Norwegian University of Life Sciences
Faculty of Biosciences
Department of Plant Sciences

Philosophiae Doctor (PhD)
Thesis 2023:50

Genetic Diversity, Selection Signatures and Nutritional Composition of Enset (*Ensete ventricosum*) from Ethiopia

Genetisk diversitet, seleksjonssignaturer og ernæringmessig sammensetning av enset (*Ensete ventricosum*) fra Etiopia

Alye Tefera Haile

Genetic Diversity, Selection Signatures and Nutritional Composition of Enset (*Ensete ventricosum*) from Ethiopia

Genetisk diversitet, seleksjonssignaturer og ernæringmessig sammensetning av enset (*Ensete ventricosum*) fra Etiopia

Philosophiae Doctor (PhD) Thesis

Alye Tefera Haile

Department of Plant Sciences
Faculty of Biosciences
Norwegian University of Life Sciences

Ås, 2023

Thesis number: 2023:50
ISSN: 1894-6402
ISBN: 978-82-575-2079-3

PhD supervisors:

Prof. Odd Arne Rognli (main supervisor)
Faculty of Biosciences, Department of Plant Sciences
Norwegian University of Life Sciences (NMBU),
NO-1432 Ås, Norway

Dr. Bizuayehu Tesfaye
College of agriculture,
School of plant and Horticultural science,
Hawassa University, Ethiopia

Prof. Trine Hvoslef-Eide
Faculty of Biosciences, Department of Plant Sciences
Norwegian University of Life Sciences (NMBU),
NO-1432 Ås, Norway

Dr. Mallikarjuna Rao Kovi
Faculty of Biosciences, Department of Plant Sciences
Norwegian University of Life Sciences (NMBU),
NO-1432 Ås, Norway

Dr. Sylvia Sagen Johnsen
Faculty of Biosciences, Department of Plant Sciences
Norwegian University of Life Sciences (NMBU),
NO-1432 Ås, Norway

Evaluation committees:

Prof. Sergio Ochatt
INRAe 17 rue Sully BP 86510,
F-21065 Dijon, France
E-mail: sergio.ochatt@inrae.fr

Ass. Prof. Jørn Henrik Sønstebø
University of South-Eastern Norway,
Gullbringvegen 36,
NO-3800 Bø, Norway
E-mail: jorn.h.sonstebo@usn.no

Committee coordinator:

Prof. Anne Kjersti Uhlen
Faculty of Biosciences, Department of Plant Sciences
Norwegian University of Life Sciences (NMBU),
NO-1432 Ås, Norway
E-mail: anne.uhlen@nmbu.no

Acknowledgements

I am very grateful to the Norwegian Agency for Development Cooperation (Norad) for funding my PhD which is part of the project "Controlling disease in sweet potato and enset in South Sudan and Ethiopia to improve productivity and livelihoods under changing climatic conditions using modern technologies" under the NORHED program (Agreement No. ETH-13/0017,2013). I would like to thank the School of Plant and Horticultural Sciences, College of Agriculture and Hawassa University for granting me a study leave. Thank you!

I would like to express my deepest appreciation to Prof. Odd Arne for being a solid point of contact throughout the project, having an open-door policy for both small questions and large discussions, and for his encouragement and keeping me calm when I felt overwhelmed. I have learned a lot from your in-depth knowledge and experience. I appreciate your prompt replies to all my queries. You believed in me and have guided me from the very beginning, you are a great teacher. I say thank you very much for all your contributions in my success. Thank you!

I would like to express my sincere gratitude to Prof. Trine Hvoslef-Eide, the main coordinator of the Enset-sweet potato NORHED project and my co- supervisor for giving me a chance to study in this project and exercise academic freedom. Thanks for the academic advice you have provided and for handling all administrative issues I always felt encouraged and energetic after our talks; you inspired me to move forward. More than anything else, thank you for the flexibility in arranging travels during the field work as well as during family visiting. I have seen your thoughtfulness and understanding particularly for mothers taking care of their children, you gave me freedom and allowing me to do my best. You are among the three power that were pushing me forward to finish this PhD. I say thank you very much for all your contributions in my success. Thank you!

I thank my co-supervisors for reviewing my manuscripts, thesis, and providing constructive comments Dr. Bizuayehu, Mallikarjuna Rao Kovi and Sylvia Sagen Johnsen for their generous input of ideas, guidance, and support. I would also like to give thanks to my local supervisor Dr. Bizuayehu, for the encouragement to pursue on the research idea at the beginning, your contribution to my work in facilitating transportation during my field work in Ethiopia and commenting my work is highly appreciated. I also Thank Dr. Andargachew, the local Norad project coordinator for facilitating transportation and other finance issue during my field work in Ethiopia. Thank you!

I would like to thank the Southern Nations, Nationalities and Peoples' Region (SNNPR) state in Ethiopia, Hawassa University, Agriculture office, and the Ethiopian Institute of Biodiversity jointly facilitated access to enset tissues. I also thank the administrative and technical staffs of Areka Agricultural Research Center, especially Dr. Zerihun, for providing me enset planting materials. I would like to thank the enset farmers, experts and development agents in Sidama, Gurage and South Omo who were cooperative and facilitated during data collection. You welcomed us to your homes, shared your food, gave us all the plant material and information. Thank you! Your contribution and hospitability is highly appreciated.

I would like to thank the Norwegian University of Life Sciences (NMBU) for all the support I got from the department of plant sciences, staff of CIGENE, LabTek staff, soil lab, fruit lab and the administrative staffs, the students information center (SiT) and the housing department (SiÅs). My country mates in Norway, made me in touch with my language and food during my study periods. I would also like to thank all my friends and colleagues at NMBU and Hawassa university for their help in one way or another. Thank you!

Last, but not least, I would not have been able to complete this without some amazing people. I am immeasurably grateful: my lovely kids, keab and Amha, for tolerance and patience, and for putting up with me being a part-time mama for such a very long time. you shouldn't have this stress at this age, but I had to decide for our better future. I love you both. Above all, I would like to share my success with my adorable children Keab and Amha who always make me to stay strong in all the difficulties and I am very much grateful with their unconditional love. I love you all. Thank you!

Most importantly, I give my heartfelt thanks to my husband Mekuye for taking care of our kids and for shouldering family responsibilities in my absence, under difficult circumstances, for his exceptional encouragement and enormous support during my study. You provided me with a steady supply of many countable and uncountable things, I found a great companion and partner in you. I love you. Thank you! I am also grateful to my sister and brother, Birtukan and Chuchu, all my entire family, my neighbours and friends for their exceptional encouragement, prayers, and invaluable support. Thank you!

Last, but the most significant, I am very gratefulness for my parents. My late mother (Oct 2011) and my late father (Nov 2020), you brought me into the world, initiated and invested in my educational journey. You never went to school, but you were respected and model hard worker farmers, traders, politicians, good entrepreneurs, and education advocates. I will never forget what you taught me; keeping domestic animals and protecting crops from monkeys and how to fight and win hardships at my early age. I thank you for investing in education of your children. You are still the reason why I am here today. You were and will be my biggest supporters and energy in my life to overcome any challenges. I have never seen anyone like you my father who appreciates technology, and I am able to experience everything because of you. I will always remember the things you have taught me and how much you loved and supported me. Now I know the reason why you were always telling me to be strong and go forward in life. Mam I will always remember your words, Problem isn't solving by talking. You were my number one supporter, all I wanted to do now is to say I love you; I miss both. GOODBYE Ababa and Emama rest in peace. Thank you for everything!

Glory be to God! እግዚአብሔር ይመስገን። Alye Tefera, March 2023

Compete with yourself and collaborate with others!

Abbreviations

AMOVA	Analysis of Molecular Variance
ANOVA	Analysis of Variance
AOAC	Association of Official Agricultural Chemists
AW_2	Wash Buffer AW2
bp	base pair
CO_2	Carbon dioxide
Ca	Calcium
CLUMPACK	Clustering Markov Packager Across K
CTAB	Cetyltrimethylammonium bromide
Cu	Copper
ddRAD	Double Digest Restriction-site Associated DNA
DNA	Deoxyribonucleic acid
df	degrees of freedom
dsDNA	double-stranded DNA
Fe	Iron
FW	Fresh Weight
F_{ST}	Fixation Index
GAE	Gallic Acid Equivalent
GC	Guanine, Cytosine
H_2O	Water
H_O	Observed Heterozygosity
iTOL	Interactive Tree Of Life
ISO	International Organization for Standardization
Mg	Magnesium
N	Nitrogen
NaCl	Sodium Chloride
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NDF	Neutral Detergent Fibre
Na	Sodium
P	Phosphorus
PC	Principal Components
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism

Dedication

Dedicated to My Mother Mulumebet W/Yonnes and My Father Tefera Haile

Table of Contents

List of papers

I. Alye Tefera Haile, Sylvia Sagen Johnsen, Mallikarjuna Rao Kovi, Trine Hvoslef-Eide, Bizuayehu Tesfaye, Odd Arne Rognli. **Comparison of different leaf preservation methods to obtain high quality DNA from enset (*Ensete ventricosum* (Welw.) Cheesman), a native and orphan food security crop in Ethiopia.** (Manuscript)

II. Alye Tefera Haile, Mallikarjuna Rao Kovi, Bizuayehu Tesfaye, Sylvia Sagen Johnsen, Trine Hvoslef-Eide, Odd Arne Rognli. **Genetic Diversity, Population Structure and Selection Signatures in Enset (*Ensete ventricosum*, (Welw.) Cheesman), an Underutilized and Key Food Security Crop in Ethiopia.** (Submitted to Genetic Resources and Crop Evolution)

III. Alye Tefera Haile, Mallikarjuna Rao Kovi, Sylvia Sagen Johnsen, Trine Hvoslef-Eide, Bizuayehu Tesfaye, Odd Arne Rognli. **Limited genetic diversity found in the Entada landrace (*Ensete ventricosum*, (Welw.) Chessman, var. Entada) from Ethiopia**. (Manuscript)

IV. Alye Haile Tefera, Trine Hvoslef-Eide, Sylvia Sagen Johnsen, Mallikarjuna Rao Kovi, Bizuayehu Tesfaye, Odd Arne Rognli. **Characterization of nutritional composition and bioactive compounds among different forms of enset (*Ensete ventricosum*, (Welw.) Chessman) from Ethiopia**. (Manuscript)

Summary

Most of the agrobiodiversity exists in the form of landraces and wild relatives in traditional smallholder-based farming systems. This diversity mainly exists among neglected and underutilized species, so called orphan crops. Moreover, climate change has a huge negative impact on agriculture resulting in uncertainty as regard agricultural productivity, particularly in food-insecure countries. Within this frame, this thesis focuses on enset (*Ensete ventricosum*), an important crop in Ethiopia which is a good example of a neglected and underutilized orphan crop. Three types of enset are present in Ethiopia, i.e., cultivated enset which is a large, perennial, and single-stemmed type propagated by induced suckers, the Entada landrace propagated by natural suckers, and wild enset propagated by seed. Enset is not well characterized at the genomic level and information on the variation in nutritional composition in diverse Ethiopian ensets is scarce. A number of cultivated and Entada enset genotypes, and some wild ensets were collected from farmers, natural habitats, and research fields in the southern regions Sidama, Gurage, and South Omo in Ethiopia. The aim was to study the genetic diversity, population structure, selection signatures, nutritional compositions, and content of bioactive compounds among and within all different form of enset.

Since most of the plant material used in these studies were collected from farms and natural habitats in remote areas, storage of plant tissue at ambient temperature for longer time was necessary. The first task was therefore to test leaf preservation methods and find a method that protected DNA from disintegration. Three methods for storage of DNA were compared, i.e., saturated NaCL-CTAB, silica gel and 96% ethanol. We found that a saturated modified NaCl-CTAB method was best for preserving the quality, integrity, and quantity of enset genomic DNA, and that the DNeasy plant mini kit approach was better for extracting genomic DNA from the preserved leaf samples than the CTAB method. High quality DNA could be extracted from samples that had been stored at ambient temperature for more than 35 days and more than 270 days at -20°C. These protocols were validated for leaf tissues of all Entada landraces, cultivated and wild enset, and used to obtain high-quality DNA for the molecular genetic analyses. Efficient, robust, and low-cost methods for preservation and storage of plant tissue sampled on collection trips to remote areas are very important and should receive further attention.

We assessed genetic diversity, population structure and selection signatures in 226 cultivated and 10 wild enset accessions (genotypes) collected from diverse enset growing regions in southern Ethiopia, using a total of 3505 SNPs markers obtained from ddRAD-sequences. Population structure and cluster analyses clearly distinguished between cultivated and wild enset. Moreover, higher levels of genetic variation were found within populations and regions (91.2 and 92.4%, respectively) than between populations and regions (8.8 and 7.6%, respectively). This result implies that the region of origin and environmental heterogeneity have little influence on the genetic variation among enset accessions. However, the genetic differentiation between regions was moderate to large (F_{ST} = 0.06–0.17). The genetic structure of enset was mainly shaped by eco-geographic factors, mode of propagation and cultivation status. Moreover, six genes potentially involved in sexual reproduction and flowering signalling, which are key processes underlying

domestication and adaptation, were under positive selection demonstrating that sexual reproduction plays an important role in shaping enset diversity.

The first study assessing the genetic diversity, selection signatures and relationship of Entada with cultivated and wild enset is presented in the thesis. A total of 117 Entada landraces were collected from three Entada growing regions in Ethiopia (Sidama, South and North Ari), and 1617 high-quality SNP markers, obtained from ddRAD-sequences, were used for the diversity studies. We found that Entada formed a completely separated clade from cultivated and wild enset. However, very little molecular variation was detected between regions (0.48%), nearly all variation was present within individuals (99.5%). Subpopulation differentiation between regions (Pairwise F_{ST}) was zero, and observed heterozygosity was very high (0.99), which is expected of plants with strict asexual propagation by suckers and limit the genetic diversity between genotypes. Prolonged clonal propagation of heterozygous genotypes from a single or few founding lineages will lead to populations with very little or no diversity between genotypes, and extremely high heterozygosity. Furthermore, eight candidate genes detected in this study to be under directional selection are involved in axillary shoot growth and they might be involved in controlling natural sucker formation in Entada. This study has generated a useful resource of information for further conservation efforts, genetic research, and breeding of Entada landraces.

Nutritional composition and bioactive compounds in the different forms of enset was compared for the first time. This was done in 14 common Enset cultivars (6 originating from Sidama and 8 from Gurage), two Entada and two wild enset genotypes, selected to represent contrasting molecular genetic diversity based on a phylogenetic analysis. We found significant nutritional variation between and within accessions and edible enset tissues leaf sheaths and corm. The highest amylose content was found in the leaf sheath of the cultivated "Mundraro" and in the corm of wild enset. The average content of amylose in the leaf sheath (45.4%) was higher than in the corm (26.6%), indicating that starch biosynthesis is more efficient in leaf sheaths than in the corm. There was no difference in amylose content between Enset cultivars of different origin. Among the cultivated ensets, the cultivar 'Kiticho' had the highest antioxidant capacity (8.4 μmol g^{-1} FW) and total phenolic content (TPC) (108 mg GAE 100 g^{-1} FW) in the leaf sheath. However, a wild genotype showed highest antioxidant capacity in leaf sheath (9.2 $\mu molg^{-1}$ FW) and corm (8.8 $\mu molg^{-1}$ FW). Entada landraces had the lowest neutral detergent fibre (NDF) and water-soluble carbohydrate (WSC) contents compared to both cultivated and wild enset, whereas wild ensets had highest content of NDF. Significant variation in mineral content was observed between corms and leaf sheaths except for Mg, P, and Na. Ca and Mg contents were significantly different among cultivated ensets. The content of the minor elements Zn and Cu was highest in the corm, while Fe was highest in the leaf sheath. These results are discussed in relation to the utilization of enset for food and feed. The variation among enset genotypes and edible enset tissues described here are important for selection of accessions for evaluation in multiple environments and for breeding purposes.

A lot of unexplored diversity is available for improving enset in Ethiopia, with patterns of diversity consistent with divergent selection on adaptive traits. Our results are useful for the conservation of genetic

resources, especially under global climate changes, and contribute to the potential discovery of functional genes and genetic mechanisms related to adaptability of enset to local climatic conditions, especially drought. This is encouraging for the potential of diversifying crops also in regions where enset is not traditionally grown, such as the food insecure dry north in Ethiopia.

Key words: *Ensete ventricosum*, Entada, ddRAD-SNPs, genetic diversity, nutritional content, bioactive compounds

Sammendrag

Det meste av agrobiodiversiteten eksisterer i form av landraser og ville slektninger i tradisjonelle småbruksbaserte dyrkingssystemer. Dette mangfoldet eksisterer hovedsakelig blant forsømte og underutnyttede arter, såkalte 'orphan crops'. Klimaendringene har dessuten en enorm negativ innvirkning på landbruket, noe som resulterer i usikkerhet når det gjelder jordbruksproduktivitet, spesielt i land med usikker matforsyning.

Innenfor denne rammen fokuserer denne oppgaven på enset (*Ensete ventricosum*), en viktig matplante i Etiopia og et godt eksempel på en forsømt og underutnyttet 'orphan crop'. Enset er ikke godt karakterisert på genomisk nivå, og informasjon om variasjon i ernæringsmessig sammensetning i forskjellige type enset fra Etiopia er mangelfull. En rekke dyrkede sorter/genotyper, samt genotyper av Entada og noen ville genotyper, ble samlet inn fra bønder, naturlige habitater og forskningsfelt i de sørlige regionene Sidama, Gurage og South Omo i Etiopia. Målet var å studere genetisk diversitet, populasjonsstruktur, seleksjonssignaturer, ernæringsmessige sammensetning og innhold av bioaktive stoffer blant og innen de ulike typer enset.

Siden det meste av plantematerialet som ble brukt i disse studiene ble samlet inn fra gårder og naturlige habitater i avsidesliggende områder, var lagring av plantevev ved romtemperatur over lengre tid nødvendig. Den første oppgaven var derfor å teste måter å konservere bladmaterialene på og finne en metode som beskyttet DNA mot nedbrytning. Tre metoder for lagring av DNA ble sammenlignet, dvs. mettet NaCL-CTAB løsning, silikagel og 96 % etanol. Vi fant at en mettet modifisert NaCl-CTAB løsning var best for å bevare kvaliteten, integriteten og kvantiteten av genomisk DNA fra enset, og at 'DNeasy Plant Mini Kit' metoden var bedre for å ekstrahere genomisk DNA fra de bevarte bladprøvene enn CTAB-metoden. DNA av høy kvalitet kunne ekstraheres fra prøver som hadde vært lagret ved romtemperatur i mer enn 35 dager og mer enn 270 dager ved -20°C. Disse protokollene ble validert for bladvev av alle Entada landraser, dyrket og vill enset, og brukt til å framskaffe DNA av høy kvalitet for de molekylærgenetiske analysene. Effektive, robuste og rimelige metoder for bevaring og lagring av plantevev som er samlet inn ved ekspedisjoner til fjerntliggende områder er svært viktig og bør få ytterligere oppmerksomhet.

Vi vurderte genetisk mangfold, populasjonsstruktur og seleksjonssignaturer i 226 kultiverte og 10 ville enset-genotyper samlet fra forskjellige regioner i Sør-Etiopia, ved å bruke totalt 3505 SNP-markører isolert fra ddRAD-sekvenser. Populasjonsstruktur og klusteranalyser skilte tydelig mellom dyrket og vill enset. Dessuten ble det funnet mye mer genetisk variasjon innen populasjoner og regioner (henholdsvis 91,2 og 92,4 %) enn mellom populasjoner og regioner (henholdsvis 8,8 og 7,6 %). Dette resultatet tyder på at regionen for opphavet til genotypene og miljømessig heterogenitet har liten innflytelse på den genetiske variasjonen blant ensetgenotypene. Imidlertid var den genetiske differensieringen mellom regioner moderat til stor (F_{ST} = 0,06–0,17). Den genetiske strukturen til enset blir hovedsakelig formet av økologiske og geografiske faktorer, formeringmåte og dyrkingsstatus. Seks gener som potensielt er involvert i seksuell reproduksjon og blomstringssignalering, nøkkelprosesser som ligger til grunn for domestisering og tilpasning, ble identifisert. Disse seks genene var under positiv seleksjon som viser at seksuell reproduksjon spiller en viktig rolle i å skape diversitet i enset.

Denne avhandlingen er den første studien som er utført av genetisk diversitet, seleksjonssignaturer og forholdet mellom landrasen 'Entada' og dyrket og vill enset. Totalt 117 Entada landraser ble samlet inn fra tre regioner i Etiopia (Sidama, Sør og Nord Ari), og 1617 høykvalitets SNP-markører, isolert fra ddRAD-

sekvenser, ble brukt til diversitetstudiene. Vi fant at Entada dannet en fullstendig adskilt grein fra dyrket og vill enset vha. en fylogenetisk analyse. Imidlertid ble det påvist svært lite molekylær variasjon mellom regioner (0,48 %), nesten all variasjon var tilstede i individer (99,5 %). Subpopulasjonsdifferensiering mellom regioner (pairwise F_{ST}) var null, og observert heterozygoti var svært høy (0,99), noe som forventes av planter med streng aseksuell forplantning med adventivskudd. Langvarig klonformering av heterozygote genotyper som stammer fra en enkelt eller få opprinnelige genotyper vil føre til populasjoner med svært lite eller ingen diversitet mellom genotyper, og ekstremt høy heterozygoti. Videre er åtte kandidatgener involvert i dannelse av adventivskudd oppdaget i denne studien, og disse kan være involvert i å kontrollere naturlig dannelse av adventivskudd i Entada. Denne studien har generert nyttig informasjon for videre bevaringsarbeid, genetisk forskning og foredling av Entada landraser.

Ernæringsmessig sammensetning og bioaktive stoffer i de forskjellige formene for enset ble sammenlignet for første gang i denne oppgaven. Dette ble gjort i 14 kjente 'sorter' av enset (6 med opprinnelse fra Sidama og 8 fra Gurage), to Entada og to ville ensetgenotyper, valgt for å representere kontrasterende molekylærgenetisk diversitet basert på den fylogenetiske analysen. Vi fant betydelig ernæringsmessig variasjon mellom og innen genotyper, og mellom de spiselige delene bladskjeder og rotknoller (corm). Det høyeste innholdet av amylose ble funnet i bladskjeder hos den dyrkede sorten 'Mundraro' og i knollen til en vill enset. Gjennomsnittlig innhold av amylose i bladskjeder (45,4 %) var høyere enn i rotknoller (26,6 %), noe som tyder på at stivelsesbiosyntesen er mer effektiv i bladskjeder enn i knoller. Det var ingen forskjell i amyloseinnhold mellom sorter av enset fra Sidama og Gurage. Blant de dyrkede ensetene hadde sorten 'Kiticho' den høyeste antioksidantkapasiteten (8,4 µmol g-1 FW) og totalt fenolinnhold (TPC) (108 mg GAE 100 g-1 FW) i bladskjeden. Imidlertid hadde en vill genotype høyest antioksidantkapasitet i bladskjeden (9,2 µmolg-1 FW) og rotknollen (8,8 µmolg-1 FW). Entada landraser hadde lavere NDF- og WSC-innhold enn både dyrket og vill enset, mens vill enset hadde høyest innhold av NDF. Signifikant variasjon i mineralinnhold ble observert mellom rotknoller og bladskjeder bortsett fra for Mg, P og Na. Ca- og Mg-innhold var signifikant forskjellig blant de dyrkede sortene. Innholdet av de sporelementene Zn og Cu var høyest i rotknollen, mens innholdet av Fe var høyest i bladsliren. Disse resultatene diskuteres i forhold til utnyttelse av enset til mat og fôr. Variasjonen mellom genotyper og vev som er beskrevet her er viktig for seleksjon av genotyper for utprøving i flere miljøer og for planteforedling.

Mye uutforsket diversitet er tilgjengelig for å forbedre enset i Etiopia, med diversitetsmønstre i samsvar med divergerende seleksjon på adaptive egenskaper. Resultatene våre er nyttige for bevaring av genetiske ressurser, spesielt under globale klimaendringer, og de kan bidra å oppdage funksjonelle gener og genetiske mekanismer knyttet til tilpasningsevnen til enset til lokale klimatiske forhold, spesielt tørke. Dette er oppmuntrende med tanke på potensialet for å utnytte diversiteten også i regioner der enset ikke tradisjonelt dyrkes, for eksempel det matusikre tørre nord i Etiopia.

Key words: *Ensete ventricosum*, Entada, ddRAD-SNPs, genetic diversity, nutritional content, bioactive compounds

1. General Introduction

Increasing global food production is one of the primary challenges to achieve food security to our world population (FAO, 2009). There is a need to produce more food and fibre with a smaller rural labour force to feed a growing population, more feedstocks for a potentially huge bioenergy market, contribute to overall development in the many agriculture-dependent developing countries, adopt more efficient and sustainable production methods and adapt to climate change. At the same time, we urgently need to cut greenhouse gas emissions from agricultural production and stop conversion of remaining forests to agricultural land. The world's population could grow to around 8.5 billion in 2030, 9.7 billion in 2050 and 10.4 billion in 2100. More than 50% of the predicted global population growth between now and 2050 is expected to happen in Africa (United Nations, 2022). There is a big gap between the amount of food we produce today, and the amount needed to feed everyone in 2050. Urbanization will continue at an accelerated pace, and about 70 percent of the world's population is expected to be urban in 2050 compared to 49 percent today (FAO, 2009). The international community committed to ending hunger and all forms of malnutrition worldwide by 2030. According to Naylor *et al*., (2004), food security can only be achieved if emphasis is given to the improvement of crops widely cultivated and used by the poor farmers and societies.

Many of the plant species that are cultivated for food are neglected and underutilized even though they play a crucial role in food security, nutrition, and income generation for rural societies. They are not recognized, if not wholly unknown, outside their origin and cultivation regions (Magbagbeola *et al*., 2010; Mayes *et al*., 2012). The reasons for this negligence vary with the producers and the agricultural experts (Dansi *et al*., 2012). In general, these crops perform better than major crops under extreme soil and climate conditions prevalent in the regions (Esfeld *et al*., 2013). To achieve food security, crop diversification is necessary. Currently, fewer than 150 crop species are commercially cultivated; 103 of them deliver up to 90% of the calories in the human diet, and of these, only four (rice, wheat, maize, and potato) provide 60% of the human energy supply (Bailey, 2016). Thus, tens of thousands of edible plant species are relatively underutilized and could be used to meet the food requirements of the growing world population (Chivenge *et al*., 2015). In addition, only producing enough food is not enough, but that food must be harvested, processed, distributed and the poor must be able to have the purchasing power to access that food resources. Therefore, any short-term improvement in food security will need to include modification (either transgenic or through conventional breeding) of neglected, underutilized, and other staple crops.

The underutilized root and tuber crops are hidden treasures of healthy nutritious food. Moreover, root and tuber crops are second in importance next to cereals as global sources of carbohydrates and play a crucial role in the human diet (Chandrasekara and Josheph Kumar, 2016). Tropical root and tuber crops are considered as the third most important crops after cereals and grain legumes (Archana *et al*., 2015). They provide a substantial part of the world's food supply and are also an important source of animal feed and

processed products for human consumption and industrial use. Root crops occupy nearly 50 million hectares of arable land and account for a global production of 560 million tonnes (Behera *et al*., 2009). Most of the world's food insecure people and subsistence farmers depend highly on root and tuber crops as their principal source of nutrition, medicine, shelter, feed, and cash income (Scott *et al*., 2000). These species produce large quantities of dietary energy and have stable yields under difficult environmental conditions and particularly during the periods of drought, famine and dry seasons (Campbe, 1987). For millions of people in the tropical humid regions of Africa, root and tuber crops occupy a position of prestige among the staple foods (Lebot, 2019). On average, root and tuber crops provide 20% of the daily per capita calorie intake for more than 640 million inhabitants of Sub-Saharan Africa, where, with the growing population, there is an increasing demand for these crops both for food and feed (Kenyon *et al*., 2006). Diversification of staple crops and the systems in which they grow is essential to make the future agriculture more sustainable, resilient, and suitable for local environments and soils (Massawe *et al*., 2016). Many root and tuber crops, except potatoes, sweet potatoes and cassava, are not yet fully explored for their genetic and nutritional diversity (Borrell *et al.,* 2019; Chandrasekara and Josheph Kumar, 2016). According to the world bank report, by the end of 2022, 685 million people could still be living in extreme poverty. Besides, the global extreme poor are concentrated in Sub-Saharan Africa (The World Bank, 2022). In addition, despite undeniable progress in reducing rates of undernourishment and improving levels of nutrition and health, nearly 800 million people are chronically hungry and over two billion do not have access to quality diets (Micha *et al*., 2020). Reducing poverty and hunger are recognized as the most important challenges globally since they are the first two of the seventeen UN Sustainability Development Goals (SDG's) (https://sdgs.un.org/)

In Ethiopia, the widely consumed root and tuber crops have been relatively neglected by research and conservation efforts (Tamiru *et al*., 2008). However, root and tuber crops are a major part of the traditional food system and income generation specifically in Southern, Western and Southwestern part of Ethiopia (Gezahegn *et al*., 2018). These crops are mainly used as food security during food shortage since they are drought-tolerant and high yielding (Wheatly *et al*., 1995). Several million people rely on root and tuber crops in Ethiopia. Among these crops, enset is one of the key food security crops in the country as a staple or co-staple food for 20 to 35 million people (Brandt *et al*., 1997; Tuffa, 2019). Besides, the enset-based farming system is a key agricultural system and farmers cultivate various enset landraces in different agroecological settings (Borrell *et al*., 2019; Tsegaye, 2002). Some studies showed that enset is the main crop of a sustainable indigenous African system that ensures food security (Brandt *et al*., 1997; Shigeta, 1992). Other minor root and tuber crops such as Anchote (*Coccinia abyssinica*), Taro (*Colocasia esculenta*), Yam (*Dioscorea alata*), Cassava (*Manihot esculenta*) and Welayta dinich (*Plectranthus edulis*) are also important in some parts of Ethiopia (Purseglove, 1985; Vavilov, 1951; Winters *et al*., 2006). Hence, an improved understanding of the production, utilization, and estimated future economic importance of these crops has potentially far-reaching implication for research and development areas at both the national and international levels.

1. 1 Origin, distribution, and botanical characteristics of enset

Enset (*Ensete ventricosum* (Welw.) Cheesman), is a large diploid (2n=2x=18) perennial monocarpic and single-stemmed herbaceous species belonging to the family Musaceae, along with banana and plantain, in the order Zingiberales and genus *Ensete* (Figure 1A-B) (Borrell *et al.,* 2019; Westphal and Stevels, 1975). Enset, often referred to as false banana, is together with banana and plantains the most known cultivated members of the plant family, with a high global and local economic and food security importance (Baker and Simmonds, 1962). The genome of *Ensete ventricosum* is approximately 547 Mb (Harrison *et al*., 2014), equivalent to the 523 Mb genome of the double-haploid (2n = 22) *Musa acuminata* genotype (D'hont *et al*., 2012). *Ensete ventricosum* is the only cultivated species of the genus *Ensete* and is cultivated exclusively for food, fibre, and animal fodder in smallholder farming systems in Ethiopia (Brandt *et al*., 1997; Guzzon and Müller, 2016).

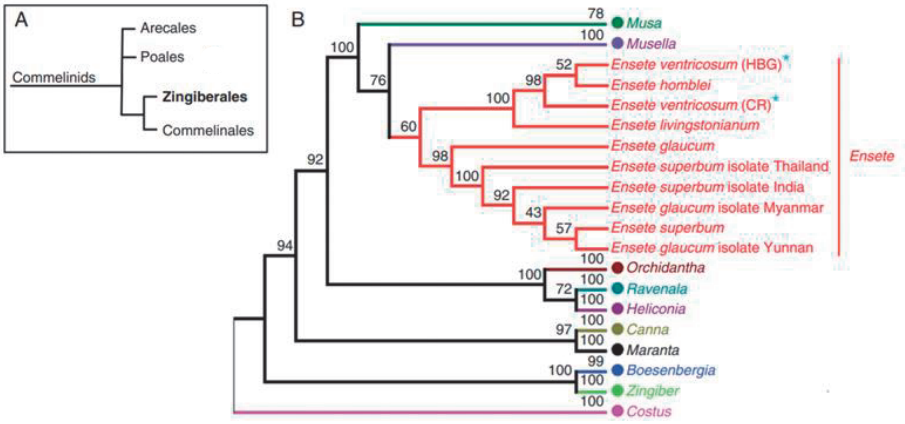


***Figure 1*:** Evolutionary relationships of the genus *Ensete* with sister groups. The genus *Ensete* is included in the Musaceae, one of eight families of the monocot order Zingiberales which together with the Commelinales is sister to the Poales (A). Evolutionary relationships of the genus *Ensete* within the Zingiberales based on ITS sequences, including collapsed sister genera within Musaceae and outgroups representing the eight families (B) (Borrell et al., 2019).

Currently, *Ensete* consists of three species with large distributions, *Ensete ventricosum* and *Ensete livingstonianum* in Africa (Figure 2A), and *Ensete glaucum* in Asia; in addition, there are five other localized endemics or near-endemic species (Borrell *et al.,* 2019). Wild species of *Ensete* are distributed throughout the central, eastern, and southern Africa as well as in Asia (Cheesman, 1947; Baker and Simmonds, 1953; Purseglove, 1985). In Africa, the wild form of *Ensete ventricosum* is widespread in tropical Africa from Ethiopia, through Kenya, Uganda and Tanzania to Mozambique and South Africa (Transvaal), and west to the Democratic Republic of Congo (Figure 2A) (Baker and Simmonds1953; Guzzon and Müller, 2016; Tsegaye, 2002). In Ethiopia wild enset is limited to the dense forests along riverbanks in the southern and

western part of the country. The wild form occurs at lower altitudes than the present area of enset cultivation in Ethiopia (Figure 2B) (Birmeta *et al*., 2004). Despite the wide distribution of wild enset, enset has been domesticated only in the Ethiopian highlands (Guzzon and Müller, 2016). Enset is mainly cultivated at the altitudes between 1500-3100 m above sea level (m a.s.l), but scattered plants can also be found at lower altitudes from 500-2400 m a.s.l (Birmeta *et al.,* 2004). However, it grows best at elevations between 1800 - 2450 m a.s.l. Further, recurrent droughts have led to the expansion of enset cultivation to other parts of the country (Ashango, 2017).

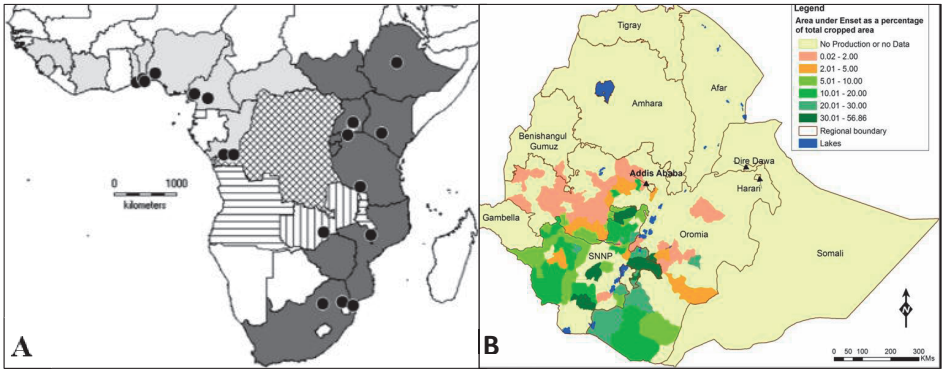


***Figure 2*:** Distribution in Africa (both cultivated and wild) (A). In light grey *Ensete livingstonianum*, in dark grey *Ensete ventricosum*, in horizontal bars countries in which *Ensete ventricosum* and *Ensete livingstonianum* co-occur (Angola and Malawi), in vertical bars the country (Zambia) in which *Ensete ventricosum* and *Ensete homblei* co-occur. In diagonal cross the country (Democratic Republic of Congo) in which all three species co-occur (Baker and Simmonds, 1953; Guzzon and Müller, 2016; Tsegaye, 2002). Regions of enset farming system in Ethiopia and area under cultivation as percentage of total cropped area (B) (Source: unpublished resource from Demeke Nigussie).

The stem of enset is thicker and larger than banana, often reaching up to 12 m height and more than 1.5 m in diameter (Birmeta *et al.,* 2004; Tsegaye, 2002). Unlike bananas, enset does not produce edible fruits, instead, it is grown for its carbohydrate-rich tissue obtained from the pseudostem, leaf sheaths and underground corm harvested 3-12 years after planting, depending on the local conditions and management efficiency (Borrell *et al.,* 2020; Brandt *et al*., 1997). Enset consists of an underground edible stem structure, known as the corm, which is 0.7–1.8 m long with a circumference of 1.5–2.5 m at maturity, a root system which is usually adventitious, a pseudostem formed from overlapping leaf sheaths that give a distinctly dilated shape at the base, and several broad leaves (Birmeta, *et al.,* 2004; Tsegaye and Struik, 2003).

