pcaH*), methane consumption and production (e.g. *pmoA* and *mcrA*) and phosphorus mineralization (*phoD*) in soil samples from three different land uses systems (native tropical forest, cashew plantation and peanut field) at Guinea-Bissau. Additionally, soil enzymatic activities, namely β-glucosidase (carbon cycle) and phosphatase activity (phosphorus cycle) were assessed. The results, a significant higher abundance of *pmoA* in the cashew field and a significantly lower phosphatase activity in the peanut field, showed that forest conversion into agricultural fields significantly affects microorganisms involved in the carbon and phosphorus cycles. Although fertilizers and mechanical tillage are rarely used, the manual tillage and initial slash-and-burn practices are sufficient to notably impact methanotrophs (part of the carbon cycle) and phosphatase activity (linked to the phosphorus cycle). This suggests that transforming native tropical forests in Guinea-Bissau into agricultural fields significantly alters the soil microbiome responsible for nutrient cycling. However, these practices do not result in substantial changes to the overall abundance of bacteria and fungi in the soil.

Keywords: qPCR, Tropical soils, Carbon, Phosphorus, Land use

Resumo

O solo contém cerca de um quarto da biodiversidade da Terra, e nele vivem microrganismos que contribuem para as várias funções de suporte de vida desempenhadas pelo solo, tais como transformação do carbono, manutenção dos ciclos dos nutrientes, formação da estrutura do solo e regulação da biodiversidade (biocontrolo), todas elas importantes para a fertilidade dos solos, e, consequentemente, para a produção de alimentos e fibras para consumo humano. Em relação aos parâmetros avaliados aquando da transformação da floresta tropical nativa, percebeu-se que a comunidade metanotrófica varia significativamente, afetando o ciclo do carbono, e que a atividade fosfatásica também, afetando o ciclo do fósforo.

As zonas tropicais, em particular as florestas tropicais, principalmente os seus solos, têm sido muito pouco estudados, especialmente quando comparados com as zonas temperadas. Daí que se saiba muito pouco sobre os efeitos da sua conversão em terrenos agrícolas. Os solos tropicais são marcadamente diferentes dos solos temperados. São frágeis e propensos à degradação, especialmente quando há perda de vegetação, caracterizados pela forte presença de óxidos de ferro e alumínio, e pouco fósforo. Os principais problemas que enfrentam são a deposição de poeiras, devido aos ventos Harmattan que trazem resíduos do deserto do Saara, as queimadas, a desflorestação, a erosão do solo, o sobrepastoreio e a seca. Mudanças no uso e gestão do solo são as maiores ameaças à sua biodiversidade e às suas funções, sendo por isso que se procurou estudá-las neste projeto.

Em conjunto com as alterações climáticas, a crescente intensificação da agricultura põe em causa a funcionalidade dos solos, contribuindo para a sua degradação e consequente perda de produtividade agrícola. Vários estudos apontam para uma perda generalizada da biodiversidade do solo com a intensificação da agricultura, o que pode ter consequências a nível da sua fertilidade uma vez que são os microrganismos os responsáveis pela reciclagem de nutrientes do solo. Os ciclos biogeoquímicos controlam a disponibilidade de nutrientes que são absorvidos pelas plantas e dependem da atividade de comunidades especificas de microrganismos do solo que transformam as várias formas de nutrientes, determinando a fertilidade dos solos e a produção agrícola.

Neste projeto procurou avaliar-se o impacto da agricultura no microbioma do solo envolvido nos ciclos do carbono e do fósforo em solos tropicais. Os usos do solo avaliados foram floresta tropical nativa, campo de cultivo de amendoim e cajual. As amostras foram recolhidas na região de Quinhámel, Guiné-Bissau, foram recolhidas 5 amostras de solo da floresta tropical nativa, 5 do cajual e 5 do amendoinzal, durante a época húmida, mais concretamente em Outubro de 2022. .

Na Guiné-Bissau, tanto o amendoim quanto o caju são culturas de elevado impacto económico. O amendoim, *Arachis hypogaea* L., é cultivado em toda a área tropical e intertropical. Cerca de 60% da produção de amendoim de África vem da África Ocidental. Para países como a Guiné-Bissau, o caju, *Anacardium occidentale* L., representa uma porção bastante significativa das receitas (94.8%) da exportação do país, sendo este um dos maiores produtores mundiais

Os processos biogeoquímicos que regulam o ciclo do carbono (C) e do fósforo (P) são ainda pouco conhecidos, principalmente devido à falta de indicadores microbianos adequados. Estudos recentes identificaram vários genes marcadores que constituem bons indicadores dos processos biogeoquímicos associados com o ciclo do C e do P. Estes genes codificam proteínas que catalisam as transformações entre as várias formas de C e P no solo e são chamados de "functional genes". Vários destes genes têm sido usados como marcadores moleculares para quantificar a abundância das comunidades microbianas envolvidas nas várias etapas dos ciclos biogeoquímicos, por sua vez constituindo indicadores das principais transformações que estão a ocorrer nesse ecossistema. Os genes funcionais avaliados foram, para o ciclo do carbono, *pmoA, GH6 cellulase, mcrA, pcaH* e *β-glus,* e, para o ciclo do fósforo, *phoD.* Estes genes codificam, respetivamente, as proteínas particlate methane monooxygenase (pMMO), celulase da família GH6, methyl coenzyme reductase (MCR), protocatechuate 3,4-dioxygenase (3,4-PCD), β-glucosidase e fosfatase alcalina (ALP). Assim, este projeto teve como objetivo avaliar como é

que a abundância dos microrganismos pertencentes a cada grupo funcional relacionado com o ciclo do C e do P se alterou quando um ecossistema natural, ou seja, a floresta tropical nativa, é convertida para uso agrícola, nomeadamente, para a produção de amendoim e caju.

Como principais técnicas, foram usadas técnicas de biologia molecular tais como, extração de DNA a partir de solo, eletroforese em gel de agarose, PCR, clonagem, sequenciação de DNA, PCR quantitativo em tempo-real (qPCR), determinação de atividades enzimáticas, entre outras. A abundância de genes taxonómicos e funcionais, bem como as atividades enzimáticas, foram quantificadas e comparadas nos três usos do solo. As atividades da fosfatase e da β-glucosidase foram quantificadas nas amostras de solo usando para-nitrofenol (p-NP) como substrato. A abundância de genes funcionais *pmoA, GH6 cellulase, mcrA* e *phoD*, bem como dos genes taxonómicos 16S rRNA *bacteriano* e ITS *fúngico*, foi quantificada por qPCR usando DNA extraído do solo. Os genes 16S rRNA *bacteriano* e ITS *fúngico* são genes taxonómicos, fornecendo informações sobre a quantidade total de bactérias e fungos no microbioma do solo. O DNA foi extraído das amostras de solo, os fragmentos de genes-alvo foram amplificados por PCR usando primers específicos selecionados a partir da literatura existente, purificados e clonados num plasmídeo. Os fragmentos dos genes alvo foram sequenciados para confirmar a sua identidade e o DNA plasmídico foi linearizado e usado para construir as retas de calibração (diluições seriadas) para quantificação de cada um dos genes-alvo por qPCR.

Relativamente aos genes taxonómicos, 16S rRNA *bacteriano (*(3.80×10^{12} a 4.42×10^{12} gene copies / g of soil) e ITS *fúngico (*3.77×10^{7} a 6.92×10^{7} gene copies / g of soil)*,* os resultados mostram não haver diferenças significativas no que diz respeito à sua abundância entre os diferentes usos do solo, o que indica que o total da população de bactérias e fungos no microbioma do solo não se alterou significativamente quando a floresta tropical nativa foi transformada em campos de amendoim e cajual. No que respeita ao ciclo do C, a atividade da β-glucosidase (0.77 µmol a 1.51 µmol pNP / g of soil / h), indicadora da taxa de degradação de celulose, não mostrou diferenças entre os três usos do solo. De igual modo, não foram também encontradas diferenças significativas na abundância do gene *GH6 cellulase* (3.38×10^{6} a 4.42×10^{6} gene copies / g of soil)*,* envolvido na degradação da celulose, o que sugere que esta degradação não é significativamente afetada pela conversão da floresta em campos agrícolas para a produção de caju e amendoim. A abundância do gene *pmoA* (5.10×10^{5} a 1.03×10^{6} gene copies / g of soil), biomarcador da comunidade metanotrófica, variou significativamente entre os microbiomas do solo da floresta nativa e do campo de amendoim, e do cajual. Observou-se maior abundância do gene *pmoA* no microbioma do solo do cajual. Este resultado parece sugerir que neste uso do solo, existe uma maior presença de bactérias metanotróficas do que na floresta nativa. O mesmo não se verificou para os organismos metanogénicos, uma vez que a abundância de *mcrA* (6.62×10^{6} a 8.82×10^{7} gene copies / g of soil) não variou significativamente entre os vários usos do solo estudados. Os resultados parecem então sugerir que o cajual se comporta como um sink de metano.

Relativamente ao ciclo do P, a atividade da fosfatase (4.14 a 6.12 µmol pNP / g of soil / h), indicativa da mineralização do P, foi afetada pelo uso do solo, diminuindo significativamente após conversão da floresta nativa em campo de produção de amendoim, o que pode indicar uma maior mineralização de P orgânico na floresta do que nas culturas agrícolas. No entanto, o mesmo não se observou com o gene *phoD* (2.24×10^{6} a 2.99×10^{6} gene copies / g of soil), bioindicador da mineralização do fósforo, o que poderá indicar que embora a atividade da enzima tenha variado significativamente, a quantidade de microrganismos com este gene não.

Resumidamente, o presente trabalho mostra que tanto o ciclo do carbono quanto o ciclo do fósforo, e os microorganismos neles envolvidos, foram significativamente afetados pelas mudanças que a floresta tropical nativa sofreu quando foi convertida nos terrenos agrícolas (campo de amendoim e caju), uma vez que tanto a abundância do gene *pmoA* quanto a atividade da fosfatase mostraram diferenças significativas entre os usos do solo.

Palavras-chave: qPCR, Solos tropicais, Carbono, Fósforo, Uso do solo.

Index

List of figures

List of tables

1. Introduction

1.1. Soil Microbiome

In 2019, Banwart *et al* defined soil functions as "flows and transformations of mass, energy, and genetic information that connect soil to the wider critical zone, transmitting the impacts of human activity at the land surface and providing a control point for beneficial human intervention". These functions include carbon transformation, nutrient cycling, formation of soil structure and biodiversity regulation (biocontrol) (FAO, 2020).

Soil biodiversity encompasses the variety of life below ground, from genes and species to their communities and ecological complexes. Soils contain more than a quarter of global biodiversity (FAO, 2020). Soil organisms vary a great deal in size, from microscopic unicellular organisms to large vertebrates, and are classified into microbes, microfauna, mesofauna, macrofauna and megafauna (Fig.1.1). Soil microorganisms are the bacteria, archaea, fungi, protozoa, viruses and algae, that live in the soil (Gilyarov, 1949; Swift *et al*., 1979). Bacteria and fungi make up hundreds to thousands of times more biomass than archaea, protozoa and viruses (Fierer, 2017). The soil microbiome varies greatly, spatially and in composition because microorganisms need to adapt to different factors that might change the composition of the community such as different substrates, physical and chemical conditions and biotic interactions, such as those with the plant community. This creates dynamic micro-ecosystems that evolve within larger macro-ecosystems, like host organisms (Berg *et al*., 2020).

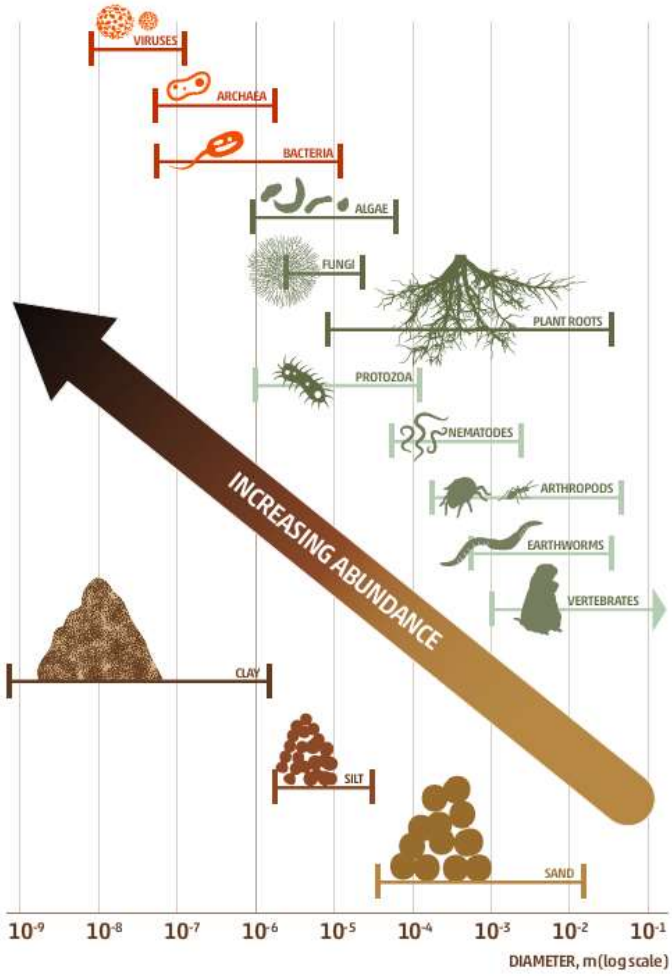


Fig.1.1 Conceptual model illustrating soil biodiversity and its relationship with spatial scales (FAO,2020).

Soil organisms are categorized into functional groups based on genetic characteristics rather than specific soil functions, because direct associations are difficult to identify (Kendzior *et al*., 2022). Functional diversity is a measure of the value and range of functional traits that exist in an ecosystem (Petchey and Gaston, 2006), which is to say that it is a measure of how many functions an ecosystem is able to fulfill. A multifunctional ecosystem is a system where its organisms perform at least two functions simultaneously (Delgado-Baquerizo *et al*., 2016). A related concept is functional redundancy (or the species redundancy hypothesis), which refers to the presence of different species able to perform the same function in the same ecosystem (Goswami *et al*., 2017). A microbiome that is functionally diverse and redundant, like that of the soil, is resilient, since it allows ecosystems to withstand species loss. So, an ecosystem with high functional redundancy is more resilient to shock, because even if some species are lost, there are others that can continue to perform their common task (Ferris and Tuomisto, 2015; Goswami *et al*., 2017). Therefore, microbial diversity seems to correlate positively with soil multifunctionality (Delgado-Baquerizo *et al*., 2016).

Approximately 30% of the Earth's land surface is covered by forest ecosystems, with very diverse soil communities. Changes in land use and agricultural management pose threats to soil biodiversity and functions. These changes include deforestation, urbanization and intensification of agricultural practices, which can lead to reductions in abundance, richness, and diversity of soil fauna and microbial diversity (FAO, 2020; Franco *et al*., 2019) especially in tropical regions where deforestation is predicted to increase (Laurance *et al*., 2014). In the tropical regions, deforestation usually involves removing plants by logging high-value wood trees and clearing low-value trees through slash and burn. This is done mostly to establish livestock farms and intensive agriculture (Garcia-Montiel *et al*., 2000). Even low-impact logging can diminish soil biodiversity and hinder ecosystem processes (de Carvalho *et al*., 2016; França *et al*., 2017).

Deforestation often causes alterations to soil physical properties, affecting resource availability and habitat suitability for many soil organisms (Neill *et al*., 1997; Garcia-Montiel *et al*., 2000). Consequently, deforestation can significantly reshape soil communities (Crowther *et a*l., 2014), often leading to a decline in specialist species and increase of generalist taxa and, therefore, a decrease in functional diversity and in the biodiversity of the microbiome (Mueller *et al*., 2016). A correlation between reduced soil biodiversity and a decline in plant species diversity has been reported, highlighting the critical role of soil biological communities in shaping plant composition. The loss of soil biodiversity due to deforestation impacts key ecosystem functions. Nutrient cycling is disrupted, resulting in reduced sequestration of carbon (C), recycling of nitrogen (N), and decomposition of organic matter. Additionally, it elevates emissions of nitrous oxide (N_2O), a potent ozone-depleting gas, and increases phosphorus (P) leaching, which aggravates even further the lack of bioavailable phosphorus in the soil. All of this significantly impairs ecosystem multifunctionality (Ravishankara *et al*., 2009; Wagg *et al.,* 2014; Revell *et al*., 2015).

When talking about microbial biomass, specifically both bacterial and fungal biomass have been reported to decrease in response to land use change (Crowther *et al*., 2014). Overall, the distinct land uses have distinct microbial community structures, both in terms of taxonomy and functionality (Mendes *et al*. 2015). It has been shown that after forest conversion microbial communities are dominated by bacteria, particularly by bacteria with a rapid turnover. Furthermore, plant-symbiont fungi were shown to decrease while soil-borne plant-pathogen fungi were shown to increase (Qu et al., 2024). Forests seem to have higher soil microbial biomass than agricultural lands, which could be related to greater plant diversity and reduced soil disturbance. Natural ecosystems exhibiting higher plant diversity and minimal soil disturbance tend to support greater soil microbial biomass (Raiesi and Beheshti, 2015; Singh *et al*., 2010, 2020) and functional redundancy, which are crucial for ecological resilience and adaptive functional shifts (Kendzior *et al*., 2022).

1.2. The uniqueness of tropical soils: the example of Guinea-Bissau soils

There are notable differences between tropical and temperate soils. Tropical soils, naturally acidic, seem to have higher clay content and bulk density, lower cation exchange capacity and water content when compared to temperate soils. Tropical soils are fragile and prone to degradation, especially when vegetation is lost, typically after deforestation. These soils have high levels of iron and aluminum oxides (giving them a characteristic red color) and low amounts of phosphorus. Many tropical soils have suffered extensive geological reworking, surface erosion and deposition, resulting in intense weathering (Minasny and Hatermink, 2011). In intensive agriculture, post-deforestation liming is often used to neutralize soil acidity, which can disrupt native microbial communities, limiting their growth, and lead to the depletion of endemic soil microbial decomposers and homogenization of soil organisms, impacting carbon sequestration, nutrient cycling, and ecosystem resilience (Fierer and Jackson, 2006; Rodrigues *et al*., 2013; de Carvalho *et al*., 2016).

In Guinea-Bissau, West Africa, the climate is tropical, hot and humid with monsoon season from June to November. The dry season is from December to May. The terrain is made up of low coastal plains with mangrove swamps, in the west of the country, that rise to low savannahs, in the east. About 71% of the soil is forest, while 29% is for permanent crops, arable soil and other uses. The soils are predominantly ferrasols, acrisols, fluvisols, leptosols and lixisols. The main problems faced are dust deposition by winds, bush fires, deforestation, soil erosion, overgrazing, and drought (FAO and ICRISAT, 2019).

Guinea-Bissau is divided into the following three agroecological zones:

i) Northern zone, with maritime climate, has lost a large amount of vegetation cover. The vegetation is made up of mangrove forests on the Atlantic lowlands flooded by tides, and savannah woodlands on the uplands. The main farming systems are freshwater swamp rice, upland cereals and peanuts (on ferralsols in the savanna woodlands), and mangrove swamp rice (on fluvisols in the mangroves).

ii) Eastern zone where vegetation is made up of savannah woodlands in the uplands (where shifting cultivation of sorghum, millet, maize, peanuts and sometimes rice is practiced) and of dry and wet grasslands in the lowlands. Permanent cultivation of rice is practiced in the freshwater swamps, where the land is very fertile. The same happens with sorghum, millet, maize and peanuts. Apart from this, the rest of the soil in this zone are, for the most part, poor.

iii) Southern zone, where there is the highest mean annual rainfall, occur the most fertile ferralsols and fluvisols, and the densest dry forest patches and thick mangrove forests. The major food crops here are rice, cashew nuts, maize, sorghum, fonio, cola, banana, citrus and palms (FAO and ICRISAT, 2019).

The savannah is a mixed grassland/ woodland ecosystem that is adjacent to the forest region. The vegetation ranges from sparse grassland to open-wooded grassland. The soils are generally well drained and possess a thin layer of organic matter. They can support limited cultivation but are prone to impoverishment. The soils in this region are characterized by moderate leaching. Most soils are old and deep, with a low nutrient-retention capacity. The wetlands in Africa are found along the coast in estuaries and deltas, as well as in inland areas. Vegetation varies from submerged forests to salty grasslands and mangrove forests. The soils on the floodplains are high in nutrients and organic matter, as well as in clay and silt (Jones *et al.*, 2013).

In Guinea-Bissau, both peanut and cashew are prevalent cultures with high economic impact. Peanut, *Arachis hypogaea,* is a leguminous plant, of high nutritional value. It is cultivated throughout the tropical and inter-tropical area (Didagbe, 2015). Around 60% of Africa's peanut production comes from West Africa (Debrah and Waliyar, 1996). The cashew tree, *Anacardium occidentale*, is an evergreen tree. Cashew has gained significant economic importance as a cash crop in many countries, where there is a high market value for the nut (Catarino *et al.,* 2015). The top five agricultural products exported by Africa include cashew nuts (94.8%) (FAO and ICRISAT, 2019). This represents a significant revenue as

a major export product for countries like Guinea-Bissau, which is one of the world's largest producers of raw cashew nuts (Catarino *et al.,* 2015). The Biombo region, where the samples for this project where collected, is situated in the west coast of Guinea-Bissau.

