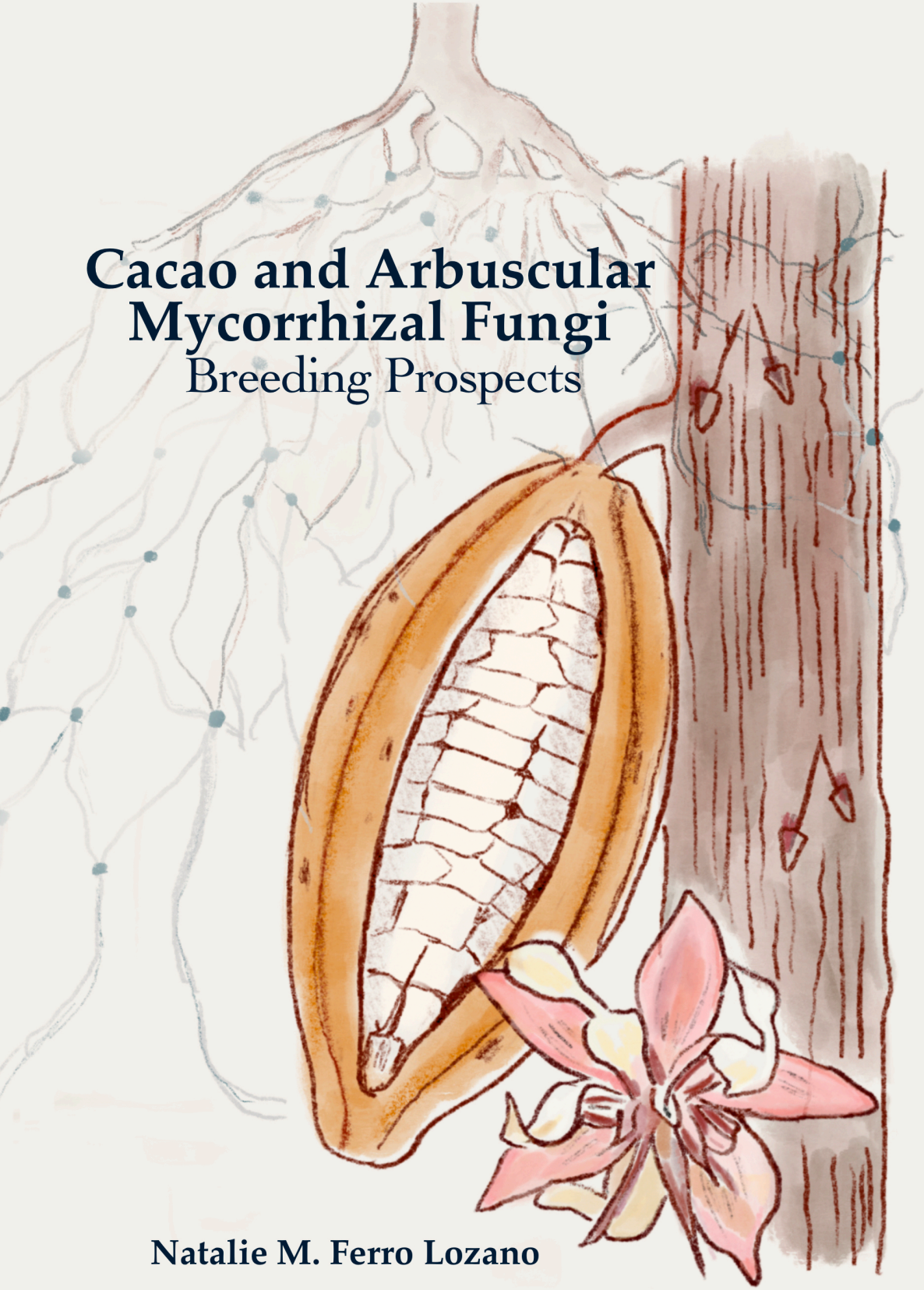
Cacao and Arbuscular Mycorrhizal Fungi

Breeding Prospects



Natalie M. Ferro Lozano

Propositions

1. Arbuscular mycorrhizal fungal community composition in cacao is more modified by scion traits than by root traits. (this thesis)

2. Breeding for mycorrhizal benefit in cacao focuses on plant selectivity for arbuscular mycorrhizal fungal community composition. (this thesis)

3. Social sciences guide the implementation of scientific discoveries in society.

4. Sustainable agricultural systems require varied locally adaptive genotypes and approaches.

5. Scientists have a duty to serve society instead of personal ambition.

6. In the current societal structure, productivity is valued over personal and family well-being.

7. Scientists and Indigenous people are the most disregarded knowledge communities on Earth.

Propositions belonging to the thesis, entitled

Cacao and Arbuscular Mycorrhizal Fungi: Breeding Prospects

Natalie M. Ferro Lozano

Wageningen, 11 May 2026

Cacao and Arbuscular Mycorrhizal Fungi: Breeding Prospects

Natalie M. Ferro Lozano

Thesis committee

Promotor

Dr *ir.* Guusje A.B. Bonnema
Associate Professor in Plant Breeding
Wageningen University & Research

Co-promotor

Prof. Dr Thomas W. Kuijper
Professor in Soil Biology
Wageningen University & Research

Dr Luis C. Mejía
Researcher
Instituto de Investigaciones Científicas y
Servicios de Alta Tecnología (INDICASAT-AIP),
Panama City, Panama

Other members
Prof. Dr Bas Zwaan, Wageningen University & Research
Dr Rafael Chan Navarrete, Olam Food Ingredients, Amsterdam
Dr Monique Weemstra, Wageningen University & Research
Prof. Dr Katie J. Field, University of Sheffield, UK

This research was conducted under the auspices of the Graduate School of Production Ecology & Resource Conservation

Cacao and Arbuscular Mycorrhizal Fungi: Breeding Prospects

Natalie M. Ferro Lozano

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr C. Kroeze,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 11 May 2026
at 10:30 a.m. in the Omnia Auditorium.

Natalie M. Ferro Lozano
Cacao and Arbuscular Mycorrhizal Fungi: Breeding Prospects
195 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2026)
With references, with summaries in English and Spanish
DOI: https://doi.org/10.18174/681165

From the love of our ancestors

To the hope of a better future

A reason to create better days

Table of Contents

Chapter 1. General Introduction 9

Chapter 2. Genotype characterization of *Theobroma cacao* germplasm from farms in Panama 37

Chapter 3. Evaluating the relative importance of cacao genetic identity, root traits, and soil properties for arbuscular mycorrhizal fungal community composition 93

Chapter 4. Arbuscular mycorrhizal fungal inocula differentially influence root traits and biomass in *Theobroma cacao* 141

Chapter 5. Discussion 171

Summary 190

Chapter 1. General Introduction



"The strength of the cacao does not come from the plant alone, nor from the fungus, nor from your hands—but from the *bond* between them. Beneath your feet, the mycelium stretches like a great river, carrying water and wisdom to the roots. In return, the cacao feeds it with sweetness. But when you strip the earth bare and drown it in potions that kill the unseen threads, the pact is broken. The blight is not a curse—it is the voice of the soil, crying out for balance."

The farmers stood in silence, the weight of the words settling upon them like rain. They had thought the land was dying because God was angry and let the demons loose. But now they understood—the land was dying because they had forgotten how to listen to it.

From that day on, they changed their ways. They planted diverse shade trees to shelter the cacao, left the soil unbroken, and fed it with fallen leaves and time. And as the mycelial threads regrew, so too did the grove—stronger than before, because now it was nurtured, not conquered.

Darwinian Evolution & Mendelian Genetics

Darwin ideated the revolutionary and yet simple concepts of adaptation, natural selection, and evolution. Each organism has a unique set of physical traits, or phenes (Pieruschka and Poorter, 2012). It was later explained by Mendel in 1866 (Abbott and Fairbanks, 2016) that traits are controlled by hereditary factors, later identified as the individual's genes. Phenes in turn define the individual's success at acquiring resources and survival in the environment, which later determine its reproductive success and thus the number of offsprings that will continue with reshuffled combinations of successful genes/phenes. The success of an individual to adapt, survive, and reproduce in its environment, is referred to as fitness. The fittest individual would then have the traits that are expected to be favourable in a particular environment and produce the highest number of offspring. So, the individual's genes interact with the natural environment to create the phenes, and these select the fittest individuals for that environment. The different phenes in a population are a representation from the genetic variation present in a population. The more genetic diversity a population has, the better chances are of this population to adapt to the changing environment. Through this process of natural selection by the environment, there is a gradual change of allelic frequency within the population between generations, that is referred to as evolution (Gregory, 2009). But that natural gradual selection by the environment, can also be done by another species instead of the environment, for instance the humans in a process called domestication.

Domestication as Ancient Plant Breeding

Domestication, selective breeding, and artificial selection are all synonyms that refer to an anthropocentric selective breeding of another species such as plants or animals to gain resources and/or services as explained by Charles Darwin in The Origin of Species (1859). This continued selection of the domesticated species results in populations with shifted mean averages of the interesting traits for the domesticator species. Crosses made by farmers and later plant breeders, have given rise to >1000 plant species that differ greatly from their wild relatives (Purugganan, 2022). Traditional plant breeding has prioritized morphological (e.g., seed or vegetative yield, colour, fruit shape) and physiological traits (e.g., plasticity and adaptation to new environments, flavour, disease resistance) for human consumption (Darwin, 1859; Driscoll et al., 2009), with most efforts focused on aboveground traits. One of the most iconic examples is maize (*Zea mays* subs. *mays*), where the domestication process formed a morphologically distinct subspecies by the hybridization of two related teosinte ancestors (*Zea mays* subs. *parviglumis & Zea mays* subs. *mexicana*; (Yang et al., 2023). Now this plant is highly successful in human-managed agrosystems, and it has expanded its cultivation area to over 200 Mha across varied agro-ecosystems and where it is used for a varied array

of uses as human and animal consumption, as well as bioindustry products like bioethanol and biodegradable packaging material among others (Saldivar and Perez-Carrillo, 2016).

Conscious selection in aboveground traits has inadvertently influenced belowground traits that suited these purposes. Despite their critical role for water/nutrient uptake and plant development (Bardgett et al., 2014; Strock and Schneider, 2022), root and rhizosphere traits remain understudied due to phenotyping challenges (Kuijken et al., 2015; Paez-Garcia et al., 2015). Few studied root traits until now include root architecture (Kell, 2011; Paez-Garcia et al., 2015), nematode/rot resistance (Boerma and Hussey, 1992), and rhizobia symbiosis in legumes (Provorov and Tikhonovich, 2003; Rengel, 2002). Root traits are especially vital in nutrient-poor soils, common in smallholder farms (Ceccarelli, 1996).

Tree crop breeding

Despite tree crops' vast diversity and local to global benefits, their long growth cycles and heterozygosity have left their potential largely untapped. There is a global estimation between 60,000 to 75,000 tree species from which 8,000 species are utilized (timber, non-timber forest products or energy) and only a fourth of these are actively managed in agroforestry systems or plantations (Beech et al., 2017; FAO, 2014; Gatti et al., 2022). Agroforestry systems and tree plantations have a significant positive impact on planetary health, food security, and climate regulation (Rosenstock et al., 2019). Nevertheless, agroforestry is not widely used and one of the challenges is breeding trees with ideal phenotypes. Tree breeding deals with long developmental times (2–5 years from seed to seed) and high heterozygosity which makes germplasm conservation difficult. Vegetative propagation methods, like grafting, partially mitigates these issues by combining a scion (donor shoot) and rootstock (root system) into a single plant (Figure 1) without genetically mixing, preserving traits clonally and halving development time. However, tree breeding remains constrained by the need to find elite material and the time required to eliminate undesirable trait linkages via backcrossing. Grafting then enables trait stacking for rootstock and scion traits separately (e.g., vigour, stress resistance) leading to more resource-efficient breeding.

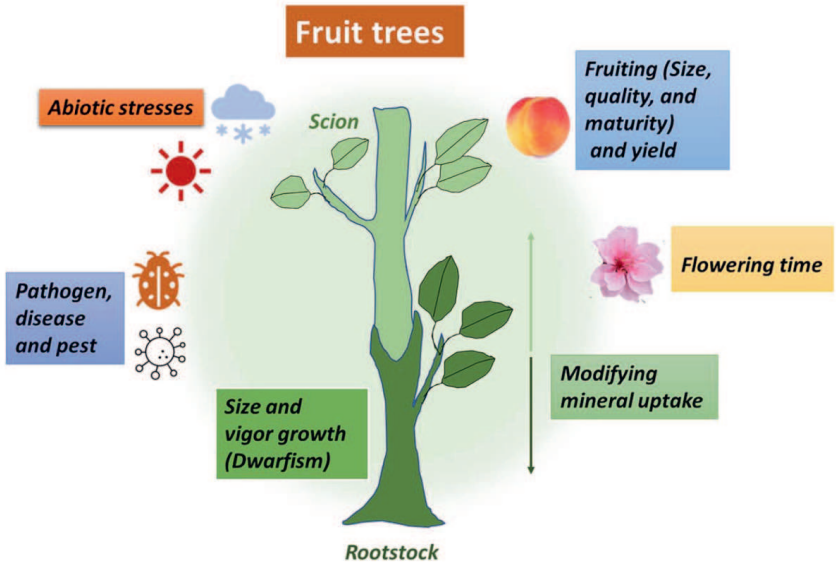


Figure 1. Desired traits for rootstocks and scions, illustrated through grafting. Note that, despite the focus on rootstocks*, they are depicted only as the lower stem—actual roots, as is often the case in publications, are curiously absent (Habibi et al., 2022).*

Interaction between Genes, Environment and Management

The selected traits (or phenes) of an individual are the result of the interaction between the genes of the individual and the environment it was grown in, and this is referred to as genotype by environment interaction (GxE) (Leon et al., 2016). Recent research highlights the additional influence of farmer management (M) to the plant phenotype from, for example, use of agrochemicals, determined crop rotations, and uses of multiple crops. This introduces a broader framework of genotype-by-environment-by-management interactions (GxExM) (Cooper et al., 2022). A greater understanding of the target environments and the management strategies used facilitates long-term sustainable plant improvement. In general, crop diversification enhances crop production and associated biodiversity which facilitates water and soil quality, and pest and disease control (Beillouin et al., 2021a). Among the diversification systems, agroforestry is the most effective at providing these ecosystems services in comparison with intercropping, cover crops, crop rotation, or mixed crops. Hence to breed for plants in an agroforestry system it is necessary to take multiple ecological factors into account, most importantly symbiotic relations.

Arbuscular Mycorrhizal Symbiosis

Life underground is driven by a complex web of organisms, from microbiota (bacteria, fungi, nematodes) to larger soil fauna (Coleman and Wall, 2015). Among these, arbuscular mycorrhizal fungi (AMF) stand out as keystone symbionts, supporting the creation of terrestrial ecosystems since the transition of plants from aquatic to land environments (Redecker et al., 2000). These fungi are

part of intimate associations with plant roots, creating a *holobiont*—a dynamic unit where host and microbes jointly influence traits like nutrient acquisition and stress resilience (Vandenkoornhuyse et al., 2015). AMF are particularly vital for plant fitness, enhancing water and nutrient uptake (Kuyper et al., 2021) while priming defences against biotic (Dowarah et al., 2022) and abiotic stresses (Qian Wang et al., 2024). Their symbiosis relies on first, the molecular dialogue which primes plant defences (Boyno et al., 2025; Jung et al., 2012) and then, the formation of specialized structures (*arbuscules*) in the cortical cells of the roots for resource exchange (Luginbuehl and Oldroyd, 2017).

Host plants have a microbial core community including a core AMF community. A core AMF community is a set of AMF species consistently found across the majority of samples within a specific habitat or host, representing the stable, foundational members of a microbial community. These core AMF communities are hypothesized to be the most ecologically and functionally relevant for that host (Neu et al., 2021). On the contrary, indicator species are species or groups that reflect the abiotic or biotic state of an environment, or indicate the biodiversity of other species, taxa, or entire communities in an area (Lawton and Gaston, 2001). The beneficial effects of mycorrhizal symbiosis (e.g., biomass, nutrient uptake, yield, or stress tolerance) provided by AMF core or indicator species can be assessed quantitatively. This measurable beneficial effect is referred to as *mycorrhizal responsiveness*. Importantly, mycorrhizal responsiveness does not always arise from AMF symbiosis in general, but from the specific interaction between a plant genotype, a particular set or individual AMF species, and the external environment. Depending on which AMF species are present in the community and the ecological strategy of the plant (Herre et al., 2005; Kiers et al., 2000), the type and degree of benefit experienced by the plant can vary substantially. Critically, AMF interactions may have played an underappreciated role in plant domestication. For the plants selected by early farmers, it is unknown how their interactions with AMF shaped their performance. Depending on the environments in which these plants were selected, the resulting genotypes may have exhibited different phenotypes under varying levels of AMF inoculum and community composition.

Agriculture and belowground microbiome

Domestication and agriculture caused a shift in the plant's relation with its associated microbiome. An important agricultural shift happened in the 1960-1970s when the need to feed an increasing population gave rise to the Green Revolution. It was characterized by three things: (1) breeding for plants, especially cereals, for increased productivity (2) under a high-input system with fertilizers, pesticides, and higher water regimes and (3) implementing these where agroclimatic systems and farmers had the best chances for achieving the maximal yield (Conway and Barbie, 1988). Each phase had signified changes in the way we grow and select the crops, which inadvertently changed the way plants relate with the soil.

Domestication and breeding can alter soil-associated microbial diversity, root traits, and performance. Reduced mycorrhizal benefits have been observed in modern cultivars across a range of crops (Hetrick et al., 1993; Li et al., 2024; Martín-Robles et al., 2018). This shift in plant-microbe relations could be for two reasons. Firstly, by breeding for increased resistance to pathogens, selected via genetic resistance, could indirectly target mutualistic partners (Toth et al., 1990). Secondly, increased selection in high-input systems for an altered root system architecture thus altering the interaction with the soil environment (Shi et al., 2024). This is exemplified in winter wheat; the root system architecture of newer cultivars has a smaller size (Fradgley et al., 2020; McGrail and McNear, 2021). Modern cotton cultivars have smaller root diameter, reducing the space for mycorrhizal colonization and increasing the need for P fertilization (Wang et al., 2023). Thus far, rhizosphere traits and their relation to microbial communities remain understudied despite their critical role on plant phenotype (Vandenkoornhuyse et al., 2015).

Root traits

Root traits have gained attention in the last years for their role in sustainable agriculture to impact plant fitness, health, and productivity (Bom et al., 2020; Saleem et al., 2018; Shi et al., 2024). As for any trait, root traits are a result from the interaction between the biotic and abiotic environment and the plant genotype. From the environment or soil perspective, research has focused on management practices that benefit positive plant-symbiont interactions or minimize adverse plant-pathogen interactions. Researchers agree that the most influential management practices that could impact mycorrhizal fungi are tillage intensity, plant field diversity, degree and type of fertilization, especially phosphorus, and potentially fungicide application. Negative impacts include high tillage intensity as it decreases AMF survival (Kabir, 2005; Säle et al., 2015), and high phosphorus fertilization as it reduces AMF diversity and abundance (Ma et al., 2023). The impact of monocultures on AMF can vary depending on the plant species. In maize, monoculture did not reduce AMF richness, suggesting a degree of plant–AMF specialization (Wachowska and Rychcik, 2023). In contrast, AMF spore viability and their ability to interact with multiple plant species were higher when grown in long-term plant mixtures compared to monocultures (Dietrich et al., 2020). Positive management practices include increased crop diversity as it increases AMF diversity (Guzman et al., 2021), and organic amendments that enhance richness of AMF communities (Liu et al., 2020). Alternatively, it has been proposed that applying mycorrhizal inoculum could increase mycorrhizal abundance. However, assessing their success is challenging, as it is difficult to monitor their persistence in the soil or their effects on native plants and microbial communities (Canfora et al., 2021; Manfredini et al., 2021; O'Callaghan et al., 2022). Nevertheless, mixtures of mycorrhizal inoculum with plant growth promoting bacteria (Gamalero et al., 2009; Noceto et al., 2021; Yu et al., 2022), organic fertilizers (Jiang et al., 2021) or vermicompost (Benaffari et al., 2022) have been shown to be beneficial in

multiple field or greenhouse trials. From the plant genotype perspective, a large part of the research focuses on breeding for root traits and their interaction with microorganisms as AMF (Hohmann and Messmer, 2017; Nerva et al., 2022; Wissuwa et al., 2009), where the root economic space framework becomes a valuable theory (Bergmann et al., 2020).

Root Economic Space Framework

The root economic space (RES) framework builds on the leaf economics spectrum proposed by Wright et al. (2004), which describes a gradient from resource-conservative to resource-acquisitive strategies. Fast-growing plant species typically exhibit low tissue density, short tissue lifespan, and high rates of resource uptake, whereas slow-growing species display the opposite traits. For roots, Bergmann et al. (2020) proposed a two-dimensional framework that incorporates both a conservation gradient and a collaboration gradient. The conservation gradient reflects the trade-off between high nitrogen content (as a proxy for resource acquisition) and high root tissue density (as a proxy for resource conservation). The collaboration gradient describes how symbiosis with AMF shapes root morphology, contrasting high specific root length (SRL) with large root diameter (D). High SRL reflects a 'do-it-yourself' strategy, where plants invest in fine roots for direct absorption, while greater root diameter provides more cortical space for AMF colonization, representing a 'collaborative' strategy. Together, these gradients help explain variation in plant growth and survival, and thus species distributions and community assembly across light, water, and nutrient gradients. Furthermore, this innovative framework provides a link between root trait variation and plants relation with AMF.

Theobroma cacao, food of the gods

Plant History

The origin of cacao (*Theobroma cacao* L.) biodiversity lies in the Upper Amazon basin, specifically along the Peru–Ecuador–Colombia frontier, where evidence of its consumption dates to 5000 BCE (Lanaud et al., 2024). Amazonian cacao comprises ten genetic groups: Amelonado, Contamana, Curaray, Guiana, Iquitos, Marañón, Nacional, Nanay, Purus (Motamayor et al., 2008), and Caquetá (Fouet et al., 2022). From this centre of diversity, cacao dispersed northward into Central America and eastward toward the Brazilian coast. Around 3000 BCE, Mesoamerican natives (in present-day Mexico, Guatemala, and Belize) semi-domesticated cacao, giving rise to the Criollo group (Motamayor et al., 2002). The pronounced phenotypic divergence of Criollo from Amazonian groups once led to the hypothesis of two domestication centres, in Central and South America (Cheesman, 1944). Domestication in isolated Mayan sinkholes (Cruz et al., 1995; Gómez-Pompa et al., 1990) increased homozygosity, reducing plant vigour, and produced a key physiological distinction: lower anthocyanin content. This trait contributed to Criollo's delicate, non-bitter flavour

but also increased its susceptibility to disease (Bekele and Phillips-Mora, 2019). Following European colonization, Criollo was disseminated across Spanish territories, but populations collapsed in the early 18th century, likely due to climatic stress during the Little Ice Age (Motilal and Sreenivasan, 2012). This collapse spurred hybridization between surviving Criollos and resilient Amazonian groups, giving rise to the Trinitario type (Bekele, 2019).

Cacao is a diploid (*2n = 2x = 20*) perennial crop that grows in the tropical belt. It is an allogamous and out-crossing plant with a complex self-incompatible system of sporophytic and gametophytic control (Lanaud et al., 2017). Still, a small percentage of plants allows selfing. From the pollination, the fruit development takes 6 months to its harvest, and from seed germination to the reproductive phase of the plant it can take from 3 to 5 years. Cacao pods generally contain 40–60 beans, which serve as the source of all cocoa-derived products. Throughout this manuscript, *cacao* will refer to the plant, whereas *cocoa* will denote its processed products. Its origin in the Amazonian forest as an understory tree resulted in naturally growing under 75%-45% of shade throughout its development (González-Orozco et al., 2020). For this reason, is usually grown in cacao agroforestry systems with secondary shade and service trees, although it can grow in monocultures in full sun (Lennon et al., 2021).

Cacao Agroforestry Systems

Cacao agroforestry systems (CAFS) exhibit two predominant scenarios. One where smallholders on small-scale farms (<10 ha) grow seed-originated rain-fed cacao, with medium to low planting density, suboptimal shade design, and modest management practices. Smallholders are responsible for 90% of the cacao production worldwide and typically CAFS are mixed within conserved forest areas (Duguma et al., 2001; Gama-Rodrigues et al., 2021; Orozco-Aguilar et al., 2024). On the other hand, medium and large farms (>100ha) grow elite genetic material with irrigation systems, in a high planting density, optimal shade design and management (Orozco-Aguilar et al., 2024). In general, the latter system has a higher yield, better pest and disease management through chemicals, access to market and thus, higher return of investment. The economic viability of CAFS is further enhanced when elite scions are grafted onto diverse rootstocks, allowing for tailored management of canopy and root systems (Schmidt et al., 2022).

Ecologically, CAFS foster intricate interactions between cacao, secondary vegetation, and soil microbiota. Shade trees favour soil fauna (Addo-Danso et al., 2024; Suárez et al., 2021), and influence AMF communities (Buyer et al., 2017; Lovera et al., 2022). While CAFS in general have significant benefits for planetary health, food security and climate resilience, poor management can increase spread of diseases (Jaimes-Suárez et al., 2022; Rosenstock et al., 2019). Important diseases in the Neotropics are *Moniliophthora perniciosa*, *M. roreri* and *Phytophthora* spp (Evans, 2007; Phillips-Mora and Wilkinson, 2007), while there are other emergent diseases as cocoa swollen-shoot virus

and *Phytophthora megakarya* in Western Africa and vascular streak dieback in Asia (Marelli et al., 2019; Ploetz, 2007). These diseases can cause complete yield losses thus increasing the attention for better CAFS practices management and cacao cultivars.

Plant breeding in cacao

Cacao breeding began in early 20th century, with the selection of genotypes with high yield in cocoa-producing countries (Bekele and Phillips-Mora, 2019). Since then, there have been initiatives to preserve the genetic diversity and developments in genetic technology to accelerate the breeding process. To preserve genetic diversity two gene banks of international importance have been developed: CATIE in Costa Rica and ICGT in Trinidad & Tobago (Bekele et al., 2024). Additionally, expeditions to the wild also serve as a great genetic reservoir to enhance genetic resources from cacao and wild relatives (González-Orozco et al., 2020). An example cultivar from early breeding programs is the famous CCN51 cultivar, moderate-high yielding with good development and resilience across a variety of climates (Jaimez et al., 2022). This cultivar originated from the crosses (IMC-67 × ICS-95) × "Canelo" (Oriente 1) in Naranjal, Ecuador, developed in response to the severe outbreak of witches' broom disease in the 1960s. Genetic resources of relevance involve the development of EST datasets (Argout et al., 2008), two sequenced genomes (cultivar: B97-61/B2 (Argout et al., 2011) and Matina 1-6 (Motamayor et al., 2013), which resulted in a 15K SNP array (Livingstone et al., 2017). These developments made it possible to conduct genome-wide association studies (GWAS) and genomic selection (GS), where low complexity traits are better selected by GWAS and quantitative traits like disease resistance are better selected using GS (McElroy et al., 2018; Navarro et al., 2017). All these developments aim to facilitate the creation of new economically and ecologically sustainable cacao varieties, considering its challenges as long generation cycles and complex self-incompatibility systems.

A cacao crop ideotype would consist of high yield during a long period, resilience to diseases and environmental stresses through harnessing the symbiotic ecological interactions present in the CAFS. Still, breeding for belowground traits in cacao has not received further attention than cadmium tolerance (Engbersen et al., 2019; Fernández-Paz et al., 2021; Galvis et al., 2023; Lewis et al., 2018) and pathogen resistance (Gutiérrez et al., 2016). Given the influence of the microbial community on plant phenotype (Vandenkoornhuyse et al., 2015), understanding the factors shaping root microbiota is essential for developing microbiome-based strategies to improve agricultural sustainability (Schmidt et al., 2022). Cacao is a highly dependent crop on mycorrhizal symbiosis (Azizah and Ragu, 1986; Chulan, 1991; Chulan and Martin, 1992). Its semi-domesticated nature within agroforestry systems has prevented cacao selection under systems where AMF do not thrive, such as high-P inputs, tillage and monocultures. This semi-domesticated nature provides a particular scenario to investigate

the mycorrhizal symbiosis and learn valuable lessons that may be useful to enhance its effects in other crops, developing agricultural technologies without compromising sustainability.

The sustainable future of chocolate

From its sacred and commercial role among elite Mesoamerican societies to its status as a global commodity, cacao has shaped human culture for millennia (Baron, 2018; Steinberg, 2002). Nowadays, annual consumption of cacao amounts to 1 kg per capita worldwide, with a peaking consumption of 11 kg per capita in Europe (Government of the Netherlands, 2025). In 2023, global production reached 5 million tons, concentrated in equatorial nations like Ivory Coast, Ghana, Indonesia, Ecuador, and Brazil (FAOSTAT, 2023a). Nevertheless, cacao production faces escalating threats from climate change and disease outbreaks, causing prices to double in 2024, a record high (Ritchie and Numbers, 2024). Additionally, increasing production in these areas goes hand in hand with unique biodiversity loss further challenging climate adaptation (Barlow et al., 2018; Norris, 2008; Soto-Navarro et al., 2021). Given these challenges, certification schemes for sustainable cacao play a critical role in ensuring responsible production.

The creation of certification schemes, such as Fairtrade and Rainforest Alliance, emerged to tackle socioeconomic and environmental issues (Dragusanu et al., 2014); however, these initiatives also encounter substantial challenges. Fairtrade certifications aim to improve cacao farmers' incomes by guaranteeing minimum prices and premiums for community development, while advocating for better labour conditions and sustainable practices. Rainforest Alliance adopts a broader approach, balancing socioeconomic and ecological benefits. Nevertheless, producers bear costly certification fees, and specialty buyers often determine premium prices through opaque criteria, limiting farmers' bargaining power (Cadby and Araki, 2021). While certifications successfully increase product prices, their benefits are inconsistent: certifications show limited impacts on actual worker wages or household incomes (Oya et al., 2018). Their socioeconomic effects outweigh their environmental benefits, with minimal influence on conservation or supply chain resilience (Blackman and Rivera, 2010; Garrett et al., 2021; Meemken et al., 2021). Recent regulations (e.g., cadmium limits, anti-deforestation laws) aim to strengthen human and environmental protections but do not resolve the numerous challenges for sustainable production (Government of the Netherlands, 2024).

Cacao in Panama

Cacao has been grown in Panama (Figure 2) by indigenous communities since 1500 BCE (Lanaud et al., 2024). In Bocas del Toro, the presence of the Ngabe Bugle indigenous community with their consumption of cocoa for its social and spiritual relevance, accounts up to now for 90% of the national cacao production (6000 ha approximately; COCABO, pers comm). Additionally, other groups have started cacao cultivation in different regions. In early 20th century, commercial

production began with the United Fruit Company importing and testing Trinitario for yield, resulting in UFC clones (Mata-Quirós et al., 2018). The commercial production was abandoned with the beginning of the World War I where the company farmers switched to textile crops more relevant for the war. The cacao plantations that were not replaced were left to the Caribbean Afro-American workers, who cultivate these until now. In a separate region further south and inland, Herrera, cacao cultivation began later in the 1970s by western settlements, motivated by government programs. Cacao genotypes from CATIE were imported between the 1970 - 1980s as part of a development program by the Ministry of Agricultural Development. These clones were mixtures of Trinitario and Criollo, selected from Costa Rica, Honduras, and Panama, and other types imported from Ecuador and Colombia. The import and proliferation of cacao cultivation increased incidence of cacao diseases requiring more on-farm management. With a simultaneous drop in the cocoa price, the cultivation of cacao was no longer of interest.

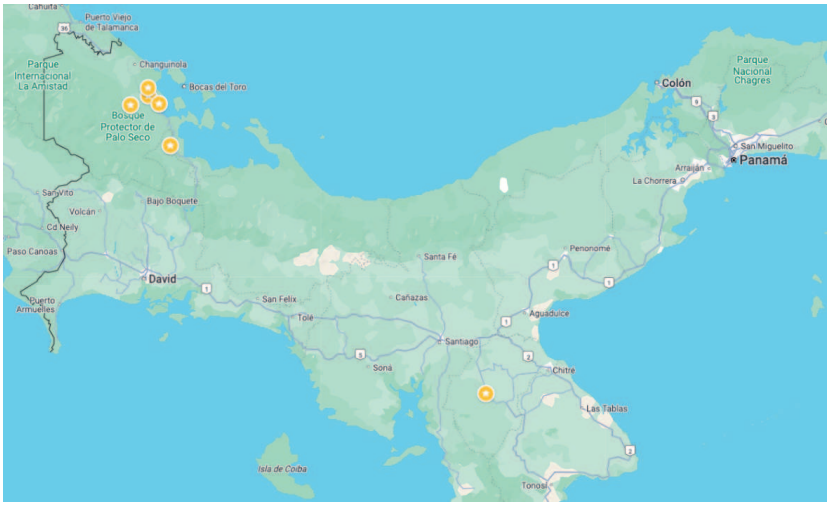


Figure 2. Fragment of the Republic of Panama, with emphasis in the cultivated studied areas.

Panamanian cacao has regained interest by farmers due to its outstanding quality in international competitions. According to the international cocoa competition, Cocoa of Excellence in Paris in 2021, Panamanian cocoa is ranked as one of the best fine-flavoured cocoa beans (Cacao of Excellence, 2021), with the challenge of having a low yield (103 kg/ha) compared to the regional average of 473 kg/ha (FAOSTAT, 2023a). Understanding the genetic diversity and population structure of the cacao farmer varieties (or germplasm) in Panama is crucial for developing and ensuring the sustainable production of the cocoa industry in the country. This can support in situ and on-farm conservation of these farmer varieties, aid in the selection and breeding of superior varieties and, when placed in the environmental context, guide decisions and resource investment in cacao farm management.

Aim of the thesis

The aim of this thesis was to investigate cacao germplasm, associated AMF communities, and the root traits that modulate their symbiosis across farms in Panama with breeding prospectives. Because plants and AMF are also shaped by agricultural management, I considered it equally important to describe the practices farmers employ. To achieve this, I studied the cacao germplasm cultivated on smallholder farms managed by Indigenous and mestizo communities in Panama, assessing the genetic composition of both above- and belowground plant tissues. I then investigated the mycorrhizal communities associated with cacao roots and examined how these are influenced by soil properties, root traits, and plant genetics. Finally, I conducted a greenhouse experiment to test for variation in mycorrhizal growth responses among cacao genotypes.

Overarching research question

How does genetic variation of rootstocks and scion, and root traits in *T. cacao* influence the AM fungal community composition (core microbiome and indicator species) and AMF colonization?

What are the root traits that facilitate the beneficial interaction with AMF?

Sub-research questions:

- How do smallholders manage trees across different regions, including direct sowing or vegetative propagation by selection of scions and grafting?
- How do management practices, like selection of neighbouring species and cacao germplasm, and environmental properties of CAFS influence AMF communities?
- What cacao genotypes are cultivated across different regions in Panama?
- What is the core AMF community in relation to cacao? Which species are indicator species for certain scion or rootstock, or specific environmental conditions?

These insights will pave the way for future research to identify the phenotypic effects that AMF species confer on different cacao genotypes, as well as the degree of specialization in their genetic interactions. Once the ecological strategies of these AMF species are understood, management practices can be tailored to favour beneficial AMF communities and promote their ecological success. Likewise, breeding programs can be guided by knowledge of plant genetics and the root traits that regulate the specificity of plant–fungal interactions. A clear understanding of existing cacao germplasm, the associated AMF communities, and the traits involved in mycorrhizal responsiveness will also enable the development of genetic marker assisted selection in cacao. Finally, linking ecological and genetic knowledge with on-farm management practices not only strengthens breeding efforts, but also opens opportunities to create social tools and resources that enhance both farm productivity and the quality of cocoa products.

Outline of the dissertation

In **Chapter 2**, I studied the genotypic characterization of the cacao trees in regions of Bocas del Toro and Herrera in Panama. I collected and genotyped both above- and belowground plant tissues to understand cacao tree management (seed propagated, grafted or rooted cuttings). There were no differences between regions or below/above ground tissues regarding heterozygosity, which broadly varied across all samples. I found that most of the cacao genotypes in Bocas del Toro are admixed, a mixture of all genetic groups, with a great predominance of Amelonado, Criollo, Nacional, IMC (Iquitos) and Nanay. In Herrera, there was a higher presence of the Criollo group, with some genotypes exhibiting 100% pure criollos. I hypothesized that in Herrera the higher elevation facilitated the conservation of pure criollos due to reported absence of *Moniliophthora spp*. Regarding management, there were slightly more trees originated from seeds across all sampled genotypes than from grafting, however Bocas del Toro had a significantly higher percentage of grafted trees in comparison with Herrera.

In **Chapter 3**, I investigated AMF community composition in cacao trees sampled from farms in Bocas del Toro and Herrera, Panama. Cacao roots were collected and analysed through Illumina MiSeq sequencing to determine AMF species composition, with *Glomus* identified as the dominant genus across all samples. While AMF diversity did not differ significantly between regions, significant variation was detected among farms, likely influenced by management practices. To complement the sequencing data, I recorded morphological root traits (specific root length, root diameter, and root tissue density) using high-quality imaging, measured root N and P concentrations through chemical analysis, and documented general farm traits such as elevation and tree density. Distance-based redundancy analysis revealed that scion genotype, root nutrient concentrations (N and P), elevation, and available P were key variables structuring AMF communities, whereas rootstock genotype and root morphological traits had no significant effect. An indicator species analysis further showed strong associations of AMF taxa with elevation, root N, root P, and scion genotype, while available P was linked to species with weaker indicator values, likely due to its narrow range of variation across farms.

In **Chapter 4**, I conducted a factorial experiment using three selfed progenies of selected clones with four soil inocula and a non-inoculated control. The progenies were derived from trees previously selected by breeder Abiel Gutierrez (IDIAP) for high yield and resistance to *Moniliophthora roreri* and *Phytophthora* spp. The four soil inocula were collected from different plantations spanning a range of soil pH and texture identified in the field survey described in Chapters 2 and 3. Six-month-old seedlings were assessed for above- and belowground biomass, as well as root traits related to the root economics space (specific root length, root diameter, nitrogen content, and root tissue density). Aboveground biomass differed significantly among AMF treatments and progenies, with a significant

interaction effect, highlighting the interplay between plant genetic composition and AMF community treatments. For root traits, diameter and nitrogen content showed genetically conserved patterns, with limited variation between progenies regardless of AMF community. In contrast, specific root length and root tissue density responded strongly to reduced AMF colonization, reflecting a pronounced environmental response independent of plant genetic differences.

Finally, in the **General Discussion**, I begin by situating the project within the social and economic context that inspired and shaped my PhD journey. I then reflect on the scientific, theoretical, and methodological lessons learned throughout the research, before concluding with broader implications from my perspective. To frame the work in a wider context, I examine current agricultural challenges and implications in transitioning from the Green Revolution to the Evergreen Revolution. I then discuss the role of agroforestry models within this transition, focusing particularly on the status of cacao agroforestry systems and their relationship to the cacao industry globally and in Panama, my home country. Over the course of this project, I realized the importance of strategic collaborations, creativity, and resourcefulness in conducting research across countries, as well as the value of methodological consistency and the theoretical insights gained from distinguishing between intra- and interspecific variation. Lastly, drawing on findings across all chapters (Figure 3), I conclude that breeding for mycorrhizal symbiosis in cacao would benefit from participatory plant breeding approaches that enhance plant selectivity for beneficial AMF communities.

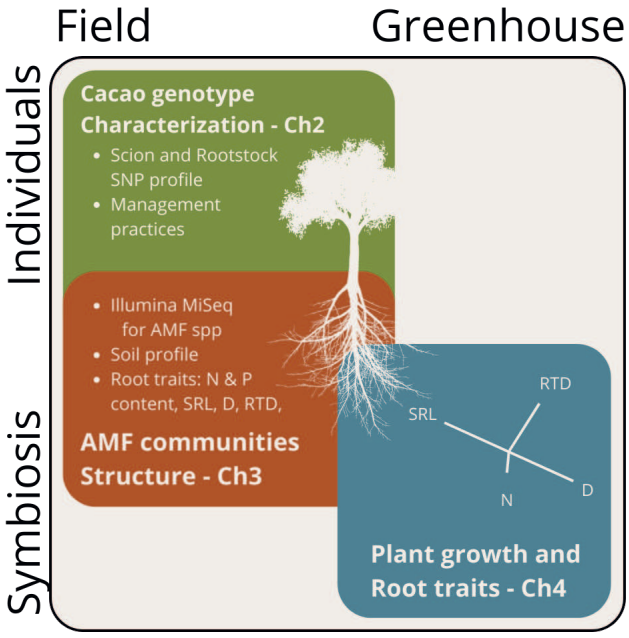


Figure 3. A schematic approach to the chapters, focusing on individuals or symbiotic understanding, whether in the field or greenhouse experiments.

References

Aavik, T., Träger, S., Zobel, M., Honnay, O., Geel, M.V., Bueno, C.G., Koorem, K., 2021. The joint effect of host plant genetic diversity and arbuscular mycorrhizal fungal communities on restoration success. Funct. Ecol. 35, 2621–2634. https://doi.org/10.1111/1365-2435.13914

Abbott, S., Fairbanks, D.J., 2016. Experiments on Plant Hybrids by Gregor Mendel. Genetics 204, 407–422. https://doi.org/10.1534/genetics.116.195198

Addo-Danso, S.D., Asare, R., Tettey, A., Schmidt, J.E., Sauvadet, M., Coulis, M., Belliard, N., Isaac, M.E., 2024. Shade tree functional traits drive critical ecosystem services in cocoa agroforestry systems. Agric., Ecosyst. Environ. 372, 109090. https://doi.org/10.1016/j.agee.2024.109090

Argout, X., Fouet, O., Wincker, P., Gramacho, K., Legavre, T., Sabau, X., Risterucci, A.M., Silva, C.D., Cascardo, J., Allegre, M., Kuhn, D., Verica, J., Courtois, B., Loor, G., Babin, R., Sounigo, O., Ducamp, M., Guiltinan, M.J., Ruiz, M., Alemanno, L., Machado, R., Phillips, W., Schnell, R., Gilmour, M., Rosenquist, E., Butler, D., Maximova, S., Lanaud, C., 2008. Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. BMC Genom. 9, 512. https://doi.org/10.1186/1471-2164-9-512

Argout, X., Salse, J., Aury, J.-M., Guiltinan, M.J., Droc, G., Gouzy, J., Allegre, M., Chaparro, C., Legavre, T., Maximova, S.N., Abrouk, M., Murat, F., Fouet, O., Poulain, J., Ruiz, M., Roguet, Y., Rodier-Goud, M., Barbosa-Neto, J.F., Sabot, F., Kudrna, D., Ammiraju, J.S.S., Schuster, S.C., Carlson, J.E., Sallet, E., Schiex, T., Dievart, A., Kramer, M., Gelley, L., Shi, Z., Bérard, A., Viot, C., Boccara, M., Risterucci, A.M., Guignon, V., Sabau, X., Axtell, M.J., Ma, Z., Zhang, Y., Brown, S., Bourge, M., Golser, W., Song, X., Clement, D., Rivallan, R., Tahi, M., Akaza, J.M., Pitollat, B., Gramacho, K., D'Hont, A., Brunel, D., Infante, D., Kebe, I., Costet, P., Wing, R., McCombie, W.R., Guiderdoni, E., Quetier, F., Panaud, O., Wincker, P., Bocs, S., Lanaud, C., 2011. The genome of *Theobroma cacao*. Nature Genet. 43, 101–108. https://doi.org/10.1038/ng.736

Azizah, C. (Hashim), Ragu, P., 1986. Growth response of *Theobroma cacao* L. seeldings to inoculation with vesicular-arbuscular mycorrhizal fungi. Plant Soil 96, 279–285. https://doi.org/10.1007/bf02374771

Bardgett, R.D., Mommer, L., de Vries, F.T., 2014. Going underground: root traits as drivers of ecosystem processes. Trends Ecol. Evol. 29, 692–699. https://doi.org/10.1016/j.tree.2014.10.006

Barlow, J., França, F., Gardner, T.A., Hicks, C.C., Lennox, G.D., Berenguer, E., Castello, L., Economo, E.P., Ferreira, J., Guénard, B., Leal, C.G., Isaac, V., Lees, A.C., Parr, C.L., Wilson, S.K., Young, P.J., Graham, N.A.J., 2018. The future of hyperdiverse tropical ecosystems. Nature 559, 517–526. https://doi.org/10.1038/s41586-018-0301-1

Baron, J.P., 2018. Making money in Mesoamerica: Currency production and procurement in the Classic Maya financial system. Econ. Anthr. 5, 210–223. https://doi.org/10.1002/sea2.12118

Beech, E., Rivers, M., Oldfield, S., Smith, P.P., 2017. GlobalTreeSearch: The first complete global database of tree species and country distributions. J. Sustain. For. 36, 454–489. https://doi.org/10.1080/10549811.2017.1310049

Beillouin, D., Ben-Ari, T., Malézieux, E., Seufert, V., Makowski, D., 2021. Positive but variable effects of crop diversification on biodiversity and ecosystem services. Glob. Chang. Biol. 27, 4697–4710. https://doi.org/10.1111/gcb.15747

Bekele, F., 2019. Trinidad and Tobago, the Birthplace of the world-famous Trinitario cocoa. https://doi.org/10.13140/rg.2.2.36603.95527

Bekele, F., Bidaisee, G., Bhola, J., Umaharan, P., 2024. Catalogue of cacaos conserved at the CRC genebank. Lima (Peru): Bioversity International/MOCCA. 40 p. https://hdl.handle.net/10568/148922

Bekele, F., Phillips-Mora, W., 2019. Cacao (*Theobroma cacao* L.). In: In: Al-Khayri, J., Jain, S., Johnson, D. (Eds.), Plant Breeding Strategies: Industrial and Food Crops, Volume 6, 409–487. Springer. https://doi.org/10.1007/978-3-030-23265-8_12

Benaffari, W., Boutasknit, A., Anli, M., Ait-El-Mokhtar, M., Ait-Rahou, Y., Ben-Laouane, R., Ahmed, H.B., Mitsui, T., Baslam, M., Meddich, A., 2022. The Native Arbuscular Mycorrhizal Fungi and Vermicompost-Based Organic Amendments Enhance Soil Fertility, Growth Performance, and the Drought Stress Tolerance of Quinoa. Plants 11, 393. https://doi.org/10.3390/plants11030393

Blackman, A., Rivera, J.E., 2010. The Evidence Base for Environmental and Socioeconomic Impacts of 'Sustainable' Certification. Discussion paper, Resources for the Future, DP10-17. https://media.rff.org/documents/RFF-DP-10-17.pdf.

Boerma, H.R., Hussey, R.S., 1992. Breeding plants for resistance to nematodes. J. Nematol. 24, 242–52.

Bom, F.J.T. van der, Williams, A., Bell, M.J., 2020. Root architecture for improved resource capture: trade-offs in complex environments. J. Exp. Bot. 71, 5752–5763. https://doi.org/10.1093/jxb/eraa324

Boyno, G., Danesh, Y.R., Çevik, R., Teniz, N., Demir, S., Durak, E.D., Farda, B., Mignini, A., Djebaili, R., Pellegrini, M., Porcel, R., Mulet, J.M., 2025. Synergistic benefits of AMF: development of sustainable plant defense system. Front. Microbiol. 16, 1551956. https://doi.org/10.3389/fmicb.2025.1551956

Buyer, J.S., Baligar, V.C., He, Z., Arévalo-Gardini, E., 2017. Soil microbial communities under cacao agroforestry and cover crop systems in Peru. Appl. Soil Ecol. 120, 273–280. https://doi.org/10.1016/j.apsoil.2017.09.009

Cadby, J., Araki, T., 2021. Towards ethical chocolate: multicriterial identifiers, pricing structures, and the role of the specialty cacao industry in sustainable development. SN Bus. Econ. 1, art. 44. https://doi.org/10.1007/s43546-021-00051-y

Canfora, L., Costa, C., Pallottino, F., Mocali, S., 2021. Trends in Soil Microbial Inoculants Research: A Science Mapping Approach to Unravel Strengths and Weaknesses of Their Application. Agriculture 11, art. 158. https://doi.org/10.3390/agriculture11020158

Ceccarelli, S., 1996. Adaptation to low/high input cultivation. Euphytica 92, 203–214. https://doi.org/10.1007/bf00022846

Cheesman, E.E., 1944. Notes on the nomenclature, classification possible and relationships of cocoa populations. Tropical Agriculture 21, 144–189.

Chulan, H.A., 1991. Effect of fertilizer and endomycorrhizal inoculum on growth and nutrient uptake of cocoa (*Theobroma cacao* L.) seedlings. Biol. Fert. Soils 11, 250–254. https://doi.org/10.1007/bf00335843

Chulan, H.A., Martin, K., 1992. The vesicular-arbuscular (VA) mycorrhiza and its effects on growth of vegetatively propagated *Theobroma cacao* L. Plant Soil 144, 227–233. https://doi.org/10.1007/bf00012879

Cobb, A.B., Duell, E.B., Haase, K.B., Miller, R.M., Wu, Y.Q., Wilson, G.W.T., 2021. Utilizing mycorrhizal responses to guide selective breeding for agricultural sustainability. Plants People Planet 3, 578–587. https://doi.org/10.1002/ppp3.10200

Cacao of Excellence, 2021 Cocoa of Excellence Awards [WWW Document]. URL https://www.cacaoofexcellence.org/2021-edition/2021-awards-winners-producers-list

Coleman, D.C., Wall, D.H., 2015. Soil Microbiology, Ecology and Biochemistry (Fourth Edition) 111–149. https://doi.org/10.1016/b978-0-12-415955-6.00005-0

Conway, G.R., Barbie, E.B., 1988. After the Green Revolution Sustainable and equitable agricultural development. Futures 20, 651–670. https://doi.org/10.1016/0016-3287(88)90006-7

Cooper, M., Messina, C.D., Tang, T., Gho, C., Powell, O.M., Podlich, D.W., Technow, F., Hammer, G.L., 2022. Predicting Genotype × Environment × Management (G × E × M) Interactions for the Design of Crop Improvement Strategies**.** Integrating Breeder, Agronomist, and Farmer Perspectives. Plant Breeding reviews 46, 467–585. https://doi.org/10.1002/9781119874157.ch8

Cruz, M. de la, Whitkus, R., Gómez-Pompa, A., Mota-Bravo, L., 1995. Origins of cacao cultivation. Nature 375, 542–543. https://doi.org/10.1038/375542a0

Cuatrecasas, 1964. Cacao and its Allies. A taxonomic revision of the Genus *Theobroma*. Contribution US Herbarium 35: 379–614.

Darwin, C.R., 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London: John Murray.

Dowarah, B., Gill, S.S., Agarwala, N., 2022. Arbuscular Mycorrhizal Fungi in Conferring Tolerance to Biotic Stresses in Plants. J. Plant Growth Regul. 41, 1429–1444. https://doi.org/10.1007/s00344-021-10392-5

Dragusanu, R., Giovannucci, D., Nunn, N., 2014. The Economics of Fair Trade. J. Econ. Perspect. 28, 217–236. https://doi.org/10.1257/jep.28.3.217

Driscoll, C.A., MacDonald, D.W., O'Brien, S.J., 2009. From Wild Animals to Domestic Pets, an Evolutionary View of Domestication. Proc. Natl. Acad. Sci. USA 106(Suppl.), 9971-9978. www.pnas.org/cgi/doi/10.1073/pnas.0901586106

Duguma, B., Gockowski, J., Bakala, J., 2001. Smallholder Cacao (*Theobroma cacao* Linn.) cultivation in agroforestry systems of West and Central Africa: challenges and opportunities. Agrofor. Syst. 51, 177–188. https://doi.org/10.1023/a:1010747224249

Engbersen, N., Gramlich, A., Lopez, M., Schwarz, G., Hattendorf, B., Gutierrez, O., Schulin, R., 2019. Cadmium accumulation and allocation in different cacao cultivars. Sci. Tot. Environ. 678, 660–670. https://doi.org/10.1016/j.scitotenv.2019.05.001

Evans, H.C., 2007. Cacao Diseases—The Trilogy Revisited. Phytopathology 97, 1640–1643. https://doi.org/10.1094/phyto-97-12-1640

FAO, The State of the World's Forest Genetic Resources. https://www.fao.org/forest-genetic-resources/assessments/first-report/en/

FAOSTAT, 2023. Crops and Livestock products, Cocoa beans, area harvested in 2023, Crops/World regions/Cocoa beans/Area harvested from pick lists. [WWW Document]. URL https://www.fao.org/faostat/en/#data/QCL (accessed 1.13.25).

Fernández-Paz, J., Cortés, A.J., Hernández-Varela, C.A., Mejía-de-Tafur, M.S., Rodriguez-Medina, C., Baligar, V.C., 2021. Rootstock-Mediated Genetic Variance in Cadmium Uptake by Juvenile Cacao (*Theobroma cacao* L.) Genotypes, and Its Effect on Growth and Physiology. Front. Plant Sci. 12, 777842. https://doi.org/10.3389/fpls.2021.777842

Fouet, O., Solorzano, R.G.L., Rhoné, B., Subía, C., Calderón, D., Fernández, F., Sotomayor, I., Rivallan, R., Colonges, K., Vignes, H., Angamarca, F., Yaguana, B., Costet, P., Argout, X., Lanaud, C., 2022. Collection of native *Theobroma cacao* L. accessions from the Ecuadorian Amazon highlights a hotspot of cocoa diversity. Plants, People, Planet 4, 605–617. https://doi.org/10.1002/ppp3.10282

Fradgley, N., Evans, G., Biernaskie, J.M., Cockram, J., Marr, E.C., Oliver, A.G., Ober, E., Jones, H., 2020. Effects of breeding history and crop management on the root architecture of wheat. Plant Soil 452, 587–600. https://doi.org/10.1007/s11104-020-04585-2

Galván, G.A., Kuyper, T.W., Burger, K., Keizer, L.C.P., Hoekstra, R.F., Kik, C., Scholten, O.E., 2011. Genetic analysis of the interaction between *Allium* species and arbuscular mycorrhizal fungi. Theor. Appl. Genet. 122, 947–960. https://doi.org/10.1007/s00122-010-1501-8

Galvis, D.A., Jaimes-Suárez, Y.Y., Molina, J.R., Ruiz, R., León-Moreno, C.E., Carvalho, F.E.L., 2023. Unveiling Cacao Rootstock-Genotypes with Potential Use in the Mitigation of Cadmium Bioaccumulation. Plants 12, art. 2941. https://doi.org/10.3390/plants12162941

Gamalero, E., Lingua, G., Berta, G., Glick, B.R., 2009. Beneficial role of plant growth promoting bacteria and arbuscular mycorrhizal fungi on plant responses to heavy metal stress. Can. J. Microbiol. 55, 501–514. https://doi.org/10.1139/w09-010

Gama-Rodrigues, A.C., Müller, M.W., Gama-Rodrigues, E.F., Mendes, F.A.T., 2021. Cacao-based agroforestry systems in the Atlantic Forest and Amazon Biomes: An ecoregional analysis of land use. Agric. Syst. 194, 103270. https://doi.org/10.1016/j.agsy.2021.103270

Gange, A.C., Ayres, R.L., 1999. On the Relation between Arbuscular Mycorrhizal Colonization and Plant "Benefit." Oikos 87, 615-621. https://doi.org/10.2307/3546829

Garrett, R.D., Levy, S.A., Gollnow, F., Hodel, L., Rueda, X., 2021. Have food supply chain policies improved forest conservation and rural livelihoods? A systematic review. Environ. Res. Lett. 16, 033002. https://doi.org/10.1088/1748-9326/abe0ed

Gatti, R.C., Reich, P.B., Gamarra, J.G.P., Crowther, T., Hui, C., Morera, A., Bastin, J.-F., de-Miguel, S., Nabuurs, G.-J., Svenning, J.-C., Serra-Diaz, J.M., Merow, C., Enquist, B., Kamenetsky, M., Lee, J., Zhu, J., Fang, J., Jacobs, D.F., Pijanowski, B., Banerjee, A., Giaquinto, R.A., Alberti, G., Zambrano, A.M.A., Alvarez-Davila, E., Araujo-Murakami, A., Avitabile, V., Aymard, G.A., Balazy, R., Baraloto, C., Barroso, J.G., Bastian, M.L., Birnbaum, P., Bitariho, R., Bogaert, J., Bongers, F., Bouriaud, O., Brancalion, P.H.S., Brearley, F.Q., Broadbent, E.N., Bussotti, F., Silva, W.C. da, César, R.G., Češljar, G., Moscoso, V.C., Chen, H.Y.H., Cienciala, E., Clark, C.J., Coomes, D.A., Dayanandan, S., Decuyper, M., Dee, L.E., Pasquel, J.D.A., Derroire, G., Djuikouo, M.N.K., Do, T.V., Dolezal, J., Đorđević, I.Đ., Engel, J., Fayle, T.M., Feldpausch, T.R., Fridman, J.K., Harris, D.J., Hemp, A., Hengeveld, G., Herault, B., Herold, M., Ibanez, T., Jagodzinski, A.M., Jaroszewicz, B., Jeffery, K.J., Johannsen, V.K., Jucker, T., Kangur, A., Karminov, V.N., Kartawinata, K., Kennard, D.K., Kepfer-Rojas, S., Keppel, G., Khan, M.L., Khare, P.K., Kileen, T.J., Kim, H.S., Korjus, H., Kumar, Amit, Kumar, Ashwani, Laarmann, D., Labrière, N., Lang, M., Lewis, S.L., Lukina, N., Maitner, B.S., Malhi, Y., Marshall, A.R., Martynenko, O.V., Mendoza, A.L.M., Ontikov, P.V., Ortiz-Malavasi, E., Camacho, N.C.P., Paquette, A., Park, M., Parthasarathy, N., Peri, P.L., Petronelli, P., Pfautsch, S., Phillips, O.L., Picard, N., Piotto, D., Poorter, L., Poulsen, J.R., Pretzsch, H., Ramírez-Angulo, H., Correa, Z.R., Rodeghiero, M., Gonzáles, R.D.P.R., Rolim, S.G., Rovero, F., Rutishauser, E., Saikia, P., Salas-Eljatib, C., Schepaschenko, D., Scherer-Lorenzen, M., Šebeň, V., Silveira, M., Slik, F., Sonké, B., Souza, A.F., Stereńczak, K.J., Svoboda, M., Taedoumg, H., Tchebakova, N., Terborgh, J., Tikhonova, E., Torres-Lezama, A., Plas, F. van der, Vásquez, R., Viana, H., Vibrans, A.C., Vilanova, E., Vos, V.A., Wang, H.-F., Westerlund, B., White, L.J.T., Wiser, S.K., Zawiła-Niedźwiecki, T., Zemagho, L., Zhu, Z.-X., Zo-Bi, I.C., Liang, J., 2022. The number of tree species on Earth. Proc. Natl. Acad. Sci. 119, e2115329119. https://doi.org/10.1073/pnas.2115329119

Gómez-Pompa, A., Flores, J.S., Fernández, A., 1990. The Sacred Cacao Groves of the Maya. Latin Amer. Antiq. 1, 247–257. https://doi.org/10.2307/972163

González-Orozco, C.E., Galán, A.A.S., Ramos, P.E., Yockteng, R., 2020. Exploring the diversity and distribution of crop wild relatives of cacao (*Theobroma cacao* L.) in Colombia. Genet. Resour. Crop Evol. 67, 2071–2085. https://doi.org/10.1007/s10722-020-00960-1

Government of the Netherlands, 2024. What requirements must cocoa meet to be allowed on the European market? [WWW Document]. URL https://www.cbi.eu/market-information/cocoa/buyer-requirements (accessed 1.13.25).

Government of the Netherlands, 2025. What is the demand for cocoa on the European market? [WWW Document]. URL https://www.cbi.eu/market-information/cocoa/what-demand (accessed 5.8.25).

