Developing Pre-Breeding Tools for the Miracle Plant [*Synsepalum dulcificum* (Schumach &Thonn.) Daniell]: Implications for Genomic Selection Strategies Optimization

by

Dèdéou Apocalypse Tchokponhoué

MSc. Biotechnology and Plant Breeding (University of Abomey-Calavi, Rep. Benin)

BSc. Crop Science (University of Abomey-Calavi, Rep. Benin)

A thesis submitted in fulfilment of the academic requirements for the degree of

Doctor of Philosophy (PhD) in Plant Breeding

Discipline of Plant Breeding

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science
University of KwaZulu-Natal
Republic of South Africa

February 2022

ABSTRACT

The miracle plant [*Synsepalum dulcificum* (Schumach & Thonn) Daniell] is an evergreen orphan shrub species originating in West Africa that produces red berries known as miracle fruits. As a unique natural source of miraculin, miracle fruits stand as a singular natural sweetener with huge applications and economic potential. Unfortunately, the lack of systematic breeding history in the species currently hinders the full exploitation of the species potential at both the local and international levels. This study was undertaken with the goal of generating information that can be used to initiate a proper breeding and large-scale cultivation program of the species in its center of origin, West Africa, with a focus on three countries (Ghana, Togo, and the Republic of Benin). The objectives of the study were: i) to determine current management practices and end-users' preferred traits to identify key breeding traits and formulate sound breeding objectives, ii) to evaluate the natural phenotypic diversity in the species to identify potential parental lines/populations for breeding purposes, iii) to assess genomic diversity and population structure in the species to reshape breeding strategies, and iv) to model the determinants of the local communities' willingness to invest in the species cultivation to articulate promotion strategies in the species.

To assess trait preferences by the miracle plant end-users (farmers, consumers, and processing companies) in the study area, semi-structured interviews and focus-group discussions were held with 300 individual respondents from various socio-cultural backgrounds and one processing company each in Benin and Ghana. The results revealed that farmers in Ghana currently manage the species better than their counterparts in Benin, with men being overall the main owners of the species. Our results offer an impetus for West Africa region-wide cultivar development, as both farmers' and consumers' preferences for breeding traits were >80% similar among the six sociolinguistic groups and the three agroecological zones considered in the study area. The preference for breeding traits was 60% similar among farmer, consumer and processing company user groups. Out of the 23 potential breeding traits identified in the study area, fruit size, fruit miraculin content, fruit yielding, early maturity, fruit edible ratio and seed part ratio represent priority breeding traits. Interestingly the traits (fruit size and miraculin content) were consistently ranked as the top two desired traits by all three end-user groups of the species. The classification conducted on 203 individual trees (accessions) sampled from seven populations in the Upper Guinea forest block (Ghana) and the Dahomey Gap (Ghana, Togo, and Benin) block revealed the existence of three natural phenotypic groups in the species. While in terms of population performance, the Volta population in the Dahomey Gap outperformed the six other populations for almost all the end-users' desired traits assessed. Cluster 3, which encompassed accessions from all seven populations sampled, presented the best performance based on end-users' preferences and, therefore, represents a potential source of elite accessions. Overall, variability ranged from low to moderate in traits evaluated [coefficient of variation: 9.65% (fruit length) – 53.34% (tree diameter at ground level] and the inferred core collection points to the Dahomey Gap as a center of diversity of the miracle plant. Despite the expected heterozygosity of H_S = 0.14 suggesting a low diversity in the miracle plant, the molecular assessment done on 322 individuals revealed a strong population differentiation in the species (F_{ST} =0.69). Populations in the Upper Guinea forest block strongly diverged from those in the Dahomey Gap block, and our data supported the existence of a barrier (which we speculated to be the Lake Volta in Ghana) disrupting connectivity between those two West African rainforest block populations of the miracle plant. Corroborating the findings of the phenotypic assessment, the genome-wide diversity analysis also defined three genetic groups. Conveying a higher resolution to the population structuring, the genomic analysis highlighted an ecoregional structuring of the miracle plant population in the study area. The DNA polymorphism and demographic history analyses indicated the Western Dahomey Gap as the likely centre of origin of the miracle plant, thus refining the previously postulated West Africa origin of the species. The Volta population was confirmed as an exceptional breeding population with regards to its high allelic richness and heterozygosity. Our findings also suggest that sociolinguistic groups in Benin valued the miracle plant more than those in Ghana, to the extent that the species diversity was affected, as fitness organs such as roots and bark were highly targeted for medicinal uses. Nonetheless, these sociolinguistic groups in Benin were also more willing to engage in the miracle plant cultivation compared with those in Ghana. However, for this to be effective, market structuring and improved variety (early fruiting) development combined with awareness raising are needed to accelerate local community engagement in the species cultivation in the entire study area.

Collectively, these results advance our knowledge of the miracle plant and form the basis for breeding initiatives in the species. The implications of these findings regarding optimization of a genomic selection strategy for accelerated genetic gains achievement in agronomic and functional traits of this species is discussed.

RESUME

Le plant miraculeux *Synsepalum dulcificum* (Schumach & Thonn.) Daniell est un arbuste sempervirent négligé originaire de l'Afrique de l'Ouest qui produit des baies rouges connues sous l'appellation fruit miraculeux. En tant qu'unique source naturelle de la « miraculine », le fruit miraculeux a de nombreuses applications associées à un énorme potentiel économique. Cependant, l'inexistence d'une histoire d'amélioration sur l'espèce freine actuellement la pleine exploitation du potentiel de l'espèce tant aux niveaux local qu'international. Ce travail a été initié dans l'optique de générer des informations pouvant servir de base à la mise en place d'un programme d'amélioration et de culture à grande échelle de l'espèce dans son aire d'origine, l'Afrique de l'Ouest, avec un accent sur trois pays à savoir le Ghana, le Togo et le Bénin. Les objectifs du travail étaient: i) d'élucider les pratiques actuelles de gestion de l'espèce et les préférences des utilisateurs finaux afin d'identifier les principaux traits d'amélioration pour formuler des objectifs d'amélioration pertinents, ii) d'explorer la diversité phénotypique naturelle de l'espèce afin d'identifier les lignées parentales/populations élites pour des fins d'amélioration, iii) d'évaluer la diversité moléculaire et la structure de la population de l'espèce afin de peaufiner les stratégies d'amélioration, et iv) de modéliser les déterminants de la volonté des communautés locales à s'investir dans la culture du plant miraculeux afin d'articuler les stratégies de promotion de l'espèce.

Pour évaluer les préférences des utilisateurs finaux (producteurs, consommateurs et transformateurs) du plant miraculeux dans le milieu d'étude, des entretiens semi-structurés et des discussions de groupe avaient été conduits avec 300 répondants appartenant à divers groupes sociolinguistiques du Bénin et du Ghana ainsi qu'avec une entreprise de transformation industrielle au Ghana. Les résultats ont révélé que les agriculteurs ghanéens gèrent actuellement mieux l'espèce que leurs homologues béninois, les hommes étant globalement les principaux propriétaires de l'espèce. Nos résultats donnent une impulsion au développement de cultivars à l'échelle de la région Ouest-Africaine, car les préférences des producteurs et des consommateurs pour les traits d'amélioration étaient > 80% similaires entre les six groupes sociolinguistiques et les trois zones agroécologiques considérés dans l'étude. La préférence pour les traits d'amélioration était similaire à 60% entre les producteurs, les consommateurs et les entreprises de transformation. Sur les 23 potentiels traits d'amélioration identifiés dans la zone d'étude, la taille des fruits, la teneur des fruits en miraculine, le rendement en fruits, la précocité de la maturité, le ratio comestible du fruit et le ratio masses graine/fruit représentaient les traits de sélection prioritaires. Fait intéressant, les caractères taille du fruit et teneur en miraculine ont été systématiquement classés comme les deux principaux caractères souhaités par les trois groupes d'utilisateurs finaux de l'espèce.

La classification effectuée sur 203 arbres individuels (accessions) échantillonnés à partir de sept populations du bloc forestier Guinéen (Ghana) et du Dahomey Gap (Ghana, Togo et Bénin) a révélé l'existence de trois groupes phénotypiques naturels au sein de l'espèce. En termes de performance, la population de la Volta dans le Dahomey Gap a surpassé les six autres populations pour presque tous les traits désirés par les utilisateurs finaux évalués. Le groupe 3, qui englobait des accessions des sept populations échantillonnées, a présenté les meilleures performances sur la base des préférences des utilisateurs finaux et, par conséquent, représente une source potentielle d'accessions élites. Dans l'ensemble, la variabilité variait de faible à modérée pour les traits évalués [coefficient de variation : 9,65% (longueur du fruit) - 53,34% (diamètre de l'arbre au niveau du sol)] et la collection noyau déduite indique que le Dahomey Gap est un centre de diversité du plant miraculeux. Malgré l'hétérozygotie attendue HS de 0,14 suggérant une faible diversité chez le plant miraculeux, l'évaluation moléculaire effectuée sur 322 individus a toutefois révélé une forte différenciation des populations de l'espèce (FST = 0,69). Les populations du bloc forestier Guinéen divergeaient fortement de celles du Dahomey Gap, et nos données soutiennent l'existence d'une barrière (que nous supposions être le lac Volta au Ghana) perturbant la connectivité entre ces deux populations du plant miraculeux dans la forêt humide tropicale de l'Afrique de l'Ouest. Corroborant les résultats de l'évaluation phénotypique, l'analyse de la diversité à l'échelle du génome a également défini trois groupes génétiques. En apportant une meilleure résolution à la structuration des populations, l'analyse génomique a mis en évidence une structuration écorégionale des populations du plant miraculeux dans la zone d'étude. Les analyses du polymorphisme de l'ADN et de l'histoire démographique ont indiqué l'Ouest du Dahomey Gap comme le centre d'origine probable du plant miraculeux, affinant ainsi l'origine précédemment postulée de l'Afrique de l'Ouest de l'espèce. La population de la Volta a été confirmée comme une population d'amélioration exceptionnelle vue sa grande richesse allélique et son hétérozygotie élevée. Nos résultats suggèrent également que les groupes sociolinguistiques au Bénin accordaient plus d'importance au plant miraculeux que ceux au Ghana, au point où la diversité génétique au sein des individus du plant miraculeux dans la zone s'en trouverait affectée. En effet, les organes tels que les racines et l'écorce sensés assurer la pérennité de l'espèce étaient fortement ciblés à des fins médicinales. Néanmoins, ces groupes sociolinguistiques au Bénin étaient également plus disposés à s'engager dans la culture de l'espèce par rapport à ceux du Ghana. Cependant, pour que cela soit effectif, la structuration du marché et le développement de variétés améliorées (à fructification précoce) combinés à une sensibilisation sont nécessaires pour accélérer l'engagement de la communauté locale dans la culture de l'espèce dans toute la zone d'étude.

Pris ensembles, ces résultats font progresser notre connaissance du plant miraculeux et constituent une base solide pour des initiatives de sélection de l'espèce. Les implications de ces résultats en termes d’optimisation d'une stratégie de sélection génomique pour la réalisation rapide de gains génétiques par rapport aux traits agronomiques et fonctionnels de cette espèce sont discutées.

DECLARATION 1 - PLAGIARISM

I, Dèdéou Apocalypse Tchokponhoué, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons’ data, pictures, graphs, or other information unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

i) Their words have been rewritten, but the general information attributed to them has been referenced.

ii) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references’ sections.

Signed:

[redacted]

Dèdéou Apocalypse Tchokponhoué

As the candidate’s supervisors, we agree to the submission of this thesis

[redacted]

Prof Julia Sibiya (Supervisor)

[redacted]

Prof Enoch. G. Achigan-Dako (Co-Supervisor)

DECLARATION 2 - PUBLICATIONS AND MANUSCRIPTS

Three chapters (Chapter 3, Chapter 4, and Chapter 6) have already been published from this thesis and one chapter is currently under review. The published manuscripts and that under review, their authors and the publication journals details read as follows:

1- **Tchokponhoué, D.A.**, Achigan-Dako, E.G., N'Danikou, S., Nyadanu, D., Kahane, R., Houéto, J., Hotegni, N.V.F., Odindo, A.O. and Sibiya, J. **2020**. Phenotypic variation, functional traits repeatability and core collection inference in *Synsepalum dulcificum* (Schumach & Thonn.) Daniell reveals the Dahomey Gap as a centre of diversity. ***Scientific Reports,*** 10 (19538). https://doi.org/10.1038/s41598-020-76103-4. [Chapter 4].

2- **Tchokponhoué, D.A.**, Achigan-Dako, E.G., N'Danikou, S., Nyadanu, D., Kahane, R., Odindo, A.O. and Sibiya, J. **2021**. Comparative analysis of management practices and end-users' desired breeding traits in the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] across ecological zones and sociolinguistic groups in West Africa. ***Journal of Ethnobiology and Ethnomedicine,*** 17 (41). https://doi.org/10.1186/s13002-021-00467-8 [Chapter 3].

3- **Tchokponhoué, D.A.,** N'Danikou, S., Fassinou Hotegni, N.V., Nyadanu, D., Kahane, R., Odindo, A.O., Achigan-Dako, E.G. and Sibiya, J. **2021**. Use patterns, knowledge diversity and drivers for the cultivation of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] in Benin and Ghana. ***Plants***, 10 (2253). https://doi.org/10.3390/plants10112253 [Chapter 6].

4- **Tchokponhoué, D.A.,** Achigan-Dako, E.G., N'Danikou, S., Nyadanu, D., Hale, L.I., Odindo, A.O. and Sibiya, J. Genome-wide diversity analysis suggests divergence among West African populations of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell], with Western Dahomey Gap as the centre of origin. [Chapter 5, Under Review].

ACKNOWLEDGEMENTS

I immensely thank God for his mercy, paving my way towards this crop and giving me the stamina required to conduct this thesis research in three West African Countries.

I would like to thank my supervisors Prof Julia Sibiya and Prof Enoch G. Achigan-Dako from the bottom of my heart for their availability and guidance throughout this research. My heartfelt gratitude to you, supervisors. Be blessed!

I place on record the financial support of the MoBreed (Enhancing training and research mobility for novel crop breeding in Africa) project, a program under the Intra – Africa Mobility scheme funded by the Education, Audiovisual and Culture Executive Agency (EACAE) of the European Union Commission to this research. My appreciation to the whole project coordinating Board. I am also indebted to the Alliance for a Green Revolution in Africa (AGRA) for the provision of funds through the Improved Masters in Cultivar Development for Africa (IMCDA) programme at UKZN and the International Foundation for Science (Grant C-6592-1) that enabled me to conduct the genome-wide diversity analysis of the species.

I am also thankful to Dr Sognigbé N'Danikou, Dr Daniel Nyadanu, Dr Rémi Kahane, Dr Nicodème V. Fassinou Hotegni, Prof Alfred Odindo, Prof Iago Hale and Prof Samson Tesfay for their comments, advice, and contributions to this thesis.

I am indebted to a few people who made my field work smoother than expected. These include Mr Jacob Houéto, Mr Emile Sogbédji, Dr Yao Dagnon, Mr Yakubu Abdulai, Mr Peter Gebra, Mr, Mr Rachidi Francisco, Mr Kocou Houéhounko, Mr Jordan Houégban, Mr Malachi Houinato, Mr Frizel Awéha, Mr Léocade Azonhoumon, Mr Borduas Gambadatoun, Mr Mathieu Ayénan and Mr Xavier Matro. Thank you so much!

Thanks also go to the numerous farmers, local consumers and miracle fruit processing companies in Benin, Togo and Ghana who generously contributed to our research with their time. Thank you for being such excellent custodians of our local knowledge.

From the field to data analyses, I am very grateful to all R package developers. Your libraries have been instrumental in attaining the results reported in this thesis. Wish to your community more inspiration.

I was very lucky to meet my colleague Aristide Houdégbé in 2011. Since then, I have experienced the truth of the saying that "brotherhood is beyond the simple family link". Thank you for also being my PhD mate for these four years, and for all the insightful discussions, we always held around our respective manuscripts.

I thank the University of KwaZulu-Natal for offering a conducive environment for my research. More greatness for you!

Special thanks to Andile Etah Mshengu for going an extra mile in assisting with administrative matters.

Much appreciation to my fellow postgraduate students; Carlos, Nelile, Emeline, Takudzwa, Athenkosi, Sharon, Nqobile, Nkosinathi, Dela, and Dr Bernice Bancole, Dr Zama Mbulawa, Mrs Clarisse Massamba and Mrs Odindo. You have been my family here in South Africa. Thanks for all the moments spent together.

I would like to thank all my colleagues and the entire staff of the Laboratory of Genetics, Biotechnology and Seed Science (GBioS) of the University of Abomey-Calavi, Republic of Benin, for all their encouragement. May God continue to strengthen our laboratory.

Thanks also go to Dr Dêêdi Sogbohossou, Mrs Lydia Sognon, Wilfried Ahouansou and Mr Hervé Kouton for their contribution to the MoBreed Project drafting and implementation.

Much appreciation to my fellows Armand, Boris, Curie, Faki, Hamed, Hermane, Issiakou, Mandzidath, Moubarakath and Micklette for their assistance and friendship.

I pay a special tribute to the Late Koffi Anglow from whom I discovered the deepest sense of wisdom. We were meeting for the first time that End, June 2019, but this did not preclude you from waiting until very late at the bus station, waiting for my arrival. You spontaneously, even without knowing me, hosted me, and fed me for days. I can still revisit in memory sequences of our very first and quite long discussion on your miracle plant plantation and am still enlightened by your vision and hope for the miracle plant. I will always have in mind your passion for the miracle plant. Rest Eternally in Peace.

My sincere appreciations to all those who either directly or indirectly contributed to this work but were not duly named, may God bless you.

DEDICATION

To my parents, for their presence and prayers.

To Aboègnonhou, my beloved, for her patience and continuous support.

To Dêêdi, my daughter, for being a unique source of joy and motivation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Rationale

Africa is endowed with rich plant diversity estimated at nearly 40,000 - 60,000 species (Linder, 2001), which holds the potential to contribute to food and nutrition security and poverty alleviation in the region. These plant resources have been used as sources of foods, medicines, fibers, and fuels while providing other services (e.g., ecosystem services) (Sinare and Gordon, 2015) and, thus, constitute a key asset for the achievement of the Sustainable Development Goals (SDGs) (Pilling et al., 2020). Nevertheless, a major bottleneck to the full exploitation of most of this diversity is the paucity of knowledge on research and investment in these species. Due to their status as locally important but poorly studied and funded crops, these species are termed orphan, minor, underutilized or neglected (Padulosi et al., 2021; Tadele, 2009).

The local importance of minor plant species was previously acknowledged (Fentahun and Hager, 2009; Kobori and Amaya, 2008), but this importance has been, today more than ever, revived. These crops have been increasingly receiving attention worldwide, as they offer many alternatives for agricultural, food and nutrition, and ecological system transformation, especially in the context of climate change (Borelli et al., 2020; Hunter et al., 2019; Mabhaudhi et al., 2019). These so-called minor species that have not undergone the domestication bottleneck still have a large reservoir of both biotic and abiotic stress resistance and yield genes that can be tapped for breeding purposes (Kibret, 2021). Consequently, how to sustainably promote orphan crops has become a central research question (Dawson et al., 2009; Nyadanu and Lowor, 2015), and an example of such a crop is the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell].

The miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] [Syn. *Richardella dulcifica* (Schumach & Thonn.) Baehni] is a West African orphan shrub species belonging to the Sapotaceae family. It is a multipurpose species whose fruits, known as miracle berries, are a unique natural source of “miraculin”, a sweetening glycoprotein (Kurihara and Beidler, 1968). Owing to the taste modifying property of the fruit, the species stands as a natural sweetener historically valued in West Africa, as our forefathers employed the fruits to consume sour foods (Inglett and May, 1968). This knowledge has been transmitted to the current generations (Oumorou et al., 2010). In parallel, miracle berries are sold on local markets and contribute substantially to household income generation (Fandohan et al., 2017), while internationally, a kilogram of the dried fruit pure powder costs up to USD 2,500 (https://www.miraclefruitfarm.com/).

This high international economic value of the species may be ascribed to the numerous modern applications of its fruits (Achigan-Dako et al., 2015). On one hand, the miracle fruits are used in cancer treatment, where they help restore taste perception in patients undergoing chemotherapy treatment (Wilken and Satiroff, 2012). On the other hand, the miracle fruits have the ability to serve as a natural sweetener and colorant in the food and beverage industry (Buckmire and Francis, 1978), a potential that seems to be under realization, as today, evidence of the interest by the Coca-Cola company for the species exists (Tafazoli et al., 2019). More importantly, the miracle fruits have the potential to be used in diabetes and overweight treatment, as the fruits can regulate insulin production (Chen et al., 2006). In parallel, the singular property of the miraculin of not being sweet by itself but being able to induce sweetness makes the miracle plant suitable for the intensive lifestyle modification approach, which is the most advocated approach (Nathan et al., 2007) for diabetes prevention and characterized by the use of a low-fat hypocaloric diet (Ramachandran et al., 2006; Tuomilehto et al., 2001). Consequently, the miracle plant has the potential to significantly lower the economic burden related to diabetes care, which is estimated, for instance, at USD 966 billion in 2021 (International Diabetes Federation, 2021). Other parts of the miracle plant (leaf, root, stem) were also reported to be key sources of unique phytochemicals sought after in the pharmaceutical industries and cosmetics (Du et al., 2014). Collectively, all these uses suggest that the miracle plant is one of the most promising orphan fruit tree crops, with numerous potential end-users, including farmers, consumers, and industrial companies.

Miraculin, a glycoprotein produced by the miracle plant has been intensively studied with topics ranging from the characterization of its structural conformation to its genetic transformation (Misaka, 2013; Paladino et al., 2008; Tafazoli et al., 2019; Takahashi et al., 1990; Theerasilp and Kurihara, 1988). Similarly, many studies have been conducted on the miracle plant nutritional potential and phytochemical content (He et al., 2016; Njoku et al., 2015; Nkwocha et al., 2014; Obafemi et al., 2019a; Obafemi et al., 2019b; Wang et al., 2011; Wang et al., 2020). However, studies tackling a better understanding of the species production, including its biology (Tchokponhoué et al., 2019a), horticulture (Demesyeux et al., 2020; Tchokponhoué et al., 2018; Tchokponhoué et al., 2019b), genetics and breeding are rather limited. Yet, proper cultivation supported by the use of elite planting materials is required to sustain the numerous utilizations of the species. Recent developments in the crop as highlighted above have shifted the interest on the miracle plant from only the miraculin to the whole plant, as all the species parts are of interest (Achigan-Dako et al., 2015).

Consequently, the transgenic breeding path laid out for the species and that specifically targets the miraculin appears today as a minimalistic approach that needs to be revised in light of the emergent genomics-assisted breeding approach to improve the set of traits desired by end-users. In addition, progress so far made in nearly ten years of recombinant miraculin expression in various organisms [e.g., tomato (*Lycopersicon esculentum* (L.) H.Karst.) (Hiwasa-Tanase et al., 2011; Ono et al., 2021); lettuce (*Lactuca sativa* L.) (Hirai et al., 2011); strawberry (*Fragaria* X *Annasa*) (Sugaya et al., 2008), carrot (*Daucus carota* L.) (Park et al., 2020; Park et al., 2021) and rice (*Oryza sativa* L.) (Kim et al., 2020)] has resulted in a yield range of 340 µg/g - 1.24 mg/g fresh weight, which remains up to 10 times lower than the natural miraculin production in miracle berries (1.30 mg/g of juice) (Demesyeux et al., 2020). In addition, the miraculin expression in transgenic plants may be lost over generations (Sun et al., 2006), while there is a high cost and technical skill required to implement the genetic transformation technology. Jointly, these observations raise questions on the cost-effectiveness of the transgenic breeding approach in the miracle plant, especially in the tropics where natural conditions are suitable for the production of the species. This, therefore, opens the room for exploring genomics-assisted breeding in the species as a way to not only improve the miraculin content but also to improve other traits of interest for the end-users.

The miracle plant is naturally distributed in West and Central Africa, from Ghana to Congo; however, the countries including Ghana, Togo and Benin have been implicated as the species center of origin (Achigan-Dako et al., 2015). The species has a wide range of uses and end-users, which suggest that many breeding traits may be of interest. Consequently, prioritizing the potential breeding traits and identifying the most salient traits will be required for proper formulation of breeding objectives to develop a widely accepted variety. Regrettably, no study has to date addressed trait preferences by end-users in this species. Furthermore, information on phenotypic variability and molecular variation, although fundamental for any breeding program, is currently lacking in the miracle plant. To the best of our knowledge, only one study, limited to southern Nigeria, employed random amplified polymorphic DNA (RAPD) markers on the species (Chibuzor et al., 2017), but no report of any phenotypic variation that can be used as a landmark existed to date. Despite the numerous applications of the species, no evidence has been provided for the species cultivation, at least from our empirical observation in Benin. Therefore, understanding the factors affecting the local communities' willingness to engage in cultivating the miracle plant species would help design sound promotion strategies for the species. Filling all these gaps is required to engage the species in the genomics-assisted breeding path.

1.2 Aim and objectives

The aim of this thesis is to generate a set of knowledge and data that can be used to build a strong breeding program in the species. The objectives of this study were four-folds:

iii) determine the end-users' preferences for breeding in the miracle plant in West Africa in order to formulate sound breeding objectives;
iv) evaluate the extent of natural phenotypic variability in the species to identify potential elite trees to use as parental lines;
v) assess the pattern of genetic variation and population structuring in the species to delimit key gene pools for use in genomic-assisted breeding schemes; and
vi) identify the use patterns and determinants for large-scale cultivation of the species to profile sustainable promotion strategies for the species in West Africa.

1.3 Research questions

To address the four research objectives, the following research questions were addressed:

i) What are the preferred traits of *S. dulcificum*' end-users, and how similar are they?
ii) What is the extent of natural variability in the miracle plant natural populations in West Africa, and how related are individuals based on their quantitative phenotypic traits?
iii) To what extent is the miracle plant gene pools delimitation in West Africa mediated by the geographical origin?
iv) What is the miracle plant use patterns in West Africa and what are the determinants of local communities to engage in its cultivation?

1.4 Thesis structure

This thesis is organized around seven chapters, including one introductory chapter (**Chapter 1**), one review chapter (**Chapter 2**), four experimental chapters (**Chapters 3-6**) and an overview chapter (**Chapter 7**). Three of the four experimental chapters have already been published, whereas the fourth is currently under review. Because Chapters 2-6 were independent and published and/or intended to be published as such, some information (e.g., general description of the species, references) was unavoidably repeated across them. In addition, because requirements for figure preparation varied from one journal to another one, figure panel numbering style varied slightly from one chapter to another, but was kept consistent within each chapter.

Chapter 1 presents the rationale, aim and objectives, and research questions of the thesis and outlines the thesis structure.

Chapter 2, which is the literature review, discusses the importance of the miracle plant and identifies the current knowledge gaps before the proposal of an in-context breeding plan.

Chapter 3, published in the "***Journal of Ethnobiology and Ethnomedicine***: https://doi.org/10.1186/s13002-021-00467-8), presents a new index (boosted management index) that was developed and used to analyse the variation in the miracle plant current management intensity in Benin and Ghana. Trait preferences and prioritization by the end-user groups in the two countries were also profiled, and levels of similarity following end-user groups, sociolinguistic groups and gender were quantified. Key breeding traits in the species were identified.

Chapter 4 is published in "***Scientific Reports***: https://doi.org/10.1038/s41598-020-76103-4". It analyses for the first time and from three West African countries, including Ghana, Togo, and Benin the natural quantitative variation in nine key functional traits of the miracle plant, including some indicated by the end-users. Most importantly, the relationship among 203 phenotypically characterized individuals was assessed, a phenotype-based core collection was inferred and clusters encompassing potentially performant individuals for future breeding programs were identified. The coefficient of repeatability was also used to approach estimates of broad-sense heritability for the fruit traits assessed.

Chapter 5 is currently under review. In this chapter, high-quality DArTseq-SNP markers were employed to understand the molecular diversity and population structure of the miracle plant in Benin, Togo and Ghana. Furthermore, the demographic history of the species in the region was depicted, and the potential origin of the species was inferred.

Chapter 6, is published in "***Plants***: https://doi.org/10.3390/plants10112253". In this chapter, how sociolinguistics in Benin and Ghana interacted with the miracle plant was revisited, and an original framework for analyzing knowledge acquisition/transmission in orphan crops in general was proposed. The key levers to foster a large-scale cultivation of the species in West Africa were also identified.

Chapter 7 is the last and presents the overall discussion of the thesis and analyses the implications of the thesis findings. The general conclusion ends the thesis.

1.5 References

Achigan-Dako, E. G., Tchokponhoué, D. A., N'Danikou, S., Gebauer, J., and Vodouhè, R. S. (2015). Current knowledge and breeding perspectives for the miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell. *Genet. Resour. Crop Evol.* **62**, 465-476.

Borelli, T., Hunter, D., Padulosi, S., Amaya, N., Meldrum, G., de Oliveira Beltrame, D. M., Samarasinghe, G., Wasike, V. W., Güner, B., and Tan, A. (2020). Local solutions for sustainable food systems: The contribution of orphan crops and wild edible species. *Agronomy* **10**, 231.

Buckmire, R., and Francis, F. (1978). Pigments of miracle fruit, *Synsepalum dulcificum*, Schum, as potential food colorants. *J. Food Sci.* **43**, 908-911.

Chen, C. C., Liu, I. M., and Cheng, J. T. (2006). Improvement of insulin resistance by miracle fruit (S*ynsepalum dulcificum*) in fructose-rich chow-fed rats. *Phytother. Res.* **20**, 987-992.

Chibuzor, I. A., Bukola, O., Adejoke, A. O., and Chidozie, O. P. (2017). Genetic assessment of the shrub *Synsepalum dulcificum* (Schumach & Thonn.) Daniell in Nigeria using the Randomly Amplified Polymorphic DNA (RAPD). *Int. J. Genet. Genomics* **4**, 45-50.

Dawson, I. K., Hedley, P. E., Guarino, L., and Jaenicke, H. (2009). Does biotechnology have a role in the promotion of underutilised crops? *Food Policy* **34**, 319-328.

Demesyeux, L., Brym, M., Valdes, D., Collazo, C., and Chambers, A. H. (2020). Yield and miraculin content of nine miracle fruit (*Synsepalum dulcificum*) morphotypes. *Euphytica* **216**, 1-12.

Du, L., Shen, Y., Zhang, X., Prinyawiwatkul, W., and Xu, Z. (2014). Antioxidant-rich phytochemicals in miracle berry (*Synsepalum dulcificum*) and antioxidant activity of its extracts. *Food Chem.* **153**, 279-284.

Fandohan, A. B., Gouwakinnou, G. N., Tovissode, C. F., Bonou, A., Djonlonkou, S. F. B., Houndelo, L. F., Sinsin, C. L. B., and Assogbadjo, A. E. (2017). Usages traditionnels et valeur économique de *Synsepalum dulcificum* au Sud-Bénin. *Bois For. Trop.* **332**, 17-30.

Fentahun, M. T., and Hager, H. (2009). Exploiting locally available resources for food and nutritional security enhancement: wild fruits diversity, potential and state of exploitation in the Amhara region of Ethiopia. *Food secur.* **1**, 207-219.

He, Z., Tan, J. S., Abbasiliasi, S., Lai, O. M., Tam, Y. J., and Ariff, A. B. (2016). Phytochemicals, nutritionals and antioxidant properties of miracle fruit Synsepalum dulcificum. *Ind. Crops Prod.* **86**, 87-94.

Hirai, T., Shohael, A. M., Kim, Y.-W., Yano, M., and Ezura, H. (2011). Ubiquitin promoter–terminator cassette promotes genetically stable expression of the taste-modifying protein miraculin in transgenic lettuce. *Plant Cell Rep.* **30**, 2255-2265.

Hiwasa-Tanase, K., Nyarubona, M., Hirai, T., Kato, K., Ichikawa, T., and Ezura, H. (2011). High-level accumulation of recombinant miraculin protein in transgenic tomatoes expressing a synthetic miraculin gene with optimized codon usage terminated by the native miraculin terminator. *Plant Cell Rep.* **30**, 113-124.

Hunter, D., Borelli, T., Beltrame, D. M., Oliveira, C. N., Coradin, L., Wasike, V. W., Wasilwa, L., Mwai, J., Manjella, A., and Samarasinghe, G. W. (2019). The potential of neglected and underutilized species for improving diets and nutrition. *Planta* **250**, 709-729.

Inglett, G., and May, J. F. (1968). Tropical plants with unusual taste properties. *Econ. Bot.* **22**, 326-331.

International Diabetes Federation (2021). IDF Diabetes Atlas.

Kibret, K. B. (2021). The potentials of African neglected and orphan crops in augmentation of African food security. *In* "Food security and safety" (O. O. Babalola, ed.), pp. 43-62. Springer.

Kim, H. K., Go, J. Y., Park, S.-Y., Kang, K. K., and Jung, Y. J. (2020). Production of miraculin protein in suspension cell lines of transgenic rice using Agrobacterium. *J. Plant Biotechnol.* **47**, 227-234.

Kobori, C. N., and Amaya, D. B. R. (2008). Uncultivated Brazilian green leaves are richer sources of carotenoids than are commercially produced leafy vegetables. *Food Nutr. Bull.* **29**, 320-328.

Kurihara, K., and Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science* **161**, 1241-1243.

Linder, H. (2001). Plant diversity and endemism in sub-Saharan tropical Africa. *J. Biogeogr.* **28**, 169-182.

Mabhaudhi, T., Chimonyo, V. G. P., Hlahla, S., Massawe, F., Mayes, S., Nhamo, L., and Modi, A. T. (2019). Prospects of orphan crops in climate change. *Planta* **250**, 695-708.

Misaka, T. (2013). Molecular mechanisms of the action of miraculin, a taste-modifying protein. *Semin. Cell Dev. Biol.* **24**, 222-225.

Nathan, D. M., Davidson, M. B., DeFronzo, R. A., Heine, R. J., Henry, R. R., Pratley, R., and Zinman, B. (2007). Impaired fasting glucose and impaired glucose tolerance: implications for care. *Diabetes Care* **30**, 753-759.

Njoku, N. E., Ubbaonu, C. N., Alagbaoso, S. O., Eluchie, C. N., and Umelo, M. C. (2015). Amino acid profile and oxidizable vitamin content of *Synsepalum dulcificum* berry (miracle fruit) pulp. *Food Sci. Nutr.* **3**, 252-256.

Nkwocha, C., Njoku, O., and Ekwueme, F. (2014). Proximate and micronutrient analyses of synsepalum dulcificum pulp. *Scient. Res. J.* **2**, 2201-2796.

Nyadanu, D., and Lowor, S. (2015). Promoting competitiveness of neglected and underutilized crop species: comparative analysis of nutritional composition of indigenous and exotic leafy and fruit vegetables in Ghana. *Genet. Resour. Crop Evol.* **62**, 131-140.

Obafemi, T. O., Olaleye, M. T., and Akinmoladun, A. C. (2019a). Antidiabetic property of miracle fruit plant (*Synsepalum dulcificum* Shumach. & Thonn. Daniell) leaf extracts in fructose-fed streptozotocin-injected rats via anti-inflammatory activity and inhibition of carbohydrate metabolizing enzymes. *J. Ethnopharmacol.* **244**, 112124.

Obafemi, T. O., Onasanya, A., Adeoye, A., Falode, J. A., Daniel, D. J., Irefo, E. F., Ojo, A. O., Fadaka, A., Afolabi, O. B., and Awe, J. O. (2019b). Protective effect of methanolic and flavonoid-rich leaf extracts of *Synsepalum dulcificum* (Danielli) on lead-acetate-induced toxicity in Wistar albino rats. *J. Appl. Pharm. Sci.* **9**, 065-072.

Ono, A., Hiwasa-Tanase, K., Nonaka, S., and Ezura, H. (2021). The accumulation of recombinant miraculin is independent of fruit size in tomato. *Plant Biotechnol.*, 20.0904 a.

Oumorou, M., Dah-Dovonon, J., Aboh, B., Hounsoukaka, M., and Sinsin, B. (2010). Contribution á la conservation de *Synsepalum dulcificum*: régénération et importance socio-économique dans le département de l'ouémé (Bénin). *Ann. Sci. Agron* **14**, 101-120.

Padulosi, S., Meldrum, G., King, E. O., and Hunter, D. (2021). NUS: What they are and why we need them more than ever. *In* "Orphan crops for sustainable food and nutrition security" (S. Padulosi, E. O. King, D. Hunter and M. S. Swaninathan, eds.), pp. 3-18. Routledge.

Paladino, A., Costantini, S., Colonna, G., Facchiano, A. M. J. B., and Communications, B. R. (2008). Molecular modelling of miraculin: structural analyses and functional hypotheses. *Biochem. Biophys. Res. Commun.* **367**, 26-32.

Park, Y.-J., Han, J.-E., Lee, H., Jung, Y.-J., Murthy, H. N., and Park, S.-Y. (2020). Large-scale production of recombinant miraculin protein in transgenic carrot callus suspension cultures using air-lift bioreactors. *AMB Express* **10**, 1-7.

Park, Y.-J., Han, J.-E., Lee, H., Lee, J.-Y., Ho, T.-T., and Park, S.-Y. (2021). Production of recombinant miraculin protein in carrot callus via Agrobacterium-mediated transformation. *Plant Cell, Tissue and Organ Cult* **145**, 615-623.

Pilling, D., Bélanger, J., Diulgheroff, S., Koskela, J., Leroy, G., Mair, G., and Hoffmann, I. (2020). Global status of genetic resources for food and agriculture: challenges and research needs. *In* "Genetic Resources", Vol. 1, pp. 4-16.

Ramachandran, A., Snehalatha, C., Mary, S., Mukesh, B., Bhaskar, A., and Vijay, V. (2006). The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). *Diabetologia* **49**, 289-297.

Sinare, H., and Gordon, L. J. (2015). Ecosystem services from woody vegetation on agricultural lands in Sudano-Sahelian West Africa. *Agric. Ecosyst. Envenviron.* **200**, 186-199.

Sugaya, T., Yano, M., Sun, H.-J., Hirai, T., and Ezura, H. (2008). Transgenic strawberry expressing the taste-modifying protein miraculin. *Plant Biotechnol.* **25**, 329-333.

Sun, H.-J., Cui, M.-l., Ma, B., and Ezura, H. (2006). Functional expression of the taste-modifying protein, miraculin, in transgenic lettuce. *FEBS Letters* **580**, 620-626.

Tadele, Z. (2009). Role of orphan crops in enhancing and diversifying food production in Africa. *In* "Afric. Technol. Dev. Forum J.", Vol. 6, pp. 9-15.

Tafazoli, S., Vo, T. D., Roberts, A., Rodriguez, C., Viñas, R., Madonna, M. E., Chiang, Y.-H., Noronha, J. W., Holguin, J. C., and Ryder, J. A. (2019). Safety assessment of miraculin using in silico and in vitro digestibility analyses. *Food Chem. Toxicol.* **133**, 1-10.

Takahashi, N., Hitotsuya, H., Hanzawa, H., Arata, Y., and Kurihara, Y. (1990). Structural study of asparagine-linked oligosaccharide moiety of taste-modifying protein, miraculin. *J. Biol. Chem.* **265**, 7793-7798.

Tchokponhoué, D., Achigan-Dako, E., N'Danikou, S., Houdégbé, A., Agossou, C., Assogba-Komlan, F., and Vodouhè, R. (2018). Regeneration ability and seedling growth in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Fruits* **73**, 13-21.

Tchokponhoué, D. A., N'Danikou, S., and Achigan-Dako, E. G. (2019a). A combination of approaches evidenced seed storage behaviour in the miracle berry *Synsepalum dulcificum* (Schumach. et Thonn.) Daniell. *BMC Plt. Biol.* **19:117**, 1-13.

Tchokponhoué, D. A., N'Danikou, S., Houéto, J. S., and Achigan-Dako, E. G. (2019b). Shade and nutrient-mediated phenotypic plasticity in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Sci Rep.* **9:5137**, 1-11.

Theerasilp, S., and Kurihara, Y. (1988). Complete purification and characterization of the taste-modifying protein, miraculin, from miracle fruit. *J. Biol. Chem.* **263**, 11536-11539.

Tuomilehto, J., Lindström, J., Eriksson, J. G., Valle, T. T., Hämäläinen, H., Ilanne-Parikka, P., Keinänen-Kiukaanniemi, S., Laakso, M., Louheranta, A., and Rastas, M. (2001).

Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N. Engl J Med* **344**, 1343-1350.

Wang, H.-M., Chou, Y.-T., Hong, Z.-L., Chen, H.-A., Chang, Y.-C., Yang, W.-L., Chang, H.-C., Mai, C.-T., and Chen, C.-Y. (2011). Bioconstituents from stems of *Synsepalum dulcificum* Daniell (Sapotaceae) inhibit human melanoma proliferation, reduce mushroom tyrosinase activity and have antioxidant properties. *J. Taiwan Inst. Chem. Engin.* **42**, 204-211.

Wang, J., Shao, S., Huang, R., and Wu, S. (2020). Chemical constituents and their antioxidant activities from the leaves of *Synsepalum dulcificum*. *Nat. Prod. Res.*, 1-6.

Wilken, M. K., and Satiroff, B. A. (2012). Pilot study of “miracle fruit” to improve food palatability for patients receiving chemotherapy. *Clin. J. Oncol. Nurs.* **16**, E173-E177.

CHAPTER 2

State of knowledge and breeding avenue for the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell]: a review

Abstract

Minor fruit tree species represent a major component of sub-Saharan African (SSA) agroecosystems and play a tremendous role in sustaining local communities' wellbeing and livelihoods. However, most of them have remained poorly documented and lack systematic breeding efforts that could upgrade their potential. This is particularly true for the West African native miracle plant (*Synsepalum dulcificum*), a promising shrub from the Sapotaceae known as a unique natural source of miraculin, a sweetening glycoprotein. This review shows that most of the studies on the miracle plant have focused on characterization of the miraculin and its transgenic expression, the nutritional properties of the fruits and leaves and their antioxidant phytochemical content. While generation of knowledge required for the proper cultivation of the species to sustain its utilization is in its infancy, information essential to initiate a proper breeding program is still lacking. Therefore, for the miracle plant, combining conventional breeding and genomic-assisted breeding will be needed to ensure a quick elite cultivar release. Achieving this would, thus, entail: i) identifying traits that define an elite cultivar in the sense of end-users, ii) assembling germplasm from the species potential centre of diversity, iii) characterizing the molecular diversity and population structure in the species, iv) building a core collection to generate a manageable active working population, v) characterizing the phenotypic diversity in the core collection against end-users defined breeding traits, vi) sub-setting the core collection for conventional breeding purposes, vii) developing a genomic selection scheme for early advancement of promising individuals, viii) identifying quantitative trait single nucleotide polymorphisms (QTSNPs) by exploiting reference genomes of sister species, ix) exploring genome editing based on QTSNPs identified, and x) fostering regional collaboration. This review also demonstrates that while points i) to viii) are currently achievable in the miracle plant, more studies related to the reference genome assembly will be needed to fully take advantage of the genome editing technology. For this latter to quickly happen, collaborative research (point (x)) needs to be seriously considered.

2.1 Background

The world is rich in biodiversity, with more than 352000 plant species reported (Royal Botanic Gardens Kew, 2021), out of which only 0.5% are cultivated (Purugganan, 2019) and nearly 7,000 species are used as food (Thrupp, 2000). This suggests that more than 70% of human food is derived from the wild. More strikingly, only four major species [maize: *Zea mays* L., rice: *Oryza sativa* L., wheat: *Triticum aestivum* L., and potato: *Solanum tuberosum* L.] provide 60% of the world's calorie consumption (N'Danikou and Tchokponhoué, 2020), while more than 65% of global crop production comes from only nine species (Pilling et al., 2020) despite the rich agricultural biodiversity available. Today, more than ever, the sustainability and resilience of this highly skewed and selective food system is challenged by the adverse effects of climate change. In such a context, realizing sustainable development goals (SDGs), such as SDG2 "zero hunger", SDG3 "good health and wellbeing", may be highly challenging unless diversification options are explored, and orphan crops hold the potential to help achieve these goals (Jamnadass et al., 2020; Mabhaudhi et al., 2019).

Additionally, known as neglected and underutilized species, specialty crops, minor crops, or future crops, orphan crops represent traditional species that are locally or regionally important and adapted but have benefited very little from researchers, funders, and policy makers (Hendre et al., 2019; Jamnadass et al., 2020; Tadele and Bartels, 2019). These crops include annuals and perennials, cereals, vegetables, legumes, tubers, roots, ornamentals, and fruits. Diverse studies have reported on orphan crops' importance in terms of contributions to food security, climate resilience, and nutritional security, with a recent study listing up to 15 different key categories of contributions. Orphan crops have a common aspect of being understudied, but this is particularly exacerbated when it comes to fruit tree crops.

Fruit tree crops represent key resources that can help alleviate poverty, mitigate climate change, and approach food and nutrition security in Africa, in addition to being an important component of the agroecosystem. For instance, the Shea tree (*Vitellaria paradoxa* C.F. Gaertn), known as the "Women Gold" (Elias and Carney, 2007), spans west to east Africa and represents an important contribution to local and international economies, with more than 18 million people involved in its exploitation (Naughton et al., 2015). Marula [*Sclerocarya birrea* (A. Rich)], well known across the Sub-Saharan region, is used not only as food, firewood, shade tree, remedies, and in the production of liqueur (Gouwakinnou et al., 2011; Nwonmu, 2006), but also has sociocultural importance. The economic importance of baobab (*Adansonia digitata* L.) has been acknowledged as the species has well-organized value chains in many African countries (Abdellrhman and Adam, 2020; Jäckering et al., 2019; Kamatou et al., 2011). Baobab leaves serve as vegetables and/or medicines in many African countries (Achigan-Dako et al., 2010; Fischer et al., 2020) while pulp consumption provides three times more potassium than banana, three times more vitamin C than oranges, nearly twice the daily recommended vitamin C for six-year-old children and 80% of their daily magnesium needs (Stadlmayr et al., 2020). Ecologically, baobab seems to represent a keystone species in the reproductive activity of *Bubalornis niger* A. Smith, the red buffalo weaver, while providing habitats to many other wild animals (Rhodes, 2009). The seeds of *Allanblackia floribunda* Oliv., a salient species from the republic of Benin to the Democratic Republic of Congo, represent key resources in the international cosmetics industry with a guaranteed market of more than US $100 million for African actors with the company Unilever PLC (Atangana et al., 2011). Consequently, leveraging these African orphan fruit tree crops through breeding holds promise to improve livelihoods in the SSA region. Nevertheless, most of them currently lack a clear breeding history, except those that have benefited from the domestication effort implemented by the World Agroforestry Centre (Mpeck et al., 2003; Tchoundjeu et al., 2006). Although there is increased effort to make reference genomes available for a number of African orphan fruit tree crops (http://africanorphancrops.org/meet-the-crops), this information can only be utilized when a proper and realistic breeding plan exists for these crops, an ideal we are still far from today in these Sub-Saharan African species.

Therefore, this review discusses the miracle plant [(*Synsepalum dulcificum* (Schumach and Daniell)], a singular African orphan fruit tree with the potential to contribute to SDG2. Here, we present the current state of knowledge on the species and its uses before proposing an in-context breeding path to help the species benefit from genomics-assisted breeding.

2.2 Description, systematics, origin, and current distribution of the miracle plant

The miracle plant also known as *Synsepalum dulcificum* (Schumach & Thonn.) Daniell or *Richardella dulcifica* (Schumach & Thonn.) Baehni is a perennial shrub species from the Sapotaceae, a family with confusing and unstable systematics (Swenson and Enderberg, 2003). It is a member of the Chrysophylleae tribe. Often reaching the size of 4 m tall at the adult stage (Achigan-Dako et al., 2015), the miracle plant has whitish to creamy flowers and produce 2–2.5 cm long and 1 cm large ovoid to oblong red berries known as miracle fruits or miracle berries. Its 4.5–7.5 cm long and 3 5–3.8 cm leaves are simple and alternate with a 4.5 – 5 mm long petiole.

The miracle plant exhibits a pinnate eucamptodromous venation whose highest order is seven (Ayensu, 1972).

The miracle plant naturally thrives in West and Central Africa, extending from the Ivory Coast to the Democratic Republic of Congo. Although the species was also reported to occur in Tanzania, we speculate this to probably be a confusion with its closest sister (*Synsepalum passarge*i (Engl.) T.D. Penn. known to originate in Tanzania) (Swenson et al., 2008). The likely origin of the miracle plant coincides in a complex of countries formed by Ghana, Togo, Benin, and Nigeria (Achigan-Dako et al., 2015; Swenson et al., 2008). However, the species is currently distributed on all continents, probably following introductions after the slave trade from the Gulf of Guinea. Today, the species has been reported in many countries, including the United States of America (USA), Brazil, China, Philippines, Malaysia, Japan, and Australia (**Fig. 2.1**). In terms of ecological requirements, the miracle plant grows on well-drained acidic soils and is an intermediate shade-tolerant species (Tchokponhoué et al., 2019b). It has been recorded in swamp forests, fields and farms (Adomou, 2005; Akoègninou et al., 2006).

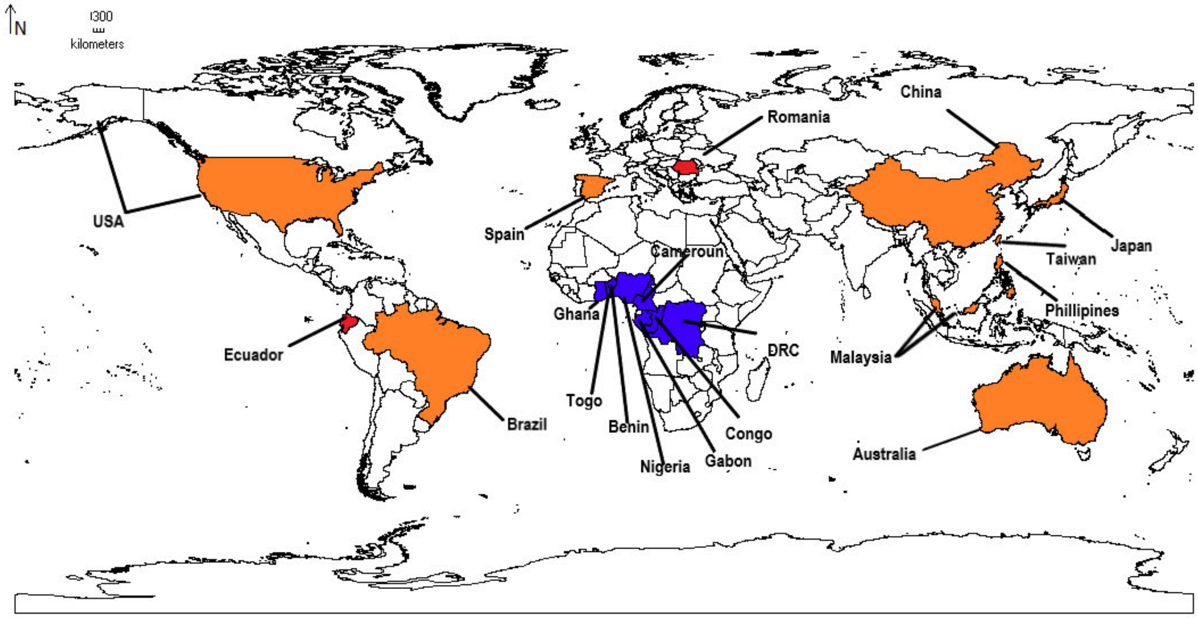


Figure 2.1. Map indicating the current distribution area of the miracle plant in the world. **Blue** highlighted countries represent the natural distribution area of the species, while **orange** shows the current introduction areas and **red** the forthcoming introduction areas.

2.3 Propagation ecology of the miracle plant

Poor knowledge of orphan crop propagation methods represents a major bottleneck for their improvement. Consequently, developing cost-effective propagation methods is required as a fundamental pre-breeding activity. The miracle plant can be propagated both sexually and asexually. Although seed-based regeneration was considered challenging (Chen et al., 2012) as the species exhibits recalcitrance (Tchokponhoué et al., 2019a), recent developments have indicated that nearly 100% seed germination can be attained if the growing medium is adequately chosen and the seed post-shedding phase is properly managed (Oumorou et al., 2010; Tchokponhoué et al., 2018). In parallel, vegetative propagation methods continue to be refined, and today apart from stem-cutting propagation (Chen et al., 2012), air layering has also proven to be effective in this species (Geoffery and Sani, 2017). Nevertheless, despite these encouraging results, vegetative propagation methods have yet to be optimized in the species given the still moderate success rate (50–60% rooting percentage). Rooting the miracle plant through micropropagation has always been a challenge, partly due to the wood anatomy (Ayensu, 1972). Alternative micropropagation methods such as tissue culture have been explored (Chen et al., 2012), with however a low rooting rate. Progress so far made suggests sexual propagation as the most efficient propagation method in the miracle plant.

2.4 Growth, phenology, and reproductive biology in the miracle plant

As in a number of Sapotaceae species [e.g., *V. paradoxa* (Sanou et al., 2004; Teklehaymanot, 2004) and African Star apple *Chrysophyllum albidum* G. (Don).(Olufemi et al., 2019)], slow growth has been confirmed in the miracle plant (Tchokponhoué et al., 2018). Experimental data by Tchokponhoué et al. (2018) indicates that the maximum height and diameter of the miracle plant at the age of six months is 5-6 cm and 2-3 mm, respectively. Such a poor growth rate lengthens the maturity time of the species, as observed by Joyner (2006) who reported that the species naturally bears fruit from the 4 h to 6th year after planting. However, proper water, nutrient and shade management holds the potential to shorten this time to fruiting to 18 to 24 months (Tchokponhoué et al., 2017; Tchokponhoué et al., 2019b).

Dynamics in crop phenology have been increasingly used as a proxy of the potential effect of climate change in plant species and are informative for breeding purposes. Nevertheless, such knowledge seems to be critically lacking in orphan crops. In situ data collected on adult individuals of the miracle plant by Oumorou et al. (2010) suggests three flowering (April – May; July – August and November – December) and three fruiting (May, August and December – February) windows for the species in Eastern Benin. Contrary to the in-situ observations, ex situ data suggests a continuous phenology for the miracle plant marked by flowering and fruiting throughout the year, especially under enhanced agronomic practices (Demesyeux et al., 2020; Tchokponhoué et al., 2017). Unfortunately, the data of both studies were limited to only one year of observation and thus, not offering information on interannual variation. Hence, it is necessary to extend observations over several years and possibly to a larger geographical area as the species distribution is quite large.

Regarding the flower development, detailed knowledge on the species phenophases and their respective lengths can be found in the studies by Xingwei et al. (2016) and Tchokponhoué et al. (2017). A total of six phenophases (S1 – S6) were distinguished with durations ranging from three days to 90 days. The six phenophases were described as follows: S1 - budding to flowering; S2 - flowering to flower bloom; S3 - flower bloom to fructification onset; S4 - fructification onset to physiological maturing; S5 - physiological maturing to fruit ripening onset; and S6 - fruit ripening onset to fruit full ripening. The proper depiction of phenophases S2 and S3 variation patterns is crucial for the success of any future crosses in the species. Indeed, although empirical observations suggest that autogamy is possible in the species (Tchokponhoué et al., 2017), the exposure pattern of the miracle plant style (stigma) coupled with frequent visits of pollinators, does not discard a concurrent allogamy in the species. Consequently, experimental data are needed to clarify the prominence of each mating system in the species.

2.5 Miracle plant cultivation and production statistics

Although the miracle plant is known to originate in West Africa, its cultivation in the region seems to still be in its infancy. In the Republic of Benin, the miracle plant is currently found as individual trees in home gardens, as previously reported by Achigan-Dako et al. (2015), and no evidence suggests the evolution of the situation since 2015; yet, experimental data have suggested the possibility of integrating the species in even complex agroforestry systems (Tchokponhoué et al., 2019b). For other countries, such as Togo, Ghana and Nigeria, a diagnosis of the current cultivation status is required. More importantly, a comparative analysis of the management practices applied to the crop, if any, in the species natural distribution areas is required. As a consequence of the current cultivation status of the species in Benin, production statistics are lacking overall, although Achigan-Dako et al. (2015) indicated that the annual production of a single individual miracle tree can reach 12 – 15 kg with substantial variation depending on the tree age, branching pattern and habitat. Similar to the situation observed in West Africa, no reliable data suggested the proper cultivation of the miracle plant in its introduction areas; the only exception was in Florida (the USA), where the crop was cultivated in a controlled environment by the “fruit farm” private company. Experimental data collected from this farm suggested that the annual production of 5–6-year-old individual miracle plant juveniles can range from 0.06 kg to 3.44 kg (Demesyeux et al., 2020). Whether comparable or differential productivity could also be expected in West Africa where the species naturally thrives, and from West African genotypes, deserves thorough investigation.

2.6 Pests and diseases in the miracle plant

The miracle plant can be infested by a number of pests and diseases. Diseases so far reported include fruit rot, leaf spot and leaf blight caused by *Pestalotiopsis synsepali*, *Colletotrichum karstii*, *C. aenigma* B. Weir & P.R. Johnst and *C. siamense* B. Weir & P.R. Johnst, fungal species, respectively (Chen et al., 2002; Truong et al., 2018) (**Table. 2.1**). For pests, a recent study conducted in the Ashanti, Eastern and Volta regions of Ghana reported a total of 15 different insect species belonging to five orders (**Fig. 2.2**), with Lepidoptera being the most represented with 10 different species (Atuahene and Osekre, 2018). Nevertheless, the Diptera represented only by *Ceratitis punctata* Wiedm was the most common causal order, responsible for more than 50% of the damage observed on the fruit. *Ceratitis punctata* spreads ovipositional punctures on both green and ripe miracle fruit, exposing them to further fungal and bacterial infections and rot. *Ceratitis punctata*-infested fruit lose firmness and color brightness before dropping some five days later. Two of the insect pests (*Oecophylla longinoda* Latr. and *Tetramolium aculeatum* Mayr.) were also reported to have beneficial effects on miracle plants, as they reduce infestation by *C. punctata* and lepidoptera. In Benin, empirical observations also suggest that the species has a range of pathogens affecting the different parts of the plant, including the seeds, fruits, leaves, stems, and roots, with the incidence likely to increase as production intensifies. Proper identification and documentation of biology, phenology and incidence of these pests are required to develop sound and eco-friendly control measures.

Table 2.1. Pests and diseases recorded on the miracle plant in Africa and Asia

Pests and diseases		Organ affected	Symptoms/Damages	Causal agent	References
Diseases	Leaf spot	Leaf	Tiny black dots at leaf margin that develop and produce big spots and chlorosis areas encompassed by a dark brown line. Ultimately results in defoliation	*Pestalotiopsis synsepali*	Chen et al. (2002)
				Colletotrichum karstii. Weir & P.R. Johnst	Truong et al. (2018)
	Leaf blight	Leaf	Small lesion at the tip of the leaf, that develops and increases in size towards the petiole. Ultimately results in defoliation	*Colletotrichum siamense*. Weir & P.R. Johnst	Truong et al. (2018)
	Fruit rot	Fruit	Dark brown lesions around wounded area of the fruit extending to the fruit that is ultimately blackened	*Colletotrichum aenigma* B. Weir& P.R.Johnst	Truong et al. (2018)
Insects	-	Unripe and ripe fruit	Ovipositional punctures, fruit rot, fruit deformation	*Ceratitis punctata* Wiedm	Atuahene and Osekre (2018)
	-	Partially ripe fruits and leaf		*Achae janata* Linn.	
	--			*Achae* spp	
	-			*Grammodes geometrica* Lat.	
	-			*Coeliades forestan* Stoil.	
	-		Boring and sucking of fruit and leaf defoliation	*Catopsilla frorellia* Fab.	
	-			*Danaus chrysipus* Linn.	
	-			*Appias sylvia* Fab.	
	-			*Bicyclus cenaes* Hewiston	
	-			*Bicyclus saftza* westwood	
	-			*Amauris nievius* Linn.	
	-	Fruit	Creamy spot of fruit	*Oecophylla longinoda* Latr.	
	-	Fruit	Creamy spot of fruit	*Tetramorium asculeatum* Mayr.	
	-	Leaf	Defoliation	*Zonocerus variegatus* L.	
	-	Bark	Root and bark feeder	*Microtermes natalensis* Hev.	

2.7 Functional properties of the miracle plant

The miracle plant is functionally important because it is the source of numerous metabolites, nutrients, and phytochemicals, some of which have unique properties. These are discussed in the following sections.

2.7.1 Miraculin and its sweetening property

Although the use of miracle fruits dates back to the 1600s in West Africa, it was only in 1852 that the unusual taste modifying property of the berries was scientifically reported (Daniel, 1852). This triggered a suite of studies (Brouwer et al., 1968; Inglett et al., 1965; Kurihara and Beidler, 1968) intended to characterize and synthesise the sweetening substance. These studies were successful in 1968 (Kurihara and Beidler, 1968), with the taste modifying substance found to be a glycoprotein called "miraculin", which is contained in the mucilaginous part of the fruits. Miraculin is a macromolecular compound, and the first estimates suggested that its molecular weight was in the average range of 42,000 Dalton (based on ultracentrifugation) to 48,000 Dalton (based on gel filtration) (Brouwer et al., 1968), while updated estimates have reported a weight of 28000 Dalton (Theerasilp and Kurihara, 1988). Aspartic acid and valine accounted for the largest amino acid parts of the purified miraculin (Theerasilp and Kurihara, 1988), which is a single polypeptide of 191 amino acids.

Miraculin has no taste by itself and is thermolabile and stable at pH values ranging from 3 to 12. Its purification yield is extremely low, as nearly 20 kg of fresh fruits are needed to obtain 1 g of pure miraculin (Brouwer et al., 1968). Although miraculin is acknowledged to induce a sweet taste perception overall, sweetness perception may be differentially expressed depending on which molecular form is at play. Indeed, the molecular modelling of miraculin revealed that the glycoprotein exists in three forms, including a monomer form (**Fig. 2.3.A**), a dimer form (**Fig. 2.3.B**) and a tetramer form (**Fig. 2.3.C,D**) (Paladino et al., 2008). While the monomer form does not elicit any sweet taste perception at any pH and even at high concentrations, the dimer form triggers sweetness perception at acidic pH. This differential activity is ascribed to the higher exposure of Hs29 and Hs59, the two histidine residues known to be responsible for the sweetening property (Igeta et al., 1991) in the dimer form compared with the monomer. Like the dimer form, the tetramer form is also active, even if it is to a lesser extent. In the miracle fruit, the miraculin is mostly present in the dimer form (Hirai et al., 2010), which is why the miracle fruit triggers sweet taste perception. At the cellular level, the miraculin is produced from the golgi apparatus outward and passes through successive cell walls to ultimately accumulate in intercellular layers (Hirai et al., 2010) and not in intercellular spaces, as previously speculated (Masuda et al., 1995).

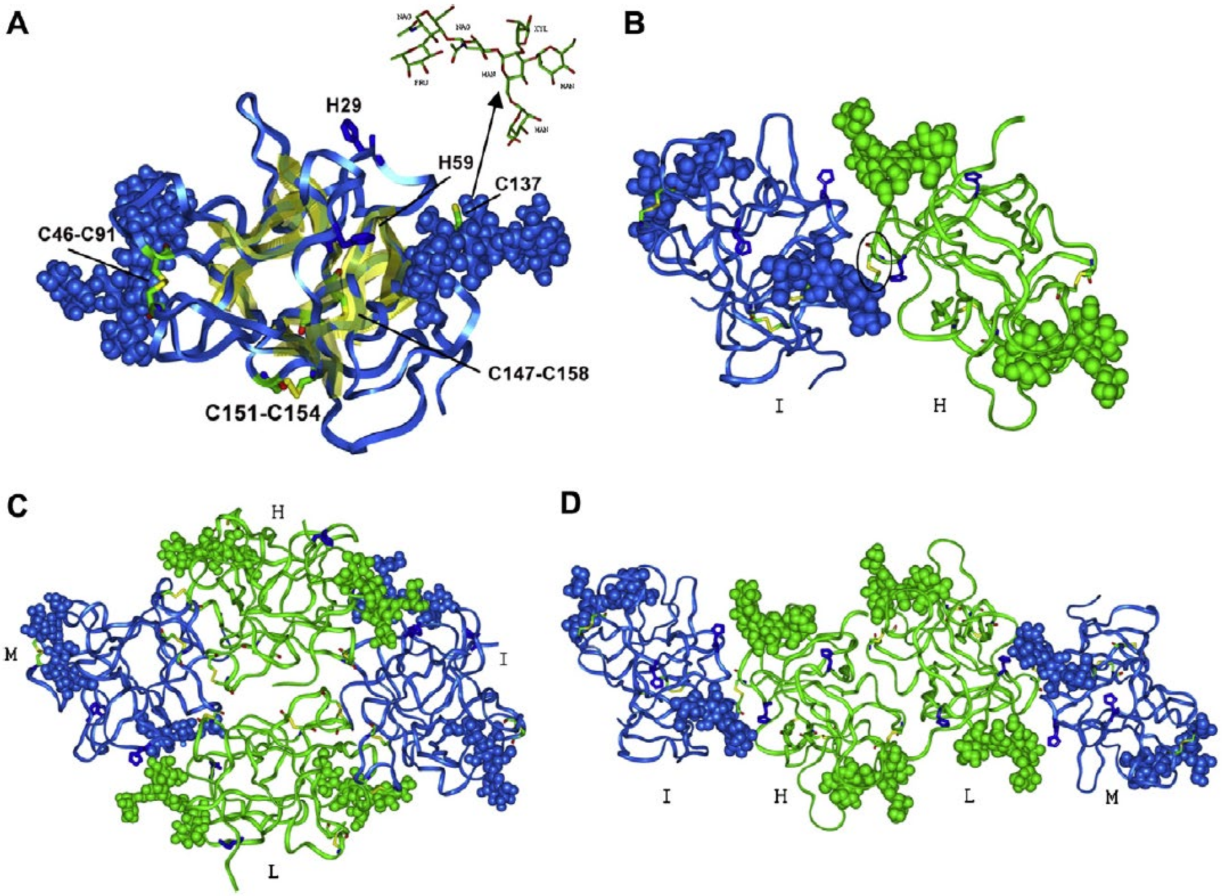


Figure 2.2. 3D molecular modelling of miraculin. **A**: Monomer form; **B**: Dimer form [each monomer (H or I) has a specific colour]; **C**: Globular tetramer form; and **D**: Linear tetramer form. Source (Paladino et al., 2008).

2.7.2 Nutrient content of the miracle plant

Apart from its miraculin content, the miracle fruit is also an important source of nutrients. The overall trend indicates that iron, calcium, and vitamin C are the major micronutrients, macronutrients, and vitamins, respectively, in the miracle fruit. According to Awotedu and Ogunbamowo (2019), the micronutrient contents (mg/kg) of the miracle fruit are as follows: iron (102), copper (89), manganese (34) and zinc (17), while the macronutrient contents are as follows: calcium (1395), magnesium (2678.8), potassium (2900) and sodium (1531). This suggests that the miracle fruit is nearly 2.5 times iron-richer, 2 times manganese-richer, and more than 20 times copper-richer than *Rubus spp.* (Iron: 21.635–32.6 mg/kg; Mn: 13.95–19.48 mg/kg;

Cu: 0.25–2.34) (Ahmad et al., 2015). Compared to various apple (*Malus domestica* Borkh. varieties, these values also reveal a higher micronutrient content of the miracle fruit (Macit et al., 2021). For the macronutrients, while the miracle fruit seems to be denser in calcium and magnesium than apple (Ca: < 700 mg/kg, Mg: < 450 mg/kg), a reverse trend was observed for the potassium content, with apple potassium content being greater than 5750 mg/kg. However, large infraspecific variations were observed for both micronutrient and macronutrient contents in the miracle fruit (**Table 2.2**), because of potential differences in genotypes, environmental conditions and harvest time, among others. Likewise, a large variation in miracle fruit ascorbic acid content (28.95 – 40.1 mg/100 g) was also noted in the literature (Du et al., 2014; He et al., 2016). However, all values observed indicated that the miracle fruit is a richer source of vitamin C compared to blueberry (considered a superfood) (Ruiz et al., 2013) and *Averrhoa carambola* (Edem et al., 2008). Conversely, values reported for various studies on miracle fruit proximate analyses seem to concur, with only a low magnitude of variation (**Table 2.3**). Qualitatively, the miracle fruit seed oil also exhibited an interesting fatty acid profile that is more diverse than that of palm oil, palm kernel oil, Jatropha seed oil, coconut oil, soybean oil, sunflower oil, groundnut oil, cotton seed oil and corn oil (Akbar et al., 2009; Edem, 2002; He et al., 2016). Interestingly, the miracle fruit seed oil stands as the second most important source of palmitic acid, behind palm oil (Edem, 2002). Palmitic acid is known to play a crucial role in adipose and intestinal organ development in infants. Therefore, how miracle fruit seed oil can be taken advantage of needs to be explored, especially in regions where access to palm oil is challenging.

Table 2.2. Miracle fruit micro and macro-nutrients content. Source (Awotedu and Ogunbamowo, 2019) and (Nkwocha et al., 2014)

		(Awotedu and Ogunbamowo, 2019)	Nkwocha (2014)
Macronutrients (mg/kg)	Calcium (Ca)	1395	100
	Magnesium (Mg)	2678.8	*
	Potassium (K)	2900	*
	Sodium (Na)	1531	*
Micronutrients (mg/kg)	Zinc (Zn	17	9.49
	Iron (Fe)	102	24.2
	Copper (Cu)	89	6.22
	Manganese (Mn)	34	*

*: Not quantified in the study.

Table 2.3. Proximate analysis of the miracle fruit

	References			
%	Awotedu and Ogunbamowo (2019)	He et al. (2016)	Njoku et al. (2016)	Nkwocha et al. (2014)
Crude protein	10.26	*	2.48	7.75
Crude Fat	2.32	0	2.1	3.26
Crude Fibre	3.63	12.5	0.57	6.24
Ash Content	3.95	1	0.87	4.36
Carbohydrate	17.44	22.5	48.85	18.84
Moisture	62.4	*	45.2	59.55

*Not reported in the study.

2.7.3 Antioxidant phytochemicals and biological activities of the miracle plant

Antioxidants are useful phytochemicals in preventing free radical-induced cell damages and are found in medicinal plants, fruits, edible macro-fungi and vegetables, among others (Zhang et al., 2015). As previously synthesized by Achigan-Dako et al. (2015) and Akinmoladun et al. (2020), the miracle plant is a tremendous source of phytochemicals, including antioxidant phytochemicals. All the plant parts were reported to be phytochemical-rich, with contents differing moderately to extensively from one plant part to another. The reported total phenolic contents (mg GC equivalent dry weight) of the miracle fruit flesh and seed are 1448 and 306.7, respectively, whereas the flavonoid contents of the same parts are 9.9 and 3.8. This suggests a higher antioxidant capacity of the miracle fruit flesh compared with the miracle fruit seed and a higher antioxidant capacity of the miracle fruit compared with well-known antioxidant rich berries such as blueberry and black berry (Du et al., 2014). The miracle fruit flesh contains a total of nine phenolic compounds, with the most represented being epicatechin (17.8 mg/100 g fresh weight); three anthocyanin compounds, with the most important being malvidin galactoside (10.1 mg/100 mg) and four carotenoids, with α-tocopherol (5.8 mg/100 g fresh weight) being the most represented (Du et al., 2014).

For the leaves, the total phenolic and flavonoid contents are 90 mg GC equiv/g and 33 mg GC equiv/g, respectively. A total of 18 different phenolic compounds were identified from the miracle plant leaves, indicating that leaves have a more diversified phenolic constitution than the fruits. The most important phenolic compounds of the leaves include quercetin-3-rhamnoside (4.13 mg/g) and gallic acid (2.73 mg/g) (Liu et al., 2021). Other phytochemical components inventoried in the leaves included dihydrodendranthemoside (A), sawaranin (B), (7S,8R)-dihydrodehydrodiconiferyl alcohol (C), (7S,8R)-dihydrodehydrodiconiferyl alcohol-9—D-glucopyranoside (D), quercetin (E), and quercetin-3-O-glucoside (F), with the miracle plant being the second natural source of compounds A and B, the first natural sources of compound A and B being *Alangium premnifolium* Ohwi (Otsuka et al., 1994) and *Polygonum chinense* L. (Zheng et al., 2018), respectively. Collectively, these phytochemicals confer various biological activities to the miracle plant, as shown via various extracts (methanol, ethanol) of the fruit, leaves and other parts. These include cholesterol-lowering (Huang et al., 2020), antidiabetic (Fazilah et al., 2019; Obafemi et al., 2017; Obafemi et al., 2019), anti-aging, anticancer (Seong et al., 2021; Wang et al., 2011), anti-convulsant (Jeremiah), hepatoprotective and anti-glycemic (Haddad et al., 2020), and anti-hyperuricaemic (Shi et al., 2016) effects.