1.3. The Terrestrial Carbon Cycle

While plants acquire their C from the atmosphere through photosynthesis, they rely on a spectrum of macro and micronutrients obtained from the soil to generate biomass and facilitate nutrient and energy transfer. Soil microorganisms play a rather important role in physically breaking down plant residues. This breakdown allows soil microorganisms to release the nutrients and the energy within the plant material (FAO, 2020). Carbon is fixed into the soil through the transformation of plant and animal detritus, and through the activity of bacteria and archaea that can fix carbon by using atmospheric CO_2 as their energy source. Carbon is either fixed or released from the soil, depending on the activity of soil organisms, and driven by soil conditions (FAO, 2020). The combined carbon content of soil bacteria rivals that of all plants and their N and phosphorus contents are a lot higher than that of plants (FAO, 2020). Soil microorganisms contribute to the stabilization of soil organic carbon by promoting the formation of soil microaggregates, stabilized by proteins like glomalin (produced by arbuscular mycorrhizal fungi), in which soil organic matter is protected from physical decomposition (Gougoulias *et al.*, 2014). In the soil, organic matter is composed of decomposed litter and dead biomass residues (necromass) (Trivedi *et al.*, 2013). The largest flows of carbon in the soil are decomposition and carbon storage, resulting from broader aggregate processes like soil respiration. Most soil microorganisms are aerobic heterotrophs (Schimel and Schaeffer, 2012).

The carbon cycle is not independent, it is coupled with other elements that are essential for life, like nitrogen (N) and phosphorus (P), that thus control the rate of microbial consumption and carbon respiration (Gougoulias *et al.*, 2014). The carbon cycle (Fig.1.2 and 1.3) is defined by the balance between photosynthesis and respiration: autotrophic organisms such as plants and photo- and chemoautotrophic microorganisms in the soil fix carbon from the atmosphere and turn it into organic carbon compounds, while autotrophs and heterotrophs send carbon back to the atmosphere, through respiration, in the form of CO_2. Heterotrophic microorganisms decompose dead organic matter from plants, animals or other microorganisms using carbon compounds as substrate for their metabolic needs. In this process, some of the carbon is retained in the microbial biomass and some is sent back to the soil as metabolites or back to the atmosphere as CO_2 (Gougoulias *et al.*, 2014). In particular, methane and CO_2 are produced by methanogens, in anoxic conditions, through methanogenesis (*mcrA* gene) in which methyl-coenzyme M and coenzyme B are converted to methane (Kou *et al.*, 2017; Tiwari *et al.*, 2018) which is then decomposed by methanotrophic microorganisms, through methane oxidation (*pmoA* gene), resulting in the production of methanol and then, depending on the type of methanotroph, transformed into other compounds (Grabarse *et al.*, 2001).

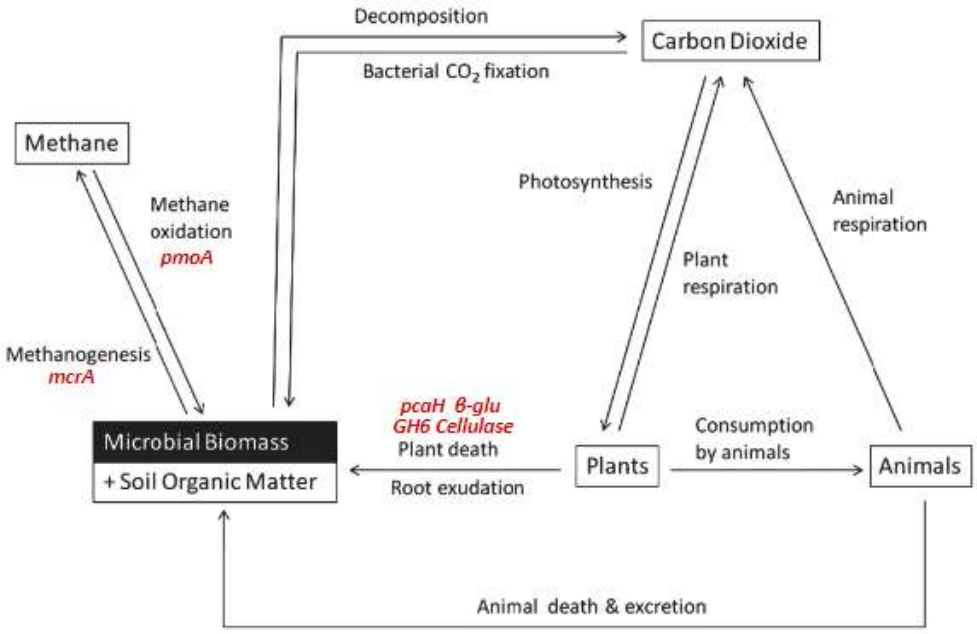


Figure 1.2 Diagram of the terrestrial carbon cycle adapted (Gougoulias *et al*., 2014). In red are the carbon cycle genes under analysis in this project (*pmoA, mcrA, pcaH, β-glu* and *GH6 cellulase*).

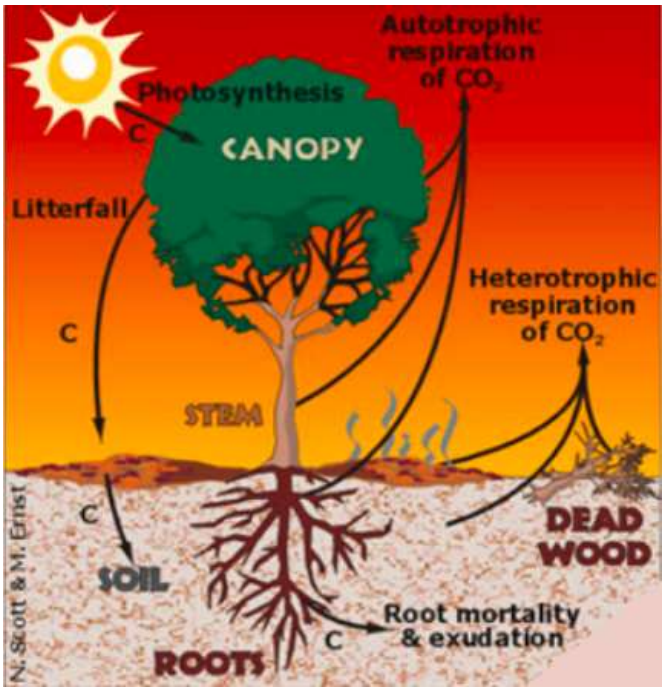


Figure 1.3 Diagram of the terrestrial carbon cycle (Luo *et al*., 2015).

Agricultural practices can cause unpredictable changes in the soil microbiome, which then in turn can cause shifts in greenhouse gases emissions as well as in soil carbon storage. This unpredictability is due to limited knowledge surrounding the soil microbiome but also due to the complexity of the interactions between the soil microbiome and biotic and abiotic factors (Kendzior *et al*., 2022). Soil holds between 55-70% of the total African terrestrial carbon pool. In tropical climates, the absence of winter or frosts to kill insects or microorganisms, combined with constant heat and humidity, accelerates decomposition. This rapid decomposition generally results in low organic carbon

levels in tropical soils (Jones *et al.*, 2013). African soils display a wide range of soil organic carbon values and distribution: deserts like the Sahara, Kalahari, and Horn of Africa have very low soil organic carbon (0 to 40 tonnes ha per 1m depth), while wetlands in countries such as Sudan, DR Congo, and Tanzania have high levels (60 to more than 150 tonnes ha per 1m depth) (Sombroek *et al.,* 1993; Williams *et al.,* 2007; Jones *et al*.,2013). Converting tropical forests to farmland reduces topsoil carbon by 20-50% because of the presence of less detritus, increased erosion (because of disturbance of soil structure and aggregates), and higher decomposition rates (due to fertilization or planting of N-fixing species) (Guo and Gifford, 2002; Murty *et al.,* 2002). Most soil organic carbon is concentrated in the topsoil, particularly vulnerable to human activities and climate change, leading to carbon loss and soil degradation.

Currently, Africa emits slightly more carbon than it sequesters, mainly due to heterotrophic respiration and organic matter decomposition, which are mostly balanced by plant photosynthesis (Jones *et al.,* 2013). Carbon emissions from land clearance and biomass burning in Africa are higher than those from fossil fuel combustion. While savannah fires contribute to emissions, their impact is mitigated by subsequent regrowth. Current data indicates a small carbon imbalance in Africa, but climatic fluctuations, particularly droughts, can significantly alter net ecosystem productivity and increase wildfire frequency, making Africa a major source of global carbon emissions (Jones *et al.,* 2013).
Due to pressures like climate variability, land use changes, population growth, industrialization, and urbanization, carbon emissions in Africa are expected to rise substantially in the coming years (Williams *et al.,* 2007). Africa contributes over 20% of global fossil fuel emissions from land use and nearly 40% of global biomass burning emissions, with its vegetation and soils accounting for about 13% and 10-12%, respectively, of the global total (IPCC, 2007; Ciais *et al*., 2011). Only in tropical forests is there near parity between carbon stored in vegetation and soil, because of continuously growing vegetation and fast recycling of dead plant matter (Jones *et al.,* 2013).

1.4. The Terrestrial Phosphorus Cycle

Phosphorus (P) is essential to life, being a necessary component of nucleic acids and phospholipids, among other macromolecules. It is essential to cellular energy processes, growth and development. It is involved in many critical biological processes, such as energy metabolism, synthesis of nucleic acids and membranes, and photosynthesis. P exists in the soil in either inorganic (Pi) or organic (Po) forms. Organic P is biologically assimilable while inorganic P is geochemically bound. Plants can only utilize organic P after its hydrolysis into orthophosphates (Hui *et al.,* 2013). Inorganic P precipitates very fast, consequently very little of it is present in the soil in its soluble form (Spohn and Kuzyakov, 2013). For soil organisms to use them, organic P compounds must suffer enzymatic hydrolysis into orthophosphate ($H_2PO_4^{2-}$ or H_2PO_4) (Fraser *et al.,* 2015). These orthophosphates are then rapidly absorbed by soil particles or immobilized by phosphorus-consuming bacteria (Fig.1.4) (Jones *et al*., 2013).

The biggest source of soil phosphorus is the weathering of continental bedrock, particularly apatite ($Ca_{10}(PO_4)_6(OH, F, Cl)_2$, the most abundant phosphorus rich mineral in continental bedrock. In the terrestrial phosphorus cycle, phosphorus is gathered in three main pools: bedrock, soil and living organisms. The weathering of bedrock happens mostly through the dissolution of phosphorus rich minerals, like apatite, into phosphate, by acids produced by microorganisms. Phosphate, in the soil solution (PO_4^3), is then available for plants to use, and is later returned to the soil by decay of litterfall and root turnover. Because of phosphate absorption by ferric iron and aluminum oxyhydroxides, soil phosphate content is maintained at a low concentration in acidic soils rich in iron and aluminum, such as those from Africa. To bypass this low concentration, plants have physiological strategies that allow them to obtain phosphorus, among them the production, by the roots or associated fungi, of chelating agents that solubilize the phosphorus absorbed by ferric iron and calcium substrates, and the production

of enzymes or acids able to solubilize phosphate near the roots. Absorption is considered the most important process controlling terrestrial phosphorus bioavailability. To minimize the loss of phosphorus, plants reabsorb most of their phosphorus before litterfall and recycle the fallen litter, this is particularly efficient in tropical soils, characteristically infertile. Here, there is almost no phosphorus in the topsoil, instead, it is in the biomass (Ruttenberg, 2001). In little or only moderately weathered soils with neutral to alkaline pH, the primary source of inorganic P is calcium phosphates. In more weathered and acidic soils, as is the case of the soils of Guinea-Bissau, the source is iron and aluminum hydroxides (Spohn and Kuzyakov, 2013).

In natural ecosystems, inorganic P may originate from the mineralization of plant litter, algae, soil organic matter and sediments (Hui *et al.,* 2013). Plants and microorganisms have developed several ways to mineralize organic P and to solubilize bound inorganic P. For instance, they can release protons and organic acids such as oxalate and citrate that solubilize bound inorganic P (Spohn and Kuzyakov, 2013). P is considered the most inaccessible and unavailable of all soil nutrients (Bergkemper *et al*., 2016). P also interacts with other essential elements like carbon (C) and nitrogen (N), regulating biological processes. The C:N:P ratio is considered an important indicator for nutrient fluxes in global circulation models (Hui *et al.,* 2013).

Phosphorus deficiency is common in areas of high rainfall, on acid, clay-rich or poor calcareous soils. Due to its high reactivity, inorganic phosphorus is rarely found as a free element, and geologically occurs as phosphate rocks (PO_4^{3-}). The tropical soils of Africa, highly weathered and iron-rich, tend to be deficient in plant-available phosphorus. The low pH along with high levels of iron and aluminum oxides, tend to immobilize phosphorus in soil particles, making it unavailable for plants. When this happens, large amounts of phosphorus fertilizer are added to the soil to increase crop yields. It is estimated that around half a million tons of phosphorus is lost every year from cultivated soil in Africa which is double Africa's annual phosphorous consumption. Given the phosphorus' essential role in living organisms, its slow natural cycle and the low solubility of natural phosphorus-containing compounds, the agricultural industry is heavily reliant on fertilizers containing concentrated phosphoric acids (H_3PO_4) (Jones *et al.*, 2013).

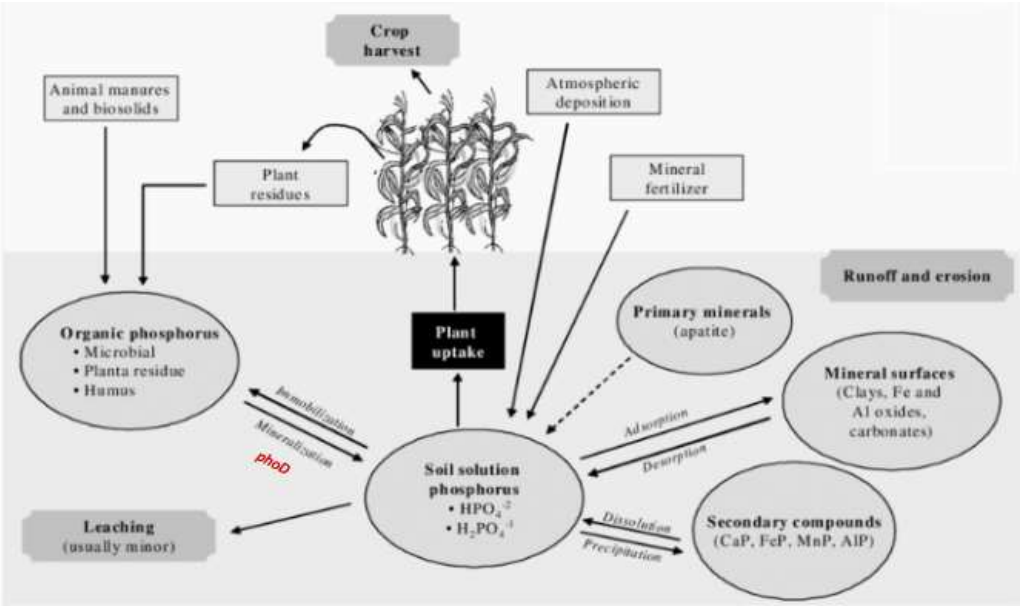


Figure 1.4 Diagram of the terrestrial Phosphorous cycle (Vendramini *et al.*, 2007). The *phoD* gene, encoding (extracellular) alkaline phosphatase (ALP), was chosen in this study to evaluate the impact of land use on the phosphorous cycle in tropical soils.

1.5. Functional Genes and Enzymatic Activities

To determine the abundance of specific functional groups of soil microbes, gene markers such as functional genes have been used. These are good indicators of their associated biogeochemical processes in the nutrient cycles and therefore of changes in the workings of these cycles when subjected to different factors, such as land use changes. Functional genes are genes that encode core metabolic proteins associated with biogeochemical processes, like the C and P nutrient cycles. Quantification of functional genes in soils allows the quantification of the potential of a soil to perform an associated process, a function and an approximate quantification of microbes that belong to that functional group. These genes reflect the diversity of microbial functional groups and their responses to environmental factors, and diversity of specific functions and changes in the microbial composition of communities involved in them. Functional genes can be used for this purpose of estimating microbial abundance by quantifying genes involved in key biogeochemical processes because, their phylogeny strongly correlates with *16S rRNA* gene phylogeny (Lammel *et al*., 2015; Imhoff, 2016; Geisen *et al*., 2019; Thiele-Bruhn *et al*., 2020).

To evaluate how land use impacts the C cycle in tropical soils, the functional genes *pcaH, pmoA*, *β-glu, GH6 cellulase* and *mcrA* were quantified by qPCR and the enzymatic activity of β-glucosidase was measured.

The *pcaH* gene encodes part of the β-subunit of the enzyme protocathecuate 3,4-dioxygenase involved in the β-ketoadipate pathway, which degrades aromatic compounds related to lignin, present in plants, by cleaving their rings. This pathway is a major catabolic route in terrestrial soils, especially in the catabolism of phenolic compounds. Furthermore, it is biologically conserved and widespread in the soil bacterial and fungal communities. Therefore, the structural genes that encode enzymes for this pathway are very similar across a wide range of organisms (Buchan *et al.*, 2001; El Azhari *et al.*, 2008).

The *pmoA* gene encodes a membrane-bound methane monooxygenase (particulate MMO or pMMO) that methanotrophs, microorganisms that use methane (CH_4) as their carbon and energy source, use to catalyze the oxidation of methane into methanol. Methanol is then converted to formaldehyde by methanol dehydrogenase (MDH), after which the metabolic pathways diverge depending on the type of methanotroph. This gene is commonly used to characterize methanotrophic communities in soil (Kou *et al.*, 2017; Tiwari *et al.,* 2018).

The *mcrA* gene encodes the alpha subunit of the methyl-coenzyme M reductase (MCR). This enzyme is only found in methanogens, except for the methane-oxidizing archaea (Steinberg and Regan, 2009). Methyl-coenzyme M reductase (MCR) catalyzes the final reaction of the pathway in which methyl-coenzyme M and coenzyme B are converted to methane and the heterodisulfide CoM-S-S-CoB, which conserves energy (Grabarse *et al.,* 2001). Because they catalyze the production of methane and carbon dioxide, methanogens are integral to carbon cycling, in anaerobic soils and sediments, during organic matter degradation (Steinberg and Regan, 2009).

Soil biology has focused on evaluating microorganisms' degrading capacity through the enzymatic activity of the enzymes they produce and release into the soil. Soil enzyme activities are sensitive, practical, easily measured, integrative and responsive to soil tillage and structure, which makes them good biological indicators of soil quality and soil ecosystem sustainability. Soil extracellular enzymes catalyze biochemical reactions involved in organic matter decomposition, nutrient mineralization, plant growth and nutrient cycling, by breaking down complex organic molecules into absorbable forms. Because of this, decomposition rates are closely related to the activity of soil enzymes. These enzymes also catalyze biochemical reactions involved in soil aggregation (Geisen *et al.,* 2019). The most used soil enzymes as biological indicators are the ones involved in the degradation of litter and hydrolases, taking part in carbon (C), nitrogen (N), phosphorus (P), and sulfur (S) cycles, such as β-glucosidase, β-galactosidase, urease, phosphatase, and arylsulphatase. Factors that influence

enzymatic activities include soil depth, type, temperature, moisture, pH, substrate availability, and management practices (Adetunji *et al.,* 2017).

Cellulose is the largest biomass source on the planet. The *β-glu* gene encodes a β-glucosidase (Cairns and Esen, 2010). β-Glucosidase is the rate-limiting enzyme of the degradation of cellulose (Zang *et al.,* 2018). Cellulose is the primary structural component of plant cell walls, and, consequently of plant debris. Cellulases are enzymes that break down cellulose, a polysaccharide composed of β-1,4 linked glucose units (Fig.1.5). This group includes endo-1,4-glucanase, exo-1,4-glucanase, and β-glucosidases (Deng and Tabatabai, 1994). These three enzymes compose the cellulase systems, present in all cellulose-degrading microorganisms, and act together to completely hydrolyze cellulose. In the final phase of cellulose degradation, β-glucosidase hydrolyzes the cellobiose residue, producing glucose (Fig.1.5), a crucial carbon source for soil microorganisms. It is this role in carbon cycling that has made β-glucosidase so commonly used for soil quality assessment. Cellulases belong to the glycoside hydrolase (GH) superfamily, which includes at least ten GH families (GH 1, 3, 5, 6, 8, 9, 12, 44, 45, and 48) (Cairns and Esen, 2010; Merlin *et al.,* 2014; Adetunji. *et al.,* 2017; Chen *et al.,* 2022). Cellulases of the GH6 family, encoded by the GH6 cellulase gene, function as endocellulases and/or exocellulases and play a key role in the early stage of cellulose degradation (Merlin *et al*., 2014). GH6 cellulases degrade cellulose into the smaller cellobiose residue that β-glucosidases further break into glucose. At the early stage, cellulose degradation catalyzed by endocellulases and/or exocellulase is essential for the stabilization of lignocellulosic wastes (Chen *et al.*, 2022).

β-glucosidase	Cellobiose hydrolysis	Glucoside + H_2O → ROH + glucose	C-cycling
Cellulase	Cellulose hydrolysis	Hydrolysis of β-1, 4 - glucan bonds	C-cycling

Figure 1.5 β-glucosidase and cellulase activity adapted from Adetunji *et al*. (2017).