Gregory, T.R., 2009. Understanding Natural Selection: Essential Concepts and Common Misconceptions. Evol.: Educ. Outreach 2, 156–175. https://doi.org/10.1007/s12052-009-0128-1

Gutiérrez, O.A., Campbell, A.S., Phillips-Mora, W., 2016. Breeding for resistance in cacao. In: Bailey, B.A., Meinhardt, L.W. (Eds), Cacao Diseases, A History of Old Enemies and New Encounters 567–609. Springer. https://doi.org/10.1007/978-3-319-24789-2_18

Guzman, A., Montes, M., Hutchins, L., DeLaCerda, G., Yang, P., Kakouridis, A., Dahlquist-Willard, R.M., Firestone, M.K., Bowles, T., Kremen, C., 2021. Crop diversity enriches arbuscular mycorrhizal fungal communities in an intensive agricultural landscape. New Phyt. 231, 447–459. https://doi.org/10.1111/nph.17306

Habibi, F., Liu, T., Folta, K., Sarkhosh, A., 2022. Physiological, biochemical, and molecular aspects of grafting in fruit trees. Hortic. Res. 9, uhac032. https://doi.org/10.1093/hr/uhac032

Hetrick, B.A.D., Wilson, G.W.T., Cox, T.S., 1993. Mycorrhizal dependence of modern wheat varieties, landraces, and ancestors. Can. J. Bot. 71, 512–518. https://doi.org/10.1139/b92-253

Hoeksema, J.D., 2010. Ongoing coevolution in mycorrhizal interactions. New Phyt. 187, 286–300. https://doi.org/10.1111/j.1469-8137.2010.03305.x

Hohmann, P., Messmer, M.M., 2017. Breeding for mycorrhizal symbiosis: focus on disease resistance. Euphytica 213, art. 113. https://doi.org/10.1007/s10681-017-1900-x

Hornstein, E.D., Sederoff, H., 2023. Back to the Future: Re-Engineering the Evolutionarily Lost Arbuscular Mycorrhiza Host Trait to Improve Climate Resilience for Agriculture. Crit. Rev. Plant Sci. 43, 1–33. https://doi.org/10.1080/07352689.2023.2256093

Jaimes-Suárez, Y.Y., Carvajal-Rivera, A.S., Galvis-Neira, D.A., Carvalho, F.E.L., Rojas-Molina, J., 2022. Cacao agroforestry systems beyond the stigmas: Biotic and abiotic stress incidence impact. Front. Plant Sci. 13, 921469. https://doi.org/10.3389/fpls.2022.921469

Jaimez, R.E., Barragan, L., Fernández-Niño, M., Wessjohann, L.A., Cedeño-Garcia, G., Cantos, I.S., Arteaga, F., 2022. *Theobroma cacao* L. cultivar CCN 51: a comprehensive review on origin, genetics, sensory properties, production dynamics, and physiological aspects. PeerJ 10, e12676. https://doi.org/10.7717/peerj.12676

Jiang, S., An, X., Shao, Y., Kang, Y., Chen, T., Mei, X., Dong, C., Xu, Y., Shen, Q., 2021. Responses of Arbuscular Mycorrhizal Fungi Occurrence to Organic Fertilizer: A meta-analysis of field studies. Plant Soil 469, 89–105. https://doi.org/10.1007/s11104-021-05153-y

Jung, S.C., Martinez-Medina, A., Lopez-Raez, J.A., Pozo, M.J., 2012. Mycorrhiza-Induced Resistance and Priming of Plant Defenses. J. Chem. Ecol. 38, 651–664. https://doi.org/10.1007/s10886-012-0134-6

Kabir, Z., 2005. Tillage or no-tillage: Impact on mycorrhizae. Can. J. Plant Sci. 85, 23–29. https://doi.org/10.4141/p03-160

Kell, D.B., 2011. Breeding crop plants with deep roots: their role in sustainable carbon, nutrient and water sequestration. Ann. Bot. 108, 407–418. https://doi.org/10.1093/aob/mcr175

Kuijken, R.C.P., Eeuwijk, Fred.A. van, Marcelis, L.F.M., Bouwmeester, H.J., 2015. Root phenotyping: from component trait in the lab to breeding. J. Exp. Bot. 66, 5389–5401. https://doi.org/10.1093/jxb/erv239

Kuyper, T.W., Wang, X., Muchane, M.N., 2021. The Interplay Between Roots and Arbuscular Mycorrhizal Fungi Influencing Water and Nutrient Acquisition and Use Efficiency, in: Rengel, Z., Djalovic, I. (Eds.), The Root Systems in Sustainable Agricultural Intensification. pp. 193–220. https://doi.org/10.1002/9781119525417.ch7

Lanaud, C., Fouet, O., Legavre, T., Lopes, U., Sounigo, O., Eyango, M.C., Mermaz, B., Silva, M.R.D., Solorzano, R.G.L., Argout, X., Gyapay, G., Ebaiarrey, H.E., Colonges, K., Sanier, C., Rivallan, R., Mastin, G., Cryer, N., Boccara, M., Verdeil, J.-L., Mousseni, I.B.E., Gramacho, K.P., Clément, D., 2017. Deciphering the *Theobroma cacao* self-incompatibility system: from genomics to diagnostic markers for self-compatibility. J. Exp. Bot. 68, 4775–4790. https://doi.org/10.1093/jxb/erx293

Lanaud, C., Vignes, H., Utge, J., Valette, G., Rhoné, B., Caputi, M.G., Nieto, N.S.A., Fouet, O., Gaikwad, N., Zarrillo, S., Powis, T.G., Cyphers, A., Valdez, F., Nunez, S.Q.O., Speller, C., Blake, M., Valdez, F.Jr., Raymond, S., Rowe, S.M., Duke, G.S., Romano, F.E., Solórzano, R.G.L., Argout, X., 2024. A revisited history of cacao domestication in pre-Columbian times revealed by archaeogenomic approaches. Sci. Rep. 14, 2972. https://doi.org/10.1038/s41598-024-53010-6

Lennon, A.M., Lewis, V.R., Farrell, A.D., Umaharan, P., 2021. Photochemical responses to light in sun and shade leaves of *Theobroma cacao* L. (West African Amelonado). Sci. Hortic. 276, 109747. https://doi.org/10.1016/j.scienta.2020.109747

Leon, N. de, Jannink, J., Edwards, J.W., Kaeppler, S.M., 2016. Introduction to a Special Issue on Genotype by Environment Interaction. Crop Sci. 56, 2081–2089. https://doi.org/10.2135/cropsci2016.07.0002in

Lewis, C., Lennon, A.M., Eudoxie, G., Umaharan, P., 2018. Genetic variation in bioaccumulation and partitioning of cadmium in *Theobroma cacao* L. Sci. Tot. Environ. 640, 696–703. https://doi.org/10.1016/j.scitotenv.2018.05.365

Li, Y., Chen, H., Gu, L., Wu, J., Zheng, X., Fan, Z., Pan, D., Li, J., Shu, W., Rosendahl, S., Wang, Y., 2024. Domestication of rice may have changed its arbuscular mycorrhizal properties by modifying phosphorus nutrition-related traits and decreasing symbiotic compatibility. New Phyt. 243, 1554-1570. https://doi.org/10.1111/nph.19901

Liu, J., Zhang, J., Li, D., Xu, C., Xiang, X., 2020. Differential responses of arbuscular mycorrhizal fungal communities to mineral and organic fertilization. MicrobiologyOpen 9, e00920. https://doi.org/10.1002/mbo3.920

Livingstone, D., Stack, C., Mustiga, G.M., Rodezno, D.C., Suarez, C., Amores, F., Feltus, F.A., Mockaitis, K., Cornejo, O.E., Motamayor, J.C., 2017. A Larger Chocolate Chip—Development of a 15K *Theobroma cacao* L. SNP Array to Create High-Density Linkage Maps. Front. Plant Sci. 8, 2008. https://doi.org/10.3389/fpls.2017.02008

Lovera, M., Cuenca, G., Fajardo, L., Cáceres, A., Guerra-Sierra, B.E., 2022. AMF Diversity in Coffee and Cacao Agroforestry Systems: Importance for Crop Productivity and Forest Conservation, In: Luga M.A., Pagano, M.C. (Eds), Mycorrhizal Fungi in South America, Biodiversity, Conservation, and Sustainable Food Production, pp. 107–127. https://doi.org/10.1007/978-3-031-12994-0_5

Luginbuehl, L.H., Oldroyd, G.E.D., 2017. Understanding the Arbuscule at the Heart of Endomycorrhizal Symbioses in Plants. Curr. Biol. 27, R952–R963. https://doi.org/10.1016/j.cub.2017.06.042

Manfredini, A., Malusà, E., Costa, C., Pallottino, F., Mocali, S., Pinzari, F., Canfora, L., 2021. Current Methods, Common Practices, and Perspectives in Tracking and Monitoring Bioinoculants in Soil. Front. Microbiol. 12, 698491. https://doi.org/10.3389/fmicb.2021.698491

Marelli, J.-P., Guest, D.I., Bailey, B.A., Evans, H.C., Brown, J.K., Junaid, M., Barreto, R.W., Lisboa, D.O., Puig, A.S., 2019. Chocolate Under Threat from Old and New Cacao Diseases. Phytopathology 109, 1331–1343. https://doi.org/10.1094/phyto-12-18-0477-rvw

Martín-Robles, N., Lehmann, A., Seco, E., Aroca, R., Rillig, M.C., Milla, R., 2018. Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. New Phyt. 218, 322–334. https://doi.org/10.1111/nph.14962

Mata-Quirós, A., Arciniegas-Leal, A., Phillips-Mora, W., Meinhardt, L.W., Motilal, L., Mischke, S., Zhang, D., 2018. Assessing hidden parentage and genetic integrity of the "United Fruit Clones" of cacao (*Theobroma cacao*) from Costa Rica using SNP markers. Breed. Sci. 68, 18057. https://doi.org/10.1270/jsbbs.18057

McElroy, M.S., Navarro, A.J.R., Mustiga, G., Stack, C., Gezan, S., Peña, G., Sarabia, W., Saquicela, D., Sotomayor, I., Douglas, G.M., Migicovsky, Z., Amores, F., Tarqui, O., Myles, S., Motamayor, J.C., 2018. Prediction of Cacao (*Theobroma cacao*) Resistance to Moniliophthora spp. Diseases via Genome-Wide Association Analysis and Genomic Selection. Front. Plant Sci. 9, 343. https://doi.org/10.3389/fpls.2018.00343

McGrail, R.K., McNear, D.H., 2021. Two centuries of breeding has altered root system architecture of winter wheat. Rhizosphere 19, 100411. https://doi.org/10.1016/j.rhisph.2021.100411

Meemken, E.-M., Barrett, C.B., Michelson, H.C., Qaim, M., Reardon, T., Sellare, J., 2021. Sustainability standards in global agrifood supply chains. Nature Food 2, 758–765. https://doi.org/10.1038/s43016-021-00360-3

Motamayor, J.C., Lachenaud, P., Mota, J.W. da S. e, Loor, R., Kuhn, D.N., Brown, J.S., Schnell, R.J., 2008. Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L). PLoS One 3, e3311. https://doi.org/10.1371/journal.pone.0003311

Motamayor, J.C., Mockaitis, K., Schmutz, J., Haiminen, N., III, D.L., Cornejo, O., Findley, S.D., Zheng, P., Utro, F., Royaert, S., Saski, C., Jenkins, J., Podicheti, R., Zhao, M., Scheffler, B.E., Stack, J.C., Feltus, F.A., Mustiga, G.M., Amores, F., Phillips, W., Marelli, J.P., May, G.D., Shapiro, H., Ma, J., Bustamante, C.D., Schnell, R.J., Main, D., Gilbert, D., Parida, L., Kuhn, D.N., 2013. The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. Genome Biol. 14, r53–r53. https://doi.org/10.1186/gb-2013-14-6-r53

Motamayor, J.C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A., Lanaud, C., 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89, 380–386. https://doi.org/10.1038/sj.hdy.6800156

Motilal, L.A., Sreenivasan, T.N., 2012. Revisiting 1727: Crop Failure Leads to the Birth of Trinitario Cacao. J. Crop Improv. 26, 599–626. https://doi.org/10.1080/15427528.2012.663734

Navarro, J.A.R., Phillips-Mora, W., Arciniegas-Leal, A., Mata-Quirós, A., Haiminen, N., Mustiga, G., III, D.L., Bakel, H. van, Kuhn, D.N., Parida, L., Kasarskis, A., Motamayor, J.C., 2017. Application of Genome Wide Association and Genomic Prediction for Improvement of Cacao

Productivity and Resistance to Black and Frosty Pod Diseases. Front. Plant Sci. 8, 1905. https://doi.org/10.3389/fpls.2017.01905

Nerva, L., Sandrini, M., Moffa, L., Velasco, R., Balestrini, R., Chitarra, W., 2022. Breeding toward improved ecological plant–microbiome interactions. Trends Plant Sci. 27, 1134–1143. https://doi.org/10.1016/j.tplants.2022.06.004

Noceto, P.-A., Bettenfeld, P., Boussageon, R., Hériché, M., Sportes, A., Tuinen, D. van, Courty, P.-E., Wipf, D., 2021. Arbuscular mycorrhizal fungi, a key symbiosis in the development of quality traits in crop production, alone or combined with plant growth-promoting bacteria. Mycorrhiza 31, 655–669. https://doi.org/10.1007/s00572-021-01054-1

Noceto, P.-A., Mathé, A., Anginot, L., Tuinen, D. van, Wipf, D., Courty, P.-E., 2024. Effect of rootstock diversity and grafted varieties on the structure and composition of the grapevine root mycobiome. Plant Soil 504, 307–332. https://doi.org/10.1007/s11104-024-06624-8

Norris, K., 2008. Agriculture and biodiversity conservation: opportunity knocks. Conserv. Lett. 1, 2–11. https://doi.org/10.1111/j.1755-263x.2008.00007.x

O'Callaghan, M., Ballard, R.A., Wright, D., 2022. Soil microbial inoculants for sustainable agriculture: Limitations and opportunities. Soil Use Manage. 38, 1340–1369. https://doi.org/10.1111/sum.12811

Orozco-Aguilar, L., Lopez-Sampson, A., Cerda, R.H., Casanoves, F., Ramirez-Argueta, O., Matute, J.D., Salazar, J.C.S., Rüegg, J., Saj, S., Milz, J., Schneidewind, U., Garces, A.M., Daza, E.B., Molina, J.R., Suarez, Y.J., Agudelo-Castañeda, G.A., Deheuvels, O., Sosa, E.B., Gómez, J.H., Jaimez, R.E., Espinoza, S.R., Bordeaux, M., Vargas, C.C., Tinoco, L., Monserrate, G.P., Flores, J.P., Morales, A.A., Arevalo-Hernandez, C.O., Gardini, E.A., Pocasangre, L.E., Araque, O., Koutouleas, A., Magaña, E.S., Dominguez, O., Arenas, P., Sotopinto, L., Salgado-Mora, M., Gama-Rodrigues, A., Gama-Rodrigues, E., Holder, A., Ramtahal, G., Umaharan, P., Muller, M.W., Mendes, F.T., Somarriba, E., 2024. CacaoFIT: the network of cacao field trials in Latin America and its contribution to sustainable cacao farming in the region. Front. Sustain. Food Syst. 8, 1370275. https://doi.org/10.3389/fsufs.2024.1370275

Oya, C., Schaefer, F., Skalidou, D., 2018. The effectiveness of agricultural certification in developing countries: A systematic review. World Dev. 112, 282–312. https://doi.org/10.1016/j.worlddev.2018.08.001

Paez-Garcia, A., Motes, C.M., Scheible, W.-R., Chen, R., Blancaflor, E.B., Monteros, M.J., 2015. Root Traits and Phenotyping Strategies for Plant Improvement. Plants 4, 334–355. https://doi.org/10.3390/plants4020334

Phillips-Mora, W., Wilkinson, M.J., 2007. Frosty Pod of Cacao: A Disease with a Limited Geographic Range but Unlimited Potential for Damage. Phytopathology 97, 1644–1647. https://doi.org/10.1094/phyto-97-12-1644

Pieruschka, R., Poorter, H., 2012. Phenotyping plants: genes, phenes and machines. Funct. Plant Biol. 39, 813–820. https://doi.org/10.1071/fpv39n11_in

Ploetz, R.C., 2007. Cacao Diseases: Important Threats to Chocolate Production Worldwide. Phytopathology 97, 1634–1639. https://doi.org/10.1094/phyto-97-12-1634

Provorov, N.A., Tikhonovich, I.A., 2003. Genetic resources for improving nitrogen fixation in legume-rhizobia symbiosis. Genet. Resour. Crop Evol. 50, 89–99. https://doi.org/10.1023/a:1022957429160

Purugganan, M.D., 2022. What is domestication? Trends Ecol. Evol. 37, 663–671. https://doi.org/10.1016/j.tree.2022.04.006

Redecker, D., Kodner, R., Graham, L.E., 2000. Glomalean Fungi from the Ordovician. Science 289, 1920–1921. https://doi.org/10.1126/science.289.5486.1920

Rengel, Z., 2002. Breeding for better symbiosis. In: Adu-Gyamfi, J.J. (Ed.), Food Security in Nutrient-Stressed Environments: Exploiting Plants' Genetic Capabilities, pp. 245–260. Springer. https://doi.org/10.1007/978-94-017-1570-6_27

Ritchie, H., 2024. The chocolate price spike: what's happening to global cocoa production? [WWW Document]. URL https://www.sustainabilitybynumbers.com/p/cocoa-prices (accessed 5.8.25).

Rosenstock, T.S., Dawson, I.K., Aynekulu, E., Chomba, S., Degrande, A., Fornace, K., Jamnadass, R., Kimaro, A., Kindt, R., Lamanna, C., Malesu, M., Mausch, K., McMullin, S., Murage, P., Namoi, N., Njenga, M., Nyoka, I., Valencia, A.M.P., Sola, P., Shepherd, K., Steward, P., 2019. A Planetary Health Perspective on Agroforestry in Sub-Saharan Africa. One Earth 1, 330–344. https://doi.org/10.1016/j.oneear.2019.10.017

Saldivar, S.O.S., Perez-Carrillo, E., 2016. Maize. Encyclopedia of Food and Health, pp. 601–609. https://doi.org/10.1016/b978-0-12-384947-2.00436-0

Säle, V., Aguilera, P., Laczko, E., Mäder, P., Berner, A., Zihlmann, U., Heijden, M.G.A. van der, Oehl, F., 2015. Impact of conservation tillage and organic farming on the diversity of arbuscular mycorrhizal fungi. Soil Biol. Biochem. 84, 38–52. https://doi.org/10.1016/j.soilbio.2015.02.005

Saleem, M., Law, A.D., Sahib, M.R., Pervaiz, Z.H., Zhang, Q., 2018. Impact of root system architecture on rhizosphere and root microbiome. Rhizosphere 6, 47–51. https://doi.org/10.1016/j.rhisph.2018.02.003

Schmidt, J.E., DuVal, A., Isaac, M.E., Hohmann, P., 2022. At the roots of chocolate: understanding and optimizing the cacao root-associated microbiome for ecosystem services. A review. Agron. Sustain. Dev. 42, 14. https://doi.org/10.1007/s13593-021-00748-2

Schmidt, J.E., DuVal, A., Puig, A., Tempeleu, A., Crow, T., 2021. Interactive and Dynamic Effects of Rootstock and Rhizobiome on Scion Nutrition in Cacao Seedlings. Front. Agron. 3, 754646. https://doi.org/10.3389/fagro.2021.754646

Shi, X., Zhao, Y., Xu, M., Ma, L., Adams, J.M., Shi, Y., 2024. Insights into plant–microbe interactions in the rhizosphere to promote sustainable agriculture in the new crops era. New Crops 1, 100004. https://doi.org/10.1016/j.ncrops.2023.11.002

Song, F., Bai, F., Wang, J., Wu, L., Jiang, Y., Pan, Z., 2020. Influence of Citrus Scion/Rootstock Genotypes on Arbuscular Mycorrhizal Community Composition under Controlled Environment Condition. Plants 9, 901. https://doi.org/10.3390/plants9070901

Song, F., Pan, Z., Bai, F., An, J., Liu, J., Guo, W., Bisseling, T., Deng, X., Xiao, S., 2015. The Scion/Rootstock Genotypes and Habitats Affect Arbuscular Mycorrhizal Fungal Community in Citrus. Front. Microbiol. 6, 1372. https://doi.org/10.3389/fmicb.2015.01372

Soto-Navarro, C.A., Harfoot, M., Hill, S.L.L., Campbell, J., Mora, F., Campos, C., Pretorius, C., Pascual, U., Kapos, V., Allison, H., Burgess, N.D., 2021. Towards a multidimensional biodiversity index for national application. Nature Sustain. 4, 933–942. https://doi.org/10.1038/s41893-021-00753-z

Steinberg, M.K., 2002. The globalization of a ceremonial tree: The case of cacao (*Theobroma cacao*) among the Mopan Maya. Econ. Bot. 56, 58–65. https://doi.org/10.1663/0013-0001(2002)056[0058:tgoact]2.0.co;2

Strock, C.F., Schneider, H.M., 2022. Root Phenes for Improving Nutrient Capture in Low-Fertility Environments**.** Plant Breeding Reviews 46, 587–625. https://doi.org/10.1002/9781119874157.ch9

Suárez, L.R., Salazar, J.C.S., Casanoves, F., Bieng, M.A.N., 2021. Cacao agroforestry systems improve soil fertility: Comparison of soil properties between forest, cacao agroforestry systems, and pasture in the Colombian Amazon. Agric., Ecosyst. Environ. 314, 107349. https://doi.org/10.1016/j.agee.2021.107349

Toth, R., Toth, D., Starke, D., Smith, D.R., 1990. Vesicular-arbuscular mycorrhizal colonization in *Zea mays* affected by breeding for resistance to fungal pathogens. Can. J. Bot. 68, 1039–1044. https://doi.org/10.1139/b90-131

Trivedi, P., Leach, J.E., Tringe, S.G., Sa, T., Singh, B.K., 2020. Plant–microbiome interactions: from community assembly to plant health. Nature Rev. Microbiol. 18, 607–621. https://doi.org/10.1038/s41579-020-0412-1

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

Vierheilig, H., Coughlan, A.P., Wyss, U., Piché, Y., 1998. Ink and Vinegar, a Simple Staining Technique for Arbuscular-Mycorrhizal Fungi. Appl. Environ. Microb. 64, 5004–5007. https://doi.org/10.1128/aem.64.12.5004-5007.1998

Wang, Q., Liu, M., Wang, Z., Li, J., Liu, K., Huang, D., 2024. The role of arbuscular mycorrhizal symbiosis in plant abiotic stress. Front. Microbiol. 14, 1323881. https://doi.org/10.3389/fmicb.2023.1323881

Wang, X., Zhang, M., Sheng, J., Feng, G., Kuyper, T.W., 2023. Breeding against mycorrhizal symbiosis: Modern cotton (*Gossypium hirsutum* L.) varieties perform more poorly than older varieties except at very high phosphorus supply levels. J. Integr. Agric. 22, 701–715. https://doi.org/10.1016/j.jia.2022.08.004

Wissuwa, M., Mazzola, M., Picard, C., 2009. Novel approaches in plant breeding for rhizosphere-related traits. Plant Soil 321, 409. https://doi.org/10.1007/s11104-008-9693-2

Yang, N., Wang, Y., Liu, X., Jin, M., Vallebueno-Estrada, M., Calfee, E., Chen, L., Dilkes, B.P., Gui, S., Fan, X., Harper, T.K., Kennett, D.J., Li, W., Lu, Y., Ding, J., Chen, Z., Luo, J.,

Mambakkam, S., Menon, M., Snodgrass, S., Veller, C., Wu, Shenshen, Wu, Siying, Zhuo, L., Xiao, Y., Yang, X., Stitzer, M.C., Runcie, D., Yan, J., Ross-Ibarra, J., 2023. Two teosintes made modern maize. Science 382, eadg8940. https://doi.org/10.1126/science.adg8940

Yu, L., Zhang, H., Zhang, W., Liu, K., Liu, M., Shao, X., 2022. Cooperation between arbuscular mycorrhizal fungi and plant growth-promoting bacteria and their effects on plant growth and soil quality. PeerJ 10, e13080. https://doi.org/10.7717/peerj.13080

Chapter 2. Genotype characterization of *Theobroma cacao* germplasm from farms in Panama

Natalie Ferro-Lozano, Dapeng Zhang, Luis Mejia, Thomas W. Kuyper, Guusje Bonnema



Chapter 2

Abstract

Cacao farms in Panama are traditionally managed, primarily by indigenous communities, with historical records dating back to 1500 BC. However, the specific cacao germplasm in use and the associated management practices remain largely undocumented. This study aims to characterize the cacao germplasm cultivated in Panamanian farms and to understand how it relates to management and environmental conditions. Despite its significant biogeographical relevance for cacao semi-domestication, Panama lacks substantial research on the germplasm utilized in its cacao farms. We addressed three research questions: (i) What germplasm is used, and how does it compare to Central and South American diversity? (ii) What does the germplasm reveal about management, exchange, and maintenance? (iii) Which environmental factors are associated with specific germplasms? We surveyed 16 farms across five regions in Bocas del Toro and in Las Minas, Herrera. Root and leaf samples from five trees per farm were analysed using a 96-SNP array. Environmental and soil data were also collected. In Bocas del Toro, cacao genotypes showed a continuous admixture of all major genetic groups, whereas genotypes in Herrera were primarily Amelonado and Criollo, including individuals with pure Criollo ancestry. Limited germplasm exchange between regions resulted in distinct genetic compositions. Farmers predominantly used seeds to establish plantations, though grafting was more common in Bocas del Toro (55%) than in Herrera (16%). A LASSO regression analysis identified total nitrogen and pH as the most important environmental variables associated with genetic differences. This study highlights the genetic diversity of cacao in Panama and its potential value for national industry development.

Introduction

Cacao (*Theobroma cacao*, 2n = 2x = 20) is a perennial crop of major economic importance, cultivated between 10ºN and 10ºS of the Equator. Its primary product, chocolate, fuels a global-industry value of 47 billion U$D in 2021 with continued growth projected (Fortune Business Insights, 2024). The genetic diversity of a plant species directly influences its potential for cultivation, serving as a critical source of genetic variation for disease resistance, yield improvement, and other agronomic traits. Cacao's centre of origin lies in the Upper Amazon Basin, at the frontiers of Peru, Ecuador, and Colombia (Lanaud et al., 2024). From there, it spread eastward toward Brazil and northward, ultimately reaching Central America where it was semi-domesticated by the Mayas and Olmecs around 3,500 years ago (Clement et al., 2010; Motamayor et al., 2002). However, as a highly heterozygous, allogamous species historically associated with American civilizations, its classification has long posed significant challenges.

Traditionally, cacao was categorized into three main groups: Criollo, Forastero, and Trinitario types. Separated by the Panamanian isthmus, Criollo types were domesticated by Mesoamerican natives in Central America, while Forastero types referred to all cacao in the Amazon Basin. These populations developed unique and consistent traits, which led to their historical classification as separate subspecies— Criollo types corresponding to *T. cacao* subsp. *cacao* and Forastero types to *T. cacao* subsp. *sphaerocarpum* (Cuatrecasas, 1964). Due to Criollo's disease susceptibility, the Cocoa Research Centre in Trinidad & Tobago hybridized Criollo with Forastero individuals, creating a new Trinitario group. This group combined the fine flavour of Criollo and the disease resistance of Forastero. With the advent of molecular techniques, recent genetic studies have provided deeper insights into the origin, evolution, and domestication of the cacao tree. Now the Criollo group is known to be another genetic group with a South American origin instead of a subspecies (Motamayor and Lanaud, 2002). Additionally, Forastero group is now composed of ten distinct cacao clusters, including Marañon, Curaray, Iquitos, Nanay, Contamana, Amelonado, Purus, Nacional, and Guiana (Motamayor et al., 2008), and more recently, Caquetá (Fouet et al., 2022).

Numerous cacao cultivars have been selected and developed as scions from these groups through plant breeding to meet agricultural and social needs such as disease resistance, yield optimization, and flavour enhancement. This effort involves global organizations like CacaoNet (2025), CIRAD (2025) and World Cocoa Foundation (2025). Cacao breeding has only recently begun, in 1920, and faces many challenges typical of long-lived perennial crops (DuVal et al., 2017). These include the long generation time, the difficulty of developing parental populations due to complex gametophytic and sporophytic self-incompatibility leading to high heterozygosity, and the recalcitrant nature of cacao seeds (Berjak and Pammenter, 2017), which cannot be stored long-term. Still, scion genotypes have been either developed by breeding and selecting or selected from the wild, with the most notable

genotypes in Latin America including CCN-51, ICS clones, TCS genotypes, CATIE selections, and IMC series.

While scions have been developed or selected for their desirable traits, rootstocks have traditionally received less attention in breeding programs. Farmers commonly use locally seed-sown trees with low productivity or quality as rootstocks during farm maintenance (Schmidt et al., 2021). Critical gaps remain in research, such as the optimal timing for rootstock renewal and whether neglecting this practice may reduce vigour or yield. Nevertheless, clones CAU37, CAU39, CAU42, PA121, PA46, PA150, Pound7, Pound12, EET399, EET400, EET96, SPA9, and especially IMC67, are recommended for use as rootstocks (Echeverri-Rodriguez, 2006; Montoya et al., 2022). A rootstock should have tolerance to climatic conditions and diseases in addition to a good vegetative vigour. As research on cacao rootstocks has expanded—particularly to address cadmium accumulation (Almeida et al., 2022; Fernández-Paz et al., 2021; Galvis et al., 2023), improve nutrient uptake efficiency (Galvis et al., 2023; N'zi et al., 2023; Schmidt et al., 2021; Setyawan et al., 2018), and enhance disease resistance (Ribeiro et al., 2016)— understanding their role within cacao agroforestry systems (CAFS) becomes increasingly relevant.

Cacao is primarily cultivated by smallholders across 11.6 million hectares over 60 countries worldwide, producing 5.6 million tons annually, with over 90% of farmers managing 2–5-hectare plots (FAOSTAT, 2023b). Cacao cultivation requires between 45% – 75% shade during its development for which it is grown in CAFS with varying numbers of other tree species, from a few timber species to a multiple array of timber and fruit tree species. Cacao trees can be propagated through seeds, root cuttings, or grafting with the latter being the most common to accelerate development by half the time, as seed-to-harvest takes up from three to five years. Smallholder farmers in Panama originally grew cacao trees from seeds. Recently, in few areas where technological development is increasing, they select trees with the best yield, resistance, and quality to use as scions when it is time to renovate their farms. These scions are then grafted onto low-productivity trees, as previously mentioned. Knowledge of both the genotypes of scion and rootstock of trees provides insight into the methods used for farm regeneration and maintenance. Although farms are recommended to be renovated every 20 years (Yacomelo-Hernández et al., 2021), farms in Panama can reach 40 years without renovation.

While regeneration practices help maintain the vigour and productivity of plantations, cacao cultivation still faces significant challenges. Diseases, pests, climate change, and international regulations threaten productivity. *Moniliophthora roreri* (frosty pod), *M. perniciosa* (witches' broom), and *Phytophthora* spp (black pod) (Evans, 2007) account for a conservative 30% yield loss, with an additional 30% lost to squirrels. Climate change exacerbates production risks by altering rainfall patterns and constraining cultivation to more humid areas (Júnior et al., 2023; Kaba et al.,

2022). Additionally, EU regulations on cadmium content (2019) and deforestation-free imports (2020) impose strict standards (Government of the Netherlands, 2024). Even though both regulations aim to ensure public and environmental health, they lack assistance to help farmers adjust to new rules or practices. Assistance could include financial aid, technical training, access to resources, and guidance on compliance with the regulations.

In Panama, these global challenges are further deepened by unique local and historical factors, shaping cacao cultivation's trajectory. Cacao has been widely cultivated across Panama, influenced by the social and spiritual traditions of indigenous communities, who have traditionally grown it for their own consumption. Indigenous communities have cultivated cacao for centuries, preserving native varieties through isolation due to limited infrastructure. Commercial production began in 1920 when the United Fruit Company (UFC) introduced Trinitario varieties, later supplemented by CATIE genotypes (1975 and 1979), from multiple Central American countries. However, increasing disease pressure and declining cocoa prices eventually caused a significant drop in cultivation for international export. Today, cacao remains vital to Panama's rural economy but faces productivity challenges. Approximately 6,000 hectares are cultivated by 1,500 farmers, predominantly (90%) indigenous growers in Bocas del Toro and 10% smallholders in Herrera (COCABO, pers. comm.). Despite winning international flavour awards (Cacao of Excellence Awards Excellence, 2021), yields average just 103.1 kg/ha—far below the central American mean of 473 kg/ha (FAOSTAT, 2023b), underscoring the need for improved practices.

Harnessing Panama's unique cacao genetic diversity could address these yield gaps. The Isthmus — a hypothesized centre of cacao semi-domestication (Lanaud et al., 2024; Motamayor et al., 2002)— acts as a geographical bottleneck likely facilitating the mixing of North and South American populations (Cuatrecasas, 1964). Additionally, the admixture of introduced and native cacao leaves the genetic ancestry of modern Panamanian cacao largely unresolved. This germplasm is predicted to be regionally distinct yet reflect a rich mixture of diverse cacao genetic groups — a result of Panama's position as the geographical frontier between Criollo and other genetic lineages. Characterizing this diversity is crucial because no study has yet analysed Panamanian germplasm, despite its potential to inform breeding programs, on-farm conservation, and climate-resilient cultivation strategies. Furthermore, linking genetic groups to local environmental factors could reveal which environmental and management practices have coevolved with these genetic groups. We hypothesized that: (i) Panamanian germplasm in farms is a varied mixture of known cacao genetic groups in the biodiversity pool; (ii) germplasm exchange between Bocas del Toro and Herrera is minimal; (iii) grafting in Bocas del Toro is more prevalent than in Herrera as a result of greater popularity in cacao cultivation and; (iv) soil and climatic characteristics differentially associate with

tree genetic backgrounds across the Atlantic coastal (Bocas del Toro) and Pacific mountainous (Herrera) regions sampled.

Materials & Methods

Sample Collection and DNA extraction

Five cacao trees were assessed from each of 16 farms from two provinces in Panama, Bocas del Toro and Herrera (Figure 1). A total of 81 trees were marked for assessment: 25 (5x5) in Las Minas, Herrera, and 56 ((5 x 10) + 6) in five different regions of Bocas del Toro (Supplementary material A). Farms 9 and 10 in Bocas del Toro were owned by companies and not local farmers, from here on referred to as commercial farms. Each cacao tree had its leaf and root tissue sampled separately (representing scion and rootstocks, or above and belowground) and dried immediately in silica gel for further DNA analysis. An individual tree is represented by two genotypes (leaf and root). The DNA extraction and subsequent genotyping were done in the USDA facilities in Maryland, USA. For both seedlings and rooted cuttings, their DNA profiles should be identical; however, if trees are grafted, one expects a different DNA profile between root DNA (representing the rootstock) and leaf DNA (representing the scion).

Furthermore, DNA fingerprint profiles from 150 well-characterized cacao samples were included as reference material (Supplementary material B). These reference profiles were previously developed by the Sustainable Perennial Crops Laboratory (SPCL) at USDA and documented in earlier studies (Cosme et al., 2016; Ji et al., 2013; Lindo et al., 2018). The reference samples encompassed individual genotypes representative of established cacao genetic populations or germplasm collections (Figure 1).

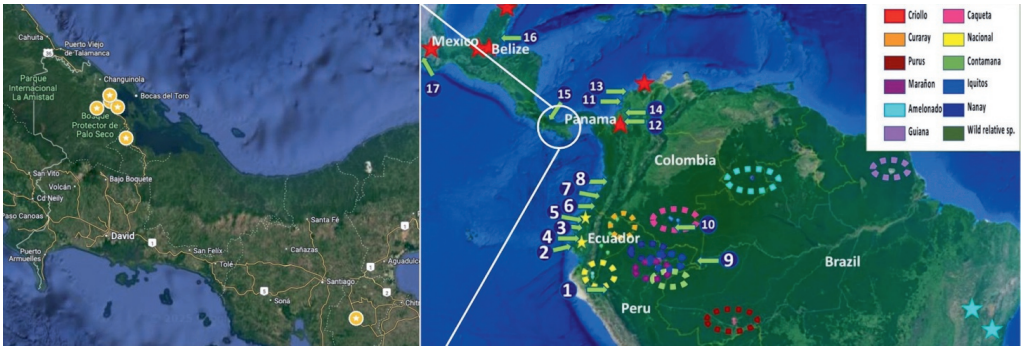


Figure 1. Left image depicts sampled Panamanian regions within the zoomed white circle. Stars represent the historical area of cultivation of Criollo (red), Amelonado (light blue) and Nacional (yellow). The wild genetic groups are represented by segmented circles, except for Criollo, which is the semi-domesticated genetic group and there is no close wild-type genetic group. Image extracted from (Lanaud et al., 2024).

SNP markers and Genotyping

The methodology used for 96-SNP genotyping was as described in (Lindo et al., 2018). SNP markers were identified in expressed sequence tags (ESTs) from cDNA sequences across various cacao organs, genotypes, and environments, representing the transcriptome across ten chromosomes (Argout et al., 2011, 2008). The 96 SNPs used for this panel were selected based on the level of polymorphism and their distribution across the ten chromosomes of cacao and were previously reported by Fang (2014) and Ji (2013). The 96 SNPs, the flanking sequences, and their distribution across the ten cacao chromosomes are listed in Supplementary material C.

SNP data analysis

Raw data were organized in Microsoft Excel 2024 (Version 16.91) by locus and sample. Descriptive statistics measuring informativeness and quality of the 96 successfully amplified SNP loci were calculated using the GenAlEx 6.5b2 program (Peakall and Smouse, 2012; Smouse et al., 2017). We analysed observed heterozygosity (H_o) per genotype (Supplementary material D), and compared profiles between Bocas and Herrera regions, and leaf and root tissues sampled from the same trees (Supplementary material E). To compare different and single-origin trees between regions we did a Chi-square. Two genotypes were classified as genetic duplicates if their SNP profiles differed by ≤6% differences (i.e. ≤11 alleles). This 6% threshold was empirically determined by analysing four known clonal pairs, representing the maximum observed variation between replicates (Genetic Duplicates Group, GDG; Supplementary Material D). Genetic duplicates found across distinct trees present clonal propagation (a scion or rooted cuttings from the same mother tree) but can also represent close kinship (closely related siblings).

To assess genetic relationships among Panamanian cacao genotypes (Supplementary material F) and reference samples (Supplementary material G), we conducted distance-based multivariate analyses, including hierarchical clustering, principal component analysis (PCA), and kinship matrices. For the hierarchical clusters the bootstrap value was represented as approximately unbiased p-value (AU) from 1000 iterations. Pairwise genetic distances were computed using Rstudio Version 2024.04.2+764 and R ver. 4.4.1 (www.r-project.org) using the method described in (James et al., 2021). Population structure was determined using model-based Bayesian clustering in *STRUCTURE v2.3.4* (Pritchard et al., 2000). The analysis included reference samples selected based on their documented history in Central America and the Caribbean cacao cultivation (Lanaud et al., 2024).

Environmental data collection

Field characterization was conducted across 16 farms. We recorded elevation, average tree age, farm type, cacao tree density, shade percentage, and neighboring species. For the latter, we identified the three nearest trees to each sampled tree, totalizing 15 trees per farm except in farm BDT11 (18

trees) (Supplementary material H). Soil analysis below each tree included texture (% sand, % silt, % clay), pH (in water), organic matter (%), nitrogen (N; %), available phosphorus (P; mg/g), C:N ratio, potassium (K; mg/l), and exchangeable cations (Ca, Mg, Al, Mn, Fe, Cu; Cmol/kg or mg/l), alongside CEC (cation exchange capacity), aluminum saturation and base saturation. Methods have been described according to Villarreal (2015). To avoid overfitting and reduce dimensionality, we applied Horn's parallel analysis (HPA) with 720 iterations (30 iterations per variable), retaining axis with an adjusted eigenvalue > 1 and variables with largest contributions (cos2). Hereafter, we removed highly collinear variables between each other with correlation matrices. From the initial 18 soil variables and 6 farm traits, 13 key environmental variables remained for further analysis (Figure 4): Elevation, percentage of sand (% sand), percentage of silt (% silt), percentage of clay (% clay), pH, organic matter percentage (OM%), total nitrogen (TN %), available P (mg/g), calcium (Cmol/kg), magnesium (Cmol/kg), aluminum (Cmol/kg), iron (mg/l), and base saturation. HPA, PCA and correlation matrices were processed using paran, FactoMineR, Hmisc and corrplot packages in R 4.4.2. The FactoMineR package z-transforms the data (average = 0, standard deviation = 1) before analysis.

Genetic and environmental modelling

We performed Ridge regression to explore potential associations between local germplasm and environmental factors (Hoerl and Kennard, 1970). Thirteen environmental variables averaged over the 16 plantations were used as predictors and the STRUCTURE ancestry-derived PC coefficients as multivariate responses. This approach was particularly suited to our data because: (1) Ridge regression handles correlated predictors while retaining all pre-selected variables, and (2) it effectively models continuous multivariate response data of this sample size. We considered values above an absolute value of 0.01 to be meaningful, as this corresponds to a 1% decrease or increase in the representation of a genetic group for each unit change in the environmental predictor.

Results

Global genetic context of Panamanian germplasm

Analysing SNP profiles of both root and leaf tissues of 80 trees from 16 plantations in Bocas del Toro and Herrera and the trees from reference populations, resulted in identification of 10 robust clusters (Figure 2, Supplementary material G). Eight clusters (AU > 95%) corresponded to the reference genotype populations: Guiana, IMC Iquitos, Nanay, Criollo, SCA Contamana, Nacional, LCT EEN Curaray, and Purus. The other 2 clusters formed subgroups within a larger cluster (AU = 79%) containing all Amelonado reference genotypes, most of the Panamanian germplasm, and 2 individuals from Nanay reference populations. The Parinari reference group had a bootstrap value below the generally accepted credibility value of 80% (AU = 74%).

Considering the STRUCTURE analysis, Panamanian germplasm was more admixed than the reference populations (Figure 2). Panamanian germplasm was grouped in the Amelonado cluster as an admixed gradient between Amelonado and Criollo groups with notable contributions of Nacional, IMC (Iquitos), and Nanay. Importantly, the germplasm from Panama forms a non-significant cluster (AU = 0%) indicating no reproducible branch support across 1000 bootstrap iterations. Only BdT10.5-R clustered within the Nanay group, and all trees in Herrera farm 14, Her15.4-0 and Her16.2-0, represented as GDG-10, clustered within the Criollo group.

Panamanian germplasm

Observed heterozygosity (H_o) across 96 SNPs ranged between 0% to 75%, with a mean of 32% (Table 1) indicating low to moderately high genetic diversity among Panamanian cacao farms (Supplementary material D). From all samples, 39 genotypes had low heterozygosity (0-20%), 99 genotypes were between 21% – 49% heterozygosity and the other 24 genotypes were above 50%. The observed heterozygosity was similar across regions, with Bocas del Toro showing a H_o of 33% and Herrera a H_o of 29%. Additionally, no significant differences were found between the aboveground (leaves) (H_o = 34%) and belowground (roots) organs of the trees (H_o = 29%).

Table 1.Summary of observed heterozygosity across Panamanian samples (Supplementary material D).

Category	Mean Ho	SD Ho	Min Ho	Max Ho
Overall	32%	18%	0%	75%
Leaf (scion)	29%	16%	0%	69%
Root (rootstock)	34%	19%	0%	75%
Bocas del Toro	33%	17%	1%	75%
Herrera	29%	20%	0%	65%

Hierarchical cluster analysis revealed low overall genetic relatedness, with 50 of 162 genotypes clustering into four well-supported groups (bootstrap >95%; Supplementary Material G). The other 112 genotypes were an admixed gradient with groups of bootstrap = 0. Population structure analysis separated the data into groups based on shared allele frequencies. This analysis showed distinct genetic backgrounds between regions (Figure 3). Genotypes from Bocas del Toro exhibited a higher admixture, dominated by Amelonado, Nacional and Nanay groups, and reduced Criollo group while in Herrera Criollo and Amelonado were prevalent (Table 2). The difference in SCA (Contamana) should be taken lightly since <5% difference among structural genetic groups is based on allelic frequencies and would require a larger difference to be realistically (not statistically) considered different. Two groups of genotypes stood out as pure Amelonado in Farm 5 or pure Criollo in farm 14 (Supplementary Material H – Figure 3). Although statistical tests for contribution of

genetic groups between regions were significant (Supplementary material I), this result was mostly due to the increased presence of Criollo in Herrera. Statistical tests for contribution of genetic groups between root and leaf tissues were not significant.

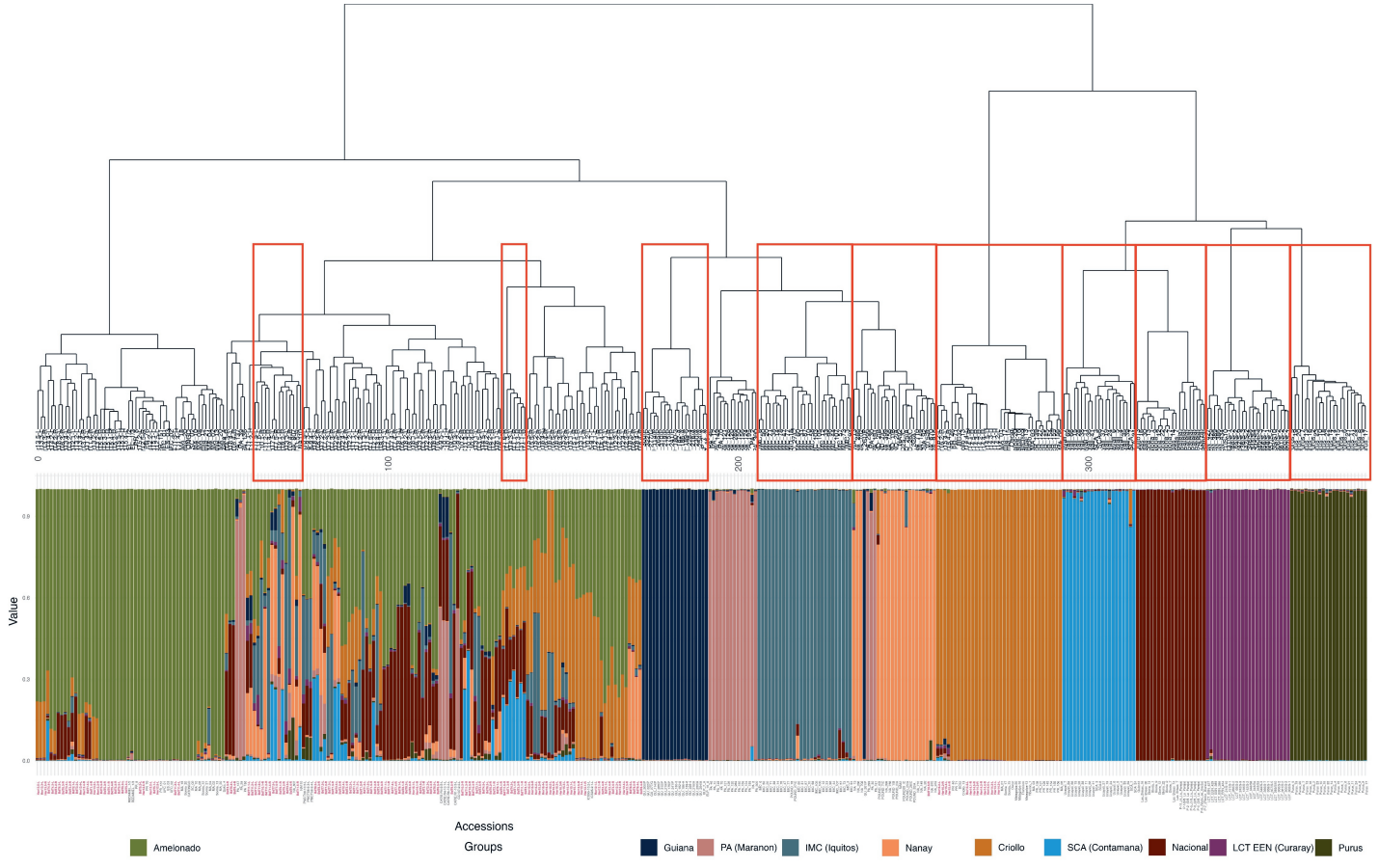


Figure 2. Hierarchical cluster and structure plot of the Panamanian germplasm and the reference populations. Below the individual labels, there are the labels that correspond to the genetic groups. Red rectangles show bootstrap values over 95%.

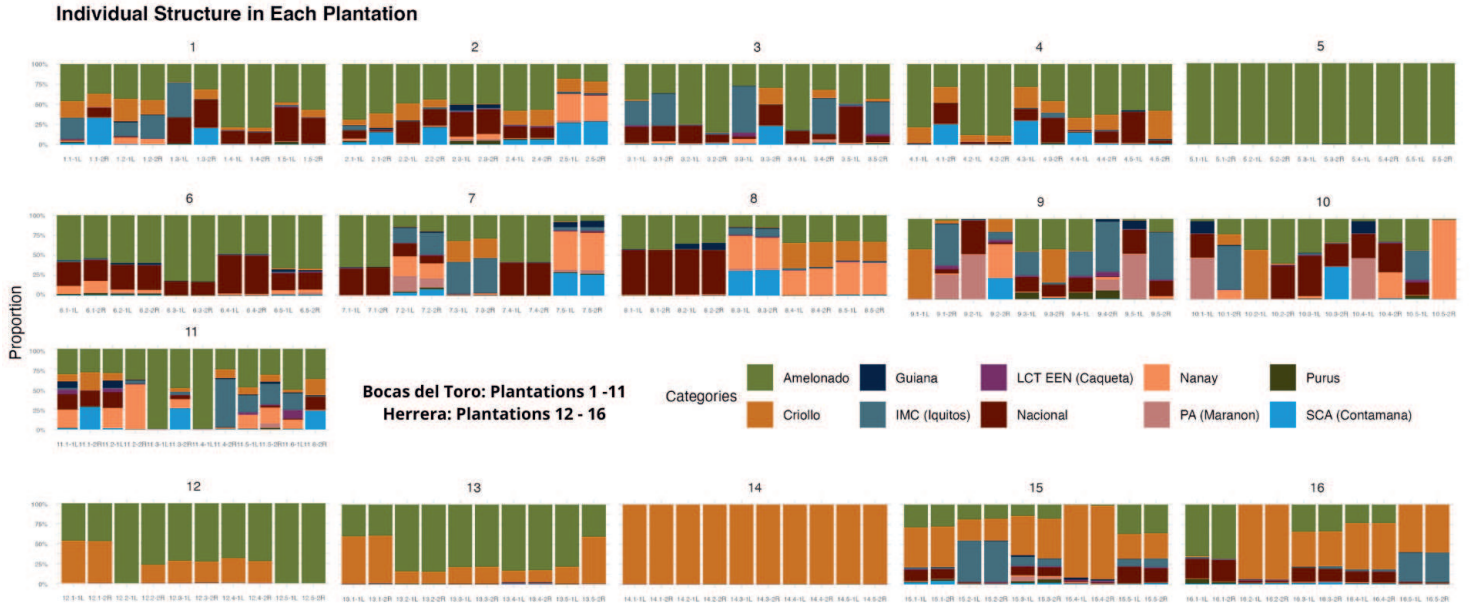


Figure 3. Panamanian samples structure analysis. For each plantation there are 10 columns, representing five pairs of above or belowground genotypes of the tree, interspersed.

Table 2. Two- tailed t-tests between Mean presence of the genetic groups across regions and across leaves or roots. The asterisk shows significance between the groups.

Region	*IMC Iquitos*	*Nacional*	*Nanay*	*Guiana*	*Purus*	*LCT (Curaray)*	*PA (Maranon)*	*Amelonado*	*SCA (Contamana)*	*Criollo****
Bocas del Toro	9.08%	16.04%	7.76%	1.28%	0.80%	0.88%	3.13%	47.89%	4.93%	***8.21%***
Herrera	4.68%	4.33%	0.21%	0.25%	0.46%	0.39%	0.41%	37.81%	0.53%	***50.89%***
Leaves	6.59%	13.26%	4.25%	1.28%	0.77%	0.84%	3.36%	46.90%	*2.06%*	20.69%
Roots	8.86%	11.60%	6.62%	0.64%	0.62%	0.62%	1.22%	42.66%	*5.09%*	22.08%

Maintenance and exchange of genotypes across farms

Genetic profiling revealed that 43% of trees were pairs of rootstock-scion with different origins (grafted), while 57% were of single genetic origin (Table 3). Across regions we observed a significantly greater number of grafted trees in Bocas del Toro (55%) than in Herrera (16%) ($p < 0.001$; Supplementary material E). Among the 81 trees, there were 46 trees of single origin, and 35 trees of different origin. Single-origin trees are either derived from rooted cuttings or from seeds, while different origin trees are from either commercially imported or farm-selected scions grafted on low-quality/productivity trees.

Table 3. Percentages of Trees that have different vs single genetic origins across regions and plantations

	Different	Single
% of Trees	43%	57%
Bocas del Toro	55%	45%
Herrera	16%	84%

We compared DNA fingerprints across the 162 leaf and root samples and identified 10 GDGs (genetic duplicate groups – Table 4). GDGs are nearly identical genotypes based on similarity in SNP-profiles, which represent clones or close kinship (Supplementary Material D). Two GDGs were composed of single genetic groups: GDG-3 of Amelonado genetic background, and GDG-10 of Criollo genetic background (Figure 2 / Figure 3). These GDGs were represented by single-genotype trees and were either rooted clonal cuttings or very closely related siblings from seeds since their heterozygosity was very low (0.01 – 0.156). Interestingly, these clones were found in different farms, and even different regions. The other GDGs were found across scions and rootstocks. GDG-1, GDG-2 and GDG-4, were clonal scions selected from the wild as part of the IDIAP initiatives to protect and improve native cacao germplasm (Gutierrez, 2020). GDG 6, represented by roots of trees in two different Bocas plantations likely represent either rooted cuttings later grafted, or two closely related siblings. GDG 5, 7, 8 and 9 are pairs of scion and rootstock from different trees. Additionally, we identified commercial scions corresponding to 6 cultivars recommended by CATIE (Phillips-Mora et al., 2012) in the commercial farms of Bocas del Toro 9 and 10. While local farms rely on readily available germplasm adapted to regional conditions, commercial farms often introduce imported cacao varieties, highlighting the balance between agricultural expansion and the preservation of local biodiversity.

Table 4. Genetic Duplicate Groups.

Genetic Duplicate Group	Region	Farm and tree	Ho	Comment
GDG-1	Bocas del Toro	11.1	0.333	Scion IDIAP 26-61
GDG-1	Bocas del Toro	11.2	0.333	Scion IDIAP 26-61
GDG-2	Bocas del Toro	11.3	0.021	Scion IDIAP 26-60
GDG-2	Bocas del Toro	11.4	0.021	Scion IDIAP 26-60
GDG-3	Bocas del Toro	5.1	0.010	Seedling
GDG-3	Bocas del Toro	5.2	0.010	Seedling
GDG-3	Bocas del Toro	5.3	0.010	Seedling
GDG-3	Bocas del Toro	5.4	0.010	Seedling
GDG-3	Bocas del Toro	5.5	0.010	Seedling
GDG-3	Herrera	12.2	0.010	Scion
GDG-3	Herrera	12.5	0.000	Seedling
GDG-4	Bocas del Toro	11.5	0.281	Scion IDIAP 26-59
GDG-4	Bocas del Toro	11.6	0.271	Scion IDIAP 26-59
GDG-5	Bocas del Toro	3.4	0.250	Scion
GDG-5	Bocas del Toro	3.2	0.229	Rootstocks
GDG-6	Bocas del Toro	3.3	0.740	Rootstocks
GDG-6	Bocas del Toro	4.1	0.708	Rootstocks
GDG-7	Bocas del Toro	4.3	0.698	Scion
GDG-7	Bocas del Toro	11.1	0.740	Rootstocks
GDG-8	Herrera	12.3	0.302	Scion
GDG-8	Herrera	12.2	0.323	Rootstocks
GDG-9	Herrera	12.4	0.333	Scion
GDG-9	Herrera	12.3	0.333	Rootstocks
GDG-10	Herrera	14.1	0.073	Seedling
GDG-10	Herrera	14.2	0.073	Seedling
GDG-10	Herrera	14.3	0.073	Seedling
GDG-10	Herrera	14.4	0.073	Seedling
GDG-10	Herrera	14.5	0.073	Seedling
GDG-10	Herrera	15.4	0.156	Seedling
GDG-10	Herrera	16.2	0.104	Seedling
ICS95	Bocas del Toro	9.1	0.438	Commercial scion
ICS95	Bocas del Toro	10.2	0.458	Commercial scion
PMCT58	Bocas del Toro	9.3	0.292	Commercial scion
PMCT58	Bocas del Toro	9.4	0.292	Commercial scion
PMCT58	Bocas del Toro	10.5	0.281	Commercial scion
CATIE-R4	Bocas del Toro	9.5	0.375	Commercial scion
CATIE-R4	Bocas del Toro	10.1	0.385	Commercial scion

Associations between environmental factors and genetic groups

After performing a PCA analysis based on the structural genetic groups, the PC1 axis separated groups of Amelonado and Criollo, while PC2 axis separated Amelonado/Criollo versus all other groups (Supplementary material J). As such, environmental associations along the PC1 axis indicate a positive relationship with Criollo and a negative one with Amelonado—or vice versa. Associations along the PC2 axis reflect a positive relationship with Criollo/Amelonado and a negative relationship with the remaining groups—or the reverse. Among environmental variables, used as predictors in the Ridge model, there was no significant difference found between regions (Figure 4). Ridge regression revealed relevant associations between 3 selected predictor environmental variables and the genetic structure of admixed trees derived from PC components, either across all plantation samples (scions and rootstocks, $n = 32$) or separately for scions ($n = 16$) and rootstocks ($n = 16$) (Table 5; Supplementary Material K). The regression models were done separately for PC1, explaining the

contrast between Criollo and Amelonado, and PC2, explaining the remaining variation between Criollo/Amelonado and the other groups. The model explained 14-62% of variance (R^2 = 0.14-0.62) in genetic group distribution across environmental gradients.

Across scion, rootstock, and combined associations, Total Nitrogen (TN) was associated to a decreased Criollo presence and increased Amelonado presence shown by PC1, while PC2 shows TN association with decreased Amelonado/ Criollo presence and increased presence of all other genetic groups. TN showed a two to ten-fold larger effect compared with all other environmental characteristics. pH showed association only in leaves to a decreased Criollo presence and increased Amelonado presence shown by PC1, but in leaves and roots showed association with increased Amelonado/ Criollo presence and decreased presence in all other groups shown by PC2. Organic matter percentage showed associations only in combined analysis to a decreased Criollo presence and increased Amelonado presence shown by PC1, and with decreased Amelonado/ Criollo presence and increased presence in all other groups shown by PC2.



Figure 4. Environmental characterization of 80 trees from 16 cacao farms through a principal component analysis.

Table 5. PCA-Ridge Model. The numbers displayed blue is where the environmental factor is related to a fractional decrease of respective genetic group, and in yellow, when the environmental factor is related to an increase in percentage of the respective genetic group. PC1 explains the variation between Amelonado and Criollo mainly, and PC2 explains the variation between Amelonado/Criollo and all other groups.

	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
PC1										
Leaves										
TN (%)	-0.0006	0.0010	0.0008	0.0000	-0.0001	0.0001	-0.0008	0.0218	0.0003	-0.0223
pH	0.0003	-0.0005	-0.0004	0.0000	0.0001	0.0000	0.0004	-0.0101	-0.0001	0.0104
Roots										
TN (%)	-0.0020	0.0026	0.0010	0.0001	-0.0001	-0.0002	-0.0002	0.0259	0.0009	-0.0279
General										
TN (%)	-0.0036	0.0048	0.0021	0.0002	-0.0003	-0.0001	-0.0013	0.0659	0.0013	-0.0689
OM %	-0.0009	0.0013	0.0006	0.0001	-0.0001	0.0000	-0.0003	0.0172	0.0003	-0.0180
PC2										
Leaves										
pH	-0.0075	-0.0171	-0.0056	-0.0020	-0.0010	-0.0006	-0.0084	0.0223	-0.0014	0.0213
TN (%)	0.0176	0.0399	0.0130	0.0046	0.0023	0.0015	0.0196	-0.0520	0.0032	-0.0496
Roots										
pH	-0.0165	-0.0118	-0.0145	-0.0008	-0.0008	-0.0010	-0.0029	0.0287	-0.0065	0.0260
TN (%)	0.0265	0.0189	0.0233	0.0013	0.0013	0.0015	0.0046	-0.0460	0.0104	-0.0417
General										
pH	-0.0343	-0.0416	-0.0299	-0.0038	-0.0025	-0.0022	-0.0149	0.0731	-0.0119	0.0680
OM %	0.0088	0.0107	0.0077	0.0010	0.0006	0.0006	0.0038	-0.0188	0.0031	-0.0175
TN (%)	0.0516	0.0625	0.0449	0.0058	0.0037	0.0034	0.0224	-0.1099	0.0179	-0.1021

Discussion

Our study confirmed that (i) the Panamanian germplasm in farms is an admixture of all known cacao genetic groups in the American gene pool; (ii) the exchange of cacao germplasm between Bocas del Toro and Herrera is minimal; (iii) grafting in Bocas del Toro is more popular compared with Herrera; and (iv) germplasm selection is correlated with soil properties.

Farm Germplasm

The heterozygosity across all samples ranged between 0% to 75%, showing the wide range of modes of reproduction and the variation among cacao trees in Panamanian farms. From all samples, 24% had low heterozygosity (0-20%) while the rest had moderate to high heterozygosity (between 21% - 75%). Farms 5 and 14, which have predominantly low-heterozygosity trees, indicate reduced allelic and genetic variation, potentially making the trees more susceptible to diseases and environmental changes. In contrast, farms with a mix of low- and high-heterozygosity trees, or those with predominantly moderate to high heterozygosity, exhibit greater allelic diversity and genetic variation, enhancing adaptability, resilience, and overall population fitness. There were no significant differences in the heterozygosity between the regions Herrera and Bocas, and between scion and rootstock. On the other hand, there were farms with only high or low heterozygosity trees (Supplementary material E).

The phylogenetic trees showed that the Panamanian germplasm formed a separate branch without credible support (a bootstrap value of 0). Lack of credible branch support typically occurs when the data lack the resolution needed to clarify this part of the tree, or the individuals are highly admixed with each other, probably being assigned randomly each time. Based on the structure analysis, we conclude that the second explanation is more likely.

The Panamanian germplasm is an admixture between Amelonado and Criollo, with a high percentage of Nacional, IMC (Iquitos) and Nanay. Material in Bocas del Toro consisted predominantly of Amelonado, Criollo, Nacional, IMC (Iquitos) and Nanay, placed in decreasing order. These last three groups in addition to Guiana, LCT-EEN (Curaray), PA (Marañon), SCA (Contamana) were significantly more present in Bocas del Toro than in Herrera. Contrastingly, Herrera genotypes had a dominance of Amelonado and Criollo genetic groups, with some genotypes detected as 100% Criollo. The varied mixture present in Bocas del Toro could partly be explained by the introduction of UF clones and CATIE clones between 1910 - 1980. The UF clones are composed of Amelonado, Criollo, and Nacional (Mata-Quirós et al., 2018) and the CATIE clones, sampled in this study (Phillips-Mora et al., 2012), are composed of Amelonado, Parinari, Nacional, Criollo, and IMC (Iquitos). Still the presence of Guiana, LCT EEN (Curaray), and Scavina (Contamana) groups

cannot be explained by this material introduction, since these are not present in the UF clones or the recommended material by CATIE.

The presence of pure Criollos trees, as found in Herrera, is rare. These Criollo trees are known for disease susceptibility caused by their lower anthocyanin content and thus less pungent biochemical composition (Elwers et al., 2009; Qin et al., 2017), as well as for their low genetic variability and homozygosity (Bekele and Phillips-Mora, 2019; Lachenaud and Motamayor, 2017). The theory proposed by experts in Panama suggests that the forests of Herrera and surrounding areas in central Panama may harbor ancient Criollo cacao trees dating back over a century. This hypothesis is well-grounded, as Panama falls within the established domestication range of Criollo cacao (Motamayor et al., 2002). However, the Criollo genetic group experienced a significant population decline during the Little Ice Age (1350–1850 AD), particularly during its peak between 1630 and 1728, resulting in a genetic bottleneck (Lozano-García et al., 2007; Motilal and Sreenivasan, 2012). The climatic fluctuations of this period, characterized by temperature drops of up to 2°C and increased cold winds, severely impacted Criollo cultivation across tropical America, contributing to economic downturns in the Spanish colonial cacao trade. These oscillating conditions likely created alternating periods of expansion and contraction for natural Criollo populations, with the species retreating to microclimate refugia during harsh periods and expanding when conditions improved (Motamayor et al., 2002). Areas like Las Minas in Herrera, adjacent to the pristine forests of Veraguas and home to the culturally isolated Ngäbe-Buglé people, may have served as critical sanctuaries for pure Criollo genotypes. The Ngäbe-Buglé, who maintain traditional cacao consumption practices and limited interaction with outside populations, likely played an unintentional yet crucial role in preserving genetic purity by minimizing cross-pollination.

Farm maintenance methods

Regardless of the cacao's origin, whether selected from native forests or introduced based on recommendations, the environment and the farmer continuously shape the selection of the best-performing cacao. Farmers' decision on how to regenerate trees influence plant development, affecting productivity, quality, resilience, and sustainability in both agricultural and ecological systems. Although farmers traditionally regenerate trees by seeds (57% of trees), grafting (43% of trees) gained popularity in Bocas del Toro (55%) while in Herrera remained low (16%). This difference may be attributed to the historic commercial cultivation of Bocas del Toro in combination with the presence of cacao technicians from IDIAP and COCABO, who support farmers with grafting training.

Grafting has the advantages of keeping valuable germplasm in terms of quality, yield, disease resistance, predictability in production, in addition to shortening the time until the first harvest by half compared to seed-raised trees. Conversely, grafting requires skills, and a further drawback is the

grafting incompatibility between scion and rootstock (Darikova et al., 2011), as well as the risk of perpetuating existing diseases.

Farms with rootstocks and grafted scions were in close contact with agricultural technicians (Farms 1, 2, 3, and 4 from Bocas del Toro) or were commercial plantations (Farms 9 and 10). In all cases, the scions were either imported genotypes or best-performing trees duplicated as scions while the rootstocks were the original material found in the farm. Farms 6, 7, and 8 from Bocas del Toro only had trees developed from seeds. From personal communication with the regional technicians in Las Minas, all the single-origin trees in Herrera are likely from seeds, although there can be few exceptions where these are from rooted scions or renewed trees by grafting a branch on the same tree. Farm 12 from Herrera was owned by a technician from MIDA who also had experience with the grafting technique, hence the grafted genotypes in his and surrounding farms. From the oral surveys, we know that clone GDG-3 from Farms BdT5 and Her12 was developed by CATIE and introduced by MIDA in 1980, corresponding to rooted cuttings of a pure Amelonado genotype. The remaining farms in Herrera (13, 14, 15, and 16) predominantly contained seed-propagated trees. Notably, these plantations included 7 trees belonging to GDG-10, varying from completely identical to less than 6% different between SNP alleles. These trees fall within the Criollo genetic cluster, which our study confirms is characterized by low heterozygosity and limited genetic variability (Bekele and Phillips-Mora, 2019). From this, we conclude that these trees are siblings sharing close kinship rather than representing clonal material.

Environmental associations

After noting significant differences in the genetic group contribution between Bocas del Toro and Herrera, and particularly, the high presence of the Criollo group in Herrera, we explored soil properties and environmental factors like elevation that can be associated with regional differences in germplasm composition. Elevation is associated with lower temperatures and lower yearly precipitation for Herrera (Supplementary material A). These two latter are key factors preventing *Moniliophthora* sporulation and germination (Leandro-Muñoz et al., 2017). The reported absence of *Moniliophthora spp* in Las Minas, Herrera, likely facilitated the cultivation of Criollo in the area. Furthermore, geological differences such as higher soil magnesium (Mg) and calcium (Ca) levels in Bocas del Toro compared with Herrera may reinforce patterns of genetic selection. Environmental variables such as organic matter, potassium, and exchangeable cations (Mn, Cu), along with farm-level traits as average tree age, farm type, cacao tree density, shade cover, and the number of neighbouring species, did not exhibit a significant influence.