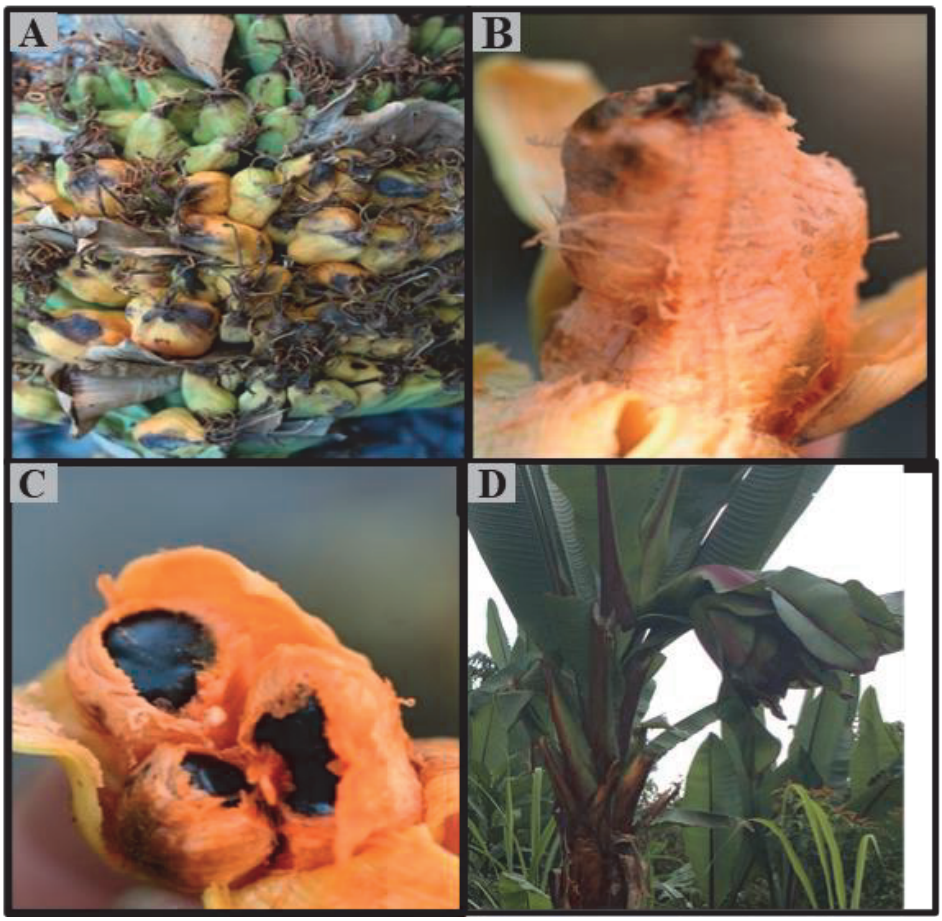


Figure 3: Physical form and the plant morphology of enset plant. Enset fruits on inflorescence (A), fruit (B), seeds inside fruits (C), mature flowering enset plant (D) (Photo taken by the author during sample collection) (Source figures A, B and C: http://www.bananas.org/f2/fruits-seeds-ensete-ventricosum-11923.html).

When the plant matures, an inflorescence is produced from the true stem that emerges through the leaf sheaths; and over time small fruits that contain large and very hard seeds develop in the inflorescence (Figure 3A-D).

1.2. Enset and its farming system in Ethiopia

Ethiopia is extremely dependent on agriculture as a predominant source of employment, income, and food security for most of its population (Tadesse *et al*., 2021). Ethiopian agricultural systems are naturally highly diverse (Borrell *et al.,* 2019). According to some authors, a total of 162 crop species are cultivated

throughout the highlands of southern, western, eastern, and central Ethiopia (Zemede and Ayele, 1995). Several ethno-linguistic groups cultivate enset as a main perennial plantation on farms, with other crop species complementing enset in the form of mixed subsistence farming systems for socio-cultural, economic, and environmental use-values. On average, enset farms grow and herd more than 10 different crops and livestock species (Sibhatu *et al*., 2015), with farmers growing many different enset landraces within one plantation (Zippel, 2005). Anthropologists, archaeologists, historians, and other scholars argue that enset was domesticated in Ethiopia as early as 10,000 years ago (Brandt *et al.,* 1997). During his travels in Africa to discover "the Source of the Nile", Bruce, (1790) collected plant species and stated that enset naturally produce and grow in every part of Abyssinia (the earliest name of Ethiopia). This indicated that enset is one of the earliest domesticated and cultivated crops in the country. Eighty percent of the enset production is concentrated in the South and Southwestern part of the country (Bezuneh *et al.,* 1967). Besides, the crop represents 65% of the total crop production in the southern regions of Ethiopia (Borrell *et al.,* 2019; Tuffa, 2019). The name "Tree Against Hunger", is common across different cultural groups cultivating enset and was first recorded in a European language by the traveling Portuguese priest Jeronimo Lobo in 1640 (Costa and Lockhart, 1984). It is an important food security crop sustaining the lives of many people (Borrell *et al*., 2019; Brandt *et al*., 1997), which was evident during the harsh famine in Ethiopia in the 1980's, where enset growing communities were not affected at all (Rahmato, 1995). Outside Ethiopia, enset is reported to have provided an emergency food in Vietnam during the Second World War. In parts of north and central Vietnam the growing point is used as a vegetable (Oyen and Lemmens, 2002). In addition, parts of *Ensete glaucum* are consumed in New Guinea, particularly the ripe fruits (Kennedy, 2009).

1.3. Food security and nutritional value of enset

Several agricultural crops are known to originate from Ethiopia (Purseglove, 1985; Vavilov, 1951; Winters *et al*., 2006). This includes coffee (*Coffea arabica*), tef (*Eragrostis tef*), safflower (*Carthamus tinctorius*), noug (*Guizotia abyssinica*), anchote (*Coccinia abyssinica* (Lam.) Cogn.), khat (Catha edulis Forsk), and enset (*Ensete ventricosum,* (Welw.) Chessman). Most research works, and subsequent improvement and conservation programs have so far focused mainly on cereals (Oli, 2006). Most of the root and tuber crops, including enset, did not get a fair share of attention by researchers and policy makers. However, enset is the major starch staple and an underexploited starch crop with significant potential in Ethiopia and beyond (Borrell *et al*., 2019). It is also stated that Ethiopia is well known for its diversity of indigenous food plants, of which 27% are cultivated vegetables by traditional farmers in home gardens, and about 29% non-cultivated traditional vegetables (Asfaw, 1997). However, chronic food insecurity in Ethiopia is estimated to be about 10% and this figure rises to more than 15% during frequent drought years causing acute food insecurity (Endalew *et al*., 2015). Enset is highly drought tolerant with a broad distribution in different agro-ecological zones and a lifesaving crop for both humans and livestock (Shigeta, 1990). Besides, enset can be harvested at any time of the year and at any stage over several years (including when it is immature), and enset-derived starch can also be stored for long periods (Birmeta *et al*., 2004). According to the

2019/20 main crop season survey, 157 mill enset plant were harvested. This gives a total yield of about 9.4 mill tonnes per year, making it one of the largest perennial food crops in the country (Central Statistical Agency, CSA, Government of Ethiopia, 2020). Also compared to all other crops in Ethiopia, enset ranks the second most produced crop species, with the fourth highest yield per hectare (Borrell *et al*., 2020). However, yield and ratio of different enset food products vary and depends on cultivated accessions (Borrell *et al*., 2020). The usable parts of the enset are the pseudostem (leaf sheaths) (Figure 4A), the corm (underground part) (Figure 4A) and leaf petioles, but the pseudostem and corm are the main edible parts from the enset plant (Tsegaye, 2002).

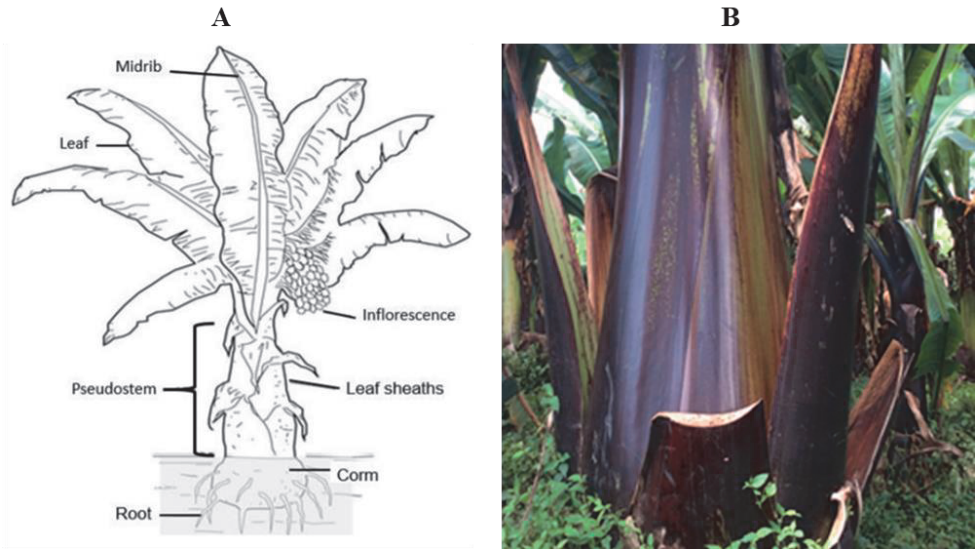


Figure 4: A: Mature enset at flowering stage (adapted from (Brandt *et al*., 1997); B: Enset plant ready for harvest (Photo taken by the author during sample collection).

The major processed foods from enset are Kocho (obtained through fermentation of decorticated leaf sheaths and corm), Bulla (a white powder produced by drying squeezed sap from scraped leaf sheaths and grated corms), and Amicho (boiled enset corm, usually from a younger plant) (Figure 4A-B) (Brandt *et al*., 1997; Yemataw *et al*., 2014). The nutritional value of enset products is comparable to other starchy products from species such as sweet potato, taro and yam (Tsegaye, 2002). The pseudostem is rich in soluble carbohydrates (80%) and starch (65%) but has low protein content (4%) (Mohammed *et al.,* 2013). However, the fat and carbohydrate contents of enset products are better than in sweet potato and yam (Jacobsen *et al.,* 2018; Tsegaye, 2002). The corm has the highest concentration of most soluble carbohydrates and starch, and least of protein, fibre, cellulose, and sugar (Mohammed *et al*., 2013). Chemical composition and the relative concentrations of certain macro and trace elements show a considerable variation in enset (Nurfeta *et al*., 2008). The enset food contains more calcium and iron than

most cereals, tubers, and root crops, and some enset landraces are believed to have medicinal value, which is utilized by the enset growing communities (Daba and Shigeta, 2016). Moreover, several bioactive compounds such as antioxidants, antiobesity, immunomodulatory activities, hypocholesterolemic, antimicrobial and antidiabetic, are present in tubers and root crops (Chandrasekara and Josheph Kumar, 2016). These bioactive compounds are important to reduce heart strokes, cancers, chronic respiratory diseases, and diabetes. The bioactive compounds such as antioxidants and total phenolic compounds of different types of enset have seldom been described (Forsido et al 2013; Desssalegn 2019). The chemical composition of food crops can vary considerably between regions within a country and among countries (Greenfield and Southgate, 2003). Similarly, the quality of enset products depends on the accessions type, environmental conditions, age of the plant and method of processing (Karssa and Papini, 2018). Characterization of the nutritional quality of enset accessions is very important for the utilization of the enset biodiversity in Ethiopia.

1.4. Genetic resources and molecular marker of enset

Ethiopia is the primary centre of origin and centre of diversity of enset (Purseglove, 1985; Vavilov, 1951). Enset farms in Ethiopia are rich in enset diversity, since farmers grow various types of enset used for specific purposes on the same farm (Shumbulo *et al*., 2012). Previous studies have characterised phenotypic variation among accessions and regions for traits such as maturity time, plant height, pseudostem height, circumference, leaf number and end-product yield in 378 accessions of enset (Yemataw *et al*., 2017). Moreover, a large part of enset diversity still exists in farmers' field (on-farm), unrepresented in *ex situ* collections (Haile, 2014; Yemataw *et al*., 2017). Thus far, very few studies were conducted to characterize enset genetic resources, both in farmers' field and germplasm collection centres.

In a biological comparison of wild and cultivated enset, the wild enset propagates naturally by seed under natural condition, however, cultivated enset in farmer fields are only propagated vegetatively by the local farmers (Borrell *et al*., 2019; Brandt *et al*., 1997; Shigeta, 1992). Based on the domestication status and modes of propagation, the three horticultural forms of enset are cultivated enset (Figure 5 A-F), Entada landraces (*Ensete ventricosum* (Welw.) Cheesman var. Entada) propagated from spontaneous suckers (Figure 5 G-H), and wild enset (Figure 5 I-J) (Bekele and Shigeta, 2011; Olango *et al*., 2015). However, wild and a few cultivated genotypes are produced from seeds, and the species appears to have an outcrossing reproductive system (Birmeta *et al*., 2004; Brandt *et al*., 1997). Enset is harvested before flowering (unlike banana) to avoid reallocation of resources from edible storage organs to the inedible inflorescences, thus seeds are not available for establishing new plants (Borrell *et al*., 2020). However, vegetative propagation is preferred due to increased vigour of suckers (Alemu and Sandford, 1991). Besides, domesticated enset seeds also have low germination ability and are highly variable (Diro *et al*., 2003; Negash, 2001; Tesfaye, 1992). Enset propagation by inducing suckers is a cultivation practice carried out every year, from late December to early February, using a mixture of accessions (Tsegaye, 2002; Zippel, 2005).

Another cultivated enset landrace know as Entada (*Ensete ventricosum* (Welw.) Cheesman var. Entada) is a single domesticated enset landrace mainly from the South Omo region in the country (Bekele and Shigeta, 2011). Entada is a naturally suckering cultigen of enset and resembles *Musa* species as regards the suckering trait (Figure 5 G-H) (Shigeta, 1990). Entada is propagated by spontaneous suckers like banana, and the local name is Intada (Bekele and Shigeta, 2011). The name indicates that the plant grows by itself or multiply by itself (Shigeta, 1992). Entada has lost the apical dominance, which is present in enset but has kept the ability to flower and set fruit (Shigeta, 1992). Generally, Entada is different from all the other cultivated and wild enset both morphologically and genetically, which do not produce and propagate with suckers (Figure 5 G-H) (Olango *et al*., 2015; Shigeta, 1992). Phenotypically Entada is short in stature, produce numerous suckers naturally, have short leaves, and the nature of the leaves are plastic. However, evidence indicates that the Entada landrace is one of the most genetically unstudied food crops in Ethiopia (Shigeta, 1992). A previous molecular diversity study of enset using SSR and SNP markers revealed that Entada belongs to the genus *Ensete* of the Musaceae family (Olango *et al*., 2015; Yemataw *et al*., 2018), but so far, no research has been conducted to investigate and document the phenotypic and genetic variation among and within Entada landraces and the relationship of Entada with other cultivated and wild enset genotypes.

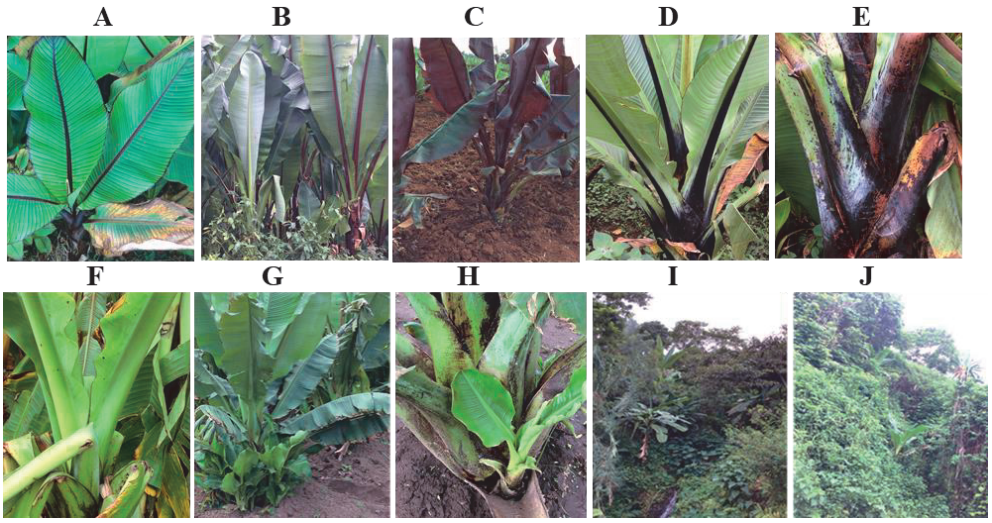


Figure 5: Phenotypic variation among sequenced enset accession from cultivated enset (A-F), Entada landraces (G-H) and wild enset (I-J) (Photos taken by the author during sample collection).

Landraces of Entada are being maintained and used by mainly the Ari people in Ethiopia (Shigeta, 1992). The major processed food from Entada is Amicho prepared from the underground corm (the underground base of the stem that serves as a storage organ). The fresh corm is cooked like potatoes and yam. It is very

important to sustain the cultivation of Entada since it is not only a food plant but also a multi-purpose crop with cultural values and an animal feed source (Shigeta, 1992).

Understanding genetic variation and geographic origins of germplasm in line with their cultivation history and ecological adaptation are essential for conservation and breeding (Solomon *et al.,* 2019). Germplasm diversity is vital to successful breeding programs. Such diversity is important for broadening the genetic base, as it increases the probability of finding more unique genes for which two parents have diverse alleles (Solomon *et al*., 2019). Conservation of genetic diversity is an important prerequisite for developing new cultivars with desirable agronomic characters. While many germplasm collections have been established worldwide, many of them face major difficulties due to large size and lack of adequate information about population structure and genetic diversity (Lee *et al*., 2016). Similarly, enset accessions have traditionally been characterized based on phenotypes, however, phenotypic descriptions are limited by the cost, time, and space required to make visual observations and measurements (Heuzé *et al*., 2017). Moreover, the vegetative propagation nature and long perennial life cycle of enset have made the programs laborious, time-consuming, and costly (Bezuneh, 2010). Although assessment of morphological variation present in enset is feasible, its use is rather limited due to the small number of phenotypic markers and the fact that they are influenced by many environmental factors (Negash *et al*., 2002).

In Ethiopia, farmers give vernacular names for each enset accession and distinguish between them based on the main observable physical or biochemical characteristics. Further farmers distinguish based on the output product and fibre content (Shumbulo *et al*., 2012). However, language differences among the enset growing communities resulted in homonyms, synonyms, and multiple duplications of landraces in conservation centres and made them overfilled and bulky to manage effectively (Bezuneh, 2010; Negash *et al*., 2002). This is beginning to change as the importance of this crop becomes better understood. Studies using molecular markers such as Amplified Fragment Length Polymorphism (AFLP) (Negash *et al*., 2002), Random Amplified Polymorphic DNA (RAPD) (Birmeta *et al*., 2004), Inter Simple Sequence Repeats (ISSR) (Tobiaw and Bekele, 2011), chloroplast DNA sequences (Bekele and Shigeta, 2011), Simple-Sequence Repeats (SSRs) (Gerura *et al*., 2019; Getachew *et al*., 2014; Olango *et al*., 2015), and Single Nucleotide Polymorphism (SNP) (Tesfamicael *et al*., 2020) have revealed that there are genetic diversity among wild and cultivated enset accessions. However, these studies were conducted in particular locations of enset growing areas in southern Ethiopia. Since Ethiopia is the centre of diversity there still are many enset rich locations that are thought to harbour huge diversity of cultivated and wild enset not yet studied and unrepresented in *ex situ* collections (Olango *et al*., 2015; Yemataw *et al*., 2017). Numerous landraces and elite cultivars adapted to diverse local habitats in Ethiopia have been developed via thousands of years of natural and artificial selection, domestication, and evolutionary processes. Therefore, it is very important to employ new molecular techniques and exploit the hitherto unstudied enset growing regions to discover and characterize new sources of variation and study the relationships forms of enset as well as relatedness with the already characterized enset diversity in other parts of Ethiopia.

Several modern molecular techniques are now being applied together with phenotypic descriptions to describe genetic diversity and relatedness among enset accessions (Birmeta *et al*., 2004; Tesfamicael *et al*., 2020). Single Nucleotide Polymorphisms (SNPs) and Simple Sequence Repeats (SSRs) are the most common DNA markers for genetic diversity studies (Tsykun *et al*., 2017). Among all DNA markers, SNPs are the most abundant and robust markers. They are feasible for automated high-throughput genotyping processing of large numbers of samples and available for multiple assay options using different technology platforms to meet the demand for genetic studies and molecular breeding of crop plants (Alkan and Eichler, 2011; Bus *et al.,* 2012; Hamilton *et al.,* 2011). The recently developed ddRAD-seq technique is used extensively for population genetic studies in a wide range of non-model organisms (Andrews *et al*., 2016; Peterson *et al*., 2012). The ddRAD sequencing provides a useful tool for marker-assisted breeding, genotype identification and determination of genomic organization and evolution in plants. It is a powerful and relatively cost-effective approach for developing numerous SNP markers and constructing high-density genetic maps (Peterson *et al*., 2012). SNPs are powerful tools to resolve the differences among extremely similar individuals and increase the accuracy of diversity estimates (Hinze *et al.,* 2017). Currently, enset is one of the indigenous orphan crops for which there is high protection and development concern in Ethiopia (Olango *et al*., 2015). However, lack of knowledge about the detail genetic diversity of this plant species complicates conservation, improvement, and utilization efforts by farmers, conservationists, and breeders (Negash *et al*., 2002). Despite its current importance, abundant diversity and huge potential, enset is probably the least studied food security crop in Africa, even if it is the second most widely grown crop in Ethiopia and serves as a key staple food crop for about 20 to 35% of the Ethiopian population. Detailed diversity studies of both cultivated and wild enset, and the Entada landrace in Ethiopia are needed (Borrell *et al.,* 2019; Tesfamicael *et al*., 2020). Furthermore, research in fields such as agronomy, breeding, pathology, and conservation is very much needed to optimize and maximize the benefits from this multi-purpose crop that is so important for Ethiopians (Borrell *et al*., 2019). Therefore, investigations of genetic diversity, population structure, potential selection signatures of different enset types, and the nutritional diversity are essential for conservation and management of germplasm resources, including identification of duplicate accessions and breeding efforts.

1.5. Research objectives and framework

This project was initiated with the general objective of addressing the magnitude and structure of genetic diversity, and chemical and nutritional composition of enset in Ethiopia for the benefits of future collection, conservation, and breeding programs. The last chapter provides concluding remarks and outlines future research needs. Since the thesis was developed as individual manuscripts already submitted or ready to be submitted to peer-reviewed journals, repetitions of introductory information in each of the chapters occurs, allowing for an independent reading.

The specific objectives of this PhD thesis were to:

1. Find optimal methods for preservation of leaf tissues of enset, collected during field expeditions and necessary to store at ambient temperature for various time spans, and DNA extraction methods to ensure integrity, quantity, and quality of genomic DNA.
2. Evaluate the efficacy and suitability of Single Nucleotide Polymorphism (SNP) markers developed from ddRAD (double-digest Restriction-Site-Associated DNA) sequencing for high-throughput genotyping and estimation of genetic diversity among and within cultivated and wild enset accessions collected from three regions in Southern Ethiopia.
3. Study the genetic diversity among accessions of the Entada landrace, relationships between Entada, cultivated and wild enset, and identify candidate genes involved in sucker formation.
4. Characterize the nutritional composition and content of bioactive compounds in plant parts used for food and feed among the most common enset clones used by farmers and compare it with different forms of enset.

2. Summary of materials and methods

2.1 Description of sampling area and plant material

In this PhD project we used enset plant material both for the genetic diversity studies and the nutritional and bioactive compounds analyses. Enset accessions were collected from three main enset culture communities, which are the densely populated enset cultivating administrative regions Sidama, Gurage and South Omo (Paper I, II and III). We collected 226 cultivated and 10 wild enset accessions originating from different geographical locations and agro-ecological zones (Paper II). The wild ensets were collected around farms, along riversides and in deep forests only in the South Omo region. Furthermore, 129 Entada landraces were collected from Sidama, South and North Ari regions (Paper III). Eighteen enset accessions (14 cultivated and 2 wild ensets, and 2 Entada landraces), were collected from the fields at Areka Agricultural Research Centre (AARC) and Hawassa university (HU) Research Center for characterizing nutritional composition and bioactive compounds of enset (Paper IV). Unfurled young, healthy, and fresh leaves were collected and used for genetic diversity studies, and five years old enset plant tissues i.e., leaf sheath and corm, for nutritional composition analyses.

2.2. Methods

2.2.1. Preservation of collected leaf materials and extraction of genomic DNA

We tested three leaf preservation (NaCl-CTAB, silica gel and ethanol) and two DNA extraction (DNeasy Plant Mini Kit and CTAB) methods for the preservation of plant materials over various time periods up to five weeks, and for obtaining high quality and quantity of genomic DNA (Paper I). We found that the saturated NaCl-CTAB solution for leaf preservation and the DNeasy Plant Mini Kit (Qiagen) for genomic DNA extraction gave the best results and these methods were selected for preparation of genomic DNA

used in the studies reported in paper II and III. Unfurled young, healthy, and fresh leaves were collected and maintained in saturated NaCl-CTAB solution with minor modification (Rogstad, 1992) to preserve genomic DNA from degradation during transportation from the field in Ethiopia to the laboratory at the Norwegian University of Life Sciences (NMBU) at Ås in Norway. The concentration was measured with a Thermo Scientific NanoDrop™ (ND). DNA quantification was performed by agarose gel electrophoresis (1%), Qubit® dsDNA BR assay kit (Q) and Quant-iT™ PicoGreen™ dsDNA assay (Paper I, II and III).

2.2.2. Double-digest Restriction-Site-Associated DNA (ddRAD) library preparation and Illumina sequencing

The ddRAD procedure used in this study was modified from Peterson *et al. (*2012*)* (paper II and III). We calculated the number of reads required for 20X coverage of restriction fragments in the 150–500 bp size range across 10 multiplexed individuals using multiple enzyme pairs, assuming 0.44 GC content, to ensure that restriction fragments could feasibly be sequenced with enough coverage on an Illumina MiSeq platform. The libraries were analyzed using an Agilent 2100 Bioanalyzer and diluted to a concentration of 35 nM for paired-end sequencing using the V2 sequencing kit on the MiSeq platform (Illumina). The sequencing was performed at NMBU.

2.2.3. SNP calling and population genetic analyses

The ddRAD sequence data obtained was quality checked using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). High quality reads were retained after trimming the bad quality reads using the Trimmomatic program (Bolger *et al.,* 2014). The SNPs were filtered using STACKS 2 (Rochette *et al*., 2019). The SNP data were used to estimate population genetic parameters and population structure of cultivated and wild enset (Paper II) and Entada landraces (Paper III). We estimated the population genetic structure using the software fastSTRUCTURE (Raj *et al*., 2014), cluster analyses including maximum-likelihood (ML) phylogenetic tree analyses were performed using PhyML 3.0 (Guindon *et al.,* 2010), and PCA analyses using TASSEL v5.2 (Bradbury *et al*., 2007). The trees were prepared and visualized using the iTOL v4 online tool (Letunic and Bork, 2019). The patterns of genetic variation and population differentiation, including heterozygosity and pairwise genetic differentiation (F_{ST}) (Weir and Cockerham, 1984), were estimated based on SNP genotypes using Arlequin v.3.5 (Excoffier and Lischer, 2010). To detect loci under directional selection, we used the hierarchical method (Excoffier *et al*., 2009), a modified approach of Beaumont and Nichols (1996) (Beaumont and Nichols, 1996), implemented in the Arlequin software package version 3.5.1.3 (Excoffier and Lischer, 2010). The putative function of genes with outlier SNPs was identified using the Gene Ontology (GO) annotation using Blast2GO software tool version 3.0 (Conesa *et al*., 2005).

2.2.4. Characterization and analyses of nutritional composition and bioactive compounds

All analyses were based on dry weight (DW), while the total phenol content and antioxidant capacity were measured based on fresh weight (FW). Samples were freeze dried and ground to pass through a 1 mm

screen (coffee grinder in stainless steel). The Megazyme K-Amyl amylose/amylopectin analysis kit was used to quantify the amylose (AM) content following the manufacturer's protocol (Megazyme, Wicklow, Ireland). The amylose content (%) was calculated as (Abs510 Amyl/Abs510 total)*66,8. The conversion factor originating from the dilutions during the protocol steps was 66.8. For determination of WSC, samples were extracted in 0.05 MNa-acetate buffer at room temperature for 18 h and filtered through paper (Randby *et al*., 2010). The carbon (C) and nitrogen (N) was determined based on the classical Pregl-Dumas method (https://www.nmbu.no/download/file/fid/48964) and the amount of aNDF (NDF on organic matter basis) was determined as described in Mertens (2002) using and Ankom200 Fiber Analyzer (Ankom Technology). The major and minor minerals samples were decomposed with ultrapure concentrated HNO3 (nitric acid) prepared by dissolving 0.25 g of sample with 5 mL of HNO3 in acid-washed Teflon tubes at 260 °C in a Milestone Ultraclave (260 °C for 20 minutes). The samples were diluted by adding 50 mL deionized water and analyzed using ICP-OES (Inductively coupled plasma atomic emission spectroscopy) (Agilent 5110 ICP-OES) and ICP-MS (Inductively coupled plasma mass spectrometry) (Agilent 8800 ICP-MS). Phosphorus (P), potassium (K), sulfur (S), calcium (Ca), and magnesium (Mg) were quantified using the ICP-OES. Sodium (Na), iron (Fe), copper (Cu) and zinc (Zn) were quantified using the ICP-MS. Dry matter content was determined gravimetrically based on sample weight loss after being heated in an oven at 120 °C for 48 hours. TPC was determined using the Folin-Ciocalteu reagent based on the procedure (Singleton *et al*., 1999). Three grams (g) of fresh homogenate sample was weighed and extracted using 30 mL methanol. The liquid sample was centrifuged, and the supernatant was mixed with Folin Ciocalteu reagent and 7.5% (w/v) sodium carbonate, sample was incubated for 15 minutes and measured at 765 nm wavelength. The results are expressed as mg gallic acid equivalent (GAE) per g fresh weight (FW). The antioxidant capacity (AO) was estimated by the method of (Benzie and Strain, 1999). Three grams of fresh homogenate sample was weighed, extracted in 30 mL methanol, centrifuged and the supernatant was mixed with acetate buffer, TPTZ (2,4,6-tri-pyridyl-s-triazine) and iron trichloride, incubated for 10 minutes and the absorbance was measured at 595 nm.

Analysis of variance was performed using the R software (version 3.6.2) and Proc GLM in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) to test differences in contents of various nutritional composition and bioactive compounds among and within different forms of enset. Multiple means comparisons were tested using the Tukey test, and differences between considered statistically significant at $P \leq 0.05$. The correlation coefficients were estimated using Pearson correlation (r).

3. Main findings and discussions

3.1. Appropriate methods for preserving and extracting DNA from enset leaves

Many leaves preservation and DNA extraction methods have been developed, most of them tailored for isolation of DNA from fresh leaf samples (Bressan *et al*., 2014; Rogstad, 1992). The first objective (Paper I) was to find a method that preserved the leaf materials at ambient temperature over various time periods

up to five weeks. The saturated NaCl-CTAB solution was found to preserve leaves better than silica gel or 96% ethanol. Besides, visible contamination was not observed on the enset leaves and this is most likely due to the bactericidal and detergent properties of CTAB (Rogstad, 1992; Thomson, 2002). Storage in ethanol did not preserve the DNA of enset leaves. DNA of samples preserved in silica gel were highly degraded. Most likely the preservation efficiency of silica gel is affected by the enset leaf secondary compounds and/or tissue characteristics (Abdel-Latif and Osman, 2017; Bressan *et al*., 2014). Moreover, the DNeasy Plant Mini Kit (Qiagen) method was found to be more efficient in extracting DNA from the preserved samples than the standard CTAB method; being faster and producing genomic DNA of higher quality (Kotchoni and Gachomo, 2009; Margam *et al*., 2010). Most likely the CTAB method is better for young and fresh leaves (Guo *et al*., 2018). The preservation protocol was validated for leaf tissues of all Entada landraces, cultivated and wild ensets. Therefore, the saturated NaCl-CTAB solution and the DNeasy Plant Mini Kit were used in this study to obtain high quality genomic DNA for the experiments described in Paper II and III.

3.2. Genetic diversity, population structure, and selection signatures in enset (Paper II)

Single nucleotide polymorphism (SNP) genotyping arrays and next-generation sequencing technologies are powerful tools to measure the level of genetic polymorphism within a population (Bianchi *et al*., 2020). In this study, we developed 3,505 high-quality SNP markers from the 236 cultivated and wild enset samples collected from various enset growing zones (Paper II). The SNP marker data was used to estimate genetic diversity, population structure, and selection signatures in cultivated and wild enset. Our results showed that the genetic structure and cluster analyses clearly distinguished between cultivated and wild enset due to the difference in propagation methods and cultivation status. The cultivated enset is propagated from induced suckers by farmers, and wild enset is known to propagate by seeds (Birmeta *et al.,* 2004; Gerura *et al.,* 2019; Olango *et al.,* 2015; Tesfamicael *et al.,* 2020). AMOVA showed much higher levels of genetic variation within populations and regions (91.2 and 92.4%, respectively) than between populations and regions (8.8 and 7.6%, respectively). This shows that the region of origin and environmental heterogeneity have little influence on the genetic variation in enset accessions. According to some studies, large genetic variation within populations is not necessarily caused by environmental heterogeneity but could be due to historical patterns of relationship (Schaal *et al*., 1998). However, the genetic differentiation between regions was moderate to large (F_{ST} = 0.06–0.17). The highest F_{ST} values were observed between wild enset (from South Omo) and cultivated ensets from Gurage and Sidama. This shows that these accessions are more isolated from one another and most likely there is no wild enset growing in these regions. However, moderate genetic differentiation was found in South Omom region. This might be due to the co-existence of wild and cultivated enset in the South Omo region. Moreover, we also found that wild enset had lower levels of heterozygosity than expected, indicating that wild enset is a sexually propagated plant within a restricted area. This might be lead to inbreeding and increased homozygosity. However, relatively high levels of heterozygosity were observed in all cultivated populations, which is consistent with the outcrossing nature of enset during sexual reproduction. Most likely the genetic structure of enset can to a

large extent be explained by eco-geographic factors, cultivation status and mode of propagation. The genetic structure of plant populations reflects the interactions of various factors, including the long-term evolutionary history of genetic drift, the reproductive system, gene flow and selection (Godoy *et al.,* 2018; Schaal *et al.,* 1998). Overall, high levels of unexplored diversity for enset improvement remain available in Ethiopia, with patterns of diversity consistent with divergent selection on adaptive traits. Furthermore, an outlier F_{ST} analysis identified 12 candidate loci that distinguished between wild and cultivated enset. Among these, six loci might be under positive selection and potentially involved in sexual reproduction and flowering signalling, which are key processes underlying domestication and adaptation (Borrell *et al.,* 2019). This demonstrates that sexual reproduction plays an important role in shaping enset diversity.

3.3. Genetic diversity among clones of the Entada landrace and its relationships with cultivated and wild enset (Paper III)

Most of the plant landraces are cultivated and maintained by smallholder farmers and private gardeners all over the world. In this context, the Entada landrace (*Ensete ventricosum,* (Welw.) Chessman, var. Entada) is probably the most typical indigenous crop in Ethiopia. Entada landrace propagated by natural suckers while the cultivated enset propagated by induced suckers. We developed 1,617 high quality polymorphic SNP markers from ddRAD sequences across 117 Entada landraces, since 12 of the 129 Entada genotypes collected turned out to be cultivated enset with suckers (see below). These markers were used to study the genetic diversity among the Entada landraces. For the comparison of Entada with cultivated and wild ensets, we used 2,823 high quality SNP markers, which were polymorphic in both cultivated enset (226 genotypes), wild enset (10 genotypes) and the Entada landraces (117). The joint phylogenetic analysis involving all enset genotypes grouped the cultivated enset, wild enset and Entada landraces into three distinct clusters. Entada forms a completely separated clade from cultivated and wild enset accessions in the phylogenetic analysis (Paper III). However, an interesting result from the joint phylogenetic analysis was that 12 of the Entada landraces clustered with cultivated enset. Morphological characters confirm that these Entada landraces indeed are cultivated enset with natural suckers. This was surprising since we expected these landraces with natural suckers to cluster in the Entada group. This illustrates that high levels of unexplored diversity for enset improvement remain available in Ethiopia.

PCA analysis showed a clear differentiation between Entada landraces from Sidama and North Ari, and nearly complete overlap of landraces from South and North Ari, indicating that Sidama landraces are more isolated from the two other regions. This is not surprising since South and North Ari are very close neighbouring regions. However, minimal molecular variation was detected between regions (0.48%), nearly all variation was present within individuals (99.5%). This study revealed that asexual propagation limit the genetic diversity between, but not within individuals. Subpopulation differentiation between regions was not found, however, observed heterozygosity was very high, which is expected of Entada plants with strict asexual propagation by spontaneous sucker (Olango *et al.,* 2015; Shigeta, 1992). In conclusion, Entada landraces with prolonged clonal growth and propagation will develop monoclonal populations with

no significant clonal diversity. Furthermore, for the first time, signatures of selection were studied using the F_{ST} outlier approach, which identified eight loci differentiating between cultivated enset and Entada landrace. Four of the loci have putative gene functions, i.e., *Lateral suppressor protein, Auxin response factor 2A, Cytokinin dehydrogenase,* and *Scarecrow-like protein 18.* The candidate genes are involved in axillary shoot growth, and they might have important influences on the natural and induced sucker formation in Entada landraces and cultivated enset, respectively.