To evaluate how the conversion of tropical forests into agricultural fields impacts the P-cycle, the abundance of the *phoD* gene and the enzymatic activity of phosphatase were quantified in the soils of the native tropical forest, peanut and cashew fields. The *phoD* gene, along with homologous *phoA* and *phoX* genes, is part of the pho regulon, related to phosphate starvation. In bacteria, these three genes encode an (extracellular) alkaline phosphatase (ALP), a phosphomonoesterase, that bacteria excrete and that catalyzes the hydrolysis of the ester-phosphate of orthophosphate monoesters like sugar phosphates, phytates and nucleotides (P compounds) meaning that it converts organic P into phosphate (Fig.1.6). This contributes to the mineralization of P compounds making P bioavailable in the soil. High levels of mineral P can inhibit pho regulon transcription and suppress the functioning of ALP (Fraser *et al.*, 2015). The activity of extracellular phosphatase in soil seems to be negatively correlated with the availability of inorganic P (Spohn and Kuzyakov, 2013).

Phosphatase	Release of PO_4^- phosphate	Phosphate ester + H_2O → ROH	P-cycling

Figure 1.6 Phosphatase activity, adapted from Adetunji *et al*. (2017).

Phosphatases are enzymes that catalyze the hydrolysis of phosphoric acid into phosphate. These enzymes are secreted by soil microorganisms. The amount of phosphatase in the soil depends on the amount of microorganisms, the amount of organic materials, mineral and organic fertilizers, and on tillage and other agricultural practices. In the soil, the main sources of phosphatases are plants and microorganisms. Because plants only use inorganic P and a large amount of soil P is organically bound,

the mineralization of organic P by microorganisms is crucial for plant nutrition. When phosphorus is lacking in the soil, plant roots and microorganisms secrete more phosphatase to intensify the solubilization and remobilization of phosphate, influencing the ability of plants to cope with phosphorus limitation. This shows that plants and microorganisms' demand for phosphorus can be linked to the production and activity of soil phosphatases. Because of this, phosphatase activity can be used as an indicator of inorganic phosphorus availability for plants and microorganisms. Phosphatases include a group of enzymes, phosphomonoesterases, that, as mentioned above, hydrolyze the ester-phosphate bonds in soil organic P. This releases phosphate into the soil which can then be used by plants. Phosphomonoesterase is the most studied among the soil phosphatases. It is active in both acidic and alkaline conditions, depending on its optimal pH, and acts on low molecular P compounds with monoester bonds, including nucleotides, sugar phosphates and polyphosphates. So, acid phosphatase is found mainly in acid soils, while alkaline phosphatase is found mainly in alkaline soils (Adetunji *et al*., 2017).

Enzymes are substrate-specific and cannot represent the full range of soil microbiological activity. Soil enzymes catalyze one specific biological reaction, so their activity cannot be equated to the whole microbial activity in the soil, which is made up of a plethora of enzymatic reactions. Quantification of enzyme activity is not able to separate active and inactive microorganisms. It only measures process rates and has limited taxonomic resolution (Geisen *et al.,* 2019).

1.6. Project Objectives and Techniques

There is little knowledge on the impacts of threats like deforestation and land conversion (specifically forest to agricultural land) on soil biodiversity and function. Most studies about the conversion of tropical forests are about their conversion into pastures. There are not as many studies about the conversion of tropical forests into agricultural land. In the particular case of the tropical regions, where Guinea-Bissau is located, in comparison to the temperate regions, this lack of knowledge is intensified by the lack of research efforts and by decreasing research capabilities concerning tropical soils. Furthermore, most of the studies that do exist focusing on tropical regions, pertain mostly to the tropical ecosystems of South America, like those in the Amazon rainforest, while very few studies focus on African tropical ecosystems. Even though it has been proven that changes to the natural ecosystems are major threats to soil functions and biodiversity, global biodiversity assessments and conservation acts have consistently been devoid of information on soil biodiversity. Which is to say these assessments have a low taxonomic resolution, omitting microorganisms, and are very limited in the geographic area they cover, focusing mostly on the temperate regions. All of this results in a very poor understanding of global biodiversity patterns, and therefore, the impact of specific threats, and also of specific measures for the management of the consequences of those threats.

This project aims to understand the impact of the conversion of Guinea-Bissau's native tropical forest into agricultural land, specifically peanut and cashew cultures, on the soil microbiome involved in the carbon and phosphorus cycles. To fulfill this objective (Table 1.1), the abundance of target genes and enzymes, involved in the C and P cycles, was measured in 3 different land uses: native tropical forest, cashew plantation and peanut field. The abundance of functional genes involved in the C and P cycles was quantified by qPCR amplification of DNA extracted from various samples from Guinea-Bissau. Also, the abundance of the *bacterial* 16S rRNA *and fungal* ITS was determined to allow the quantification of the total amount of bacteria and fungi in the soil. The abundance of taxonomic, C- and P-functional genes in the soil from the 3 land uses was quantified by qPCR using specific primers and SYBR-Green chemistry. First, DNA was extracted from the soil samples, and the target gene fragments were isolated, and PCR amplified, purified and cloned. The fragments were sequenced to confirm the nucleotide sequences. The plasmid DNA was linearized by enzymatic digestion and used for establishing the standard curves for the qPCR of each functional or taxonomic gene. DNA from each soil sample

was qPCR amplified and the gene copy number in each soil type was calculated using the standard curve. Phosphatase and glucosidase activity were measured in soil samples from the native tropical forest and the peanut and cashew cultures.

Table 1.1 Parameters evaluated in this project

<table>
<tr><td rowspan="5">Land Use</td><td rowspan="5">Native Tropical Forest

Cashew Field

Peanut field</td><td rowspan="3">Gene abundance (copy number/g of soil)</td><td>Taxonomic Genes</td><td colspan="2">Bacterial 16S rRNA
Fungal ITS
Archaeal 16S rRNA</td></tr>
<tr><td rowspan="2">Functional Genes</td><td>Carbon Cycle</td><td>β-glu
GH6 cellulase
pmoA
mcrA
pcaH</td></tr>
<tr><td>Phosphorus Cycle</td><td>phoD</td></tr>
<tr><td rowspan="2">Enzymatic activity (µmol pNP/ g of soil / h)</td><td>β-glucosidase</td><td colspan="2"></td></tr>
<tr><td>Phosphatase</td><td colspan="2"></td></tr>
</table>

2. Materials and Methods

2.1. Soil Sampling

Soil samples were collected in October 2022, during the wet season in Guinea-Bissau in the Biombo region, near the city of Quinhámel. The collection was made at 3 different land uses (Fig 2.1), in the same geographical area and therefore presenting no major edaphoclimatic variations. These sites were: native tropical forest (11º53'37.7''N and 15º49'11.7'W); cashew plantation with mature trees (around 30 years old) where there was no use of fertilizers or soil tillage (11º53'59.8''N and 15º49'35.2''W); and peanut field where there was manual tillage of the soil but no use of synthetic fertilizers (11º53'36.9''and 15º49' 73.5''W). The climate of the region is tropical, humid and hot, with monsoon type rainy season, June to November, and dry season, December to May. Five plots of 10 m x 10 m were established on each site and the soil was collected at each corner and at the center. The 5 soil samples collected in each plot were combined into a pooled sample totalizing 5 representative soil samples (biological replicates) in each site. Approximately 20 g of soil were collected per sample, in depths from 0 cm to 20 cm. The soil samples were packed into sterile 15 mL Falcon tubes and immediately stored in a cooler box, and then stored at -20 ºC, on the same day. Spatula for soil collection was washed with ethanol between each soil sampling.

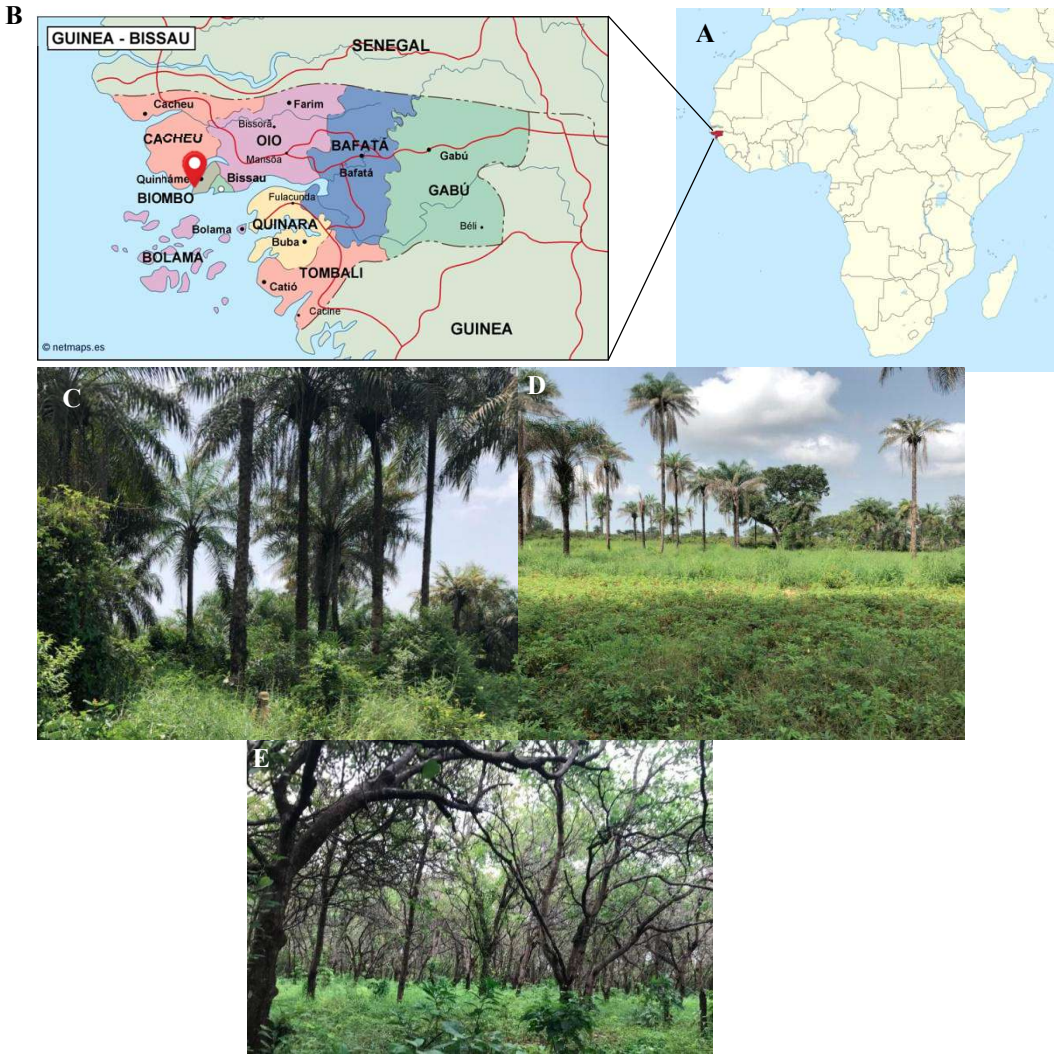


Figure 2.1 Guinea-Bissau location in the African continent (A) and Biombo region under study identified as red symbol. Land uses studied: tropical forest (C), peanut field (D), and cashew plantation (E) in Guinea-Bissau from where soil was sampled. Images: A) TUBS; B) netmaps.es; C-D) Photos by Mónica Sebastiana and Filipa Monteiro.

2.2. DNA Extraction from Soil

DNA was extracted from 5 replicated soil samples from the 3 land-uses: forest, cashew plantation and peanut field, totalizing 15 soil samples. DNA extraction (0.150 g soil) was performed using the DNeasy PowerSoil Pro Kit (Qiagen), following the manufacturer's protocol with the following modifications: washing with the C5 solution was done 3 times and the column was incubated for 5 minutes before each centrifugation. These modifications were implemented to minimize the presence of contaminants at the end of the DNA extraction. This kit was developed for isolating microbial genomic DNA from environmental samples such as soils which are rich in humic acids that can inhibit PCR reactions. First, soil samples are added to a bead-beating tube for homogenization, followed by chemical cell lysis. The genomic DNA is then collected on a silica membrane within a spin column, where it is washed from contaminants and eluted, ready for downstream applications, such as PCR, qPCR and next-generation sequencing (NGS) (DNeasy PowerSoil Pro Kit Handbook, 2021). A Nanodrop 1000 (ThermoFisher) was used to obtain soil DNA concentration values, and Abs 260/280 and Abs 260/230 ratios were analyzed to estimate DNA purity. DNA samples were analyzed by gel electrophoresis (80 V constant) using 1% agarose gels, TAE buffer and Greensafe nucleic acid stain.

2.3. Construction of qPCR Standard Curves for qPCR of Taxonomic and Functional Genes

The abundance of taxonomic genes, and genes involved in C and P cycling on the soil samples was quantified by real-time quantitative PCR (qPCR), using standard curves of gene copy numbers (absolute quantification) prepared from serial dilutions of plasmids containing the genes of interest, as described in the next sections.

2.3.1. PCR Amplification and Purification of Taxonomic and Functional Genes

Taxonomic and functional gene sequences were amplified from soil DNA samples using gene specific primers selected from the available literature (Table 2.1). Each PCR reaction contained the following components: 0.75 µL of DNA Taq Polymerase (New England Biolabs); 2.5 µL of 10x standard Taq reaction buffer; 1 µL of Forward Primer (10 µM); 1 µL of Reverse Primer (10 µM); 1 µL of dNTPs (10 µM); 1 µL of BSA (10 mg/µL) and 5 µL of soil DNA, in a final volume of 25 µL. BSA was used to enhance amplification yield in PCR and to help stabilize the reaction. Thermocycling conditions are shown in Table 2.2. First, conventional PCR programs were used for all the genes, but for the *pmoA* gene and the *GH6 cellulase* gene more than one DNA band was amplified, so a touch-down PCR program was adopted to increase PCR specificity (Table 2.2). The PCR products were analyzed on 1% agarose gels and the specific size DNA bands were cut from the gels using a scalpel and a U.V. transilluminator equipment. DNA was purified from the excised agarose gel bands with the QIAquick Gel Extraction Kit (Qiagen), according to the manufacturer's instructions. Briefly, gel bands were dissolved in a buffer with a pH indicator to ensure the optimal pH for DNA binding. The mixture was then added to a QIAquick spin column, where nucleic acids bind to the silica membrane in high-salt conditions. After washing away impurities, pure DNA was eluted using low salt buffer. This purification process effectively removes primers, nucleotides, enzymes, mineral oil, salts, agarose, ethidium bromide, and other contaminants, allowing the DNA to be used for subsequent applications (Qiagen, 2024). DNA concentration was measured using a Nanodrop equipment. The purified gene fragments were sequenced using Sanger technology (STAB VIDA) and the DNA sequences were searched (Blast N) against the NCBI database (www.ncbi.nlm.nih.gov) for confirmation of the specific gene identity.

Table 2.1 PCR primers (nucleotide sequences, mean amplicon length and annealing temperature) used for amplification of taxonomic and functional genes from soil DNA.

Target gene	**Gene Type**	**Nutrient cycle**	**Primers (5' – 3')**	**Amplicon size** (bp)	**Annealing temperature** (ºC)	**Reference**
bacterial 16S rRNA	Taxonomic	-	Eub341F CCTACGGGAGGCAGCAG Eub534R ATTACCGCGGCTGCTGGC	193	58	Muyzer *et al*. (1993)
archaeal 16S rRNA	Taxonomic	-	344f ACGGGGYGCAGCAGGCGCGA 958r YCCGGCGTTGAMTCCAATT	600	62	Lindsay *et al.* (2017)
fungal ITS	Taxonomic	-	ITS4 TCCTCCGCTTATTGATATGC ITS5 GGAAGTAAAAGTCGTAACAAGG	600	55	White *et al.* (1990)
pmoA	Functional	Carbon	A189 GGNGACTGGGACTTCTGG Mb661 GGTAARGACGTTGCNCCGG	491	52	Costello and Lidstrom (1999)
mcrA	Functional	Carbon	mlas GGTGGTGTM GGDTTCACMCARTA mcrArev CGTTCATBGCGTAGTTVGGRTAGT	500	55	Steinberg and Regan (2008)
pcaH	Functional	Carbon	PCAHf GAGRTSTGGCARGCSAAY PCAHr CCGYSSAGCACGATGTC	395	60	El Azhari*. et al*., (2008)
β-glu	Functional	Carbon	BgluF2 TTCYTBGGYRTCAACTACTA BgluR4 CCGTTYTCGGTBAYSWAGA	180	53	Cañizares *et al*. (2011)
β-glu	Functional	Carbon	GH1BF CCTACCAGATYCARGG	N/A	50	Zang *et al*. (2018)

			GH1BR GAGGAAGRTCCCARTG			
β-glu	Functional	Carbon	GH3EF GGTGGTCGCRRYTGGGA GH3ER CCAGGCATCGGWCATRTC	N/A	56	Zang *et a*l. (2018)
GH6 cellulase	Functional	Carbon	Cell2F ACCTGCCCGRCCGYGACT Cell2R GAGSGARTCSGGCTCRAT	150	62 *	Merlin (2014)
phoD	Functional	Phosphorus	phoD-FW TGTTCCACCTGGGCGAYWMIATHTAYG phoD-RW CGTTCGCGACCTCGTGRTCRTCCCA	208	60	Eder *et al*. (1996)
phoD	Functional	Phosphorus	ALPS-F730 CAGTGGGACGACCACGAGGT ALPS-1101 GAGGCCGATCGGCATGTCG	371	57	Sakurai *et al*. (2008)

*The annealing temperature described in the literature is 64 ºC. However, after performing a gradient PCR followed by an agarose electrophoresis gel, it was determined that at 62 ºC the amplification was highest.

Table 2.2 Thermal conditions for soil DNA amplification by PCR

Target gene	Primers	Thermal conditions	Reference
bacterial 16S rRNA	Eub34F Eub53R	94 ºC for 15 min 50 cycles: 94 ºC for 30 s, 58 ºC for 20 s and 72 ºC for 30 s 72 ºC for 10 min	Adapted from Smits *et al.* (2004) and Muyzer *et al.* (1993)
archaeal 16S rRNA	344f 958r	95 °C for 5 min 35 cycles: 95 °C for 1 min, 62 ºC for 1 min, 72 °C for 1.5 min. 72°C for 10 min	Lindsay *et al*. (2017)
fungal ITS	ITS4 ITS5	96 ºC 3 min 35 cycles: 94 ºC for 30 s, 55 ºC for 30 s, 72 ºC for 1 min 72 ºC for 7 min	Adapted from Khan and Bhadauria (2019)
pmoA	A189 Mb661	95 ºC for 5 min 12 cycles of: 95 ºC for 45 s; touchdown with a 1 ºC decrease per cycle, starting at 62 ºC for 45 s and ending at 50 ºC, 72 ºC for 45 s 40 cycles of: 95 ºC for 30 s, 50 ºC for 50 s, 72 ºC for 45 s. 72ºC for 15 min	Adapted from Kou *et al*. (2017) and Costello and Lidstrom (1999)
mcrA	mlas mrcA rev	95 °C for 3.5 min 40 cycles: 95 °C for 30 s, 55 °C for 45 s, 72 °C for 30 s 72 °C for 7 min.	Steinberg and Regan (2009)
pcaH	PCAHf PCAHr	94 °C for 5 min, 50 cycles: 94 °C for 1 min, 60 °C for 1 min ,72 °C for 1 min, 72 °C for 7 min	Adapted from El Azhari. *et al*. (2007)
GH6 cellulase	Cell2F Cell2R	95 ºC for 5 min 10 cycles: 95 ºC for 30 s; touchdown with a 1 ºC decrease per cycle, starting at 74 ºC for 45 s and ending at 62 ºC, 72 ºC for 45 s 40 cycles: 95 ºC for 30 s, 62 ºC for 45 s, 72 ºC for 45 s 72 ºC for 15 min	Adapted from Merlin *et al*. (2014)
β-glu	BgluF2 BgluR4	95 ºC for 5 min 50 cycles :95 ºC for 30 s, 5 3ºC for 1min, 72 ºC for 1 min 72 ºC for 10 min	Adapted from Cañizares *et al*. (2011)
	GH1BF GH1BR	95 °C for 30 s 45 cycles: 95 °C for 5 s, 50 °C for 30 s, 72 °C for 1 min 72 ºC for 10 min	Adapted from Zang *et al.* (2018)
	GH3EF GH3ER	95 °C for 30 s, 45 cycles: 95 °C for 5 s, 56 °C for 30 s, 72 °C for 1 min 72 ºC for 10 min	Adapted from Zang. *et al.* (2018)

phoD	phoD-FW phoD-RW	95 °C for 7 min 30 cycles: 95 °C for 1 min, 60 °C for 1 min, 72 °C for 45 s 72 °C for 7 min	Adapted from Bergkemper *et al*. (2016) and Eder *et al.* (1996)
	ALPS-F730 ALPS-1101	95°C for 3 min 30 cycles: 95°C for 30 s, 57°C for 30 s, and 72°C for 30 s 72°C 10 min	Sakurai *et al.* (2008);

2.3.2. Cloning and Transformation

The purified gene fragments were cloned into the pJET1.2/blunt plasmid, using the sticky end cloning protocol from the CloneJET PCR Cloning Kit (Thermo Scientific), as follows. First, the purified PCR products were blunted with 1 µL of DNA blunting enzyme, in a solution of 10 µL of 2x reaction buffer and up to 7 µL of MiliQ water, for a total volume of 18 µL. This was centrifuged and incubated at 70 ºC for 5 minutes and then chilled on ice. To set up the ligation reaction, to these 10 µL, 1 µL of both T4 DNA ligase and pJET1.2/blunt (50 ng/ µL) were added, for a total of 20 µL of ligation mixture, which was then centrifuged and incubated at 20 ºC for 5 minutes and then used for *E. coli* transformation. The ligation mixture (10 µL) was used to transform 100 µL of competent *E. coli* cells (One Shot TOP10 Competent Cells, Invitrogen) using the chemical transformation (heat shock) procedure, according to manufacturer's instructions. After adding the ligation mixture, the cells were incubated on ice for 30 minutes and then at 42 ºC for exactly 30 seconds, after which 800 µL of pre-warmed SOC medium were added. These microcentrifuge tubes were incubated at 37 ºC for at least 1hour and 30 minutes. The tubes were centrifuged for 10 minutes at 3000 rpm, most of the supernatant was removed and the resulting pellet was resuspended and spread into LB medium plates with ampicillin (100 ng/mL). Following bacterial growth overnight at 37 °C, bacterial colonies (5-10 colonies) were picked and analyzed by colony PCR using the vector sequencing primers (CloneJET PCR Cloning Kit, Thermo Scientific). First, a swab of the plated colonies was taken and submerged in 50 µL of MiliQ water in a microcentrifuge tube, which was microwaved for 5 seconds to lyse *E. coli* cells. Colony PCR reactions consisted of 10 µL of Dream Taq Green PCR Master Mix(2x); 0.4 µL of pJET forward sequencing primer (10 µM); 0.4 µL of pJET reverse sequencing primer (10 µM); 5 µL of lysed bacteria, in a total reaction volume of 20 µL. The PCR thermal conditions were: 95 °C for 3 minutes, 25 cycles of 94 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 1 min (CloneJET PCR Cloning Kit User Guide, 2024). The PCR products were analyzed by agarose gel electrophoresis and those showing a DNA band with the specific amplicon size were selected as positive. Bacterial colonies from positive clones were grown on liquid LB medium (10 mL) with ampicillin (100 ng/ µL), overnight at 37 °C with agitation (225 rpm).