2.8 Miracle plant uses

The miracle plant is a multipurpose species with more than 60 documented uses (Ekpo et al., 2008; Fandohan et al., 2017; Oumorou et al., 2010). The most salient use of the species is its utilization as a sweetener (Fandohan et al., 2017; Oumorou et al., 2010) due to the above elicited taste modifying property. In addition, the fruit, leaves, and roots are sold on local markets in Benin and provide substantial revenue. A kilogram of fresh fruit is sold at USD 3.5 – 6.5 in Benin (Fandohan et al., 2017; Oumorou et al., 2010). All the non-edible parts of the species, including the seeds, flowers, leaves, roots and bark have medical applications (Achigan-Dako et al., 2015). The leaves are used to treat diabetes, hypothermia, enuresis, malaria, and cough, while the roots are used to treat cough, tuberculosis, and gonorrhoea (Ekpo et al., 2008; Fandohan et al., 2017; Oumorou et al., 2010). The seed serves in stomach-ache, obesity and anaemia treatment, whereas the stem is used in making agricultural tool handles (e.g., hoes), while the twigs are used as vegetal toothbrush (Oumorou et al., 2010). Most of these local uses were recorded in Benin, and documenting how other sociolinguistic groups in the species distribution ranges valued the species could offer insight into the regional importance of the species and help tailor adequate conservation and promotion strategies, as the species is currently endangered in Benin (Adomou, 2005).

If fresh miracle fruit costs less than USD 5/kg in its native region of West Africa, it fetches the astounding price of USD 2,500 per kg of pure dried powder (https://www.miraclefruitfarm.com/). As a result of its taste modifying properties and its richness in phytochemicals, the species has found applications in the medical, cosmetics and food and beverage industries. Like the leaves, the fruit is used to treat and/or to prevent diabetes, as it helps regulate insulin production (Chen et al., 2006). Similarly, the sweetening property is exploited in cancer treatment to help restore taste perception in patients undergoing chemotherapy (Wilken and Satiroff, 2012). The miracle fruit is also used in the yoghurt industry, where it serves as an encapsulation material to extend and improve the viability and stability of probiotic strains (Fazilah et al., 2019). The red exocarp has a colouring ability that is exploited to colour food, whereas the pulp is used to sweeten beverages (Buckmire and Francis, 1978). Recent developments even suggest an interest of the Coca-Cola company in exploiting this colouring and sweetening ability (Tafazoli et al., 2019). The potential applications of the numerous phytochemicals in the species suggest its promising value for drug development (Liu et al., 2021). The seed produces an oil used in cosmetics to improve hand motor skills and to treat women's hair breakage (Del Campo et al., 2017; Gorin et al., 2018).

2.9 Cytogenetics and genomic resources

Cytogenetics remains one of the least explored scientific fields in the miracle plant, and to our knowledge, the only attempts to cytologically characterize the species dates back to the 1980s. Using samples from Cameroon, Arends (1976) reported the number of chromosomes to be $2n = 26$ in the miracle plant. Later, Pilz (1981) indicated $2n = 28$ for the species, raising doubt about the exact number of chromosomes. However, considering the basic number of chromosomes ($2n = 26$) in other species of the genus, we may speculate $2n = 26$ to be the most plausible number of chromosomes for the miracle plant, although *de novo* counting including samples from a larger geographical coverage will be needed to clarify the situation and test for any origin-supported infraspecific variation in the species. At present, the indicated number of chromosomes for the species by the Chromosomes Count Database is $2n = 26$ (Rice et al., 2015) and can be adopted as such.

Conversely, a complete chloroplast genome has been assembled for the miracle plant (Niu et al., 2020). This genome has a size estimate of 158,462 bp and is made up of a large single-copy region, a small single-copy region and two inverted repeat regions of 88,256 bp, 18,669 bp, and 25,958 bp. The genome has a GC content of 36.87% and a total of 133 genes, including 88 protein-coding genes, 37 transfer RNA (tRNA) genes and eight ribosomal RNA (rRNA) genes. This very first mega genomic resource in the miracle plant opens room for extensive phylogeographic studies in the species, as the species has been spreading across the world without a clear knowledge of its route of colonization to date.

Information on the nuclear genome size is still lacking in the miracle plant; and constrains access of the miracle plant to some genomics facilities and platform as sequencing depth often depends on it. A better knowledge of the genome size of the miracle plant will open room for the proper exploitation of genomics resources existing in its sister species, such as *Argania spinosa* and *V. paradoxa,* for which a draft (Khayi et al., 2020) and reference (Hale et al., 2021) genome, respectively, exist. While such information can be generated using the flow cytometry technique, especially in the case of limited resources (Kim et al., 2015; Schmuths et al., 2004; Temsch and Greilhuber, 2001), an ambitious approach could rather target complete reference genome assembly to take advantage of the k-mer approach (Mgwatyu et al., 2020).

2.10 Germplasm collection in the miracle plant

Germplasm represents a key and strategic resource in any plant improvement program. Although such a resource was previously lacking in the miracle plant (Achigan-Dako et al., 2015), collaborative effort made by researchers of the Laboratory of Genetics, Biotechnology and Seed Science (GBioS)at the University of Abomey-Calavi (Republic of Benin) and the Cocoa Research Institute of Ghana (CRIG) over the last five years has made it possible to assemble a collection of 430 accessions from Ghana, Togo, Benin, and Nigeria, the four countries known to constitute the species centre of origin. To the best of our knowledge, this is the first worldwide collection for the miracle plant and constitutes a crucial resource to tap into, for improving the species. The collection is currently held at the GBioS and will be partly duplicated at CRIG.

2.11 Genetic diversity

Our knowledge of the miracle plant genetic diversity is still scanty. While no report exists on the species natural phenotypic diversity, only one study using Random Amplified Polymorphic DNA (RAPD) markers reported molecular diversity (Chibuzor et al., 2017) in southern Nigeria. The conclusion of the study revealed a low genetic diversity. This calls for a thorough investigation, at least in the species centre of origin. Currently, two morphotypes of the miracle fruit based on the exocarp colour can be distinguished: a red morphotype that is more common and a yellow morphotype reported in Nigeria (Njoku et al., 2015). Based on that information and given that red fruit skin colour is linked to anthocyanin pigments and yellow colour is attributed to flavonoid pigments (Buckmire and Francis, 1976), we speculate on a large variation of these two phytochemicals in the miracle fruit. Likewise, whether this exocarp based differentiation is also accompanied by any other phenotypic differentiation needs to be clarified. A recent investigation of the species diversity based on the miraculin content, in Florida (the USA), distinguished a total of nine different morphotypes, out of which the "imperial" morphotype was the most promising one (Demesyeux et al., 2020). Such a morphotype can be considered in future breeding programs. With the recent biotechnological advance, numerous sequencing platforms such as RADseq (Karam et al., 2015), Kompetitive Allele Specific PCR (Majeed et al., 2019) and DArTseq (Kilian et al., 2003) among others, have been employed to generate SNPs used to depict molecular diversity in plant species. In *Synsepalum dulcificum*, the only marker so far used is the Random Amplified Polymorphism DNA (RAPD) and employing up to date markers such as SNPs will tremendously improve the breeding pace.

2.12 Current breeding approaches

To our knowledge, no evidence of any ongoing conventional, mutation or genomics-assisted breeding currently exists in the miracle plant. This is supported by the lack of basic foundational knowledge on natural phenotypic diversity of species. Similarly, no study has so far reported on combining abilities and gene action in the species, yet such information is crucial in the choice of parental lines to cross. A mutation breeding study was initiated through the identification of optimal mutagen dosage (Chambers et al., 2018), but how the mutants phenotypically diverged is yet to be reported. Conversely, transgenic breeding has been intensively developed since 2011 with the miraculin as a target trait and successfully expressed in a number of species including tomato (*Lycopersicon esculentum* (L.) H. Karst.) (Hiwasa-Tanase et al., 2011; Ono et al., 2021); lettuce (*Lactuca sativa* L.) (Hirai et al., 2011); strawberry (Fragaria X Annasa) (Sugaya et al., 2008), carrot (*Daucus carota* L.) (Park et al., 2020; 2021) and rice (*Oryza sativa* L.) (Kim et al., 2020). Of all these crops, tomato showed the best miraculin yield, and recent environmental safety trials conducted on genetically modified miraculin-accumulating tomato (GM tomato) concluded an absence of risk for GM tomato production (Hiwasa-Tanase et al., 2021), thus opening room for large scale cultivation of commercial exploitation of GMO tomato in Japan. However, recombinant miraculin expression differed from the miracle plant improvement, the latter having in scope all functional traits (given the diversity of use of the miracle plant with).

2.13 Proposed step-by-step guide towards miracle plant improvement in West Africa

Ghana, Togo, Benin, and Nigeria are the countries implicated in the miracle plant centre of diversity, origin and uses, end-users in West Africa and will, therefore, mostly benefit from an improved variety of the species. These four countries can constitute the miracle plant breeding mega-environment.

2.13.1 End-users preferred traits definition: step 1

An important part of modern miracle plant applications lies in the exploitation of its taste-modifying property, which is attributed to miraculin. The miraculin has been extensively studied, and its exogenous expression in other species placed it as a potential candidate trait for breeding. However, in the current context, this can be done in countries such as the USA and Japan, but whether West Africa also perceives the trait as key needs to be deciphered. Moreover, a common justification for the transgenic expression of the miraculin is the supposedly low productivity of the miracle plant (Ono et al., 2021; Park et al., 2021), which seems more like a speculation as there is no data establishing the fact. Nevertheless, breeding for highly productive miracle plants will probably benefit West African farmers. However, despite the two traits, breeding priorities need to be established as these can vary with region and end-users' preferences. Hence, it is necessary to employ a participatory approach involving key end-user groups of the species to identify and prioritize breeding traits to focus on. Benin and Ghana are hotspots of species utilization in West Africa, and trait preference exercises can prioritize these two countries and then extend to Togo and Nigeria. Key end-user groups to consider include producers, consumers, and processors.

2.13.2 Germplasm collection: step 2

Germplasm is a key resource in plant breeding. A large collection of 350 accessions of the miracle plant has already been constituted and could be tapped for traits' improvement. Nevertheless, other natural distribution areas of the species (e.g., Cameroon, Central Africa, Democratic Republic of Congo) are yet to be represented in this collection. Consequently, further prospection and collection activities in Central Africa need to be considered. For now, the current germplasm of 350 individuals can be utilised.

2.13.3 Molecular diversity and population structure profiling: step 3

Knowledge of genetic variation organization is crucial in planning breeding strategies in tree crops. Historically, markers such as RAPD, Amplified Fragment Length Polymorphisms (AFLP) and Single Sequence Repeats (SSR) have been used to depict molecular diversity in several tree crop species (De la Rosa et al., 2003; Martinez-Armental et al., 2003; Sanou et al., 2005), but the genomic era makes it possible to employ advanced sequencing technologies and platforms to yield SNP markers that are most suitable to reveal the genetic variation pattern at an unpreceded resolution and cost-effectively (Sansaloni et al., 2011; You et al., 2018). In this context, employing SNP markers to understand the molecular variation of the miracle plant at its centre of diversity, is highly desired. Because the miracle plant currently lacks a reference genome, sequencing technology such as the Diversity Array Technology sequencing (DArTseq) platform (Kilian et al., 2003) and SNP calling pipeline such as the genotype-by-sequencing (GBS) SNP-Calling Reference Optional Pipeline (GBS-SNP-CROP) (Melo et al., 2016) may be recommended to depict the diversity and infer population structure in the species. To that end, involving samples from the whole natural distribution range of the species spanning from Ghana to the Democratic Republic of Congo could offer a complete view of the population structure pattern in the species. We recommend molecular assessment with a minimum size of 350 individuals. In practice this germplasm size may be expanded depending on the available resources, but for the purpose of this breeding plan, we use a hypothetical size of 350 individuals as initial germplasm size.

2.13.4 Core collection development as an active working sample: step 4

Germplasm collection, maintenance and phenotypic characterization are key in any plant breeding program. But in a recalcitrant tree crop species such as the miracle plant, living collection establishment, the only sustainable and most cost-effective manner to actively maintain the species germplasm, and evaluation of a large population is laborious and tedious. Consequently, the development of a core collection intended to capture the maximum diversity in the least redundant possible accessions represents a nice approach, especially in cases of limited space and resources. Core collection can be developed using either phenotypic data and/or molecular (marker) data, giving rise to three possible core collection types. In the breeding strategy of the miracle plant, since SNP markers will have been generated during population structure assessment in step 3, we recommend prioritizing the molecular marker-based core collection even if a phenotype-based core collection can also be developed if in situ phenotypic data were collected at a certain point. Considering the standard that a core collection should be made of approximately 10–30% of the initial germplasm size (Hu et al., 2000; Ortiz et al., 1998), we expect the development of a SNP-based core collection of approximately 105 highly diverse accessions (considering an initial germplasm size of 350 accessions, for instance) to constitute an active working population in the miracle plant breeding strategy. At least three core collection development methods should be tested from the wealth of existing methods including the advanced M strategy in the PowerCore software (Kim et al., 2007), the GenoCore (Jeong et al., 2017), the coreHunter in the coreHunter 3 package (De Beukelaer et al., 2018), the least distance stepwise sampling strategy (Liu et al., 2020) and the best one selected to retain the final core sample.

2.13.5 Phenotypic diversity assessment: step 5

The SNP-based core collection of 105 individuals obtained in step 4 will be extensively phenotyped using a common garden approach in two countries with contrasting environmental conditions [e.g., Ghana (humid) and Benin (moderately dry)] and in at least three localities per country. Key traits to consider should include all farmers' desired breeding traits in addition to the miracle fruit ionome (micronutrients, macronutrients, and trace elements) and the fruit major phenols, anthocyanins, and carotenoids. Based on all traits evaluated, a multi-trait stability index (a measure of overall performance and stability for all traits combined) (Olivoto and Lúcio, 2020) will be computed per accession and the top ten accessions selected. In addition, broad-sense heritability estimates and genetic correlations could be computed for all traits.

2.13.6 Combining ability and gene action study: step 6

The top ten accessions selected from the phenotypic assessment in step 5 will be used in a formal conventional breeding scheme for a proper study of combining ability, gene action, heterosis and narrow sense heritability for each trait of interest. To that end, a full diallel mating design without selfing will help generate 90 full sib-families to be evaluated in the same environment as in Step 4 with parental lines. Here, early fruiting will be triggered through proper water and nutrient management as per Tchokponhoué et al. (2017) to speed up the conventional breeding process. Based on the gene action, selection strategies will be adjusted, and crosses exhibiting the highest heterosis for each trait will be considered (if gene action permits it), and hybrids will be selfed on two to four generations for final evaluation before the registration process.

2.13.7 Genomic selection: step 7

While miracle plant conventional breeding will focus on exploiting the core collection of 105 individuals, the large possibilities offered by genomic selection to rapidly select elite individuals based on their Genomic Estimated Breeding Values (GEBV) will be exploited using the whole initial germplasm. Genomic selection is known to work with at least two populations: a training population (TP) and a candidate population (CP). Here, the training population, to phenotype, will be made of the core collection of 105 accessions (C1) of step 4 plus a second core collection built on the remaining set of 245 accessions (350 - 105). Considering the same criteria of 30% germplasm size as core collection size, the second collection size will be made of C2 = 75 accessions to give rise to a TP size of 180 accessions (TP = C1 + C2) and then a candidate population of 170 individuals. The training population will be evaluated in six environments, three in Benin and three in Ghana. Models calibrated based on phenotypic and marker data in TP will be used to estimate GEBV in CP and the best 25 accessions selected for multilocation trials in Ghana and Benin. The multi-trait stability index will be employed to identify the top 10 accessions that will be either selfed and released or engaged in a round of full diallel mating design to generate a set of 90 new individuals whose GEBV will also be computed and involved in another selection cycle.

2.13.8 Genome-wide association study for quantitative trait nucleotide discovery: step 8

A genome-wide association study in the miracle plant will also make use of the largely phenotyped TP (180 accessions) in step 7 and their marker data to identify SNPs associated with breeding traits of interest. Since the miracle plant currently lacks a reference genome, the available reference genome of its sister species *V. paradoxa* or *Argania spinosa* can be exploited. However, based on the AOCC priority crops, the African star apple (*Chrysophyllum albidum* G. Don) seems to be the best sister species to use in lieu of the miracle plant’s own genome because of the similarity of their number of chromosomes (2n = 26), but also their high phylogenetic relatedness.

2.13.9 Exploring genome editing option: step 9

Genome editing represents a powerful biotechnology tool to accelerate breeding of orphan crops such as the miracle plant. Unfortunately, the application of this tool is often constrained in most of orphan crops due to the lack of a reference genome and the absence of an effective plant transformation system (Hua et al., 2019; Lemmon et al., 2018). While the miracle plant currently lacks a reference genome, the bottleneck that constitutes the nonexistence of an effective transformation system has been unlocked since 2011 with the transgenic expression of miraculin in tomato (Hiwasa-Tanase et al., 2011). This suggests that the only barrier for the miracle plant to access gene editing is the lack of a reference genome. Thus, effort should be targeted towards a reference genome, as this will undoubtedly make room for unprecedented discoveries in this miracle plant.

2.13.10 Giving price to collaboration: cross-cutting step

Despite being a West African native species, the miracle plant has become a worldwide promising orphan crop and is attracting increased attention. To accelerate knowledge generation on this crop and attain region-wide elite planting materials, there is a necessity to collaborate at least at the regional level. A mere illustration of the importance of such collaboration is the currently available collection for the species. Expanding this collection with accessions from central Africa will be much easier through a collaborative effort. In addition, a regional collaboration will speed up breeding activities and help attract more funding that will be crucial in further stages of the breeding process. Finally, there is urgent need to engage the miracle plant breeding community regarding human resources capacity building to effectively keep pace and take advantage of the fast-evolving biotechnological era. A crucial example is the genome editing through which we foresee the possibility of a rapid fruit size improvement in the species.

2.14 Conclusion

The miracle plant through its multiple applications has the potential to sustainably contribute to the West African socio-economic growth and help the region approach SDGs. For this to be effective, it is necessary to engage active research on the species in the sub-region, including the development of elite planting material, the promotion of large scale-cultivation and the development of the species value chain, among others. This thesis tackles aspects pertaining to the species breeding by generating first-order information that is required to implement a sound breeding program of the species in the West African region and calls on the necessity to speed up the entry of the miracle plant in the era of genomics.

2.15 References

Abdellrhman, T. E.-K., and Adam, Y. O. (2020). Value chain analysis of baobab (*Adansonia digitata* L.) Fruits in Blue Nile State, Sudan *Agric. For. J.* **4**, 131-138.

Achigan-Dako, E. G., Pasquini, M. W., Assogba Komlan, F., N'danikou, S., Yédomonhan, H., Dansi, A., and Ambrose-Oji, B. (2010). Traditional vegetables in Benin. *Cotonou, Benin: Institut National des Recherches Agricoles du Bénin, Imprimeries du CENAP*.

Achigan-Dako, E. G., Tchokponhoué, D. A., N'Danikou, S., Gebauer, J., and Vodouhè, R. S. (2015). Current knowledge and breeding perspectives for the miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell. *Genet. Resour. Crop Evol.* **62**, 465-476.

Adomou, A. (2005). Vegetation patterns and environmental gradients in Benin, University of Wageningen, Wageningen, Netherlands.

Ahmad, M. Z., Masood, S., Sultana, S., Hadda, T. B., Bader, A., and Zafar, M. (2015). Antioxydant and nutraceutical value of wild medicinal Rubus berries. *Pak. J. Pharm. Sci.* **28**, 241-247.

Akbar, E., Yaakob, Z., Kamarudin, S. K., Ismail, M., and Salimon, J. (2009). Characteristics and composition of Jatropha curcas seed oil from Malaysis and its potential as biodiesel feedstock. *Eur. J. Sci. Res.* **29**, 396-403.

Akinmoladun, A. C., Adetuyi, A. R., Komolafe, K., and Oguntibeju, O. O. (2020). Nutritional benefits, phytochemical constituents, ethnomedicinal uses and biological properties of Miracle fruit plant (*Synsepalum dulcificum* Shumach. & Thonn. Daniell). *Heliyon* **6**, e05837.

Akoègninou, A., Van der Burg, W., and Van der Maesen, L. J. G. (2006). "Flore analytique du Bénin," Backhuys Publishers.

Arends, J. (1976). Somatic chromosome numbers of some African Sapotaceae. *Acta Bot. Neerlendica* **25**, 449-457.

Atangana, A. R., van der Vlis, E., Khasa, D. P., van Houten, D., Beaulieu, J., and Hendrickx, H. (2011). Tree-to-tree variation in stearic and oleic acid content in seed fat from *Allanblackia floribunda* from wild stands: potential for tree breeding. *Food Chem.* **126**, 1579-1585.

Atuahene, M. F., and Osekre, E. A. (2018). Insect pests of sweet berry (*Richardella dulcifica*) (Schunach & Thonn.) Baehni in Ghana. *Int. J. Sci. Technol. Res.* **7**.

Awotedu, O. L., and Ogunbamowo, P. O. (2019). Nutritional , antitnutritional and phytochemiclas profile of the leaves anf fruits of *Synsepalum dulcificum*. *Am. J. Biol. Chem.* **7**.

Ayensu, E. S. (1972). Morphology and anatomy of *Synsepalum dulcificum* (Sapotaceae). *Bot. J. Linn. Soc.* **65**, 179-187.

Brouwer, J. N., Van Der Wel, H., Francke, A., and Hennig, G. J. (1968). Miraculin, the sweetening protein from the miracle fruit. *Nature* **220**, 373-374.

Buckmire, R., and Francis, F. (1976). Anthocyanins and flavonols of miracle fruit, *Synsepalum dulcificum*, Schum. *J. Food Sci.* **41**, 1363-1365.

Buckmire, R., and Francis, F. (1978). Pigments of miracle fruit, *Synsepalum dulcificum*, Schum, as potential food colorants. *J. Food Sci.* **43**, 908-911.

Chambers, A. H., Demexyeux, L., Moon, P., and Fu, Y. (2018). Optimization of miracle fruit (Synsepalum dulcificum) seed germination and mutagenesis. *Afr. J. Food Sci. Technol.* **9**.

Chen, C. C., Liu, I. M., and Cheng, J. T. (2006). Improvement of insulin resistance by miracle fruit (S*ynsepalum dulcificum*) in fructose-rich chow-fed rats. *Phytother. Res.* **20**, 987-992.

Chen, X. W., Abdullah, T. L., Abdullah, N. A. P., and Hassan, S. A. (2012). Rooting response of miracle fruit (*Synsepalum dulcificum*) softwood cuttings as affected by indole butyric acid. *Am. J. Agric. Biol. Sci.* **7**, 442-46.

Chen, Y., Wei, G., and Chen, W. (2002). New species of Pestalotiopsis. *Mycosystema* **21**.

Chibuzor, I. A., Bukola, O., Adejoke, A. O., and Chidozie, O. P. (2017). Genetic assessment of the shrub *Synsepalum dulcificum* (Schumach & Thonn.) Daniell in Nigeria using the Randomly Amplified Polymorphic DNA (RAPD). *Int. J. Genet. Genomics* **4**, 45-50.

Daniel, W. F. (1852). On the *Synsepalum dulcificum*, De Cand. or, miraculuous berry of Western Africa. *Pharm. J.* **11**, 445-448.

De Beukelaer, H., Davenport, G. F., and Fack, V. (2018). Core Hunter 3: flexible core subset selection. *BMC Bioinformatics* **19:203**, 1-12.

De la Rosa, R., Angiollilo, A., Guerrero, C., Pellegrini, M., Rallo, L., Besnard, G., Berville, A., Martin, A., and Baldoni, I. (2003). A first linkage map of olive (*Olea europeae* L.) cultivars using RAPD, AFLP, RFLP and SSR markers. *Theor. Appl. Genet.* **106**, 1273-1282.

Del Campo, R., Zhang, Y., and Wakeford, C. (2017). Effect of miracle fruit (*Synsepalum dulcificum*) seed oil (MFSO®) on the measurable improvement of hair breakage in women with damaged hair: a randomized, double-blind, placebo-controlled, eight-month trial. *J. Clin. Aesthet. Dermat.* **10**, 39-48.

Demesyeux, L., Brym, M., Valdes, D., Collazo, C., and Chambers, A. H. (2020). Yield and miraculin content of nine miracle fruit (*Synsepalum dulcificum*) morphotypes. *Euphytica* **216**, 1-12.

Du, L., Shen, Y., Zhang, X., Prinyawiwatkul, W., and Xu, Z. (2014). Antioxidant-rich phytochemicals in miracle berry (*Synsepalum dulcificum*) and antioxidant activity of its extracts. *Food Chem.* **153**, 279-284.

Edem, C. A., Dosunmu, M. I., Ebong, A. C., and Jones, M. (2008). Determination of proximate composition of ascorbic acid and heavy metal content of star fruit (*Averrhoa carambola*). *Glob. J. Pure Appl. Sci.* **14**, 193-195.

Edem, D. (2002). Palm oil: biochemical, physiological, nutritional, hematological and toxilogica aspects: a review. *Plant Foods Hum. Nutr.* **57**, 319-341.

Ekpo, B. A., Bala, D. N., Essien, E. E., and Adesanya, S. A. (2008). Ethnobotanical survey of Akwa Ibom state of Nigeria. *J. Ethnopharmacol.* **115**, 387-408.

Elias, M., and Carney, J. (2007). African shea butter: a feminized subsidy from nature. *Afr. J. Int. Afr. Inst.* **77**, 37–62.

Fandohan, A. B., Gouwakinnou, G. N., Tovissode, C. F., Bonou, A., Djonlonkou, S. F. B., Houndelo, L. F., Sinsin, C. L. B., and Assogbadjo, A. E. (2017). Usages traditionnels et valeur économique de *Synsepalum dulcificum* au Sud-Bénin. *Bois For. Trop.* **332**, 17-30.

Fazilah, N. F., Hamidon, N. H., Ariff, A. B., Khayat, M. E., Wasoh, H., and Halim, M. (2019). Microencapsulation of Lactococcus lactis Gh1 with gum Arabic and *Synsepalum dulcificum* via spray drying for potential inclusion in functional yogurt. *Molecules* **24**, 1422.

Fischer, S., Jäckering, L., and Kehlenbeck, K. (2020). The Baobab (*Adansonia digitata* L.) in Southern Kenya–A Study on Status, Distribution, Use and Importance in Taita–Taveta County. *Environ. Manag.* **66**, 305-318.

Geoffery, M., and Sani, H. (2017). Promotion of adventitious root formation of miracle fruit (*Synsepalum dulcificum* daniell) through stem cuttings and air layering technique. *Transactions Sci. Technol.* **4**, 1-7.

Gorin, S., Wakeford, C., Zhang, G., Sukamtoh, E., Matteliano, C. J., and Finch, A. E. (2018). Beneficial effects of an investigational wristband containing *Synsepalum dulcificum* (miracle fruit) seed oil on the performance of hand and finger motor skills in healthy subjects: A randomized controlled preliminary study. *Phytother. Res.* **32**, 321-332.

Gouwakinnou, G. N., Lykke, A. M., Assogbadjo, A. E., and Sinsin, B. (2011). Local knowledge, pattern and diversity of use of *Sclerocarya birrea*. *J Ethnobiol Ethnomed* **7**, 1-9.

Haddad, S. G., Mohammad, M., Raafat, K., and Saleh, F. A. (2020). Antihyperglycemic and hepatoprotective properties of miracle fruit (*Synsepalum dulcificum*) compared to aspartame in alloxan-induced diabetic mice. *J. Integr. Med.* **18**, 514-521.

Hale, I., Ma, X., Melo, A. T., Padi, F. K., Hendre, P. S., Kingan, S. B., Sullivan, S. T., Chen, S., Boffa, J.-M., Muchugi, A., Danquah, A., Barnor, M. T., Jamnadass, R., van de Peer, Y., and van Deynze, A. (2021). Genomic resources to guide improvement of the shea tree. *Front. Plant Sci.* **12**.

He, Z., Tan, J. S., Abbasiliasi, S., Lai, O. M., Tam, Y. J., and Ariff, A. B. (2016). Phytochemicals, nutritionals and antioxidant properties of miracle fruit Synsepalum dulcificum. *Ind. Crops Prod.* **86**, 87-94.

Hendre, P. S., Muthemba, S., Robert Kariba, R., Alice Muchugi, A., Fu, Y., Chang, Y., Song, B., Liu, H., Liu, M., Liao, X., Sahu, S. K., Wang, S., Li, L., Haorong Lu, H., Shufeng Peng, S., Shifeng Cheng, S., Xun Xu, X., Huanming Yang, H., Jian Wang, J., Xin Liu, X., Anthony Simons, A., Shapiro, H.-Y., Mumm, R. H., Van Deynze, A., and Jamnadass, R. (2019). African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops. *Planta* **250**, 989–1003.

Hirai, T., Sato, M., Toyooka, K., Sun, H.-J., Yano, M., and Ezura, H. (2010). Miraculin, a taste-modifying protein is secreted into intercellular spaces in plant cells. *Journal of plant physiology* **167**, 209-215.

Hirai, T., Shohael, A. M., Kim, Y.-W., Yano, M., and Ezura, H. (2011). Ubiquitin promoter–terminator cassette promotes genetically stable expression of the taste-modifying protein miraculin in transgenic lettuce. *Plant Cell Rep.* **30**, 2255-2265.

Hiwasa-Tanase, K., Nyarubona, M., Hirai, T., Kato, K., Ichikawa, T., and Ezura, H. (2011). High-level accumulation of recombinant miraculin protein in transgenic tomatoes expressing a synthetic miraculin gene with optimized codon usage terminated by the native miraculin terminator. *Plant Cell Rep.* **30**, 113-124.

Hiwasa-Tanase, K., Yano, T., Kon, T., Terakawa, T., and Ezura, H. (2021). Environmental risk assessment of transgenic miraculin-accumulating tomato in a confined field trial in Japan. *Plant Biotechnol.* **21**.

Hu, J., Zhu, J., and Xu, H. (2000). Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theor. Appl. Genet.* **101**, 264-268.

Hua, K., Zhang, J. S., Botella, J. R., Ma, C. L., Kong, F. J., Liu, B. H., and Zhu, J. K. (2019). Perspectives on the application of genome-editing technologies in crop breeding. *Mol. Plant* **12**, 1047 -1059.

Huang, W., Chung, H. Y., Xuan, W., Wang, G., and Li, Y. (2020). The cholesterol-lowering activity of miracle fruit (Synsepalum dulcificum). *J. Food Biochem.* **44**, e13185.

Igeta, H., Tamura, Y., Nakaya, K., Nakamura, Y., and Kurihara, Y. (1991). Determination of disulfide array and subunit structure of tastemodifying protein, miraculin, . *Biochim. Biophys. Acta* **1079** 303–307.

Inglett, G., Dowling, B., Albrecht, J., and Hoglan, F. (1965). Taste modifiers, taste-modifying properties of miracle fruit (*Synsepalum dulcificum*). *J. Agric. Food Chem.* **13**, 284-287.

Jäckering, L., Fischer, S., and Kehlenbeck, K. (2019). A value chain analysis of baobab (*Adansonia digitata* L.) products in Eastern and Coastal Kenya. *JARTS* **120**, 91-104.

Jamnadass, R., Mumm, R. H., Hale, I., Hendre, P., Muchugi, A., Dawson, I. K., Powell, W., Graudal, L., Yana-Shapiro, H., Simons, A. J., and Van Deynze, A. (2020). Enhancing African orphan crops with genomics. *Nat. Genet.* **52**, 356–360.

Jeong, S., Kim, J.-Y., Jeong, S.-C., Kang, S.-T., Moon, J.-K., and Kim, N. (2017). GenoCore: A simple and fast algorithm for core subset selection from large genotype datasets. *PLoS ONE* **12**.

Joyner, G. (2006). The miracle fruit. *In* "Quandong magazine of the west Australian nut and tree crop association" (P. Scott, ed.), pp. 15. West Australian Nut and Tree Crop Association, Subiaco, West Australia.

Karam, M. J., Lefevre, F., Dagher-Kharrat, B. M., Pinosio, S., Vendramin, G. G. (2015). Genomic exploration and molecular marker development in a large and complex conifer genome using RADseq and mRNAseq. *Mol. Ecol. Res*. **3**, 601-612.

Kamatou, G., Vermaak, I., and Viljoen, A. (2011). An updated review of Adansonia digitata: A commercially important African tree. *S. Afr. J. Bot.* **77**, 908-919.

Khayi, S., Gaboun, F., Pirro, S., Tatusova, T., El Mousadik, A., Ghazal, H., and Mentag, R. (2020). Complete chloroplast genome of *Argania spinosa*: structural organization and phylogenetic relationships in Sapotaceae. *Plants* **9**.

Kilian, A., Huttner, E., Wenzl, P. E., Jacoud, D., Carling, J., Caig, V., Evers, M., Heller-Uszynska, K. A., Cayla, C., Patarapuwadol, C., and al., e. (2003). The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. *In* "International congress in the wake of the double helix: from the green revolution to the gene revolution", pp. 443-461, Bologna, Italy.

Kim, H. K., Go, J. Y., Park, S.-Y., Kang, K. K., and Jung, Y. J. (2020). Production of miraculin protein in suspension cell lines of transgenic rice using Agrobacterium. *J. Plant Biotechnol.* **47**, 227-234.

Kim, K.-W., Chung, H.-K., Cho, G.-T., Ma, K.-H., Chandrabalan, D., Gwag, J.-G., Kim, T.-S., Cho, E.-G., and Park, Y.-J. (2007). PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* **23**, 2155-2162.

Kim, S., Han, M., and Rayburn, A. L. (2015). Genome size and seed mass analyses in Cicer arietinum (Chickpea) and wild Cicer species. *HortScience* **50**, 1751-1756.

Kurihara, K., and Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science* **161**, 1241-1243.

Lemmon, Z. H., Reem, N. T., Dalrymple, J., Soyk, S., Swartwood, K. E., Rodrgıguez-Leal, D., van Eck, J., and Lippman, Z. B. (2018). Rapid inprovement od domestication traits in an orphan crop by genome editing. *Nature Plants* **4**.

Liu, M., Hu, X., Wang, X. W., Zhang, J., Peng, X., Hu, Z., and Liu, Y. (2020). Constructinga core collection of the medicinl plant *Angelica biserrata* using genetic and metabolic data. *Front. Plant Sci.* **11**.

Liu, Y.-G., L, B., Fu, Q., Zhang, X.-M., Ma, F.-Y., and Hu, Y. H. (2021). Miracle fruit leaf extract: antioxidant activity evaluation, constituentidentification, and medical applications. *Anal. Lett.* **54**, 2211-2226.

Mabhaudhi, T., Chimonyo, V. G. P., Hlahla, S., Massawe, F., Mayes, S., Nhamo, L., and Modi, A. T. (2019). Prospects of orphan crops in climate change. *Planta* **250**, 695-708.

Majeed, U., Darwish, E., Rehman, Zhang, X. (2019). Kompetitive allele specific PCR: a singleplex genotyping platformand its application. *J. Agric. Sci.* **11**(1), 11-20.

Macit, I., Aydin, E., Tas, A., and Gondogdu, M. (2021). Fruit quality properties of the local apple varieties of Anatolia. *Sustainability* **13**.

Martinez-Armental, J., Sozzi, G. O., Sanchez-perez, R., Rubio, M., and Gradziel, T. M. (2003). New approaches to Prunus tree crop breeding. *Food, Agric. Env.* **1**, 52-63.

Masuda, Y., Nirasawa, S., Nakaya, K., and Kurihara, Y. C. (1995). CloningandsequencingofacDNA encoding ataste-modifyingprotein,miraculin. *Gene* **161**, 175–177.

Melo, A. T., Bartaula, R., and Hale, I. (2016). GBS-SNP-CROP: a reference-optional pipeline for SNP discovery and plant germplasm characterization using variable length, paired-end genotyping-by-sequencing data. *BMC Bioinformatics* **17**, 29.

Mgwatyu, Y., Stander, A. A., Ferreira, S., Williams, W., and Hesse, U. (2020). Rooibos (*Aspathus linearis*) genome size estimation using flow cytometry and k-mer analyses. *Plants* **9**.

Mpeck, N., Asaah, M. L., Tchoundjeu, Z. E. K., and Atangana, A. R. (2003). Strategies for the domestication of *Ricinodendron heudelotii*: Evaluation of variability in natural populations from Cameroon. *Food Agric. Env. and Env.* **13 & 14**

257–262.

N'Danikou, S., and Tchokponhoué, D. A. (2020). Plant domestication for enhanced food security. *In* "Zero Hunger" (W. L. Filho, A. M. Azul, L. Brandli, P. G. Ozuyar and T. Wall, eds.), pp. 644-654. Springer Nature, Switzerland.

Naughton, C. C., Lovett, P. N., and Mihelcic, J. R. (2015). Land suitability modeling of shea (*Vitellaria paradoxa*) distribution across sub-Saharan Africa. *Appl. Geogr.* **58**, 217–227.

Niu, Y.-F., Ni, S.-B., and Liu, J. (2020). Complete cholroplast genome of Synsepalum dulcificum: a magical plant that modifies sour flavors to sweet. *Mitochondrional DNA Part B* **5**, 3052-3053.

Njoku, N. E., N., U. C., S.O., A., I.M., A., and C.N., E. (2016). Proximate, anti-nutritional and phytochemical composition of the yellow variety of the *Synsepalum dulcificum* (Miracle fruit) berry. *Amer. J. Food Sci. Technol.* **4**, 102-108.

Njoku, N. E., Ubbaonu, C. N., Alagbaoso, S. O., Eluchie, C. N., and Umelo, M. C. (2015). Amino acid profile and oxidizable vitamin content of *Synsepalum dulcificum* berry (miracle fruit) pulp. *Food Sci. Nutr.* **3**, 252-256.

Nkwocha, C., Njoku, O., and Ekwueme, F. (2014). Proximate and micronutrient analyses of synsepalum dulcificum pulp. *Scient. Res. J.* **2**, 2201-2796.

Nwonmu, F. O. C. (2006). The socio-cultural and economic relevance of the marula tree and its susitainability in South Africa. *Africa Insight* **36**, 249-265.

Obafemi, T., Akinmoladun, A., Olaleye, M., Agboade, S. O., and Onasanya, A. A. (2017). Antidiabetic potential of methanolic and flavonoid-rich leaf extracts of *Synsepalum dulcificum* in type 2 diabetic rats. *J-AIM* **8**, 238-246.

Obafemi, T. O., Olaleye, M. T., and Akinmoladun, A. C. (2019). Antidiabetic property of miracle fruit plant (*Synsepalum dulcificum* Shumach. & Thonn. Daniell) leaf extracts in fructose-fed streptozotocin-injected rats via anti-inflammatory activity and inhibition of carbohydrate metabolizing enzymes. *J. Ethnopharmacol.* **244**, 112124.

Olivoto, T., and Lúcio, A. D. C. (2020). metan: An R package for multi-environment trial analysis. *Meth. Ecol. Evol.* **11**, 783-789.

Olufemi, O. O., Peter, Y. O., Ayooluwa, O. A., and Olabisi, O. C. (2019). Growth medium and soil amendment influence on seedling growth responses of African star apple (*Chrysophyllum albidum*). *J. Bot. Res.* **1**, 16-26.

Ono, A., Hiwasa-Tanase, K., Nonaka, S., and Ezura, H. (2021). The accumulation of recombinant miraculin is independent of fruit size in tomato. *Plant Biotechnol.*, 20.0904 a.

Ortiz, R., Ruiz-Tapia, E. N., and Mujica-Sanchez, A. (1998). Sampling strategy for a core collecton of Peruvian quinoa germplasm. *Theor. Appl. Genet.* **96**, 475-483.

Otsuka, K., Kamada, K., Ogimi, C., Hirata, E., Takushi, A., and Takeda, Y. (1994). Alangionosides A and B , ionol glycosydes from leaves of *Alangium premnifolium*. *Phytochemistry* **35**, 1331-1334.

Oumorou, M., Dah-Dovonon, J., Aboh, B., Hounsoukaka, M., and Sinsin, B. (2010). Contribution á la conservation de *Synsepalum dulcificum*: régénération et importance socio-économique dans le département de l'ouémé (Bénin). *Ann. Sci. Agron* **14**, 101-120.

Paladino, A., Costantini, S., Colonna, G., and Facchiano, A. M. (2008). Molecular modelling of miraculin: structural analyses and functional hypotheses. *BBRC* **367**, 26-32.

Park, Y.-J., Han, J.-E., Lee, H., Jung, Y.-J., Murthy, H. N., and Park, S.-Y. (2020). Large-scale production of recombinant miraculin protein in transgenic carrot callus suspension cultures using air-lift bioreactors. *AMB Express* **10**, 1-7.

Park, Y.-J., Han, J.-E., Lee, H., Lee, J.-Y., Ho, T.-T., and Park, S.-Y. (2021). Production of recombinant miraculin protein in carrot callus via Agrobacterium-mediated transformation. *Plant Cell, Tissue and Organ Cult* **145**, 615-623.

Pilling, D., Bélanger, J., Diulgheroff, S., Koskela, J., Leroy, G., Mair, G., and Hoffmann, I. (2020). Global status of genetic resources for food and agriculture: challenges and research needs. *In* "Genet. Resour.", Vol. 1, pp. 4-16.

Pilz, G. E. (1981). Sapotaceae of Panama. *Ann. Mo. Bot. Gard.* **68**, 172-203.

Purugganan, M. D. (2019). Evolutionary insights into the nature of plant doestication. *Curr. Biol.* **29**, R705-R714.

Rhodes, B. (2009). ""Testing the Baobab's mettle: an evaluation of the suistanability and economic potential of harvesting *Adansonia digitata*." University of North Carolina Chapel Hill.

Rice, A., gLICk, L., Abadi, S., Einhorn, M., KOPelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., and Mayrose, I. (2015). The chromosome count database (CCDB). *New Phytol.* **206**, 19-26.

Royal Botanic Gardens Kew (2021). World's ckeclist of vascular Plants, version. 2.

Ruiz, A., Hermosin-Gutirez, I., Vergara, C., von Baer, D., Zapata, M., and Hitschfeld, A. (2013). Anthocyanin profiles in south Patagonian Wild berries by HPLC-DAD-ESI/MS/MS. *Food Res. Int.* **51**, 706-713.

Sanou, H., Kambou, S., Teklehaymanot, Z., Dembele, M., Yossi, H., Sina, S., Djingdia, L., and Bouvet, J.-M. (2004). Vegetative propagation of *Vitellaria paradoxa* by grafting. *Agr. Syst.* **60**.

Sanou, H., Lovett, N. P., and Bouvet, J.-M. (2005). Comparison of quantitative and molecular variation in agroforestry populations of the shea tree (*Vitellaria paradoxa* C.F. Gaertn) in Mali. *Mol. Ecol.* **14**.

Sansaloni, C. P., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., and Kilian, A. (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: Genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proc.* **5**.

Schmuths, H., Meister, A., Horres, R., and Bachmann, K. (2004). Genome size variation among accessions of *Arabidopsis thaliana*. *Annals of Botany* **93**, 317-321.

Seong, J., Oyong, G. G., and Cabrea, E. C. (2021). *Synsepalum dulcificum* extracts exhibit cytotoxic activity on human colorectal cancer cells and upregulated c-fos and c-jun early apoptotic gene expression. *Asian Pacific J. Trop. Biomed.* **16**.

Shi, Y.-C., Lin, K.-S., Jhai, Y.-F., Lee, B.-H., Han, Y., Cui, Z., Hsu, W.-H., and Wu, S.-C. (2016). Miracle fruit (Synsepalum dulcificum) exhibits as a novel anti-hyperuricaemia agent. *Molecules* **21**, 140.

Stadlmayr, B., Wanangwe, J., Waruhiu, C. G., Jamnadass, R., and Kehlenbeck, K. (2020). Nutritional composition of baobab (Adansonia digitata L.) fruit pulp sampled at different. *J. Food Comp. Anal.* **94**.

Sugaya, T., Yano, M., Sun, H.-J., Hirai, T., and Ezura, H. (2008). Transgenic strawberry expressing the taste-modifying protein miraculin. *Plant Biotechnol.* **25**, 329-333.

Swenson, U., and Enderberg, A. A. (2003). Phylogeny, character evolution, and classification of Sapotaceae (Ericales). *Cladistics* **21**, 101-130.

Swenson, U., Richardson, J. E., and Bartish, I. V. (2008). Multi-gene phylogeny of the pantropical subfamily Chrysophylloideae (Sapotaceae): evidence of generic polyphyly and extensive morphological homoplasy. *Cladistics* **24**, 1006–1031.

Tadele, Z., and Bartels, D. (2019). Promoting orphan crops research and development. *Planta* **250**, 675-676.

Tafazoli, S., Vo, T. D., Roberts, A., Rodriguez, C., Viñas, R., Madonna, M. E., Chiang, Y.-H., Noronha, J. W., Holguin, J. C., and Ryder, J. A. (2019). Safety assessment of miraculin using in silico and in vitro digestibility analyses. *Food Chem. Toxicol.* **133**, 1-10.

Tchokponhoué, D., Achigan-Dako, E., N'Danikou, S., Houdégbé, A., Agossou, C., Assogba-Komlan, F., and Vodouhè, R. (2018). Regeneration ability and seedling growth in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Fruits* **73**, 13-21.

Tchokponhoué, D. A., N'Danikou, S., Hale, I., Van Deynze, A., and Achigan-Dako, E. G. (2017). Early fruiting in *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell juveniles induced by water and inorganic nutrient management. *F1000Research* **6**.

Tchokponhoué, D. A., N'Danikou, S., and Achigan-Dako, E. G. (2019a). A combination of approaches evidenced seed storage behaviour in the miracle berry *Synsepalum dulcificum* (Schumach. et Thonn.) Daniell. *BMC Plt. Biol.* **19:117**, 1-13.

Tchokponhoué, D. A., N'Danikou, S., Houéto, J. S., and Achigan-Dako, E. G. (2019b). Shade and nutrient-mediated phenotypic plasticity in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Sci Rep.* **9:5137**, 1-11.

Tchoundjeu, Z., Asaah, E. K., Anegbeh, P. O., Degrande, A., Mbile, P., and Facheux, C. (2006). Putting participatory domestication into practice in West and Central Africa. *For. Trees Livelihoods* **16**, 53–69.

Teklehaymanot, Z. (2004). Exploiting the potential of indigenous agroforestry trees: Parkia biglobosa and Vitellaria paradoxa in Sub-Saharan Africa. *In* "New Vistas in Agroforestry. Advaces in Agrofrestry" (P. K. R. NAir, M. R. Rao and L. E. Buck, eds.). Springer, Dordrecht.

Temsch, E. M., and Greilhuber, J. (2001). Genome size in *Arachis duranensis*: a critical study. *Genome* **44**, 826-830.

Theerasilp, S., and Kurihara, Y. (1988). Complete purification and characterization of the taste-modifying protein, miraculin, from miracle fruit. *J. Biol. Chem.* **263**, 11536-11539.

Thrupp, L. A. (2000). Linking agricultural biodiversity and food security: the valuable role of agrobiodiversity for sustainable agriculture. *International affairs* **76**, 265-281.

Truong, H. H., Sato, T., Ishikawa, S., Minoshima, A., Nishimura, T., and Hirooka, Y. (2018). Three colletotrichum species rsponsible for anthracnose on *Synsepalum dulcificum* (Miracle fruit). *Int. J. Phytopathol.* **7**, 89-101.

Wang, H.-M., Chou, Y.-T., Hong, Z.-L., Chen, H.-A., Chang, Y.-C., Yang, W.-L., Chang, H.-C., Mai, C.-T., and Chen, C.-Y. (2011). Bioconstituents from stems of *Synsepalum dulcificum* Daniell (Sapotaceae) inhibit human melanoma proliferation, reduce mushroom tyrosinase activity and have antioxidant properties. *J. Taiwan Inst. Chem. Engin.* **42**, 204-211.

Wilken, M. K., and Satiroff, B. A. (2012). Pilot study of "miracle fruit" to improve food palatability for patients receiving chemotherapy. *Clin. J. Oncol. Nurs.* **16**, E173-E177.

Xingwei, C., Abdullah, T. L., Taheri, S., Abdullah, N. A. P., and Hassan, S. A. (2016). Flower ontogenesis and fruit development of *Synsepalum dulcificum*. *HortScience* **51**, 697-702.

You, Q., Yang, X., Peng, Z., Xu, L., and Ang, J. (2018). Development and application of a high throuptut genotyping tool for polyploidy crops: Single Nucleotide Polymorphism (SNP) array. *Front. Plant Sci.* **9**.

Zhang, Y.-J., Gan, R.-Y., Li, S., Zhou, Y., Li, A.-N., Xu, D.-P., and Li, H.-B. (2015). Antioxydant phytochemicals for the prevention and treatment of chronic diseases. *Molecules* **20**, 21138-21156.

Zheng, H.-C., Lu, Y., and Chen, D.-F. (2018). Anticomplement compounds from *Polygonum chinense*. *Bioorg. Med. Chem. Lett.* **28**, 1495-1500

CHAPTER 3 [1]

Comparative analysis of management practices and end-users' desired breeding traits in the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] across ecological zones and sociolinguistic groups in West Africa

Abstract

Understanding end-users preferred breeding traits and plant management practices is fundamental in defining sound breeding objectives and implementing a successful plant improvement programme. Since such knowledge is lacking for *Synsepalum dulcificum*, a worldwide promising orphan fruit tree species, we assessed the interrelationships among socio-demography, ecology, management practices, diversity and ranking of desired breeding traits by end-users of the species (farmers, final consumers, and processing companies) in West Africa. Semi-structured interviews, field-visits and focus groups were combined to interview a total of 300 farmers and final consumers belonging to six sociolinguistic groups sampled from three ecological zones of Benin and Ghana. One processing company in Ghana was also involved. Data collected included socio-demographic characteristics; crop management systems and practices; and preferences of farmers, final consumers and processing companies and ranking of breeding traits. Men (86.33% of respondents) were the main holders of *S. dulcificum* in the study area. The three most frequent management practices observed in the species included weeding, fertilization, and pruning, which were applied by 75.66%, 27.33% and 16.66% of respondents, respectively. The management intensity index varied significantly across ecological zones, sociolinguistic groups, and instruction level ($p < 0.001$) but was not affected by gender ($p > 0.05$). General multigroup similarity indices for both farmer-desired traits and final consumer-desired traits were high across ecological zones ($C_S^T \geq 0.83$). Nevertheless, respondents from the Guineo-Congolian (Benin) and the Deciduous forest (Ghana) zones expressed higher agreement in the ranking of desired-breeding traits. Preference for breeding traits was 60% similar among farmers, final consumers,

---

[1] Chapter published as: **Tchokponhoué, D. A.,** Achigan-Dako, E. G., N'Danikou, S., Nyadanu, D., Kahane, R., Odindo, A. O., & Sibiya, J. (2021). Comparative analysis of management practices and end-users' desired breeding traits in the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] across ecological zones and sociolinguistic groups in West Africa. Journal of Ethnobiology and Ethnomedicine, 17:41. **https://doi.org/10.1186/s13002-021-00467-8**

and processors. The key breeding traits desired by these end-users included in descending order of importance large fruit size, early fruiting, high fruit yielding (for farmers); large fruit size, high fruit miraculin content, fruit freshness (for final consumers); and high fruit miraculin content, large fruit size, high fruit edible ratio (for processing companies). This study revealed stronger variations in current management practices across ecological zones than across sociolinguistic groups. Top two traits to consider in breeding varieties of *S. dulcificum* to meet various end-users' expectations in West Africa include fruit size and fruit miraculin content. These results constitute a strong signal for a region-wide promotion of the resource.

3.1 Introduction

Synsepalum dulcificum is a slow-growing and long-living West African native fruit tree species belonging to the Sapotaceae family. The species can grow up to a height of 4 meters at maturity with a crown diameter ranging from 1 to 5 m (**Fig. 3.1A-B**) (Achigan-Dako et al., 2015). Its mature and ripe fruits are 10 mm wide and 20-25 mm long oblong to ovoid-shaped red berries (**Fig. 3.1 C-E**) (Ayensu, 1972; Lim, 2013) named “miracle berry”. The fruit is on average 10.5 mm-wide and 18.8 mm-long. The species is thought to combine both autogamy and allogamy, though the preponderant mating system is yet to be determined (Achigan-Dako et al., 2015). *Synsepalum dulcificum* is mainly propagated by seeds, which are recalcitrant (Tchokponhoué et al., 2018; Tchokponhoué et al., 2019a). Alternative propagation methods such as cuttings, and layering are not commonly used because of the difficult adventitious rooting in the species.

Figure 3.1. Traits variation in *Synsepalum dulcificum*. Tree size variation (A-B); Fruit size variation (C) and fruit shape (D: oblong fruit, E: Ovoid fruit) variation

Globally classified as a Least Concern species (link 1), *S. dulcificum* is threatened and endangered in West Africa (Adomou, 2005; Fandohan et al., 2017), its centre of origin and diversity, but is considered as one of the most valuable neglected berry crops because of its potential worldwide. The miracle berry is a unique natural source of miraculin; a sweetening glycoprotein that changes sour taste to sweet (Kurihara and Beidler, 1968) and serves as a natural sweetener in the food and beverage industries (Buckmire and Francis, 1978; Tafazoli et al., 2019). It is also used in diabetes and cancer treatments (Chen et al., 2006) as well as in cosmetics for the treatment of hair breakage (Diserud and Ødegaard, 2007). However, *S. dulcificum* remains poorly documented from many points of view; for instance, information relating to its management practices, production constraints and trade chain does not exist. Internationally, the miracle berry fetches a high price (over USD 2,500 / kg) (link 2, link 3), but income generation from the species by farmers at the grass-root level is yet to be properly documented. Similarly, knowledge on end-users' preferred traits, though fundamental to breeding of the species, is still lacking.

Participatory elicitation of breeding traits preference as a premix to implementing sound breeding programmes in plant species has been the centre of interest of many studies tackling a diversity of crop commodities including cereals (Mengistu et al., 2019; Sibiya et al., 2013), legumes and pulses (Banla et al., 2018; Coulibaly et al., 2020; Horn et al., 2015), vegetable, tuber and root crops (Agre et al., 2017; Bechoff et al., 2018; Muhinyuza et al., 2012; Placide et al., 2015). Noticeably, the core of these works was on annual and bi-annual crops; and except for the works of the World Agroforestry Centre in its effort to promote the domestication of indigenous fruit tree species [e.g. bush mango, *Irvingia gabonensis* (Aubry-Lecomte ex. O'Rorke) Baill.; safou tree, *Dacryodes edulis* (G. Don) HJ Lam; and vegetable tallow tree (*Allanblackia floribunda* Oliv.)] (Leakey et al., 2012), studies on participatory breeding in African indigenous fruit tree crops are limited (Ekué et al., 2010). In addition, these participatory studies mostly focused on "farmers" and did not integrate other user target groups such as final consumers or even processors, whose preferences are equally important and could somehow affect the definition of breeding objectives. This focus on farmers only, consequently constitutes a limitation that has also been highlighted by Hussein (2017) and Anja et al. (2017). These authors emphasized on the necessity to expand trait preference evaluation exercise to various user groups to decipher as much as possible diversity in preferences.

Trends in preference for breeding traits by farmers indicated the influence of numerous factors including ecological conditions and socio-demographic factors such as ethnicity, gender and landholding (Coulibaly et al., 2020; Placide et al., 2015; Weltzien et al., 2019), among others. In Nigeria for instance, while women emphasized the processing traits (e.g. easy to peel) in cassava (*Manihot esculenta* Crantz), men were more interested in agronomic traits (Teeken et al., 2018). The same authors also reported that traits such as "early maturity" and "high yielding" were more frequently sought-after by farmers in Southeast Nigeria compared with those in the Southwest, whereas the reverse was true for traits such as “cooking time”. A comparison of preferences for groundnut (*Arachis hypogaea* L.) revealed that while farmers in Ethiopia put a particular emphasis on traits such as early maturity, shell yield and drought resistance, those in Togo mainly targeted pod yield, pod size and oil yield (Abady et al., 2019; Banla et al., 2018). In contrast, a high similarity was observed between Benin and Togo farmers for trait preferences in groundnut (Banla et al., 2018; Loko et al., 2020). In ackee (*Blighia sapida* K.D. Koenig), a minor multipurpose fruit tree species, preference for fruit-traits varied among Benin sociolinguistic groups; the *Batoonuu* group considered exclusively the fruit size, the *Natemba* group preferred the aril colour, while the *Otamari* indicated the aril size as their trait of interest (Ekué et al., 2010). Instances of sociolinguistic group-specific preferred breeding traits were also reported in Kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal and Baudet] (Coulibaly et al., 2020), groundnut (Loko et al., 2020) and African locust bean (*Parkia biglobosa* Jacques) (Houndonougbo et al., 2020), among other crop species. It has also been shown that production constraints account for an important part in articulation and ranking of preferred breeding traits by farmers (Weltzien and Christinck, 2017). Consequently, the intensity of management practices defined as a set of actions directly or indirectly implemented by farmers to ensure the availability and the sustainability of plant production (Casas et al., 2008) is likely to shape farmers’ preference for breeding traits.

Building on the above-mentioned limitations, this study was undertaken to assess the interrelationships between socio-demographic factors, ecological conditions, management practices, diversity and preference for breeding traits of three different end-user groups of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.)] in West Africa that included the farmers, the final consumers and the processing companies. This study conducted in Ghana and Benin addressed the following questions: i) What are the ongoing management practices in *S. dulcificum* and how are they influenced by socio-demographic factors and agro-ecological conditions? ii) How diverse are end-users’ desired breeding traits in the species and to what extent are they influenced by socio-demographic variables and agro-ecological conditions? and iii) How similar are desired breeding traits among farmers, final consumers and industrial processors? The answers to these questions will pave the way for defining sound breeding objectives to meet multiple actors’ preferences for the species in West Africa.

3.2 Materials and methods

3.2.1 Study area

This study was conducted from April to December 2019 in Benin and Ghana (**Fig. 3.2**), two West African countries indicated as the centre of origin of *S. dulcificum* (Achigan-Dako et al., 2015). In Benin, *S. dulcificum* is confined to only one (Guineo-Congolian zone) out of the three ecological zones of the country, whereas in Ghana it is found in three (Evergreen forest, Deciduous forest, and Transitional zones) out of the six ecological zones of the country. Six regions (three per country) were chosen based on the presence of the species. While in Benin all the three regions (Zou, Mono, Couffo) were part of the Guineo-Congolian ecological zone, those of Ghana (Volta region, Eastern region and Western region) belonged to two distinct ecological zones: the Evergreen ecological zone represented by the Western region, and the Deciduous forest ecological zone represented by the Eastern and Volta regions.

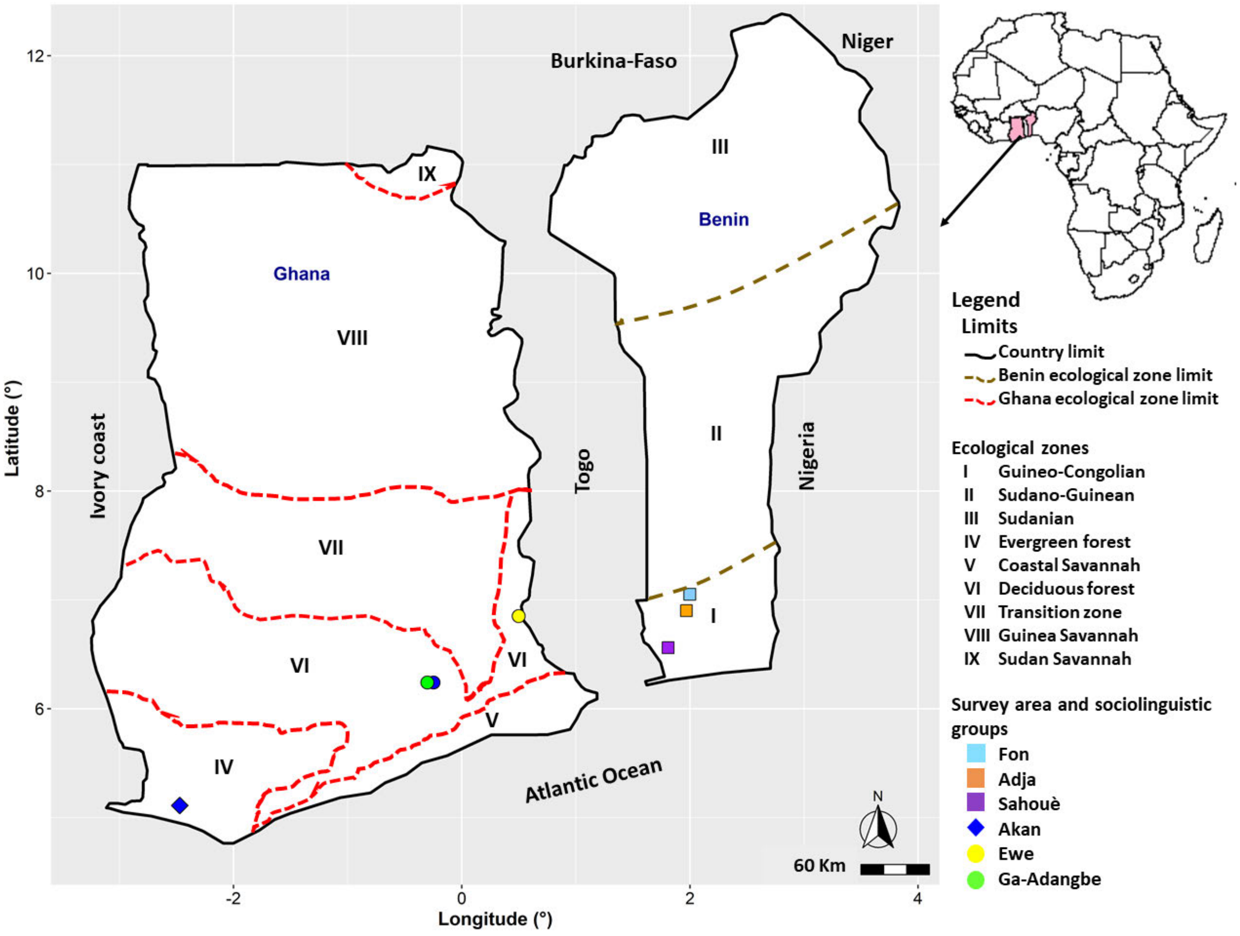


Figure 3.2. Map showing the study area

The Guineo-Congolian zone (GC) in Benin is characterized by a sub-equatorial climate with a bimodal rainfall pattern (Adomou et al., 2006). The rainfall ranges from 900 to 1,300 mm with an average of 1,200 mm while the annual temperature ranges between 25 and 39 °C with an average of 28 °C. The region experiences a relative humidity ranging between 69 and 97% (Adomou, 2005; Fandohan et al., 2011), and dominant soil types include ferralitic and ferruginous soils. The Deciduous forest ecological zone (DF) in Ghana also has a bimodal rainfall pattern with an annual rainfall varying from 1,200 to 1,600 mm with an average of 1,500 mm (Abbam et al., 2018). The mean annual temperature in this ecological region is 26.4°C and the soil is predominantly of eutric nitosol type (Obeng, 1971). In the Evergreen forest ecological zone (EF) of Ghana, the rainfall ranges from 800 to 2,800 mm with an average of 2,200 mm (Abbam et al., 2018). The mean annual temperature is 26.1°C and predominant soil types are acrisols and ferralsols (Abbam et al., 2018; Asravor et al., 2019; Obeng, 1971). In terms of moisture gradient, EF is a moist zone, DF a moderately moist, and GC a dry zone.

3.2.2 Interview sampling method

3.2.2.1 Individual respondents (farmers and final consumers)

Six sociolinguistic groups (three in Benin and three in Ghana) were selected for this study. In Benin, these groups included the *Fon*, *Adja* and *Sahouè*, while in Ghana, the three groups were *Akan*, *Ga-adangbe* and *Ewe*. The *Ewe* and *Ga-adangbe* groups inhabit the Deciduous forest ecological zone, while the *Akan* group was found in both Deciduous and Evergreen Forest ecological zones. The study focused on informants who had the species on their farms to gather information on management practices and preferences (Tongco, 2007). Consequently, we combined convenience sampling and snowball techniques, two commonly used non-probability sampling methods (N'Danikou et al., 2015; Sogbohossou et al., 2018), to select the respondents. Snowball was used to identify potential respondents who own and or cultivate the species. The final respondents then included those who i) had at least 5 years' experience in the management or cultivation of the species, ii) had their trees already in reproductive phase (bearing fruits), iii) were miracle berry-consumers and iv) had formally given their consent to participate in the study after the aim of the study was explained to them. These filters were imposed to ensure that respondents fully addressed the research questions.

3.2.2.2 Industrial processing company

To identify *S. dulcificum* processing companies in Benin and Ghana, we conducted online research, crossed-checked with the information provided by owners/farmers. They were asked to list any company they traded their product with or were aware of. Internet search was conducted in the Google search engine with the following research terms: 1) “Miracle:berry / Sweet:berry / *Synsepalum dulcificum*; Company; Ghana / Benin”, 2) “Miracle:berry / Sweet:berry / *Synsepalum dulcificum* + Company + Ghana / Benin”, 3) “Miracle:berry / Sweet:berry / *Synsepalum dulcificum*:Company:Ghana / Benin”. The identified processing companies were contacted afterwards, and those who gave their consent were included in the study.

3.2.3 Data collection

3.2.3.1 Individual respondents (farmers and final consumers)

Semi-structured interviews based on a questionnaire were combined with field visits to collect data on individual respondents. Interviews were conducted using the respondents preferred language, which was either Adja, Fon, or Sahouè in Benin, and Twi, Ewe, or Ga in Ghana. To facilitate the communication between interviewers and interviewees, each interviewer (where necessary) was accompanied by a well-trained local guide (who understood both the interviewee-spoken language and French/English) to facilitate the questions/answers translation. Direct interviews using either English or French were done, where necessary.

The data collected were related to i) the socio-demographic background of the informants, ii) *S. dulcificum* ownership (individual trees: miracle plant trees found isolated (not in a group) versus plantation: in which the miracle plant trees were found in a group (>100 trees in the context of this study) planted in a clear layout marked by a regular spacing among trees, iii) the production system in which the species has evolved [home garden: a production system near dwellings, and somehow well controlled by the owner for the target products versus on-farm: an open agricultural production space usually farther away from the dwellings and larger than a home garden (Niñez, 1987)], iv) the relative importance of *S. dulcificum* in the production system, v) cropping practices *of S. dulcificum*, vi) *S. dulcificum* production objectives, vii) *S. dulcificum* trade chain and profitability, viii) the informant’s awareness of grown *S. dulcificum* morphotypes and varieties, ix) the farmers’ desired-traits for an improved variety of *S. dulcificum* and x) the final consumer-desired breeding traits for an improved variety of *S. dulcificum*.

To characterize the production system in which *S. dulcificum* is involved and the relative importance of this latter in the whole system, *S. dulcificum*-based production systems were visited, and the top five most important species based on all possible functional attributes held (food, medicine, market value etc.) were elicited with the farmer. Then, the pairwise comparison technique was used to score each of the five species mentioned by the farmer on a scale ranging from 0 (the least important crop) to 5 (the most important one). The same listing of the five most important attributes and their subsequent scoring in a pairwise comparison scheme was also applied to i) the farmer-desired traits for an improved variety of *S. dulcificum*, and ii) the final consumer-desired breeding trait for an improved variety. Farmer's income from the miracle berry was estimated from the last harvest season of 2018. The estimation covered only the year 2018 because the study was conducted in different months of 2019 and including 2019 selling events would introduce a bias since a group of respondents were interviewed earlier in the year, and another later. Besides, more than 1-year recalling back of revenue could have been difficult for most of the respondents.

The taxonomic identification of the species, indicated as important by the respondent, was first done *in-situ* by the first author using available flora resources including the analytical flora of Benin (Akoègninou et al., 2006), the trees, shrubs and lianas of West African dry zones manual (Arbonnier, 2004), and the plant list (http://www.theplantlist.org/) online resource (The Plant List, 2013) . The identifications were afterwards confirmed through a formal taxonomic identification at the National Herbarium of Benin.

3.2.3.2 Industrial processing companies

A focus group interview was held with four members holding different positions in one processing company to collect information related to the company genesis, company production objectives and desired breeding traits for a variety suitable to each of the company's production line as well as the pairwise ranking of these desired traits.

3.2.4 Data analysis

The R environment version 3.6.2. (R Core Team, 2019) was used for data analyses.

Descriptive statistics (mean, standard error, range and frequency) were computed on the socio-demographic characteristics of respondents using the functions *descript ()* and *crosstab ()* of the package 'misty' (Yanagida, 2020). The major plant species in the farmer's production system, as well as the relative importance of *S. dulcificum* in the system, were assessed using descriptive statistics computed on scores obtained from the pairwise ranking. A χ^2 test or Fisher exact test (to account for observation count < 5) was used to test dependence between *S. dulcificum* holding systems and categorical socio-demographic factors (gender, ethnicity, instruction levels) on one hand, and ecological zone on the other. To characterize the importance of management practices in the species, we followed the framework developed by Blancas et al. (2013) to propose a new index, the boosted management intensity index (BMI), which is adapted from the management intensity index (MI) developed by Sogbohossou et al. (2018). This new index is based on ten key indicators of management practices (**Table 3.1**) and incorporates weighing coefficients to reflect better the situation in perennial species where certain practices are to be repeated over time. The coefficients are proposed to consider elapsed time since some management practices (e.g., pruning, weeding) were applied for the last time. For a specific respondent *j* the BMI is computed following the **Equation (3.1)** as:

$$\mathrm{BMI_j} = \sum_{i=1}^{m} \mathrm{IL_i} + \sum_{i=1}^{m} \mathrm{C_i ML_i} + \sum_{i=1}^{m} \mathrm{C_i UI_i} + \mathrm{PSD} + \mathrm{RH} + \sum_{i=1}^{m} \mathrm{UT_i} + \mathrm{TH} + \mathrm{TI} + \mathrm{LF} + \mathrm{MF} \quad \textbf{(Eqt. 3.1)}$$

where BMI_j is the index computed for the respondent *j*, *m* is the total number of management practices constituting a specific management variable, C_i is the weighing coefficient associated to a management practice *i*, IL_i is the score for the practice *i* of the variable “establishment labour”, ML_i is the score for the practice *i* of the variable “management labour”, UI_i is the score related to the use of input, *PSD* is the score related to the distance to the cultivation/production site, *RH* is the score related to the reaction to an unauthorized harvest, *UT* is the score related to the use of tools in the species management, *TH* is the score related to the type of harvest, *TI* is the score related to the time invested in the species management, *LF* is the score related to the labour force invested in the species and *MF* is the score related to the management form.

The variation of BMI following ecological zone and socio-demographic factors was analysed using a generalized linear model with a Poisson or quasi-Poisson (to account for overdispersion) error structure. The effect of the holding system on income generation was analysed using a Wilcoxon test due to normality assumption violation. We used the *prop.test*-based binomial test to assess the difference in fruit production objective following ecological zone and sociolinguistic group.

Table 3.1. Variables used in the computation of the boosted management intensity index (BMI)

Variables	State of variables and codified scores	Weighing variable		Weighing coefficient
Installation labour (IL)	Land clearing (1); pegging and lining (1); holing (1); base manuring (1)	None	-	1
Management labour (ML)	Weeding (3); pruning (3); irrigation (3); fertilization (3)	Time since last application	≤ 1 year	1
Use of inputs (UI)	Do not use agrochemicals (0); use agrochemicals (3)		2 - 10 years	2/3
			> 10years	1/3
Distance to cultivation site (PSD)	≤ 100 m (1); up to 1 km (2); up to 5 km (3); more than 5 km (4)			
Reaction to unauthorized harvest (RH)	No reaction (0); yes, admonition applies (1); yes, monetary sanction (2); yes, complain to authority (3)	None	-	1
Use of tool (UT)	Manual (1); hook (2); knife (3); machete (1)	None	-	1
Type of harvest (TH)	Opportunistic (1); planned (2)	None	-	1
Time invested (TI)	Min (1); hours (2); days (3)	None	-	1
Labour force invested (LF)	No staff hiring (1); staff hiring (2)	None	-	1
Management form (MF)	Collection from the wild (1); protection on farm (2); protection in home garden (3); seedling transplantation (4); seed sowing (5)	None	-	1

-: Not applicable

The similarity of farmers' and final consumers' desired breeding traits among ecological zones, gender, and sociolinguistic groups on one hand, and the similarity of traits preference among farmers, final consumers and processors on the other were analysed using the general multi-group similarity index ($\mathrm{C}_{\mathrm{S}}^{\mathrm{T}}$) computed following Diserud and Ødegaard (2007) as:

$$\mathrm{C}_{\mathrm{S}}^{\mathrm{T}} = \frac{T}{T-1}\left(\frac{\sum_{i<j} a_{ij} - \sum_{i<j<k} a_{ijk} + \sum_{i<j<k<l} a_{ijkl} - \sum_{i<j<k<l<m} a_{ijklm} + \cdots}{\sum_i a_i}\right) \qquad \textbf{(Eqt. 3.2)}$$

where *T* is the total number of groups for which the index is computed, a_i is the number of traits listed by the group A_i, i =1, 2, 3,…,T; a_{ij} is the number of traits shared by groups A_i and A_j; a_{ijk} is the number of traits shared by A_i, A_j, A_k etc.

$\mathrm{C}_{\mathrm{S}}^{\mathrm{T}}$ value ranges from 0 (no similarity in traits preference) to 1 (total similarity in traits preference). An illustration of the calculation of this index is provided in **Box 3.1.**

Box 3-1. Calculation of the general similarity index for more than two groups

Let case 1 be a study comparing traits preference in a species "A" by farmers from three different socio-cultural groups (1, 2 and 3) and whose preferred traits are elicited as follow:

A1: [a, b, c, d, e]; **A2**: [a, d, e, f, h] and **A3**: [b, g, h, j].