The largest association was with total nitrogen (TN), where it showed between a 2- to 10-fold increased effect compared with all other environmental characteristics. TN, pH, and OM% were positively associated with Amelonado and negatively with Criollo, probably highlighting the

nutritional requirements and thus agronomic practices associated with Amelonado in contrast with Criollo. Criollo displays a characteristic low vigour (Bekele and Phillips-Mora, 2019) and so, slower growth and development. TN was also associated with reduced Amelonado/Criollo presence and increased presence of genetic background from all other groups, whereas pH exhibited the opposite trend. These patterns were consistent across scions, rootstocks, and combined samples. The positive association between pH and the dominant groups (Criollo/Amelonado) across tissues may reflect the link between higher pH and greater nutrient availability, leading to more fertile soils and enhanced plant nutrient absorption (Fernandez and Hoeft, 1992; McCauley et al., 2009).

While clear associations were observed between soil properties (e.g., TN, pH) and specific genetic groups, it is important to acknowledge that the distribution of genetic diversity was uneven across regions. Bocas del Toro exhibited greater heterogeneity, with all genetic groups represented, whereas Herrera was dominated by Amelonado and Criollo. These regional differences likely reflect both historical planting practices and environmental gradients (e.g., elevation), which could partially confound the associations detected. Nonetheless, the consistency of soil-genotype trends across tissues suggests that soil conditions may indeed influence, or be shaped by, the success and persistence of genetic backgrounds.

Conclusion

Germplasm of Panama's most productive regions is predominantly composed of Amelonado and Criollo groups, alongside Nacional, IMC Iquitos, and Nanay. Surprisingly, we encountered 100% pure Criollo germplasm, vestiges of the original semi-domesticated cacao cultivated by Mesoamerican cultures. The potential discovery of pure Criollo trees in Herrera not only underscores the resilience of this genetic group but also opens new avenues for research into climate-resilient cacao cultivation and the conservation of heirloom varieties. Elevation appeared as a potential key factor for Criollo presence, while Amelonado association with total nitrogen provides valuable information for optimizing on-farm germplasm selection and management. As grafting techniques gain popularity, they could drive agricultural technology development in these areas. The germplasm characterized in these areas holds great value for regional breeding programs, climate adaptation strategies, and preservation of genetic heritage. Furthermore, the presence of these genetic groups has renewed socioeconomic interest in cacao cultivation opening new research into the organoleptic and manufacturing qualities of the resulting chocolate. These trees are more than crops; they are living bridges between history, resilience, and the future of sustainable cacao.

Supplementary Material

A. Plantations surveyed across the Panamanian Isthmus, with climatic data for both regions

Code	Region	Farm	Town	District	Altitude (masl)	Latitude	Longitude	Secondary Spp	Type	Topography (1 to 3)	Drainage (1 to 3)	Tree density (#/ht)	Tree density (#/ht)	Size (ht)	Average age of trees
BdT_1	Bocas del Toro	1	Milla 7	La Gloria	51	9°21'4"	82°26'29'	7	Cacao Forest	3	3	300	Low	1	50+
BdT_2	Bocas del Toro	2	Milla 7	La Gloria	12	9°19'20"	82°26'48"	6	Cacao Forest	3	3	1000	High	3	15
BdT_3	Bocas del Toro	3	Rio Oeste Arriba	Almirante	39	9°15'23"	82°25'15"	6	Cacao Plantation	3	2	600	Low	2.9	31
BdT_4	Bocas del Toro	4	Rio Oeste Arriba	Almirante	35	9°15'39	82°25'22"	7	Cacao Plantation	1	3	600	Low	1.5	40
BdT_5	Bocas del Toro	5	Soledad	Valle Risco	118	9°15'19"	82°27'23"	4	Forest	2	2	1000	High	4.5	50+
BdT_6	Bocas del Toro	6	Nuduobidi	Valle Risco	81	9°15'21"	82°28'6"	4	Cacao Plantation	1	3	750	Medium	1.5	40
BdT_7	Bocas del Toro	7	Alto conejo	Hirondai	308	8°58'53"	82°20'26"	6	Cacao Plantation	2		2000	Very high	6	5
BdT_8	Bocas del Toro	8	Filo Verde, Tuwai	Hirondai	62	9°0'39"	82°19'51"	3	Permaculture	1	2	933-800	Medium	1	22
BdT_9	Bocas del Toro	9	Quebrada Limon	Almirante	6	9°16'54"	82°25'14"	2	Commercial Plant	2	2	1111	High	3.32	11
BdT_10	Bocas del Toro	10	Rio Uyama	Los Higuerones, Cauchero	4	9°9'91"	82°19'0"	1	Commercial Plant	1	2	1111	High	21.45	10
BdT_11	Bocas del Toro	11	Nuevo Paraiso	Almirante	15	9°16'50"	82°24'49'	5	Exp Garden	1	1	600	Low	0.5	10
Her_12	Herrera	12	Las Minas	Cabecera, Las Minas	338	7°47"51"	80°45'3"	6	Garden	2	3	150	Low	1	7
Her_13	Herrera	13	El Ciprian	Las Minas	404	7°41'3"	80°45'11"	9	Exp Garden	1	1	20	Low	0.25	10
Her_14	Herrera	14	El Nanzal	Las Minas	449	7°44'13"	80°44'21"	6	Forest	3	2	20	Low	1	6
Her_15	Herrera	15	La desbarranca da	Chepo, Las Minas	546	7°39'53"	80°47'16"	6	Cacao Plantation	2	3	200	Low	20	13
Her_16	Herrera	16	Tres puntas	Las	600	7°47'53"	80°48'2"	8	Garden	2	2	200	Low	0.5	6

Average values of precipitation, temperature and relative humidity extracted from the Meteorological Institute of Panama (IMH) for reference.

	Region	Altitude (masl)	Average Precipitation (mm/month)	Max Daily Temperature (°C)	Relative Humidity (%)
Average	**Bocas del Toro**	**66.5**	**283**	**25.6**	**85.8**
	Herrera	**467**	**188**	**23.8**	**88.9**

B. Name, source, and group of cacao (*Theobroma cacao* L.) reference accessions used in analysis.

Accession	Source/ Provider	Group	Genetic Group
La Gloria 1	INIAP[a], Ecuador	Nacional	Nacional
La Gloria 2	INIAP, Ecuador	Nacional	Nacional
La Gloria 4	INIAP, Ecuador	Nacional	Nacional
La Gloria 7	INIAP, Ecuador	Nacional	Nacional
La Gloria 9	INIAP, Ecuador	Nacional	Nacional
La Gloria 12	INIAP, Ecuador	Nacional	Nacional
La Gloria 13	INIAP, Ecuador	Nacional	Nacional
La Gloria 20	INIAP, Ecuador	Nacional	Nacional
La Gloria 22	INIAP, Ecuador	Nacional	Nacional
La Gloria 24	INIAP, Ecuador	Nacional	Nacional
La Gloria 25	INIAP, Ecuador	Nacional	Nacional
Las Brisas 3	INIAP, Ecuador	Nacional	Nacional
Las Brisas 4	INIAP, Ecuador	Nacional	Nacional
Las Brisas 7	INIAP, Ecuador	Nacional	Nacional
Las Brisas 14	INIAP, Ecuador	Nacional	Nacional
Las Brisas 16	INIAP, Ecuador	Nacional	Nacional
Las Brisas 22	INIAP, Ecuador	Nacional	Nacional
Las Brisas 25	INIAP, Ecuador	Nacional	Nacional
Las Brisas 26	INIAP, Ecuador	Nacional	Nacional
Las Brisas 30	INIAP, Ecuador	Nacional	Nacional
IMC 10	ICGT[b], Trinidad	UAF	IMC
IMC 103	ICGT, Trinidad	UAF	IMC
IMC 105	ICGT, Trinidad	UAF	IMC
IMC 12	ICGT, Trinidad	UAF	IMC
IMC 14	ICGT, Trinidad	UAF	IMC
IMC 20	ICGT, Trinidad	UAF	IMC
IMC 3	ICGT, Trinidad	UAF	IMC
IMC 31	ICGT, Trinidad	UAF	IMC
IMC 33	ICGT, Trinidad	UAF	IMC
IMC 38	ICGT, Trinidad	UAF	IMC
IMC 47	ICGT, Trinidad	UAF	IMC
IMC 60	ICGT, Trinidad	UAF	IMC
IMC 71	ICGT, Trinidad	UAF	IMC
IMC 73	ICGT, Trinidad	UAF	IMC
IMC 77	ICGT, Trinidad	UAF	IMC
IMC 94	ICGT, Trinidad	UAF	IMC
IMC 96	ICGT, Trinidad	UAF	IMC

POUND 18	ICGT, Trinidad	UAF	IMC
POUND 27/C	ICGT, Trinidad	UAF	IMC
POUND 21/B	ICGT, Trinidad	UAF	IMC
NA 13	ICGT, Trinidad	UAF	NA
NA 286	ICGT, Trinidad	UAF	NA
NA 399	ICGT, Trinidad	UAF	NA
NA 43	ICGT, Trinidad	UAF	NA
NA 670	ICGT, Trinidad	UAF	NA
NA 7/10	ICGT, Trinidad	UAF	NA
NA 702	ICGT, Trinidad	UAF	NA
NA 756	ICGT, Trinidad	UAF	NA
NA 79	ICGT, Trinidad	UAF	NA
NA 807	ICGT, Trinidad	UAF	NA
NA 916	ICGT, Trinidad	UAF	NA
NA 92	ICGT, Trinidad	UAF	NA
POUND 1/B	ICGT, Trinidad	UAF	NA
POUND 2/A	ICGT, Trinidad	UAF	NA
POUND 2/B	ICGT, Trinidad	UAF	NA
POUND 7/C	ICGT, Trinidad	UAF	NA
POUND 26/C	ICGT, Trinidad	UAF	NA
POUND 32/A	ICGT, Trinidad	UAF	NA
POUND 25/A	ICGT, Trinidad	UAF	NA
POUND25 /C	ICGT, Trinidad	UAF	NA
PA 12	ICGT, Trinidad	UAF	PA
PA 121	ICGT, Trinidad	UAF	PA
PA 124	ICGT, Trinidad	UAF	PA
PA 126	ICGT, Trinidad	UAF	PA
PA 13	ICGT, Trinidad	UAF	PA
PA 136	ICGT, Trinidad	UAF	PA
PA 137	ICGT, Trinidad	UAF	PA
PA 149	ICGT, Trinidad	UAF	PA
PA 150	ICGT, Trinidad	UAF	PA
PA 156	ICGT, Trinidad	UAF	PA
PA 169	ICGT, Trinidad	UAF	PA
PA 191	ICGT, Trinidad	UAF	PA
PA 195	ICGT, Trinidad	UAF	PA
PA 2	ICGT, Trinidad	UAF	PA
PA 289	ICGT, Trinidad	UAF	PA
PA 299	ICGT, Trinidad	UAF	PA
PA 30	ICGT, Trinidad	UAF	PA
PA 300	ICGT, Trinidad	UAF	PA
PA 39	ICGT, Trinidad	UAF	PA
PA 70	ICGT, Trinidad	UAF	PA
Ucayali 11	ICT[c], Peru	UAF	SCA
Ucayali 15	ICT, Peru	UAF	SCA
Ucayali 19	ICT, Peru	UAF	SCA
Ucayali 24	ICT, Peru	UAF	SCA
Ucayali 28	ICT, Peru	UAF	SCA
Ucayali 35	ICT, Peru	UAF	SCA
Ucayali 36	ICT, Peru	UAF	SCA
Ucayali 37	ICT, Peru	UAF	SCA
Ucayali 38	ICT, Peru	UAF	SCA

Ucayali 39	ICT, Peru	UAF	SCA
Ucayali 4	ICT, Peru	UAF	SCA
Ucayali 41	ICT, Peru	UAF	SCA
Ucayali 43	ICT, Peru	UAF	SCA
Ucayali 5	ICT, Peru	UAF	SCA
Ucayali 6	ICT, Peru	UAF	SCA
Ucayali 66	ICT, Peru	UAF	SCA
Ucayali 69	ICT, Peru	UAF	SCA
Ucayali 7	ICT, Peru	UAF	SCA
SCA 6	ICGT, Trinidad	UAF	SCA
SCA 24	ICGT, Trinidad	UAF	SCA
BC32 Belize	SPCL, USDA ARS[d]	Criollo	Criollo
PR 127	TARS[e], USDA ARS	Criollo	Criollo
PR 129	TARS, USDA ARS	Criollo	Criollo
PR 139	TARS, USDA ARS	Criollo	Criollo
PR 145	TARS, USDA ARS	Criollo	Criollo
PR 154	TARS, USDA ARS	Criollo	Criollo
PR 155	TARS, USDA ARS	Criollo	Criollo
PR 156	TARS, USDA ARS	Criollo	Criollo
PR 46	TARS, USDA ARS	Criollo	Criollo
PR 71	TARS, USDA ARS	Criollo	Criollo
Francisco	TechnoServe, Nicaragua	Criollo	Criollo
Matagalpa 13	TechnoServe, Nicaragua	Criollo	Criollo
Matagalpa 16	TechnoServe, Nicaragua	Criollo	Criollo
Matagalpa 20	TechnoServe, Nicaragua	Criollo	Criollo
Oscar	TechnoServe, Nicaragua	Criollo	Criollo
Tiburcio 1	TechnoServe, Nicaragua	Criollo	Criollo
Tiburcio 2	TechnoServe, Nicaragua	Criollo	Criollo
Tiburcio 3	TechnoServe, Nicaragua	Criollo	Criollo
Tiburcio 4	TechnoServe, Nicaragua	Criollo	Criollo
Tiburcio 5	TechnoServe, Nicaragua	Criollo	Criollo
ES 01 Belize	SPCL, USDA ARS	Amelonado	Amelonado
ES 02 Belize	SPCL, USDA ARS	Amelonado	Amelonado
ES 03 Belize	SPCL, USDA ARS	Amelonado	Amelonado
SC 02 Belize	SPCL, USDA ARS	Amelonado	Amelonado
UFC 01 Belize	SPCL, USDA ARS	Amelonado	Amelonado
UFC 02 Belize	SPCL, USDA ARS	Amelonado	Amelonado
FC 04 Belize	SPCL, USDA ARS	Amelonado	Amelonado
VC 03 Belize	SPCL, USDA ARS	Amelonado	Amelonado
PR 151	TARS, USDA ARS	Amelonado	Amelonado
PR 76	TARS, USDA ARS	Amelonado	Amelonado
PR 8	TARS, USDA ARS	Amelonado	Amelonado
REDAMEL 1/3	ICGT, Trinidad	Amelonado	Amelonado
Madagas 24	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 16	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 8	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 2	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 13	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 21	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 42	SPCL, USDA ARS	Amelonado	Amelonado
Madagas 5	SPCL, USDA ARS	Amelonado	Amelonado
PR 157	TARS, USDA ARS	Trinitario	Trinitario

PR 28	TARS, USDA ARS	Trinitario	Trinitario
PR 9	TARS, USDA ARS	Trinitario	Trinitario
RIM 189	CATIE[f], Costa Rica	Trinitario	Trinitario
ICS 6	ICGT, Trinidad	Trinitario	Trinitario
ICS 40	ICGT, Trinidad	Trinitario	Trinitario
ICS 83	ICGT, Trinidad	Trinitario	Trinitario
ICS 89	ICGT, Trinidad	Trinitario	Trinitario
ICS 93	ICGT, Trinidad	Trinitario	Trinitario
ICS 1	ICGT, Trinidad	Trinitario	Trinitario

[a]INIAP = Instituto Nacional Autónomo de Investigaciones Agropecuarias

[b] ICGT = International Cocoa Genebank in Trinidad

[c]ICT = Instituto de Cultivos Tropicales

[d]SPCL, USDA ARS = Sustainable Perennial Crops Laboratory, United States Department of Agriculture, Agricultural Research Service

[e]TARS = Tropical Agricultural Research Station

[f]CATIE = Centro Agronómico Tropical de Investigación y Enseñanza

C. SNP codes and flanking sequences for the 96 polymorphic markers used in genotyping Jamaican and reference cacao accessions.

SNP ID	LG[b]	Flanking sequences and SNP position
TcSNP11	9	GTCAAATGTTGCTGATCTTATTTGATGCTGTTTTGGGTTGGCTAGATAAATGGCTACTTG '**[T/G]**AGGTGACAAAGTTTTAAGGTCTGGTGCCCTTACAGTTGGAAGGGGTTTTTGTTCCTGCAC
TcSNP17	2	CAACTCTAATAACAGTGTTTCAAGAGAAAACCCTGAAGACCCAGATCAAGAATCACCCCA **[A/G]**GAAACAACTTCCCAACGTGAAAGAATAATCCTTATAAATCCATTCACTCAAGGCATGGTG
TcSNP19	3	GAAGAGAAAAAGATGAAGAAGAAGATTGCTCTCGAAACTCCAGAAACCGAGAAGAAGATG **[A/C]**AGAAGCTGAAGGAGCCGAGGCTAAAAGAGGAGGACAATGATTCGGAGTTGAAGAAGAGTA
TcSNP23	8	AGGTTGTGGGTGCAAGGAAAGATGAATTGCAACAGAAAATAACCTCGCATTCTAGCATGC **[T/C]**TGCTTGATGCAGAGAAGCATGAATCAAGGTTTAGTTGTACTCCTTGAAGTCTGATTTCTG
TcSNP25	9	TTTATGTTACTAGTTTTTGTTGGGTGGTTGTCTGTTTCTTGTTTGCATTATTTTGTGGAC **[C/G]**GGACCTTAGTATATCCTGCTTGTTTGCTCTGTTTTTTGTTCAGTAGTCTGGGTTGTTAGT
TcSNP28	5	CGGTGTACAAGGCATTATGGAGGGAGGCCAAAGAGTGTATTTATGTTGAACCAGATGAGG **[T/G]**TGAGCATGGCAAAGGTGTGTTCTGGTATAGGAATAAGCTTTAAATTTGCATGTAGGAAAG
TcSNP32	4	TTCCACTTTTGTCAATATAGATCAGTCATGGATGATATTGTGCAAAAATAATACCATAGT **[A/T]**CTAAGAAATCATGCACATTTTTGTAAATAGGAAGAATCTGCCATGGTAATGTTGTGTTGA
TcSNP42	1	TTCTTGGAGTCAAACCAACACCCAAGAAGGAACTTCTGGGTATTCCTGTTATAACACATG **[A/G]**CAAATCCTTAAAGGTGCCAACTAAGTTCGATGCTAGAACAGCTTGGCCACAATGTAGCAC
TcSNP60	4	AACATCAACCAAAAAGCTTTTTATTGGAGGTCTTTCGTACGGAACTGATGATCAGACTCT **[T/C]**AAGGAAGCATTTTCCGGATTTGGGGATGTCACTGAAGCAAGGATTATCATTGATAGAGAC
TcSNP122	2	AGATCAAAGAAGCTCCCAAACGTGAGGATGAGAAACAACAGTGAAGTCACTAATCCTCGT **[T/C]**GTAGTATTTTACTTCTCTGGTTGTTTTTGTATGGAGAAGTTGCTTCGTCCCMCAATCGCT
TcSNP139	8	AAATCTAATACTAGTAACATTCTAGACCGATATTTATGTCTAGGGAGAAAGACATYTTGA **[T/G]**TTTGGCTTGGTGGTTCATGTTGATTTTTCTGTTGATACCAAATCAATCAGACCATCGATG
TcSNP144	10	ACCGACACAGGATTCACCGTGAATGGGGTCAACTATGAAGGCAGCTTGCTATGTGTGGGA **[A/C]**ATTTGCTCACTTCTTGGGTTCCCAATAAGTTTTCTCAAATTACTCCCGACAGCTTATCTA
TcSNP150	5	ATACCCTTCGACGACTACTGCTGACTTGGGAACACCCACCAGCGTTTTGCCGGCAACATC **[T/G]**GGTGACTGGTCTGATTTTTCCTCTAATTTTTATCTAGAACTTTGTCAAGCGTTTAAAATA
TcSNP151	8	TGAAGCTGGAGGGATTGTGGAGAGTGTTGGTGAGGGTGTGACTGATCTCCAACCAGGTGA **[T/C]**CATGTTCTCCCCATCTTTACTGGAGAATGCAAGGAGTGTCGCCACTGCTTGTCAGAAGAA
TcSNP160	5	TGTGGTGAAGCGTGAAGATCTTTGGATCACCTCCAAACTCTGGAACACTAATCATGCACC **[A/T]**GAAGATGTACAAGAGGCATTGAATTCTACATTGGAAGACTTACATCTGAACTATCTAGAT

TcSNP172	9	TCTATCATAAGGTCTTTTTTCAGGTGACTCTTCTTTGGATGGACATTTGTCTTTTGTTAG
		[A/G]TGCTTTGATTTTGATCTTTATTAGTAGGTTTTACAATCTACTTGTGGATCAAACTCCAAA
TcSNP193	9	GCATAGAAATTCTACGACAGCTCCCTCAATTTTCCTTGTTATCAAAGTTTCCTGTTTAAT
		[A/C]TCGTTTTCCATTCCTGAAAATGTCTCTGATTCCAAGCTTCTTTGGCAGCCGCCGCAGCAA
TcSNP205	1	TACATTCTGCACACCCTGAGGCGATAGGCTCACAAGCTATTGACCTTCAATTCTCCATTT
		[A/G]TTCTTGGAGGGAAATGCGTCATATTAAATAAGACATCTTACCAACAGCAGTTGAATAAGA
TcSNP242	9	AGCAATCTTCCTCCGCCAAACAAGCCATCAATCAAACCCTCGTTCCTCTGTCTTCGCCTT
		[T/C]CTCTTCGTTGCCGTCATCGTCATCGATCTCATCATCTCCATTTTCGGGTTCCTTGCCAAT
TcSNP269	8	AACAAAAGAAATCAAACTGTTCCATTTCAAAAGGAATAGAGAATAAGATAAACTACGATT
		[A/G]GGAAAGAACAAAACCTAAATATATGATAGCCCAAAATTCCTGACCACTTTGCATATTTAT
TcSNP277	4	CAGCCGTGATGAATACACAAAGGAGAGATTGGAGGCCATAAATGATGAGTTCAAGCTGTA
		[T/C]CGTTGCCATACGATATTGAACTGTGCTCGTGCTTGCCCTAAGGGCCTGAACCCAGGAAAG
TcSNP309	6	ATTCACTCGAGTAAGAAAAAAATTTACATTTTAAGTACAAGAAATTTCCCAAGATGCCTC
		[T/C]CTTCCACAAACAATATTCAGTCTAAAGCACACCGTCTTCCGTACCATCCCAGTGTTCGTT
TcSNP316	2	CAAGAGTCGGTTCAAGAAATCGTACGGGTCGCAGGAGGAGCACGATTACAGGTTCAAGGT
		[A/T]TTTCAGGATAACTTGAGACGTGCGGCGCGTCATCAGAAGCTTGACCCGTCTGCGAGTCAC
TcSNP372	4	GGAAATTCTGAAGATTGTGAAGGAGAACTTCGATTTCAGACCTGGAATGATTTCTATCAA
		[A/T]CTTGATCTTAAGAGGGGTGGCAATGGCAGGTTTTTGAAGACTGCTGCCTATGGGCATTTT
TcSNP389	10	CTGTTGCAACAAGATGGATGCCACAACCCCCAAATACTCAAAGGCAAGGTATGATGAAAT
		[T/C]GTGAAGGAAGTTTCTTCATACTTGAAGAAGGTTGGATACAACCCTGACAAGATTCCCTTT
TcSNP395	4	CTCTCTTTCTCTCTCACTTTTGCTGTCATTAACATAATCATTTCTGCATTTGTGAAAGCT
		[T/C]ATAATTTAATCTCTACCAATGGCTGACCGTGACCGCCCTCACCAGATTCAGGTTCACCAA
TcSNP413	3	CGCAGTTTTCTCAGTTTCGGGCAATGATTCTAGTTATATCATGAAACTAAGTTCTTAGGT
		[T/C]GTTATTATGAGCCTATGCATACAGCATACGTAATGCGGAACTGTAGTTGATACTTTATGT
TcSNP418	1	CAATTCCCTTTCCGGAAAAATCCCACAACAGATAGGGAATCTCCAGAAGCTCACTGTCTT
		[A/G]AACCTCGCAGACAATAAGCTAAGTGGAGAGATCCCAAGCTCCTTAGTTCAACTCTCCAGT
TcSNP427	9	AGCATGGCACTCAGCTGGAACTTTCGATGTTAAGACCAGGACCGGAGGTCCATTCGGAAC
		[T/G]ATGAAGCAGCCTGCTGAGCTGGCACATGGTGCTAACAACGGTCTCGATATTGCTGTTAGG
TcSNP437	2	CGTAAACAAGCAGCTTTCGCTCACGGGCTTCAAGACAAGGCTCACGGTATCATCCCCATC
		[A/T]CAGTTATCTGGCTCGTGCCACAGTACGCCCTCCATGGAATAGCTGAAGCTTTCATGTCGA
TcSNP497	6	AATCTTTGTCAAGACCCTGACCGGCAAAACCATAACCTTGGAAGTTGAAAGCTCCGACAC
		[T/C]ATCGACAACGTGAAGGCCAAGATCCAGGACAAAGAAGGAATTCCCCCGGACCAACAAAGG
TcSNP510	1	GCTCGGTTTACCAGTGGCGGCAACGGTGAACTGTGCGGACAACACGGGAGCCAAGAACCT
		[C/G]TACATCATTTCAGTGAAAGGAATCAAAGGGCGTCTCAACAGGCTACCATCGGCTTGCGTC

TcSNP521	2	GATTCTGTGAACTATAGCACCAAGGAAAAGAAGGCCAACCTTAAAGAAAATCCTTTAATA
		[A/C]ATACTACTGACTGGTATATGTATCAGAAACAAAAAAAGAAATAGTAGAAATTCCAATGGC
TcSNP524	5	CTTGGAACCGTTCAAAGGATCATCTTGGGGAGTGTGAGCAACTACGTTATGACACATGCT
		[T/C]CTTGCCCTGTTACCATCGTTAAGGAATCAAGCTCCATGAAGCACTGATAAGTTCATATAT
TcSNP529	1	ATAGGCCTTGAATGACTTCATCTCCTCTGTCACCTTCACAAGCTCAACAAAAGGCTTCTC
		[A/C]AGAAAAATAGGCATGTATGGTCCTTGCACTTGGGTTGCAGTTGCAAGCTCCTCGGCAGTA
TcSNP534[a]	3	TGTAAAGAGTGCCATTGTTCAAGTGGATGCAGCGCCTTTTAAGCAATGGTACCTTCAGCA
		[T/C]TATGGTGTCGACATTGGGAGGAAGAAGAAAACAGCAGCTAAGAAGGAAGCTACTGAGGAA
TcSNP546	10	ACAACATTGCGCTCAAGGGTCTTGCCAAGTTCTTCAAGGAATCGAGCTTAGAAGAGAGAG
		[A/G]GCATGCTGAGAAACTGATGAAATACCAGAACAAACGAGGTGGAAAAGTGAAGCTGCAGTC
TcSNP547	8	AGAGCCTTTCGCTTCCTCATCTTTCTCTCTGCGACTGCGTTTCAGTGTTGGACTGAGTTG
		[A/G]AAGGAGGATGGCGGTTCGTAGGGACAAAGGGCAGTCAATCCGCGGATCTCGAATTGCTAT
TcSNP560	10	GAGGCAGAGAAATACAAGGCCGAGGATGAGGAGCACAAGAAGAAGGCGGAGGCCAAGAAT
		[T/G]CTTTGGAGAACTACGCCTATAACATGAGGAACACTGTCAAGGATGAGAAGATTGGCTCAA
TcSNP561	6	GCCAATTGTCCCTGTCAGGACAAAGAAGGCATATGAAACAATTACCATGTACCCAAAGTA
		[A/G]AGGATGCCTGAAACCAGCTTTGTGATCTCCAACTTGGTAAAGAAGTAGAAAACTGAGTAC
TcSNP563	9	CATGCTCTTCAGGAGGCCTCTCAACTATCAGCAGCAACAGGAGAATCAGGCCCAGCAAGC
		[T/C]ATGCTTGCTAAGTAATTAATATTAGTGTTCATTAATCAGTATCCAGAAAATGGATGAAAA
TcSNP577	5	CAATTGCAAACATCCGCCATGATGATCATGATGATGATAATGATGAACAGAAAACAAGAT
		[C/G]AACGGCCACAAATTTAACACAAATTACAAGGGAAAAAAAGGGGAGAAAATARAAAACTCA
TcSNP588	4	ATCTCGATCTGATCTGTTTCTGGGAGGAGGATATGAAGAAGTTGACGGCGACGAGGATTT
		[A/C]AAGGCGGAGTTCTTATGTCCGTTTTGCGCTGAGGATTTCGACATTGTTGGGCTTTGTTGC
TcSNP591	1	GGGCGTCGAGAAGACTTCGGCTGCTATTAAGGTCTTGAAACAGATCGGGGCTTATGCGGA
		[A/C]GTCGAGAAGGCGAAAGATAGCCAGGCCATTCGTCCCGGAAAGGGCAAAATGAGGAATCGC
TcSNP595	3	TAGGCATGTATACTATATTAATTGAATTTAGAAACTACGTGTGCCATGATTAAATGTAAT
		[A/G]TATGCAAGCCTAGATGTGGCTCTGAGGTAGGGCATGGCTTGTTGTTATACGGATATATGT
TcSNP606	7	ATGTGTACAAGTTCACATGTTGAGTGAAAACTAGAATTTCTGTGTGTTAAGCTTGTGCCT
		[A/G]CATAAAAAGGATAATTGTACAAAATGTTAAAAAGAATAATTTGTGCTTTGCAACGTTTAA
TcSNP619	6	GAACAAAAATTGTAAATTAATCTTATGCATGGAGTAATGACCCCACAGCTTTGCAACACC
		[T/C]CAAAAATGGTGGTTGTCCCTTTGTACATAATGATTCTGGGATGAATCATTTGTCTTTGTA
TcSNP632	9	CAACCGTGAGAACTACATCAATGAGCTAATTAGGAGAATGAATTAAACTTGGCCTTGTTT
		[T/G] TTAATTTAGTAGTTGAGGTATTTTGGCTTATTTGAAAGGTTCTGATCKTAATTTCTTCCG
TcSNP645	5	TGAAGGCACTGGAGCCCTTTGTGAATTGGCTGGAGGAAGCAGAAGAAGAGGAATAAGTGC
		[A/G]ATAACAATAATATTATGATCACTTTGAAAAGAAAAGGGGTGAACCCTGTACTTGTTCCTT

TcSNP653	10	AAATGGTTTATACTAGTACAACAGTAACAAGAAAAGAGAAAATATAACATTAAAAATGAA
		[A/G]TTTCAAAAACAATTAAGGAAAAAAAGGGATATCCGACCTGAAGCACACACCCTCAGAAAC
TcSNP674	10	GGAGGTTGAAGAGGAAGAGAAGAAAGATGAGACAAAGATCTAAGTAGTCTTCTTTAGAGC
		[T/C]AGTAATTGAGTCTTCTGTCTTTTAGATTTCAGTTGTTTTATGGGAACCTTTTTGTTAAAG
TcSNP689	3	CATATTTGTGTTTGGCAGTGGGGTTAGCCTTAGTCGTTTGGTCGGTCTCGAATGGTGATG
		[A/G]GGTTGCCGCTGCTGACCGAATGGTGGGGAGAAGTTCCAATTCGTTTCAGCAGATGAGAAG
TcSNP709	1	GGCATTTGCAGATGAAGTTGGCATACCTTTTATGGAAACCAGTGCAAAAGATGCCACAAA
		[T/C]GTGGAACAGGCTTTCATGGCCATGGCTGCTTCCATCAAGGATAGAATGGCAAGCCAACCA
TcSNP723	10	AAATTGGAACCCCACCTCAAACCATGCTCATGGCCATGGATACTAGCAATGATGCTGCTT
		[T/G]GTTACCTTGCTCAGGCTGCCTTGGCTGCTCTTCTAGTGTTTTCGACAATGCTAAATCAAC
TcSNP731	10	ATTTACAAGGCTAAGGTGAAGAAGAATGGCACACTAGTTGCATTTGGGGAAAGGTT
		[A/G]CTAGACCATGGTAACAGTGGCGTTATGTGCTAAGTAAGTAACCTCCCCTA
TcSNP750	6	ACCTCTAGCCTTAAGCTCTGCAAGCCGCCTTGACAAATCATCCTCATCAACCTTCTCCTG
		[T/C]GTCTTGGTTGACACAGCATGAGCAGCTGGCTGTGGCAGCCCAACAGAGACCTCCAACCCA
TcSNP751	5	AGCATTTCACCATGGCCATTTTCTTGCTTCCTTGCTTGCTTAAGCAGCTTGAGGGGTTCC
		[T/G]TTCGTCTGTGGATGCAACTGGAGTTCCAAGTCATCTTTCCCGTCTTTCTCCCTAGCCTTT
TcSNP791	7	AGTTGAGTTCAATCTTTGTTTCAGAATGGCAGACGAGTTGGTTTTGTTGGATTTCTGGCC
		[A/T]AGTCCATTTGGGATGAGAGTGAGGATTGCTTTGGCTGAGAAAGGGATAAAACATGAGTAC
TcSNP836	2	AAGGGTAATCCCATGTGGAATCAGCTATCAGTGCCATCTTCCACCATGTACTCATGGGAC
		[T/C]CCAACTCAACCTACATTCATGAGCCTCCATACTTCAAGAGTATGACCATGGAGCCTCCTG
TcSNP852	3	TGTTCCTTTACATCACCATCTTGACTGTCATGGGTGTGTCTCAGTCTAAGACCAAGTGCA
		[C/G]TACTGTGGGTATTCAAGGCATTGCTTGGGCCTTTGGAGGCATGATCTTTGCTCTTGTCTA
TcSNP872	4	AGTGATGTTGCACAGGTCACAAAGCTTCTCAACCTTGTTGCCAGTCAATCCAAGAGGTCG
		[C/G]AGATCCCAAAGAACAAGGTGGTTCTCAGTTCCCCCAGTTACAAGCTTGTATCCCTTGCTC
TcSNP886	4	GCGAGTGCCGAGTAGGCTCAGGAGGATCAAGTAGGGAGAAATAAGCGGGCGGTTTGAGTC
		[T/C]ACTAGAGACTCTCTCTAGGGGAGCTTTACTAGTATTATGTTTGTCACTTATGAGGTAAAG
TcSNP891	2	CAGAAAGCACACTTTGCACAAGGTTACACAGTATAAGAAGGGTAAGGATAGTTTGGCTGC
		[T/C]CAAGGGAAACGCCGTTATGATCGCAAACAATCAGGTTATGGTGGGCAGACCAAACCAGTG
TcSNP894	6	TCTGCCTCTCCACGTTCCCTTCGCTCCTTTCTCCCGCTCCGAAAAGCTTGGCCGAATCGC
		[C/G]GACTTCACCCGCTCTTTCCCTTCCTCGAACGCCAACCCATCCAACCGGCCTTCCTCTGCA
TcSNP915	1	TAAGGATCCTGAAGGTCTTCGCATATTTTACTATCTGGTTCAGGACTTGAAGTGCTTCGT
		[A/G]TTCTCTCTCATCTCACTCCACTTCAAGATAAAGCCTATTTAAAGAATGATAATCGTCTGC
TcSNP917	10	TATGGAAACTGGGTAAAGGCTCAAGAATGGGCGGCAATTCCTGGTGGACAAGCCACTTAC
		[T/C]TGGGCAACCCTGTTCCATGGGGTACCCTCCCTATCATCTTGGTCATCGAATTCTTCGCCA

TcSNP929	3	TGAAATGGAAATGAAAATACAAAAAACCAAGTCATGTTCAATAGCAAGTGATGTCCTTTT
		[C/G]TCTGCTGATTATAACAGAAGTTGTGTTAATTTTACACTTAGTTGATGTGGTTTTAAAGGG
TcSNP944	7	TGACATGCCCCCACTGGAGGATGCTGATGCAGAGGCCGAGGGCAGCAAGATGGAGGAAGT
		[T/C]GACTAAGCCTGACATRTTTTAATGTTTGAACTTTATCTTTATGTAGTAGTCTCATTTTAA
TcSNP953	4	GGGTCTCTAAATGGCACTTGAGGTTGCCAGAGATAAAGCTTCATTCATGCTCCAAGCCAA
		[A/T]TCTTTGGGCATTCCAGGGTCAAACACCTCAAAAACAAGTTTCTTTATGGCTCAAATCATG
TcSNP994	6	AACAAACCGAAATATATTCTAGGGAAACTTTGCATTTGCAACCCCTATTTGACTTTCGAT
		[T/C]TGCTCAGATCTGCTTTTGAATCCTCAACTTCTGAATAGTCAAYATGGGAGGAGGGTTCAG
TcSNP996	3	AGGCCGGCTCGAGAACCGGCTGTCGCTCGAGTCCGACATAGCCTCCGTGAAACCAACCCA
		[T/C]GTCTTCAATGCCGCTGGAGTCACCGGGCGTCCAAACGTCGACTGGTGCGAATCCCATAAG
TcSNP998	5	AACAAGAACAAGGGAATGGGGACAATATGATCCCACTAGACATGATGGCTTTTCCTCCTC
		[A/G]TAGTAGTGAGAAAGGAATAACATAGAATCTAAGCTGTCACAGATGTCTCCACTGCAACCT
TcSNP999	8	ACTGGGAGTAAACAAAGCAAAACCAAATTTATTGGAGTAAATTATGGGCAAGTGGC
		[T/C]GACAACCTACCGACCGGCAGCCACTGCAAAGCTCAGCACTAACAAAAG
TcSNP1034	4	CTTCAGCCAGCACTCAGCAACTAGGAAGTGAGCATTGCAGTTTTTGTTCTTCATTTTTCT
		[T/C]GCCTCCTGGTTTGAATTGCCCCGGTTTGCTTGTTGAGTCACTTCATCTTAAGTTCATAGA
TcSNP1060	2	CAGTTCTTGATTTTGAGATTGAAGTCCTAACCTTGCACTGAGTACTTGGCACTTGGTAAC
		[T/C]GTTGGAAGAATATGTAGCCTGAATAAGTCACGTGCCATGAACATATATGGATATTATCGG
TcSNP1062	3	ATGCAAGAGGGAACTCTGGAGAGTTCAGTATGCTCTAAGCCGCATTCGAAATGCTGCAAG
		[A/G]GACCTTCTGACCCTTGATGAGAAGAACCCACGTCGGATCTTTGAGGGTGAGGCTCTTCTT
TcSNP1069	10	TCTCAATATTCCCAAGCTAATAAGAGCTTGTGATGAGCAGCAGCACTGGAAGGAACTTAC
		[C/G]TATTTGTATATCCAATATGATGAGTTTGATAATGCTGCTACTACTGTGATGAACCACTCT
TcSNP1075	1	TTTCATGGCAGCTCATGCACTTATAATTTATGATTGCAACTCCCAACTGATACATAAATG
		[A/T]GTAGTAGGCCTGTATAGATGAATTACTCTCGAAACAAAGCATCTTCGAGAGTGTCCATGT
TcSNP1096	1	CATTATACTTAGGATCAATGTTAAGGCGGAAATGGGCTTCAGCTGACTCAACAAACTTTG
		[T/C]ACTTGTCATTTCCTTTAACAATGATATTGCTGTCTTTAGATCATACTCCTTTTTGTTTTC
TcSNP1108	4	TGAGAGAGAAGCAGAGCAAGGAACAAGATCAAAGCTTTCAGCTCTAACACGACCAACAAT
		[A/T]ACTCTACCGCCAAGACCAGCCATGGATGGTTTATTTTCAAGTGGTTCAGGGCTAAGTCCA
TcSNP1126	7	CGTAAGGGAACTTACAAGAGTGAGGACATTGATATTGTGGCACTAAAGAACTAC
		[A/T]TGGATGCAGTACTTTGGTGAGATTGGTATTGGCACACCCACAGAACTACTGTGA
TcSNP1136	2	ACTTAAATATGAGTATCAGGACTAATATGGAAGGTGTTTACAAGCTAAATTGTGTTTTGG
		[C/G]CATGACCAGGACGCATGAGAGGTTCTCATTTCAACACTTTTGCAATTTCAGTACTTATTT
TcSNP1149	3	TCCTGCAGTTCCAAAAATCGCAATCCCCCTTGCTCTTTTCTACTCTTCTAGAGACTTCTC
		[T/C]TTGTTATTTCTTTCCTGTTCTCTCTTCTCTTGCTTTCCCCCTGCAATGGCGTCTGATCTG

TcSNP1159	2	ATTAAAGGTGAGCAGGAGACAGAAGGAGGTACTTGTGTTTGGCTTTAGCTTATTGGAA
		[A/C]GGAGGAGAAGAGKTTTGGTTGATTTGTTGGAGTTAAGAGGGAAGGGTAAGAAAGGAGGAA
TcSNP1165	2	GATCAATTGCTTAAACATATTTGCATAGCATTATAGATATTTTCAAAACCATGATGGAGG
		[T/C]TGCCAAGGTTTTCCTAGCTCTTACTTGTCTCATGGCCTTAGCCTTGGTCCTCCCCTCCCA
TcSNP1175	4	TAAGGCTACATTCGTGGTGTTTGTCACGGACGGATATGGGTACGCTCAAATGGCTTGCCC
		[T/G]CATCTCTCCAGACAGAGCCAGGGATCCCAAAGTGGAAGGCAAGACAGAAGAGAACAAGAA
TcSNP1194	4	TATCTTAATGGTCAAAGCCTCTCTTTTCTTGTTACAACCAGTGATGGCCGCAGTGTTGTC
		[T/G]CTTACAATGTTGCTCCAGCCGGTTGGTCCTTTGGACAGACTTACGCTGGAGGGCAGTTTA
TcSNP1253	9	CACTTGCCACTCAAGTTCCTCACTAAAGCATTGAAACCAGCATCAGAAGTCTGATTTCTA
		[T/G]ATTACCAGCACTTATCAAACTTCTTTCAATCAGGATAGGTGAGTAAAGAAATGAGGCGCT
TcSNP1270	7	ATATTTGAGTTCTGTTTGATGTTTACTACAAGACCTGCCTCACTTCTTCCCTCGCAGCTT
		[T/C]GTACCAAGAGCCCTYTCCAATTTACCAATTATCTGTTGATTAGGGATGGCTTTTCCAGAT
TcSNP1350	1	TCCAAAATTTTTTATCACTATATGATGGACCATATAGCCTAAATAAATATAGCAAAAATG
		[A/C]ATATCACAACAAATTATATGGCTGGCTTGCAAAAAGACTATAAGGGCTTGTGGCTAGTTC
TcSNP1383	7	GCATTCAGAAGGGTCCAGTTCGTGGCATTTCACTCAAGCTACAAGAAGAAGAGCGTGAGC
		[A/G]TCGCATGGACTTTGTTCCTGACGAGTCAGCTATTAACACTGATGAAATCAAGGTTGACAA
TcSNP1392	10	GGATGACATGCCAAAGCCTGGTCAGCCTGGTCAAGCTCCTCGTGCCGGTGCTGGAAATCC
		[T/C]GCACGTCCTTAAGGCCTTCAAAGCTTCAGACGAGTTTCTATATTTTGTTATGCTCTAGAA
TcSNP1414	9	TGTGACTTACGGTTACCCCAATCTCAAGAGCGTGAGGGAGTTGATTTACAAAAGAGGTTA
		[T/C]GGGAAGTTGAACAAGCAGCGTGTTGCTTTGACTGACAATGAAATCATTGAGCAGGCTCTG
TcSNP1439	9	AAGATTACAACAATAGGACCCCACGCAGCTATTAACAAACTTACCAACCCAAACCTGCAA
		[T/C]ACTTTTGCTTATTCTCCCCAAAACCTACATAGCAAATGGACCTACCTATA
TcSNP1442	9	AAAAGTGATGGAGGAAAGGGAAAGAGATGGGTGGGAGTAGAGATGGCCTTTGGTGTTTCC
		[T/C]TATGGATTTTCACTCTACATCAAAGCTTTGGTGTGTGATGGAATTTACAGCTACTGTTAT
TcSNP1484[a]	6	CAACCGTGAGAACTACATCAATGAGCTAATTAGGAGAATGAATTAAACTTGGCCTTGTTT
		[T/G]TTAATTTAGTAGTTGAGGTATTTTGGCTTATTTGAAAGGTTCTGATCKTAATTTCTTCCG

[a]TcSNP – marker excluded from the analysis
[b]LG – Linkage group
[c] [N/N]- SNP

D. Genetic duplicates groups and observed heterozygosity

The accessions nomenclature first refers to location (BdT for Bocas del Toro, and Her for Herrera), followed by the plantation and tree number, and lastly mentioning genotype of leaf tissue (L), root tissue (R), or 0, when it is a single origin tree (L=R). Genetic duplicates between scions and rootstocks represent clones, used as rootstock or scion. Genetic duplicates from single origin trees are either clones, planted as rooted cuttings, or could be very closely related siblings, from seeds of highly homozygous trees. Whether clones or closely related siblings, these unique SNP signatures are referred to as genetic duplicates group (GDG). Genetic duplicate groups have varying colors to highlight them, and commercial varieties imported to Panama are shaded in grey.

Accessions	GDG	H_o	Accessions	GDG	H_o	Accessions	GDG-	H_o	Accessions	GDG	H_o
BdT1.1-L	BdT1.1-L	46.9%	BdT5.3-0	GDG-3	1.0%	Her14.1-0	GDG-10	7.3%	BdT3.1-R	BdT3.1-R	25.0%
BdT1.2-L	BdT1.2-L	49.0%	BdT5.4-0	GDG-3	1.0%	Her14.2-0	GDG-10	7.3%	BdT3.2-R	GDG-5	22.9%
BdT1.3-L	BdT1.3-L	30.2%	BdT5.5-0	GDG-3	1.0%	Her14.3-0	GDG-10	7.3%	BdT3.3-R	GDG-6	74.0%
BdT1.4-0	BdT1.4-0	29.2%	BdT6.1-0	BdT6.1-0	34.4%	Her14.4-0	GDG-10	7.3%	BdT3.4-R	BdT3.4-R	44.8%
BdT1.5-L	BdT1.5-L	40.6%	BdT6.2-0	BdT6.2-0	36.5%	Her14.5-0	GDG-10	7.3%	BdT3.5-R	BdT3.5-R	32.3%
BdT10.1-L	CATIE-R4	38.5%	BdT6.3-0	BdT6.3-0	21.9%	Her15.1-0	Her15.1-0	56.3%	BdT4.1-R	GDG-6	70.8%
BdT10.2-L	ICS95	45.8%	BdT6.4-0	BdT6.4-0	2.1%	Her15.2-0	Her15.2-0	37.5%	BdT4.3-R	BdT4.3-R	40.6%
BdT10.3-L	CATIE-R1	37.5%	BdT6.5-0	BdT6.5-0	30.2%	Her15.3-0	Her15.3-0	57.3%	BdT4.4-R	BdT4.4-R	40.6%
BdT10.4-L	CATIE-R6	38.5%	BdT7.1-0	BdT7.1-0	20.8%	Her15.4-0	GDG-10	15.6%	BdT4.5-R	BdT4.5-R	38.5%
BdT10.5-L	PMCT58	28.1%	BdT7.2-0	BdT7.2-0	35.4%	Her15.5-0	Her15.5-0	50.0%	BdT9.1-R	BdT9.1-R	36.5%
BdT11.1-L	GDG-1	33.3%	BdT7.3-0	BdT7.3-0	29.2%	Her16.1-0	Her16.1-0	35.4%	BdT9.2-R	BdT9.2-R	45.8%
BdT11.2-L	GDG-1	33.3%	BdT7.4-0	BdT7.4-0	29.2%	Her16.2-0	GDG-10	10.4%	BdT9.3-R	BdT9.3-R	59.4%
BdT11.3-L	GDG-2	2.1%	BdT7.5-0	BdT7.5-0	34.4%	Her16.3-0	Her16.3-0	65.6%	BdT9.4-R	BdT9.4-R	28.1%
BdT11.4-L	GDG-2	2.1%	BdT8.1-0	BdT8.1-0	32.3%	Her16.4-0	Her16.4-0	46.9%	BdT9.5-R	BdT9.5-R	39.6%
BdT11.5-L	GDG-4	28.1%	BdT8.2-0	BdT8.2-0	36.5%	Her16.5-0	Her16.5-0	45.8%	Her12.2-R	GDG-8	32.3%
BdT11.6-L	GDG-4	27.1%	BdT8.3-0	BdT8.3-0	34.4%	BdT1.1-R	BdT1.1-R	75.0%	Her12.3-R	GDG-9	33.3%
BdT2.1-L	BdT2.1-L	33.3%	BdT8.4-0	BdT8.4-0	37.5%	BdT1.2-R	BdT1.2-R	31.3%	Her12.4-R	Her12.4-R	40.6%
BdT2.2-L	BdT2.2-L	45.8%	BdT8.5-0	BdT8.5-0	32.3%	BdT1.3-R	BdT1.3-R	59.4%	Her13.5-R	Her13.5-R	0.0%
BdT2.3-0	BdT2.3-0	30.2%	BdT9.1-L	ICS95	43.8%	BdT1.5-R	BdT1.5-R	51.0%			
BdT2.4-0	BdT2.4-0	37.5%	BdT9.2-L	CC137	36.5%	BdT10.1-R	BdT10.1-R	25.0%			
BdT2.5-0	BdT2.5-0	36.5%	BdT9.3-L	PMCT58	29.2%	BdT10.2-R	BdT10.2-R	42.7%			
BdT3.1-L	BdT3.1-L	29.2%	BdT9.4-L	PMCT58	29.2%	BdT10.3-R	BdT10.3-R	65.6%			
BdT3.2-L	BdT3.2-L	33.3%	BdT9.5-L	CATIE-R4	37.5%	BdT10.4-R	BdT10.4-R	50.0%			
BdT3.3-L	BdT3.3-L	29.2%	Her12.1-0	Her12.1-0	57.3%	BdT10.5-R	BdT10.5-R	10.4%			
BdT3.4-L	GDG-5	25.0%	Her12.2-L	GDG-3	1.0%	BdT11.1-R	GDG-7	74.0%			
BdT3.5-L	BdT3.5-L	33.3%	Her12.3-L	GDG-8	30.2%	BdT11.2-R	BdT11.2-R	20.8%			
BdT4.1-L	BdT4.1-L	31.3%	Her12.4-L	GDG-9	33.3%	BdT11.3-R	BdT11.3-R	63.5%			
BdT4.2-0	BdT4.2-0	19.8%	Her12.5-0	GDG-3	0.0%	BdT11.4-R	BdT11.4-R	34.4%			
BdT4.3-L	GDG-7	69.8%	Her13.1-0	Her13.1-0	51.0%	BdT11.5-R	BdT11.5-R	30.2%			
BdT4.4-L	BdT4.4-L	34.4%	Her13.2-0	Her13.2-0	25.0%	BdT11.6-R	BdT11.6-R	69.8%			
BdT4.5-L	BdT4.5-L	35.4%	Her13.3-0	Her13.3-0	12.5%	BdT2.1-R	BdT2.1-R	39.6%			
BdT5.1-0	GDG-3	1.0%	Her13.4-0	Her13.4-0	15.6%	BdT2.2-R	BdT2.2-R	56.3%			
BdT5.2-0	GDG-3	1.0%	Her13.5-L	Her13.5-L	17.7%						

E. Genetic differences (in %) in SNP profiles between scion and rootstock.

The trees, their variability as the difference in percentage of the SNP profiles between scion and rootstock and the origin of the tree. Different SNP profiles between scion and rootstock suggests the tree has a grafted origin, while similar profiles suggest a single origin tree from seeds or rooted cuttings. Values below 6% indicate a single origin, while values above 6% indicate different origins.

Accessions	Variability	Origin	Accessions	Variability	Origin	Accessions	Variability	Origin
BdT1.1	24.48%	Different	BdT7.5	1.04%	Single	Her14.3	1.04%	Single
BdT1.2	13.02%	Different	BdT8.1	0.00%	Single	Her14.4	1.04%	Single
BdT1.3	28.65%	Different	BdT8.2	2.08%	Single	Her14.5	1.04%	Single
BdT1.4	1.04%	Single	BdT8.3	1.04%	Single	Her15.1	2.08%	Single
BdT1.5	10.42%	Different	BdT8.4	1.04%	Single	Her15.2	3.13%	Different
BdT2.1	31.77%	Different	BdT8.5	1.04%	Single	Her15.3	1.04%	Single
BdT2.2	27.08%	Different	BdT9.1	39.06%	Different	Her15.4	1.04%	Single
BdT2.3	1.04%	Single	BdT9.2	29.69%	Different	Her15.5	1.04%	Single
BdT2.4	3.13%	Different	BdT9.3	31.77%	Different	Her16.1	1.04%	Single
BdT2.5	3.13%	Different	BdT9.4	26.56%	Different	Her16.2	1.04%	Single
BdT3.1	35.42%	Different	BdT9.5	29.17%	Different	Her16.3	1.04%	Single
BdT3.2	19.79%	Different	BdT10.1	33.85%	Different	Her16.4	0.00%	Single
BdT3.3	41.67%	Different	BdT10.2	32.81%	Different	Her16.5	1.04%	Single
BdT3.4	28.65%	Different	BdT10.3	32.29%	Different			
BdT3.5	23.44%	Different	BdT10.4	32.29%	Different			
BdT4.1	44.27%	Different	BdT10.5	29.17%	Different			
BdT4.2	3.65%	Single	BdT11.1	41.15%	Different			
BdT4.3	40.10%	Different	BdT11.2	25.00%	Different			
BdT4.4	32.29%	Different	BdT11.3	35.94%	Different			
BdT4.5	39.06%	Different	BdT11.4	36.98%	Different			
BdT5.1	1.56%	Single	BdT11.5	23.44%	Different			
BdT5.2	1.56%	Single	BdT11.6	38.54%	Different			
BdT5.3	1.56%	Single	Her12.1	2.08%	Single			
BdT5.4	1.56%	Single	Her12.2	18.75%	Different			
BdT5.5	1.56%	Single	Her12.3	9.90%	Different			
BdT6.1	2.08%	Single	Her12.4	22.92%	Different			
BdT6.2	1.04%	Single	Her12.5	2.08%	Single			
BdT6.3	3.13%	Different	Her13.1	2.08%	Single			
BdT6.4	1.04%	Single	Her13.2	1.04%	Single			
BdT6.5	2.08%	Single	Her13.3	1.04%	Single			
BdT7.1	1.04%	Single	Her13.4	1.04%	Single			
BdT7.2	3.13%	Different	Her13.5	36.46%	Different			
BdT7.3	2.08%	Single	Her14.1	1.04%	Single			
BdT7.4	1.04%	Single	Her14.2	1.04%	Single			

The summary of the different and single origin across trees, farms per region and each farm. Following the chi-square test for proportions between different and single-origin trees between regions.

% of Trees	Different 43%	Single 57%
Farms of Bocas del Toro	55%	45%
BdT 1	80.00%	20.00%
BdT 2	40.00%	60.00%
BdT 3	100.00%	0.00%
BdT 4	80.00%	20.00%
BdT 5	0.00%	100.00%
BdT 6	0.00%	100.00%
BdT 7	0.00%	100.00%
BdT 8	0.00%	100.00%
BdT 9	100.00%	0.00%
BdT 10	100.00%	0.00%
BdT 11	100.00%	0.00%
Farms of Herrera	16%	84%
Her 12	60.00%	40.00%
Her 13	20.00%	80.00%
Her 14	0.00%	100.00%
Her 15	0.00%	100.00%
Her 16	0.00%	100.00%

	Observed	**Expected**	**Difference**	**Squared Difference**	**Squared Difference / Expected**
Bocas del Toro (Different)	31	24	7	49	2.04
Bocas del Toro (Single)	25	32	-7	49	1.53
Herrera (Different)	4	11	-7	49	4.45
Herrera (Single)	21	14	7	49	3.50
				Chi-Square	11.53
				p-value	*0.00095634*

F. Panamanian Germplasm

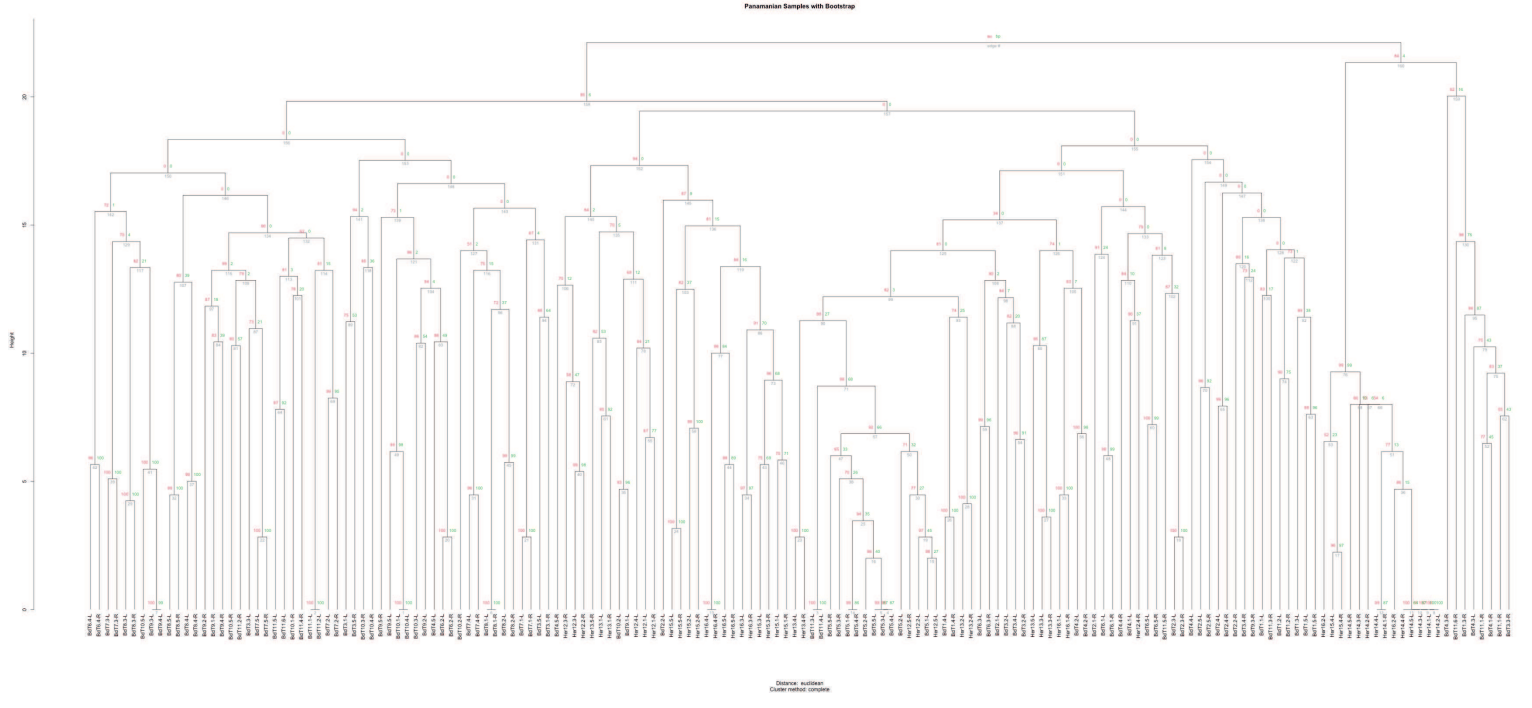


Figure S1. Hierarchical clustering with bootstrap values based on the Euclidean distances are shown. Numbers in grey represent the number of the node, whereas the one in red are the approximately unbiased probability (au) reflecting percentage p-values supporting that node and the numbers in green are the bootstrap probability (bp) reflecting how consistently this bootstrap appears.

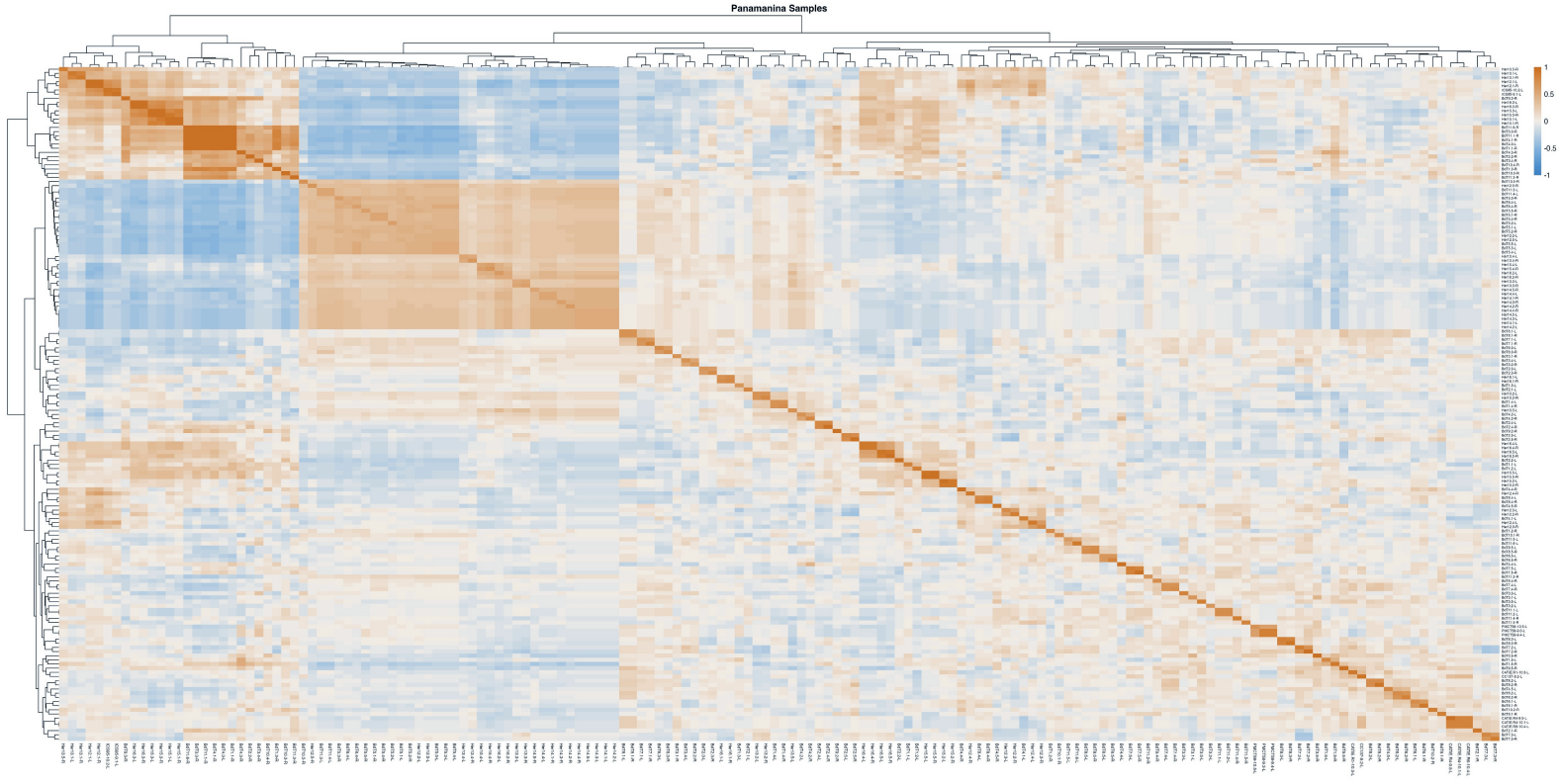


Figure S2. Kinship Matrix of Panamanian Germplasm Highlighted in the upper left corner are two sets of trees from both regions that are highly similar to each other, while the remaining genotypes show greater admixture. The first square in the far upper left represents the Criollo group, and the second square next to it represents the purer Amelonado group. All other genotypes are admixed.