2.3.3. Purification of Plasmid DNA

Plasmid DNA was extracted from the positive bacterial cultures by firstly removing 1 mL of the liquid culture to a microcentrifuge tube and centrifugating it at 8000 rpm for 2 minutes, then removing the supernatant and repeating the process until at least 5 mL of the liquid culture had gone through this process and a pellet had been formed. Afterwards, using the GeneJET Plasmid Miniprep Kit (Thermofisher), the pellet was resuspended, lysed and neutralized with the kit's solutions. After centrifugation, the clear lysate was applied to a silica column, where DNA molecules selectively bind under high salt conditions. After binding, contaminants were removed through washing, and the pure plasmid DNA was eluted using elution buffer. Plasmid DNA concentration was then measured using a Nanodrop. To verify the presence of the target inserts in the plasmid vector, plasmid DNA was Sanger sequenced using the pJET1.2 vector sequencing primers (STAB VIDA Lda, Costa da Caparica,

Portugal). Sequences were searched against the NCBI database (BLASTN) to confirm that the correct gene had been cloned.

2.3.4. Plasmid Linearization and Determination of Gene Copy Number

The plasmid DNA was linearized with the restriction enzyme HindIII. The reaction mixture contained 1 µL of HindIII-HF (Takara Bio), 2 µL of Cutsmart buffer and 1 ng of plasmid DNA, in a total reaction volume of 20 µL. The reaction took place at 37 ºC, for 15 to 20 minutes and was stopped by inactivating the enzyme at 80 ºC for 20 minutes. The linearization was confirmed by gel electrophoresis. After linearization, plasmid concentrations were obtained using a Nanodrop and the gene copy number was calculated using the following equation (Fig.2.2):

$$copy\ number/\mu l = \left(DNA\ (ng/\mu l) \times 6,022 \times 10^{23} \right) \div \left(length(bp) \times 1 \times 10^{9} \times 660 \right)$$

6,022 x 10^{23}: Avogadro's number
$1x10^{9}$: Conversion factor
660: Average mass of one base pair of dsDNA (g / mol)
Length: pJET 1.2 plasmid length + insert length (base pairs)

Figure 2.2 Equation for gene copy number (gene copies / µL)

The plasmid DNAs were 10-fold serial diluted and used to produce the qPCR standard curves for the quantification of the taxonomic and functional genes on the soil samples (next section).

2.4. Real-Time Quantitative PCR (qPCR)

qPCR was used to quantify the gene abundance (copy number) of each taxonomic and functional gene in soil samples from the 3 land uses (tropical forest, cashew plantation and peanut field). qPCR reactions were performed on a Step One Real Time PCR System (Applied Biosystems), using SYBR Green chemistry and gene specific primers (Table 2.1). Primer pairs were selected according to the available literature. Each qPCR reaction contained: 5 µL of 2X SensiFAST SYBR Hi-ROX (Bioline); 0.2 µL of forward primer (10 µM); 0.2 µL of reverse primer (10 µM.); 1 µL of BSA (10 mg/µL); 2 µL of sample DNA in a total volume of 10 µL. Master mixes were prepared in a PCR chamber previously irradiated with U.V. light. Reaction parameters in the Step One qPCR equipment were: StepOne ™ Instrument (48 wells), Quantitation-Standard Curve, SYBR Green Reagents and Standard ramp speed. To increase specificity (amplification of only one PCR product) a touchdown qPCR protocol was used, according to the thermocycling conditions described on Table 2.3. Before qPCR of *bacterial* 16S rRNA gene, *GH6 cellulase* gene*, mcrA* gene and *phoD* gene*,* the soil DNA samples were diluted with miliQ water (*bacterial* 16S rRNA 1:1000; *GH6 cellulase, mcrA* and *phoD* 1:2;1:4;1:6 and 1:7) due to reaction inhibition resulting from excess DNA. The standard curve plasmid 10-fold serial dilutions were amplified in the same qPCR plate as the DNA samples extracted from the soil. Two qPCR replicated reactions (technical replicates) were performed for each biological replicated DNA soil sample. A no-template control (DNA replaced by MiliQ water) was included in each qPCR plate. Because degenerate primers and detection using SybrGreen can result in the production of artifacts, such as primer-dimers, melting curves were analyzed to ensure single-product formation and when multiple products formed PCR products were analyzed by agarose gel electrophoresis (as described earlier). Standard curves were obtained by plotting the Ct (threshold cycle) value for each point of the plasmid serial dilution against the log10 of the corresponding gene copy number. The copy number of the taxonomic and functional genes was calculated using the soil DNA Ct values and the standard curve. First, the logarithm of the number of gene copies per µL was calculated using the equation obtained from the standard curve. This number was multiplied by the dilution factor applied to the samples and the result was further multiplied by 50 µL since this was the volume of DNA extracted from each of the soil samples. Finally, this was

divided by the mass of soil used (g) to obtain the total number of gene copies present in 1 g of soil. The replicates were then averaged to obtain the total number of gene copies of each gene in 1 g of soil from each of the land uses (native tropical forest, cashew culture and peanut culture).

Table 2.3 Thermocycling conditions for qPCR

Target gene	**Primers**	**qPCR thermocycling conditions**	**Reference**
bacterial 16S rRNA	Eub341F Eub534R	Initial denaturation of 5 min at 95 ºC, followed by 1 min at 97 ºC, 10 cycles of: 30 s at 95 ºC; touchdown with a 1 ºC decrease per cycle, starting at 68 ºC for 45 s and ending at 58 ºC, which is the annealing temperature; and 45 s at 72 ºC of elongation. This was followed by 40 cycles of: 30 s at 95 ºC, 45 s at 58 ºC and 45 s at 72 ºC. Final elongation of 15 min at 72 ºC.	Adapted from Smits *et al.* (2004) and Muyzer. *et al*. (1993)
fungal ITS	ITS4 ITS5	Initial denaturation of 3 min at 96 ºC, followed by 1 min at 97 ºC, 10 cycles of: 30 s at 94 ºC; touchdown with a 1 ºC decrease per cycle, starting at 65 ºC for 30 s and ending at 55 ºC, which is the annealing temperature; and 1 min at 72 ºC of elongation. This was followed by 40 cycles of: 30 s at 94 ºC, 30 s at 55 ºC and 1 min at 72 ºC. Final elongation of 7 min at 72 ºC	Adapted from Khan.and Bhadauria (2018)
GH6 cellulase	Cell2F Cell2R	Initial denaturation of 5 min at 9 5ºC, followed by 1 min at 97 ºC, 10 cycles of: 30 s at 95 ºC; touchdown with a 1 ºC decrease per cycle, starting at 74 ºC for 45 s and ending at 62 ºC, which is the annealing temperature; and 45 s at 72 ºC of elongation. This was followed by 40 cycles of:30 s at 95 ºC, 45 s at 62 ºC and 45 s at 72 ºC. Final elongation of 15 min at 72 ºC.	Adapted from Merlin *et al.* (2014)
pmoA	A189 Mb661	Initial denaturation of 5 min at 9 5ºC, followed by 1 min at 97 ºC, 12 cycles of: 45 s at 9 5ºC; touchdown with a 1 ºC decrease per cycle, starting at 62 ºC for 45 s and ending at 50 ºC, which is the annealing temperature; and 45 s at 72 ºC of elongation. This was followed by 40 cycles of: 30 s at 95 ºC, 50 s at 50 ºC and 45 s at 72 ºC. Final elongation of 15 min at 72 ºC.	Adapted from Kou *et al.* (2017) and Costello and Lidstrom (1999)
mcrA	mlas mrcA rev	Initial denaturation of 3 min at 95°C, followed by 1 min at 97ºC, 10 cycles of: 30 s at 95°C, touchdown with a 1 ºC decrease per cycle, starting at 65ºC for 45 s and ending at 55 ºC, which is the annealing temperature; and 30 s at 72 ºC of elongation. This was followed by 40 cycles of: 30 s at 95ºC, 45 s at 55°C, and 30 s at 72°C, and a final extension of 10 min at 72°C	Adapted from Steinberg and Regan (2009)
phoD	ALPS-F730 ALPS-1101	Initial denaturation of 3 min at 95 ºC, followed by 1 min at 97 ºC, 10 cycles of: 30 s at 95 ºC; touchdown with a 1 ºC decrease per cycle, starting at 67 ºC for 30 s and ending at 57 ºC, which is the annealing temperature; and 30 s at 72 ºC of elongation. This was followed by 40 cycles of: 30 s at 95 ºC, 30 s at 57 ºC and 30 s at 72 ºC. Final elongation of 10 min at 72 ºC.	Sakurai *et al*. (2008)

2.5. β-Glucosidase and Phosphatase Soil Enzymatic Activity

Soil β-glucosidase and phosphatase enzymatic activity were measured according to the protocol from the UIUC Soils Lab (Soil Labs SOP, University of Illinois, 2021), using para-nitrophenyl (pNP) linked substrates that release para-nitrophenol (pNP) when hydrolyzed by soil enzymes. For the β-glucosidase activity the substrate used was the para-nitrophenyl-B-d-glucopyranoside (pNPG), whereas for the phosphatase activity it was the para-nitrophenyl-phosphate (pNPP). In these assays, soil and pNP-linked substrates were added to buffer or water, incubated at 37 °C for 1- 2 hours and then the reaction was terminated by adding a base, followed by quantification of the released pNP by Abs reading at 410 nm. For the phosphatase activity, 1 g of soil from each biological replicate was transferred to 50 mL Falcon tubes followed by 5 mL of 20 mM pNPP. For the β-glucosidase activity, 5 mL of 10 mM pNPG were added to another set of tubes with soil. Soil-only controls were prepared in the same way but with no substrate (pNPP/pNPG) added, adding water instead to account for dissolved organic matter (DOM) interference. Substrate-only controls were prepared in the same way but with no soil added, only 5 ml of pNPP/pNPG. The tubes were capped and top wrapped with parafilm to avoid evaporation during incubation, and swirled for 1 min. The tubes were incubated at 37 ºC for 1 hour in a water bath. A standard curve of p-nitrophenol (p-NP) was prepared (Table 2.4). The 0 mM p-NP (only water) was also used as no-soil no-substrate control (true blank). To terminate the enzymatic reaction, 4 mL of 0.1 M THAM (Tris, or trishydroxymethylaminomethane) and 1 mL 2.0 M $CaCl_2$ were added to samples, blanks and standards which were left to sit for approximately 5 minutes until clear aliquot was evident. 1 mL of supernatant of the resting reaction solution was transferred into microcentrifuge tubes that were centrifuged (14,000 rpm for 1 minute and 45 seconds). 200 µL of the supernatant were transferred to a 96-well microplate and the absorbance was measured at 410 nm. The absorbance readings of the samples and controls were converted to pNP concentrations (mM) based on the calibration curve obtained from the pNP standards, and the potential enzymatic activity (µmol pNP g^{-1} soil hr^{-1}) was obtained using the following equation (Fig 2.4):

$$Potential\ activity = \frac{(a \times b \times c \times b - d \times b) \times 0.01 \times 1000}{1g\ of\ soil \times 1h}$$

a: Sample concentration (mM)
b: Dilution
c: Hydrolysis blank concentration (mM)
d: DOM blank concentration (mM)

Figure 2.3 Equation for potential enzymatic activity (µmol pNP g^{-1} soil hr^{-1}) adapted (Soil Labs SOP, University of Illinois, 2021).

Table 2.4 - Preparation of standard curve of p-nitrophenol (pNP).

pNP 5 mM (mL)	**Water** (mL)	**Final pNP** (mM)
0.00	5.00	0.00
0.05	4.95	0.05
0.10	4.90	0.10
0.20	4.80	0.20
0.40	4.60	0.40
0.60	4.40	0.60
0.80	4.20	0.80

2.5. Statistical Analysis

Statistical analysis was performed to test the effect of land use (forest, cashew plantation, peanut field) on the target genes and enzymatic activities involved in the carbon and phosphorus cycles, using the IBM SPSS statistics software (version 29.0 SPSS Inc., USA). First, the gene abundances and enzymatic activities were analyzed for normality and homogeneity of variance using the Shapiro-Wilk and the Levene's test, respectively. ANOVA was performed to determine if the difference between the means of the land uses was statistically significant ($p<0.05$) and, if it was significant, a Tukey post hoc test was performed to determine between which land uses there were statistical differences. When the variances were not homogenous, a Kruskall Wallis non-parametric test was performed.

3. Results and Discussion

3.1. Soil DNA Extraction

). The average concentration l $4.32x10^{-6}$ g of DNA/g of soil for the forest samples, $2.93x10^{-6}$g of DNA/g of soil for the peanut soil samples and $4.56x10^{-6}$g of DNA/g of soil for the cashew soil samples. Abs 260/280 and Abs 260/230 ratios were measured (Table 3.1). In soil samples the most common contaminants extracted with DNA are humic acids, containing phenolic groups. These groups can denature other molecules by bonding to amides or oxidizing, forming quinones that bond to proteins and DNA. These humic acids can also inhibit DNA polymerase and chelate magnesium, an essential cofactor for this enzyme. The presence of inhibitors is especially problematic for qPCR since this method is considered to be more sensitive than the standard PCR, (Venturini *et al.*, 2020). Since the 260/280 ratio values obtained were within the 1.4 – 1.8 range, while the 260/230 ratios varied between 0.7–1.5, these values are in accordance with those reported for soil samples extracted with commercial kits from various biotech companies (Venturini *et al.*, 2020) and indicate that the DNA samples had low levels of inhibitors [e.g. proteins (abs at 280nm), humic acids, urea, lipids, guanidine, EDTA, salts, phenols and polysaccharides (abs at 230nm)]. Therefore, the low levels of inhibitors make our soil DNA samples adequate for qPCR analysis.

Table 3.1 Average Nanodrop results (concentration in ηg/µl, Abs 260/280 and 260/230 ratios) for each land use, with the corresponding standard deviation e

Land use	Average Soil DNA Concentration (ng/ul) ± SD	Average Abs 260/280 ± SD	Average Abs 260/280 ± SD
Forest	11.23 ± 4.35	1.69 ± 0.33	1.15 ± 0.58
Peanut Field	9.2 ± 4.84	1.58 ± 0.34	0.54 ± 0.64
Cashew Field	13.2 ± 4,19	1.70 ± 0.32	0.69 ± 0.35

Analysis by agarose gel electrophoresis (Fig 3.1) shows a band of high molecular weight corresponding to the genomic DNA extracted from soil samples.

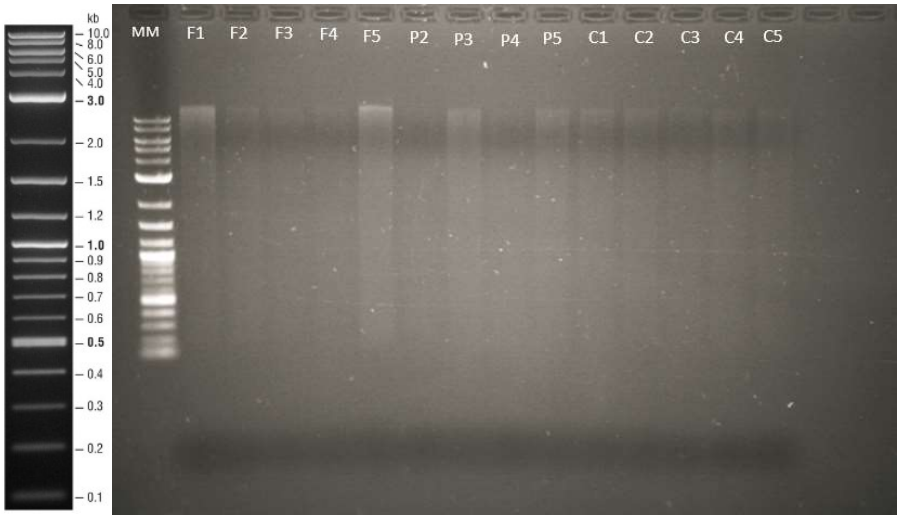


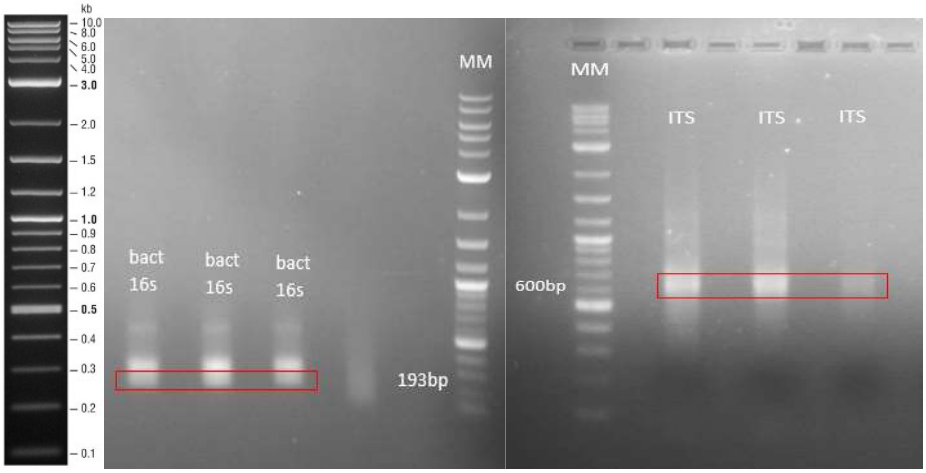
Figure 3.1 Agarose gel electrophoresis of DNA (~50ng) extracted from soil samples. MM: molecular marker; F: forest; P: peanut field; C: cashew field.

3.2. qPCR Standard Curves of Taxonomic and Functional Genes

3.2.1. Amplification of Target Genes from Soil DNA

After soil DNA extraction, PCR was performed to amplify the target genes using specific primer pairs found in the literature. DNA bands with the specific amplicon sizes were obtained for the *fungal* ITS (600bp), the *bacterial* 16S rRNA (193bp) and the *archaeal 16S rRNA* (600 bp) genes (Fig 3.2). For the functional genes, DNA bands were obtained for the *pmoA* (491bp), *mcrA* (500bp), *GH6 cellulase* (150bp) and *phoD* (371bp) genes (Fig 3.2). For some genes (*GH6 cellulase, phoD, mcrA*) nonspecific DNA bands were detected on the gel and a touch-down PCR program was tested, which resulted in more specific amplifications (Fig 3.2), except for the *mcrA* gene, which showed an additional DNA band with lower molecular weight. No amplification was obtained for the *pcaH* and the *β-glu* genes from the soil DNA samples. Therefore, for the *pcaH* and *β-glu* genes no additional experiments were performed.

The *pcaH* gene is involved in the β-ketoadipate pathway that degrades aromatic compounds derived from lignin, by cleaving their rings. It was to be expected that it would be found in all the land uses. Because of this, it is hypothesized that the primer pair available (PCAHf/PCAHr) was not suitable for amplification of this gene or the PCR conditions were not optimal. It could also be that the *pcaH* gene was simply not present on the samples, however, that seems less likely since this pathway is a major catabolic route in terrestrial soils especially in what concerns the catabolism of phenolic compounds. Furthermore, this gene is encoded in the chromosomes and is biologically conserved and widespread in the soil bacteria and fungi (Buchan *et al*., 2001). The *β-glu* gene encodes a β-glucoside which is present in all domains of life, where it plays an essential role in the removal of non-reducing terminal glucosyl residues from saccharides and glycosides. Since this gene is present in all domains of life and is especially important for the conversion of biomass by microorganisms (Cairns and Esen, 2010) it was, like the *pcaH* gene*,* to be expected that it would be easily amplified from the DNA extracted from soil samples. Therefore, it was hypothesized that the primer pair selected (BgluF2/BgluR4) was not suitable and/or the PCR conditions were not ideal. Two additional primer pairs were tested (GH1BF/GH1BR and GH3EF/GH33R) but still no amplification of the *β-glu* gene was achieved. Considering these results, it was assumed that the microorganisms with this gene may not be present in sufficient amounts to be detected in the soils sampled. The *β-glu* gene was replaced in this study by a gene encoding a cellulase of the GH6 family, which even though not a β-glucosidase, is present in the same cellulase complex (Merlin *et al.,* 2014).