The objective is to calculate the general multi-group similarity index of preferred traits among these three socio-cultural groups.

For this example, T = 3, $a_1 = 5$, $a_2 = 5$ and $a_3 = 4$, $a_{12} = 3$, $a_{13} = 1$, $a_{23} = 1$, $a_{123} = 0$

$$\mathbf{C}_{\mathbf{S}}^{\mathbf{3}} = \frac{T}{T-1}\left(\frac{\sum_{i<j} a_{ij} - \sum_{i<j<k} a_{ijk}}{\sum_i a_i}\right) = \frac{3}{3-1}\left(\frac{[a_{12}+a_{13}+a_{23}]-a_{123}}{a_1+a_2+a_3}\right) = \frac{3}{3-1}\left(\frac{[3+1+1]-0}{5+5+4}\right)$$

$$\mathbf{C}_{\mathbf{S}}^{\mathbf{3}} = \frac{3}{2}\left(\frac{5}{14}\right)$$

$$\mathbf{C}_{\mathbf{S}}^{\mathbf{3}} = \frac{15}{28}$$

The similarity index of preferred traits among these three socio-cultural groups is 0.53, which reflects a moderate preference similarity.

The agreement in the ranking of farmer-desired breeding traits across ecological zones and across sociolinguistic groups was assessed with the Kendall-W coefficient of concordance computed using the function *Kendall ()* of the 'irr' package (Gamer et al., 2019). The same coefficient was also computed in the case of final consumer-desired traits and to analyse the concordance in the ranking pattern of desired traits by the three user groups (farmers, final consumers, and processors) in this study.

3.3 Results

3.3.1 Socio-demographic characteristics of respondents

A total of 300 individual respondents were interviewed in this study. The distribution of the sampling size across ecological zones and sociolinguistic groups as well as the socio-demographic background of these respondents are detailed in **Appendix 3.1**, Additional file 1. The proportion of women (13.67%) involved in *S. dulcificum* management was significantly lower than that of men (86.33%) ($\chi^2 = 313.9$, df =1, $p < 0.0001$). The goodness of fit test indicated that this ratio 14:86 of women/men holding the species was statistically constant across sociolinguistic groups ($\chi^2 \leq 0.96$, df = 1, $p > 0.05$) and ecological zones ($\chi^2 \leq 0.39$, df = 1, $p > 0.05$). The holders of *S. dulcificum* were mainly autochthons ($\chi^2 = 420.01$, df = 1, $p < 0.0001$) with 67% of them educated; and approximately 40% of them reached the level of secondary school ($\chi^2 = 189.16$, df = 4, $p < 0.0001$). Almost all respondents were married. The youngest respondent was 25 years old, an *Adja*, whereas the oldest was 102 years old from the *Ga-adangbe* sociolinguistic group. On average, the respondents were 55.47 ± 0.91 years old. Household size ranged from 2 to 30 with an average of 7.25 ± 0.24 members per household and differed significantly across ecological zones ($p < 0.0001$) and sociolinguistic groups ($p < 0.0001$).

3.3.2 Profile of miracle berry processing companies

Two companies, based in Ghana, namely “Sweet Life Group Ghana Limited” and “Miracle fruit processing Ghana Limited” invested in miracle berry processing in the study area. Miracle Fruit Processing Ghana Ltd, formerly known as MB Group Ghana and created in 2015, is well known by farmers (100% of plantation owners) in Ghana. On the contrary, no farmer mentioned the Sweet Life Group Ghana Ltd company, which is still operating. This latter company declined to participate in the study; so, only information from one company was captured in this study.

3.3.3 *Synsepalum dulcificum* holding systems

Across the three ecological zones, the miracle plant was found in two production systems, namely home gardens, and on-farm (**Fig. 3.3A-D**). While in the Guineo-Congolian zone the species was exclusively observed as individual trees (**Fig. 3.3A**), in the Deciduous and

Evergreen forest zones it was present as either individual trees (**Fig. 3.3B**) and/or a commercial plantation (**Fig.3.3C-D**).

Figure 3.3. *Synsepalum dulcificum* in various production systems. Individual tree in home garden (**A**); Individual tree on farm (**B**); Plantation in home garden (**C**) Plantation on farm (**D)**

Owners of individual trees of *S. dulcificum* represented 97.66% of respondents. The total number of individual trees they owned varied from 1 to 36 with an average of 2.79 ± 0.21 trees. Gender did not affect the number of individual trees of *S. dulcificum* owned ($p = 0.427$). In contrast, this number differed significantly among sociolinguistic groups ($p = 0.006$), with the *Ga-adangbe* (4.57 ± 0.64 trees) and the *Ewe* (3.30 ± 0.54 trees) owning two-fold more trees than the *Fon* (1.60 ± 0.12 trees). In parallel, respondents from the Deciduous forest zone (3.41 ± 0.32 tree) owned more trees than those from the Guineo-Congolian zone (1.84 ± 0.32 trees) ($p < 0.001$).

Within the Deciduous forest and Evergreen forest ecological zones, the main holding system depended on the sociolinguistic group affiliation (Fisher exact, $p = 0.001$), with the *Akan* being the only group owning plantations of *S. dulcificum*. Likewise, there was a significant association between instruction level and miracle plant holding system (Fisher exact, p =

0.002), and only educated respondents held plantations of *S. dulcificum*. Though only men owned *S. dulcificum* plantations, the test of independence indicated insignificant association between gender and holding system (Fisher exact, $p = 0.59$). We recorded a total of 17 plantations in this study, out of which 12 were located on farm. The mapping of the distribution of these 12 on-farm plantations is presented in **Appendix 3.2**. The total number of plantations held per owner varied from 1 to 4 and with an average of 2.47 ± 0.36 plantations per owner. The size of a *S. dulcificum* plantation ranged from 0.16 to 89.03 ha in the study area, whereas the per-farmer cumulated acreage of *S. dulcificum* plantations ranged from 0.60 to 109.26 ha, out of which only 0–8.09 ha (representing approximately 8% increase relatively to the existing plantation acreage before 2015) were installed following the advent of the miracle fruit processing company, which indicated that most of the plantations were established before the company arrival. This illustrated that the company arrival did not boost or has not impacted yet the species plantation expansion. All the plantations were established using seeds.

3.3.4 *Synsepalum dulcificum* relative importance in production systems

A total of 62 plant species from 32 families (**Appendix 3.3**) were recorded in the respondents' production systems across the three ecological zones. Eighteen and 24 out of these 62 species had citation frequencies higher than 5% in home garden and on-farm production systems, respectively. The most important species for respondents (**Table 3.2**) were significantly dominated by perennials in the home-garden system ($\chi^2 = 9.00$, df = 1, $p = 0.002$), whereas annual species were as represented as perennial species in the on-farm production system ($\chi^2 = 2.08$, df = 1, $p = 0.14$). In general, *S. dulcificum* appeared poorly ranked in respondents' production system, though it was perceived as more important in the home gardens than on-farm. Nevertheless, a disaggregated analysis indicated that *S. dulcificum* importance was differentially perceived across ecological zones and sociolinguistic groups. In the Evergreen forest ecological zone, *S. dulcificum* ranked 2nd and 3rd in home garden and on-farm production systems, respectively, whereas in Deciduous forest and Guineo-Congolian zones it was not in the top five species (**Table 3.3**). Similarly, the species was more integrated into *Akan*'s and *Sahouè*'s production systems than in any other sociolinguistic group (see Additional file 2). In contrast, both men and women ranked *S. dulcificum* out of the top 5 most important species (**Appendix 3.4**). Most important species included cocoa (*Theobroma cacao* L.), maize (*Zea mays* L.), oil palm (*Elaeis guineensis* Jacq.) and cassava (*Manihot esculenta* Crantz).

Table 3.2. Diversity and relative importance of frequently involved species in *Synsepalum dulcificum*-based production system

Species*	On-farm		Species*	Home-garden	
	Importance score	Rank		Importance score	Rank
***Theobroma cacao* L.**	4.06	1st	***Theobroma cacao* L.**	4.39	1st
Zea mays L.	4.00	2nd	***Elaeis guineensis* Jacq.**	3.61	2nd
***Elaeis guineensis* Jacq.**	3.46	3rd	***Musa parasidica* L.**	3.55	3rd
Manihot esculenta Crantz.	3.30	4th	***Coffea canephora* L.**	3.43	4th
***Cola nitida* (Vent.) Schott & Endl.**	3.09	5th	*Manihot esculenta* Crantz.	3.27	5th
Solanum aethiopicum L.	3.00	6th	***Musa sapientum* L.**	3.12	6th
Musa sapientum L.	2.55	7th	*Colocasia esculenta* (L.) Schott	3.00	7th
***Citrus sinensis* (L.) Osbeck**	2.40	8th	*Discorea alata* L.	3.00	8th
Capsicum sp.	2.31	9th	***Chrysophyllum albidum* G.Don**	2.79	9th
Vigna unguiculata (L.) Walp	2.29	10th	***Cocos nucifera* L.**	2.78	10th
***Tectona grandis* L.f.**	2.23	11th	***Citrus lemon* L.**	2.66	11th
***Irvingia gabonensis* (Aubry-Lecomte ex O'Rorke)**	2.19	12th	***Citrus sinensis* (L.) Osbeck**	2.66	12th
Dioscorea alata L.	2.18	13th	***Moringa oleifera* (Gaetn.) Dunal**	2.42	13th
***Musa parasidica* L.**	2.12	14th	***Annona muricata* L.**	2.38	14th
***Acacia auriculiformis* Benth.**	2.00	15th	***Mangifera indica* L.**	2.25	15th
***Annona muricata* L.**	2.00	15th	***Persea americana* Mill.**	2.22	16th
***Chrysophyllum albidum* G.Don**	2.00	15th	***Synsepalum dulcificum* (Schumach &Thonn.) Daniell**	1.70	17th
***Xylopia aethiopica* (Dunal) A. Rich.**	2.00	15th	*Ananas comosus* (L.) Merr.	1.28	18th
Colocasia esculenta (L.) Schott	1.94	19th	-	-	-
Persea americana Mill.	1.87	20th	-	-	-
***Synsepalum dulcificum* (Schumach &Thonn.) Daniell**	1.83	21st	-	-	-

Species*	On-farm		Species*	Home-garden	
	Importance score	Rank		Importance score	Rank
Lycopersicum esculentum L.	1.80	22nd	-	-	-
Arachis hypogaea L.	1.64	23rd	-	-	-
***Cocos nucifera* L.**	1.45	24th	-	-	-

-: Not applicable

Table 3.3. Per ecological zone- disaggregated importance of species involved in *Synsepalum dulcificum*-based production systems

Production system	Guineo-Congolian zone			Deciduous forest zone			Evergreen forest zone		
	Species	IS*	Rank	Species	IS*	Rank	Species	IS*	Rank
Farm	*Zea mays* L.	4.40	1st	*Theobroma cacao* L.	4.10	1st	*Manihot esculenta* Crantz.	5.00	1st
	Elaeis guineensis Jacq.	3.73	2nd	*Manihot esculenta* Crantz.	3.58	2nd	*Hevea brasiliensis* (Willd. Ex A. Juss.) Mull.Arg.	3.50	2nd
	Manihot esculenta Crantz.	2.96	3rd	*Zea mays* L.	3.42	3rd	*Synsepalum dulcificum* (Schumach &Thonn.) Daniell	3.40	3rd
	Capsicum sp.	2.42	4th	*Cola nitida* (Vent.) Schott & Endl.	3.10	4th	*Psidium guajava* L.	3.00	4th
	Citrus sinensis L.	2.41	5th	*Solanum aethiopicum* L.	3.00	5th	*Theobroma cacao* L.	3.00	5th
Home-garden	*Musa sapientum* L.	3.38	1st	*Theobroma cacao* L.	4.41	1st	*Musa sapientum* L.	5.00	1st
	Elaeis guineensis Jacq.	3.37	2nd	*Elaeis guineensis* Jacq.	3.73	2nd	*Synsepalum dulcificum* (Schumach &Thonn.) Daniell	4.70	2nd
	Citrus lemon L.	3.30	3rd	*Musa parasidica* L.	3.55	3rd	*Psidium guajava* L.	4.00	3rd
	Citrus sinensis (L.) Osbeck	3.12	4th	*Manihot esculenta* Crantz.	3.50	4th	*Musa parasidica* L.	3.50	4th
	Cocos nucifera L.	2.80	5th	*Coffea canephora* L.	3.43	5th	*Citrus sinensis* (L.) Osbeck	3.00	5th

***IS:** importance score.

3.3.5 *Synsepalum dulcificum* management practices

3.3.5.1 Types and sources of planting materials

In the study area, respondents mainly used two types of planting material for the establishment of *S. dulcificum*: seeds, seedlings, or a combination of both. Overall, respondents used seedlings (55.21%) more frequently than seeds (44.79%) (χ^2 = 4.60, df = 1, *p = 0.03*). While the choice of planting material depended neither on gender (χ^2 = 1.20, df = 2, *p = 0.55*) nor on sociolinguistic group affiliation (χ^2 = 15.37, df = 10, *p = 0.14*), we observed that inhabiting the Evergreen forest ecological zone strongly conditioned the use of seedlings as the only planting material (χ^2 = 14.95, df = 4, *p = 0.004).* Likewise, seeds more frequently used in home-garden, and seedlings for on-farm production (χ^2 = 2.60, df = 1, *p = 0.08*).

These planting materials were obtained by farmers from four different sources including research centres, markets, immediate neighbour (sourcing neighbour located at < 5-km radius far from the respondent) and far-off neighbour (sourcing neighbour located at ≥ 5-km radius far from the respondent). Immediate neighbour was by far the commonest means of sourcing *S. dulcificum* planting material (χ^2 = 174.29, df = 3, *p < 0.0001*). Seeds obtained from immediate neighbour or far-off neighbour were systematically sown without any clear selection criteria, while seedlings, which were not produced *per se*, but rather just uprooted beneath mother trees and transplanted were selected based on their vigour. The relative importance of each source disaggregated by ecological zone and sociolinguistic group is presented in **Fig. 3.4**. A research centre as source of planting material was only reported by respondents in Deciduous and Evergreen forest ecological zones, while purchase from the market only occurred in the Guineo-Congolian ecological zone. There, only the *Adja* and *Fon* respondents bought planting materials from the market, whereas in Ghana, only the *Akan* and *Ga-adangbe* accessed planting materials from a research centre. The *Ewe* and *Sahouè* respondents only accessed both seeds and seedlings from neighbours as gifts.

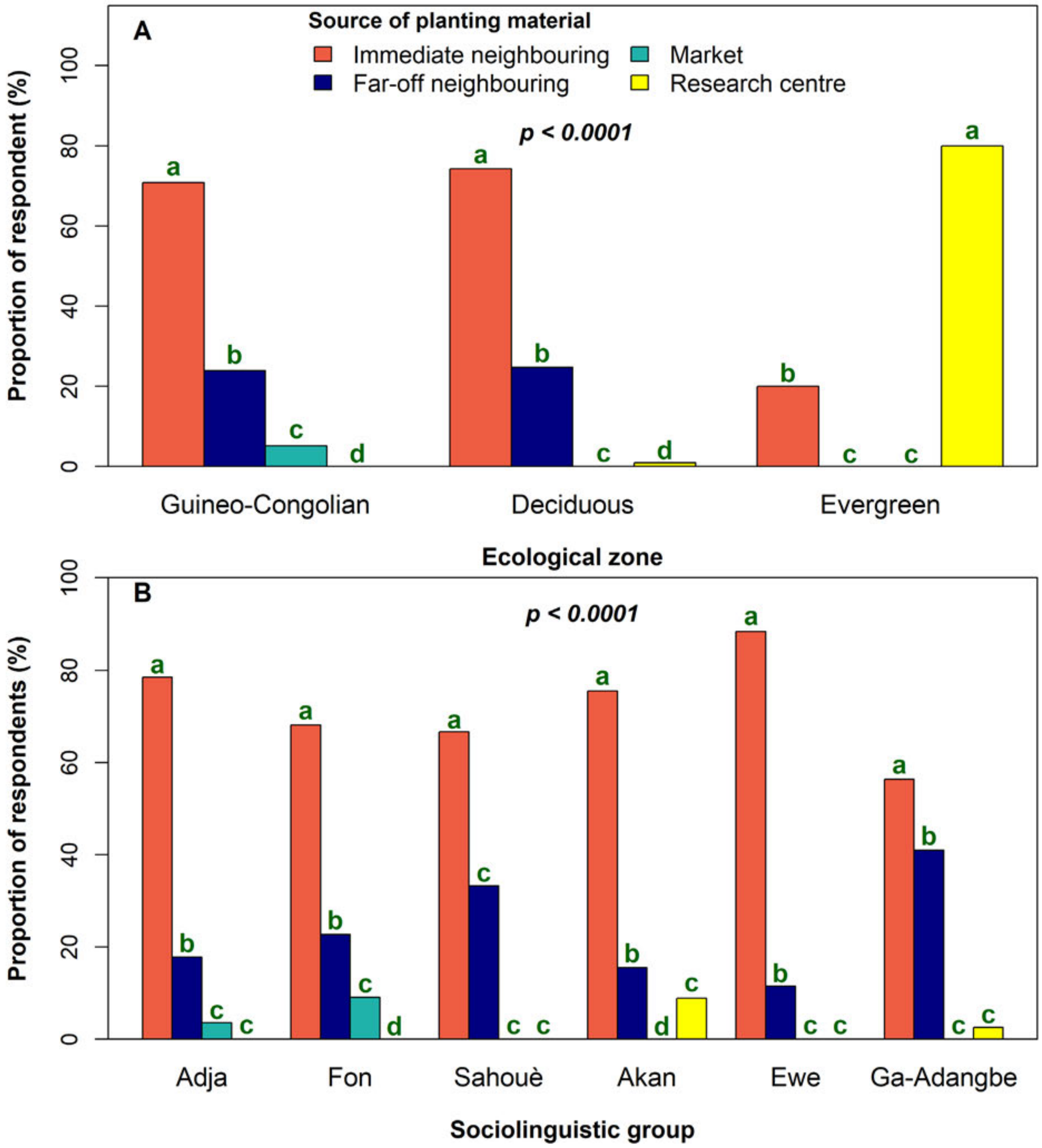


Figure 3.4. Relative importance of *Synsepalum dulcificum* planting materials' source across ecological zones (**A**; Guineo-Congolian: Guineo-Congolian zone; Deciduous: Deciduous forest and Evergreen: Evergreen forest) and sociolinguistic groups (**B**).

3.3.5.2 Variation of management intensity of *Synsepalum dulcificum*

The ten variables summarizing the set of management practices in *S. dulcificum* are indicated in **Table 1**. The combination of all these variables indicated that the boosted management intensity index (BMI) of *S. dulcificum* was on average 14.61 ± 0.24 and ranged from 7.00 to 34.00. This management intensity was significantly affected by the ecological zone and some socio-demographic factors. Management intensity index was two-fold higher in the Evergreen forest ecological zone (BMI = 29.0 ± 1.70) than in Deciduous forest (BMI = 14.54 ± 0.30) and Guineo-Congolian (BMI = 14.19 ± 0.32) zones (**Fig. 3.5A**), while the species management by women (BMI = 14.46 ± 0.57) was as intense as by men (BMI = 14.63 ± 0.26) (**Fig. 3.5B**). Regarding the influence of sociolinguistic groups, the *Akan* managed more intensely the species (BMI = 16.88 ± 0.86) than any other sociolinguistic groups, whereas the lowest management intensity index was obtained with the *Sahouè* sociolinguistic group (BMI = 13.26 ± 0.59) (**Fig. 3.5C**). Instruction level also exerted a significant effect on the management intensity (**Fig. 3.5D**) though increased instruction level did not necessarily reflect in higher management intensity. Plantations (BMI = 30.14 ± 1.14) were also better managed compared with individual trees (BMI = 14.24 ± 0.20) ($p < 0.0001$). There was no correlation between respondents' age and management intensity (r = 0.02, $p = 0.69$) on one hand, and between experience in the species management and management intensity (r = –0.03, $p = 0.57$) on the other.

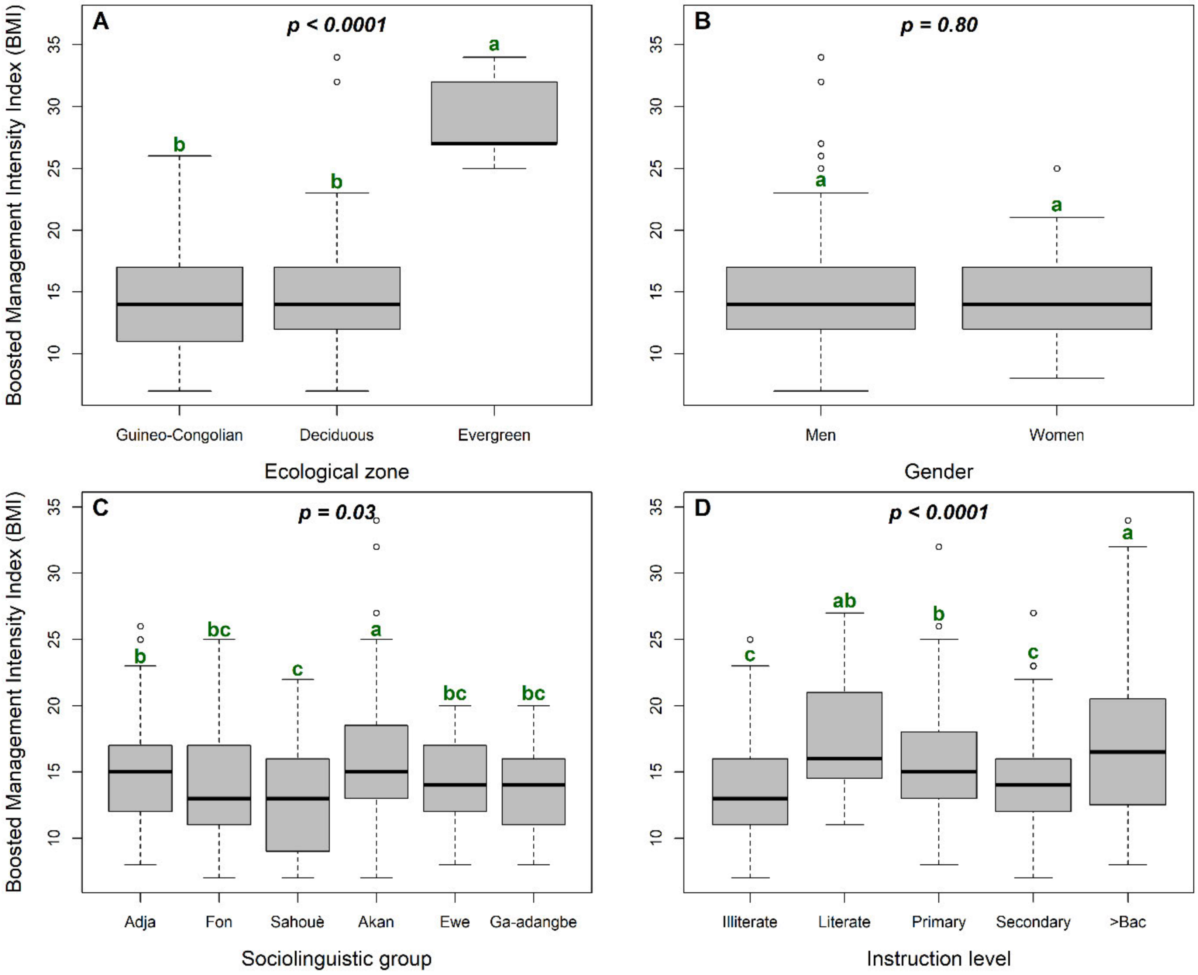


Figure 3.5. Variation of management intensity index in *S. dulcificum* following ecological zones (**A**), gender (**B**), sociolinguistic group (**C**) and instruction level (**D**).

3.3.5.3 Farmers' awareness of morphotypes in *Synsepalum dulcificum*

For 100% of respondents, there is no variety of *S. dulcificum*. However, 6.33% of respondents indicated the existence of different morphotypes that they essentially distinguished through the difference in fruit size and fruit exocarp colour. Where morphotypes were reported, 100% of respondents differentiated big fruit sized from small fruit sized morphotypes, whereas only one out of the 300 respondents (from the Deciduous forest ecological zone of Ghana, and the *Ga-adangbe* sociolinguistic group) indicated the existence of a yellow morphotype in addition to the ordinary known red morphotype.

3.3.6 Farmer-desired breeding traits in *Synsepalum dulcificum*

Among the respondents, 94.60% expressed the desire to have a variety /an improved variety of *S. dulcificum*. Those not desiring any improved variety indicated that an improved variety will not be as "powerful" as the actual landraces, specifically referring to a reduction in the species phytochemical compounds.

In total, 18 breeding traits (**Table 3.4**) were mentioned by farmers as of interest in any new/improved variety. Overall, the top five farmer-desired breeding traits were: big fruit size > early fruiting > high fruit yielding > dwarf tree > high fruit miraculin content. We recorded one women-specific preferred trait, that was, "uniformity of fruit ripening", and four men-specific preferred traits, which included: "extended on-tree fruit shelf life", "high fruit edible ratio", "low fruit shedding", and "long biological productivity period".

The general multi-group similarity indices for farmer-desired breeding traits across the three ecological zones and six sociolinguistic groups were $C_S^3 = 0.84$ and $C_S^6 = 0.94$, respectively. Regarding the ecological zones, the Deciduous forest and Guineo-Congolian ecological zones exhibited a higher similarity for farmer-desired breeding traits ($C_S^2 = 0.94$) than any other pair of ecological zones ($C_S^2 = 0.56$ for the pair Guineo-Congolian and Evergreen forest zones; $C_S^2 = 0.61$ for the pair Deciduous forest and Evergreen forest zones). Likewise, the highest pairwise similarity index between sociolinguistics groups was obtained for *Akan* - *Ewe* ($C_S^2 = 0.96$) and the lowest one for *Ga-adangbe –Sahouè* $\left(C_S^2 = 0.74\right)$ (**Table 3.5**).

Respondents in Guineo-Congolian and Deciduous forest ecological zones exhibited a perfect agreement in the ranking pattern of their desired traits (Kendall-W = 1.00). For respondents of these two ecological zones, big fruit size, early fruiting and high fruit yielding were the top three desired traits for an improved variety of *S. dulcificum,* whereas their counterparts of Evergreen forest ecological zone, instead targeted resistance to biotic stresses, early fruiting, and high fruit yielding as traits of interest (**Table 3.6**).

Table 3.4. Similarity of farmers-desired breeding traits for *Synsepalum dulcificum* across sociolinguistic groups

Traits	Cumulated score	Rank	Reasons (relative citation frequency)
Big fruit size	202	1st	For more clients and an easy selling (0.57); a better use efficiency due to higher pulp mass (0.19); for more weight for more income (0.19); easy harvesting (0.02); for performant offspring (0.03)
Early fruiting	191	2nd	Early benefit (consumption, sales) from the species (1)
High fruit yielding	166	3rd	For more income (0.82); for a diversity of utilization (0.18)
Dwarf tree	83	4th	Easy harvesting (0.92); easier integration in agroforestry systems (0.2)
High fruit miraculin content	69	5th	Attract more client (0.63); landmark characteristics necessary to enjoy the fruit (0.37)
Fruit freshness	53	6th	To attract the buyer/consumer (0.69); For a safety of consumption (0.31)
Large tree crown	31	7th	High productivity (0.56); fruit less exposed to stealing (0.1); provide shelter to rest (0.23); easy management operation (0.06); escaping children disturbance (0.05)
Long fruit shelf life	22	8th	Increased storability for a staggered utilization/commercialization (1)
Biotic stress resistance	18	9th	Healthy and productive tree (0.57); longevity of the tree (0.25); to have pest-free and clean fruits (0.18)
Fast growing	15	10th	Reduced waiting time to fruiting (1)
High fruiting frequency	12	11th	For more revenue (0.55); for more fruits at the end of the year (0.36); for a permanent availability of the fruits (0.09)
Abiotic stress tolerance	7	12th	Maintenance of fruiting performance (0.68); better water use efficiency (0.16); high survival rate (0.16),
Big stem diameter	7	13th	Higher resistance to abiotic stress (0.57); high production (0.28); withstanding to children disturbance (0.15)
Extended on-tree fruit shelf life	7	14th	For a reduced on-farm post ripening loss (1)
Long lasting production	5	15th	For a long-lasting benefit from it (1)
Low shedding	4	16th	For a high production (1)
High fruit edible ratio	2	17th	For more juice from the fruit (1)
Uniformity of ripening	1	18th	For a one-once grouped harvest (1)

Table 3.5. Similarity of farmers-desired breeding traits for *Synsepalum dulcificum* across sociolinguistic groups

	Adja	Fon	Sahouè	Akan	Ewe	Ga-adangbe
Adja		0.97	0.93	0.60	1.00	0.90
Fon	0.84		0.97	0.65	0.97	0.91
Sahouè	0.82	0.74		0.70	0.90	0.70
Akan	0.86	0.78	0.77		0.75	0.70
Ewe	0.89	0.81	0.86	0.96		0.90
Ga-adangbe	0.84	0.75	0.74	0.85	0.88	

Table 3.6. Ecological zone-based variation of the top five farmers-desired breeding traits for an improved variety of *Synsepalum dulcificum*

Traits	Guineo-Congolian zone		Deciduous forest zone		Evergreen forest zone	
	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank
Big fruit size	78	1st	121	1st	2	4th
Early fruiting	73	2nd	113	2nd	5	1st
High fruit yielding	51	3rd	111	3rd	4	3rd
High fruit miraculin content	26	-	63	4 h	-	-
Dwarf tree	26	4th	55	5 h	2	4th
Biotic stress resistance		-	-	-	5	1st
Large tree crown	17	5th	-	-	-	-
Overall ranking agreement	Kendall-W = 0.6					

Traits such as big fruit size, early fruiting, high fruit yielding and dwarf tree were consistently ranked among the first four most important farmer-desired breeding traits (**Table 3.7**) by the sociolinguistic groups of Benin i.e., *Adja*, *Fon* and *Sahouè*. These sociolinguistic groups also exhibited a very high agreement in their ranking pattern (Kendall-W > 0.90, **Table 3.5**). In contrast, a lower agreement was observed among Ghana sociolinguistic groups, the maximum concordance coefficient of desired traits was 0.90 (**Table 3.5**). For instance, for *Ewe* and *Ga-adangbe,* big fruit size was the most important trait desired for an improved variety, while their counterpart A*kan* rather targeted a high fruit yielding variety (**Table 3.7**).

3.3.7 Local consumers' preferences for breeding traits in *Synsepalum dulcificum*

The 300-interviewed miracle berry final consumers in Benin and Ghana together listed a total of nine desired traits for an improved variety adapted for consumption. Out of these traits only five, which also represented the top five desired traits have a citation frequency higher than 5% (**Table 3.8**). Overall, desired traits were 94.56% similar across the three ecological zones and 83.60% identical across the six sociolinguistic groups. The highest pairwise similarity index of final consumers' desired breeding traits was obtained between respondents of Deciduous forest and Guineo-Congolian zones ($\mathrm{C}_{\mathrm{S}}^{2} = 0.87$), while the highest agreement in traits ranking pattern was observed between respondents from Deciduous forest and Evergreen forest ecological zones (Kendall-W = 0.90). The lowest agreement coefficient (Kendall-W = 0.40) of these traits ranking was also observed between final consumers of Guineo-Congolian and Evergreen forest ecological zones. Regarding the sociolinguistic groups, *Adja* and *Sahouè* final consumers preferred similar traits, while *Adja* and *Ga-adangbe* final consumers, on one hand, and *Sahouè* and *Ga-adangbe* final consumers, on the other hand, exhibited the most divergent preference for desired-breeding traits (**Table 3.9**).

Based on the ecological zone, all final consumers concurred on "big fruit size", "high fruit miraculin content" and "fruit freshness" as the three most desired traits to consider in breeding an improved miracle berry variety (**Table 3.10**). Fruit shape and low potency were the two "men-specific" traits recorded in this study. All the final consumers belonging to Benin's sociolinguistic groups ranked big fruit size as the most desired trait in an improved miracle berry variety, whereas the Ghanaian sociolinguistic groups indicated "high fruit miraculin content" as their most desired trait in a variety of *S. dulcificum* (**Table 3.11**). Nevertheless, the final consumers of five out of the six sociolinguistic groups investigated in this study ranked "big fruit size", "high fruit miraculin content" and "fruit freshness" as their top three preferred traits for an ideal variety of miracle plant. This was reflected in the high overall concordance coefficient of 0.78 (**Table 3.11**).

Table 3.7. Sociolinguistic groups-based variation of the top five farmers-desired breeding traits for an improved variety of *Synsepalum. dulcificum*

Traits	Adja		Fon		Sahouè		Akan		Ewe		Ga-adangbe	
	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank
Big fruit size	28	1st	27	1st	23	2nd	38	3rd	42	1st	43	1st
Early fruiting	20	2nd	27	1st	26	1st	42	2nd	38	2nd	38	3rd
High fruit yielding	18	3rd	21	3rd	12	3rd	43	1st	32	3rd	40	2nd
Dwarf tree	14	4th	6	4th	6	4th	18	4th	20	5th	19	4th
Large tree crown	7	5th	-		6	4th	-	-	-	-	-	-
Fruit freshness	-		5	5th	-	-	-	-	-	-	18	5th
High fruit miraculin content	-		-	-	-	-	17	5th	31	4th	-	-
Overall ranking agreement	Kendall-W = 0.69											

-: Not applicable

Table 3.8. Final consumers-desired breeding traits for an improved variety of *Synsepalum dulcificum*

Traits	Cumulated score	Rank	Reasons (Citation frequency)
High miraculin content	368	1st	To enjoy the fruit (1)
Big fruit size	350	2nd	For a quick satisfaction (0.74), to attract consumers (0.26)
Fruit freshness	293	3rd	To attract consumer (0.64), for a safety of consumption (0.36)
Long fruit shelf life	91	4th	Better storability for a staggered consumption over the time (1)
High fruit edible ratio	48	5th	More juice from the fruit (1)
Fruit colour sharpness	41	6t	To make the fruit more attractive for the consumer (1)
Low potency	8	7th	To enjoy the taste of other foods after the consumption of miracle berry (1)
Firmness	6	8th	For a better storability (1)
Fruit shape	1	9th	Determined the fruit attractiveness (1)

Table 3.9. Pair-wise similarity index across sociolinguistic groups for all consumer-preferred traits for an improved variety of *Synsepalum dulcificum*. Values in the lower diagonal are pair-wise similarity indices of desired-traits and those in the upper diagonal are Kendall-W concordance coefficients for traits ranking

	Adja	Fon	Sahouè	Akan	Ewe	Ga-adangbe
Adja		1.00	0.76	0.70	0.90	0.90
Fon	0.92		0.97	0.70	0.90	0.90
Sahouè	1.00	0.92		0.65	0.81	0.81
Akan	0.76	0.85	0.76		0.90	0.95
Ewe	0.85	0.80	0.85	0.80		1.00
Ga-adangbe	0.66	0.769	0.66	0.92	0.71	

Table 3.10. Ecological zone-based variation of the top five final consumers-desired breeding traits for an improved variety of *Synsepalum dulcificum*

Traits	Guineo-Congolian zone		Deciduous forest zone		Evergreen forest zone	
	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank
Big fruit size	134	1st	211	3rd	5	3rd
High fruit miraculin content	78	2nd	284	1st	6	2nd
Fruit freshness	57	3rd	229	2nd	7	1st
High fruit edible ratio	41	4th				
Long fruit shelf life	36	5th	50	4th	5	3rd
Fruit colour sharpness	-	-	5	5th	-	-
Fruit firmness	-	-			2	5th
Overall ranking agreement	Kendall-W = 0.56					

Table 3.11. Sociolinguistic groups-based variation of the top five final consumers-desired breeding traits for an improved variety of *Synsepalum dulcificum*

Traits	Adja		Fon		Sahouè		Akan		Ewe		Ga-adangbe	
	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank	Cumulated score	Rank
Big fruit size	49	1st	51	1st	34	1st	57	3rd	85	2nd	74	2nd
High fruit miraculin content	22	2nd	41	2nd	15	3rd	104	1st	109	1st	77	1st
Fruit freshness	17	3rd	25	3rd	15	3rd	92	2nd	71	3rd	73	3rd
High fruit edible ratio	11	4th	-	-	18	2nd	-	-	-	-	-	-
Long fruit shelf life	11	4th	13	5th	12	5th	24	4th	13	4th	19	4th
Colour sharpness	-	-	20	4th	-	-	-	-	5	5th	-	-
Low potency	-	-	-	-	-	-	3	5th	-	-	2	5th
Overall ranking agreement	Kendall-W = 0.78											

-: Not applicable

-

3.3.8 Processor-desired breeding traits in *Synsepalum dulcificum*

At the end of the focus group discussion held with sections of managers from the Miracle Fruit Processing Ghana Ltd company, a total of five main preferred-breeding traits were listed in order of importance as: “high fruit miraculin content”, “big fruit size”, “high fruit edible ratio” and “fruit freshness” (for the miracle fruit powder production line) and “high seed portion” (for the dry seed production line).

3.3.9 Consistency among end-users desired breeding traits in *Synsepalum dulcificum*

The consistency analysis of the top five desired breeding traits among farmers, final consumers, and industrial processors (**Table 3.12**) revealed a good level of similarity ($C_S^3 = 0.60$) among the three end-user groups although agreement in ranking of these desired traits was in general very low (Kendall-W = 0.11). However, the pairwise similarity index between these three end-user groups revealed that final consumers’ and industrial processors’ top five desired breeding traits were highly similar (80% similarity index), and almost in perfect agreement on the ranking pattern (Kendall-W = 0.90). Conversely, desired-breeding traits for an improved variety were divergent from farmers and final consumers, on one hand, to farmers and processors, on the other hand. Either way, the most important traits for all three groups included big fruit size and high fruit miraculin content.

Table 3.12. Consistency parameters among miracle berry end-users for preferred breeding traits in *Synsepalum dulcificum* in West Africa

End-users preferred breeding traits and ranking			Rank	Overall ranking agreement
Farmers	Processors	Consumers		
Big fruit size	High fruit miraculin content	High fruit miraculin content	1st	0.11
Early fruiting	Big fruit size	Big fruit size	2nd	
High fruit yielding	High fruit edible ratio	Fruit freshness	3rd	
Dwarf tree	Fruit freshness	Long fruit shelf life	4th	
High fruit miraculin content	High seed portion	High fruit edible ratio	5th	

Pair-wise similarity index (lower diagonal) and ranking concordance coefficient (upper diagonal) of preferred trait between end-user groups

	Farmers	Processors	Consumers
Farmers		0.00	0.00
Processors	0.40		0.90
Consumers	0.40	0.80	

3.4 Discussion

3.4.1 Respondents' socio-demographic profile, tree holding system and relative importance of *Synsepalum dulcificum* in current production systems

Men were the main respondents in this study as they were the most involved in *S. dulcificum* management. This comes without surprise as it is known that management of perennial species is mainly carried out by men (Avohou et al., 2012). Indeed, planting perennial species is culturally bound with land ownership (Neef, 2001) and it is true that in West Africa, men are the prominent land owners (Quisumbing et al., 1999). For instance, in this study, men had on average 9.16 ± 2.77 ha, whereas women only had 2.53 ± 0.03 ha. However, the more frequent involvement of men in *S. dulcificum* management is not translated into a higher number of trees held compared with women. This finding aligned with that of Fandohan et al. (2017) who did not detect any differences in the number of *S. dulcificum* trees possessed by men and women in Benin, but is contrary to observation in other perennial species [e.g. bitter kola (*Garcinia kola* Heckel)] (Dadjo et al., 2020) where men owned more trees than women. This implies that if given the opportunity, women can contribute equally as men to conserve and manage *S. dulcificum,* as their management intensity index is also statistically comparable to that of men in this study.

Our findings also corroborated results of Fandohan et al. (2017) on the variation in the number of trees possessed following sociolinguistic group affiliation. Moreover, we detected an ecological gradient signal in the number of trees possessed, with respondents in drier ecological zones owning fewer trees compared with those in moister ecological zones. This confirms the crucial role of water availability and rainfall in the establishment and development of the species (Tchokponhoué et al., 2017). We could also speculate that the more market-oriented production in Ghana (moister ecological zones than in Benin) compared with Benin where the market is still emerging could have provided incentives for Ghanaian farmers to protect/grow more trees, though less than 5% had commercial plantations.

Crops listed by respondents as being important in their production system mostly included those already well established due to their high potential to contribute to food security (e.g. maize, cassava (Rosenthal and Ort, 2012)) and substantial income/cash generation (e.g. cocoa, coffee (Bandanaa et al., 2016)) or to stand as multi-purpose species (e.g. palm oil combining high market and food values). The poor ranking of the species despite its staggering market value might be explained by its low caloric value coupled with an incapacity of the current production system to expose the crop potential. As revealed by our findings, only a low proportion of respondents currently have the species in plantation while plantation size is still low. Then, it appears necessary to increase awareness of *S. dulcificum* potential to encourage more producers to meet the growing demand for the species. To that, unravelling socio-economic, biological and cultural drivers for sustainable cultivation of the species is required.

3.4.2 *Synsepalum dulcificum* management practices

Quality planting material is crucial for a successful establishment and subsequent development of tree species (Degrande et al., 2013). In this study, although seeds and seedlings were co-used by respondents, most of them established their trees/plantations using seedlings. This preference for seedlings transplanting was also reported in *B. sapida* (Ekué et al., 2010) and *G. kola* (Dadjo et al., 2020) two other minor orphan tree crops, and partly appeared in the case of *S. dulcificum* as a strategy to skirt the difficulty to germinate its recalcitrant seeds (Tchokponhoué et al., 2019a). According to farmers, using the seed compels one to quickly sow it, which is not the case with seedlings that have been harvested beneath *S. dulcificum* plants or orchards of which planting can be postponed and carried out later. Besides, using seedlings offered the advantage to select highly vigorous individuals, thus ensuring better growth and productivity. This could be the main reason why in the Evergreen forest ecological zone, which is dominated by plantations of *S. dulcificum,* respondents mainly used seedlings, sourced from the Plant Genetic Resources Research Institute, Bunso (a national research centre previously interested in promoting *S. dulcificum* in Ghana). As also observed in other crops such as *Oryza sativa* L.(Suvi et al., 2020), *Solanum tuberosum* L. (Echodu et al., 2019), *M. esculenta* (Teeken et al., 2018) and *M. geocarpum* (Coulibaly et al., 2020), exchange of plant materials prevailed in *S. dulcificum* among farmers and in all sociolinguistic groups, with sometimes the exchange taking place between farmers from distant districts. Contrary to observations in cassava where planting material exchange between farmers is partly through monetary means (Teeken et al., 2018), planting material exchange in the case of *S. dulcificum* occurred only as a gift, probably because most of the respondents perceived the species as not important in their production system. The quasi-absence or inactivity of a research centre focusing on the crop could have also favoured the magnitude of landraces exchanged among farmers. However, the *Adja* and *Fon* were the only farmers who accessed planting materials from the market. This could result from them being the only two sociolinguistic groups that transact the fruit on local markets.

Our results revealed that management intensity varied following ecological conditions and sociolinguistic groups as reported by (Blancas et al., 2013), but also following farmer's instruction level. Overall, management practices are easier and more rigorously applied when environmental conditions are conducive. As indicated above, the Evergreen forest ecological zone has a very high annual rainfall, which is favourable to the species. Besides, respondents in that ecological zone being in contact with a research centre for the acquisition of planting materials, they could have also likely benefited from technical advice enabling them to accumulate knowledge to apply more rigorously management practices. Indeed, specific practices such as pegging, lining and pruning were applied by 60–100% of respondents in the Evergreen zone against only 11.5 – 26.1% of those in Deciduous forest and Guineo-Congolian zones. Furthermore, being educated offers a competitive advantage in terms of understanding potential, ability to take initiative, to innovate and to try new technologies (Banla et al., 2018; Morris et al., 1999). This might explain why the educated respondents in general exerted a higher management intensity. The *Akan* represented the only sociolinguistic group that owned plantations; and maintaining a high level of productivity in these plantations made it necessary for them to exert a higher maintenance effort, which would have translated into the highest management intensity index observed in this sociolinguistic group.

The lack of improved *S. dulcificum* variety confirmed the absence of breeding programmes for the crop in West Africa. In such a context, farmer's knowledge represents a good asset to establish a sound and sustainable one. Indeed, farmers' knowledge of morphological variation in the species is in line with literature reports. For instance, the small fruit- and big fruit-sized morphotypes distinguished by the farmers in this study were previously reported by Coronel et al. (2009).

3.4.3 End-users' preferences

This study highlighted a diversity of traits desired by *S. dulcificum* end-users and revealed that farmers desired more breeding traits than final consumers, while the farmers and final consumers combined desired more traits than industrial processors. This indicated a specialization tendency for trait preference along the species value-chain. Within either farmers' or consumers' groups, the high similarity of desired traits across not only ecological zones but also sociolinguistic groups implied that a unique variety will work for each group at the level of the West African sub-region. More importantly, both final consumers and farmers in ecologically closer zones tended to have a more similar preference for breeding traits. The trend in final consumer-desired breeding traits aligned with the ethnocentrism hypothesis which predicts that ethnically close final consumers exhibited a similar choice pattern towards specific produces (Hersche, 1994).

It is reported that farmers ranked their preference for breeding traits to meet different scenarios including the necessity to mitigate or overcome their production constraints (Marimo et al., 2020). Here, the most important farmer-desired traits in *S. dulcificum* were towards ensuring a high profitability. Indeed, early fruiting reduces waiting time and ensures a quick acquisition of the marketable produce –the fruit–, which is better harvested on short trees and whose size conditioned the yield, this latter together with the taste (miraculin content) being determinant in attracting clients. Fruit tree species have a long juvenility and early fruiting is a strongly desired trait for all of them. Bhargava and Srivastava (2019) argued that early fruiting also helps reduce management period and cost and thus increases profitability. As reported in *M. esculenta* (Teeken et al., 2018) and *Musa* spp (Marimo et al., 2020), gender-specific desired traits also exist in *S. dulcificum* and seemed to illustrate gender role in the species' value chain. For instance, the "uniformity of ripening" was only mentioned by women who are known as the main harvesters/collectors and sellers of the fruit on local markets. For them, a uniform ripening could ensure a one-once and cost-efficient harvest as well as a grouped selling. Contrary to women, men do not sell the fruit in markets, but rather wait for retailers, processors or final consumers to come to them, hence their desire to have the ripe fruits to be on the trees over a longer period. This justifies the importance of the trait "extended on-tree fruit shelf life" that is specific to men. The significance of some traits poorly ranked today might increase in the future especially when it will come to large-scale cultivation. This is the case of the uniformity or ripening, a trait deemed of paramount importance in mechanized fruit harvesting (Gallardo et al., 2018).

For final consumers, a preferred variety of *S. dulcificum* should encompass traits that will contribute to attracting them and maximizing their satisfaction. Ahead of such traits is the miraculin content, driver of the landmark attribute of the miracle berry, that is, its ability to induce sweetness. This creates an impetus for exploitation of miraculin coupled with the fruit size and the fruit freshness. As indicated by our results, the ranking of these three traits at the overall level also corroborates their ranking across ecological zones and sociolinguistic groups, thus strengthening their high importance for final consumers.

3.5 Implications for future research and breeding

The moderate similarity of breeding traits among the three end-user groups suggested the necessity to foresee variety development per functionally similar group. Consequently, in the light of pairwise-similarity index obtained in this study, final consumers and industrial processing could be targeted when developing a specific variety and farmers exclusively when developing a different one. However, the three groups of end-users shared the most valuable desired traits in the species which include big fruit size and high fruit miraculin content.

Breeding perennial plant species is a long-lasting and demanding process in which each step is to be carefully conducted. In the case of *S. dulcificum,* the clear identification of end-users' preferred traits at the beginning of the process constitutes a robust basis for subsequent steps. In this study, we identified the key traits that can be incorporated in any breeding programme targeting the species improvement in West Africa. In absence of improved varieties in the species, the first necessity is to develop materials that significantly outperform current landraces for the elicited traits. This requires the proper evaluation of the existing diversity for an informed choice of parental lines to be advanced. *Synsepalum dulcificum* being a tree species, the very first step in this attempt is to identify potential elite trees, taking into consideration end-users preferred traits and an in situ phenotypic characterization of the species in its centre of diversity could offer first insights. Besides, information on how end-users' desired traits are correlated could inform on the relevance of the development of index selection for parallel and multi-traits selection. Furthermore, a genome-wide scan of the phenotypically characterized groups will help establish sound breeding populations towards varietal development and ensure a sustainable management of the remaining genetic resource diversity. Given its perennial nature, *S. dulcificum* will also benefit from the implementation of genomic selection as a chief- approach in selection strategies.

3.6 Conclusion

This study is the first of its kind to analyse management practices and breeding traits preference by end-users in *S. dulcificum*. Our findings provide first-order information that will feed the on-going pre-breeding process of the miracle berry in West Africa. From the holding system to the revenue generated by the crop to major end-users' preferred traits, this study compiled important information that will contribute to approaches to develop improved varieties and to enhance the crop's value chain.

In particular, we found that management practices were relatively more advanced in the Evergreen forest ecological zone than in the two other ecological zones of the study, with the *Akan* applying more rigorously ideal practices. While individual tree ownership was, in general, the dominant holding system of the species in West Africa, a prominence of plantations was observed in the Evergreen zones of Ghana. *Synsepalum dulcificum* has the potential to become one of the main cash crops in West Africa; but for this to be effective, development of improved material for sustainable cultivation is desired. The most influential traits desired by end-users include early fruiting, high fruit yielding, big fruit size, long fruit shelf-life, tree dwarfism, high fruit edible ratio, high fruit miraculin content and fruit freshness. Consequently, any breeding programmes targeting the species improvement in West Africa should prioritize these traits to meet multi-actor expectations. Because *S. dulcificum* is currently poorly cultivated, there is a necessity to promote/value the crop and explore factors that serve as drivers to trigger its sustainable cultivation in West Africa to combine the preservation of a fragile biodiverse environment with the growing demand and enable grass-root populations to benefit from the species potential.

3.7 References

Abady, S., Shimelis, H., and Janila, P. (2019). Farmers' perceived constraints to groundnut production, their variety choice and preferred traits in eastern Ethiopia: implications for drought-tolerance breeding. *J. Crop Improv.* **33**, 505-521.

Abbam, T., Johnson, F. A., Dash, J., and Padmadas, S. S. (2018). Spatiotemporal variations in rainfall and temperature in Ghana over the twentieth century, 1900–2014. *Earth Space Sci.* **5**, 120-132.

Achigan-Dako, E. G., Tchokponhoué, D. A., N'Danikou, S., Gebauer, J., and Vodouhè, R. S. (2015). Current knowledge and breeding perspectives for the miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell. *Genet. Resour. Crop Evol.* **62**, 465-476.

Adomou, A. (2005). Vegetation patterns and environmental gradients in Benin, University of Wageningen, Wageningen, Netherlands.

Adomou, A. C., Sinsin, B., and Van der Maesen, L. J. G. (2006). Phytosociological and chorological approaches to phytogeography: a meso-scale study in Benin. *Syst. Geograph. Plants* **76**, 155-178.

Agre, A. P., Bhattacharjee, R., Dansi, A., Becerra Lopez-Lavalle, L. A., Dansi, M., and Sanni, A. (2017). Assessment of cassava (*Manihot esculenta* Crantz) diversity, loss of landraces and farmers preference criteria in southern Benin using farmers' participatory approach. *Genet. Resour. Crop Evol.* **64**, 307-320.

Akoègninou, A., Van der Burg, W., and Van der Maesen, L. J. G. (2006). "Flore analytique du Bénin," Backhuys Publishers.

Anja, C., Weltzien, E., Rattunde, F., and Ashby, J. (2017). "Gender differentiation of farmer preferences for varietal traits in crop improvement: evidence and issues."

Arbonnier, M. (2004). "Trees, shrubs and lianas of West African dry zones. CIRAD."

Asravor, J., Wiredu, A. N., Siddig, K., and Onumah, E. E. (2019). Evaluating the environmental-technology gaps of rice farms in distinct agro-ecological zones of Ghana. *Sustainability* **11**, 2072.

Avohou, H. T., Vodouhe, R. S., Dansi, A., Kpeki, B., and Bellon, M. (2012). Ethnobotanical factors influencing the use and management of wild edible plants in agricultural environments in Benin. *Ethnobot. Res. Appl.* **10**, 571-592.

Ayensu, E. S. (1972). Morphology and anatomy of *Synsepalum dulcificum* (Sapotaceae). *Bot. J. Linn. Soc.* **65**, 179-187.

Bandanaa, J., Egyir, I. S., and Asante, I. (2016). Cocoa farming households in Ghana consider organic practices as climate smart and livelihoods enhancer. *Agric. Food Sec.* **5**, 29.

Banla, E. M., Dzidzienyo, D. K., Beatrice, I. E., Offei, S. K., Tongoona, P., and Desmae, H. (2018). Groundnut production constraints and farmers' trait preferences: a pre-breeding study in Togo. *J. Ethnobiol. Ethnomed.* **14**, 75.

Bechoff, A., Tomlins, K., Fliedel, G., Becerra Lopez-Lavalle, L. A., Westby, A., Hershey, C., and Dufour, D. (2018). Cassava traits and end-user preference: relating traits to consumer liking, sensory perception, and genetics. *Crit. Rev. Food Sci.* **58**, 547-567.

Bhargava, A., and Srivastava, S. (2019). Plant breeding. *In* "Participatory plant breeding: Concept and Applications" (A. Bhargava and S. Srivastava, eds.). Springer Nature, Singapore.

Blancas, J., Casas, A., Perez-Salicrup, D., Caballero, J., and Vega, E. (2013). Ecological and socio-cultural factors influencing plant management in Nahuatl communities of the Tehuacan Valley, Mexico. *J. Ethnobiol. Ethnomed.* **9**, 1.

Buckmire, R., and Francis, F. (1978). Pigments of miracle fruit, *Synsepalum dulcificum*, Schum, as potential food colorants. *J. Food Sci.* **43**, 908-911.

Casas, A., Rangel, S., Torres, I., Pérez-Negrón, E., Solís, L., Parra, F., Delgado, A., Blancas, J., Farfán, B., and Moreno, A. (2008). In situ management and conservation of plant resources in the Tehuacán- Cuicatlán Valley, Mexico: an ethnobotanical and ecological approach. *In* "Current topics in ethnobotany" (U. De Albuquerque and M. Ramos, eds.), Kerala: Research Signpost.

Chen, C. C., Liu, I. M., and Cheng, J. T. (2006). Improvement of insulin resistance by miracle fruit (S*ynsepalum dulcificum*) in fructose-rich chow-fed rats. *Phytother. Res.* **20**, 987-992.

Coronel, R. E., Sotto, R. C., and Rabara, R. C. (2009). The dwarf and round-fruited miracle fruit [*Synsepalum dulcificu*m (Schum. & Thonne) Daniell]. *Philip. J. Crop Sci.* **34**, 108-111.

Coulibaly, M., Agossou, C. O., Akohoué, F., Sawadogo, M., and Achigan-Dako, E. G. (2020). Farmers' preferences for genetic resources of kersting's groundnut [*Macrotyloma geocarpum* (Harms) Maréchal and Baudet] in the production systems of Burkina Faso and Ghana. *Agronomy* **10**, 371.

Dadjo, C., Nyende, A. B., Salako, K. V., Hounkpèvi, A., and Assogbadjo, A. E. (2020). Socio–economic factors determining conservation and cultivation of *Garcinia kola* Heckel—A medicinal plant extinct in the wild in Benin. *Econ. Bot.* **74**, 115-125.

Degrande, A., Tadjo, P., Takoutsing, B., Asaah, E., Tsobeng, A., and Tchoundjeu, Z. (2013). Getting trees into farmers' fields: success of rural nurseries in distributing high quality planting material in Cameroon. *Small-Scale For.* **12**, 403-420.

Diserud, O. H., and Ødegaard, F. (2007). A multiple-site similarity measure. *Biol. Lett.* **3**, 20-22.

Echodu, R., Edema, H., Wokorach, G., Zawedde, C., Otim, G., Luambano, N., Ateka, E. M., and Asiimwe, T. (2019). Farmers' practices and their knowledge of biotic constraints to sweetpotato production in East Africa. *Physiol. Mol. Plant. Pathol.* **105**, 3-16.

Ekué, M. R., Sinsin, B., Eyog-Matig, O., and Finkeldey, R. (2010). Uses, traditional management, perception of variation and preferences in ackee (*Blighia sapida* KD Koenig) fruit traits in Benin: implications for domestication and conservation. *J. Ethnobiol. Ethnomed.* **6**, 12.

Fandohan, A. B., Gouwakinnou, G. N., Tovissode, C. F., Bonou, A., Djonlonkou, S. F. B., Houndelo, L. F., Sinsin, C. L. B., and Assogbadjo, A. E. (2017). Usages traditionnels et valeur économique de *Synsepalum dulcificum* au Sud-Bénin. *Bois For. Trop.* **332**, 17-30.

Fandohan, B., Assogbadjo, A. E., Kakaï, R. G., and Sinsin, B. (2011). Geographical distribution, tree density and fruit production of *Tamarindus indica* L.(Fabaceae) across three ecological regions in Benin. *Fruits* **66**, 65-78.

Gallardo, R. K., Zhang, Q., Dossett, M., Polashock, J. J., Rodriguez-Saona, C., Vorsa, N., Edger, P. P., Ashrafi, H., Babiker, E., and Finn, C. E. (2018). Breeding trait priorities of the blueberry industry in the United States and Canada. *HortScience* **53**, 1021-1028.

Gamer, M., Jim Lemon, J., Fellows, I., and Singh, P. (2019). Package ‘irr’. *In* "Various Coefficients of Interrater Reliability and Agreement".

Hersche, J. (1994). Ethnocentric tendencies, marketing strategy and import purchase behavior. *Int. Mark. Rev.* **11**, 4-16.

Horn, L., Shimelis, H., and Laing, M. (2015). Participatory appraisal of production constraints, preferred traits and farming system of cowpea in the northern Namibia: implications for breeding. *Legume Res.* **38**, 691-700.

Houndonougbo, J., Kassa, B., Salako, V., Idohou, R., Assogbadjo, A., and Kakaï, R. G. (2020). Perceived variation of fruit traits, and preferences in African locust bean [*Parkia biglobosa* (Jacq.) Benth.] in Benin: implications for domestication. *Genet. Resour. Crop Evol.* **67**, 1-15.

Hussein, S. (2017). New variety design and product profiling. *In* "The business of plant breeding" (G. J. Persley and V. M. Anthony, eds.), pp. 85. CABI, Switzerland.

Kurihara, K., and Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science* **161**, 1241-1243.

Leakey, R. R., Weber, J. C., Page, T., Cornelius, J. P., Akinnifesi, F. K., Roshetko, J. M., Tchoundjeu, Z., and Jamnadass, R. (2012). Tree domestication in agroforestry: progress in the second decade (2003–2012). *In* "Agroforestry-the future of global land use" (D. Garrity, ed.), pp. 145-173. Springer, Dordrecht.

Lim, T. (2013). *Synsepalum dulcificum*. *In* "Edible medicinal and non-medicinal plants" (T. Lim, ed.), Vol. 6, pp. 146-150. Springer, Dordrecht.

Loko, Y. L. E., Montcho, D., Zandjanakou-Tachin, M., Orobiyi, A., Toffa, J., Hounmakou, E., Gavoedo, D., and Dansi, A. (2020). Farmers' management of peanut (*Arachis hypogaea* L.) diversity, their varietal preference traits and uses in Southern and Central Benin. *J. Crop Sci. Biotechnol.*, 1-14.

Marimo, P., Caron, C., Van den Bergh, I., Crichton, R., Weltzien, E., Ortiz, R., and Tumuhimbise, R. (2020). Gender and trait preferences for banana cultivation and use in sub-saharan Africa: A Literature Review. *Econ. Bot.* **74**, 226-241.

Mengistu, G., Shimelis, H., Laing, M., and Lule, D. (2019). Assessment of farmers' perceptions of production constraints, and their trait preferences of sorghum in western Ethiopia: implications for anthracnose resistance breeding. *Acta Agr. Scand. B Soil Plant* **69**, 241-249.

Morris, M. L., Tripp, R., and Dankyi, A. (1999). "Adoption and impacts of improved maize production technology: A case study of the Ghana Grains Development Project," Rep. No. 1405-7735. CIMMYT, Mexico.

Muhinyuza, J. B., Shimelis, H., Melis, R., Sibiya, J., and Nzaramba, M. N. (2012). Participatory assessment of potato production constraints and trait preferences in potato cultivar development in Rwanda. *Int. J. Dev. Sustain.* **1**, 358-380.

N'Danikou, S., Achigan-Dako, E. G., Tchokponhoue, D. A., Agossou, C. O., Houdegbe, C. A., Vodouhe, R. S., and Ahanchede, A. (2015). Modelling socioeconomic determinants for cultivation and in-situ conservation of *Vitex doniana* Sweet (Black plum), a wild harvested economic plant in Benin. *J. Ethnobiol. Ethnomed.* **11**, 28.

Neef, A. (2001). Land tenure and soil conservation practices-evidence from West Africa and Southeast Asia. *In* "10th International Soil Conservation Organization Conference", pp. 125-130, Purdue Univeristy and the USDA-ARS-National Soil Erosion Research Laboratory.

Niñez, V. (1987). Household gardens: Theoretical and policy considerations. *Agr. Syst.* **23**, 167-186.

Obeng, H. B. (1971). National soil map of Ghana. Catrographic Section of the Soil Research Institute.

Placide, R., Shimelis, H., Laing, M., and Gahakwa, D. (2015). Farmers' perceptions, production and productivity constraints, preferences, and breeding priorities of sweetpotato in Rwanda. *HortScience* **50**, 36-43.

Quisumbing, A. R., Payongayong, E. M., Aidoo, J., and Otsuka, K. (1999). "Women's land rights in the transition to individualized ownership: implications for the management of tree resources in Western Ghana."

R Core Team (2019). "R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,Vienna, Austria. URL https://www.R-project.org/," Vienna, Austria.

Rosenthal, D. M., and Ort, D. R. (2012). Examining cassava's potential to enhance food security under climate change. *Trop. Plant Biol.* **5**, 30-38.

Sibiya, J., Tongoona, P., Derera, J., and Makanda, I. (2013). Smallholder farmers' perceptions of maize diseases, pests, and other production constraints, their implications for maize breeding and evaluation of local maize cultivars in KwaZulu-Natal, South Africa. *Afr. J. Agr. Res.* **8**, 1790-1798.

Sogbohossou, E. D., Achigan-Dako, E. G., van Andel, T., and Schranz, M. E. (2018). Drivers of management of spider plant (*Gynandropsis gynandra*) across different socio-linguistic groups in Benin and Togo. *Econ. Bot.* **72**, 411-435.

Suvi, W. T., Shimelis, H., and Laing, M. (2020). Farmers' perceptions, production constraints and variety preferences of rice in Tanzania. *J. Crop Improv.*, 1-18.

Tafazoli, S., Vo, T. D., Roberts, A., Rodriguez, C., Viñas, R., Madonna, M. E., Chiang, Y.-H., Noronha, J. W., Holguin, J. C., and Ryder, J. A. (2019). Safety assessment of miraculin using in silico and in vitro digestibility analyses. *Food Chem. Toxicol.* **133**, 1-10.

Tchokponhoué, D., Achigan-Dako, E., N'Danikou, S., Houdégbé, A., Agossou, C., Assogba-Komlan, F., and Vodouhè, R. (2018). Regeneration ability and seedling growth in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Fruits* **73**, 13-21.

Tchokponhoué, D. A., N'Danikou, S., Hale, I., Van Deynze, A., and Achigan-Dako, E. G. (2017). Early fruiting in *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell juveniles induced by water and inorganic nutrient management. *F1000Research* **6**.

Tchokponhoué, D. A., N'Danikou, S., and Achigan-Dako, E. G. (2019). A combination of approaches evidenced seed storage behaviour in the miracle berry *Synsepalum dulcificum* (Schumach. et Thonn.) Daniell. *BMC Plt. Biol.* **19:117**, 1-13.

Teeken, B., Olaosebikan, O., Haleegoah, J., Oladejo, E., Madu, T., Bello, A., Parkes, E., Egesi, C., Kulakow, P., and Kirscht, H. (2018). Cassava trait preferences of men and women farmers in Nigeria: implications for breeding. *Econ. Bot.* **72**, 263-277.

The Plant List (2013). Version 1.1. Published on the Internet; http://www.theplantlist.org/ (accessed from April -December 2019).

Tongco, M. D. C. (2007). Purposive sampling as a tool for informant selection. *Ethnobot. Res. Applic.* **5**, 147-158.

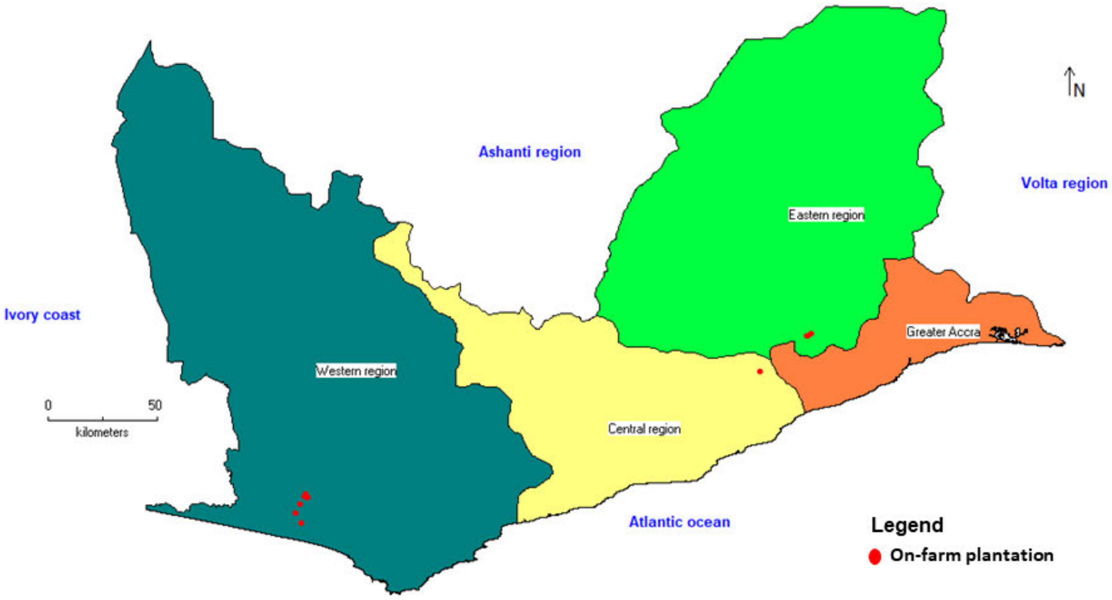
Weltzien, E., and Christinck, A. (2017). Participatory breeding: developing improved and relevant crop varieties with farmers. *In* "Agricultural Systems: Agroecology and rural innovation for development" (S. Snapp and B. Pound, eds.), pp. 259-301. Elsevier.

Weltzien, E., Rattunde, F., Christinck, A., Isaacs, K., and Ashby, J. (2019). Gender and farmer preferences for varietal traits: evidence and issues for crop improvement. *Plant Breed. Rev.* **43**, 243-278.

Yanagida, T. (2020). Package ‘misty’.

Appendix 3.1: Socio-demographic profile of individual respondents

Country		Benin			Ghana			Total N = 300
Ecological zone		Guineo-Congolian			Deciduous + Evergreen forest	Deciduous forest		
Sociolinguistic group		Adja (n = 46)	Fon (n = 62)	Sahouè (n = 38)	Akan (n = 52)	Ewe (n = 51)	Ga-adangbe (n = 51)	
Gender	Female (%)	1.33	3.00	1.67	2.67	2.67	2.33	13.67
	Male (%)	14.00	17.66	11.00	14.67	14.33	14.67	86.33
Instruction level	Illiterate (%)	8.00	10.67	6.33	1.00	0.67	3.67	30.34
	Literate (%)	1.00	0.67	0.00	0.67	0.00	0.33	2.67
	Primary (%)	3.00	5.33	3.33	2.66	3.00	4.67	21.99
	Secondary (%)	3.33	3.00	2.33	11.33	11.67	8.00	39.66
	≥ BAC (%)	0.00	1.00	0.67	1.67	1.67	0.33	5.34
Marital status	Single (%)	0.33	0.33	0.33	0.67	2.00	1.33	4.99
	Married (%)	15.00	18.33	11.67	14.00	13.33	12.67	85.00
	Widower (%)	0.00	2.00	0.67	2.67	1.67	3.00	10.01
Migration	Autochthon (%)	14.33	18.67	12.67	14.67	15.66	16.00	92.00
	Allochthon (%)	1.00	2.00	0.00	2.67	1.33	1.00	08.00
Age (year)	Range	25 - 88	26 - 89	30 - 85	29 - 83	26 - 87	30 - 102	25 - 102
	Mean ± se	54.46 ± 2.31	52.10 ± 2.16	57.95 ± 2.72	54.65 ± 1.91	56.90 ± 2.06	58.11 ± 2.55	56 .00 ± 01
Landholding (ha)	Range	0.40 - 25	0.04 - 30	0.40 - 40.00	0.16 - 3.51	0.40 - 7.02	0.80- 16.18	0.40 - 7.02
	Mean ± se	3.16 ± 0.59	3.42 ± 0.57	4.74 ± 1.15	13.33 ± 5.92	20.13 ± 2.63	4.35 ± 0.39	8.14 ± 2.41
Household size	Range	01 - 30	01 - 16	02 - 19	01 - 20	01 - 13	0 - 19	00 - 30
	Mean ± se	10.70 ± 0.98	6.55 ± 0.39	7.95 ± 0.58	06.42 ± 0.45	5.92 ± 0.41	6.63 ± 0.49	07.25 ± 0.24



Appendix 3.2. Distribution map of the twelve on-farm plantations of the miracle plant (*Synepalum dulcificum*) in Ghana. The map only focused on Ghana because no plantation was recorded in Benin and Togo list of plant species recorded in production systems involving *Synsepalum dulcificum* in the study area.