3.4. Characterization of nutritional composition and bioactive compounds in Enset (Paper IV)

The current study is the first assessment of the nutritional variation among different forms of enset. Enset (*Ensete ventricosum,*) is grown for its carbohydrate-rich food obtained from the pseudostem, leaf sheaths and underground corm harvested 3-12 years after planting (Brandt *et al.,* 1997). According to farmers, corm size, tissue quality for extracting starch, root structure for harvestability, drought, frost and disease tolerance are traits that are variable among clonal genotypes (Asfaw, 2002; Tsegaye and Struik, 2001). We performed chemical analyses and estimated the contents of amylose (AM), neutral detergent fiber (NDF), antioxidant capacity (AO), total phenol content (TPC), water-soluble carbohydrates (WSC), dry matter (DM), and major minerals and minor elements in cultivated enset (14 common cultivars), 6 originating from Sidama and 8 from Gurage), wild enset (2 clones) and Entada landraces (2 clones) originating from Sidama.

We detected statistically significant nutritional differences between accessions and edible parts of enset tissues, i.e., leaf sheaths and corm. On average, higher amylose content was observed in leaf sheaths than in corm, suggesting that the starch biosynthesis was more efficient in leaf sheaths than in the corm. Although not statistically significant, genotypes from Sidama had higher amylose content than Gurage genotypes in both tissues, most pronounced in corms (P=0.07). The highest amylose content was found in the leaf sheath of the Mundraro cultivar (64.1%) from Sidama and in the corm of one of the wild ensets (62.6%), while the lowest content was 22.0% in the leaf sheath of cv. Astara and 20.9% in the corm of cv. Agade, both cultivars from Gurage. The Mundraro genotype from Sidama is especially interesting since it has highest amylose content of all cultivated ensets both in leaf sheath and corm. The differences in amylose content are most certainly due to genotypic variation, at least for variation between the cultivated ensets which had been grown in the same field for many years and were sample at the same age. The differences are also due to variation in the main edible tissues (Gebre-Mariam *et al*., 1996; Moorthy, 2002; Seung, 2020). Surprisingly, except for two enset accessions, the average amylose content in the leaf sheath (36.3-64.1%) is higher than in banana (24.4-40.7%) (Fontes *et al*., 2017; Ravi and Mustaffa, 2013; Waliszewski *et al*., 2003). Starch from accessions with the highest amylose content can be used as potential sources for maltodextrin and glucose syrup production.

The WSC content varied significantly between tissues (highest in leaf sheaths), cultivation status and enset vs. Entada in leaf sheath. WSC content in the corm did not vary significantly between different enset genotypes and geographical regions, while WSC contents in the leaf sheaths differed significantly between

different enset types, being higher in cultivated enset than in wild enset and Entada. The common cultivar 'Kiticho' stands out since it has the highest WSC content in leaf sheaths and is also among the highest in WSC of the corm. We found a moderate positive correlation (r=0.49) between WSC content in the leaf sheaths and amylose content, but negative correlation (-0.18) in the corm. Since starch is a polymer of glucose units and glucose is on fraction of WSC, we would expect a positive relationship between carbohydrate and starch/amylose contents. The negative correlation in the corm is thus difficult to explain.

The amount of fibre, measured by NDF, is important for the feed value of enset leaves, and low NDF content means that the leaves have high energy concentration and provides a good forage for ruminants. The Entada landraces had lowest NDF content in the leaf sheaths, wild enset highest and of the cultivated ensets, Mundraro and Kiticho had significantly higher NDF content than the other genotypes. Thus, the Entada genotypes should be very promising as feed sources providing that the yield is satisfactory. The NDF content in the present study was lower than the values reported by Nurfeta *et al*., (2009). The NDF content of wild enset in the present study is comparable with that of banana pseudostem (Carmo *et al*., 2018). The differences in NDF contents is due to variation in genotype, and roots and tubers crops generally contain a substantial amount of dietary fibre (Chandrasekara and Josheph Kumar, 2016).

The large positive correlations between NDF content on the one hand, and WSC and amylose in the leaf sheaths is somewhat surprising. It is likely that high WSC content create a good source for starch deposition, and WSC is also providing the building blocks (glucose and other sugars) of cellulose and hemicellulose, which are main components of NDF. The positive correlation can only be explained by the composition of NDF, i.e., it must be made up of a low proportion of lignin and mostly hemicellulose. Also, the NDF content among the cultivated enset genotypes was low compared to other studies, which could have affected these correlations.

Significant variation of mineral content was observed between corm and leaf sheath of enset, except for contents of Mg, P and Na. Mineral contents did not vary according to the geographical region, except for Mg in corms which was significantly highest among Sidama genotypes. Some enset genotypes are believed to have medicinal value, maybe due to large amounts of Ca and P (Daba and Shigeta 2016), and they are used for that purpose by enset farmers. The wild enset had much larger content of Ca than the cultivated ensets in both tissues, which could be a genotypic effect but also an environmental effect since the wild ensets were not grown in the same experimental field as the cultivated ensets. However, another study (only on cultivated enset) showed lower Ca contents in the leaf sheath and corm than the values reported in this study (Nurfeta et al., 2008). The content of the minor minerals S, Fe, Zn, and Cu was significantly higher in corms than in leaf sheaths, except for Fe. The highest Fe content was found in leaf sheaths of wild enset and the lowest in the corms of the Entada landraces. All accessions had higher Cu content in the corm than in the leaf sheath. These results are in agreement with previous reports (Debebe et al., 2012). We provide further evidence that enset is an essential source of Zn and Cu, and that corm contains much higher levels of these minor elements than leaf sheaths.

As expected, very strong positive correlations were observed between TPC and AO contents both in leaf sheaths and corms. The average TPC and AO contents in leaf sheaths and corms were not significantly different, but within tissues, there were significant differences both within and between different types of enset. Of the cultivated enset, Kiticho had highest TPC and AO contents in leaf sheaths, while Agade had highest contents in corms, but lowest contents in the leaf sheath. This demonstrates very different characteristics of these two commonly grown genotypes. However, genotype Wild1 had significantly highest antioxidant capacity in leaf sheaths (9.2 $\mu molg^{-1}$ FW) and Wild2 in corms (8.8 $\mu molg^{-1}$ FW) of all enset genotypes. This demonstrates that there are valuable untapped genetic resources among wild ensets that can be utilized for improving food quality of enset products. In general, the content of bioactive compounds in tissues of Entada genotypes is average, and Entada does not seem to be promising sources of antioxidants. Except for two of the common cultivated ensets, the others have higher TPC contents than reported for enset by Forsido et al., (2013). The quantity of TPC content will vary with genotype and tissue part (Chung et al., 2008; Patthamakanokporn et al., 2008). Like banana, the enset plant has different content of total phenol in different edible parts (Kandasamy and Aradhya, 2014). Edible enset tissue parts can serve as an equally good source of dietary polyphenols as other Ethiopian staple carbohydrate foods, as also shown by Forsido et al., (2013). Further, the total phenolic content of most of the enset accessions are very high compared with other tuber vegetables reported in the literature (Cornago et al., 2011; Shan et al., 2005).

The enset genotypes differed considerably in dry matter content, for most of them, DM content was much higher in corms than in leaf sheaths. This has also been found in other studies (Debebe et al., 2012; Negash, 2002; Nurfeta et al., 2008). There are, however, some genotypes, i.e., Medasho, Ado, Badedet, and one of the wild ensets (Wild2), that has very similar DM content in the two tissues. Wild2 has very low DM content, and this type would an interesting genotype for feeding ruminants during the dry season in areas where the water supply is scarce, as pointed out by Nurfeta et al. (2008).

4. Conclusions

Enset (*Ensete ventricosum*) is a large perennial herbaceous plant distributed across central, eastern, and southern parts of Africa. Despite of its wide distribution in Africa, it has only been cultivated and domesticated in the southern part of Ethiopia. Surprisingly, it is not much known outside a narrow zone of cultivation in southern Ethiopia considering the fact that it is a staple food for 20 to 35 million people through its starch rich main edible parts, the corm and the pseudostem. Moreover, the large majority of enset diversity existing in farmers' field are unrepresented in *ex situ* collections. Genetic diversity assessment is crucial for characterization and conservation of accessions, and for utilization in breeding programs. In this PhD work, baseline information on the genetic and nutritional diversities of enset collected from different enset cultivation areas in the South and South-Western parts of Ethiopia is presented.

Preservation of plant tissue sampled from remote farmers' fields, and rapid extraction of DNA of high quality and quantity for utilization in molecular studies is challenging. For genetic diversity studies, researchers usually collect plant materials from different locations, and it is critical to preserve the biological samples so that the integrity of DNA is preserved for the longest possible time. We found that the classical saturated NaCl-CTAB method was the best method for preserving leaf tissues and secure integrity and good quality of the extracted genomic DNA for further downstream applications. The preservation protocol was validated for leaf tissues of all cultivated, wild and Entada landraces. Efficient, robust, and low-cost methods for preservation and storage of plant tissue sampled on collection trips to remote areas are very important and should receive further attention.

In conclusion, our genetic diversity and population structure studies on enset genotypes, managed and maintained by different farming communities in Southern Ethiopia, found a significant subdivision between cultivated and wild enset and a large genetic variation within populations, indicating a heterogeneous collection. Most of the genetic variability exists within geographical regions and very little between regions, and enset from Sidama and South Omo are more genetically diverse than enset from Gurage. Relatively high levels of heterozygosity within populations are consistent with the outcrossing nature of enset during sexual reproduction, which has played an important role in shaping the genetic structure of enset. This was further corroborated by an outlier F_{ST} analysis that identified six genes linked to genes involved in sexual reproduction and flowering signalling, which are key processes underlying domestication and adaptation. Extensive and unexplored genotype diversity is present among enset in Ethiopia, encouraging further exploration and preservation of genotypes with desirable traits. Our results are useful for the conservation of genetic resources, especially under global climate changes, and contribute to the potential discovery of functional genes and genetic mechanisms related to adaptability of enset to local climatic conditions, especially drought. This is encouraging for the potential of diversifying crops also in regions where enset is not traditionally grown, such as the food insecure dry north in Ethiopia.

We describe, for the first time, genetic diversity among and within the Entada landrace, a variant of enset that has lost apical dominance present in enset. It is therefore propagated by natural suckers as banana. Entada formed a completely separated clade and group based on phylogenetic and PCA analyses of Entada, cultivated and wild enset genotypes. However, little molecular variation was detected between landraces from the different regions. Despite this, the PCA analysis differentiated between genotypes from Sidama and the Ari region. Observed heterozygosity was extremely high, which is expected in plants with strict asexual propagation. Prolonged clonal propagation of heterozygous genotypes from a single or few founding lineages will lead to populations with very little or no diversity between genotypes, and extremely high heterozygosity. Signatures of directional selection were identified at eight loci using an F_{ST} analysis across cultivated enset and Entada landraces. These genes are involved in axillary shoot growth and might be involved in controlling natural sucker formation in Entada. Overall, this study has generated a useful resource of information for further conservation efforts, genetic research, and breeding of Entada landraces.

We analysed the composition of chemical and bioactive compounds of fresh leaf sheaths and corms of fourteen common enset cultivars, originating from the Sidama and Gurage regions, and in addition two wild ensets and two Entada genotypes from Sidama. These genotypes were selected to secure diversity and was based on the relationships revealed in the phylogenetic study. Differences in chemical composition between accessions, regions, and the edible parts leaf sheaths and corms were significant. On average, higher amylose content was observed in leaf sheaths than in corm, and highest amylose content was found in the leaf sheath of cv. Mundraro from Sidama (64.1%) and in the corm of one of the wild ensets (62.6%). Genotypes/cultivars from Sidama have higher amylose content in both plant parts than ensets from Gurage. The amylose content of some of the enset genotypes is very high compared to other root and tuber crops like potato and cassava, which has amylose contents around 20%. These high-amylose genotypes could be used as sources for production of maltodextrin and glucose syrup. The Entada landraces had the lowest NDF, while wild enset had highest NDF content. NDF is important for the feed value of enset leaves, and low NDF content means that the leaves have high energy concentration and provides a good forage for ruminants. The content of major and trace mineral elements, and bioactive compounds, show different variability regarding variations between tissues, genotypes, and regions of origin. The content and availability of these nutritional elements are essential for people that has enset as their major staple food source. The variations demonstrated in this study can be used to select and develop enset cultivar with specific nutritional compositions, and also provide information to farmers which enset cultivars would create optimal combinations on their farms.

Enset has a history of being an Ethiopian indigenous food security crop during difficult times, e.g., during severe drought. We hope that our study will contributes to changing the enset production and consumption systems. This study demonstrates that enset has high value, not only as source of starch, but also because of its balanced content of minerals and bioactive compounds. There is a great potential to further utilize the significant diversity present among the different types of enset, also for introducing enset as a more food secure crop for the food insecure regions in the dry north of Ethiopia.

5. Future prospects and needs

- Wide-scale efforts on natural resource preservation and maintenance of the existing Entada landraces, cultivated and wild enset germplasm in Ethiopia are also essential. Moreover, the genetic variation needs to be assessed using morphological, agronomic, and molecular traits to develop a strategy to conserve and utilize the existing genetic diversity.
- Further research is needed to expand this study to other regions of Ethiopia that are not covered by the present study and give more comprehensive coverage to wild enset, as well as consider elite genotypes across central, eastern, and southern parts of Africa to investigate the available diversity thoroughly.
- More research is needed to investigate the cultural, socio-economic, and gender-associated aspects of enset cultivation to assess and understand the dynamics of enset biodiversity.

- No study has been conducted so far on the distribution, management, modes of reproduction and flowering. Therefore, it is essential that the interdisciplinary approaches such us combining ethnobotanical with molecular methods is needed for documentation of the indigenous knowledge system.
- Enset is an underexploited starch-stable crop with a higher potential in Ethiopia and beyond. Thus, further research is required to upgrade the nutritional content and the processing technology for all types of enset and making the devices available to the end users.
- Awareness and knowledge about all types of enset in Ethiopia are very insignificant. Thus, we suggest more studies and policymakers' attention toward maintaining the existing enset farming system and wider distribution and adoption in the other areas of the country, especially the drought-prone regions in the north of Ethiopia.
- In the end, high-throughput genotyping and phenotyping techniques should be used for the description and utilization of enset genetic resources, for development of improved enset types with broad adaptation to diverse climates and farming systems by novel breeding methods.

References

Abdel-Latif, A., Osman, G. (2017). Comparison of three genomic DNA extraction methods to obtain high DNA quality from maize. *Plant methods*, *13*, 1-9. https://doi.org/DOI:10.1186/s13007-016-0152-4

Akinnagbe, A., Ji, X. B., Yang, M., Ewald, D. (2011). Ethanol pretreatment increases DNA yields from dried tree foliage. *Conservation Genetics Resources*, *3*, 409-411. https://DOI:10.1007/s12686-010-9367-2

Alemu, K., Sandford, S. (1991). Enset in North Omo region. Farmer's research technical pamphlet, No.1. Farm Africa, Addis Ababa, Ethiopia

Alkan, C., Coe, B. P., Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, *12*(5), 363-376. https://doi.org/doi:10.1038/nrg2958

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81. https://doi.org/doi:10.1038/nrg.2015.28

Archana, M., Vimala, B., Bala, N., Chakrabarti, S., George, J., Gowda, H. (2015). Underutilized tropical tuber crops with hidden treasure of food, nutrition and medicine. *International Journal of Tropical Agriculture*, *33*(4 (Part IV)), 3803-3815.

Asfaw, B. T. (2002). *Studies on landrace diversity, in vivo and in vitro regeneration of enset (Enset ventricosum Welw.)*. Köster: Milan, Lombardy, Italy.

Asfaw, Z. (1997). Conservation and use of traditional vegetables in Ethiopia. In Traditional African Vegetables: Proceedings of the IPGRI International Workshop on Genetic Resources of Traditional Vegetables in Africa. Conservation and Use. ICRAF-HQ, Nairobi. Institute of Plant Genetic and Crop Plant Research, Rome (pp. 57-65).

Ashango, T. B. (2017). EFFECT OF CORM AND CORM PIECES ON REGENERATION AND MULTIPLICATION OF ENSET (ENSETE VENTRICOSUM (WELW.) CHEESMAN). *International Journal of Research-GRANTHAALAYAH*, *5*(5), 281-299. https://doi.org/10.29121/granthaalayah.v5.i5.2017.1860

Bailey, A. (2016) Mainstreaming agrobiodiversity in sustainable food systems: Scientific foundations for an agrobiodiversity index - Summary. Rome (Italy): Bioversity International, 32 p.

Baker, R., Simmonds, N. (1953). The genus Ensete in Africa. *Kew Bulletin*, *8*(3), 405-416. https://doi.org/10.2307/4115529

The World Bank, T. W. (2022). *Poverty and Shred Prosperity 2022: Correcting Course*. The World Bank. https://doi.org/10.1596/978-1-4648-1893-6

Beaumont, M. A., Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*(1377), 1619-1626. https://doi.org/10.1098/rspb.1996.0237

Behera, K. K., Maharana, T., Sahoo, S., Prusti, A. (2009). Biochemical quantification of protein, fat, starch, crude fibre, ash and dry matter content in different collection of greater yam (Dioscorea alata L.) found in Orissa. *Nature Science*, *7*(7), 24-32.

Bekele, E., Shigeta, M. (2011). Phylogenetic relationships between Ensete and Musa species as revealed by the trnT trnF region of cpDNA. *Genetic Resources and Crop Evolution*, *58*(2), 259-269. https://doi.org/10.1007/s10722-010-9568-2

Benzie, I. F., Strain, J. (1999). [2] Ferric reducing/antioxidant power assay: direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. *Methods in Enzymology*, *299*, 15-27. https://doi.org/10.1016/S0076-6879(99)99005-5

Bezuneh, T. (2010,). Technological challenges of sustainable enset farming system: For enhancing the production of food/fiber and industrial outputs. Proceedings of Enset Research and Development Experiences in Ethiopia, Wolkite, Ethiopia. 19–20 August 2010

Bezuneh, T., Feleke, A., Bayie, R. (1967). The cultivation of the genus Ensete in Ethiopia. Soil and Crop Science Society of Florida, Florida.

Bianchi, D., Brancadoro, L., De Lorenzis, G. (2020). Genetic Diversity and Population Structure in a Vitis spp. Core Collection Investigated by SNP Markers. *Diversity*, *12*(3), 103. https://doi.org/10.3390/d12030103

Birmeta, G., Nybom, H., Bekele, E. (2004). Distinction between wild and cultivated enset (Ensete ventricosum) gene pools in Ethiopia using RAPD markers. *Hereditas*, *140*(2), 139-148. https://doi.org/10.1111/j.1601-5223.2004.01792.x

Bolger, A. M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/doi.org/10.1093/bioinformatics/btu170

Borrell, J. S., Biswas, M. K., Goodwin, M., Blomme, G., Schwarzacher, T., Heslop-Harrison, J. S. P., Wendawek, A. M., Berhanu, A., Kallow, S., Janssens, S., Molla, E. L., Davis, A. P., Woldeyes, F., Willis, K., Demissew, S., Wilkin, P. (2019). Enset in Ethiopia: a poorly characterized but resilient starch staple. *Annals of Botany*, *123*(5), 747-766. https://doi.org/10.1093/aob/mcy214

Borrell, J. S., Goodwin, M., Blomme, G., Jacobsen, K., Wendawek, A. M., Gashu, D., Lulekal, E., Asfaw, Z., Demissew, S., Wilkin, P. (2020). Enset-based agricultural systems in Ethiopia: A systematic review of production trends, agronomy, processing and the wider food security applications of a neglected banana relative. *Plants, People, Planet*, *2*(3), 212-228. https://doi.org/DOI:10.1002/ppp3.10084

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633-2635. https://doi.org/doi:10.1093/bioinformatics/btm308

Brandt, S., Spring, A., Hiebsch, C., McCabe, J., Tabogie, E., Diro, M., Wolde-Michael, G., Yntiso, G., Shigeta, M., Tesfaye, S. 1997. *"The "Tree Against Hunger": Enset-based agricultural systems in Ethiopia"*. American Association for the Advancement of Science,Washington, DC, USA.

Bressan, E. A., Rossi, M. L., Gerald, L. T., Figueira, A. (2014). Extraction of high-quality DNA from ethanol-preserved tropical plant tissues. *BMC Research Notes*, *7*(1), 268. https://doi.org/10.1186/1756-0500-7-268

Bruce, J. (1790). *Select specimens of natural history, collected in travels to discover the source of the nile, in egypt, arabia, abyssinia, and nubia* (Vol. 5). J. Ruthven.

Bus, A., Hecht, J., Huettel, B., Reinhardt, R., Stich, B. (2012). High-throughput polymorphism detection and genotyping in Brassica napus using next-generation RAD sequencing. *BMC Genomics*, *13*(1), 1-11.

Campbe, B. (1987). The use of wild fruits in Zimbabwe. *Econ Bo*, *41*(3), 375-385. https://doi.org/doi.org/10.1007/BF02859054

Chandrasekara, A., Josheph Kumar, T. (2016). Roots and tuber crops as functional foods: a review on phytochemical constituents and their potential health benefits. *International journal of food science 2016*, 15. https://doi.org/10.1155/2016/3631647

Cheesman, E. (1947). Classification of the bananas: the genus Ensete Horan. *Kew bulletin*, *2*, 97-106. https://doi.org/10.2307/4109206

Chivenge, P., Mabhaudhi, T., Modi, A. T., Mafongoya, P. (2015). The potential role of neglected and underutilised crop species as future crops under water scarce conditions in Sub-Saharan Africa. *International journal of environmental research and public health*, *12*(6), 5685-5711. https://doi.org/10.3390/ijerph120605685

Chung, Y. C., Chiang, B. H., Wei, J. H., Wang, C. K., Chen, P. C., Hsu, C. K. (2008). Effects of blanching, drying and extraction processes on the antioxidant activity of yam (Dioscorea alata). *International Journal of Food Science and Technology*, *43*(5), 859-864. https://doi.org/10.1111/j.1365-2621.2007.01528.x

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674-3676. https://doi.org/10.1093/bioinformatics/bti610

Central Statistical Agency (CSA), Government of Ethiopia. 2020. Agricultural sample survey 2019/20 (2012 E.C.) Report on area and production of major crops (Private Peasant Holdings, Meher Season). Statistical Bulletin 587 (Vol. I). Addis Ababa, Ethiopia

D'hont, A., Denoeud, F., Aury, J.-M., Baurens, F.-C., Carreel, F., Garsmeur, O., Noel, B., Bocs, S., Droc, G., Rouard, M. (2012). The banana (Musa acuminata) genome and the evolution of monocotyledonous plants. *Nature*, *488*(7410), 213-217.

Daba, T., Shigeta, M. (2016). Enset (Ensete ventricosum) production in Ethiopia: its nutritional and socio-cultural values. *Agriculture Food Sciences Research*, *3*(2), 66-74. https://doi.org/10.20448/journal.512/2016.3.2/512.2.66.74

Dansi, A., Vodouhè, R., Azokpota, P., Yedomonhan, H., Assogba, P., Adjatin, A., Loko, Y., Dossou-Aminon, I., Akpagana, K. (2012). Diversity of the neglected and underutilized crop species of importance in Benin. *The scientific world journal*, *2012*, 19. https://doi.org/doi:10.1100/2012/932947

Debebe, A., Chandravanshi, B., Wondimu, T. (2012). Metallic nutrients in enset (Ensete ventricosum) corm cultivated in Wolliso and Wolkite Towns in Ethiopia. *Ethiopian Journal of Science*, *35*(2), 71-80.

Dessalegn, E. (2019). Effect of Traditional Fermentation Time on Total Phenolic Content and In vitro Antioxidant Activity of Kocho: Ethiopian Fermented Food from Enset (Ensete ventricusum welw. Chessman*). Food Science and Quality Management*, 91: 44-51, https://doi.org/10.7176/FSQM/91-05

Diro, M., Van Staden, J., Bornman, C. (2003). In vitro regeneration of Ensete ventricosum from zygotic embryos of stored seeds. *South African Journal of Botany*, *69*(3), 364-369.

United Nations. (2022). World population prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO. 3. 2022.

Endalew, B., Muche, M., Tadesse, S. (2015). Assessment of food security situation in Ethiopia: a review. *Asian Journal of Agricultural Research*, *9*(2), 55-68.

Esfeld, K., Uauy, C., Tadele, Z. (2013). Application of TILLING for orphan crop improvement. *Biotechnology of neglected and underutilized crops*, 83-113. https://doi.org/DOI:10.1007/978-94-007-5500-0

Excoffier, L., Hofer, T., Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*(4), 285-298. https://doi.org/10.1038/hdy.2009.74

Excoffier, L., Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

FAO. (2009). Global agriculture towards 2050. High Level Expert Forum-How Feed World.

Fontes, S. d. M., Cavalcanti, M. T., Candeia, R. A., Almeida, E. L. (2017). Characterization and study of functional properties of banana starch green variety of Mysore (Musa AAB-Mysore). *Food Science and Technology*, *37*, 224-231. https://doi.org/doi.org/10.1590/1678-457X.18916

Forsido, S. F., Rupasinghe, H. V., Astatkie, T. (2013). Antioxidant capacity, total phenolics and nutritional content in selected Ethiopian staple food ingredients. *Tropical Resource Mangement Papers*, *64*(8), 915-920. https://doi.org/10.3109/09637486.2013.806448

Gebre-Mariam, T., Abeba, A., Schmidt, P. (1996). Isolation and physico-chemical properties of enset starch. *Starch-Stärke*, *48*(6), 208-214. https://doi.org/10.1002/star.19960480603

Gebremariam, S. (1996). Enset research in Ethiopia 1976-1984. Proceedings from the International Workshop on Enset, Addis Abeba (Ethiopia), 13-20 Dec 1993.

Gerura, F. N., Meressa, B. H., Martina, K., Tesfaye, A., Olango, T. M., Nasser, Y. (2019). Genetic diversity and population structure of enset (Ensete ventricosum Welw Cheesman) landraces of Gurage zone, Ethiopia. *Genetic Resources and Crop Evolution*, *66*(8), 1813-1824. https://doi.org/10.1007/s10722-019-00825-2

Getachew, S., Mekbib, F., Admassu, B., Kelemu, S., Kidane, S., Negisho, K., Djikeng, A., Nzuki, I. (2014). A look into genetic diversity of enset (Ensete ventricosum (Welw.) cheesman) using transferable microsatellite sequences of banana in Ethiopia. *J. Crop Improv*, *28*(2), 159-183. https://doi.org/doi.org/10.1080/15427528.2013.861889

Gezahegn, B., Gebire, W., Bamud, K. (2018). Identification of major crop production constraints in debub ari district, south omo zone, southern Ethiopia. *Horticult Int J*, *2*(6), 383-388. https://doi.org/DOI:10.15406/hij.2018.02.00081

Godoy, F. M. d. R., Lenzi, M., Ferreira, B. H. D. S., Silva, L. V. D., Zanella, C. M., Paggi, G. M. (2018). High genetic diversity and moderate genetic structure in the self-incompatible, clonal Bromelia hieronymi (Bromeliaceae). *Botanical Journal of the Linnean Society*, *187*(4), 672-688. https://doi.org/doi.org/10.1093/botlinnean/boy037

Greenfield, H., Southgate, D. A. (2003). *Food composition data: production, management, and use*. Food and Agriculture Organization of the United Nations. https://doi.org/10.1007/978-1-4615-3544-7

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307-321. https://doi.org/10.1093/sysbio/syq010

Guo, Y., Yang, G., Chen, Y., Li, D., Guo, Z. (2018). A comparison of different methods for preserving plant molecular materials and the effect of degraded DNA on ddRAD sequencing. *Plant Diversity*, *40*, 106-116. https://doi.org/10.1016/j.pld.2018.04.001

Guzzon, F., Müller, J. V. (2016). Current availability of seed material of enset (Ensete ventricosum, Musaceae) and its Sub-Saharan wild relatives. *Genetic resources*, *63*(2), 185-191. https://doi.org/DOI:10.1007/s10722-015-0355-y

Haile, M. (2014). Cluster analysis for evaluation of genetic diversity in Enset (*Enset ventricosum* (Welw.): Cheesman) clones at Areka Condition. J Plant Sci, 2(1), 55-69.

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite North American potato germplasm. *BMC Genomics*, *12*(1), 302. https://doi.org/doi.org/10.1186/1471-2164-12-302

Harrison, J., Moore, K. A., Paszkiewicz, K., Jones, T., Grant, M. R., Ambacheew, D., Muzemil, S., Studholme, D. (2014). A draft genome sequence for ensete ventricosum, the drought-tolerant "tree against hunger". *Agronomy*, *4*(1), 13-33. https://doi.org/10.3390/agronomy4010013

Hinze, L. L., Hulse-Kemp, A. M., Wilson, I. W., Zhu, Q.-H., Llewellyn, D. J., Taylor, J. M., Spriggs, A., Fang, D. D., Ulloa, M., Burke, J. J. (2017). Diversity analysis of cotton (Gossypium hirsutum L.) germplasm using the CottonSNP63K Array. *BMC Plant Biology*, *17*(1), 37. https://doi.org/DOI:10.1186/s12870-017-0981-y

Jacobsen, K., Blomme, G., Tawle, K., Muzemil, S., Yemataw, Z. (2018). Dietary diversity associated with different enset [Ensete ventricosum (Welw.) Cheesman]-based production systems in Ethiopia. *Fruits*. https://doi.org/10.17660/th2018/73.6.6

Karssa, T., Papini, A. (2018). Effect of clonal variation on quality of Kocho, traditional fermented food from enset (Ensete ventricosum), Musaceae. *Int. J. Food Sci. Nutr. Eng*, *8*, 79-85. https://doi.org/DOI:10.5923/j.food.20180803.04

Kennedy, J. (2009). Bananas and people in the homeland of genus Musa: not just pretty fruit. *Ethnobotany Research*, *7*, 179-197.

Kenyon, L., Anandajayasekeram, P., Ochieng, C., Ave, C., Uk, K. (2006). A synthesis/lesson-learning study of the research carried out on root and tuber crops commissioned through the DFID RNRRS research programmes between 1995 and 2005 (DFID, *United Kingdom*.

Kotchoni, S. O., Gachomo, E. W. (2009). A rapid and hazardous reagent free protocol for genomic DNA extraction suitable for genetic studies in plants. *Molecular Biology Reports*, *36*(6), 1633-1636. https://doi.org/10.1007/s11033-008-9362-9

Lebot, V. (2019). *Tropical root and tuber crops*. Cabi.

Lee, H.-Y., Ro, N.-Y., Jeong, H.-J., Kwon, J.-K., Jo, J., Ha, Y., Jung, A., Han, J.-W., Venkatesh, J., Kang, B.-C. (2016). Genetic diversity and population structure analysis to construct a core collection from a large Capsicum germplasm. *BMC genetics*, *17*(1), 142. https://doi.org/doi.org/10.1016/j.plaphy.2015.05.010

Letunic, I., Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*, 256-259. https://doi.org/doi:10.1093/nar/gkz239

Costa, M. G. d., Lockhart, D. M. A. (1984). *The Itinerario of Jeronimo Lobo*. London : Hakluyt Society.

Magbagbeola, J., Adetoso, J., Owolabi, O. (2010). Neglected and underutilized species (NUS): a panacea for community focused development to poverty alleviation/poverty reduction in Nigeria. *Journal of Economics and International Finance*, *2*(10), 208-211

.

Margam, V. M., Gachomo, E. W., Shukle, J. H., Ariyo, O. O., Seufferheld, M. J., Kotchoni, S. O. (2010). A simplified arthropod genomic-DNA extraction protocol for polymerase chain reaction (PCR)-based specimen identification through barcoding. *Molecular Biology Reports*, *37*(7), 3631-3635. https://doi.org/10.1007/s11033-010-0014-5

Massawe, F., Mayes, S., Cheng, A. (2016). Crop diversity: an unexploited treasure trove for food security. *Trends in Plant Science*, *21*(5), 365-368.

Mayes, S., Massawe, F., Alderson, P., Roberts, J., Azam-Ali, S., Hermann, M. (2012). The potential for underutilized crops to improve security of food production. *Journal of Experimental Botany*, *63*(3), 1075-1079. https://doi.org/10.1016/j.tplants.2016.02.006

Mertens, D. R. (2002). Gravimetric determination of amylase-treated neutral detergent fiber in feeds with refluxing in beakers or crucibles: collaborative study. *Journal of AOAC International*, *85*(6), 1217-1240. https://doi.org/doi.org/10.1093/jaoac/85.6.1217

Micha, R., Mannar, V., Afshin, A., Allemandi, L., Baker, P., Battersby, J., Bhutta, Z., Chen, K., Corvalan, C., Di Cesare, M. (2020). *2020 Global nutrition report: action on equity to end malnutrition* (1916445276).

Mohamed, M., Alsadon, A., Al Mohaidib, M. (2010). Corn and potato starch as an agar alternative for Solanum tuberosum micropropagation. *African Journal of Biotechnology*, *9*(1). https://doi.org/10.4314/AJB.V9I1

Mohammed, B., Gabel, M., Karlsson, L. M. (2013). Nutritive values of the drought tolerant food and fodder crop enset. *African Journal of Agricultural Research*, *8*(20), 2326-2333. https://doi.org/10.5897/AJAR12.1296

Moorthy, S. N. (2002). Physicochemical and functional properties of tropical tuber starches: a review. *Starch-Stärke*, *54*(12), 559-592. https://doi.org/10.1002/1521-379X(200212)54:12<559::AID-STAR2222559>3.0.CO;2-F

Naylor, R. L., Falcon, W. P., Goodman, R. M., Jahn, M. M., Sengooba, T., Tefera, H., Nelson, R. J. (2004). Biotechnology in the developing world: a case for increased investments in orphan crops. *Food Policy*, *29*(1), 15-44.

Negash, A. (2001). *Diversity and conservation of enset (Ensete ventricosum Welw. Cheesman) and its relation to household food and livelihood security in South-western Ethiopia*. Wageningen University and Research.

Negash, A., Tsegaye, A., van Treuren, R., Visser, B. (2002). AFLP analysis of enset clonal diversity in south and southwestern Ethiopia for conservation. *Crop Science*, *42*(4), 1105-1111. https://doi.org/10.2135/cropsci2002.1105

Nurfeta, A., Tolera, A., Eik, L. O., Sundstøl, F. (2008). Yield and mineral content of ten enset (Ensete ventricosum) varieties. *Tropical animal health and production*, *40*(4), 299-309. https://doi.org/10.1007/s11250-007-9095-0

Olango, T. M., Tesfaye, B., Pagnotta, M. A., Pè, M. E., Catellani, M. (2015). Development of SSR markers and genetic diversity analysis in enset (Ensete ventricosum (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. *BMC Genetics*, *16*(1), 98. https://doi.org/DOI:10.1186/s12863-015-0250-8

Oli, M. T. (2006). *Assessing diversity in yams (Dioscorea spp.) from Ethiopia based on morphology, AFLP markers and tuber quality, and farmers' management of landraces*. Cuvillier Verlag. Göttingen.

Oyen, L., Lemmens, R. (2002). *Plant resources of tropical Africa. Precursor*. Programme PROTA.

Patthamakanokporn, O., Puwastien, P., Nitithamyong, A., Sirichakwal, P. P. (2008). Changes of antioxidant activity and total phenolic compounds during storage of selected fruits. *Journal of Food Composition Analysis*, *21*(3), 241-248. https://doi.org/10.1016/j.jfca.2007.10.002

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, K. M., McGinty, R. C., Couture, G., Pehrsson, P. R., McKillop, K., Fukagawa, N. K. (2021). Dietary fiber, starch, and sugars in bananas at different stages of ripeness in the retail market. *PloS One*, *16*(7), e0253366.

Purseglove, J. W. (1985). *Tropical Crops: Monocotyledons* (5th ed.). Longman.

Rahmato, D. (1995). Resilience and vulnerability: enset agriculture in southern Ethiopia. *Journal of Ethiopian studies*, *28*(1), 23-51.

Raj, A., Stephens, M., Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *197*(2), 573-589. https://doi.org/doi:10.1534/genetics.114.164350/-/DC1

Randby, Å. T., Nørgaard, P., Weisbjerg, M. R. (2010). Effect of increasing plant maturity in timothy-dominated grass silage on the performance of growing/finishing Norwegian Red bulls. *Grass and Forage Science*, *65*(3), 273-286. https://doi.org/10.1111/j.1365-2494.2010.00745.x

Ravi, I., Mustaffa, M. M. (2013). Starch and amylose variability in banana cultivars. *Indian Journal of Plant Physiology*, *18*(1), 83-87. https://doi.org/DOI:10.1007/s40502-013-0014-2

Reid, R., Horvath, D. (1980). Soil chemistry and mineral problems in farm livestock. A review. *Animal Feed Science and Technology*, *5*(2), 95-167. https://doi.org/10.1016/0377-8401(80)90002-4

Rochette, N. C., Rivera-Colón, A. G., Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737-4754. https://doi.org/DOI:10.1111/mec.15253

Rogstad, S. H. (1992). Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon*, *41*, 701-708. https://doi.org/10.2307/1222395

Schaal, B., Hayworth, D., Olsen, K. M., Rauscher, J., Smith, W. (1998). Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, *7*(4), 465-474. https://doi.org/doi.org/10.1046/j.1365-294x.1998.00318.x

Scott, G. J., Rosegrant, M. W., Ringler, C. (2000). Global projections for root and tuber crops to the year 2020. *Food Policy*, *25*(5), 561-597. https://doi.org/10.1016/S0306-9192(99)00087-1

Seung, D. (2020). Amylose in starch: towards an understanding of biosynthesis, structure and function. *New Phytologist*, *228*(5), 1490-1504. https://doi.org/10.1111/nph.16858

Shigeta, M. (1990). FOLK IN-S/TU CONSERVATION OF ENSETE (ENSETE VENTRICOSUM (WELW) EE CHEESMAN); TOWARDS THE. *African Study Monographs*, *10*(3), 93-107.