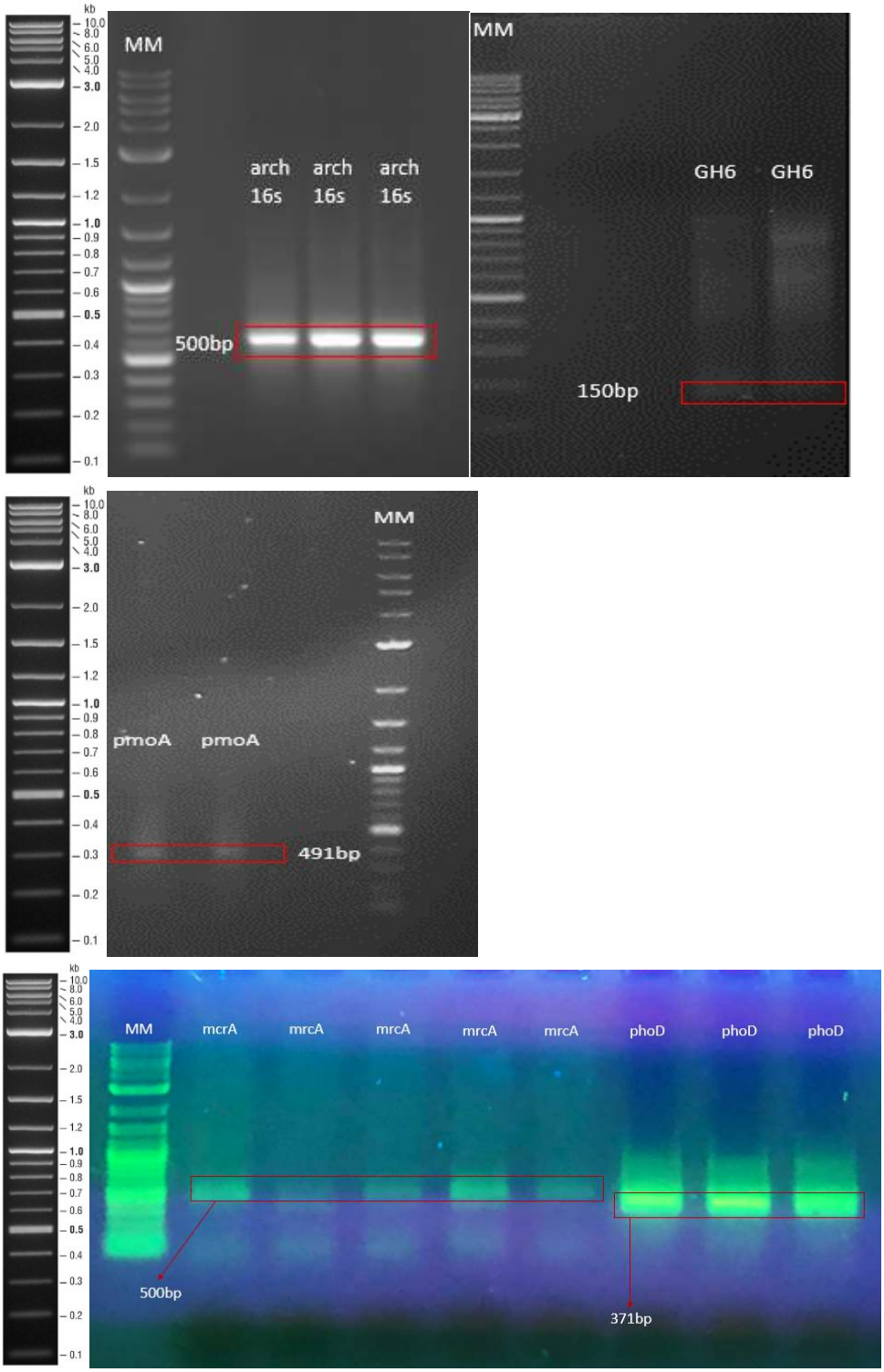


Figure 3.2 Agarose gel electrophoresis of taxonomic and functional gene amplicons, amplified of a pool of soil DNA samples.

3.2.2. Cloning and Transformation

. All the gene fragments, except for *archaeal* 16S rRNA, produced a DNA band with the expected amplicon size, showing that the cloning of the gene fragments had been successful (Fig 3.3). In the case of *archaeal* 16S rRNA, a band was indeed produced but not of the expected amplicon size (600bp) indicating that this gene fragment was not successfully cloned. The *E. coli* cells were indeed

transformed with the pJET 1.2 plasmid, since there was bacterial growth. However, the target gene fragment did not appear to be there. Therefore, this gene was not further analyzed.

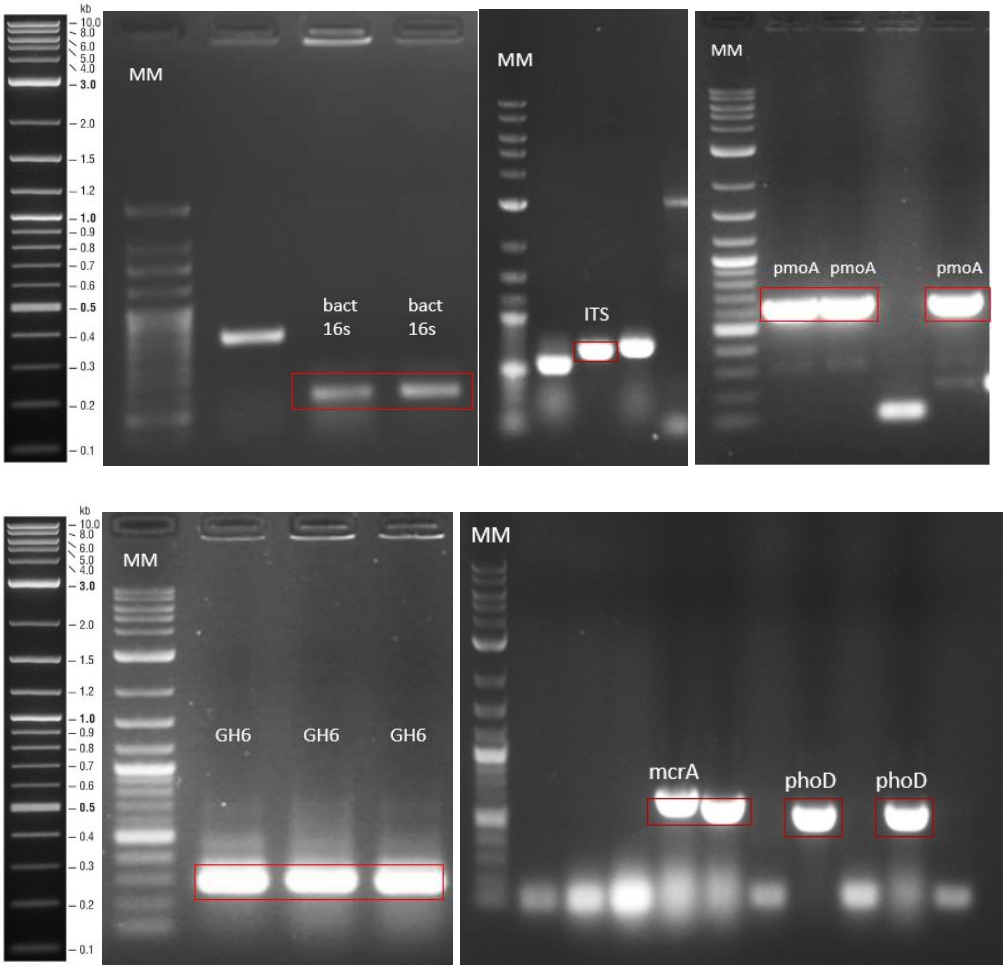


Figure 3.3 Agarose gel electrophoresis of colony PCR products, colony PCR amplification with cloning vector primer pair

To confirm that the right sequences had been cloned, plasmid DNA from each gene was extracted by miniprep and sequenced. The sequences (Supplementary Table 5.1) were searched (BLASTN and BLASTX) against the NCBI database. Matches were obtained for all the sequenced gene fragments, confirming the specificity of the primers used.

3.2.3. Linearization of Plasmid DNA

The plasmid DNA of each functional (*GH6 cellulase, pmoA, mcrA and phoD)* and taxonomic gene (*bacterial* 16S rRNA and *fungal* ITS) was linearized by digestion with the restriction enzyme HindIII. The digestion was confirmed by agarose electrophoresis gel (Fig 3.4)

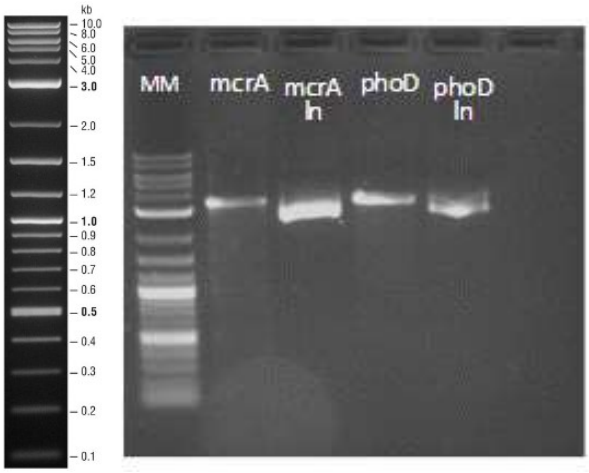


Figure 3.4 Agarose gel eletrophoresis of the purified plasmid DNA of *mcrA* and *phoD*, post digestion with the restriction enzyme HindIII. "ln" stands for "linear"

3.2.4. Determination of Plasmid Copy Number

After plasmid linearization, the plasmid concentrations were obtained (Figure 3.5):

Table 3.2 Number of copies/µl of each target gene obtained post digestion of the purified plasmid DNA.

Target genes	**Copy number** (gene copies)
bacterial 16S rRNA	$1{,}35 \times 10^{10}$
ITS	$1{,}05 \times 10^{10}$
GH6 cellulase	$1{,}34 \times 10^{10}$
pmoA	$1{,}11 \times 10^{10}$
mcrA	$1{,}16 \times 10^{10}$
phoD	$1{,}83 \times 10^{10}$

The plasmid DNAs were 10-fold serial diluted to produce the qPCR standard curves for quantification of the abundance of each taxonomic and functional gene in the soil from the native tropical forest, peanut field and cashew plantation.

3.2. Quantification of Taxonomic and Functional Gene Abundance

qPCR was performed to quantify gene abundance (copy number/ g soil) of each target gene in soil from the three land uses (native tropical forest, cashew plantation and peanut field). The performance of the qPCR standard curves is shown on Table 3.3.

Table 3.3 qPCR R^2 and efficiency, standard curve linearity and detection limit.

Target Genes	R^2	Linearity	Efficiency
bacterial 16S rRNA	0.994	$10^8 – 10^6$	57%
Fungal ITS	0.999	$10^5 – 10^3$	89%
GH6 cellulase	0.988	$10^5 – 10^2$	99%
pmoA	0.997	$10^4 – 10^1$	87%
mcrA	0.998	$10^7 – 10^4$	79%
phoD	0.991	$10^4 – 10^1$	99%

The 10-fold dilutions were linear from 10^8 to 10^1 gene copies, with the R^2 value ranging from 0.988 to 0.999, and the qPCR efficiency from 57% to 99%.

For the *bacterial* 16S rRNA, the melting temperature (Fig 3.5) of the cloned gene fragments and DNA extracted from soil samples varied between 83 ºC and 85 ºC. Melting curve analysis shows that most of the gene fragments amplified from soil DNA present only one peak corresponding to a single PCR product. However, some DNA soil samples showed melting curves with an additional peak. Therefore, the qPCR reactions were analyzed by agarose gel electrophoresis (Fig 3.6). The results showed that for all the soil samples, only one DNA band was produced with the expected size (193 bp), and therefore, there was only one amplification product.

For the *fungal* ITS, the melting temperature (Fig 3.5) of the cloned gene fragments and DNA extracted from soil samples varied between 83 ºC and 91 ºC. Several DNA soil samples showed melting curves with more than one peak. However, the qPCR reactions' analyses by agarose gel electrophoresis (Fig 3.6) showed that for most of the soil samples, only one DNA band was produced with the expected size (600 bp). For some samples an additional band of a higher molecular weight was visible in the forest samples. It was hypothesized that the variability in the fragments' melting temperature reflects the higher diversity of ITS sequences among the microorganisms present at these soil sample. Since this variation was visible mainly in the forest samples it may be indicative that forests are the ecosystems with the highest soil biodiversity. The melting curve of the negative control showed a peak, however, in the electrophoresis gel no DNA band corresponding to the specific size (600 bp) was detected suggesting that some primer dimer could have formed in the absence of the template DNA (NTC).

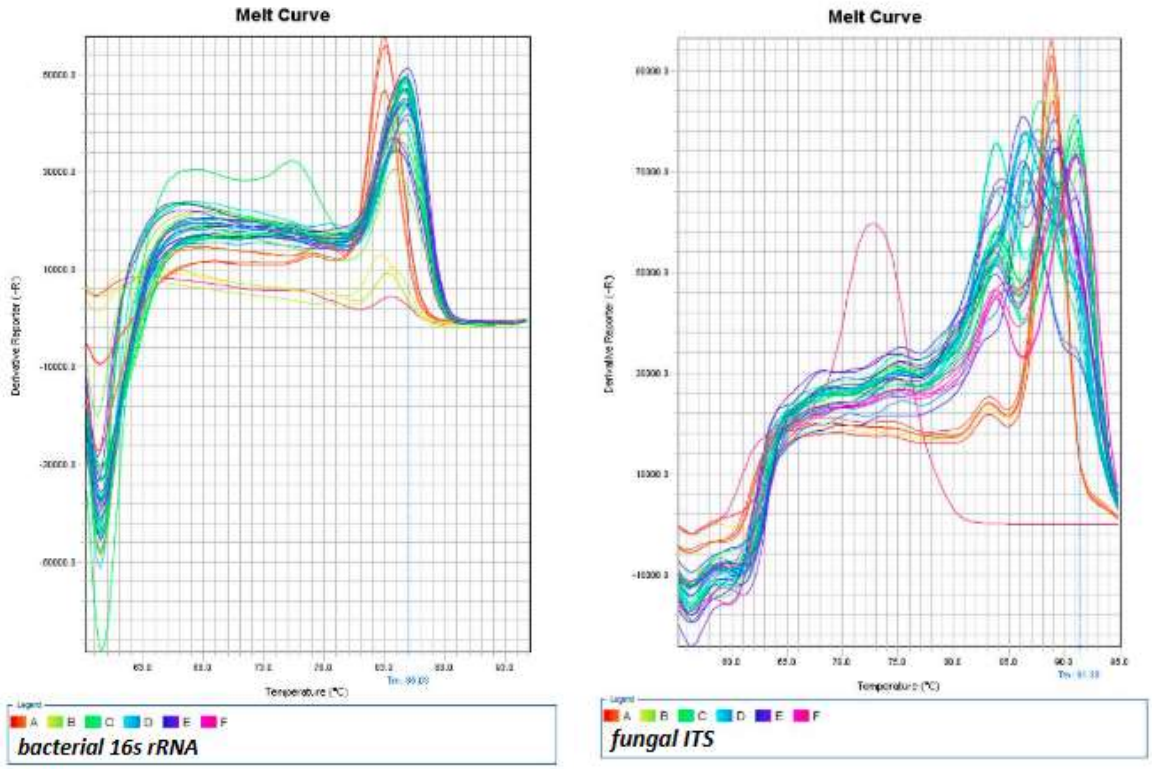
The melting temperature (Fig 3.5) of the cloned *GH6 cellulase* gene fragments and DNA extracted from soil samples varied between 87 ºC and 89 ºC. Melting curve analysis shows that only one PCR product was amplified from the template DNA samples (Fig. 3.7). Like for the *fungal* ITS gene, a dissociation peak was detected in the no template control, however, at the electrophoresis gel (Fig 3.6) no DNA band corresponding to the specific size (150 bp) was detected.

The melting temperature (Fig 3.5) of cloned *pmoA* gene fragments and DNA extracted from soil samples varied between 86 ºC and 90 ºC. Some fragments amplified from DNA soil samples showed melting curves with an additional peak. However, the analysis by agarose gel electrophoresis (Fig 3.6) showed that only one DNA band was produced with the expected size (491 bp). Like for the *fungal ITS* gene and *GH6 cellulase* gene, a dissociation peak was observed in the negative control, however, in the electrophoresis gel no DNA band corresponding to the specific size (491 bp) was detected.

The melting temperature (Fig 3.5) of cloned *mcrA* gene fragments and DNA extracted from soil samples varied between 87 ºC and 90 ºC. Most DNA soil samples showed melting curves with more

than one peak (Fig. 3.5). The analysis by agarose gel electrophoresis (Fig 3.6) showed that one DNA band was produced with the expected size (500 bp), and an additional band of lower molecular weight (approximately 150 bp) was also detected. It was detected in all the samples, including the serial plasmid dilutions used for the standard curve and in the negative control. Because of this it was assumed to be the result of primer dimer.

The melting temperature (Fig 3.5) of cloned *phoD* gene fragments and DNA extracted from soil samples varied between 90 ºC and 91ºC. Melting curve analysis shows that only one PCR product was amplified (Fig. 3.5), further confirmed by the agarose gel analysis (Fig 3.6).