Appendix 3.3. Complete list of plant species recorded in production systems involving *Synsepalum dulcificum* in the study area

Botanical family	Species (**Voucher specimen code**)	Production systems
Anacardiaceae	*Mangifera indica* L. (**DAT01**)	Home-garden, farm
Annonaceae	*Annona muricata* L. (**DAT13**)	Home-garden, farm
	Monodora myristica (Gaetn.) Dunal (**DAT11**)	Home-garden, farm
	Xylopia aethiopica (Dunal) A. Rich. (**DAT04**)	Home-garden, farm
Apocynaceae	*Picralima nitida* (Stapf.) T.& H.Durand (**DAT30**)	Home-garden, farm
Araceae	*Colocasia esculenta* (L.) Schott (**DAT16**)	Home-garden, farm
Arecaceae	*Borassus aethiopum* Mart. (**DAT07**)	Home-garden
	Cocos nucifera L. (**DAT08**)	Home-garden, farm
	Elaeis guineensis Jacq. (**DAT09**)	Home-garden, farm
Asteraceae	*Vernonia amygdalina* Delile (**DAT10**)	Home-garden
Bignoniaceae	*Crescentia cujete L.* (**DAT03**)	Home-garden
	Newbouldia laevis (P.Beauv.) Seem (**DAT12**)	Home-garden
Brassicaceae	*Brassica oleracea L.* (**DAT02**)	Farm
Bromeliaceae	*Ananas comosus* (L.) Merr. (**DAT19**)	Home-garden, farm
Caricaceae	*Carica papaya* L. (**DAT14**)	Home-garden, farm
Clusiaceae	*Garcinia kola* Heckel (**DAT05**)	Home-garden, farm
Combretaceae	*Anogeissus leiocarpa* (DC.) Guill. & Perr. (**DAT06**)	Home-garden
	Terminalia superba Engl. & Diels (**DAT17**)	Farm
Dioscoreaceae	*Dioscorea alata* L. (**DAT21**)	Home-garden, farm
	Discorea dumetorum (Kunth) Pax (**DAT20**)	Farm
Euphorbiaceae	*Hevea brasiliensis* (Willd. Ex A. Juss.) Mull.Arg. (**DAT18**)	Farm
	Manihot esculenta Crantz. (**DAT22**)	Home-garden, farm
	Manihot glaziovii Mull. Arg. (**DAT35**)	Home-garden
Irvingiaceae	*Irvingia gabonensis* (Aubry-Lecomte ex O'Rorke) Baill. (**DAT24**)	Home-garden, farm
Lamiaceae	*Tectona grandis* L.f. (**DAT27**)	Home-garden, farm
Lauraceae	*Persea americana* Mill. (**DAT37**)	Home-garden, farm
Leguminosae	*Acacia auriculiformis* Benth. (**DAT25**)	Farm
	Arachis hypogaea L. (**DAT28**)	Home-garden, farm
	Caesalpinia bonduc (L.) Roxb. (**DAT29**)	Home-garden
	Glycine max (L.) Merr. (**DAT05**)	Farm
	Phaseolus lunatus L. (**DAT31**)	Farm
	Pterocarpus santalinoides DC. (**DAT32**)	Home-garden
	Vigna unguiculata (L.) Walp (**DAT51**)	Home-garden, farm
Malvaceae	*Abelmoschus esculentus* (L.) Moench (**DAT34**)	Home-garden, farm
	Cola nitida (Vent.) Schott & Endl. (**DAT23**)	Home-garden, farm
	Theobroma cacao L. (**DAT36**)	Home-garden, farm
Meliaceae	*Azadirachta indica* A.Juss. (**DAT26**)	Home-garden
	Khaya senegalensis (Desv.) A. Juss (**DAT38**)	Home-garden, farm
Moraceae	*Artocarpus altilis* (Parkinson ex F.A.Zorn) Fosberg (**DAT50**)	Home-garden

Botanical family	Species (**Voucher specimen code**)	Production systems
	Milicia excelsa (Welw.) C.C.Berg (**DAT49**)	Farm
Moringaceae	*Moringa oleifera* (Gaetn.) Dunal (**DAT41**)	Home-garden
Musaceae	*Musa parasidica* L. (**DAT58**)	Home-garden, farm
	Musa sapientum L. (**DAT62**)	Home-garden, farm
Myrtaceae	*Eucalyptus camaldulensis* Dehnh. (**DAT44**)	Home-garden
	Psidium guajava L. (**DAT45**)	Home-garden, farm
Oxalidaceae	*Averoa carrambola* L. (**DAT52**)	Home-garden
Passifloraceae	*Passiflora edulis* Sims (**DAT47**)	Farm
Poaceae	*Bambusa vulgaris* Schrad. (**DAT48**)	Home-garden
	Saccharum officinale L. (**DAT40**)	Home-garden
	Zea mays L. (**DAT39**)	Home-garden, farm
Rubiaceae	*Coffea canephora* L. (**DAT53**)	Home-garden, farm
Rutaceae	*Citrus limon* L. (**DAT46**)	Home-garden, farm
	Citrus sinensis (L.) Osbeck (**DAT33**)	Home-garden, farm
Sapotaceae	*Chrysophyllum albidum* G.Don (**DAT60**)	Home-garden, farm
	Synsepalum dulcificum (Schumach & Thonn.) Daniell (**DAT55**)	Home-garden, farm
Solanaceae	*Capsicum spp* (**DAT56**)	Home-garden, farm
	Lycopersicum esculentum L. (**DAT57**)	Home-garden, farm
	Solanum aethiopicum L. (**DAT42**)	Home-garden, farm
	Solanum macrocarpon L. (**DAT59**)	Home-garden, farm
	Solanum melongena L. (**DAT54**)	Home-garden
Zingiberaceae	*Afromomum melegueta* K.Schum. (**DAT61**)	Farm
	Zingiber officinale Roscoe (**DAT43**)	Farm

Appendix 3.4. Per gender-disaggregated importance of species involved in *Synsepalum dulcificum*-based production system

Production system	Men			Women		
	Species	IS*	Rank	Species	IS*	Rank
Farm	*Theobroma cacao* L.	4.09	1st	*Zea mays* L.	3.88	1st
	Zea mays L.	3.98	2nd	*Theobroma cacao* L.	3.77	2nd
	Elaeis guineensis Jacq.	3.50	3rd	*Manihot esculenta* Crantz.	3.50	3rd
	Manihot esculenta Crantz.	3.30	4th	*Elaeis guineensis* Jacq.	3.09	4th
	Solanum aethiopicum L.	3.13	5th	*Acacia auriculiformis* Benth.	3.00	5th
Home-garden	*Theobroma cacao* L.	4.35	1st	*Theobroma cacao* L.	4.50	1st
	Elaeis *guineensis* Jacq.	3.75	2nd	*Musa parasidica* L.	4.00	2nd
	Musa parasidica L.	3.46	3rd	*Persea americana* Mill.	4.00	3rd
	Musa sapientum L.	3.08	4th	*Chrysophyllum albidum* G.Don	3.33	4th
	Cocos nuciferae L.	2.85	5th	*Colocasia esculenta* (L.) Schott	3.33	5th

*IS: Importance score.

CHAPTER 4 [2]

Phenotypic variation, functional traits repeatability and core collection inference in *Synsepalum dulcificum* (Schumach & Thonn.) Daniell reveals the Dahomey Gap as a centre of diversity

Abstract

The miracle plant *Synsepalum dulcificum* is a multipurpose natural sweetener and a promising West African orphan fruit shrub candidate for genetic improvement. Unfortunately, basic knowledge such as phenotypic variation and inheritance estimates required for implementing a breeding program are still lacking. A set of 203 accessions were sampled in two habitats from seven populations spread across the Dahomey Gap (DG) and the Upper Guinea forest (UG) in West Africa. The phenotypic diversity and allometric relationships among functional traits were analysed; the broad-sense heritability was estimated for fruit-traits, and a mini-core collection was developed in the species. Quantitative variation in tree- and fruit-traits was recorded, and multivariate analyses were performed to assess relationships among accessions, whereas heritability was estimated using the coefficient of repeatability. Tree-traits observed in *S. dulcificum* were more variable than fruit-traits. While habitat-type only affected tree-traits, the provenance population significantly affected both fruit- and tree-traits, with the UG populations outperforming the DG ones. Significant correlations were observed among fruit-traits on one hand, and among tree-traits on the other hand, whereas poor correlations were observed between tree- and fruit-traits. The multivariate analysis grouped accessions in three clusters. Promising individuals for high fruit mass and pulp-dense genotypes' selection were identified within clusters. Repeatability estimates for fruit-traits ranged from 0.015 (edible ratio) to 0.88 (fruit mass). The Core Hunter algorithm enabled the extraction of 41 individuals as robust representatives of the initial set of 203 accessions, and the mapping of this core collection suggested Dahomey Gap as a centre of diversity of the species. These original findings offer opportunities, not only for the genetic improvement of end-users' desired traits in *S. dulcificum,* but also for targeted ex-situ conservation in the species.

---

[2] Chapter published as: **Tchokponhoué, D. A.,** Achigan-Dako, E. G., N'Danikou, S., Nyadanu, D., Kahane, R., Houéto J., Fassinou Hotegni, N.V., Odindo, A. O., & Sibiya, J. (2020). Phenotypic variation, functional traits repeatability and core collection inference in *Synsepalum dulcificum* (Schumach & Thonn.) Daniell reveals the Dahomey Gap as a centre of diversity. *Scientific Reports*, 10:19538. https://doi.org/10.1038/s41598-020-76103-4

4.1 Introduction

The miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell [Syn. *Richardella dulcifica* (Schumach & Thonn.) Baehni] (**Figs. 4.1a-b**) – Sapotaceae – is a West African native shrub that produces a red fruit known as "miracle berry" (**Figs. 4.1c-d**). The berry is a unique natural source of miraculin (Kurihara and Beidler, 1968), a glycoprotein contained in the pulp that induces sweetness. *Synsepalum dulcificum* has numerous other applications ranging from traditional to modern uses (Achigan-Dako et al., 2015). Traditionally, all the non-edible parts of the species are involved in the treatment of many ailments including malaria, enuresis, coughing, and tooth decay among others (Fandohan et al., 2017; Oumorou et al., 2010). Currently, the miracle berry is considered as a reliable alternative to synthetic sugar (Andrade et al., 2019; Rodrigues et al., 2016) and helps to control obesity (Swamy et al., 2014) and diabetes (Chen et al., 2006; Han et al., 2019; Obafemi et al., 2017) in modern pharmaceuticals. It also serves as a food additive with its red skin and whitish pulp being, respectively, used for colouring and sweetening beverages and foods (Buckmire and Francis, 1978). Applications of the species in cosmetics relate to the seed (**Fig. 4.1e**) oil used in hair breakage control and the improvement of hand and finger motor skills (Del Campo et al., 2017; Gorin et al., 2018). In West Africa, particularly in Benin, the miracle fruit is also sold on the open market and thus contributes to improving household livelihoods through income generation (Fandohan et al., 2017).

Despite its growing market value (link1, link2), *S. dulcificum* is on the red list of the International Union for Conservation of Nature and classified as an endangered species in Benin (Adomou, 2005), with a risk of total depletion if no action is taken. Unfortunately, no conservation effort at either the local, regional and global scale has been so far reported in the species. Significant advances have been made in the development of farmer-friendly regeneration protocols (Tchokponhoué et al., 2018) and cultivation practices (Tchokponhoué et al., 2017; Tchokponhoué et al., 2019b) for production of *S. dulcificum* as a horticultural crop. However, foundation research for the genetic improvement and gene pool conservation of the species is still scanty. With an increasing interest in *S. dulcificum* (Choi and Garza, 2020; Tafazoli et al., 2019) worldwide, the economic potential of the shrub can be exploited if elite cultivars with improved traits are developed. However, the lack of basic knowledge on the genetics of the species is perceived as a bottleneck to the development and implementation of a relevant breeding program. For instance, the extent of genetic variation in *S. dulcificum* is elusive, and until now, there has been no study on the phenotypic diversity of the species. In their report on the molecular diversity of the species, Chibuzor et al. (2017) did not include any quantitative phenotypic variation. While ex-situ evaluation is a common approach for the phenotypic characterization of annual and biannual plant genetic resources (Sogbohossou et al., 2019), in-situ evaluation is important in any preliminary study of the morphological variation in perennial species (Singh et al., 2020; Sun et al., 2019). This has been shown in many fruit tree species including black raspberry (*Rubus occidentalis* L.) (Yazdanpour et al., 2018), sumac (*Rhus coriaria* (L.) Kuntze) (Fereidoonfar et al., 2018), ber (*Ziziphus* spp.) (Norouzi et al., 2017), wild almond (*Prunus scoparia* L.) (Khadivi-Khub and Anjam, 2014) and the African bush mangoes (*Irvingia* spp.) (Vihotogbé et al., 2013), among others. This approach also provides primary information to shape a gene pool conservation strategy (Ouborg, 2010), and sounds relevant when it comes to an orphan and threatened species such as the miracle plant.

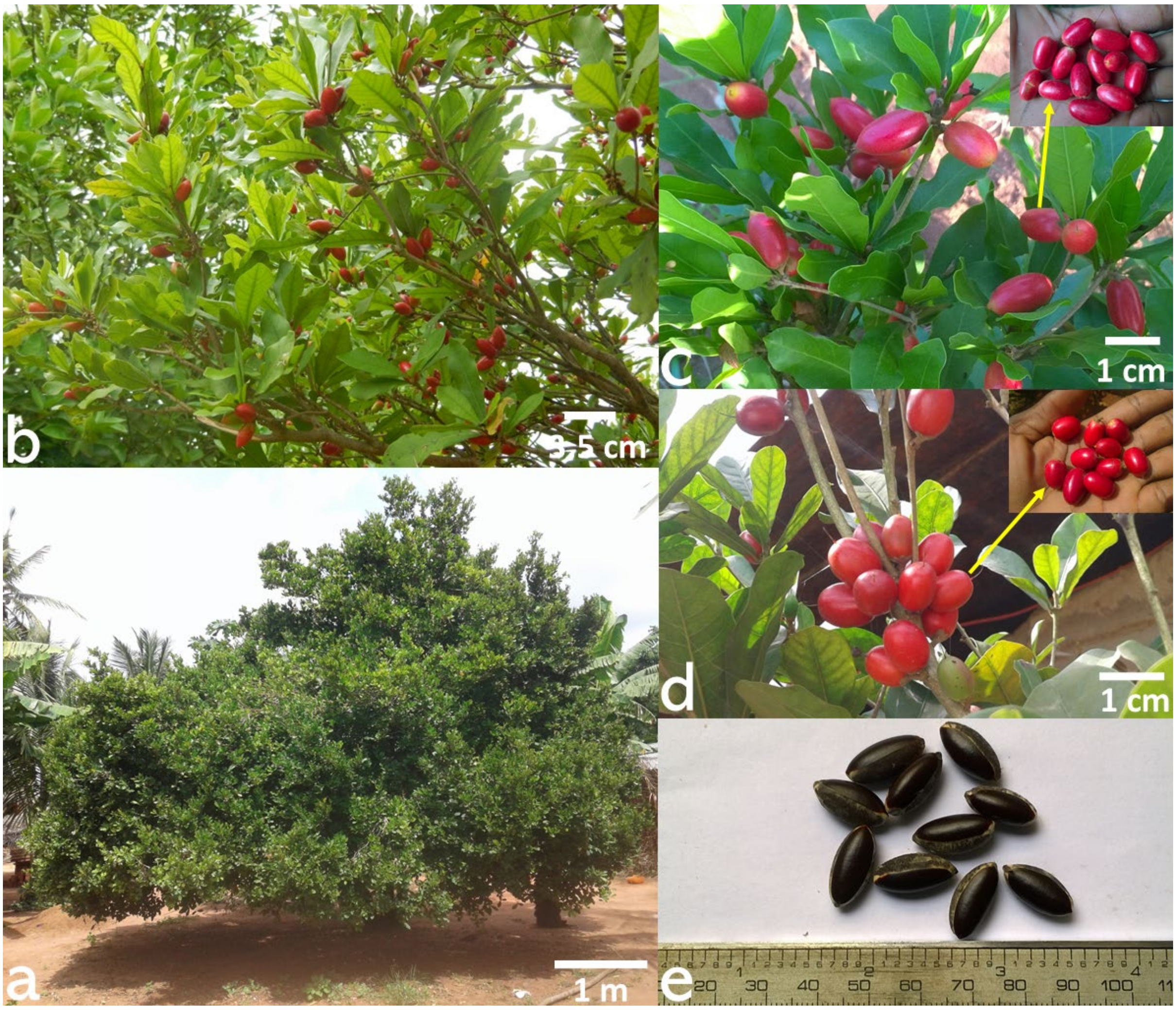


Figure 4.1. *Synsepalum dulcificum*, the studied species and some of its key organs. (**a**) A 110-year-old tree in home garden in Mission-Tove, Togo; (**b**) *S. dulcificum* branches bearing ripe fruits (miracle berries); (**c**) Oblong-shaped miracle berries; (**d**) Ovoid-shaped miracle berries; and (**e**) *S. dulcificum seeds*.

Given that the fruit is the most valued and economically important organ in *S. dulcificum,* fruit and seed traits, thus, represent good candidates for genetic diversity studies and any selection. Interestingly, fruits and related traits are highly prioritized by end-users as shown in the previous chapter. Predicting genetic gain for a target trait requires prior knowledge of its heritability (Martínez-García et al., 2017), and such information is currently unavailable in *S. dulcificum*. In perennial species, the coefficient of repeatability, which is the correlation between repeated measurements of the same individual over time and/or space is considered a reliable estimator of broad-sense heritability (Falconer, 1960). It has been used for fruit-traits heritability estimation in several tropical, subtropical and temperate tree species including mangabeira (*Hancornia speciose* Gomes) (Fonseca et al., 2018), akebia [*Akebia trifoliate* (Thumb.) Koidz.] (Zou et al., 2018), shea tree (*Vitellaria paradoxa* C.F. Gaertn.) (Sanou et al., 2006) and peach tree [*Prunus persica* (L.) Batch] (Albuquerque et al., 2004), among others. This metric appears relevant in the case of *S. dulcificum* to inform on the potential inheritance of fruit traits.

Core collection development emerged as an approach to optimize plant genetic resources conservation and utilization but is still in its infancy for perennial and fruit tree species (Belaj et al., 2012; Le Cunff et al., 2008; Mahmoodi et al., 2019). More importantly, no evidence exists for such resources in African indigenous underutilized fruit tree species, despite its potential to rationalize orphan fruit tree breeding process by helping to skirt the necessity to evaluate multi-traits on a large sample of individuals. Because *Synsepalum dulcificum* exhibits recalcitrant seeds (Tchokponhoué et al., 2019a), cryopreservation appears to be the most adequate way for the maintenance of its genetic resources (Bharuth et al., 2020; Edesi et al., 2020). Unfortunately, cryopreservation is currently non-existent in West Africa. The only cost-effective alternative ways to conserve the species ex-situ is to establish it as living collections, whose size will be tremendously reduced, and the management much eased with the development of a core collection.

The current natural distribution area of *S. dulcificum* ranges from Ivory Coast to the Democratic Republic of Congo (link). This area covers four major ecological regions of the African lowland rain forest including the Upper Guinea forest block (UG), the Dahomey Gap (DG), the Lower Guinea forest block (LG) and the Congolian forest block (CB) (Leal, 2004). However, the species has been reported to originate from a complex of countries including Ghana, Togo and Benin (Inglett and May, 1968; Juhé-Beaulaton, 2014; Swenson et al., 2008), which encompasses UG and DG.

Therefore, the objectives of the present study were to: i) assess the phenotypic diversity of *S. dulcificum* among genetic resources from Benin, Togo, and Ghana in West Africa; ii) estimate the heritability of some major fruit traits, and iii) develop a core collection in the species. The overall goal of this study is to design and implement a relevant breeding program for the miracle plant in West Africa in order to sustain the utilization of the species. The study was guided by the following research questions: i) what is the extent of quantitative variations in the miracle berry tree- and fruit-traits in its centre of origin and how are these traits affected by the ecological conditions?, ii) what is the clustering pattern among the miracle plant accessions?, ii) how heritable are the miracle berry traits?, and iv) what is the structure of the phenotype-based core collection in the species in the distribution area of West Africa including Benin, Togo and Ghana?.

4.2 Materials and methods

4.2.1 Study area for sampling

The data collection focused on the Upper Guinea forest block (UG) and the Dahomey Gap (DG) (**Fig. 4.2**) and was expected to capture maximum diversity. The DG is a savannah corridor disrupting the zonal West African rain forest. It was created during the late Holocene following an abrupt climatic change between 4,500 and 3,400 cal. years BP (Salzmann and Hoelzmann, 2005). It is considered as a geographical barrier between UG demarcated at the West through the Volta river, and LG demarcated at the East through the Weme river (Booth, 1958). The DG zone experiences 1,000 - 1,200 mm annual rainfall. The vegetation is a mixture of savannahs, gallery forests, and fragments of swamp forests, as compared to that of UG, dominated by rain forest with > 2,000 mm annual rainfall (White, 1983). The UG block corresponds to the rain forest belt extending from Sierra Leone to Ghana (Leal, 2004). A total of seven collection sites were investigated including four (Oueme, Zou, Mono and Volta) in DG and three (Eastern, Central and Western) in UG, giving rise to seven populations (**Fig. 4.2**). A population of *S. dulcificum* was defined as a set of accessions possibly interbreeding and randomly distributed in an agroforestry or tree-based production system within the same area or environment. Specifically, two distant populations of *S. dulcificum* (e.g. Mono and Volta, Eastern and Western) were separated by a minimum geographical distance of 50 km whereas two adjacent populations of *S. dulcificum* were separated by a maximum distance of 15 km coupled with the existence of a natural barrier such as a mountain (as observed between the adjacent populations Central and Eastern, in Ghana) or a river (as observed between the adjacent populations Zou and Oueme, and between the adjacent populations Mono and Zou, in Benin).

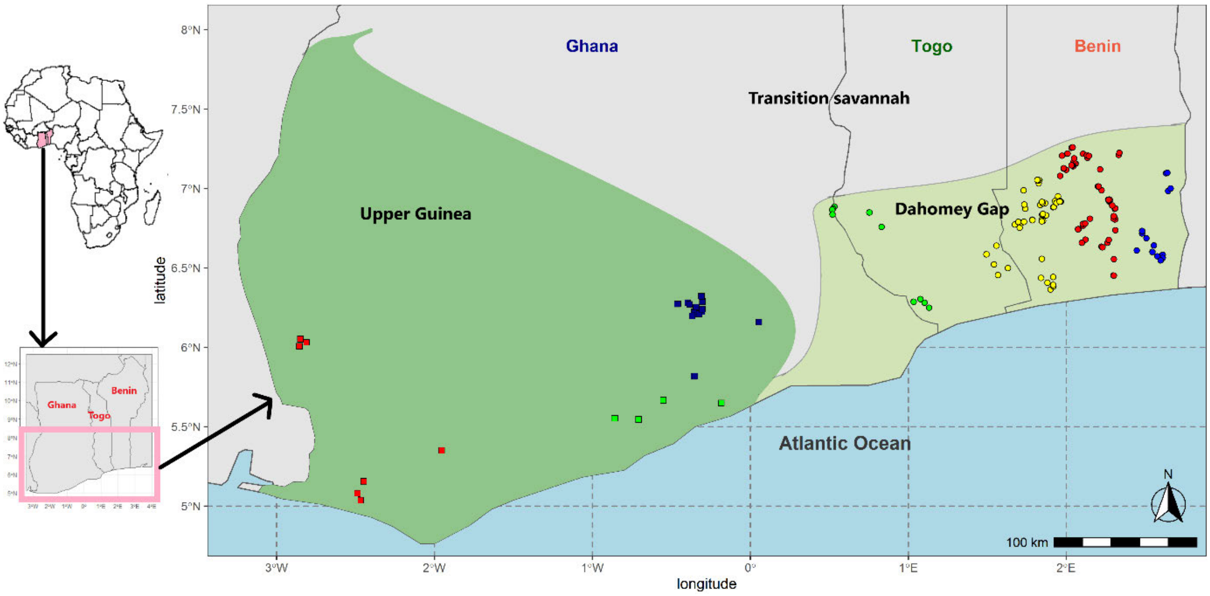


Figure 4.2. Study area showing the spatial distribution of the 203 accessions of Synsepalum dulcificum sampled in West Africa. Accessions are coloured according to the seven population sites under study. Filled shapes: Upper Guinea forest block populations (🟥 :Western, 🟩:Central, 🟦: Eastern); and Empty shapes: Dahomey Gap populations (🟢: Volta, 🟡: Mono, 🔴: Zou, 🔵: Oueme). Map generated using the R environment (https://www.r-project.org/) (Version 3.6.2).

From September 2016 to August 2019, a total of 203 accessions (individual trees) of *S. dulcificum* were sampled from the seven populations, with the number of accessions sampled per population varying from 4 to 66 (Oueme: 19, Zou: 67, Mono: 67, Volta:11, Eastern: 25 Central: 4, and Western: 10). The sampled accessions were all of reproductive age (*i.e.*, having borne fruits at least once), and bearing fruits during the field sampling. The minimum distance of 100 m was observed between two sampled shrubs to reduce occurrence of closely related individuals.

4.2.2 Data collection

For each accession, data on the habitat type, tree-traits, fruit-traits, and the tree age (where possible) were recorded. The habitat was recorded as either a home garden, which is here defined according to Niñez (1987) as a production system characterized by its proximity to dwellings, and over which the owner has a certain control on the target products; or as a farm, which is defined as an open agricultural production space, and usually farther away from the dwellings than a home garden. The tree age was estimated by the owner (the planter) where applicable; otherwise, one of his closest relatives (e.g., brother, child, or grandchild) gave the information.

Three tree-traits and six fruit-traits were measured per accession (**Table 4.1**). The tree-traits included: 1- Plant height (PlHgt), 2- Tree diameter at ground level (DiamGr), and 3- Tree crown diameter (CrDiam). The fruit traits included: 4- Fruit length (AvFL), 5- Fruit width (AvFW), 6- Fresh fruit mass (AvFM), 7- Seed mass (AvSM), 8- Edible ratio (EdRt) which is the pulp-to-fruit ratio, and 9- Fruit shape index (FrSh).

Table 4.1. Quantitative variation in *Synsepalum dulcificum* tree- and fruit traits (n = sample size for the measurement, Min: Minimum, Max: Maximum, SD: Standard deviation; cv: coefficient of variation).

Traits		Traits code	n	Min	Max	Mean	SD	cv (%)
Tree-traits	Tree diameter at ground level (cm)	DiamGr	203	4.28	68.78	19.12	10.20	53.34
	Tree height (m)	PlHgt	203	0.86	7.5	3.47	1.27	36.59
	Tree crown diameter (m)	CrDiam	203	0.75	8.8	4.24	1.62	38.2
Fruit-traits	Fruit length (mm)	AvFL	1,015	13.04	26.41	18.86	1.82	9.65
	Fruit width (mm)	AvFW	1,015	7.39	18.3	10.59	1.28	12.08
	Individual fruit mass (g)	AvFM	8,875	0.43	2.03	1.09	0.28	25.68
	Individual seed mass (g)	AvSM	8,875	0.16	0.6	0.35	0.07	20
	Edible ratio (%)	EdRt	8.875	45.25	83.01	66.99	7.17	10.7
	Fruit shape index	FrSh	1,015	1.06	2.39	1.79	0.18	10.05

The plant height was measured from the soil surface to the tallest point of the shrub using a tape measure. The crown diameter was obtained by averaging the measures taken in the south-north and east-west directions. The diameter at ground level was obtained using the **Equation (4.1):**

$$DiamGr = C_{Gr}/\pi \qquad \textbf{(Eqt. 4.1)}$$

where DiamGr is the tree diameter at ground level, and C_{Gr} is the circumference at ground level measured with a tape measure.

The fruit-traits were recorded in-situ on 20 to 50 mature, fully ripe (red) and pests-free fruits that were randomly collected from the four sides (South, East, West and North) of each tree and bulked in replicates of 10 fruits each. A total of 8,875 single fruits were assessed for this study.

The fresh mass of each replicate (FM_{10}) was measured (± 0.01 g) using a portable electronic scale (OHAUS, model: PA512, Port Melbourne, Australia). The mean value of the replicates was then divided by 10 to obtain an average fruit mass (AvFM) of the accession. The same procedure was used to obtain the average seed mass (AvSM) of each accession after the fruits were depulped, while being kept in their replicate lot. The Edible Ratio (EdRt) was determined on each replicate, and the mean EdRt of an accession was the average value of all replicates of the accession. The EdRt was computed following the **Equation (4.2)**:

$$EdRt = (FM_{10} - SM_{10})/FM_{10} \qquad \textbf{(Eqt. 4.2)}$$

where FM_{10} is the fresh mass of 10 fruits, and SM_{10} is the fresh mass of the seeds of the same 10 fruits.

The fruit length and width were measured on five fruits randomly sampled from the different replicates per accession, using a digital Vernier calliper (± 0.01 mm). The mean values of all replicates per accession were used as the average fruit length (AvFL) and width (AvFW) of the accession. The fruit shape index (FrSh) of each accession was determined from five fruits following the **Equation (4.3)**:

$$FrSh = AvFL/AvFW \qquad \textbf{(Eqt. 4.3)}$$

4.2.3 Data analysis

All the analyses were performed using the R language Version 3.6.2 (R Core Team, 2019).

4.2.3.1 Tree morphology and fruit-trait variation

Descriptive statistics (minimum, maximum, mean, standard deviation, coefficient of variation) were computed to understand the overall variation in the species. The effect of the habitat-type on the tree height, diameter at ground level, crown diameter, fruit mass, fruit length, fruit width, fruit shape index and seed mass were assessed using a t-test, or a Mann Whitney's test, where relevant. Likewise, the effect of the provenance population of accessions on the same variables was analysed using a type-II analysis of variance (ANOVA) to account for sample size imbalance among populations, or a Kruskal-Wallis test, where relevant. The Edible Ratio was assessed using a generalized linear model with a quasi-binomial error structure to account for overdispersion. Where the effect of the population was significant, a contrast analysis was employed to depict the influence of the ecological region (Upper Guinea vs Dahomey Gap) on the different parameters tested.

4.2.3.2 Relationship among tree morphology and fruit-traits

Relationships among tree- and fruit-traits were sought, and their significance were tested using Pearson's or Spearman's correlation tests, as appropriate. The correlation strength and significance were illustrated using the *chart.Correlation ()* function of the "PerfomanceAnalytics" package (Peterson et al., 2018) . To depict the effect of the ecological region on the relationship among functional fruit-traits in the species, allometric regressions were used with sequential tests for slope equality and elevation shift between DG and UG accessions for various regression lines. This was done using the *sma ()* function of the "smatr" package (Warton et al., 2012), whereas regression lines were drawn using the *ggPredict ()* function of the "ggiraphExtra" package (Moon, 2016).

To assess the relationships among and between accessions and tree- and fruit-traits, a series of multivariate analyses were conducted. First, a Principal Component Analysis (PCA) was performed using the *PCA ()* function implemented in the "FactoMineR" package (Lê et al., 2008) to retain the most meaningful components. Second, the clustering tendency in the dataset was tested using the Hopkins statistics "H" through the *hopkins ()* function implemented in the "clustertend" package (YiLan and RuTong, 2015). A value of H close to 0.5 indicated that the dataset was not clusterable, whereas a value of H close to 0 indicated a clusterable dataset. In the latter case, a hierarchical clustering on the retained principal components (HCPC) was performed using the *HCPC ()* function of the "FactoMineR" package. The graphical outputs were visualized using the function *fviz_cluster ()* of the "factoextra" package (Kassambara and Mundt, 2017).

4.2.3.3 Estimation of repeatability

For each of the six fruit-traits, the repeatability estimation was conducted using the linear mixed effect models framework implemented in the "rptR" package (Stoffel et al., 2017). To that, the provenance population was set as a fixed effect and the tree/accession as a random effect. The repeatability of the edible ratio was estimated using a generalized linear mixed-effect model with the *rptProportion ()* function whereas repeatability estimates of other fruit traits (fruit mass, seed mass, fruit length, fruit width and fruit shape index) were determined using a generalized linear mixed-effects model fitted by restricted maximum likelihood (REML) with the *rptGaussian ()* function. The interest in using the "rptR" package over the classical approach was mainly its ability to deal with non-gaussian data (e.g., the edible ratio in our case that is a proportion data). The significance of each repeatability was tested using the Likelihood Ratio Test (LRT), and the uncertainties (standard error: S.E. and Confidence Interval: C.I.) associated to the repeatability estimates were determined using a parametric bootstrapping (n = 1,000 data samples).

4.2.3.4 Development and evaluation of the core collection

A Core Hunter phenotype data was developed using the *phenotype ()* function implemented in Core Hunter version 3 of the "Core Hunter" package (De Beukelaer et al., 2018). Thereafter the *samplecore ()* function of the same package was applied to the Core Hunter data previously generated to develop the core set of accessions using the average-entry-to-nearest-entry distance optimization objective supported by the Gower distance measure. The quality of the core collection developed was evaluated by first computing the Coincidence Rate of range (CR%) (**Equation 4.4**), the Variable Rate (VR%) (**Equation 4.5**), the Variance Difference percentage (VD%) (**Equation 4.6**), and the Mean Difference percentage (MD%) (**Equation 4.7**) between the whole collection and the core collection, following Kim et al. (2007) and Hu et al. (2000):

$$\mathrm{CR\%} = \frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}} \frac{\mathrm{R}_{\mathrm{C_i}}}{\mathrm{R}_{\mathrm{W_i}}} \times 100 \qquad \textbf{(Eqt. 4.4)}$$

Where $\mathrm{R}_{\mathrm{C_i}}$ is the range of the core collection for the trait i, and $\mathrm{R}_{\mathrm{W_i}}$ is the range of the whole collection for the trait i;

$$\mathrm{VR\%} = \frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}} \frac{\mathrm{CV}_{\mathrm{C_i}}}{\mathrm{CV}_{\mathrm{W_i}}} \times 100 \qquad \textbf{(Eqt. 4.5)}$$

where $\mathrm{CV}_{\mathrm{C_i}}$ is the coefficient of variation of the core collection for the trait i, and $\mathrm{CV}_{\mathrm{W_i}}$ is the coefficient of variation of the whole collection for the trait i;

$$V\mathrm{D\%} = \frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}} \frac{\left|\mathrm{V}_{\mathrm{W_i}} - \mathrm{V}_{\mathrm{C_i}}\right|}{\mathrm{V}_{\mathrm{C_i}}} \times 100 \qquad \textbf{(Eqt. 4.6)}$$

where $\mathrm{V}_{\mathrm{C_i}}$ is the variance of the core collection for the trait i, and $\mathrm{V}_{\mathrm{W_i}}$ is the variance of the whole collection for the trait i;

$$\mathrm{MD\%} = \frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}} \frac{\left|\mathrm{M_{W_i}} - \mathrm{M_{C_i}}\right|}{\mathrm{M_{C_i}}} \times 100 \qquad \textbf{(Eqt. 4.7)}$$

where $\mathrm{M_{C_i}}$ is the mean of the core collection for the trait i, and $\mathrm{M_{W_i}}$ is the mean of the whole collection for the trait i.

Then, the existence of significant difference of phenotypic traits was tested between the whole collection and the core collection, using a t-test (for seed mass), a Wilcoxon's test (for fruit mass, fruit length, fruit width, tree stem diameter, tree height, tree crown diameter and fruit shape index), or a generalized linear model with a quasi-binomial error distribution (for edible ratio).

Additional packages such as "rworldxtra" (South, 2012), "ggplot2" (Wickham, 2016) (for basic graphics management), "rnaturalearth" (South, 2017) and "sf" (Pebesma, 2018) (for mapping and geo-referencing) were also used.

4.3 Results

4.3.1 Quantitative variation in tree- and fruit-traits

Tree-traits generally exhibited larger variation than fruit-traits (**Table 4.1**). The phenotypic coefficients of variation ranged from 36.59% (tree height) to 53.34% (tree diameter at ground level) for tree-traits, and from 9.65% (fruit length) to 25.68% (fruit mass) for fruit-traits. The effect of the population provenance was significant ($p < 0.05$) on all the assessed traits except on tree diameter ($p = 0.211$) and tree crown diameter ($p = 0.337$) (**Figs. 4.3a-i**). In particular, trees were taller in the Upper Guinea forest (UG) than in the Dahomey Gap (DG) ($p = 0.001$, **Fig. 4.3b**). Likewise, accessions from the UG produced heavier, longer and larger fruits with higher edible ratio than those from the DG ($p < 0.0001$, **Figs. 4.3d-h**), whereas fruit shape index ($p < 0.0001$, **Fig. 4.3i**) was greater in accessions from DG (oblong fruit) than from UG (round fruit). Among the screened populations, the Volta population produced bigger fruits and heavier seeds ($p < 0.0001$).

Contrary to the provenance population, habitat-type had lower effect on most studied traits (**Figs. 4.4a-i**), only significantly affecting tree-traits (tree height, $p = 0.004$, and crown diameter, $p = 0.02$), with relatively bigger shrubs in home gardens than in farms. Tree diameter at ground level ($p = 0.07$) was relatively greater in shrubs on farms than in those on home gardens. The age estimates of the studied trees ranged from 9 to 150 years and was on average 57.69 ± 2.04 years. Insignificant age difference was found between individuals on farms (57.75 ± 3.44 years) and those in-home gardens (57.66 ± 2.56 years) ($p = 0.56$) (**Appendix 4.1**).

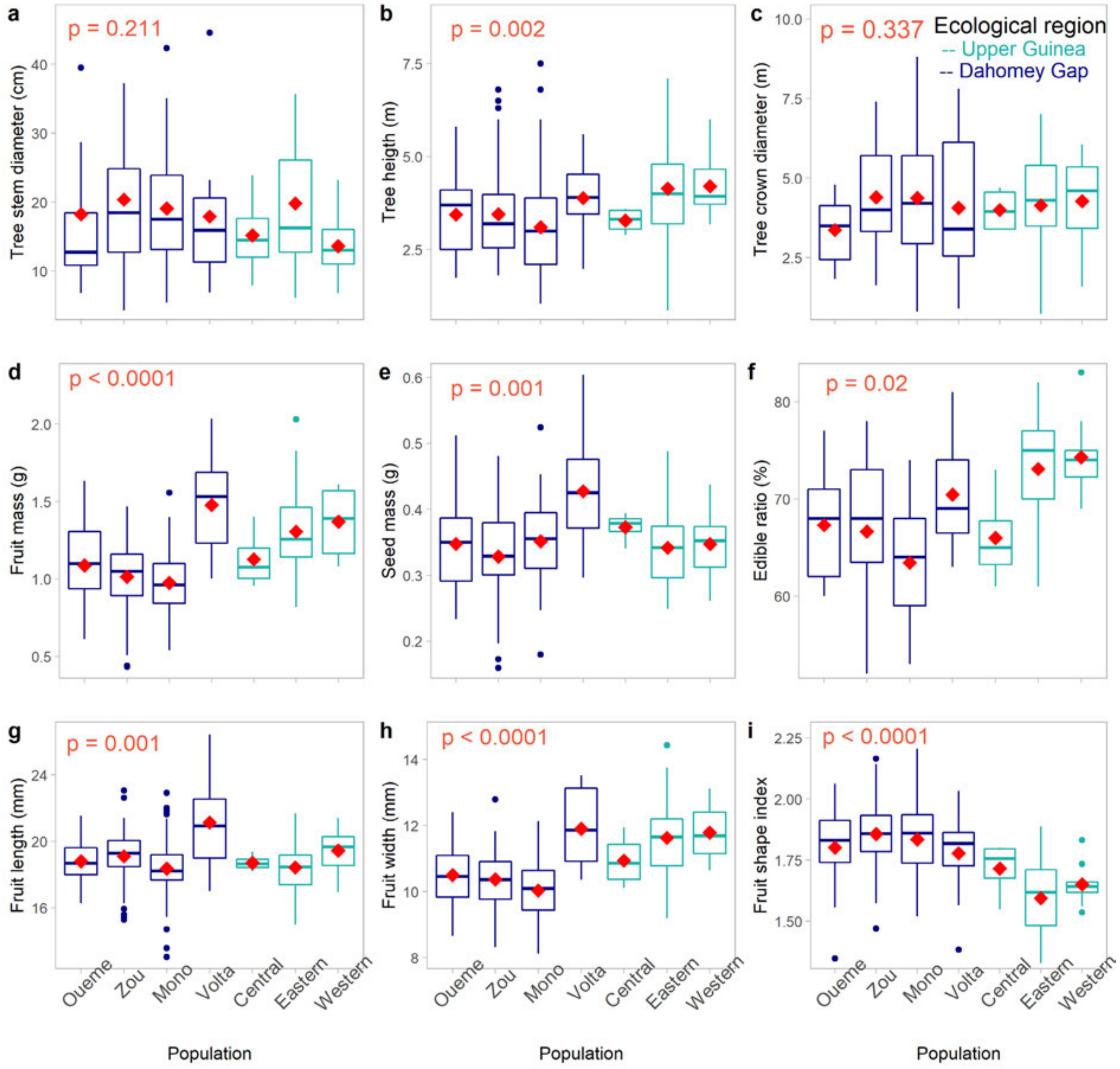


Figure 4.3. Quantitative variation of *Synsepalum dulcificum* tree- and fruit-traits across the seven sampled populations. Median values are in bold, red diamond shapes represent mean values, dots above and below boxplots are outliers, and lower and upper tails represent minimum and maximum values, respectively.

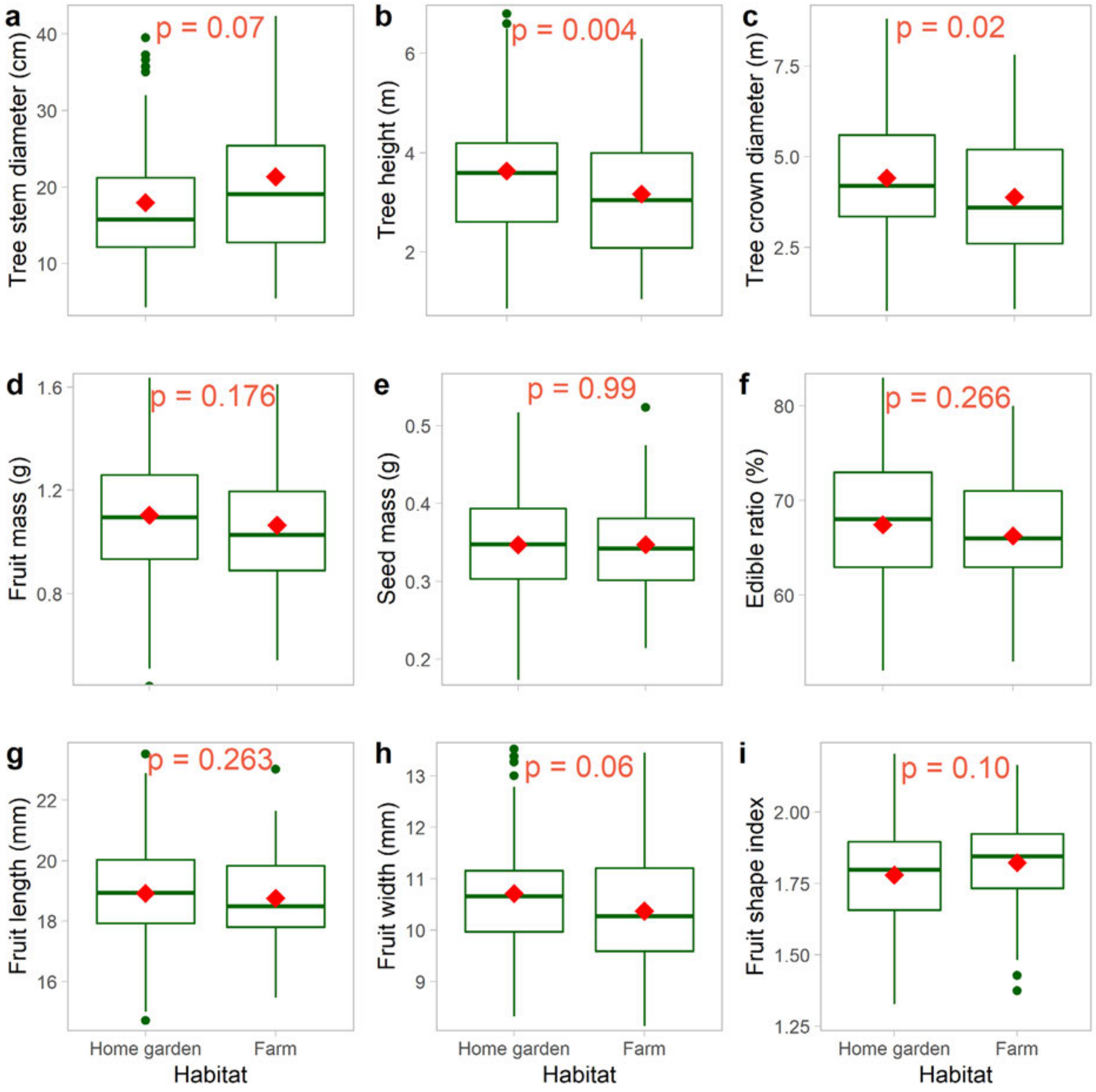


Figure 4.4. Quantitative variation of *Synsepalum dulcificum* tree- and fruit-traits across the habitat-types under study. Median values are in bold, red diamond shapes represent mean values, dots above and below boxplots are outliers, and lower and upper tails represent minimum and maximum values, respectively.

4.3.2 Relationships among tree- and fruit-traits

Coefficients of correlation ranged from -0.00286 to 0.83 (**Fig. 4.5**). All tree-traits were positively and significantly correlated, and the strongest relationship was observed between tree height and crown diameter (r = 0.62, $p < 0.001$). The strongest relationship among fruit-traits was found between fruit mass and fruit width (r = 0.83, $p < 0.001$). However, fruit-traits were not uniformly correlated. Fruit mass was significantly correlated with all other fruit-traits (r ≥ 0.49, *p < 0.01*) except with fruit shape index (r = - 0.4), and fruit shape index was only correlated with fruit length (r = 0.34, *p < 0.001*). Edible ratio was positively correlated with fruit mass (r = 0.63, *p < 0.001*), but negatively with the seed mass (r = -0.26, *p < 0.001*). Overall, there was a poor relationship between tree-traits and fruit-traits in *S. dulcificum*, and the highest correlation was observed between tree height and fruit width (r = 0.31, *p < 0.01*) (**Fig. 4.5**).

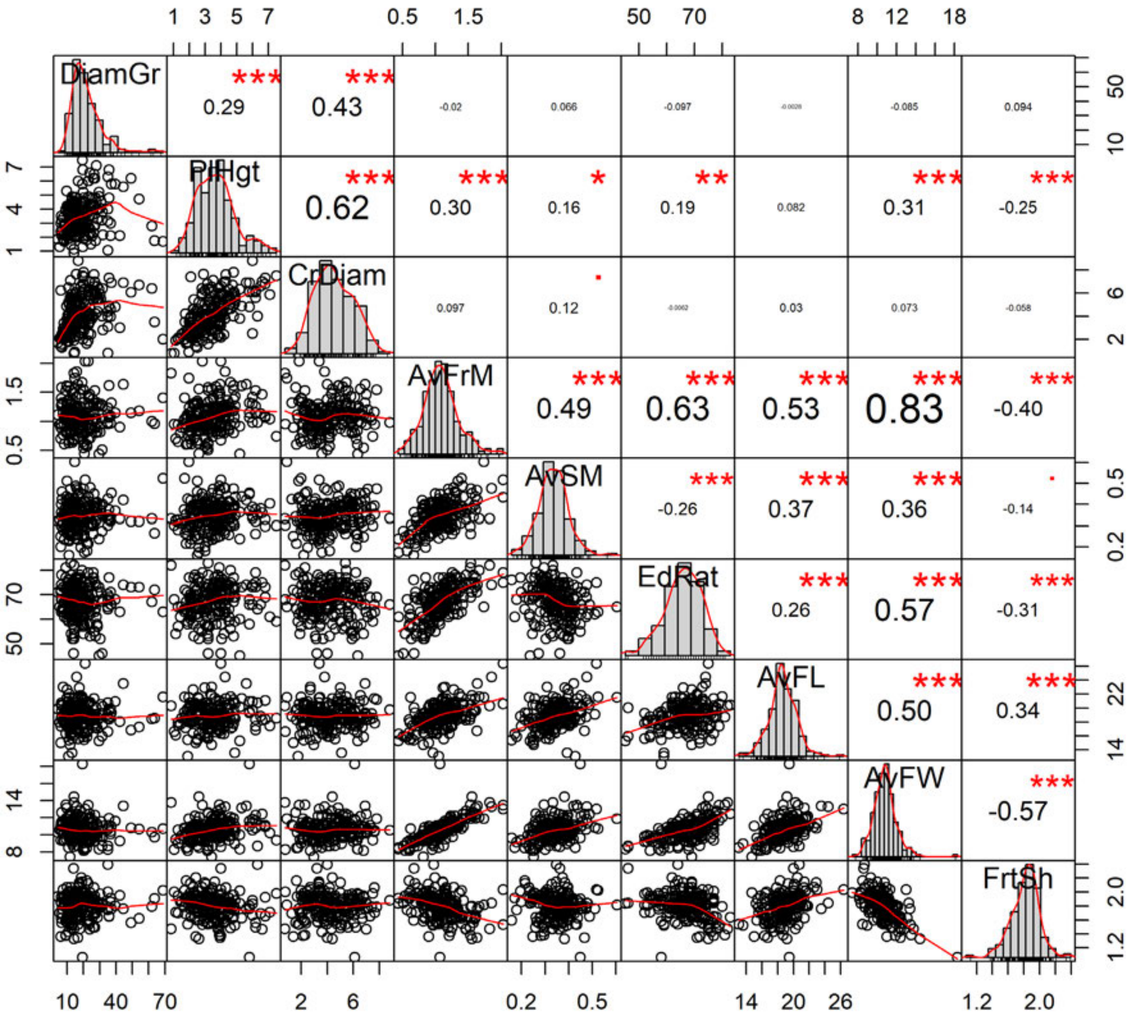


Figure 4.5. Mixed correlogram based on nine *Synsepalum dulcificum* tree- and fruit-traits (Traits list detailed in Table 4. 1). On the main diagonal histograms are presented for each trait. Figures at the lower triangle present bivariate scatterplots, with a fitted line. Upper triangle contains correlation coefficients (r) with significance value from a correlation test. Magnification of r values is proportional to the correlation strength and magnification of the p-value value symbol is proportional to the significance level: *** for p < 0.001, ** for p < 0.01, * for p < 0.05, ▪for p < 0.1.

Similarly, out of all the traits assessed, the tree age estimate was only correlated, and moderately to the tree diameter (r = 0.53, *p < 0.001*) and weakly to the tree crown diameter (r = 0.3, *p < 0.001*) (**Appendix 4.2**).

The ecological conditions seemed to influence the relationships observed between fruit-traits (**Figs. 4.6a-i**).

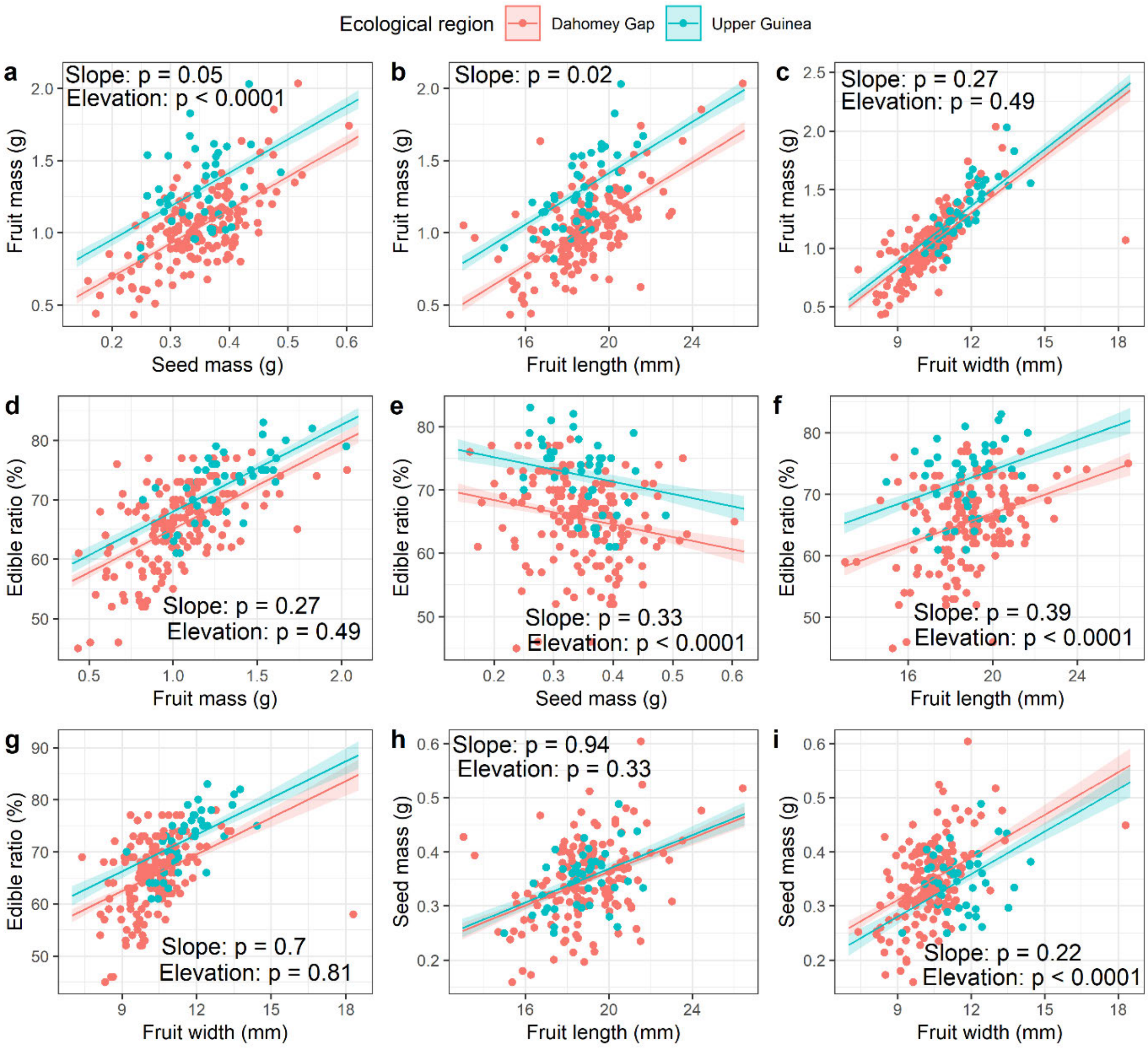


Figure 4.6. Allometric relationships among selected functional fruit-traits in *Synsepalum dulcificum* across the Dahomey Gap (in red) and the Upper Guinea forest (in blue) ecological regions. Slope: Test on difference between slopes; Elevation: Test on elevation shift in regression lines.

For example, stronger correlations were observed between fruit mass and seed mass (**Fig. 4.6a**) on one hand, and between fruit mass and fruit length (**Fig. 4.6b**), on the other hand, for accessions from the UG compared with those from the DG. However, the correlation between fruit mass and fruit width was not influenced by ecological conditions (**Fig. 4.6c**). Similarly, correlations between edible ratio and fruit mass (**Fig. 4.6d**), edible ratio and fruit width (**Fig. 4.6g**), and between seed mass and fruit length (**Fig. 4.6h**) were not affected by ecological conditions. Nonetheless, a highly significant effect of the ecological region was noted on correlations between edible ratio and seed mass (**Fig. 4.6e**), edible ratio and fruit length (**Fig. 4.6f**) and between seed mass and fruit width (**Fig. 4.6i**). In this case, an increase in fruit width resulted in a greater seed mass gain in accessions from DG than those from UG, whereas a decrease in seed mass yielded a greater gain in edible ratio in accessions from UG than those from the DG. A greater correlation between edible ratio and fruit length was observed under UG ecological conditions than under DG conditions.

4.3.3 Relationships among accessions of *S. dulcificum*

From the principal component analysis, it emerged that the first four components explained a significant proportion (83.3%) of the total variance (**Figs. 4.7a-c**) and these were retained for further analyses. The most correlated traits to the first component (35.1% of the total variability) were fruit mass ($r = 0.93$, $p < 0.0001$), fruit width ($r = 0.91$, $p < 0.0001$) and edible ratio ($r = 0.61$, $p < 0.0001$). This first component is that of the “fruit size”. Similarly, the crown diameter ($r = 0.83$, $p < 0.0001$), tree height ($r = 0.69$, $p < 0.0001$) and tree diameter ($r = 0.61$, $p < 0.0001$) were the most correlated to the second component, which accounted for 19.6% of the total variability. This component is that of “tree-shape”. The fruit shape index ($r = 0.756$, $p < 0.001$) and fruit length ($r = 0.71$, $p < 0.001$) were the most represented variables on the third component “the fruit shape component” that accounted for 16.8% of the total variability. No variables significantly correlated to the component 4 (11.8% of the total variance).

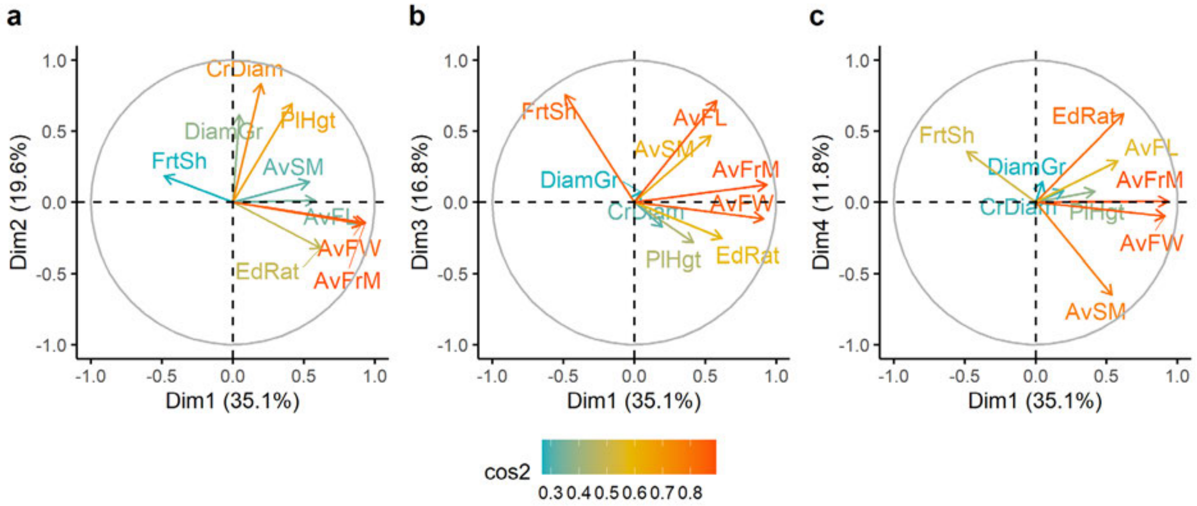


Figure 4.7. Correlation circle indicating *Synsepalum dulcificum* tree- and fruit-traits projection on the first two components: (**a**); first and third components: (**b**); and first and fourth components: (**c**). Traits list is detailed in Table 4.1).

The Hopkins statistics computed on the data was H = 0.23 (< 0.5), indicating the existence of significant clusters in the dataset. The hierarchical clustering on the first four components produced three clusters (**Fig. 4.8**). Clusters 1, 2 and 3 were made up of 46.8%, 26.5% and 26.7% of the total number of accessions, respectively and a χ^2 test on the clustering structure indicated a significant representation of accessions from the Upper Guinea in cluster 3 (*p = 0.001*) (**Appendix 4.3**). The most determinant quantitative characteristics of each cluster are presented in **Table 4.2**.

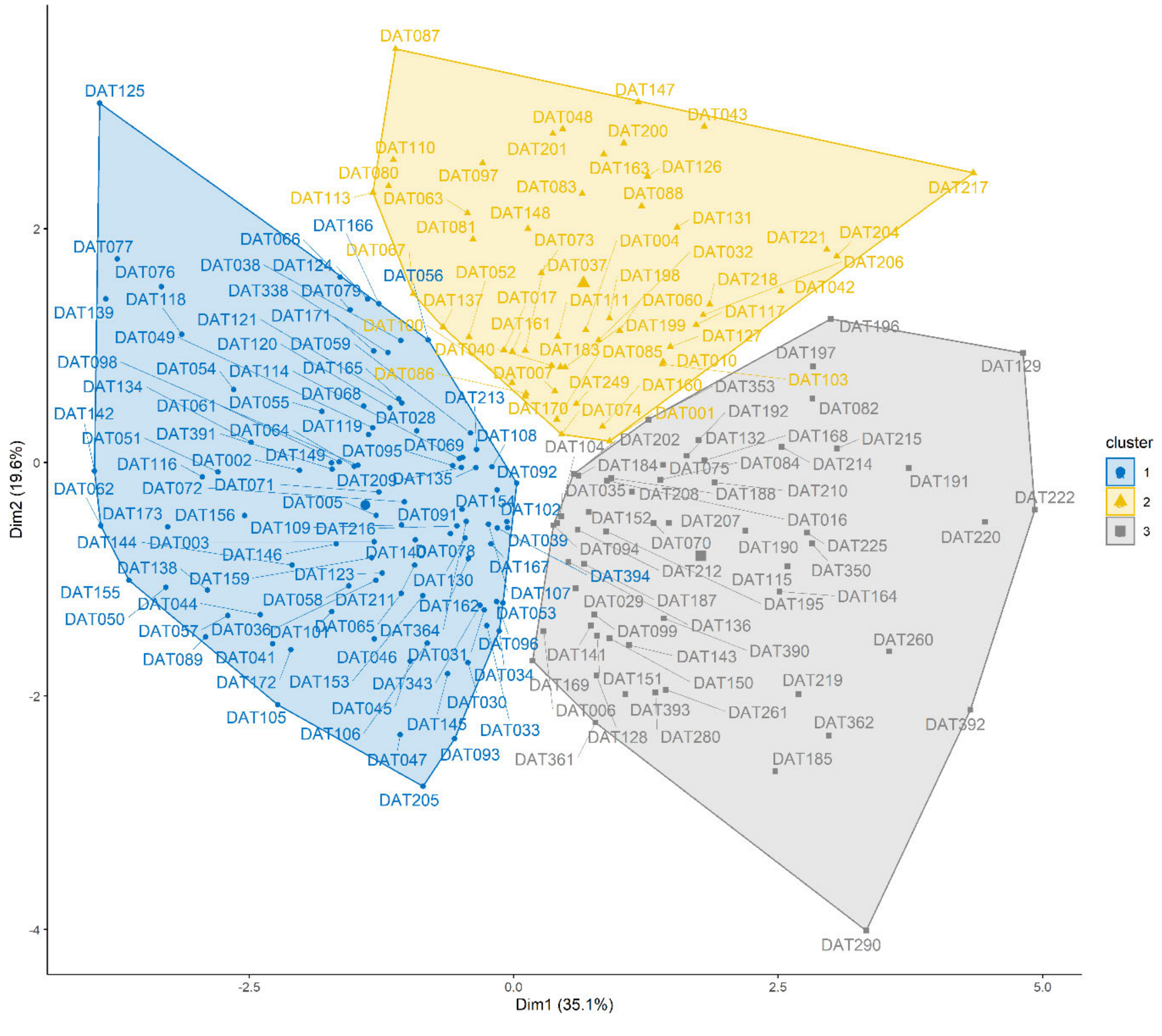


Figure 4.8. Factor map showing *Synsepalum dulcificum* accessions grouping in clusters

Cluster 1 was significantly characterized by all the nine traits studied, and mainly encompassed the least performing accessions observed in the study (**Table 4.2**). These accessions were of weak architecture, producing poor standard fruits that are markedly oblong in shape. Cluster 2 was significantly described by fruit length and seed mass in addition to tree traits (**Table 4.2**). Accessions in this cluster had a strong tree architecture and a high seed mass. Cluster 3 was characterized by individuals with a weak tree architecture, but with the highest performing fruit-traits. In part, accessions of cluster 3 had the highest fruit mass, fruit width, edible ratio, and seed mass (**Table 4.2**) and are the elite trees in this study.

Table 4.2. Phenotypic differentiation among *S. dulcificum* accession clusters. Figures in bold indicate values that are significantly greater than the overall means, while the rest of the figures indicate values that are significantly lower than the overall means for all accessions. NA: Quantitative trait not significantly describing the cluster.

Quantitative traits		Clusters		
		1	2	3
Tree-traits	DiamGr	17.14 ± 0.88	**26.7 ± 12.21**	15.42 ± 6.21
	PlHgt	2.81 ± 0.09	**4.59 ± 1.30**	NA
	CrDiam	3.60 ± 0.12	**5.97 ± 1.20**	3.68 ± 1.36
Fruit-traits	AvFL	17.99 ± 0.16	**19.45 ± 1.42**	**19.79 ± 1.78**
	AvFW	9.76 ± 0.07	**NA**	**11.89 ± 1.21**
	AvFM	0.88 ± 0.02	**NA**	**1.37 ± 0.22**
	AvSM	0.31 ± 0.01	**0.37 ± 0.05**	**0.37 ± 0.06**
	EdRt	64.01 ± 0.69	**NA**	72.32 ± 5.34
	FrSh	**1.84 ± 0.02**	**NA**	1.67 ± 0.19

4.3.4 Fruit traits' repeatability in *S. dulcificum*

Repeatability coefficients for the six fruit-traits under study are presented in **Table 4.3**. All the repeatability estimates were highly significant ($p < 0.0001$) with most traits exhibiting high repeatability. Thus, the trait with the highest potential broad-sense heritability based on the repeatability estimates was fruit mass, whereas the one with the least broad-sense heritability was the edible ratio.

Table 4.3. Estimation of repeatability (adjusted) coefficient for six fruit-traits in *S. dulcificum* (S.E.: Standard error).

Fruit-traits	Repeatability (r_{adj})	Uncertainty*		Likelihood ratio test
		S.E.	Confidence interval	
AvFL	0.745	0.023	[0.70, 0.786]	< 0.0001
AvFW	0.666	0.028	[0.608, 0.717]	< 0.0001
AvFrM	0.883	0.013	[0.855, 0.906]	< 0.0001
AvSM	0.821	0.019	[0.784, 0.857]	< 0.0001
EdRt	0.015	0.002	[0.01, 0.018]	< 0.0001
FrSh	0.58	0.03	[0.527, 0.647]	< 0.0001

*Determined using 1,000 parametric bootstraps.

4.3.5 Mapping the core collection of *S. dulcificum*

Based on the phenotypic data from the 203 accessions evaluated, the Core Hunter algorithm returned as core collection a total of 41 accessions. The spatial distribution of these 41 accessions is presented in **Figure 4.9**. This core set is made up of eight accessions from the Upper Guinea forest and 33 from the Dahomey Gap. The repartition of the core set accessions in the three clusters obtained above indicated that Clusters 1, 2 and 3 were represented by 15, 13 and 13 accessions, respectively, with the Upper Guinea ecological region having mainly accessions of Cluster 3 against the Dahomey Gap region encompassing accessions from all the three clusters.

The coincidence rate of range (CR) per trait varied from 87.57 to 100% for an overall CR of 97.6% (**Table 4.4**). Similarly, all the variable rates (VR) per trait were higher than 100%, whereas the variance difference percentage between the whole collection and the core collection was 46.95%. More captivatingly, the difference between the whole collection and the core collection was not significant for any of the nine phenotypic traits evaluated ($0.26 < p < 0.99$). This observation was supported by the low mean difference percentage between the core and the whole collection, which was 3.23% (**Table 4.4**).

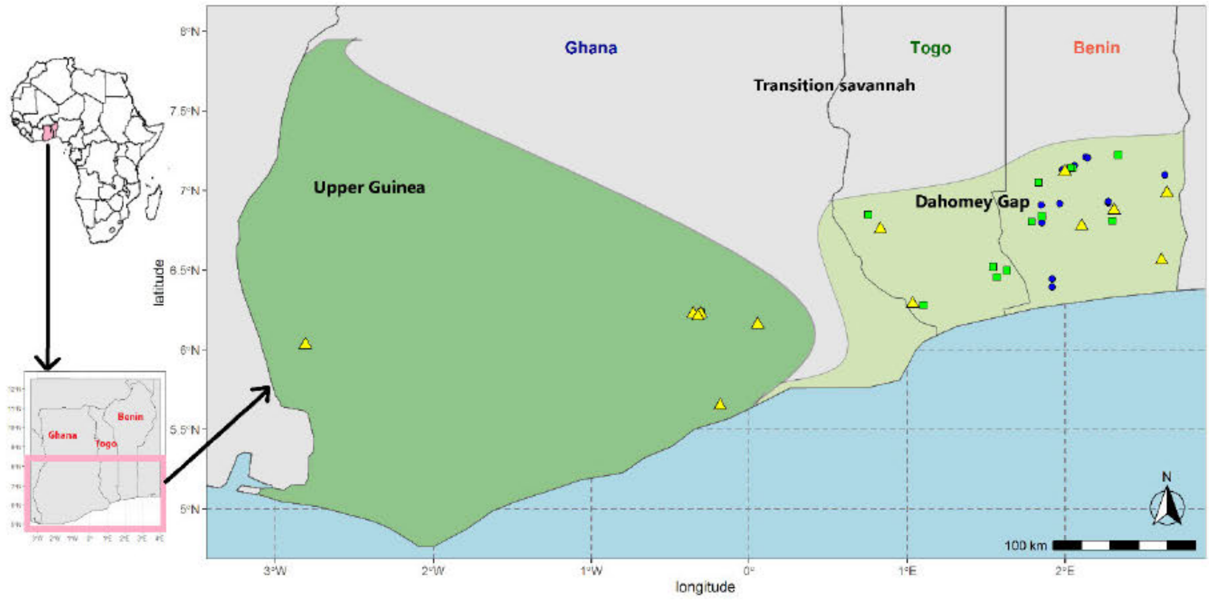


Figure 4.9. Spatial distribution of the core set (41 accessions: 8 accessions for UG and 33 for DG) extracted from 203 accessions of *Synsepalum dulcificum* in West Africa. Accessions are coloured following their belonging to the three clusters identified in this study: ●: Cluster 1 (least performant accessions with regards to the evaluated traits); 🟩: Cluster 2 (elite accessions for tree-traits); △: Cluster 3 (potential elite accessions for fruit-traits). Map generated using the R environment (https://www.r-project.org/) (Version 3.6.2).

4.4 Discussion

The study reports for the first-time, the phenotypic variation, heritability estimates for functional fruit traits and the development of a core collection for *S. dulcificum* in West Africa.

The miracle plant *S. dulcificum* is known to be an endangered species in Benin where it has been mostly reported in swamp forests (Adomou, 2005). However, from this current study, although not a purely ecological one, it is worth noting that no accession was found in the swamp forest in Benin, but rather in farms and mostly in home gardens. In Togo, another portion of the Dahomey Gap, the species was notably absent from forests, especially from sacred forests where it had a rarity index of 0.98. This trend suggests that the predominant habitat of the species in the Dahomey Gap might have shifted from the natural environment to human-close environment such as home gardens, probably due to anthropogenic pressures such as land clearing for agricultural expansion. This seemed to align with the idea of home gardens as a biodiversity hotspot for endangered species (Salako et al., 2014).

Figure 4.10. Quality attributes of the core collection (Core) extracted form the whole collection (Whole).

Traits	Range		CRi (%)	CV		VRi (%)	Mean		MDi (%)	Variance		VDi (%)	Means comparison (Whole vs Core)
	Whole	Core		Whole	Core		Whole	Core		Whole	Core		
DiamGr	64.50	62.68	97.17	53.34	64.05	120.07	19.12	22.33	14.37	104.17	204.63	49.08	*p = 0.26*
PlHgt	6.64	6.64	100.00	36.59	45.88	125.38	3.47	3.70	6.21	1.613	2.87	43.9	*p = 0.61*
CrDiam	8.05	7.05	87.57	38.2	47.92	125.44	4.23	4.23	0.00	2.61	4.11	36.37	*p = 0.99*
AvFL	13.37	12.83	95.96	9.65	14.06	145.69	18.86	18.95	0.47	3.31	7.1	53.35	*p = 0.89*
AvFW	10.91	10.91	100.00	12.08	18.58	153.80	10.59	10.82	2.12	1.64	4.04	59.46	*p = 0.52*
AvFM	1.60	1.60	100.00	25.68	32.06	124.84	1.09	1.14	4.38	0.078	0.13	41.66	*p = 0.43*
AvSM	0.44	0.43	97.72	20	27.07	135.35	0.35	0.35	0.00	0.004	0.008	46.7	*p = 0.82*
EdRt	37.76	37.76	100.00	10.7	13.53	126.44	66.99	67.66	0.99	0.59	0.83	38.67	*p = 0.44*
FrSh	1.33	1.33	100.00	10.05	15.07	149.95	1.79	1.78	0.56	0.03	0.07	53.42	*p = 0.97*
			97.6			*134.1*			*3.23*			*46.95*	-

CV: Coefficient of variation, **CRi**: Coincidance range for the trait i**; VRi**: variable range for the trait i; **VDi**: Variance difference for the trait i; **MDi**: Mean difference percentage for the trait i.

In the Upper Guinea forest, the current absence of the species from forests might not only illustrate an in-situ anthropogenic pressure on the resource but could also reflect a land conversion scenario in which forests were degraded in favour of tree plantation systems. This hypothesis is supported by previous reports indicating that the Upper Guinea forest has been experiencing serious degradations, which have already resulted in the loss of nearly 80% of the original forest cover (van Rompaey, 1993). Almost all the farms in which the species was recorded in the Upper Guinea forest zone were characterized by a tree plantation system where *S. dulcificum* was associated with cocoa (*Theobroma cacao* L.), rubber tree [*Hevea brasiliensis* (Willd. ex A.Juss.) Müll.Arg.], orange tree (*Citrus* spp) or *Musa* spp.

To date, no information exists on the variation of quantitative traits in *S. dulcificum*. In this study, using the populations sampled from the species' centre of origin, trends of traits variations were similar to those observed in sister Sapotaceae species such as *V. paradoxa* (Sanou et al., 2006) from tropical Africa and *Argania spinosa* (L.) Skeels (Metougui et al., 2017) from Northern Africa. However, the magnitude of the variation observed in *S. dulcificum* is greater than in *V. paradoxa* and *A. spinosa* for tree-traits and lower in *S dulcificum* than in *V. paradoxa* and *A. spinosa* for fruit-traits. Considering tree-traits variation for instance, the difference in magnitude between *V. paradoxa* and *S. dulcificum* might be attributed to the degree of heterogeneity of the habitats of each species. Fundamentally, *V. paradoxa* is found in similar habitat-types (i.e., wild and parkland), whereas *S. dulcificum* is found in more diversified habitats including farms and home gardens where it undergoes different management intensities. In home gardens, *S. dulcificum* benefits from management practices such as watering and occasional fertilization compared to farms where there is no special management provided to the species. This is likely to increase the magnitude of the differences observed in tree-traits between these two habitats. Information on tree age is rarely reported in studies tackling phenotypic diversity in perennial fruit trees (Abdulai et al., 2017; Omondi et al., 2019; Tsobeng et al., 2020). In this study, the age estimates for *S. dulcificum* confirmed the prominence of long-living species in the Sapotaceae family (Karambiri et al., 2016). The absence of differences of age between home garden trees and those found on farms ruled out the hypothesis that tree age is a driver of tree-traits variation between these two habitats in *S. dulcificum*. This is furthermore supported by the overall poor correlation between tree age estimate and tree-traits.

This study provided for the first-time information on *S. dulcificum* tree and crown diameter variation and expanded knowledge on the height growth potential of the species. *Synsepalum dulcificum* is a long-living microphanerophyte (Adomou, 2005) whose maximum height was reported to be around four meters (Achigan-Dako et al., 2015). However, in favourable conditions this height may reach up to seven meters and the crown diameter up to eleven meters. Such phenotypes are found mainly in home gardens since individuals on farms are prone to high human pressures including harvesting of leaves and branches as well as debarking of stem for medicinal purposes (Fandohan et al., 2017), thus impeding their normal growth.