G. Panamanian and Reference Germplasm

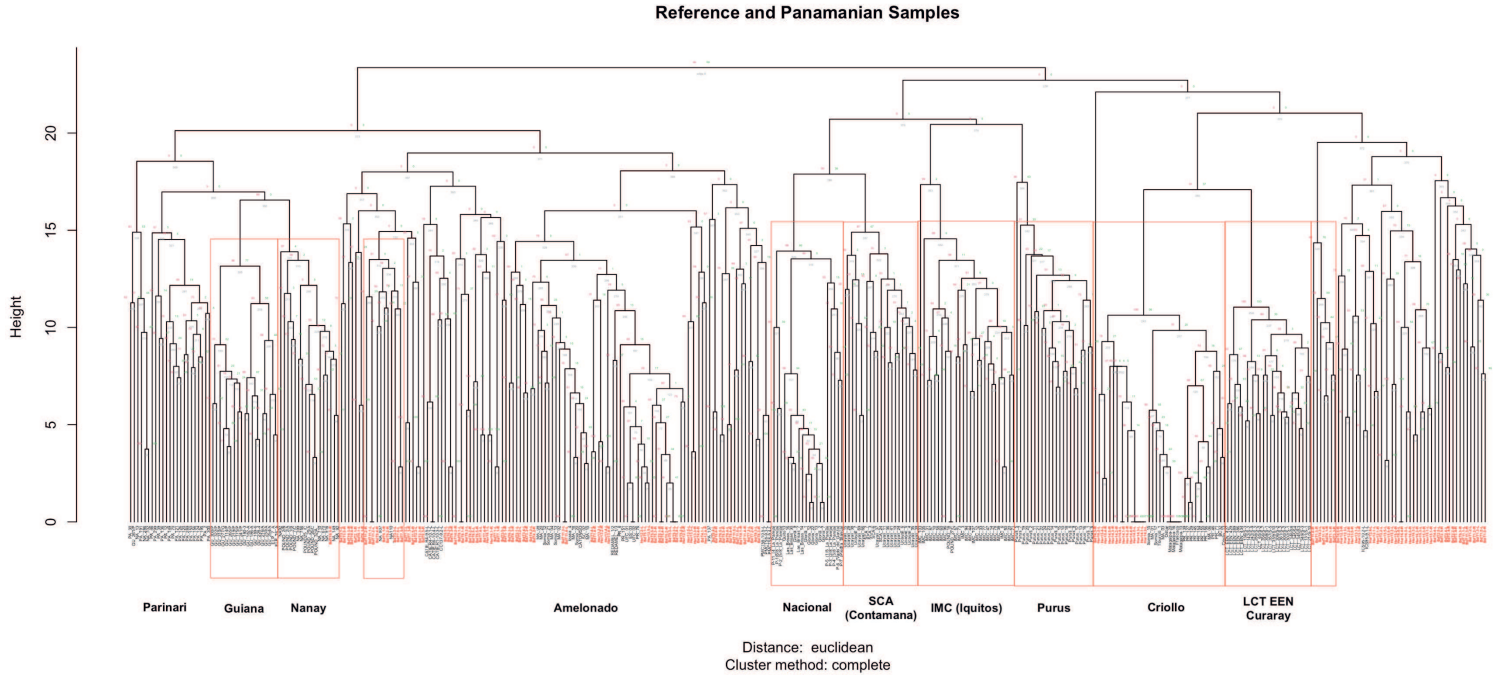


Figure S3. Hierarchical clustering with bootstrap values based on the Euclidean distances are shown. Numbers in grey represent the number of the node, whereas the ones in red are the approximately unbiased probability (au) reflecting percentage p-values supporting that node and the ones in green are the bootstrap probability (bp) reflecting how consistently this bootstrap appears. Red rectangles show groups with bootstrap values over 90%. The red labels are used for the Panamanian germplasm, and the black labels are the reference populations.

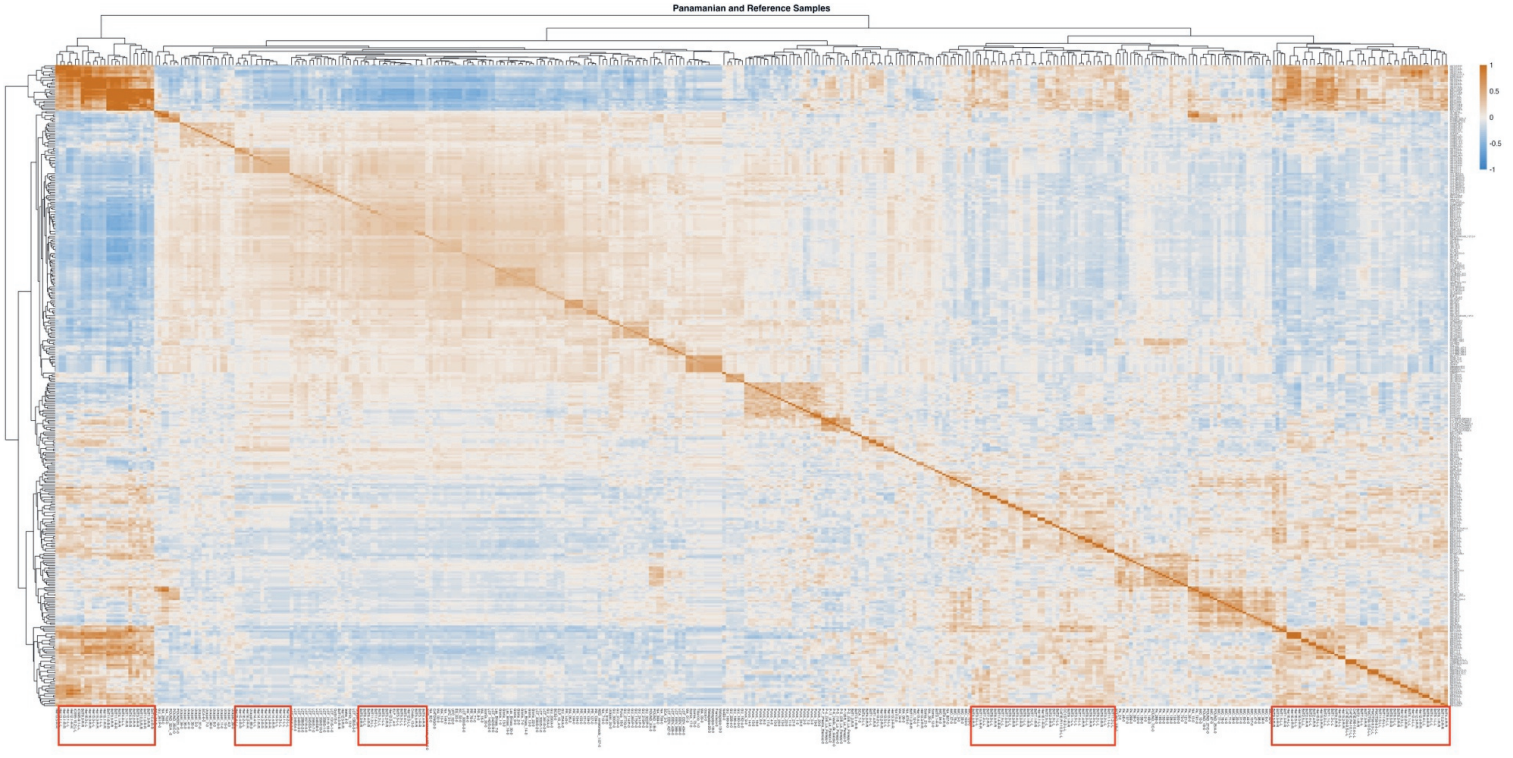


Figure S4. Kinship Matrix of Panamanian Germplasm (signaled inside red squares) and Reference Germplasm.

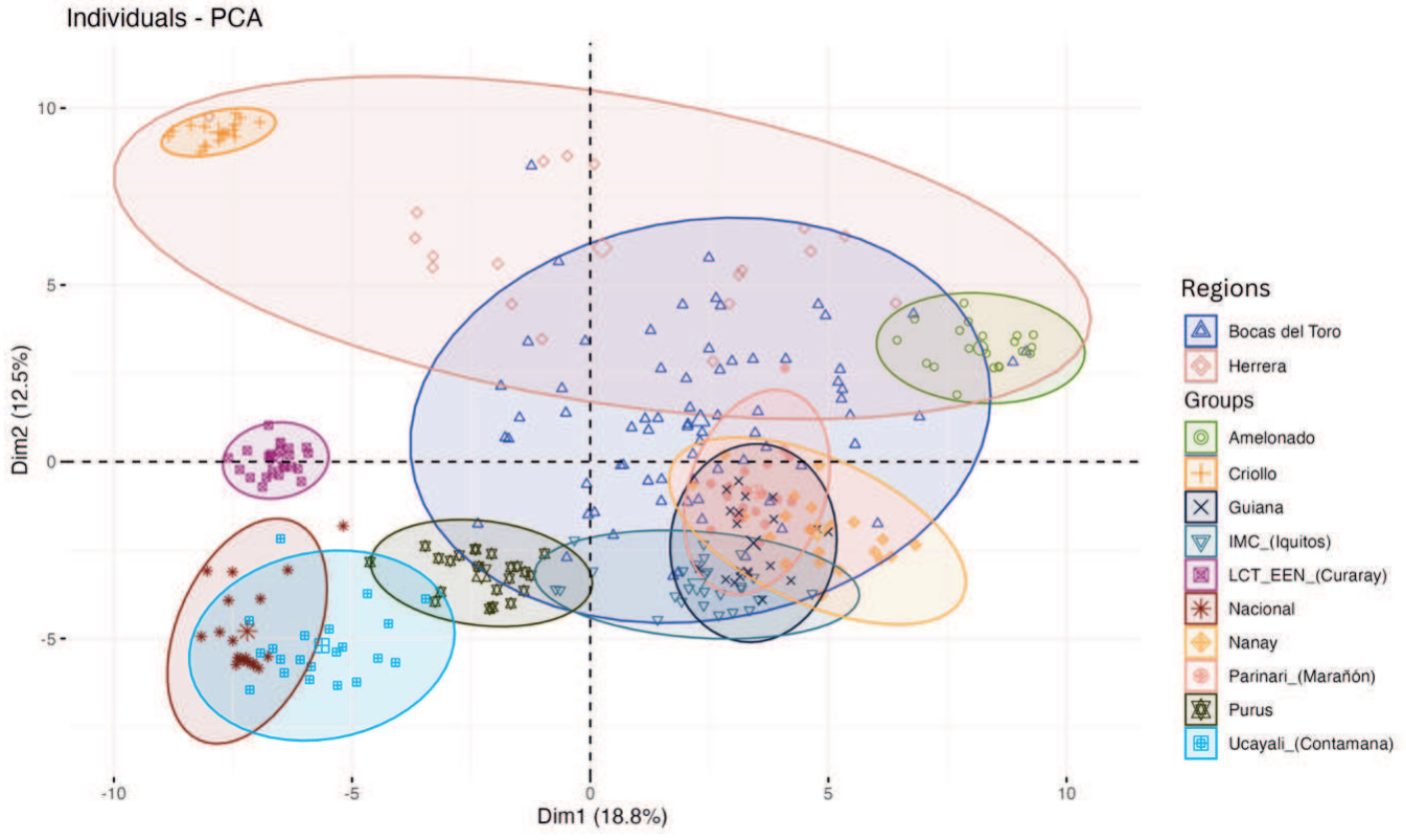


Figure S5. Principal component analysis of the Panamanian and reference germplasm. Most of the material from Panamanian farms is either a varied mixture between the cacao genetic groups of Criollo and Amelonado, typically called Trinitario in the past, or a varied mixture between multiple different groups except Ucayali, LCT EEN Curaray.

H. Secondary tree species

The most and least frequently occurring neighbouring tree species in cacao farms. The first table summarizes all 16 surveyed sites in Panama, while the second table shows the results for each individual farm.

Species	Timber / Medicinal / Edible	Neighbouring Individuals	Overall Presence Percentage
Theobroma cacao	Edible	91	39.0%
Musa spp	Edible	44	18.8%
Cordia alliodora	Timber	19	8.1%
Handroanthus	Timber	9	3.8%
Coffea spp	Edible	8	3.4%
Bactris gasipaes	Edible	7	3.0%
Gustavia superba	Edible	7	3.0%
Cocos nucifera	Edible	4	1.7%
Artocarpus communis	Edible / Medicinal	3	1.3%
Citrus sinensis	Edible	3	1.3%
Crescentia cujete	Medicinal	3	1.3%
Guazuma ulmifolia	Medicinal	3	1.3%
Bixa orellana	Edible / Medicinal	2	0.9%
Citrus aurantiifolia	Fruit	2	0.9%
Bombacopsis quinata	Timber	2	0.9%
Persea americana	Edible	2	0.9%
Psidium spp	Edible / Timber	2	0.9%
Swietenia macrophylla	Timber	2	0.9%
Virola spp	Timber	2	0.9%
Annona muricata	Edible	1	0.4%
Carica papaya	Edible	1	0.4%
Cecropia spp	Medicinal / Timber	1	0.4%
Ceratonia siliqua	Timber	1	0.4%
Citrus paradisi	Edible	1	0.4%
Curcuma longa	Edible	1	0.4%
Dyphisa americana	Timber	1	0.4%
Enterolobium cyclocarpum	Medicinal / Timber / Edible	1	0.4%
Gliricidia sepium	Timber	1	0.4%
Herrania spp	NA	1	0.4%
Mangifera indica	Edible	1	0.4%
Nephelium lappaceum	Edible	1	0.4%
Pterocarpus officinalis	Timber	1	0.4%
Socratea exorrhiza	Timber	1	0.4%
Spondias mombin	Edible	1	0.4%
Syzygium malaccense	Edible	1	0.4%
Terminalia cattapa	Timber	1	0.4%
Vochysia ferruginea	Timber	1	0.4%
Unknown		9	3.9%
Total		233	100.0%

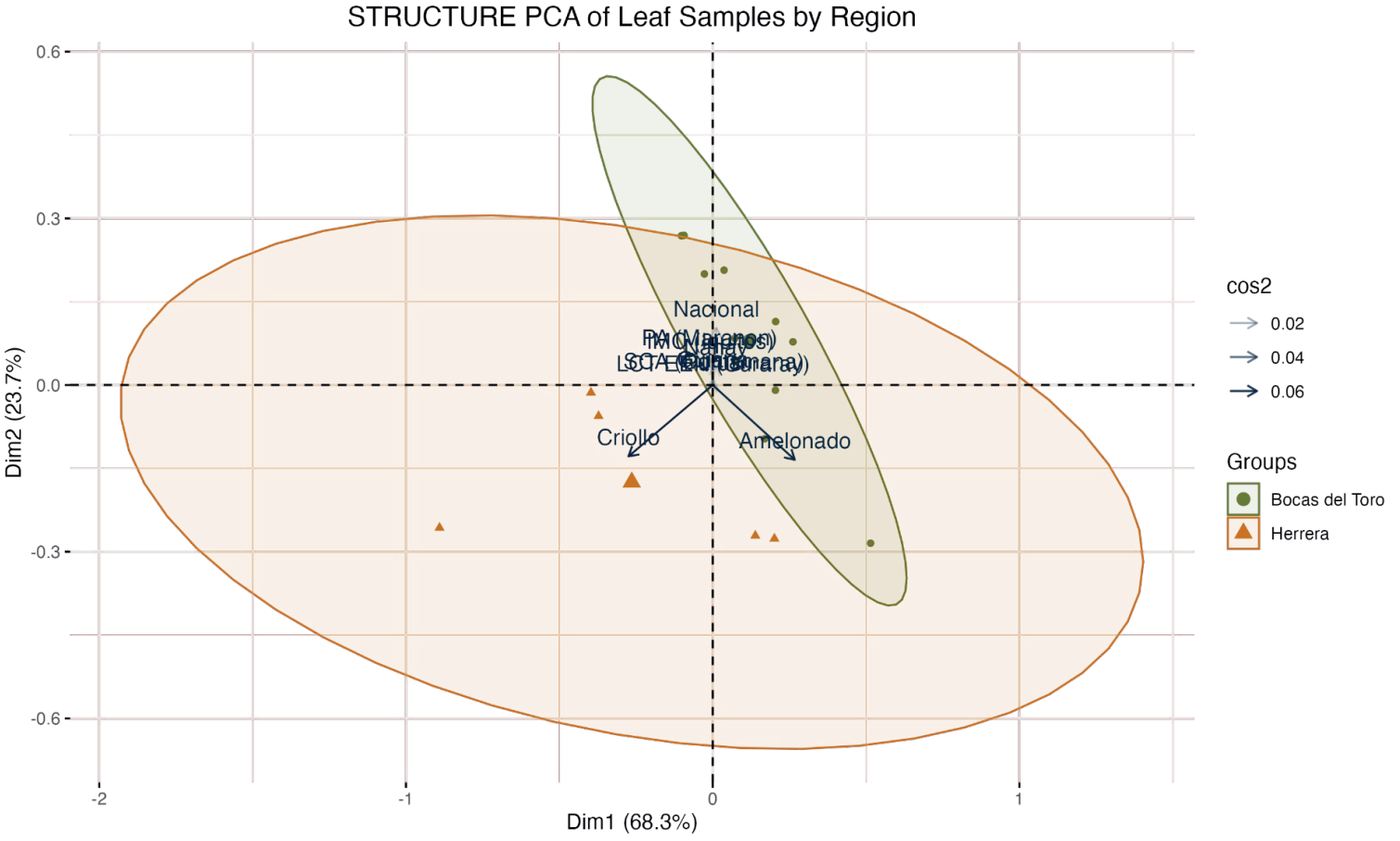
Plantations	Top tree	2nd tree	3rd tree
BdT_1	Musa spp: 5	Cordia alliodora: 3	Gustavia superba: 3
BdT_2	Cordia alliodora: 6	Gustavia superba: 2	Unknown: 2
BdT_3	Cordia alliodora: 3	Musa spp: 2	Gustavia superba: 1
BdT_4	Musa spp: 6	Artocarpus communis: 1	Curcuma longa: 1
BdT_5	Cordia alliodora: 2	Musa spp: 2	Pterocarpus officinalis: 1
BdT_6	Musa spp: 5	Cordia alliodora: 2	Guazuma ulmifolia: 2
BdT_7	Musa spp: 4	Cordia alliodora: 2	Bactris gasipaes: 1
BdT_8	Musa spp: 2	Bactris gasipaes: 1	Cordia alliodora: 1
BdT_9	Tabebuia spp: 1	Unknown: 1	Annona muricata: 0
BdT_10	Tabebuia spp: 2	Annona muricata: 0	Artocarpus communis: 0
BdT_11	Musa spp: 6	Bactris gasipaes: 1	Cocos nucifera: 1
Her_12	Artocarpus communis: 2	Bixa orellana: 2	Carico papaya: 1
Her_13	Musa spp: 3	Crescentia cujete: 2	Tabebuia spp: 2
Her_14	Coffea spp: 7	Pachira quinata: 2	Psidium spp: 2
Her_15	Musa spp: 4	Citrus paradisi: 1	Dyphisa americana: 1
Her_16	Musa spp: 3	Citrus aurantiifolia: 2	Citrus sinensis: 2

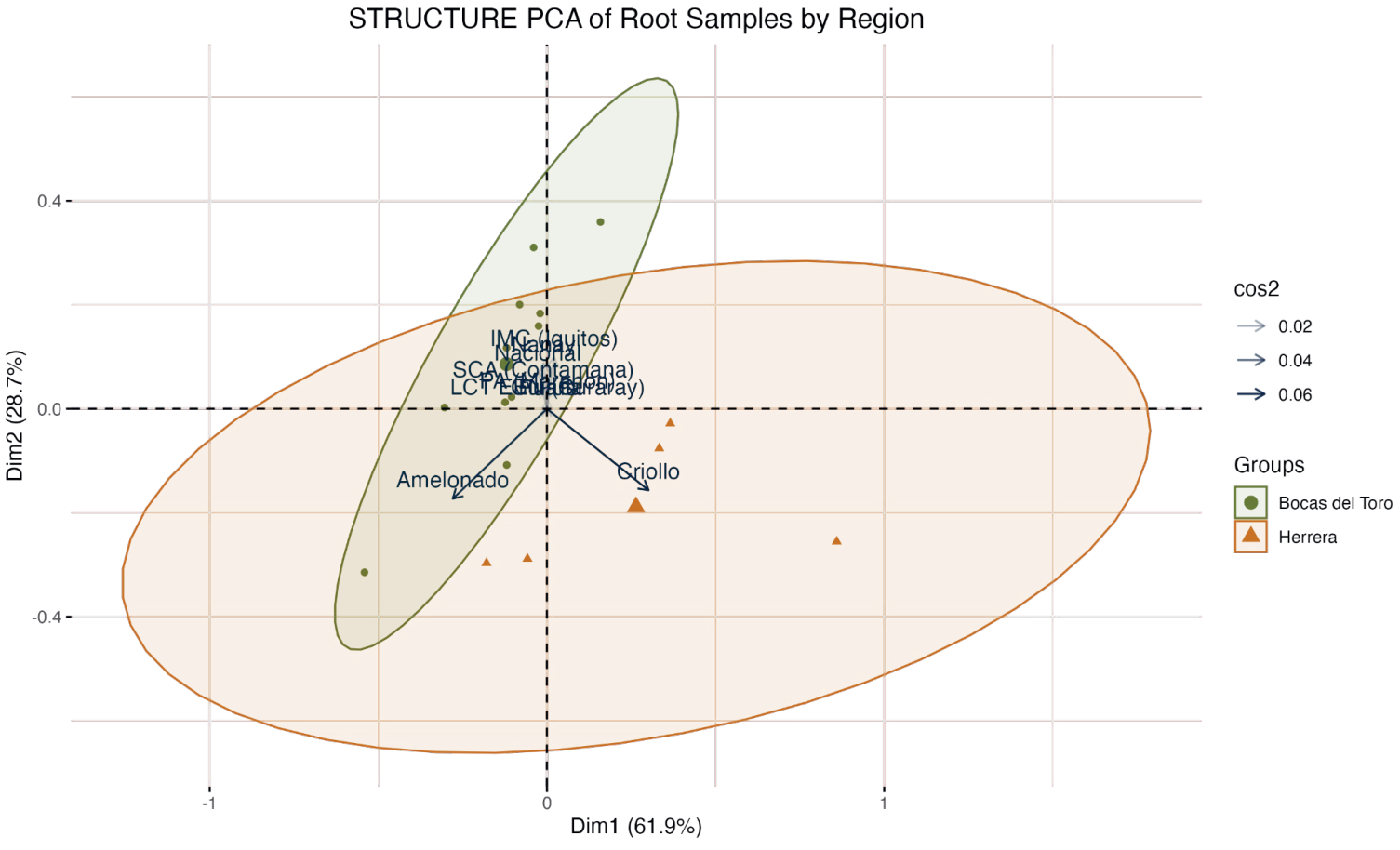
I. Genetic structure differences between regions and plant parts

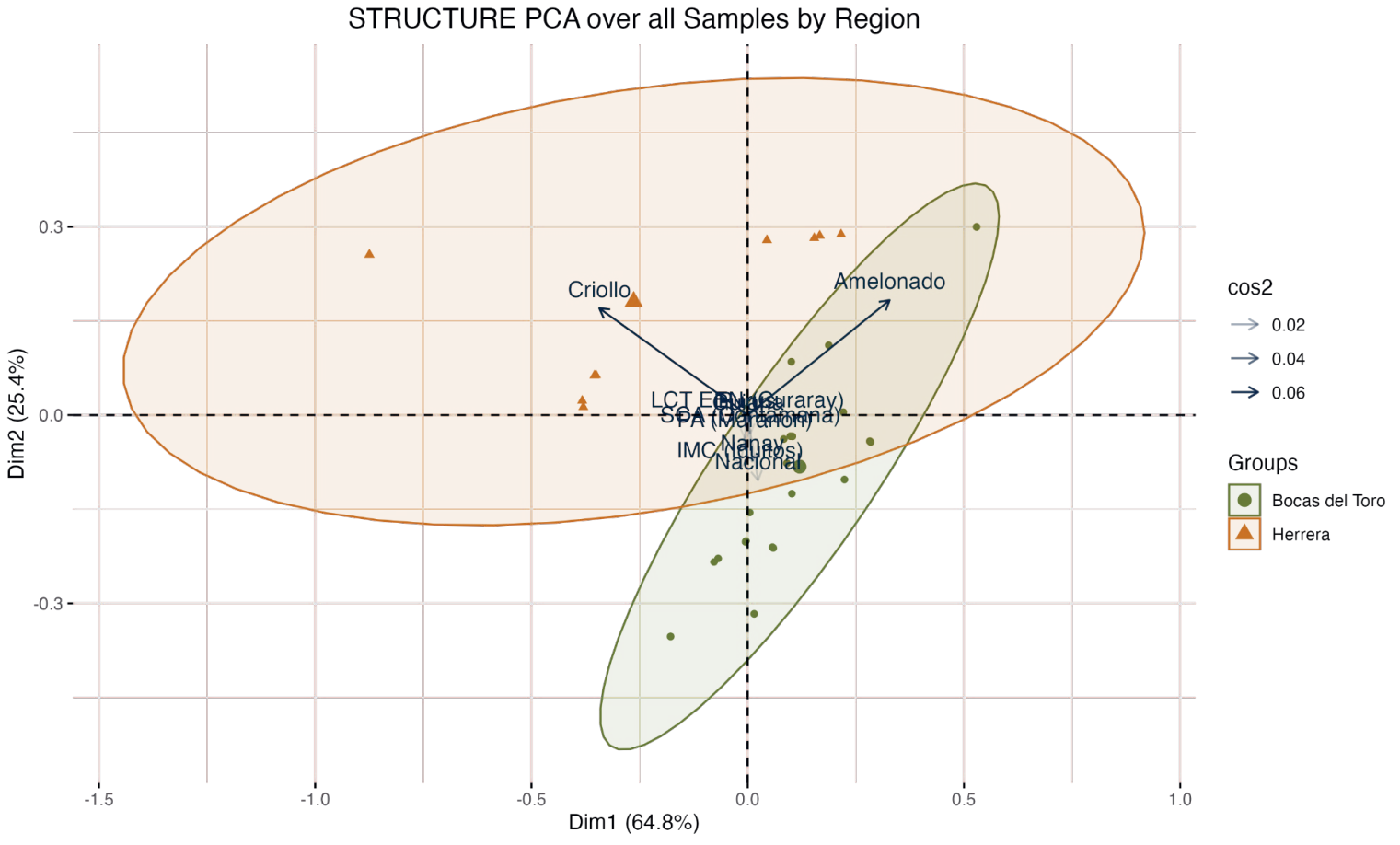
Permutational multivariate analysis of variance (PERMANOVA) under a reduced model (adonis2, Bray–Curtis distance, 999 permutations) showed that plant genetic structure significantly differed between region, explaining 18.3% of the variation ($F = 35.88$, $p = 0.001$). In contrast, plant genetic structure did not significantly differ between plant tissue, explaining only 0.6% of the variation ($F = 0.92$, $p = 0.453$).

Factor	Df	Sum of Squares	R^2	F.Model	Pr(>F)
Region	1	5.7335	0.1832	35.876	0.001
Residual	160	25.5707	0.8169		
Total	161	31.3043	1.0000		
ab.bel	1	0.1796	0.0057	0.9233	0.453
Residual	160	31.1247	0.9943		
Total	161	31.3043	1.0000		

J. PCA derived from the STRUCTURE analysis




STRUCTURE PCA of Root Samples by Region
Dim2 (28.7%)
Dim1 (61.9%)
0.4
0.0
-0.4
-1
0
1
IMC (Iquitos)
Nacional
SCA (Contamana)
Amelonado
Criollo
cos2
0.02
0.04
0.06
Groups
Bocas del Toro
Herrera


STRUCTURE PCA over all Samples by Region
Criollo
Amelonado
Nanay
Nacional
0.3
0.0
-0.3
-1.5
-1.0
-0.5
0.0
0.5
1.0
Dim2 (25.4%)
Dim1 (64.8%)
cos2
0.02
0.04
0.06
Groups
Bocas del Toro
Herrera

K. PCA-Ridge Model

We used as genetic response variables the principal components 1 and 2 from the structure-resultant PCAs and the relevant environmental factors to perform a PCA-LASSO model. The numbers shaded blue is where the environmental factor is related to an increase in percentage of that genetic group, and in yellow, when the environmental factor is related to a decrease in percentage of that genetic group. PC1 explains the variation between Amelonado and Criollo mainly, and PC2 explains the variation between Amelonado/Criollo and all other groups.

For scions

PC1 explained 55.2% of the genetic structure variation and had a goodness of fit (R^2) of 17.25%

R^2 = 17.25 %	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
TN (%)	-0.0006	0.0010	0.0008	0.0000	-0.0001	0.0001	-0.0008	0.0218	0.0003	-0.0223
Al (Cmol/kg)	-0.0001	0.0002	0.0002	0.0000	0.0000	0.0000	-0.0002	0.0050	0.0001	-0.0051
Mg (Cmol/kg)	-0.0001	0.0002	0.0002	0.0000	0.0000	0.0000	-0.0002	0.0043	0.0001	-0.0044
Ca (Cmol/kg)	0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	-0.0001	0.0017	0.0000	-0.0017
OM %	0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	-0.0001	0.0017	0.0000	-0.0017
P (mg/l)	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0013	0.0000	-0.0013
Base Saturation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	-0.0003
%Sand	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	-0.0002
Fe (mg/l)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	-0.0002
%Silt	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Elevation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0001
%Clay	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0008	0.0000	0.0008
pH	0.0003	-0.0005	-0.0004	0.0000	0.0001	0.0000	0.0004	-0.0101	-0.0001	0.0104

PC2 explained 24.3% of the genetic structure variation and had a goodness of fit (R^2) of 40.98%.

R^2 = 40.98 %	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
pH	-0.0075	-0.0171	-0.0056	-0.0020	-0.0010	-0.0006	-0.0084	0.0223	-0.0014	0.0213
%Silt	-0.0002	-0.0004	-0.0001	0.0000	0.0000	0.0000	-0.0002	0.0005	0.0000	0.0005
P (mg/l)	-0.0001	-0.0003	-0.0001	0.0000	0.0000	0.0000	-0.0002	0.0004	0.0000	0.0004
Base Saturation	-0.0001	-0.0003	-0.0001	0.0000	0.0000	0.0000	-0.0002	0.0004	0.0000	0.0004
%Clay	-0.0001	-0.0003	-0.0001	0.0000	0.0000	0.0000	-0.0001	0.0004	0.0000	0.0004
Elevation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Fe (mg/l)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
%Sand	0.0001	0.0002	0.0001	0.0000	0.0000	0.0000	0.0001	-0.0003	0.0000	-0.0003
Ca (Cmol/kg)	0.0002	0.0004	0.0001	0.0000	0.0000	0.0000	0.0002	-0.0005	0.0000	-0.0005
Mg (Cmol/kg)	0.0002	0.0005	0.0002	0.0001	0.0000	0.0000	0.0003	-0.0007	0.0000	-0.0007
OM %	0.0017	0.0040	0.0013	0.0005	0.0002	0.0001	0.0019	-0.0052	0.0003	-0.0049
Al (Cmol/kg)	0.0023	0.0053	0.0017	0.0006	0.0003	0.0002	0.0026	-0.0069	0.0004	-0.0066
TN (%)	0.0176	0.0399	0.0130	0.0046	0.0023	0.0015	0.0196	-0.0520	0.0032	-0.0496

For rootstocks

PC1 explained 49.9% of the genetic structure variation and had a goodness of fit (R^2) of 14.73%.

R^2 = 14.73 %	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
TN (%)	-0.0020	0.0026	0.0010	0.0001	-0.0001	-0.0002	-0.0002	0.0259	0.0009	-0.0279
Mg (Cmol/kg)	-0.0002	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0025	0.0001	-0.0027
Al (Cmol/kg)	-0.0001	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0018	0.0001	-0.0020
OM %	-0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0015	0.0001	-0.0016
Ca (Cmol/kg)	-0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0013	0.0000	-0.0014
P (mg/l)	-0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0000	-0.0007
Base Saturation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	-0.0003
%Silt	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	-0.0002
Fe (mg/l)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	-0.0002
%Sand	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	-0.0002
Elevation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0001
%Clay	0.0001	-0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0009	0.0000	0.0009
pH	0.0005	-0.0007	-0.0003	0.0000	0.0000	0.0001	0.0001	-0.0071	-0.0003	0.0077

PC2 explained 24.1% of the genetic structure variation and had a goodness of fit (R^2) of 37.47%.

R^2 = 37.47 %	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
pH	-0.0165	-0.0118	-0.0145	-0.0008	-0.0008	-0.0010	-0.0029	0.0287	-0.0065	0.0260
%Silt	-0.0004	-0.0003	-0.0003	0.0000	0.0000	0.0000	-0.0001	0.0006	-0.0001	0.0006
Base Saturation	-0.0002	-0.0001	-0.0002	0.0000	0.0000	0.0000	0.0000	0.0003	-0.0001	0.0003
Elevation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
%Clay	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Fe (mg/l)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000
P (mg/l)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	-0.0001
%Sand	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	-0.0002	0.0001	-0.0002
Ca (Cmol/kg)	0.0004	0.0003	0.0004	0.0000	0.0000	0.0000	0.0001	-0.0008	0.0002	-0.0007
Mg (Cmol/kg)	0.0009	0.0006	0.0008	0.0000	0.0000	0.0001	0.0002	-0.0015	0.0003	-0.0014
OM %	0.0030	0.0022	0.0027	0.0002	0.0001	0.0002	0.0005	-0.0053	0.0012	-0.0048
Al (Cmol/kg)	0.0052	0.0037	0.0046	0.0003	0.0002	0.0003	0.0009	-0.0090	0.0020	-0.0081
TN (%)	0.0265	0.0189	0.0233	0.0013	0.0013	0.0015	0.0046	-0.0460	0.0104	-0.0417

Across all samples

PC1 explained 52.4% of the genetic structure variation and had a goodness of fit (R^2) of 42.59%.

R^2 = 42.59 %	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
TN (%)	-0.0036	0.0048	0.0021	0.0002	-0.0003	-0.0001	-0.0013	0.0659	0.0013	-0.0689
OM %	-0.0009	0.0013	0.0006	0.0001	-0.0001	0.0000	-0.0003	0.0172	0.0003	-0.0180
Mg (Cmol/kg)	-0.0003	0.0004	0.0002	0.0000	0.0000	0.0000	-0.0001	0.0050	0.0001	-0.0052
P (mg/l)	-0.0002	0.0003	0.0001	0.0000	0.0000	0.0000	-0.0001	0.0035	0.0001	-0.0037
Ca (Cmol/kg)	-0.0001	0.0002	0.0001	0.0000	0.0000	0.0000	-0.0001	0.0026	0.0001	-0.0027
%Silt	-0.0001	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0023	0.0000	-0.0024
Fe (mg/l)	-0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0013	0.0000	-0.0013
pH	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	-0.0001
%Sand	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Base Saturation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Elevation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0004	0.0000	0.0005
%Clay	0.0002	-0.0002	-0.0001	0.0000	0.0000	0.0000	0.0001	-0.0034	-0.0001	0.0035
Al (Cmol/kg)	0.0003	-0.0004	-0.0002	0.0000	0.0000	0.0000	0.0001	-0.0056	-0.0001	0.0059

PC2 explained 23% of the genetic structure variation and had a goodness of fit (R^2) of 61.93%.

R^2 = 61.93 %	IMC (Iquitos)	Nacional	Nanay	Guiana	Purus	LCT EEN (Curaray)	PA (Maranon)	Amelonado	SCA (Contamana)	Criollo
pH	-0.0343	-0.0416	-0.0299	-0.0038	-0.0025	-0.0022	-0.0149	0.0731	-0.0119	0.0680
%Silt	-0.0003	-0.0003	-0.0002	0.0000	0.0000	0.0000	-0.0001	0.0006	-0.0001	0.0006
Elevation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
%Clay	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
Base Saturation	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	-0.0001
%Sand	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0001	-0.0003	0.0000	-0.0002
P (mg/l)	0.0003	0.0004	0.0003	0.0000	0.0000	0.0000	0.0001	-0.0007	0.0001	-0.0006
Fe (mg/l)	0.0003	0.0004	0.0003	0.0000	0.0000	0.0000	0.0001	-0.0007	0.0001	-0.0007
Mg (Cmol/kg)	0.0007	0.0008	0.0006	0.0001	0.0000	0.0000	0.0003	-0.0015	0.0002	-0.0014
Ca (Cmol/kg)	0.0018	0.0022	0.0016	0.0002	0.0001	0.0001	0.0008	-0.0039	0.0006	-0.0036
Al (Cmol/kg)	0.0020	0.0024	0.0018	0.0002	0.0001	0.0001	0.0009	-0.0043	0.0007	-0.0040
OM %	0.0088	0.0107	0.0077	0.0010	0.0006	0.0006	0.0038	-0.0188	0.0031	-0.0175
TN (%)	0.0516	0.0625	0.0449	0.0058	0.0037	0.0034	0.0224	-0.1099	0.0179	-0.1021

References

Almeida, N.M. de, Almeida, A.-A.F. de, Santos, N. de A., Nascimento, J.L. do, Neto, C.H. de C., Pirovani, C.P., Ahnert, D., Baligar, V.C., 2022. Scion-rootstock interaction and tolerance to cadmium toxicity in juvenile *Theobroma cacao* plants. Sci. Hortic. 300, 111086. https://doi.org/10.1016/j.scienta.2022.111086

Argout, X., Fouet, O., Wincker, P., Gramacho, K., Legavre, T., Sabau, X., Risterucci, A.M., Silva, C.D., Cascardo, J., Allegre, M., Kuhn, D., Verica, J., Courtois, B., Loor, G., Babin, R., Sounigo, O., Ducamp, M., Guiltinan, M.J., Ruiz, M., Alemanno, L., Machado, R., Phillips, W., Schnell, R., Gilmour, M., Rosenquist, E., Butler, D., Maximova, S., Lanaud, C., 2008. Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. BMC Genom. 9, 512. https://doi.org/10.1186/1471-2164-9-512

Argout, X., Salse, J., Aury, J.-M., Guiltinan, M.J., Droc, G., Gouzy, J., Allegre, M., Chaparro, C., Legavre, T., Maximova, S.N., Abrouk, M., Murat, F., Fouet, O., Poulain, J., Ruiz, M., Roguet, Y., Rodier-Goud, M., Barbosa-Neto, J.F., Sabot, F., Kudrna, D., Ammiraju, J.S.S., Schuster, S.C., Carlson, J.E., Sallet, E., Schiex, T., Dievart, A., Kramer, M., Gelley, L., Shi, Z., Bérard, A., Viot, C., Boccara, M., Risterucci, A.M., Guignon, V., Sabau, X., Axtell, M.J., Ma, Z., Zhang, Y., Brown, S., Bourge, M., Golser, W., Song, X., Clement, D., Rivallan, R., Tahi, M., Akaza, J.M., Pitollat, B., Gramacho, K., D'Hont, A., Brunel, D., Infante, D., Kebe, I., Costet, P., Wing, R., McCombie, W.R., Guiderdoni, E., Quetier, F., Panaud, O., Wincker, P., Bocs, S., Lanaud, C., 2011. The genome of *Theobroma cacao*. Nat. Genet. 43, 101–108. https://doi.org/10.1038/ng.736

Bekele, F., Phillips-Mora, W., 2019. Advances in Plant Breeding Strategies: Industrial and Food Crops, Volume 6, 409–487. https://doi.org/10.1007/978-3-030-23265-8_12

Berjak, P., Pammenter, N.W., 2017. Recalcitrance. In: Encyclopedia of Applied Plant Sciences (Second Edition), 532–539. https://doi.org/10.1016/b978-0-12-394807-6.00048-4

CacaoNet, 2025. Partners and networks in the Global Strategy [WWW Document]. URL https://www.cacaonet.org/partners-and-networks-in-the-global-strategy?utm_source=chatgpt.com (accessed 1.20.25).

CIRAD, 2025. Cocoa roadmap summary: The road to sustainable cocoa growing [2022-2032].

Clement, C.R, Cristo-Araújo, M.D., D'Eeckenbrugge, G.C., Pereira, A.A., Picanço-Rodrigues, D., 2010. Origin and Domestication of Native Amazonian Crops. Diversity 2, 72–106. https://doi.org/10.3390/d2010072

Cacao of Excellence, 2021 Cocoa of Excellence Awards [WWW Document]. URL https://www.cacaoofexcellence.org/2021-edition/2021-awards-winners-producers-list

Cosme, S., Cuevas, H.E., Zhang, D., Oleksyk, T.K., Irish, B.M., 2016. Genetic diversity of naturalized cacao (*Theobroma cacao* L.) in Puerto Rico. Tree Genet. Genomes 12, 88. https://doi.org/10.1007/s11295-016-1045-4

Cuatrecasas, 1964. Cacao and its Allies. A taxonomic revision of the Genus *Theobroma*. Contribution US Herbarium 35: 379–614.

Darikova, J.A., Savva, Y.V., Vaganov, E.A., Grachev, A.M., Kuznetsova, G.V., 2011. Grafts of Woody Plants and the Problem of Incompatibility Between Scion and Rootstock (a review). J. Sib. Fed. Univ. Biol. 4, 54–63. https://doi.org/10.17516/1997-1389-0185

DuVal, A., Gezan, S.A., Mustiga, G., Stack, C., Marelli, J.-P., Chaparro, J., Livingstone, D., Royaert, S., Motamayor, J.C., 2017. Genetic Parameters and the Impact of Off-Types for *Theobroma cacao* L. in a Breeding Program in Brazil. Front. Plant Sci. 8, 2059. https://doi.org/10.3389/fpls.2017.02059

Echeverri-Rodriguez, J., 2006. El injerto en la producción de cacao órganico. Manejo Integrado de Plagas y Agroecología 78, 21–23. https://doi.org/10.18356/f4dd1c2b-es

Elwers, S., Zambrano, A., Rohsius, C., Lieberei, R., 2009. Differences between the content of phenolic compounds in Criollo, Forastero and Trinitario cocoa seed (*Theobroma cacao* L.). Eur. Food Res. Technol. 229, 937–948. https://doi.org/10.1007/s00217-009-1132-y

Evans, H.C., 2007. Cacao Diseases—The Trilogy Revisited. Phytopathology 97, 1640–1643. https://doi.org/10.1094/phyto-97-12-1640

Fang, W., Meinhardt, L.W., Mischke, S., Bellato, C.M., Motilal, L., Zhang, D., 2014. Accurate Determination of Genetic Identity for a Single Cacao Bean, Using Molecular Markers with a Nanofluidic System, Ensures Cocoa Authentication. J. Agric. Food Chem. 62, 481–487. https://doi.org/10.1021/jf404402v

FAOSTAT, 2023. Crops and Livestock products, Cocoa beans, area harvested in 2023, Crops/World regions/Cocoa beans/Area harvested from pick lists. [WWW Document]. URL https://www.fao.org/faostat/en/#data/QCL (accessed 1.13.25).

Fernández-Paz, J., Cortés, A.J., Hernández-Varela, C.A., Mejía-de-Tafur, M.S., Rodriguez-Medina, C., Baligar, V.C., 2021. Rootstock-Mediated Genetic Variance in Cadmium Uptake by Juvenile Cacao (*Theobroma cacao* L.) Genotypes, and Its Effect on Growth and Physiology. Front. Plant Sci. 12, 777842. https://doi.org/10.3389/fpls.2021.777842

Fortune Business Insights, 2024. Cocoa and Chocolate Market Size, Share & COVID-19 Impact Analysis, By Type (Cocoa Ingredients (Butter, Liquor, Powder) and Chocolate (Dark, Milk, White, and Filled)), By Application (Food & Beverage, Cosmetics, Pharmaceuticals, and Others), and Regional Forecast, 2022-2029 [WWW Document]. URL https://www.fortunebusinessinsights.com/industry-reports/cocoa-and-chocolate-market-100075 (accessed 10.17.24).

Fouet, O., Solorzano, R.G.L., Rhoné, B., Subía, C., Calderón, D., Fernández, F., Sotomayor, I., Rivallan, R., Colonges, K., Vignes, H., Angamarca, F., Yaguana, B., Costet, P., Argout, X., Lanaud, C., 2022. Collection of native *Theobroma cacao* L. accessions from the Ecuadorian Amazon highlights a hotspot of cocoa diversity. Plants, People, Planet 4, 605–617. https://doi.org/10.1002/ppp3.10282

Galvis, D.A., Jaimes-Suárez, Y.Y., Molina, J.R., Ruiz, R., León-Moreno, C.E., Carvalho, F.E.L., 2023. Unveiling Cacao Rootstock-Genotypes with Potential Use in the Mitigation of Cadmium Bioaccumulation. Plants 12, 2941. https://doi.org/10.3390/plants12162941

Government of the Netherlands, 2024. What requirements must cocoa meet to be allowed on the European market? [WWW Document]. URL https://www.cbi.eu/market-information/cocoa/buyer-requirements (accessed 1.13.25).

Gutierrez, A., 2020. Morphological Characterization of Three Promising Creole Genotypes of *Theobroma cacao* L., in Panama. Ciencia Agropecuaria 30, 150-169 [in Spanish]. http://www.revistacienciaagropecuaria.ac.pa/index.php/ciencia-agropecuaria/article/view/134/98.

Hoerl, A.E., Kennard, R.W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 12, 55–67. https://doi.org/10.1080/00401706.1970.10488634

James, G., Witten, D., Hastie, T., Tibshirani, R., 2021. An Introduction to Statistical Learning, with Applications in R, Springer Texts in Statistics. https://doi.org/10.1007/978-1-0716-1418-1_3

Ji, K., Zhang, D., Motilal, L.A., Boccara, M., Lachenaud, P., Meinhardt, L.W., 2013. Genetic diversity and parentage in farmer varieties of cacao (*Theobroma cacao* L.) from Honduras and Nicaragua as revealed by single nucleotide polymorphism (SNP) markers. Genet. Resour. Crop. Evol. 60, 441–453. https://doi.org/10.1007/s10722-012-9847-1

Júnior, J.O. de S., Silveira, F.G.F. da, Pereira, B.L. da S., Mateus-Rodríguez, J.F., Baligar, V.C., Ahnert, D., 2023. Edaphic and physiographic attributes and death of cacao trees in drought years. Agron. J. 115, 1069–1084. https://doi.org/10.1002/agj2.21215

Kaba, J.S., Asare, A.Y., Andoh, H., Kwashie, G.K.S., Abunyewa, A.A., 2022. Toward Sustainable Cocoa (*Theobroma cacao* L) Production: The Role of Potassium Fertilizer in Cocoa Seedlings Drought Recovery and Survival. Int. J. Fruit Sci. 22, 618–627. https://doi.org/10.1080/15538362.2022.2092932

Lachenaud, P., Motamayor, J.C., 2017. The Criollo cacao tree (*Theobroma cacao* L.): a review. Genet. Resour. Crop Evol. 64, 1807–1820. https://doi.org/10.1007/s10722-017-0563-8

Lanaud, C., Vignes, H., Utge, J., Valette, G., Rhoné, B., Caputi, M.G., Nieto, N.S.A., Fouet, O., Gaikwad, N., Zarrillo, S., Powis, T.G., Cyphers, A., Valdez, F., Nunez, S.Q.O., Speller, C., Blake, M., Valdez, F.Jr., Raymond, S., Rowe, S.M., Duke, G.S., Romano, F.E., Solórzano, R.G.L., Argout, X., 2024. A revisited history of cacao domestication in pre-Columbian times revealed by archaeogenomic approaches. Sci. Rep. 14, 2972. https://doi.org/10.1038/s41598-024-53010-6

Leandro-Muñoz, M.E., Tixier, P., Germon, A., Rakotobe, V., Phillips-Mora, W., Maximova, S., Avelino, J., 2017. Effects of microclimatic variables on the symptoms and signs onset of *Moniliophthora roreri*, causal agent of Moniliophthora pod rot in cacao. PLoS ONE 12, e0184638. https://doi.org/10.1371/journal.pone.0184638

Lindo, A.A., Robinson, D.E., Tennant, P.F., Meinhardt, L.W., Zhang, D., 2018. Molecular Characterization of Cacao (*Theobroma cacao*) Germplasm from Jamaica Using Single Nucleotide Polymorphism (SNP) Markers. Trop. Plant Biol. 11, 93–106. https://doi.org/10.1007/s12042-018-9203-5

Lozano-García, Ma. del S., Caballero, M., Ortega, B., Rodríguez, A., Sosa, S., 2007. Tracing the effects of the Little Ice Age in the tropical lowlands of eastern Mesoamerica. Proc. Natl. Acad. Sci. 104, 16200–16203. https://doi.org/10.1073/pnas.0707896104

Mata-Quirós, A., Arciniegas-Leal, A., Phillips-Mora, W., Meinhardt, L.W., Motilal, L., Mischke, S., Zhang, D., 2018. Assessing hidden parentage and genetic integrity of the "United Fruit Clones" of cacao (*Theobroma cacao*) from Costa Rica using SNP markers. Breed. Sci. 68, 18057. https://doi.org/10.1270/jsbbs.18057

Montoya, T.O., Ramírez, A.M.H., Vasquez, T. de la H., Trujillo, A.I.U., 2022. Propagation of IMC67 Plants, Universal Cacao (*Theobroma cacao* L.) Rootstock via Somatic Embryogenesis. Int. J. Fruit Sci. 22, 78–94. https://doi.org/10.1080/15538362.2021.2023067

Motamayor, J.C., Lachenaud, P., Mota, J.W. da S. e, Loor, R., Kuhn, D.N., Brown, J.S., Schnell, R.J., 2008. Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L). Plos One 3, e3311. https://doi.org/10.1371/journal.pone.0003311

Motamayor, J.C., Lanaud, C., 2002. Molecular Analysis of the Origin and Domestication of *Theobroma cacao* L., in: Engels, J.M.M., Rao, V.R., Brown, A.H.D., Jackson, M.T. (Eds.), Managing Plant Genetic Diversity. Proceedings of an International Conference, Kuala Lumpur, Malaysia, 12-16 June 2000. CABI Publishing, pp. 77–87. https://doi.org/10.1079/9780851995229.0077

Motamayor, J.C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A., Lanaud, C., 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89, 380–386. https://doi.org/10.1038/sj.hdy.6800156

Motilal, L.A., Sreenivasan, T.N., 2012. Revisiting 1727: Crop Failure Leads to the Birth of Trinitario Cacao. J. Crop Improv. 26, 599–626. https://doi.org/10.1080/15427528.2012.663734

N'zi, J.-C., Koné, I., M'bo, K.A.A., Koné, S., Kouamé, C., 2023. Successful grafting elite cocoa clones (*Theobroma cacao* L.) as a function of the age of rootstock. Heliyon 9, e18732. https://doi.org/10.1016/j.heliyon.2023.e18732

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Phillips-Mora, W., Arciniegas-Leal, A., Mata-Quirós, A., Motamayor-Arias, J.C., 2012. Catalogue of cacao clones selected by CATIE for commercial plantings. https://www.cacaonet.org/fileadmin/templates/CacaoNet/Uploads/publications/CatalogueofClones_ENGLISH.pdf

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155, 945–959. https://doi.org/10.1093/genetics/155.2.945

Qin, X.-W., Lai, J.-X., Tan, L.-H., Hao, C.-Y., Li, F.-P., He, S.-Z., Song, Y.-H., 2017. Characterization of volatile compounds in Criollo, Forastero, and Trinitario cocoa seeds (*Theobroma cacao* L.) in China. Int. J. Food Prop. 20, 2261–2275. https://doi.org/10.1080/10942912.2016.1236270

Ribeiro, M.A.Q., Almeida, A.-A.F. de, Alves, T.F.O., Gramacho, K.P., Pirovani, C.P., Valle, R.R., 2016. Rootstock × scion interactions on *Theobroma cacao* resistance to witches' broom: photosynthetic, nutritional and antioxidant metabolism responses. Acta Physiol. Plant. 38, 73. https://doi.org/10.1007/s11738-016-2095-9

Schmidt, J.E., DuVal, A., Puig, A., Tempeleu, A., Crow, T., 2021. Interactive and Dynamic Effects of Rootstock and Rhizobiome on Scion Nutrition in Cacao Seedlings. Front. Agron. 3, 754646. https://doi.org/10.3389/fagro.2021.754646

Setyawan, B., Puspitasari, N., Susilo, A.W., Sari, I.A., 2018. Rootstock Characteristics of Three Combinations of *Theobroma cacao* L. Crosses on Different Water Availability. Pelita Perkeb. (a Coffee Cocoa Res. J.) 34, 137–145. https://doi.org/10.22302/iccri.jur.pelitaperkebunan.v34i3.328

Smouse, P.E., Banks, S.C., Peakall, R., 2017. Converting quadratic entropy to diversity: Both animals and alleles are diverse, but some are more diverse than others. PLoS ONE 12, e0185499. https://doi.org/10.1371/journal.pone.0185499

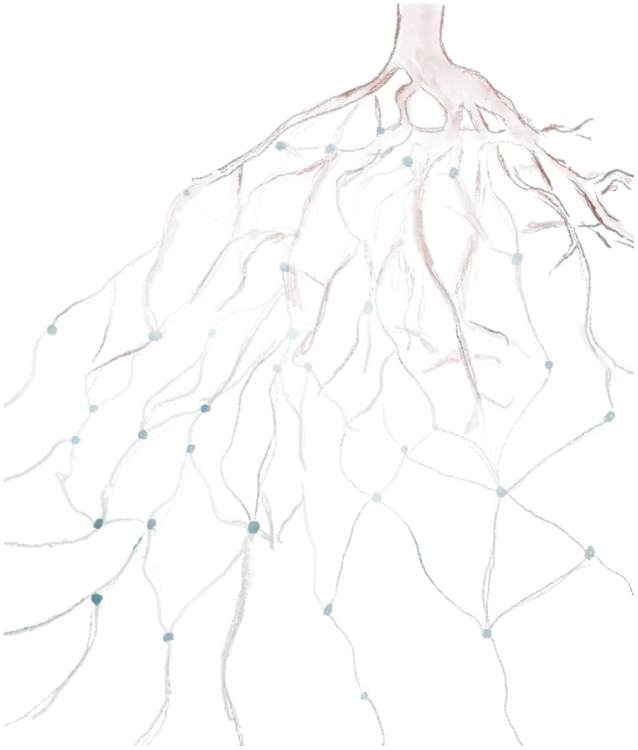
Villarreal, J., Name, B., Barahona, L.A., 2015. Tecnicas analiticas del laboratorio de suelos. Instituto de Investigación Agropecuaria de Panamá. Panamá.

World Cocoa Foundation, 2025. Home [WWW Document]. URL https://worldcocoafoundation.org/ (accessed 1.20.25).

Yacomelo-Hernández, M.J., Córdoba-Gaona, O. de J., Redondo-Herrera, V.M., Hernández-Arredondo, J.D., Arenas-Rubio, I., Pérez-Artiles, L., 2021. Renovation and rehabilitation practices of unproductive cacao plantations: Strategies to maintain production in Colombia. Rev. Fitotec. Mex. 44, 625. https://doi.org/10.35196/rfm.2021.4.625

Chapter 3. Evaluating the relative importance of cacao genetic identity, root traits, and soil properties for arbuscular mycorrhizal fungal community composition

Natalie Ferro-Lozano, Indira Quintero, Luis Mejia, Guusje Bonnema, Thomas W. Kuyper



Abstract

Arbuscular mycorrhizal fungal (AMF) communities in a plant species are shaped by plant genetics, root traits, and environmental variables such as elevation, soil properties, and plant community. Despite AMF's known benefits, their agricultural potential remains underexplored. Cacao (*Theobroma cacao*) is a mycotrophic agroforestry crop. Separately harnessing rootstock and scion genetic variation associated with AMF community composition could contribute to more sustainable cacao agroforestry systems. This study aimed to: (1) characterize AMF communities associated within cacao roots, and (2) evaluate how these communities are shaped by scion and rootstock genotypes, root traits, soil properties, and presence of conspecific or heterospecific neighbouring trees. We sampled roots and soils from 81 trees in 16 Panamanian farms in the provinces of Bocas del Toro and Herrera. We analysed scion and rootstock genotypes, root traits (the four traits of the root economics space, plus root phosphorus concentration), AMF communities, and soil characteristics. Our results show that *Glomus* was the dominant genus across all samples. Scion genotype, root nutrient concentration (N and P), elevation, and available P emerged as key variables structuring AMF communities, while rootstock genotype and root morphological traits did not. There was high variation in species composition between individual farms. These findings highlight the role of host and environmental variables in AMF composition, informing strategies for sustainable cacao cultivation through optimized plant-fungal interactions.

Introduction

Arbuscular mycorrhizal fungi (AMF) are ancient plant symbionts that enabled the colonization of terrestrial ecosystems by enhancing plant nutrient uptake and priming immune defences. Despite their ecological importance, AMF remain understudied—particularly in tropical regions—since most research has focused on temperate ecosystems (Alexander and Selosse, 2009; Kuyper et al., 2004). As a result, key aspects of this symbiosis remain unresolved, including the diversity of AMF species within a single plant species, the regulation in the interaction between the plant and fungal community, and how environmental variables shape AMF communities. Knowledge about these interactions is however key, as optimized AMF–plant relationships offer promising solutions for sustainable agriculture in low-input and organic systems, typical in tropical ecosystems. Agricultural success in the tropics often mirrors tropical forest structure by focusing on nutritionally high-value crops, maintaining local plant diversity, and integrating perennials to sustain soil fertility and optimize resources (Altieri, 2004). This approach positions agroforestry systems as the most effective agricultural strategy in the tropics.

Cacao (Theobroma cacao) agroforestry systems (CAFS) are among the most widespread tropical perennial cropping systems, with approximately 6,000 hectares cultivated by around 1,400 smallholder families in Panama alone. Traditionally, farmers rely on the use of biological and structural diversity over time and space, resulting in low-input practices with favourable energy *efficiency* (Altieri, 2004). Farmers integrate a variety of local livestock and plant species, including secondary tree species that provide timber, fruits, and key ecosystem functions such as shade, soil fertility, and erosion control. Other agroforestry practices such as grafting—combining rootstock and scion genotypes—are also employed to preserve genetic diversity and enhance plant development and productivity. However, these traditional practices have often been disrupted in recent decades, as smallholders face increasing pressure to maximize short-term yields under conditions of low cocoa prices. This has led to a growing reliance on costly external inputs, such as synthetic fertilizers and pesticides, resulting in nutrient imbalances, pest and disease outbreaks, and declining yields (Cardoso and Kuyper, 2006). Addressing these challenges requires renewed emphasis on increasing plant genetic diversity within and among species, and improving soil management to support sustainable AMF symbiosis, enhancing ecosystem services.

AMF communities play a vital role in tropical agroforestry systems supporting productivity and ecosystem functioning (Cardoso and Kuyper, 2006). Locally, AMF expand the volume of soil exploited by the plant, while alleviating toxicity by heavy metals. AMF also interact with beneficial microbes—such as phosphate-solubilizing, nitrogen-fixing, and growth-promoting bacteria—to enhance crop resilience and pathogen resistance (L. Wang et al., 2024). Globally, AMF shape plant

community structure and composition (Schroeder et al., 2020) and contribute to carbon sequestration (Hawkins et al., 2023; Rillig et al., 2001).

Additionally, composition of AMF communities is shaped by both biotic and abiotic factors, including plant genetics, root traits, soil conditions, and surrounding vegetation. In tropical ecosystems, key drivers consist of plant community dynamics, soil properties as pH and available P, and environmental factors, including precipitation, temperature and elevation (Broeckhoven et al., 2025; Jemo et al., 2018; Mai et al., 2023; Pereira et al., 2020a; Rodríguez-Echeverría et al., 2017; Sousa et al., 2018, 2022; Thanni et al., 2022; Veras et al., 2024, 2025; Yang et al., 2025; Zhang et al., 2021). Although less frequently studied, significant plant effects on richness of AMF communities and composition have been associated with root traits such as specific root length (SRL), root diameter, root phosphorus and calcium content (Deveautour et al., 2018; Ramana et al., 2023; Sweeney et al., 2021; Wan et al., 2021). Since it is unknown to what extent plant genetics, root traits or plant community dynamics drives AMF community composition, distinguishing between scion and rootstock genetic identities is therefore essential. If AMF composition is driven by the plant's carbon supply or nutrient demand, scion identity may be more influential than rootstock identity. Conversely, if root morphological traits and their interactions with the rhizosphere dominate (Bergmann et al., 2020), rootstock identity is expected to correlate more strongly with AMF community structure than scion identity.

This study aims to characterize AMF communities within cacao roots in Panamanian agroforestry systems and to assess how scion and rootstock genotypes, root traits, and environmental variables influence their composition. To address this, we sampled cacao roots from five trees in each of 16 plantations across two regions in Panama. We SNP-profiled above- and belowground plant material and measured soil and root properties. Specifically, we asked: (1) What is the diversity of AMF across farms and regions, and are there AMF taxa commonly associated with cacao*?* (2) What is the relative importance of plant genetics, root morphology, and environmental variables for AMF community structure? (3) Which virtual taxa (VT), a molecular proxy for AMF species, are associated with specific environmental variables?

Materials and Methods

Study sites and design

We studied AMF communities in two cacao-growing regions of Panama, viz., Bocas del Toro and Herrera. Within these regions, sixteen plantations were selected. These plantations represent a wide range of cacao-growing conditions in terms of elevation, coastal proximity, temperature, and precipitation. These plantations grow mostly cacao landraces and selections from the wild (Chapter 2), with a few modern cultivars as CATIE R4, R1, R6, PMCT58, ICS95, and CC137. Five cacao trees

were assessed from each of the farms; in one farm six trees were assessed. A total of 81 trees were selected: 25 (5x5) in Las Minas, Herrera and 56 ((5 x 10) + 6) in five different regions of Bocas del Toro (Supplementary material A).

Sample Collection

Of each cacao tree, we sampled leaf and root tissue separately for assessment of genetic identity of the above- and below-ground tissues of the tree. For trees raised from seeds, root and leaf DNA profiles should be identical, however for grafted trees, DNA profiles between root (representing the rootstock) and leaf (representing the scion) should be different. Leaf tissue was placed in a box with silica gel for immediate drying and subsequent plant DNA analysis. Root tissue was divided into three subsamples. The first subsample was dried for determination of mycorrhizal colonization, root nitrogen and phosphorus quantification, and plant root DNA analysis. Plant DNA extraction and subsequent genotyping was done in the USDA facilities in Maryland, USA. A second root subsample was wrapped in a wet paper towel and stored at 5°C until further analysis for image scanning and root trait assessment. The third root subsample was stored in 95% ethanol for mycorrhizal fungal community analysis.

Plant SNP markers and Genotyping

The methodology used for 96-SNP genotyping was as described by Lindo et al. (2018). SNP markers were identified comparing expressed sequence tags (ESTs), generated from various cacao organs and environments, between different genotypes (Argout et al., 2011, 2008). The 96 SNPs, nicely distributed across the 10 cacao chromosomes, used for this panel were selected from this dataset for their ability to distinguish cacao genetic groups, as previously reported by (Fang et al., 2014; Ji et al., 2013). The 96 SNPs, the flanking sequences and their distribution across the ten cacao chromosomes are listed in Chapter 2. These SNPs were analysed in a PCA to use the principal components as independent genetic variables to explain the association with AMF community composition (Supplementary material B). Hereafter, we identified the individual SNPs contributing significantly to the principal components (PC contribution > above 0.5).