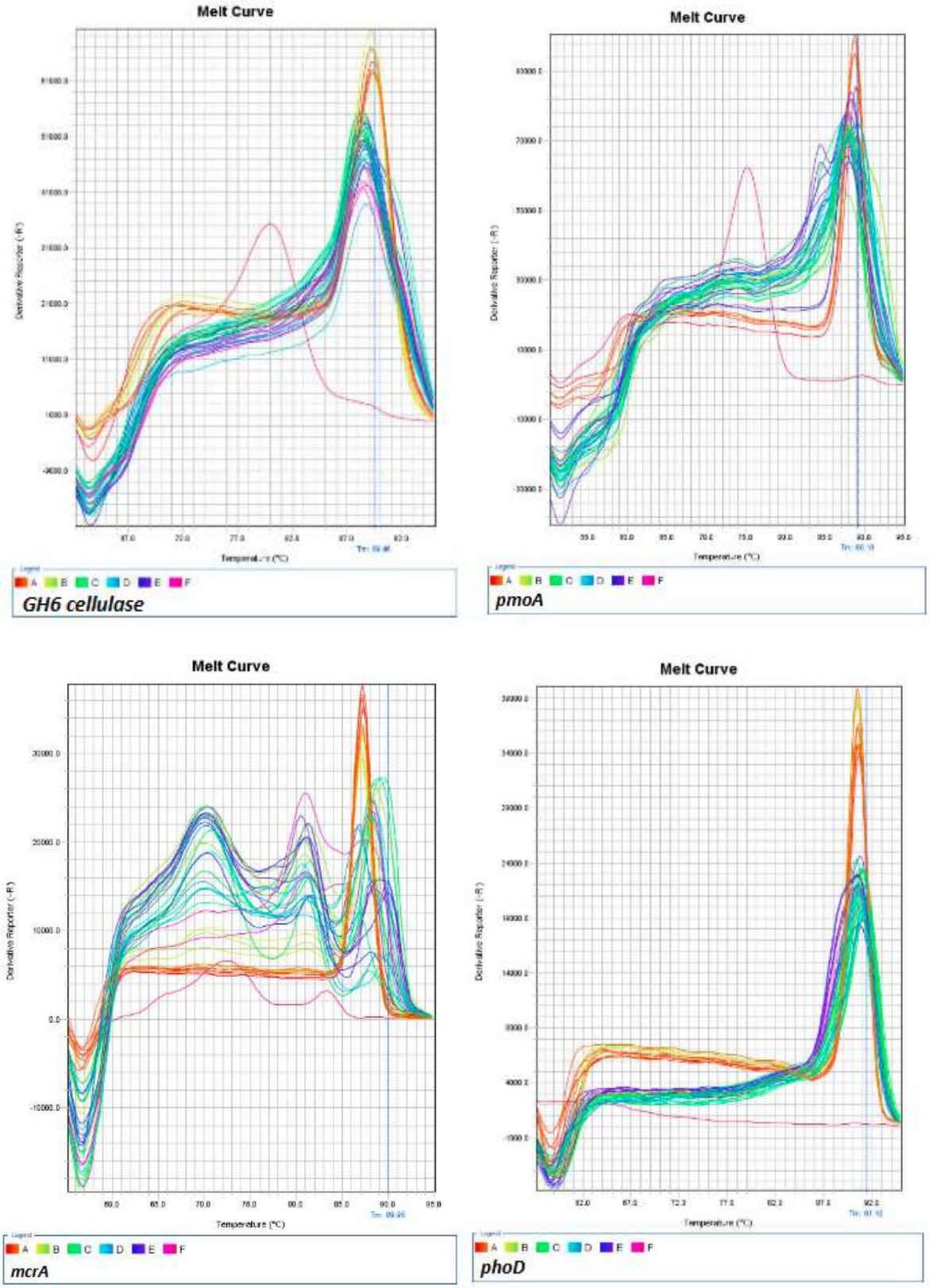
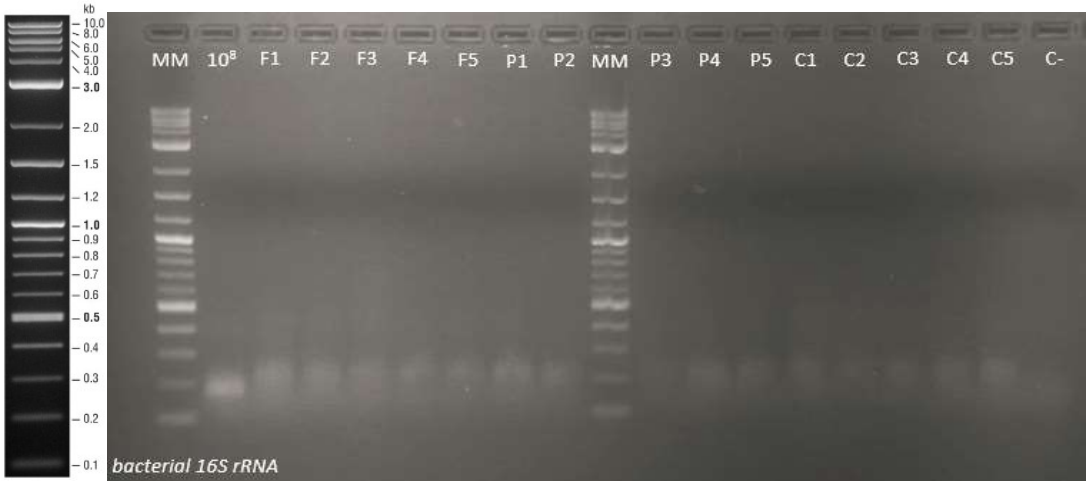
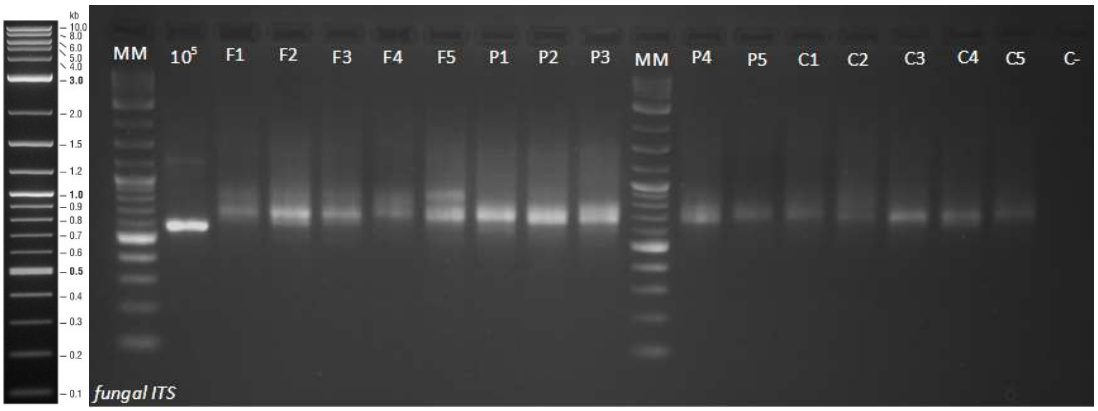
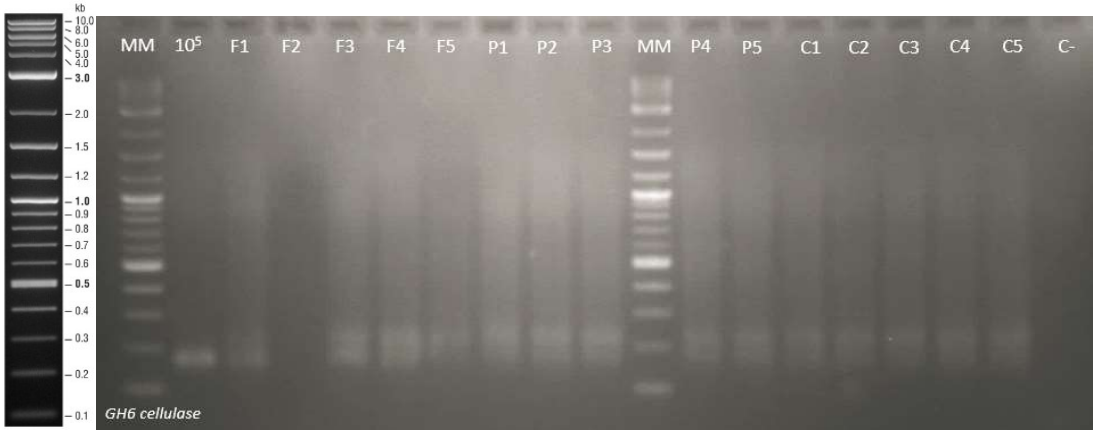
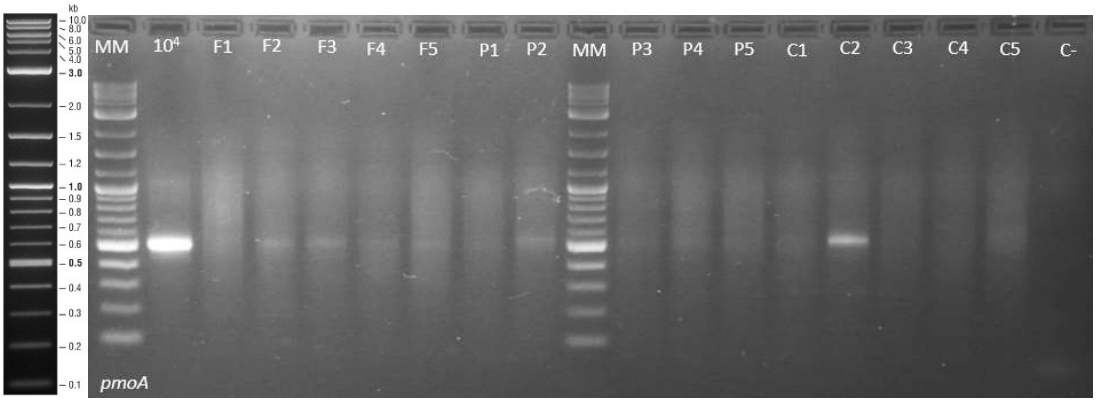


Figure 3.5 Second derivative dissociation curve relating the SYBR Green fluorescence emission to the denaturation temperature in the qPCR of the taxonomic and functional genes.


kb
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.2
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
MM
10⁸
F1
F2
F3
F4
F5
P1
P2
MM
P3
P4
P5
C1
C2
C3
C4
C5
C-
bacterial 16S rRNA


kb
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.2
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
MM
10⁵
F1
F2
F3
F4
F5
P1
P2
P3
MM
P4
P5
C1
C2
C3
C4
C5
C-
fungal ITS


kb
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.2
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
MM
10⁵
F1
F2
F3
F4
F5
P1
P2
P3
MM
P4
P5
C1
C2
C3
C4
C5
C-
GH6 cellulase


kb
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.2
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
MM
10⁴
F1
F2
F3
F4
F5
P1
P2
MM
P3
P4
P5
C1
C2
C3
C4
C5
C-
pmoA

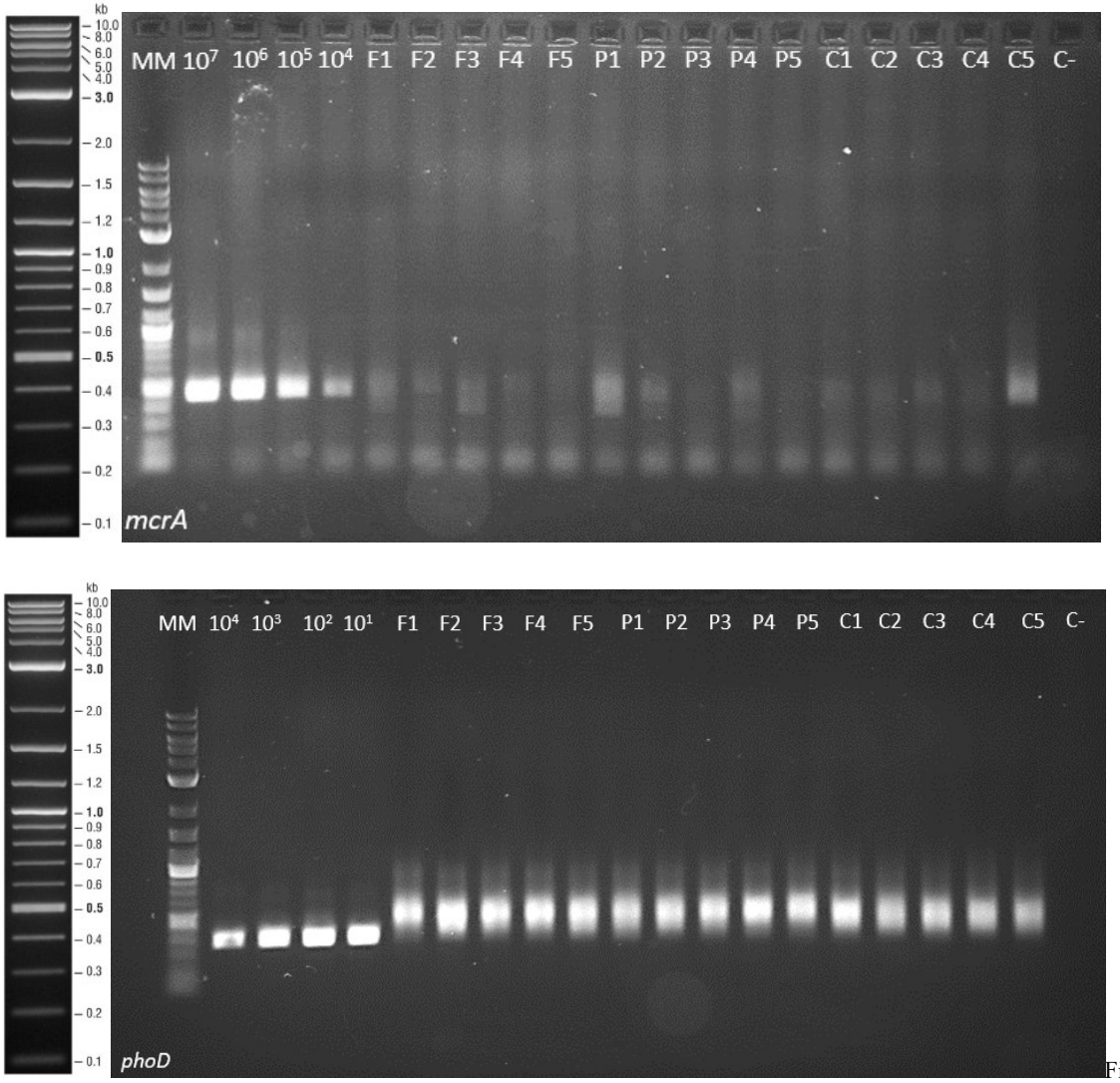
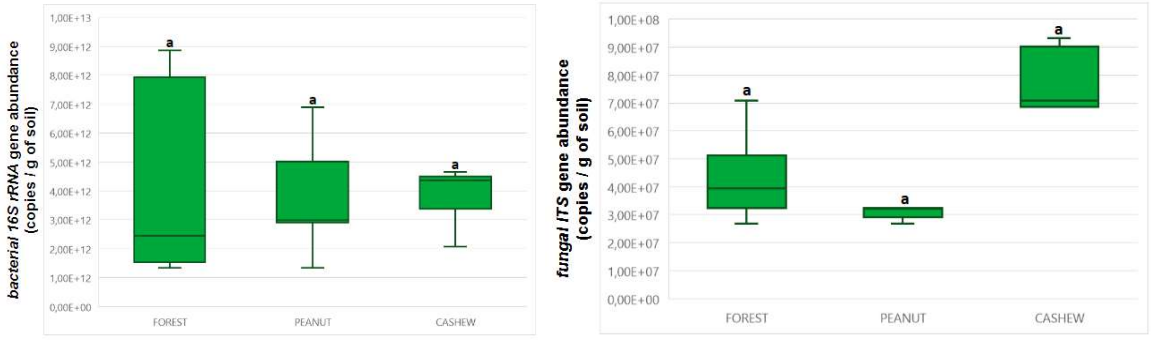


Figure 3.6 Agarose gel electrophoresis of the real-time PCR products obtained from the five replicates of the three land uses for the taxonomic and functional genes.

Figure 3.7 shows the abundance of taxonomic and functional genes in the soils from the 3 land uses (forest, peanut and cashew).



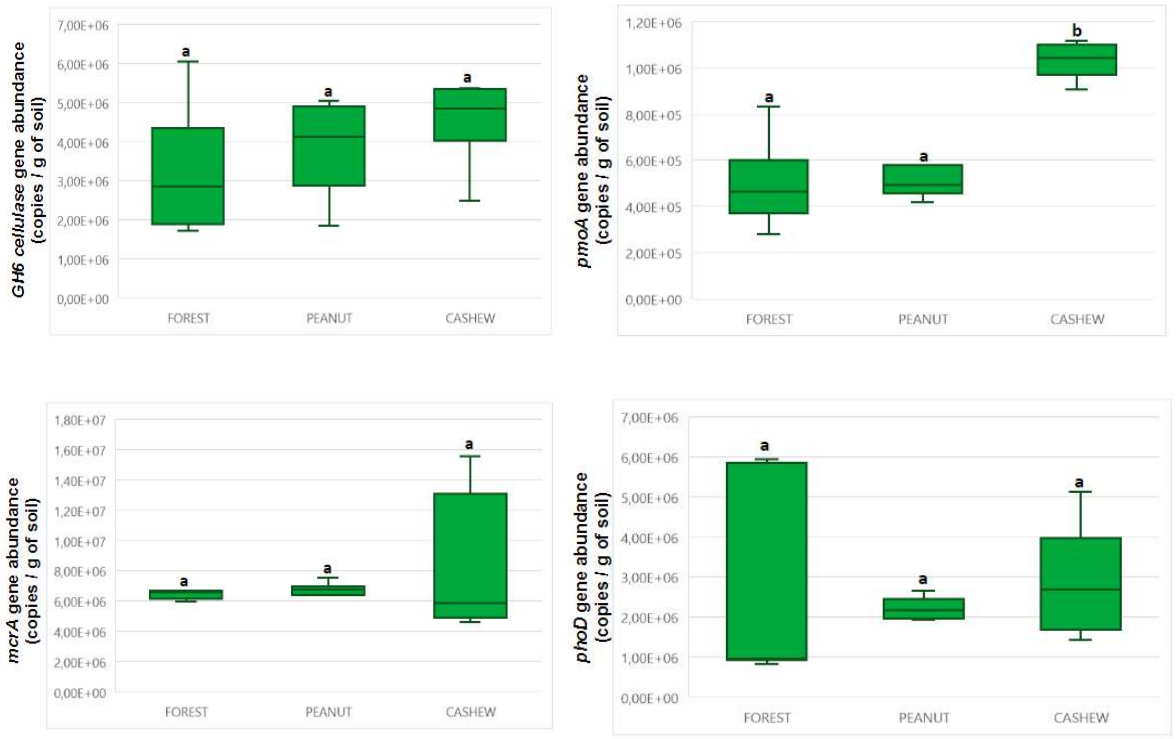


Figure 3.7 Gene abundance in soil from different land uses (forest, peanut, cashew). Different letters indicate significant differences at $p < 0.05$ according to ANOVA (n=5).

For both *bacterial* 16S rRNA gene (3.80×10^{12} to 4.42×10^{12} gene copies / g of soil) and *fungal* ITS gene (3.77×10^{7} to 6.92×10^{7} gene copies / g of soil) there were no significant differences in gene abundance between the three different land uses (Fig 3.7), indicating that the total amounts of bacteria and fungi in the soil microbiome did not change significantly when the native tropical forest was converted into peanut or cashew fields. However, nothing can be said about microbial diversity with just these results. It is known that deforestation, like the one that happens preceding the transformation of the native tropical forest into agricultural land, leads to the disruption of resource availability and habitat suitability for soil organisms, including soil microorganisms. This often results in the decline of specialist species, decreasing functional diversity and increasing the presence of generalist species. So, even if the total amount of microorganisms does not change, their biodiversity is known to change, even going as far as decreasing (Qu *et al.*, 2024). The fact that the total amount of soil microorganisms did not change significantly between the three land uses might be because the agricultural techniques used to transform the native tropical forest into a cashew field and a peanut field, like manual tillage, did not disturb the soil as much as other techniques, like mechanical tillage, that have been shown to decrease the total amount of soil microorganisms in the converted agricultural land (Singh *et al*., 2020).

Cellulase is involved in the carbon cycle, taking part in the degradation of carbon compounds. There were no significant differences in the amount of the *GH6 cellulase* gene between the three soil uses (Fig 3.7), suggesting that cellulose degradation was not significantly affected when the native forest was converted into peanut or cashew cultures. These results showed that the number of microorganisms with *GH6 cellulase* gene (3.38×10^{6} to 4.42×10^{6} gene copies / g of soil) was not significantly affected by the conversion of native tropical forest into peanut and cashew plantation fields. The results did not show, however, which cellulase containing organisms were present, what genera or species, and in what amount. Merlin *et al*. (2014) have also quantified the abundance of the *GH6 cellulase* gene and found that, within the scope of their samples, it was highest in a grassland, and that there were significant differences in *GH6 cellulase* gene abundance in their tested soils. The soil samples tested in the current

project correspond only to forest and agricultural land so a direct comparison cannot be made. However, taking into consideration their findings, it could be assumed that the native tropical forest would have a significantly higher amount of *GH6 cellulase* gene than what was obtained for the cashew and peanut fields. Since that was not observed, it was hypothesized that the practice of manual tillage in the agricultural fields was not enough to significantly lower the abundance of *GH6 cellulase* gene.

The abundance of the *pmoA* gene (5.10×10^5 to 1.03×10^6 gene copies / g of soil) was significantly affected by the land use, being more abundant in the soil of the cashew culture (Fig 3.7). This seems to suggest that in this soil microbiome there was a more abundant methanotrophic community, relative to the forest or the peanut culture. These results could suggest that in the soil microbiome of the cashew culture there might be a more intensive methane consumption, making the cashew plantations acting as a methane sink. Our results seem to be in contradiction with other literature reports, where it has been found that in the native forest there is a higher amount of methanotrophs, in comparison with agricultural fields (Tiwari *et al.,* 2018), possibly due to more intensive practices than the ones used in the agricultural cultures in this study. Methane oxidation has been reported to reduce with the increase in disturbance of the soil, namely between forest soils and agricultural lands. (Tiwari *et al.*, 2018). Smaller amounts of methanotrophs observed in agricultural soils, in comparison to forest soils, have also been suggested to happen due to an inhibitory effect caused by high amounts of nitrogen, either because of the production of nitrogen rich cultures, the use of fertilizers rich in nitrogen, or both (Tiwari *et al*., 2018). Since in the cultures studied here no synthetic fertilizers were used, there might not have been nitrogen related inhibition and therefore more copies of the *pmoA* gene were present in the soil of the cashew culture instead of in the soil of the native tropical forest.

The abundance of the *mcrA* gene (6.62×10^6 to 8.82×10^7 gene copies / g of soil) was not significantly affected by the land use, but abundances were higher in the soil of the cashew culture (Fig 3.7). This seems to suggest that the methanogenic community did not suffer significant changes with native forest conversion. It is interesting that although there were no significant changes in abundance, the maximum number of gene copies per gram of soil was registered in the cashew culture, where the methanotrophic (*pmoA*) community did seem to significantly increase. This seems to be in accordance with some of the existing literature, although there does not seem to be a consensus. Alves *et al.* (2022) reported that the abundance of methane oxidation genes, like *pmoA*, suffer a reduction with forest conversion to pasture but methanogenesis genes, like *mcrA*, do not. However, they observed changes in the composition of the methanogenic microbial community, namely an increase in abundance and richness.

The abundance of the *phoD* gene (2.24×10^6 to 2.99×10^6 gene copies / g of soil) was not significantly affected by the land use, but abundances were higher in the soil of the native tropical forest (Fig 3.7). This seems to be in accordance with findings by other authors, such as Neal *et al*. (2017), that reported no significant differences in *phoD* gene abundance between land uses (grassland, bare fallow soil and arable soil), although higher abundances of *phoD* gene were reported in the grassland. It has also been reported (Azene *et al*., 2023) that natural forests have significantly higher abundance of the *phoD* gene than other land uses (artificial forests, farmlands and shrublands). It is important to note that the land uses analyzed were not the same as the ones analyzed here so no definitive conclusion can be made. Azene *et al.* (2023) also refers to the fact that there is no consensus about the relation between the abundance of the *phoD* gene and alkaline phosphatase (ALP) activity, since both a positive relation and no relation at all have both been previously reported. Part of the pho regulon, the *phoD* gene encodes part of an (extracellular) ALP. The results of the current study might suggest that there is no significant difference in the amount of this enzyme between the land uses.

It is important to note that all molecular methods have their limitations. Among them are the fact that DNA-based methods can amplify DNA of dead or inactive organisms (for example fungal spores). When it comes to qPCR specifically, some disadvantages include the fact that the quality of the

data obtained depends on the matrix from which the DNA is extracted, in this case soil, and the associated contaminants, leading to, possibly, PCR inhibition; the use of small soil samples, in this case about 0.150 g, which might not provide an accurate representation of the soil microbial community in the non-homogenous soil; and that it relies strongly on the specificity of the primers used. Equally important is to highlight the advantages of using qPCR, the reasons why it was chosen for this project, namely being highly standardized, sensitive, selective and reproducible; cheaper than other methods; and the large number of targets it can analyze (Geisen *et al*., 2019; Thiele-Brhun. *et al*., 2020).