Ayensu (1972) successfully described the anatomy and morphology of *S. dulcificum* but made no mention of the fruit and seed morphometry. In an attempt to fill this gap, Lim (2013) in his botanical description of the species, indicated the miracle fruit being roughly 1.0 cm wide for a length ranging from 2.0 to 2.5 cm. Nevertheless, this current study revealed greater variations for both fruit length (1.30 – 2.64 cm) and fruit width (0.73 -1.83 cm). Given the significant effect the provenance population exhibited on fruit-traits in this study, differences with Lim (2013)'s findings could be attributed to environmental variation. On the other hand, a difference in sampling effort coupled with potential genetic variation among the individuals measured might also explain the differences observed in the reported fruit width and length. While Lim (2013) did not indicate the number of fruits used to compute the statistics, in this study nearly 8,875 fruits were measured.

Due to the increasing interest in *S. dulcificum,* this study purposively targeted the fruit and its components. For instance, while the edible ratio is of great interest in medicinal, pharmaceutical, food and beverage industries (Tafazoli et al., 2019), the seed mass is of interest in pharmaceutical and cosmetics (Del Campo et al., 2017; Gorin et al., 2018; Huang et al., 2020). Consequently, knowledge of the tree-to tree or population-to population variation for these specific traits would be useful for elite tree selection by the breeders and decision making for investment by stakeholders. This study reported for the first time the extent of the edible ratio and seed mass variation in the miracle berry and showed the influence of the ecological conditions on the fruit traits in general. An average edible ratio of roughly 67% was observed in *S. dulcificum,* which is one of the highest when compared to other tropical Africa indigenous fruit species such as *V. paradoxa* (61.33% - 62.0% of edible ratio) (Gwali et al., 2012; Sanou et al., 2006), *Balanites aegyptica* (L.) Delile. (44.9 - 50.65%) (Ahmed et al., 2020), and subtropical species such as *A. trifoliata* (23.52-27.63%) (Zou et al., 2019). More importantly, under favourable conditions, the edible ratio in *S. dulcificum* can increase up to 72% as observed for the set of individuals in the Upper Guinea forest block, a region characterized by higher rainfall and less sunlight exposure. This region was also favourable for heavier fruit production than the Dahomey Gap, which was characterized by drier conditions with savannah as main vegetation type. This strengthened arguments in favour of the previously highlighted beneficial effect of watering as well as the detrimental effect of sunlight exposure on *S. dulcificum* growth, reproductive and fruiting performances (Tchokponhoué et al., 2017; Tchokponhoué et al., 2019a). The influence of climatic conditions (e.g., rainfall) on fruit traits in perennial species was also previously reported in *Afzelia Africana* Smith (Houehanou et al., 2019) in which individuals in wet conditions performed better than those in dry conditions.

The poor correlation observed in this study between *S. dulcificum* tree-traits and fruit-traits is in line with findings in other African indigenous fruit trees species (e.g., *V. paradoxa* (Gwali et al., 2012; Sanou et al., 2006), *Sclerocaria birrea* (A. Rich) Hocchst. (Gouwakinnou et al., 2011)), and suggests that tree architecture cannot be used to predict fruiting performance in adult individuals of *S. dulcificum*. Depending on the habitat, different pressures may interfere with the standard species growth, thus changing the trajectory of the relationship between tree morphology traits and any other group of traits. On the contrary, the significant and positive correlation between fruit mass and fruit width, and between fruit mass and the edible ratio, combined with the negative correlation between the edible ratio and seed mass are of significant interest for the development of a selection strategy in breeding of improved varieties of *S. dulcificum*. Such associations make it possible to simultaneously improve the fruit size and the edible ratio while improving the fruit mass. This could translate into an increased benefit and profitability for the food and beverages industries, which are more interested in the edible part of the fruit. On the other hand, it might be difficult to effectively select for seed mass using the fruit mass to benefit the cosmetics industry, thus a direct selection for high seed mass yielding genotypes would be required. Selection of dual-purpose miracle berry genotypes (e.g. for simultaneous use in both cosmetics and food and beverages industries) could be explored by developing a selection index that assigns weights of economic value for each trait (seed mass and edible ratio) (Cotterill and Dean, 1990). More importantly, the influence of the ecological conditions on the relationships between most of the functional fruit traits in *S. dulcificum* suggests the possibility to conduct environment-specific traits selection and production. It further suggests the need to involve multi-environments or locations in evaluation and development of improved varieties of *S. dulcificum*.

Genetic gain is fundamental in plant breeding program and is a function of the heritability, the phenotypic variation and the selection intensity (Zobel and Talbert, 1984). A joint analysis of the variability (variance) and the heritability (coefficient of repeatability) of the six functional fruit-traits in this study indicated that under similar selection pressure, a greater genetic gain may be achieved in fruit mass, seed mass and fruit width against a lower gain in the edible ratio. In fact, by nature, breeding for traits expressed in a ratio is considered more laborious than simple traits (Zou et al., 2019), and this is well illustrated by the edible ratio of the miracle fruit that has an extremely low repeatability estimate (R = 0.015). On the contrary, this trait seemed to be more heritable in other fruit tree species, such as *A. trifoliate* (Zou et al., 2018) where the repeatability estimate for edible ratio was as high as 0.98. This suggests a potential species-specific pattern in the heritability of this important trait. Fruit mass and seed mass had high heritability estimates, which suggests that these traits could be easily improved in the miracle berry once suitable testing environments are selected. While those traits were also found to be highly heritable in *A. trifoliate* (Zou et al., 2018), they were less heritable in *V. paradoxa* (Sanou et al., 2006) and *Allanblackia floribunda* Oliver (Atangana et al., 2011) among other indigenous fruit tree species.

Whether high heritability combined with large phenotypic variability can ensure consequent genetic gain, a good baseline breeding population is desired to accelerate genetic gain (Shelbourne, 1992). The development of a performing breeding population in tree breeding first relies on elite tree identification (Leakey and Page, 2006). Findings from this study highlighted the existence of a diversity of breeding population pools in the miracle plant that can be tailored to specific breeding objectives. Fundamentally, it emerged that the Upper Guinea (UG) represented a source of promising parental lines that can be immediately exploited to initiate a breeding program in the miracle plant.

A core collection was developed, which in terms of size is approximately 20% of the size of the whole collection evaluated. This proportion is in line with the recommended size for a good core collection (10-30% of the initial collection) (Bhattacharjee et al., 2007; Escribano et al., 2007). Historically, four parameters have been used to evaluate the internal quality of a core collection: the coincidence range (CR), the mean difference percentage (MD), the variance difference (VD) and the variable rate (VR) (Hu et al., 2000). In this study, the core collection exhibited a coincidence rate CR that is by far higher than the threshold of CR > 80% required for a core collection to optimally represent an initial whole collection. While it was indicated that a well-constructed core collection should have no more than 20% of the traits deviating significantly from the whole collection, none of the traits in the core collection significantly deviated from the whole collection. The MD was also extremely low, suggesting that the core is a good representative of the whole set of accessions. Most important, the core collection encompassed individuals from all the sampled populations in this study, and all the clusters identified with a quasi-similar representativeness. Large VD and VR are preferred for a good collection (Mahmoodi et al., 2019) as they illustrate how much diversity of the whole collection is maintained in the core collection. An exceptionally high variable rate (> 100%) and a VD of 46.95% were observed, which together suggests that further diversity can be obtained from the core collection compared to the whole collection. Combined, these characteristics suggest that the regional core collection compiled is robust enough to represent the whole collection and can consequently be considered as a reliable working sample in implementing ex-situ conservation measures and developing a breeding population for association study and further genomic selection implementation (Duan et al., 2017) in the species. Practically, this core collection will ease and accelerate further evaluations (multi-location trials, genomic evaluation, metabolomic evaluation) and consequently shorten time to cultivar release. Besides, this core collection to be maintained as a living collection will help secure the reservoir of useful alleles the breeding process might need to tap in the future while favouring their co-evolution with environmental factors. An analysis of the core collection constitution indicated that all the three clusters identified in this study were represented in the Dahomey Gap zone whereas mainly one cluster was present in the Upper Guinea zone. This suggests the Dahomey Gap to be a centre of *S. dulcificum* diversity while providing an opportunity to properly elucidate phylogeographical relationships of the Dahomey Gap populations with the Upper Guinea, the Lower Guinea and the Congolian populations of the species.

The geographical scope of this study did not permit to integrate the potential variability in the yellow morphotype, which is known to be found only in Nigeria (Njoku et al., 2015). Consequently, further research is needed at a broader scale, involving at least Eastern Nigeria, to depict the full picture of the phenotypic diversity of the species at the West African level.

4.5 Conclusion

This study evaluated the phenotypic variation in the miracle plant *S. dulcificum*, from its centre of origin, West Africa. A prominent effect of the provenance population was found on most of the phenotypic traits, with a better performance from the Upper Guinea forest accessions compared to their counterparts from the Dahomey Gap. Tree-traits were more variable than fruit-traits. Traits such as fruit mass, edible ratio and fruit width were highly correlated and exhibited high heritability estimates implying that simultaneous selection of the traits can be done. In addition, *S. dulcificum* individuals likely to serve as elite parental lines in the selection and development of high-yielding and pulp-dense cultivars were also identified. Finally, the successful development of a core collection of 41 accessions from the initial set of 203 accessions suggested the Dahomey Gap as a centre of diversity of the species. These findings pave the way for focused ex-situ conservation measures implementation as well as a cost-effective breeding program design for the species.

4.6 Reference

Abdulai, I., Krutovsky, K. V., and Finkeldey, R. (2017). Morphological and genetic diversity of shea trea (*Vitellaria paradoxa*) in the savannah regions of Ghana. *Genet. Res. Crop. Evol.* **64**, 1253-1268.

Achigan-Dako, E. G., Tchokponhoué, D. A., N'Danikou, S., Gebauer, J., and Vodouhè, R. S. (2015). Current knowledge and breeding perspectives for the miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell. *Genet. Resour. Crop Evol.* **62**, 465-476.

Adomou, A. (2005). Vegetation patterns and environmental gradients in Benin, University of Wageningen, Wageningen, Netherlands.

Ahmed, A. A. O., Kita, A., Nemś, A., Miedzianka, J., Foligni, R., Abdalla, A. M. A., and Mozzon, M. (2020). Tree-to-tree variability in fruits and kernels of a *Balanites aegyptiaca* (L.) Del. population grown in Sudan. *Trees* **34**, 111–119.

Albuquerque, A. S., Bruckner, C. H., Cruz, C. D., Salomão, L. C. C., and Neves, J. C. L. (2004). Repeatability and correlations among peach physical traits. *Crop Breed. Appl. Biot.* **4**, 441-445.

Andrade, A. C., Martins, M. B., Rodrigues, J. F., Coelho, S. B., Pinheiro, A. C. M., and Bastos, S. C. (2019). Effect of different quantities of miracle fruit on sour and bitter beverages. *LWT* **99**, 89-97.

Atangana, A. R., van der Vlis, E., Khasa, D. P., van Houten, D., Beaulieu, J., and Hendrickx, H. (2011). Tree-to-tree variation in stearic and oleic acid content in seed fat from *Allanblackia floribunda* from wild stands: potential for tree breeding. *Food Chem.* **126**, 1579-1585.

Ayensu, E. S. (1972). Morphology and anatomy of *Synsepalum dulcificum* (Sapotaceae). *Bot. J. Linn. Soc.* **65**, 179-187.

Belaj, A., del Carmen Dominguez-García, M., Atienza, S. G., Urdíroz, N. M., De la Rosa, R., Satovic, Z., Martín, A., Kilian, A., Trujillo, I., and Valpuesta, V. (2012). Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DArTs, SSRs, SNPs) and agronomic traits. *Tree Genet. Genomes* **8**, 365-378.

Bharuth, V., Naidoo, C., Pammenter, N. W., Lamb, J. M., and Moodley, T. (2020). Responses to chilling of recalcitrant seeds of *Ekebergia capensis* from different provenances. *S. Afr. J. Bot.* **130**, 8-24.

Bhattacharjee, R., Khairwal, I., Bramel, P. J., and Reddy, K. (2007). Establishment of a pearl millet [*Pennisetum glaucum* (L.) R. Br.] core collection based on geographical distribution and quantitative traits. *Euphytica* **155**, 35-45.

Booth, A. (1958). The Niger, the Volta and the Dahomey Gap as geographic barriers. *Evolution* **12**, 48-62.

Buckmire, R., and Francis, F. (1978). Pigments of miracle fruit, *Synsepalum dulcificum*, Schum, as potential food colorants. *J. Food Sci.* **43**, 908-911.

Chen, C. C., Liu, I. M., and Cheng, J. T. (2006). Improvement of insulin resistance by miracle fruit (S*ynsepalum dulcificum*) in fructose-rich chow-fed rats. *Phytother. Res.* **20**, 987-992.

Chibuzor, I. A., Bukola, O., Adejoke, A. O., and Chidozie, O. P. (2017). Genetic assessment of the shrub *Synsepalum dulcificum* (Schumach & Thonn.) Daniell in Nigeria using the Randomly Amplified Polymorphic DNA (RAPD). *Int. J. Genet. Genomics* **4**, 45-50.

Choi, S. E., and Garza, J. (2020). Effects of different miracle fruit products on the sensory characteristics of different types of sour foods by descriptive analysis. *J. Food Sci.* **85**, 36-49.

Cotterill, P. P., and Dean, C. A. (1990). "Successful tree breeding with index selection," CSIRO, Division of Forestry and Forest Products.

De Beukelaer, H., Davenport, G. F., and Fack, V. (2018). Core Hunter 3: flexible core subset selection. *BMC Bioinformatics* **19:203**, 1-12.

Del Campo, R., Zhang, Y., and Wakeford, C. (2017). Effect of miracle fruit (*Synsepalum dulcificum*) seed oil (MFSO®) on the measurable improvement of hair breakage in women with damaged hair: a randomized, double-blind, placebo-controlled, eight-month trial. *J. Clin. Aesthet. Dermat.* **10**, 39-48.

Duan, H., Cao, S., Zheng, H., Hu, D., Lin, J., Cui, B., Lin, H., Hu, R., Wu, B., and Sun, Y. (2017). Genetic characterization of Chinese fir from six provinces in southern China and construction of a core collection. *Sci. Rep.* **7**, 1-10.

Edesi, J., Tolonen, J., Ruotsalainen, A. L., Aspi, J., and Häggman, H. (2020). Cryopreservation enables long-term conservation of critically endangered species *Rubus humulifolius*. *Biodivers. Conserv.* **29**, 303-314.

Escribano, P., Viruel, M., and Hormaza, J. (2007). Establishment of a core collection to optimise the conservation of cherimoya (*Annona cherimola* Mill.) genetic resources using SSR information. *In* "XII EUCARPIA Symposium on Fruit Breeding and Genetics 814", pp. 67-70.

Falconer, D. S. (1960). "Introduction to quantitative genetics," Oliver And Boyd; Edinburgh; London.

Fandohan, A. B., Gouwakinnou, G. N., Tovissode, C. F., Bonou, A., Djonlonkou, S. F. B., Houndelo, L. F., Sinsin, C. L. B., and Assogbadjo, A. E. (2017). Usages traditionnels et valeur économique de *Synsepalum dulcificum* au Sud-Bénin. *Bois For. Trop.* **332**, 17-30.

Fereidoonfar, H., Salehi-Arjmand, H., Khadivi, A., and Akramian, M. (2018). Morphological variability of sumac (*Rhus coriaria* L.) germplasm using multivariate analysis. *Ind. Crops Prod.* **120**, 162-170.

Fonseca, C. E. L. d., Morais, F. M. d., Gonçalves, H. M., Aquino, F. d. G., and Rocha, F. S. (2018). Repeatability of fruit traits from two *Hancornia speciosa* populations from the core region of the Brazilian Cerrado. *Pesqui. Agropecu Bras.* **53**, 710-716.

Gorin, S., Wakeford, C., Zhang, G., Sukamtoh, E., Matteliano, C. J., and Finch, A. E. (2018). Beneficial effects of an investigational wristband containing *Synsepalum dulcificum* (miracle fruit) seed oil on the performance of hand and finger motor skills in healthy subjects: A randomized controlled preliminary study. *Phytother. Res.* **32**, 321-332.

Gouwakinnou, G. N., Assogbadjo, A. E., Lykke, A. M., and Sinsin, B. (2011). Phenotypic variations in fruits and selection potential in *Sclerocarya birre*a subsp. birrea. *Sci. Hortic. Amsterdam* **129**, 777-783.

Gwali, S., Nakabonge, G., Okullo, J. B. L., Eilu, G., Nyeko, P., and Vuzi, P. (2012a). Morphological variation among shea tree (*Vitellaria paradoxa* subsp. *nilotica*)'ethnovarieties' in Uganda. *Genet. Resour. Crop. Evol.* **59**, 1883-1898.

Gwali, S., Nakabonge, G., Okullo, J. B. L., Eilu, G., Nyeko, P., and Vuzi, P. (2012b). Morphological variation among shea tree (*Vitellaria paradoxa* subsp. *nilotica*)'ethnovarieties' in Uganda. *Genet. Resour. Crop sEvol.* **59**, 1883-1898.

Han, Y. C., Wu, J. Y., and Wang, C. K. (2019). Modulatory effects of miracle fruit ethanolic extracts on glucose uptake through the insulin signaling pathway in C2C12 mouse myotubes cells. *Food Sci. Nutr.* **7**, 1035-1042.

Houehanou, T. D., Prinz, K., Hellwig, F., Assogbadjo, A. E., Gebauer, J., Kakaï, R. L. G., and Sinsin, B. (2019). Morphological trait variation and relationships of *Afzelia africana* Sm. caused by climatic conditions and anthropogenic disturbance in Benin (West Africa). *Genet. Resour. Crop Evol.* **66**, 1091-1105.

Hu, J., Zhu, J., and Xu, H. (2000). Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theor. Appl. Genet.* **101**, 264-268.

Huang, W., Chung, H. Y., Xuan, W., Wang, G., and Li, Y. (2020). The cholesterol-lowering activity of miracle fruit (Synsepalum dulcificum). *J. Food Biochem.* **44**, e13185.

Inglett, G., and May, J. F. (1968). Tropical plants with unusual taste properties. *Econ. Bot.* **22**, 326-331.

Juhé-Beaulaton, D. (2014). Le fruit miracle (*Synsepalum dulcificum*): des voyageurs sur la côte ouest africaine aux laboratoires pharmaceutiques. *In* "ATM Collections de Muséums sources d'histoire".

Karambiri, M., Elias, M., Vinceti, B., and Grosse, A. (2016). Exploring local knowledge and preferences for shea (*Vitellaria pradoxa*) ethnovarieties in Soutwest Burkina-Faso through a gender and ethnic lens. *For. Trees Livelihoods* **26**, 13-28.

Kassambara, A., and Mundt, F. (2017). Factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0. 4.

Khadivi-Khub, A., and Anjam, K. (2014). Morphological characterization of *Prunus scoparia* using multivariate analysis. *Plant Syst. Evol.* **300**, 1361-1372.

Kim, K.-W., Chung, H.-K., Cho, G.-T., Ma, K.-H., Chandrabalan, D., Gwag, J.-G., Kim, T.-S., Cho, E.-G., and Park, Y.-J. (2007). PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* **23**, 2155-2162.

Kurihara, K., and Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science* **161**, 1241-1243.

Le Cunff, L., Fournier-Level, A., Laucou, V., Vezzulli, S., Lacombe, T., Adam-Blondon, A.-F., Boursiquot, J.-M., and This, P. (2008). Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. *subsp. sativa*. *BMC Plt. Biol.* **8:31**, 1-12.

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1-18.

Leakey, R., and Page, T. (2006). The ‘ideotype concept’and its application to the selection of cultivars of trees providing agroforestry tree products. *For. Trees Livelihoods* **16**, 5-16.

Leal, M. E. (2004). The African rain forest during the Last Glacial Maximum an archipelago of forests in a sea of grass, Wageningen University.

Lim, T. (2013). *Synsepalum dulcificum*. *In* "Edible medicinal and non-medicinal plants" (T. Lim, ed.), Vol. 6, pp. 146-150. Springer, Dordrecht.

Mahmoodi, R., Dadpour, M. R., Hassani, D., Zeinalabedini, M., Vendramin, E., Micali, S., and Nahandi, F. Z. (2019). Development of a core collection in Iranian walnut (*Juglans regia L.*) germplasm using the phenotypic diversity. *Sci. Hortic. Amsterdam* **249**, 439-448.

Martínez-García, P. J., Famula, R. A., Leslie, C., McGranahan, G. H., Famula, T. R., and Neale, D. B. (2017). Predicting breeding values and genetic components using generalized linear mixed models for categorical and continuous traits in walnut (*Juglans regia*). *Tree Genet. Genomes* **13**, 109.

Metougui, M. L., Mokhtari, M., Maughan, P. J., Jellen, E. N., and Benlhabib, O. (2017). Morphological variability, heritability and correlation studies within an argan tree population (*Argania spinosa* (L.) Skeels) preserved in situ. *Int. J. Agric. For.* **7**, 42-51.

Moon, K.-W. (2016). Interactive plot. *In* "Learn ggplot2 Using Shiny App" (K.-W. Moon, ed.), pp. 295-347. Springer, Switzerland.

Niñez, V. (1987). Household gardens: Theoretical and policy considerations. *Agr. Syst.* **23**, 167-186.

Norouzi, E., Erfani-Moghadam, J., Fazeli, A., and Khadivi, A. (2017). Morphological variability within and among three species of *Ziziphus* genus using multivariate analysis. *Sci. Hortic. Amsterdam* **222**, 180-186.

Obafemi, T., Akinmoladun, A., Olaleye, M., Agboade, S. O., and Onasanya, A. A. (2017). Antidiabetic potential of methanolic and flavonoid-rich leaf extracts of *Synsepalum dulcificum* in type 2 diabetic rats. *J-AIM* **8**, 238-246.

Omondi, M., Rimberia, F. K., Wainaina, C. M., Mukundi, J. B. N., Orina, J., Gebauer, J., and Kehlenbeck, K. (2019). Fruit morphological diversity and productivity of baobab (*Adansonia digitata* L.) in coastal and lower eastern Kenya. *For. Trees Livelihoods* **28**, 266-280.

Ouborg, N. J. (2010). Integrating population genetics and conservation biology in the era of genomics. *Biol. Lett.* **6**, 3-6.

Oumorou, M., Dah-Dovonon, J., Aboh, B., Hounsoukaka, M., and Sinsin, B. (2010). Contribution á la conservation de *Synsepalum dulcificum*: régénération et importance socio-économique dans le département de l'ouémé (Bénin). *Ann. Sci. Agron* **14**, 101-120.

Pebesma, E. (2018). Simple features for R: standardized support for spatial vector data. *R J.* **10**, 439-446.

Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., Zivot, E., Cornilly, D., Hung, E., Lestel, M., and Balkissoon, K. (2018). 'PerformanceAnalytics': Econometric Tools for Performance and Risk Analysis. R Team Cooperation.

R Core Team (2019). "R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,Vienna, Austria. URL https://www.R-project.org/," Vienna, Austria.

Rodrigues, J. F., da Silva Andrade, R., Bastos, S. C., Coelho, S. B., and Pinheiro, A. C. M. (2016). Miracle fruit: An alternative sugar substitute in sour beverages. *Appetite* **107**, 645-653.

Salako, V. K., Fandohan, B., Kassa, B. B., Assogbadjo, A. E., Idohou, A. F. R., Gbedomon, R. C., Chakeredza, C., Dulloo, M. E., and Glele-Kakaı, R. (2014). Home gardens: an assessment of their biodiversity and potential contribution to conservation of threatened species and crop wild relatives in Benin. *Genet. Resour. Crop. Evol.* **61**, 313.

Salzmann, U., and Hoelzmann, P. (2005). The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *Holocene* **15**, 190-199.

Sanou, H., Picard, N., Lovett, P. N., Dembélé, M., Korbo, A., Diarisso, D., and Bouvet, J.-M. (2006). Phenotypic variation of agromorphological traits of the shea tree, *Vitellaria paradoxa* CF Gaertn., in Mali. *Genet. Resour. Crop Evol.* **53**, 145-161.

Shelbourne, C. (1992). Genetic gains from different kinds of breeding population and seed or plant production population. *S. Afr. For. J.* **160**, 49-65.

Singh, K., Sharma, Y. P., and Gairola, S. (2020). Morphological characterization of wild Rosa L. germplasm from the Western Himalaya, India. *Euphytica* **216:41**, 1-24.

Sogbohossou, E. D., Kortekaas, D., Achigan-Dako, E. G., Maundu, P., Stoilova, T., Van Deynze, A., de Vos, R. C., and Schranz, M. E. (2019). Association between vitamin content, plant morphology and geographical origin in a worldwide collection of the orphan crop *Gynandropsis gynandra* (Cleomaceae). *Planta* **250**, 933-947.

South, A. (2012). Rworldxtra: Country boundaries at high resolution. *R package version 1*.

South, A. (2017). rnaturalearth: World map data from Natural Earth. *R package version 0.1. 0*.

Stoffel, M. A., Nakagawa, S., and Schielzeth, H. (2017). rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Meth. Ecol. Evol.* **8**, 1639-1644.

Sun, W., Yuan, X., Liu, Z.-J., Lan, S., Tsai, W.-c., and Zou, S.-Q. (2019). Multivariate analysis reveals phenotypic diversity of *Euscaphis japonica* population. *PLoS ONE* **14**.

Swamy, K. B., Hadi, S. A., Sekaran, M., and Pichika, M. R. (2014). The clinical effects of Synsepalum dulcificum: a review. *J. Med. Food.* **17**, 1165-1169.

Swenson, U., Richardson, J. E., and Bartish, I. V. (2008). Multi-gene phylogeny of the pantropical subfamily Chrysophylloideae (Sapotaceae): evidence of generic polyphyly and extensive morphological homoplasy. *Cladistics* **24**, 1006–1031.

Tafazoli, S., Vo, T. D., Roberts, A., Rodriguez, C., Viñas, R., Madonna, M. E., Chiang, Y.-H., Noronha, J. W., Holguin, J. C., and Ryder, J. A. (2019). Safety assessment of miraculin using in silico and in vitro digestibility analyses. *Food Chem. Toxicol.* **133**, 1-10.

Tchokponhoué, D., Achigan-Dako, E., N'Danikou, S., Houdégbé, A., Agossou, C., Assogba-Komlan, F., and Vodouhè, R. (2018). Regeneration ability and seedling growth in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Fruits* **73**, 13-21.

Tchokponhoué, D. A., N'Danikou, S., Hale, I., Van Deynze, A., and Achigan-Dako, E. G. (2017). Early fruiting in *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell juveniles induced by water and inorganic nutrient management. *F1000Research* **6**.

Tchokponhoué, D. A., N'Danikou, S., and Achigan-Dako, E. G. (2019a). A combination of approaches evidenced seed storage behaviour in the miracle berry *Synsepalum dulcificum* (Schumach. et Thonn.) Daniell. *BMC Plt. Biol.* **19:117**, 1-13.

Tchokponhoué, D. A., N'Danikou, S., Houéto, J. S., and Achigan-Dako, E. G. (2019b). Shade and nutrient-mediated phenotypic plasticity in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Sci Rep.* **9:5137**, 1-11.

Tsobeng, A., Akem, M., M-L., A., Muchugi, A., Degrande, A., Tchoundjeu, Z., Jamnadass, R., and Na'a, F. (2020). Tree-to-tree variation in fruit of three populations of *Trichoscypha acuminata* in Cameroon. *Sci. Afr.* **7**, 1-12.

van Rompaey, R. S. (1993). Forest gradients in West Africa: a spatial gradient analysis. PhD thesis, Wageningen.

Vihotogbé, R., van den Berg, R. G., and Sosef, M. S. (2013). Morphological characterization of African bush mango trees (Irvingia species) in West Africa. *Genet. Resour. Crop Evol.* **60**, 1597-1614.

Warton, D. I., Duursma, R. A., Falster, D. S., and Taskinen, S. (2012). smatr 3–an R package for estimation and inference about allometric lines. *Methods Ecol. Evol.* **3**, 257-259.

White, F. (1983). The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa by F White. *Natural Resources Research Report XX, UNESCO, Paris, France*, 1876-1895.

Wickham, H. (2016). "ggplot2: elegant graphics for data analysis," Springer.

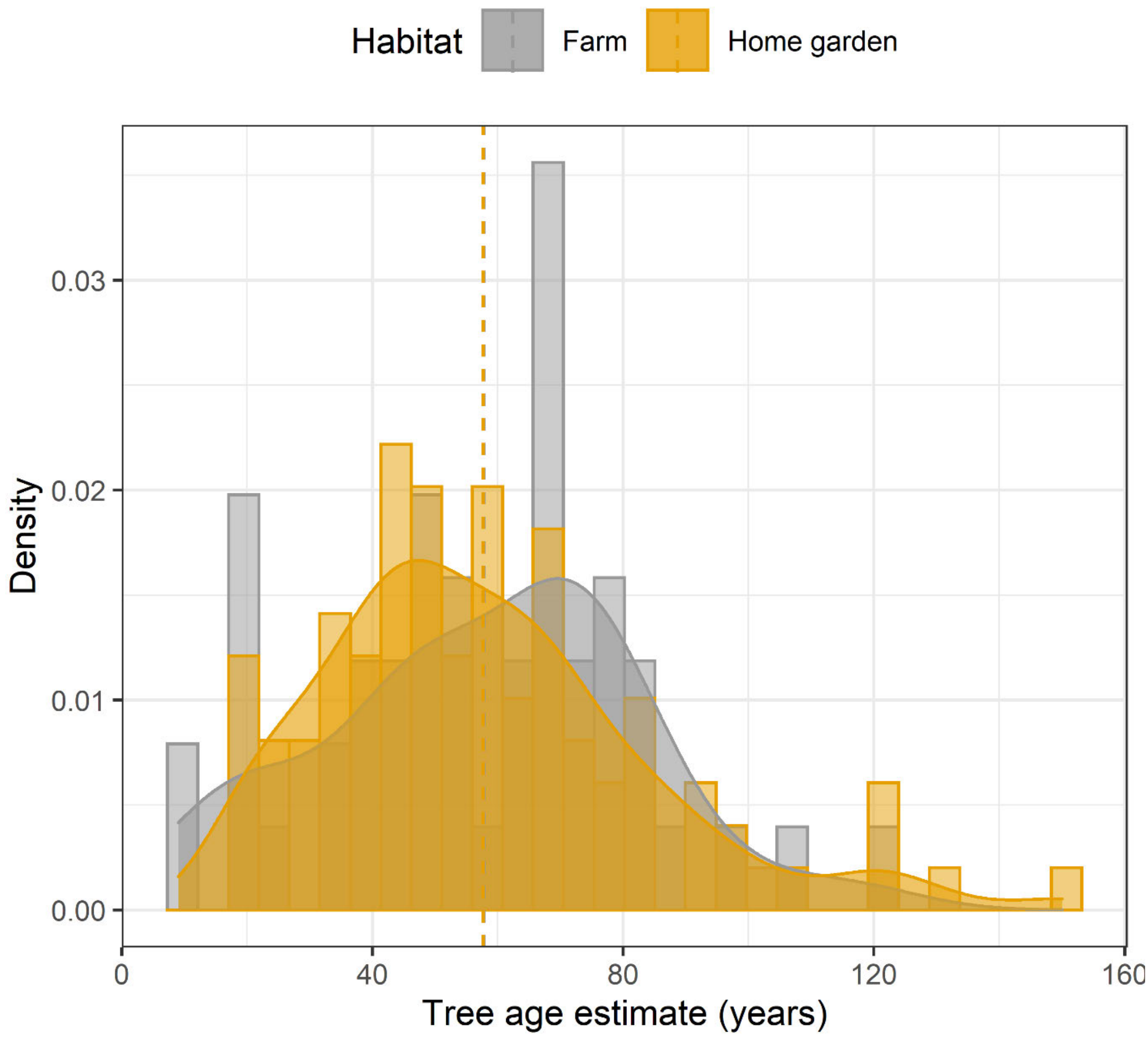
Yazdanpour, F., Khadivi, A., and Etemadi-Khah, A. (2018). Phenotypic characterization of black raspberry to select the promising genotypes. *Sci. Hortic. Amsterdam* **235**, 95-105.

YiLan, L., and RuTong, Z. (2015). clustertend: Check the Clustering tendency. *R package version* **1**.

Zobel, B., and Talbert, J. (1984). "Applied forest tree improvement," John Wiley & Sons.

Zou, S., Yao, X., Zhong, C., Zhao, T., and Huang, H. (2018). Genetic analysis of fruit traits and selection of superior clonal lines in *Akebia trifoliate* (Lardizabalaceae). *Euphytica* **214:111**, 1-9.

Zou, S., Yao, X., Zhong, C., Zhao, T., and Huang, H. (2019). Effectiveness of recurrent selection in *Akebia trifoliata* (Lardizabalaceae) breeding. *Sci. Hortic. Amsterdam* **246**, 79-85.

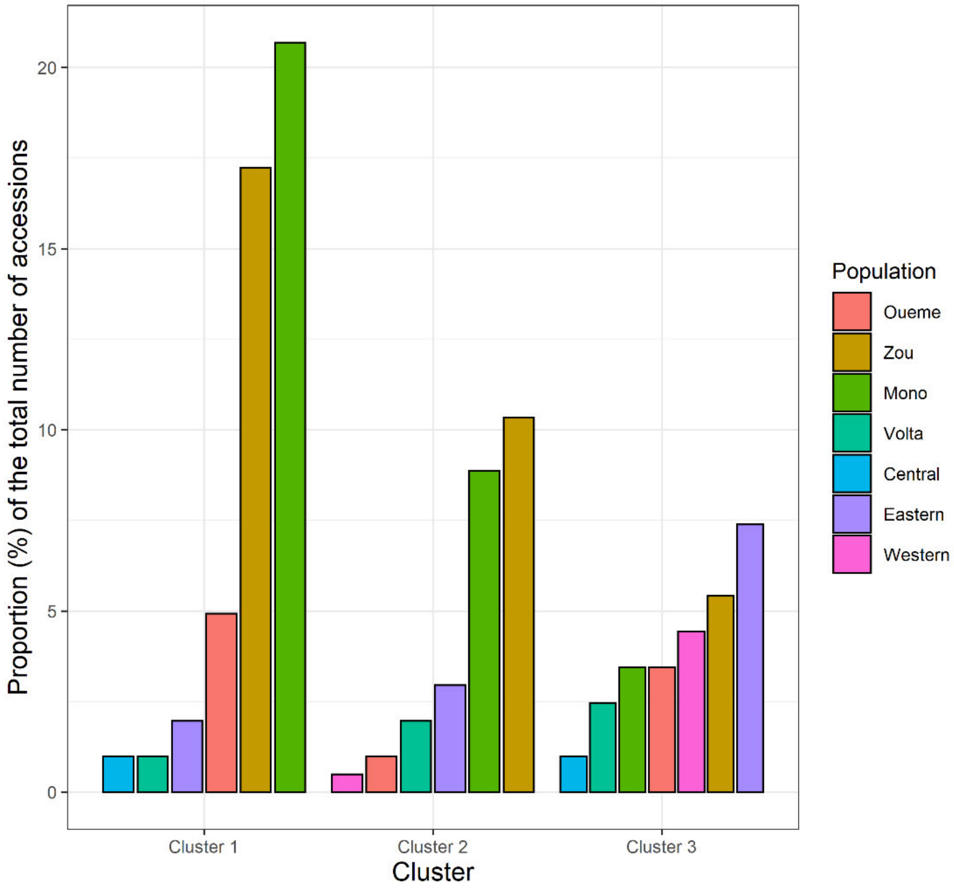


Appendix 4.1. Variation of tree age estimate for *Synsepalum dulcificum* trees in home garden and on farm. Coloured vertical dashed lines (here almost confounded) represented mean age estimate following each habitat type.

Appendix 4.2. Spearman correlation between tree age estimate and tree- and fruit-traits in *Synsepalum dulcificum*

Tree-and fruit-traits	Tree age estimate
Tree diameter at ground level (cm)	0.53***
Tree height (m)	0.1ns
Tree crown diameter (m)	0.3***
Fruit length (mm)	-0.07ns
Fruit width (mm)	-0.21**
Individual fruit mass (g)	-0.10ns
Individual seed mass (g)	-0.03ns
Edible ratio (%)	-0.09ns
Fruit shape index	0.17*

***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, ns: Non-significant.



Appendix 4.3. Relative contribution of *Synsepalum dulcificum* provenance populations to various clusters.

CHAPTER 5 [3]

Genome-wide diversity analysis suggests divergence among West African populations of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell], with Western Dahomey Gap as the center of origin

Abstract

Although the miracle plant (*Synsepalum dulcificum*) is viewed as one of the most economically promising orphan berry crops worldwide, its genetic improvement and sustainable conservation are hindered by a lack of understanding of its evolutionary history and current population structure. Here we report, for the first time, the application of genome-wide SNP genotyping to a diverse panel of *S. dulcificum* accessions to specifically depict the genetic diversity and population structure of the species in the Dahomey Gap and Upper Guinea regions in an effort to infer its evolutionary history and origin. Our findings suggest low overall genetic diversity but strong population divergence within the species. STRUCTURE and Principal Coordinate Analysis together distinguished three genetic groups corresponding to the Upper Guinea (**UG**), Western Dahomey Gap (**WDG**) and Central Dahomey Gap (**CDG**) regions. Application of Monmonier's algorithm in the Barriers software reveals the existence of a natural barrier disrupting the connectivity between the Upper Guinea and Dahomey Gap populations. The WDG population consistently exhibited the highest levels of nucleotide and haplotype diversity, while the CDG population exhibited the lowest. Analyses of Tajima's D, Fu's Fs and Achaz Y* statistics, combined with the mismatch distribution curves, suggest that while both UG and CDG populations likely experienced recent expansions, WDG population is at equilibrium. These findings suggest a geographical structuring of genetic variation, which supports the conclusion of the Western Dahomey Gap being the likely centre of origin of *S. dulcificum*. These results provide the most significant insights to guide informed breeding population development and the design of sound conservation strategies for the miracle plant.

---

[3] Chapter **under review** as: **Tchokponhoué, D.A.,** Achigan-Dako, E. G., N'Danikou, S., Nyadanu, D., Hale, I., Odindo, A. O., & Sibiya, J. Genome-wide diversity analysis suggests divergence among West African populations of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell], with Western Dahomey Gap as the centre of origin.

5.1 Introduction

Synsepalum dulcificum (Schumach & Thonn.) Daniell [Syn. *Richardella dulcifica* (Schumach & Thonn.) Baehni], also known as the miracle plant, is a slow-growing evergreen fruit tree species (Tchokponhoué et al., 2018) in the Sapotaceae family. A long living monoecious species that rarely reaches a height exceeding 8 m (Tchokponhoué et al., 2020), *S. dulcificum* produces red berries known as miracle berries or miracle fruits that contain recalcitrant seeds (Tchokponhoué et al., 2019). The miracle plant is the only known natural source of miraculin (Kurihara and Beidler, 1968), a taste-modifying glycoprotein contained in the mucilaginous part of the fruit. Coupled with the novel sweetening property of miraculin, the miracle berry's otherwise high micronutrient and antioxidant content (Du et al., 2014) has led to its use in a wide range of culinary and pharmaceutical applications, the latter including both diabetes and cancer treatments (Buckmire and Francis, 1978; Wilken and Satiroff, 2012). The seed oil improves hand motor skills and prevents hair breakage in women (Del Campo et al., 2017; Gorin et al., 2018), while various other parts of the plant (e.g., root, leaves, and bark) are excellent sources of numerous health-promoting phytochemicals, such as vanillic acid, epi-syringaresinol, β-sitosterol, and lupeol (Chen et al., 2010a, b; Masson, 2014). On the international market, dried miracle fruit fetches an astounding price of USD 2,500 per kg of pure powder (Link 1 , Link 2). Given the diversity of its uses and its high value, *S. dulcificum* stands as one of the most significant and promising berry crops worldwide, even more so in light of its recent formal authorization in the EU market (Turck et al., 2021).

Compared with more heavily invested berry crops, such as strawberry (*Fragaria* × *ananassa*) (Barbey et al., 2020; Denoyes et al., 2016; Nelson et al., 2021), blueberry (*Vaccinium corymbosum*) (Die and Rowland, 2013; Qi et al., 2021), cranberry (*Vaccinium macrocarpon*) (Covarrubias-Pazaran et al., 2016; Diaz-Garcia et al., 2021), and blackberry (*Rubus* spp.) (Foster et al., 2019; Garcia-Seco et al., 2015), all of which benefit from long breeding histories and extensive genomic resources, *S. dulcificum* is an orphan crop with no systematic history of breeding or readily available genomic resources to facilitate its improvement. Even the most basic information on genetic diversity, though fundamental to implementing a sound breeding program, is scanty and limited to a single low-resolution assessment of a geographically narrow collection (Chibuzor et al., 2017). Although three distinct phenotypic groups are reported in the species in West Africa (Tchokponhoué et al., 2020), such groups are yet to be corroborated by molecular evidence. To date, the only reported molecular evaluation in the species (Chibuzor et al., 2017) is based on random amplified polymorphic DNA (RAPD) markers, a platform that today is considered obsolete in light of the advent and now common application of reduced-representation library sequencing approaches, such as Genotyping-by-Sequencing (GBS; (Elshire et al., 2011), restriction site-associated DNA sequencing (Rad-seq; (Miller et al., 2007), Diversity Array Technology Sequencing (DArTSeq; (Sansaloni et al., 2011)), and others. The advantage of such methods lies in their ability to efficiently generate genome wide sets of high-quality single nucleotide polymorphisms (SNPs) suitable for high resolution molecular diversity assessments, especially in crops currently lacking a reference genome (Feng et al., 2019; Melo et al., 2016; Poland and Rife, 2012) such as *S. dulcificum*. More importantly, the RAPD study's conclusions were based on only 40 individuals sampled exclusively in Southern Nigeria, whereas the species exhibits a much larger natural distribution, spanning West and Central Africa. Overall, there is a clear opportunity and need to bring modern sequencing-based assessment methods to a larger sample of the species, not only in terms of geographical coverage but also numbers, in order to achieve the resolution necessary for reliable diversity and population structure assessment.

The miracle plant naturally occurs in the West African rainforest, a dense tropical humid vegetation block that extends, west to east, from Guinea to the Democratic Republic of the Congo (Leal, 2004). The forest vegetation in this region experienced past climatic fluctuations during the Holocene period that ultimately induced its disruption into three ecologically distinct blocks: two forested blocks, the western Upper Guinea and eastern Lower Guinea, and one intervening savannah block known as the Dahomey Gap. Currently perceived as an open savanna, the Dahomey Gap successively experienced: i) a shift from its original semi-evergreen forest state to a Sudano-Guinean savannah zone during the dry Glacial Stages (8,400–4,500 cal. yr BP); ii) forest re-establishment, a consequence of rewetting under the environmental conditions of the late Interglacial and early Holocene (3,300–1300 cal. yr BP); and iii) establishment of the persisting open savannah seen today (Dupont and Weinelt, 1996). It is widely acknowledged that past climatic fluctuations, like those observed in Dahomey Gap, can drive significant changes in plant species distribution, diversification, evolutionary dynamics, and speciation (Demenou et al., 2018; Hagen et al., 2021; Plana et al., 2004; Welker, 2017). However, the extent of such influences has been rarely explored under tropical conditions (Hagen et al., 2021; Teixeira et al., 2021). The current distribution of the miracle plant in these three West African forest blocks offer a unique opportunity to gain insight into the effects of past climate changes on the genetic variation and population reconfiguration of a slow-growing tropical tree species.

By applying genome-wide, reduced representation sequencing to a large and geographically diverse collection of *S. dulcificum* germplasm, this study seeks to understand the basic population structure of the species, particularly as it relates to the Upper Guinea-Dahomey Gap ecological divergence. Beyond this, the objective of the study is to infer the evolutionary history of the species, including its likely center of origin, and shed light on the current patterns of genetic diversity to guide conservation and breeding efforts.

5.2 Materials and methods

5.2.1 Study area

This study specifically focused on the Upper Guinea and Dahomey Gap blocks of the West African rainforest (**Fig. 5.1**), as these two areas were previously suggested as the potential center of origin and diversity of the miracle plant (Swenson et al., 2008; Tchokponhoué et al., 2020). Also known as the Guinean forest block, the Upper Guinea block extends from Guinea to eastern Ghana (Leal, 2004) and is dominated by semi-evergreen and evergreen rainforests experiencing annual rainfalls of 1,600–2,000 mm and 2,000–4,551 mm, respectively (Duminil et al., 2013). The Dahomey Gap is a ca. 200 km broad savannah corridor that was created during the late Holocene, following an abrupt climate change 4500 and 3400 cal. years BP (Salzmann and Hoelzmann, 2005). The annual rainfall in DG is between 1,000–1,200 mm. Indicated as a biogeographical barrier between the Upper Guinea and Lower Guinea blocks, the Dahomey Gap encompasses the coasts of Benin, Togo, and Ghana and is the most significant distributional disjunction for forest taxa in West Africa (Fuchs and Bowie, 2015).

5.2.2 Tree sampling

From September 2016 to August 2020, a total of 322 individual trees (hereafter, accessions) of *S. dulcificum* were sampled from seven populations (**Table 5.1**) spread across the Dahomey Gap and Upper Guinea, all within the political boundaries of the countries of Benin, Togo, and Ghana. A population of miracle plant was defined as a set of individuals possibly interbreeding and randomly distributed within a defined agroecosystem (Tchokponhoué et al., 2020). To be considered distinct, two populations of *S. dulcificum* had to be separated by a minimum distance of 15 km. Two of the seven populations, namely the “Volta” and “Mono” populations, were cross-country populations. The “Volta” population extends from western Togo to the Volta region of Ghana, while the Mono population stretches from western Benin to eastern Togo.

The number of accessions sampled per population ranged from 14 to 74 (**Table 5.1**). Within a population, sampled individuals were at least 100 m from each other to reduce the risk of collecting closely related individuals. Each sampled accession was georeferenced (Longitude, Latitude, Altitude) using a Garmin eTrex10® GPS.

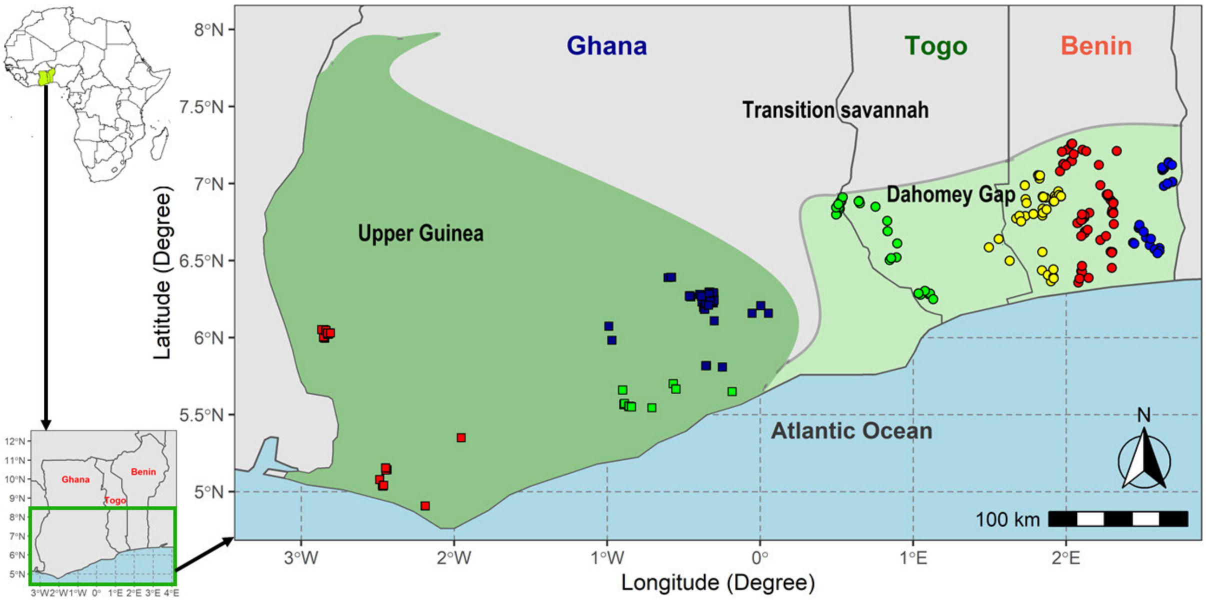


Figure 5.1. Map showing the miracle plant (*Synsepalum dulcificum*) accessions sampled in the Upper Guinea and Dahomey Gap blocks of the West African tropical rainforest for this study. The locations of individual accessions are indicated by squares (Upper Guinea block) or circles (Dahomey Gap block). The three Upper Guinea populations are the Western (red), Central (green) and Eastern (blue) populations. The four Dahomey Gap populations are the Volta (green), Mono (yellow), Zou (red), and Oueme (blue) populations.

5.2.3 Leaf tissue sampling and genomic data generation

For each accession, three to five young and pest-free leaf samples were collected and directly dried using silica gel to preserve the DNA for further use. Organized into collection kits consisting of racks of 96 tubes, 15 - 20 mg of silica-gel dried leaf tissue from each accession was shipped to SEQART AFRICA (Nairobi, Kenya), formerly Integrated Genotyping Support and Service (IGSS), an African branch of Diversity Arrays Technology Pty. Ltd. (Canberra, Australia), for DNA extraction, library preparation and genotyping. DNA extraction was based on the company's in-house protocol, and genotyping was based on the DArTseq™ technology (Sansaloni et al., 2011), combining complexity reduction methods and Next Generation Sequencing (NGS) platforms. Complexity reduction was achieved using a standard two enzyme system (*Pst1* and *MseI*), which was followed by high-density sequencing on the Illumina Hiseq 250 platform.

Table 5.1. Distribution of miracle plant (*Synsepalum dulcificum*) accessions sampled per ecological region and population

Ecology	Country	Population	Population code	Longitudinal range	Latitudinal range	Altitude range	Sample (n)
Upper Guinea				LW 002°51'55.9" — LW 000°03'12.7"	LN 05°20'59.6"— LN 06°23'34.7"	9 m — 579 m	110
	Ghana	Western	WEST	LW 002°51'55.9" — LW 001°57'17.6"	LN 05°20'59.6" — LN 06°03'07"	9 m — 276 m	30
		Central	CENT	LW 000°53'59.2" — LW 000°11' 3.8"	LN 5°39'00.3" — LN 05°39'34.6"	70 m — 151 m	14
		Eastern	EAST	LW 000°59'23.4" — LW 000°03'12.7"	LN 05°58'57" — LN 06°23'34.7"	37 m — 579 m	66
Dahomey Gap			—	LE 000°29'45.6" — LE 002°41'39.9"	LN 06°14'54.8 — LN 07°15'30.4"	-2 m — 570 m	212
	Ghana and Togo	Volta (Volta 1, 2 3)	VOLT	LE 000°29'45.6" — LE 001°07'50.7"	LN 06°14'54.8 — LN 06°50'45.8"	34 m — 570 m	40
	Togo and Benin	Mono	MON	LE 001°29'40.7" — LE 001°57'58.5"	LN 06°21'49.59" — LN 07°03'06.8"	-2 m — 220 m	74
	Benin	Zou	ZOU	LE 001°57'38.1" — LE 02°20'9.559"	LN 06°21'28.4" — LN 7°15'30.4"	11 m — 247 m	67
		Oueme	OUE	LE 002°26'17.2" — LE 002°41'39.9"	LN 06°32'46.6" — LN 07°08'16.6"	11 m — 138 m	31

5.2.4 Bioinformatics and genetic analyses

Raw read processing and SNP calling were done using Diversity Array Technology's propriety analytical pipelines (Walters et al., 2020), incorporating technical replicates as a measure of reproducibility at each locus. The generated SNPs were further quality-filtered prior to their use in downstream analyses. First, the "dartR" package (Gruber et al., 2018) (V 1.9.9) was employed to filter loci that: i) deviated from Hardy-Weinberg equilibrium at a threshold of 5%; ii) had a call rate < 85% across the population; iii) had a minor allele frequency (MAF) < 5%; and iv) had an average repeatability of alleles at a locus (reproducibility rate) < 95%. Missing cells in the genotyping matrix obtained after this first round of filtering were imputed based on the Ensemble Method algorithm as implemented by the "Optimal Imputation V. 1.0.5" plugin on the KDcompute platform (https://kdcompute.seqart.net/kdcompute/plugins). The resultant imputed matrix was then re-filtered based on MAF (loci with MAF < 5% culled) to obtain the final genotyping matrix for use in all downstream analyses. Finally, the "dartR" package was used to generate file formats suitable for analyses in other relevant packages.

5.2.4.1 Genetic diversity, population divergence, and isolation by distance

Genetic diversity in the species was characterized by computing various overall diversity indices, including observed heterozygosity (Ho), expected heterozygosity (Hs), inbreeding coefficient (F_{IS}), and total gene diversity (Ht), all using the *gl.basic.stats()* function in the "dartR" package. Each statistic was tested for significant departure from 0 with the function *overallTest()* in the "strataG" package (Archer et al., 2017). Ho, Hs, F_{IS} and allelic richness (Ar) were also computed within populations to assess population-to-population differences in genetic diversity. Ho and Hs were computed with the functions *gl.report.heterozygosity()* in "dartR", F_{IS} with the function *basic.stats()* in the "hierfstat" package (Goudet et al., 2015), while Ar was computed using the function *allel.rich()* in the "PopGenReport" package (Adamack and Gruber, 2014). Differences in of these statistics among populations was tested using a Kruskal-Wallis test. Significance of the pairwise differences in Hs between populations was tested using the function *gl.test.heterozygosity()* with 1,000 replications of re-randomization in "dartR". The existence of private alleles [allele present in only one population (Kalinowski, 2004)] was checked using the function *privates_alleles()* in the "poppr" package (Kamvar et al., 2021), whereas the proportion of shared alleles per pairs of populations was computed with the function *pairwise.propShared()* in the "PopGenReport" package.

We further tested population divergence in the species by computing an overall differentiation index (Fstg) as well as pairwise genetic differentiation index (Fstp) between pairs of populations following Weir and Cockerman (1984). Fstg was computed and tested for significance with *overallTest()*, whereas Fstp was obtained with the function *pairwise.wcfst ()*

in the “STAMPP” package (Pembleton et al., 2013) and tested for significance using 1,000 bootstraps across loci. Finally, we conducted an isolation by distance (IBD) test to understand the degree to which geographical distance correlates with genetic distance in the species. To accomplish this, a Mantel test was computed between the natural logarithm of geographical distance and the genetic distance quantity [Fstp / (1-Fstp)] (Gruber et al., 2018) as implemented in the function *gl.ibd()* in “dartR”. The significance of the Mantel correlation was assessed using 1,000 permutations.

5.2.4.2 Population structure and genetic barrier detection

Population structuring was explored using four approaches. Firstly, a hierarchical Analysis of Molecular Variance (AMOVA) was conducted to partition the genetic variation among ecological regions, populations, and individuals. AMOVA was conducted using the function *poppr.amova()* in “poppr”, and the significance of each source of variation was evaluated using a bootstrap method (1,000 permutations) through the function *randtest()* in “poppr”. Secondly, a principal coordinate analysis (PCoA) was performed using the function *gl.pcoa()*, with the bivariate plot of the results visualized using the function *gl.pcoa.plot()*, both functions in “dartR”. Thirdly, a Bayesian clustering analysis was conducted using STRUCTURE (Pritchard et al., 2000). Three independent runs, with the number of clusters K varying from 2 to 10, were executed with a run length of 10,000 Markov chain Monte Carlo replicates after a burn-in period of 10,000 iterations, based on a correlated-allele frequency model. The resulting output was transferred into STRUCTURE HARVESTER (Earl, 2012), where the standard Delta K method (Evanno et al., 2005) identification was used to identify the optimal number of genetic clusters (K) in the dataset. Based on the optimal number of clusters identified, a bar plot of assigned genotypes was built using the “Pophelper” package (Francis, 2017). Finally, population structure was investigated via consideration of the phylogenetic relationships among accessions, as evidenced by the neighbor joining model (Saitou and Nei, 1987) based on the Tajima-Nei evolutionary distance (Tajima and Nei, 1984). The phylogeny model was tested using 1,000 bootstrap replications in the software MEGA X (Kumar et al., 2018), and the generated tree was exported as a Newick file and edited using the Interactive Tree of Life (ITOL, V.5) (Letunic and Bork, 2021). To test for the existence of physical/genetic barrier(s) disrupting genetic connectivity among the studied populations, Monmonier’s algorithm was implemented in the Barriers software (V.2.2) (Manni et al., 2004).

5.2.4.3 Nucleotide polymorphism, neutrality statistics, and evolutionary histories of inferred genetic populations

The genetic clusters or sub-populations inferred from the analyses of population structure were used to build new datasets for DNA polymorphism analysis. FASTA files were generated for each genetic cluster using the function *gl2FASTA()* in “dartR” and used for analysis in the software package DnaSP (V.6.12.03) (Rozas et al., 2017). Specifically, the cluster-based polymorphism data were evaluated via five parameters, namely the number of segregating sites (S), the mean number of pairwise nucleotide differences (K), the number of haplotypes (h), the nucleotide diversity (π), and the haplotype diversity (Hd) (Rozas et al., 2017). The hypothesis of selective neutrality was tested by computing a suite of neutrality statistics, namely Tajima’s D, Fu’s Fs, Achaz’s Y*, and Ramos-Onsins and Rozas’s R_2* (Fu and Li, 1993; Ramos-Onsins and Rozas, 2002; Tajima, 1989). P-values associated with these indices were obtained through a coalescent simulation based on 1,000 bootstrap replications. A mismatch distribution (distribution of the observed pairwise nucleotide site differences and the expected values) analysis was also performed for each cluster to assay the sudden population expansion hypothesis, still in DnaSP (V.6.12.03).

5.3 Results

5.3.1 Genotyping

A total of 16,514 raw DArTseq SNPs was called using the sequencing data from the 322 sampled accessions, a set that was further narrowed to 2,704 high confidence SNPs (**Appendix 5.1**) after filtering. The polymorphism information content (PIC) of these 2,704 markers ranged from 0.09 to 0.49 (mean = 0.41), with a PIC value > 0.4 for more than 75% of the markers. The minor allele frequency (MAF) of the final set of markers ranged from 0.051 to 0.5 (mean = 0.32), with more than 70% of markers having an MAF above 0.3.

5.3.2 Genetic diversity within and among *S. dulcificum* populations in West Africa

Overall estimates of observed heterozygosity (H_o), expected heterozygosity (H_s), total gene diversity (H_t) and inbreeding (F_{IS}) in the species were 0.02, 0.14, 0.44 and 0.79, respectively. $H_{S.}$ H_T and F_{IS} significantly different from 0 ($p < 0.001$). Population estimates of these statistics, are shown in **Table 5.2**, along with associated p-values showing that all differ significantly among populations. The Volta population in the Dahomey Gap consistently exhibited the highest estimates of diversity, while the Central population in the Upper Guinea block exhibited the lowest. This signal of higher diversity within the Volta population compared to other populations is further strengthened by the highest mean allelic richness exhibited by this population (Ar_{VOL} = 1.96). In contrast, the lowest allelic richness was observed in the Western population in the Upper Guinea block (Ar_{WEST} = 1.31) (**Table 5.2**).

Table 5.2. Genetic diversity indices for the seven populations of miracle plant (*Synsepalum dulcificum*) sampled in this study. H_O: Observed heterozygosity; H_S: Expected heterozygosity; F_{IS}: Inbreeding coefficient; Ar: Mean Allelic richness. See Table 1 for full population names.

Population	H_O	H_S	F_{IS}	A_r
WEST	0.018	0.072	0.75	1.31
CENT	0.015	0.072	0.792	1.36
EAST	0.034	0.189	0.818	1.72
VOLT	0.048	0.334	0.854	1.96
MON	0.022	0.097	0.774	1.42
ZOU	0.017	0.082	0.788	1.34
OUE	0.047	0.156	0.697	1.56
P-values	< 0.0001	< 0.0001	< 0.0001	< 0.0001

As shown in **Table 5.3**, pairwise comparisons of expected heterozygosity revealed significant differences for all population pairs ($p < 0.0001$), except for the pair Central-Zou ($p = 0.12$). No private alleles were detected for any population, but the proportion of shared alleles between pairs of populations ranged widely, from 21.1% (West-Zou) to 97.1% (Mono-Zou) (**Table 5.3**).

Table 5.3. Pairwise differences in expected heterozygosity (lower diagonal) and pairwise allele share (upper diagonal) among sampled populations. Values in lower diagonal are differences in expected heterozygosity between pairs of populations, with associated p-values (determined based on 1,000 replications of re-randomization) in brackets. Bold values indicate significance. See Table 1 for full population names.

	WEST	CENT	EAST	VOLT	MON	ZOU	OUE
WEST		92.84	90.53	39.34	22.04	21.10	26.24
CENT	**-0.012 (0)**		90.16	39.94	23.38	22.51	27.51
EAST	**0.106 (0)**	**-0.118 (0)**		46.25	27.77	26.64	32.41
VOLT	**0.246 (0)**	**-0.258 (0)**	**-0.140 (0)**		77.77	76.55	81.55
MON	**0.015 (0)**	**-0.028 (0)**	**0.090 (0)**	**-0.230 (0)**		97.14	90.82
ZOU	**0.008 (0.01)**	-0003 (0.12)	**0.115 (0)**	**0.255 (0)**	**0.024 (0)**		80.91
OUE	**0.072 (0)**	**-0.084 (0)**	**0.034 (0)**	**-0.174 (0)**	**-0.056 (0)**	**0.08 (0)**	

5.3.3 Population divergence and isolation by distance

The calculated global differentiation index (Fst_g) of 0.65 for this collection of *S. dulcificum* accessions differed significantly from 0 ($p < 0.001$). All pairwise differentiation index (Fst_p) was also significant, ranging from 0.01 to 0.90 ($p < 0.0001$) (**Table 5.4**). The neighboring Mono and Zou populations in the Dahomey Gap were the least differentiated populations (Fst_p = 0.01), while the Central (Upper Guinea) and Zou (Dahomey Gap) populations were found to exhibit the highest level of divergence (Fst_p = 0.9). Fst_p ranged from 0.01 to 0.31 for populations within the Dahomey Gap versus 0.08 to 0.19 in the Upper Guinea block (**Table 5.4**). The Mantel correlation revealed significant positive correlation (r = 0.71, $p = 0.003$) between geographical and genetic distance, indicating a measurable effect of isolation by distance in the West African population of *S. dulcificum* sampled in this study.

Table 5.4. Pairwise differentiation index (Fstp) among sampled populations of miracle plant (*Synsepalum dulcificum*) in the study area. All values are significant at p < 0.000. See Table 1 for full population names.

	WEST	CENT	EAST	VOLT	MON	ZOU	OUE
WEST							
CENT	**0.19**						
EAST	**0.09**	**0.08**					
VOLT	**0.65**	**0.60**	**0.58**				
MON	**0.88**	**0.88**	**0.81**	**0.27**			
ZOU	**0.90**	**0.90**	**0.82**	**0.31**	**0.01**		
OUE	**0.84**	**0.82**	**0.75**	**0.16**	**0.11**	**0.13**	

5.3.4 Population structure and genetic barriers

As shown in the table of hierarchical AMOVA results (**Table 5.5**), significant population structure was observed in the species, with 76.3% of the molecular variation attributed to differences between the Dahomey Gap and Upper Guinea ecological regions. Only 3.9% of the molecular variation could be assigned to differences among populations within ecological regions, while 15.8% was observed among accessions within populations. The percentage of molecular variance explained attributable to within-accession variation (4.0%) is only slightly larger than that observed among populations.

Table 5.5. Table of results from the hierarchical AMOVA, based on two ecological regions (Upper Guinea and Dahomey Gap), seven populations, 322 individual accessions and 2,704 SNP markers.

Source of variation	Df	Sum Square	Mean Square	Variance component	% variation	p
Between ecological regions	1	221724.2	221724.2	750.1	76.3	0.001
Among populations within regions	5	18167.7	3633.5	38.5	3.9	0.001
Among accessions within populations	315	110241.1	350.0	155.5	15.8	0.001
Within accessions	322	12545.5	39.0	39.0	4.0	0.001
Total	643	362678.5	564.0	983.1	100.0	

The first axis of the principal coordinate analyses, explained 71.2% of the genetic variation and effectively discriminated the accessions based on their ecoregional origin, with populations collected in the Upper Guinea block (WEST, CENT, EAST) clustering together, while those from the western and central areas of the Dahomey Gap block (MON, ZOU, OUE) forming another cluster (**Fig. 5.2)**. According to this analysis, the Volta population (VOLT), collected along the Ghana-Togo border in the western Dahomey Gap, represents a transition of sorts, with sub-populations differentially aligning with the two main regional clusters. The second axis, explaining only 4% of the observed variation, appears effective mainly in discriminating accessions within the Mono and Zou populations, two populations already established to be highly similar to one another (**Table 5.4**).

Bringing a slightly higher resolution to the West African *S. dulcificum* population structure, the STRUCTURE analysis indicated K = 3 as the optimal number of clusters (**Appendix 5.1**). At K = 2, the two observed clusters corresponded perfectly with the Upper Guinea and Dahomey Gap blocks, reiterating the ecological structuring of molecular variation suggested by the analyses above. Optimization of K, however, resulted in a further split of the Dahomey Gap cluster into two sub-clusters, ultimately leading to the detection of three distinct genetic groups (**Appendix 5.2, Fig. 5.3)**. The set of 111 accessions comprising Cluster 1 contained 108 out of the 110 accessions collected in the Upper Guinea block in Ghana, plus two accessions collected in VOLT2 sub-population and one accession in VOLT3 sub-population. This cluster represented the Upper Guinean group (UG). The relatively small Cluster 2 contained 11 accessions collected all from the VOLT1 sub-population in the Western part of the Dahomey Gap block. This Cluster defined the Western Dahomey Gap genetic group (WDG). Cluster 3, the largest one with 196 accessions, was composed of 195 accessions collected from VOLT2 and VOLT3 sub-populations of the VOLT population and MON, ZOU and OUE populations in the Central part of Dahomey Gap block and one accession from the EAST population in the Upper Guinea block. This cluster is the Central Dahomey Gap genetic group (CDG). Three other accessions (DAT347, DAT348, DAT349) from the OUE population in the Dahomey Gap block were disputed by CDG and WDG groups, as was the accession DAT278 from the EAST population in the Upper Guinea block disputed by UG and WDG. These four accessions then formed an admixed set.

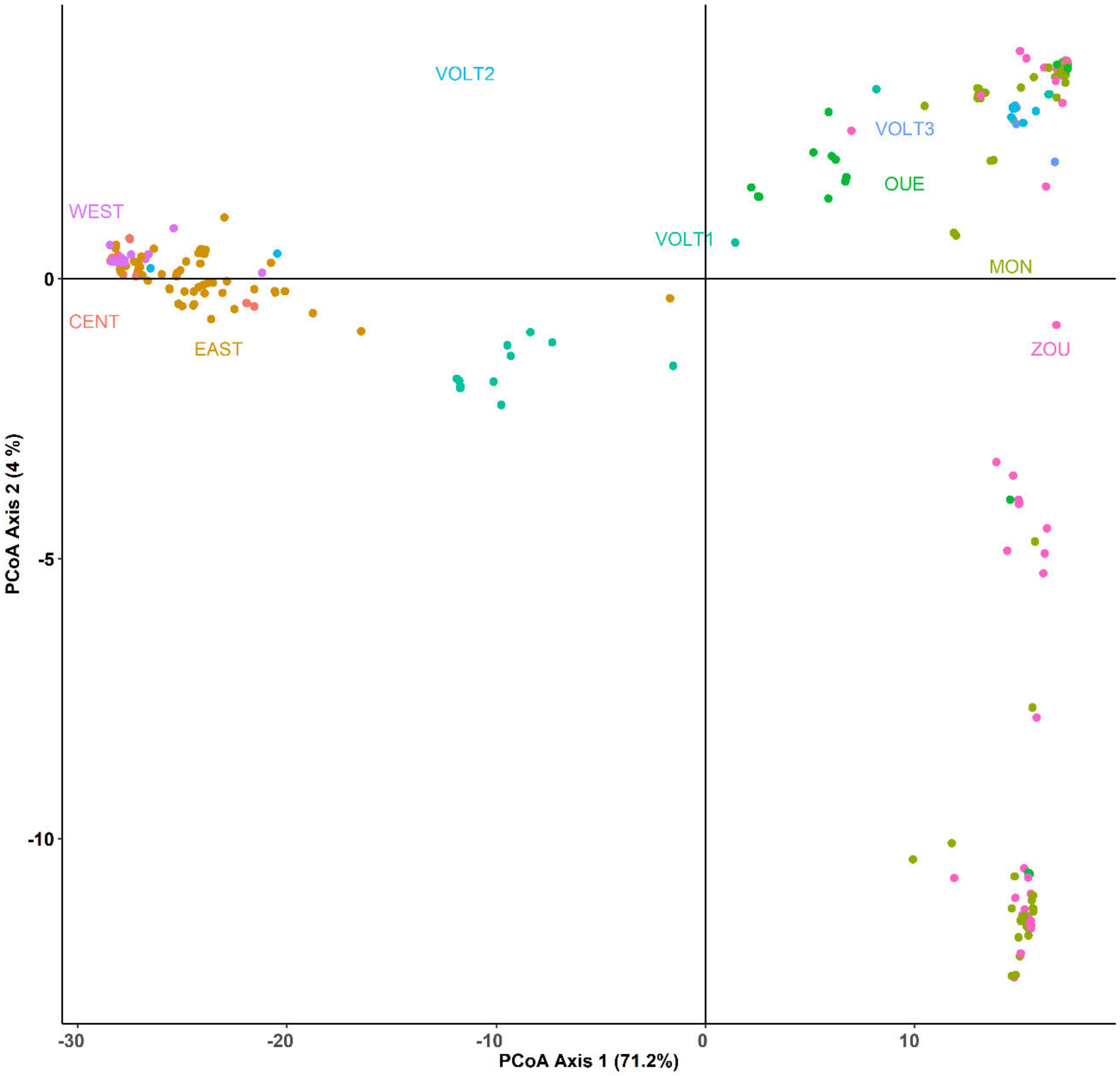


Figure 5.2. Biplot of principal coordinate analysis based on 2,704 markers and 322 accessions of miracle plant (*Synsepalum dulcificum*) sampled from the Upper Guinea block (WEST: Western population; CENT: Central population; EAST: Eastern population) and the Dahomey Gap block (VOLT1: Volta 1 sub-population; VOLT2: Volta 2 sub-population; VOLT3, Volta 3 sub-population; MON: Mono population; ZOU: Zou population; and OUE: Oueme population).

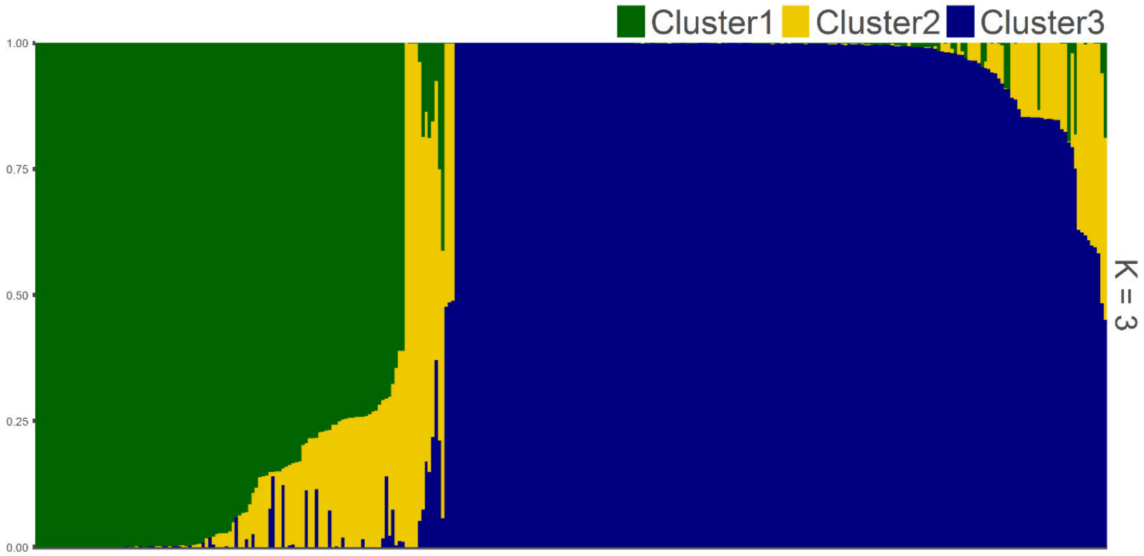


Figure 5.3. Bar plot showing the structuring pattern of the 322 accessions of miracle plant (*Synsepalum dulcificum*) collected in the Upper Guinea and the Dahomey Gap at K = 3 (optimal number of clusters detected by STRUCTURE analysis). Cluster 1 = Upper Guinea genetic group (UG), Cluster 2 = Western Dahomey Gap genetic group (WDG), and Cluster 3: Central Dahomey Gap genetic group (CDG).