Root Trait Assessment

Dried fine roots of each tree were cut into 1.5 cm root fragments, placed in tissue cassettes, and hydrated overnight in distilled water before staining. The staining procedure used sequential solutions of (1) KOH 10% for 16 hrs at 45°C, (2) H_2O_2 3% for 30 min at room temperature, (3) HCl 1M for 15 min at room temperature, and lastly (4) 0.05% trypan blue in lactoglycerol for 45 min at 45°C. In between the solutions, the samples were washed thrice with distilled water. All samples were placed in lactoglycerol for 3-4 days at room temperature to eliminate the excess of trypan blue. After this,

we analysed roots in microscope slides for mycorrhizal colonization (MC) through the Trouvelot method (Trouvelot et al., 1986). We assessed the percentage of root length colonized by AMF (M%).

Fresh fine roots, defined as roots of less than 1 mm wide and representing the three lower root orders, were thoroughly washed to remove all soil. Around three root fragments per tree were submerged in a shallow layer of water in a transparent tray (120 mm by 120 mm by 17 mm) and carefully separated with wooden chopsticks so they did not touch. These roots were then scanned under high resolution (Epson Perfection V850 Pro). After scanning, fine roots were dried for three days at 60°C and subsequently weighed to determine root dry weight of scanned roots (sRDW). Root scans were analysed using WinRhizo software (Regent Instruments, 2000) to determine average root diameter (D), specific root length (SRL), and root tissue density (RTD). Additionally, root N content and root P content were analysed through the Kjeldahl method and colorimetrically, using the molybdenum blue method.

Fungal DNA Extraction and Sequencing

Roots stored in ethanol 95% were thoroughly washed, lyophilized for 7 days, and homogenized with the aid of liquid nitrogen. Later, 0.2 g of the ground material was used for DNA extraction using a Nucleospin Plant II kit (Macherey-Nagel, Duren, Germany) according to the manufacturer's instructions. The quantity and quality of the DNA extracts were assessed using NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Dreieich, Germany). To decrease variations in the PCR process, genomic DNA was amplified in triplicate (Polz & Cavanaugh, 1998, Luo 2020) using nested polymerase chain reaction (PCR) and then pooled for Illumina sequencing. The following protocol was adapted from (Wahdan et al., 2021). The first reaction was performed with GlomerWT0 (5′-CGAGDWTCATTCAAATTTCTGCCC-3′) (Wubet et al., 2006) and Glomer1536 (5′-AATARTTGCAATGCTCTATCCCCA-3′) (Morgan and Egerton-Warburton, 2017; Wubet et al., 2006) primer pair. The second reaction [amplification of a 560 bp fragment of the small subunit (SSU/18S) of the rDNA gene region] was performed using NS31 (5′-TTGGAGGGCAAGTCTGGTGCC-3′) (Morgan and Egerton-Warburton, 2017) attached with P5 Illumina adapter and AML2 (5′-GAACCCAAACACTTTGGTTTCC-3′) primers (Lee et al., 2008) attached with P7 Illumina adapter. All PCRs were conducted using the proofreading Kapa HiFi polymerase (2X KAPA HiFi HotStart ReadyMix, Kapa Biosystems, Boston, MA, United States). The PCR product was purified using Agencourt AMPure® XP beads (Beckman Coulter Inc., Indianapolis, IN, USA). Indexing of the purified amplicons was done using the Nextera index kit (A and B set) (Illumina, San Diego, CA, USA). Finally, amplicon libraries were quantified by Qubit 1X dsDNA HS Assay Kit (Invitrogen) and pooled to give equimolar representation of each. Illumina MiSeq sequencing was performed with 2 x 250bp cartridges at the Molecular Fungal Lab (Genomic Lab),

Institute of Scientific Research and High Technology Services of Panama (INDICASAT-AIP for its Spanish acronym) in Panama City, Panama.

Bioinformatics workflow

We used the DADA2 package (Callahan et al., 2016) in R software version 4.5.0 (R-Development-Core-Team, 2025) for sequence processing. Sequences corresponding to the forward and reverse primers were trimmed from the demultiplexed raw reads, using cutadapt (Martin, 2011), to produce 12,349,470 reads from 80 samples. One sample did not include enough material to go through the triplicate process and was discarded. Sequences were quality-trimmed and filtered with default parameters by the dadasnake pipeline (Weißbecker et al., 2020) and a quality threshold of 20. Sequencing reads with quality scores less than 20 (99% confidence interval) were truncated after two consecutive base calls. Truncated reads that were less than 75% of their original length were removed from further analyses. Consequently, as input reads were 250 bp, the resulting data set contained reads with a minimum 187 bp. The average sequence length between the primers was larger (~540 bp) than twice the amplicon length with the Illumina MiSeq, 2 × 250-bp sequencing technology, and thus did not allow for overlap and could not be assembled (Supplementary Material C). For that reason, we conducted all downstream analyses using only the 250-bp forward reads. This choice for using forward reads was based on Davison et al. (2012) where truncating NS31/AML2 reads from 400 to 170 bp resulted in a near identical capacity to capture AMF diversity compared with the complete 400 bp fragment. This is because most taxonomically informative characters occur in the 5′-most 170 bases.

After having filtered, denoised and removed chimeras, we obtained 1,363,906 high-quality reads clustered into 713 amplicon sequence variants (ASVs). The taxonomic identification of each ASV was performed by aligning it against the NCBI GenBlast to sift out all sequences belonging to plant and other non-mycorrhizal fungal DNA, then to AMF virtual taxa (VT) from the MaarjAM database (Öpik et al., 2010) using BLAST (Altschul et al., 1990) at 97% similarity. ASVs with BLAST and MaarjAM matches below 97%, were treated as potentially new VT. These ASVs were subjected to phylogenetic analysis to enable highly similar ASVs to be grouped as one novel VT. Virtual taxa are considered as species hypotheses, for easiness in this manuscript, we will treat VT as species. A total of 645,609 of sequences belonging to Glomeromycota (the clade constituting the AMF) were retrieved from all samples with a maximum of 47,711 and a minimum of 46 reads per sample. For normalization, the dataset was rarefied to the minimum number of 1,000 reads per sample as done in Wahdan (2021), resulting in 76 samples with 683 ASVs. Four samples with low sequencing depth (less than 1000 reads) were discarded. As analyses were done per tree, this exclusion did not impact site-level representation. The rarefaction curve (Supplementary Material D) indicated that, although

high-quality sequencing was obtained, sequencing depth was insufficient to fully capture the biodiversity present.

Environmental Data Collection

Measured plantation characteristics included elevation, average tree age, farm type, cacao tree density, shade percentage, and taxonomic identity of the three neighbouring trees per sampled tree, split in neighbouring *Theobroma* and other species. Soil analysis below each tree included texture (% sand, % silt, % clay), pH (in water), organic matter (%), nitrogen (N; %), available phosphorus (P; mg/g), C:N (carbon to nitrogen ratio), potassium (K; mg/l), and exchangeable cations (Ca, Mg, Al, Mn, Fe, Cu; Cmol/kg or mg/l), alongside CEC (cation exchange capacity), Al saturation, and base saturation (Villarreal et al., 2015). To identify key plantation and soil variables, as in Chapter 2, we applied Horn's parallel analysis, PCA cos^2 values, and correlation matrices, retaining significant axes and variables while removing collinearity. From the 18 measured soil variables and 6 farm traits, 13 environmental variables remained for further analysis: Elevation, percentage of sand (% sand), percentage of silt (% silt), percentage of clay (% clay), pH, organic matter percentage (OM%), total nitrogen (TN %), available P (mg/g), soil magnesium (Cmol/kg), soil aluminium (Cmol/kg), soil calcium (Cmol/kg), soil iron (mg/l), and number of neighbouring *Theobroma*.

Statistical Analysis and Software

All statistical analyses were performed using RStudio Version 2024.04.2+764 and R ver. 4.4.1 (www.r-project.org). We created a 'phyloseq' object with the mycorrhizal fungal community composition and the metadata associated to each of the samples. The phyloseq object was normalized to relative abundance for taxonomic visualization and rarefied for determination of alpha and beta diversity using the R package phyloseq (McMurdie and Holmes, 2013). Alpha diversity was assessed using Observed VTs, Chao1, and Shannon index. Observed index accounts for number of species, Chao1 estimates the number of species considering that rare species may have been missed, and Shannon index accounts for both species' richness and evenness (Colwell, 2009). Beta diversity was visualized through Principal Coordinate Analysis (PCoA) and tested for homogeneity of dispersion to validate the permutational analysis of variance (PERMANOVA). This was done to test for significant differences between mycorrhizal fungal communities from regions and farms using the 'adonis' function from the vegan package with 999 permutations. All plots were created with the R package ggplot2 (Wickham, 2016). PCA was performed with the R package FactoMineR (Lê et al., 2008) and their visualizations with the R package ggbiplot (Vu et al., 2024).

Distance-based redundancy analysis (db-RDA) was used to evaluate whether variation in the fungal community structure could be significantly explained by plant germplasm, root traits, and/or environmental variables. To account for the non-Euclidean nature and non-normality of several soil

data, dbRDA was performed on a Bray-Curtis dissimilarity matrix using the *capscale* function from the 'vegan' R package (Legendre and Anderson, 1999; Oksanen et al., 2022). Separate dbRDA models were constructed to examine the influence of different groups of predictor variables: (i) PCA loadings from above and belowground plant SNP genotyping (scion_PC1, scion_PC2, scion_PC3, rootstock_PC1, rootstock_PC2, rootstock_PC3); (ii) five root functional traits (root diameter, root tissue density, root N, root P, and mycorrhizal colonization (%); and (iii) thirteen soil environmental variables previously mentioned; and (iv) a full RDA model that included all genetic, root, and environmental variables together to test which variables stood out as key factors in the mycorrhizal fungal community composition. For the full model, scion_PC1 and rootstock_PC1 were excluded based on variance inflation factors (VIF) calculated using the *vif.cca* function in 'vegan'. Prior to analysis, all predictor variables were evaluated for skewness. Log transformations were applied to organic matter, soil magnesium, soil aluminium, and root phosphorus, while available phosphorus and root tissue density were square root transformed. All variables were subsequently standardized via z-transformation. The significance of each model, its constrained axes, and individual predictors was tested via PERMANOVA using the *anova.cca* function with 9,999 permutations.

We used indicator species analysis (ISA) to identify VT from the mycorrhizal fungal communities exhibiting strong associations with the relevant variables identified in the RDAs. The default ISA settings were used, based on relative abundance. The variables were grouped into three groups, the lowest, middle, and upper tertile (Table 1). This analysis was done using the 'indicspecies' R package, with 9,999 permutations (Cáceres and Legendre, 2009).

Table 1. Group tertile division for each variable used in the Indicator Species Analysis.

	PC2 Scion	Elevation	Available P	RootN	RootP	Mycorrhizal Colonization
Low	[-1.996, -0.507]	[6, 51]	[1.00E-06, 0.01]	[6.49, 9.15]	[0.53, 1.01]	[8.03, 24.13]
Med	[-0.507, 0.376]	[51, 338]	[0.01, 0.04]	[9.15, 11.77]	[1.01, 1.72]	[24.13, 32.2]
High	[0.376, 1.677]	[338, 600]	[0.04, 0.59]	[11.77, 16.32]	[1.72, 4.33]	[32.2, 55.17]

Results

From 76 rarefied samples, we identified 683 ASVs that belonged to 82 virtual taxa (VTs) of AMF (Supplementary material E). A total of 95% of the ASVs were assigned to the genus *Glomus* (69 VTs), while the remaining 4% were distributed among *Acaulospora (5 VTs)*, *Paraglomus (3 VTs), Claroideoglomus (2 VTs), Diversispora (2 VTs),* and *Redeckera (1 VT)*. Additionally, 54 ASVs (approximately 1%) could not be taxonomically classified due to sequence similarity below 97% compared to the MaarjAM reference database. These unidentified taxa were present in 73 of the 76 samples, and compared between each other at 97% similarity, these suggest at least 8 new VT (Supplementary material F).

Diversity across samples

We identified 72 VT in Bocas del Toro and 55 VT in Herrera. Of these, 47 VT were shared between both regions, while Bocas del Toro harboured 25 - and Herrera 8 unique VT respectively (see Supplementary Material G). Although PERMANOVA indicated significant differences in community composition between Bocas del Toro and Herrera, unequal within-region variation suggests this result may reflect dispersion rather than true compositional differences. Inspection of the VT list indicated that Herrera formed a subset of diversity from Bocas del Toro, as shown in the Bray-Curtis diagram in Figure 1. The average AMF VT richness for Bocas del Toro (farms 1 – 11) was 11.2 VT and for Herrera (farms 12 - 16) was 12.9, but the difference was not significant. Low VT richness was observed in farm #9 (average 3 VT; other farms between 8 and 16 VT).

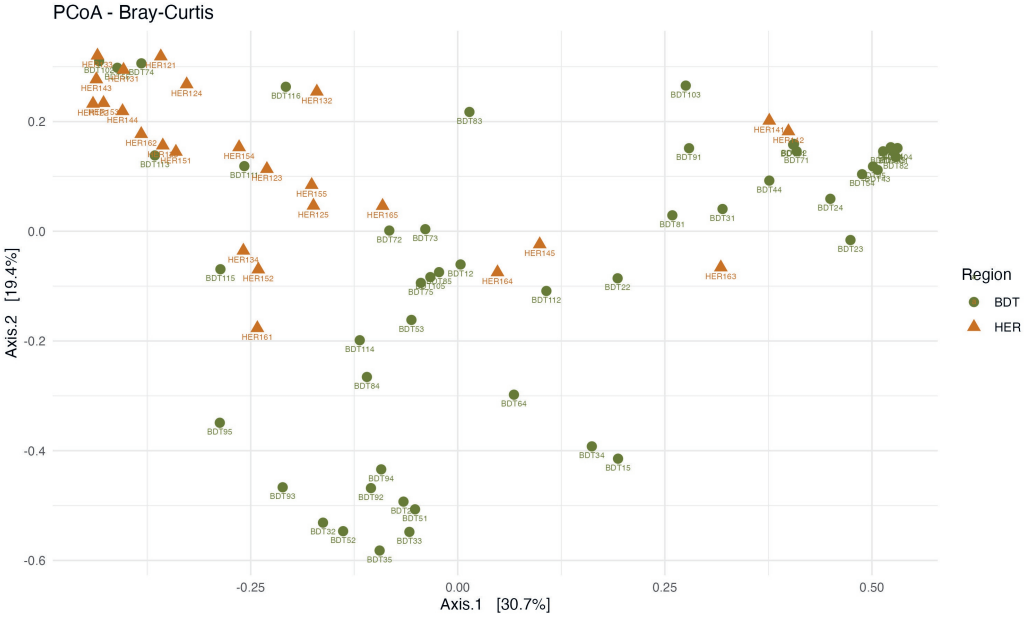


Figure 1. PCoA plot of AMF Virtual Taxa composition in Bocas del Toro (circles) and Herrera (triangles). Unequal within-region variation shows dispersion rather than true compositional differences drive differences between regions. No statistical differences in composition were found between both regions.

Significant differences in diversity between the farms did exist (Figure 2). In Bocas del Toro, farm 10 (and to a lesser extent farm 1 and 11) held the lowest observed VT richness with also the lowest number of rare VT. In general, all farms retained a high number of rare species, as Chao1 index indicated (Figure 2). Farms 15 and 16 from Herrera had the most evenly distributed VT.

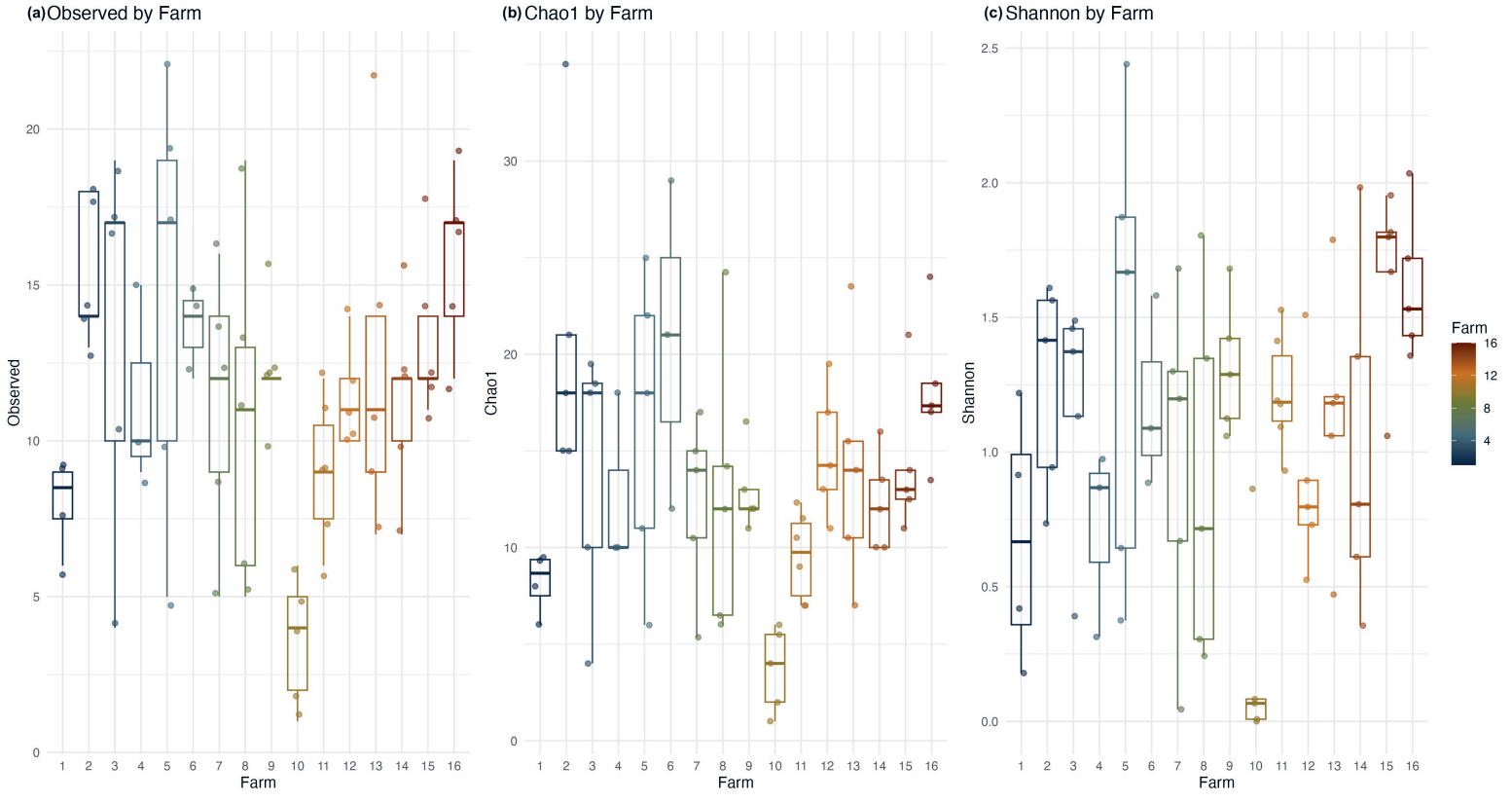


Figure 2. Alpha diversity values for each farm. The alpha indexes shown are (a) Observed number of VT, (b) Chao1, and (c) Shannon diversity index.

Core Mycobiome

The core mycobiome, defined as the VT present in over 75% of the samples regardless of their abundance, consisted of six VT, all belonging to the *Glomus* genus (Table 2). VTX00092 and VTX00069 were most abundant, representing 26% and 22% of total read numbers. Four other VT of the core microbiome were present in a lower abundance (11% or less).

Table 2. Core mycobiome with their relative abundance derived from total reads with their pertaining genus.

Genus	Species	Relative Abundance	Frequency in Samples
Glomus	VTX00092	26%	100%
Glomus	VTX00069	22%	98%
Glomus	VTX00397	11%	98%
Glomus	VTX00270	8%	92%
Glomus	VTX00113	5%	80%
Glomus	VTX00122	4%	86%

Variables influencing AMF community composition

Principal component analysis summarized cacao germplasm variance based on profiling of 96 biallelic SNP loci across scions and rootstocks (Supplementary material B): scion_PC1 (22.7%), scion_PC2 (9.5%), scion_PC3 (5.8%), rootstock_PC1 (19%), rootstock_PC2 (10%), and rootstock_PC3 (5.5%). In a different PCA aiming to understand how genetic, root, and environmental variables relate to each other, without considering AMF community composition, many variables were correlated (Figure 3 - Supplementary material H/I). The genetic variables scion_PC1 and rootstock_PC1 are highly correlated, indicating parallel SNP variation for the SNPs loaded on those axes and reflecting the large number of trees raised from seeds, where the above and belowground genotype is identical. Contrary, scion_PC2 and rootstock_PC2 were inversely correlated, indicating SNPs that differ between scion and rootstock, while scion_PC3 and rootstock_PC3 were not correlated. For environmental variables, elevation and soil magnesium were negatively correlated, indicating in farms at higher elevation soils are lower in Mg. The pairs root P and available P, and root N and total nitrogen, were highly correlated. Lastly, SRL was negatively correlated with RTD, as well as pH and Aluminium, following their mathematical (in case root diameter is highly invariant) or chemical relation respectively.

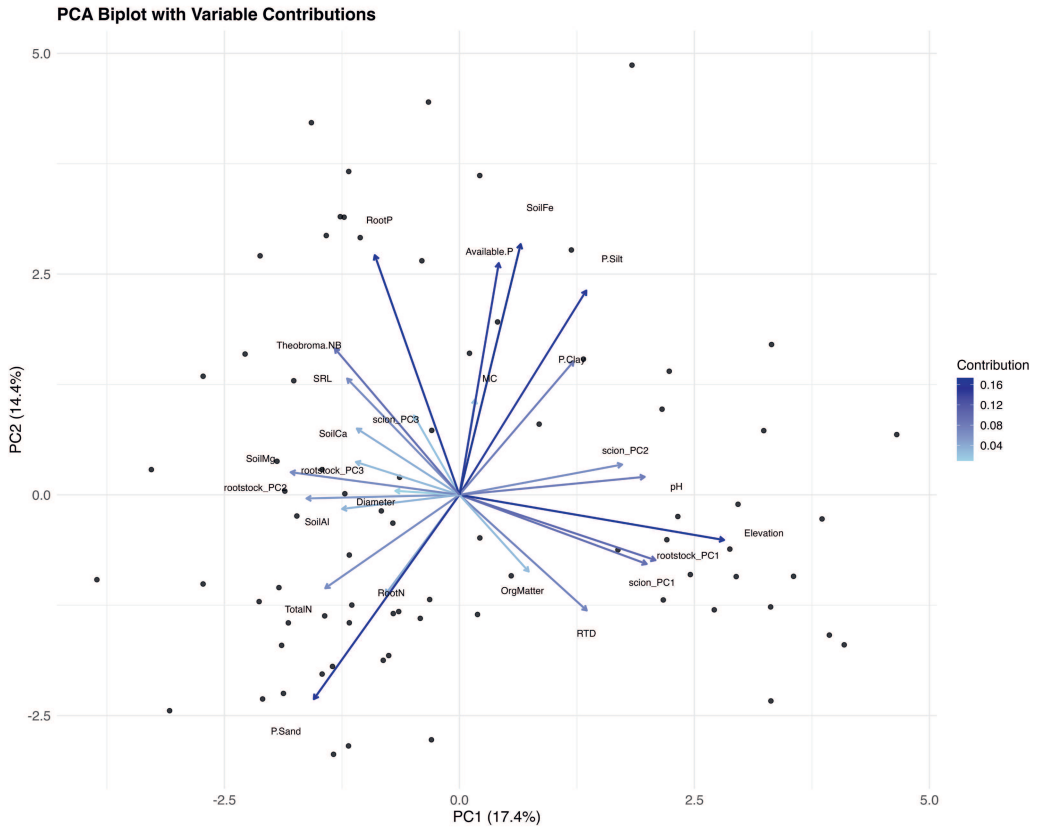


Figure 3. PCA between all influencing variables, for each arrow, colours indicate contribution, and direction indicates correlation between them. The PCA loadings for each axis are in Supplementary material H. Acronyms: SRL - specific root length, RTD – root tissue density, MC – mycorrhizal colonization percentage, P. Silt – soil percentage of silt, P. Clay – soil percentage of clay, and P. Sand – soil percentage of sand.

For the redundancy analysis (Table 3 – Supplementary material J), aiming to evaluate whether variation in fungal community structure could be significantly explained by plant germplasm, root traits, and/or environmental variables, we performed three separate models: for genetic variables ($p = 0.0175$), root traits ($p = 0.0015$), and soil properties ($p = 0.0001$). For genetic variables, the scion PC2 was the only variable with a significant contribution to explain the variance of AMF community composition. We identified 63 alleles contributing the most to the Scion_PC2 variation (Supplementary material K). For root traits, root nitrogen, root phosphorus, and mycorrhizal colonization were significantly associated with the AMF community composition. For environmental properties, elevation and available P showed significant effects on the AMF community. In the full model with all variables (genetic, root morphological and physiological, and environmental properties - $p = 0.0001$), root phosphorus and available P stood out as the main factors associated with AMF community composition.

Table 3. Redundancy Analysis Models to identify main drives for mycorrhizal community. In the left columns are the individual RDA models considering either genetic, root or soil variables. At the right column is the full RDA model considering all variables simultaneously. Significant variables and models are denoted by bold letters.

Model & Traits	Variance	*F* value (*P* value)	Model & Traits	Variance	*F* value (*P* value)
Genetic (p = 0.0175)			**All together (p = 0.0001)**		
Scion_PC1	0.267	0.766 (0.635)	scion_PC2	0.006	1.543 (0.121)
Scion_PC2	**0.682**	**1.959 (0.045)**	scion_PC3	0.010	0.454 (0.940)
Scion_PC3	0.145	0.416 (0.960)	rootstock_PC2	0.004	1.010 (0.384)
Rootstock_PC1	0.157	0.452 (0.933)	rootstock_PC3	0.006	0.990 (0.410)
Rootstock_PC2	0.186	0.535 (0.871)	RootN	0.007	0.685 (0.722)
Rootstock_PC3	0.493	1.416 (0.163)	**RootP**	**0.007**	**2.212 (0.029)**
Residual	24.008		Diameter	0.000	0.527 (0.884)
Roots (p = 0.0015)			RTD	0.000	0.692 (0.718)
Root N	**1.012**	**2.958 (0.005)**	MC	0.000	0.650 (0.764)
Root P	**1.017**	**2.971 (0.004)**	Elevation	0.009	1.025 (0.383)
Diameter	0.392	1.144 (0.301)	P. Sand	0.007	
RTD	0.460	1.343 (0.184)	P. Silt	0.011	
MC	**0.985**	**2.880 (0.007)**	P. Clay	0.008	
Residual	23.948		pH	0.012	1.128 (0.311)
Soil (p = 0.0001)			Org Matter	0.009	0.975 (0.426)
Elevation	**0.000**	**2.107 (0.036)**	Total N	0.000	1.101 (0.329)
P. Sand	0.000		**Available P**	**0.000**	**2.815 (0.007)**
P. Silt	0.000		Soil Mg	0.000	1.381 (0.179)
P. Clay	0.016		SoilAl	0.000	1.789 (0.071)
pH	0.009	1.305 (0.214)	SoilCa	0.007	0.594 (0.819)
OrgMatter	0.007	1.093 (0.331)	SoilFe	0.021	1.451 (0.153)
TotalN	0.007	1.044 (0.371)	Theobroma.NB	0.012	1.856 (0.066)
Available P	**0.010**	**2.319 (0.021)**	Residual	0.010	
SoilMg	0.011	1.519 (0.128)			
SoilAl	0.000	1.880 (0.059)			
SoilCa	0.000	0.818 (0.581)			
SoilFe	0.000	1.291 (0.218)			
Theobroma.NB	0.000	1.331 (0.195)			
Residual	0.478				

Indicator Species Analysis

We performed an Indicator Species Analysis (ISA) to identify taxa associated with significant variables previously identified in the RDA, including Scion genotype, Root Nitrogen (N), Root Phosphorus (P), Mycorrhizal colonization (%M), Elevation, and Available Phosphorus (P) (Figure 4 – Supplementary material L). In total, 22 unique VT (VTX00113, ASV 388, VTX00115, VTX00069, VTX00248, VTX00186, VTX00280, VTX00093, VTX00247, VTX00096, VTX00185, VTX00412, VTX00215, VTX00399, VTX00175, VTX00103, VTX00342, VTX00126, VTX00209, VTX00111,

VTX00120, VTX00364) were significantly associated with at least one variable (IndVal > 0.3, $p \leq 0.05$). For scion_PC2 variable, six VT (VTX00113, VTX00069, VTX00280, VTX00215, VTX00175, VTX00342) were significantly associated, in addition to a new VT (ASV 388). The relative abundance ranged between 22% and 0.02% and frequency ranged between 99% - 12% for these scion-specific taxa (Figure 5 - Supplementary material M). Elevation, root phosphorus, root nitrogen, and scion variables were associated with the largest number of AMF species showing significant indicator values, indicating their strong association with AMF community composition.

Several species showed strong associations with multiple variables. For instance, elevation and scion shared three species. VTX00113 showed high indicator values for elevation (IndVal = 0.84, p = 0.000) and intermediate values for scion (IndVal = 0.63, p = 0.012); ASV 388 showed high indicator values for elevation (IndVal = 0.71, p = 0.007) and intermediate values for scion (IndVal = 0.55, p = 0.047); and VTX00175 showed low indicator values for both elevation (IndVal = 0.47, p = 0.004) and scion (IndVal = 0.45, p = 0.011). A higher indicator value to elevation and only moderate association with scion genotype suggests that these AMF species are generalists across these plant genotypes but are ecologically more constrained or selected by elevation or factors for which elevation is a proxy (e.g., geological history). Furthermore, root nitrogen and elevation shared three species. VTX00115 was significantly associated with high indicator values for elevation (IndVal = 0.71, p = 0.002), and root N (IndVal = 0.69, p = 0.004); VTX00186 was significantly associated with intermediate indicator values for root N (IndVal = 0.54, p = 0.000) and elevation (IndVal = 0.44, p = 0.037); and VTX00399 was significantly associated with lower indicator values for elevation (IndVal = 0.47, p = 0.04), and root N (IndVal = 0.48, p = 0.03). From these, VTX00115 was also significantly associated with high indicator values for root P (IndVal = 0.65, p = 0.04) and VTX00186 with intermediate indicator values for available P (IndVal = 0.6, p = 0.000), and root P (IndVal = 0.48, p = 0.007). Another species with significant associations between root N (IndVal = 0.45, p = 0.012), root P (IndVal = 0.46, p = 0.012) and available P (IndVal = 0.53, p = 0.000) was VTX00412.

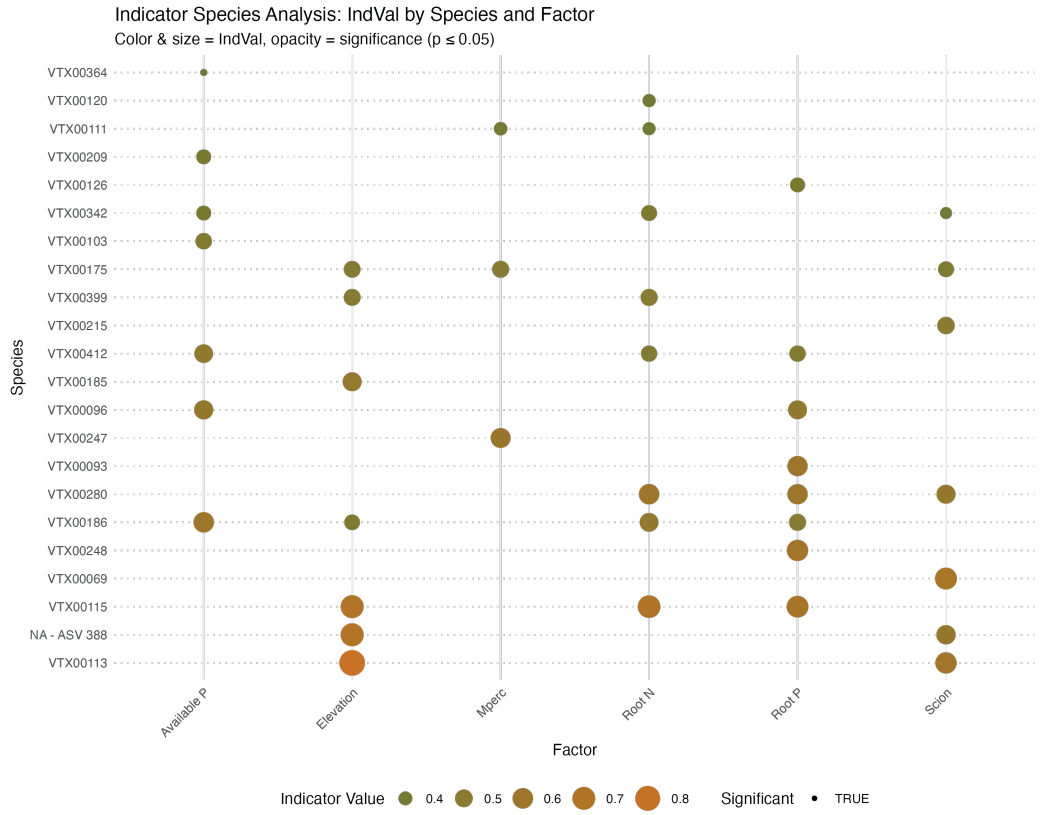


Figure 4. Dot plot showing indicator virtual taxa (VT) significantly associated with soil, cacao genetics, and environment variables identified in the redundancy analysis (Supplementary material L). Each dot represents the Indicator Value for a VT–variable combination. Larger, darker dots indicate higher Indicator Values and lower associated p-values. Dots are approximately ordered by increasing Indicator Value, with higher values appearing toward the bottom of the plot. Mperc is fractional mycorrhizal colonization, scion refers to PC2 of scion diversity based on 96 profiled SNPs, P is phosphorus and N is nitrogen.

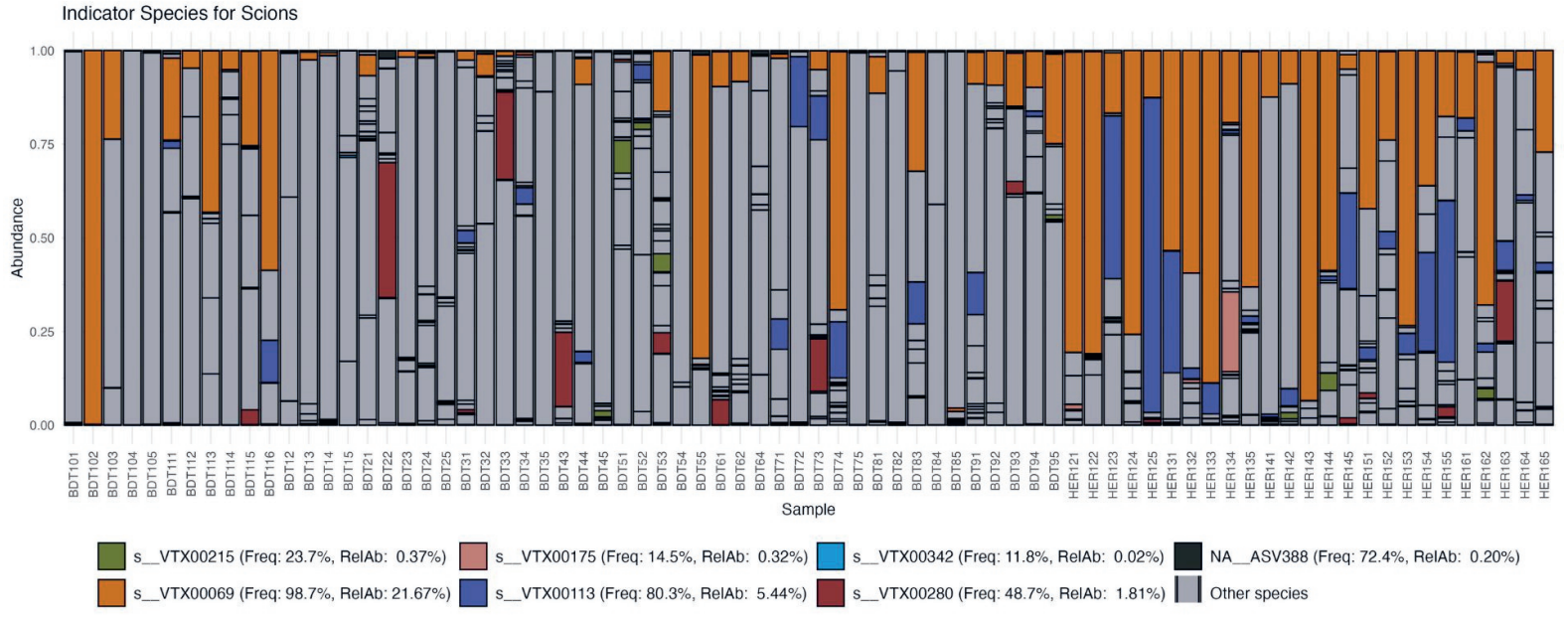


Figure 5. Indicator VT linked to scion_PC2 (Supplementary material M). Legend shows for each species their frequency across samples (Freq), and relative abundance (RelAb).

Discussion

We found a significant association between root nitrogen (N), root phosphorus (P), and scion germplasm with AMF community composition. The scion and rootstock represent distinct plant compartments, with the scion governing aboveground functions such as leaf development, photosynthesis, and fruit production, while the rootstock regulates belowground processes including nutrient and water uptake, root architecture, and rhizosphere conditions that facilitate symbiotic interactions. Both compartments influence systemic signalling, with root N and P reflecting both the scion's metabolic demands and the rootstock's capacity to interact with mycorrhizal partners. Nitrogen and phosphorus are key elements transferred between the photosynthetic and mycorrhizal partners (Kuyper et al., 2021; Rui et al., 2022). Considering cacao genetics PC variables, PC2 shows a negative correlation between scion and rootstock, indicating influential contrasting SNP allelic patterns affecting AMF composition. In contrast, PC1 is composed of SNPs with similar allelic composition between scion and rootstock, and PC3 of SNPs that are uncorrelated between each other. Notably, the SNP alleles significantly contributing to the scion genetic variable (scion_PC2) signal multiple areas in the genome suggesting a polygenic trait. These findings may serve to identify genomic regions associated with the regulation of mycorrhizal community composition (Supplementary material K).

Evidence across scientific literature regarding the relative influence of scion versus rootstock on AMF communities remains mixed. The scion influence observed in this study contrasts with previous findings in cacao, where the rootstock exerted a stronger influence on bacterial and fungal diversity than the scion (Schmidt et al., 2021). However, this contrasting study assessed ITS sequences, and ITS primers are known to be biased against AMF species (Lekberg et al., 2018), causing their diminished signals in studies of overall fungal diversity due to their typically low abundance (<5%) within the total fungal community (Hawksworth and Lucking, 2017). In a different plant system, citrus, Song et al. (2015) reported a greater influence of the scion, whereas in a subsequent study they found the rootstock to be more impactful (Song et al., 2020). The difference likely reflects that Song et al. (2015) included genotypes from diverse field environments, while Song et al. (2020) controlled for environmental variation by using a single, uniformly managed field.

Regarding environmental variables, elevation and available P were significantly associated with AMF community composition. Elevation can be considered as a proxy for other environmental variables, including geographical location (average altitude of farms in Bocas del Toro was 67 masl, compared with 467 masl in Herrera), temperature (including soil temperature), and precipitation as reflected in soil pH and base saturation, which in turn can alter AMF community composition (Looby et al., 2016). In our results, pH and total N had significantly positive and negative correlations with

elevation (Supplementary material K), and both factors can exert a major influence on AMF community composition (Oehl et al., 2010).

In a global long-term meta-analysis, soil P availability was the main predictor for AMF diversity (Ma et al., 2023). Long-term studies included in this global meta-analysis (Oehl et al., 2004; Wang et al., 2020) covered a broader range of soil available phosphorus (7.59-136 mg/L) due to different fertilization or management treatments, compared to our organically managed plots (0-22 mg/g). Despite this narrower P range, available soil P still had an impact on AMF community composition. For example, Farm 10 in Bocas del Toro showed the lowest AMF diversity, coinciding with higher soil P levels (6.4 mg/L) relative to the other farms (2.72 mg/L on average). Both Farms 9 and 10, managed by a commercial company, transitioned to organic practices 3 and 2.5 years ago, respectively, following conventional fertilizer use. This legacy of fertilization may partly explain the reduced AMF diversity at Farm 10, as high phosphorus availability is known to suppress AMF colonization and richness (Ketchiemo et al., 2024; Z. Qin et al., 2020).

AMF species with significant indicator values (IndVal) serve as bioindicators for these genetic, root, and environmental conditions. Scion germplasm, root P, root N, and elevation had AMF species with high IndVal, while available P and fractional mycorrhizal colonization had intermediate IndVal. Consequently, we can hypothesize scion, nutrient dynamics, and elevation (as proxy for more environmental factors) as the main ecological associations with AMF community composition. Three VT (VTX00113, ASV 388, VTX00175) had high IndVal associated with elevation and intermediate IndVal associated with scion, suggesting a plant filter from the available AMF pool. This shows the association between elevation and scion germplasm, where Herrera on average had a higher altitude than Bocas del Toro, and also Herrera has more trees with Criollo ancestry in contrast with the admixed population of Bocas del Toro. Four indicator species (VTX00115, VTX00186, VTX00280 and VTX00412) had similar IndVal for root P and N in different categories, additionally to either elevation, available P or scion genotype (PC2). It is expected that these AMF species are specialized in nutrient uptake across spatial variation of soil nutrient availability and elevation that influence root nutrient status.

In this study, AMF richness was dominated by the genus Glomus, represented by 69 VT. Glomus has been widely reported as the prevalent AMF in cacao systems (Lovera et al., 2022; Rincón et al., 2021; Valgaz et al., 2022). A recent review of South American cacao farms based on spore morphology documented 67 species, of which 37% belonged to Glomeraceae, compared to 84% observed here (Lovera et al., 2022). The dominance of Glomeraceae likely reflects their colonization ability, ecological adaptability, and the large number of described species in this family. In contrast, Lovera et al. found higher abundances of Acaulosporaceae and Gigasporaceae. Their higher proportion of Acaulosporaceae may relate to the lower soil pH in some South American cacao

plantations, while Gigasporaceae were notably absent from Colombian sites. Other spore-based studies in cacao also identified Glomus and Acaulospora as dominant, but reported lower richness, between 11 and 49 species (Cyriak et al., 2025; Djezou et al., 2025; Kadanga et al., 2024; Rincón et al., 2021; Valgaz et al., 2022). However, spore-based surveys may not reflect the active symbiotic community, as spores are stress-induced, dormant structures rather than active root symbionts (Smith and Read, 2008). Soil-based sequencing can similarly overestimate diversity, since roots filter available AMF taxa (Guo et al., 2022). By focusing on root-associated AMF with molecular tools, this study more accurately captured the symbionts actively colonizing cacao roots.

From the VT identified in this study, some were previously reported in Barro Colorado Island (Husband et al., 2002a, 2002b). When compared with the MaarjAM database (Supplementary Material N), 26% of the detected VT had been previously reported, while 74% appear to be newly documented for this region—highlighting the underexplored nature of AMF communities in tropical cacao systems. From the previously reported VT, four are part of the core mycobiome (VTX00069, VTX00122, VTX00113, and VTX00092) reported from Barro Colorado Island in Panama and Monteverde in Costa Rica. Among these, VTX00069, VTX00113 and VTX00092 are globally widespread, having been detected in 5 to 7 different biomes and across 35 to 53 countries. VTX00122, on the other hand, is less commonly reported at the global scale, occurring in only 4 to 5 biomes and in 17 to 27 countries to date. The other VT species from the core mycobiome, VTX00397 and VTX00270, are also less globally common and newly reported for the region.

Diversity in Herrera was a subset of the diversity in Bocas del Toro, despite similar VT richness and diversity. This difference in species diversity may be caused by more intensive sampling in Bocas del Toro (11 farms) than in Herrera (five farms), especially given that stochasticity (dispersal limitation) is a major factor in structuring AMF communities (Caruso, 2018). Additional factors as elevation, which represents a proxy for environmental factors and geological history with impacts on dispersal limitation can explain absence of certain AMF in Herrera. Regarding cacao germplasm, Bocas del Toro had across farms more admixed germplasm, whereas germplasm in Herrera was more uniform. Host genotype is known to modulate mycorrhizal interactions, influencing colonization levels and nutrient exchange, potentially shifting species abundances (Guigard et al., 2023; Marrassini et al., 2025). Management practices also varied: farms in Herrera were more intensively managed, with frequent pruning and lower shade levels, while those in Bocas del Toro had higher shade and less pruning—reducing photosynthesis but also lowering temperatures.

Conclusion

Most AMF studies have focused on temperate regions, even though these fungi hold great potential for advancing agriculture in tropical ecosystems. Overall, AMF remain understudied,

leaving key questions unresolved—such as how plant genetic identity, particularly grafted scion and rootstock genotypes, and their interaction with environmental factors influence root traits and shape AMF communities. In this study, we investigated AMF diversity in association with *Theobroma cacao* and identified main factors associated with differences in AMF community structure. *Glomus* was the dominant genus across all samples. Scion genotype, root nutrient concentrations (N and P), elevation, and available soil P emerged as significant factors potentially structuring AMF communities, whereas rootstock genotype and root morphological traits had no detectable effect. These findings underscore the importance of host and environmental factors in AMF composition and suggest opportunities to improve sustainable cacao cultivation through optimized plant–fungal interactions. Future research should explore (1) how scion-driven metabolic demand or carbon allocation influences AMF communities under varying environmental constraints, and (2) how to optimize scion germplasm to support efficient AMF symbiosis in low-input tropical agriculture.

Acknowledgments

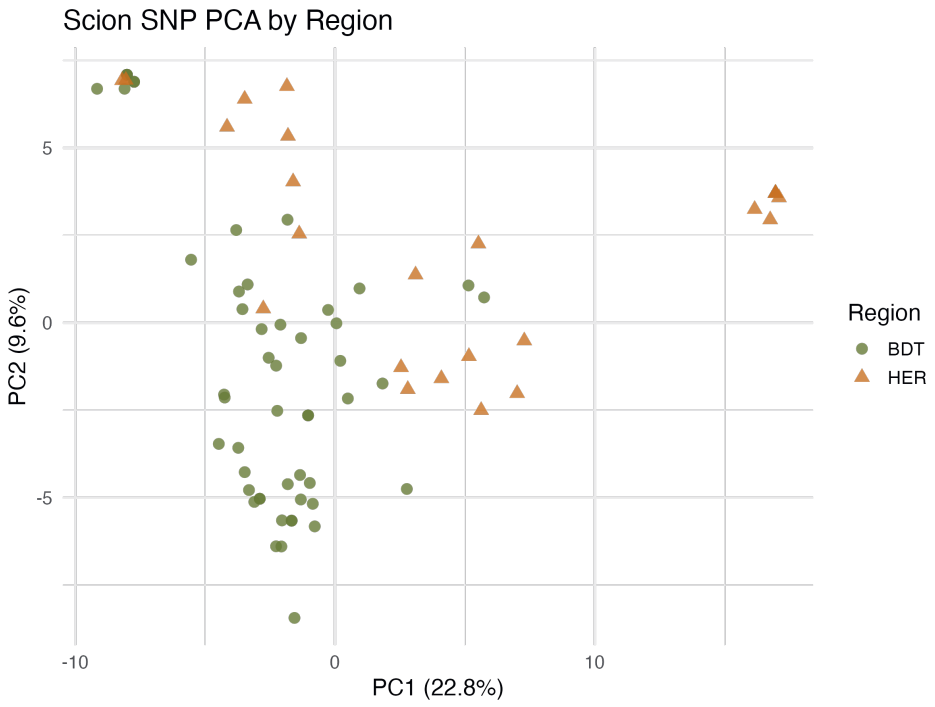
Many thanks to Mareli Sanchez Julia, Phillip J de Lorimier and Hernan Capador for the help, support and guidance through the AMF community analysis.

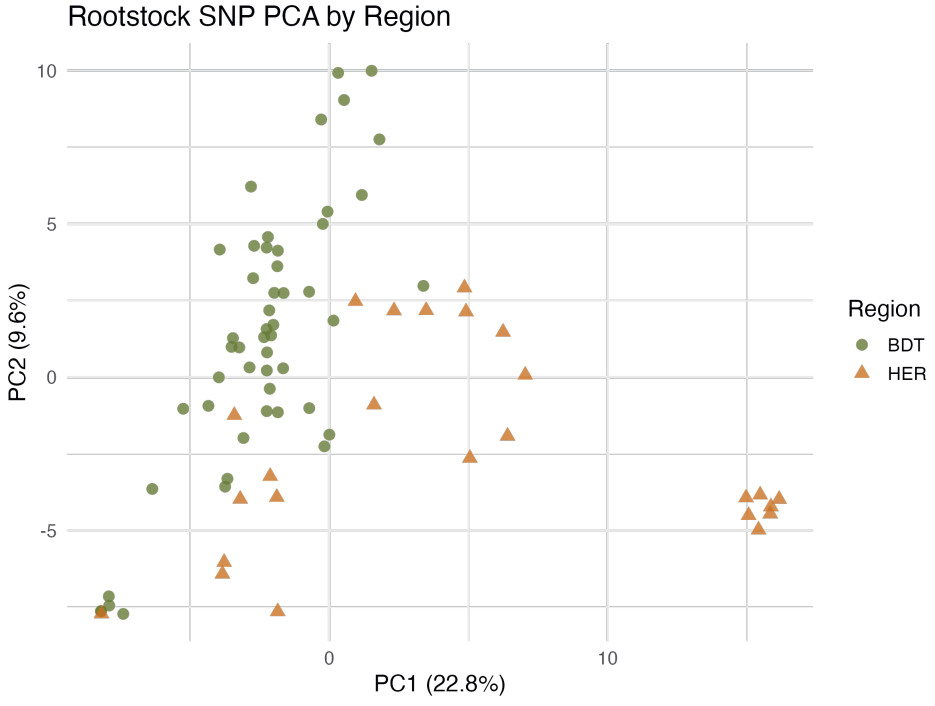
Supplementary material

A. Plantations surveyed across the Panamanian Isthmus

Code	Region	Farm	Town	District	Altitude (masl)	Latitude	Longitude	Secondary Spp	Type	Topography (1 to 3)	Drainage (1 to 3)	Tree density (#/ht)	Tree density (#/ht)	Size (ht)	Average age of trees
BdT_1	Bocas del Toro	1	Milla 7	La Gloria	51	9°21'4"	82°26'29'	7	Cacao Forest	3	3	300	Low	1	50+
BdT_2	Bocas del Toro	2	Milla 7	La Gloria	12	9°19'20"	82°26'48"	6	Cacao Forest	3	3	1000	High	3	15
BdT_3	Bocas del Toro	3	Rio Oeste Arriba	Almirante	39	9°15'23"	82°25'15"	6	Cacao Plantation	3	2	600	Low	2.9	31
BdT_4	Bocas del Toro	4	Rio Oeste Arriba	Almirante	35	9°15'39	82°25'22"	7	Cacao Plantation	1	3	600	Low	1.5	40
BdT_5	Bocas del Toro	5	Soledad	Valle Risco	118	9°15'19"	82°27'23"	4	Forest	2	2	1000	High	4.5	50+
BdT_6	Bocas del Toro	6	Nuduobidi	Valle Risco	81	9°15'21"	82°28'6"	4	Cacao Plantation	1	3	750	Medium	1.5	40
BdT_7	Bocas del Toro	7	Alto conejo	Hirondai	308	8°58'53"	82°20'26"	6	Cacao Plantation	2		2000	Very high	6	5
BdT_8	Bocas del Toro	8	Filo Verde, Tuwai	Hirondai	62	9°0'39"	82°19'51"	3	Permaculture	1	2	933-800	Medium	1	22
BdT_9	Bocas del Toro	9	Quebrada Limon	Almirante	6	9°16'54"	82°25'14"	2	Commercial Plant	2	2	1111	High	3.32	11
BdT_10	Bocas del Toro	10	Rio Uyama	Los Higuerones, Cauchero	4	9°9'91"	82°19'0"	1	Commercial Plant	1	2	1111	High	21.45	10
BdT_11	Bocas del Toro	11	Nuevo Paraiso	Almirante	15	9°16'50"	82°24'49'	5	Exp Garden	1	1	600	Low	0.5	10
Her_12	Herrera	12	Las Minas	Cabecera, Las Minas	338	7°47"51"	80°45'3"	6	Garden	2	3	150	Low	1	7
Her_13	Herrera	13	El Ciprian	Las Minas	404	7°41'3"	80°45'11"	9	Exp Garden	1	1	20	Low	0.25	10
Her_14	Herrera	14	El Nanzal	Las Minas	449	7°44'13"	80°44'21"	6	Forest	3	2	20	Low	1	6
Her_15	Herrera	15	La desbarranca da	Chepo, Las Minas	546	7°39'53"	80°47'16"	6	Cacao Plantation	2	3	200	Low	20	13
Her_16	Herrera	16	Tres puntas	Las	600	7°47'53"	80°48'2"	8	Garden	2	2	200	Low	0.5	6

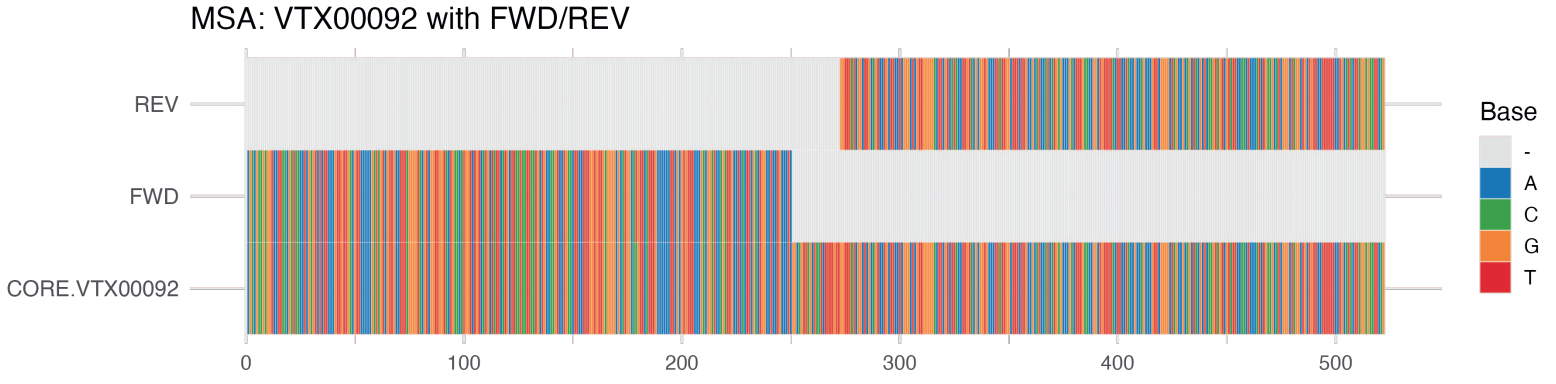
B. Plant SNP PCA with regions indicated by different colours



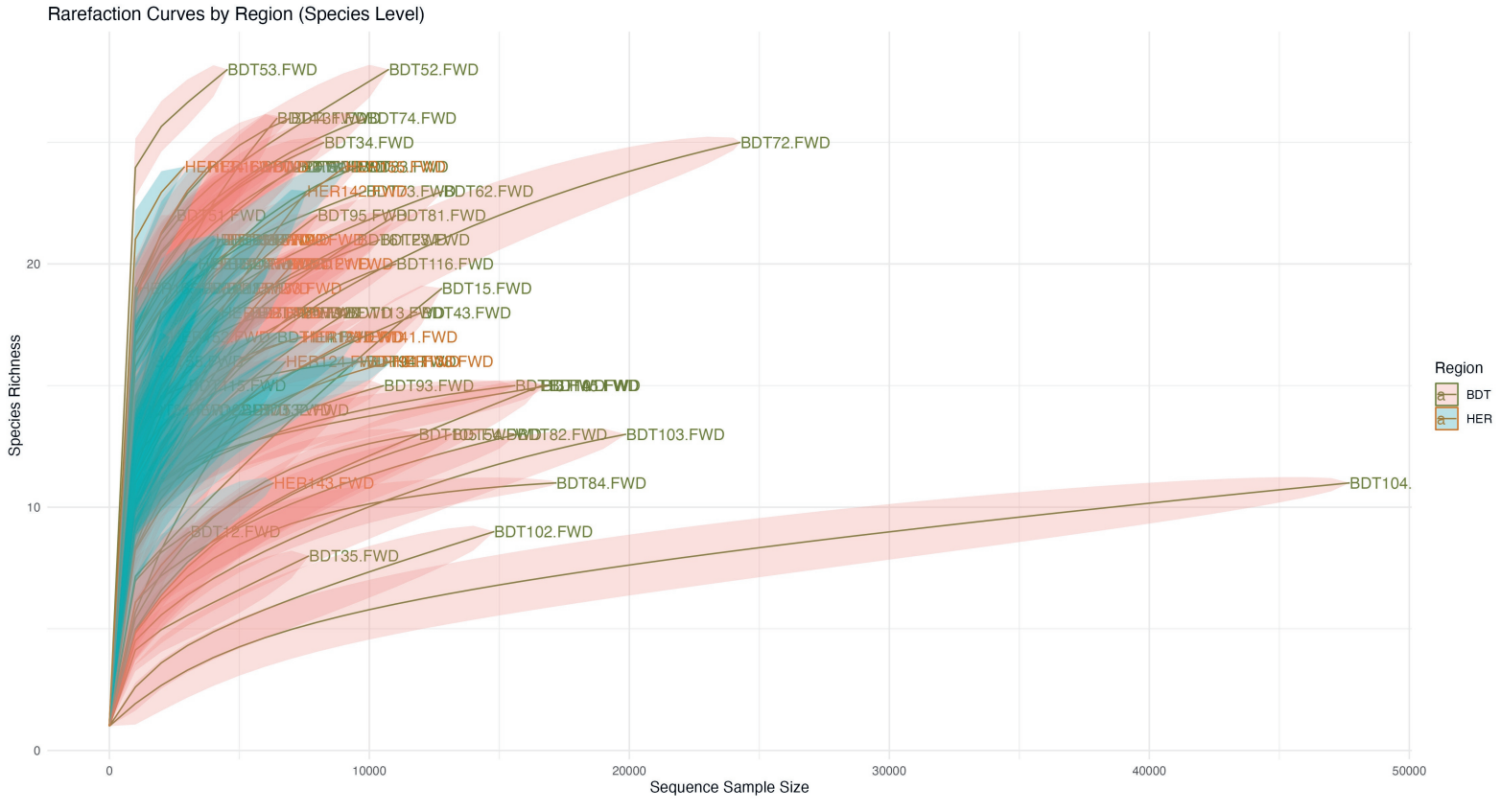


C. Example alignments of forward and reverse reads

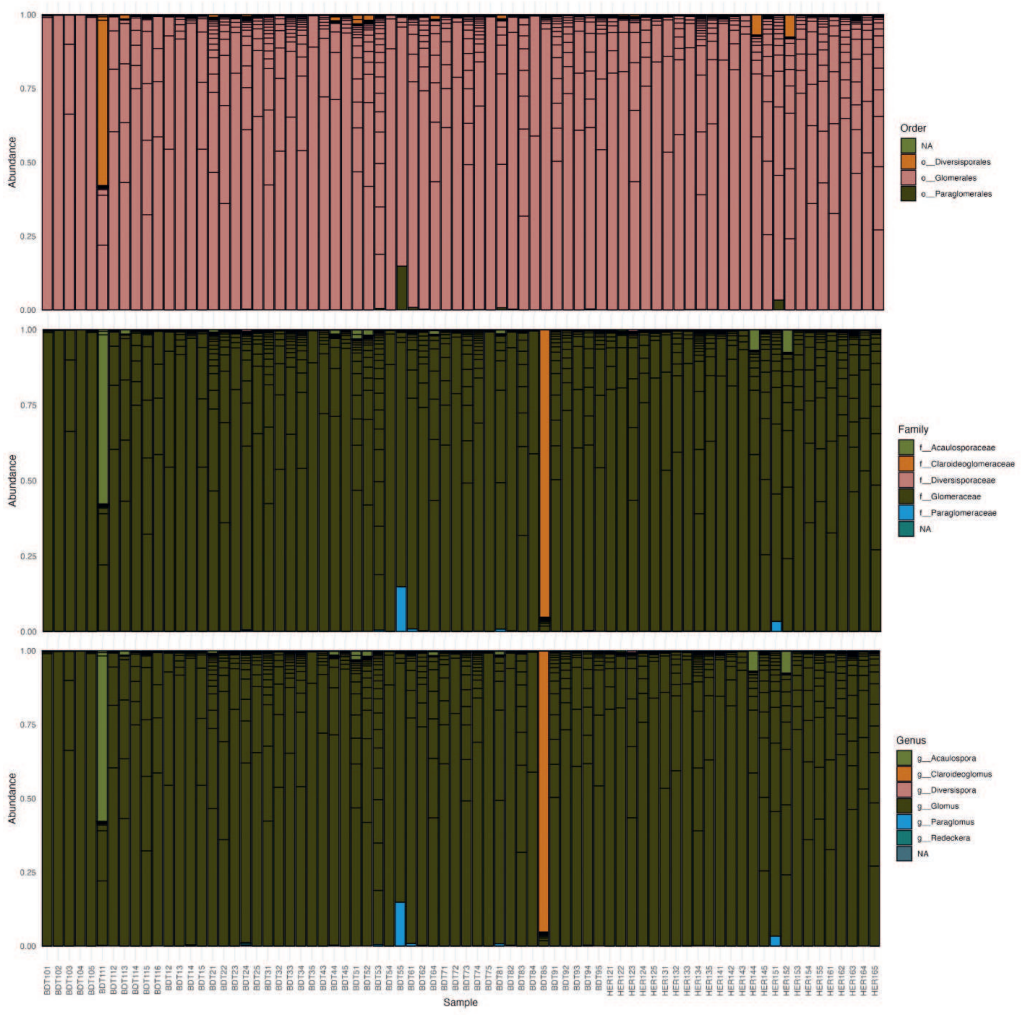
Assigned to Virtual Taxon sequences from the MaarjAM database showing approximately 20 base regions separating reads and preventing successful assembly. Numbers represent position in MAFFT alignment of MaarjAM Virtual Taxon database to which the sequence reads were aligned.