3.3. Soil Enzyme Activities

Soil enzymes catalyze several biochemical reactions which result in the transformation of organic matter, and the release of inorganic matter for plant growth and nutrient cycling (Adetunji *et al.*, 2017). To supplement the data obtained from qPCR, an enzymatic essay to test the activity of glucosidase and phosphatase in the three land uses was performed. The results are shown in figure 3.8.

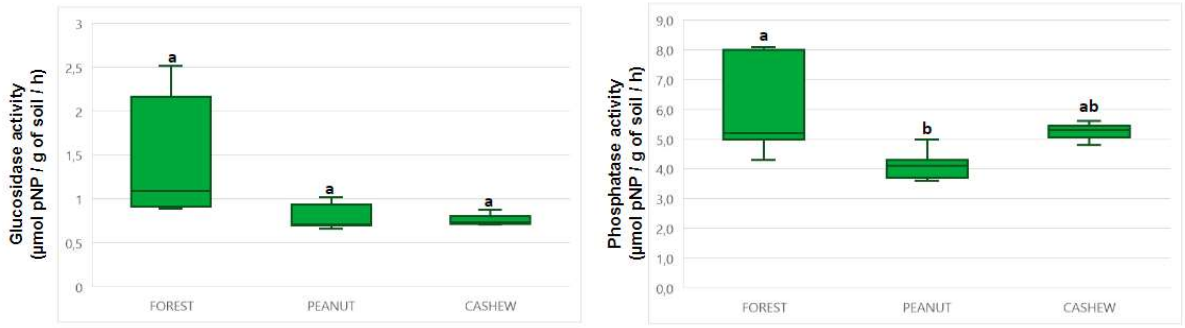


Figure 3.8 Glucosidase and phosphatase activity in soil from different land uses (forest, peanut, cashew). Letters indicate significant differences at $p < 0.05$ according to ANOVA (n=3-5).

The highest average activity of glucosidase, 1.51 µmol pNP / g of soil / h, was obtained experimentally in the forest soil, and the lowest average, 0.77 µmol pNP / g of soil / h, in the cashew field. The land use change did not significantly affect glucosidase activity in the soil microbiome (Fig 3.8). Glucosidase activity has been previously found to be significantly higher in forests than in agricultural land, where it was concluded that land conversion had disturbed the growth of the glucosidase producing organisms. It was also found that, in converted land with conventional (mechanical) tillage, adoption of a no-till system and a reduction in tillage frequency could increase β-glucosidase activity, where microbial biomass, carbon and nitrogen also increased. The increase in β-glucosidase activity was due to improvement in microbial biomass and substrate availability as well as a reduction in soil disturbance because of the reduction in tillage intensity (Adetunji *et al.,* 2017). This could explain the lack of significant differences between the three land uses in the present study since the method of tillage used, manual tillage, was much less intensive than conventional tillage.

The highest average phosphatase activity, 6.12 µmol pNP / g of soil / h, was observed in the forest soil, whereas the peanut field soil showed the lowest average value, 4.14 µmol pNP / g of soil / h (Fig 3.8), showing a significant difference between these two land uses. Phosphatases are responsible for the mineralization of organic phosphorus into phosphate, making it bioavailable, allowing it to be used by roots or soil microbes. This may show that when the native forest was converted into peanut culture, phosphatase activity, mineralization of organic phosphorus, decreased. Therefore, phosphorus was less bioavailable, there was less phosphate available for the plants. These findings are in accordance with some of the previous literature. It has been shown that phosphatase activity is higher in forests than in managed or disturbed land. When comparing forests and agricultural land, the highest phosphatase activities were observed in the forests. This higher phosphatase activity seems to be related to a higher availability of organic matter and nutrients such as nitrogen. This is because phosphatase is a protein,

and nitrogen is needed for protein synthesis. Forests are rich in organic matter, including organic forms of nitrogen and phosphorus, that can be used by soil microorganisms (Margalef *et al.*, 2017). Phosphatase activity showed significant differences between the land uses, but *phoD* gene abundance, and therefore, phosphatase ALP abundance, did not, perhaps because in the soil there were also other phosphatases, not only ALP, whose activities were in fact affected by land use change. A more likely possibility is that, even though the primers used to amplify and quantify the *phoD* gene were degenerate, and, therefore, able to amplify a wider range of *phoD* gene sequences, they were not able to amplify the total diversity of *phoD* gene sequences. So, since both the *pmoA* gene and phosphatase activity showed significant differences between the land uses, it would seem that both the carbon cycle and the phosphorus cycle, respectively, and the microorganisms in them involved, were significantly affected by the changes the native tropical forest suffers to become agricultural profitable plantations. In this case, by becoming cashew and peanut cultures.

Even though there was little to no use of synthetic fertilizers and mechanical tillage when turning the native tropical forest into cashew and peanut cultures, the manual tillage and initial slash and burn were enough to significantly affect the methanotrophs involved in the carbon cycle and the phosphatase activity involved in the phosphorus cycle. Besides, peanut is an annual crop and its cultivation involves a periodic preparation of soil by slash and burn, which would create an imbalance of soil microbiome between peanut planting and adjacent areas outside peanut plantation, that would need to be assessed. Also, cashew is a perennial crop and is cultivated with no soil agricultural practices since cashew orchard establishment. However, from this work it can be said that the conversion of the native tropical forest in Guinea-Bissau into peanut and cashew fields significantly affected the soil microbiome involved in the nutrient cycles of carbon and phosphorus.

However, these techniques did not cause significant alterations to the total amount of bacteria and fungi present in the soil microbiome of the tropical native forest of Guinea-Bissau.

3.4. Future Perspectives

In the future, more extensive studies on the presence or absence of the *β-glu* and *pcaH* genes in tropical soils are needed. We were expecting to be able to easily amplify these genes and we could not. Perhaps there is a need for the design of new primer pairs for these genes.

In this study it was observed that the total quantity of bacteria and fungi in the soil microbiome remained relatively stable following the conversion of native tropical forest into peanut and cashew fields. However, this observation does not provide insights into potential changes in the overall diversity of the soil microbiome or shifts in specific microbial species. Further research is essential to explore these aspects in greater detail.

4.Bibliography

Adetunji, A. T., Lewu, F. B., Mulidzi, R., & Ncube, B. (2017). The biological activities of β-glucosidase, phosphatase and urease as soil quality indicators: a review. *Journal of Soil Science and Plant Nutrition*, *17*(3), 794–807.

Alves, K. J., Pylro, V. S., Nakayama, C. R., Vital, V. G., Taketani, R. G., Santos, D. G., Rodrigues, J. L. M., Tsai, S. M., & Andreote, F. D. (2022). Methanogenic communities and methane emissions from enrichments of Brazilian Amazonia soils under land-use change. *Microbiological Research*, *265*. https://doi.org/10.1016/j.micres.2022.127178

Azene, B., Zhu, R., Pan, K., Sun, X., Nigussie, Y., Gruba, P., Raza, A., Guadie, A., Wu, X., & Zhang, L. (2023). Land use change alters phosphatase enzyme activity and phosphatase-harboring microbial abundance in the subalpine ecosystem of southeastern Qinghai-Tibet Plateau, China. *Ecological Indicators*, *153*. https://doi.org/10.1016/j.ecolind.2023.110416

Banwart S.A., Nikolaidis N.P., Zhu Y.G., Peacock C.L. & Sparks D.L. (2019). Soil Functions: Connecting Earth's Critical Zone. *Annual Review of Earth and Planetary Sciences, 47*, 333-359.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., … Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, *8*(1), 103. https://doi.org/10.1186/s40168-020-00875-0

Bergkemper, F., Kublik, S., Lang, F., Krüger, J., Vestergaard, G., Schloter, M., & Schulz, S. (2016). Novel oligonucleotide primers reveal a high diversity of microbes which drive phosphorous turnover in soil. *Journal of Microbiological Methods*, *125*, 91–97. https://doi.org/10.1016/j.mimet.2016.04.011

Buchan, A., Neidle, E. L., & Moran, M. A. (2001). Diversity of the Ring-Cleaving Dioxygenase Gene pcaH in a Salt Marsh Bacterial Community. *Applied and Environmental Microbiology*, *67*(12). https://doi.org/10.1128/AEM.67.12.5801-5809.2001

Cairns, J. R. K., & Esen, A. (2010). β-Glucosidases. *Cellular and Molecular Life Sciences, 67*(20). https://doi.org/10.1007/s00018-010-0399-2

Cañizares, R., Benitez, E., & Ogunseitan, O. A. (2011). Molecular analyses of β-glucosidase diversity and function in soil. *European Journal of Soil Biology*, *47*(1), 1–8. https://doi.org/10.1016/j.ejsobi.2010.11.002

Carvalho, T. S. C. de, Jesus, E. da C., Barlow, J., Gardner, T. A., Soares, I. C., Tiedje, J. M., & Moreira, F. M. de S. (2016). *Land Use Intensification in the Humid Tropics Increased Both Alpha and Beta Diversity of Soil Bacteria. Ecology*. *97*(10), 2760–2771.

Catarino, L., Menezes, Y., & Sardinha, R. (2015). Cashew cultivation in Guinea-Bissau – risks and challenges of the success of a cash crop. *Scientia Agricola*, *72*(5), 459–467. https://doi.org/10.1590/0103-9016-2014-0369

Chen, Y., Zhang, Y., Shi, X., Xu, L., Zhang, L., & Zhang, L. (2022). The succession of GH6 cellulase-producing microbial communities and temporal profile of GH6 gene abundance during vermicomposting of maize stover and cow dung. *Bioresource Technology*, *344*, 126242. https://doi.org/10.1016/J.BIORTECH.2021.126242

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B.,

Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C. (2014). Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, *Biogeosciences, 11,* 3547–3602, https://doi.org/10.5194/bg-11-3547-2014

Costello, A. M., & Lidstrom, M. E. (1999). Molecular characterization of functional and phylogenetic genes from natural populations of methanotrophs in lake sediments. *Applied and Environmental Microbiology*, *65*(11), 5066–5074. https://doi.org/10.1128/aem.65.11.5066-5074.1999

Crowther, T. W., Maynard, D. S., Leff, J. W., Oldfield, E. E., McCulley, R. L., Fierer, N., & Bradford, M. A. (2014). Predicting the responsiveness of soil biodiversity to deforestation: a cross-biome study. *Global Change Biology*, *20*(9), 2983–2994. https://doi.org/10.1111/gcb.12565

Debrah, S. K., & Waliyar, F. (1996). Groundnut Production and Utilization in Africa; Past Trends; Projections and Opportunities for Increased Production. *The 5th Regional Groundnut Workshop for West Africa*.

Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar, D., Berdugo, M., Campbell, C. D., & Singh, B. K. (2016). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nature Communications*, *7*(1), 10541. https://doi.org/10.1038/ncomms10541

Delgado-Baquerizo, M., Giaramida, L., Reich, P.B., Khachane, A.N., Hamonts, K., Edwards, C., Singh, B.K., et al. (2016). Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *Journal of Ecology, 104*(4), 936-946

Deng, S. P., & Tabatabai, M. A. (1994). Cellulase activity of soils. *Soil Biology and Biochemistry*, *26*(10), 1347–1354. https://doi.org/10.1016/0038-0717(94)90216-X

Didagbe, O. Y., Houngnandan, P., Dedehouanou, H., Sina, H., & Bello, D. O. (2015). Characterization of the peanut production systems in their main agroecological regions in Benin. *European Scientific Journal*, *11*(33).

Eder, S., Shi, L., Jensen, K., Yamane, K., Hulett, F.M., (1996). A Bacillus subtilis secreted phosphodiesterase/alkaline phosphatase is the product of a Pho regulon gene, phoD. *Microbiology* 142, 2041–2047.

El Azhari, N., Bru, D., Sarr, A., & Martin-Laurent, F. (2008). Estimation of the density of the protocatechuate-degrading bacterial community in soil by real-time PCR. *European Journal of Soil Science*, *59*(4), 665–673. https://doi.org/10.1111/j.1365-2389.2008.01029.x

FAO and ICRISAT. (2019). Climate-Smart Agriculture in Guinea-Bissau. CSA Country Profiles for Africa Series. International Center for Tropical Agriculture (CIAT); International Crops Research Institute for the Semi-Arid Tropics (ICRISAT); Food and Agriculture Organization of the United Nations (FAO). Rome, Italy. 23p.

FAO, ITPS, GSBI, CBD and EC. (2020). State of knowledge of soil biodiversity - Status, challenges and potentialities, Report 2020. Rome, FAO. https://doi.org/10.4060/cb1928en

FAO, ITPS, GSBI, SCBD and EC. (2020). State of knowledge of soil biodiversity – Status, challenges and potentialities, Summary for policy makers. Rome, FAO. https://doi.org/10.4060/cb1929en

Ferris, H., & Tuomisto, H. (2015). Unearthing the role of biological diversity in soil health. *Soil Biology and Biochemistry*, *85*, 101–109. https://doi.org/10.1016/j.soilbio.2015.02.037

Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, *15*, 579.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3). https://doi.org/10.1073/pnas.0507535103

França, F. M., Frazão, F. S., Korasaki, V., Louzada, J., & Barlow, J. (2017). Identifying thresholds of logging intensity on dung beetle communities to improve the sustainable management of Amazonian tropical forests. *Biological Conservation*, *216*, 115–122. https://doi.org/10.1016/j.biocon.2017.10.014

Franco, A. L. C., Sobral, B. W., Silva, A. L. C., & Wall, D. H. (2019). Amazonian deforestation and soil biodiversity. *Conservation Biology*, *33*(3), 590–600. https://doi.org/10.1111/cobi.13234

Fraser, T. D., Lynch, D. H., Bent, E., Entz, M. H., & Dunfield, K. E. (2015). Soil bacterial phoD gene abundance and expression in response to applied phosphorus and long-term management. *Soil Biology and Biochemistry*, *88*, 137–147. https://doi.org/10.1016/J.SOILBIO.2015.04.014

Garcia-Montiel, D. C., Neill, C., Melillo, J., Thomas, S., Steudler, P. A., & Cerri, C. C. (2000). Soil Phosphorus Transformations Following Forest Clearing for Pasture in the Brazilian Amazon. *Soil Science Society of America Journal*, *64*(5). https://doi.org/10.2136/sssaj2000.6451792x

Geisen, S., Briones, M. J. I., Gan, H., Behan-Pelletier, V. M., Friman, V. P., de Groot, G. A., Hannula, S. E., Lindo, Z., Philippot, L., Tiunov, A. V., & Wall, D. H. (2019). A methodological framework to embrace soil biodiversity. *Soil Biology and Biochemistry,136* https://doi.org/10.1016/j.soilbio.2019.107536

Gilyarov, M. (1949). *Characteristic Features of soil as a habitat and its significance in the evolution of insects*.

Goswami, M., Bhattacharyya, P., Murkherjee, I., & Tribedi, P. (2017). Functional diversity: an important measure of ecosystem functioning. *Adv Microbiol*, *7*, 82–93.

Gougoulias, C., Clark, J. M., & Shaw, L. J. (2014). The role of soil microbes in the global carbon cycle: Tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems. *Journal of the Science of Food and Agriculture* ,94 (12), 2362–2371 https://doi.org/10.1002/jsfa.6577

Grabarse, W., Mahlert, F., Duin, E. C., Goubeaud, M., Shima, S., Thauer, R. K., Lamzin, V., & Ermler, U. (2001). On the mechanism of biological methane formation: Structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding. *Journal of Molecular Biology*, *309*(1). https://doi.org/10.1006/jmbi.2001.4647

Guo, L. B., & Gifford, R. M. (2002). Soil carbon stocks and land use change: A meta analysis. *Global Change Biology*, *8*(4). https://doi.org/10.1046/j.1354-1013.2002.00486.x

Hui, D., Mayes, M. A., & Wang, G. (2013). Kinetic parameters of phosphatase: A quantitative synthesis. *Soil Biology and Biochemistry*, *65*. https://doi.org/10.1016/j.soilbio.2013.05.017

Imhoff, J. (2016). New dimensions in microbial ecology—functional genes in studies to unravel the biodiversity and role of functional microbial groups in the environment. *Microorganisms*, *4*(2), 19. https://doi.org/10.3390/microorganisms4020019

IPCC, (2007). Fourth Assessment Report (AR4). Working Group II Report "Impacts, Adaptation and Vulnerability" - Chapter 9: Africa. IPCC, Cambridge University Press, Cambridge, UK pp. 433-46

Jones, A., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Dewitte, O., Gallali, T., Hallet, S., Jones, R., Kilasara, M., Le Roux, P., Micheli, E., Montanarella, L., Spaargaren, O., Thiombiano, L., Van Ranst, E., Yemefack, M., & Zougmoré, R. (2013). *Soil Atlas of Africa*. https://doi.org/10.2788/52319

Kendzior, J., Raffa, D. W., & Bogdanski, A. (2022). *A review of the impacts of crop production on the soil microbiome: Innovations and policy recommendations to address environmental degradation, climate change and human health*.

Khan, A. M., & Bhadauria, S. (2019). Molecular characterization of keratin degrading fungi isolated from semi-arid soil by PCR using ITS4 and ITS5 primers. *Journal of King Saud University - Science*, *31*(4), 1418–1423. https://doi.org/10.1016/j.jksus.2018.04.014

Kou, Y., Li, J., Wang, Y., Li, C., Tu, B., Yao, M., & Li, X. (2017). Scale-dependent key drivers controlling methane oxidation potential in Chinese grassland soils. *Soil Biology and Biochemistry*, *111*, 104–114. https://doi.org/10.1016/j.soilbio.2017.04.005

Lammel, D. R., Feigl, B. J., Cerri, C. C., & Nüsslein, K. (2015). Specific microbial gene abundances and soil parameters contribute to C, N, and greenhouse gas process rates after land use change in southern Amazonian soils. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.01057

Laurance, W., Sayer, J., & Cassman, K. (2014). Agricultural Expansion and Its Impacts on Tropical Nature. *Trends in Ecology and Evolution*, *29*(2), 107–116.

Lindsay, M. R., Anderson, C., Fox, N., Scofield, G., Allen, J., Anderson, E., Bueter, L., Poudel, S., Sutherland, K., Munson-McGee, J. H., Van Nostrand, J. D., Zhou, J., Spear, J. R., Baxter, B. K., Lageson, D. R., & Boyd, E. S. (2016). Microbialite response to an anthropogenic salinity gradient in Great Salt Lake, Utah. *Geobiology*, *15*(1), 131–145. https://doi.org/10.1111/gbi.12201

Luo, Y., Keenan, T. F., & Smith, M. (2015). Predictability of the terrestrial carbon cycle. *Global Change Biology*, *21*(5), 1737–1751. https://doi.org/10.1111/gcb.12766

Margalef, O., Sardans, J., Fernández-Martínez, M., Molowny-Horas, R., Janssens, I. A., Ciais, P., Goll, D., Richter, A., Obersteiner, M., Asensio, D., & Peñuelas, J. (2017). Global patterns of phosphatase activity in natural soils. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-01418-8

Mendes IC, Kappes C, Ono FB, Sousa DMG, Reis-Junior FB, Lopes AAC, Semler TD, Zancanaro L. (2015). Qualidade biológica do solo: por que e como avaliar. *Rondonópolis: Boletim de Pesquisa da Fundação MT, 2017*, 98-105.