Shigeta, M. (1992). *The ethnobotanical study of ensete (Ensete ventricosum) in the southwestern Ethiopia* Kyoto University]. Japan. http://hdl.handle.net/2433/86243

Shumbulo, A., Gecho, Y., Tora, M. (2012). Diversity, challenges and potentials of enset (Ensete ventricosum) production: in case of Offa Woreda, Wolaita Zone, Southern Ethiopia. *Food Sci Q Manage*, *7*, 24-31.

Sibhatu, K. T., Krishna, V. V., Qaim, M. (2015). Production diversity and dietary diversity in smallholder farm households. *Proceedings of the National Academy of Sciences*, *112*(34), 10657-10662. https://doi.org/10.1073/pnas.1510982112

Simmonds, N. W. (1962). *The evolution of the bananas*. Longmans, Green, London.

Singleton, V. L., Orthofer, R., Lamuela-Raventós, R. M. (1999). [14] Analysis of total phenols and other oxidation substrates and antioxidants by means of folin-ciocalteu reagent. In *Methods in Enzymology* (Vol. 299, pp. 152-178). Elsevier.

Solomon, A. M., Han, K., Lee, J.-H., Lee, H.-Y., Jang, S., Kang, B.-C. (2019). Genetic diversity and population structure of Ethiopian Capsicum germplasms. *PloS One*, *14*(5), e0216886. https://doi.org/doi.org/10.1371/journal.pone.0216886

Song, F.-L., Gan, R.-Y., Zhang, Y., Xiao, Q., Kuang, L., Li, H.-B. (2010). Total phenolic contents and antioxidant capacities of selected Chinese medicinal plants. *International Journal of Molecular Sciences*, *11*(6), 2362-2372. https://doi.org/10.1016/S0076-6879(99)99017-1

Tadesse, B., Tilahun, Y., Bekele, T., Mekonen, G. (2021). Assessment of challenges of Crop Production and Marketing in Bench-Sheko, Kaffa, Sheka, and West-Omo Zones of Southwest Ethiopia. *Heliyon*, e07319. https://doi.org/10.1016/j.heliyon.2021.e07319

Tamiru, M., Becker, H. C., Maass, B. L. (2008). Diversity, distribution and management of yam landraces (Dioscorea spp.) in Southern Ethiopia. *Genetic Resources and Crop Evolution*, *55*(1), 115-131. https://link.springer.com/content/pdf/10.1007%2Fs10722-007-9219-4.pdf

Tesfamicael, K. G., Gebre, E., March, T. J., Sznajder, B., Mather, D. E., Lopez, C. M. R. (2020). Accumulation of mutations in genes associated with sexual reproduction contributed to the domestication of a vegetatively propagated staple crop, enset. *J bioRxiv*. https://doi.org/doi.org/10.1038/s41438-020-00409-7

Tesfaye, M. (1992). The effect of soaking, temperature and other pretreatments on the germination of " ensete" seed. *In Proceedings of the International Seed Testing Association*.

Thomson, J. A. (2002). An improved non-cryogenic transport and storage preservative facilitating DNA extraction from 'difficult'plants collected at remote sites. *Telopea*, *9*(4), 755-760. https://doi.org/10.7751/telopea20024013

Tobiaw, D. C., Bekele, E. (2011). Analysis of genetic diversity among cultivated enset (Ensete ventricosum) populations from Essera and Kefficho, southwestern part of Ethiopia using inter simple sequence repeats (ISSRs) marker. *African Journal of Biotechnology*, *10*(70), 15697-15709. https://doi.org/10.5897/AJB11.885

Tsegaye, A. (2002). *On indigenous production, genetic diversity and crop ecology of enset (Ensete ventricosum (Welw.) Cheesman).* Wageningen University and Research.

Tsegaye, A., Struik, P. C. (2001). Enset (Ensete ventricosum (Welw.) Cheesman) kocho yield under different crop establishment methods as compared to yields of other carbohydrate-rich food crops. *Netherlands Journal of Agricultural Science*, *49*(1), 81-94. https://doi.org/10.1016/s1573-5214(01)80017-8

Tsegaye, A., Struik, P. C. (2003). Growth, radiation use efficiency and yield potential of enset (Ensete ventricosum) at different sites in southern Ethiopia. *Annals of Applied Biology*, *142*, 71-81. https://edepot.wur.nl/18168

Tsykun, T., Rellstab, C., Dutech, C., Sipos, G., Prospero, S. (2017). Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus Armillaria cepistipes. *Heredity*, *119*(5), 371-380. https://doi.org/10.1038/hdy.2017.48

Tuffa, A. C. (2019). Value chain and nutritional analyses of Warqe food products in relation to post-harvest losses. *Swedish University of Agricultural Sciences, Uppsala, Sweden.*

Vavilov, N. I. (1951). *The origin, variation, immunity and breeding of cultivated plants* (Vol. 72). LWW.

Waliszewski, K. N., Aparicio, M. A., Bello, L. s. A., Monroy, J. A. (2003). Changes of banana starch by chemical and physical modification. *Carbohydrate polymers*, *52*(3), 237-242. https://doi.org/10.1016/S0144-8617(02)00270-9

Weir, B. S., Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358-1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Weiwei, Z., Ting, L., Yan, H., Wenjing, L., Qingfang, B., Chongwei, J., Lingli, L., Xianyong, L. (2020). Effect of nitrogen limitation on antioxidant qualities is highly associated with genotypes of lettuce (Lactuca sativa L.). *Pedosphere*, *30*(3), 414-425. https://doi.org/10.1016/S1002-0160(19)60833-7

Weng, X., Huang, Y. (2014). Relationship structure-antioxidant activity of hindered phenolic compounds. Grasas y Aceites, 65(4), e051-e051. doi: http://dx.doi.org/10.3989/gya.0225141

Westphal, E., Stevels, J. M. C., Stevels, J. M. C. 1975. *Agricultural systems in Ethiopia* (Vol. 826). Wageningen: Centre for Agricultural Publishing and Documentation.

Wheatly, C., Scoot, G. J., Wiersema, S. (1995). *Adding value to root and tuber crops: A manual on product development* . CIAT.

Winters, P., Cavatassi, R., Lipper, L. (2006). Sowing the seeds of social relations: The role of social capital in crop diversity. https://doi.org/10.22004/ag.econ.289051

Wojdyło, A., Oszmiański, J., Czemerys, R. (2007). Antioxidant activity and phenolic compounds in 32 selected herbs. *Food Chemistry*, *105*(3), 940-949. https://doi.org/10.1016/j.foodchem.2007.04.038

Yemata, G. (2020). Ensete ventricosum: A Multipurpose Crop against Hunger in Ethiopia. *The Scientific World Journal*, *2020*. https://doi.org/10.1155/2020/6431849

Yemataw, Z., Chala, A., Ambachew, D., Studholme, D. J., Grant, M. R., Tesfaye, K. (2017). Morphological variation and inter-relationships of quantitative traits in enset (Ensete ventricosum (Welw.) Cheesman) germplasm from south and south-western Ethiopia. *Plants*, *6*(4), 56. https://doi.org/10.3390/plants6040056

Yemataw, Z., Mohamed, H., Diro, M., Addis, T., Blomme, G. (2014). Ethnic-based diversity and distribution of enset (Ensete ventricosum) clones in southern Ethiopia. *Journal of Ecology and the Natural Environment*, *6*(7), 244-251. https://doi.org/DOI:10.5897/JENE2014.0450

Yemataw, Z., Muzemil, S., Ambachew, D., Tripathi, L., Tesfaye, K., Chala, A., Farbos, A., O'Neill, P., Moore, K., Grant, M. (2018). Genome sequence data from 17 accessions of Ensete ventricosum, a staple food crop for millions in Ethiopia. *Data in brief*, *18*, 285-293. https://doi.org/doi.org/10.1016/j.dib.2018.03.026

York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., Albersheim, P. (1986). Isolation and characterization of plant cell walls and cell wall components. In *Methods in Enzymology* (Vol. 118, pp. 3-40). Elsevier. https://doi.org/10.1016/0076-6879(86)18062-1

Zemede, A., Ayele, N. (1995). Home-gardens in Ethiopia: characteristics and plant diversity. *SINET, an Ethiopian Journal of Science*, *18*(2), 235-266.

Zippel, K. (2005). Diversity over time and space in Enset landraces (Ensete ventricosum) in Ethiopia. In Huber, B. A., SinclairB. J., Lampe, K. -H. (Eds.), African Biodiversity, 423–438. Boston, MA: Springer. https ://doi.org/10.1007/b105089

Paper I

Comparison of different leaf preservation methods to obtain high quality DNA from enset (*Ensete ventricosum*), a native and orphan food security crop in Ethiopia

Alye Tefera Haile[1, 2*], Sylvia Sagen Johnsen[1], Mallikarjuna Rao Kovi[1], Trine Hvoslef-Eide[1], Bizuayehu Tesfaye[2], Odd Arne Rognli[1*]

[1]Faculty of Biosciences, Department of Plant Sciences, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway

[2]School of Plant and Horticultural Science, College of Agriculture, Hawassa University, P.O. Box 05, Awassa, Ethiopia

*Corresponding Author's E-mail: odd-arne.rognli@nmbu.no

Abstract

Enset (*Ensete ventricosum*) is a staple food for more than 20 million Ethiopians and only cultivated in the native indigenous farming systems of Ethiopia. In contrast to other cultivated species in the Musaceae family, enset has been relatively little studied at the molecular level. Application of advanced molecular genetic techniques requires rapid extraction of DNA of high quality and quantity. Fresh, lyophilized tissues, as well as tissues stored in liquid nitrogen are mainly preferred to avoid DNA degradation, thus most of the DNA extraction protocols recommend these types of tissues as starting material. However, such sample processing techniques are difficult to utilize in many developing countries and at collection sites of many endemic plant species, underutilized or orphan crop species like enset. These situations necessitate the development of alternative protocols for leaf preservation and optimized methods for isolating high-quality DNA from dried or preserved leaf samples. In this study, three different leaf preservation and two DNA extraction methods were compared. Fresh young leaf tissue was preserved using the minor modified saturated NaCl-CTAB solution, silica gel or 96% ethanol at ambient temperature for more than 35 days. Subsequently, DNA was extracted using either the DNeasy Plant Mini Kit or the CTAB method. As compared to silica gel and 96% ethanol, the minor modified saturated NaCl-CTAB solution preserved the quality, quantity, and integrity of enset genomic DNA. This method consistently produced genomic DNA of high-quality and quantity at affordable cost. The DNeasy Plant Mini Kit method was found to be more efficient than the standard CTAB method, being faster and producing genomic DNA of higher quality. Using saturated NaCl-CTAB solution is an accessible, efficient, scalable, and inexpensive way to preserve enset leaves during collection and transportation. The preservation protocol was validated for leaf tissues of all cultivated and wild enset, and Entada landraces. Genomic DNA of high quality and quantity was obtained from preserved enset leaves, which can be used for further downstream applications including PCR and sequencing.

Keywords:

Ensete ventricosum, leaf preservation, NaCl-CTAB, Silica gel, Ethanol, DNA extraction

Introduction

Rapid extraction of high-quality and pure DNA is a prerequisite step for implementing the most advanced molecular techniques used in modern biological research project such as developing molecular markers, genetic mapping, sequencing, and marker-assisted selection (Guo *et al.,* 2018; Pervaiz *et al.,* 2011). The isolation of pure and intact genomic DNA of high quality and quantity is a limiting factor in many plants molecular genetic studies, mainly for the next generation sequencing platforms, where high-molecular-weight DNA, free of contaminants is required (Abdel-Latif and Osman, 2017; Vaillancourt and Buell, 2019). Unlike animals and microbes, the DNA extraction methods need to be adjusted to fit each plant species and even to each plant tissue, because of the presence of secondary metabolites. (Sangwan *et al*., 1998; Sönmezoğlu and Terzi, 2019). Various protocols have been developed for the extraction of genomic DNA from plants but a universal application has not been developed (Abdel-Latif and Osman, 2017; Varma *et al.,* 2007). The main cause of variability and modification in DNA extraction protocols such as the CTAB protocol, is the composition of cell walls and intra- and extracellular components (Pervaiz *et al.,* 2011). For genetic diversity studies, in many cases, researchers usually collect plant materials from different locations or even from different countries. It is critical to preserve biological samples using methods that maintain the integrity of DNA for the longest possible time, especially for endemic plant species mostly present in remote areas that are difficult to collect (Doyle and Dickson, 1987). Taxa from temperate zones, botanical gardens and major cultivated crops are often well sampled, but species found in the tropics, particularly in remote areas, are poorly represented (Harris, 1993).

Genomic DNA can be extracted from fresh, dried, or preserved plant tissues using various extraction methods. DNA extraction from plants is generally compromised by excessive contamination of secondary metabolites (Sahu *et al*., 2012). Therefore, extraction of large quantities of high-quality DNA from plant tissues can be difficult in some species due to the presence of large amounts of phenolic compounds, high levels of DNases and large amounts of organelle DNA (Varma *et al.,* 2007). Young and fresh plant tissues are commonly used as sources of high molecular weight genomic DNA as they contain small amounts of secondary metabolites, less polysaccharides and is easier to grind and isolate DNA from compared to older plant tissues (Abu Almakarem *et al*., 2012; Nickrent, 1994). It is possible to use fresh plant material when the laboratory is close to the research sites, the greenhouse, or the growth chamber. However, using fresh tissue samples are often not practical because the collection sites are located far away from the laboratory, leading to the need for preservation of collected samples and transport to the laboratory where the DNA extraction will be performed. Degradation of genomic DNA and other biochemical processes begin immediately after the tissue is removed from the plant (Abu Almakarem *et al*., 2012). Fresh, dehydrated, or lyophilized tissues, as well as tissues stored in the liquid nitrogen, are mainly preferred to avoid DNA degradation. Nevertheless, such sample processing is impossible to perform in many developing countries and at locations of many underutilized tropical plant species (Bressan *et al*., 2014). In these cases, plant tissue samples are usually preserved and stored at ambient temperature until brought back to the laboratory or cold storage, where already the DNA might be degraded. Therefore, many ecological and biodiversity studies performed in remote areas need different methods to preserve and store leaf materials, prior to molecular analyses to prevent DNA degradation (Bainard *et al.,* 2010).

Enset (*Ensete ventricosum*) is only cultivated in the native indigenous farming systems in Ethiopia (Brandt *et al*., 1997; Guzzon and Müller 2016). It is a staple food for over 20 to 35 million people through its starch rich main edible parts, the corm and the pseudostem (Borrell *et al*., 2019, Tuffa, 2019). Effective preservation methods are essential when sampling tissues from the remote farmers' fields for investigating molecular diversity or other molecular studies. Further, the leaf preservation methods should be compatible with the requirements of commercial and airline shippers (Chase and Hills, 1991). Dried leaf samples and rapid desiccation is a good alternative for storing herbarium samples in the field, however, preserving herbarium samples in silica gel or by drying has resulted in significant degradation of genomic DNA obtained from herbarium specimens (Chase and Hills, 1991; Doyle and Dickson, 1987; Liston *et al.,* 1990; Nickrent, 1994; Pyle and Adams, 1989). Hydrate leaf preservation methods such as ethanol and saturated NaCl-CTAB solution have also been tested on samples from various plant species and the results have been inconsistent (Abu Almakarem *et al*., 2012; Doyle and Dickson, 1987; Rogstad, 1992). Besides, poor leaf preservation method can lead to DNA degradation and co-precipitation of PCR inhibitors (Bainard *et al.,* 2010; Rogstad, 1992). However, the saturated NaCl-CTAB solution is an important method for the collection of plant samples from remote areas (Rogstad, 1992). This method has been proved to be significant for the preservation of DNA samples of taxonomically diverse species (Nickrent, 1994; Rogstad, 1992; Štorchová *et al*., 2000). The advantage of this method is that the dry ingredients, both NaCl and the CTAB chemicals, are easy to obtain and transport to remote areas in developing countries (Rogstad, 1992). However, few studies have been conducted using this method for leaf preservation and for studying its effects on the quality and quantity of genomic DNA. The quality and quantity of the DNA from the samples also depend on the duration of sampling and storage, and the plant species (Doyle and Dickson, 1987; Guo *et al.,* 2018).

Most of the DNA extraction protocols, recommends fresh leaf samples for genomic DNA isolation, but this is impossible when the samples are collected in remote geographical areas. These situations necessitate the development of appropriate protocols for leaf preservation and optimized methods for isolation of high-quality DNA from preserved or dried leaf samples. The main objectives of this study were: (1) to find an optimal method for preserving leaf tissues of enset stored at ambient temperature for various time spans; (2) to examine how the preservation methods affects both the quality and quantity of the genomic DNA; and (3) to study the effect of DNA extraction methods on DNA quality and quantity.

Materials and Methods

Plant samples and treatments

Enset leaf tissues used for this study were collected from enset plants grown in farmer fields in the Southern Nations, Nationalities and Peoples' Region (SNNPR) state in Ethiopia. The duration of the storage of collected samples varies from 7 to 35 days at ambient temperature (AT) (**Table 1**). Young fresh leaf material, free from visible fungal and insect damage, was selected from each enset genotype. The collected samples were divided into three parts and preserved using three different methods, i.e., 1) in saturated NaCl-CTAB solution (hexadecyltrimethylammonium bromide), 2) in indicator silica gel, and 3) in ethanol (96%). The saturated NaCl-CTAB solution was prepared following the original protocol (Rogstad, 1992),

with minor modifications. Briefly, 550 g NaCl was added to 1 L of tap water, boiled, cooled at ambient temperature, and mixed thoroughly until the salt precipitated. Then, 35 g of CTAB was added gradually with irregular intervals mixing, until the solution became viscous. 35-40 mL of the prepared solution was aliquoted into 50 mL falcon tubes and used for the preservation of tissue samples. A pair of scissors was used to remove leaf samples from the mother plants. The scissors were cleaned with ethanol (96%) between independent samples. Fresh cigar-leaf samples harvested from each enset genotype were stored immediately in the 50 mL tubes containing the saturated NaCl-CTAB preservation buffer. Samples were then placed in a black plastic bag and stored in a dark room at ambient temperature. The second preservation method used orange indicating silica gel (https://www.agmcontainer.com/920013). The silica gel was applied at 10:1 gram ratio for effective leaf preservation (Chase and Hills, 1991). The leaf tissue samples were placed in small tea bags, the tea bags were stapled, and the tea bags were transferred to individual plastic bags containing 50-60 gram of silica gel (Goldberg and Weintraub, 2001). The third preservation method used ethanol (96%). Approximately 2.5 gram of leaf samples were stored in 15 mL Falcon tubes containing 8 mL of 96% ethanol following the protocol (Bressan *et al*., 2014). All samples were stored at ambient temperatures during field collection and transportation from the remote farmer fields to the laboratory at the Norwegian University of Life Sciences, Ås, Norway.

Sample preparation for DNA extraction

Upon arrival at the laboratory, the minor modified saturated NaCl-CTAB solution was washed off thoroughly with deionized water and excess water wiped off the leaves with dry white wipes (Kimberly-Clark™ Professional Kimtech Science™) (**Figure 2b**). Leaf samples were put in liquid nitrogen, ground quickly using a pestle and mortar, and the ground powder transferred into 2 mL microcentrifuge tubes (Eppendorf A.G., Hamburg, Germany) (**Figure 3a**). Leaf samples preserved in silica gel were removed from the tea bags (**Figure 2c**) and pulverized using a pestle and mortar in liquid nitrogen (**Figure 3b**). Leaf samples preserved in ethanol were washed with deionized water, dried with white wipes (**Figure 2d**) and pulverized with a pestle and mortar in liquid nitrogen (**Figure 3c**). Pestles and mortars were washed and dried before starting each sample preparation, and all pulverized leaf samples stored at -80 ^{o}C until further analyses. For DNA extraction, 100 mg of pulverized leaf material of each sample preserved in NaCl-CTAB and ethanol, and 20 mg of each sample dried in silica gel were used.

DNA extraction and detections

Total genomic DNA was extracted from the preserved and dried leaf materials using two different DNA extraction methods. The DNeasy Plant Mini Kit (2016) was used following the manufacturer's instructions (QIAGEN, Hilden, Germany); the other method was the CTAB (cetyltrimethylammonium bromide) procedure (Kovi *et al.,* 2011; Murray and Thompson, 1980). However, the CTAB method did not show promising results for the preserved and dried enset leaf samples. The DNeasy Plant Mini Kit (2016) method was used for DNA extraction by varying the volume of the AE buffer (from 50 µL to 100 µL) and testing various amounts of tissue (50-100 mg) to find the optimum amount of starting material. After removal of the final AW2 washing buffer, the spin column was transferred into a new 2 ml collection tube and

centrifuged again for 1 minute at 20,000 x g to remove leftover AW2 washing buffer from the spin column, which can affect downstream applications. Finally, we used a volume of 50 µL AE elution buffer with 100 mg fresh preserved enset leaf samples (**Table 2**, **Figure 4**, Sample ID 1(d)). The DNA of some accessions were extracted from preserved samples after a few days, whereas DNA for others were extracted following longer preservation times up to nine months at -20 °C to compare the quality and quantity of DNA (**Table 1**).

Qualitative and quantitative analysis of extracted DNA

The concentration, purity (A260/A280 ratio), and absorbance ratio at 260–280 nm (A260/A230 ratio) were measured with a Thermo Scientific NanoDrop™ (ND) 1000 Spectrophotometer (Thermo Fisher Scientific, USA), using 1 µL of each sample. According to DNeasy® Plant handbook (2020), a purified DNA has an A260/A280 ratio of 1.7–1.9, indicating high purity of the DNA. The quality and integrity of total genomic DNA was detected using 1% (w/v) agarose gel electrophoresis. Finally, the DNA concentration was measured with the Qubit® dsDNA BR assay kit (Q) (**Table 5**).

Digestion with restriction enzyme

To check whether the genomic DNA extracted could be successfully digested with restriction endonucleases the DNA was digested with the restriction enzyme *EcoRI*. This is important for downstream applications like PCR and sequencing. The digested DNA was checked by electrophoresis on a 1% agarose gel (**Figure 8a, b**).

Results

Comparisons of leaf preservation methods

After few days of storage at ambient temperature there were clear differences between the leaf preservation methods as well as the physical appearance of the intact leaf samples (color) (**Figure 2**) and pulverized (powder texture) samples (**Figure 3**). Moreover, visible biological contamination and infection were not observed on the leaves preserved by all these three methods. Samples preserved with the minor modified saturated NaCl-CTAB solution maintained their original leaf color, i.e., the leaves remained green with no browning contrary to the other methods (**Figure 2b, 3a**). Leaf samples preserved in silica gel or ethanol did not maintain their visible physical characteristics except for a few samples (**Figure 2**). Samples stored in ethanol (96%) changed their color and formed a sticky and clay-mud like powder following pulverization (**Figure 3c**).

Comparison of DNA extraction methods

Total genomic DNA was extracted from preserved and dried leaves using the DNeasy plant mini kit and the CTAB extraction methods. Agarose gel electrophoresis and NanoDrop showed that total genomic DNA extracted using the DNeasy plant mini kit method gave DNA of high concentration and purity (**Figure 4-8**) (**Table 3**), whereas DNA extracted using the CTAB method showed comparatively lower concentration and purity (**Table 3**).

DNA quality and quantity from preserved and dried enset leaves

The DNA quality and quantity were measured from preserved and dried tissue using NanoDrop and Qubit. Overall, the DNA extracted from samples preserved in saturated NaCl-CTAB were of high quality with A260/A280 ratios ranging from 1.70 to 2.01 and concentrations of DNA ranging from 50.8 to 222.1 ng/µL (**Table 3**). The storage temperature (ambient or -20 ℃) had no effect on the quality and quantity of DNA from leaves preserved in saturated NaCl-CTAB (**Table 2-5**) (**Figure 5, 7, 8**). The DNA extracted from samples preserved with different methods appeared as distinct bands separated on the gel at their corresponding high molecular weights with little evidence of shearing and absence of RNA contamination (**Figure 5, 7**). However, DNA obtained from dried leaf samples preserved in silica gel and the hydrated leaf samples preserved in ethanol (96%) stored for 7 days at ambient temperature were highly degraded and not visible on the gel (**Figure 5, 7**). Thus, the DNA from these samples were not analyzed further. Furthermore, we performed digestion of the DNA with the restriction enzyme *EcoRI*, to further validate that the DNA extracted from leaves preserved using the saturated NaCl-CTAB method can be used in down-stream analysis like sequencing and molecular marker development (**Figure 8a, b**).

Discussion

One of the main advantages of the saturated NaCl-CTAB solution for preserving enset leaves is the use of common and inexpensive reagents, scalability, and its simplicity of usage. In addition, it was easy to transport samples in this solution from farmer fields in Ethiopia to Norway (Bhattacharjee *et al*., 2009; Rogstad, 1992). To the best of our knowledge, no other studies have been investigating the effects of different enset leaf preservation and DNA extraction methods on the quality, quantity and integrity of DNA extracted from the preserved and dried enset leaf material. In this study, we determined that DNA could be successfully extracted from enset leaves preserved using the minor modified saturated NaCl-CTAB solution for longer periods (over 35 days) at ambient temperature. The NaCl-CTAB solution preserved both the physical properties of the enset leaf samples and high quality and quantity of genomic DNA could be extracted after 7 to 35 days at ambient temperature and from samples stored more than nine months at -20 ℃ (**Figure 5, 7**) (**Table 4**). One of the reasons for slowing down the DNA degradation process using the NaCl-CTAB method is attributable to the high salt concentration, which partially dehydrate the leaf tissues. Furthermore, CTAB interacts with nucleic acids, proteins and carbohydrates to slow down the DNA degradation processes (Abu Almakarem *et al.,* 2012). Another reason that the saturated NaCl-CTAB solution might be more suitable for enset leaf preservation than silica gel is the fact that enset has hard leathery leaves with few stomata and high contents of water and fiber (Rogstad, 1992). In addition, the solution and the falcon tube protect the enset leaves from shaking and from physical damaged during field collection and transportation, which are important to minimize DNA degradation (Rogstad, 1992). Besides, visible contamination was not observed on the enset leaves and this is most likely due to the bactericidal and detergent properties of CTAB, as seen in previous studies, (Rogstad, 1992; Thomson, 2002). Further, the saturated NaCl-CTAB leaf preservation method also facilitates cleaning of leaves during the removal of the preservation solution before DNA extraction (Thomson, 2002). On the contrary, it has been shown that other plant species such as *Nardus stricta* L. (Poaceae), with thick cuticles, were not preserved well using

the saturated NaCI-CTAB preservation method (Štorchová *et al*., 2000). This indicates that the rate of degradation varies among plant species and leaf preservation methods (Chase and Hills, 1991; Rogstad, 1992; Štorchová *et al*., 2000; Till *et al.,* 2015). However, enset leaves stored in other chemicals such as ethanol (96%) did not preserve the DNA even after only seven days at ambient temperature (**Figure 5, 7**). This might be because ethanol did not facilitate and induce leaf lysis, cell wall disruption, and deactivation of DNAases during field collection and transportation (Akinnagbe *et al.,* 2011; York *et al.,* 1986). Similarly, Pyle and Adams, (1989) found that preservation of spinach leaves in 95% ethanol for as little as 24 h resulted in significant DNA degradation. On the contrary, Bressan *et al*., (2014) found that *Jatropha curcas* and other tropical species can be successfully preserved in ethanol for up to 30 days.

This study showed that the DNA quality of the samples preserved in silica gel were highly degraded compared to the samples preserved in the minor modified saturated NaCl-CTAB solution (**Figure 5, 7**). Most likely the preservation efficiency of silica gel is affected by the enset leaf secondary compounds or tissue characteristics (Savolainen *et al*., 1995), as most tropical plant species contain considerable amounts of secondary compounds (Abdel-Latif and Osman, 2017; Bressan *et al*., 2014). As pointed out, enset has hard leathery leaves with high contents of water and fiber, and the natural shape of the young cigar leaf. Another reason might be that enset is a highly drought tolerant species and will not easily desiccate by the silica gel (Adams *et al*., 1999; Borrell *et al*., 2019). Other studies have reported DNA degradation in other plant tissues preserved in silica gel due to the accumulation of phenolic compounds, which interfere with the quality of the isolated DNA (Akinnagbe *et al.,* 2011; Bainard *et al.,* 2010; Bhattacharjee *et al*., 2009). However, preservation in silica gel works in some plant species such as grasses and small herbs (Chase and Hills, 1991; Laulier *et al*., 1995; Margam *et al*., 2010). All this shows that different plant species require specific leaf preservation methods and the efficiency of each leaf preservation method can vary among and within plant species (Chase and Hills, 1991; Rogstad, 1992; Štorchová *et al.,* 2000).

Regarding the DNA extraction methods, the DNeasy Plant Mini Kit (Qiagen) is more suitable for large number of DNA samples with limited time compared to the CTAB method (Kotchoni and Gachomo, 2009; Margam *et al.,* 2010). In the present study, we used both solvent (CTAB) and solid (Qiagen kit) phase DNA extraction (SPE) methods (TOPÇU *et al.,* 2016). The total genomic DNA extracted using the DNeasy plant mini kit method gave DNA of higher concentrations and purities than DNA extracted using the CTAB method (**Table 3**). This indicates that most likely the CTAB method is better for young and fresh leaves than preserved leaf samples (Guo *et al.,* 2018). The second reason might be that the DNA extracted using the DNeasy Plant Mini Kit method most likely is free from secondary metabolites that interfere with the quality and quantity of DNA (Sika *et al.,* 2015). Most secondary compounds in plant tissues affect the disruption of tissues, interferes with the DNA extraction, reduce the DNA quality and inhibit subsequent molecular analyses like PCR and sequencing (Bressan *et al*., 2014; Varma *et al.,* 2007).

Overall, we obtained DNA of high quality and quantity when we used the NaCl-CTAB preservation method combined with DNeasy Plant Mini Kit (Qiagen) extraction method from leaf samples of enset stored for variable length of time and at various temperatures (ambient and -20 °C). We also tested that the DNA obtained was suitable for downstream applications (**Figure 8a, b**). Further, we observed little

contamination by polysaccharides and proteins as indicated by the ratios of A260/A230 and A260/A280 (**Table 2-5**). When we compared the effect of different leaf preservation methods, i.e., saturated NaCl-CTAB solution, silica gel and ethanol, we found that only the saturated NaCl-CTAB solution had slight degradation after 7 to 35 days of storage at ambient temperature, whereas the other two methods showed significant degradation after the same storage time (**Figure 5, 7**). The difference between these three leaf preservation methods became clear after seven days storage times at ambient temperature (**Figure 5, 7**). This indicates that saturated NaCl-CTAB solution can preserve enset DNA without causing severe degradation. The spectrophotometric measurements with NanoDrop tend to give higher readings of DNA quality and quantity of DNA from tissue preserved in silica gel and ethanol (**Table 4**), most likely due to contamination with the molecules absorbed at 260 nm or the interferences of proteins, and the NanoDrop reading degraded DNA (Ponti *et al*., 2018). Moreover, DNA purity can be severely affected by various components of sample matrices such as polysaccharides, lipids, and polyphenols or extraction chemicals like CTAB (Vinson *et al.,* 2018).

Conclusions

An efficient leaf preservation and DNA extraction method for enset leaf material is described. Based on the results, the minor modified saturated NaCl-CTAB leaf preservation method was found to be a better field preservation method for maintaining freshness, and integrity, quality and quantity of DNA of enset samples than preservation in silica gel or 96% ethanol. Further, this method makes the transportation of the samples from remote areas easy. The method consistently produces high yield and high-quality genomic DNA of enset at an affordable cost. Also, we found that the DNeasy plant mini kit approach performed better in extracting high quality and quantity of enset genomic DNA than the CTAB method. The high-quality genomic DNA extracted using this method, was used for further downstream applications including PCR and sequencing. Therefore, our results provided useful suggestions for preservation methods and DNA extraction methods.

Acknowledgement

The authors thank the Norwegian Agency for Development Cooperation (NORAD) for funding this study, which is part of the project 'Controlling Disease in Sweet potato and Enset in South Sudan and Ethiopia to Improve Productivity and Livelihoods under Changing Climatic Conditions using Modern Technologies' under the NORHED program (Agreement no. ETH-13/0017).

Conflict of interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Abdel-Latif, A., Osman, G. 2017. Comparison of three genomic DNA extraction methods to obtain high DNA quality from maize. *Plant methods*, *13*, 1-9. https://doi.org/DOI 10.1186/s13007-016-0152-4

Abu Almakarem, A. S., Heilman, K. L., Conger, H. L., Shtarkman, Y. M., Rogers, S. O. (2012). Extraction of DNA from plant and fungus tissues in situ. BMC Research Notes, 5(1), 1-11. https://doi.org/doi:10.1186/1756-0500-5-266

Adams, R. P., Zhong, M., Fei, Y. 1999. Preservation of DNA in plant specimens: inactivation and re-activation of Dnases in field specimens. *Molecular Ecology*, *8*(4), 681-683. https://doi.org/10.1046/j.1365-294X.1999.84600.x

Akinnagbe, A., Ji, X. B., Yang, M., Ewald, D. 2011. Ethanol pretreatment increases DNA yields from dried tree foliage. *Conservation Genetics Resources*, *3*(3), 409-411. https://doi.org/DOI:10.1007/s12686-010-9367-2

Bainard, L., Klironomos, J., Hart, M. 2010. Differential effect of sample preservation methods on plant and arbuscular mycorrhizal fungal DNA. *Journal of Microbiological Methods*, *82*(2), 124-130. https://doi.org/10.1016/j.mimet.2010.05.001

Bhattacharjee, R., Ferguson, M., Gedil, M., Dumet, D., Ingelbrecht, I. 2009. Field collection, preservation and large scale DNA extraction procedures for cassava (Manihot esculenta Crantz.). *African journal of biotechnology*, *8*(15).

Borrell, J. S., Biswas, M. K., Goodwin, M., Blomme, G., Schwarzacher, T., Heslop-Harrison, J. S. P., Wendawek, A. M., Berhanu, A., Kallow, S., Janssens, S., Molla, E. L., Davis, A. P., Woldeyes, F., Willis, K., Demissew, S., Wilkin, P. 2019. Enset in Ethiopia: a poorly characterized but resilient starch staple. *Annals of Botany*, *123*(5), 747-766. https://doi.org/10.1093/aob/mcy214

Brandt, S., Spring, A., Hiebsch, C., McCabe, J., Tabogie, E., Diro, M., Wolde-Michael, G., Yntiso, G., Shigeta, M., Tesfaye, S. 1997. *"The "Tree Against Hunger": Enset-based agricultural systems in Ethiopia"*. American Association for the Advancement of Science,Washington, DC, USA.

Bressan, E. A., Rossi, M. L., Gerald, L. T., Figueira, A. 2014. Extraction of high-quality DNA from ethanol-preserved tropical plant tissues. *BMC Research Notes*, *7*(1), 268. https://doi.org/10.1186/1756-0500-7-268

Chase, M. W., Hills, H. H. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon*, *40*(2), 215-220. https://doi.org/10.2307/1222975

Doyle, J. J., Dickson, E. E. 1987. Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, *36*(4), 715-722. https://doi.org/10.2307/1221122

Goldberg, L., Weintraub, S. 2001. Regulations Change for Cobalt Indicating Silica Gel. *AIC News*, *26*(1), 14.

Guo, Y., Yang, G., Chen, Y., Li, D., Guo, Z. 2018. A comparison of different methods for preserving plant molecular materials and the effect of degraded DNA on ddRAD sequencing. *Plant Diversity*, *40*, 106-116. https://doi.org/10.1016/j.pld.2018.04.001

Guzzon, F., Müller, J. V. 2016. Current availability of seed material of enset (Ensete ventricosum, Musaceae) and its Sub-Saharan wild relatives. *Genetic Resources and Crop Evolution*, *63*(2), 185-191. https://doi.org/10.1007/s10722-015-0355-y

Harris, S. A. 1993. DNA analysis of tropical plant species: an assessment of different drying methods. *Plant Systematics and Evolution*, *188*(1-2), 57-64.

Kotchoni, S. O., Gachomo, E. W. 2009. A rapid and hazardous reagent free protocol for genomic DNA extraction suitable for genetic studies in plants. *Molecular Biology Reports*, *36*(6), 1633-1636. https://doi.org/10.1007/s11033-008-9362-9

Kovi, M. R., Zhang, Y., Yu, S., Yang, G., Yan, W., Xing, Y. 2011. Candidacy of a chitin-inducible gibberellin-responsive gene for a major locus affecting plant height in rice that is closely linked to Green Revolution gene sd1. *Theoretical applied genetics*, *123*(5), 705-714. https://doi.org/10.1007/s00122-011-1620-x

Laulier, M., Pradier, E., Bigot, Y., Periquet, G. 1995. An easy method for preserving nucleic acids in field samples for later molecular and genetic studies without refrigerating. *Journal of Evolutionary Biology*, *8*(5), 657-663. https://doi.org/10.1046/j.1420-9101.1995.8050657.x

Liston, A., Rieseberg, L. H., Adams, R. P., Do, N., Ge-lin, Z. 1990. A method for collecting dried plant specimens for DNA and isozyme analyses, and the results of a field test in Xinjiang, China. *Annals of the Missouri Botanical Garden*, *77*(4), 859-863.

Margam, V. M., Gachomo, E. W., Shukle, J. H., Ariyo, O. O., Seufferheld, M. J., Kotchoni, S. O. 2010. A simplified arthropod genomic-DNA extraction protocol for polymerase chain reaction (PCR)-based specimen identification through barcoding. *Molecular biology reports*, *37*(7), 3631-3635. https://link.springer.com/content/pdf/10.1007%2Fs11033-010-0014-5.pdf

Murray, M., Thompson, WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic acids research*, *8*(19), 4321-4326. https://doi.org/10.1093/nar/8.19.4321

Nickrent, D. L. 1994. From field to film: rapid sequencing methods for field-collected plant species. *Biotechniques*, *16*(3), 470-475.