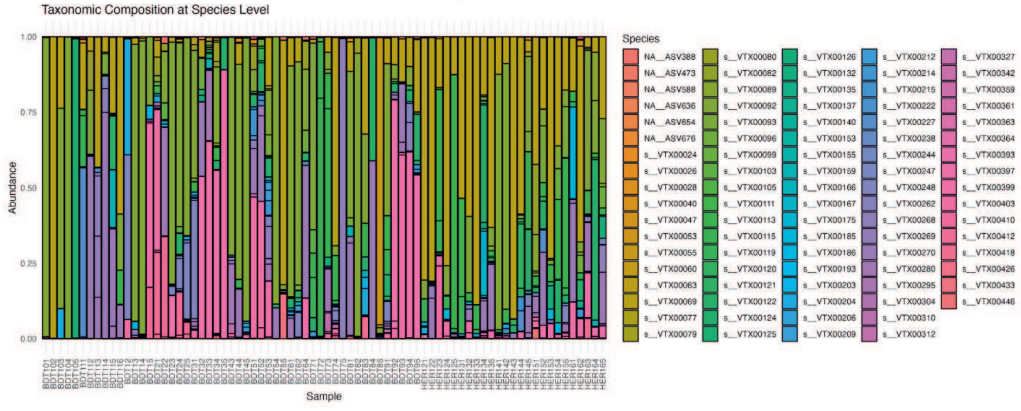


D. Rarefaction Curve



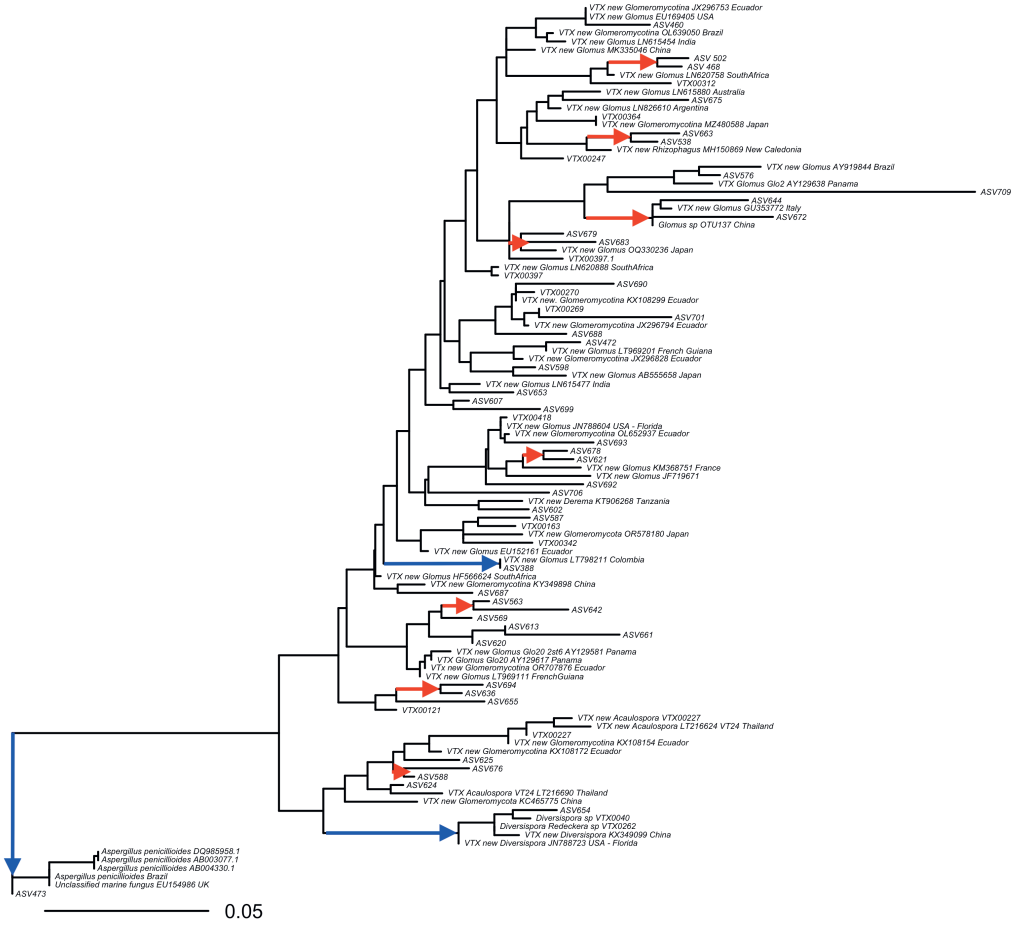
E. Taxonomic information of each sample



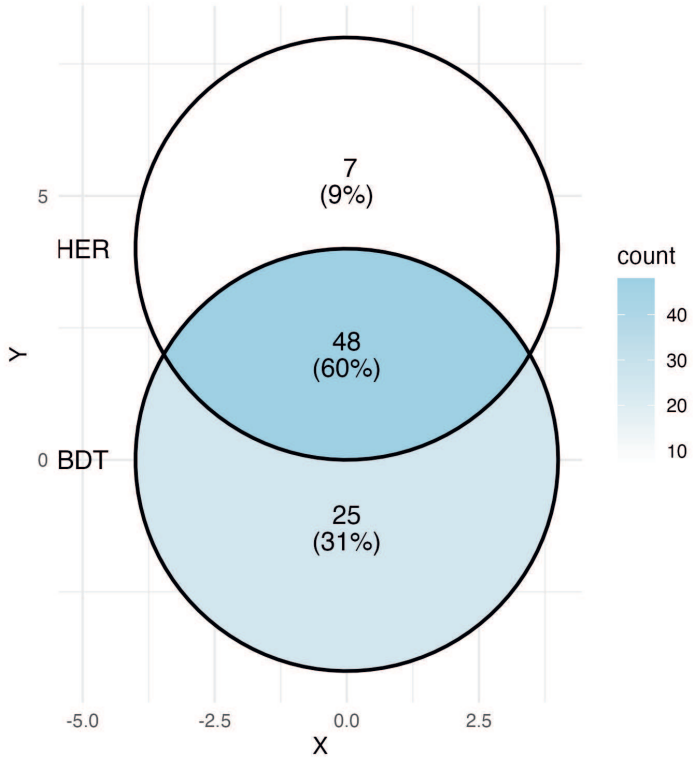


F. Phylogenetic tree of unidentified sequences

Maximum Likelihood tree of the unknown sequences obtained in this study. Red arrows indicate ASVs that are highly similar both phylogenetically and at the sequence level (>97% similarity), likely representing a novel VT. Blue arrows highlight particular ASVs: ASV473 is not a Glomeromycotan ASV; ASV654 likely belongs to Diversispora; and ASV388, closely related to VTX LT798211 from Colombia, is widespread across cacao samples and shows a significant association with the scion_PC2 SNP profile.



G. VT shared between cocoa growing regions in Panama



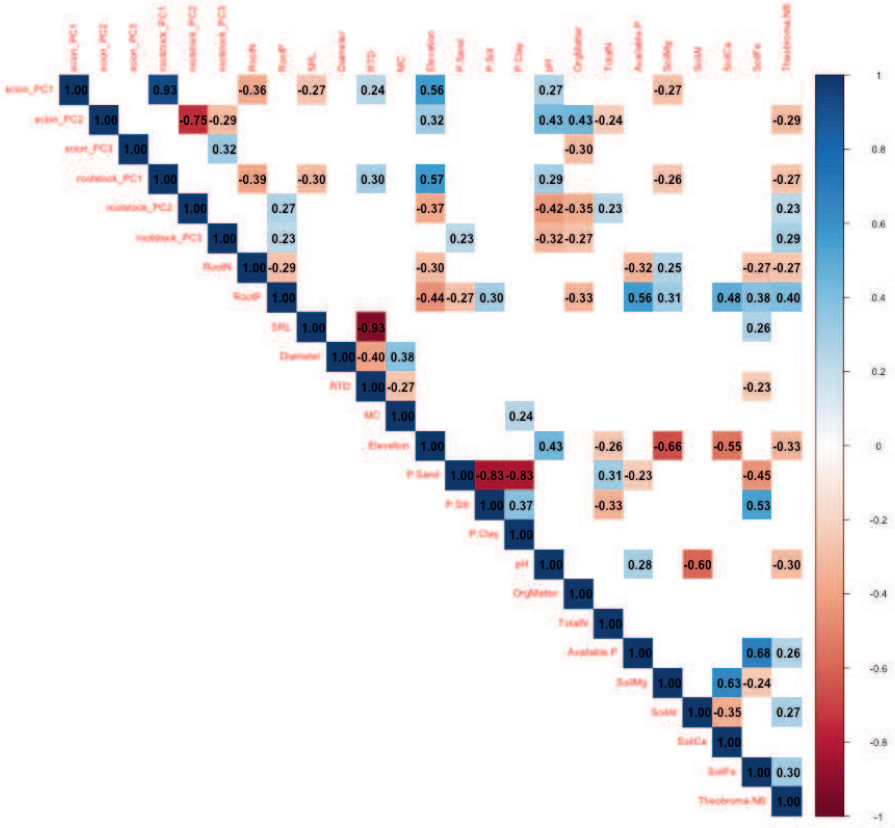
H. Influencing Variables PCA

Loadings and contributions of each variable to the principal components.

Loadings			Contributions		
Variable	PC1	PC2	Variable	PC1	PC2
scion_PC1	0.285	-0.112	scion_PC1	8.1	1.25
scion_PC2	0.247	0.049	scion_PC2	6.11	0.24
scion_PC3	-0.069	0.128	scion_PC3	0.48	1.65
rootstock_PC1	0.298	-0.106	rootstock_PC1	8.88	1.12
rootstock_PC2	-0.232	-0.006	rootstock_PC2	5.4	0
rootstock_PC3	-0.157	0.053	rootstock_PC3	2.48	0.28
RootN	-0.112	-0.161	RootN	1.26	2.58
RootP	-0.129	0.388	RootP	1.65	15.04
SRL	-0.171	0.188	SRL	2.91	3.54
Diameter	-0.097	0.006	Diameter	0.95	0
RTD	0.193	-0.187	RTD	3.74	3.5
MC	0.024	0.155	MC	0.06	2.39
Elevation	0.402	-0.073	Elevation	16.17	0.53
P. Sand	-0.221	-0.33	P. Sand	4.89	10.92
P. Silt	0.193	0.33	P. Silt	3.72	10.92
P. Clay	0.174	0.216	P. Clay	3.02	4.67
pH	0.282	0.029	pH	7.93	0.08
OrgMatter	0.105	-0.124	OrgMatter	1.11	1.54
TotalN	-0.204	-0.151	TotalN	4.16	2.28
Available.P	0.06	0.375	Available.P	0.36	14.06
SoilMg	-0.256	0.037	SoilMg	6.57	0.13
SoilAl	-0.178	-0.023	SoilAl	3.18	0.05
SoilCa	-0.156	0.107	SoilCa	2.43	1.15
SoilFe	0.093	0.406	SoilFe	0.87	16.45
Theobroma.NB	-0.189	0.237	Theobroma.NB	3.59	5.61

I. Correlation matrix between genetic, root and environmental variables

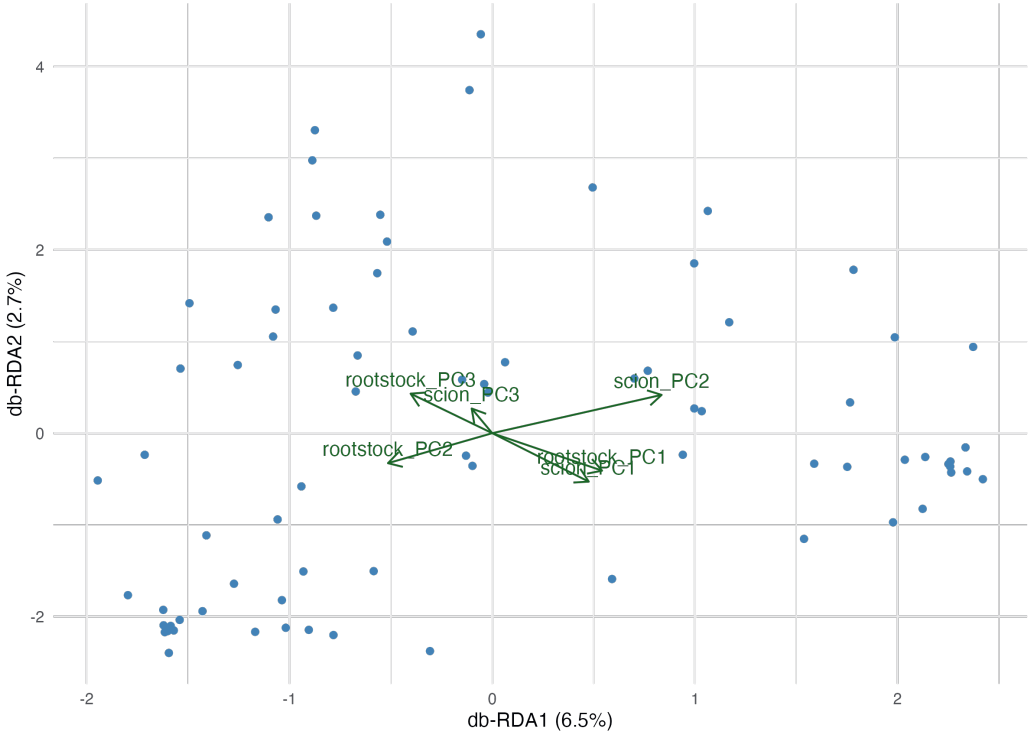
Only correlations that are statistically significant ($p < 0.05$, based on a Pearson correlation test) are shown. The colours indicate the direction and magnitude of the correlation coefficients.



J. Distance-based Redundancy Analysis

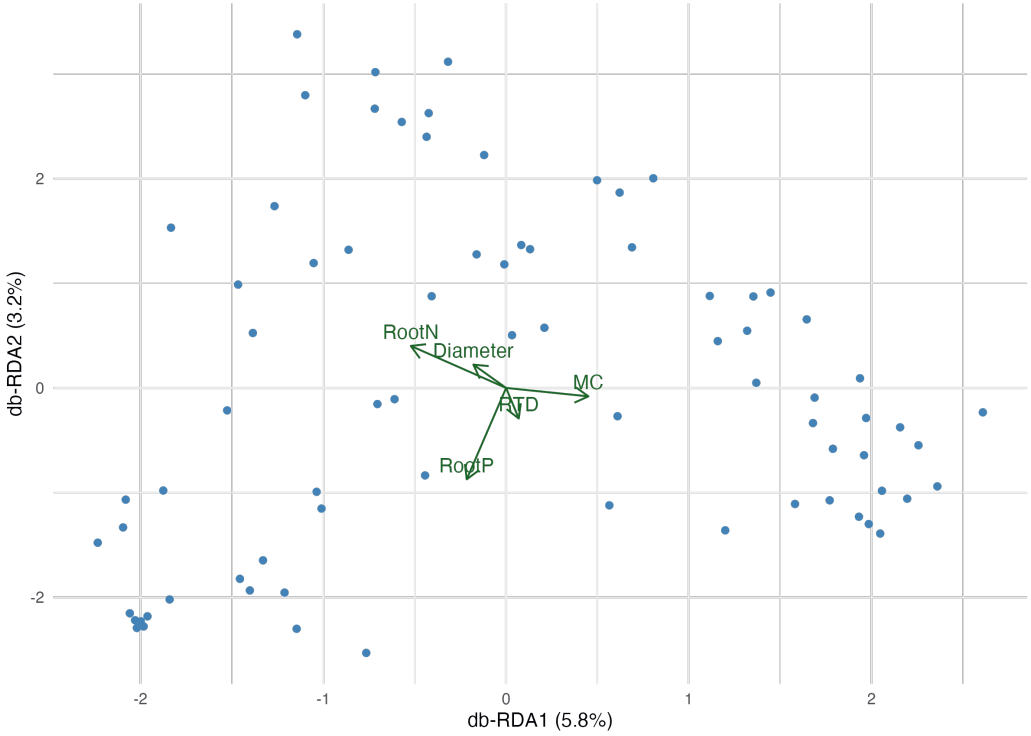
RDA plots and tables showing variable loadings on the canonical axes and the significance of the axes. Significant axes and variables are highlighted in bold with an asterisk. Panel 1 presents a db-RDA on PCA loadings from above- and belowground plant SNP genotyping (scion_PC1 to scion_PC3, rootstock_PC1 to rootstock_PC3); Panel 2 shows a db-RDA on five root functional traits (root diameter, root tissue density, root N, root P, and mycorrhizal colonization [%]); Panel 3 depicts a db-RDA on thirteen soil and environmental variables; and Panel 4 shows a full RDA model including all genetic, root, and environmental variables, highlighting the key factors associated with the mycorrhizal fungal community composition.

1. db-RDA on Genetic Components



Variable	**db-RDA1***	db-RDA2
F value (*P value*)	**5.067 (0.014)**	2.110 (0.529)
scion_PC1	0.4759	-0.5287
scion_PC2*	**0.8353**	0.4180
scion_PC3	-0.1016	0.2693
rootstock_PC1	0.5417	-0.4090
rootstock_PC2	-0.5145	-0.3273
rootstock_PC3	-0.4028	0.4315

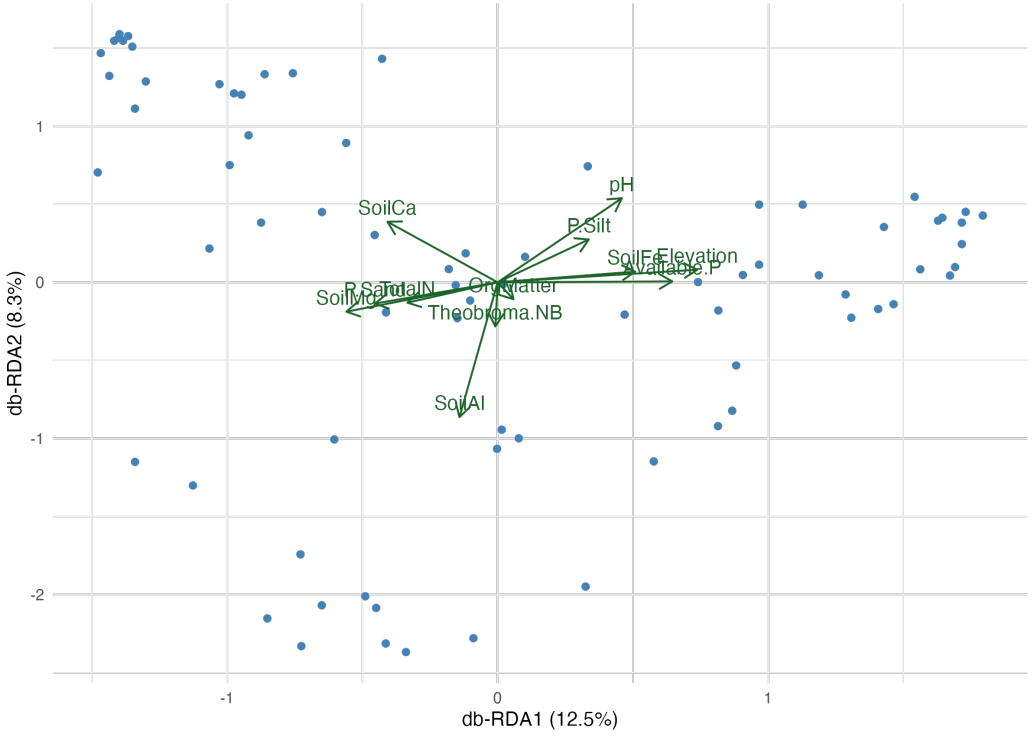
2. db-RDA on Root Properties



Variable	**db-RDA1***	db-RDA2
F value (*P value*)	**4.638 (0.015)**	2.524 (0.205)
RootN*	**-0.5214**	0.4016
RootP*	**-0.2154**	-0.8730
Diameter	-0.1789	0.2227
RTD	0.0693	-0.2920
MC*	**0.4505**	-0.0804

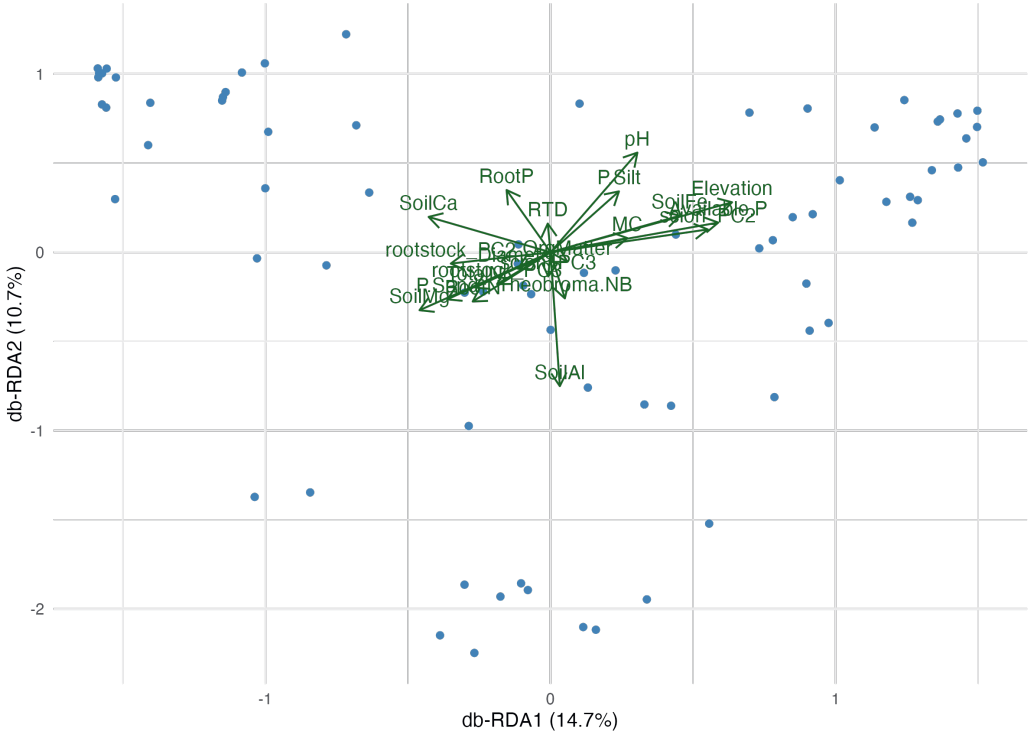
3. db-RDA on Soil Properties

The variable for percentage of clay (P. Clay) did not appear among the biplot scores because it had negligible (close to zero) contribution to the constrained ordination; consequently, the analysis did not assign it a vector in the ordination space. Nevertheless, it was part of the model and term significance analysis.



Variable	**db-RDA1***	**db-RDA2***
F value (*P value*)	**11.38 (0.000)**	**7.540 (0.002)**
Elevation*	**0.7374**	**0.0807**
P. Sand	-0.4530	-0.1361
P. Silt	0.3370	0.2742
pH	0.4590	0.5381
OrgMatter	0.0585	-0.1080
TotalN	-0.3339	-0.1304
Available.P*	**0.6456**	**0.0055**
SoilMg	-0.5591	-0.1880
SoilAl	-0.1411	-0.8632
SoilCa	-0.4073	0.3875
SoilFe	0.5085	0.0684
Theobroma.NB	-0.0083	-0.2828

4. Combined db-RDA with all 21 variables



Variable	db-RDA1*	db-RDA2*	Variable	db-RDA1*	db-RDA2*
	13.75	**10.03**		**13.75**	**10.03**
F value (P value)	**(0.000)**	**(0.000)**	F value (P value)	**(0.000)**	**(0.000)**
scion_PC2	0.5514	0.1282	P. Sand	-0.3620	-0.2641
scion_PC3	-0.0080	-0.1357	P. Silt	0.2400	0.3416
rootstock_PC2	-0.3497	-0.0635	pH	0.3046	0.5592
rootstock_PC3	-0.1874	-0.1787	OrgMatter	0.0592	-0.0537
RootN	-0.2729	-0.2762	TotalN	-0.2647	-0.1979
RootP*	**-0.1541**	**0.3492**	**Available.P***	**0.5882**	**0.1666**
Diameter	-0.1140	-0.0925	SoilMg	-0.4594	-0.3250
RTD	-0.0097	0.1611	SoilAl	0.0326	-0.7511
MC	0.2697	0.0794	SoilCa	-0.4275	0.1984
Elevation	0.6366	0.2805	SoilFe	0.4530	0.2050
			Theobroma.NB	0.0503	-0.2589

K. Allelic contribution to the Scion PCA

Each variable is coded by origin (S = scion, R = rootstock), SNP position, and allele (A/B). Linking group and contribution to each PC are shown, ordered by descending contribution to significant PC2.

Linking Group	Variable	PC1 contrib	PC2 contrib	Linking Group	Variable	PC1 contrib	PC2 contrib
3	S-1062.B	0.135	3.200	1	S-1096.B	0.016	1.077
7	S-1270.B	0.049	3.146	6	S-894.A	0.015	1.038
6	S-561.B	0.096	3.018	6	S-619.A	0.481	0.995
1	S-529.A	0.157	2.975	9	S-632.B	0.607	0.967
5	S-998.A	0.441	2.345	4	S-953.B	0.033	0.957
9	S-172.B	0.069	2.322	3	S-534.A	0.468	0.912
10	S-917.A	0.059	2.294	9	S-11.A	0.341	0.908
8	S-547.A	0.032	2.165	9	S-1442.B	0.552	0.885
10	S-653.A	0.479	1.984	1	S-1075.B	0.003	0.878
9	S-1439.A	0.490	1.870	6	S-309.B	0.480	0.862
9	S-25.A	0.073	1.693	2	S-521.A	0.019	0.861
10	S-546.A	0.111	1.656	2	S-1136.A	1.088	0.832
9	S-25.B	0.086	1.602	8	S-999.A	0.131	0.824
7	S-791.B	0.032	1.566	5	S-645.A	0.938	0.813
9	S-193.B	0.463	1.532	9	S-1414.A	0.069	0.794
4	S-886.A	0.524	1.513	3	S-689.A	1.066	0.755
2	S-836.A	0.003	1.512	3	S-1062.A	0.034	0.726
9	S-563.A	0.552	1.502	1	S-1350.A	0.000	0.714
1	S-205.A	0.535	1.447	2	S-1159.B	1.205	0.709
2	S-891.B	0.030	1.408	8	S-151.B	0.469	0.692
3	S-413.A	0.486	1.400	8	S-151.A	0.464	0.685
6	S-750.A	0.041	1.378	2	S-316.B	0.505	0.685
3	S-1149.A	0.085	1.315	9	S-172.A	0.007	0.676
7	S-1270.A	0.031	1.297	1	S-510.A	0.511	0.674
8	S-269.A	0.552	1.293	8	S-547.B	0.016	0.657
6	S-561.A	0.067	1.282	2	S-521.B	0.026	0.654
5	S-150.A	0.568	1.244	8	S-139.B	0.148	0.651
8	S-999.B	0.009	1.233	7	S-1126.A	0.814	0.608
10	S-560.A	0.446	1.155	7	S-944.A	0.957	0.591
10	S-546.B	0.002	1.096	9	S-427.A	0.284	0.539
9	S-1253.A	0.484	1.095	10	S-144.B	0.808	0.508
1	S-1350.B	0.071	1.086				

L. Indicator Species Analysis

VT significantly associated to each significant variable identified by the db-RDA models are presented in the following table, ordered by decreasing indicator value. For each VT, the corresponding category of the variable (Low, Medium, or High) is also indicated.

Factor	Species	Low	Med	High	IndVal	p.value
Elevation	VTX00113	0	1	1	0.8426	0.0001
Available P	VTX00186	1	0	0	0.5976	0.0001
Available P	VTX00412	1	0	0	0.5306	0.0004
Root N	VTX00186	0	0	1	0.5367	0.0006
Elevation	VTX00115	0	1	1	0.7067	0.002
Root N	VTX00115	1	1	0	0.6913	0.0039
Available P	VTX00103	1	0	0	0.4629	0.0041
Mperc	VTX00175	0	0	1	0.4797	0.0044
Elevation	VTX00175	0	0	1	0.4729	0.0046
Root N	VTX00342	0	0	1	0.4472	0.0056
Elevation	VTX00185	0	1	1	0.5421	0.0067
Elevation	NA - ASV 388	1	1	0	0.7109	0.0068
Root P	VTX00186	0	1	0	0.4781	0.007
Scion	VTX00215	0	0	1	0.4929	0.0078
Available P	VTX00342	1	0	0	0.4226	0.0091
Available P	VTX00209	1	0	0	0.4226	0.0094
Scion	VTX00069	0	0	1	0.6597	0.0105
Scion	VTX00175	0	0	1	0.4480	0.0114
Root P	VTX00412	0	1	0	0.4597	0.0121
Root N	VTX00412	0	0	1	0.4575	0.0126
Scion	VTX00113	0	0	1	0.6319	0.0129
Mperc	VTX00111	0	0	1	0.3928	0.0141
Available P	VTX00096	1	1	0	0.5494	0.0178
Mperc	VTX00247	1	1	0	0.5779	0.018
Root P	VTX00126	1	0	0	0.4232	0.0187
Root N	VTX00111	0	0	1	0.3836	0.0236
Scion	VTX00280	0	1	0	0.5422	0.0256
Root P	VTX00248	1	1	0	0.6326	0.0257
Root P	VTX00096	1	0	0	0.5406	0.0285
Root N	VTX00399	0	1	1	0.4814	0.0288
Root N	VTX00280	0	1	1	0.5968	0.0303
Elevation	VTX00186	1	1	0	0.4428	0.037
Root P	VTX00280	1	1	0	0.5927	0.038
Elevation	VTX00399	1	1	0	0.4761	0.0391
Root P	VTX00115	1	0	1	0.6535	0.0424
Root P	VTX00093	1	0	1	0.5890	0.0426
Scion	NA - ASV 388	0	1	0	0.5539	0.0476
Scion	VTX00342	1	0	0	0.3677	0.0481
Available P	VTX00364	0	0	1	0.3321	0.0484
Root N	VTX00120	0	1	0	0.3859	0.0495

M. AMF Indicator VTs occurrence scion genotype across samples

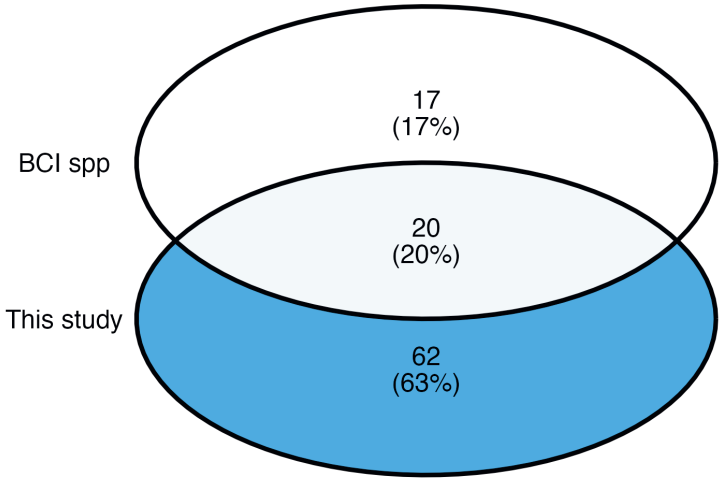
Table shows the relative abundances of each VT per sample. Values above 1% are highlighted in green.

Sample ID	VTX00069	VTX00113	VTX00280	VTX00215	VTX00175	NA ASV388	VTX00342
BDT101	0.002	0.000	0.000	0.000	0.000	0.000	0.000
BDT102	0.998	0.000	0.000	0.000	0.000	0.000	0.000
BDT103	0.237	0.000	0.000	0.000	0.000	0.000	0.000
BDT104	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BDT105	0.002	0.000	0.000	0.000	0.000	0.000	0.000
BDT111	0.219	0.019	0.000	0.000	0.000	0.002	0.000
BDT112	0.047	0.000	0.000	0.000	0.000	0.000	0.000
BDT113	0.431	0.000	0.000	0.000	0.000	0.000	0.000
BDT114	0.051	0.001	0.000	0.000	0.000	0.000	0.000
BDT115	0.252	0.001	0.039	0.000	0.000	0.002	0.000
BDT116	0.586	0.112	0.000	0.000	0.000	0.000	0.000
BDT12	0.000	0.002	0.000	0.000	0.000	0.003	0.000
BDT13	0.021	0.000	0.000	0.000	0.000	0.001	0.001
BDT14	0.008	0.000	0.000	0.000	0.000	0.000	0.000
BDT15	0.000	0.000	0.000	0.000	0.000	0.000	0.006
BDT21	0.056	0.000	0.000	0.000	0.000	0.003	0.001
BDT22	0.000	0.000	0.360	0.000	0.000	0.022	0.000
BDT23	0.017	0.000	0.001	0.000	0.000	0.001	0.000
BDT24	0.010	0.000	0.004	0.000	0.000	0.004	0.001
BDT25	0.000	0.000	0.001	0.000	0.000	0.003	0.002
BDT31	0.025	0.033	0.009	0.000	0.000	0.001	0.000
BDT32	0.059	0.000	0.000	0.000	0.000	0.009	0.000
BDT33	0.013	0.000	0.233	0.000	0.000	0.002	0.002
BDT34	0.008	0.044	0.000	0.000	0.000	0.004	0.000
BDT35	0.000	0.000	0.000	0.000	0.000	0.002	0.000
BDT43	0.001	0.000	0.199	0.000	0.000	0.000	0.000
BDT44	0.070	0.029	0.000	0.000	0.000	0.000	0.000
BDT45	0.001	0.000	0.000	0.017	0.000	0.003	0.000
BDT51	0.007	0.001	0.000	0.087	0.000	0.007	0.000
BDT52	0.001	0.041	0.000	0.018	0.003	0.002	0.000
BDT53	0.161	0.004	0.055	0.048	0.000	0.002	0.000
BDT54	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BDT55	0.809	0.001	0.000	0.001	0.000	0.011	0.000
BDT61	0.093	0.000	0.067	0.000	0.000	0.003	0.000
BDT62	0.081	0.000	0.000	0.001	0.000	0.002	0.000
BDT64	0.001	0.000	0.001	0.000	0.000	0.011	0.000
BDT71	0.012	0.081	0.001	0.000	0.000	0.007	0.000

Sample ID	VTX00069	VTX00113	VTX00280	VTX00215	VTX00175	NA ASV388	VTX00342
BDT72	0.001	0.187	0.000	0.000	0.000	0.001	0.000
BDT73	0.050	0.117	0.140	0.000	0.000	0.001	0.000
BDT74	0.690	0.150	0.000	0.000	0.000	0.002	0.000
BDT75	0.002	0.000	0.000	0.000	0.000	0.000	0.000
BDT81	0.097	0.000	0.000	0.000	0.000	0.003	0.000
BDT82	0.001	0.000	0.000	0.000	0.000	0.000	0.000
BDT83	0.318	0.111	0.000	0.000	0.000	0.003	0.000
BDT84	0.000	0.000	0.000	0.000	0.000	0.003	0.000
BDT85	0.009	0.004	0.001	0.000	0.000	0.002	0.000
BDT91	0.085	0.112	0.000	0.000	0.000	0.004	0.000
BDT92	0.092	0.000	0.000	0.000	0.000	0.001	0.000
BDT93	0.143	0.000	0.033	0.000	0.000	0.007	0.000
BDT94	0.096	0.014	0.000	0.000	0.000	0.002	0.000
BDT95	0.239	0.000	0.000	0.012	0.000	0.005	0.000
HER121	0.802	0.000	0.000	0.000	0.014	0.000	0.000
HER122	0.807	0.004	0.000	0.000	0.000	0.002	0.000
HER123	0.161	0.435	0.000	0.000	0.003	0.000	0.000
HER124	0.758	0.003	0.000	0.000	0.000	0.000	0.000
HER125	0.125	0.840	0.009	0.000	0.000	0.000	0.000
HER131	0.534	0.325	0.000	0.000	0.000	0.000	0.000
HER132	0.594	0.028	0.000	0.000	0.009	0.000	0.000
HER133	0.887	0.082	0.000	0.000	0.000	0.000	0.000
HER134	0.192	0.009	0.004	0.003	0.213	0.000	0.000
HER135	0.629	0.018	0.000	0.000	0.003	0.001	0.000
HER141	0.124	0.008	0.000	0.000	0.000	0.000	0.000
HER142	0.088	0.046	0.000	0.017	0.000	0.000	0.000
HER143	0.934	0.000	0.000	0.000	0.000	0.000	0.000
HER144	0.587	0.011	0.000	0.046	0.000	0.000	0.000
HER145	0.039	0.255	0.016	0.000	0.000	0.001	0.000
HER151	0.422	0.032	0.014	0.000	0.000	0.001	0.000
HER152	0.237	0.047	0.000	0.000	0.000	0.001	0.000
HER153	0.734	0.055	0.000	0.000	0.000	0.000	0.000
HER154	0.360	0.264	0.000	0.000	0.000	0.000	0.000
HER155	0.175	0.430	0.027	0.000	0.000	0.001	0.000
HER161	0.176	0.034	0.000	0.000	0.000	0.001	0.000
HER162	0.648	0.023	0.000	0.027	0.000	0.008	0.000
HER163	0.034	0.078	0.160	0.000	0.000	0.001	0.000
HER164	0.051	0.014	0.000	0.000	0.000	0.000	0.000
HER165	0.271	0.023	0.002	0.000	0.000	0.000	0.000

N. Shared VT between VT previously reported from Barro Colorado Island and this study

VTs reported from Barro Colorado Island in the MaarjAM database were compared with those found in this study using the following Venn diagram.



References

Alexander, I., Selosse, M., 2009. Mycorrhizas in tropical forests: a neglected research imperative. New Phyt. 182, 14–16. https://doi.org/10.1111/j.1469-8137.2009.02798.x

Altieri, M.A., 2004. Linking ecologists and traditional farmers in the search for sustainable agriculture. Front. Ecol. Environ 2, 35–42. https://doi.org/10.1890/1540-9295(2004)002[0035:leatfi]2.0.co;2

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/s0022-2836(05)80360-2

Argout, X., Fouet, O., Wincker, P., Gramacho, K., Legavre, T., Sabau, X., Risterucci, A.M., Silva, C.D., Cascardo, J., Allegre, M., Kuhn, D., Verica, J., Courtois, B., Loor, G., Babin, R., Sounigo, O., Ducamp, M., Guiltinan, M.J., Ruiz, M., Alemanno, L., Machado, R., Phillips, W., Schnell, R., Gilmour, M., Rosenquist, E., Butler, D., Maximova, S., Lanaud, C., 2008. Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. BMC Genom. 9, 512. https://doi.org/10.1186/1471-2164-9-512

Argout, X., Salse, J., Aury, J.-M., Guiltinan, M.J., Droc, G., Gouzy, J., Allegre, M., Chaparro, C., Legavre, T., Maximova, S.N., Abrouk, M., Murat, F., Fouet, O., Poulain, J., Ruiz, M., Roguet, Y., Rodier-Goud, M., Barbosa-Neto, J.F., Sabot, F., Kudrna, D., Ammiraju, J.S.S., Schuster, S.C., Carlson, J.E., Sallet, E., Schiex, T., Dievart, A., Kramer, M., Gelley, L., Shi, Z., Bérard, A., Viot, C., Boccara, M., Risterucci, A.M., Guignon, V., Sabau, X., Axtell, M.J., Ma, Z., Zhang, Y., Brown, S., Bourge, M., Golser, W., Song, X., Clement, D., Rivallan, R., Tahi, M., Akaza, J.M., Pitollat, B., Gramacho, K., D'Hont, A., Brunel, D., Infante, D., Kebe, I., Costet, P., Wing, R., McCombie, W.R., Guiderdoni, E., Quetier, F., Panaud, O., Wincker, P., Bocs, S., Lanaud, C., 2011. The genome of *Theobroma cacao*. Nat. Genet 43, 101–108. https://doi.org/10.1038/ng.736

Bergmann, J., Weigelt, A., Plas, F. van der, Laughlin, D.C., Kuyper, T.W., Guerrero-Ramirez, N., Valverde-Barrantes, O.J., Bruelheide, H., Freschet, G.T., Iversen, C.M., Kattge, J., McCormack, M.L., Meier, I.C., Rillig, M.C., Roumet, C., Semchenko, M., Sweeney, C.J., Ruijven, J. van, York, L.M., Mommer, L., 2020. The fungal collaboration gradient dominates the root economics space in plants. Sci Adv 6, eaba3756. https://doi.org/10.1126/sciadv.aba3756

Broeckhoven, I., Devriese, A., Honnay, O., Merckx, R., Bruno, V., 2025. Impact of agricultural systems on arbuscular mycorrhizal fungi community composition in robusta coffee roots in the Democratic Republic of Congo. Mycorrhiza 35, 30. https://doi.org/10.1007/s00572-025-01204-9

Cáceres, M.D., Legendre, P., 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90, 3566–3574. https://doi.org/10.1890/08-1823.1

Cardoso, I.M., Kuyper, T.W., 2006. Mycorrhizas and tropical soil fertility. Agric. Ecosyst. Environ. 116, 72–84. https://doi.org/10.1016/j.agee.2006.03.011

Caruso, T., 2018. Disentangling the factors shaping arbuscular mycorrhizal fungal communities across multiple spatial scales. New Phyt. 220, 954–956. https://doi.org/10.1111/nph.15212

Colwell, R.K., 2009. Biodiversity: Concepts, Patterns, and Measurement. The Princeton Guide to Ecology, p. 257–263. https://doi.org/10.1515/9781400833023.257

Cyriak, A.Y.F., Alban, M.K.A., Mamadou, C., Daouda, K., Christophe, K., 2025. Mycorrhizal Interaction between Companion Trees and Cocoa Trees (*Theobroma cacao* L.) in Traditional Agroforestry Systems in Côte d'Ivoire. Microbiol. Res. J. Int. 35, 25–44. https://doi.org/10.9734/mrji/2025/v35i31548

Davison, J., Öpik, M., Zobel, M., Vasar, M., Metsis, M., Moora, M., 2012. Communities of Arbuscular Mycorrhizal Fungi Detected in Forest Soil Are Spatially Heterogeneous but Do Not Vary throughout the Growing Season. PLoS ONE 7, e41938. https://doi.org/10.1371/journal.pone.0041938

Deveautour, C., Donn, S., Power, S.A., Bennett, A.E., Powell, J.R., 2018. Experimentally altered rainfall regimes and host root traits affect grassland arbuscular mycorrhizal fungal communities. Mol. Ecol. 27, 2152–2163. https://doi.org/10.1111/mec.14536

Djezou, K.M., Lammoglia, S.-K., Droh, G., Jagoret, P., Claude, N.J., 2025. Diversity and Abundance of Arbuscular Mycorrhizal Fungi in Cocoa (*Theobroma cacao* L.) Rhizospheres Under Contrasting Soil Conditions in Côte D'Ivoire. Under review. SSRN. https://doi.org/10.2139/ssrn.5237048

Fang, W., Meinhardt, L.W., Mischke, S., Bellato, C.M., Motilal, L., Zhang, D., 2014. Accurate Determination of Genetic Identity for a Single Cacao Bean, Using Molecular Markers with a Nanofluidic System, Ensures Cocoa Authentication. J. Agric. Food Chem. 62, 481–487. https://doi.org/10.1021/jf404402v

Guigard, L., Jobert, L., Busset, N., Moulin, L., Czernic, P., 2023. Symbiotic compatibility between rice cultivars and arbuscular mycorrhizal fungi genotypes affects rice growth and mycorrhiza-induced resistance. Front. Plant Sci. 14, 1278990. https://doi.org/10.3389/fpls.2023.1278990

Guo, Y., Bei, Q., Dzomeku, B.M., Martin, K., Rasche, F., 2022. Genetic diversity and community composition of arbuscular mycorrhizal fungi associated with root and rhizosphere soil of the pioneer plant *Pueraria phaseoloides*. iMeta 1, e51. https://doi.org/10.1002/imt2.51

Hawkins, H.-J., Cargill, R.I.M., Nuland, M.E.V., Hagen, S.C., Field, K.J., Sheldrake, M., Soudzilovskaia, N.A., Kiers, E.T., 2023. Mycorrhizal mycelium as a global carbon pool. Curr. Biol. 33, R560–R573. https://doi.org/10.1016/j.cub.2023.02.027

Hawksworth, D.L., Lücking, R., 2017. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. Microbiology Spectrum 5, art. 1128. https://doi.org/10.1128/microbiolspec.funk-0052-2016

Husband, R., Herre, E.A., Turner, S.L., Gallery, R., Young, J.P.W., 2002. Molecular diversity of arbuscular mycorrhizal fungi and patterns of host association over time and space in a tropical forest. Mol. Ecol. 11, 2669–2678. https://doi.org/10.1046/j.1365-294x.2002.01647.x

Husband, R., Herre, E.A., Young, J.P.W., 2002. Temporal variation in the arbuscular mycorrhizal communities colonising seedlings in a tropical rainforest. FEMS Microb. Ecol. 42, 131-136. https://doi.org/10.1016/s0168-6496(02)00323-9

Jemo, M., Dhiba, D., Hashem, A., Abd_Allah, E.F., Alqarawi, A.A., Tran, L.-S.P., 2018. Mycorrhizal fungal community structure in tropical humid soils under fallow and cropping conditions. Sci. Rep. 8, art. 17061. https://doi.org/10.1038/s41598-018-34736-6

Ji, K., Zhang, D., Motilal, L.A., Boccara, M., Lachenaud, P., Meinhardt, L.W., 2013. Genetic diversity and parentage in farmer varieties of cacao (*Theobroma cacao* L.) from Honduras and Nicaragua as revealed by single nucleotide polymorphism (SNP) markers. Genet. Resour. Crop Evol. 60, 441–453. https://doi.org/10.1007/s10722-012-9847-1

Kadanga, P., Tchabi, A., Ablede, K.A., M'Boumba, E.B., Ametefe, K.E., Kondow, M., 2024. Diversity of Arbuscular Mycorrhizal Fungi Associated with Cocoa (*Theobroma cacao* L.) Agroforestry Systems in Togo. Curr. J. Appl. Sci. Technol. 43, 30–44. https://doi.org/10.9734/cjast/2024/v43i114442

Ketchiemo, F.T., Fotso, B., Djabou, A.S.M., Evina, V.J.E., Essambita, J.Y., Tang, F.M.E., Niemenak, N., 2024. Soil Physico-Chemical Properties and Different Altitudes Affect Arbuscular Mycorrhizal Fungi Abundance and Colonization in Cacao Plantations of Cameroon. Am. J. Plant Sci. 15, 57–82. https://doi.org/10.4236/ajps.2024.152005

Kuyper, T.W., Cardoso, I.M., Onguene, N.A., Murniati, M., Noordwijk, M. van, 2004. Managing mycorrhiza in tropical multispecies agroecosystems. in: Noordwijk, M. van, Cadisch, G, Ong, C.K. (Eds), Below-ground interactions in tropical agroecosystems: concepts and models with multiple plant components 243–261. https://doi.org/10.1079/9780851996738.0243

Kuyper, T.W., Wang, X., Muchane, M.N., 2021. The Interplay Between Roots and Arbuscular Mycorrhizal Fungi Influencing Water and Nutrient Acquisition and Use Efficiency, in: Rengel, Z., Djalovic, I. (Eds.), The Root Systems in Sustainable Agricultural Intensification. pp. 193–220. https://doi.org/10.1002/9781119525417.ch7

Lê, S., JulieJosse, Husson, F., 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25, 1-18. https://doi.org/10.18637/jss.v025.i01.

Legendre, P., Anderson, M.J., 1999. Distance-Based Redundancy Analysis: Testing Multispecies Responses in Multifactorial Ecological Experiments. Ecol. Monogr. 69, 1–24. https://doi.org/10.2307/2657192

Lekberg, Y., Vasar, M., Bullington, L.S., Sepp, S., Antunes, P.M., Bunn, R., Larkin, B.G., Öpik, M., 2018. More bang for the buck? Can arbuscular mycorrhizal fungal communities be characterized adequately alongside other fungi using general fungal primers? New Phyt. 220, 971–976. https://doi.org/10.1111/nph.15035

Lindo, A.A., Robinson, D.E., Tennant, P.F., Meinhardt, L.W., Zhang, D., 2018. Molecular Characterization of Cacao (*Theobroma cacao*) Germplasm from Jamaica Using Single Nucleotide Polymorphism (SNP) Markers. Trop. Plant Biol. 11, 93–106. https://doi.org/10.1007/s12042-018-9203-5

Looby, C.I., Maltz, M.R., Treseder, K.K., 2016. Belowground responses to elevation in a changing cloud forest. Ecol. Evol. 6, 1996–2009. https://doi.org/10.1002/ece3.2025

Lovera, M., Cuenca, G., Fajardo, L., Cáceres, A., Guerra-Sierra, B.E., 2022. AMF Diversity in Coffee and Cacao Agroforestry Systems: Importance for Crop Productivity and Forest Conservation, In: Luga M.A., Pagano, M.C. (Eds), Mycorrhizal Fungi in South America,

Biodiversity, Conservation, and Sustainable Food Production, pp. 107–127. https://doi.org/10.1007/978-3-031-12994-0_5

Ma, X., Xu, X., Geng, Q., Luo, Y., Ju, C., Li, Q., Zhou, Y., 2023. Global arbuscular mycorrhizal fungal diversity and abundance decreases with soil available phosphorus. Glob. Ecol. Biogeogr. 32, 1423–1434. https://doi.org/10.1111/geb.13704

Mai, S., Mao, H., Jiang, Y., Huang, T., Yang, Q., Xing, G., Wang, X., Yang, H., Liu, W., 2023. Characteristics of the soil arbuscular mycorrhizal fungal community along succession stages in tropical forest and its driving factors. Front. Environ. Sci. 10, 1110655. https://doi.org/10.3389/fenvs.2022.1110655

Marrassini, V., Ercoli, L., Paredes, A.V.A., Pellegrino, E., 2025. Positive response to inoculation with indigenous arbuscular mycorrhizal fungi as modulated by barley genotype. Agron. Sustain. Dev. 45, 21. https://doi.org/10.1007/s13593-025-01016-3

McMurdie, P.J., Holmes, S., 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE 8, e61217. https://doi.org/10.1371/journal.pone.0061217

Morgan, B.S.T., Egerton-Warburton, L.M., 2017. Barcoded NS31/AML2 Primers for Sequencing of Arbuscular Mycorrhizal Communities in Environmental Samples. Appl. Plant Sci. 5, 1700017. https://doi.org/10.3732/apps.1700017

Oehl, F., Laczko, E., Bogenrieder, A., Stahr, K., Bösch, R., Heijden, M. van der, Sieverding, E., 2010. Soil type and land use intensity determine the composition of arbuscular mycorrhizal fungal communities. Soil Biol. Biochem. 42, 724–738. https://doi.org/10.1016/j.soilbio.2010.01.006

Oehl, F., Sieverding, E., Mäder, P., Dubois, D., Ineichen, K., Boller, T., Wiemken, A., 2004. Impact of long-term conventional and organic farming on the diversity of arbuscular mycorrhizal fungi. Oecologia 138, 574–583. https://doi.org/10.1007/s00442-003-1458-2

Oksanen, J., Simpson, G.L., Blanchet, G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Borman, T., Carvalho, G., Chirico, M., Caceres, M.D., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., Martino, C., McGlinn, D., Ouellette, M.-H., Cunha, E.R., Smith, T., Stier, A., Braak, C.J.F.T., Weedon, J., 2022. vegan: Community Ecology Package [WWW Document]. URL https://cran.r-project.org/web/packages/vegan/index.html

Öpik, M., Vanatoa, A., Vanatoa, E., Moora, M., Davison, J., Kalwij, J.M., Reier, Ü., Zobel, M., 2010. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). New Phyt. 188, 223–241. https://doi.org/10.1111/j.1469-8137.2010.03334.x

Pereira, C.M.R., López-García, Á., Silva, D.K.A. da, Maia, L.C., Frøslev, T.G., Kjøller, R., Rosendahl, S., 2020. Tropical forest type influences community assembly processes in arbuscular mycorrhizal fungi. J. Biogeogr. 47, 434–444. https://doi.org/10.1111/jbi.13739

Qin, Z., Zhang, H., Feng, G., Christie, P., Zhang, J., Li, X., Gai, J., 2020. Soil phosphorus availability modifies the relationship between AM fungal diversity and mycorrhizal benefits to

maize in an agricultural soil. Soil Biol. Biochem. 144, 107790. https://doi.org/10.1016/j.soilbio.2020.107790

Ramana, J.V., Tylianakis, J.M., Ridgway, H.J., Dickie, I.A., 2023. Root diameter, host specificity and arbuscular mycorrhizal fungal community composition among native and exotic plant species. New Phyt. 239, 301–310. https://doi.org/10.1111/nph.18911

Rillig, M.C., Wright, S.F., Nichols, K.A., Schmidt, W.F., Torn, M.S., 2001. Large contribution of arbuscular mycorrhizal fungi to soil carbon pools in tropical forest soils. Plant Soil 233, 167–177. https://doi.org/10.1023/a:1010364221169

Rincón, C., Droh, G., Villard, L., Masclaux, F.G., N'guetta, A., Zeze, A., Sanders, I.R., 2021. Hierarchical spatial sampling reveals factors influencing arbuscular mycorrhizal fungus diversity in Côte d'Ivoire cocoa plantations. Mycorrhiza 31, 289–300. https://doi.org/10.1007/s00572-020-01019-w

Rodríguez-Echeverría, S., Teixeira, H., Correia, M., Timóteo, S., Heleno, R., Öpik, M., Moora, M., 2017. Arbuscular mycorrhizal fungi communities from tropical Africa reveal strong ecological structure. New Phyt. 213, 380–390. https://doi.org/10.1111/nph.14122

Rui, W., Mao, Z., Li, Z., 2022. The Roles of Phosphorus and Nitrogen Nutrient Transporters in the Arbuscular Mycorrhizal Symbiosis. Int. J. Mol. Sci. 23, 11027. https://doi.org/10.3390/ijms231911027

Schmidt, J.E., DuVal, A., Puig, A., Tempeleu, A., Crow, T., 2021. Interactive and Dynamic Effects of Rootstock and Rhizobiome on Scion Nutrition in Cacao Seedlings. Front. Agron. 3, 754646. https://doi.org/10.3389/fagro.2021.754646

Schroeder, J.W., Dobson, A., Mangan, S.A., Petticord, D.F., Herre, E.A., 2020. Mutualist and pathogen traits interact to affect plant community structure in a spatially explicit model. Nature Commun. 11, 2204. https://doi.org/10.1038/s41467-020-16047-5

Smith, S.E., Read, D., 2008. Mycorrhizal Symbiosis (Third Edition). Academic Press.

Song, F., Bai, F., Wang, J., Wu, L., Jiang, Y., Pan, Z., 2020. Influence of Citrus Scion/Rootstock Genotypes on Arbuscular Mycorrhizal Community Composition under Controlled Environment Condition. Plants 9, 901. https://doi.org/10.3390/plants9070901

Song, F., Pan, Z., Bai, F., An, J., Liu, J., Guo, W., Bisseling, T., Deng, X., Xiao, S., 2015. The Scion/Rootstock Genotypes and Habitats Affect Arbuscular Mycorrhizal Fungal Community in Citrus. Front Microbiol 6, 1372. https://doi.org/10.3389/fmicb.2015.01372

Sousa, N.M.F., Roy, J., Hempel, S., Rillig, M.C., Maia, L.C., 2022. Precipitation and temperature shape the biogeography of arbuscular mycorrhizal fungi across the Brazilian Caatinga. J. Biogeogr. 49, 1137–1150. https://doi.org/10.1111/jbi.14376

Sousa, N.M.F., Veresoglou, S.D., Oehl, F., Rillig, M.C., Maia, L.C., 2018. Predictors of Arbuscular Mycorrhizal Fungal Communities in the Brazilian Tropical Dry Forest. Microb. Ecol. 75, 447–458. https://doi.org/10.1007/s00248-017-1042-7

Sweeney, C.J., Vries, F.T. de, Dongen, B.E. van, Bardgett, R.D., 2021. Root traits explain rhizosphere fungal community composition among temperate grassland plant species. New Phyt. 229, 1492–1507. https://doi.org/10.1111/nph.16976

Thanni, B., Merckx, R., Bauw, P.D., Boeraeve, M., Peeters, G., Hauser, S., Honnay, O., 2022. Spatial variability and environmental drivers of cassava—arbuscular mycorrhiza fungi (AMF) associations across Southern Nigeria. Mycorrhiza 32, 1–13. https://doi.org/10.1007/s00572-021-01058-x

Trouvelot, A., Kough, V., Gianinazzi-Pearson, V., 1986. Mesure du taux de mycorhization VA d'un systeme radiculaire. Recherche de methods d'estimation ayant une signification fonctionnelle. In: Gianinzaai-Perason, V., Gianinazzi, S. (eds), Physiological and Genetical Aspects of Mycorrhizae, pp217-221. INRA, Paris.

Valgaz, A.P.F. de, Naranjo-Morán, J., Román, G.R., Oviedo-Anchundia, J., Torres, M.R., Barcos-Arias, M., 2022. Discovering the Diversity of Arbuscular Mycorrhizal Fungi Associated with Two Cultivation Practices of *Theobroma cacao*. Diversity 14, 651. https://doi.org/10.3390/d14080651

Veras, J.S. da N., Escobar, I.E.C., Mendes-Alvarenga, R.L., Santos, V.M. dos, Silva, D.K.A. da, Silva, J.A. da, Barros, M.F., Maia, L.C., 2024. Rainfall and soil properties drive the temporal dynamics of arbuscular mycorrhizal fungal assemblages in a seasonally dry tropical forest. Acta Oecol. 123, 104000. https://doi.org/10.1016/j.actao.2024.104000

Veras, J.S. da N., Mendes-Alvarenga, R.L., Rodrigues, G.R.L. de S., Escobar, I.E.C., Maia, L.C., 2025. Spatial and environmental factors influence AMF assemblages across the Brazilian Atlantic Forest - a biodiversity hotspot. Eur. J. For. Res. 144, 121–131. https://doi.org/10.1007/s10342-024-01745-2

Villarreal, J., Name, B., Barahona, L.A., 2015. Tecnicas analiticas del laboratorio de suelos. Instituto de Investigación Agropecuaria de Panamá. Panamá.

Vu, V.Q., Friendly, M., Tavadyan, A., 2024. ggbiplot: A Grammar of Graphics Implementation of Biplots [WWW Document]. URL https://cran.r-project.org/web/packages/ggbiplot/index.html

Wahdan, S.F.M., Reitz, T., Heintz-Buschart, A., Schädler, M., Roscher, C., Breitkreuz, C., Schnabel, B., Purahong, W., Buscot, F., 2021. Organic agricultural practice enhances arbuscular mycorrhizal symbiosis in correspondence to soil warming and altered precipitation patterns. Environ. Microbiol. 23, 6163–6176. https://doi.org/10.1111/1462-2920.15492

Wan, X., Chen, X., Huang, Z., Chen, H.Y.H., 2021. Contribution of root traits to variations in soil microbial biomass and community composition. Plant Soil 460, 483–495. https://doi.org/10.1007/s11104-020-04788-7

Wang, L., George, T.S., Feng, G., 2024. Concepts and consequences of the hyphosphere core microbiome for arbuscular mycorrhizal fungal fitness and function. New Phyt. 242, 1529–1533. https://doi.org/10.1111/nph.19396

Wang, Q., Ma, M., Jiang, X., Guan, D., Wei, D., Cao, F., Kang, Y., Chu, C., Wu, S., Li, J., 2020. Influence of 37 Years of Nitrogen and Phosphorus Fertilization on Composition of Rhizosphere Arbuscular Mycorrhizal Fungi Communities in Black Soil of Northeast China. Front. Microbiol. 11, 539669. https://doi.org/10.3389/fmicb.2020.539669

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wubet, T., Wei, M., Kottke, I., Oberwinkler, F., 2006. Two threatened coexisting indigenous conifer species in the dry Afromontane forests of Ethiopia are associated with distinct arbuscular mycorrhizal fungal communities. Botany 84, 1617–1627. https://doi.org/10.1139/b06-121

Yang, C., Teng, Z., Jin, Z., Ouyang, Q., Lv, L., Hou, X., Hussain, M., Zhu, Z., 2025. Structure and composition of arbuscular mycorrhizal fungal community associated with mango. Front. Plant Sci. 16, 1578936. https://doi.org/10.3389/fpls.2025.1578936

Zhang, J., Quan, C., Ma, L., Chu, G., Liu, Z., Tang, X., 2021. Plant community and soil properties drive arbuscular mycorrhizal fungal diversity: A case study in tropical forests. Soil Ecol. Lett. 3, 52–62. https://doi.org/10.1007/s42832-020-0049-z

Chapter 4. Arbuscular mycorrhizal fungal inocula differentially influence root traits and biomass in *Theobroma cacao*

Natalie Ferro-Lozano, Ray Hendriks, Oscar Valverde-Barrantes, Guusje Bonnema, Thomas W. Kuyper



Abstract

One of the most relevant mutualistic root-symbiotic microbial groups is the group of arbuscular mycorrhizal fungi (AMF). They prime plant immune response and provide an overall improved nutrient status by supplying phosphorus and water, often the limiting factors for photosynthesis. A new perspective valuable for plant crop development would be to breed for beneficial mutualistic symbiosis, for which it is important to understand how each partner modifies the structure that supports the symbiosis. To achieve this understanding, we evaluated selfed progenies of cacao clones for root traits relevant to the symbiosis, such as diameter (D), root tissue density (RTD), specific root length (SRL), and root N content (N), while evaluating the mycorrhizal responsiveness of the plant in terms of biomass production and mycorrhizal root colonization. Three cacao clones were previously selected from agroforestry systems for beneficial above-ground traits such as high fruit yield and bean quality. We grew selfed progenies from these clones with inocula from four different AMF communities and a non-inoculated control. These AMF communities were from different farms in Panamá, varying in elevation and proximity to the ocean as well as pH and soil texture, which influence mycorrhizal community composition. Regarding the aboveground biomass production, we found significant main effects from AMF and progenies treatments, and significant interaction between mycorrhizal communities and progenies. Root N and D were significantly different between progenies and mycorrhizal treatments, while RTD and SRL were only significantly different between mycorrhizal treatments, particularly mycorrhizal treatments and non-inoculated control. Our findings strengthen the value of considering root traits when developing new plant cultivars in association with AMF.

Introduction

Roots and their associated microbiome are known to have a critical role in plant nutrient acquisition and contribute to tolerance against abiotic and biotic stress (Hohmann and Messmer, 2017). In ecology, there have been advances in understanding the effects of root trait variation, and their interaction with abiotic factors and microbial communities. Despite this knowledge, root traits and the root-associated microbiome have been overlooked by traditional plant breeders who have preferred optimizing aboveground traits as optimal yield, flowering time, height, and resistance to pests and diseases. This lack of attention has been further aggravated by the logistical difficulties associated with root measurements and the intrinsic complexity of belowground processes. The challenge is to translate this understanding of root traits and their ecological interactions into actionable breeding strategies that optimize plant–microbiome partnerships.

Understanding root trait variation is a simple way to generalize ecosystem strategies over plant species (Weigelt et al., 2021). Building on Wright's (2004) proposal of a leaf economic spectrum, which highlights a one-dimensional gradient from acquisitive 'fast' plants to conservative 'slow' ones, it has become clear that roots traits are better classified in a two-dimensional expanded framework. The root economics space (RES) framework, devised by Bergmann et al. (2020) through a global meta-analysis, highlights a similar conservation gradient additional to a collaboration gradient. The conservation gradient associates negatively root tissue density (RTD) representing tissue investment and slow growth, versus nitrogen content (N), representing root metabolic activity and faster growth. The collaboration gradient considers the symbiosis with arbuscular mycorrhizal fungi (AMF), based on the negative correlation between root diameter (D) and specific root length (SRL). Root D is a proxy for root cortical area, which is the space where roots can harbor AMF communities, while a large SRL reflects the plants' economic investment in its own root exploration. The global meta-analysis revealed that plants that prioritize their root system development (do-it-yourself strategy) have a lower D and higher SRL. In contrast, plants with a more collaborative approach (an 'outsourcing' strategy') exhibit the opposite trend. This innovative framework provides a link between root trait variation and how much plants depend on their relationship with AMF.

AMF are key plant symbionts that can mitigate agricultural challenges and strengthen plant development. The mycorrhizal fungal mycelium extends the reach of plant roots, enhancing nutrient acquisition, particularly for phosphorus and immobile cations like zinc, copper, and iron. As well, mycorrhizal fungi also contribute to soil structure and stability, preventing erosion and enhancing water retention, vital for maintaining soil health under climate change (Kuyper and Jansa, 2023). Thus, these symbiotic relationships boost plant resilience to various stressors, such as drought and pathogens, reducing the need for chemical interventions (Frew et al., 2024). By fostering a more robust and resource-efficient plant-microbe partnership, either by plant breeding or management,

mycorrhizal symbiosis can contribute significantly to sustainable agriculture, promoting ecosystem health, reducing environmental footprint, and improving overall crop productivity in a manner that aligns with long-term ecological balance (Pascual et al., 2015; Soto-Navarro et al., 2021).

There are various approaches using AMF to enhance sustainable agriculture in tropical ecosystems, including steering microbial communities through management and breeding for mutualistic symbiosis. The first approach considers agricultural practices that support beneficial AMF communities, such as reduced use of fungicides and phosphorus-fertilizer application, minimal tillage, biodiversity maintenance, crop rotation, and organic amendments (Verbruggen et al., 2012; Qiming Wang et al., 2024). The second approach focuses on plant breeding for enhanced mycorrhizal mutualism. Selecting plants for mycorrhizal responsiveness under low-input systems has been proposed, though it can result in suboptimal growth under conditions of mycorrhizal inoculum limitation (Cobb et al., 2021; Kaeppler et al., 2000). The RES framework also helps identify root traits linked to AMF interactions. D and N tend to be more genetically conserved, while RTD and SRL vary in their plastic response towards the environment (Kramer-Walter et al., 2016; Ostonen et al., 2007; Valverde-Barrantes et al., 2017a). Nevertheless, breeding programs rarely incorporate plant responses to AMF, largely due to uncertainties. Plant benefits from mycorrhizal symbiosis depend on three factors: (i) plant species, characterized by root traits; (ii) environmental context, including soil fertility and water availability; and (iii) plant genotype and fungal species interaction (Watts-Williams et al., 2019). For cacao (*Theobroma cacao)*, an agroforestry crop of ecological and economic significance, integrating AMF responsiveness into breeding programs could represent a promising yet underexplored strategy to improve resilience and productivity.