Merlin, C., Besaury, L., Niepceron, M., Mchergui, C., Riah, W., Bureau, F., Gattin, I., & Bodilis, J. (2014). Real-time PCR for quantification in soil of glycoside hydrolase family 6 cellulase genes. *Letters in Applied Microbiology*, *59*(3), 284–291. https://doi.org/10.1111/lam.12273

Minasny, B., & Hartemink, A. E. (2011). Predicting soil properties in the Tropics. *Earth-Science Reviews*, *106*(1–2), 52–62. https://doi.org/10.1016/j.earscirev.2011.01.005

Mueller, R. C., Rodrigues, J. L. M., Nüsslein, K., & Bohannan, B. J. M. (2016). Land use change in the Amazon rain forest favours generalist fungi. *Functional Ecology*, *30*(11), 1845–1853. https://doi.org/10.1111/1365-2435.12651

Murty, D., Kirschbaum, M. U. F., Mcmurtrie, R. E., & Mcgilvray, H. (2002). Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. *Global Change Biology*, *8*(2). https://doi.org/10.1046/j.1354-1013.2001.00459.x

Muyzer, G., de Waal, E. C., & Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rrna. *Applied and Environmental Microbiology*, *59*(3), 695–700. https://doi.org/10.1128/aem.59.3.695-700.1993

Neal, A. L., Rossmann, M., Brearley, C., Akkari, E., Guyomar, C., Clark, I. M., Allen, E., & Hirsch, P. R. (2017). Land-use influences phosphatase gene microdiversity in soils. *Environmental Microbiology*, *19*(7). https://doi.org/10.1111/1462-2920.13778

Neill, C., Piccolo, M. C., Cerri, C. C., Steudler, P. A., Melillo, J. M., & Brito, M. (1997). Net nitrogen mineralization and net nitrification rates in soils following deforestation for pasture across the southwestern Brazilian Amazon Basin landscape. *Oecologia*, *110*(2). https://doi.org/10.1007/s004420050157

Petchey, O. L., & Gaston, K. J. (2006). Functional diversity: back to basics and looking forward. *Ecology Letters*, *9*(6), 741–758. https://doi.org/10.1111/j.1461-0248.2006.00924.x

QIAGEN. (2021). DNeasy® PowerSoil® Pro Kit Handbook. Retrieved from https://www.qiagen.com

QIAGEN. (2024). *QIAquick Gel Extraction Kit*. Retrieved November 25, 2024, from https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/qiaquick-gel-extraction-kit

QIAGEN. (2020). *QIAquick® Spin Handbook*. Retrieved from https://www.qiagen.com

Qu, X., Li, X., Bardgett, R. D., Kuzyakov, Y., Revillini, D., Sonne, C., Xia, C., Ruan, H., Liu, Y., Cao, F., Reich, P. B., & Delgado-Baquerizo, M. (2024). Deforestation impacts soil biodiversity and ecosystem services worldwide. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(13), e2318475121. https://doi.org/10.1073/PNAS.2318475121/SUPPL_FILE/PNAS.2318475121.SD01.XLSX

Revell L.E., Tummon F., Salawitch R.J., Stenke A., Peter T. (2015). The changing ozone depletion potential of N_2O in a future climate. *Geophysical research letters, 42*(22), 10047-10055.https://doi.org/10.1002/2015GL065702

Raiesi, F., & Beheshti, A. (2015). Microbiological indicators of soil quality and degradation following conversion of native forests to continuous croplands. *Ecological Indicators*, *50*, 173–185. https://doi.org/10.1016/j.ecolind.2014.11.008

Ravishankara, A. R., Daniel, J. S., & Portmann, R. W. (2009). Nitrous Oxide (N_2O): The Dominant Ozone-Depleting Substance Emitted in the 21st Century. *Science*, *326*(5949), 123–125. https://doi.org/10.1126/science.1176985

Rodrigues, J.L.M., Pellizari, V.H., Mueller, R., Baek, K., Jesus, E. de C., Paula, F.S., Mirza, B., et al. (2013). Conversion of the Amazon Rainforest to Agriculture Results in Biotic Homogenization of Soil Bacterial Communities. *Proceedings of the National Academy of Sciences, 110*(3), 988–93. https://doi.org/10.1073/pnas.1220608110

Ruttenberg, K. C. (2001). Phosphorus Cycle. *Encyclopedia of Ocean Sciences: Second Edition*, 401–412. https://doi.org/10.1016/B978-012374473-9.00277-0

Sakurai, M., Wasaki, J., Tomizawa, Y., Shinano, T., & Osaki, M. (2008). Analysis of bacterial communities on alkaline phosphatase genes in soil supplied with organic matter. *Soil Science and Plant Nutrition*, *54*(1), 62–71. https://doi.org/10.1111/j.1747-0765.2007.00210.x

Schimel, J. P., & Schaeffer, S. M. (2012). Microbial control over carbon cycling in soil. *Frontiers in Microbiology, 3*(SEP). https://doi.org/10.3389/fmicb.2012.00348

Singh, A. K., Jiang, X. J., Yang, B., Wu, J., Rai, A., Chen, C., Ahirwal, J., Wang, P., Liu, W., & Singh, N. (2020). Biological indicators affected by land use change, soil resource availability and seasonality in dry tropics. *Ecological Indicators*, *115*. https://doi.org/10.1016/j.ecolind.2020.106369

Singh, B. K., Bardgett, R. D., Smith, P., & Reay, D. S. (2010). Microorganisms and climate change: terrestrial feedbacks and mitigation options. *Nature Reviews Microbiology*, *8*(11), 779–790. https://doi.org/10.1038/nrmicro2439

Smits, T. H. M., Devenoges, C., Szynalski, K., Maillard, J., & Holliger, C. (2004). Development of a real-time PCR method for quantification of the three genera Dehalobacter, dehalococcoides, and Desulfitobacterium in microbial communities. *Journal of Microbiological Methods*, *57*(3), 369–378. https://doi.org/10.1016/j.mimet.2004.02.003

SOP: Enzyme assays (pNP). 2021. Soils Lab, University of Illinois Urbana-Champaign. Urbana, IL. Accessed at: https://margenot.cropsciences.illinois.edu/methods-sops/

Sombroek, W. G., Nachtergaele, F. O., & Hebel, A. (1993). Amounts, dynamics and sequestering of carbon in tropical and subtropical soils. *Ambio*, *22*, 417–426.

Spohn, M., & Kuzyakov, Y. (2013). Distribution of microbial- and root-derived phosphatase activities in the rhizosphere depending on P availability and C allocation - Coupling soil zymography with 14C imaging. *Soil Biology and Biochemistry*, *67*. https://doi.org/10.1016/j.soilbio.2013.08.015

Steinberg, L. M., & Regan, J. M. (2009). mcrA-targeted real-time quantitative PCR method to examine methanogen communities. *Applied and Environmental Microbiology*, *75*(13), 4435–4442. https://doi.org/10.1128/AEM.02858-08

Swift, M. J., Heal, O. W., & Anderson, J. M. (1979). *Decomposition in terrestrial ecosystems*. University of California Press.

Thermo Fisher Scientific. (2013). *One Shot™ TOP10 Chemically Competent E. coli Manual* (Publication No. MAN0000950). Retrieved November 25, 2024, from https://assets.thermofisher.com/TFS-Assets/LSG/manuals/oneshottop10_man.pdf

Thermo Fisher Scientific. (2014). *GeneJET Plasmid Miniprep Kit User Guide* (Publication No. MAN0012655). Retrieved November 25, 2024, from https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012655_GeneJET_Plasmid_Miniprep_UG.pdf

Thermo Fisher Scientific. (2015). *CloneJET PCR Cloning Kit*. Retrieved from https://www.thermofisher.com

Thermo Fisher Scientific. (2015). *CloneJET PCR Cloning Kit User Guide* (Publication No. MAN0012966). Retrieved November 25, 2024, from https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012966_CloneJET_PCR_Cloning_40rxn_UG.pdf

Thermo Fisher Scientific. (2024). *CloneJET PCR Cloning Kit* (Cat. No. K0503). Retrieved November 25, 2024, from https://www.thermofisher.com/order/catalog/product/K0503

Thiele-Bruhn, S., Schloter, M., Wilke, B.-M., Beaudette, L. A., Martin-Laurent, F., Cheviron, N., Mougin, C., & Römbke, J. (2020). Identification of new microbial functional standards for Soil Quality Assessment. *SOIL*, *6*(1), 17–34. https://doi.org/10.5194/soil-6-17-2020

Tiwari, S., Singh, C., & Singh, J. S. (2018). Land use changes: a key ecological driver regulating methanotrophs abundance in upland soils. *Energy, Ecology and Environment*, *3*(6), 355–371. https://doi.org/10.1007/s40974-018-0103-1

Trivedi, P., Anderson, I. C., & Singh, B. K. (2013). Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for global prediction. *Trends in Microbiology*, *21*(12), 641–651.

Vendramini, J. M. B., Silveira, M. L. A., Dubeux Jr., J. C. B., & Sollenberger, L. E. (2007). Environmental impacts and nutrient recycling on pastures grazed by cattle. *Revista Brasileira de Zootecnia*, *36*(suppl), 139–149. https://doi.org/10.1590/s1516-35982007001000015

Venturini, A. M., Nakamura, F. M., Gontijo, J. B., da França, A. G., Yoshiura, C. A., Mandro, J. A., & Tsai, S. M. (2020). Robust DNA protocols for tropical soils. *Heliyon*, *6*(5). https://doi.org/10.1016/j.heliyon.2020.e03830

Wagg, C., Bender, S. F., Widmer, F., & van der Heijden, M. G. A. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences*, *111*(14), 5266–5270.White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols*, 315–322. https://doi.org/10.1016/b978-0-12-372180-8.50042-1

Williams, C. A., Hanan, N. P., Neff, J. C., Scholes, R. J., Berry, J. A., Denning, A. S., & Baker, D. F. (2007). Africa and the global carbon cycle. *Carbon Balance and Management, 2*(1). https://doi.org/10.1186/1750-0680-2-3

Zang, X., Liu, M., Fan, Y., Xu, J., Xu, X., & Li, H. (2018). The structural and functional contributions of β-glucosidase-producing microbial communities to cellulose degradation in composting. *Biotechnology for Biofuels*, *11*(1). https://doi.org/10.1186/s13068-018-1045-8

5. Appendix

Supplementary Table 5.1 Plasmid DNA sequence from each gene, extracted by miniprep.

Target Gene with Primer Forward	Plasmid DNA Sequence
bacterial 16S rRNA with pJET1.2 Forward	AAACGGGGGGTTGGGGGNCGGAAAGTCGAGGCTGAGTTTTTAGCAAGATGGGGGAATATTGC ACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGAAGAAGGCCTTCGGGTTGTAAAG TACTTTCAGCGGGGAGGAAGGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAA GAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATATCTTTCTAGAAGATCTCCTACAATA TTCTCAGCTGCCATGGAAAATCGATGTTCTTCTTTTATTCTCTCAAGATTTTCAGGCTGTATATTA AAACTTATATTAAGAACTATGCTAACCACCTCATCAGGAACCGTTGTAGGTGGCGTGGGTTTTC TTGGCAATCGACTCTCATGAAAACTACGAGCTAAATATTCAATATGTTCCTCTTGACCAACTTTA TTCTGCATTTTTTTTGAACGAGGTTTAGAGCAAGCTTCAGGAAACTGAGACAGGAATTTTATTA AAAATTTAAATTTTGAAGAAAGTTCAGGGTTAATAGCATCCATTTTTTGCTTTGCAAGTTCCTCA GCATTCTTAACAAAAGACGTCTCTGTTGACATGTTTAAAGTTTAAACCTCCTGTGTGAAATTATT ATCCGCTCATAATTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCAATTGCTTTCCAGTCGGGAAAC CTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGAGGAGAGGCGGTTTGCGTATTGGG CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAAC ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT CCATAGGCTCCGCCCCCCTGACAAGCATCACAAAAATCGACGCTAAAGTCAGAGGTGGCGAA ACCCGACGGGACTATAAAGATCCCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG TTCCAACCCTGCCCCTTACCGAAAACCTTTCCCCCTTTCTCCTTCGGGAAGCGTGGCGCTTTTT NAAAGCTAAG
fungal ITS with pJET1.2 Forward	NNNNNGNAGNNCNNNNNNNNNNNNNNNTCGAGTTTTTCAGCAAGATCCTCCGCTTATTGAT ATGCTTAAGTTCAGCGGGTATCCCTACCTGATTCGAGGTCAACCTTTGGTATAATGGCTTCTGGC AGGCTCCGTCCGGGCTCCTAGAGCGAGAAGTATCTACTACGCTCGGGGCCGTCCGACACCGCC ACTGATTTTGAAGCGCGCCGTGGGCCGGCGGCGCTCAACTCCAAGCTGAGTGCTTGAGGGCT ATAATGACGCTCGAACAGGCATACCCTCCGGAATACCAGAGGGTGCAATGTGCGTTCAAAGAT TCGATGATTCACTGAATTCTGCAATTCACATTACTTATCGCATTTCGCTGCGTTCTTCATCGATGC CAGAACCAAGAGATCCGTTGTTGAAAGTTTTTAACTATTTAATTTTGCTCAGACAACACTAACT TTACGAGTTGTGGTCCTCTGGCAGACGCTGGCCAGCCGGAGCCGGCGGCCCCAAAGGCGGGC CTGCCAAAGCAACAGAGTATAATAGACAGCGGGTTGGAGATCCGCCCGAAGGCGAGTCCTCT GTAATGATCCTTCCGCACGTTCACCTACGGAAACCTTGTTACGACTTTTACTTCCATCTTTCTAG AAGATCTCCTACAATATTCTCACCTGCCAGGGGGGAGGGGGGNCGTAGGAAAGAAGTGGAAC AAAACTTTGCAGCTTGTTCGTG
GH6 cellulase with pJET1.2 Forward	CAGCAAGATGAGGGAGTCGGGCTCGATGTAGGTGATGATCTTGAGGTTCGCGTACTTCGAGTC GCTCATGATCGACGCGATGGGGTCGATGTACTCGGTCTCGTAGCGGCTGAGATCCCCCGCGCC GAGCTCGCCGTTGGAGGCCAGGGCGGCGCAGTCGCGGCCGGGCAGATCTTTCTAGAAGATCT CCTACAATATTCTCAGCTGCCATGGAAAATCGATGTTCTTCTTTTATTCTCTCAAGATTTTCAGG CTGTATATTAAAACTTATATTAAGAACTATGCTAACCACCTCATCAGGAACCGTTGTAGGTGGCG TGGGTTTTCTTGGCAATCGACTCTCATGAAAACTACGAGCTAAATATTCAATATGTTCCTCTTGA CCAACTTTATTCTGCATTTTTTTTGAACGAGGTTTAGAGCAAGCTTCAGGAAACTGAGACAGG AATTTTATTAAAAATTTAAATTTTGAAGAAAGTTCAGGGTTAATAGCATCCATTTTTTGCTTTGC AAGTTCCTCAGCATTCTTAACAAAAGACGTCTCTTTTGACATGTTTAAAGTTTAAACCTCCTGT GTGAAATTATTATCCGCTCATAATTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAAGCCT GGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCAATTGCTTTCCA GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGAGAGGCGGTTT GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGG CGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGC TGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC GCTCTCCTGTTCGAACCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAACGTG GCGCTTTCTCATAACTCACGCTGTAGGGATCTCAGTTCGGGGAAGGCGTTCCCTCCAACTGGG CTGGTGGCCAAACCCCCCGTTCAGCCGAACCGCTGGCCCTTATCCGGGAACTTTCGCTTGAAT CCAACCCGGTA
pmoA with pJET1.2 Forward	TTCAGCAAGATCCGGAGCAACGTCTTTACCGAAGGTGCGCATCGTGCCGCGCTCGATGATCCG GATGTATTCGGGCATTGAGGTGCGGACATAGTGATAGCCGATGAGATCGGCAACCGTCATCAGC TGGCCATATTGCTCGCTCGGCTGATGGAACTGCGCCAGGATGACCCAATTATTCGGATACATCA ACAGCGAGAAGCCCATCGAACCGACGATCATCGTGATGATCCAGCTCTTGGAGAGCAGCAGC ACGATATCAAGGAACAACGCCTGCGGAATGAGCGAGGTTGGCCAAATGAGGTTGACGGGATA GTAGGTCCAGCCCCAGAAGTTGATGTAGCGGTCCAGCCATTCGCCGACGAGCAGCGCCAGGC ACAGGAAGGTCGCGCCAAACGGCATCCGGTAGTGGTTCCAGAAGAAGTATTGGGCGGCGGCC

	GGGAAGGTCACGAGCAGGATCGGCACAATGGTTGGCCAAAGGCGGCGGTCTTTCCAGTCGAC CCAGAAGTCCCAGTCACCATCTTTCTAGAAGATCTCCTACAATATTCTCAGCTGCCATGGAAAA TCGATGTTCTTCTTTTATTCTCTCAAGATTTTCAGGCTGTATATTAAAACTTATATTAAGAACTAT GCTAACCACCTCATCAGGAACCGTTGTAGGTGGCGTGGGTTTTCTTGGCAATCGACTCTCATGA AAACTACGAGCTAAATATTCAATATGTTCCTCTTGACCAACTTTATTCTGCATTTTTTTTGAACG AGGTTTAGAGCAAGCTTCAGGAAACTGAGACAGGAATTTTATTAAAAATTTAAATTTTGAAGA AAGTTCAGGGTTAATAGCATCCATTTTTTGCTTTGCAAGTTCCTCAGCATTCTTAACAAAAGAC GTCTCTTTTGACATGTTTAAAGTTTAAACCTCCTGTGTGAAATTATTATCCGCTCATAATTCCAC ACATTATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA CATTAATTGCGTTGCGCTCACTGCCAATTGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCA TTAATGAATCGGCCAACGCGCGGGGAAAAGGCGGTTTGCGTTTTGGGCGCTCT
mcrA with pJET1.2 Forward	NNNNNNNNGNAGNNNNNNNNNNCNCGANTTTTTCGCNAGATGATGACATCCTGGACGACTA CTGCTACTGGGCCGCTGACCTGATCAAGTCCAAGTACGGTGGGCTGTGCAAGAGCAAGCCATC CATGGAACTCATGGAGAAGCTGGGCACTGAGGTCGCTTCCTACGCTCTCGAGATGTACGAGAG GTACCCCGCCGCTATGGAAACCCACTTCGGTGGATCCCAGAGGGCAACGGTATCTGCAGCAGC GACCGGTATCGCTTGCGCTATGGCAACCGGAAACGCCGACTTCGGTGTCAACGGCTGGTACCA GTGCATGCTCTACCACAAGGAAAGGACCGGCAGACTCGGCTTCTACGGTTACGACCTGCAGG ACCAGTGCGGTTCCTCCAACTCCTTCGCATACAGGAGCGACGAGGGCCTGCCCATGGAACTAA GAGGCCCGAACTATCCCAACTACGCAATGAACGATCTTTCTAGAAGATCTCCTACAATATTCTC AGCTGCCATGGAAAATCGATGTTCTTCTTTTATTCTCTCAAGATTTTCAGGCTGTATATTAAAAC TTATATTAAGAACTATGCTAACCACCTCATCAGGAACCGTTGTAGGTGGCGTGGGTTTTCTTGG CAATCGACTCTCATGAAAACTACGAGCTAAATATTCAATATGTTCCTCTTGACCAACTTTATTCT GCATTTTTTTTGAACGAGGTTTAGAGCAAGCTTCAGGAAACTGAGACAGGAATTTTATTAAAA ATTTAAATTTTGAAGAAAGTTCAGGGTTAATAGCATCCATTTTTTGCTTTGCAAGTTCCTCAGCA TTCTTAACAAAAGACGTCTCTTTTGACATGTTTAAAGTTTAAACCTCCTGTGTGAAATTATTATC CGCTCATAATTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATG AGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCAATTGCTTTCCAGTCGGGAAACCTG TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGNTTTGCGTATTGGGCGC TCTTCCNCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGGGAGCGGGATCA GNTCACTCAAAGGCGGTAATACGGTTNTCCACGAATCAGGGGATACGCAGGAAGAACATGGG AGCAAAGGCCACCAAAGGCCAGAACCGTAAAAGGCCNNNTGGTGGCGTTTTCAA
phoD with pJET1.2 Forward	AAAAAAAGGGGGAGGGCTTCTNAATNCCTCTANNCGAGTTTTAGCAGATGAGGCCGATCGGC ATGTCGCTGGCAATCACCTTCCACGTCGCCTTGCTGGCCGGCAACTTCGCCTTTAGCCACGCCA ATTGTTCCGGGCCGAGGTAGGCGGCTTCGGGGCTCATCTGCGGCTGGCGGTTGGCCGTGTTCG GGCCGCGGTAGCTGCGCTGGTCCAGCACAAATACTTCCAGCGACGGCCCGTAGGGAATGGTGC GATAGACGCGCTCCGAAGCCAGGGCGTCGGGCCGGATCGGCACGTAGTCGAAAAACGCCCGC TTGGCTCGCGCGGCCAGCAGCGACACGCTCTTGACCGTGTATTTCGCGTCGTCGAGCATCTGG CCGGGGAACCAGTTGTTGCGCACCTCGTGGTCGTCCCACTGATCTTTCTAGAAGATCTCCTACA ATATTCTCAGCTGCCATGGAAAATCGATGTTCTTCTTTTATTCTCTCAAGATTTTCAGGCTGTATA TTAAAACTTATATTAAGAACTATGCTAACCACCTCATCAGGAACCGTTGTAGGTGGCGTGGGTT TTCTTGGCAATCGACTCTCATGAAAACTACGAGCTAAATATTCAATATGTTCCTCTTGACCAACT TTATTCTGCATTTTTTTTGAACGAGGTTTAGAGCAAGCTTCAGGAAACTGAGACAGGAATTTTA TTAAAAATTTAAATTTTGAAGAAAGTTCAGGGTTAATAGCATCCATTTTTTGCTTTGCAAGTTCC TCAGCATTCTTAACAAAAGACGTCTCTTTTGACATGTTTAAAGTTTAAACCTCCTGTGTGAAAT TATTATCCGCTCATAATTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTG CCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCAATTGCTTTCCAGTCGGG AAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGC GGTATCAGCTCACTCAAAGGCGGGAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA AAACTGTGAGCAAAAGGCCAGCAAAGGCCGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT TCCTAGGGTCCGCCCCCCGGACANGCTCCCAAAATCGACNCTCAGTCAGAGGGGCGAAACCC GAAGGACTTAAAAAAACAAG

Supplementary Table 5.2 Best nanodrop results (concentration in ng/µl, Abs 260/280 and 260/230 ratios) for each of the 5 sample points of each of the 3 land uses

Samples	**Concentration (ηg/µl)**	**Abs 260/280**	**Abs 260/230**
Forest 1	16.7	1.89	1.94
Forest 2	8.8	1.95	1.60
Forest 3	9.0	1.89	1.44
Forest 4	13.5	1.55	0.87
Forest 5	16.8	1.71	1.93
Peanut 1	6.5	1.66	1.40
Peanut 2	9.7	2.29	1.94
Peanut 3	8.5	1.50	1.88
Peanut 4	9.0	1.87	0.58
Peanut 5	10.2	1.85	0.66
Cashew 1	12.7	2.00	0.77
Cashew 2	11.9	2.15	0.47
Cashew 3	9.4	1.95	0.52
Cashew 4	18.6	2.33	0.98
Cashew 5	15.8	1.84	0.47