Pervaiz, Z., Turi, N., Khaliq, I., Rabbani, M., Malik, S. 2011. Methodology: a modified method for high-quality DNA extraction for molecular analysis in cereal plants. *Genetics and molecular research*, *10*(3), 1669-1673. https://doi.org/10.4238/vol10-3gmr1346

Ponti, G., Maccaferri, M., Manfredini, M., Kaleci, S., Mandrioli, M., Pellacani, G., Ozben, T., Depenni, R., Bianchi, G., Pirola, G. M. 2018. The value of fluorimetry (Qubit) and spectrophotometry (NanoDrop) in the quantification of cell-free DNA (cfDNA) in malignant melanoma and prostate cancer patients. *Clinica Chimica Acta*, *479*, 14-19. https://doi.org/10.1016/j.cca.2018.01.007

Pyle, M. M., Adams, R. P. 1989. In situ preservation of DNA in plant specimens. *Taxon*, 576-581. https://doi.org/doi.org/10.2307/1222632

Rogstad, S. H. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon*, *41*, 701-708. https://doi.org/10.2307/1222395

Sahu, S. K., Thangaraj, M., Kathiresan, K. N. 2012. DNA extraction protocol for plants with high levels of secondary metabolites and polysaccharides without using liquid nitrogen and phenol. *International Scholarly Research Notices*, *2012*. https://doi.org/10.5402/2012/205049

Sangwan, N., Sangwan, R., Kumar, S. 1998. Isolation of genomic DNA from the antimalarial plant Artemisia annua. *Indian Academy of Sciences*, *16*(4), 365-365. https://doi.org/10.1023/A:1007545102889

Savolainen, V., Cuénoud, P. S. R. Martinez, MDP Crèvecoeur, M., and Manen, J.-F. 1995. The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. *Plant Systematics and Evolution*, *197*(1-4), 87-98. https://doi.org/10.1007/BF00984634

Sika, K. C., Kefela, T., Adoukonou-Sagbadja, H., Ahoton, L., Saidou, A., Baba-Moussa, L., Baptiste, L. J., Kotconi, S. O., Gachomo, E. W. 2015. A simple and efficient genomic DNA extraction protocol for large scale genetic analyses of plant biological systems. *Plant Gene*, *1*, 43-45. https://doi.org/10.1016/j.plgene.2015.03.001

Sönmezoğlu, Ö. A., Terzi, B. 2019. Comparison of DNA extraction protocols for PCR-based techniques in wheat. *vrupa Bilim ve Teknoloji Dergisi*(17), 860-865. https://doi.org/10.31590/ejosat.646491

Štorchová, H., Hrdličková, R., Chrtek Jr, J., Tetera, M., Fitze, D., Fehrer, J. 2000. An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *Taxon*, *49*, 79-84. https://doi.org/10.2307/1223934

Thomson, J. A. 2002. An improved non-cryogenic transport and storage preservative facilitating DNA extraction from 'difficult'plants collected at remote sites. *Telopea*, *9*(4), 755-760. https://doi.org/10.7751/telopea20024013

Till, B. J., Jankowicz-Cieslak, J., Huynh, O. A., Beshir, M. M., Laport, R. G., Hofinger, B. J., Till, B. J., Jankowicz-Cieslak, J., Huynh, O. A., Beshir, M. M. 2015. Sample collection and storage. In *Low-Cost Methods for Molecular Characterization of Mutant Plants* (pp. 9-11). Springer, Cham. https://doi.org/10.1007/978-3-319-16259-1_3

TOPÇU, A., AŞIR, S., TÜRKMEN, D. 2016. DNA Purification by Solid Phase Extraction SPE Methods. https://doi.org/DOI:10.15671/HJBC.20164420568

Tuffa, A. C. 2019. Value chain and nutritional analyses of Warqe food products in relation to post-harvest losses. *Swedish University of Agricultural Sciences, Uppsala, Sweden.*

Vaillancourt, B., Buell, C. R. 2019. High molecular weight DNA isolation method from diverse plant species for use with Oxford Nanopore sequencing. *BioRxiv*, 783159. https://doi.org/doi.org/10.1101/783159

Varma, A., Padh, H., Shrivastava, N. 2007. Plant genomic DNA isolation: an art or a science. *Biotechnology Journal*, *2*(3), 386-392. https://doi.org/10.1002/biot.200600195

Vinson, C., Mangaravite, E., Sebbenn, A., Lander, T. 2018. Using molecular markers to investigate genetic diversity, mating system and gene flow of Neotropical trees. *Brazilian Journal of Botany*, *41*(2), 481-496. https://doi.org/10.1007/s40415-018-0472-x

York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., Albersheim, P. 1986. Isolation and characterization of plant cell walls and cell wall components. *Methods in Enzymology*, *118*, 3-40. https://doi.org/10.1016/0076-6879(86)18062-1

Tables

Table 1. Enset *(Ensete ventricosum* (Welw.) Cheesman*)* sample location and leaf preservation information

Accessions name	Cultivation region	No. in Figure 1-8	Storage method	Storage time (days)	
				AT	-20 ℃ with NaCl-CTAB
Tsela	South Omo	1	NaCl-CTAB solution	21	14
Ado	Sidama	2	NaCl-CTAB solution and silica gel	7	-
Chacho	Sidama	3	NaCl-CTAB solution and silica gel	7	-
Dado	Sidama	4	NaCl-CTAB solution and silica gel	7	-
Borbancho	Sidama	5	NaCl-CTAB solution	7	-
Bolanicho	Sidama	6	NaCl-CTAB solution	7	-
Chacho	Sidama	7	NaCl-CTAB solution	14	30
Wild	South Omo	8	NaCl-CTAB solution	14	30
Wild	South Omo	9	NaCl-CTAB solution	21	60
Mcho	Sidama	10	NaCl-CTAB solution	35	>270
Tunakecho	Sidama	11	NaCl-CTAB solution	35	>270
Kerta	South Omo	12	Ethanol (96%)	7	-
Golaa	South Omo	13	Ethanol (96%)	7	-
Entada	South Omo	14	Ethanol (96%)	7	-
Entada	South Omo	15	Ethanol (%96)	7	-
Wusasi	South Omo	16	NaCl-CTAB solution	21	14
Golaa	South Omo	17	NaCl-CTAB solution	21	14
Buuka	South Omo	18	NaCl-CTAB solution	21	14
Zokuma	South Omo	19	NaCl-CTAB solution	21	14
Dusak	South Omo	20	NaCl-CTAB solution	21	14
Siknda	South Omo	21	NaCl-CTAB solution	21	14
Mono	South Omo	22	NaCl-CTAB solution	21	14
Solka	South Omo	23	NaCl-CTAB solution	21	14
Entada	South Omo	24	NaCl-CTAB solution and Ethanol (96%)	7	-
Entada	South Omo	25	NaCl-CTAB solution and Ethanol (96%)	7	-

[AT] stands for ambient temperature

Table 2. Quantity and quality of genomic DNA obtained using different amounts of tissue and elution buffer volumes, quantified by NanoDrop

Sample ID	Weight (mg)	Volume of elution buffer (µL)	DNA quantity (ng/µL)	A260/280	A260/230
1(a)	125	100	55.00	1.78	2.33
1(b)	100	100	55.90	1.77	2.26
1(c)	100	75	62.80	1.78	2.39
1(d)	100	50	103.10	1.78	2.02
1(e)	80	80	42.50	1.79	2.70

[1] sample ID and accession name corresponding to number is shown in Table 1

Table 3. Comparison of DNA quantity and quality of samples extracted with different extraction methods, quantified by NanoDrop (ND)

Extraction methods	Storage methods	DNA quantity (ng/µL)	A260/280	A260/230
Minor modified DNeasy plant mini kit	Modified saturated NaCl-CTAB	50.80 – 222.10	1.70 - 2.01	2.09 - 2.57
CTAB method	Modified saturated NaCl-CTAB	7.50 - 9.70	1.99 - 2.47	-0.67 - 7.93

Table 4. Control of DNA quantity and quality among different leaf preservation methods and seven days stored at ambient temperature, quantified by NanoDrop (ND)

Sample ID	Storage methods	DNA quantity (ng/µL)	A260/280	A260/230
2s	Silica Gel (orange color)	87.70	1.82	2.31
3s		72.77	1.76	1.94
4s		66.02	1.87	1.98
2c	Saturated NaCl-CTAB	72.98	1.73	2.02
3c		43.68	1.70	2.04
4c		72.80	1.70	1.79
24e	Ethanol (96%)	32.37	1.72	2.04
25e		58.92	1.79	2.08
24c	Saturated NaCl-CTAB	97.39	1.78	2.27
25c		54.62	1.71	2.32

[S] indicate silica gel, [C] indicate saturated NaCl-CTAB and [e] indicate ethanol (96%) leaf preservation method and numbers correspond to accessions names shown in Table 1

Table 5. Quantity and quality of the total genomic DNA extracted with DNeasy Plant Mini Kit and preserved in the minor modified NaCl-CTAB solution quantified by NanoDrop (ND) and Qubit (Q)

	DNA quantity		DNA quality	
Sample ID	Q (µg/ml)	ND (ng/µL)	A260/280	A260/230
16	0.56	71.27	1.77	2.46
17	0.36	54.02	1.76	2.35
18	0.45	59.00	1.79	2.39
19	0.66	98.08	1.81	2.41
20	0.75	89.93	1.79	2.18
21	0.57	64.80	1.82	2.54
22	0.38	57.48	1.82	2.17
23	1.21	106.70	1.79	2.23

Figures

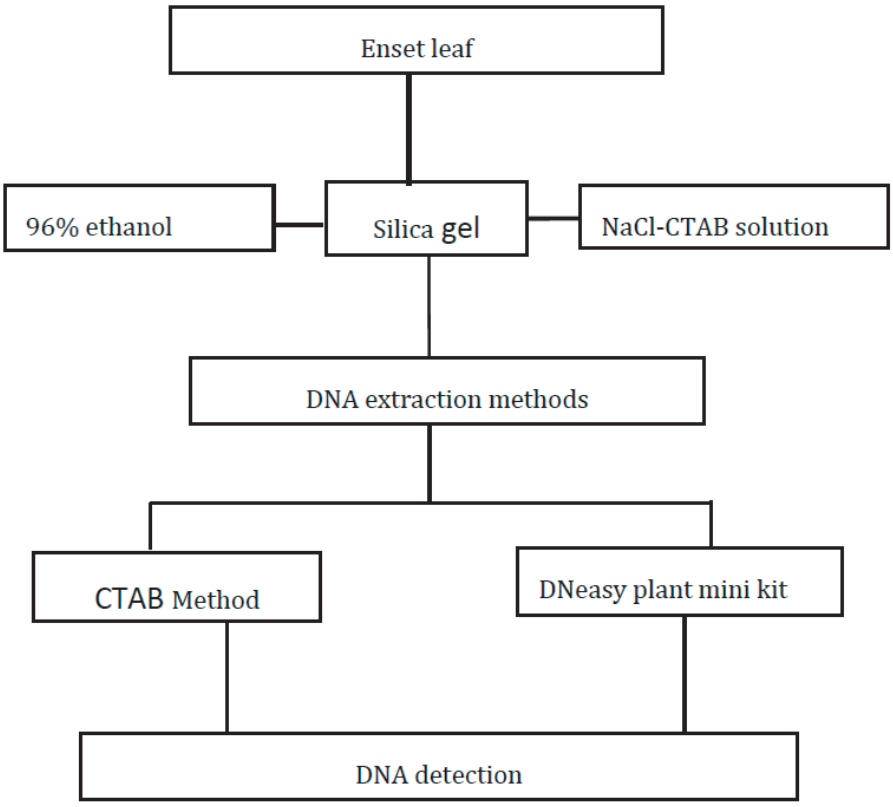


Figure 1. *Flowchart of workflow design*

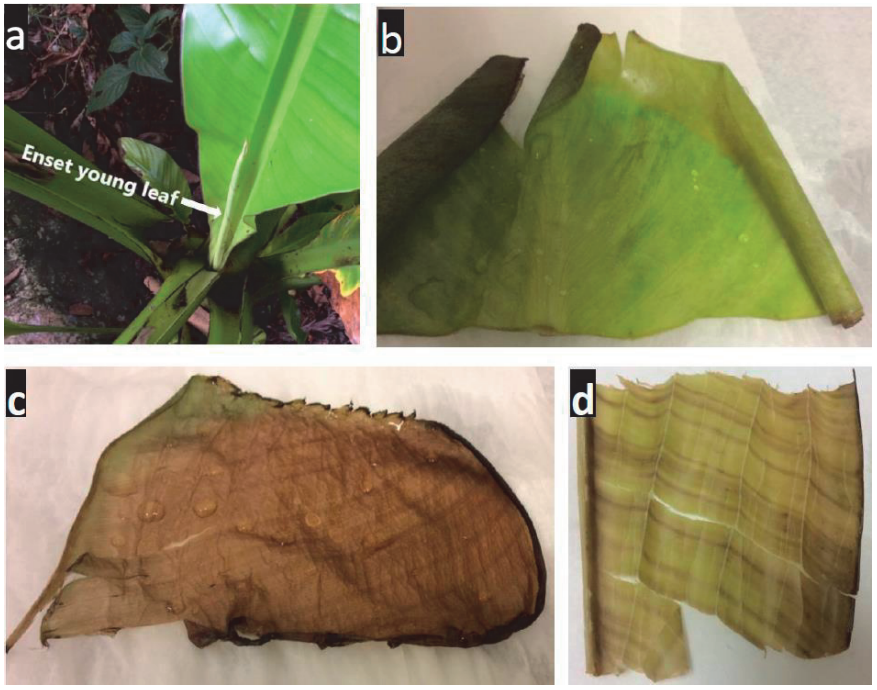


Figure 2. *Enset plant and preserved enset leaves after seven days of storage at ambient temperature. The external structure of the enset plant and enset young leaf used for preservation and DNA extraction (**a**) ; preserved in saturated NaCl-CTAB solution (**b**), preserved in silica gel (**c**), preserved in ethanol (96%) (**d**)*

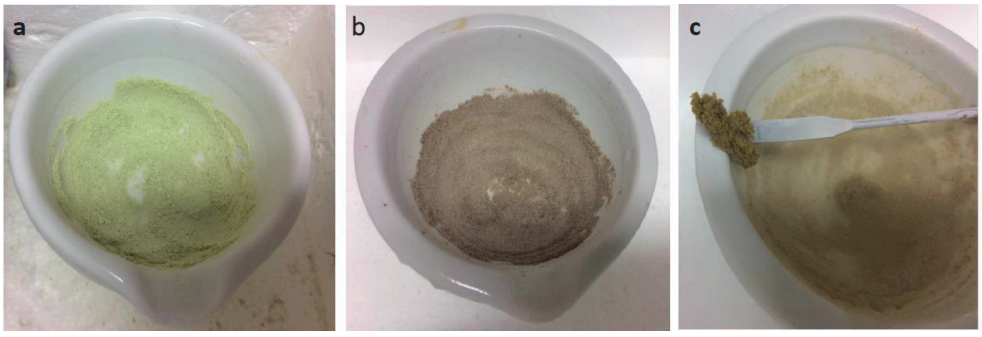


Figure 3. *Pulverized enset leaf samples after seven days of storage at ambient temperature. Preserved in saturated NaCl-CTAB solution (**a**), silica gel (**b**) and ethanol (96%) (**c**)*

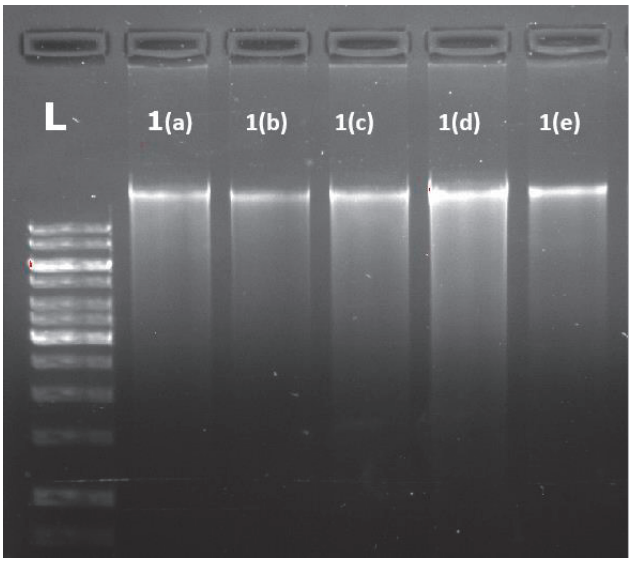


Figure 4. *Agarose gel electrophoresis of total genomic DNA extracted with the DNeasy Plant Mini Kit. Lanes 1(a) to 1(e) correspond to the samples in Table 2 with different combinations of amount of leaf material and elution buffer (AE), i.e. 1(a) 125 mg/100 µL, 1(b) 100mg/100 µL, 1(c) 100mg/75µL, 1(d) 100mg/50µL, 1(e) 80mg/80µL. L, 1kb, DNA molecular weight ladder (Thermo). Accession that corresponds to the number is shown in Table 1*

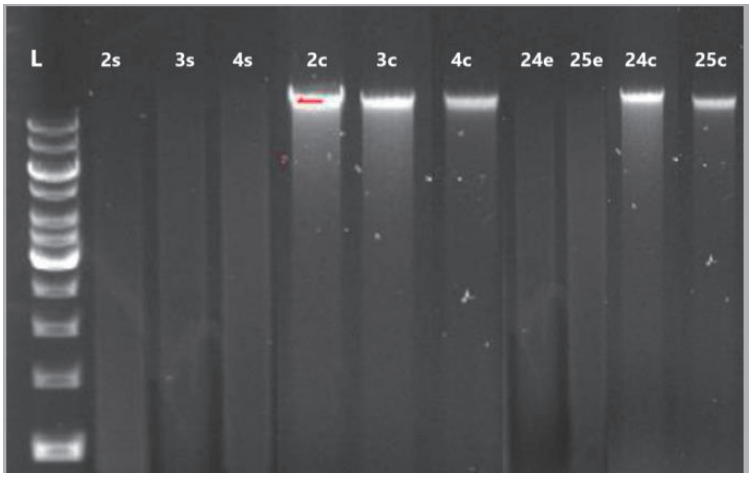


Figure 5. *Agarose gel (1%) electrophoresis of total genomic DNA extracted from preserved leaf by the DNeasy Plant Mini Kit following different leaf preservation methods but the same genotype and the same storage period (seven days) at ambient temperature (AT). Samples are stained with 1 µL Red safe; L: 1kb DNA molecular weight ladder (Thermo), s: silica gel, c: NaCl-CTAB solution and e: ethanol (96%). Accession names that correspond to the numbers are described in Table 1*

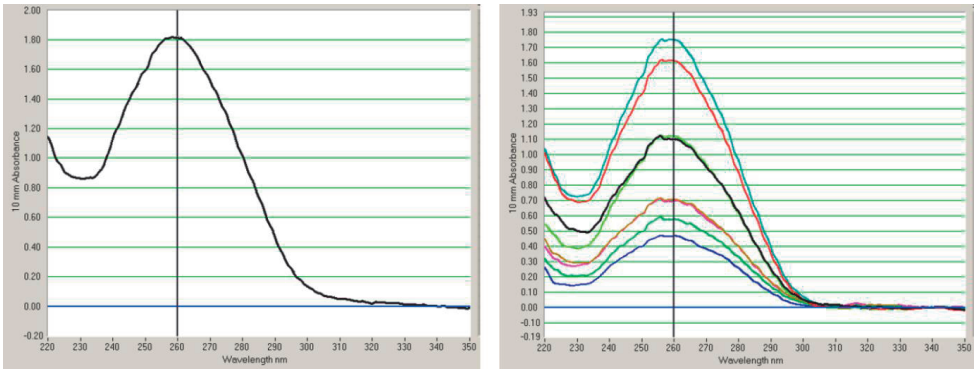


Figure 6. *Nano-Drop measurement profile of genomic DNA extractions from preserved leaves in saturated NaCl-CTAB solution. DNA extracted by using DNeasy Plant Mini Kit method. Scanned on NanoDrop from 220 to 350 nm*

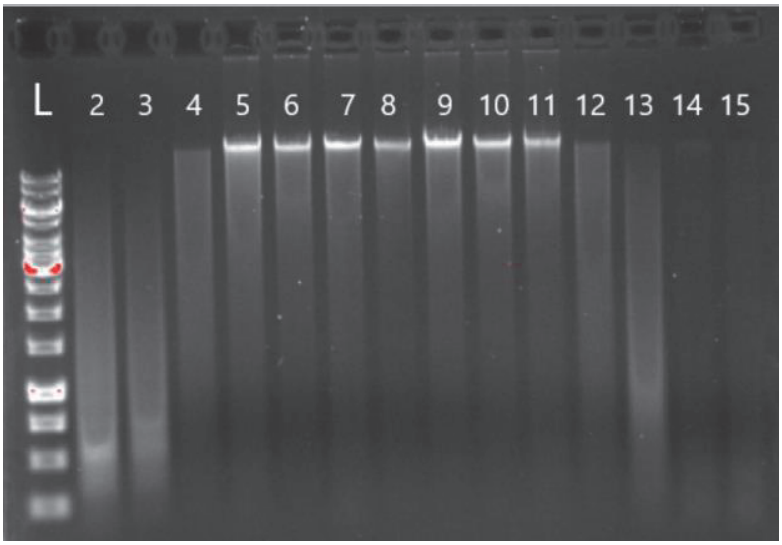


Figure 7. *Agarose gel (1%) electrophoresis of total genomic DNA extracted from preserved leaves by the DNeasy Plant Mini Kit following different duration of preservation. Lanes 2-4: Silica gel preserved leaf after seven days at ambient temperature (AT), Lanes 5-6: 7 days preserved in NaCl-CTAB at AT, Lanes 7-8: preserved for 14 days at AT and stored 30 days at -20 ℃ in NaCl-CTAB, Lane 9: preserved for 21 days at AT and stored 60 days at -20 ℃ in NaCl-CTAB, Lanes 10-11: preserved for 35 days at AT and stored >270 days -20 ℃ in NaCl-CTAB, Lanes 12-15: ethanol (96%) preserved leaf after 7 days AT, L: 1kb, DNA molecular weight ladder (Thermo). Samples are stained with 1 µL Red safe. Accessions corresponding to the numbers (2-15) are described in Table 1.*

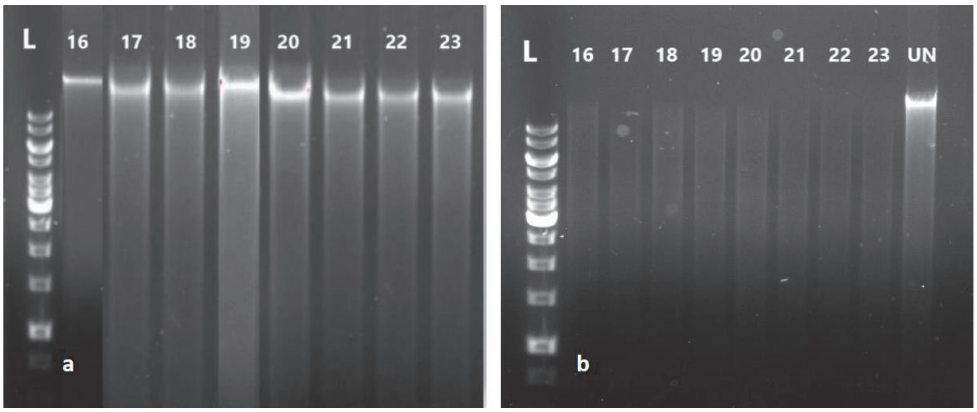


Figure 8. *Restriction uncut and cut enset leaf DNA from the minor modified saturated NaCl-CTAB leaf preservation method and the DNA yield quantified by NanoDrop (ND) and Qubit (Q) (Table 5). Total genomic DNA uncut with restriction EcoRI enzymes (**a**) and the same DNA cut with EcoRI enzyme (**b**). UN, uncut enset leaf total genomic DNA. L, 1 kb, DNA molecular weight ladder (Thermo). Name that corresponds to the numbers are shown in Table 1.*

Paper II

Genetic Diversity, Population Structure and Selection Signatures in Enset (*Ensete ventricosum*, (Welw.) Cheesman), an Underutilized and Key Food Security Crop in Ethiopia

Alye Tefera Haile[1, 2*], Mallikarjuna Rao Kovi[1], Sylvia Sagen Johnsen[1], Bizuayehu Tesfaye[2], Trine Hvoslef-Eide[1], Odd Arne Rognli[1*]

[1]Faculty of Biosciences, Department of Plant Sciences, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway

[2]School of Plant and Horticultural Science, College of Agriculture, Hawassa University, P.O. Box 05, Awassa, Ethiopia

* Corresponding Author's E-mail: odd-arne.rognli@nmbu.no

Abstract

Enset (*Ensete ventricosum*) is a multipurpose, drought-tolerant and a key food security crop, which is the staple food for peoples in the South and Southwestern parts of Ethiopia. Despite its importance, enset is an orphan crop as regards genetic research and breeding. In this study, we characterized genetic diversity, population structure and selection signatures in 226 cultivated and 10 wild enset accessions collected from diverse enset growing regions of Ethiopia using 3505 high-quality SNP markers obtained from ddRAD-sequences. The population structure and cluster analyses clearly distinguished between cultivated and wild enset. AMOVA revealed much higher levels of genetic variation within populations and regions (91.2 and 92.4%, respectively) than between populations and regions (8.8 and 7.6%, respectively). This shows that the region of origin and environmental heterogeneity have little influence on the genetic variation. However, the genetic differentiation between regions was moderate to large (F_{ST} = 0.06–0.17). The genetic structure of enset was mainly shaped by eco-geographic factors, mode of propagation and cultivation status. Six genes potentially involved in sexual reproduction and flowering signalling, which are key processes underlying domestication and adaptation, were under positive selection demonstrating that sexual reproduction plays an important role in shaping enset diversity. A lot of unexplored diversity is available for improving enset in Ethiopia, with patterns of diversity consistent with divergent selection on adaptive traits. This diversity also shows potential for introducing enset as a more food secure crop for the food insecure regions in the dry north of Ethiopia.

Keywords:

Ensete ventricosum, ddRAD, population structure, SNPs, food security

Introduction

Enset (*Ensete ventricosum*), commonly known as false banana or Abyssinian banana, is a perennial diploid (2n=18), monocarpic species belonging to the family Musaceae in the genus *Ensete (*Westphal *et al*., 1975*)*. Enset, banana and plantains are the most important cultivated members of the family; all with high global and local economic as well as food security importance (Baker and Simmonds 1953; Simmonds, 1962). The genus *Ensete* consists of three species with extensive geographical distribution, *E. ventricosum* and *E. livingstonianum* in Africa and *E. glaucum* in Asia, and five other localized endemics or near-endemic species (Borrell *et al*., 2019). Although wild species of *Ensete* are distributed throughout Central, Eastern, and Southern Africa, as well as Asia, enset (*Ensete ventricosum*) is the only cultivated species of the genus *Ensete* and its domestication and utilization as a food and fiber crop is so far restricted to Ethiopia. Anthropologists, archaeologists, historians and other scholars have argued that enset was domesticated in Ethiopia as early as 10,000 years ago (Brandt *et al*.,1997). During his travels in the 18th century, Bruce (1790) stated that enset was naturally produced and grown in every part of Abyssinia (now Ethiopia) if provided with enough heat and water. This indicates that enset is one of the early cultivated crops in Ethiopia. Enset is highly drought tolerant with a wide agroecological distribution and is cultivated only with household-produced inputs (Brandt *et al*., 1997; Tsegaye and Struik, 2002). It is unknown whether its wide distribution across a range of altitudes involves genetic or phenotypic adaptation (Tsegaye, 2002).

Eighty percent of the enset production is concentrated in the southern and southwestern part of Ethiopia (Bezuneh *et al*., 1967). The crop represents 65% of the total crop production in the southern regions of the country and serves as staple and co-staple food for about 20 to 35% of the Ethiopian population, largely in the south and southwestern part of the country (Borrell *et al*., 2020; Spring *et al*., 1996). Furthermore, it is used for several other purposes, such as animal feed, fiber, construction material and in traditional medicine. This became evident through the great famine in Ethiopia in the years 1888 to 1892 (Tobiaw and Bekele, 2011), and is the reason why enset is called "*The Tree Against Hunger*" (Brandt *et al*., 1997; Costa and Lockhart, 1984),1984). The crop best grows at cooler, higher altitudes and is found mostly between 1200–3100 m above sea level (Brandt *et al.,* 1997). Enset plants grow 4-8 m, sometimes up to 11 m height. Cultivated enset are propagated vegetatively, while wild enset reproduces through seeds (Birmeta *et al.,* 2004; Borrell *et al.,* 2019; Tsegaye and Struik, 2001). Harvest regularly occurs after 4 to 6 years after transplanting, but age at harvest varies from as early as three years up to twelve years after transplanting (Borrell *et al*., 2020). According to the 2019/20 main crop season survey, about 158 mill enset plants were harvested with a total yield of about 9.45 mill tons: making it one of the largest perennial food crops in the country (Central Statistical Agency, CSA, Government of Ethiopia, 2020). The cultivation of enset provides farmers and their families with security against hunger. Although the normal harvesting time is 4-6 years after planting, if other crops fail, they have their enset plants and can harvest any time. This is an important aspect of introducing enset to other, more food insecure regions in Ethiopia, particularly in the dry north.

Ethiopia is the center of origin of many plant species, including enset (Engels and Hawkes, 1991). The presence of wild and cultivated enset indicates that Ethiopia is the primary center of origin and center of

diversity (Purseglove, 1985; Vavilov, 1951). Ethnic groups in Ethiopia recognize and exploit various enset landraces. Regions in Ethiopia with diverse cultural history have rich biodiversity (Tsegaye, 2002). Enset-based farming system is a major agricultural system and farmers cultivate many enset landraces across various climatic and agroecological systems (Borrell *et al*., 2019). Several studies of genetic diversity of specific enset accessions from local regions using molecular markers such as Amplified Fragment Length Polymorphism (AFLP) (Negash *et al*., 2002; Tesfamicael *et al*., 2020), Random Amplified Polymorphic DNA (RAPD) (Birmeta *et al.,* 2004), Inter Simple Sequence Repeats (ISSR) (Tobiaw and Bekele, 2011), chloroplast DNA sequences (Bekele and Shigeta, 2011), Simple-Sequence Repeats (SSRs) (Gerura *et al*., 2019; Getachew *et al*., 2014; Olango *et al.,* 2015) and Single Nucleotide Polymorphism (SNPs) (Tesfamicael *et al.,* 2020) have been conducted and revealed genetic diversity among and within wild and cultivated enset accessions. The cultivated enset is vegetatively propagated, genetic divergence among clones may be minimal and could be difficult to detect using these marker types (McKey *et al*., 2010). Moreover, different molecular markers have different properties and will reveal different aspects of genetic diversity (Karp and Edward, 1997). The investigations mentioned above were conducted in certain enset growing areas in the southern and southwestern part of the country. Since Ethiopia is the center of diversity, many enset rich locations harboring large amounts of diversity of cultivated and wild enset are yet to be studied and is not represented in *ex situ* collections. Enset clones have traditionally been characterized phenotypically, however, phenotypic description is limited by the cost, time and space required to make visual observations and measurements (Hinze *et al.,* 2017).

Several new molecular techniques are now being applied together with phenotypic descriptions to investigate genetic diversity and relatedness in enset accessions (Birmeta *et al.,* 2004). Single nucleotide polymorphisms (SNPs) and simple sequence repeats (SSRs) are the most common DNA markers for genetic diversity studies (Tsykun *et al*., 2017). Among all DNA markers, SNPs are the most abundant and robust markers, and they are suitable for automated high-throughput genotyping of large numbers of samples. A range of assay options using different technology platforms to meet the demand for genetic studies and molecular breeding of crop plants are also available (Alkan *et al*., 2011; Bus *et al*., 2012; Hamilton *et al.,* 2011). The recently developed double-digest restriction-site associated DNA (ddRAD) technique is used extensively for population genetic studies in a wide range of non-model organisms (Andrews *et al*., 2016; Peterson *et al*., 2012). ddRAD sequencing provides a useful tool for marker-assisted breeding, genotype identification and determination of genomic organization and evolution in plants. It is a powerful and relatively cost-effective approach for developing numerous SNP markers and constructing high-density genetic maps (Peterson *et al.,* 2012). SNP markers are powerful tools for estimating genetic similarities and diversity. SNPs are able to resolve the differences among extremely similar individuals and increase the accuracy of diversity estimates (Hinze *et al.,* 2017).

Despite the abundance, diversity and ecological importance of enset, the species is not well characterized at the genomic level and has been far less studied than other cultivated species in the family Musaceae (Borrell *et al*., 2019). More detailed diversity studies of both cultivated and wild enset accessions in Ethiopia are needed to meet future needs, including diversification of crops in more vulnerable regions in

Ethiopia. Novel sources of genetic diversity need to be identified, characterized, incorporated into breeding programs, and utilized for the development of non-redundant core collections for conservation and breeding. In this study, SNP markers were developed and used to understand the population divergence of cultivated and wild enset. Understanding the genetic basis of enset domestication provides a valuable foundation for enset conservation and genetic improvement. The objectives of the present study were: (1) to evaluate the efficacy and suitability of SNP markers developed from ddRAD sequencing for high-throughput genotyping of enset; (2) to assess population structure, genetic diversity, and relationships among and within cultivated and wild enset accessions, and (3) to identify candidate genes potentially subjected to domestication and selection.

Materials and Methods

Sampling Area

The Southern Nations, Nationalities and Peoples' Region (SNNPR) state has a total area of 117,506 km2, with altitudes ranging from 378 to 4,201 meters above sea level (masl) (Abebe, 2005). Enset accessions were collected from three main enset culture communities, which are densely populated enset cultivating administrative regions (Sidama, Gurage and South Omo). The wild enset were collected around farms, along riversides and in deep forests. The three collection regions were deliberately chosen based on their enset production potential in SNNPR, where more than two-thirds of the country's enset production is located (Zeberga *et al*., 2014). We collected 226 cultivated and 10 wild enset accessions originating from different geographical locations and agroecological zones (**Table 1**; **Table S1**). The major ethnic regions cultivating enset and the study areas in Southern Ethiopia are show in **Figure 1**.

Preparation of NaCl-CTAB Preservation and Samples Collection

The saturated NaCl-CTAB solution was used to preserve the enset leaf samples upon collection, as described by Rogstad, (1992) with minor modifications. Briefly, 550 g NaCl was added to 1 L of water, boiled, and cooled at ambient temperature, and mixed thoroughly until the salt precipitated. Then, 35 g of CTAB was added gradually with intermittent irregular intervals mixing, until the solution became viscous. 35-40 mL of the prepared solution was aliquoted into 50 mL Falcon tubes and used for preservation of tissue samples. A pair of scissors was used to remove leaf samples from the mother plants, and the scissors were cleaned with ethanol (96%) between independent samples. Fresh cigar-leaf samples harvested from each enset accession were stored immediately in the 50 mL tubes containing the saturated NaCl-CTAB preservation solution. Samples were then placed in a black plastic bag and stored in a dark room at ambient temperature to preserve genomic DNA from degradation during transportation from the farmer fields in Ethiopia to the laboratory in Norway.

DNA Extraction

Young, healthy, and fresh leaves were stored in NaCL-CTAB solution for genomic DNA extraction. DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). DNA quality and quantity were determined using a NanoDrop spectrophotometer (Thermo Fisher, Inc.) and agarose gel electrophoresis

(1%). DNA concentrations were determined using the Qubit® dsDNA BR assay kit (Life Sciences) and Quant-iT™ PicoGreen™ (Life Sciences) dsDNA assay.

Double-digest Restriction-Site-Associated DNA (ddRAD) Library Preparation and Illumina Sequencing

We calculated the number of reads required for 20X coverage of restriction fragments in the 150–500 bp size range across 10 multiplexed individuals using multiple enzyme pairs, assuming 0.44 GC content, to ensure that restriction fragments could feasibly be sequenced with enough coverage on an Illumina MiSeq platform. The ddRAD procedure used in this study was modified from Peterson *et al. (*2012*)* (For further ddRAD information and the complete protocol, see Supplementary information, **Table 2-7; Table S1**). 500 ng of each DNA sample was double digested using *EcoR1* HF (the "rare cutter"– recognizes a six bases motif, i.e., 5′-GAATTC), and *MseI* (the "frequent cutter"– recognizes a four bases motif, i.e., 5′-TTAA) restriction endonucleases, and adapters ligated to the digested fragments. Each DNA sample with a unique P1 barcode, and a P2 barcode common for all samples. Samples containing unique P1 barcodes were pooled, and the Sage Science Blue Pippin system (www.sagescience.com) was used to select fragments of about 500 bp. Size-selected libraries was bound to Dynabeads® M-270 Streptavidin magnetic beads (Invitrogen), to eliminate fragments without the P2 adapter, and the libraries amplified by PCR using Phusion™ Polymerase kit (Invitrogen) and index-marked primers for further tagging of the samples. The libraries were analyzed using an Agilent 2100 Bioanalyzer and diluted to a concentration of 35 nM for paired-end sequencing using the V2 sequencing kit on the MiSeq platform (Illumina). The sequencing was performed at the Norwegian University of Life Sciences, Norway.