The neighbor joining (NJ) tree as shown in **Fig. 5.4** improved the resolution of the ecoregional clustering pattern outlined by the STRUCTURE analysis by not only identifying three major groups, but also by reintegrating the four admixed accessions in their original regions of collection. Based on the NJ tree, the largest group (Group 3, **Fig. 5.4**) encompassed all the individuals of the STRUCTURE-inferred CDG genetic group (196 accessions) plus the three admixed accessions from the Oueme population and one accession that escaped from the WDG group. Consequently, the STRUCTURE-defined WDG group with 11 accessions was now refined to 10 accessions based on the NJ analysis (Group 2, **Fig. 5.4**). These two groups (Group 2 and 3) reconstituted the Dahomey Gap group at K = 2. The intermediate group (Group 2, **Fig. 5.4**, 112 accessions) had the same constitution as the STRUCTURE-defined UG genetic group, with however the additional admixed accession that originated in the Upper Guinea but indicated to belong to the WDG genetic group by the STRUCTURE analysis.

Combined, STRUCTURE, PCoA and NJ analyses inferred the following stable genetic groups: CDG group (included 196 accessions all from the Central part of the Dahomey Gap block), UG group (111 accessions, all from the Upper Guinea block) and WDG group (10 accessions all from the western part of the Dahomey Gap block, and which geographically lies between the Upper Guinea and Central Dahomey Gap) that are considered in the next section analyses.

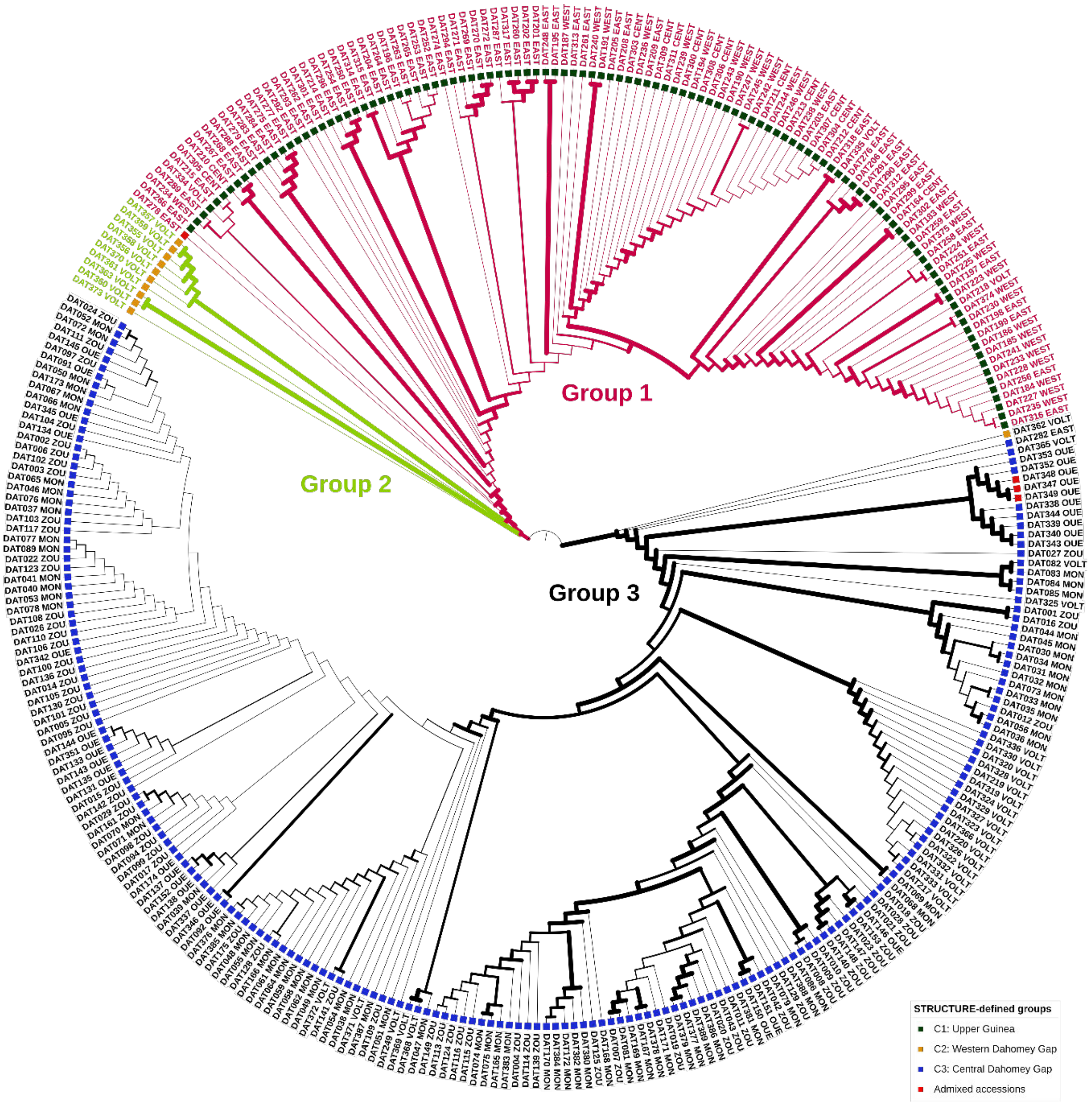


Figure 5.4. Neighbour joining tree-based phylogenetic relationship among the 322 accessions of miracle plant (*Synsepalum dulcificum*). Accessions are coloured based on their belonging to NJ-defined groups whereas associated markers' colour indicated STRUCTURE grouping (with red markers indicating admixed accessions). Branches' widths are proportional to the bootstrap values in a 1,000 replicates trial. Each accession name is composed of the collecting code and the accession initial population of origin (WEST: Western population; EAST: Eastern population, CENT: Central population, VOLT: Volta population, MON: Mono, ZOU: Zou population and OUE: Oueme population).

The Monmonier's algorithm as shown in **Fig. 5.5** demonstrated the existence of a physical barrier between the Upper Guinea and the Dahomey Gap populations, which is located between the Volta and Eastern populations.

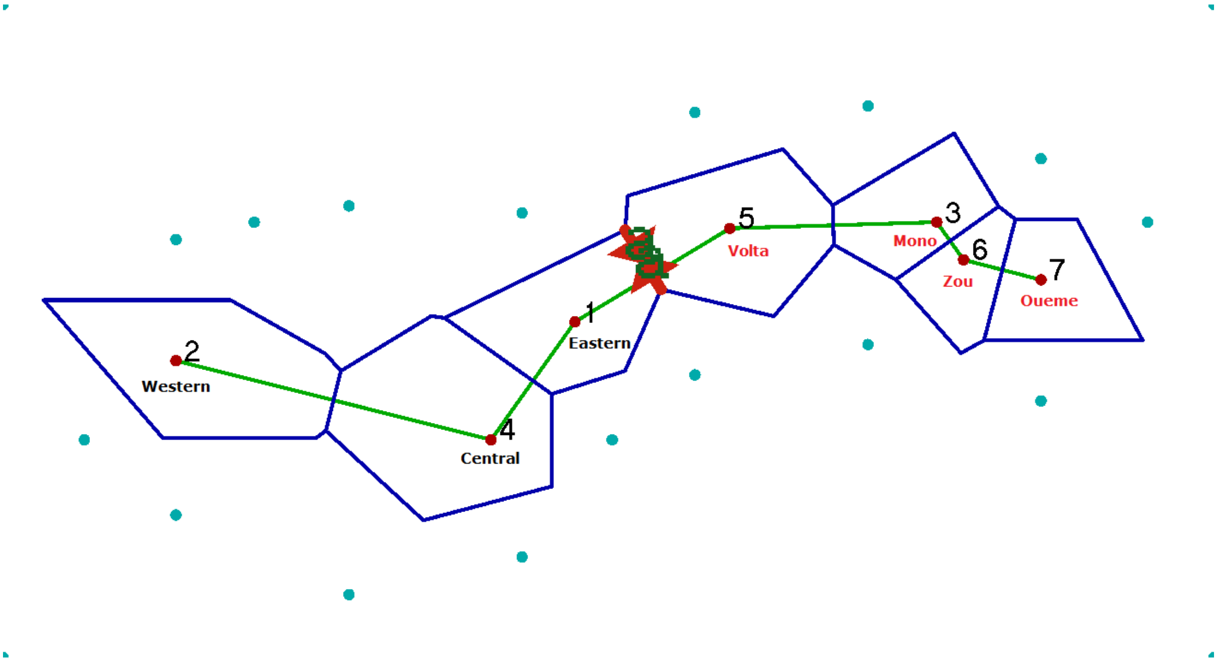


Figure 5.5. Genetic barrier (double red arrows) disrupting connectivity between Upper Guinea and Dahomey Gap populations of miracle plant (*Synsepalum dulcificum*) as indicated by the Barriers (V.2.2.) software. Upper Guinea populations are in black font and Dahomey Gap ones in red font.

5.3.5 Nucleotide polymorphism and evolutionary history based of inferred genetic groups

Diversity indices computed for the inferred genetic clusters are presented in **Table 5.6**. Nucleotide diversity was overall low in the species (π = 58.36 ± 5 x 10^{-5}) but exhibited a six-fold to 50-fold variation across the genetic group. The Western Dahomey Gap group (WDG) had the highest nucleotide diversity (π_{WDG} = 156 ± 20 x 10^{-5}), whereas the Central Dahomey Gap group (CDG) presented the lowest one (π_{CDG} = 2 ± 0 x 10^{-5}). The same trend was also observed in haplotype diversity with the Western Dahomey Gap group exhibiting the highest haplotype diversity ($\mathrm{Hd}_{\mathrm{WDG}}$ = 1 ± 0.04) and the Central Dahomey Gap group the lowest one ($\mathrm{Hd}_{\mathrm{CDG}}$= 0.76 ± 0.02). Contrary to nucleotide diversity, the average haplotype diversity in the species was high (Hd = 0.9 ± 0.01). The Upper Guinea genetic group consistently presented an intermediate level of diversity (**Table 5.6**).

Table 5.6. Estimates of nucleotide diversity indices for STRUCTURE-inferred genetic groups of miracle plant (*Synsepalum dulcificum*) in West Africa. UG: Upper Guinea, WDG: Western Dahomey Gap and CDG: Central Dahomey Gap.

Genetic group	Ecological region	S	k	(π ± sd) x 10^{-5}	h	Hd ± sd
Cluster 1 (UG)	Upper Guinea	418	37.74	21 ± 2.00	62	0.95 ± 0.01
Cluster 2 (WDG)	Dahomey Gap	764	280.62	156 ± 20	10	1.00 ± 0.04
Cluster 3 (CDG)	Dahomey Gap	72	04.08	2 ± 0.00	36	0.76 ± 0.02

S: Number of segregating sites, k: Average number of pairwise nucleotide difference, h = number of haplotypes, Hd: haplotype diversity and **π:** nucleotide diversity

The neutrality test results shown in Table 7 indicated that Tajima's D, Fu's Fs and Achaz Y* statistics were all negative for UG and CDG, but all positive for WDG except Achaz Y*. None of Tajima's D, Fu's Fs and Achaz Y* were significant in WDG (D = 0.19, *p = 0.61*; Fs = 1.76, *p = 0.5*; and Y* =-0.08*, p = 0.47*), while they were all very significant in CDG (D = -2.03, *p = 0.002* and Fs = -14.51*, p = 0.002* and Y* =-1.63*, p = 0.02*), thus suggesting a significant departure from neutrality in the Central Dahomey Gap genetic population. By contrast, only Tajima's D (D = -1.75, *p = 0.01*) and Achaz Y* (Y* = -1.94, *p = 0.005*) were significant in UG. As for the Ramos-Onsins and Rozas's R^2 statistics, it was low and significant in both UG and CDG, but high and not significant in WDG (**Table 5.7**). The mismatch distribution curves (**Fig. 5.6**) exhibited a unimodal pattern in CDG (**Fig. 5.6C**), a bimodal pattern in UG (**Fig. 5.6A**), and a multimodal pattern in WDG (**Fig. 5.6B**).

Table 5.7. Neutrality test statistics on inferred genetic groups (populations) of miracle plant (*Synsepalum dulcificum*) in the study area. Values in bold are significant. UG: Upper Guinea, WDG: Western Dahomey Gap; CDG: Central Dahomey Gap.

Ecological zone	Upper Guinea		Dahomey Gap			
Genetic group	Group 1 (UG)	*p*	Group 2 (WDG)	*p*	Group 3 (CDG)	*p*
Tajima's D (D)	**-1.75***	0.01	0.19	0.61	**-2.03****	0.002
Fu's FS (FS)	-3.582	0.25	1.76	0.50	**-14.51****	0.002
Achaz Y* (Y)	**-1.94****	0.005	-0.08	0.47	**-1.63****	0.026
Ramos-Onsins and Rozas's R^2	0.04	0.007	0.15	0.54	**0.03***	0.016

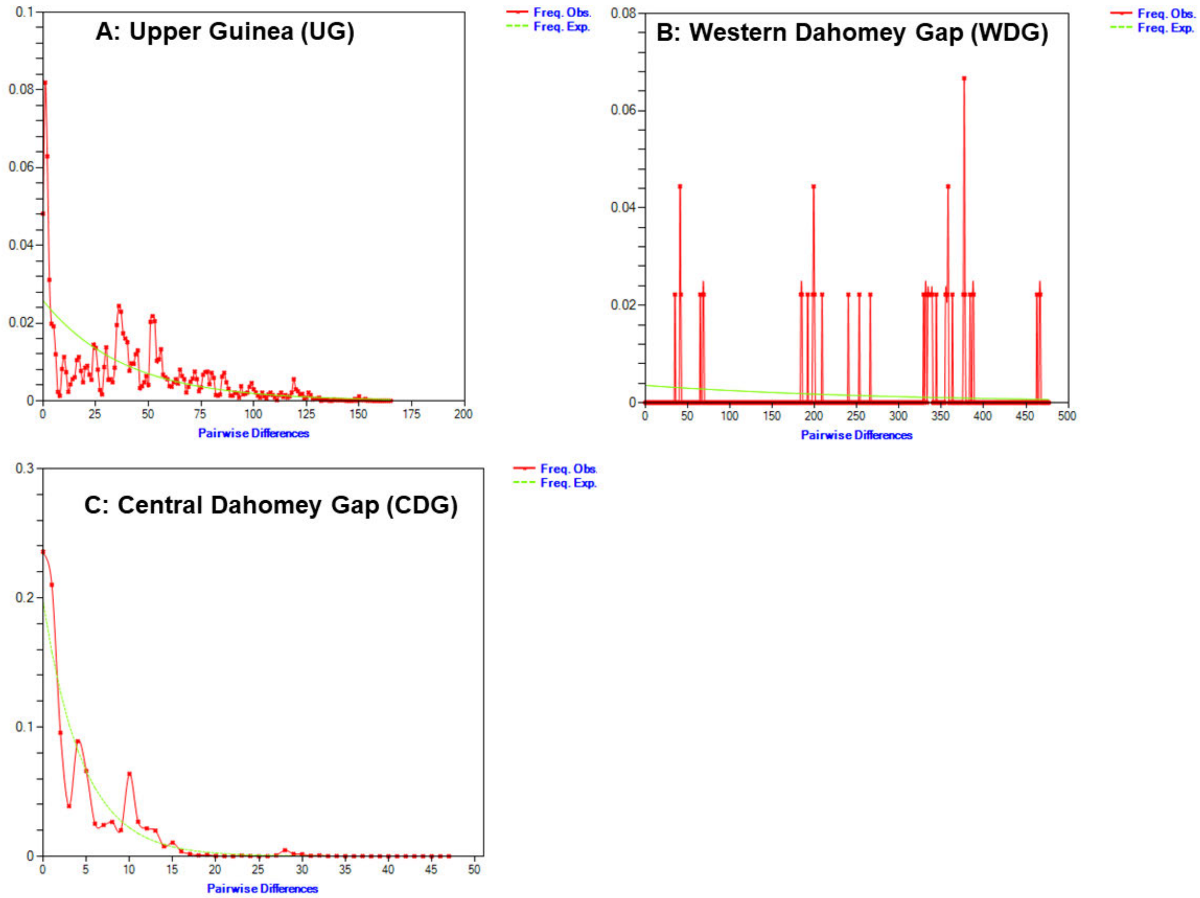


Figure 5.6. Mismatch distribution curves for genetic populations of miracle plant (*Synsepalum dulcificum*) detected in the Upper Guinea and the Dahomey Gap blocks.

5.4 Discussion

Knowledge in the Sapotaceae family remained scanty from many points of view (Ayensu, 1972), particularly in terms of genetic improvement. Although commendable efforts have been observed these last years for the systematic phenotypic characterization of the family species (Gwali et al., 2012; Sanou et al., 2006; Tchokponhoué et al., 2020), genomics study are particularly rare in the family, with *Vitellaria paradoxa* (Hale et al., 2021) and *Argania spinosa* (Khayi et al., 2020; Slimane et al., 2020) remaining the most advanced species of the family in the genomics field, and simple sequence repeat (SSR) markers being the most employed to date in the family (Mouhaddab et al., 2017; Pakhrou et al., 2017; Sanou et al., 2005). This study, the first to employ SNP markers to understand the genetic variation organization in *Synsepalum dulcificum* is expected to pave the way for systematic genomics-assisted-breeding of the species.

SNPs are salient across plant genomes and less prone to mutation compared to other types of markers (e.g., SSR markers) and consequently represent first-order markers in the study of plant genomic diversity (Carrasco et al., 2018; Wei et al., 2021). The use of reasonable amount of quality SNP is crucial for an effective resolution of the diversity organization. Here, the 2,704 SNPs used were highly informative as revealed by their average PIC value.

Based on these SNP markers, an extremely low genomic diversity was shown in the species (see **Table 5.2**), a result that corroborated the conclusion of Chibuzor et al. (2017) who employed RAPD markers to study the species diversity in Southern Nigeria. This low diversity is accompanied by a high inbreeding coefficient (F_{IS}), which is indicative of a high inbreeding that could have resulted from a reduction of the species effective population size (Szczecińska et al., 2016) given that *S. dulcificum* is locally endangered in the study area (Adomou, 2005). This extent of inbreeding could also reflect a predominance of autogamy in the species. Presence of autogamy was experimentally detected in the species by Tchokponhoué et al. (2017), but whether autogamy is predominant over allogamy is yet to be clarified. The high inbreeding coefficient ($F_{IS} = 0.79$) suggesting that accessions investigated are near the inbred line state (Gbedevi et al., 2021) offers an impetus to quickly advance breeding population development for agronomic and functional traits improvement in the species. Despite the overall low diversity in the species, Volta population in particular, and Eastern and Oueme populations to some extent can provide useful individuals for breeding program of the species as these populations exhibited the highest expected heterozygosity and mean allelic richness.

A Fst below 0.05 indicates a low differentiation, while Fst values in the ranges 0.05 – 0.15, 0.15 – 0.25 and > 0.25 suggest moderate, high and very high differentiation, respectively (Wright, 1978). Relative to other perennial species [e.g., *Garcinia kola* (Heckel): Fst < 0.09, (Dadjo et al., 2020); *Elaeis spp*: Fst = 0.32, (Pereira et al., 2020); *Macademia spp*: Fst = 0.4, (O'Connor et al., 2019) and *Malus sieversii* (Lebed.): Fst < 0.16, (Richards et al., 2009)], the singular Fst (Fstg = 0.69) for S. *dulcificum* estimated in this study signals an extremely poor gene flow within the species. Such an extent of differentiation can be first explained by the significant isolation by distance observed in the species. Indeed, as highlighted by the pairwise differentiation index (**Table 5.4)**, geographically closer populations were genetically more connected. The AMOVA results which highlighted a predominant geographical structuring of the molecular variation gave more insight into this poor gene flow in the species. Factors including the life cycle, the reproductive system, and the seed dispersal, among others, are known to affect genetic variation organization in plant species (Nybom, 2004). Perennial and outcrossing species are expected to retain higher variation within population and exhibit a high heterozygosity and a low differentiation, while annual and selfing species are more variable among populations with a high differentiation and a low diversity (Andriamihaja et al., 2021; Kouam et al., 2012). In the perennial predominantly autogamous *S. dulcificum*, our observation matches this prediction since more genetic variation was partitioned among ecological zones that could have constituted mega-populations of interest in the species.

This hypothesis of ecologically defined mega-population is supported by the results of the structure analysis, which indicated a primary segregation (at K = 2) of the studied accessions and populations into two gene pools ─Upper Guinea and Dahomey Gap─, among which gene flow is restricted by the existence of a physical barrier as evidenced by the Monmonier's algorithm (**Fig. 5.5**). The geographical location of this identified barrier that is between the Volta region and the Eastern region seemed to strongly coincide with the lake Volta, which was previously pinpointed as a strong biogeographical barrier between the Dahomey Gap and Upper Guinea populations in many living organisms (Booth, 1958; Nicolas et al., 2010), and henceforth in *S. dulcificum* as suggested by the pairwise differentiation index (**Table 5.4**). In addition, the lake Volta barrier was indicated to have evolutionary divergence implications on populations it separated (Moreau, 1969). Even if a similar trend of differentiation was reported between Dahomey Gap and Upper Guinea populations in many other species [e.g., *Pentadesma butyracea* Sabine' (Ewédjè, 2012) and *Distemonanthus benthaminus* (Demenou et al., 2016)], it was to a lesser extent. Three genetic pools were identified in *S. dulcificum* with the isolation between Upper Guinea and Dahomey Gap gene pools maintained, and a further split of the Dahomey Gap gene pool revealed two further clusters. This splitting pattern reveals an overall higher diversity in the Dahomey Gap compared with Upper Guinea; and this is confirmed by the principal coordinate analysis through a denser clustering pattern of individuals in Upper Guinea versus a scattered one for those in the Dahomey Gap. Although three widely spread populations were considered in the Upper Guinea, they ultimately all fitted into a single cluster, a trend that was also previously documented in many species collected from the region, either only from Ghana [e.g., *D. benthaminus* (Demenou et al., 2016)] or from several countries [e.g., *Terminalia superba* collected from Guinea to Ghana (Demenou et al., 2018)]. This finding might illustrate that the Guinean block of the West African forest has been relatively stable with a tightly maintained connectivity among its populations, and seems to be in tune with the hypothesis of the Upper Guinea being an independent biogeographic region (Linder et al., 2012).

The neighbor joining tree result aligned with that of STRUCTURE on three genetic cluster in the species. Two clusters in the Dahomey Gap corresponding with the Central Dahomey Gap (CDG) and Western Dahomey Gap (WDG) and a single cluster in the Upper Guinea (UG). Individuals forming the CDG genetic group were all sampled at the core of the Dahomey Gap where dominant vegetation is savannah. Individuals constituting the WDG came from the

Ghana - Togo highland ─a semi-forested region considered as a crossroad between the Upper Guinea and the Lower Guinea (Decher et al., 2021).

Genomic diversity in living organism is indicated to decrease from centers of origin outward (Rosenbom et al., 2014; Yuan et al., 2008). The DNA polymorphism analysis suggested an exceptionally higher nucleotide and haplotype diversity in WDG compared with CDG and UG. This points to the WDG as the likely center of origin of *S. dulcificum*. Because WDG genetic population mainly came from the Ghana-Togo highland, a region known as a hotspot of diversity in West Africa (Myers et al., 2000), WDG being the center of origin of *S. dulcificum* was therefore not surprising. This finding stands as the first molecular evidence to refine the origin of the species as previous speculations vaguely suggested a complex of countries including Ghana, Togo, Benin and Nigeria to form the potential center of origin (Achigan-Dako et al., 2015; Swenson et al., 2008).

Several statistics have been employed to infer living organisms' evolutionary history, with Tajima's D and Fu's Fs being the most known ones (Fu, 1997). However, Fu's Fs and Ramos-Onsins and Rozas's R^2 were evidenced as being the most powerful in detecting population growth, with the first one being adapted for large populations, and R^2 being more suitable for small sized populations (Ramos-Onsins and Rozas, 2002). Unfortunately, both are prone to unbiased estimation when sequencing errors occurs, and Achaz Y* statistic was then proposed (Achaz, 2008) as an alternative. This study combined all these four statistics for the sake of an unambiguous population evolutionary history inference. Tajima's D, Fu's Fs and Achaz Y* were all positive and non-significant in WDG, and thus suggest that the Western Dahomey Gap is a stable population at demographic equilibrium. This conclusion was further supported by the high and non-significant value of Ramos-Onsins and Rozas' R^2 statistic and the multimodal pattern of the mismatch distribution curve in WDG (**Fig. 5.6B**). Conversely, a signature of a recent population expansion is detected in the Central Dahomey Gap as Tajima's D, Fu's Fs and Achaz Y* are all negative and very significant for this genetic population. Concurrent to this conclusion are the small positive R^2 statistics and the unimodal pattern of the mismatch distribution curve (Rosly et al., 2013) in CDG (**Fig. 5.6C**). In the Upper Guinea, a rather complex demographic history characterized by a possibly past population expansion that is restricted to some areas was inferred from the negative but non-significant Tajima's D and Fu's Fs values (Rosly et al., 2013). This complex demographic history is also supported by the small, positive, and significant R^2 statistics and the bimodal pattern of the mismatch distribution observed in the Upper Guinea (**Fig. 5.6A**). Overall, these various statistics yielded consistent conclusions within each genetic cluster and pointed to differential demographic histories in the investigated populations. The WDG population is mainly located in the Ghana-Togo highland region whose heterogenous geology might have buffered the influence of historic climatic oscillations, thus contributing to its demographic equilibrium.

The inferred evolutionary history combined with the genomic diversity pattern in the species highlighted that the Central Dahomey Gap Cluster that experienced a recent population expansion exhibited the lowest diversity, while the stable Western Dahomey Gap population unsurprisingly had the highest diversity. This is a strong signal that the past climate change through its successive vegetation reconfigurations has markedly impacted the genomic diversity and the population structuring in the West African population of *Synsepalum dulcificum*.

5.5 Conclusion

Knowledge on molecular variation is crucial for planning breeding and conservation strategies but lacked in the orphan *Synsepalum dulcificum*. This study, the first comprehensive one to employ SNP markers in S*. dulcificum*, provides insight into the species genetic variation, population structuring, evolutionary history, and origin. The natural genomic diversity of *S. dulcificum* is low, with however highly differentiated populations, probably due to the lake Volta that is acting as a biophysical barrier between the Dahomey Gap and the Upper Guinea gene pools. Nevertheless, the Volta population in the Dahomey Gap harbors a moderate diversity ($H_S = 0.334$) that can be quickly tapped in developing breeding populations in the species. The genomic variation in the species is highly mediated by ecological conditions, with clear footprints from the past climate changes. Noticeably, *S. dulcificum* individuals in the Central Dahomey Gap gene pool experienced a recent demographic expansion, while the Upper Guinea gene pool underwent a complex demographic change marked by an expansion that was restricted to some local regions. This study demonstrated the Western Dahomey Gap, a region that has its population at demographic equilibrium, as the likely center of origin of *Synsepalum dulcificum*, thus calling for the necessity to undertake a thorough germplasm collection in this region for the ex-situ safeguarding of useful alleles as the species is currently endangered in the whole West Africa. Although this study did not report any data related to morphological distinctiveness in the species, the evidence of a physical barrier as well as the existence of highly differentiated gene pools on both sides, coupled with some empirical morphological differences make us speculate on an ongoing allopatric speciation in the species, which deserves further examination at a larger spatial scale integrating samples from not only the Upper Guinea and Dahomey Gap, but also from the Lower Guinea (Stretching from Nigeria to Central Africa).

5.6 References

Achaz, G. (2008). Testing for neutrality in samples with sequencing errors. *Genetics* **179**, 1409-1424.

Achigan-Dako, E. G., Tchokponhoué, D. A., N'Danikou, S., Gebauer, J., and Vodouhè, R. S. (2015). Current knowledge and breeding perspectives for the miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell. *Genet. Resour. Crop Evol.* **62**, 465-476.

Adamack, A. T., and Gruber, B. (2014). PopGenReport: simplifying basic population genetic analyses in R. *Meth. Ecol. Evol.* **5**, 384-387.

Adomou, A. (2005). Vegetation patterns and environmental gradients in Benin, University of Wageningen, Wageningen, Netherlands.

Andriamihaja, C. F., Ramarosandratana, A. V., Grisoni, M., Jeannoda, V. H., and Besse, P. (2021). Drivers of population divergence and species differentiation in a recent group of indigenous orchids (*Vanilla spp.*) in Madagascar. *Ecol. Evol.* **11**, 2681-2700.

Archer, F. I., Adams, P. E., and Schneiders, B. B. (2017). stratag: An R package for manipulating, summarizing and analysing population genetic data. *Mol. Ecol. Resour.* **17**, 5-11.

Ayensu, E. S. (1972). Morphology and anatomy of *Synsepalum dulcificum* (Sapotaceae). *Bot. J. Linn. Soc.* **65**, 179-187.

Barbey, C., Hogshead, M., Schwartz, A. E., Mourad, N., Verma, S., Lee, S., Whitaker, V. M., and Folta, K. M. (2020). The genetics of differential gene expression related to fruit traits in strawberry (*Fragaria× ananassa*). *Front. Genet.* **10**.

Booth, A. (1958). The Niger, the Volta and the Dahomey Gap as geographic barriers. *Evolution* **12**, 48-62.

Buckmire, R., and Francis, F. (1978). Pigments of miracle fruit, *Synsepalum dulcificum*, Schum, as potential food colorants. *J. Food Sci.* **43**, 908-911.

Carrasco, B., González, M., Gebauer, M., García-González, R., Maldonado, J., and Silva, H. (2018). Construction of a highly saturated linkage map in Japanese plum (*Prunus salicina* L.) using GBS for SNP marker calling. *PloS ONE* **13**.

Chen, C., Wang, Y., and Wang, H. (2010a). Chemical constituents from the leaves of *Synsepalum dulcificum*. *Chem. Nat. Compd.* **46**, 495-495.

Chen, C., Wang, Y., and Wang, H. (2010b). Chemical constituents from the roots of *Synsepalum dulcificum*. *Chem. Nat. Compd.* **46**, 448-449.

Chibuzor, I. A., Bukola, O., Adejoke, A. O., and Chidozie, O. P. (2017). Genetic assessment of the shrub *Synsepalum dulcificum* (Schumach & Thonn.) Daniell in Nigeria using the Randomly Amplified Polymorphic DNA (RAPD). *Int. J. Genet. Genomics* **4**, 45-50.

Covarrubias-Pazaran, G., Diaz-Garcia, K., Schlautman, b., Deutsch, j., Salazar, W., Hernandez-Ochoa, M., Grygleski, E., Steffan, S., Lorizzo, M., Polashock, J., Vorsa, N., and Zalapa, J. (2016). Exploiting genotyping by sequencing to characterize the genomic structure of the American cranberry through high-density linkage mapping. *BMC Genom.* **17**.

Dadjo, C., Nyende, A. B., Yao, N., Kiplangat, N., and Assogbadjo, A. E. (2020). Genome-wide genetic diversity and population structure of *Garcinia kola* (Heckel) in Benin using DArT-Seq technology. *PLoS ONE* **15**.

Decher, D., Norris, R. W., Abedi-Lartey, M., Oppong, J., Hutterer, R., Weinbrenner, M., Koch, M., Podsiadlowski, L., and Kilpatrick, C. W. (2021). A survey of small mammals in the Volta Region of Ghana with comments on zoogeography and conservation. *Zoosystema* **43**, 253-281.

Del Campo, R., Zhang, Y., and Wakeford, C. (2017). Effect of miracle fruit (*Synsepalum dulcificum*) seed oil (MFSO®) on the measurable improvement of hair breakage in women with damaged hair: a randomized, double-blind, placebo-controlled, eight-month trial. *J. Clin. Aesthet. Dermat.* **10**, 39-48.

Demenou, B. B., Doucet, J.-L., and Hardy, O. J. (2018). History of the fragmentation of the African rain forest in the Dahomey Gap: insight from the demographic history of *Terminalia superba*. *Heredity* **120**, 547–561.

Demenou, B. B., Pineiro, R., and Hardy, O. J. (2016). Origin and history of the Dahomey Gap separating West and Central African rain forest: insight from the phylogeography o the legume tree *Distemonanthus benthamianus*. *J. Biogeogr.* **43**, 1020-1031.

Denoyes, B., Amaya, I., Liston, A., Tennessen, J., Ashman, T.-L., Whitaker, V., Hytönen, T., van de Weg, E., Osorio, S., and Folta, K. (2016). Genomics tools available for unravelling mechanisms underlying agronomical traits in strawberry with more to come. *In* "VIII International Strawberry Symposium 1156", pp. 13-24.

Diaz-Garcia, L., Garcia-Ortega, L. F., González-Rodríguez, M., Delaye, L., Iorizzo, M., and Zalapa, J. (2021). Chromosome-level genome assembly of the American cranberry (*Vaccinium macrocarpon* Ait.) and its wild relative *Vaccinium microcarpum*. *Front. Plant Sci.* **12**.

Die, J. V., and Rowland, L. J. (2013). Advent of genomics in blueberry. *Mol. Breed.* **32**, 493-504.

Du, L., Shen, Y., Zhang, X., Prinyawiwatkul, W., and Xu, Z. (2014). Antioxidant-rich phytochemicals in miracle berry (*Synsepalum dulcificum*) and antioxidant activity of its extracts. *Food Chem.* **153**, 279-284.

Duminil, J., Brown, R. P., Ewédjè, E.-E. B. K., Mardulyn, P., Doucet, J.-L., and Hardy, O. J. (2013). Large-scale pattern of genetic differentiation within African rainforest trees:

insights on the roles of ecological gradients and past climate changes on the evolution of *Erythrophleum spp* (Fabaceae). *BMC Evol. Biol.* **13**.

Dupont, L. M., and Weinelt, M. (1996). Vegetation history of the savanna corridor between the Guinean and the Congolian rain forest during the last 15000 years. *Veg. Hist. Archaebot.* **5**, 273-292.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359-361.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. (2011). A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PloS One* **6**.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611-2620.

Ewédjè, E. (2012). Biologie de la reproduction, phylogéographie et diversité de l’arbre à beurre *Pentadesma butyracea* Sabine (Clusiaceae)-implications pour sa conservation au Bénin, Université Libre de Bruxelles, Bruxelles.

Feng, J., Zhao, S., Li, M., Zhang, C., Qu, H., Li, Q., Li, J., Lin, Y., and Pu, Z. (2019). Genome-wide genetic diversity detection and population structure analysis in sweetpotato (*Ipomoea batatas*) using RAD-seq. *Genomics* **112**, 1978-1987.

Foster, T. M., Bassil, N. V., Dossett, M., Worthington, M. L., and Graham, J. (2019). Genetic and genomic resources for *Rubus* breeding: A roadmap for the future. *Hortic. Res.* **6**.

Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* **17**, 27-32.

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.

Fu, Y. X., and Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709.

Fuchs, J., and Bowie, R. C. (2015). Concordant genetic structure in two species of woodpecker distributed across the primary West African biogeographic barriers. *Mol. Phylogenet. Evol.* **88**, 64-74.

Garcia-Seco, D., Zhang, Y., Gutierrez-Mañero, F. J., Martin, C., and Ramos-Solano, B. (2015). RNA-Seq analysis and transcriptome assembly for blackberry (*Rubus sp. Var. Lochness*) fruit. *BMC Genom.* **16**.

Gbedevi, K. M., Boukar, O., Ishikawa, H., Abe, A., Ongom, P. O., Unachukwu, N., Rabbi, I., and Fatokun, C. (2021). Genetic Diversity and Population Structure of Cowpea

[*Vigna unguiculata* (L.) Walp.] Germplasm Collected from Togo Based on DArT Markers. *Genes* **12**.

Gorin, S., Wakeford, C., Zhang, G., Sukamtoh, E., Matteliano, C. J., and Finch, A. E. (2018). Beneficial effects of an investigational wristband containing *Synsepalum dulcificum* (miracle fruit) seed oil on the performance of hand and finger motor skills in healthy subjects: A randomized controlled preliminary study. *Phytother. Res.* **32**, 321-332.

Goudet, J., Jombart, T., and Goudet, M. J. (2015). Package ‘hierfstat’. *In* "R package version 0.04-22. Retrieved from http://www. r-project. org, http://github. com/jgx65/hierfstat".

Gruber, B., Unmack, P. J., Berry, O. F., and Georges, A. (2018). dartR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* **18**, 691-699.

Gwali, S., Nakabonge, G., Okullo, J. B. L., Eilu, G., Nyeko, P., and Vuzi, P. (2012). Morphological variation among shea tree (*Vitellaria paradoxa* subsp. *nilotica*)‘ethnovarieties’ in Uganda. *Genet. Resour. Crop. Evol.* **59**, 1883-1898.

Hagen, O., Alexander Skeels, A., Onstein, R. E., Jetz, W., and Pellissier, L. (2021). Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proc. Natl. Acad. Sci. U.S.A.* **118**.

Hale, I., Ma, X., Melo, A. T., Padi, F. K., Hendre, P. S., Kingan, S. B., Sullivan, S. T., Chen, S., Boffa, J.-M., Muchugi, A., Danquah, A., Barnor, M. T., Jamnadass, R., van de Peer, Y., and van Deynze, A. (2021). Genomic resources to guide improvement of the shea tree. *Front. Plant Sci.* **12**.

Kalinowski, S. T. (2004). Counting alleles with raredaction: Private alleles and hierarchical sampling designs. *Cons. Genet.* **5**, 539-543.

Kamvar, Z. N., Tabima, J. F., Everhart, S. E., Brooks, J. C., Krueger-Hadfield, S. A., and Sotka, E. (2021). Package ‘poppr’.

Khayi, S., Gaboun, F., Pirro, S., Tatusova, T., El Mousadik, A., Ghazal, H., and Mentag, R. (2020). Complete chloroplast genome of *Argania spinosa*: structural organization and phylogenetic relationships in Sapotaceae. *Plants* **9**.

Kouam, E. B., Pasquet, R. S., Campagne, P., Tignegre, J.-B., Thoen, K., Gaudin, R., Ouedraogo, J. T., Salifu, A. B., Muluvi, G. M., and Gepts, P. (2012). Genetic structure and mating system of wild cowpea populations in West Africa. *BMC Plt. Biol.* **12**.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547.

Kurihara, K., and Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science* **161**, 1241-1243.

Leal, M. E. (2004). The African rain forest during the Last Glacial Maximum an archipelago of forests in a sea of grass, Wageningen University.

Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleid Acids Res.* **49**, W293-W296.

Linder, H. P., de klerk, H. M., Born, J., Burges, N. D., Fjeldas, J., and Rahbek, C. (2012). The partitioning of Africa: statistically defined biogeographical regions in sub-saharan Africa. *J. Biogeogr.* **39**, 1189-1205.

Manni, F., Guerard, E., and Heyer, E. (2004). Geographic patterns of (genetic,morphological, linguistic) variation: how much barriers can be detected by using Monmonier's algorithm. *Hum. Biol.* **76**, 173-190.

Masson, L. (2014). Phenolic acids as natural antioxidants. *In* "Phytochemicals of neutraceutical importance" (D. Praksh and G. Sharma, eds.), pp. 196-207. CAB International, Wallingford.

Melo, A. T., Bartaula, R., and Hale, I. (2016). GBS-SNP-CROP: a reference-optional pipeline for SNP discovery and plant germplasm characterization using variable length, paired-end genotyping-by-sequencing data. *BMC Bioinformatics* **17**, 29.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., and Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* **17**, 240-248.

Moreau, R. E. (1969). Climate changes ad the distribution of forest vertabrates in West Africa. *Zoology* **158**.

Mouhaddab, J., Msanda, F., Filali-Maltouf, A., Belkadi, B., Ferradouss, A., El Modafar, C., Koraichi, S. I., and El Mousadik, A. (2017). Using microsatellite markers to map genetic diversity and population structure of an endangered Moroccan endemic tree (*Argania spinosa* L. Skeels) and development of a core collection. *Plant Gene* **10**, 51-59.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A., and Kent, J. (2000). Biodiversity hostpots for conservation priorities. *Nature* **403**, 852-858.

Nelson, J. R., Verma, S., Bassil, N. V., Finn, C. E., Hancock, J. F., Cole, G. S., Knapp, S. J., and Whitaker, V. M. (2021). Discovery of three loci increasing resistance to charcoal rot caused by *Macrophomina phaseolina* in octoploid strawberry. *G3* **11**, 1-10.

Nicolas, V., Akpatou, B., Wendelen, W., Peterhans, J., Olayemi, J., Decher, J., Missoup, A. D., Denys, C., Barriere, P., Cruaud, C., and Colyn, M. (2010). Molecular and morphometric variation in two siblings species of the genus Praomys (Rodentia:Muridae): implications for biogeography. *Zool. J. Linnean Soc.* **160**, 397-419.

Nybom, H. (2004). Comparision of different nuclear markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* **13**, 1143 - 1155.

O'Connor, K., Kilian, A., Hayes, B., Hardner, C., Nock, C., Baten, A., Alam, M., and Topp, B. (2019). Population structure, genetic diversity and linkage disequilibrium in a macadamia breeding population using SNP and silicoDArT markers. *Tree Genet. Genomes* **15**, 1-16.

Pakhrou, O., Medraoui, L., Yatrib, C., Alami, M., Filali-Maltouf, A., and Belkadi, B. (2017). Assessment of genetic diversity and population structure of an endemic Moroccan tree (*Argania spinosa* L.) based in IRAP and ISSR markers and implications for conservation. *Physiol. Mol. Biol. Plants* **23**, 651-661.

Pembleton, L. W., Cogan, N. O., and Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Resour.* **13**, 946-952.

Pereira, V. M., Filho, J. A. F., Leao, A. P., Vargas, L. H. G., de Farias, M. P., Rios, S. A., da Cunha, R. N. V., Formighieri, E. F., Alves, A. A., and Junior, M. T. S. (2020). American oil palm from Brazil: Genetic diversity, population structure, and core collection. *Crop Sci.* **60**, 3212-3227.

Plana, V., Gascoigne, A., Forrest, L. L., Harris, D., and Toby Pennington, R. T. (2004). Pleistocene and pre-Pleistocene Begonia speciation in Africa. *Mol. Philogenet. Evol.* **31**, 449–461.

Poland, J., and Rife, T. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* **5**, 92-102.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Qi, X., Ogden, E. L., Bostan, H., Sargent, D. J., Ward, J., Gilbert, J., Lorizzo, M., and Rowland, L. J. (2021). High-density linkage map construction and qtl identification in a diploid blueberry mapping population. *Front. Plant. Sci.* **12**.

Ramos-Onsins, S. E., and Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* **19**, 2092-2100.

Richards, C. M., Volk, G. M., Reilly, A. A., Henk, A. D., Lockwood, D. R., Reeves, P. A., and Forsline, P. L. (2009). Genetic diversity and population structure in Malus sirversii, a wild progenitor species of domestocated apple. *Tree Genet. Genomes* **5**, 339-347.

Rosenbom, S., Costa, V., Al-Araimi, N., Kefena, E., Abdel-Moneim, A. S., Abdalla, M. A., Bakhiet, A., and Beja-Pereira, A. (2014). Genetic diversity of donkey populations from the putative centers of domestication. *Anim. Genet.* **46**, 30-36.

Rosly, H. A.-A. M., Nor, S. A. M., Yahya, K., and Naim, D. M. (2013). Mitochondrial DNA diversity of mud crab *Scylla olivacea* (Portunidae) in Peninsular Malaysia: a preliminary assessment. *Mol. Biol. Rep.* **40**, 6407-6418.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., and Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299-3302.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.

Salzmann, U., and Hoelzmann, P. (2005). The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *Holocene* **15**, 190-199.

Sanou, H., Lovett, N. P., and Bouvet, J.-M. (2005). Comparison of quantitative and molecular variation in agroforestry populations of the shea tree (*Vitellaria paradoxa* C.F. Gaertn) in Mali. *Mol. Ecol.* **14**.

Sanou, H., Picard, N., Lovett, P. N., Dembélé, M., Korbo, A., Diarisso, D., and Bouvet, J.-M. (2006). Phenotypic variation of agromorphological traits of the shea tree, *Vitellaria paradoxa* CF Gaertn., in Mali. *Genet. Resour. Crop Evol.* **53**, 145-161.

Sansaloni, C. P., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., and Kilian, A. (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: Genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proc.* **5**.

Slimane, K., Elhouda, A. N., Fatima, G., Pirro, S., Oussama, B., Gonzalo, C. M., Lightfoot, D. A., Turgay, U., Bouchra, C., and Redouane, M. (2020). First draft genome assembly of the Argane tree (*Argania spinosa*). *F1000Res.* **7**.

Swenson, U., Richardson, J. E., and Bartish, I. V. (2008). Multi-gene phylogeny of the pantropical subfamily Chrysophylloideae (Sapotaceae): evidence of generic polyphyly and extensive morphological homoplasy. *Cladistics* **24**, 1006–1031.

Szczecińska, M., Sramko, G., Wołosz, K., and Sawicki, J. (2016). Genetic diversity and population structure of the rare and endangered plant species *Pulsatilla patens* (L.) Mill in East Central Europe. *PLoS ONE* **11**.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.

Tajima, F., and Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* **1**, 269-285.

Tchokponhoué, D., Achigan-Dako, E., N'Danikou, S., Houdégbé, A., Agossou, C., Assogba-Komlan, F., and Vodouhè, R. (2018). Regeneration ability and seedling growth in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Fruits* **73**, 13-21.

Tchokponhoué, D. A., Achigan-Dako, E. G., N'Danikou, S., Nyadanu, D., Kahane, R., Houéto, J., Hotegni Fassinou, N. V., Odindo, A. O., and Sibiya, J. (2020). Phenotypic variation, functional traits repeatability and core collection inference in *Synsepalum* *dulcificum* (Schumach & Thonn.) Daniell reveals the Dahomey Gap as a centre of diversity. *Sci. Rep.* **10**, 1-17.

Tchokponhoué, D. A., N'Danikou, S., Hale, I., Van Deynze, A., and Achigan-Dako, E. G. (2017). Early fruiting in *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell juveniles induced by water and inorganic nutrient management. *F1000Research* **6**.

Tchokponhoué, D. A., N'Danikou, S., and Achigan-Dako, E. G. (2019). A combination of approaches evidenced seed storage behaviour in the miracle berry *Synsepalum dulcificum* (Schumach. et Thonn.) Daniell. *BMC Plt. Biol.* **19:117**, 1-13.

Teixeira, H., Montade, V., Salmona, J., Metzger, J., Bremond, L., Kasper, T., Daut, G., Rouland, S., Ranarilalatiana, S., Rakotondravony, R., Chikhi, L., Behling, H., and Radespiel, U. (2021). Past environmental changes affected lemur population dynamics prior to human impact in Madagascar. *Commun. Biol.* **4**.

Turck, D., Castenmiller, J., De Henauw, S., Hirsch-Ernst, K. I., Kearney, J., Maciuk, A., Mangelsdorf, I., and McArdle, H. J. (2021). Safety of dried fruits of *Synsepalum dulcificum* as a novel food pursuant to Regulation (EU) 2015/2283. *EFSA J.* **19**.

Walters, R. J., Robinson, T. P., Byrne, M., Wardell-Johnson, G. W., and Neville, P. (2020). Contrasting patterns of local adaptation along climatic gradients between a sympatric parasitic and autotrophic tree species. *Mol. Ecol.* **29**, 3022-3037.

Wei, X., Shen, F., Zhang, Q., Liu, N., Zhang, Y., Xu, M., Liu, S., Zhang, Y., Ma, X., and Liu, W. (2021). Genetic diversity analysis of Chinese plum (*Prunus salicina* L.) based on whole-genome resequencing. *Tree Genet. Genomes* **17**, 1-10.

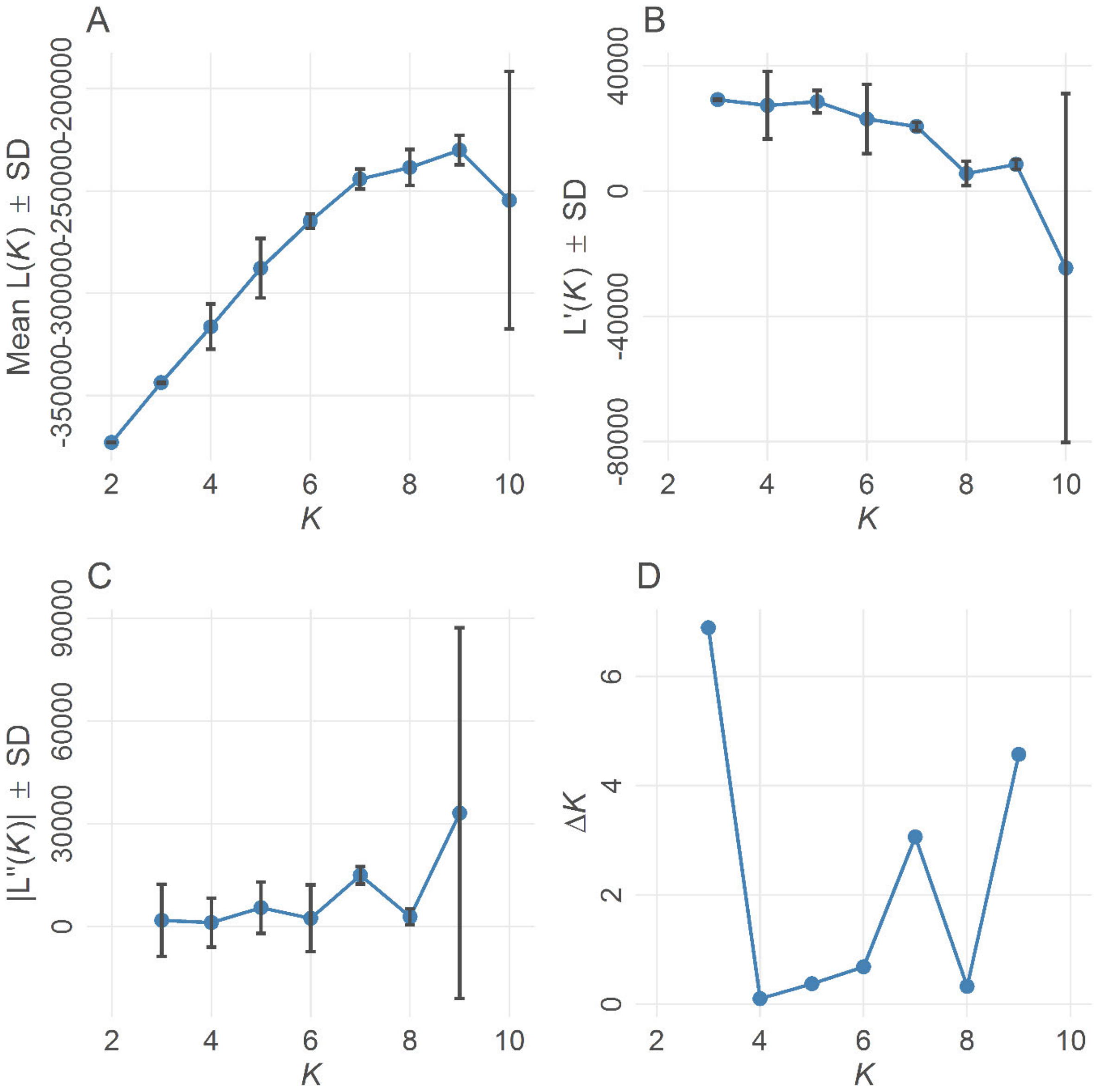
Weir, B. S., and Cockerman, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Welker, B. H. (2017). "The History of our tribe: Hominini," Open SUNY Textbooks, Geneseo, USA.

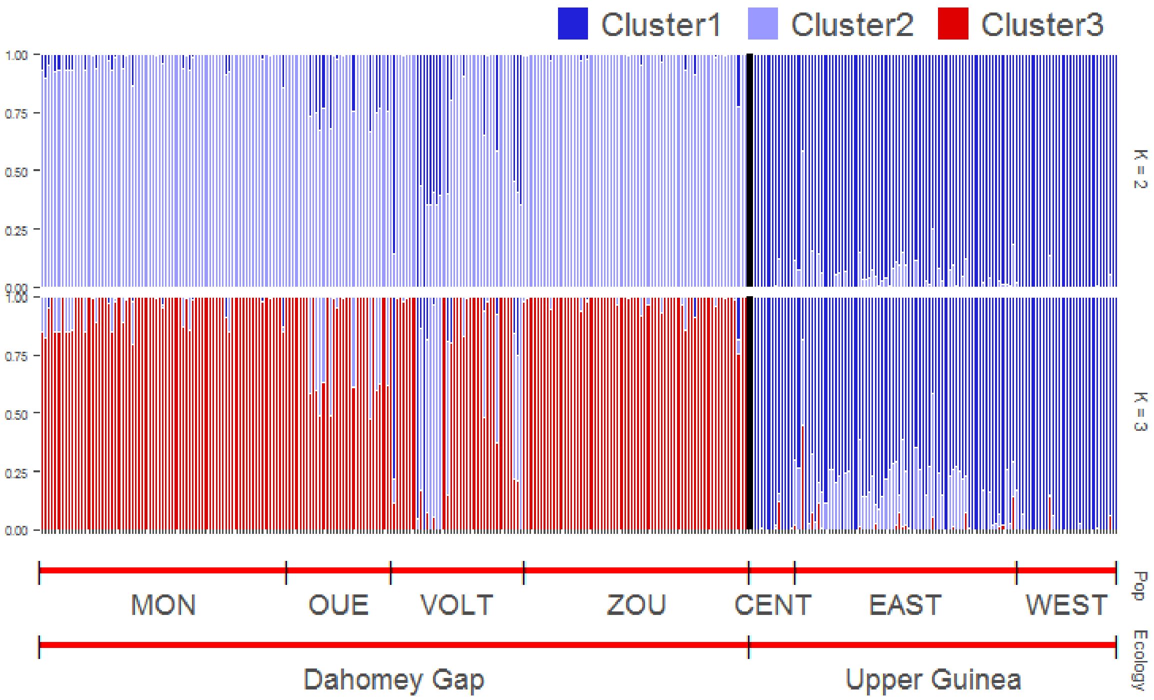
Wilken, M. K., and Satiroff, B. A. (2012). Pilot study of “miracle fruit” to improve food palatability for patients receiving chemotherapy. *Clin. J. Oncol. Nurs.* **16**, E173-E177.

Wright, S. (1978). "Evolution and genetics of populations: variability within and among natural population," University of Chicago press.

Yuan, Z.-H., Chen, X. S., Zhang, C. Y., He, T. M., Feng, J. R., and Feng, T. (2008). Population genetic structure in apricot (*Armeniaca* Mill.) revealed by Fluorescent-AFLP markers. *Acta Hortic. Sin.* **35**, 319-328.



Appendix 5.1. Evanno plots indicating the mean likelihood L(K) and variance per K (A), the mean rate of change of the likelihood distribution (B), the absolute value of the 2nfd order rate of change of the likelihood distribution (C) and Delta K = mean (|L” (K) / sd (L(K)) (D) from structure on 322 accessions and 2704 SNPs in miracle plant (*Synsepalum dulcificum*).



Appendix 5.2. Combined bar plots showing population splitting pattern from K = 2 to K= 3 in miracle plant (*Synsepalum dulcificum*). West: Western population; CENT: Central population; EAST: Eastern population; VOLT: Volta population; MON: Mono population; ZOU: Zou population and OUE: Oueme population.

CHAPTER 6 [4]

Use patterns, knowledge diversity and drivers for the cultivation of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] in Benin and Ghana

Abstract

Despite the growing interest in the miracle plant worldwide due to its numerous applications, the threats and the wild harvest of the species hamper its sustainable utilisation. Moreover, traditional knowledge so far documented on the species is limited to a narrow geographical coverage of its natural distribution range, which is West and Central Africa. This study analysed the use variation and knowledge acquisition pattern of the miracle plant among West African sociolinguistic groups and deciphered the drivers of populations' willingness and readiness to engage in cultivating the species. Semi-structured interviews were conducted with 510 respondents purposively selected from nine sociolinguistic groups in Benin and Ghana using the snowball sampling approach. Information was collected on respondents' socio-demographic profile, miracle plant ownership, plant parts used and preparation methods, knowledge of the species bioecology, perceived threats on the species, willingness to cultivate, maximum acreage to allocate to the species and maximum price to pay for a seedling. Descriptive statistics, generalized linear models, classification and regression tree models were used for data analysis. The miracle plant ownership mode depended on the age category. Sociolinguistic affiliation, level of schooling, migratory status and religion significantly affected the number of trees owned. We recorded 76 uses belonging to six use categories. The overall use-value of the miracle plant significantly varied according to the respondent sociolinguistic affiliation, main activity and religion. Men were the main source of knowledge and knowledge is mainly acquired along the family line. Knowledge related to food and social uses was mostly acquired from parents and people of the same generation, while magico-therapeutic and medicinal use-related knowledge were inherited from parents and grandparents. Sociolinguistic affiliation, awareness of taboos and market availability were the most important drivers of respondent willingness to cultivate the miracle plant. While the respondent's level of schooling and perception of plant growth rate determined the maximum acreage, they were willing to allocate to the species in cultivation schemes, their main activity,

---

[4] Chapter published as: **Tchokponhoué, D. A.**, N'Danikou, S., Fassinou Hotegni, N. V., Nyadanu, D., Kahane, R., Odindo, A. O. & Sibiya, J. (2021). Use patterns, knowledge diversity and drivers for the cultivation of the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] in Benin and Ghana. Plants, 10(11), 2253. **https://doi.org/10.3390/plants10112253**

sociolinguistic affiliation and knowledge of the species time to fruiting drove the maximum purchase price they were willing to offer for a seedling of the species. Our findings provide key information for the promotion of miracle plant cultivation in the study area.

6.1 Introduction

The miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell [Syn: *Richardella dulcifica* (Schumach & Thonn.) Baehni] is a slow-growing West African native tree species belonging to the Sapotaceae family (Tchokponhoué et al., 2018). It naturally thrives in West and Central Africa where it grows on well-drained acidic soils with pH ranging between 4.5. and 5.8 (Achigan-Dako et al., 2015). The species is an intermediate shade-tolerant species (Tchokponhoué et al., 2019) that grows well in warm, wet and humid areas. In West Africa (e.g., Benin and Ghana), the species is found in gallery forests, home gardens, backyards, fallows and farms (Adomou, 2005; Tchokponhoué et al., 2021), and is semi-domesticated, benefiting from weeding, fertilization and pruning as main management practices (Tchokponhoué et al., 2021). *Synsepalum dulcificum* exhibits the highest significance for human wellbeing among the nearly thirty species in the *Synsepalum* genus (Rokni et al., 2019). The species is well-known for being a unique natural source of “miraculin”, a taste modifying glycoprotein contained in the miracle fruit pulp, which has a sweetening effect on any sour food (Kurihara and Beidler, 1968).

The miracle plant has numerous modern and traditional applications in the food and beverage, cosmetics and pharmaceutical industries (Achigan-Dako et al., 2015). In the food and beverage industry, the fruit is used as a non-nutritive natural sweetener and beverage colourant (Buckmire and Francis, 1978), a reliable substitute to synthetic sugar in lemonade and juice (Rodrigues et al., 2016) and an ingredient in functional yoghurt preparation (Fazilah et al., 2020). In cosmetics, the seed oil is used to prevent hair breakage (Del Campo et al., 2017) and improved finger motor skills (Gorin et al., 2018). It is also utilized in the prevention and treatment of diabetes and cancer, two non-communicable diseases with heavy socio-economic burdens (Bommer et al., 2018; Bray et al., 2018). The miracle fruit improves insulin resistance and synthesis [15,16] and restores taste perception in patients undergoing chemotherapy treatment (Wilken and Satiroff, 2012). Consequently, the species could gain expansions of its role in the global health system as projections anticipate 237 million and 6.9 million new cases of diabetes and cancer respectively, by 2045–2050 (International Diabetes Federation, 2019; Pilleron et al., 2021). In addition, high demand for the species in the food and beverage industry is expected due to the increasing pressure and demand for non-nutritive sweeteners. Economically, the miracle fruit has a reeling market value in the USA where a kilogram of the pure powder costs USD 2500 USD (https://www.miraclefruitfarm.com/supplements) and increasing effort is being made for its formal approval in the European Union market (Menéndez-Rey et al., 2021).

Historically, in West Africa, the miracle fruit has assisted in the consumption of sour foods, unripe fruits and to sweeten local beverages (Inglett and May, 1968), whereas all the non-edible parts of the plant are reported to have nearly 64 medicinal applications (e.g., malaria, tuberculosis and cough treatments) (Fandohan et al., 2017). Despite its occurrence in more than 10 West and Central African countries, the only proper ethnobotanical investigations conducted on the species were limited to some sociolinguistic groups in Benin (Fandohan et al., 2017; Oumorou et al., 2010) and only sparsely documented information exists on its uses from Nigeria, Togo and Côte d’Ivoire (Ekpo et al., 2008; Malan and Neuba, 2021). Consequently, there is a paucity of information on the species’ traditional ecological knowledge, that is, the body of knowledge accumulated on the species by local communities through history, by means of direct experience and contact with nature and transmitted from generation to generation (Agbani et al., 2018; Si, 2020). The information is necessary to gauge the sustainability of exploitation and contribute to profiling adequate conservation measures (Agbo et al., 2020; Leiper et al., 2018), especially in the centre of origin of the species. Existing reports on the miracle plant reveal that local populations hold a strong knowledge of medicinal applications (Oumorou et al., 2010). Key to the sustainability of these uses is the understanding of the mode of knowledge acquisition and or transmission. However, these have not been documented in the literature.

Agricultural expansion, population growth and overharvesting of plant parts for medicinal purposes are among the threats to the miracle plant in its centre of origin (Adomou, 2005; Fandohan et al., 2017). In parallel, current exploitation of the species in West Africa mainly rely on just a few tree stands in farmers’ backyards, home gardens and farms (Tchokponhoué et al., 2021). Taken together, these two observations suggest that the current supply is unlikely to meet the growing local and international demands for products of the species (e.g., leaves, roots, seeds and fruits) (Gorin et al., 2018; Menéndez-Rey et al., 2021; Wilken and Satiroff, 2012) and to hold its promise as a lever of West Africa economic growth. Cultivation was suggested as a sustainable alternative to semi-wild or wild harvest of plants when demand outgrows natural populations capacities (N’Danikou et al., 2015). Exploring options for cultivation (as opposed to semi-wild or wild harvest) is therefore desirable in anticipating the predicted increased demand for the potential products that can be derived from the species. For this to be effective, a clear understanding of the factors shaping farmers’ decision making to invest in the cultivation of the species is required. Lessons learnt from previous studies investigating determinants for the cultivation of perennial plants by local populations in West Africa concluded on the prominence of ethnicity, gender, age, instruction level and knowledge of tree biology, as key drivers (Dadjo et al., 2020; Gandji et al., 2018; N'Danikou et al., 2015). *Synsepalum dulcificum* being of the same nature, these factors are expected to play an important role in farmers' decision making to cultivate.

The objectives of this study were, therefore, to analyse the use variation, assess the knowledge acquisition pattern across sociolinguistic groups, gender, level of schooling and age categories and depict factors affecting farmers' decision to engage in the cultivation of *S. dulcificum*. Specifically, the following research questions were addressed: (i) Do differences in the respondents' socio-demographic background affect ownership of trees and if so how? (ii) To what extent do the socio-demographic factors affect the species use pattern in West Africa? (iii) What is the knowledge acquisition pattern in the miracle plant, and does it harbour any peculiarities? (iv) What are the most critical drivers of the local population's willingness and readiness to engage in the cultivation of the miracle plant, and how do they interact with each other?

6.2 Materials and methods

6.2.1 Study area

This study was conducted from April to December 2019 in Benin (Latitudes 6°–12°50′ N, Longitudes 1°–3°70′ E) and Ghana (Latitudes 4.7°–11°30′ N, Longitudes 2°50′ W–1°40′ E) (**Fig. 6.1**), two West African countries known to belong to the centre of origin of the miracle plant. In Benin, the species is confined to the Guineo-Congolian region, in the Northern Guinea zone of West Africa. The region is characterised by a humid tropical climate with an annual rainfall, temperature and relative humidity ranging between 900 and 1300 mm; 25 °C and 29 °C and 69 and 97%, respectively. The dominant vegetation in the region includes savannah, woodland mosaic and relics of dry semi-deciduous forest on ferralitic and ferruginous soils (Achigan-Dako et al., 2011; Adomou, 2005). The region is mainly dominated by Kwa- and Yoroboid-speaking groups with more than 10 sociolinguistic groups including Fon, Sahouè, Adja, Aizo, Nagot and Holli, among others (Heldmann, 2008). In Ghana, the miracle plant is found in the Transitional and the Deciduous forest (Tchokponhoué et al., 2021), shared between the Northern and Southern Guinea ecological zone of West Africa and in the Tropical Rainforest zone of West Africa, which are relatively moister than the Guineo-Congolian zone of Benin. The annual precipitation in these regions ranges from 800 mm to 2800 mm with a mean annual temperature between 26.1–26.4 °C (Abbam et al., 2018). These regions are dominated by deciduous tropical forests and lush forested vegetation cover, whereas the main soil types include nitosols, acrisols and ferralsols (Obeng, 1971). Major sociolinguistic groups include Akan, Ewe and Ga-adangbe.

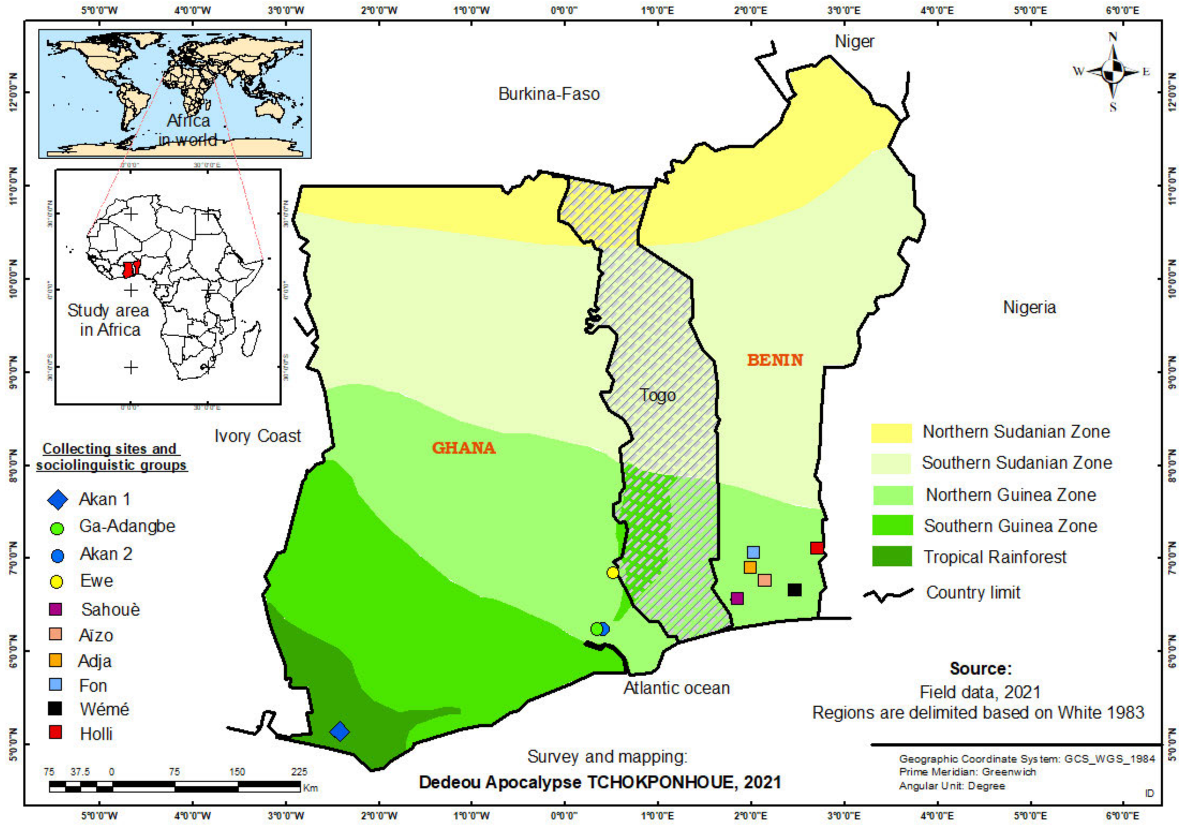


Figure 6.1. Map showing the study area.

6.2.2 Respondents' sampling

Nine sociolinguistic groups (six in Benin and three in Ghana, **Table 6.1**) that use the species (Fandohan et al., 2017; Tchokponhoué et al., 2021) were targeted in this study. As we were only interested in the actual use-value of the species, and because the theoretical knowledge might not be systematically translated into actual use value (De Los Angeles La Torre-Cuadros and Islebe, 2003), respondents were purposively sampled using the snowball technique. A respondent comprised a person who: (i) was able to identify the species including recognizing the miracle fruit from a photograph and knowing its local names, (ii) effectively used at least one part of the miracle plant in his/her lifetime, and (iii) gave his/her prior and informed consent to participate in the study. In total, 510 respondents across the groups, with an average between 51 and 87 per sociolinguistic group were interviewed (**Table 6.1**).

Table 6.1. Proportion of the respondents falling within different socio-demographic characteristics in Benin and Ghana.

		Ghana			Benin						
Factor	Modalities	Akan (n = 53)	Ewe (n = 53)	Ga-adangbe (n = 51)	Adja (n = 54)	Aizo (n = 56)	Fon (n = 87)	Holli (n = 55)	Sahouè (n = 50)	Wémé (n = 51)	Total (N = 510)
Gender	Men	84.91	84.91	86.27	92.59	92.86	88.51	98.18	88.00	94.12	90.00
	Women	15.09	15.09	13.73	7.41	7.14	11.49	1.82	12.00	5.80	10.00
Age category	Young: <30 years	5.66	3.78	0.00	7.41	1.79	6.90	56.36	0.00	0.00	9.00
	Adult: 30–59 years	54.72	52.83	56.86	61.11	48.21	60.92	38.18	50.00	58.82	54.00
	Old: ≥60 years	39.62	43.39	43.14	31.48	50.00	32.18	5.46	50.00	41.18	37.00
Level of schooling	No-schooling	5.66	3.77	21.57	51.85	51.79	47.13	76.36	50.00	31.37	38.00
	Literate	3.77	0.00	1.96	7.41	0.00	2.30	7.27	0.00	1.96	03.00
	Primary	13.21	16.98	27.45	16.67	25.00	28.73	14.55	24.00	37.25	22.00
	Secondary	67.93	69.81	47.06	24.07	23.21	17.24	1.82	20.00	23.54	33.00
	>Bac *	9.43	9.44	1.96	0.00	0.00	4.60	0.00	6.00	5.88	04.00
Main activity	Farming	84.91	94.34	98.04	85.19	71.43	54.02	87.27	90.00	72.55	82.00
	Handcraft	9.43	3.77	1.96	7.41	10.71	25.29	1.82	2.00	9.80	08.00
	Teaching	1.89	1.89	0.00	0.00	3.57	6.90	0.00	6.00	1.97	03.00
	Trading	3.77	0.00	0.00	3.70	3.57	9.20	0.00	0.00	7.84	03.00
	Traditional healing	0.00	0.00	0.00	3.70	10.72	4.59	10.91	2.00	7.84	04.00
Religion	Christian	98.11	98.11	100.00	38.89	55.36	44.83	7.27	34.00	62.75	60.0
	Indigenous	0.00	0.00	0.00	61.11	44.64	55.17	92.73	66.00	27.45	39.00
	Muslim	1.89	1.89	0.00	0.00	0.00	0.00	0.00	0.00	9.80	01.00
Migratory status	Autochthon	83.02	92.45	94.12	92.59	96.43	87.36	100.00	100.00	92.16	93.00
	Allochthon	16.98	7.55	5.88	7.41	3.57	12.64	0.00	0.00	7.84	07.00

6.2.3 Data collection

Face-to-face semi-structured interviews based on a questionnaire were used for the data collection. A verbal agreement was obtained from traditional authorities of local communities prior to administering the questionnaire. Interviews were conducted in the respondent's preferred languages, which were either his/her local language, French, or English. Direct interviews using either English or French were done only when the respondents desired so. Otherwise, each interviewer (where necessary) was accompanied by a well-trained local guide (who understood both French/ English and the interviewee-spoken language) to ease the questions/answers translation. The data collected were related to (i) the socio-demographic information (gender, age, sociolinguistic affiliation, religion, migratory status, land ownership, level of schooling and main activity) of the respondents, (ii) the local names of the miracle plant and their meanings, (iii) the miracle plant ownership status, mode of acquisition and number of trees owned, (iv) the plant parts used and use forms, (v) the method of preparation for each indicated use, (vi) the source of knowledge for each use mentioned, (vii) the respondents' perception of the conservation status of the species and the perceived threatening factors if any, (viii) the knowledge of the species' bioecology, (ix) the taboos (prohibitions imposed by social custom or as a protective measure (Colding and Folke, 2001) and superstitions (beliefs not based on reason or knowledge (Ng et al., 2010)) on the species, (x) the respondents' willingness to engage into cultivating the species, (xi) the maximum acreage the respondent is ready to allocate for the cultivation of the species (when answer to the question (x) is yes) and (xii) the respondent-proposed maximum purchase price of a seedling of the species.