Breeding for enhanced AMF symbiosis, particularly in cacao, could address its agroforestry cultivation challenges and improve productivity in smallholder systems. Approximately 90% of the cacao is produced by resource-poor smallholder farmers with low yields caused by diseases and pests. These challenges are augmented when cacao is managed organically to access better fair-trade prices (Cadby and Araki, 2021; Dragusanu et al., 2014). In addition, cacao's heterozygous, allogamous nature complicates breeding efforts, having only resulted in semi-domestication by the Olmec and Maya civilizations (present-day Mexico, Belize, and Guatemala) (Motamayor et al., 2002). Given cacao's strong dependence on mycorrhizal symbiosis (Cuenca et al., 1990; Lovera et al., 2022; Paguntalan et al., 2020), improving this interaction through farm management and breeding strategies presents a promising pathway to enhance plant health and productivity. Cacao studies assessing root traits have ignored differences among varieties, as well as how root trait variation may regulate mycorrhizal benefits. Also genotype interaction has generally been overlooked, by exploring plant genotype or fungal species independently. These are crucial aspects when considering breeding for mycorrhizal benefit.

This study aimed to look at root trait variation between cacao genotypes (by using selfed progenies), fungal community properties (different AMF treatments) and their interaction (by measuring plant vigor and development). Our hypotheses were (i) selfing progenies differ in root traits, which drive variation in AMF colonization and growth; (ii) AMF community identity alters root traits and growth; (iii) Growth is maximized by matching progeny traits favoring colonization and AMF communities that deliver high benefit.

Material & Methods

Plant Material

The parental genotypes used, AS-CP 26-59, AS-CP 26-60, and AS-CP 26-61 (in Chapter 2 as trees Genotype 5, Genotype 7 and Genotype 4, respectively), were previously selected from Panamanian wild cacao populations by breeders (Engr. Abiel Gutiérrez, pers. comm.) from the Panamanian Institute for Agricultural Innovation (IDIAP, by its acronym in Spanish). These were selected for high performance in yield, disease resistance, and quality, while belowground traits, including mycorrhizal symbiosis, had not been assessed before. The parental genotypes were grafted in the experimental garden of IDIAP located in Almirante, Bocas del Toro, Panamá (9° 17' 47" N, 82° 24' 30" W). The levels of heterozygosity (H_o) of AS-CP 26-59, AS-CP 26-60, and AS-CP 26-61 were 3.13%, 28.65% and 34.38% based on profiles of 96 biallelic selected SNPs in Chapter 2. Each genotype has a different genetic background (Lanaud et al., 2024), as analyzed previously by STRUCTURE in Chapter 2. AS-CP 26-59 is composed dominantly by Amelonado; AS-CP 26-60 is composed by a mixture between Amelonado, IMC (Iquitos) and Nanay, and lastly, AS-CP 26-61 is composed by Amelonado, Nanay, and Nacional (Supplementary material A). The heterozygosity and origin composition are particularly important as to what level of variation can be expected within and between selfed progenies.

Growth Conditions

Flowers of each genotype were self-fertilized (November 15th, 2021), resulting in genotype-derived selfed progenies AS-CP 26-59, AS-CP 26-60, and AS-CP 26-61 (hereafter respectively: Progenies S_1-59, S_1-60, S_1-61). Five months after self-fertilizing the flowers of genotypes AS-CP 26-59, AS-CP 26-60, and AS-CP 26-61 (April 15th, 2022), the resulting cacao pods were harvested. Seeds were immediately prepared for germination by removing the mucilaginous layer with sterile sawdust, disinfected with 0.5% sodium hypochlorite for 3 min, and rinsed with sterile water. Seeds were then left to germinate for 21 days in growth chambers in a sterilized substrate composed of one part river sand and two parts dark soil rich in organic matter (autoclaved twice with two days in between (Habte and Osorio, 2001). Nine days later, the seedlings were transplanted into individual

5-L plastic pots (h: 20 cm; ⌀: 22 cm) and moved to the greenhouse for continuous growth over six months. The pots were filled with pasteurized substrate after steaming at 80°C for two hours before cooling down overnight, turning it over, and steaming again at 80°C for two hours. Plastic pots were filled with 4.5 L of this pasteurized substrate, and 500 mL (10% of total soil) of soil inocula containing mycorrhizal live mycelium or pasteurized soil for the control treatment (CT) was added in direct contact with seedlings. Control pots were watered with 100 mL of microbial wash and the AMF inoculated pots were watered with a similar amount of tap water. Pots were left without irrigation for the next ten days to ensure the establishment and prevent the washing out of the AMF and other microbes. For the rest of the growing months, the plants were watered similarly, but water quantities varied slightly compensating for higher transpiration of bigger plants.

For the AMF treatments, we used AMF-containing soils from two farms in Bocas del Toro (B4, B6) and two farms in Herrera (H3, H7), next to the non-inoculated control (Supplementary material B, Table S1). We took soil inoculum from the upper 20 cm of soil, one meter distant from the sampled cacao trees. To ensure variation between AMF inocula, soils were selected from farms that differed considerably in pH and soil texture, as these characteristics are major factors determining mycorrhizal community composition (Alguacil et al., 2016; Jansa et al., 2014). Soils were stored in plastic containers at 4ºC to minimize the growth of other microorganisms (Habte & Osorio, 2001) until inoculating plants. We made a microbial wash of the four soils to ensure the differences in plant growth observed are only due to the mycorrhizal and not bacterial communities (Wang et al., 2018). We separately soaked 250 ml of the four mycorrhizal soils in 2 L filtered water for two hours. The suspensions were then filtered by two-layers coffee filter paper and a 50 µm sieve, ensuring that most AMF spores remained behind, while bacteria could pass through. We mixed the four filtrates to obtain a homogeneous microbial wash of all soils combined, used for the control treatment. Although it is possible for spores smaller than 50 µm to be present, there was a severe limitation in the control treatments (CT) for mycorrhizal symbiosis. The pasteurization treatment may equally not have been 100% successful, resulting in some mycorrhizal inoculum in the control treatment.

Seedlings were grown on tables in a greenhouse at the Smithsonian Tropical Research Institute in Gamboa (9°07'10.8"N 79°42'06.0"W). The experiment was conducted with a randomized complete block design (RCBD) with two factors (three selfed progenies, five mycorrhizal treatments) and 25 replicates. Individual plants were randomly allocated to a position within the block and randomly rearranged within the block every four weeks.

Seedling growth and trait assessment

After six months, the seedlings were harvested by separating roots, leaves, and stems. Leaves of individual seedlings were photographed to analyze leaf area using ImageJ. Roots were thoroughly

washed to remove all soil. Thereafter, all lateral roots were separated from the main root. Then, leaves, stems, and a subsample of roots were dried separately for three days at 60 °C before weighing to determine the dry weight of leaves (LDW) and stems (SDW) which together yield aboveground dry weight (ADW). A subsample of these fine roots was wrapped in a wet paper towel and stored at 5°C until further use. Within every treatment, we selected the five plants with the highest aboveground biomass and the five with the lowest for root trait imaging analysis, resulting in an analysis of 10 out of the 25 plants per treatment combination. Around three fine-root fragments from these plants under 1 mm diameter, and representing first-order, second-order and third-order fine roots were submerged in a shallow layer of water in a transparent tray (120 mm by 120 mm by 17 mm) and carefully separated with wooden chopsticks, so they did not touch. These were then scanned under high resolution (Epson Perfection V850 Pro). After scanning, fine roots and the rest of the (coarse) roots were dried separately for three days at 60°C and subsequently weighed to determine root dry weight of scanned roots (sRDW) and total root dry weight (RDW). Root scans were analyzed using WinRhizo software (Regent Instruments, 2000) to determine average root diameter (D), specific root length (SRL), and root tissue density (RTD). WinRhizo automatically generated root length (RL) and root diameter (D). SRL was calculated by dividing the total scanned RL by the root dry weight (sRDW), and RTD was calculated, treating roots as cylinders, using equation 1 relating to diameter and SRL. Additionally, root N content was analyzed through the Kjeldahl method.

$$Equation\ 1: \quad RTD = 4\ /\ (SRL\ x\ D^2\ x\ \pi)$$

Mycorrhizal Colonization

The dried fine roots of five random plants per treatment were cut into 1.5 cm root fragments, placed in tissue cassettes, and hydrated overnight in distilled water before staining. The staining procedure used sequential solutions of (1) KOH 10% for 16 hrs. at 45°C, (2) H_2O_2 3% for 30 min at room temperature, (3) HCl 1M for 15 min at room temperature, and lastly (4) 0.05% trypan blue in lactoglycerol for 45 min at 45°C. In between the solutions, the samples were washed thrice with distilled water. All samples were placed in lactoglycerol for 3-4 days at room temperature to eliminate the excess of trypan blue. After this, we analyzed roots in microscope slides for mycorrhizal colonization (MC) through the Trouvelot method (Trouvelot et al., 1986). Here we report the percentage of root length colonized by mycorrhizal fungi (M%).

Statistical Analysis

For above and belowground biomass and root traits, we tested for interaction between mycorrhizal treatment and genotype progenies using a type III analysis of variance (ANOVA). When interaction was insignificant, we used a type II ANOVA and Tukey's honestly significant difference (HSD) test (Supplementary material C). Block effects were considered and found significant. We

performed bivariate correlations (Supplementary material D) and principal component analysis (PCA) to explore the relationships among root traits, grouped by AMF treatments (including or not the control) and progenies. We conducted two separate PCAs, with the first one excluding the control (as the focus was on possible differential performance of different AMF inocula) and the second one including it. As cacao plants are dependent on the AMF symbiosis and probably never occur in the field under non-mycorrhizal conditions, we report the first PCA in the main Results and the second PCA in Supplementary Material E. An additional PCA with the above and belowground biomass was done to observe the differences among AMF treatments (Supplementary material F). We used boxplots to illustrate the main differences between biomass and root traits (SRL, D, RTD, and N) using mycorrhizal treatments and progenies as main factors.

All statistical analyses were performed using RStudio Version 2024.04.2+764 and R ver. 4.4.1 (www.r-project.org). The plots were created with the R package ggplot2 (Wickham, 2016). The AIC model was performed with the R package AICcmodavg (Mazerolle, 2023). Tukey's HSD test was performed with the R package multcompView (Graves et al., 2015). PCA was performed with the R package FactoMineR (Lê et al., 2008) and their visualizations with the R package ggbiplot (Vu et al., 2024). Bivariate correlations were calculated with the R package Hmisc (Harrell, 2023) and visualized through the R package PerformanceAnalytics (Peterson, 2004).

Results

Mycorrhizal Colonization

The control was not completely free of mycorrhizal colonization, but colonization percentage was significantly lower than in the mycorrhizal treatments (*p-value* < 0.0001 – Figure 1 & Supplementary material C). There was a significant difference between AMF treatments without including the CT (*p-value* = 0.002). Within progenies, S1-59 and S_1-61 did not present significant differences in mycorrhizal colonization between AMF treatments, but these were present in S_1-60, suggesting the latter progeny has a variable response to AMF percentage of colonization.

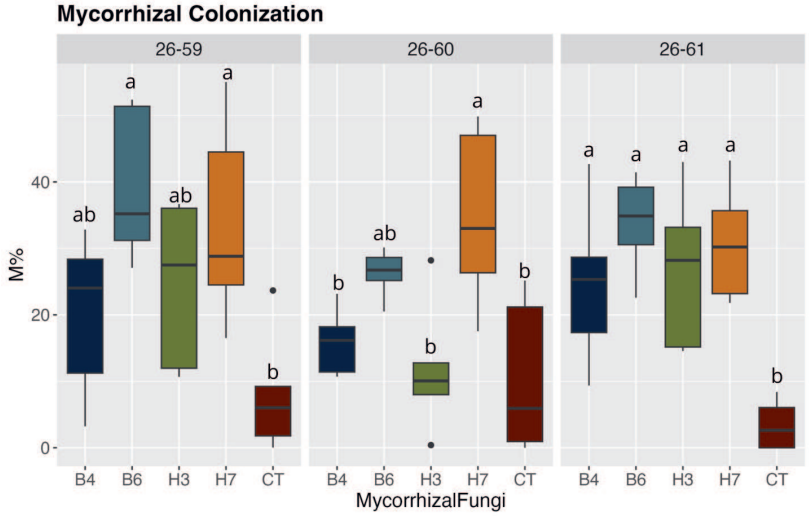


Figure 1 Fractional *mycorrhizal root colonization across three different plant progenies in function of five mycorrhizal treatments. For every progeny, boxes sharing the same letter are not significantly different (P > 0.05), as determined by Tukey's test.*

Above and belowground biomass

Above and belowground biomass of the cacao seedlings inoculated with different AMF communities were significantly different across the selfed progenies (S_1s) and AMF treatments (Figure 2 – Supplementary Material C). Control treatments exhibited significantly lower biomass than mycorrhizal treatments, both aboveground (~150%) and belowground (~77%). Progeny S_1-59 had significantly lower biomass (p=0.03) than S_1-60 and S_1-61. A significant interaction between AMF and progenies (p=0.04) was observed. Plants exposed to the H3 mycorrhizal communities were significantly larger in S_1-60 than in S_1-59 and S_1-61.

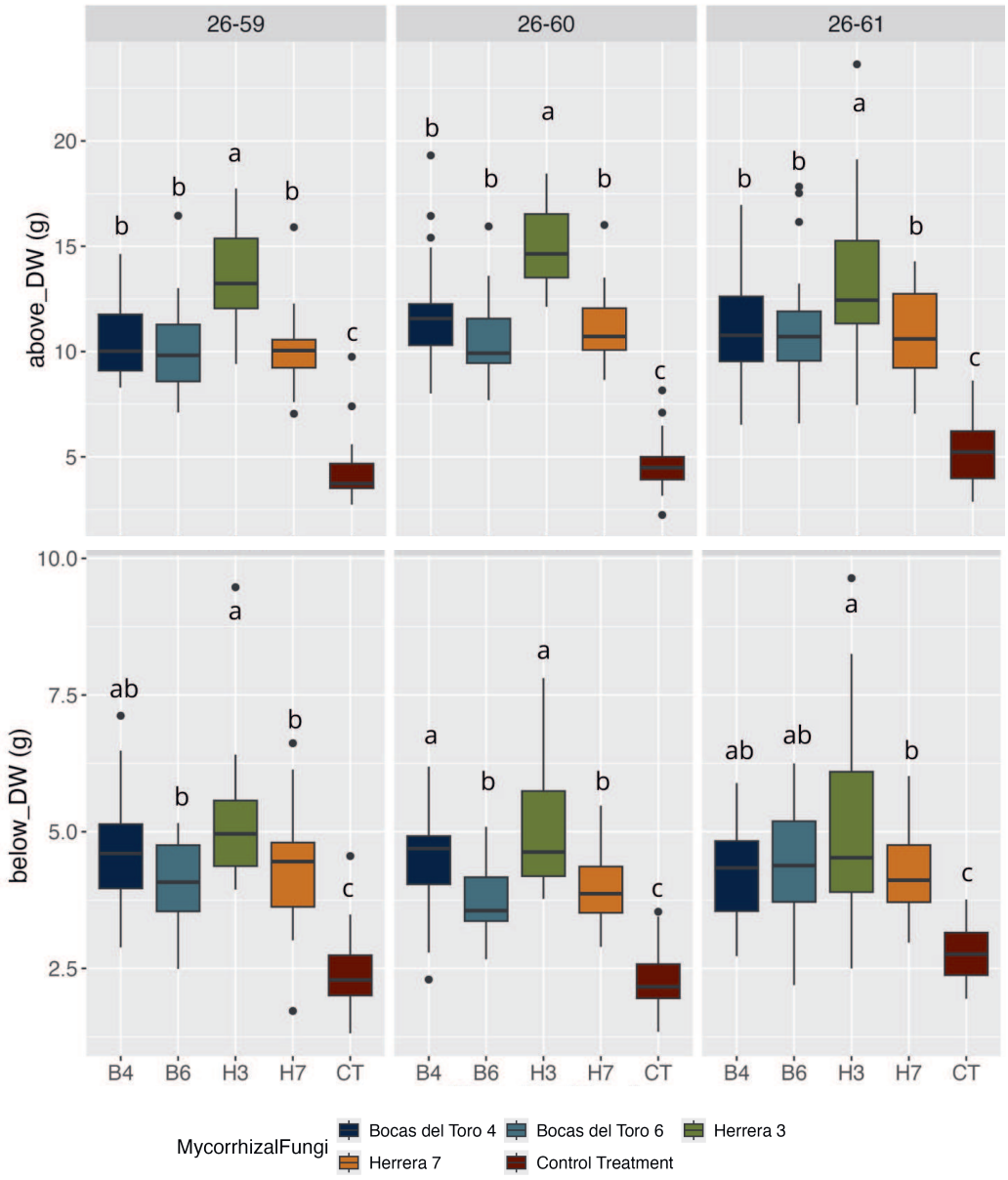


Figure 2 Above and Belowground biomass across three different plant progenies in function of five mycorrhizal treatments. For every genotype, boxes sharing the same letter are not significantly different (P > 0.05), as determined by Tukey's test.

Root Traits

The root nitrogen concentration (N – Figure 3a) differed significantly between different mycorrhizal treatments, but not between the inoculated and non-inoculated treatments. Root N content averaged over all three S_1s was significantly larger in treatment H3 than that in B4, B6, and CT. Inoculated plants had significantly higher RTD and significantly lower SRL than the control

plants (Figure 3b – 3d). In S_1-59 and S_1-60, RTD of mycorrhizal plants were significantly greater than of the CT, but in S1-61, only treatment H3 had significantly larger RTD than CT. Root diameter (D – Figure 3c) in treatment H3 and H7 in S_1-60 and treatment H7 in S1-61 was significantly greater than CT. Intraspecific variation, as expressed through the coefficient of variation (CV), in D was the lowest among the four root traits (CV = 8%), while SRL was the highest (CV = 29%). Intermediate values were observed for RTD (CV = 20%) and Root N (CV = 14%).

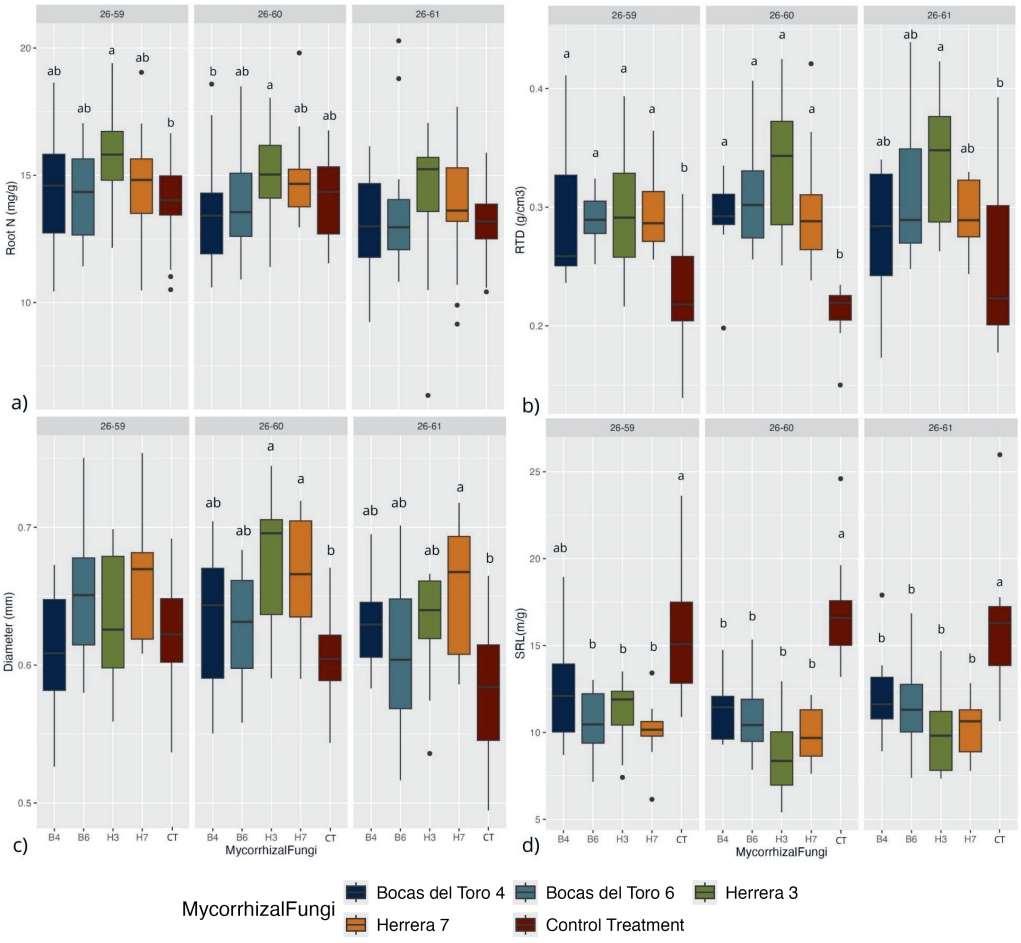


Figure 3. Root traits across treatments. (a) Root N (b) Root Tissue Density (RTD) (c) Diameter (D) (d) Specific Root Length (SRL). Boxes sharing letters are not significantly different, as determined by Tukey's test within progenies.

Correlations between Root traits, Plant Biomass, and Mycorrhizal Colonization

A summary of these correlations is provided in Figure 4. The only significant effect associated with MC was with RTD (r=-0.51*). Root diameter positively correlated with aboveground ($p < 0.05$) and belowground biomass ($p < 0.01$); root N correlated negatively with aboveground biomass, but positively with belowground biomass. Additionally, SRL showed a significant but small negative

correlation with belowground biomass (r=-0.17). There were significant negative correlations between root D and SRL (-0.65***) and between SRL and RTD (r=-0.67***), whereas root N was not significantly correlated with the three other root traits. The principal components explained 46.7% in PC1 while PC2 and PC3 were relatively equal, 27.6% and 24.1% respectively (Figure 5 – Supplementary Material E - Table S4). PC1 vs. PC2 partly mimicked the collaboration gradient (strong negative correlation between SRL and D). However, on PC1 vs. PC3 the collaboration gradient was more clearly indicated. In both cases there was a significant negative correlation between RTD and SRL. The conservation gradient (RTD vs. N) was only present in PC1 vs. PC3.

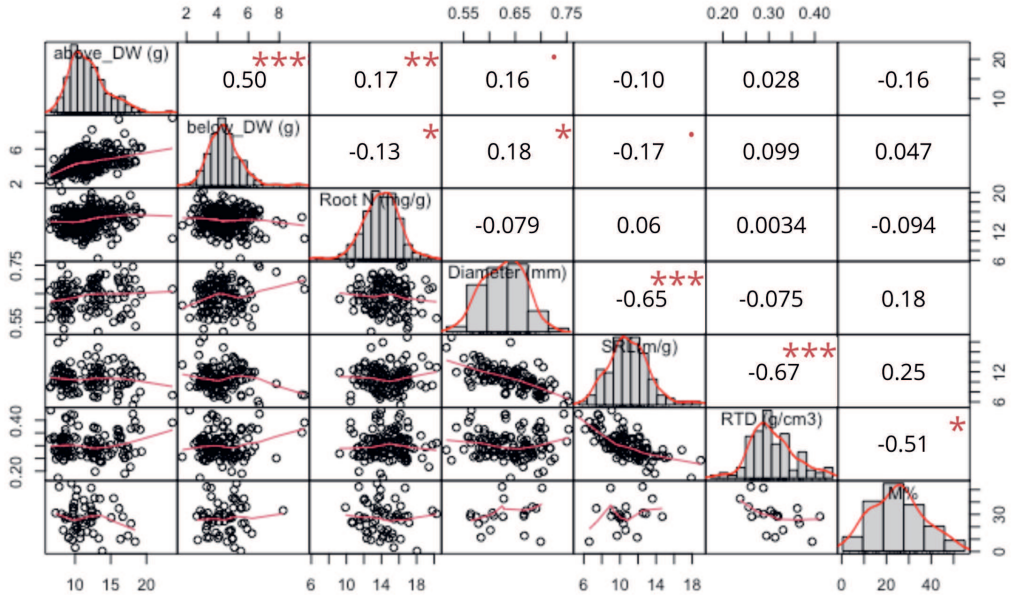


*Figure 4. Correlations considering only mycorrhizal treatments. A significance of p = 0.0000 is denoted by '***', p<0.001 by '**', p<0.01 by '*', and p<0.05 by '·'.*

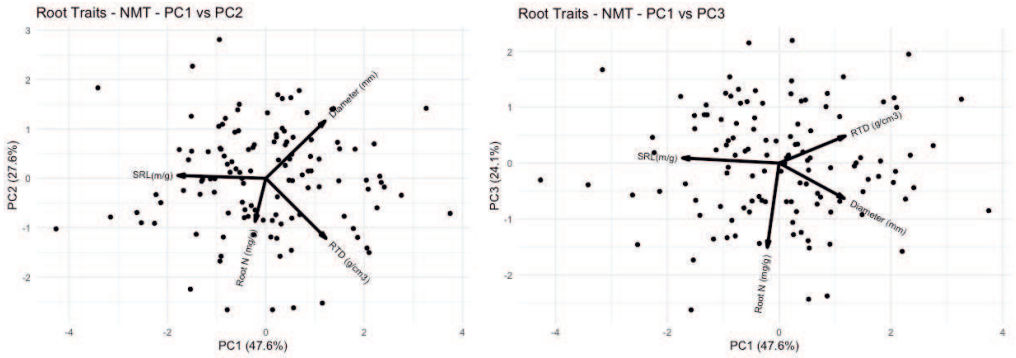


Figure 5. Principal component analysis for root traits excluding CT (a) PC1 vs PC2 (b) PC1 vs PC3.

Discussion

Our study showed that (i) selfed progenies (S_1s) from various genotypes vary in root traits and growth performance; (ii) there is a strong mycorrhizal effect on biomass and some root traits; (iii) AMF communities differ in their influence on root traits and plant performance; (iv) there is no significant relation between D and M% (but still a positive correlation between D and above and belowground biomass); (v) the best plant growth response (S_1-60 H3) was observed in progeny with root traits that favored AMF colonization, and with an AMF community that generally provided the largest growth enhancement; and (vi) the intraspecific RESpace differed from the interspecific RESpace (Bergmann et al., 2020).

Cacao genetic variation and plant performance

Plant genetic variation is crucial to variation in their mycorrhizal response and benefit, as much of this response depends on plant genotype and fungal species interaction. In this experiment, we accounted for the genetic background of the plants, which is often overlooked in studies evaluating variation in mycorrhizal response. Developing seed-progenies was essential to adequately phenotype root traits and ensure homogeneous and natural root development across treatments. Given that cacao is a highly heterozygous and allogamous species, the absence of homozygous cacao lines precludes the use of seed-derived genetic controls. By addressing genotypic variation, our approach provides a more robust framework for understanding mycorrhizal plant response.

The parental genotypes differ in their genetic background and their levels of heterozygosity, so we expected both inter and intra variation in the segregation across the seed-derived progenies. The differences in heterozygosity were not associated to a phenotypic variation contrast between progenies, as variation between S1-59 (parental genotype 3% heterozygosity) plants was similar to that of the other two progenies, indicating low heritability for the studied traits. Only aboveground biomass and root N showed significant differences between S_1s ($p<0.05$), where S_1-59 progeny had significantly lower biomass development, and S_1-61 had significantly lower root N content in contrast with the other two. The lower biomass of S_1-59 could indicate a degree of inbreeding depression provided its low heterozygosity. Root diameter showed marginally significant differences ($p= 0.09$), where S_1-60 and S_1-61 were significantly different. This is consistent with a study by Valverde-Barrantes et al. (2017a) who showed that root N and diameter are root traits more genetically influenced than SRL and RTD.

A few advantages of using these parents were (i) their optimal traits in cacao growth (yield, quality, resistance), (ii) their wild origin, which conferred a naturally selected adaptation for enhanced performance within agroforestry systems, and (iii) their diverse genetic backgrounds. An agroforestry system involves intercropping techniques and low-input organic management, both practices that

facilitate AMF symbiosis (Cardoso et al., 2017). Selection for cacao with augmented performance in a low-input agroforestry system could explain the similar effects of growth performance and root traits in response to mycorrhizal treatments. Breeding cacao for improved mycorrhizal symbiosis in such systems could be particularly advantageous, as AMF contributes to efficient nutrient acquisition and pathogen protection, reducing reliance on synthetic fertilizers and agrochemicals. Management practices that promote mycorrhizal symbiosis, such as reduced tillage, minimal use of phosphorus fertilizers and fungicides, and incorporating crop rotations, can further enhance these benefits (Verbruggen and Kiers, 2010).

Mycorrhizal versus control treatments effects

Cacao tree seedlings depend on mycorrhizal association for growth, nutrition, and resistance to biotic and abiotic stress (Lovera et al., 2022). This dependence could explain the overall significant impact of AMF treatments on biomass development. Moreover, inoculation with AMF influenced RTD and SRL. AMF treatments showed increased RTD in contrast to the non-inoculated control. A possible explanation is that mycorrhizal symbiosis enhances overall biomass quality, and structural integrity through the lignification of the roots (Hu-zhe et al., 2005), thus causing a higher RTD (and lower SRL). When excluding the control treatment, the correlation between M% and RTD became negative, showing that among mycorrhizal treatments, higher fractional colonization increases the cortex fraction in comparison with the stele. Furthermore, a higher SRL is associated with plants developing their root system for their own nutrient acquisition (do-it-yourself) instead of investing in an outsourcing strategy through a mycorrhizal fungal partner (Bergmann et al., 2020).

AMF communities differ on how root traits and biomass respond

The high source of variation coming from the mycorrhizal treatments could be due to quantitative differences in the amount of mycorrhiza (inoculum potential) or qualitative differences (AMF community composition). It is unlikely that the inoculum potential influenced the results, as it was assumed to be consistent across all treatments, and no correlation was observed between fractional mycorrhizal colonization and plant performance. Alternatively, qualitative differences are shown by differential responses to AMF inocula, highlighting the potential benefit of a particular AMF community in association with cacao. This interaction in the symbiosis changes the magnitude of the benefit received by the plant regarding availability of nutrients and water, as well as plant resilience to (a)biotic stress (Kuyper et al., 2021) influencing biomass (M. Qin et al., 2020; Wu et al., 2022). Community composition data from the collected sites (Chapter 3) and pots, in case the seedlings selectively promote certain AMF species, will be available soon (Ferro Lozano et al., in prep). Among the mycorrhizal treatments, the H3 inoculum resulted in the highest biomass, with interactive significant effects (Supplementary material C). Progeny S_1-60 showed the highest growth

performance, with the highest D, the lowest SRL but also the lowest MC. This suggests that breeding cacao for high D would increase mycorrhizal benefit, as also found in Wang et al. (2023), and also implies that for cacao performance, AMF community composition is more relevant than MC. Root N and root D are differentially influenced by mycorrhizal treatments. Nitrogen content in the roots is typically used as a proxy for nutrient acquisition and demand (Reich et al., 2008). This could be related to mycorrhizal species in the H3 treatment that positively influenced N acquisition strategy of the root system for nitrogen, resulting in a higher biomass development in the plant. For root D, its variation across treatments was the lowest, suggesting some genetic limitation, possibly under high environmental selective pressure due to mycorrhizal dependency, according to (Janos, 2007).

Root Economics Space (RES) Framework

Regardless of whether non-inoculated control is included, the RES framework from this data set differs from (Bergmann et al., 2020) as it showed a stronger negative correlation between SRL and RTD than between SRL and D. We also did not observe a significant correlation between root N and any other root trait. The primary reason for the differences between the interspecific RES framework and the observed tradeoffs observed in cacao is lack of information about intraspecific variation in root traits. Bergmann et al. (2020) used mean values of root traits for each species, under the assumption that interspecific variation is more important than intraspecific variation. While that assumption may be correct when dealing with the full phylogenetic diversity of plants, one should not forget that there can also be a relatively high intraspecific versus interspecific variation for root chemical traits and SRL (Westerband et al., 2021) associated with complex plastic responses of the root systems to the soil environmental changes.

In the RES framework, there is a strong negative correlation between D and SRL, and smaller negative correlations between RTD and SRL, and RTD and D. (Ma et al., 2018) found that in thin-rooted species, such as cacao (reported by authors for global reviews (Bergmann et al., 2020; Guerrero-Ramírez et al., 2021)), low RTD can increase root exploration (SRL). As the root D or cortex area expands to accommodate AMF, RTD can decrease, because the cortex has a lower tissue density than the stele. Because root D was hardly variable, variation in SRL was more likely driven by RTD. As shown by Equation 1, SRL negatively correlates with both D and RTD; therefore, with strong phylogenetic constraints on D, a stronger (negative) correlation between SRL and RTD than between SRL and D is to be expected In this dataset the intraspecific variation in D (CV = 8%) was the lowest among the four root traits evaluated, while SRL was the highest (CV = 29%), followed by RTD (CV = 20%). This is different from Bergmann where, considering mean values for each species, D shows 10-fold variation versus RTD with a 3-5-fold variation. Since root D and N exhibit a higher phylogenetic conserved signal and SRL and RTD are more environmentally plastic traits (Ostonen et

al., 2007; Valverde-Barrantes et al., 2017a), it follows that in this study SRL and RTD are more responsive to presence of AMF as part of their environmental interaction, compared with root N and D.

Mycorrhizal symbiosis is a ubiquitous partnership among plants. However, the interaction between plant genotypes and the AMF community composition largely affects the beneficial mycorrhizal response of the plant (Berger and Gutjahr, 2021). This highlights the value of intraspecific variation in root traits regulating the mycorrhizal symbiosis. Although intraspecific variation can dramatically influence the soil microbiome composition (Roches et al., 2018), the role of intraspecific variation is underestimated in ecosystem functioning, especially when altering the interactions of microbial communities (Stahlhut et al., 2023)

Conclusions

It is crucial to understand through which mechanisms plants and microbial partners influence each other if we aim to explore the potential use of soil microbial communities in agriculture and (agro-)forestry. This study highlights the overall dependence of cacao plants on the mycorrhizal symbiosis and, moreover, the degree of interactions between cacao genotypes and mycorrhizal fungal communities. The H3 AMF treatment could have a species composition that favors cacao biomass development under our experimental conditions. These results show that mycorrhizal communities can provide vigorous, faster-developing tree seedlings. This interaction, mediated by the root system, shows an intricated plastic root response (SRL and RTD) and more genetically conserved (N and D) responses. While SRL seems negatively correlated to mycorrhizal growth response, RTD and D show positive correlations. The best growth response is S_1-60 combined with H3 showing a higher RTD, lower SRL, and higher D, but also with lower mycorrhizal colonization. Future research should focus on (a) identifying optimal combinations of environmental conditions and mycorrhizal partnerships, as can be essential for advancing soil management in sustainable agriculture, and (b) the mechanisms regulating the symbiosis between the mycorrhizal plant and its mycorrhizal fungal partner as it could greatly expand the potential of mycorrhizal symbiosis in plant breeding.

Acknowledgements

Engineers and workers at Agricultural Research Institute of Panama (IDIAP by its Spanish acronym), especially Dir. Ismael Camargo and Engr. Abiel Gutierrez for their collaboration with the used cacao genotypes and crosses. Isis Ochoa and Rivieth Castillo from STRI for all their help in the greenhouse and lab logistics. Dr. Luis C. Mejia for all his support and guidance.

Supplementary material

A. Plant material origins

For each of the parental accessions used, the percentages of their genetic origins are shown.

Genetic Groups	Accession AS-CP 26-59	Accession AS-CP 26-60	Accession AS-CP 26-61
IMC (Iquitos)	0.211	0.000	0.014
Nacional	0.004	0.001	0.196
Nanay	0.141	0.001	0.234
Guiana	0.010	0.001	0.088
Purus	0.004	0.000	0.002
LCT EEN (Curaray)	0.068	0.001	0.049
PA (Maranon)	0.004	0.001	0.006
Amelonado	0.491	0.994	0.307
SCA (Contamana)	0.007	0.001	0.018
Criollo	0.062	0.001	0.089

B. Map of Panama with Soil location can physicochemical characteristics

Red circled dots denote the locations of field sampling for mycorrhizal communities. These are in the Caribbean and Pacific sides, in Bocas del Toro and Herrera, respectively.

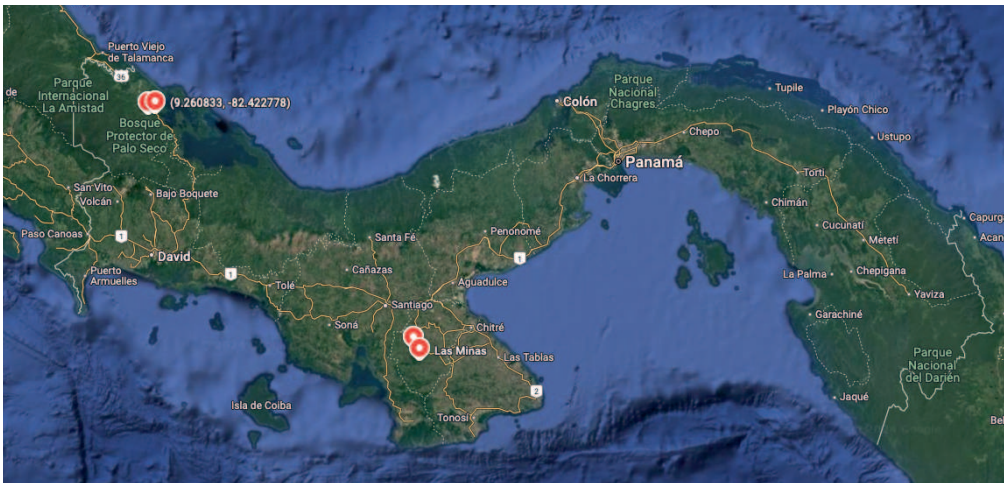


Table S4. Plantation's geographic location, elevation, and differential soil properties influencing mycorrhizal composition.

Treatment	B4	B6	H3	H7
Town and Region	Rio Oeste Arriba, Almirante, Bocas del Toro, Panama	Nuduobidi, Valle Risco, Bocas del Toro, Panama	Las Minas, Las Minas, Herrera, Panama	Tres Puntas, Las Minas, Herrera, Panama
GPS	9°15'39 N, 82°25'22" W	9°15'21" N, 82°28'6" W	7°47"51" N, 80°45'3" W	7°47'53" N, 80°48'2" W
Elevation	35 masl	81 masl	338 masl	600 masl
pH	5.68	4.84	6.48	5.42
Sand-Silt-Clay	68.8-24-7.2	48-30.4-21.6	46.4-31.2-22.4	52.8-26.4-20.8

C. Across treatments

ANOVA results

All the ANOVA tables include the non-inoculated control except for fractional mycorrhizal colonization, for which there are two tables, including and excluding control. For the aboveground ANOVA biomass, there was an interaction between AMF and progenies ($p < 0.04$), for which a type III ANOVA was kept, including the block effect, which was statistically significant. For all other response variables, a type II ANOVA was used, since there were no significant interaction effects.

Parameters	Units	Progenies		Mycorrhizal fungi		Table		Progenies x Mycorrhiza	
Including CT		*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Aboveground dry weight	g	3.64	0.02733*	102.28	<2e-16***	12.28	<2e-16***	2.0811	0.03711*
Belowground dry weight	g	2.31	0.1	102.523	<2e-16***	6.297	2.36e-16***		
Root N content	mg/g	10.7	2.98e-05***	8.78	9.35e-07***	3.28	7.09e-07***		
Root tissue density	g/cm3	1.2	0.305	16.683	7.02e-11***	1.398	0.123		
SRL	m/g	0.55	0.576567	39.404	<2e-16***	2.815	0.000115***		
Root diameter	mm	2.41	0.0939·	6.851	5.36e-05***	1.402	0.1207		
Mycorrhizal colonization	%	2.1	0.134	13.443	3.45e-07***	0.905	0.597		
Root:shoot ratio		15.9	2.5e-07***	69.92	<2e-16***	11.26	<2e-16***		

Response	MC (%)				
Excluding CT	Df	SS	Mean Sq.	F	p - value
Progenies	2	662	331	2.54	0.09628·
Mycorrhizal Fungi	3	2351	783.6	6.013	0.00258**
Table	25	2737	109.5	0.84	0.66876
Residuals	29	3779	130.3		

Averages with Tukey's test

Values sharing the same letter are not significantly different as determined by Tukey's test. When letters are missing, there were no significant differences shown.

Table S5. Mycorrhizal treatment differences.

Mycorrhizal Fungi	above DW (g)	below DW (g)	Root: Shoot	Root N (mg/g)	Diameter (mm)	SRL (m/g)	RTD (g/cm3)	M%	SLA (cm^2/g)
B4	11.1 (b)	4.4 (b)	0.40 (c)	13.6 (b)	0.623 (bc)	11.9 (b)	0.286 (b)	20.2 (b)	38.7 (a)
B6	10.6 (b)	4.0 (b)	0.39 (bc)	13.9 (b)	0.630 (bc)	10.9 (bc)	0.303 (ab)	33.1 (a)	38.4 (a)
H3	13.9 (a)	5.0 (a)	0.37 (b)	15.0 (a)	0.646 (ab)	9.9 (c)	0.324 (a)	21.1 (b)	38.5 (a)
H7	10.6 (b)	4.2 (b)	0.40 (bc)	14.4 (ab)	0.662 (a)	10.0 (c)	0.296 (ab)	33.1 (a)	38.1 (a)
CT	4.7 (c)	2.5 (c)	0.55 (a)	13.8 (b)	0.603 (c)	16.2 (a)	0.228 (c)	7.8 (c)	36.8 (b)

Table S6. Progeny differences.

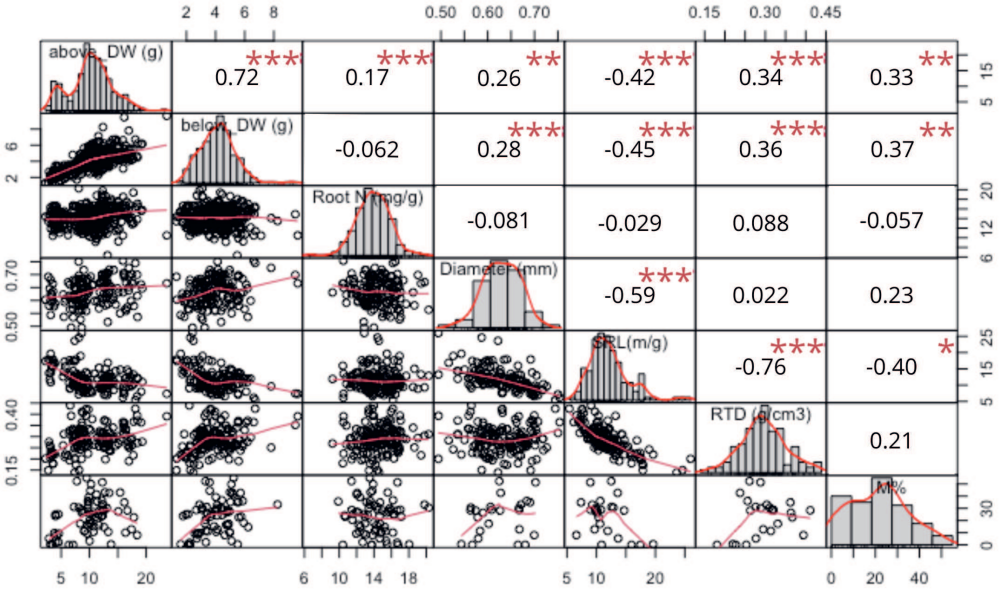
Progenies	above DW (g)	below DW (g)	Root: Shoot	Root N (mg/g)	Diameter (mm)	SRL (m/g)	RTD (g/cm3)	M%	SLA (cm2/g)
26-59	9.7 (b)	4.1	0.451	14.6 (a)	0.636 (ab)	11.9	0.279	25.1	383.9 (a)
26-60	10.6 (a)	3.9	0.396	14.3 (a)	0.641 (a)	11.5	0.289	19.4	367.9 (b)
26-61	10.3 (a)	4.1	0.427	13.6 (b)	0.621 (b)	12.0	0.294	24.7	391.5 (a)

Table S7. Interaction effects. P-values for aboveground effects in H3 treatment are lower than 0.05. For Root N, Diameter and SRL, p-values are lower than 0.1.

Mycorrhizal Fungi	Progenies	above DW (g)	below DW (g)	Root: Shoot	Root N (mg/g)	Diameter (mm)	SRL (m/g)	RTD (g/cm3)	M%	SLA (cm^2/g)
B4	26-59	10.6	4.6	0.437	14.3	0.609	12.4	0.291	19.9	39.2
	26-60	11.8	4.5	0.390	13.4	0.630	11.3	0.292	15.9	37.4
	26-61	11.0	4.2	0.393	13.0	0.631	12.2	0.275	24.6	39.4
B6	26-59	10.1	4.0	0.409	14.2	0.653	10.5	0.290	39.4	38.4
	26-60	10.5	3.7	0.369	13.9	0.626	10.9	0.307	26.2	37.8
	26-61	11.0	4.3	0.399	13.4	0.612	11.4	0.311	33.7	38.9
H3	26-59	13.4 (b)	5.1	0.394	15.6 (a)	0.631(ab)	11.2 (a)	0.297	24.5	38.3
	26-60	15.0 (a)	5.0	0.341	15.1 (ab)	0.677 (a)	8.8 (b)	0.335	11.8	36.9
	26-61	13.4 (b)	4.9	0.374	14.4 (b)	0.629 (b)	9.9 (ab)	0.339	26.8	40.3
H7	26-59	10.1	4.3	0.434	14.6	0.663	10.0	0.296	33.8	38.8
	26-60	11.2	3.9	0.360	14.8	0.667	9.8	0.299	34.7	36.0
	26-61	10.7	4.2	0.405	13.9	0.656	10.2	0.294	30.8	39.6
NC	26-59	4.3	2.4	0.583	14.1	0.625	15.5	0.222	8.1	37.1
	26-60	4.6	2.3	0.520	14.0	0.603	17.0	0.211	10.4	35.8
	26-61	5.2	2.7	0.564	13.1	0.579	16.3	0.250	3.4	37.4

D. Correlations including control treatments (CT)

There is a strong positive correlation of biomass with mycorrhizal colonization (0.33** for above and 0.37** for belowground), showing the strong response to mycorrhiza by cacao.



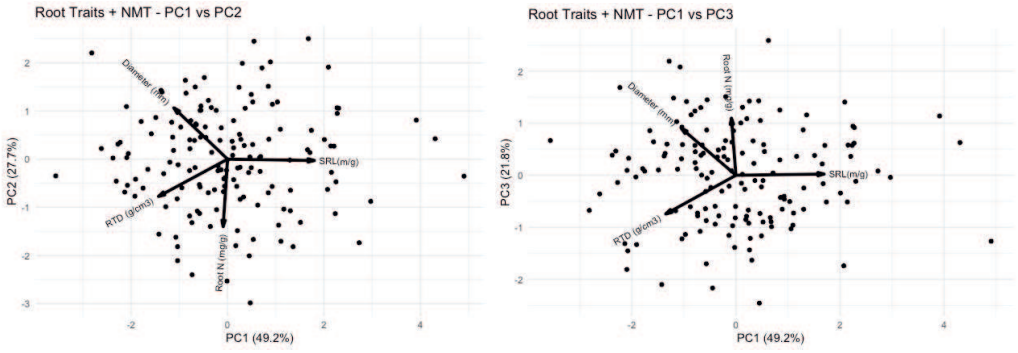
E. Intraspecific PCA

PCA analysis of cacao plants from a greenhouse experiment with complete information for the four core root traits. The Eigenvalue, the proportion of variance explained by each principal component (PC), and the contributions of each root traits for those PC are displayed.

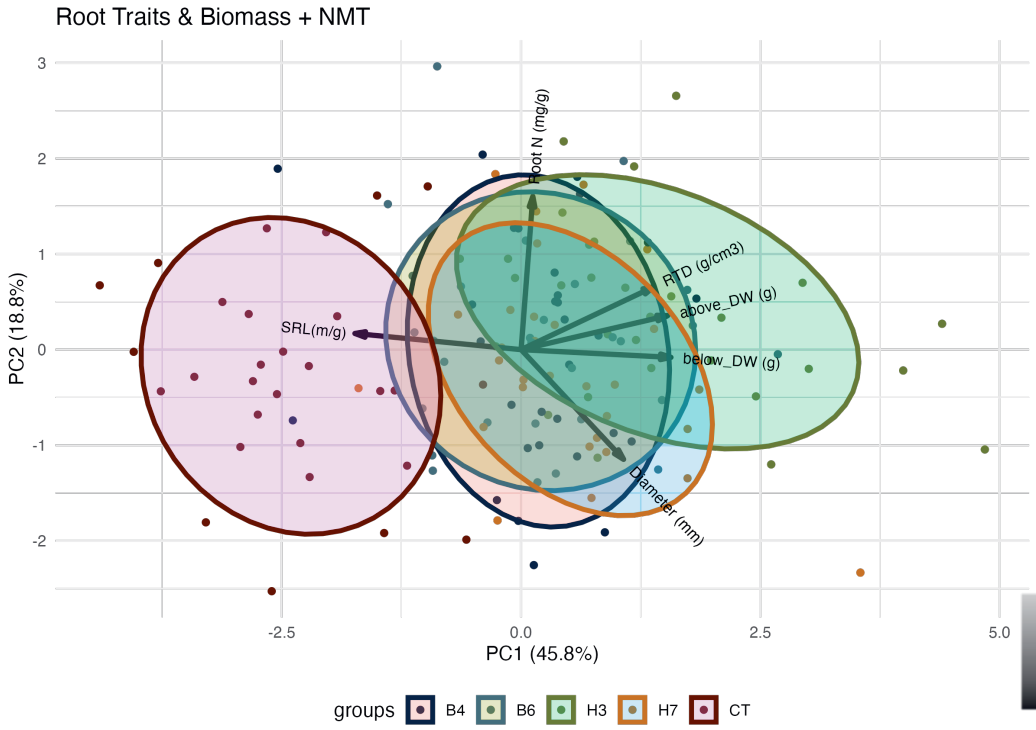
		PC1	PC2	PC3	PC4
	Eigenvalue	1.90425488	1.10555694	0.96233926	0.02784892
Excluding	Variance (%)	47.60	27.63	24.05	0.69
CT	Root N (mg/g)	0.7727806	21.22275126	78.0030843	0.001383894
(n=120)	Root D (mm)	23.4368900	37.53624951	13.6811160	25.345744514
Figure 3	SRL (m/g)	51.6177611	0.09133522	0.2790834	48.011820269
	RTD (g/cm^3)	24.1725683	41.14966401	8.0367164	26.641051323
	Eigenvalue	1.96753892	1.10976249	0.87284165	0.04985694
Including	Variance	49.188473	27.744062	21.821041	1.246424
CT	Root N (mg/g)	0.1283408	53.0945224	46.7724487	0.004688075
(n=149)	Root D (mm)	18.980237	30.6862510	31.7037722	18.629739782
Figure S3	SRL (m/g)	49.526046	0.02088815	0.0206089	50.432457064
	RTD (g/cm^3)	31.365376	16.1983383	21.5031702	30.933115078

Root D – average root diameter, SRL – specific root length, RTD – root tissue density, and Root N – root nitrogen content.

Principal component analysis including control treatments. (a) PC1 vs PC2; and (b) PC1 vs PC3. When including CT in PC1 vs PC2, there is a robust negative correlation between SRL and D and between SRL and RTD. Moreover, root N is also independent of all other root traits. When using PC1 vs PC3, the same negative correlations between SRL and RTD/D arise, but now, Root N positively correlates with D. SRL – specific root length (m/g), D – diameter (mm), Root N – nitrogen content in roots (mg/g), and RTD – root tissue density (g/cm3).



F. PCA of measured traits across mycorrhizal treatments



References

Alguacil, M. del M., Torres, M.P., Montesinos-Navarro, A., Roldán, A., 2016. Soil Characteristics Driving Arbuscular Mycorrhizal Fungal Communities in Semiarid Mediterranean Soils. Appl. Environ. Microbiol. 82, 3348–3356. https://doi.org/10.1128/aem.03982-15

Berger, F., Gutjahr, C., 2021. Factors affecting plant responsiveness to arbuscular mycorrhiza. Curr. Opin. Plant Biol. 59, 101994. https://doi.org/10.1016/j.pbi.2020.101994

Bergmann, J., Weigelt, A., Plas, F. van der, Laughlin, D.C., Kuyper, T.W., Guerrero-Ramirez, N., Valverde-Barrantes, O.J., Bruelheide, H., Freschet, G.T., Iversen, C.M., Kattge, J., McCormack, M.L., Meier, I.C., Rillig, M.C., Roumet, C., Semchenko, M., Sweeney, C.J., Ruijven, J. van, York, L.M., Mommer, L., 2020. The fungal collaboration gradient dominates the root economics space in plants. Sci Adv 6, eaba3756. https://doi.org/10.1126/sciadv.aba3756

Cadby, J., Araki, T., 2021. Towards ethical chocolate: multicriterial identifiers, pricing structures, and the role of the specialty cacao industry in sustainable development. SN Bus. Econ. 1, 44. https://doi.org/10.1007/s43546-021-00051-y

Cardoso, E.J.B.N., Nogueira, M.A., Zangaro, W., 2017. Importance of Mycorrhizae in Tropical Soils. In: de Azevedo, J., Quecine (eds), Diversity and Benefits of Microorganisms from the Tropics pp. 245–267. Springer. https://doi.org/10.1007/978-3-319-55804-2_11

Cobb, A.B., Duell, E.B., Haase, K.B., Miller, R.M., Wu, Y.Q., Wilson, G.W.T., 2021. Utilizing mycorrhizal responses to guide selective breeding for agricultural sustainability. Plants People Planet 3, 578–587. https://doi.org/10.1002/ppp3.10200

Cuenca, G., Herrera, R., Meneses, E., 1990. Effects of VA mycorrhiza on the growth of cacao seedlings under nursery conditions in Venezuela. Plant Soil 126, 71–78. https://doi.org/10.1007/bf00041370

Dragusanu, R., Giovannucci, D., Nunn, N., 2014. The Economics of Fair Trade. J. Econ. Perspect. 28, 217–236. https://doi.org/10.1257/jep.28.3.217

Frew, A., Weinberger, N., Powell, J.R., Watts-Williams, S.J., Aguilar-Trigueros, C.A., 2024. Community assembly of root-colonizing arbuscular mycorrhizal fungi: beyond carbon and into defence? ISME J. 18, wrae007. https://doi.org/10.1093/ismejo/wrae007

Graves, S., Piepho, H.-P., Selzer, L., Dorai-Raj, S., 2015. multcompView: Visualizations of Paired Comparisons [WWW Document]. URL https://cran.r-project.org/web/packages/multcompView/index.html (accessed 7.30.24).

Guerrero-Ramírez, N.R., Mommer, L., Freschet, G.T., Iversen, C.M., McCormack, M.L., Kattge, J., Poorter, H., Plas, F., Bergmann, J., Kuyper, T.W., York, L.M., Bruelheide, H., Laughlin, D.C., Meier, I.C., Roumet, C., Semchenko, M., Sweeney, C.J., Ruijven, J., Valverde-Barrantes, O.J., Aubin, I., Catford, J.A., Manning, P., Martin, A., Milla, R., Minden, V., Pausas, J.G., Smith, S.W., Soudzilovskaia, N.A., Ammer, C., Butterfield, B., Craine, J., Cornelissen, J.H.C., Vries, F.T., Isaac, M.E., Kramer, K., König, C., Lamb, E.G., Onipchenko, V.G., Peñuelas, J., Reich,

P.B., Rillig, M.C., Sack, L., Shipley, B., Tedersoo, L., Valladares, F., Bodegom, P., Weigelt, P., Wright, J.P., Weigelt, A., 2021. Global root traits (GRooT) database. Glob. Ecol. Biogeogr. 30, 25–37. https://doi.org/10.1111/geb.13179

Habte, M., Osorio, W., 2001. Arbuscular Mycorrhizas: Producing and Applying Arbuscular Mycorrhizal Inoculum. College of Tropical Agriculture and Human resources, University of Hawaii, Honolulu. http://hdl.handle.net/10125/25589

Harrell, F.E.J., 2023. Hmisc: Harrell Miscellaneous [WWW Document]. URL https://hbiostat.org/R/Hmisc/

Hu-zhe, Z., Chun-lan, C., Yu-ting, Z., Dan, W., Yu, J., Yong, K.K., 2005. Active changes of lignification-related enzymes in pepper response to *Glomus intraradices* and/or *Phytophthora capsici*. J. Zhejiang Univ. Sci. B 6, 778–786. https://doi.org/10.1007/bf02842437

Janos, D.P., 2007. Plant responsiveness to mycorrhizas differs from dependence upon mycorrhizas. Mycorrhiza 17, 75–91. https://doi.org/10.1007/s00572-006-0094-1

Jansa, J., Erb, A., Oberholzer, H., Šmilauer, P., Egli, S., 2014. Soil and geography are more important determinants of indigenous arbuscular mycorrhizal communities than management practices in Swiss agricultural soils. Mol. Ecol. 23, 2118–2135. https://doi.org/10.1111/mec.12706

Kaeppler, S.M., Parke, J.L., Mueller, S.M., Senior, L., Stuber, C., Tracy, W.F., 2000. Variation among Maize Inbred Lines and Detection of Quantitative Trait Loci for Growth at Low Phosphorus and Responsiveness to Arbuscular Mycorrhizal Fungi. Crop Sci. 40, 358–364. https://doi.org/10.2135/cropsci2000.402358x

Kramer-Walter, K.R., Bellingham, P.J., Millar, T.R., Smissen, R.D., Richardson, S.J., Laughlin, D.C., 2016. Root traits are multidimensional: specific root length is independent from root tissue density and the plant economic spectrum. J. Ecol. 104, 1299–1310. https://doi.org/10.1111/1365-2745.12562

Kuyper, T.W., Jansa, J., 2023. Arbuscular mycorrhiza: advances and retreats in our understanding of the ecological functioning of the mother of all root symbioses. Plant Soil 489, 41-88. https://doi.org/10.1007/s11104-023-06045-z

Kuyper, T.W., Wang, X., Muchane, M.N., 2021. The Interplay Between Roots and Arbuscular Mycorrhizal Fungi Influencing Water and Nutrient Acquisition and Use Efficiency, in: Rengel, Z., Djalovic, I. (Eds.), The Root Systems in Sustainable Agricultural Intensification. pp. 193–220. https://doi.org/10.1002/9781119525417.ch7

Lanaud, C., Vignes, H., Utge, J., Valette, G., Rhoné, B., Caputi, M.G., Nieto, N.S.A., Fouet, O., Gaikwad, N., Zarrillo, S., Powis, T.G., Cyphers, A., Valdez, F., Nunez, S.Q.O., Speller, C., Blake, M., Valdez, F.Jr., Raymond, S., Rowe, S.M., Duke, G.S., Romano, F.E., Solórzano, R.G.L., Argout, X., 2024. A revisited history of cacao domestication in pre-Columbian times revealed by archaeogenomic approaches. Sci. Rep. 14, 2972. https://doi.org/10.1038/s41598-024-53010-6

Lê, S., JulieJosse, Husson, F., 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25, 1-18. https://doi.org/10.18637/jss.v025.i01.

Lovera, M., Cuenca, G., Fajardo, L., Cáceres, A., Guerra-Sierra, B.E., 2022. AMF Diversity in Coffee and Cacao Agroforestry Systems: Importance for Crop Productivity and Forest Conservation, In: Luga M.A., Pagano, M.C. (Eds), Mycorrhizal Fungi in South America, Biodiversity, Conservation, and Sustainable Food Production, pp. 107–127. https://doi.org/10.1007/978-3-031-12994-0_5

Ma, Z., Guo, D., Xu, X., Lu, M., Bardgett, R.D., Eissenstat, D.M., McCormack, M.L., Hedin, L.O., 2018. Evolutionary history resolves global organization of root functional traits. Nature 555, 94–97. https://doi.org/10.1038/nature25783

Mazerolle, M.J., 2023. AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). https://cran.r-project.org/web/packages/AICcmodavg/AICcmodavg.pdf

Motamayor, J.C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A., Lanaud, C., 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89, 380–386. https://doi.org/10.1038/sj.hdy.6800156

Ostonen, I., Püttsepp, Ü., Biel, C., Alberton, O., Bakker, M.R., Lõhmus, K., Majdi, H., Metcalfe, D., Olsthoorn, A.F.M., Pronk, A., Vanguelova, E., Weih, M., Brunner, I., 2007. Specific root length as an indicator of environmental change. Plant Biosyst. 141, 426–442. https://doi.org/10.1080/11263500701626069

Paguntalan, D.P., Jr, I.E.B., Aggangan, N.S., 2020. Cacao – Arbuscular Mycorrhizal Fungi Association: A Review on Its Increased Potential in Improving the Ecology of Cacao Plantations in the Philippines. Plant Soil 183, 315–322. https://doi.org/10.1007/bf00011447

Pascual, U., Termansen, M., Hedlund, K., Brussaard, L., Faber, J.H., Foudi, S., Lemanceau, P., Jørgensen, S.L., 2015. On the value of soil biodiversity and ecosystem services. Ecosyst. Serv. 15, 11–18. https://doi.org/10.1016/j.ecoser.2015.06.002

Peterson, B.G., 2004. PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis [WWW Document]. URL https://cran.r-project.org/web/packages/PerformanceAnalytics/index.html

Qin, M., Zhang, Q., Pan, J., Jiang, S., Liu, Y., Bahadur, A., Peng, Z., Yang, Y., Feng, H., 2020. Effect of arbuscular mycorrhizal fungi on soil enzyme activity is coupled with increased plant biomass. Eur. J. Soil Sci. 71, 84–92. https://doi.org/10.1111/ejss.12815

Reich, P.B., Tjoelker, M.G., Pregitzer, K.S., Wright, I.J., Oleksyn, J., Machado, J., 2008. Scaling of respiration to nitrogen in leaves, stems and roots of higher land plants. Ecol. Lett. 11, 793–801. https://doi.org/10.1111/j.1461-0248.2008.01185.x

Roches, S.D., Post, D.M., Turley, N.E., Bailey, J.K., Hendry, A.P., Kinnison, M.T., Schweitzer, J.A., Palkovacs, E.P., 2018. The ecological importance of intraspecific variation. Nature Ecol. Evol. 2, 57–64. https://doi.org/10.1038/s41559-017-0402-5

Soto-Navarro, C.A., Harfoot, M., Hill, S.L.L., Campbell, J., Mora, F., Campos, C., Pretorius, C., Pascual, U., Kapos, V., Allison, H., Burgess, N.D., 2021. Towards a multidimensional biodiversity index for national application. Nature Sustain. 4, 933–942. https://doi.org/10.1038/s41893-021-00753-z

Stahlhut, K.N., Conway, M., Mason, C.M., Bauer, J.T., 2023. Intraspecific variation in mycorrhizal response is much larger than ecological literature suggests. Ecology 104, e4015. https://doi.org/10.1002/ecy.4015

Trouvelot, A., Kough, V., Gianinazzi-Pearson, V., 1986. Mesure du taux de mycorhization VA d'un systeme radiculaire. Recherche de methods d'estimation ayant une signification fonctionnelle. In: Gianinzaai-Perason, V., Gianinazzi, S. (eds), Physiological and Genetical Aspects of Mycorrhizae, pp217-221. INRA, Paris.