Sequence Data Analysis and SNP Calling

The GBS data obtained was quality checked using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). High quality reads were retained after trimming the bad quality reads using Trimmomatic program (Bolger *et al*., 2014). The raw paired-end sequence reads obtained from MiSeq were quality checked after removing the adapters and barcodes. The clean paired-end reads were used to call the SNPs using the STACKS 2 pipeline (Rochette *et al*., 2019). The SNPs were filtered based on the following criteria: (1) variant should be bi-allelic SNPs, (2) SNPs having more than 20% missing information were excluded, (3) genotypes having more than 20% missing information were excluded, and (4) markers with minor allele frequency (MAF) > 0.05 were retained.

Population Structure Analysis

Population groups were inferred using the fastSTRUCTURE software (Raj *et al.,* 2014). Twenty independent test runs were conducted allowing K to vary from 1 to 20. The optimal value of K for these runs was then determined using the ChooseK function. The script ChooseK, included with the fastSTRUCTURE package, was used to choose the number of subpopulations that maximize the marginal likelihood. The cluster membership matrices of the fastSTRUCTURE outputs were visualized using structure selector tool (Li *et*

al., 2018). Following the assignment of individuals to populations, the program package CLUMPACK (Kopelman *et al.,* 2015) was used to summarize the structure results into structure plots.

Genetic Diversity Analysis

For genetic diversity analysis, subpopulations were defined as the number of clusters produced by fastSTRUCTURE at K = 12. Genetic variation among and within populations, observed (H_O) and expected (H_E) heterozygosity, and pairwise fixation index (F_{ST}) for the subpopulations (Weir and Cockerham, 1984) was estimated by analysis of molecular variance (AMOVA) using Arlequin v.3.5 (Excoffier and Lischer 2010). Significance ($P<0.05$) of the F_{ST} values were estimated using 1023 permutations. F_{ST} results were interpreted using the same standard as in (Pino Del Carpio *et al.*, 2011; Hartl *et al.,* 1997; Wright, 1978).

Phylogenetic Trees and PCA Analyses

To examine the relationship between cultivated and wild enset accessions, Principal Component Analysis (PCA) was performed using TASSEL v5.2 (Bradbury *et al.,* 2007) and maximum-likelihood (ML) phylogenetic tree analyses performed using PhyML 3.0 (Guindon *et al.,* 2010). The trees were prepared and visualized using the iTOL v4 online tool (Letunic and Bork, 2019). PCAs were graphically summarized using scatter plots. Populations were named according to the passport data denoting geographical origin.

F_{ST} Outlier Tests for Detecting SNP Loci Under Selection

To detect loci under directional selection, we used the hierarchical method (Excoffier *et al*., 2009), a modified approach of Beaumont and Nichols (1996), implemented in the Arlequin software package version 3.5.1.3 (Excoffier and Lischer, 2010). We employed an hierarchical island model based on 2 groups (cultivated and wild enset) with 50,000 simulations to generate the joint distribution of F_{ST} versus heterozygosity. Loci that fall out of the 99% confidence intervals of the distribution were identified as outliers being putatively under selection. The putative function of genes with outlier SNPs was identified using the Gene Ontology (GO) annotation using the Blast2GO software tool version 3.0 (Conesa *et al*., 2005).

Results

SNP Discovery and Filtering

Following sequencing of the double digest RAD fragments, data processing and SNP filtering, SNPs with high heterozygosity (> 80%) were removed and a total number of 3505 high-quality SNPs were identified among the 236 enset accessions.

Genetic Structure

The genetic structure analysis using fastSTRUCTURE suggests that the most likely number of subpopulations is 12, i.e., the model complexity that maximizes likelihood is 12 (likelihood=-0.82) and the highest peak shows K = 12 as optimal (**Figure 2A**). The results of the fastSTRUCTURE analysis are shown in **Figure 2B**. The 10 wild enset accessions from South Omo make up a distinct group which is stable at all levels above K=7. The 62 cultivated ensets from Gurage separated from Sidama and South Omo accessions

and seems to make up a rather unique subpopulation, while the cultivated enset accessions from Sidama (72) and South Omo (92) represent many subpopulations.

PCA and Phylogenetic Relationships

Results from the principal component analysis (PCA) are presented in **Figure 3**. The PCA showed that some of the populations were more clearly separated while others were clustered more closely. The first three components described 20, 18 and 9% of the total variance, respectively. PC1, with some overlap, separates Gurage accessions from accessions of the other regions, while PC2 separates the South Omo accessions in two clusters, one of them overlapping with the Sidama cluster. However, the Sidama and South Omo accessions in this cluster are partly separated by PC3, with the wild accessions clustering, as expected, with the South Omo subcluster. The phylogenetic analysis grouped the enset accessions into different clusters, to a large degree reflecting geographical origins and cultivation status (**Figure 4**). The wild accessions formed a clearly distinguished clade from the cultivated enset accessions (**Figure 4B**). Generally, few accessions of cultivated enset tended to have longer branches (**Figure 4B**). Interestingly, twelve accessions (19.35%) collected from the Sidama region clustered into Gurage. However, surprisingly no Gurage accessions clustered into Sidama in this study. Besides, four accessions collected from Sidama and two from Gurage clustered into South Omo. Some accessions have the same names in different regions, e.g., Gena, Astara and Mazia, however, they are certainly different accessions since they cluster in different clades in the phylogenetic tree (**Figure 4**). The phylogenetic analysis confirms the results of the structure analysis; the most genetically unique accessions, apart from the wild accessions, are the Gurage accessions, while accessions from Sidama seems to have a mixed ancestry, whereas the South Omo accessions clearly represents two genetically diverse subgroups.

Genetic Diversity and Pairwise Population Differentiation

The results of the AMOVA analysis are presented in **Table 2**. Most of the genetic variation (91.2 %) is within the enset accessions, and very little (8.8%) between cultivated and wild accessions. Analysis of the genetic differentiation between the geographic regions showed that 92.4% of the genetic variation was within and only 7.6% between geographic regions (**Table 2**). Observed (H_O) heterozygosity was slightly higher in Sidama and the wild group (H_O= 0.33 and 0.32, respectively) than in Gurage and South Omo (both H_o=0.31), while expected (H_E) heterozygosity was slightly lower than observed in Sidama and Gurage and higher than observed in South Omo and the wild group (**Table 3)**. Generally, the molecular variation was highest within the wild accessions and lowest within cultivated enset from Gurage. Pairwise population differentiation (F_{ST}) showed, as expected, that the largest subpopulation division is between the wild accession group and the cultivated (moderate to large differentiation, F_{ST}=0.14-0.17), with the largest differentiation between the wild and the Gurage group. Between the cultivated enset groups, the largest differentiation is between Gurage and South Omo (F_{ST}=0.10), while differentiation between Sidama and the other two groups are smaller (F_{ST}=0.06-0.07) (**Table 4)**.

Signatures of Selection and Functional Analyses

Signatures of balancing and directional selection were identified at 35 loci among cultivated and wild accessions using the hierarchical method (Excoffier *et al*., 2009) (**Figure 5**). Putative balancing selection was detected at 23 loci and directional selection was detected at 12 loci (**Figure 5**). Among the 12 loci, six loci have putative gene functions, while the other six loci have unknown gene functions. Putative functions of these six loci are described in **Table 5**.

Discussion

Population Structure and Differentiation between Wild and Cultivated Enset

In this study, a high-throughput sequencing technology was used to explore genetic diversity, population structure, and selection signatures in cultivated and wild enset accessions collected across the center of origin and domestication in Ethiopia. The ancestral admixture and phylogenetic analyses showed a clear separation between wild and cultivated enset (**Figure 2B, 4**). Most probably this separation between wild and cultivated enset populations can be attributed to the difference in propagation methods (Birmeta *et al.,* 2004; Gerura, *et al*., 2019; Olango *et al.,* 2015; Tesfamicael *et al*., 2020). It is interesting to note that cultivated enset accessions collected from regions where wild enset grows showed higher admixture and weaker clustering than those collected from regions where wild enset does not grow. This could be due to higher enset diversity in that specific region with wild enset and indicates exchange of genetic material by crossing between cultivated and wild enset. Besides, the phylogenetic tree analysis showed that populations from adjacent regions like Sidama and Gurage formed a polyphyletic group, which was not the case with distantly located populations e.g., populations from Sidama and South Omo (**Figure 4B**). This genetic structure could be explained by a combination of genetic drift locally and the founder population. However, the analyses showed admixture of very few accessions irrespective of their origins whether the accessions were located isolated far apart or close like Sidama and Gurage (**Figure 4)**. Remarkably, some accessions collected from Sidama clustered with Gurage accessions, suggesting that these accessions are most possibly of Gurage origin. However, no Gurage accessions clustered with accessions from Sidama. In addition, four accessions from Sidama and two from Gurage clustered with South Omo accessions. Taken together, this indicates human sharing and exchange of some clonal materials among and within regions (Gerura *et al*., 2019; Getachew *et al*., 2014). As pointed out earlier, a few accessions have the same vernacular names in different regions, e.g., Gena, Astara and Mazia. However, they are certainly different accessions genetically based on their SNP profiles, and they have not been exchanged by humans even if they have the same vernacular name.

The phylogenetic tree showed long branches for the wild population from South Omo and for a few cultivated enset accessions too (**Figures 4A, 4B**), suggesting high rates of nucleotide substitution and consequently high diversity. Furthermore, the phylogenetic tree revealed a relatively close association between South Omo and Sidama enset populations (**Figures 4A, 4B**) and lowest F_{ST} value was found between these two populations (**Table 4**). Thus, showing that Sidama and South Omo populations have

close relationship with each other which might be due to possible vicariant evolutionary event from a single common ancestor through the fragmentation of their common ancestor's range or historical relationship (Schaal *et al*., 1998).

Values of the fixation index (F_{ST}) above 0.15 indicate significant differentiation between populations (Frankham *et al.,* 2002). In this study, we observed that significant divergence between enset populations. The wild population showed moderate to large genetic differentiation from the cultivated populations from the regions, while there was relatively small differentiation between the cultivated populations. Cultivated enset is only propagated vegetatively and farmers harvest enset before seed set, while wild enset are propagated exclusively by sexual reproduction (Birmeta *et al.,* 2004; Brandt *et al.,* 1997). As a result of this, gene flow between cultivated and wild enset is probably very limited. Besides, the natural distribution of wild enset, as well as the farming and management practices of cultivated accessions have an impact (Birmeta *et al.,* 2004; Olango *et al.,* 2015). Further, limited exchange of genetic material by humans or natural factors may be considered as the main reasons for the larger genetic differentiation observed between wild and cultivation populations (Birmeta *et al.,* 2004; Gerura *et al*., 2019; Tesfamicael *et al*., 2020).

Importantly, moderate genetic differentiation was found between wild and cultivated enset from South Omo. This might be due to the co-existence of wild and cultivated enset in the South Omo region, where farmers introduce wild accessions into the cultivation areas and hence genetic exchange occurs between cultivated and wild populations of enset in this particular region (Shigeta, 1992). In contrast, the highest F_{ST} value (0.17) was observed between wild enset and accessions from Gurage (**Table 4**). This shows that these accessions are more isolated from one another; most likely there is no wild enset growing in the Gurage region. Similar results can be seen from the population structure and phylogenetic analyses. The Gurage accessions are separated and formed a single cluster on their own far from the wild enset cluster (**Figure 3, 4**). Another reason is that Gurage maybe has a different cultural and ethnic origin. This indicates that there is unique genetic diversity within the Gurage accessions, which is not related to the geographical distance to the other regions investigated in our study. Besides, Sidama and wild enset populations showed higher differentiation from one another. Most probably accessions from the regions are not currently breeding with one another and there is no sharing of planting materials. Concerning cultivated enset, accessions from Gurage and South Omo show low connectivity (**Figure 3, 4**). This might be due to a distinct genetic profile within Gurage and South Omo accessions and possibly no frequent exchange of accessions between the two regions. Our SNPs data indicate that the cultivated and wild enset accessions are very divergent. Besides, the principal component and phylogenetic tree analysis grouped the 236 enset accessions into four major clusters, where the wild individuals clustered separately. Other enset diversity studies have also reported a high level of genetic differentiation between cultivated and wild enset accessions (Birmeta *et al.,* 2004; Gerura *et al*., 2019; Olango *et al.,* 2015; Tobiaw and Bekele, 2011). Also, geographic form of genetic structure was observed with consistent distinct grouping of cultivated enset accessions from Sidama, Gurage and South Omo. This knowledge of population structure and genetic

diversity between cultivated and wild enset accessions is crucial for future studies and breeding for new introductions.

Genetic Diversity Within and Across Populations

The large regional variation in agroecological conditions, different cultures and management relatively large geographic distances between the different enset growing regions within the country should result in large genetic variation among regions. However, multiple lines of evidence show that the level of genetic variation among regions (geographical areas) is low. For instance, AFLP analysis of 192 enset accessions from six growing regions showed a limited proportion of variation among growing regions (11-13 %), but a considerable variation within regions (87-89%) (Tesfamicael *et al*., 2020). Earlier studies also found limited variation among growing regions compared to within regions, i.e., 13% using AFLPs (Tesfamicael *et al*., 2020), 4.8% using AFLPs (Negash *et al*., 2002) and 16% using SSRs (Olango *et al.,* 2015). These values indicated that the high proportion of genetic variation within regions is a general feature of the enset species.

In the current study, the low genetic structuring among regions of enset that were observed both by the average pairwise F_{ST} values and AMOVA indicate that allele sharing between regions is high. The AMOVA analysis showed that the level of genetic variation among regions is limited (7.6%) and very high within regions (92.4%) (**Table 2**). This is also evident from the low F_{ST} values observed between the cultivated enset accessions from the different growing regions (**Table 4**). These results show that genetic variation in enset accessions are less affected by the region of origin (Schaal *et al*., 1998), but has rather been shaped by a long history of extensive human exchange of clonal materials among regions, and different communities may select different sources of the germplasm to suit their specific cultural needs (Gerura *et al*., 2019; Getachew *et al*., 2014; Negash *et al*., 2002). Furthermore, there has also most probably been extensive exchanges of clones particularly between highland and lowland regions because farmers in the latter area believe that suckers imported from the mountain areas grow better than those raised locally (Tesfaye and Lüdders, 2003). Because of the large genetic variation among accessions within regions, clonal selection based on desirable traits may be effective for most of the natural populations in Ethiopia. In addition, the large genetic variation within regions may be partly explained by gene flow and common origin of the populations. According to some studies, large genetic variation within populations is not necessarily caused by environmental heterogeneity but could be due to historical patterns of relationship (Schaal *et al*., 1998).

In this study was detected 3505 SNPs markers from 236 (226 cultivated and 10 wild) enset accessions. This number of SNPs might be considered low relative to the 5011 SNPs detected from 141 (120 cultivated and 21 wild enset) studied by (Tesfamicael *et al*., 2020). Moreover, the observed heterozygosity (H_O) and expected heterozygosity (H_E) is low (**Table 3**) compared to studies using other DNA marker systems such as ISSR (Getachew *et al*., 2014) and SSRs (Gerura *et al*., 2019; Olango *et al*., 2015), but higher than with AFLP markers, which revealed lower observed and expected heterozygosity in cultivated and in wild enset populations (Tesfamicael *et al*., 2020). However, it is difficult to make direct comparisons between previous

and the present study, due to differences in the number and types of the studied enset accessions and different SNP calling and filtering parameters applied. Other reasons may be that we did not include sources of genetic variation from other enset growing regions which might contain additional allelic variation. Also, different molecular markers have different properties and will scan different regions of the genome (Karp and Edward, 1997).

In this context, it is interesting that wild enset had lower levels of heterozygosity than expected, indicating that wild enset is a sexually propagated plant within a restricted area, which will limit gene flow and lead to inbreeding and increased homozygosity (**Table 3**) (Birmeta *et al.,* 2004; Shigeta, 1992). Moreover, the suitable habitats for wild enset has been sharply declining in Ethiopia because of population growth and deforestation, and the geographical range of wild enset is more limited, possibly due to more specific ecological requirements or alternatively loss of habitat (Birmeta *et al.,* 2004; Olango *et al.,* 2015). This reduction in effective population size might have contributed to the observed lower heterozygosity in wild enset due to the increased chances of inbreeding. This differs from what has been reported based on SSR markers (Olango *et al*., 2015).

However, relatively high levels of heterozygosity were observed in all cultivated populations (**Table 3**), which is consistent with the outcrossing nature of enset during sexual reproduction (Olango *et al*., 2015). Enset might have improved phenotypes through heterosis, so that growers favor heterozygous varieties during selective propagation practices (Oztolan-Erol *et al*., 2021). Further, the current levels of enset diversity reflect frozen variation; that is variation which arose through sexual reproduction in an ancestral population (Chapman *et al*., 2000). In addition, occasional gene flow from wild enset and possibly from other enset species can occur too (Birmeta *et al.,* 2004). Other possible causes of this type of clonal variation might be somatic mutations, introduction of new variation from outside of the cultivated populations, and introduction of new landraces from other regions (Shigeta, 1990; Tsykun *et al*., 2017). Another possible cause might be the perennial and highly clonally propagated species that are highly selected for adaptability and productivity under cultivation, and different pollination mechanisms (Birmeta *et al.,* 2004; Negash *et al*., 2002; Yemataw *et al*., 2016). According to Shank (1994) considerable clonal variation is present within enset for characters associated with growth and adaptation.

Above and beyond, the highlands of southern Ethiopia form the geographical center of enset cultivation (Vavilov, 1997). According to Harlan (1951), high altitude areas have high concentrations of diverse and unique landraces and can be designated as microcenters of enset diversity. All such factors in combination or alone have resulted in a high degree of genetic diversity in the presently studied enset accessions. The most important point is that most likely differences in genetic diversity among regions are important for farmers; different accessions contribute to the high diversity that is observed at each site and provides strong evidence for selection by humans. Enset diversity in Ethiopia may thus be extensive but it is not effectively utilized, as the available germplasm is poorly known (Borrell *et al*., 2019).

Genetic Signatures for Differential Selection Between Cultivated and Wild Enset

Little is known about the genetic makeup and population differentiation between cultivated and wild enset. Knowledge about the genetic adaptation of enset is essential for breeding strategies. A central aim of evolutionary biology is to understand the molecular basis for adaptive differences between populations (Lotterhos and Whitlock, 2014). Higher genetic population differentiation for adaptive SNP than neutral SNP is expected if adaptation to local environments is the principal source of genetic differentiation (De Villemereuil and Gaggiotti, 2015). F_{ST} outlier approaches has been applied to many crops, such as tomato (Sim *et al.,* 2011), perennial ryegrass (Kovi *et al*., 2015), soybean (Li *et al*., 2014), European beech (Laura *et al*., 2018), banana (Hinge *et al.,* 2022) and common bean (Papa *et al.,* 2007) for identifying adaptive differentiation. Markers detected in these crops have been mapped to the genomic regions with known QTL/genes related to domestication.

Wild enset propagates by seed under natural condition, while cultivated enset is propagated only vegetatively by local farmers (Borrell *et al*., 2019; Brandt *et al*., 1997; Shigeta, 1992). Most probably the genetic differences between wild and cultivated enset populations can be attributed to the different reproduction systems (Birmeta *et al.,* 2004; Gerura *et al*., 2019; Olango *et al.,* 2015; Tesfamicael *et al*., 2020). Continued vegetative propagation during cultivation can lead to loss of sexual reproduction capacity (Denham *et al*., 2020), thus flowering, seed development, seed size, numbers of viable seeds per fruit and per infructescence are important traits that differentiate cultivated and wild enset (Borrell *et al*., 2019; Brandt *et al*., 1997).

In the present study, we identified 12 candidate loci putatively under positive selection based on F_{ST} values displaying differentiation higher than the 99% limit of the confidence interval (**Figure 5, Table 5**). Among them, six loci, i.e., E-2488, E-3078, E-298, E-1617, E-3031 and E-3091, might be directly under selection. SNP annotation showed the putative functions of all these candidate loci (**Table 5**) are involved in different biological processes, including sexual reproduction and flowering signaling in plants, which are key players in domestication and adaptation (Borrell *et al*., 2019). E-2488 was identified as a SAUR-like auxin-responsive protein. Small auxin-upregulated RNAs (SAURs) is the largest family of early auxin responsive genes in higher plants regulating a wide range of cellular, physiological, and developmental processes (Ren and Gray, 2015; Zhang *et al*., 2021). Most of the *SAUR* genes, which are part of auxin response factors (ARF) regulate cell elongation, at least in the seedlings (Sun *et al*., 2016). Further, Hu *et al.* (2015) showed higher expression of *MaARF* genes at initial days of flowering than at later stages, suggesting crucial roles of the *ARF* genes in early banana fruit development. E-3078 was identified as an isoflavone synthase gene (IFS), which plays a natural role in plant defense and root nodulation. Manipulating the expression of IFS in legumes showed improved pathogen and stress responses (Jung *et al.,* 2000). E-298 was detected as a DNA binding with one finger (Dof) protein, which is a plant-specific transcription factor having multiple roles, such as seed maturation and germination (Ruta *et al*., 2020). Further, Dof proteins involved in the growth and development of banana reproductive organs (Dong *et al*., 2016; Venkatesh and Park, 2015). E-1617 was identified as a serine/threonine-protein kinase (STK). STKs are involved in various

developmental processes like cell proliferation, modification of cell shape and apoptosis. Proteomic studies in somatic embryo development in banana, showed that serine /threonine- protein kinase (spot 17) was found to be highly expressed in mature somatic embryos and these proteins are associated with pattern formation and tissue specification during embryonic developmental process (Kumaravel *et al.,* 2020). E-3031 was identified as histone acetyltransferases (HATs), which plays critical roles in the regulation of chromatin structure and gene expression. Genetic analysis and cytological study revealed that the double mutation induced severe defects in the formation of male and female gametophyte, resulting in an arrest of mitotic cell cycle at early stages of gametogenesis (Latrasse *et al*., 2008), thus showing their crucial roles in cell division. The final SNP, E-3091 was associated with R2R3-MYB transcription factor. These transcription factors have been shown to play regulatory roles during plant development, and responses to biotic and abiotic stress in banana (Pucker *et al*., 2020). Further, MYB genes *MaMYB4*, an R2R3-MYB repressor transcription factor, negatively regulates the biosynthesis of anthocyanin in banana (Deng *et al*., 2021) and also *MaMYB3* is involved in fruit ripening through modulation of starch degradation (Fan *et al*., 2018). Moreover, two of the genes, serine/threonine-protein kinase and MYB transcription factor identified in our study were also detected in a similar study of enset (Tesfamicael *et al*., 2020).

Conclusions

In conclusion, our genetic diversity and population structure studies on enset genotypes from in Ethiopia found a significant subdivision between cultivated and wild enset and a large genetic variation within populations, indicating a heterogeneous collection. The genetic structure of enset was mainly shaped by eco-geographic factors, mode of propagation and cultivation status. The molecular marker analysis shows that most of the genetic variability exists within geographical regions and very little between regions. The results also imply that sexual reproduction plays an important role in shaping the genetic structure of enset populations. Enset from Sidama and South Omo are more genetically diverse than enset from Gurage. The diversity within the regions is not the same and different accessions contribute to the large diversity observed at each site. This study reports extensive and unexplored enset genotype diversity present in Ethiopia, encouraging further exploration and preservation of genotypes with desirable traits inside the known centers of origin. Furthermore, we identified 12 candidate loci under positive selection, at least 6 of these genes are most probably linked to genes involved in sexual reproduction and flowering signalling that are key processes underlying domestication and adaptation. These novel findings are useful for the conservation of genetic resources, especially under global climate changes, and contribute to the potential discovery of functional genes and genetic mechanisms related to adaptability of enset to local climatic conditions, especially drought. This is encouraging for the potential of diversifying crops also in regions where enset is not traditionally grown, such as the food insecure dry north.

Acknowledgements

The authors thank the Norwegian Agency for Development Cooperation (NORAD) for funding this study, which is part of the project 'Controlling Disease in Sweet potato and Enset in South Sudan and Ethiopia to

Improve Productivity and Livelihoods under Changing Climatic Conditions using Modern Technologies' under the NORHED program (Agreement no. ETH-13/0017). The authors would like to thank Øyvind Jørgensen and Anne Guri Marøy for excellent technical support.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Abebe, T. 2005. *Diversity in homegarden agroforestry systems of Southern Ethiopia* [Dissertation, Wageningen University,The Netherlands]. Wageningen University.

Alkan, C., Coe, B. P., Eichler, E. E. 2011. Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, *12*(5), 363-376. https://doi.org/doi:10.1038/nrg2958

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., Hohenlohe, P. A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81. https://doi.org/doi:10.1038/nrg.2015.28

Baker, R., Simmonds, N. 1953. The genus Ensete in Africa. *Kew Bulletin*, *8*(3), 405-416.

Beaumont, M. A., Nichols, R. A. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*(1377), 1619-1626. https://doi.org/doi.org/10.1098/rspb.1996.0237

Bekele, E., Shigeta, M. 2011. Phylogenetic relationships between Ensete and Musa species as revealed by the trnT trnF region of cpDNA. *Genetic Resources and Crop Evolution*, *58*(2), 259-269. https://doi.org/10.1007/s10722-010-9568-2

Bezuneh, T., Feleke, A., Bayie, R. 1967. The cultivation of the genus Ensete in Ethiopia. In *Soil Crop Sci Soc Fla Proc* (Vol. 27, pp. 133-141).

Birmeta, G., Nybom, H., Bekele, E. 2004. Distinction between wild and cultivated enset (Ensete ventricosum) gene pools in Ethiopia using RAPD markers. *Hereditas*, *140*(2), 139-148. https://doi.org/10.1111/j.1601-5223.2004.01792.x

Bolger, A. M., Lohse, M., Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/doi.org/10.1093/bioinformatics/btu170

Borrell, J. S., Biswas, M. K., Goodwin, M., Blomme, G., Schwarzacher, T., Heslop-Harrison, J. S. P., Wendawek, A. M., Berhanu, A., Kallow, S., Janssens, S., Molla, E. L., Davis, A. P., Woldeyes, F., Willis, K., Demissew, S., Wilkin, P. 2019. Enset in Ethiopia: a poorly characterized but resilient starch staple. *Annals of Botany*, *123*(5), 747-766. https://doi.org/10.1093/aob/mcz057

Borrell, J. S., Goodwin, M., Blomme, G., Jacobsen, K., Wendawek, A. M., Gashu, D., Lulekal, E., Asfaw, Z., Demissew, S., Wilkin, P. 2020. Enset-based agricultural systems in Ethiopia: A systematic review of production trends, agronomy, processing and the wider food security applications of a neglected banana relative. *Plants, People, Planet*, *2*(3), 212-228. https://doi.org/10.1002/ppp3.10084

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633-2635. https://doi.org/doi:10.1093/bioinformatics/btm308

Brandt, S., Spring, A., Hiebsch, C., McCabe, J., Tabogie, E., Diro, M., Wolde-Michael, G., Yntiso, G., Shigeta, M., Tesfaye, S. (1997). *The "Tree Against Hunger": Enset-based agricultural systems in Ethiopia*. (Vol. 56). American Association for the Advancement of Science, Washington, DC, USA.

Bus, A., Hecht, J., Huettel, B., Reinhardt, R., Stich, B. 2012. High-throughput polymorphism detection and genotyping in Brassica napus using next-generation RAD sequencing. *BMC Genomics*, *13*(1), 1-11. https://doi.org/doi:10.1186/1471-2164-13-281

Chapman, H. M., Parh, D., Oraguzie, N. 2000. Genetic structure and colonizing success of a clonal, weedy species, Pilosella officinarum (Asteraceae). *Heredity*, *84*(4), 401-409.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674-3676. https://doi.org/10.1093/bioinformatics/bti610

Central Statistical Agency (CSA), Government of Ethiopia. 2020. Agricultural sample survey 2019/20 (2012 E.C.) Report on area and production of major crops (Private Peasant Holdings, Meher Season). Statistical Bulletin 587 (Vol. I). Addis Ababa, Ethiopia

De Villemereuil, P., Gaggiotti, O. E. 2015. A new F_{ST}-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, *6*(11), 1248-1258. https://doi.org/doi.org/10.1111/2041-210X.12418

Deng, G.-M., Zhang, S., Yang, Q.-S., Gao, H.-J., Sheng, O., Bi, F.-C., Li, C.-Y., Dong, T., Yi, G.-J., He, W.-D. 2021. MaMYB4, an R2R3-MYB repressor transcription factor, negatively regulates the biosynthesis of anthocyanin in banana. *Frontiers in plant science*, *11*, 2023. https://doi.org/doi.org/10.3389/fpls.2020.600704

Denham, T., Barton, H., Castillo, C., Crowther, A., Dotte-Sarout, E., Florin, S. A., Pritchard, J., Barron, A., Zhang, Y., Fuller, D. Q. 2020. The domestication syndrome in vegetatively propagated field crops. *Annals of Botany*, *125*(4), 581-597. https://doi.org/doi.org/10.1093/aob/mcz212

Dong, C., Hu, H., Xie, J. 2016. Genome-wide analysis of the DNA-binding with one zinc finger (Dof) transcription factor family in bananas. *Genome*, *59*(12), 1085-1100. https://doi.org/doi.org/10.1139/gen-2016-008

Engels, J., Hawkes .1991. The Ethiopian gene centre and its genetic diversity. *Plant genetic resources of Ethiopia* (pp. 23-41). Cambridge University Press.

Excoffier, L., Hofer, T., Foll, M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*(4), 285-298. https://doi.org/10.1038/hdy.2009.74

Excoffier, L., Lischer, H. E. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fan, Z. Q., Ba, L. J., Shan, W., Xiao, Y. Y., Lu, W. J., Kuang, J. F., Chen, J. Y. 2018. A banana R2R3-MYB transcription factor MaMYB3 is involved in fruit ripening through modulation of starch degradation by repressing starch degradation-related genes and MabHLH6. *The Plant Journal*, *96*(6), 1191-1205. https://doi.org/doi.org/10.1111/tpj.14099

Frankham, R., Briscoe, D. A., Ballou, J. D. 2002. *Introduction to conservation genetics*. Cambridge university press. https://doi.org/doi.org/10.1017/CBO9780511808999

Gerura, F. N., Meressa, B. H., Martina, K., Tesfaye, A., Olango, T. M., Nasser, Y. 2019. Genetic diversity and population structure of enset (Ensete ventricosum Welw Cheesman) landraces of Gurage zone, Ethiopia. *Genetic Resources and Crop Evolution*, *66*(8), 1813-1824. https://doi.org/10.1007/s10722-019-00825-2

Getachew, S., Mekbib, F., Admassu, B., Kelemu, S., Kidane, S., Negisho, K., Djikeng, A., Nzuki, I. 2014. A look into genetic diversity of enset (Ensete ventricosum (Welw.) cheesman) using transferable microsatellite sequences of banana in Ethiopia. *J. Crop Improv*, *28*(2), 159-183. https://doi.org/doi.org/10.1080/15427528.2013.861889

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307-321. https://doi.org/10.1093/sysbio/syq010

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite North American potato germplasm. *BMC Genomics*, *12*(1), 302. https://doi.org/doi.org/10.1186/1471-2164-12-302

Harlan, J. R. 1951. Anatomy of gene centers. *The American Naturalist*, *85*(821), 97-103. https://doi.org/doi.org/10.1086/281657

Hartl, D. L., Clark, A. G., Clark, A. G. 1997. *Principles of population genetics* (3rd ed., Vol. 116). Sinauer associates Sunderland, MA.

Hinge, V. R., Shaikh, I. M., Chavhan, R. L., Deshmukh, A. S., Shelake, R. M., Ghuge, S. A., Dethe, A. M., Suprasanna, P., Kadam, U. S. 2022. Assessment of genetic diversity and volatile content of commercially grown banana (Musa spp.) cultivars. *Scientific Reports*, *12*(1), 1-16. https://doi.org/doi.org/10.1038/s41598-022-11992-1

Hinze, L. L., Hulse-Kemp, A. M., Wilson, I. W., Zhu, Q.-H., Llewellyn, D. J., Taylor, J. M., Spriggs, A., Fang, D. D., Ulloa, M., Burke, J. J. 2017. Diversity analysis of cotton (Gossypium hirsutum L.) germplasm using the CottonSNP63K Array. *BMC Plant Biology*, *17*(1), 37. https://doi.org/doi:10.1186/s12870-017-0981-y

Hu, W., Zuo, J., Hou, X., Yan, Y., Wei, Y., Liu, J., Li, M., Xu, B., Jin, Z. 2015. The auxin response factor gene family in banana: genome-wide identification and expression analyses during development, ripening, and abiotic stress. *Frontiers in Plant Science*, *6*, 742. https://doi.org/10.3389/fpls.2015.00742

Jung, W., Yu, O., Lau, S.-M. C., O'Keefe, D. P., Odell, J., Fader, G., McGonigle, B. 2000. Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. *Nature Biotechnology*, *18*(2), 208-212. https://doi.org/https://doi.org/10.1038/72671

Karp, A. 1997. *Molecular tools in plant genetic resources conservation: a guide to the technologies* (No. 2). Bioversity International.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., Mayrose, I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179-1191. https://doi.org/doi:10.1111/1755-0998.12387

Kovi, M. R., Fjellheim, S., Sandve, S. R., Larsen, A., Rudi, H., Asp, T., Kent, M. P., Rognli, O. A. n. . 2015. Population structure, genetic variation, and linkage disequilibrium in perennial ryegrass populations divergently selected for freezing tolerance. *Frontiers in plant scie*, *6*, 929. https://doi.org/doi.org/10.3389/fpls.2015.00929

Kumaravel, M., Uma, S., Backiyarani, S., Saraswathi, M. S. 2020. Proteomic analysis of somatic embryo development in Musa spp. cv. Grand Naine (AAA). *Scientific Reports*, *10*(1), 4501. https://doi.org/10.1038/s41598-020-61005-2

Latrasse, D., Benhamed, M., Henry, Y., Domenichini, S., Kim, W., Zhou, D.-X., Delarue, M. 2008. The MYST histone acetyltransferases are essential for gametophyte development in Arabidopsis. *BMC Plant Biology*, *8*(1), 121. https://doi.org/10.1186/1471-2229-8-121

Laura, C.-A., Matthias, A., Markus, M., Christoph, S., Reiner, F., Krutovsky, K. V. 2018. Genetic variation and signatures of natural selection in populations of European beech (Fagus sylvatica L.) along precipitation gradients. https://doi.org/DOI:10.1007/s11295-018-1297-2

Letunic, I., Bork, P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*, 256-259. https://doi.org/doi:10.1093/nar/gkz239

Li, Y.-H., Reif, J. C., Jackson, S. A., Ma, Y.-S., Chang, R.-Z., Qiu, L.-J. 2014. Detecting SNPs underlying domestication-related traits in soybean. *BMC Plant Biology*, *14*(1), 1-8. https://doi.org/doi:10.1186/s12870-014-0251-1

Li, Y. L., Liu, J. X. 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources*, *18*(1), 176-177. https://doi.org/DOI:10.1111/1755-0998.12719

Costa, M.G.D.A., Lockhart, D.M. (1984). *The Itinerario of Jeronimo Lobo*. London: Hakluyt Society.

Lotterhos, K. E., Whitlock, M. C. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, *23*(9), 2178-2192. https://doi.org/doi.org/10.1111/mec.12725

McKey, D., Elias, M., Pujol, B., and Duputié, A. 2010. The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, *186*(2), 318-332. https://nph.onlinelibrary.wiley.com/doi/full/10.1111/j.1469-8137.2010.03210.x

Negash, A., Tsegaye, A., van Treuren, R., Visser, B. 2002. AFLP analysis of enset clonal diversity in south and southwestern Ethiopia for conservation. *Crop Science*, *42*(4), 1105-1111. https://doi.org/10.2135/cropsci2002.1105

Olango, T. M., Tesfaye, B., Pagnotta, M. A., Pè, M. E., Catellani, M. 2015. Development of SSR markers and genetic diversity analysis in enset (Ensete ventricosum (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. *BMC Genetics*, *16*(1), 98. https://doi.org/DOI:10.1186/s12863-015-0250-8

Oztolan-Erol, N., Helmstetter, A. J., Inan, A., Buggs, R. J., & Lucas, S. J. 2021. Unraveling Genetic Diversity amongst European Hazelnut (Corylus avellana L.) Varieties in Turkey. *Frontiers in plant science*, *12*, 1250. https://doi.org/10.3389/fpls.2021.661274

Papa, R., Bellucci, E., Rossi, M., Leonardi, S., Rau, D., Gepts, P., Nanni, L., Attene, G. . 2007. Tagging the signatures of domestication in common bean (Phaseolus vulgaris) by means of pooled DNA samples. *Annals of Botany*, *100*(5), 1039-1051. https://doi.org/doi.org/10.1093/aob/mcm151

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pino Del Carpio, D., Basnet, R. K., De Vos, R. C., Maliepaard, C., Visser, R., & Bonnema, G. 2011. The patterns of population differentiation in a Brassica rapa core collection. Theor Appl Genet 122(6), 1105-1118. https://doi.org/DOI:10.1007/s00122-010-1516-1

Purseglove, J. W. 1985. *Tropical Crops: Monocotyledons* (5th ed.). Longman.

Raj, A., Stephens, M., Pritchard, J. K. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Gnetics*, *197*(2), 573-589. https://doi.org/doi:10.1534/genetics.114.164350/-/DC1

Ren, H., Gray, W. M. 2015. SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Molecular plant*, *8*(8), 1153-1164. https://doi.org/doi.org/10.1016/j.molp.2015.05.003

Rochette, N. C., Rivera-Colón, A. G., Catchen, J. M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737-4754. https://doi.org/DOI:10.1111/mec.15253

Rogstad, S. H. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon*, *41*, 701-708. https://doi.org/10.2307/1222395

Ruta, V., Longo, C., Lepri, A., De Angelis, V., Occhigrossi, S., Costantino, P., & Vittorioso, P. 2020. The DOF transcription factors in seed and seedling development. *Plants*, *9*(2), 218. https://doi.org/doi.org/10.3390/plants9020218

Schaal, B., Hayworth, D., Olsen, K. M., Rauscher, J., Smith, W. 1998. Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, *7*(4), 465-474. https://doi.org/doi.org/10.1046/j.1365-294x.1998.00318.x

Shank, R. 1994. The enset culture: a technical report on Enset ventricosum or 'False Banana'. In: Addis Ababa: United Nations-Emergencies Unit for Ethiopia.