6.2.4 Data analysis

The analyses were performed in the R environment (V. 3.6.2) (R Core Team, 2019). The socio-demographic profile of respondents was summarized using descriptive statistics (mean, standard error of the mean, frequency) computed with the functions *descript ()* and *crosstab ()* of the package "misty" (Yanagida, 2020). Content analyses were conducted to assess the species local names variation across sociolinguistic groups. The difference in the preponderance of the recorded miracle plant acquisition modes was tested using a Chi-square multiple proportion comparison test with the function prop. test *()* of base R. Similarly, the dependence between socio-demographic characteristics and the miracle plant acquisition mode was tested using either a Chi-square test or a Fisher's exact test (to account for contingency table containing values lower than 5) of independence. To test if the number of trees owned were affected by the respondents' socio-demographic characteristics, we used a Poisson- or quasi-Poisson- (to account for overdispersion) fitted generalised linear model (glm). To assess the respondent's knowledge of the species bioecology, we firstly used a Chi-square test to assess the existence of any association between miracle plant habitat and the studied sociolinguistic groups. Secondly, a multiple proportion comparison test was used to understand the relative importance of different propagation modes reported for the species in the study area. The same test was also used to depict the perception of the respondents on the species growth trend, whereas the difference in the species time to fruiting as indicated by respondents from various sociolinguistic groups was tested using a Poisson-fitted glm. Medicinal uses were classified into body systems following the International Classification of Primary Care 2 (ICPC2, 2005) and in use categories. We used the function *ethnoChord ()* of the "EthnobotanyR" package (Whitney, 2019) to graphically illustrate the association between sociolinguistic groups and use categories. To assess the extent of species valuation by each respondent, the species use value per respondent (UVj) was computed following **Equation (6.1):**

$$\mathrm{UV_j} = \textstyle\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{UR_i} \qquad \textbf{(Eqt. 6.1)}$$

Where UR_i is a specific use report by the respondent j and n is the total number of use reports mentioned.

From UV_J, the total use-value of the species was computed afterwards following Phillips and Gentry (1993) using **Equation (6.2)**:

$$\mathrm{UV} = \mathrm{UV_j}/\mathrm{N} \qquad \textbf{(Eqt. 6.2)}$$

where N is the total number of respondents.

The variation of UV following the different socio-demographic factors was analysed by fitting a glm with a Poisson/quasi-Poisson-error structure. The use value and the number of trees owned were correlated using Spearman/Pearson correlation, while the statistical difference of correlations among attributes of each sociodemographic factor was tested using a Fisher's R-to-transformation. To assess the specificity of plant parts for their use in a particular category or to treat a particular body system, we computed the informant agreement ratio (IAR) following Bakwaye et al. (2013) using **Equation (6.3)**:

$$\mathrm{IAR} = (\mathrm{nr_i} - \mathrm{na_i})/(\mathrm{nr_i} - 1) \qquad \textbf{(Eqt. 6.3)}$$

where nr_i is the total number of citations recorded for a given use category or body system i and na_i is the total number of different plant parts that are employed in this specific use category or body system i.

To analyse the knowledge acquisition pattern in the species, we developed a framework called "KMTO" that brought together four knowledge attributes including (i) the source or kernel of the knowledge "K", (ii) the knowledge mutation form "M", (iii) the knowledge acquisition type "T" and (iv) the knowledge acquisition order "O". In this model, the source of the knowledge K had two states: "internal" where the respondent acquired the knowledge along his/her family line (e.g., from a nephew, the grandfather, a cousin, etc.) and "external" where the knowledge is acquired from a source other than the respondent's family (e.g., the community or a friend). The knowledge mutation form "M" is either a "transition" where the knowledge transmitter and the respondent have the same sex (men from men or women from women) or a "transversion" where the transmitter and the respondents were of different sex (men from women or women from men). The acquisition type "T" has three states: "vertical" where the knowledge was acquired from another generation (e.g., respondent acquiring the knowledge from his father or grandmother), "horizontal" where the knowledge was acquired from the same generation (e.g., respondent acquiring the knowledge from a sister or a cousin or a person of the same generation) and "transversal" in which the knowledge was acquired through a self-learning experiment (e.g., knowledge acquired from a book). The knowledge transmission order "O" is defined as either "1" (respondent acquiring the knowledge from a direct progenitor: father or mother), "2" (the respondent acquired the knowledge for the grandfather/ grandmother) or "0" (characteristic of horizontal and transversal knowledge acquisition type). A graphical illustration of this KMTO framework is presented in **Fig. 6.2**. Following this framework, specific knowledge can be acquired following 36 different paths. We compared the overall importance of each knowledge attribute state using a multiple proportions comparison with the function *prop.test ()*. A Chi-square test or a Fisher.exact test of independence was used to test the association between the predominant state of each knowledge attribute and the species use category.

To identify the drivers of respondents' decision making to engage or not in the cultivation and depict the interactions among these factors we built a classification tree model using the function *rpart ()* with the "*class*" method in the "rpart" package (Therneau et al., 2015). When the respondent is favourable for the cultivation, the drivers of cultivation readiness assessed through the maximum acreage to allocate and the maximum price to pay for a seedling were assessed by building a regression tree model, still in the "rpart" package, but with the "*anova*" method this time. Finally, the function *plot.rpart ()* of the "rpart.plot" package (Milborrow and Milborrow, 2020) was used to visualize trees generated for the classification and regression tree models.

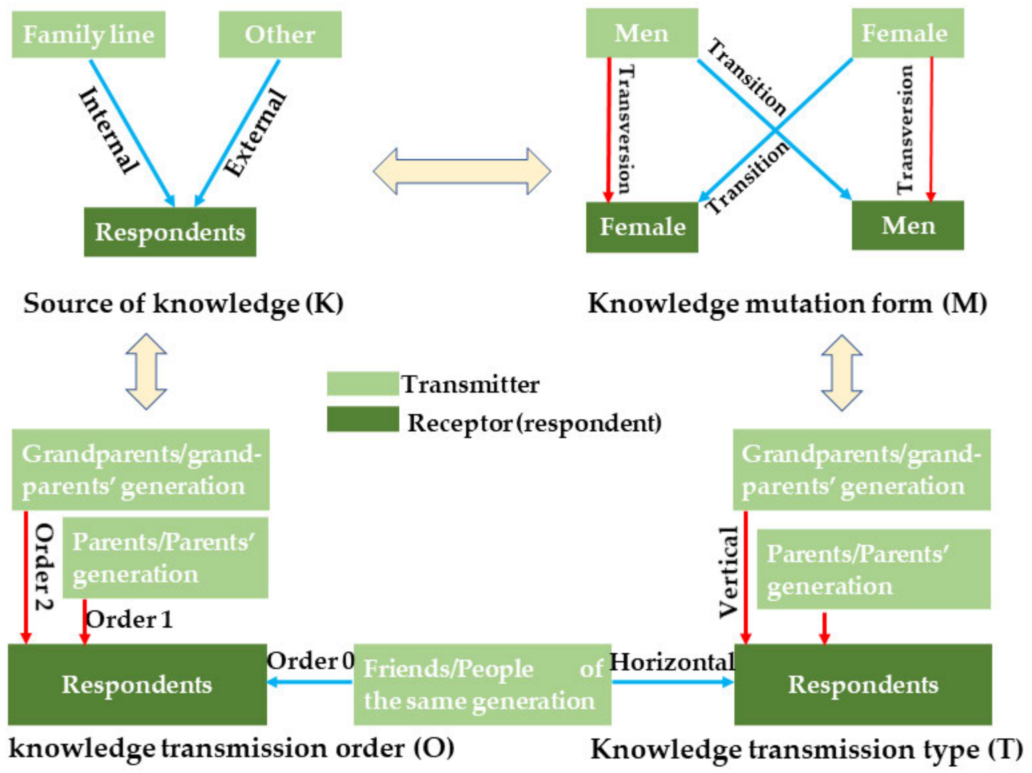


Figure 6.2. Graphical illustration of the Knowledge Mutation Type Order (KMTO) analysis framework.

6.3 Results

6.3.1 Socio-demographic characteristics of respondents

The proportion of women (10%) involved in the study was significantly lower than that of men (90%) ($\chi^2 = 649.6, df = 1, p < 0.001$) (**Table 6.1**). More than 50% of respondents were 30–59 years old, whereas the young (less than 30-years old) were the least represented. Respondents were on average 55 ± 0.68 years old with a higher proportion of autochthon compared with allochthon ($\chi^2 = 649.6, df = 1, p < 0.001$), and most of them did not school with farming as the main activity. Most of the respondents were Christians or practised indigenous religions (e.g., the Thron deity in Benin).

6.3.2 *Synsepalum dulcificum* nomenclature, ownership pattern and bioecology

The miracle plant is referred to as Sinssi by the *Adja*; Sislè by the *Aizo*, *Fon* and *Sahouè;* Agbanyun by the *Holli*; Sièsiè by the *Wémé* (Benin); Etimea by the *Akan*; Eliku by the *Ewe*, and Atanmanmi by the *Ga-adangbe* (Ghana). Several other name variants and their meanings were recorded in the two countries (**Table 6.2**).

Table 6.2. Diversity and meaning of the miracle plant (*Synsepalum dulcificum*) nomenclature in Benin.

Country	Sociolinguistic Groups	Local Names	Meaning
Ghana	Akan	Etimea*, Asaaba, Aswan	Something sweet
	Ewe	Elẽ*, Eliku*, Elindidi	Something naturally sweet
	Ga-adangbe	Atanmanmi*	Something sweet
Benin	Adja	Sinssi*	-Something sweet -Something simultaneously sweet and bitter -Hard and long-living element
	Aizo	Sièsiè*, Sisrè	Something sweet
	Fon	Sislè*, Sisrè*	A naturally sweet fruit Something whose fruits are sweet but the leaves and root are bitter The co-spouses No other taste than sweet in contact with the tongue
	Holli	Agbanyun*	Honey's calabash
	Sahouè	Sisrè*	Something sweet
	Wémé	Sièsiè*	Something sweet every time

*Dominant name in each sociolinguistic.

A total of 366 respondents (nearly 72%) owned miracle plant trees. Four ownership modes were recorded including (i) legacy where the respondent inherited the trees from his/her father or grandfather, (ii) self-established where the trees were planted by the respondent himself/herself, (iii) the combination of legacy and self-established and (iv) the spontaneous establishment where the species emerged on the respondents' land likely following natural dispersion. Nearly 50% of the respondents inherited the trees ($\chi^2 = 88.427$, $df = 3$, $p < 0.0001$), while self-establishment of trees was recorded in 40% of respondents. A very significant dependence was observed between the miracle plant ownership mode and the age category ($p = 0.004$), the sociolinguistic membership ($p = 0.008$) and the migratory status ($p = 0.002$) of the respondent. While the elderly respondents generally planted the miracle plant themselves, adults and youth mainly inherited them. Likewise, *Akan* and *Ga-adangbe* mostly established their own trees, whereas respondents in other sociolinguistic groups mainly inherited them. Allochthons mainly planted the trees, while autochthons inherited them. No association was observed between the miracle plant ownership mode and gender, level of schooling, or main activity ($p > 0.05$).

The Poisson-modelled generalised linear model indicated that the number of miracle trees owned was significantly affected by the respondent's sociolinguistic affiliation, migratory status, level of schooling and religion (**Fig. 6.3**). For instance, the *Akan* held nearly seven-fold more trees (14.15 ± 4.42 trees) than the *Holli* (1.68 ± 0.38 trees) (**Fig. 6.3B**, *t-value* = 4.24, $p < 0.0001$). Similarly, respondents with a high level of schooling tended to own more trees than those with a lower level of schooling (**Fig. 6.3D**, $t = -3.62$, $p = 0.0003$), while allochthons had on average more miracle plant trees than autochthons (**Fig. 6.3F**, $t = -3.90$, $p = 0.0001$). Christians also owned more trees than Muslim and Indigenous religious practitioners (**Fig. 6.3G**, $t = -2.76$, $p = 0.006$). Age category only exerted a marginal effect on the number of trees owned (**Fig. 6.3C**, $t = 1.24$, $p = 0.06$), whereas neither the gender (**Fig. 6.3A**) nor the main activity (**Fig. 6.3E**) significantly affected the number of trees held by the respondents ($p > 0.05$).

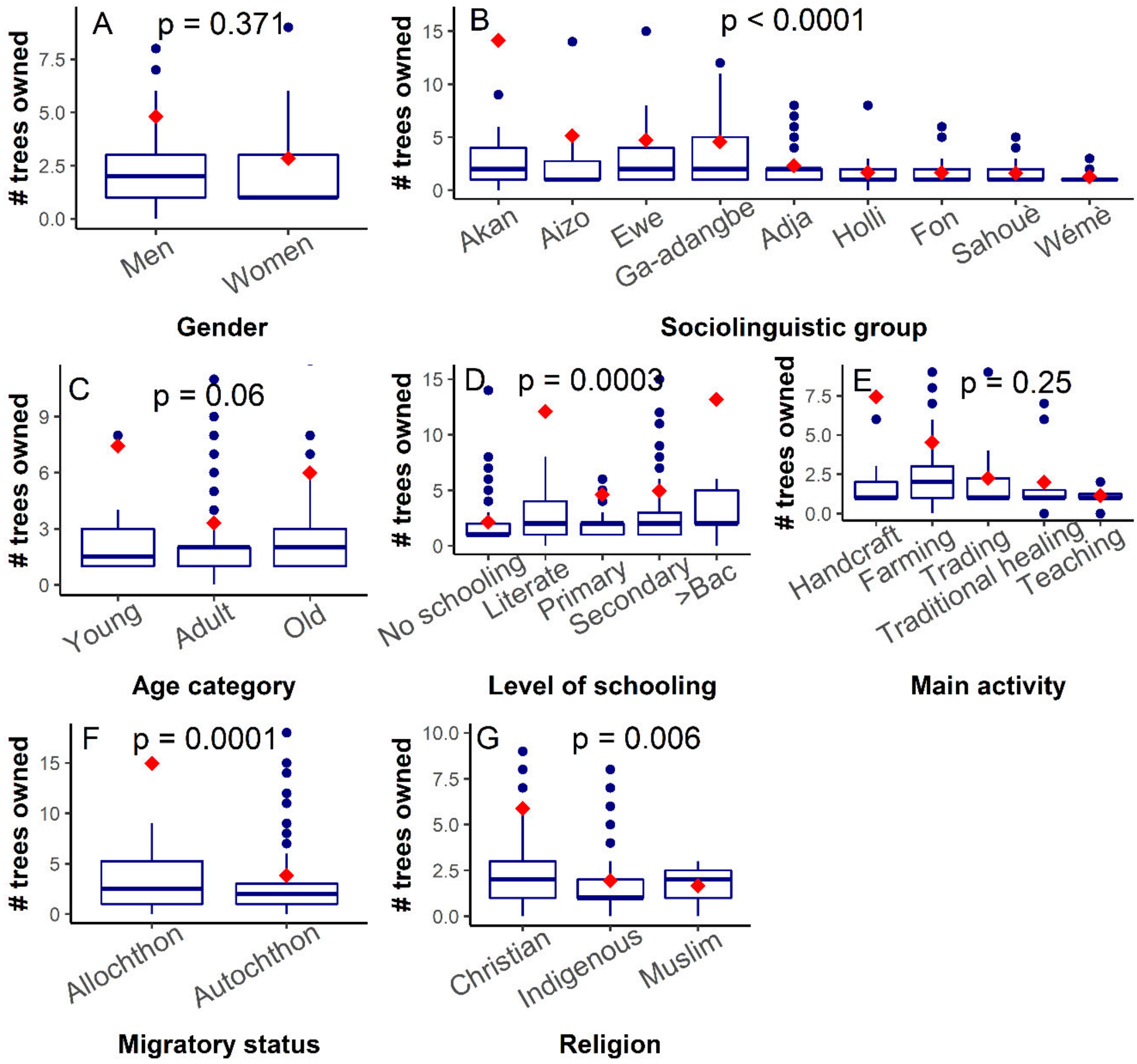


Figure 6.3. Variation of the number of miracle plant (*Synsepalum dulcificum*) trees owned following various socio-demographic factors in Benin and Ghana. Median values are in bold, red diamond shapes represent mean values, dots above and below boxplots are outliers, and lower and upper tails represent minimum and maximum values, respectively.

The miracle trees were recorded in two types of habitats, namely home gardens and farms. The species main habitat was significantly associated to respondents' sociolinguistic affiliation ($\chi^2 = 61.26, df = 8, p < 0.004$) with *Adja*, *Akan*, *Ewe* and *Sahouè* mainly having the trees on farms while *Aizo, Fon, Ga-adangbe*, *Wémè* and *Holli* had them in home gardens, though the species was more frequently found on farms (54%) than in home gardens (46%) ($\chi 2 = 3.9696, df = 1, p = 0.04$). Four possible propagation techniques were listed by the respondents, but direct seed sowing and seedling transplanting were the most popular ones (**Fig. 6.4A**). Most of the respondents indicated that the miracle plant is a slow-growing species (**Fig. 6.4B**) due to the time to first fruiting averaging 5.5 years. *Akan* and *Ga-Adangbe* reported a time to first fruiting that was significantly shorter than other groups (**Fig. 6.4C**).

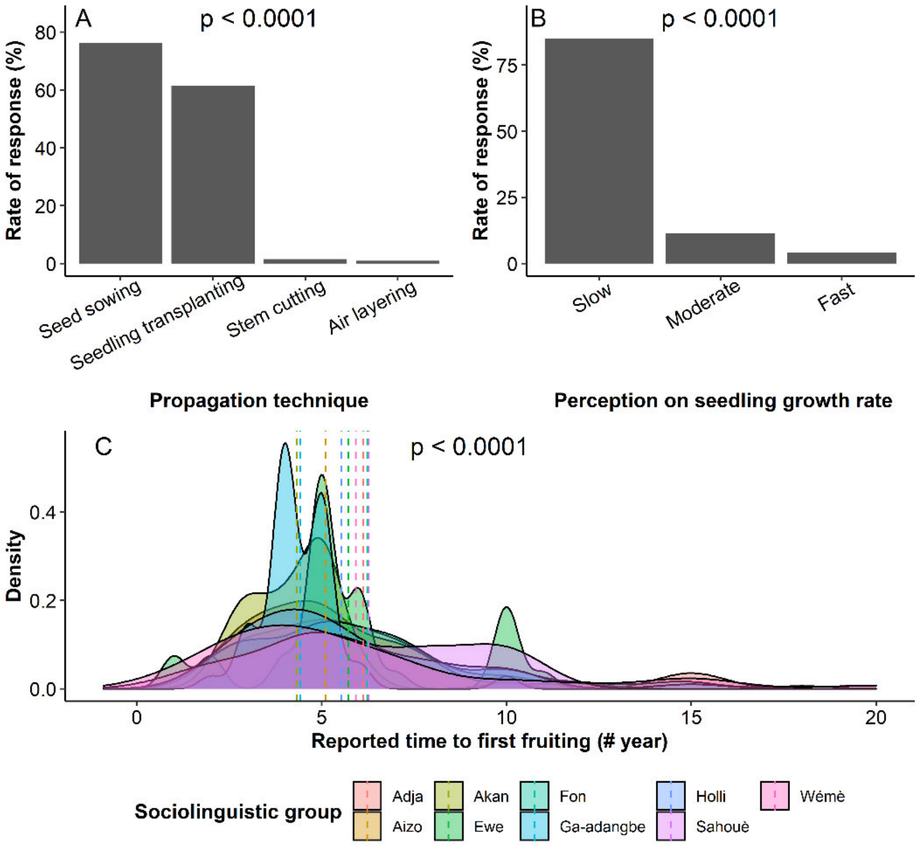


Figure 6.4. Respondent's knowledge of the miracle plant (*Synsepalum dulcificum*) regeneration, growth and development in Benin and Ghana. (**A**) Reported propagation techniques employed in the species propagation; (**B**) Perception on the species growth pace and (**C**) Reported time to first fruiting by respondents according to sociolinguistic groups (Each coloured density curve represents the distribution of times to first fruiting of the miracle plant tree within a specific sociolinguistic group, and coloured dotted lines represent the average time to first fruiting within each sociolinguistic group).

6.3.3 Use patterns and knowledge acquisition in *Synsepalum dulcificum*

Seventy-six uses grouped into six categories namely food, medicines, magico-spiritual, sales, social and firewood were recorded for the species (**Table 6.3**). The “medicinal” use category had 56 different uses followed by the “magico-spiritual” use category with 11 uses. A single-use type represented each of the food and fuelwood categories. Magico-spiritual uses were only reported in Benin, whereas the use for firewood was reported only by the *Ewe* in Ghana (**Fig. 6.5**).

The overall use-value of the miracle plant was UV = 2.45 ± 0.06. Its disaggregation per use category revealed that the use-value of the food category was significantly greater than that of other use categories (**Fig. 6.6A**). Likewise, the fruit had a greater UV compared with all other plant parts (**Figure 6.6B**). The species UV varied significantly across the sociolinguistic group, the main activity and the religion ($p < 0.0001$) and only marginally following the gender ($p = 0.08$) and the level of schooling ($p = 0.06$) (**Figs. 6.7A-H**). Sociolinguistic groups in Benin used the species more than sociolinguistic groups in Ghana. The *Wémé* in Benin exploited the species better than any other sociolinguistic groups (**Fig. 6.7B**). Likewise, the species use value was greater for traditional healers (**Fig. 6.7E**) and indigenous religious practitioners (**Fig. 6.7H**) than for respondents in other socio-professional categories or practising other religions. Conversely, no significant effect of age category, migratory status and species ownership status (**Figs. 6.7C,F,G**) was detected on the use-value ($p > 0.05$). The correlation between use-value and number of trees was overall weak and non-significant ($rho = -0.01$, $p = 0.89$), but differed significantly among attributes of all the studied socio-demographic factors (**Appendix 6.1**). For instance, the lowest correlation between use-value and number of trees owned for the sociolinguistic affiliation was obtained for the *Ewe* ($r = -0.04$, $p = 0.74$), whereas the highest one was observed within the *Akan* community ($r = 0.54$, $p < 0.0001$) ($Z = 4.86$, $p = 0.0001$). Similarly, the relationship between use-value and number of trees owned was stronger for respondents with a primary level of schooling ($r = 0.3$, $p = 0.005$) than for respondents with any other level of schooling ($r < 0.08$, $p > 0.37$) ($Z = 3.66$, $p = 0.01$) on one hand, and greater for allochthons ($r = 0.61$, $p = 0.002$) compared with autochthons ($r = -0.001$, $p = 0.98$) ($Z = 0.97$, $p < 0.0001$) on the other hand. Conversely, this association was extremely low for both men and women ($r < 001$, $p > 0.05$) ($Z = 0.97$, $p = 0.9$).

Table 6.3. Use category, body systems and informant agreement on the documented uses for the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana.

Category	Sub-Category/body systems	Use	Number of use report	Plant parts involved	IAR*
Food	Sweetener	Sweetener	497	Fruits	1.00
Sales	Commercialization	Commercialization	112	Fruits, leaves, roots, seeds and twigs	0.96
Social	Handcraft	Chewing stick	117	Twigs	1.00
	Tool	Hoe handle	1	Twigs	NA
	Harmony	Conflict resolution, against alcoholism, miraculin activity disruption, accelerating the walk of the new-born, Improving voice	10	Fruits, leaves, roots, seeds, and twigs	0.55
Magico-spiritual	Lucky charm	Lucky charm	95	Flower, fruit, leaves, roots and twigs	0.95
	Woe induction	Woe induction	21	Seeds, leaves, fruit and twigs	0.85
	Bewitchment	Bewitchment	15	Leaves, roots, seeds and twigs	0.78
	Protection	Protection	13	Leaves, roots and twigs	0.83
	Love attraction	Attract love	12	Fruits, leaves and twigs	0.81
	Wealth attraction	Wealth attraction	8	Fruit, leaves, roots and twigs	0.57
	Ritualistic	Ritualistic	12	Fruit, leaves and roots	0.81
Medicinal	Blood	Anaemia	3	Leaves	1.00
	Circulatory	Haemorrhoid, heart pain and hypertension	22	Bark, leaves, roots, and seeds	0.86
	Digestive	Absence of taste perception, angina, caries, jaundice, diarrhoea, hepatitis, purgative, sore throat, stomach-ache,	90	Bark, fruit, leaves, roots, seeds and twigs	0.94

Category	Sub-Category/body systems	Use	Number of use report	Plant parts involved	IAR*
		tooth decay, tooth pain, ulcer, and vomiting			
	Endocrine, metabolic, and nutritional	Loss of appetite for food, diabetes, and overweight	18	Fruit, leaves, roots and seeds	0.82
	Eye	Eye pain	2	Fruit and leaves	0.00
	Female genital system	Cyst, menstrual pain, and menstrual irregularity	9	Leaves and roots	0.875
	General health and unspecified	Excessive fatigue, cancer, fainting, fever, malaria, measle	68	Bark, fruit, leaves, roots and seeds	0.94
	Male genital system	Male impotency and manhood stimulation	35	Bark, fruit, leaves, roots, seeds and twigs	0.85
	Musculoskeletal	Back pain, Chest pain and hip-ache	6	Leaves and roots	0.80
	Neurological	Headache and lack of reactivity to stimuli	6	Fruit and leaves	0.80
	Pregnancy and childbearing	Delivery, lactation stimulation, weaning and miscarriage	20	Bark, fruit, leaves and roots	0.84
	Psychological	Lack of sexual appetite of women and Memory-aid	20	Bark, flower, fruit, leaves and roots	0.78
	Respiratory	Cough, respiratory troubles	10	Bark, fruit, leaves and roots	0.66
	Skin	Wound healing and boil	6	Bark, leaves and roots	0.60
	Urinary	Presence of blood in urine, Enuresis and Kidney ailment	26	Leaves and roots	0.96
Firewood	Firewood	Firewood	1	Branches/twigs	NA*

*IAR: Informant agreement ratio.

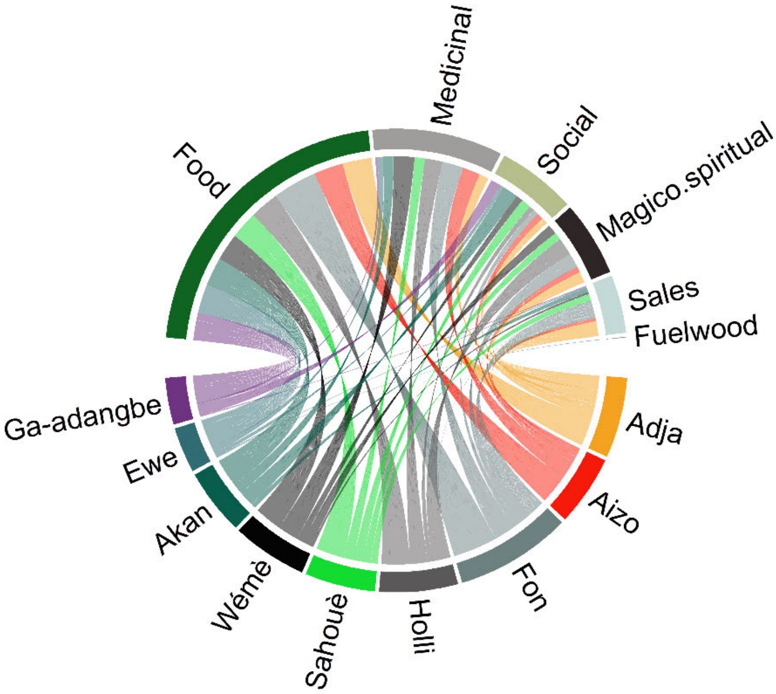


Figure 6.5. Association between sociolinguistic groups and use category in the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana. Each coloured line joining a specific sociolinguistic group and a use category represent one use report.

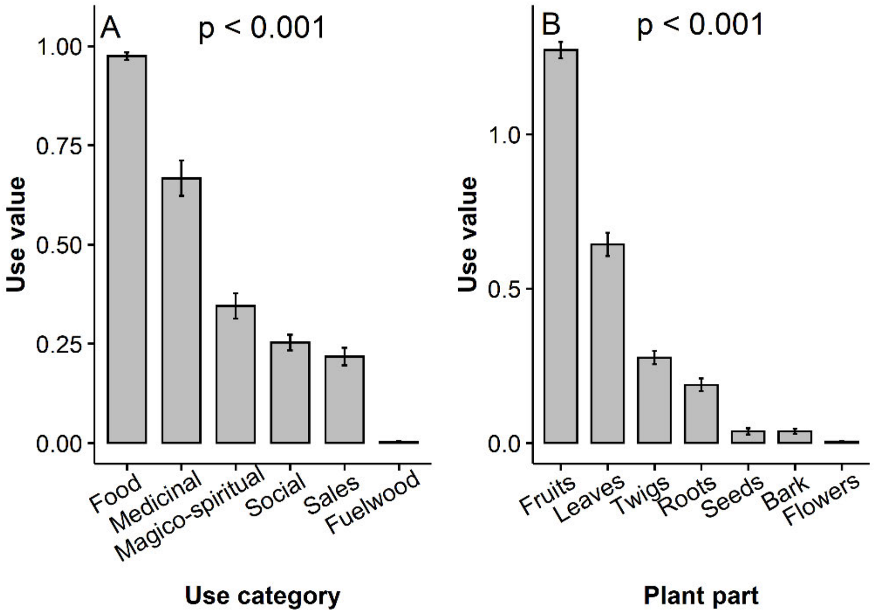


Figure 6.6. Use-value of *Synsepalum dulcificum* per use category and plant parts in Benin and Ghana. Bars represent average use values and error bars represent standard errors or means (SEM).

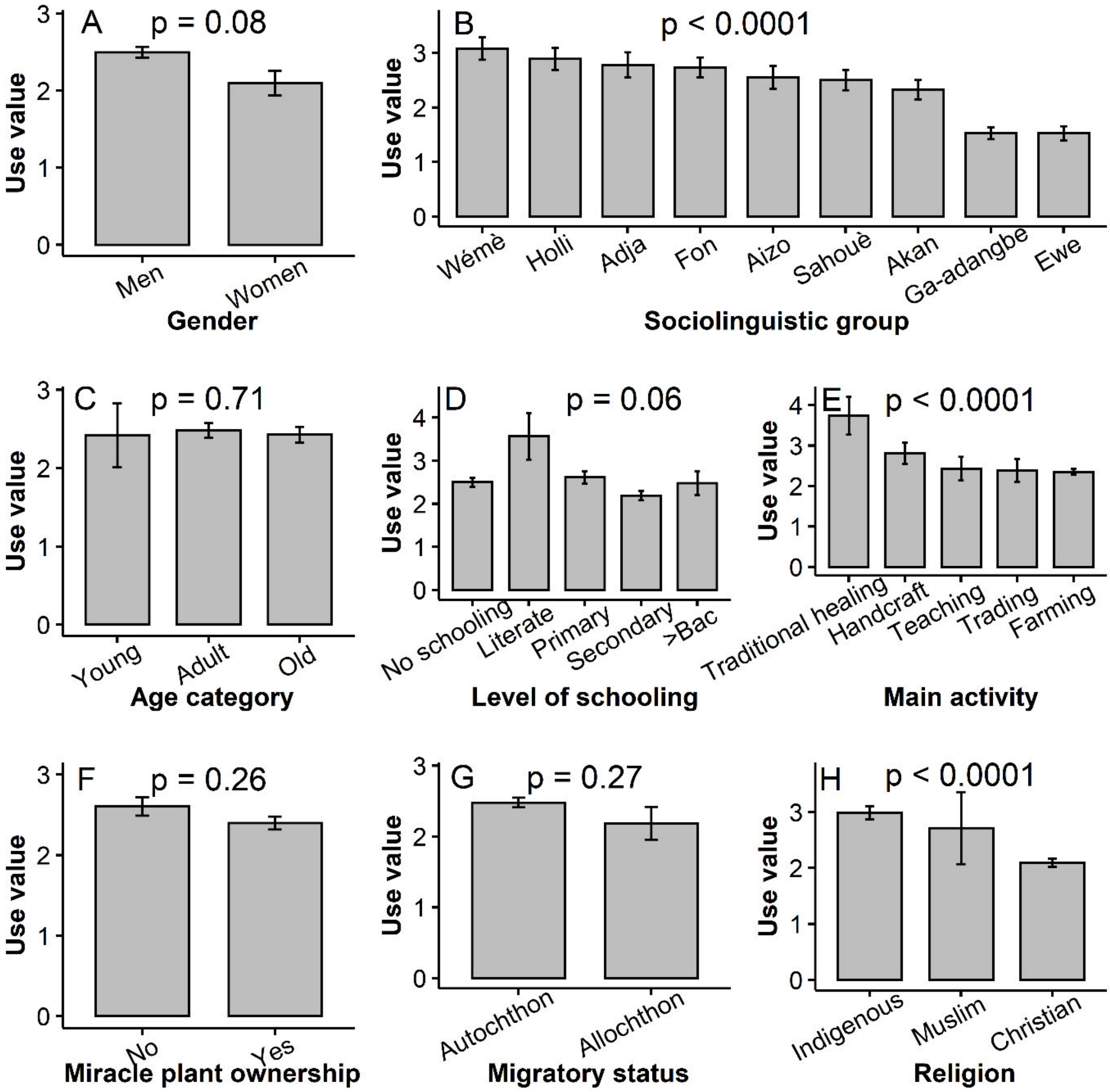


Figure 6.7. Variation of the miracle plant (*Synsepalum dulcificum*) use-value following various socio-demographic variables in Benin and Ghana. Bars represent average use values and error bars represent standard errors of means (SEM).

All the miracle plant parts including flowers, fruits, seeds, leaves, twigs, bark and roots are employed in medicinal use with 16 various body systems treated (**Table 6.3**). The top five, most frequently treated ailments, in order of importance included: (i) the digestive system, mainly stomach ache and tooth pain; (ii) the general health system, dominated by malaria, fever and asthenia; (iii) the male genital system to stimulate erectile functions and treat male impotency; (iv) the urinary system, mainly enuresis and haematuria; and (v) the circulatory systems, commonly high blood pressure and haemorrhoids. Apart from the bark, all the other plant parts also had magico-spiritual applications, predominantly to bring luck.

The highest consensus among respondents was around the use of the fruit as a sweetener (IAR = 1) and the leaves in blood system ailment treatments (e.g., anaemia) (IAR = 1). These were followed by the use of twigs as chewing sticks (IAR = 0.97), treatment of urinary system ailments (e.g., enuresis) (IAR = 0.96) and in the treatment of general health and digestive systems disorders (e.g., malaria and stomach ache) (IAR = 0.94) (**Table 6.3**).

For both medicinal and magico-spiritual uses, the plant parts are employed either fresh/raw or dry and with various methods of preparations (**Appendix 6.2**). Decoction, grinding and direct use were the main preparation methods employed when using the plant parts. When used for medicinal purposes the preparations were administered orally, whereas the powder obtained from the grinding of plant parts was used as lapping powder or applied to skin scars followed by incantations in magico-spiritual utilisations. For the preparation, the miracle plant parts were used either alone or associated with other elements or other plant species. **Table 6.4** details some key medicinal and magico-spiritual uses of the species, the preparation modes and the dosage.

Table 6.4. Detailed description for some selected medicinal and magico-spiritual applications of the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana.

Medicinal/magico-spiritual application	Description
Lucky charm for traders and candidates to exams (Benin)	Chewing of the fresh leaves of *S. dulcificum* together with a fruit of *Cola nitida* (Vent.) Schott & Endl. as needed in the morning
	Grinding the miracle plant leaves and use the powder as an ingredient to prepare a soap that will be used every day
	The preparation obtained by grinding together the dry miracle plant roots, leaves and fruits + dry leaves of *Arbrus precatorius* L.+ *Garcinia kola* (Heckel) is shaped in small balls that are kept under the tongue
	Prepare an infusion of the mixture of the leaves of the miracle plant and seeds of *Afromomum melegueta* [Roscoe] K. Schum to which some sugar or honey is added. The infusion is drunk when needed
Haemorrhoid (Benin + Ghana)	Infusion of the root of the miracle plant + other non-disclosed plants + kaolin to be taken when needed
	The decoction of the miracle plant's leaves together with the bark associated with other non-disclosed plants is taken 3 times per day for a week.

Medicinal/magico-spiritual application	Description
Cough (Benin)	The decoction of the miracle plant bark mixed with the leaves of *Casuarina equisitifolia* L. is drunk daily until the disease stopped
Enuresis (Benin)	Add residues of a hen's nest to the miracle plant leaves and make a decoction of it that will be drunk as long as needed
Headache (Benin + Ghana)	Dry and grind altogether the leaves of the miracle plant with the seeds of *Afromomum melegueta* [Roscoe] K. Schum. The powder obtained is then topically applied to the skin scars made on the forehead of the sick person.
Ulcer (Benin)	Decoction of the leaves of the miracle plant + leaves of *Heterotis rotundifolia* (Sm.) Jacq-Fél. to be taken when the pain is felt
Stomach-ache (Benin + Ghana)	Pour ground dry seeds of miracle fruit in warm water and drink it. Drink the alcohol-based decoction of the mixture of the miracle fruit, leaves and roots.

The analysis of the knowledge acquisition pattern following the KMTO framework revealed that in general respondents acquired knowledge of the miracle plant from an internal source with a first-order transitionally vertical path marked by the prominence of the father as the main source of knowledge (**Fig. 6.8**). While the knowledge acquisition source was not linked to the use category (**Fig. 6.9A**, $\chi^2 = 8.72, df = 4, p = 0.06$), the knowledge mutation form ($\chi^2 = 49.13, df = 12, p < 0.0001$), knowledge acquisition type ($\chi^2 = 9.627, df = 8, p < 0.0001$) and prevailing acquisition order ($\chi^2 = 10.12, df = 8, p < 0.0001$) (**Fig. 6.9B–D**) were conversely all dependent on the use category. Besides the men transition that represented the primary knowledge mutation form, women's knowledge of food, sales and magico-spiritual use was secondarily acquired from men, while knowledge of medicinal and social uses held by men secondarily came from women. More interestingly, knowledge of medicinal and magico-spiritual uses was inherited more from grandparents than from transmitters of the same generation, whereas knowledge related to the food and social uses was obtained more from transmitters of the same generation than from grandparents.

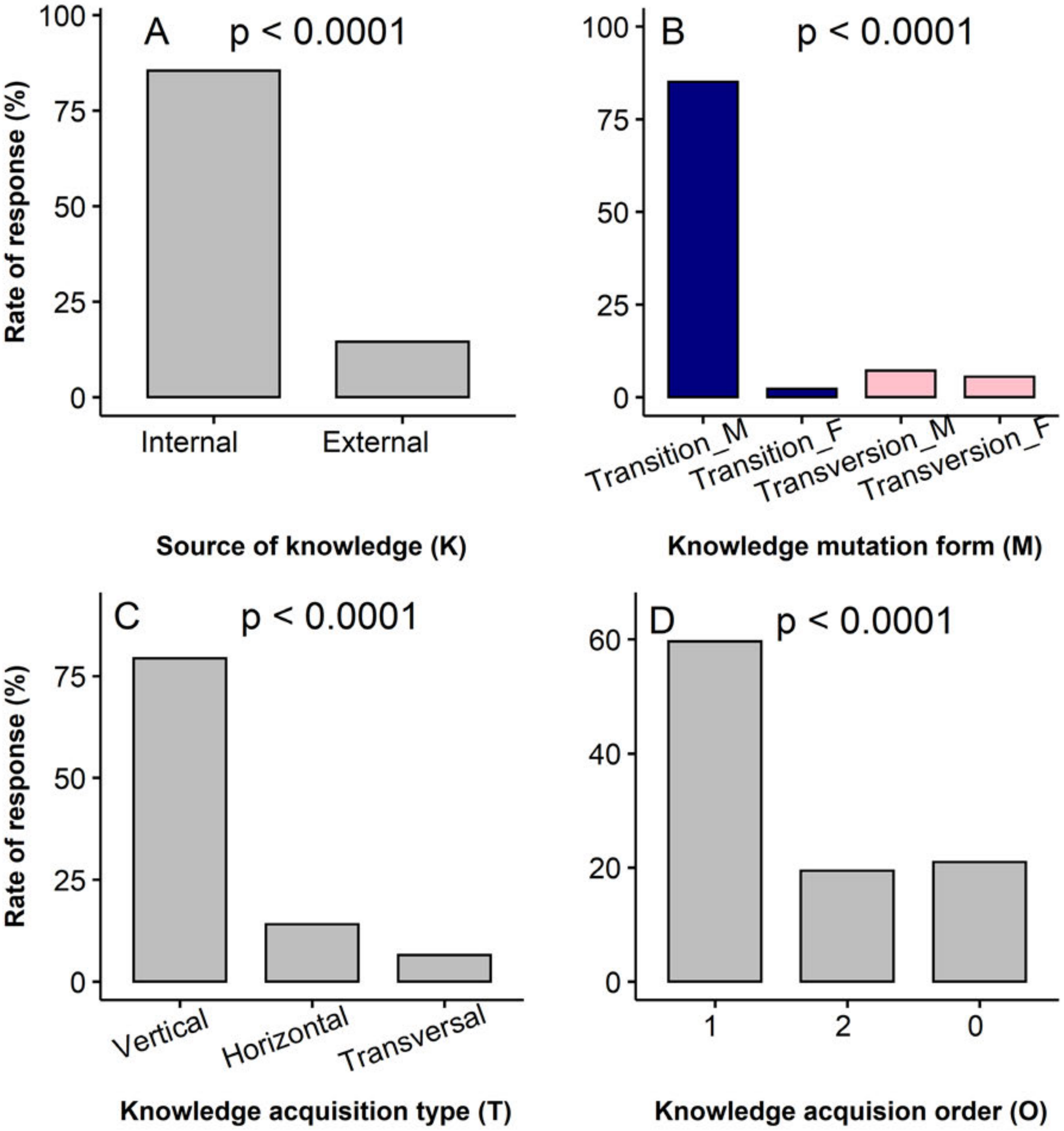


Figure 6.8. Relative importance of the knowledge acquisition components following the Kernel-Mutation-Type-Order (KMTO) analysis framework in the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana.

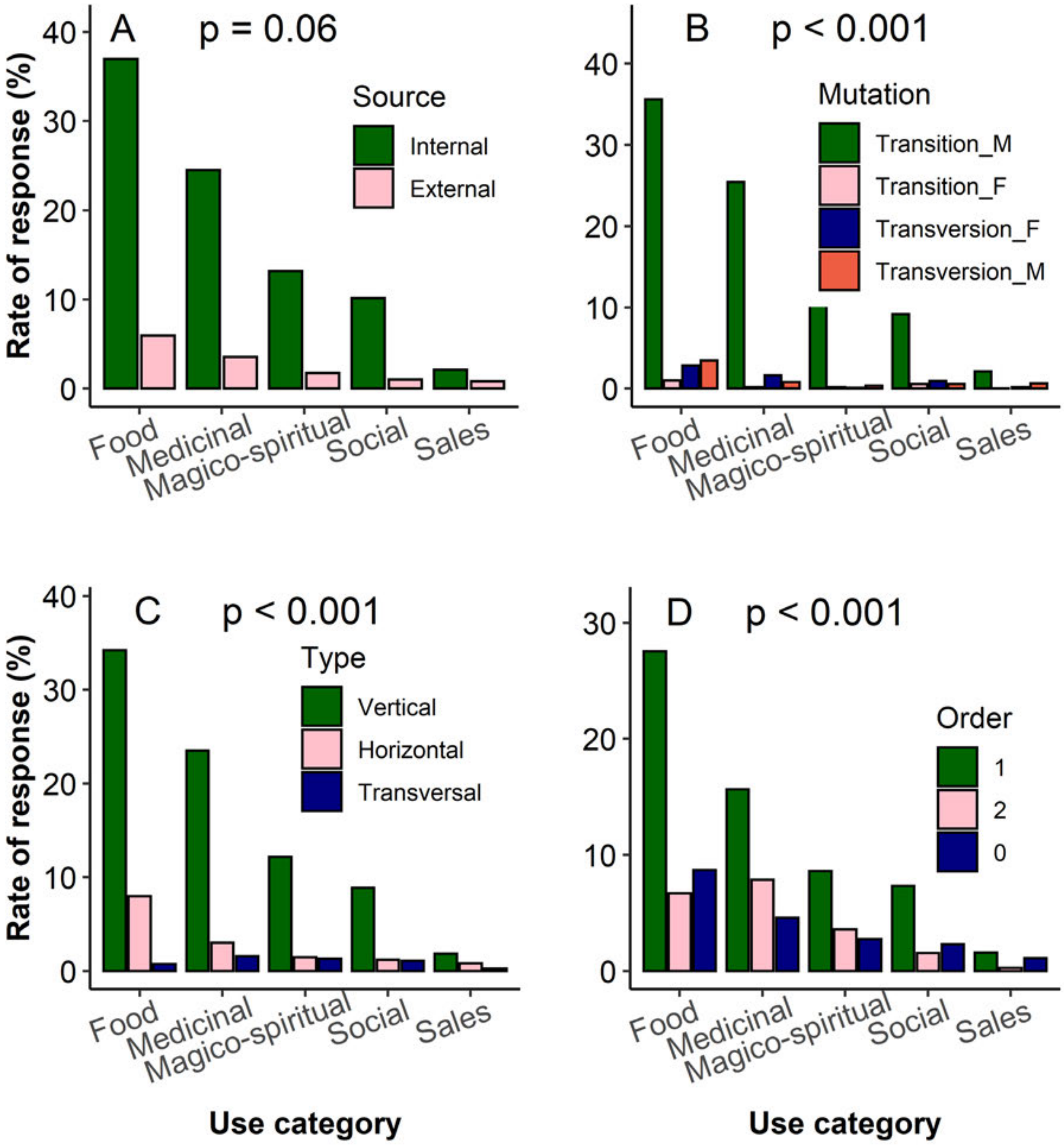


Figure 6.9. Association between use category and knowledge acquisition path based on the KMTO framework in the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana. (A): Source of knowledge, (B): knowledge mutation form, (C): Knowledge acquisition type and (D): Knowledge acquisition order.

6.3.4 Perception of threats, taboos and superstitions on *Synsepalum dulcificum*

A total of 506 out of the 510 respondents shared their perception of the availability of the species in their environment. Three availability classes: “Decline”, “Stability” and “Increase” were recorded for the species. A significantly higher proportion of respondents (85%) indicated the depletion of the species compared with those indicating stability (6% of respondents) and increase (9% of respondents) ($\chi^2 = 189.3$, $df = 2$, $p < 0.001$). The perception of the prominence of each availability class depended on the sociolinguistic affiliation ($\chi^2 = 68.052$, $df = 16$, $p < 0.001$), the migratory status ($\chi^2 = 15.374$, $df = 2$, $p = 0.0004$), the religion ($\chi^2 = 26.644$, $df = 4$, $p < 0.001$) and the level of schooling ($\chi^2 = 31.489$, $df = 8$, $p = 0.0001$). For instance, the class “Increase” was predominant over the class “depletion” for Ghana sociolinguistic groups, while the reverse held true for Benin sociolinguistic groups. The major causes of this depletion as reported by the respondents included the erosion of the crop-related knowledge, crop negligence, agricultural expansion, the seasonal bushfires, the lack of cultivation initiatives in the crop and overharvesting (especially of the roots and leaves by medicinal plant vendors) (**Fig. 6.10**).

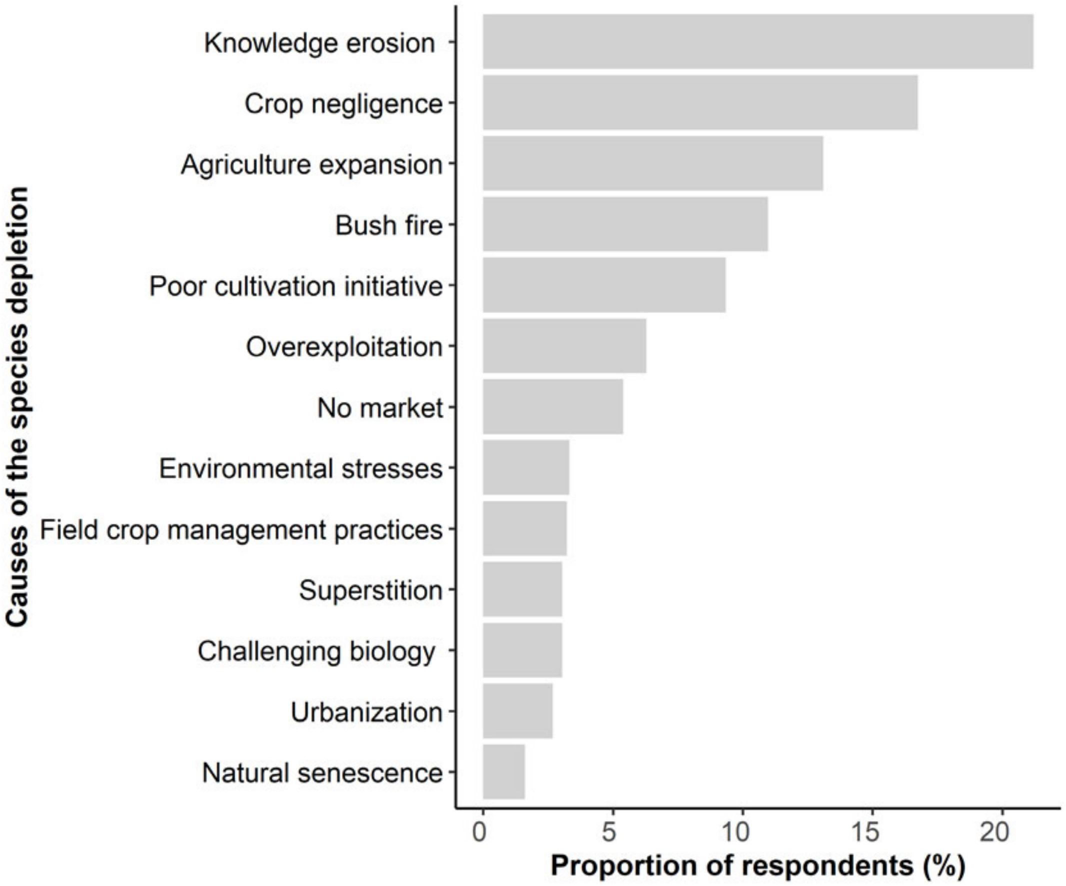


Figure 6.10. Threatening factors of the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana.

Ten taboos and six superstitions were reported in the miracle plant (**Appendix 6.3**). While both taboos and superstitions existed in all sociolinguistic groups sampled in Benin, they were only reported by Ga-adangbe out of the three sociolinguistic groups investigated in Ghana. Additionally, both taboo ($\chi^2 = 23.05$, $df = 1$, $p < 0.0001$) and superstition ($\chi^2 = 168.3$, $df = 1$, $p < 0.0001$) were more frequently reported by sociolinguistic groups in Benin than those in Ghana. The commonest taboo on the miracle plant was related to the prohibition to set fire close to the tree. As for the superstition, the most frequently reported one was that "*someone who plants the species will die or one of his parents will die before the plant starts bearing fruits*".

6.3.5 Determinants of willingness and readiness to cultivate *Synsepalum dulcificum*

The classification tree model revealed that six out of the 14 candidate variables interacted to significantly determine farmers' decision to engage in the cultivation of the miracle plant in Benin and Ghana (**Fig. 6.11**). These variables included the sociolinguistic affiliation, the existence of taboos on the crop, the respondent's perception of market availability, the level of schooling, the perception of the time to fruiting and the existence of superstition on the crop. While the *Akan* and *Ga-adangbe* were, in general, not willing to engage in the cultivation of the miracle plant, the decision of their counterparts *Adja*, *Aizo*, *Ewe*, *Fon*, *Holli*, *Sahouè* and *Wémé* was determined by the existence of taboos on the crop, with those reporting taboos more willing to invest in the species cultivation. In absence of taboos, the willingness of *Adja*, *Aizo*, *Ewe*, *Fon*, *Holli*, *Sahouè* and *Wémé* to cultivate the species was conditioned by the interaction between their perception on the existence of a market and their level of schooling. Among those indicating the existence of a market, respondents whose maximum level of schooling was the secondary school were, in general, more willing to cultivate the species than those who had higher or lower levels of schooling. In this latter group, the *Aizo*, *Ewe*, *Fon* who reached the maximum level of primary school were willing to engage in the species cultivation, whereas the least and most educated *Aizo*, *Ewe*, *Fon* respondents were not willing to do so. For the *Adja*, *Holli*, *Sahouè* and *Wémè* who perceived the availability of a market and who had a level of schooling other than the secondary level, the decision to cultivate the miracle plant was determined by their perception of time to fruiting in the species. Those indicating the species current fruiting time to be acceptable or fast were unsurprisingly willing to engage in the species cultivation, while their counterparts who perceived the time to fruiting as very long accepted to cultivate the species only when there is no superstition on the crop. The *Aizo* and *Wémè* who did not report any known taboos and were not aware of any available market for the species were not ready to cultivate the species. Among the *Adja*, *Ewe*, *Fon, Holli*, *Sahouè* who did not report any taboos, nor perceived any available market for the miracle plant, only the *Ewe* and *Fon* who reached the primary school level were willing to cultivate the miracle plant.

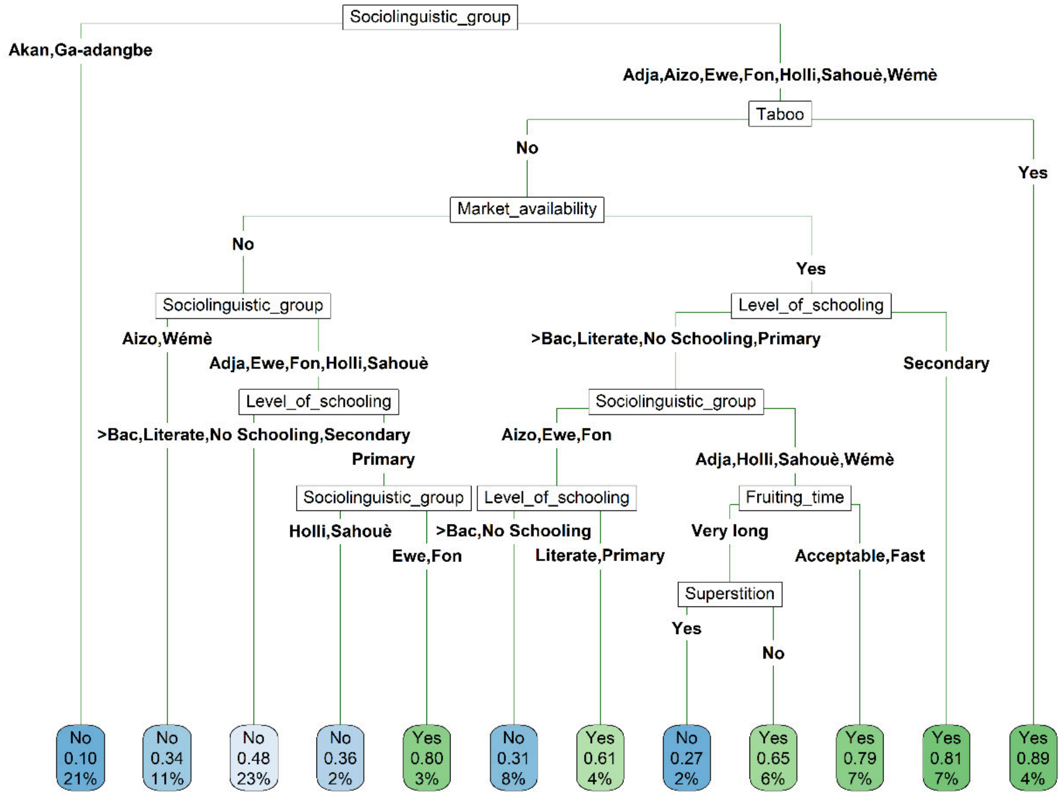


Figure 6.11. Classification tree depicting the determinants of willingness to cultivate the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana. Each node successively presents the predicted class [willing to cultivate (Yes) or not willing to cultivate (No) the miracle plant tree], the predicted probability of willingness to cultivate and the percentage of observation at the node.

A total of 45% of respondents in this study indicated their willingness to cultivate *S. dulcificum* and the extent of their readiness to do so was evaluated through the maximum acreage they were ready to allocate to the crop and the maximum price they were ready to pay to acquire an ordinary seedling of a miracle plant.

The regression tree model revealed that the only two factors driving plot acreage allocation to the miracle plant by a respondent were by order of importance the respondents' level of schooling and their perception of the growth rate of the crop (**Fig. 6.12**). While in general respondents were only ready to allocate on average 0.43 ha of their land for the cultivation of the miracle plant, those reaching the highest level of schooling (University level) were ready to allocate four times more acreage i.e., 1.8 ha to the crop, whereas the respondents who did not school were ready to allocate on average a maximum of 0.2 ha. For the respondents who had intermediate levels of schooling (primary and secondary levels), the acreage to allocate for the cultivation of the miracle plant was conditioned by their perception of the species growth rate. Intermediate level-educated respondents who perceived a poor growth rate (slow and moderate growth rate) in the species were ready to allocate on average 0.4 ha to the crop, versus two times more acreage i.e., 0.83 ha for their counterparts who perceived the species growth rate as fast.

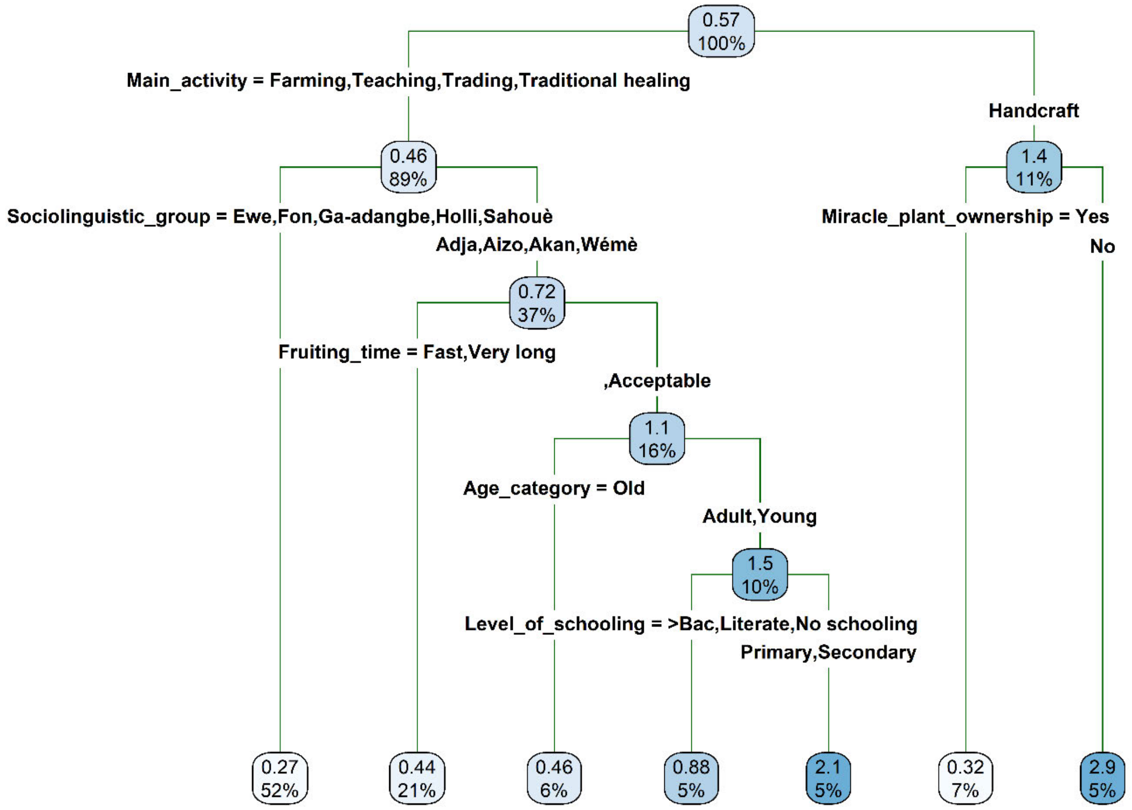


Figure 6.12. Regression tree depicting the drivers of acreage (ha) allocation as an indicator of respondents' readiness to cul-tivate the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana. Each node successively presents the pre-dicted value of the acreage respondents are willing to allocate for the species cultivation and the percentage of observation at the node.

The miracle plant seedling purchase price as proposed by respondents varied on average from USD 0.27 to USD 2.9 and was mainly driven by the respondents' main activity, the current miracle plant ownership status of the respondent, the sociolinguistic affiliation, the perception on the time to fruiting, the respondents' age and level of schooling (**Fig. 6.13**). While respondents were in general willing to pay on average the amount of USD 0.57 to acquire a seedling of the miracle plant, handcraft makers were willing to pay nearly three times higher price for a seedling and five times higher price when they did not previously own any miracle plants. Conversely, having previously owned a miracle tree decreased to an average of USD 0.4 the amount at which handcraft makers were ready to purchase a seedling of the crop. In contrast to the handcraft makers, the farmers, traders, traditional healers and teachers were only willing to pay an average of USD 0.46 for a seedling of the miracle plant. However, the young and adult *Adja, Aizo*, *Akan* and *Wémé,* who perceived the time to fruiting of the miracle plant as acceptable and who had an intermediate level of schooling (primary and secondary) were willing to pay up to USD 2.1 to acquire one seedling. In contrast, extreme levels of schooling decreased to USD 0.88 the amount this same group of respondents were ready to pay to acquire a miracle plant seedling. As far as the older respondents were concerned, they were only willing to pay an average of 0.41 USD for one seedling.

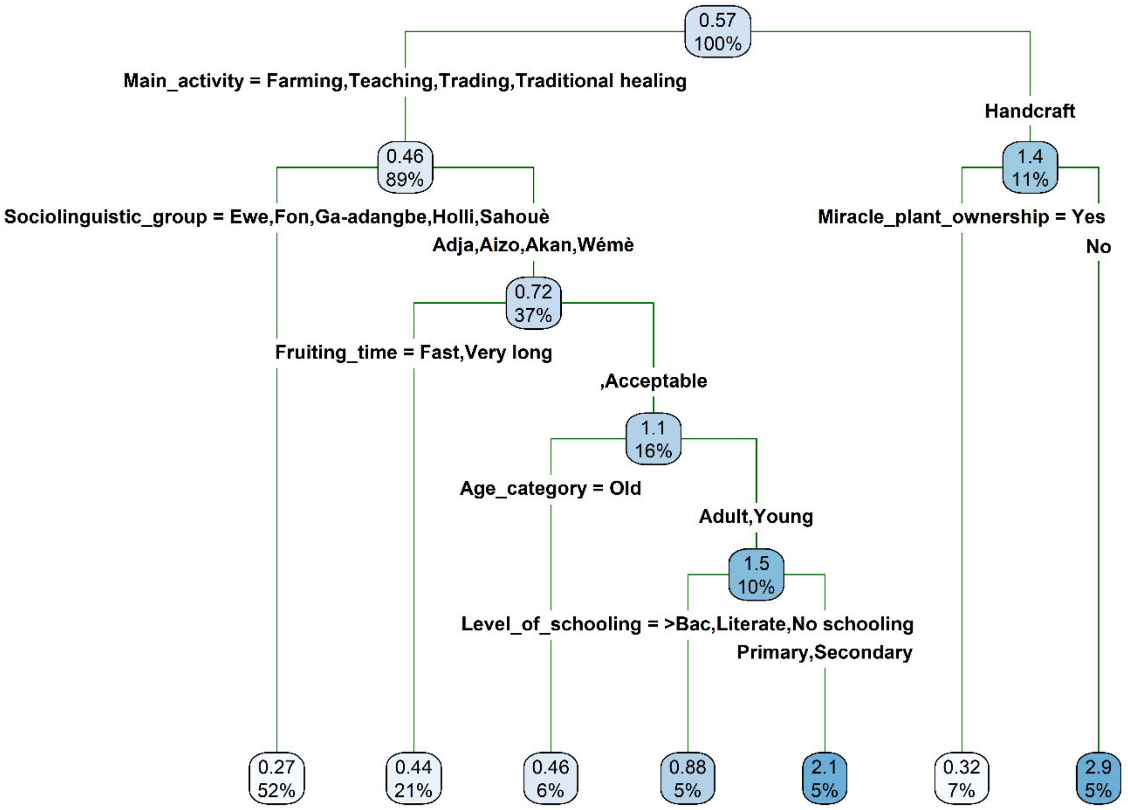


Figure 6.13. Regression tree depicting the drivers of the maximum purchase price (USD) of a miracle plant (*Synsepalum dulcificum*) seedling as a measure of respondents' readiness to cultivate the species in Benin and Ghana. Each node successively presents the predicted value of the acreage respondents are willing to allocate for the species cultivation and the percentage of observation at the node.

6.4 Discussion

6.4.1 *Synsepalum dulcificum* nomenclature, ownership pattern and bioecology

Folk names used by the local population to designate plant species often reflect a diversity of attributes that may relate to the plant's habitats, uses, morphology and biological characteristics, among others (Agbo et al., 2020; Gouwakinnou et al., 2011; Mongalo and Makhafola, 2018; Teka et al., 2020). The local names recorded for the miracle plant (**Table 6.2**) illustrated not only the tree anatomy and longevity but also the functional attributes. Previous ethnobotanical studies suggested that the commonly used names for the miracle plant in Benin reflected the sweetening activity of the fruit (Fandohan et al., 2017). Additionally, the meaning "*No other taste than sweet in contact of the tongue*" (**Table 6.2**) given by the *Fon* perfectly illustrated the mechanism of action of the "miraculin", which is the glycoprotein of the fruit responsible for the sweetening activity. One of the scientifically postulated mechanisms of action of the miraculin is that it binds to the membrane surface of the tongue's taste cells to trigger the sweetness perception (Kurihara and Beidler, 1969). Thus, sweetness is only perceived when the fruit is in contact with the tongue receptors. The name "Agbanyun" used by the *Holli*, which literally means "honey calabash" was to claim that the miracle fruit holds the paroxysm of sweetness and this strongly aligned with the scientific evidence suggesting the miracle fruit to be 400,000 times sweeter than sucrose (Kurihara and Beidler, 1968). A proper study of plant folk names then holds the potential to help discover detailed plant functions.

Most of the miracle plant trees owned by the respondents were inherited. This strengthened observations by Fandohan et al. (2017) in Benin on the species, and confirmed the legacy as an important ownership mode in semi-domesticated plant species (Dadjo et al., 2020). Nevertheless, we observed in this study an influence of the respondents' age on the predominant ownership mode whereby young and adults representing 63% of respondents (**Table 6.1**) inherited more of the species, while the elderly respondents (37% of respondents, **Table 6.1**) mainly planted the tree themselves. This prominence of inheritance currently translated an overall poor active conservation of the species and thus raised concerns regarding the future of the species in West Africa since the "older generation" is passing.

Although women held on average three miracle plant trees and men on average five (**Fig. 6.3A**), this difference was not significant; however, it is concurrent with previous findings which have reported an absence of variation in the total number of miracle plant trees held by men and women in Benin (Fandohan et al., 2017). This similarity in the number of trees owned might partly be attributed to the fact that both women and men are increasingly being exposed to similar social beliefs and norms (e.g., superstition). Because socio-cultural norms strongly differed among sociolinguistic groups and that the perception of such norms and beliefs also varied among religions (Ekué et al., 2010), the observed significant difference among sociolinguistic groups and among religions for the number of miracle plant owned did not come as a surprise (Fandohan et al., 2017). Indeed, the sociolinguistic groups from Ghana practising Christianity at more than 98% (**Table 6.1**) held by far more miracle plant trees than Benin sociolinguistic groups whose respondents mainly practised indigenous religions. Because indigenous religions are demarcated from Christianity through the prominence of socio-cultural beliefs and norms, we suspect that these norms (e.g., superstition) likely preclude either form of miracle plant ownership. This hypothesis is furthermore strengthened by the fact that superstition, which mostly existed in indigenous religions, in Benin in particular, was confirmed as one of the threats factors for the species. An example of such superstition is that "*Someone who plants a miracle plant will die or one of his parents will die before the tree starts bearing fruits*". This particular superstition was also reported in Eastern Africa on tamarind (*Tamarindus indica*) (Ebifa-Othieno et al., 2017) and sheds light on why the young and adults have not planted the trees themselves. Our results on the species ownership by sociolinguistic groups revealed that the *Holli* held between one and eight miracle plant trees, which is new since no previous studies reported miracle plant trees ownership by this community (Fandohan et al., 2017; Oumorou et al., 2010).

The miracle plant trees were mainly found in home gardens and on farms in this study. This aligned with findings of previous investigations reporting gallery forests, cultivated farms, home gardens and fallows as possible miracle plant habitats (Adomou, 2005; Fandohan et al., 2017; Oumorou et al., 2010). Our findings also expand our knowledge of the species habitat in West Africa by highlighting that while home-gardens persist as the major habitat of the species in most of Benin's sociolinguistic groups, cultivated farms predominated in Ghana.

Although the respondents for the miracle plant currently employ four different propagation techniques, only those based on the seed are the most utilized (**Fig. 6.4A**). This could be ascribed to the easiness of the transplanting action and the high germination rate of the seeds (Tchokponhoué et al., 2018), contrarily to the cuttings that exhibited a recalcitrant adventitious rooting (Chen et al., 2012). Interestingly, more than 80% of the respondents perceived the miracle plant as a slow-growing species (**Fig. 6.4B**), a knowledge that aligned with the experimental conclusions of Tchokponhoué et al. (2018) in the species. Similarly, the overall five years (**Fig. 6.4C**) reported for the age to first fruiting fitted in the time frame of 4–7 years reported for the time to fruiting in natural conditions (Joyner, 2006).

6.4.2 Use patterns and knowledge acquisition in *Synsepalum dulcificum*

Our study expanded the local applications of the miracle plant from 64 uses (Fandohan et al., 2017) to 76 uses. Therefore, through our study, miracle plant showed more documented uses compared with marula (*Sclerocarya birrea*, 20 uses) (Gouwakinnou et al., 2011), baobab (*Adansonia digitata* L., 38 uses) (De Caluwé et al., 2009), sweet detar (*Detarium microcarpum* Guill. & Perr., 42 uses) (Agbo et al., 2020) and the red kapok tree (*Bombax costatum* Pellegr. & Vuillet, 46 uses) (Assogba et al., 2017). The analysis of the importance of each use category suggested that the miracle plant is more medicinal than a food multipurpose species. Though the miracle plant medicinal value was previously indicated (Fandohan et al., 2017), this study has updated the medicinal uses and categorized them into body systems (**Table 6.3**), making easier and more reliable comparisons with other species. The 16 body systems recorded for the miracle plant alone quantitatively compare well to the number of body systems covered by a set of 309 medicinal plants studied by Ribeiro et al. (2017) in Brazil. Likewise, the miracle plant covered more body systems than did a set of 105 medicinal taxa studied in Portugal (Vinagre et al., 2019). These illustrated how deep the medicinal breadth of the miracle plant is and calls for more pharmacological and pharmacognostic investigations to set the scientific bases of the numerous folk utilisations of the species. Good illustrations in this vein exist Obafemi et al. (2017) and Obafemi et al. (2019) whose findings enlightened, for instance, the use of the miracle plant leaves by the local population in the treatment of diabetes. Conclusions of many studies indicated that the commonly treated disorders by typical medicinal plants were related to the digestive, circulatory and respiratory body systems (Koura et al., 2011; Ribeiro et al., 2017), a trend that is confirmed in this study with the digestive system that was the most treated body system by the miracle plant; then followed by the urinary and circulatory systems. The three most important use categories recorded are concurrent with the findings of Fandohan et al. (2017) on the miracle plant and of (Assogba et al., 2017) on the red kapok tree.

Contrarily to the “age, gender and dynamics of knowledge” hypothesis, which suggested that age category or gender affects an individual’s plant knowledge (Hanazaki et al., 2013; Voeks, 2007) and which was confirmed in a number of species (e.g., *D. microcarpum* (Agbo et al., 2020), *T. indica* (Ebifa-Othieno et al., 2017), *Borassus aethiopum* Mart. (Salako et al., 2018)), we observed that the miracle plant exhibited similar use-values for both men and women on one hand, and for young, adult and older respondents on the other hand (**Figs. 6.7A,C**). The absence of differences of use-value among the age categories could indicate that an effective knowledge transmission system from the elderly people to the young is in place in the case of the miracle plant, while the low number of women involved in this study compared to men could partly explain the similar use-value observed between men and women. Alternatively, the comparative advantage that women could have over men due to the medicinal plant status of the miracle plant (de Albuquerque et al., 2011) could have been buffered by the concomitant prominence of magico-spiritual use in the species, a use category mainly reported by men. Conversely, we observed a significant variation of the use-value among sociolinguistic groups (**Fig. 6.7B**), which concurred not only with previous findings in the species (Fandohan et al., 2017) but also with the trend in many other species such as the African locust bean *Parkia biglobosa* (Jacq.) G.Don (Koura et al., 2011), *A. digitata* (De Caluwé et al., 2009), *D. microcarpum* (Agbo et al., 2020) and *Gardenia erubescens* tapf & Hutch. (Ouédraogo et al., 2019). The low use-values recorded for the species in the Ghana sociolinguistic groups versus the high use values obtained in Benin sociolinguistic groups could have arisen from the discrepancy of use categories among countries. Indeed, the “Magico-spiritual use category” that counted 11 different uses was for instance only recorded within sociolinguistic groups in Benin (**Fig. 6.5**).