Valverde-Barrantes, O.J., Freschet, G.T., Roumet, C., Blackwood, C.B., 2017. A worldview of root traits: the influence of ancestry, growth form, climate and mycorrhizal association on the functional trait variation of fine-root tissues in seed plants. New Phyt. 21, 1562–1573. https://doi.org/10.1111/nph.14571

Verbruggen, E., Kiers, E.T., 2010. Evolutionary ecology of mycorrhizal functional diversity in agricultural systems. Evol. Appl. 3, 547–560. https://doi.org/10.1111/j.1752-4571.2010.00145.x

Verbruggen, E., Kiers, E.T., Bakelaar, P.N.C., Röling, W.F.M., Heijden, M.G.A. van der, 2012. Provision of contrasting ecosystem services by soil communities from different agricultural fields. Plant Soil 350, 43–55. https://doi.org/10.1007/s11104-011-0828-5

Vu, V.Q., Friendly, M., Tavadyan, A., 2024. ggbiplot: A Grammar of Graphics Implementation of Biplots [WWW Document]. URL https://cran.r-project.org/web/packages/ggbiplot/index.html

Wang, Q., Chai, Q., Dou, X., Zhao, C., Yin, W., Li, H., Wei, J., 2024. Soil Microorganisms in Agricultural Fields and Agronomic Regulation Pathways. Agronomy 14, art. 669. https://doi.org/10.3390/agronomy14040669

Wang, X., Zhang, M., Sheng, J., Feng, G., Kuyper, T.W., 2023. Breeding against mycorrhizal symbiosis: Modern cotton (*Gossypium hirsutum* L.) varieties perform more poorly than older varieties except at very high phosphorus supply levels. J. Integr. Agric. 22, 701–715. https://doi.org/10.1016/j.jia.2022.08.004

Wang, X.-X., Wang, X., Sun, Y., Cheng, Y., Liu, S., Chen, X., Feng, G., Kuyper, T.W., 2018. Arbuscular Mycorrhizal Fungi Negatively Affect Nitrogen Acquisition and Grain Yield of Maize in a N Deficient Soil. Front. Microbiol. 9, art. 418. https://doi.org/10.3389/fmicb.2018.00418

Watts-Williams, S.J., Emmett, B.D., Levesque-Tremblay, V., MacLean, A.M., Sun, X., Satterlee, J.W., Fei, Z., Harrison, M.J., 2019. Diverse *Sorghum bicolor* accessions show marked variation in growth and transcriptional responses to arbuscular mycorrhizal fungi. Plant, Cell Environ. 42, 1758–1774. https://doi.org/10.1111/pce.13509

Weigelt, A., Mommer, L., Andraczek, K., Iversen, C.M., Bergmann, J., Bruelheide, H., Fan, Y., Freschet, G.T., Guerrero-Ramírez, N.R., Kattge, J., Kuyper, T.W., Laughlin, D.C., Meier, I.C., Plas, F., Poorter, H., Roumet, C., Ruijven, J., Sabatini, F.M., Semchenko, M., Sweeney, C.J., Valverde-Barrantes, O.J., York, L.M., McCormack, M.L., 2021. An integrated framework of plant form and function: the belowground perspective. New Phyt. 232, 42–59. https://doi.org/10.1111/nph.17590

Westerband, A.C., Funk, J.L., Barton, K.E., 2021. Intraspecific trait variation in plants: a renewed focus on its role in ecological processes. Ann. Bot. 127, 397–410. https://doi.org/10.1093/aob/mcab011

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., Villar, R., 2004. The worldwide leaf economics spectrum. Nature 428, 821–827. https://doi.org/10.1038/nature02403

Wu, S., Shi, Z., Chen, X., Gao, J., Wang, X., 2022. Arbuscular mycorrhizal fungi increase crop yields by improving biomass under rainfed condition: a meta-analysis. PeerJ 10, e12861. https://doi.org/10.7717/peerj.12861

Chapter 5. Discussion



Research Foundations and Public Value

Green and Evergreen Revolution

The Green Revolution of the 1960s–1970s was designed to feed a rapidly growing population and make farming possible on a large scale. It relied on specialization in monocultures, improved cereal varieties which was later expanded to other crops, and heavy use of fertilizers, pesticides, and irrigation (Conway and Barbie, 1988; Falcon, 1970). Advances in machinery, cereal crop breeding, and chemicals (Power and Follett, 1987) encouraged farmers to focus on just a few high-yielding crops, allowing more efficient planting, management, and marketing. While this brought short-term benefits, it has also created serious environmental and social problems. Mechanization has compacted soils and caused erosion, reducing their organic matter content. Because inorganic fertilizers are applied in amounts that do not reflect short-term plant demand, they are frequently applied in excess, which damages soil health and causes eutrophication. Pesticides (insecticides, fungicides, and herbicides) not only target pests but also harm beneficial organisms and biodiversity. At the same time, relying on a narrow range of crop varieties makes crops more vulnerable to disease and further reduces the diversity of birds and pollinators. Together, these challenges illustrate the urge to transform the current practices and shift toward an "evergreen revolution" that produces food more sustainably (Struik and Kuyper, 2017).

The evergreen revolution shifts focus to practices that 'achieve productivity in perpetuity without ecological harm' (Swaminathan, 2006). Organic fertilizers and crop rotations build up soil organic matter, improve nutrient cycling, and support biodiversity, while soil testing is only used to help maintain fertility and prevent nutrient imbalances (Francis, 2020). Soil biodiversity, enhanced through microbial amendments, stimulate uptake and / or use efficiency of the nutrients already present, creating a positive feedback loop for soil health, fertility, and plant growth (Wei et al., 2024). Biopesticides, such as *Trichogramma*, not only protect crops but also offer pro-women, pro-poor, and pro-nature solutions since their cultivation is often carried out by women in community groups (Fliert and Proost, 1999; Kesavan and Malarvannan, 2010). Protecting biodiversity is vital in evergreen revolutions, supporting biotechnological advances through species that serve as indicators, aid bioremediation, safeguard water quality, and yield metabolites for health and crop improvement (Swaminathan, 2003). Cropping systems like regenerative agriculture restore soil health (Giller et al., 2021; Schreefel et al., 2020), while climate-smart agriculture adapts farming to climate change (Singh and Singh, 2017). Mixed and intercropping approaches, which increase crop species diversity or increase genetic diversity within one crop species, positively affect ecological functions. In a global meta-analysis, these diverse cropping systems that are based on increasing plant genetic diversity within or between crops enhances crop production, associated biodiversity, and several supporting

and regulating ecosystem services (Beillouin et al., 2021b). Breeding crops that optimally benefit from these systems will further boost this agricultural shift (Bourke et al., 2021; Haug et al., 2021). Among these cropping systems, agroforestry stands out as the leading approach for preserving local germplasm, biodiversity, and planetary health (Rosenstock et al., 2019; Swaminathan, 2012).

Agroforestry systems (AFS) combine trees, crops, and sometimes livestock, to provide social, economic, ecological, and environmental benefits. For smallholders, they diversify income by allowing farmers to cultivate fruit and timber trees, increase food security through more balanced diets, and create new jobs along extended value chains (Mukhlis et al., 2022). These opportunities can especially benefit marginalized groups, including women in rural areas. AFS also offer sustainable bioenergy options such as biofuels, biogas, and electricity (Sharma et al., 2016). Ecologically, they enhance biodiversity across trophic levels, providing functions that are essential for smallholder farming. A greater diversity of insects increases pollination efficiency and fruit production, while natural biocontrol agents regulate pest populations. However, these benefits depend on choosing optimal plant–tree combinations, since poor design can also favour pests and diseases. Plant diversity across annual and perennial crops contributes to water security, stable microclimates, and overall ecosystem resilience. AFS also improve soil health through nitrogen fixation by leguminous trees, nutrient recycling from decomposed biomass, and nutrient capture from deeper soil layers by tree roots (Nair, 2011). These deep-rooted trees additionally enhance soil structure and microbial activity, creating healthier, more fertile soils (Fahad et al., 2022). Despite these benefits, agroforestry systems remain under-researched and underfunded.

From the Global Cacao Market to Local Agroforestry Systems in Panama

The global development of sustainable cacao agroforestry systems (CAFS) is constrained by what is known as an “hourglass” structure in the supply chain. At one end are the millions of smallholders, and at the other end are the billions of global consumers, with fewer than ten companies controlling most of the trade in between. Three major firms, Olam Food Ingredients, Barry Callebaut, and Cargill, dominate cocoa supply and utilities. Smallholders produce about 90% of the world’s cacao, with Ivory Coast (40%), Ghana (11%), and Indonesia (11%) as the leading exporters (FAOSTAT, 2023b). Smallholder farms receive far less research, development, and policy support than large-scale farms. A Scopus search on 5 August 2025 using the queries “smallholder” and “large-scale farms” (peer-reviewed articles, reviews, book chapters, and books in English) showed research on large-scale farms is roughly twice as prevalent, reflecting funding priorities, agribusiness interests, and the ease of working with uniform systems. In contrast, the diversity of smallholder agriculture—fragmented plots, mixed crops, and varied socio-economic contexts—poses greater research challenges, reinforcing its underrepresentation. This imbalance risks widening inequalities, skewing

policy, and limiting relevant innovations. In plant breeding, commercial demand often favours broadly adapted varieties, including soybeans bred for symbiosis with specific rhizobia, which are frequently sold with inoculum coatings. In contrast, smallholders need locally tailored genotypes, making collaborative and participatory breeding approaches more suitable. This global imbalance highlights the importance of examining how the cacao industry in Panama develops.

In Panama, cacao agroforestry systems (CAFS) sustain many livelihoods yet remain underdeveloped due to local social and economic challenges. Although Panama is classified as a high-income country (Fig 1 - World Bank, 2024), reducing inequalities, particularly among women, Afro-descendants, and Indigenous people, who are the main cocoa producers, remains a major challenge (World Bank, 2025). Agriculture is the country's largest employer (IMF, 2023), with cacao cultivation led by the Ngäbe-Buglé in Bocas del Toro, who cultivate over 6,000 hectares. Productivity remains four to five times lower than in other working sectors, due to largely unknown germplasm and an estimated 30% loss of total potential yield (COCABO, personal communication, 2022). Panama lies within cacao's centre of origin, where it has been cultivated for millennia (Lanaud et al., 2024), but very little local germplasm research exists. The only initiative before 2020 consisted of the Ministry of Environment (MIDA) selecting three high-yielding, disease-resistant wild accessions (Gutierrez, 2020). Moreover, smallholders earn only $0.02 per kilogram of fermented and dried cacao, sold through cooperatives that hold annual export contracts with fixed prices and bonuses for additional production. Despite this, earnings remain far below the 2025 international price of $7.37 per kilogram for fermented and dried cacao (YCharts, 2025). Most farms span 2–3 hectares, producing about 90 kg/ha/year, yielding minimal income that limits investment in technology. Farmers often attempt to sell to local buyers at higher prices or expand production to increase yield. Strengthening the value chain while developing technologies that harness agroforestry biodiversity, such as mycorrhizal symbiosis, offers a promising path to improve both farmer livelihoods and the sustainability of Panama's cacao industry.

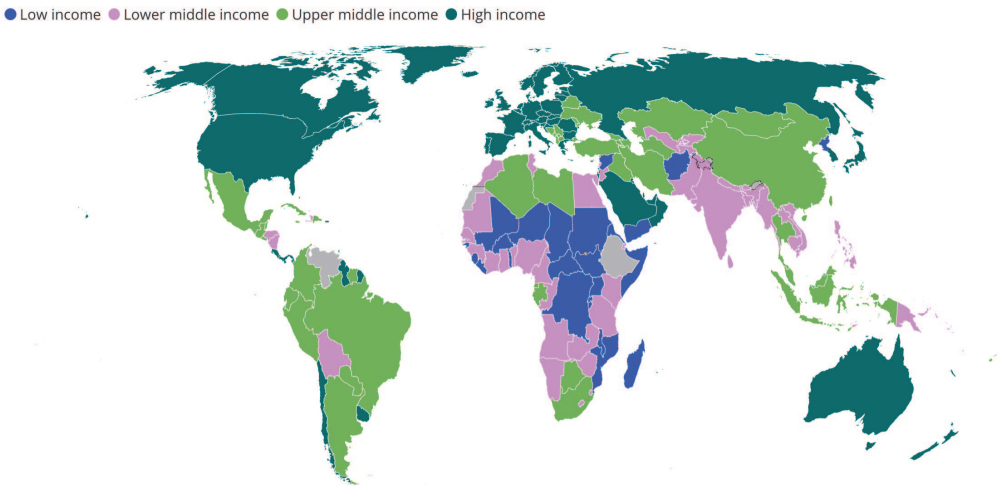


Figure 1. A map of World Bank high-income economies, as of 2023, where blue-colored countries, including Panama, are classified as high-income countries (World Bank, 2024)

Innovations in Cacao Cultivation: Optimizing Cacao–AMF Symbiosis

Cacao is a perennial crop particularly dependent on mycorrhizal symbiosis, creating opportunities to enhance productivity through these associations. Mycorrhizal symbiosis occurs between plant roots and specific soil fungi, with different fungal species providing a range of benefits in exchange for photosynthates. These benefits include improved nutrient uptake (Kuyper et al., 2021) and the priming of abiotic and biotic defenses (Boyno et al., 2025). Cacao grown without mycorrhizal fungi exhibits nutrient deficiencies, reduced vigour (Azizah and Ragu, 1986; Chulan and Martin, 1992; Cuenca et al., 1990; Paguntalan et al., 2020), and compromised immunity (Herre et al., 2007). Mycorrhizal symbiosis is particularly relevant in low-input-, organically managed systems. These systems are common in CAFS, where resource-poor farmers often rely on organic practices out of necessity, but also embrace them as a pathway to access FairTrade premiums (Cadby and Araki, 2021; Dragusanu et al., 2014). Optimizing cacao–arbuscular mycorrhizal relationships, while improving cacao germplasm, could provide a viable solution to current agricultural challenges, provided we understand the environmental conditions and genetic factors that promote mutualistic symbiosis.

Modern plant breeding increasingly integrates genetics, physiology, and environment to enhance crop performance under mycorrhizal symbiosis. Cacao, a semi-domesticated, highly heterozygous, and allogamous perennial, remains slow and complex to breed. Grafting offers a “double” perspective on genotypes, supporting separate study of above- and belowground effects on symbiosis, such as species composition, colonization, and functional benefits (Hohmann and Messmer, 2017). Traditional breeding has emphasized yield and quality, yet these traits also shape rootstock selection

and root traits critical for sustainable production and AMF management. Rootstocks should balance vegetative and generative growth, improve sink–source relationships, and integrate traits that support microbial symbiosis and disease resistance (López-Marín et al., 2017). Root traits and soil quality also influence microbial community structure and diversity in cacao agroforestry systems (Schmidt et al., 2024). The cacao root microbiome is now studied within a genetic × environment × management (G×E×M) framework (Schmidt et al., 2022). High-throughput screening of plant traits and root microbiomes identified heritable, field-relevant variation linked to microbial functions and root morphology (Wille et al., 2020). Advances in genome assembly, resequencing, and SNP technologies allow integration of genomic and environmental data for early-stage selection (Flanagan and Jones, 2019; Livingstone et al., 2017). Early screening in target environments captures key plant–microbe interactions, reducing the need to raise trees to maturity. Together, these approaches can accelerate breeding of cacao varieties optimized for productive, sustainable, and microbially enriched agroforestry systems. Early screening in target environments captures key plant–microbe interactions, reducing the need to raise trees to maturity. Together, these approaches can accelerate breeding of cacao varieties optimized for productive, sustainable, and microbially enriched agroforestry systems.

From Ideas to Evidence: The Arc of My PhD Research

After reflecting on the socioeconomic and agricultural context of my home country, in 2019 I successfully applied for a PhD scholarship to improve *Theobroma cacao* germplasm by studying its interaction with arbuscular mycorrhizal fungi (AMF), focusing on the characterization of plant genetic composition and root traits that impact on and respond to AMF. Before addressing this, it was essential to characterize both above- and belowground plant genetics, root traits, AMF community composition, and their interactions. In this thesis, I aim to answer the following questions:

Overarching research question

- **How does genetic variation of rootstocks and scion, and root traits in *T. cacao* influence the AM fungal community composition (core microbiome and indicator species) and AMF colonization?**
- **What are the root traits that facilitate the beneficial interaction with AMF?**

Sub-research questions:

- What cacao genotypes are cultivated across different regions in Panama?
- How do smallholders manage trees across different regions, including direct sowing or vegetative propagation by selection of scions and grafting?
- How do management practices, like selection of neighbouring species and cacao germplasm, and environmental properties of CAFS influence AMF communities?

- What is the core AMF community in relation to cacao? Which AMF species are indicator species for certain scion or rootstock, or specific environmental conditions?

To have a successful project, I actively sought collaborators across several international universities, including Wageningen University, and engaged local partners such as governmental agencies (e.g., IDIAP – Institute for Agricultural Innovation of Panama) and research institutions (e.g., Indicasat-AIP – Panama Institute of Scientific Research and High Technology Services AIP, and STRI – Smithsonian Tropical Research Institute). Partnering with governmental agencies was crucial to identify cacao-growing regions and to establish connections with local farmers. Each research institute brought specific strengths: Indicasat-AIP provided molecular expertise in AMF and plant DNA sequencing, while STRI contributed ecological expertise and access to greenhouse facilities. These collaborations significantly strengthened the quality of the project and were central to its success. The outcome of this research would not have been possible without strategic networking and partnerships.

Fostering creativity and resourcefulness was essential throughout the methodological process. Conducting research in the Global South involves realities vastly different from those in the Global North. While researchers in developed countries often benefit from established support systems, including technicians, workflows, and ready resources, much of Latin America operates in a more competitive and individualistic environment due to scarce science funding (Ciocca and Delgado, 2017). Most logistical, technical, and administrative tasks must be learned, coordinated, and executed independently. I managed the logistics, equipment, and materials for both the field survey and greenhouse experiments, including navigating COVID-related travel restrictions during fieldwork with my international supervisors. That visit proved invaluable, reshaping and enriching our understanding of cacao cultivation and the experimental design. For the molecular work, I coordinated the procurement of all reagents and materials for PCR extractions and Illumina sequencing, working closely with experts to ensure readiness. While routine for many researchers, in Latin America these tasks demand time, effort, and risk costly errors—challenges that European researchers encounter to a lesser extent. In practice, nothing unfolds exactly as planned, so constant re-planning without losing sight of the experiment's core aim became the most valuable skill I had to develop. Overall, this project reinforced the value of strategic collaboration, adaptability, and persistence in strengthening the methodology, showing that its success does not rely solely on technical competence.

Overview of main findings

In Chapter 2, I found that the cacao genetic population in Bocas del Toro region was more admixed, with large representation of Amelonado, Nacional, IMC (Iquitos), and Nanay, compared

with the cacao population of Herrera, mostly dominated by Amelonado and Criollo. Interestingly, in Herrera there were individuals with pure Criollo ancestry, a rare finding given its high disease susceptibility. The Criollo genetic group was domesticated by the Mesoamerican indigenous cultures and stands out for its delicate cacao flavour and low astringency. Across regions, smallholders traditionally develop trees through seeds (57% single tree origin vs 43% grafted trees). In Bocas del Toro however, with its historical cacao cultivation, 55% of farmers select best-performing scions and use these to renovate trees in their farms, whereas only 20% of farmers in Herrera graft their trees. From discussions with farmers I discovered, surprisingly, that rootstocks were not selected but used randomly. For environmental associations, total soil nitrogen showed the strongest association to decreased Criollo presence and increased Amelonado presence across scion, rootstock, and combined data.

In the Chapter 3, I examined arbuscular mycorrhizal fungal communities associated with cacao trees. In all I recorded at least 82 virtual taxa (a proxy for species), of which six, predominantly from the genus *Glomus*, were present in over 85% of samples, constituting the core mycobiome. While diversity did not differ significantly across regions, it varied markedly between farms. Notably, one farm in Bocas del Toro showed the lowest biodiversity indices, likely due to fertilization practices that increased soil available P. AMF community composition was significantly associated with scion genotype, root N, root P, elevation, and available P, but not by rootstock genotype nor with root morphological traits. These results may be explained by the metabolic role of the scion in AMF community composition, while root N and P reflect nutrient demand and the rootstock's capacity to support effective symbioses. Elevation, as a proxy for climatic and geological factors, and available P, a consistent determinant of AMF structure, emerged as key environmental association with AMF composition. Indicator species analysis revealed several virtual taxa strongly associated with specific scion genotypes. Lastly, these findings suggest that targeted management practices—particularly organic amendments that modulate N and P availability—can shape mycorrhizal community composition in cacao systems.

A key issue for me in the interpretation of the relationship between AMF communities and root physiological traits (root N and root P) was the variation in root N and P concentrations, complicating cross-study comparisons. In my results, root N and P concentrations seemed realistic, but the resulting N:P ratios were partly implausibly low. Nutrient stoichiometry generally predicts a positive correlation between root N and root P for balanced plant nutrition. This positive correlation was confirmed in my data series 1 and 3 ($r = +0.39$, $P = 0.007$), but series 2 showed a biologically unlikely strong negative correlation between root N and root P concentrations ($r = –0.57$, $P = 0.0004$). Consulting the lab manager revealed no analytical issues, so I accepted the measurements, which produced a PCA where root N and root P had opposite loadings (Figure 2).

My data also showed that a higher amount of soil available P had a negative effect on AMF species diversity. My results fitted with the meta-analysis by Ma et al. (2023). Integrating my data into their study proved impossible because, although the paper concluded that available P was the best predictor of AMF abundance, it did not report the actual quantities. Additionally, a cursory inspection revealed at least four different methods for assessing soil P (P-Olsen, P-Colwell, P-Mehlich, P-lactate), which could introduce bias since they measure different soil P pools. These issues underscore the need for detailed methodological scrutiny and consensus before interpreting such data.

In Chapter 4, I demonstrated that cacao selfed progenies varied in root traits and growth, and that plant responses to AMF were influenced by both plant genotype and AMF treatment. While root N and diameter showed genetic differentiation among selfed progenies, biomass, specific root length and root tissue density were more responsive to AMF treatments than to plant genetic background alone. Mycorrhizal inocula differed in their effects; in particular, AMF treatment H3 (from a Herrera plantation) produced seedlings with the highest biomass, suggesting that AMF community composition can be more important than colonization percentage for plant performance. I also found interactive effects between the selfed progenies and the mycorrhizal community treatments indicating plant genotype x AMF community interaction, denoted by a larger biomass growth across H3 treatment in progeny S_1-60.

Chapter 3 provided the original inoculum source for AMF treatments in Chapter 4. By linking AMF community composition in inocula from field (Chapter 3) and greenhouse experiments (Chapter 4, data to be analysed), I can identify AMF taxa consistently maintained across conditions. This distinction allows us to define the core AMF mycobiome, i.e. taxa present in most or all plants, and indicator species that respond to specific host developmental stages or environmental conditions. Because AMF communities likely shift throughout plant development (Herre et al., 2005; Husband et al., 2002a, 2002b) depending on the mutualistic symbiosis and environmental adaptation (Pereira et al., 2020b; Souza and Freitas, 2017), integrating these datasets will provide insight into how host metabolic pathways, as suggested by Chapter 3, shape the recruitment of functional AMF species.

Root traits such as diameter (D), specific root length (SRL), and root tissue density (RTD) followed an intraspecific Root Economics Space (RES) pattern distinct from interspecific patterns. Field survey data presented in Chapter 3 (Figure 2) confirmed that D varied far less than SRL and RTD, consistent with reduced intraspecific variation in D due to strong phylogenetic constraints (Valverde-Barrantes et al., 2017b). SRL and RTD played a major role determining root trait variance and mycorrhizal interaction. If D is effectively constant, as it is in cacao, SRL and RTD will be negatively correlated by mathematical necessity, even though AMF effects on RTD remain unexplained. I suggest a possible explanation, i.e. that AMF colonization causes increased lignification (Hu-zhe et al., 2005). Consequently, the intraspecific RES does not always mirror the

interspecific RES (Bergmann et al., 2020). This conclusion may be relevant when matching root traits of different plant species or different plant varieties in mixed and intercropping systems geared towards optimizing mycorrhizal benefit.

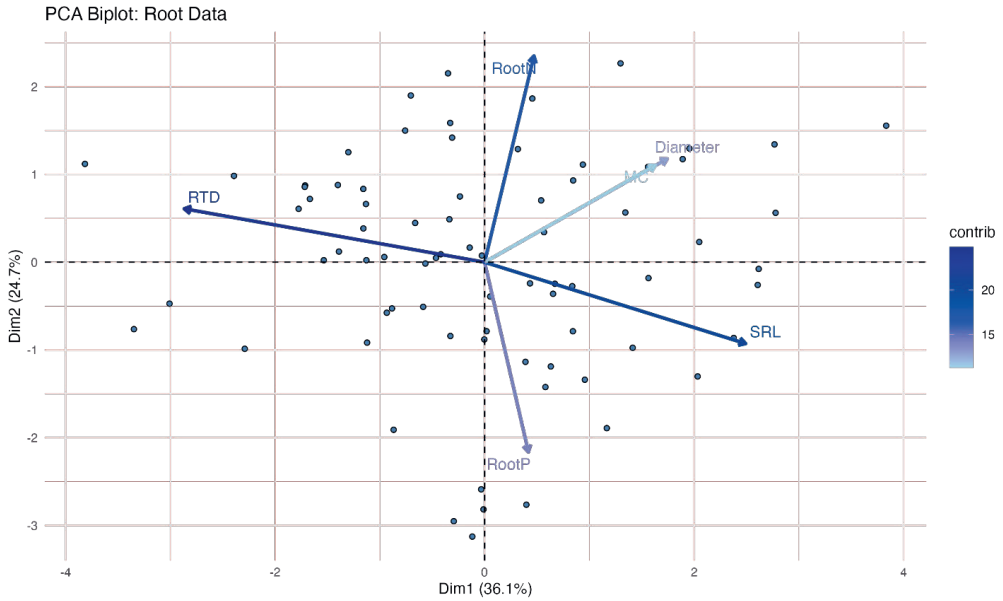


Figure 2. Root economic space framework based on the field study survey. Intensity of colors denote contribution of each variable to the overall variation.

In my research, field surveys were essential to assess the germplasm used in Panamanian agroforestry systems. Any valuable work on cacao cultivation must begin with the varieties grown by farmers or, in some cases, those still found in the wild. The field surveys revealed that scion genetic composition and environmental conditions strongly influenced AMF community structure, with little impact from root morphological traits. Greenhouse experiments, in turn, showed root diameter as critical for mycorrhizal colonization and thus dependence, while traits such as specific root length (SRL) and root tissue density (RTD) mediated physiological responses to symbiosis. Both greenhouse experiments and field surveys have their advantage and disadvantages. Surveys provided a broad view of the factors associated with AMF community composition, while greenhouse experiments offered complementary insights under controlled conditions, allowing clearer interpretation of results. Although greenhouse trials can be extended over time, they remain costly and lack ecological realism. Field experiments are more relevant to CAFS but are also expensive to maintain long-term, especially given land rental and management costs. A promising alternative is to collaborate directly with farmers through participatory approaches, such as plant breeding programs (see Conclusions), which combine scientific rigor with local knowledge.

Conclusions and Implications

When breeding for mycorrhizal symbiosis, three main targets can be considered: optimal colonization, greater plant selectivity for AMF communities, or improved responsiveness (Kuyper et al., 2021). Optimal colonization is attractive because fractional colonization often correlates positively with plant performance, although excessively high colonization can reduce growth under P-limited conditions (Graham and Abbott, 2000). Colonization is closely linked to root diameter (D) (Reinhardt and Miller, 1990; Valverde-Barrantes et al., 2016; Wang et al., 2022), a genetically conserved trait (Valverde-Barrantes et al., 2017b). Within the RES framework, SRL also falls along the collaborative gradient with AMF, and according to Valverde-Barrantes et al. (2017), it shows intermediate values in terms of genetic conservation. Results from both field and greenhouse experiments in this project confirm that cacao exhibits minimal variation in root diameter across progenies, while traits like specific root length (SRL) and root tissue density (RTD) vary and reflect functional aspects of the symbiosis. In the greenhouse H3 treated progenies showed, however, significant differences in both diameter (D) and specific root length (SRL), with the largest biomass growth corresponding to plants with the largest root diameter and lowest SRL values. This makes wider root diameter and smaller SRL a promising selection target. However, because SRL and RTD are strongly environmentally influenced, they likely have low heritability and are not very reliable breeding targets.

An important consideration is how increased root cortical space for colonization influences AMF symbiosis. Breeding in cereals has indirectly modified root diameter (Li et al., 2024; Martín-Robles et al., 2018; Wang et al., 2023) and root size (Fradgley et al., 2020; McGrail and McNear, 2021), thereby decreasing colonization patterns. In these crops, selecting for wider root diameter may offer stronger plant genetic control over space available for mycorrhizal colonization, since it likely has higher heritability. In cotton, changes in root diameter corresponded with decreased mycorrhizal response at moderate P supply (Wang et al., 2023). However, higher AMF colonization does not necessarily translate into larger presence of beneficial AMF species (VTs) or optimal mycorrhizal response. Likely by having more space for mycorrhizal colonization, plant selectivity over mycorrhizal species would become more critical.

Cacao breeding could target enhanced plant selectivity for more beneficial mycorrhizal species or stronger metabolic sanctions against less beneficial fungi. Although AMF are often described as nonspecific symbionts, many studies show substantial selectivity, either in the plant–AMF species combination or in the degree of benefit a fungus provides, with different AMF species conferring distinct plant benefits. My PhD research supports both perspectives: (1) field survey data revealed a clear scion genotype association with AMF community composition, suggesting a filtering effect by plant genetic composition, and (2) greenhouse experiments showed that the H3 AMF treatment

provided the highest biomass gains, particularly in selfed progeny S1-60. I hypothesize that biomass growth could be associated with beneficial effects for cacao yield potential. Further work characterizing AMF species consistently associated with cacao across samples, or identifying the specific fungi within the H3 community, could clarify which AMF species are most beneficial. This opens two complementary avenues: first, exploring plant genetic differences that drive preferential association with more beneficial AMF species, which could inform breeding for improved selectivity; and second, developing management practices that promote the dominance of favourable AMF communities in the field.

When breeding for mycorrhizal responsiveness it is necessary to define which plant trait is measured as mycorrhizal response. Mycorrhizal response can be involved in different developmental processes as biomass accumulation, vigour, nutrient uptake, tolerance to abiotic defences, as salinity or heavy metal tolerance, and strengthening biotic defences as immune priming against pathogens or change of volatile compounds against pests (Delavaux et al., 2017)**.** Mycorrhizal responsiveness is composed of two concepts: dependence and benefit. Dependence refers to the ability of plants to cope without AMF symbiosis, and benefit refers to degree in which a host plant benefits from AMF (Berger and Gutjahr, 2021). Breeding plants with improved responsiveness to AMF, i.e., expressing these traits more strongly when colonized (benefit), raises the concern that such plants may perform poorly under non-mycorrhizal conditions (dependence) (Hetrick et al., 1993; Kaeppler et al., 2000)**.** However, this downside is largely illusory: if a characteristic of sustainable agroecosystems is to have increased soil biodiversity and AMF richness, then the non-mycorrhizal condition is rare. This reinforces the value of breeding for mycorrhizal responsiveness within a broader ecological framework. Having considered the main avenues in breeding cacao for mycorrhizal symbiosis, future efforts targeting different aspects of mycorrhizal symbiosis or other agronomically relevant traits should ultimately be embedded within collaborative and participatory breeding approaches.

For developing new cacao cultivars, a more inclusive breeding framework such as participatory plant breeding (PPB) can be effective, given that cacao is mostly grown by smallholders. PPB, through decentralized breeding, enables the development of diverse varieties tailored to smallholder farms by directly selecting or breeding cacao trees within the target environment to enhance adaptation and performance in specific ecological contexts (Noru et al., 2024). These target environments reflect diverse management practices, including low-input and organic systems, as well as farmers' choices of neighbouring species. The latter is particularly relevant, since plant community composition is known to shape AMF community structure (Mai et al., 2023). Furthermore, Chapter 3 showed no significant differences in AMF diversity associated with cacao across regions. Breeders then incorporate farmers' feedback on plant growth and development alongside their own evaluations, and together these perspectives guide selection and crossing. Varieties emerging from

PPB programs often display variation, as farmers prioritize stability and suitability for their local conditions. Moreover, PPB offers better access to the rich genetic diversity present across the Americas—in both farms and forests—while also preserving older varieties maintained as landraces or farm clones. Since cacao is largely cultivated in CAFS by smallholders throughout the tropical belt, incorporating farmers' intuitive knowledge and practical experience into the evaluation and selection process is essential for creating resilient and successful germplasm.

Future directions

This dissertation has provided insights of the current cacao scenario in Panama and breeding prospects in symbiosis with arbuscular mycorrhizal fungi. From a technological perspective, it examined both the germplasm cultivated on farms and its associated value, as well as the soil biology underlying production—specifically, the arbuscular mycorrhizal fungal communities actively associating with the mycorrhizal cacao plant, and the plant and environmental factors associated with symbiosis, both in terms of quantity (mycorrhizal colonization) and quality (community structure). From a social and economic perspective, the personal realities of cacao farmers' livelihoods and the dynamics of the value chain that shape their harvests, should be considered. From this, I conclude that cacao farmers in Panama require a stronger social and economic structure that enables them to develop and implement available technologies for the germplasm, farm and value chain. Such a structure would support cacao breeding to improve yield and product quality, while also ensuring a fair value chain that provides access to dignified living wages.

I suggest developing plant breeding and mycorrhizal technologies to increase farm productivity, starting with the improvement of germplasm through participatory plant breeding (PPB) and management practices that foster a beneficial soil microbiome. As previously explained, PPB is a collaborative venture between breeders, farmers, and other stakeholders to evaluate, select, and cross valuable germplasm for the entire value chain, choosing varieties in the same environments where they will be grown. This approach enhances both plant performance and the value of the harvest for buyers. Effective and efficient breeding requires the inclusion of smallholders and their environments, as well as consideration of the traits they recognize as indicators of success. Further research into the associated mycorrhizal communities, characterizing their biology and the benefits they provide to cacao plants, could inform farm-specific management practices. Low-input organic agriculture, particularly agroforestry systems, holds great potential for promoting mycorrhizal symbiosis, offering resilient, long-term, and ecologically sustainable solutions.

Lastly, for social structures that support cacao industry development, it is necessary to optimize harvest times, fermentation processes, and the value chain to improve bean quality and market price. Farmers often lack the technology and theoretical knowledge to control temperature and humidity during fermentation, resulting in sub-optimal bean quality. This could be addressed through regional

processing centres where farmers can outsource fermentation or through training programs that provide both theoretical knowledge and practical equipment. Establishing a dignified value chain is crucial for all cacao farmers globally, requiring fewer intermediaries between producer and consumer. In Panama, this could be facilitated through efficient logistics, including proper storage and automated platforms. Although national production is modest, around 6 tons of cacao beans, a small volume compared to other producing countries, Panama's consistent success in international cacao awards positions it strongly for boutique markets. Thus, Panamanian cacao has greater commercial potential as a luxury boutique product rather than bulk cacao, with specialized chocolatiers as the most suitable buyers. By partnering with the right chocolatiers, Panamanian cacao can be marketed as single-origin chocolate, emphasizing the unique flavour profiles shaped by its terroir and plant germplasm.

References

Azizah, C. (Hashim), Ragu, P., 1986. Growth response of *Theobroma cacao* L. seedlings to inoculation with vesicular-arbuscular mycorrhizal fungi. Plant Soil 96, 279–285. https://doi.org/10.1007/bf02374771

Bergmann, J., Weigelt, A., Plas, F. van der, Laughlin, D.C., Kuyper, T.W., Guerrero-Ramirez, N., Valverde-Barrantes, O.J., Bruelheide, H., Freschet, G.T., Iversen, C.M., Kattge, J., McCormack, M.L., Meier, I.C., Rillig, M.C., Roumet, C., Semchenko, M., Sweeney, C.J., Ruijven, J. van, York, L.M., Mommer, L., 2020. The fungal collaboration gradient dominates the root economics space in plants. Sci Adv 6, eaba3756. https://doi.org/10.1126/sciadv.aba3756

Bourke, P.M., Evers, J.B., Bijma, P., Apeldoorn, D.F. van, Smulders, M.J.M., Kuyper, T.W., Mommer, L., Bonnema, G., 2021. Breeding Beyond Monoculture: Putting the "Intercrop" Into Crops. Front. Plant Sci. 12, 734167. https://doi.org/10.3389/fpls.2021.734167

Boyno, G., Danesh, Y.R., Çevik, R., Teniz, N., Demir, S., Durak, E.D., Farda, B., Mignini, A., Djebaili, R., Pellegrini, M., Porcel, R., Mulet, J.M., 2025. Synergistic benefits of AMF: development of sustainable plant defense system. Front. Microbiol. 16, 1551956. https://doi.org/10.3389/fmicb.2025.1551956

Cadby, J., Araki, T., 2021. Towards ethical chocolate: multicriterial identifiers, pricing structures, and the role of the specialty cacao industry in sustainable development. SN Bus. Econ. 1, 44. https://doi.org/10.1007/s43546-021-00051-y

Chulan, H.A., Martin, K., 1992. The vesicular-arbuscular (VA) mycorrhiza and its effects on growth of vegetatively propagated *Theobroma cacao* L. Plant Soil 144, 227–233. https://doi.org/10.1007/bf00012879

Ciocca, D.R., Delgado, G., 2017. The reality of scientific research in Latin America; an insider's perspective. Cell Stress Chaperones 22, 847–852. https://doi.org/10.1007/s12192-017-0815-8

Conway, G.R., Barbie, E.B., 1988. After the Green Revolution: Sustainable and equitable agricultural development. Futures 20, 651–670. https://doi.org/10.1016/0016-3287(88)90006-7

Cuenca, G., Herrera, R., Meneses, E., 1990. Effects of VA mycorrhiza on the growth of cacao seedlings under nursery conditions in Venezuela. Plant Soil 126, 71–78. https://doi.org/10.1007/bf00041370

Dragusanu, R., Giovannucci, D., Nunn, N., 2014. The Economics of Fair Trade. J. Econ. Perspect. 28, 217–236. https://doi.org/10.1257/jep.28.3.217

Fahad, S., Chavan, S.B., Chichaghare, A.R., Uthappa, A.R., Kumar, M., Kakade, V., Pradhan, A., Jinger, D., Rawale, G., Yadav, D.K., Kumar, V., Farooq, T.H., Ali, B., Sawant, A.V., Saud, S., Chen, S., Poczai, P., 2022. Agroforestry Systems for Soil Health Improvement and Maintenance. Sustainability 14, 14877. https://doi.org/10.3390/su142214877

Falcon, W.P., 1970. The Green Revolution: Generations of Problems. Am. J. Agric. Econ. 52, 698–710. https://doi.org/10.2307/1237681

FAOSTAT, 2023. Crops and Livestock products, Cocoa beans, area harvested in 2023, Crops/World regions/Cocoa beans/Area harvested from pick lists. [WWW Document]. URL https://www.fao.org/faostat/en/#data/QCL (accessed 1.13.25).

Fliert and Proest ***

Fradgley, N., Evans, G., Biernaskie, J.M., Cockram, J., Marr, E.C., Oliver, A.G., Ober, E., Jones, H., 2020. Effects of breeding history and crop management on the root architecture of wheat. Plant Soil 452, 587–600. https://doi.org/10.1007/s11104-020-04585-2

Francis, C., 2020. Organic Farming: The Ecological System. Agronomy Monogr. https://doi.org/10.2134/agronmonogr54

Giller, K.E., Hijbeek, R., Andersson, J.A., Sumberg, J., 2021. Regenerative Agriculture: An agronomic perspective. Outlook Agric. 50, 13–25. https://doi.org/10.1177/0030727021998063

Graham, J.H., Abbott, L.K., 2000. Wheat responses to aggressive and non-aggressive arbuscular mycorrhizal fungi. Plant Soil 220, 207–218. https://doi.org/10.1023/a:1004709209009

Gutierrez, A., 2020. Morphological Characterization of Three Promising Creole Genotypes of *Theobroma cacao* L., in Panama. Ciencia Agropecuaria 30, 150-169 [in Spanish]. http://www.revistacienciaagropecuaria.ac.pa/index.php/ciencia-agropecuaria/article/view/134/98.

Haug, B., Messmer, M.M., Enjalbert, J., Goldringer, I., Forst, E., Flutre, T., Mary-Huard, T., Hohmann, P., 2021. Advances in Breeding for Mixed Cropping – Incomplete Factorials and the Producer/Associate Concept. Front. Plant Sci. 11, 620400. https://doi.org/10.3389/fpls.2020.620400

Herre, E.A., Mejía, L.C., Kyllo, D.A., Rojas, E., Maynard, Z., Butler, A., Bael, S.A.V., 2007. Ecological implications of anti-pathogen effects of tropical fungal endophytes and mycorrhizae. Ecology 88, 550–558. https://doi.org/10.1890/05-1606

Hetrick, B.A.D., Wilson, G.W.T., Cox, T.S., 1993. Mycorrhizal dependence of modern wheat varieties, landraces, and ancestors. Can. J. Bot. 71, 512–518. https://doi.org/10.1139/b92-253

Hohmann, P., Messmer, M.M., 2017. Breeding for mycorrhizal symbiosis: focus on disease resistance. Euphytica 213, art. 113. https://doi.org/10.1007/s10681-017-1900-x

Hu-zhe, Z., Chun-lan, C., Yu-ting, Z., Dan, W., Yu, J., Yong, K.K., 2005. Active changes of lignification-related enzymes in pepper response to *Glomus intraradices* and/or *Phytophthora capsici*. J. Zhejiang Univ. Sci. B 6, 778–786. https://doi.org/10.1007/bf02842437

Kaeppler, S.M., Parke, J.L., Mueller, S.M., Senior, L., Stuber, C., Tracy, W.F., 2000. Variation among Maize Inbred Lines and Detection of Quantitative Trait Loci for Growth at Low Phosphorus and Responsiveness to Arbuscular Mycorrhizal Fungi. Crop Sci. 40, 358–364. https://doi.org/10.2135/cropsci2000.402358x

Kesavan, P.C., Malarvannan, S., 2010. Green to evergreen revolution: ecological and evolutionary perspectives in pest management. Curr. Sci. 99, 908–914.

Kuyper, T.W., Wang, X., Muchane, M.N., 2021. The Interplay Between Roots and Arbuscular Mycorrhizal Fungi Influencing Water and Nutrient Acquisition and Use Efficiency, in: Rengel, Z., Djalovic, I. (Eds.), The Root Systems in Sustainable Agricultural Intensification. pp. 193–220. https://doi.org/10.1002/9781119525417.ch7

Lanaud, C., Vignes, H., Utge, J., Valette, G., Rhoné, B., Caputi, M.G., Nieto, N.S.A., Fouet, O., Gaikwad, N., Zarrillo, S., Powis, T.G., Cyphers, A., Valdez, F., Nunez, S.Q.O., Speller, C., Blake, M., Valdez, F.Jr., Raymond, S., Rowe, S.M., Duke, G.S., Romano, F.E., Solórzano, R.G.L., Argout, X., 2024. A revisited history of cacao domestication in pre-Columbian times revealed by archaeogenomic approaches. Sci. Rep. 14, 2972. https://doi.org/10.1038/s41598-024-53010-6

Li, Y., Chen, H., Gu, L., Wu, J., Zheng, X., Fan, Z., Pan, D., Li, J., Shu, W., Rosendahl, S., Wang, Y., 2024. Domestication of rice may have changed its arbuscular mycorrhizal properties by modifying phosphorus nutrition-related traits and decreasing symbiotic compatibility. New Phyt. 243, 1554-1570. https://doi.org/10.1111/nph.19901

Livingstone, D., Stack, C., Mustiga, G.M., Rodezno, D.C., Suarez, C., Amores, F., Feltus, F.A., Mockaitis, K., Cornejo, O.E., Motamayor, J.C., 2017. A Larger Chocolate Chip—Development of a 15K *Theobroma cacao* L. SNP Array to Create High-Density Linkage Maps. Front. Plant Sci. 8, 2008. https://doi.org/10.3389/fpls.2017.02008

Martín-Robles, N., Lehmann, A., Seco, E., Aroca, R., Rillig, M.C., Milla, R., 2018. Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. New Phyt. 218, 322–334. https://doi.org/10.1111/nph.14962

McGrail, R.K., McNear, D.H., 2021. Two centuries of breeding has altered root system architecture of winter wheat. Rhizosphere 19, 100411. https://doi.org/10.1016/j.rhisph.2021.100411

Mukhlis, I., Rizaludin, M.S., Hidayah, I., 2022. Understanding Socio-Economic and Environmental Impacts of Agroforestry on Rural Communities. Forests 13, 556. https://doi.org/10.3390/f13040556

Nair, P.K.R., 2011. Agroforestry Systems and Environmental Quality: Introduction. J. Environ. Qual. 40, 784–790. https://doi.org/10.2134/jeq2011.0076

Noru, R.S.R., Thomas, B., Maraskole, S.K., S, V.P., Panotra, N., Rajesh, J., Karthickraja, A., Kumar, V., 2024. Participatory Plant Breeding: A Pathway to Sustainable and Resilient Agriculture. J. Adv. Biol. Biotechnol. 27, 1293–1306. https://doi.org/10.9734/jabb/2024/v27i81253

Paguntalan, D.P., Jr, I.E.B., Aggangan, N.S., 2020. Cacao – Arbuscular Mycorrhizal Fungi Association: A Review on Its Increased Potential in Improving the Ecology of Cacao Plantations in the Philippines. Plant Soil 183, 315–322. https://doi.org/10.1007/bf00011447

Power, J.F., Follett, R.F., 1987. Monoculture. Scientific American 256, 78–87. https://doi.org/10.2307/24979342

Reinhardt, D.R., Miller, R.M., 1990. Size classes of root diameter and mycorrhizal fungal colonization in two temperate grassland communities. New Phyt. 116, 129–136. https://doi.org/10.1111/j.1469-8137.1990.tb00518.x

Rosenstock, T.S., Dawson, I.K., Aynekulu, E., Chomba, S., Degrande, A., Fornace, K., Jamnadass, R., Kimaro, A., Kindt, R., Lamanna, C., Malesu, M., Mausch, K., McMullin, S., Murage, P., Namoi, N., Njenga, M., Nyoka, I., Valencia, A.M.P., Sola, P., Shepherd, K., Steward, P., 2019. A Planetary Health Perspective on Agroforestry in Sub-Saharan Africa. One Earth 1, 330–344. https://doi.org/10.1016/j.oneear.2019.10.017

Schmidt, J., Addo-Danso, S.D., Asare, R., Tettey, A., Isaac, M.E., 2024. Soil quality reflects microbial resource availability and drives rhizosphere microbiome variation in Ghanaian cocoa farms. Appl. Soil Ecol. 198, 105378. https://doi.org/10.1016/j.apsoil.2024.105378

Schmidt, J.E., DuVal, A., Isaac, M.E., Hohmann, P., 2022. At the roots of chocolate: understanding and optimizing the cacao root-associated microbiome for ecosystem services. A review. Agron. Sustain. Dev. 42, 14. https://doi.org/10.1007/s13593-021-00748-2

Schreefel, L., Schulte, R.P.O., Boer, I.J.M. de, Schrijver, A.P., Zanten, H.H.E. van, 2020. Regenerative agriculture – the soil is the base. Glob. Food Secur. 26, 100404. https://doi.org/10.1016/j.gfs.2020.100404

Sharma, N., Bohra, B., Pragya, N., Ciannella, R., Dobie, P., Lehmann, S., 2016. Bioenergy from agroforestry can lead to improved food security, climate change, soil quality, and rural development. Food Energy Secur. 5, 165–183. https://doi.org/10.1002/fes3.87

Singh, R., Singh, G.S., 2017. Traditional agriculture: a climate-smart approach for sustainable food production. Energy, Ecol. Environ. 2, 296–316. https://doi.org/10.1007/s40974-017-0074-7

Swaminathan, M.S., 2003. Bio-diversity: an effective safety net against environmental pollution. Environ. Pollut. 126, 287–291. https://doi.org/10.1016/s0269-7491(03)00241-0

Swaminathan, M.S., 2006. An Evergreen Revolution. Crop Sci. 46, 2293–2303. https://doi.org/10.2135/cropsci2006.9999

Swaminathan, M.S., 2012. Agroforestry for an Evergreen Revolution. Agroforestry - The Future of Global Land Use. Adv. Agrofor. 9, 7–10. https://doi.org/10.1007/978-94-007-4676-3_2

Valverde-Barrantes, O.J., Freschet, G.T., Roumet, C., Blackwood, C.B., 2017. A worldview of root traits: the influence of ancestry, growth form, climate and mycorrhizal association on the functional trait variation of fine-root tissues in seed plants. New Phyt. 215, 1562–1573. https://doi.org/10.1111/nph.14571

Valverde-Barrantes, O.J., Horning, A.L., Smemo, K.A., Blackwood, C.B., 2016. Phylogenetically structured traits in root systems influence arbuscular mycorrhizal colonization in woody angiosperms. Plant Soil 404, 1–12. https://doi.org/10.1007/s11104-016-2820-6

Wang, X., Zhang, M., Sheng, J., Feng, G., Kuyper, T.W., 2023. Breeding against mycorrhizal symbiosis: Modern cotton (*Gossypium hirsutum* L.) varieties perform more poorly than older varieties except at very high phosphorus supply levels. J. Integr. Agric. 22, 701–715. https://doi.org/10.1016/j.jia.2022.08.004

Wang, Y., Li, Z., Wang, S., Wang, W., Wang, N., Gu, J., 2022. Variations in Arbuscular Mycorrhizal Colonization Associated with Root Diameter and Hypodermis Passages Cells across Temperate and Tropical Woody Species. Forests 13, 140. https://doi.org/10.3390/f13020140

Wei, X., Xie, B., Wan, C., Song, R., Zhong, W., Xin, S., Song, K., 2024. Enhancing Soil Health and Plant Growth through Microbial Fertilizers: Mechanisms, Benefits, and Sustainable Agricultural Practices. Agronomy 14, 609. https://doi.org/10.3390/agronomy14030609

WorldBank, 2024. The World by Income and Region [WWW Document]. URL https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html

WorldBank, 2025. The World Bank in Panama [WWW Document]. URL https://www.worldbank.org/en/country/panama/overview

YCharts, 2025. Cocoa Price (I:CPNM) [WWW Document]. URL https://ycharts.com/indicators/world_cocoa_price (accessed 8.16.25).

Summary

This story starts with invisible threads in the soil, that are connected to almost all plants through the roots. These beings are called Mycorrhizal Fungi, in particular these ones I'll tell you about are called Arbuscular Mycorrhizal Fungi (AMF), because of the arbuscules, tree-like structures they make inside the root. These beings associate with almost(!) all plants on Earth, since 400 million years ago, and even helped change the atmosphere of the Earth by helping plants colonize all land ecosystems (that's right, before there were only aquatic plants). These beings help the plants in many ways. By extending the reach of their roots they can acquire more nutrients and water that the plant needs, and the relationship itself prepares the plant to fight off natural enemies. I study these beings and their association with the cacao trees. These cacao trees only grow in the tropics, in jungle-like ecosystems, and produce cacao beans, from which a wonderful treat is produced. This treat, chocolate, nourishes the heart and the soul. People love so much eating chocolate, that they eat in many forms, mixtures with other flavors, and even make sculptures with them. But now, there is a problem with the cacao cultivation.

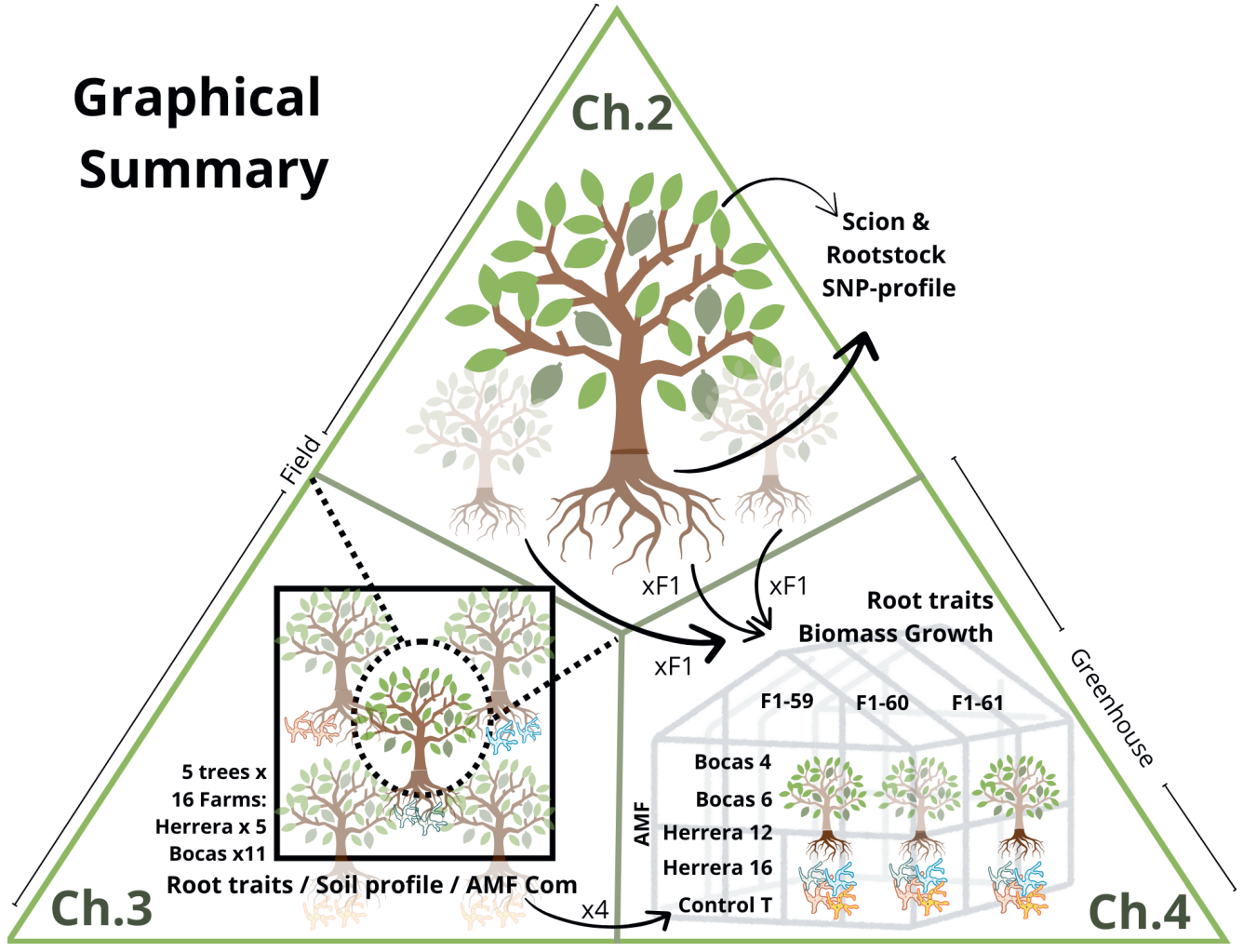
The humans, our kind, have stopped caring for the Earth, in the quest of survival, or at least most of them. They, or should I say 'we', found a way of producing a lot of food, but it meant to cut down a lot of forest, use toxic chemicals to kill the insects that want to eat all the food we produce, and give mineral nutrients in excess to the plants so they are extra-large. This way of producing food, in addition to many other ways of living, have increased the temperature on our Earth and now the climate is changing… Seasons are shifting, and plants, along with other organisms on earth, are trying to find a new rhythm of living with the Earth. In finding this new rhythm, trees lose their fruits, because weather conditions are not ideal, either by too much or too little rain or heat. This has caused a scarcity of cocoa and therefore of chocolate. I thought… may be in these helpful relationships of plants with the AMF we can find the key to help the plants find the new rhythm with the Earth. Me and others with me, went on a journey to answer some questions, and think of what could help.

During the first chapter of this journey, we went to these cacao farms and investigated the cacao trees growing in two different regions of Panama: Bocas del Toro and Herrera. In our expedition we looked at the genetic code in these trees and found that trees in Bocas del Toro were of very mixed ancestry, with all ten major genetic groups, while in Herrera there were mainly two groups, Amelonado and Criollo. We also learned if these farmers sowed seeds or used a technique called grafting, to develop new trees. Grafting is to mix two different trees, one aboveground (scion) with another belowground (rootstock), into one. More than half of the trees in Bocas del Toro and only a fourth in Herrera were grafted. We noticed that nitrogen and the level of acidity (pH) in the soil were associated with genetic differences across all trees.

As we moved to the next chapter, we thought about the AMF. Which are the species associated with these trees? Does the cacao tree genetics, root architecture, or even the environment shapes these AMF communities? We found the genus *Glomus* to be dominant across all samples, and we did not see differences in the AMF communities between regions, even though they varied in many environmental characteristics. The aboveground (scion) genetics of the plant, root nutrient (nitrogen and phosphorus) concentration, elevation, and available phosphorus in the soil were associated with shaping the species in AMF communities. We also found specific AMF species associated with these key variables.

In the last chapter of our journey, we grew some of the cacao trees from the first chapter, selfed progenies of three productive and sturdy trees to be precise, with four AMF soils from the second chapter and with a non-AMF soil, to observe how the root traits and growth of the seedling trees, changed with different AMF communities. We found great differences in plant growth between AMF soils, plant progenies and interactive effects between these, so some progenies (S_1-60) responded stronger than others to a particular AMF treatment (H3). For root traits, root nitrogen and root diameter were different between progenies and AMF, and for root tissue density and specific root length we found these to be different only between AMF soils, particularly with and without AMF.

After this whole journey, we had a lot to reflect about. We learned not only about cacao and AMF, but also about the people and culture involved. We understood the need farmers have for stronger social and economic structure that supports them to develop and apply available technologies for their plants, farm and resulting product. This new structure would guide cacao breeding, working with the farmers, to improve yield and product quality, while also ensuring fair wages that provide a dignified lifestyle for farmers.


Graphical Summary
Ch.2
Scion & Rootstock SNP-profile
Field
xF1
xF1
xF1
Root traits Biomass Growth
Greenhouse
F1-59
F1-60
F1-61
Bocas 4
Bocas 6
Herrera 12
Herrera 16
Control T
AMF
5 trees x 16 Farms: Herrera x 5 Bocas x11
Root traits / Soil profile / AMF Com
x4
Ch.3
Ch.4

Resumen

Esta historia comienza con hilos invisibles en el suelo, que están conectados a casi todas las plantas a través de las raíces. Estos seres se llaman hongos micorrízicos, y en particular sobre los que les voy a contar se llaman hongos micorrízicos arbusculares (HMA), debido a los arbusculos, estructuras similares a árboles que forman dentro de la raíz. Estos seres se asocian con casi (!) todas las plantas de la Tierra desde hace 400 millones de años, e incluso ayudaron a cambiar la atmósfera de la Tierra al apoyar a las plantas a colonizar todos los ecosistemas terrestres (sí… antes solo había plantas acuáticas). Estos seres ayudan a las plantas de muchas maneras. Al extender el alcance de sus raíces, pueden adquirir más nutrientes y agua que la planta necesita, y la relación en sí misma prepara a la planta para luchar contra sus enemigos naturales. Yo estudio estos seres y su asociación con los árboles de cacao. Estos árboles solo crecen en los trópicos, en ecosistemas similares a la selva, y producen granos de cacao, a partir de los cuales se elabora un manjar maravilloso. Este manjar, el chocolate, nutre el corazón y el alma. A todo el mundo le encanta tanto comer chocolate que lo consumen de muchas formas, mezclado con otros sabores, e incluso hacen esculturas con él. Pero ahora hay un problema con el cultivo del cacao.

Los humanos, nuestra especie, hemos dejado de cuidar la Tierra en nuestra búsqueda por sobrevivir, o al menos la mayoría de nosotros. Ellos, o debería decir «nosotros», encontramos una forma de producir mucha comida, pero eso significó talar muchos bosques, usar químicos tóxicos para matar a los insectos que se comen toda la comida que producimos y dar nutrientes minerales en exceso a las plantas para que crezcan más grandes. Esta forma de producir alimentos, además de muchas otras cosas en nuestro estilo de vida, ha aumentado la temperatura de nuestra Tierra y ahora el clima está cambiando... Las estaciones cambian y las plantas, junto con otros organismos de la Tierra, están tratando de encontrar un nuevo ritmo de vida con la Tierra. En este intento, los árboles pierden sus frutos, ya que las condiciones climáticas no son ideales, por (exceso o falta de) lluvia y calor. Esto ha provocado una escasez de cacao y, por lo tanto, de chocolate. Pensé... tal vez en estas relaciones beneficiosas de las plantas con los hongos micorrízicos arbusculares podamos encontrar la clave para ayudar a las plantas a encontrar su nuevo ritmo con la Tierra. Otras personas que me acompañaban y yo emprendimos un viaje para responder a algunas preguntas y pensar en lo que podría ayudar.

Durante la primera etapa de este viaje, visitamos estas fincas de cacao e investigamos los árboles de cacao que crecen en dos regiones diferentes de Panamá: Bocas del Toro y Herrera. En nuestra expedición, analizamos el código genético de estos árboles y descubrimos que los árboles de Bocas del Toro tenían orígenes muy variados, con los diez grupos genéticos principales, mientras que en Herrera había principalmente dos grupos, el Amelonado y el Criollo. También averiguamos si estos

agricultores sembraban semillas o utilizaban una técnica llamada injerto para desarrollar nuevos árboles. El injerto consiste en mezclar dos árboles diferentes, uno sobre el suelo (vástago) y otro bajo el suelo (portainjerto), en uno solo. Más de la mitad de los árboles de Bocas del Toro y solo una cuarta parte de los de Herrera eran injertados. Observamos que el nitrógeno y el nivel de acidez (pH) del suelo estaban relacionados con las diferencias genéticas entre todos los árboles.

Al pasar al siguiente capítulo, pensamos en los hongos micorrízicos arbusculares (HMA). ¿Cuáles son las especies asociadas a estos árboles? ¿La genética del árbol del cacao, la arquitectura de sus raíces o incluso el ambiente dan forma a estas comunidades de HMA? Descubrimos que el género *Glomus* era el dominante en todos los árboles y no observamos diferencias en las comunidades de HMA entre las regiones, a pesar de que estas variaban en muchas características ambientales. La genética aérea (vástago) de la planta, la concentración de nutrientes en las raíces (nitrógeno y fósforo), la altitud y el fósforo disponible en el suelo estaban relacionados con la composición de las especies en las comunidades de HMA. También encontramos especies específicas de HMA asociadas a estas variables clave.

En el último capítulo de nuestro viaje, sembramos árboles de cacao del primer capítulo, descendientes autopolinizados de tres árboles productivos y robustos, con cuatro suelos que contenían HMA del segundo capítulo y un suelo sin HMA, para observar cómo cambiaban las características de las raíces y el crecimiento de los árboles jóvenes. Encontramos grandes diferencias en el crecimiento de las plantas entre los suelos con HMA, las progenies de las plantas y los efectos interactivos entre ellos. O sea, algunas progenies (S1-60) respondieron con mayor intensidad que otras a un tratamiento específico con HMA (H3). En cuanto a las características de las raíces, el nitrógeno y el diámetro de las raíces diferían entre las progenies y los HMA, y en cuanto a la densidad del tejido radicular y la longitud específica de las raíces, observamos que solo diferían entre los suelos con HMA y sin HMA.

Después de todo este viaje, tuvimos mucho en qué reflexionar. Aprendimos no solo sobre el cacao y los hongos micorrízicos arbusculares, sino también sobre las personas y la cultura involucradas. Comprendimos la necesidad de que los agricultores cuenten con una estructura social y económica más sólida que los apoye en el desarrollo y la aplicación de las tecnologías disponibles para sus plantas, sus granjas y los productos resultantes. Esta nueva estructura guiaría la creación de nuevas líneas de cacao, trabajando con los agricultores, para mejorar el rendimiento y la calidad del producto, al tiempo que garantizaría precios justos que proporcionen un estilo de vida digno a los agricultores.

The research described in this thesis was financially supported by National Secretariat of Science, Technology and Innovation (SENACYT).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover and art design by Arielys Lorreine Martinez Villalaz