Shigeta, M. 1990. Folk in-situ conservation of ensete (Ensete ventricosum (WEL W.) E.E. Cheesman): towards the interpretation of indigenous agricultural science of the Ari. southwestern Ethiopia. *African Study Monographs*, *10*(3), 93-107.

Shigeta, M. 1992. *The ethnobotanical study of ensete (Ensete ventricosum) in the southwestern Ethiopia* Kyoto University]. Japan. http://hdl.handle.net/2433/86243

Sim, S., Robbins, M., Van Deynze, A., Michel, A., Francis, D. 2011. Population structure and genetic differentiation associated with breeding history and selection in tomato (Solanum lycopersicum L.). *Heredity*, *106*(6), 927-935. https://doi.org/10.1038/hdy.2010.139

Simmonds, N. W. 1962. *The Evolution of the Bananas*. Longmans, Green, London.

Spring, A., Haile, B., Tesfaye, S., Abebe, Y., Amaldegn, A., Wolde-Michae, G., Tabogie, E., Surur, O., Tsegaye, A., & Shimeles, S. 1996. Enset farming systems in southern region, Ethiopia: Report on a rapid rural appraisal in Gurage, Hadiya, and Sidama zones Deutsche Gesellschaft for Technische Zusammenarbeit (GTZ), Addis Ababa, Ethiopia

Sun, N., Wang, J., Gao, Z., Dong, J., He, H., Terzaghi, W., Wei, N., Deng, X. W., Chen, H. 2016. Arabidopsis SAURs are critical for differential light regulation of the development of various organs. *Proceedings of the National Academy of Sciences*, *113*(21), 6071-6076. https://doi.org/https://doi.org/10.1073/pnas.160478211

Tesfamicael, K. G., Gebre, E., March, T. J., Sznajder, B., Mather, D. E., Lopez, C. M. R. 2020. Accumulation of mutations in genes associated with sexual reproduction contributed to the domestication of a vegetatively propagated staple crop, enset. *J bioRxiv*. https://doi.org/doi.org/10.1038/s41438-020-00409-7

Tesfaye, B., Lüdders, P. 2003. Diversity and distribution patterns of enset landraces in Sidama, Southern Ethiopia. *Genetic Resources and Crop Evolution*, *50*(4), 359-371.

Tobiaw, D. C., Bekele, E. 2011. Analysis of genetic diversity among cultivated enset (Ensete ventricosum) populations from Essera and Kefficho, southwestern part of Ethiopia using inter simple sequence repeats (ISSRs) marker. *African Journal of Biotechnology*, *10*(70), 15697-15709. https://doi.org/10.5897/AJB11.885

Tsegaye, A. 2002. *On indigenous production, genetic diversity and crop ecology of enset (Ensete ventricosum (Welw.) Cheesman).* Wageningen University and Research.

Tsegaye, A., Struik, P. C. 2001. Enset (Ensete ventricosum (Welw.) Cheesman) kocho yield under different crop establishment methods as compared to yields of other carbohydrate-rich food crops. *Netherlands Journal of Agricultural Science*, *49*(1), 81-94. https://doi.org/10.1016/s1573-5214(01)80017-8

Tsegaye, A., Struik, P. G. 2002. Analysis of enset (Ensete ventricosum) indigenous production methods and farm-based biodiversity in major enset-growing regions of southern Ethiopia. *Experimental Agriculture*, *38*(3), 291-315. https://doi.org/10.1017/s0014479702003046

Tsykun, T., Rellstab, C., Dutech, C., Sipos, G., Prospero, S. 2017. Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus Armillaria cepistipes. *Heredity*, *119*(5), 371-380. https://doi.org/10.1038/hdy.2017.48

Vavilov, N. 1997. Five Continents (eds.) LE Rodin. *Semyon Reznik and Paul Stapleton, International Board for Plant Genetic Resources, Rome, ITALY*, 1-197.

Vavilov, N. I.(1951. *The origin, variation, immunity and breeding of cultivated plants* (Vol. 72, No. 6, p. 482). LWW.

Venkatesh, J., Park, S. W. 2015. Genome-wide analysis and expression profiling of DNA-binding with one zinc finger (Dof) transcription factor family in potato. *Plant Physiology and Biochemistry*, *94*, 73-85. https://doi.org/doi.org/10.1016/j.plaphy.2015.05.010

Weir, B. S., Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358-1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Westphal, E., Stevels, J. M. C., & Stevels, J. M. C. 1975. *Agricultural systems in Ethiopia* (Vol. 826). Wageningen: Centre for Agricultural Publishing and Documentation.

Wright, S. 1978. Variability within and among populations. *Evolution and the Genetics of Populations*.

Yemataw, Z., Tesfaye, K., Taddese, T., Dejene, T., Muzemil, S., Shashego, Z., Ambachew, D. 2016. Genetic variation for corm yield and other traits in Ethiopian enset (Ensete ventricosum (Welw.) Cheesman). *Journal of Plant Breeding and Crop Science*, *8*(9), 150-156. https://doi.org/doi.org/10.5897/JPBCS2015.0552

Zeberga, A., Yemataw, Z., Sinebo, W., Musemil, S. 2014. On farm cultivar diversity of enset (Ensete ventricosum W.) in Southern Ethiopia. *Journal of Agriculture and Development*, *4*, 62-85.

Zhang, H., Yu, Z., Yao, X., Chen, J., Chen, X., Zhou, H., Lou, Y., Ming, F., Jin, Y. 2021. Genome-wide identification and characterization of small auxin-up RNA (SAUR) gene family in plants: evolution and expression profiles during normal growth and stress response. *BMC Plant Biology*, *21*(1), 1-14. https://doi.org/doi.org/10.1186/s12870-020-02781-x

Tables

Table 1. Enset (*Ensete. ventricosum* (Welw.) Cheesman) plant materials from Ethiopia used for genetic diversity analyses

Cultivation status	Number of accessions	Geographical origin and location	Sources
Cultivated	72	Sidama	Farmers
Cultivated	62	Gurage	Farmers
Cultivated	92	South Omo	Farmers
Wild	10	South Omo	Farmers and forest area

Table 2. Analysis of Molecular Variance (AMOVA) among and within populations and regions of cultivated and wild enset.

Source	df	Sum of squares	Variance component	Total variation (%)
Cultivated and wild enset				
Among populations	3	16495.01	51.	8.79
Within populations	468	224622.88	532.71	91.20
Total	471	241117.89	584.11	
Growing regions				
Among regions	2	12822.99	44.13	7.64
Within regions	449	215813.23	533.83	92.35
Total	451	228635.22	577.96	

Table 3. Genetic diversity among cultivated and wild enset populations based on 3505 SNP markers

	Enset populations			
	Sidama	Gurage	South Omo	Wild
Observed heterozygosity (H_O)	0.33	0.31	0.31	0.32
Expected heterozygosity (H_E)	0.32	0.30	0.32	0.36

Table 4. Average pairwise population differentiation (F_{ST}).

Regions	Sidama	Gurage	South Omo
Gurage	0.07		
South Omo	0.06	0.10	
Wild	0.15	0.17	0.14

Table 5. Functional annotation of outlier SNPs potentially involved in domestication of enset.

SNP ID	Gene name	Gene function	Reference
E-2488	SAUR-like auxin-responsive protein family	This family plays important roles in flowering, plant growth and regulation of plant architecture.	https://www.uniprot.org/uniprot/Q3E901
E-3078	Isoflavone synthase gene	Controls isoflavone accumulation and is most expressed in the developing seed.	https://www.uniprot.org/uniprot/Q9M6D6
E-298	DNA binding with one finger (Dof) proteins	Shift in flowering time in the landraces through regulation of *CONSTANS*.	https://www.uniprot.org/uniprot/F4IJM6
E-1617	Serine/threonine-protein kinase	Regulates flowering time by modulating the photoperiod pathway.	https://www.uniprot.org/uniprot/Q944Q0
E-3031	Histone acetyltransferase	Plays a critical role in floral meristem development.	https://www.uniprot.org/uniprot/Q9AR19
E-3091	R2R3-MYB transcription factor	Regulate plant growth and development.	https://www.uniprot.org/uniprot/Q9SAM2

Figures

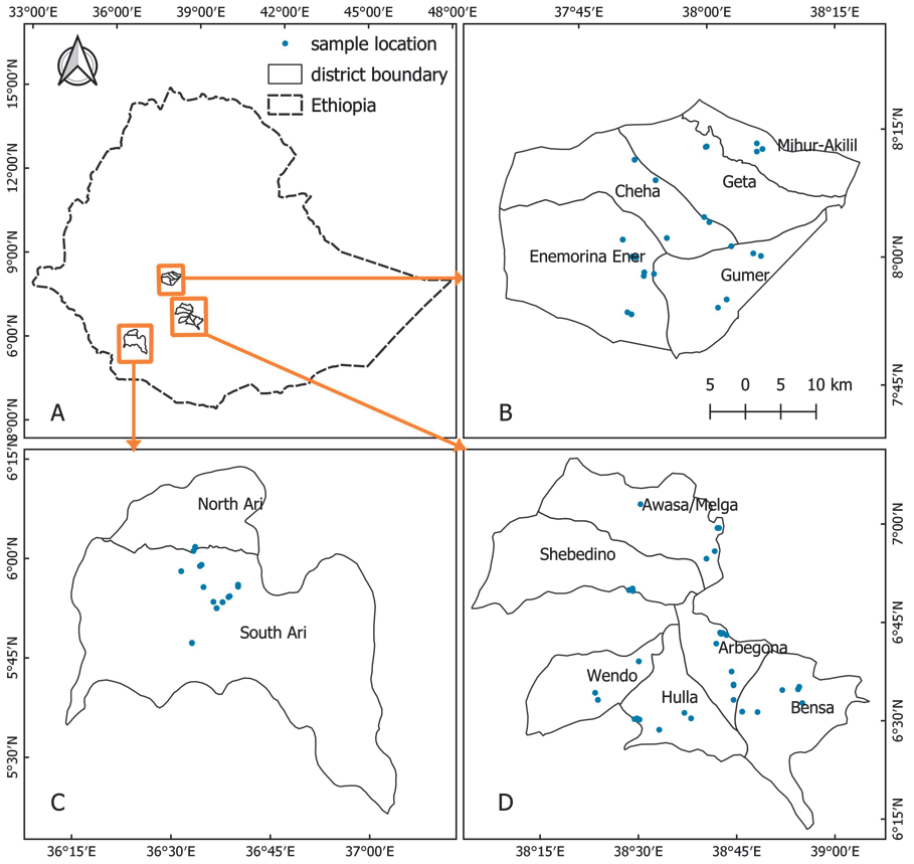


Figure 1. A: *An overview of the study districts in Ethiopia and their detailed locations; B: Gurage; C: South Omo; D: Sidama.*

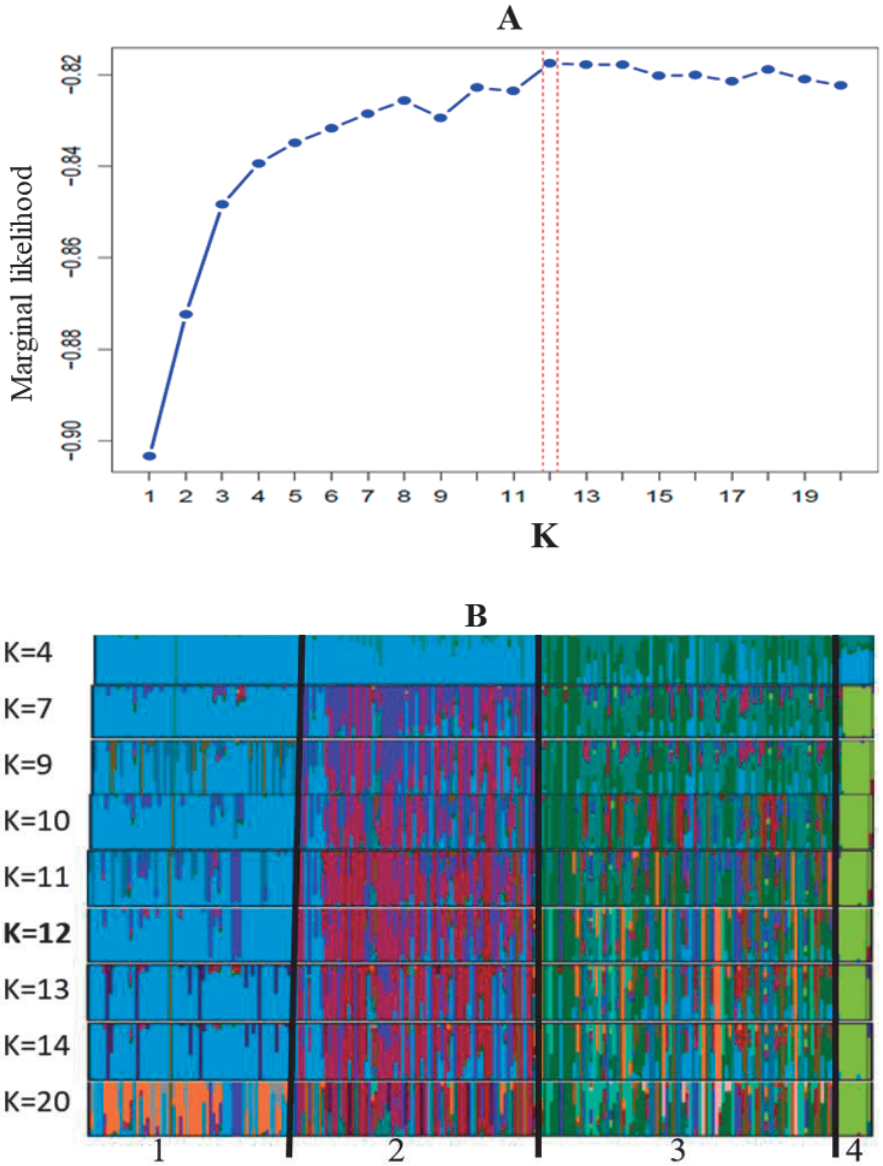


Figure 2. *Population structure; A: Model complexity that maximizes likelihood is 12 and the highest peak shows on the graph the best K = 12; B: Population structure based on fastSTRUCTURE output resulting in K = 12 being the most likely number of genetic clusters, where each cluster is represented as a different shade and each bar represents an individual within each geographic region/cultivation status; colors represent the groups identified. 1: Gurage (cultivated enset); 2: Sidama (cultivated enset); 3: South Omo (cultivated enset); 4: wild enset (from South Omo).*

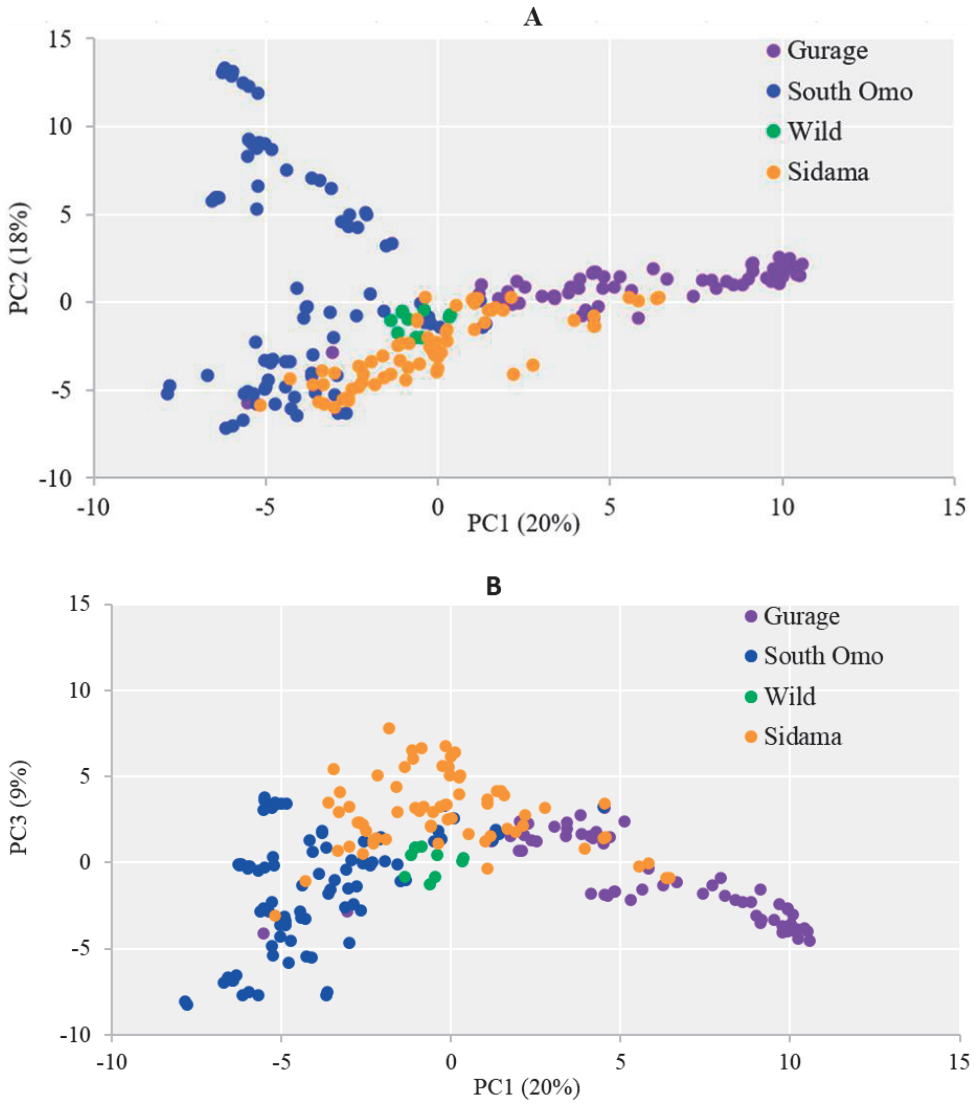
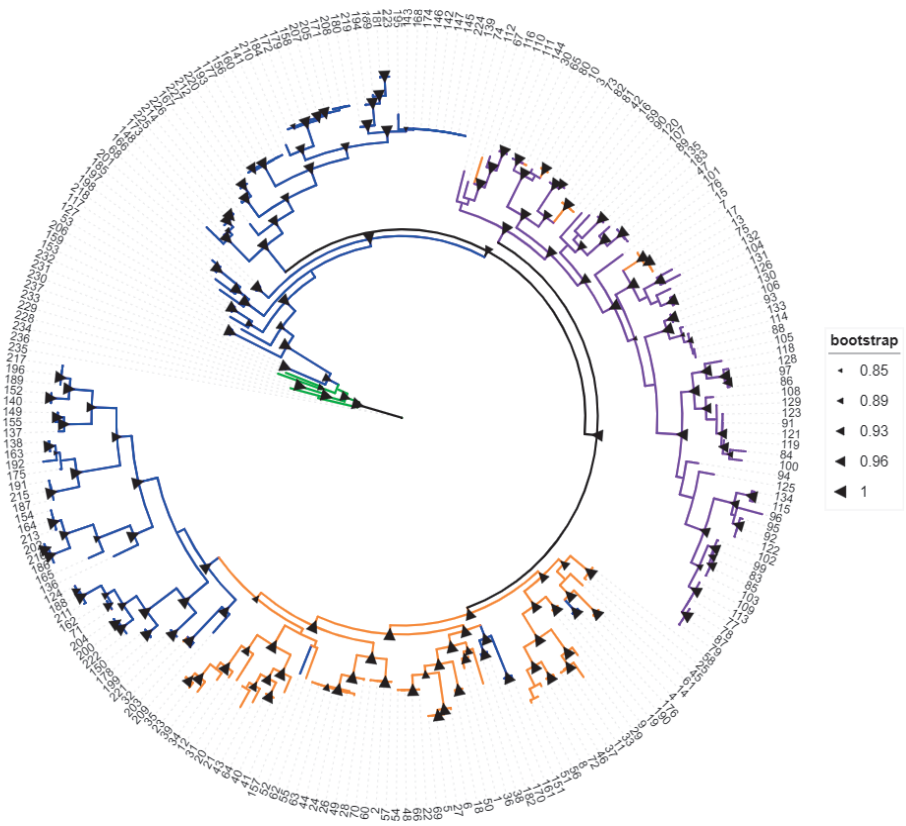


Figure 3. *A: Principal component analysis (PCA) plots of PC1 and PC2; B: PC1 and PC3. The percentages in brackets indicate the variance explained by the different PCs.*

A



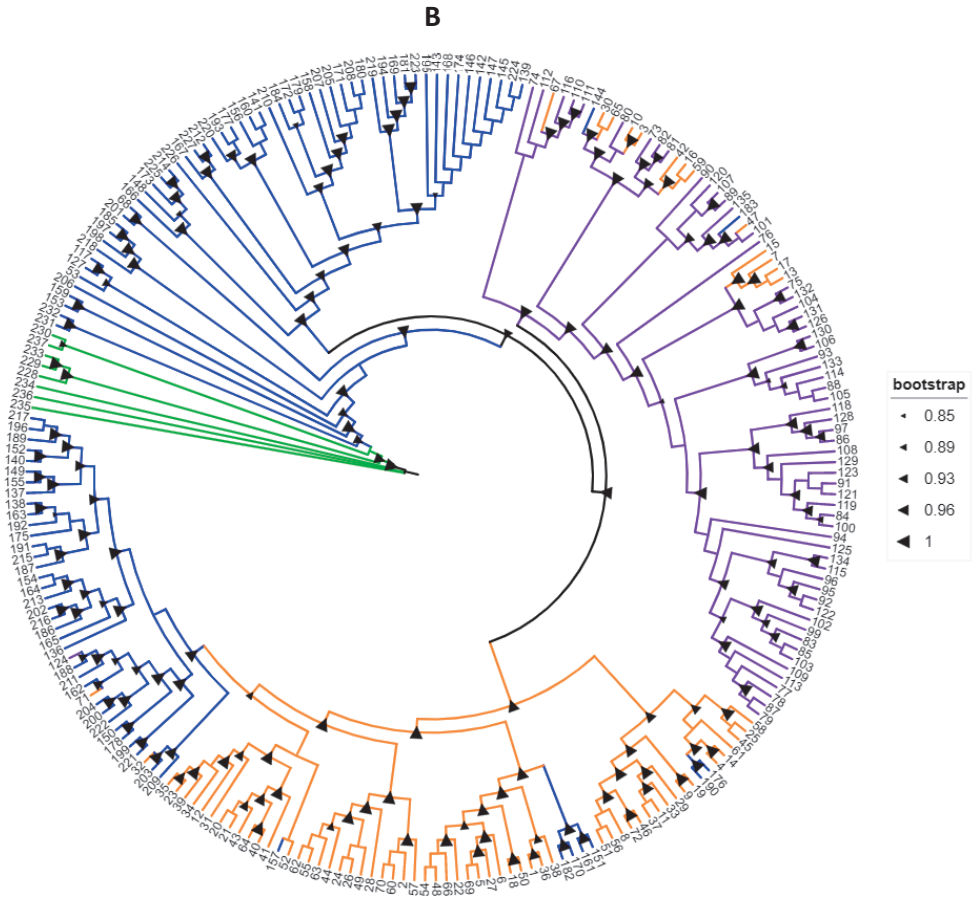


Figure 4. *Phylogenetic tree; A: Maximum-likelihood phylogenetic tree with branch length displayed; B: Topological view of the maximum-likelihood phylogenetic tree. Accessions are numbered as in Supplementary Table 1 and colored according to their geographical origin and cultivation status, i.e., South Omo: blue; Sidama: orange; Gurage: purple; wild enset: green.*

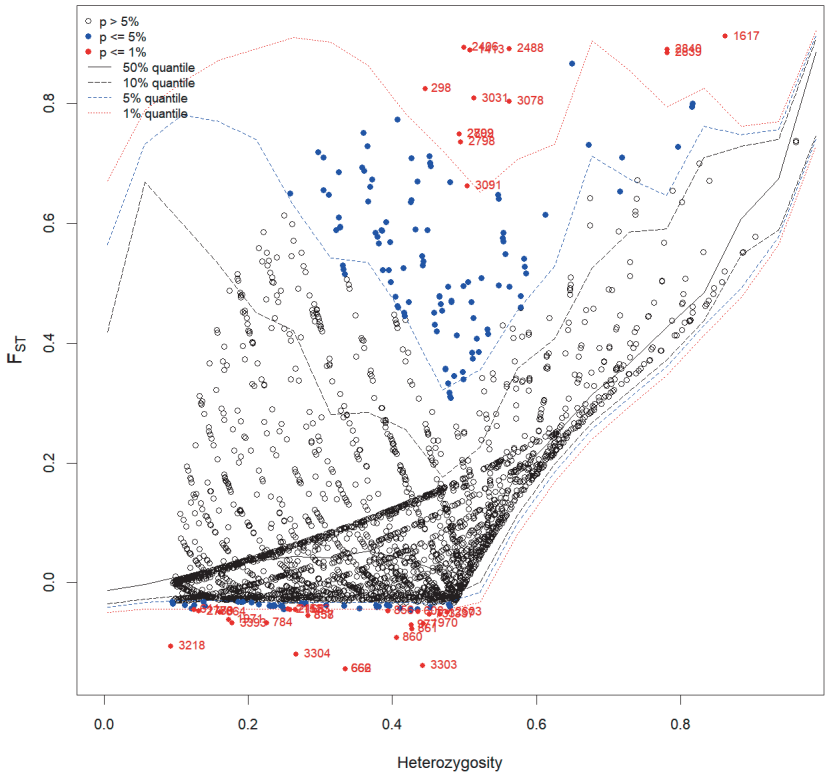


Figure 5. *Candidate loci under selection were identified using F_{ST} based outlier approach (Hierarchical structure model using Arlequin 3.5). F_{ST}: locus–specific genetic divergence among the populations; heterozygosity: measure of heterozygosity per locus. Loci significant at the 1% level are indicated by red dots.*

Supplementary material

Supplementary Table 1. Descriptions of the 226 cultivated and 10 wild enset (*Ensete ventricosum* (Welw.) Cheesman) genotypes used for genetic analysis with their corresponding geographical locations in Ethiopia.

No.	Local landrace names	Biological status	Mode of propagation	Region	Geographic coordinates	
1	Medasho	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
2	Kiticho	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
3	Ganticha	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
4	Lalamo	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
5	Gulumo	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
6	Birra	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
7	Silitte	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
8	Goloma	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
9	Gadime	Cultivated	Induced sucker	Sidama	06 43 26.1	038 42 28.7
10	Maldea	Cultivated	Induced sucker	Sidama	06 43 16.6	038 42 37.4
11	Chacho	Cultivated	Induced sucker	Sidama	06 43 16.6	40 42 37.4
12	Astara	Cultivated	Induced sucker	Sidama	06 43 16.0	038 42 43.8
13	Shawite	Cultivated	Induced sucker	Sidama	06 43 24.0	038.42.55.0
14	Serana	Cultivated	Induced sucker	Sidama	06 43 04 .5	038 43 25.7
15	Dado	Cultivated	Induced sucker	Sidama	06 43 04 .5	038 43 25.7
16	Kulle	Cultivated	Induced sucker	Sidama	06 43 04 .5	038 43 25.7
17	Shawa	Cultivated	Induced sucker	Sidama	06 41 46.0	038 41 54.5
18	Awulcho	Cultivated	Induced sucker	Sidama	06 37 28.8	038 44 14.1
19	Duwane	Cultivated	Induced sucker	Sidama	7 37 28.8	39 44 14.1
20	Lemicho	Cultivated	Induced sucker	Sidama	06 35 29.5	038 44 29.6
21	Tunakecho	Cultivated	Induced sucker	Sidama	06 35 25.2	038 44 29.6
22	Made	Cultivated	Induced sucker	Sidama	7 35 25.2	39 44 29.6
23	Nino	Cultivated	Induced sucker	Sidama	06 30 12,5	038 30 10.0
24	Kishicha	Cultivated	Induced sucker	Sidama	06 30 11.7	038 29 55.3
25	Gora	Cultivated	Induced sucker	Sidama	06 30 21,4	038 29 50,7
26	Arsho	Cultivated	Induced sucker	Sidama	06 30 13,6	038 29 49,3
27	Dargicha	Cultivated	Induced sucker	Sidama	06 30 14.5	038 29 24.6
28	Sharite	Cultivated	Induced sucker	Sidama	06 28 36.2	038 33 10.7
29	Keshicha	Cultivated	Induced sucker	Sidama	06 28 36.2	038 33 10.7
30	Boricho	Cultivated	Induced sucker	Sidama	06 31 11 .6	038 37 02.6
31	Agena	Cultivated	Induced sucker	Sidama	06 30 20.9	038 38 04.1
32	Banijo	Cultivated	Induced sucker	Sidama	06 30 20.9	038 38 04.1
33	Torora	Cultivated	Induced sucker	Sidama	06 59 22.8	038 42 17.0
34	kanko	Cultivated	Induced sucker	Sidama	06 59 22.8	038 42 17.0
35	Garawicho	Cultivated	Induced sucker	Sidama	06 59 22.8	038 42 17.0
36	Maziya	Cultivated	Induced sucker	Sidama	06 59 22.8	038 42 17.0
37	Shilo	Cultivated	Induced sucker	Sidama	06 59 22.8	038 42 17.0

38	Kishe	Cultivated	Induced sucker	Sidama	06 59 23.8	038 42 07.5
39	Linto	Cultivated	Induced sucker	Sidama	06 59 23.8	038 42 02.7
40	Bula	Cultivated	Induced sucker	Sidama	06 54 42.2	038 40 23.4
41	Lenbacho	Cultivated	Induced sucker	Sidama	06 55 51.1	038 41 37.5
42	Amboma	Cultivated	Induced sucker	Sidama	06 35 09.8	038 54 33.6
43	Umisho	Cultivated	Induced sucker	Sidama	06 35 09.8	038 54 33.6
44	Wanikore	Cultivated	Induced sucker	Sidama	06 35 09.8	038 54 33.6
45	Adama	Cultivated	Induced sucker	Sidama	06 35 09.8	038 54 33.6
46	Ayidara	Cultivated	Induced sucker	Sidama	06 35 09.8	038 54 33.6
47	Wani-wasa	Cultivated	Induced sucker	Sidama	06 35 09.8	038 54 33.6
48	Mundraro	Cultivated	Induced sucker	Sidama	06 34 46.9	038 54 21.3
49	Derasicho	Cultivated	Induced sucker	Sidama	06 34 42 .3	038 51 57.0
50	Duwiramo	Cultivated	Induced sucker	Sidama	06 34 46.9	038 54 21.3
51	Gusalo	Cultivated	Induced sucker	Sidama	06 32 40.9	038 55 01.2
52	Shonbo	Cultivated	Induced sucker	Sidama	06 31 22.4	038 45 49.8
53	Sirriro	Cultivated	Induced sucker	Sidama	06 31 22.4	038 45 49.8
54	Kitacho	Cultivated	Induced sucker	Sidama	06 31 18 .0	038 48 09.4
55	Borbancho	Cultivated	Induced sucker	Sidama	06 33 09.1	038.23.49.15
56	Botate	Cultivated	Induced sucker	Sidama	06 33 09.1	038 23 49.15
57	Bolanicho	Cultivated	Induced sucker	Sidama	06 33 09.1	038 44 29.6
58	Ado	Cultivated	Induced sucker	Sidama	06 43 16.6	39 42 37.4
59	Hekeche	Cultivated	Induced sucker	Sidama	06 34 16.5	038 23 24.0
60	Ganna	Cultivated	Induced sucker	Sidama	06 34 16.5	038 23 24.0
61	Demala	Cultivated	Induced sucker	Sidama	06 39 04.9	038 30 05.2
62	Gamechela	Cultivated	Induced sucker	Sidama	06 39 04.9	038 30 05.2
63	Mcho	Cultivated	Induced sucker	Sidama	06 39 04.9	038 30 05.2
64	Medicha	Cultivated	Induced sucker	Sidama	06 50 08.8	038 29 06.3
65	Sediso	Cultivated	Induced sucker	Sidama	06 50 08.8	038 29 06.3
66	Birbo	Cultivated	Induced sucker	Sidama	06 50 08.8	038 29 06.3
67	Kanbatcha	Cultivated	Induced sucker	Sidama	06 50 08.8	038 29 06.3
68	Gena	Cultivated	Induced sucker	Sidama	06 50 08.8	038 29 06.3
69	Hahu	Cultivated	Induced sucker	Sidama	06 49 49.0	038 29 10.3
70	Malgicha	Cultivated	Induced sucker	Sidama	06 49 49.0	038 29 10.3
71	Fatane	Cultivated	Induced sucker	Sidama	06 49 56.7	038 28 35.2
72	Kanda	Cultivated	Induced sucker	Sidama	06 49 56.5	038 28 34.2
73	Marye	Cultivated	Induced sucker	Gurage	08 00 23.6	038 05 26.3
74	Kone	Cultivated	Induced sucker	Gurage	08 00 23.6	038 05 26.3
75	Ehire	Cultivated	Induced sucker	Gurage	08 00 06.9	038 06 20.7
76	Benegn	Cultivated	Induced sucker	Gurage	08 00 06.9	038 06 20.7
77	Gezewode	Cultivated	Induced sucker	Gurage	08 01 14.8	038 02 53.4
78	Ashekit	Cultivated	Induced sucker	Gurage	08 01 14.8	038 02 53.4
79	Benet	Cultivated	Induced sucker	Gurage	08 01 14.8	038 02 53.4
80	Kembatye	Cultivated	Induced sucker	Gurage	07 55 02.5	038 02 20.2
81	Ayitkoket	Cultivated	Induced sucker	Gurage	07 55 02.5	038 02 20.2
82	Sinewot	Cultivated	Induced sucker	Gurage	07 55 02.5	038 02 20.2
83	Tegaded	Cultivated	Induced sucker	Gurage	07 54 03.1	038 01 18.6

84	Ameno	Cultivated	Induced sucker	Gurage	08 12 39.8	038 06 34.0
85	Yirgeye	Cultivated	Induced sucker	Gurage	08 12 39.8	038 06 34.0
86	Zigeziwe	Cultivated	Induced sucker	Gurage	08 13 19.8	038 05 51.8
87	Derewetiye	Cultivated	Induced sucker	Gurage	08 13 19.8	038 05 51.8
88	Boresiye	Cultivated	Induced sucker	Gurage	08 12 57.9	037 59 58.2
89	Dere	Cultivated	Induced sucker	Gurage	08 12 57.9	037 59 58.2
90	Weka	Cultivated	Induced sucker	Gurage	08 12 54.2	037 59 51.2
91	Yibye	Cultivated	Induced sucker	Gurage	08 04 40.3	037 59 40.8
92	Keneykeki	Cultivated	Induced sucker	Gurage	08 04 40.3	037 59 40.8
93	Yedabiriye	Cultivated	Induced sucker	Gurage	08 04 04.5	038 00 19.1
94	Yekechireye	Cultivated	Induced sucker	Gurage	08 08 59.1	037 54 00.9
95	Yegara amerate	Cultivated	Induced sucker	Gurage	08 08 59.1	037 54 00.9
96	Gimbuwe	Cultivated	Induced sucker	Gurage	07 57 48.1	037 52 35.5
97	Tererye	Cultivated	Induced sucker	Gurage	07 57 48.1	037 52 35.5
99	Zober	Cultivated	Induced sucker	Gurage	07 58 13.3	037 52 40.1
100	Eegetiye	Cultivated	Induced sucker	Gurage	07 57 59.9	037 53 47.7
101	Ede-mert	Cultivated	Induced sucker	Gurage	07 53 17.8	037 51 10.4
102	Shewredi	Cultivated	Induced sucker	Gurage	07 53 17.8	037 51 10.4
103	Ayiher	Cultivated	Induced sucker	Gurage	07 53 17.8	037 51 10.4
104	Wenadiye	Cultivated	Induced sucker	Gurage	07 53 32.9	037 50 43.3
105	Sapara	Cultivated	Induced sucker	Gurage	07 59 56.6	037 51 41.7
106	Egendiye	Cultivated	Induced sucker	Gurage	07 59 56.6	38 51 41.7
107	Guarye	Cultivated	Induced sucker	Gurage	07 59 56.6	39 51 41.7
108	Agade	Cultivated	Induced sucker	Gurage	07 59 56.6	40 51 41.7
109	Nechiwe	Cultivated	Induced sucker	Gurage	07 59 56.6	41 51 41.7
110	Kuashkuashiye	Cultivated	Induced sucker	Gurage	07 59 56.6	39 51 41.7
111	Derea	Cultivated	Induced sucker	Gurage	07 59 56.6	40 51 41.7
112	Lemar	Cultivated	Induced sucker	Gurage	07 59 56.6	41 51 41.7
113	Gunbura	Cultivated	Induced sucker	Gurage	07 59 56.6	39 51 41.7
114	Gimbowo	Cultivated	Induced sucker	Gurage	07 59 56.6	40 51 41.7
115	Badedet	Cultivated	Induced sucker	Gurage	07 59 56.6	41 51 41.7
116	Gezwed	Cultivated	Induced sucker	Gurage	08 00 03.41	037 51 17.0
117	Orid	Cultivated	Induced sucker	Gurage	08 00 03.41	037 51 17.0
118	Amerat	Cultivated	Induced sucker	Gurage	08 00 03.41	037 51 17.0
119	Bosere	Cultivated	Induced sucker	Gurage	08 00 03.41	037 51 17.0
120	Bereziye	Cultivated	Induced sucker	Gurage	08 00 03.41	037 51 17.0
121	Fereziye	Cultivated	Induced sucker	Gurage	08 00 03.41	037 51 17.0
122	Kanchiwe	Cultivated	Induced sucker	Gurage	08 02 12.5	037 55 20.8
123	Yeshirfiriye	Cultivated	Induced sucker	Gurage	08 02 12.5	037 55 20.8
124	Yedem-Arti	Cultivated	Induced sucker	Gurage	08 02 12.5	037 55 20.8
125	Wusmaer	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8
126	Kibnar	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8
127	Eweredi	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8
128	Keswo	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8
129	Astara	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8
130	Sherteye	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8