Because sociolinguistic groups in Ghana owned more trees than those in Benin (**Fig. 6.3B**), one could have in the light of the ‘availability hypothesis’ predicted that sociolinguistic groups in Ghana will use the species more than their counterpart in Benin. However, there was a negative and very weak correlation between the use-value and the number of trees owned, which may indicate that the hypothesis does not work for the miracle plant. However, the same availability hypothesis’ also helps predict that because the respondents having the miracle plant in their home garden are closer to the resources than their counterparts holding the species in their farms, they will consequently better use it. In this study, the miracle plant was more used when it was found in home gardens (UV = 2.59) than when cultivated on farms (UV = 2.22). Furthermore, a negative and significant correlation was observed between the use-value and distance of respondent’s dwelling to the trees, thus validating the hypothesis in this latter case. This dichotomy of conclusion strengthened the recommendations of Gaoue et al. (2017) to conceptualize the meaning of “availability” when it comes to testing the “availability hypothesis”. The greater use-value of the miracle plant for the traditional healers compared with the other socio-professional categories came without surprise since the species strongly served for both medicinal and magico-spiritual uses, two fields in which the traditional healers are reputed having broader knowledge than any other socio-professional categories (Nanyingi et al., 2008). The same also applies to the indigenous religion practitioners, who, being more involved than Muslims and Christians in magico-spiritual events and medicinal uses valued the miracle plant more. As far as the level of schooling was concerned, being highly educated provides less opportunity to be in contact with nature (N’Danikou et al., 2015) and hence narrows traditional use-related knowledge. This may explain the difference in use-value obtained between respondents who schooled (UV = 2.42) and those who did not (UV = 3.03).

We analysed the knowledge acquisition pattern in this study by employing a framework that helped depict the complex interaction between the transmitters and the receivers from a pedigree and a gender point of view. As previously reported in many species studied in Africa and worldwide (Giday et al., 2009; Mattalia et al., 2020), knowledge acquisition in the miracle plant also occurred mainly within the family network (**Fig. 6.8**) and orally, which overall means a poor knowledge flow among families within communities and an inefficient acquisition /transmission system since unwritten knowledge is prone to transmission bias due to interpretability. This paucity of knowledge exchange seemed to be particularly marked when it comes to sophisticated knowledge such as medicinal or magico-spiritual uses and this partly helped understand why the huge medicinal use reports in the species were not systematically translated into a high medicinal use-value. Indeed, medicinal and magico-spiritual uses are of high value and then preferably kept secret within families (Giday et al., 2009). This prominence of knowledge sharing within the family source unsurprisingly supported the dominance of the vertical knowledge acquisition type observed in this study (Mattalia et al., 2020). In terms of knowledge mutation, knowledge retained by fathers was preferentially passed to sons, but also knowledge retained by women was passed onto a higher proportion to sons compared to daughters (**Fig. 6.9**), probably to sustain the knowledge maintenance within the family line. Indeed, daughters are called to link to other families through marriage and it is likely that once married they will share the knowledge acquired within their family source with their husbands. So, this fear of girls sharing (even unintentionally) their family knowledge—sometimes considered secret—with their husbands' family precluded parents to transmit detailed or complex knowledge to them (Kidane et al., 2014).

6.4.3 Perception of threats, taboos, and superstitions on *Synsepalum dulcificum*

Although the global assessment by the International Union for Conservation of Nature (IUCN) indicated the miracle plant as a Least Concern species, there has been growing evidence supporting a local depletion of the species in West Africa (Adomou, 2005; Fandohan et al., 2017; Oumorou et al., 2010); including this study's findings as most of the respondents confirmed the number of miracle plants stands in their environment has been sharply decreasing. This study reported 13 various threatening factors for the species in West Africa (**Fig. 6.10**), which included all the six factors previously documented by Fandohan et al. (2017). Unsurprisingly, agricultural expansion and bush fires, which were commonly indicated as threats to biodiversity (Gouwakinnou et al., 2011; Kidane et al., 2014; N'Danikou et al., 2015) were among the most important causes of the miracle plant depletion. The significance of these two causes in the miracle plant is inflated by the low caloric value of the species, which led to its negligence by the population in favour of known staple species, coupled with an overall poor growth rate making it more vulnerable to bush fires. The 56 various medicinal uses recorded in the species (Table 3) mainly targeted the root and the leaves, two crucial organs for the species fitness, but which are overharvested. Such overharvesting inextricably leads to density reduction; delayed growth and reduced seed production with ultimately a likely impact on the species community (Cruse-Sanders and Hamrick, 2004). The overexploitation of various plant parts might partly explain why a reduced number of miracle plant trees was recorded in Benin where medicinal uses of the species were mostly concentrated. The most important threatening factor as indicated by respondents in this study was the erosion of knowledge of the species, a cause that came second in the study of Fandohan et al. (2017). Out of the 76 uses recorded, only one (the fruit used as a sweetener) was saliently known in the studied population, indicating that the knowledge of the 75 other uses which are mostly medicinal and magico-spiritual in nature were heterogeneously distributed. Because the conservation through use hypothesis pinpointed that people are more likely to protect a plant resource when they draw a substantial benefit from it (Newton, 2008); the loss of knowledge of the miracle plant values if no sensitization or awareness-raising actions was taken will continue to threaten the survival of the species in West Africa.

Taboos represented informal institutions whereby social norms lead human behaviour and contributed to biodiversity conservation (Colding and Folke, 2001). Colding and Folke (2001) distinguished six categories of taboos (segment, temporal, method, life history, specific-species and habitats taboos); and the prohibition to set fire close to the tree, to urinate on the tree or to touch the root are indirect measures to ensure better protection of the species and avoid the extraction of some sensitive plant parts (e.g., roots). All these falls in the Colding and Folke (2001)'s life history taboo category. Likewise, the miracle plant bark and roots have medicinal applications and are often harvested using cutlass. Therefore, the prohibition of having cutlass close to species helps indirectly to prevent roots and bark extraction and foster better growth and productivity. All these align with the taboos as a conservation strategy hypothesis, which predicts that certain plant species are protected using taboos handed down by generations (Colding and Folke, 2001; Quiroz and Van Andel, 2015). However, the perceived threatened status of the species despite these protection measures seemed to suggest that the enforcement of such life-history taboos is not stringent enough to slow down or reverse the pace of erosion in the current knowledge of this species.

6.4.4 Determinants of willingness and readiness to cultivate *Synsepalum dulcificum*

Scaling up the cultivation of the miracle plant is crucial for diversifying and sustaining the numerous uses and applications of the species. In this study, we employed classification and regressions tree models (**Figs 6.11–13**) to identify the most important factors that can serve as levers to achieve this objective. Sociolinguistic affiliation was the most important determinant of the respondent's willingness to cultivate the species. Although this factor was previously indicated as a determinant of Moringa' cultivation in Benin Gandji et al. (2018), its importance was to a lesser extent. In this study, the *Akan* and *Ga-adangbe* who least used the miracle plant clearly are not willing to invest in its cultivation. This aligned with the idea that the knowledge of plant use is crucial for cultivation decisions (N'Danikou et al., 2015), though not sufficient to exclusively condition it, as many other factors are likely to interact with it. Two of these factors namely market availability and level of schooling were also previously indicated as important cultivation drivers in other plant species (N'Danikou et al., 2015; Oeba et al., 2012). This study has shown for the first time that taboos, previously pinpointed as a conservation driver, could also act as a key cultivation driver, especially when backgrounds of the species' use clearly exist. Indeed, the various taboos reported in this study contributed to reducing pressure (debarking, branches removal) on the miracle plant, thus fostered an optimal growth of the species, which when coupled with a sound knowledge of use importance guaranteed true benefits for the owner. More importantly, a closer look into our data clearly showed a significantly higher use-value of the species for respondents reporting taboos (UV = 3.37) compared with those not reporting any taboos (UV = 2.41) ($p < 0.003$). Because taboos were conceptualized as a conservation strategy (Colding and Folke, 2001) and that cultivation also serves this conservation purpose in addition to foster a higher use of the species, it came without surprise that participants reporting taboos were willing to cultivate the species, potentially as a measure to sustain and expand the benefits from it. As for market availability, it serves as a financial incentive that offered the possibility to trade the harvest and then to make a substantial profit. Consequently, all the respondents who perceived the existence of markets for the miracle plant were willing to engage in its cultivation except for those who did not school or had a high level of schooling and those who reported the existence of superstition in the species. The decision of respondents who did not school not to engage in the cultivation even when the market is available might be explained by their lack of technical knowledge to apply sound management practices that ensure successful cultivation and or to access various networks to trade their product. As for respondents who schooled, we speculated that their decision not to cultivate could be linked to a lack of time for farming activities as they are likely involved in other formal employment. While previous studies established that willingness to cultivate tree species increased with the level of schooling (N'Danikou et al., 2015; Oeba et al., 2012), we rather identified in this study that intermediate-schooled people were the key group to target in promoting the miracle plant cultivation in West Africa. To the best of our knowledge, no study reported superstition in the miracle plant. Most superstitions reported on the crop ultimately predicted uncomfortable situations or early death for anyone engaged in the cultivation of the species. Because the cost opportunity of benefitting financially from the species at the expense of one’s life is very negligible, the existence of superstitions precluded cultivation, unless the respondent is aware of the early fruiting possibility in the species, knowledge that apparently disrupted the belief in the superstition. A closer look into the most dreadful superstition —"*who plants the miracle plant dies or one of his parents died before it bears fruits”*— recorded in this study revealed that it was likely meant by our forefathers to illustrate the late maturing of the miracle plant but has been misinterpreted. This consequently calls for two urgent actions. Firstly, there is a need for awareness creation to explain to people that they will not die because they have planted the miracle plant tree, but just that the species is slow growing. Secondly, this calls for breeders to accelerate research to shorten the growth cycle to develop early fruiting varieties. Besides, the miracle plant is a rare species only owned by a low proportion of the local population, we expected the tree ownership to also trigger willingness for cultivation as a measure to sustain current uses and benefits from the species, which is not the case in this study.

If knowledge of factors that influence the decision to cultivate a plant species is important in devising promotion strategies, modelling factors that affect how much effort the population are ready to deploy to achieve the cultivation objective is crucial for tangible actions. Because land and planting material are two important production factors, this study used the maximum acreage respondents were ready to allocate to the miracle plant and the maximum purchase price of a seedling as proxies of the respondents’ cultivation effort. Our findings suggest that level of schooling and knowledge of the species biology are of paramount importance in the population readiness to cultivate the miracle plant. While acreage allocation increased with the level of schooling, the proposed purchase price for the seedling rather decreased as the level of schooling becomes extreme (very high or too low). By default, higher schooling is associated with wealth in West Africa and the fact that highly educated respondents were ready to allocate more acreage is explained by the fact that they owned more land (45.2 ha on average) compared with respondents with intermediate and low level of schooling (4.20 ha on average). Contrarily to land ownership that represented a past investment, seedling purchase for cultivation rather stands as an actual or future investment, which respondents with extreme levels of schooling are less ready to make. This observation might be explained by the resource-poor status of respondents with low level of schooling, while in respondents with high level of schooling, it rather strengthened their previously indicated poor willingness to cultivate the miracle plant. Conversely, respondents with intermediate level of schooling in addition to being more willing to cultivate the species also translated it well in terms of both past and future investments for species cultivation. Our results also suggest that providing such a group of respondents with for instance fast-growing miracle plant genotypes would induce more than 100% increase in the acreage they will be willing to allocate for the cultivation of the species (**Fig. 6.12**). Activity category previously indicated as a factor affecting willingness to cultivate kola and Moringa in Benin (Dadjo et al., 2020; Gandji et al., 2018) was also a key determinant of cultivation effort in the miracle plant, with handcraft makers proposing the highest purchase price (USD 1.4) for one seedling. In addition to owning more trees compared with other socio-professional groups, handcraft makers were also the second socio-professional group to use the species most and seek to sustain and expand their current benefits from the species. This might explain why they were ready to pay higher to acquire the miracle plant. For instance, the miracle plant stem (wood) due to its structural quality was reported to be excellent in making agricultural tools (Oumorou et al., 2010), serving as roof and wall poles (Wilfred et al., 2006)and also suitable for sculpture shaping. Interestingly, those among the handcraft makers who did not own the species were even willing to pay more to acquire a seedling. Conversely, the elderly being unlikely to benefit over a long period of the species cultivation, were unsurprisingly only willing to pay USD 0.46 to acquire the species versus USD 1.5 for the young and adults.

6.5 Conclusion

The present study revealed that the miracle plant is a multi-purpose medicinal plant species with a diversity of local names, mostly known for its food value in Benin and Ghana. The species has a higher cultural importance for sociolinguistic groups in Benin compared with their counterparts in Ghana, with magico-spiritual uses being a key segregating component. Youth and adults were as knowledgeable as elderly respondents were, while traditional healers and indigenous religion practitioners valued the species the most. Men were the main source of knowledge and knowledge is mainly acquired along the family line. Our findings also suggested a heterogenous pattern of knowledge acquisition whereby soft knowledge (e.g., food use and social use) was mostly acquired from parents and people of the same generation, while sophisticated knowledge (e.g., magico-therapeutic and medicinal uses) was inherited from parents and grandparents. The predominant ownership mode of the species supported its overall depleting status, which calls for the necessity to take more active conservation measures to sustain utilization. To that end, we demonstrated that socio-cultural, economic and biological factors are key levers to consider when engaging West African (e.g., Benin and Ghana) local populations into the cultivation of the miracle plant. From a cultivation perspective, respondents were overall willing to pay up to 1.9 USD to acquire a seedling and to allocate up to 2.8 ha of their land. These figures were mainly modulated by the sociolinguistic affiliation, the market availability, the instruction level, the knowledge of the species biology, the activity category as well as social beliefs (e.g., taboo and superstition). We established that while taboo could serve as a serious driver for miracle plant cultivation, superstition clearly precludes it, hence the necessity for more investigation of social belief systems in other orphan crops. Our findings represent key decision-making tools to hasten the promotion of the miracle plant in the study area, and in West Africa in general, specifically suggesting respondents with an intermediate level of schooling as key targets to achieve this objective. Our findings also provide evidence supporting a depletion of *S. dulcificum* in West Africa caused mainly by agricultural expansion and bush fire, while the global assessment by IUCN indicated the miracle plant as a Least Concern Species. This calls for the necessity to review the status of the species and to undertake adequate measures for *in-situ* and *ex-situ* conservations.

6.6 Reference

Abbam, T., Johnson, F. A., Dash, J., and Padmadas, S. S. (2018). Spatiotemporal variations in rainfall and temperature in Ghana over the twentieth century, 1900–2014. *Earth Space Sci.* **5**, 120-132.

Achigan-Dako, E. G., N'Danikou, S., Assogba-Komlan, F., Ambrose-Oji, B., Ahanchede, A., and Pasquini, M. W. (2011). Diversity, geographical, and consumption patterns of traditional vegetables in sociolinguistic communities in Benin: implications for domestication and utilization. *Econ. Bot.* **65**, 129.

Achigan-Dako, E. G., Tchokponhoué, D. A., N'Danikou, S., Gebauer, J., and Vodouhè, R. S. (2015). Current knowledge and breeding perspectives for the miracle plant *Synsepalum dulcificum* (Schumach & Thonn.) Daniell. *Genet. Resour. Crop Evol.* **62**, 465-476.

Adomou, A. (2005). Vegetation patterns and environmental gradients in Benin, University of Wageningen, Wageningen, Netherlands.

Agbani, P. O., Kafoutchoni, K. M., Salako, K. V., Gbedomon, R. C., Kégbé, A. M., Karen, H., and Sinsin, B. (2018). Traditional ecological knowledge-based assessment of threatened woody species and their potential substitutes in the Atakora mountain chain, a threatened hotspot of biodiversity in Northwestern Benin, West Africa. *J. Ethnobiol. Ethnomed.* **14**, 21.

Agbo, R. I., Vihotogbé, R., and Missihoun, A. A. e. a. (2020). Indigenous knowledge of *Detarium microcarpum* Guill. & Perr. (Caesalpiniaceae) and implication for conservation in Benin (West Africa). *Environ. Dev. Sustain.* **22**, 6261–6285.

Assogba, G. A., Fandohan, A. B., Salako, V. K., and Assogbadjo, A. E. (2017). Usages de *Bombax costatum* (Malvaceae) dans les terroirs riverains de la Réserve de biosphère de la Pendjari, République du Bénin. *Bois For. Trop.* **333**, 17-29.

Bakwaye, F. N., Termote, C., Kibungu, K., and Van Damme, P. (2013). Identification and local importance of medicinal plants used in the Mbanza-Ngungu region, Democratic Republic of Congo. *Bois For. Trop.*, 63-77.

Bommer, C., Sagalova, V., Heesemann, E., Manne-Goehler, J., Atun, R., Bärnighausen, T., Davies, J., and Vollmer, S. (2018). Global economic burden of diabetes in adults: projections from 2015 to 2030. *Diabetes Care* **41**, 963-970.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA-Cancer J. Clin.* **68**, 394-424.

Buckmire, R., and Francis, F. (1978). Pigments of miracle fruit, *Synsepalum dulcificum*, Schum, as potential food colorants. *J. Food Sci.* **43**, 908-911.

Chen, X. W., Abdullah, T. L., Abdullah, N. A. P., and Hassan, S. A. (2012). Rooting response of miracle fruit (*Synsepalum dulcificum*) softwood cuttings as affected by indole butyric acid. *Am. J. Agric. Biol. Sci.* **7**, 442-46.

Colding, J., and Folke, C. (2001). Social taboos:“invisible” systems of local resource management and biological conservation. *Ecol. Appl.* **11**, 584-600.

Cruse-Sanders, J. M., and Hamrick, J. (2004). Genetic diversity in harvested and protected populations of wild American ginseng, *Panax quinquefoliu*s L.(Araliaceae). *Am. J. Bot.* **91**, 540-548.

Dadjo, C., Nyende, A. B., Salako, K. V., Hounkpèvi, A., and Assogbadjo, A. E. (2020). Socio–economic factors determining conservation and cultivation of *Garcinia kola* Heckel—A medicinal plant extinct in the wild in Benin. *Econ. Bot.* **74**, 115-125.

de Albuquerque, U. P., Soldati, G. T., Sieber, S. S., Ramos, M. A., de Sá, J. C., and de Souza, L. C. (2011). The use of plants in the medical system of the Fulni-ô people (NE Brazil): A perspective on age and gender. *J. Ethnopharmacol.* **133**, 866-873.

De Caluwé, E., De Smedt, S., Assogbadjo, A., Samson, R., Sinsin, B., and Van Damme, P. (2009). Ethnic differences in use value and use patterns of baobab (A*dansonia digitata* L.) in northern Benin. *Afr. J. Ecol.* **47**, 433-440.

De Los Angeles La Torre-Cuadros, M., and Islebe, G. A. T. (2003). Traditional ecological knowledge and use of vegetation in southeastern Mexico: a case study from Solferino, Quintana Roo. *Biodivers. Conserv.* **12**, 2455-2476.

Del Campo, R., Zhang, Y., and Wakeford, C. (2017). Effect of miracle fruit (*Synsepalum dulcificum*) seed oil (MFSO®) on the measurable improvement of hair breakage in women with damaged hair: a randomized, double-blind, placebo-controlled, eight-month trial. *J. Clin. Aesthet. Dermat.* **10**, 39-48.

Ebifa-Othieno, E., Mugisha, A., Nyeko, P., and Kabasa, J. D. (2017). Knowledge, attitudes and practices in tamarind (T*amarindus indica* L.) use and conservation in Eastern Uganda. *J. Ethnobiol. Ethnomed.* **13**, 1-13.

Ekpo, B. A., Bala, D. N., Essien, E. E., and Adesanya, S. A. (2008). Ethnobotanical survey of Akwa Ibom state of Nigeria. *J. Ethnopharmacol.* **115**, 387-408.

Ekué, M. R., Sinsin, B., Eyog-Matig, O., and Finkeldey, R. (2010). Uses, traditional management, perception of variation and preferences in ackee (*Blighia sapida* KD Koenig) fruit traits in Benin: implications for domestication and conservation. *J. Ethnobiol. Ethnomed.* **6**, 12.

Fandohan, A. B., Gouwakinnou, G. N., Tovissode, C. F., Bonou, A., Djonlonkou, S. F. B., Houndelo, L. F., Sinsin, C. L. B., and Assogbadjo, A. E. (2017). Usages traditionnels et valeur économique de *Synsepalum dulcificum* au Sud-Bénin. *Bois For. Trop.* **332**, 17-30.

Fazilah, N., Zani, N., Wasoh, H., Ariff, A., and Halim, M. (2020). Influence of miracle fruit (*Synsepalum dulcificum*) extract and microencapsulated Lactococcus lactis Gh1 on the antioxidant activity and probiotic viability of functional yogurt. *Int. Food Res. J.* **27**.

Gandji, K., Salako, V. K., Fandohan, A. B., Assogbadjo, A. E., and Kakaï, R. L. G. (2018). Factors determining the use and cultivation of Moringa oleifera Lam. in the Republic of Benin. *Econ. Bot.* **72**, 332-345.

Gaoue, O. G., Coe, M. A., Bond, M., Hart, G., Seyler, B. C., and McMillen, H. (2017). Theories and major hypotheses in ethnobotany. *Econ. Bot.* **71**, 269-287.

Giday, M., Asfaw, Z., Woldu, Z., and Teklehaymanot, T. (2009). Medicinal plant knowledge of the Bench ethnic group of Ethiopia: anethnobotanical investigation. *J. Ethnobiol. Ethnomed.* **5**, 1-10.

Gorin, S., Wakeford, C., Zhang, G., Sukamtoh, E., Matteliano, C. J., and Finch, A. E. (2018). Beneficial effects of an investigational wristband containing *Synsepalum dulcificum* (miracle fruit) seed oil on the performance of hand and finger motor skills in healthy subjects: A randomized controlled preliminary study. *Phytother. Res.* **32**, 321-332.

Gouwakinnou, G. N., Lykke, A. M., Assogbadjo, A. E., and Sinsin, B. (2011). Local knowledge, pattern and diversity of use of *Sclerocarya birrea*. *J. Ethnobiol. Ethnomed.* **7**, 1-9.

Hanazaki, N., Herbst, D. F., Marques, M. S., and Vandebroek., I. (2013). Evidence ofthe shifting baseline syndrome in ethnobotanical research. *J. Ethnobiol. Ethnomed.* **9**.

Heldmann, M. (2008). Ethnic groups in Benin. *In* "IMPETUS Atlas Benin. Research results" (M. Judex and H.-P. Thamm, eds.), Vol. 2007, pp. 109-110. Department of Geography, University of Bonn, Germany.

ICPC2 (2005). International classification of primary care 2: revised edition.

Inglett, G., and May, J. F. (1968). Tropical plants with unusual taste properties. *Econ. Bot.* **22**, 326-331.

International Diabetes Federation (2019). IDF Diabetes Atlas,. Brussels, Belgium.

Joyner, G. (2006). The miracle fruit. *In* "Quandong magazine of the west Australian nut and tree crop association" (P. Scott, ed.), pp. 15. West Australian Nut and Tree Crop Association, Subiaco, West Australia.

Kidane, B., van Andel, T., van der Maesen, L. J. G., and Asfaw, Z. (2014). Use and management of traditional medicinal plants by Maale and Ari ethnic communities in southern Ethiopia. *J. Ethnobiol. Ethnomed.* **10**, 46.

Koura, K., Ganglo, J. C., Assogbadjo, A. E., and Agbangla, C. (2011). Ethnic differences in use values and use patterns of *Parkia biglobosa* in Northern Benin. *J. Ethnobiol. Ethnomed.* **7**, 1-12.

Kurihara, K., and Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science* **161**, 1241-1243.

Kurihara, K., and Beidler, L. M. (1969). Mechanism of the action of taste-modifying protein. *Nature* **222**, 1176–1179.

Leiper, I., Zander, K. K., Robinson, C. J., Carwadine, J., Moggridge, B. J., and Garnett, S. T. (2018). Quantifying current and potential contributions of Australian indigenous peoples to threatened species management. *Conserv. Biol.* **32**, 1038-1047.

Malan, D. F., and Neuba, D. F. R. (2021). Wild edible plants in the Ehotilé, a fishing people around Aby lagoon (eastern littoral of Côte d'Ivoire): Knowledge and availability. *J. Appl. Nat. Sci.* **13**, 59-70.

Mattalia, G., Stryamets, N., Pieroni, A., and Sõukand, R. (2020). Knowledge transmission patterns at the border: ethnobotany of Hutsuls living in the Carpathian Mountains of Bukovina (SW Ukraine and NE Romania). *J. Ethnobiol. Ethnomed.* **16**, 41.

Menéndez-Rey, A., González-Martos, R., Ye, P., Quiroz-Troncoso, J., Alegría-Aravena, N., Sánchez-Díez, M., Maestu-Unturbe, C., Bensadon-Naeder, L., and Ramírez-Castillejo, C. (2021). Quantification of lectins in *Synsepalum dulcificum* and comparison with reference foods. *Food Chem.* **352**, 129341.

Milborrow, S., and Milborrow, M. S. (2020). Package ‘rpart. plot’.

Mongalo, N. I., and Makhafola, T. J. (2018). Ethnobotanical knowledge of the lay people of Blouberg area (Pedi tribe), Limpopo Province, South Africa. *J. Ethnobiol. Ethnomed.* **14**, 46.

N’Danikou, S., Achigan-Dako, E. G., Tchokponhoue, D. A., Agossou, C. O., Houdegbe, C. A., Vodouhe, R. S., and Ahanchede, A. (2015). Modelling socioeconomic determinants for cultivation and in-situ conservation of *Vitex doniana* Sweet (Black plum), a wild harvested economic plant in Benin. *J. Ethnobiol. Ethnomed.* **11**, 28.

Nanyingi, M. O., Mbaria, J. M., Lanyasunya, A. L., Wagate, C. G., Koros, K. B., Kaburia, H. F., Munenge, R. W., and Ogara, W. O. (2008). Ethnopharmacological survey of Samburu district, Kenya. *J. Ethnobiol. Ethnomed.* **4**, 1-12.

Newton, A. C. (2008). Conservation of tree species through sustainable use: how can it be achieved in practice? *Oryx* **42**, 195-205.

Ng, T., Chong, T., and Du, X. (2010). The value of superstitions. *J. Econ. Psychol.* **31**, 293–309.

Obafemi, T., Akinmoladun, A., Olaleye, M., Agboade, S. O., and Onasanya, A. A. (2017). Antidiabetic potential of methanolic and flavonoid-rich leaf extracts of *Synsepalum dulcificum* in type 2 diabetic rats. *J-AIM* **8**, 238-246.

Obafemi, T. O., Olaleye, M. T., and Akinmoladun, A. C. (2019). Antidiabetic property of miracle fruit plant (*Synsepalum dulcificum* Shumach. & Thonn. Daniell) leaf extracts in fructose-fed streptozotocin-injected rats via anti-inflammatory activity and inhibition of carbohydrate metabolizing enzymes. *J. Ethnopharmacol.* **244**, 112124.

Obeng, H. B. (1971). National soil map of Ghana. Catrographic Section of the Soil Research Institute.

Oeba, V. O., Otor, S. C. J., Kung'u, J. B., and Muchiri, M. N. (2012). Modelling determinants of tree planting and retention on farm for improvement of forest cover in central Kenya. *ISRN Forestry* **2012**, 867249.

Ouédraogo, K., Dimobe, K., Zerbo, I., Etongo, D., Zare, A., and Thiombiano, A. (2019). Traditional knowledge and cultural importance of *Gardenia erubescens* Stapf & Hutch. in Sudanian savanna of Burkina Faso. *J. Ethnobiol. Ethnomed.* **15**, 1-14.

Oumorou, M., Dah-Dovonon, J., Aboh, B., Hounsoukaka, M., and Sinsin, B. (2010). Contribution á la conservation de *Synsepalum dulcificum*: régénération et importance socio-économique dans le département de l'ouémé (Bénin). *Ann. Sci. Agron* **14**, 101-120.

Phillips, O., and Gentry, A. H. (1993). The useful plants of Tambopata, Peru: I. Statistical hypotheses tests with a new quantitative technique. *Econ. Bot.* **47**, 15-32.

Pilleron, S., Soto-Perez-de-Celis, E., Vignat, J., Ferlay, J., Soerjomataram, I., Bray, F., and Sarfati, D. (2021). Estimated global cancer incidence in the oldest adults in 2018 and projections to 2050. *Int. J. Cancer* **148**, 601-608.

Quiroz, D., and Van Andel, T. (2015). Evidence of a link between taboos and sacrifices and resource scarcity of ritual plants. *J. Ethnobiol. Ethnomed.* **11**, 1-11.

R Core Team (2019). "R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,Vienna, Austria. URL https://www.R-project.org/," Vienna, Austria.

Ribeiro, R. V., Bieskia, I. G. C., Baloguna, S. O., and Martinsa, D. T. O. (2017). Ethnobotanical study of medicinal plants used by Ribeirinhos in the North Araguaia microregion, Mato Grosso, Brazil. *J. Ethnopharmacol.* **205**, 69–102.

Rodrigues, J. F., da Silva Andrade, R., Bastos, S. C., Coelho, S. B., and Pinheiro, A. C. M. (2016). Miracle fruit: An alternative sugar substitute in sour beverages. *Appetite* **107**, 645-653.

Rokni, S., Wursten, B., and Darbyshire, I. (2019). *Synsepalum chimanimani* (Sapotaceae), a new species from the Chimanimani Mountains of Mozambique and Zimbabwe, with notes on the botanical importance of this area. *PhytoKeys* **133**, 115.

Salako, K. V., Moreira, F., Gbedomon, R. C., Tovissodé, F., Assogbadjo, A. E., and Kakaï, R. L. G. (2018). Traditional knowledge and cultural importance of *Borassus aethiopum* Mart. in Benin: interacting effects of socio-demographic attributes and multi-scale abundance. *J J. Ethnobiol. Ethnomed.* **14**.

Si, A. (2020). Patterns in the transmission of traditional ecological knowledge: a case study from Arnhem Land, Australia. *J. Ethnobiol. Ethnomed.* **16**, 52.

Tchokponhoué, D., Achigan-Dako, E., N'Danikou, S., Houdégbé, A., Agossou, C., Assogba-Komlan, F., and Vodouhè, R. (2018). Regeneration ability and seedling growth in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Fruits* **73**, 13-21.

Tchokponhoué, D. A., Achigan-Dako, E. G., N'Danikou, S., Nyadanu, D., Kahane, R., Odindo, A. O., and Sibiya, J. (2021). Comparative analysis of management practices and end-users' desired breeding traits in the miracle plant [*Synsepalum dulcificum* (Schumach & Thonn.) Daniell] across ecological zones and sociolinguistic groups in West Africa. *J. Ethnobiol. Ethnomed.* **17**.

Tchokponhoué, D. A., N'Danikou, S., Houéto, J. S., and Achigan-Dako, E. G. (2019). Shade and nutrient-mediated phenotypic plasticity in the miracle plant *Synsepalum dulcificum* (Schumach. & Thonn.) Daniell. *Sci Rep.* **9:5137**, 1-11.

Teka, A., Asfaw, Z., Demissew, S., and Van Damme, P. (2020). Medicinal plant use practice in four ethnic communities (Gurage, Mareqo, Qebena, and Silti), south central Ethiopia. *J. Ethnobiol. Ethnomed.* **16**, 1-12.

Therneau, T., Atkinson, B., Ripley, B., and Ripley, M. B. (2015). Package 'rpart'. *Available online: cran. ma. ic. ac. uk/web/packages/rpart/rpart. pdf (accessed on 20 April 2016)*.

Vinagre, C., Vinagre, S., and Carrilho, E. (2019). The use of medicinal plants by the population from the protected Landscape of "Serra de Montejunto", Portugal. *J. Ethnobiol. Ethnomed.* **15**, 1-30.

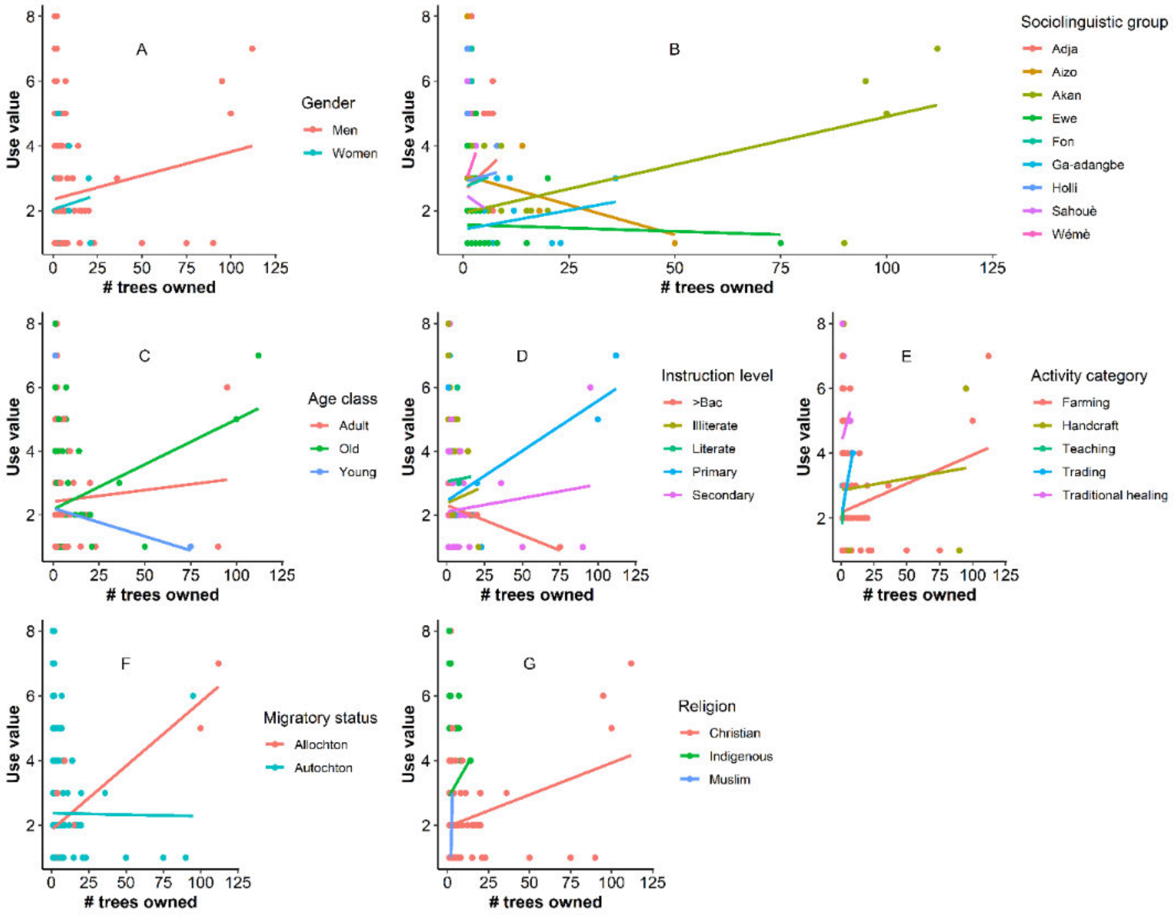
Voeks, R. A. (2007). Are women reservoirs of traditional plant knowledge? Gender, ethnobotany and globalization in northeast Brazil. *Singapore J Trop Geogr.* **28**, 7-20.

Whitney, C. (2019). EthnobotanyR: calculate quantitative ethnobotany indices. Version.

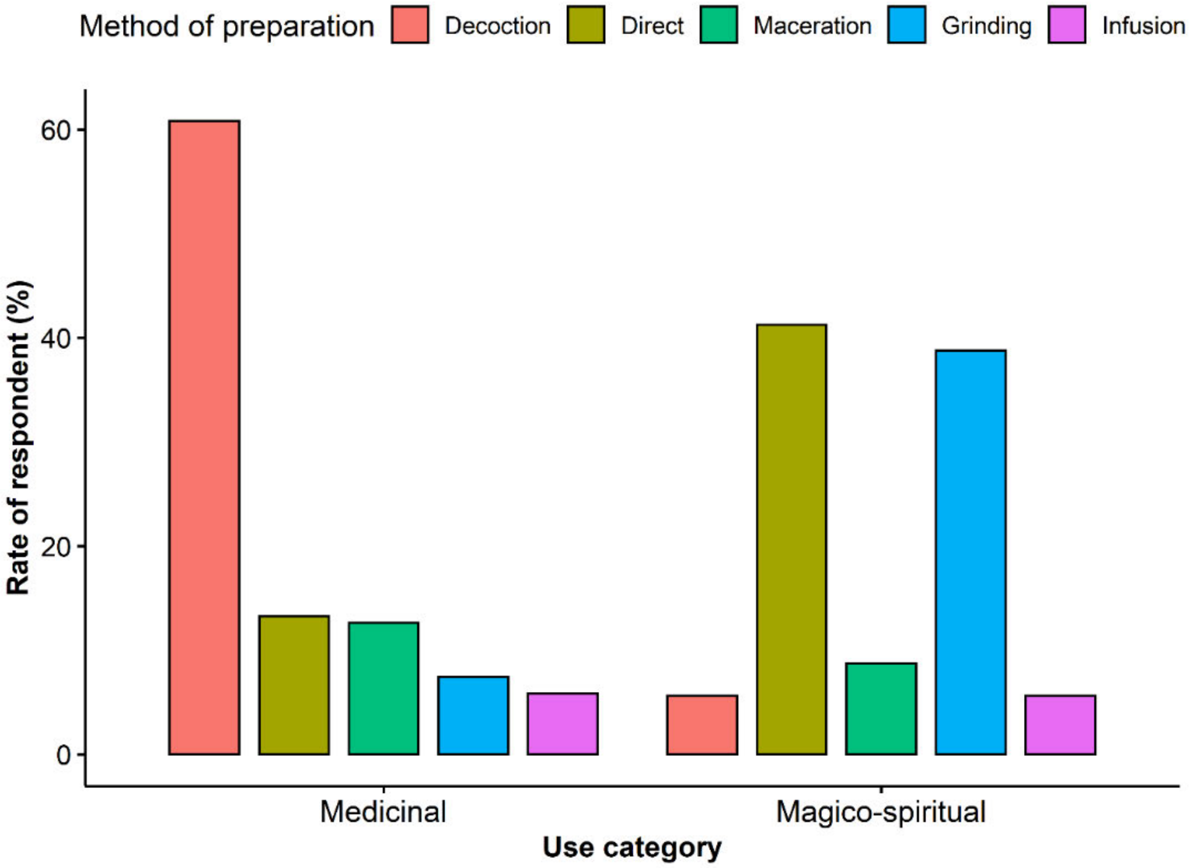
Wilfred, P., Madoffe, S. S., and Luoga, E. J. (2006). Indigenous plant uses and use values in Uluguru mountains, Morogoro, Tanzania. *J East Afric. Nat. History* **95**, 235-240.

Wilken, M. K., and Satiroff, B. A. (2012). Pilot study of "miracle fruit" to improve food palatability for patients receiving chemotherapy. *Clin. J. Oncol. Nurs.* **16**, E173-E177.

Yanagida, T. (2020). Package 'misty'.



Appendix 6.1. Regressions lines illustrating the association between use-value and number of miracle plant trees for attributes of different socio-demographic factors in Benin and Ghana.



Appendix 6.2. Relative importance of different preparation methods (all plant parts combined) for medicinal and magico-spiritual applications of the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana.

Appendix 6.3. Taboos and superstitions reported in the miracle plant (*Synsepalum dulcificum*) in Benin and Ghana. Taboos and superstitions in bold were recorded in Ghana while those in regular font were recorded in Benin

Taboos*	Concerned sociolinguistic group	Relative frequency (%)	Superstitions*	Relative frequency (%)	Concerned sociolinguistic group
It is prohibited to bring fire close to the miracle plant tress	Adja Aizo	35.28	Someone who plants the species will die or one of his parents will die before the plant starts bearing fruits	83.70	Adja Aizo
It is forbidden to bring salt close to the tree	Fon	17.60	Only the elderly people plant the species	9.30	Fon
It is prohibited to bring cutlass close to the species	Holli	5.89	**Chewing the stick of the miracle plant attracts mishaps**	2.33	Holli
It is prohibited to hand-harvest the leaves	Sahouè	5.89	**Having a miracle plant tree at home attracts problems and mishaps**	2.33	Sahouè
Prohibition to touch the species root	Wémé	5.89	**Having a miracle plant tree makes your problems last forever**	1.17	Ga-adangbe
Prohibition to climb the tree	Ga-adangbe	5.89	**The miracle plant trees are home to witches**	1.17	
It is forbidden to sweep under the tree or to make a broom touch the tree		5.89	_		
Prohibition to pee on the tree		5.89	_		
Menstruating women do not approach the tree		5.89	_		
A pregnant woman does not approach the tree		5.89	_		

*Taboos and superstitions in bold were recorded in Ghana while those in regular font were recorded in Benin.

CHAPTER 7

General discussion and conclusion

7.1 General discussion

Despite its status as a neglected and underutilized species, *Synsepalum dulcificum* has attracted increased interest in recent decades, and the necessity to meet the growing demand for species products supports the need to develop improved cultivars. This thesis is the first comprehensive work to engage the species on the breeding path. Our findings advance the knowledge on this minor tree crop and have practical implications for the future of the species breeding in West Africa.

7.1.1 Key breeding traits for the miracle plant in West Africa identified

Proper identification of priority traits allows the formulation of sound breeding objectives, which are crucial for developing a widely accepted variety (Mazanec et al., 2021; Mtshali et al., 2021), especially in perennial species known to have an intrinsically long breeding cycle. A total of 23 key breeding traits were identified for *S. dulcificum* from three main user groups in West Africa, highlighting the large spectrum of preferences. These traits are likely to constitute the basis of cultivar development in the species. However, the prioritization bias towards traits such as fruit size and fruit miraculin content observed in all three user groups showed a strong tendency towards quality attribute traits in the species. This was confirmed with the commercial exploitation of the miracle fruit in Florida (United States of America), where larger fruits with high miraculin were also sought after (Demesyeux et al., 2020).

Additionally, highly desired traits included early maturity, high yield and tree dwarfism, which are specific to farmers; fruit colour sharpness and shelf-life highly rated by consumers; edible ratio suggested by both consumers and processing companies; and seed part ratio exclusive to the latter ones. These priority traits for *S. dulcificum* breeding aligned with key breeding traits in other major horticultural fruit tree crops, such as peach [*Prunus persica* (L.) Batsch], apple (*Malus x domestica* Borkh.) and grape (*Vitis vinifera* L.), which were related to fruit size, appearance, taste and texture, among others (Cirilli et al., 2021; Li et al., 2020; Zhao et al., 2017). Breeding an ideal variety of the miracle plant combining these key traits could consequently benefit from experiences already accumulated in these heavily bred crops.

Traits related to abiotic and biotic stresses appeared weakly prioritized by end-users, probably because of the currently poor cultivation status of the species. However, based on the taste and the appearance of the miracle fruit, we can speculate that pest pressure would be unavoidable and could even rise as cultivation intensity increases. Hence, it is necessary to integrate resistance/tolerance to biotic and abiotic stresses as an additional trait to breed for, especially in the climate change context.

7.1.2 Phenotypic groups for the miracle plant in West Africa and variability of end-users most desired fruit traits

Nine traits were phenotypically evaluated and these included three of the top end-users' desired traits in West Africa: fruit mass, fruit edible ratio and seed part ratio. The profiling of the variability of these nine traits showed low to moderate levels of variation, with fruit mass (the top priority breeding trait in the species) being the most variable trait. This, combined with the high repeatability, suggested the possibility of realizing genetic gain for this key trait.

The overall high and positive phenotypic correlation of fruit mass with other priority fruit traits, such as the edible ratio, if confirmed through a genetic correlation, could offer the possibility of breeding for these traits simultaneously. Likewise, breeding efficiently for economic traits of the species, such as edible ratio and seed part ratio that are negatively correlated would require the development of an index selection. Since the economic importance of these two traits for the food processing and cosmetics industries is known, their improvement could benefit from economic selection index development. For such selection indices to be extensively applied in the species, understanding the correlation pattern of the fruit mass to other highly desired traits such as fruit miraculin content, fruiting time, yielding, shelf life and height must be considered in prospective studies in the species. Although a previous study estimated fruit and miraculin yield in the miracle plant juvenile (Demesyeux et al., 2020), no indication of the nature and strength of the correlation between these two traits was reported. Yet, of particular importance is the quantification of the natural variation in miraculin content for future breeding efforts, as this is a landmark trait of the species. The large distribution of the species in West Africa, a region with contrasting ecological conditions (White, 1983), offers a high probability of identifying promising individuals for this trait. Three clusters were detected for the nine traits evaluated on 322 West African individual trees sampled from seven populations belonging to two ecological regions, namely Upper Guinea (with three populations: West, Central, Eastern) and Dahomey Gap (with four populations: Volta, Mono, Zou and Oueme). From theses clusters, promising individuals could be selected for various breeding purposes.

7.1.3 Gene pools detected for the miracle plant

Phenotypic and molecular assessments are two complementary approaches for plant genetic resources characterization. Even though the Upper Guinea individuals, in general, seem to outperform their counterparts of the Dahomey Gap, the phenotypic evaluation also revealed that the Volta population (Dahomey Gap) performed the best for four (fruit mass, seed mass, fruit length and fruit width) out of the six fruit traits evaluated. Similarly, this Volta population was also demarcated from other populations by its moderate genetic diversity, which contrasts with the overall low diversity observed in the species. This is indicative of beneficial genetic variation within the Volta population that can be exploited. Both phenotypic and molecular approaches concurred on three genetic clusters in West Africa on the one hand, and on the centre of diversity of the species on the other hand.

Nevertheless, clusters' constitution differed significantly. While the molecular evaluation clustered the accessions based on the ecological region of provenance, clusters generated from phenotypic evaluation were all heterogeneous, each encompassing individuals from all investigated populations (West, Central, Eastern, Volta, Mono, Zou and Oueme). This discrepancy observed in the cluster's constitution may be attributed to the "in situ" rather than "a common garden" approach used for the phenotypic assessment, which may have inflated the environmental residual. We speculated that a common-garden phenotypic evaluation might generate more concurring results. However, because molecular assessment provides more reliable information, as it is not influenced by the environment, the SNP-marker based cluster constitution was considered sufficiently sound for the species.

The three gene pools observed strongly diverged and this suggests a poor tradition of germplasm exchange between communities in the Dahomey Gap and those in Upper Guinea, which was probably exacerbated by the existence of a natural barrier between the two ecological regions. Conversely, germplasm exchange may have been intense among populations belonging to the same ecological regions, as illustrated, for instance, by the almost null differentiation between the Western and Eastern populations (Fst = 0.01, Chapter 4) in Upper Guinea despite the high geographical distance between them. The three gene pools delimited by the molecular analyses included the Upper Guinea gene pool (UG), Western Dahomey Gap (WDG) and Central Dahomey Gap (CDG). The low natural diversity observed in the species also opens room for investigating mutation breeding in the species to induce variability (Chambers et al., 2018).

7.1.4 Relationship between miracle plant traditional knowledge distribution and genetic diversity in the Dahomey Gap and Upper Guinea

Indigenous communities hold strong knowledge about their resources (Reyes-García et al., 2013), and such knowledge is shared through social networks such as social affiliation (sociolinguistic group membership), organizational affiliation (organization membership), and societal kinship (e.g., marriage), among others (Delêtre et al., 2011). In the African context, germplasm flow among communities occurred with associated traditional knowledge, as indicated by Kiptot et al. (2006). Consequently, we expected *ceteris paribus* to have a low gap in traditional knowledge held by communities connected by the miracle plant germplasm exchange. This thesis documented a total of 76 uses for the miracle plant, with local communities in the Dahomey Gap valuing the species more significantly than their counterparts in the Upper Guinea, and strikingly. Magico-spiritual use category was only reported by communities in the Dahomey Gap. This gap in magico-spiritual knowledge apart from partly stemming from the difference in sociocultural backgrounds may have also been inflated by the lack of germplasm exchange between communities in these two regions.

The three genetic populations delimited by the molecular analysis namely Upper Guinea (UG), Western Dahomey Gap (WDG) and Central Dahomey Gap (CDG) are geographically laid on three regions demarcating three culturally distinct groups. The Upper Guinea cultural group (UG_g) encompasses the *Akan* and *Ga-adangbe* sociolinguistic groups. The Western Dahomey Gap cultural group (WDG_g) mainly has the *Ewe* sociolinguistic group, whereas the *Aïzo, Sahouè, Adja, Fon, Wémé* and *Holli* sociolinguistic groups make up the Central Dahomey Gap cultural group (CDG_g), while genetic diversity declined in the species following this order: $\boldsymbol{\pi}$Western Dahomey Gap > $\boldsymbol{\pi}$Upper Guinea > $\boldsymbol{\pi}$Central Dahomey Gap. The traditional knowledge held by cultural groups in the corresponding geographical regions declined in the reverse order $UV_{CDG\_g} > UV_{UG\_g} > UV_{WDG\_g}$. This inverse relationship between genetic diversity and use value suggests that the high exploitation of the miracle plant in the Central Dahomey Gap could have contributed to significantly narrowing the species genetic diversity in the region. This hypothesis is highly plausible, as locally, and in the CDG, the bark, roots and leaves, which are known to ensure species fitness, are abundantly harvested for magico-spiritual uses and folk medicines. Previous studies established that an exploitation targeting organs such as foliage or bark, which are not directly involved in plant species survival, is unlikely to affect genetic diversity or only affect it at a slower pace (Gaoue et al., 2014) compared with exploitation targeting, for instance, roots that irreversibly erase individuals from their gene pools (Cruse-Sanders and Hamrick, 2004; Mooney, 2007). In the case of the miracle plant, both types of organs are exploited, and this raises serious concerns on the future of the species if no active action is taken.

7.1.5 Raising awareness and genetic improvement in the cultivation of the miracle plant

Due to its numerous applications, the miracle plant is highly sought-after both locally and internationally. Therefore, it is important that proper cultivation be encouraged in the local communities to reverse its depletion in the natural populations. This study showcased the prominence of sociolinguistic affiliation, market availability and knowledge of cultural beliefs as factors affecting population willingness to cultivate the species. As reported in Chapter 2, communities in the *Akan* sociolinguistic group are currently the main miracle plant plantation owners in the study area, and as such, they were expected to be willing to engage in large-scale cultivation of the species. Unexpectedly, this was not the case, probably because of their overall poor access to international markets and locally, the species is not well valued. Currently, the miracle plant market seems to be a closed one, with access conditioned, for instance, by the level of Western education, while most of the species owners are lowly educated. This importance of market availability in the decision to cultivate the species was later confirmed by some sociolinguistic groups in Dahomey Gap who would like to economically depend on the crop from which they are mostly obtaining only a non-market interest (Fandohan et al., 2017).

The study showed that the belief in social norms such as taboos by sociolinguistic groups in the CDG has served as an additional incentive to cultivate the species since taboos are conceptualized to serve as a conservation strategy, a purpose also targeted by cultivation (Colding and Folke, 2001). Conversely, the belief in superstitions that wrongly relate the species cultivation to the grower’s early death results from a complex interaction between the real plant biology (late maturing species) and an attempt by the forefathers to translate it. This belief needs to be disrupted through awareness raising on the species potential and early maturing cultivar development. In addition, having an early fruiting variety would motivate farmers to cultivate the species in the face of superstitions. This result ultimately matches the breeding trait prioritization reported in Chapter 2, which revealed early fruiting as one of farmers' top three desired traits.

7.2 Implications of findings for genomic selection strategy optimization in the miracle plant

Since *Synsepalum dulcificum* is a perennial species, genomic selection (GS) currently stands as the most significant breeding strategy to realize accelerated genetic gains. Considering the set of information generated in this thesis, it is suggested that a multi-trait genomic selection scheme supported by selection index development to hasten the improvement of functional traits such as fruit size, miraculin content, early maturity, yield, and edible ratio in the species be implemented. Population structuring and sub-structuring have a major impact on GS accuracy (Guo et al., 2014; Olatoye et al., 2020; Windhausen et al., 2012). Considering the different gene pools detected and their size, running two parallel GS schemes on the Upper Guinea diversity panel and on the Dahomey Gap, panel is recommended. With its too low size, the Edge Dahomey Gap diversity panel would not, in its current state, be involved in a GS scheme. Rather, it seems well suited to serve in a diallel cross scheme to help study combining ability, gene action, narrow-sense heritability, and the extent of genetic correlation among various traits of interest in the species. Indeed, this Edge Dahomey Gap diversity panel is mainly composed of individuals of the Volta population, which are known to be high performers. In addition, proper experimental quantification of the narrow-sense heritability is required for all traits to establish selection indices in the species and feed different GS scenarios (Lenz et al., 2020). Nevertheless, individuals in the full-sib families generated from this EDG panel can later be genotyped and utilized in the GS scheme.

The accuracy of GS prediction is affected by a number of factors, including the genetic architecture of traits of interest, the training population size and its constitution, and its relationship with the candidate population set (Rincent et al., 2012; Windhausen et al., 2012). Strategies for optimizing the training population constitution have been largely investigated (Asoro et al., 2011; Hickey et al., 2014; Verges and Van Sanford, 2020; Zhao et al., 2012) with various analytical frameworks and tools suggested. Basically, a number of packages, including the “STPGA” (Akdemir, 2017), “TSDFGS” (Ou et al., 2019), “TrainSel” (Akdemir et al., 2021) and approaches such as the coefficient of determination mean (CDmean) procedure (Rincent et al., 2012) and the predictive error variance (PEV) (Tiede and Smith, 2018), have been indicated to help construct a good training population set (TRS). The miracle plant can benefit from the “TrainSel” package, which has a large set of parametrization options. Interestingly, “TrainSel” can combine TRS optimization with a particular experimental design, thus solving two problems simultaneously.

Alternatively, since the rationale in constituting a training population rests on the inclusion of as many diverse individuals to phenotype as possible, the development of a core collection within each diversity panel to serve as a training population also stands as a good approach that can be explored. Regarding the training population size, no one-size-fits-all suggestion exists and current trends in the literature rely on a sufficiently large size for the training population (Jiang et al., 2020), which can, however, be relaxed when genotypes whose performance is to be predicted are close to the training set (Larkin et al., 2019). This is the case within the various miracle plant diversity panels. Consequently, the training set would be optimized from a calibration set composed of 60% of individuals within each diversity panel, which practically corresponds to 117 and 67 individuals, respectively, for the calibration set in the Dahomey Gap and Upper Guinea diversity panels.

Under the genomic selection scheme, the function of trait phenotyping has evolved from being simply informative to becoming highly predictive, as it is important in genetic model construction and validation (Lenz et al., 2020). However, in perennial species, to the timing of phenotyping a trait has been debated in regard to genomic selection (Ballesta et al., 2018; Kainer et al., 2018; Resende Jr et al., 2012), as juvenile-age expression may not differ from mature-age expression. In the case of the miracle plant, for instance, while phenotyping a trait such as time to fruiting is relatively simple, when to phenotype traits such as fruit mass, fruit yield and miraculin content is currently unknown; and in absence of a unified framework can be challenging. Nevertheless, apart from time to flowering and time to fruiting, we can recommend, based on our long experience with the species, to phenotype the training population set at various ages of 3 years, 5 years, and 10 years. Experimentally, this could provide enough data to test for any-age specific accuracy in GS prediction. For each diversity panel, testing environments could include sites in both the Dahomey Gap and Upper Guinea (e.g., Benin, Togo, and Ghana) in case of collaborative research or different environments within each ecological region (in the case of limited resources).

Depending on the genetic architecture of the target traits (Viana et al., 2017), estimates of genomic breeding values can be based on a diversity of historically known methods, including the ridge regression BLUP (RR-BLUP) method, Bayesian methods (De Los Campos et al., 2009; Gianola, 2013; Meuwissen et al., 2001; Pérez and de Los Campos, 2014), or emerging machine learning methods (Montesinos-López et al., 2021). Although RR-BLUP is simple to implement, it has, however, the disadvantage of applying the same penalization factor to all markers, a situation that may practically be inappropriate, as some markers are linked to quantitative trait loci, while others are not, a limitation that is overcome by the Bayes approaches (Viana et al., 2017). Consequently, Bayesian methods can be applied to miracle plant diversity panels. Candidate Bayes methods include Bayesian ridge regression, Bayes B2, Bayes C and Bayesian lasso, and the best method selected using the deviance information criterion (DIC) (Spiegelhalter et al., 2002) and the posterior probability ($PM_t|I$) to be calculated for each model as follows:

$$\mathrm{DIC} = \mathrm{D}(\overline{\boldsymbol{\theta}}) + 2\mathrm{p}_{\mathrm{D}} \qquad \textbf{Eqt. (7.1)}$$

$$\boldsymbol{p(M_t|l)} = \frac{\exp(-\Delta_t/2)}{\sum_{t=1}^{4}\exp(-\Delta_t/2)} \qquad \textbf{Eqt. (7.2)}$$

where $D(\bar{\theta})$ is a point estimate of the deviance obtained by substituting the parameters by their posterior means estimates in the likelihood function; p_D is the effective number of parameters in the model; and Δ_t is the DIC difference between model t and the model presenting the lowest DIC. The best Bayesian model will be the one with the highest posterior probability. The relative efficiency of the best single trait model will be compared to that of multi-trait genomics models. In particular, the double-stage Bayesian multi-output regressor stacking (BMORS) model (Montesinos-López et al., 2019) will be employed because it is computationally faster than the usual multi-trait multi-environment model. Based on the best model, the top 10 individuals in the candidate population with the highest estimated breeding values will be identified to constitute the output in the first GS cycle. Iteratively, this set of 10 accessions/genotypes will be involved in diallel-cross schemes whose F2 will be genotyped (100-150 individuals depending on available resources) and involved in a new cycle of genomic estimated breeding value determination to select a new batch of top ten individuals. The suggested optimized breeding plan for the species is synthesized in **Fig. 7.1**.

7.3 Conclusion

The results generated in this study represent a major contribution towards “deorphanizing” the miracle plant (*Synsepalum dulcificum*) in the Dahomey Gap and Upper Guinea in West Africa. A total of 23 key breeding traits were identified from farmers, consumers, and processing companies in the study area. From these, the top nine priority traits deserving urgent improvement include fruit size, fruit miraculin content, fruit yield, maturity time, tree size, shelf-life, edible ratio, seed part ratio and fruit colour sharpness. Exclusively for farmers in West Africa, a variety possessing the following key traits: large fruit size, early fruiting, high yielding and short tree size can be boldly envisaged. Likewise, a fresh, miraculin-dense and large fruit-sized variety would be the ideal variety for the region’s consumers. From the perspective of aligning with the major preferences of all three end-user groups, including processing companies, and for the sake of market competitiveness, an ideal variety for the West African sub-region would combine early maturity, high yield, large fruit size, high edible ratio and miraculin content. Given the overall low natural molecular variation in the species, developing such a variety would require generation of diversity through successive crossing schemes, considering the observed phenotypic variability and extent of heritability.

Fundamentally, both phenotypic and molecular variation concurred on the Volta population as an ideal primary source of elite parental lines in improving the species through conventional breeding. The species centre of origin was also detected in the Western Dahomey Gap, a region geographically intermediate between the Togo and Ghana borders that entails the actual Volta population. Three major gene pools separated by a physical barrier located between the Volta region and the eastern populations currently exist in the species in Dahomey and Upper Guinea: two gene pools in the Dahomey Gap and one in Upper Guinea. Two of these three major gene pools, including the Upper Guinea gene pool and the Central Dahomey Gap gene pool, can be readily involved in genomic selection schemes for accelerated genetic gain in the above detected key traits.

The use pattern analysis of the species documented 76 different uses with a predominance of local medicinal applications, targeting mainly roots and leaves, two organs involved in plant species fitness, and a higher use value for sociolinguistic groups in Benin. This raises questions on the future of this endangered crop, thus calling for the necessity to explore large-scale cultivation options coupled with the development of ex situ conservation measures. The modelling of the cultivation determinants suggests that sociolinguistic groups in Benin are currently more willing to invest in the miracle plant cultivation than their counterparts in Ghana. However, we demonstrated that the development of early fruiting varieties combined with raising awareness and the development of the market would foster an increased overall engagement of West African local communities in the cultivation of this crop. Therefore, the central governments of Benin, Togo, Ghana and Nigeria would be encouraged to consider promoting this unique berry crop as it holds the potential to serve as a top cash crop. Consequently, more funding should be availed to breeding research to accelerate the pace of elite variety development release in the species.

Through this thesis, the largest ever-known collection of *Synsepalum dulcificum* with more than 300 accessions from the species diversity hotspot has been constituted and is currently held at the Laboratory of Genetics, Biotechnology and Seed Science of the University of Abomey-Calavi. This represents a good asset for any future genome sequencing project to accelerate trait improvement in the species. There is, thus, a necessity to form a multidisciplinary regional breeding consortium around this crop, while permanently involving the various end-users' group, to leverage its potential contribution to the global population wellbeing.

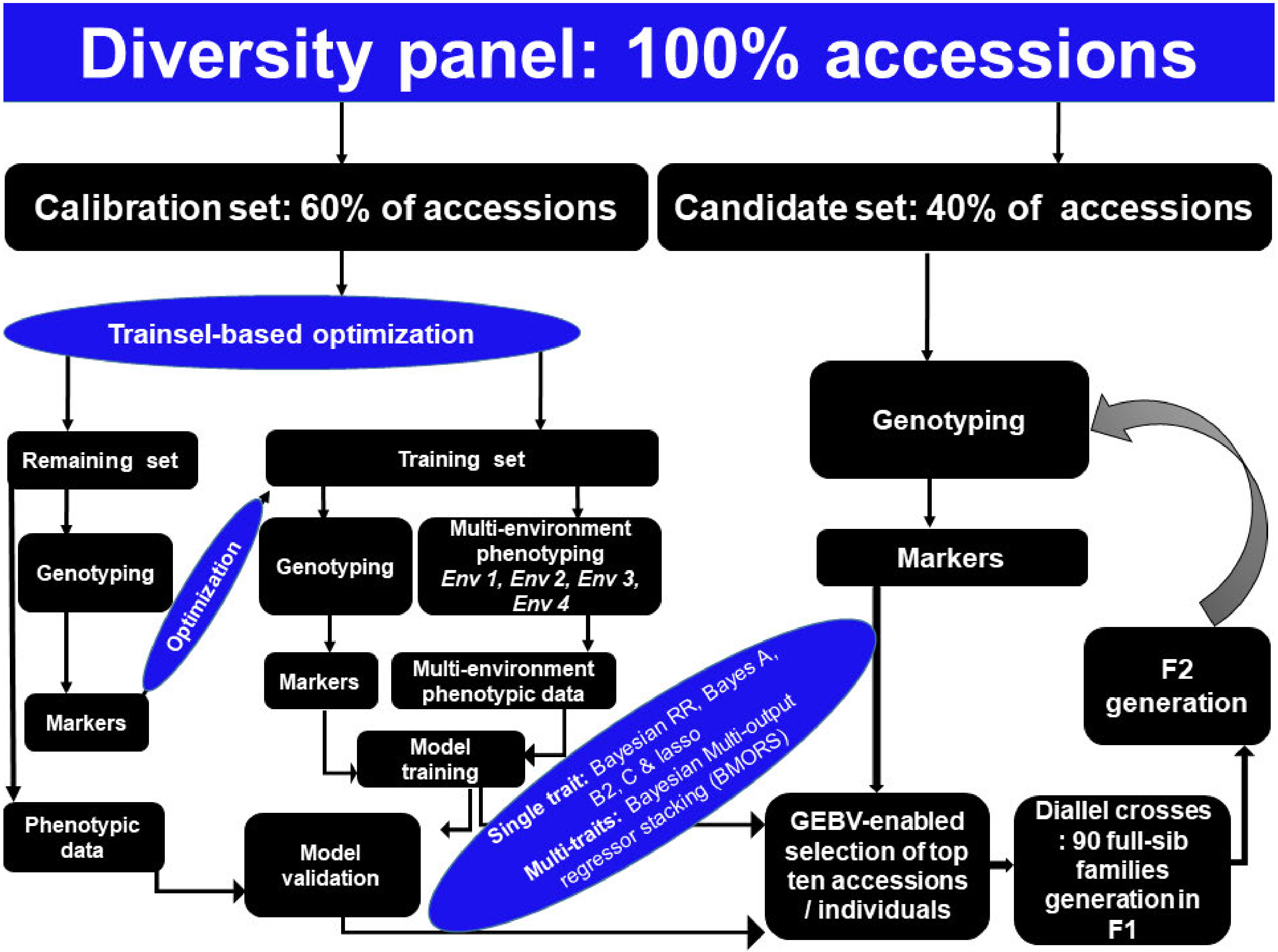


Figure 7.1. Optimized genomic selection scheme for the miracle *Synsepalum dulcificum*. The diversity panel sizes are 196 and 111, respectively for the Dahomey Gap and Upper Guinea gene pools.

7.4 References

Akdemir, D. (2017). Selection of training populations (and other subset selection problems) with an accelerated genetic algorithm (STPGA: An R-package for selection of training populations with a genetic algorithm). *arXiv preprint 1702.08088*.

Akdemir, D., Rio, S., and Sánchez, I. Y. (2021). TrainSel: an R package for selection of training populations. *Front. Genet.* **12**, 607.

Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., and Jannink, J.-L. (2011). Accuracy and training population design for genomic selection on quantitative traits in elite North American oats. *Plant Genome* **4**, 132.

Ballesta, P., Serra, N., Guerra, F. P., Hasbún, R., and Mora, F. (2018). Genomic prediction of growth and stem quality traits in *Eucalyptus globulus* Labill. at its southernmost distribution limit in Chile. *Forests* **9**, 779.

Chambers, A. H., Demesyeux, L., Moon, P., and Fu, Y. (2018). Optimizing of miracle fruit (*Synsepalum dulcificum*) seed germination and mutagenesis. *Afr. J. Food Sci. Technol.* **9**, 25-31.

Cirilli, M., Baccichet, I., Chiozzotto, R., Silvestri, C., Rossini, L., and Bassi, D. (2021). Genetic and phenotypic analyses reveal major quantitative loci associated to fruit size and shape traits in a non-flat peach collection (P. persica L. Batsch). *Hort. Res.* **8**, 1-17.

Colding, J., and Folke, C. (2001). Social taboos:“invisible” systems of local resource management and biological conservation. *Ecol. Appl.* **11**, 584-600.

Cruse-Sanders, J. M., and Hamrick, J. (2004). Genetic diversity in harvested and protected populations of wild American ginseng, *Panax quinquefoliu*s L.(Araliaceae). *Am. J. Bot.* **91**, 540-548.

De Los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K., and Cotes, J. M. (2009). Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics* **182**, 375-385.

Delêtre, M., McKey, D. B., and Hodkinson, T. R. (2011). Marriage exchanges, seed exchanges, and the dynamics of manioc diversity. *Proc. Natl. Acad. Sci. USA* **108**, 18249-18254.

Demesyeux, L., Brym, M., Valdes, D., Collazo, C., and Chambers, A. H. (2020). Yield and miraculin content of nine miracle fruit (*Synsepalum dulcificum*) morphotypes. *Euphytica* **216**, 1-12.

Fandohan, A. B., Gouwakinnou, G. N., Tovissode, C. F., Bonou, A., Djonlonkou, S. F. B., Houndelo, L. F., Sinsin, C. L. B., and Assogbadjo, A. E. (2017). Usages traditionnels et valeur économique de *Synsepalum dulcificum* au Sud-Bénin. *Bois For Trop* **332**, 17-30.

Gaoue, O. G., Lemes, M. R., Ticktin, T., Sinsin, B., and Eyog-Matig, O. (2014). Non-timber forest product harvest does not affect the genetic diversity of a tropical tree despite negative effects on population fitness. *Biotropica* **46**, 756-762.

Gianola, D. (2013). Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* **194**, 573-596.

Guo, Z., Tucker, D. M., Basten, C. J., Gandhi, H., Ersoz, E., Guo, B., Xu, Z., Wang, D., and Gay, G. (2014). The impact of population structure on genomic prediction in stratified populations. *Theor. Appl. Genet.* **127**, 749-762.

Hickey, J. M., Dreisigacker, S., Crossa, J., Hearne, S., Babu, R., Prasanna, B. M., Grondona, M., Zambelli, A., Windhausen, V. S., and Mathews, K. (2014). Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation.

Jiang, S., Cheng, Q., Yan, J., Fu, R., and Wang, X. (2020). Genome optimization for improvement of maize breeding. *Theor. Appl. Genet.* **133**, 1491-1502.

Kainer, D., Stone, E. A., Padovan, A., Foley, W. J., and Külheim, C. (2018). Accuracy of genomic prediction for foliar terpene traits in *Eucalyptus polybractea*. *G3* **8**, 2573-2583.

Kiptot, E., Franzel, S., Hebinck, P., and Richards, P. (2006). Sharing seed and knowledge: farmer to farmer dissemination of agroforestry technologies in western Kenya. *Agrof. Syst.* **68**, 167-179.

Larkin, D. L., Lozada, D. N., and Mason, R. E. (2019). Genomic selection—Considerations for successful implementation in wheat breeding programs. *Agronomy* **9**, 479.

Lenz, P. R., Nadeau, S., Mottet, M. J., Perron, M., Isabel, N., Beaulieu, J., and Bousquet, J. (2020). Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evol. Appl.* **13**, 76-94.

Li, Z., Gallardo, R. K., McCracken, V. A., Yue, C., Gasic, K., Reighard, G., and McFerson, J. R. (2020). U.S. Southeastern peach growers preferences for fruit size and external color versus resistance to brown rot disease. *HortTechnology* **30**, 576-584

Mazanec, R., Grayling, P., Doran, J., Spencer, B., and Turnbull, P. (2021). Genetic parameters and potential gains from breeding for biomass and cineole production in three-year-old *Eucalyptus polybractea* progeny trials. *Austral. For.* **84**, 13-24.

Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819-1829.

Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Cuevas, J., Montesinos-López, J. C., Gutiérrez, Z. S., Lillemo, M., Philomin, J., and Singh, R. (2019). A Bayesian genomic multi-output regressor stacking model for predicting multi-trait multi-environment plant breeding data. *G3* **9**, 3381-3393.

Montesinos-López, O. A., Montesinos-López, A., Pérez-Rodríguez, P., Barrón-López, J. A., Martini, J. W., Fajardo-Flores, S. B., Gaytan-Lugo, L. S., Santana-Mancilla, P. C., and Crossa, J. (2021). A review of deep learning applications for genomic selection. *BMC Genomics* **22**, 1-23.

Mooney, E. H. (2007). Genetic and evolutionary consequences of harvest in American ginseng, *Panax quinquefolius* L.(Araliaceae), West Virginia University, West Virginia, USA.

Mtshali, T. F., Mapholi, O. N., Ncube, K. T., Dzomba, E. F., Matelele, T. C., Chokoe, T., Mphahlele, T., Muchadeyi, F. C., and Hadebe, K. (2021). Goat farmers production objectives and trait preferences in the North West province of South Africa: An approach to identify selection criteria for community-based breeding program. *Int. J. Livestock Prod.* **12**, 64-75.

Olatoye, M. O., Clark, L. V., Labonte, N. R., Dong, H., Dwiyanti, M. S., Anzoua, K. G., Brummer, J. E., Ghimire, B. K., Dzyubenko, E., and Dzyubenko, N. (2020). Training population optimization for genomic selection in miscanthus. *G3* **10**, 2465-2476.

Ou, J.-H., and Liao, C.-T. (2019). Training set determination for genomic selection. *Theor. Appl. Genet.* **132**, 2781-2792.

Pérez, P., and de Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**, 483-495.

Resende Jr, M., Munoz, P., Acosta, J., Peter, G., Davis, J., Grattapaglia, D., Resende, M., and Kirst, M. (2012). Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytol.* **193**, 617-624.

Reyes-García, V., Molina, J. L., Calvet-Mir, L., Aceituno-Mata, L., Lastra, J. J., Ontillera, R., Parada, M., Pardo-de-Santayana, M., Rigat, M., and Vallès, J. (2013). "Tertius gaudens": germplasm exchange networks and agroecological knowledge among home gardeners in the Iberian Peninsula. *J. Ethnobiol. Ethnomed.* **9**, 1-11.

Rincent, R., Laloë, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., Rodriguez, V. M., Moreno-Gonzalez, J., Melchinger, A., and Bauer, E. (2012). Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics* **192**, 715-728.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Series B Stat Methodol.* **64**, 583-639.

Tiede, T., and Smith, K. P. (2018). Evaluation and retrospective optimization of genomic selection for yield and disease resistance in spring barley. *Mol. Breed.* **38**, 1-16.

Verges, V. L., and Van Sanford, D. A. (2020). Genomic selection at preliminary yield trial stage: Training population design to predict untested lines. *Agronomy* **10**, 60.

Viana, A. P., Silva, F. H. D. L., Gloria, L. S., Ribeiro, R. M., Krause, W., and Boechat, M. S. B. J. E. (2017). Implementing genomic selection in sour passion fruit population. . *Euphytica* **213**, 1-13.

White, F. (1983). The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa by F White. *Natural Resources Research Report XX, UNESCO, Paris, France*, 1876-1895.

Windhausen, V. S., Atlin, G. N., Hickey, J. M., Crossa, J., Jannink, J.-L., Sorrells, M. E., Raman, B., Cairns, J. E., Tarekegne, A., and Semagn, K. (2012). Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. *G3* **2**, 1427-1436.

Zhao, S., Yue, C., Luby, J., Gallardo, K., McCracken, V., McFerson, J., and Layne, D. R. (2017). US peach producer preference and willingness to pay for fruit attributes. *HortScience* **52**, 116-121.

Zhao, Y., Gowda, M., Liu, W., Würschum, T., Maurer, H. P., Longin, F. H., Ranc, N., and Reif, J. C. (2012). Accuracy of genomic selection in European maize elite breeding populations. *Theor. Appl. Genet.* **124**, 769-776.