131	Gumbre	Cultivated	Induced sucker	Gurage	08 02 01.5	037 50 08.8
132	Gezod	Cultivated	Induced sucker	Gurage	08 11 22.2	037 51 32.6
133	Emnye	Cultivated	Induced sucker	Gurage	08 11 22.2	037 51 32.6
134	Anqefuye	Cultivated	Induced sucker	Gurage	08 11 22.2	037 51 32.6
135	Boyiche	Cultivated	Induced sucker	Gurage	08 11 22.2	037 51 32.6
136	Gacha	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
137	Chelak	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
138	Geschawul	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
139	Sikar/Sikar	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
140	Kechak	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
141	Kuchi	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
142	Bosar	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
143	Ankmar	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
144	Atsarakay	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
145	Maza	Cultivated	Induced sucker	South Omo	06 02 4.17	037 24 16.7
146	Kawazrr	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
147	Molebaba	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
148	Agino	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
149	Ochaa	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
150	Odaret	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
151	Kumucha	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
152	Godra	Cultivated	Induced sucker	South Omo	06 84 4.49	037 02 41.7
153	Antsa	Cultivated	Induced sucker	South Omo	05 56 03.8	036 40 08.8
154	Siknda	Cultivated	Induced sucker	South Omo	05 56 03.8	036 40 08.8
155	Dalee	Cultivated	Induced sucker	South Omo	05 55 45.4	036 40 08.7
156	Sepa	Cultivated	Induced sucker	South Omo	05 55 45.4	036 40 08.7
157	Gofa	Cultivated	Induced sucker	South Omo	05 56 03.8	036 40 08.8
158	Monet	Cultivated	Induced sucker	South Omo	05 55 45.4	036 40 08.7
159	Kundkush	Cultivated	Induced sucker	South Omo	05 54 19.3	036 38 53.1
160	Beheken	Cultivated	Induced sucker	South Omo	05 54 19.3	036 38 53.1
161	Kumicha	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
162	Tsela	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
163	Wergnmeche	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
164	Arisig	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
165	Zokuma	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
166	Buuka	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
167	Dusak	Cultivated	Induced sucker	South Omo	05 54 10.1	036 38 41.5
168	Gaya	Cultivated	Induced sucker	South Omo	06 01 44.0	036 33 40.6
169	Karta	Cultivated	Induced sucker	South Omo	06 01 44.0	036 33 40.6
170	Bebilus	Cultivated	Induced sucker	South Omo	06 01 44.0	036 33 40.6
171	Mehaka	Cultivated	Induced sucker	South Omo	06 01 44.0	036 33 40.6
172	Garacha	Cultivated	Induced sucker	South Omo	06 01 44.0	036 33 40.6
173	Moset	Cultivated	Induced sucker	South Omo	06 01 08 .9	036 33 25.8
174	Jolak	Cultivated	Induced sucker	South Omo	06 01 08 .9	036 33 25.8
175	Arfa	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2
176	Gena	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2

177	Selta	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2
178	Adaret	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2
179	Asa	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2
180	Chelike	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2
181	Shuferak	Cultivated	Induced sucker	South Omo	05 52 30.4	036 36 52.2
182	Kerta	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
183	Shoka	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
184	Mono	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
185	Zika	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
186	Damet	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
187	Solka	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
188	Kaket	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
189	Golaa	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
190	Bublin	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
191	Salibeli	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
192	Wusasi	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
193	Nodifals	Cultivated	Induced sucker	South Omo	05 53 23 .9	036 37 46.9
194	Kakisa	Cultivated	Induced sucker	South Omo	05 53 28.6	036 36 25.9
195	Bekelo	Cultivated	Induced sucker	South Omo	05 53 28.6	036 36 25.9
196	Ollact	Cultivated	Induced sucker	South Omo	05 53 28.6	036 36 25.9
197	Gufirak	Cultivated	Induced sucker	South Omo	-	-
198	Sikarak	Cultivated	Induced sucker	South Omo	-	-
199	Zergina	Cultivated	Induced sucker	South Omo	05 47 17.9	036 33 10.1
200	Sheha	Cultivated	Induced sucker	South Omo	05 47 17.9	036 33 10.1
201	Tefana	Cultivated	Induced sucker	South Omo	05 47 13.4	036 33 10.1
202	Dama	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
203	Kachak	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
204	Baysametecho	Cultivated	Induced sucker	South Omo	06 84 49.1	037 02 41.67
205	Kayssidak	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
206	Tibla	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
207	Makka	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
208	Molla	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
209	Alka	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
210	Gammi	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
211	Sikki	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
212	Kumlla	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
213	Shiya	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
214	Garchii	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
215	Silver	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
216	Dammi	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
217	Gollect	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
218	Skimma	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
219	Wobajalk	Cultivated	Induced sucker	South Omo	06 84 44.0	037 02 41.67
220	Selti	Cultivated	Induced sucker	South Omo	05 59 01.3	036 34 36.1
221	Zergana	Cultivated	Induced sucker	South Omo	05 59 01.3	036 34 36.1
222	Shehana	Cultivated	Induced sucker	South Omo	05 59 01.3	036 34 36.1

223	Sheferek	Cultivated	Induced sucker	South Omo	05 59 01.3	036 34 36.1
224	Berga	Cultivated	Induced sucker	South Omo	05 55 40.8	036 34 54.3
225	Gerchet	Cultivated	Induced sucker	South Omo	05 55 40.8	036 34 54.3
226	Zoda	Cultivated	Induced sucker	South Omo	05 58 05.4	036 31 34.9
227	Dempa	Cultivated	Induced sucker	South Omo	05 58 05.4	036 31 34.9
228	Gella	Wild	Botanical seed	South Omo	05 58 50.0	036 34 21.4
229	Gella	Wild	Botanical seed	South Omo	05 58 50.0	036 34 21.4
230	Gella	Wild	Botanical seed	South Omo	05 58 49.9	036 34 21.4
231	Gella	Wild	Botanical seed	South Omo	05 59 01.3	036 34 36.1
232	Gella	Wild	Botanical seed	South Omo	05 59 01.3	036 34 36.1
233	Gella	Wild	Botanical seed	South Omo	05 58 05.4	036 31 34.9
234	Gella	Wild	Botanical seed	South Omo	06 01 44.0	036 33 40.6
235	Gella	Wild	Botanical seed	South Omo	06 01 44.0	036 33 40.6
236	Gella	Wild	Botanical seed	South Omo	-	-
237	Gella	Wild	Botanical seed	South Omo	-	-

Supplementary Table 2. Annealing adapters

No.	Annealing adapters	No.	Annealing adapters	No.	Annealing adapters
1	GCATG_EcoRI_P1.1	33	GGATA_EcoRI_P1.1	65	ATTAC_EcoRI_P1.2
2	AACCA_EcoRI_P1.1	34	GGCCA_EcoRI_P1.1	66	CATAT_EcoRI_P1.2
3	CGATC_EcoRI_P1.1	35	GGCTC_EcoRI_P1.1	67	CGAAT_EcoRI_P1.2
4	TCGAT_EcoRI_P1.1	36	GTAGT_EcoRI_P1.1	68	CGGCT_EcoRI_P1.2
5	TGCAT_EcoRI_P1.1	37	GTCCG_EcoRI_P1.1	69	CGGTA_EcoRI_P1.2
6	CAACC_EcoRI_P1.1	38	GTCGA_EcoRI_P1.1	70	CGTAC_EcoRI_P1.2
7	GGTTG_EcoRI_P1.1	39	TACCG_EcoRI_P1.1	71	CGTCG_EcoRI_P1.2
8	AAGGA_EcoRI_P1.1	40	TACGT_EcoRI_P1.1	72	CTGAT_EcoRI_P1.2
9	AGCTA_EcoRI_P1.1	41	TAGTA_EcoRI_P1.1	73	CTGCG_EcoRI_P1.2
10	ACACA_EcoRI_P1.1	42	TATAC_EcoRI_P1.1	74	CTGTC_EcoRI_P1.2
11	AATTA_EcoRI_P1.1	43	TCACG_EcoRI_P1.1	75	CTTGG_EcoRI_P1.2
12	ACGGT_EcoRI_P1.1	44	TCAGT_EcoRI_P1.1	76	GACAC_EcoRI_P1.2
13	ACTGG_EcoRI_P1.1	45	TCCGG_EcoRI_P1.1	77	GAGAT_EcoRI_P1.2
14	ACTTC_EcoRI_P1.1	46	TCTGC_EcoRI_P1.1	78	GAGTC_EcoRI_P1.2
15	ATACG_EcoRI_P1.1	47	TGGAA_EcoRI_P1.1	79	GCCGT_EcoRI_P1.2
16	ATGAG_EcoRI_P1.1	48	TTACC_EcoRI_P1.1	80	GCTGA_EcoRI_P1.2
17	ATTAC_EcoRI_P1.1	49	GCATG_EcoRI_P1.2	81	GGATA_EcoRI_P1.2
18	CATAT_EcoRI_P1.1	50	AACCA_EcoRI_P1.2	82	GGCCA_EcoRI_P1.2
19	CGAAT_EcoRI_P1.1	51	CGATC_EcoRI_P1.2	83	GGCTC_EcoRI_P1.2
20	CGGCT_EcoRI_P1.1	52	TCGAT_EcoRI_P1.2	84	GTAGT_EcoRI_P1.2
21	CGGTA_EcoRI_P1.1	53	TGCAT_EcoRI_P1.2	85	GTCCG_EcoRI_P1.2
22	CGTAC_EcoRI_P1.1	54	CAACC_EcoRI_P1.2	86	GTCGA_EcoRI_P1.2
23	CGTCG_EcoRI_P1.1	55	GGTTG_EcoRI_P1.2	87	TACCG_EcoRI_P1.2
24	CTGAT_EcoRI_P1.1	56	AAGGA_EcoRI_P1.2	88	TACGT_EcoRI_P1.2
25	CTGCG_EcoRI_P1.1	57	AGCTA_EcoRI_P1.2	89	TAGTA_EcoRI_P1.2
26	CTGTC_EcoRI_P1.1	58	ACACA_EcoRI_P1.2	90	TATAC_EcoRI_P1.2
27	CTTGG_EcoRI_P1.1	59	AATTA_EcoRI_P1.2	91	TCACG_EcoRI_P1.2
28	GACAC_EcoRI_P1.1	60	ACGGT_EcoRI_P1.2	92	TCAGT_EcoRI_P1.2
29	GAGAT_EcoRI_P1.1	61	ACTGG_EcoRI_P1.2	93	TCCGG_EcoRI_P1.2
30	GAGTC_EcoRI_P1.1	62	ACTTC_EcoRI_P1.2	94	TCTGC_EcoRI_P1.2
31	GCCGT_EcoRI_P1.1	63	ATACG_EcoRI_P1.2	95	TGGAA_EcoRI_P1.2
32	GCTGA_EcoRI_P1.1	64	ATGAG_EcoRI_P1.2	96	TTACC_EcoRI_P1.2

Supplementary Table 3. Adapter_P1-EcoRI

Name	OligoSequence
GCATG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATG
AACCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCA
CGATC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATC
TCGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAT
TGCAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCAT
CAACC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACC
GGTTG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTG
AAGGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGA
AGCTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTA
ACACA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACACA
AATTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATTA
ACGGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGGT
ACTGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGG
ACTTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTC
ATACG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATACG
ATGAG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGAG
ATTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTAC
CATAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATAT
CGAAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAAT
CGGCT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGCT
CGGTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGTA
CGTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTAC
CGTCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTCG
CTGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGAT
CTGCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGCG
CTGTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTC
CTTGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGG
GACAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACAC
GAGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGAT
GAGTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTC
GCCGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCGT
GCTGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGA
GGATA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGATA
GGCCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCCA
GGCTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTC
GTAGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGT
GTCCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCCG
GTCGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCGA
TACCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCG
TACGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACGT
TAGTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGTA
TATAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTATAC
TCACG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCACG
TCAGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCAGT
TCCGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCCGG
TCTGC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGC
TGGAA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGAA
TTACC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTACC
GCATG_EcoRI_P1.2	/5Phos/AATTCATGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AACCA_EcoRI_P1.2	/5Phos/AATTTGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Supplementary Table 3. Continued …

CGATC_EcoRI_P1.2	/5Phos/AATTGATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCGAT_EcoRI_P1.2	/5Phos/AATTATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGCAT_EcoRI_P1.2	/5Phos/AATTATGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CAACC_EcoRI_P1.2	/5Phos/AATTGGTTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGTTG_EcoRI_P1.2	/5Phos/AATTCAACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AAGGA_EcoRI_P1.2	/5Phos/AATTTCCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AGCTA_EcoRI_P1.2	/5Phos/AATTTAGCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACACA_EcoRI_P1.2	/5Phos/AATTTGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AATTA_EcoRI_P1.2	/5Phos/AATTTAATTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACGGT_EcoRI_P1.2	/5Phos/AATTACCGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTGG_EcoRI_P1.2	/5Phos/AATTCCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTTC_EcoRI_P1.2	/5Phos/AATTGAAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATACG_EcoRI_P1.2	/5Phos/AATTCGTATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATGAG_EcoRI_P1.2	/5Phos/AATTCTCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATTAC_EcoRI_P1.2	/5Phos/AATTGTAATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CATAT_EcoRI_P1.2	/5Phos/AATTATATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGAAT_EcoRI_P1.2	/5Phos/AATTATTCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGCT_EcoRI_P1.2	/5Phos/AATTAGCCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGTA_EcoRI_P1.2	/5Phos/AATTTACCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTAC_EcoRI_P1.2	/5Phos/AATTGTACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTCG_EcoRI_P1.2	/5Phos/AATTCGACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGAT_EcoRI_P1.2	/5Phos/AATTATCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGCG_EcoRI_P1.2	/5Phos/AATTCGCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGTC_EcoRI_P1.2	/5Phos/AATTGACAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTTGG_EcoRI_P1.2	/5Phos/AATTCCAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GACAC_EcoRI_P1.2	/5Phos/AATTGTGTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGAT_EcoRI_P1.2	/5Phos/AATTATCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGTC_EcoRI_P1.2	/5Phos/AATTGACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCCGT_EcoRI_P1.2	/5Phos/AATTACGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCTGA_EcoRI_P1.2	/5Phos/AATTTCAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGATA_EcoRI_P1.2	/5Phos/AATTTATCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCCA_EcoRI_P1.2	/5Phos/AATTTGGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCTC_EcoRI_P1.2	/5Phos/AATTGAGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTAGT_EcoRI_P1.2	/5Phos/AATTACTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCCG_EcoRI_P1.2	/5Phos/AATTCGGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCGA_EcoRI_P1.2	/5Phos/AATTTCGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACCG_EcoRI_P1.2	/5Phos/AATTCGGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACGT_EcoRI_P1.2	/5Phos/AATTACGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TAGTA_EcoRI_P1.2	/5Phos/AATTTACTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TATAC_EcoRI_P1.2	/5Phos/AATTGTATAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCACG_EcoRI_P1.2	/5Phos/AATTCGTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCAGT_EcoRI_P1.2	/5Phos/AATTACTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCCGG_EcoRI_P1.2	/5Phos/AATTCCGGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCTGC_EcoRI_P1.2	/5Phos/AATTGCAGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGGAA_EcoRI_P1.2	/5Phos/AATTTTCCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TTACC_EcoRI_P1.2	/5Phos/AATTGGTAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Supplementary Table 4. Adapter_P2-MspI

Name	OligoSequence
MspI_P2.1	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
MspI_P2.2	/5Phos/CGAGATCGGAAGAGCGAGAACAA

Supplementary Table 5. Adapter_P1-flex

Name	OligoSequence
GCATG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATGCATG
AACCA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCACATG
CGATC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATCCATG
TCGAT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGATCATG
TGCAT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCATCATG
CAACC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACCCATG
GGTTG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTGCATG
AAGGA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGACATG
AGCTA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTACATG
ACACA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACACACATG
AATTA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATTACATG
ACGGT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGGTCATG
ACTGG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGGCATG
ACTTC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTCCATG
ATACG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATACGCATG
ATGAG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGAGCATG
ATTAC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTACCATG
CATAT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATATCATG
CGAAT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAATCATG
CGGCT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGCTCATG
CGGTA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGTACATG
CGTAC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTACCATG
CGTCG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTCGCATG
CTGAT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGATCATG
CTGCG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGCGCATG
CTGTC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTCCATG
CTTGG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGGCATG
GACAC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACACCATG
GAGAT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGATCATG
GAGTC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTCCATG
GCCGT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCGTCATG
GCTGA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGACATG
GGATA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGATACATG
GGCCA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCCACATG
GGCTC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTCCATG
GTAGT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGTCATG
GTCCG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCCGCATG
GTCGA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCGACATG
TACCG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCGCATG
TACGT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACGTCATG
TAGTA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGTACATG
TATAC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTATACCATG
TCACG_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCACGCATG

Supplementary Table 5. Continued …

TCAGT_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCAGTCATG
TCTGC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGCCATG
TGGAA_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGAACATG
TTACC_flex_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTACCCATG
GCATG_flex_P1.2	/5Phos/CATGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AACCA_flex_P1.2	/5Phos/TGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGATC_flex_P1.2	/5Phos/GATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCGAT_flex_P1.2	/5Phos/ATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGCAT_flex_P1.2	/5Phos/ATGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CAACC_flex_P1.2	/5Phos/GGTTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGTTG_flex_P1.2	/5Phos/CAACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AAGGA_flex_P1.2	/5Phos/TCCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AGCTA_flex_P1.2	/5Phos/TAGCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACACA_flex_P1.2	/5Phos/TGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AATTA_flex_P1.2	/5Phos/TAATTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACGGT_flex_P1.2	/5Phos/ACCGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTGG_flex_P1.2	/5Phos/CCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTTC_flex_P1.2	/5Phos/GAAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATACG_flex_P1.2	/5Phos/CGTATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATGAG_flex_P1.2	/5Phos/CTCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATTAC_flex_P1.2	/5Phos/GTAATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CATAT_flex_P1.2	/5Phos/ATATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGAAT_flex_P1.2	/5Phos/ATTCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGCT_flex_P1.2	/5Phos/AGCCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGTA_flex_P1.2	/5Phos/TACCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTAC_flex_P1.2	/5Phos/GTACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTCG_flex_P1.2	/5Phos/CGACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGAT_flex_P1.2	/5Phos/ATCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGCG_flex_P1.2	/5Phos/CGCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGTC_flex_P1.2	/5Phos/GACAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTTGG_flex_P1.2	/5Phos/CCAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GACAC_flex_P1.2	/5Phos/GTGTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGAT_flex_P1.2	/5Phos/ATCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGTC_flex_P1.2	/5Phos/GACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCCGT_flex_P1.2	/5Phos/ACGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCTGA_flex_P1.2	/5Phos/TCAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGATA_flex_P1.2	/5Phos/TATCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCCA_flex_P1.2	/5Phos/TGGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCTC_flex_P1.2	/5Phos/GAGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTAGT_flex_P1.2	/5Phos/ACTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCCG_flex_P1.2	/5Phos/CGGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCGA_flex_P1.2	/5Phos/TCGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACCG_flex_P1.2	/5Phos/CGGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACGT_flex_P1.2	/5Phos/ACGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TAGTA_flex_P1.2	/5Phos/TACTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TATAC_flex_P1.2	/5Phos/GTATAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCACG_flex_P1.2	/5Phos/CGTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCAGT_flex_P1.2	/5Phos/ACTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCCGG_flex_P1.2	/5Phos/CCGGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCTGC_flex_P1.2	/5Phos/GCAGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGGAA_flex_P1.2	/5Phos/TTCCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TTACC_flex_P1.2	/5Phos/GGTAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Supplementary Table 6. Adapter_P2-flex

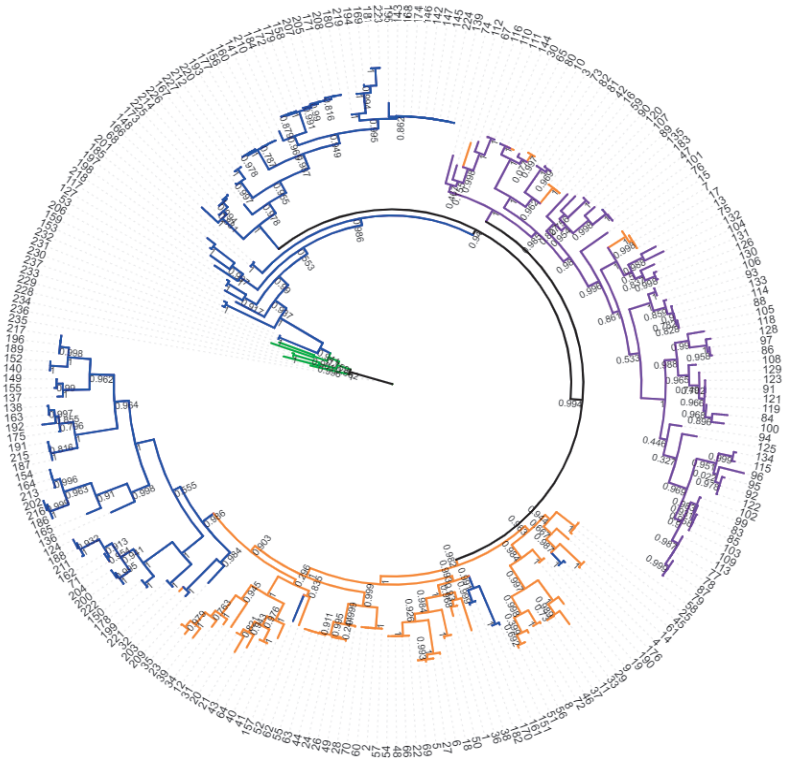
Name	OligoSequence
flex_P2.1	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
flex_P2.2	/5Phos/AATTAGATCGGAAGAGCGAGAACAA

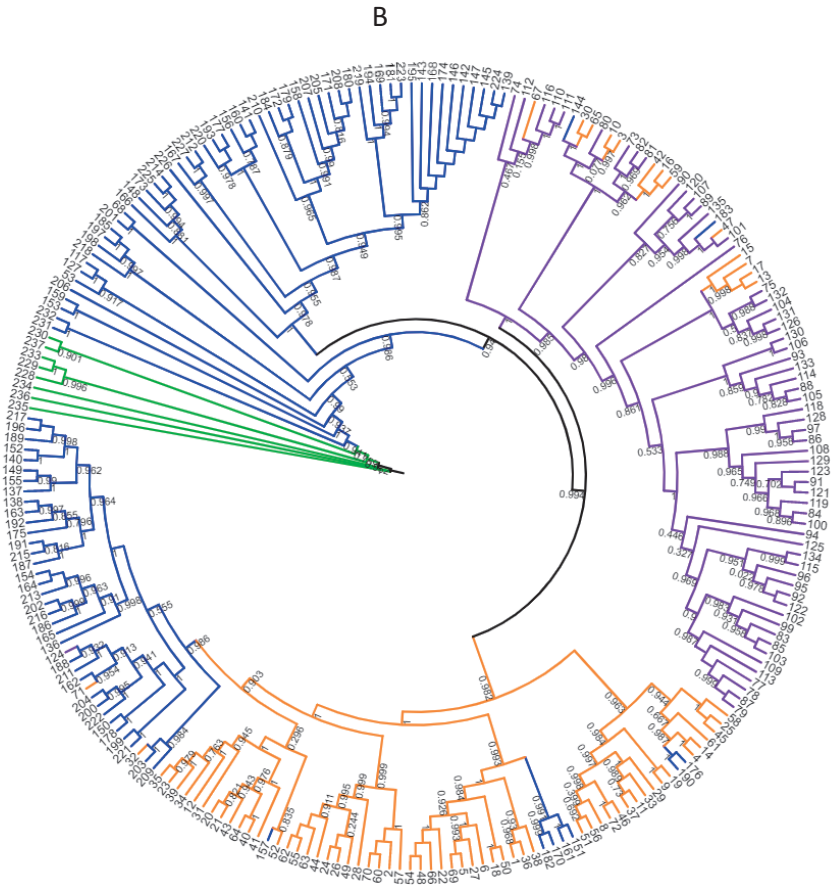
Supplementary Table 7. PCR-primers

Name	OligoSequence
PCR1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG
PCR2_Idx_1_ATCACG	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_2_CGATGT	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_3_TTAGGC	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_4_TGACCA	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_5_ACAGTG	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_6_GCCAAT	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_7_CAGATC	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_8_ACTTGA	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_9_GATCAG	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_10_TAGCTT	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_11_GGCTAC	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_12_CTTGTA	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGC

Figures

A





Supplementary Figure 1. *Phylogenic tree; A: Maximum-likelihood phylogenetic tree with branch length and bootstrap values.; B: Topological view of the maximum-likelihood phylogenetic tree. Accessions are numbered following the Supplementary Table 1 and colored according to their geographical origins and cultivation status: Blue; South Omo: Orange; Sidama: Purple; Gurage: Green; Wild enset.*

Paper III

Limited genetic diversity found in the Entada landrace (*Ensete ventricosum*, (Welw.) Chessman, var. Entada) from Ethiopia

Alye Tefera Haile[1, 2*], Mallikarjuna Rao Kovi[1], Sylvia Sagen Johnsen[1], Trine Hvoslef-Eide[1], Bizuayehu Tesfaye[2], Odd Arne Rognli[1*]

[1]Faculty of Biosciences, Department of Plant Sciences, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Ås, Norway

[2]School of Plant and Horticultural Science, College of Agriculture, Hawassa University, P.O. Box 05, Hawassa, Ethiopia

*Corresponding Author's E-mail: odd-arne.rognli@nmbu.no

Abstract

The Entada landrace (*Ensete ventricosum* (Welw.) Chessman, var. Entada) is probably the most unique indigenous crop in Ethiopia, being maintained and utilized by the Ari people in the South of Ethiopia. Here we describe genetic diversity, signatures selection and relationship of Entada with cultivated and wild enset using 117 Entada landraces collected from three Entada growing regions in Ethiopia (Sidama, South and North Ari). A total number of 1617 high-quality SNP markers, obtained from ddRAD-sequences, were used for the diversity studies. Entada formed a completely separated clade and group in the phylogenetic and PCA analyses of Entada, cultivated and wild enset genotypes. The principal component analysis (PCA) clearly differentiated between landraces from Sidama and North Ari, however, no differentiation was observed between South and North Ari. Very little molecular variation was detected between regions (0.48%), and nearly all variation was present within individuals (99.5%). Subpopulation differentiation between regions (pairwise F_{ST}) was zero, and observed heterozygosity was 0.99, which is expected of plants with strict asexual propagation. Prolonged clonal propagation of heterozygous genotypes from a single or few founding lineages will lead to populations with very little or no diversity between genotypes, and extremely high heterozygosity, as found in this study (H_O=0.99). Signatures of directional selection were identified at eight loci based on an F_{ST} outlier analysis. The candidate genes detected are involved in axillary shoot growth and might be involved in controlling natural sucker formation in Entada.

Keywords:

Entada (*Ensete ventricosum* (Welw.) Cheesman var. Entada), SNP markers, ddRAD, Outlier SNPs, Genetic diversity

Introduction

Most of the plant landraces are cultivated and maintained by smallholder farmers and private gardeners all over the world. Numerous landraces have originated as a result of agriculture and horticulture over the past 10,000 years (Zeven, 1998). Landraces are commonly considered as locally adapted and endemic to a specific area. Many of the plant species that are cultivated for food are neglected and underutilized despite their crucial role in the food security, nutrition and income generation of rural societies (Magbagbeola *et al*., 2010). Rural communities prepare food and different products such as medicine, shelter, feed, and fuel from different orphan crops. Mostly, underutilized crops make up a significant part of the diet of rural households, typically during periods of drought, famine and dry seasons (Campbe, 1987).

Landraces are geographically and ecologically distinctive populations (Brown, 1978), which are highly diverse containing a mixture of genotypes (Hawkes, 1983). They often have comparative advantages over commercial cultivars because they have been selected to survive stressful conditions and can be cultivated using low input and/or organic cultivation methods. Moreover, there is an increasing interest in finding new food sources to alleviate malnutrition. The local landraces constitute valuable germplasm for plant breeding and it is important to conserve the genetic diversity present in landraces (Kölliker *et al*., 2003).

In Ethiopia, Entada landrace (*Ensete ventricosum* (Welw.) Cheesman var. Entada) is probably the most unique and understudied indigenous landrace. Entada is an enset landrace that has lost apical dominance present in enset; therefore, it is propagated by natural suckers similar to banana (Shigeta, 1990). The local or farmers name "Intada" actually indicates that the plant grow or multiply by itself (Shigeta, 1992). Entada has the ability to flower and set fruits, but it is different from all other cultivated and wild enset both morphologically and genetically, which do not produce and propagate with suckers (Olango *et al*., 2015; Shigeta, 1992). Landraces of Entada are being maintained and utilized mainly by the Ari people in the Southern regions of Ethiopia (Shigeta, 1992). The major processed foods from Entada are Amicho prepared from the underground corm (the underground base of the stem that serves as a storage organ). The fresh corm is cooked like potato and yam. It is a multi-purpose crop with cultural values in religious ceremonies and as a feed source in addition to being a food. Due to its tolerance to adverse factors such as drought and different soil types, this landrace is considered an alternative crop for areas with extreme growing conditions (Brandt *et al*., 1997; Shigeta, 1992). However, so far there is no genetic diversity studies among and within Entada landrace (Olango *et al*., 2015; Shigeta, 1992).

A previous molecular diversity study of enset using SSR (microsatellites) and SNP (single nucleotide polymorphism) markers revealed that Entada belongs to the genus *Ensete* of the Musaceae family (Olango *et al*., 2015; Yemataw *et al*., 2018). Currently, several new molecular techniques are being applied together with phenotypic descriptions to investigate genetic diversity and relatedness in enset accessions (Birmeta, *et al*., 2004). Single nucleotide polymorphisms (SNPs) and simple sequence repeats (SSRs) are the most common DNA markers for genetic diversity studies (Tsykun *et al*., 2017). Among all DNA markers, SNPs are abundant and robust markers, which are well suited for automated high-throughput genotyping of large numbers of samples. Besides, SNPs are able to resolve the differences among extremely similar individuals and increase the accuracy of diversity estimates (Hinze *et al*., 2017).

Double digest restriction-site associated DNA (ddRADseq) is one of the reduced representation sequencing methods for SNP discovery at a genome-wide scale. It includes digesting the DNA with two restriction enzymes to allow greater control of the genomic regions sampled for sequencing and more reproducible recovery of sequenced regions. Specific size-selected fragments are generated and sequenced (Peterson *et al.*, 2012). ddRADseq eliminates random the shearing step of the original RAD protocol (Baird *et al*., 2008). The application of the ddRADseq technology has been successfully applied in many plant species such as tomato (Esposito *et al*., 2020), strawberry (Davik *et al*., 2015), northern red oak (Konar *et al.,* 2017), and Oriental thuja (*Platycladus orientalis*) (Jin *et al*., 2019). Here, we report on the development and utilization of the first set of SNP markers developed from ddRAD sequences to study genetic diversity and relatedness among Entada landraces, cultivated and wild enset collected from the Entada growing regions in Ethiopia. To facilitate progress in future Entada conservation and breeding, it is essential to discover and characterize Entada and understand its relationships to cultivated and wild enset populations. In the present study we applied SNP markers to: (1) study genetic diversity and relationship of Entada landraces with cultivated and wild enset, (2) determine the effects of clonal propagation on genetic diversity, and (3) identify candidate genes involved in sucker formation in Entada.

Materials and Methods

Plant Samples and Treatments

Leaf tissue from 129 Entada genotypes were collected from South Ari (91), North Ari (17) and Sidama (21) regions in Ethiopia (**Figure 1**). The saturated NaCl-CTAB solution was used to preserve the Entada leaf samples upon collection, as described by Rogstad (1992), with minor modifications. Briefly, 550 g NaCl was added to 1 L of water, boiled, and cooled at ambient temperature, and mixed thoroughly until the salt precipitated. Then, 35 g of CTAB was added gradually with intermittent irregular intervals mixing, until the solution became viscous. 35-40 mL of the prepared solution was aliquoted into 50 mL Falcon tubes and used for preservation of tissue samples. A pair of scissors was used to remove leaf samples from the mother plants, and the scissors were cleaned with ethanol (96%) between independent samples. Fresh cigar-leaf samples harvested from each genotype were stored immediately in the 50 mL tubes containing the saturated NaCl-CTAB preservation buffer. Samples were then placed in a black plastic bag and stored in a dark room at ambient temperature, to preserve genomic DNA from degradation during transportation from the farmer fields in Ethiopia to the laboratory in Norway. Upon arrival at the laboratory, the saturated NaCl-CTAB solution was washed off thoroughly with deionized water and excess water wiped off the leaves with dry white wipes (Kimberly-Clark™ Professional Kimtech Science™). Leaf samples were put in liquid nitrogen, ground quickly using a pestle and mortar, and the ground powder transferred into 2 mL microcentrifuge tubes (Eppendorf A.G., Hamburg, Germany). Pestles and mortars were washed and dried before starting each sample preparation, and all pulverized leaf samples stored at -80 ˚C until further analyses. For DNA extraction, 100 mg of pulverized leaf material of each sample preserved in NaCl-CTAB was used.

DNA Extraction and Quantification

DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). Genomic DNA quality and quantity were checked by 1% (w/v) agarose gel electrophoresis and using a NanoDrop spectrophotometer. Lastly, the DNA concentration was measured with the Qubit® dsDNA BR assay kit (Q) and Quant-iT™ PicoGreen™ (Life Sciences) dsDNA assay.

Double-digest Restriction-site-associated DNA (ddRAD) Library Preparation and llumina Sequencing

The ddRAD procedure used in this study was modified from (Peterson *et al.*, 2012). We calculated the number of reads required for 20X coverage of restriction fragments in the 150–500bp size range across 10 multiplexed individuals using multiple enzyme pairs, assuming a GC content of 0.44, to ensure that restriction fragments could feasibly be sequenced with enough coverage on an Illumina MiSeq platform. 500ng of each DNA sample was double digested using *EcoR1* and *MseI* restriction endonucleases, and unique P1 barcode adapters were ligated to the digested fragments from each sample while a common P2 barcode adapter was ligated to fragments from all samples (For information about adapters and primers, see Supplementary tables 1-3). Samples containing unique P1 barcodes were pooled, and the Sage Science Blue Pippin system (www.sagescience.com) was used to select fragments of about 500bp. Size-selected libraries were bound to Dynabeads® M-270 Streptavidin magnetic beads (Invitrogen) to eliminate fragments without the P2 adapter, and the libraries amplified by PCR using Phusion™ Polymerase kit (Invitrogen) and index-marked primers for further tagging of the samples (Supplementary Table 3). The libraries were analyzed using an Agilent 2100 BioAnalyser and diluted to a concentration of 35nM for sequencing using the V2 sequencing kit on the MiSeq platform (Illumina). The sequencing was performed at the Norwegian University of Life Sciences, Ås, Norway.

Sequence Data Analysis and SNP Calling

The ddRAD sequence data obtained was quality checked using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). High quality reads were retained after trimming the bad quality reads using the Trimmomatic program (Bolger *et al*., 2014). The SNP calling was performed using the STACKS 2 program (Rochette *et al*., 2019). The obtained SNPs were further quality filtered according to the following criteria: (1) variants should be bi-allelic SNPs, (2) SNPs having more than 20% missing information were excluded, (3) genotypes having more than 20% missing information were excluded, and (4) markers with minor allele frequency (MAF; MAF > 0.05) were retained. We identified 1617 high quality SNPs across 117 Entada landraces, since 12 of the 129 Entada genotypes collected turned out to be cultivated enset with suckers (see results section).

Genetic Diversity and Cluster Analyses

Genetic variation among and within populations/regions, observed (H_o) and expected (H_e) heterozygosity, and pairwise fixation indices (F_{ST}) for the subpopulations (Weir and Cockerham, 1984) was estimated by analysis of molecular variance (AMOVA) using Arlequin v.3.5 (Excoffier and Lischer, 2010). To examine the relationship among Entada landraces, and cultivated and wild enset, principal component analyses (PCA) were performed using TASSEL v5.2 (Bradbury *et al*., 2007) and maximum-likelihood (ML) phylogenetic

tree analyses performed using PhyML 3.0 (Guindon *et al*., 2010). The trees were prepared and visualized using the iTOL v4 online tool (Letunic and Bork, 2019). PCAs were graphically summarized using scatter plot.

F_{ST} Outlier Analysis to Detect Candidate Genes for Sucker Development

The primary distinction between the cultivation of enset and Entada lies in the growth of suckers. While Entada generates them naturally, enset necessitates their induction. To identify potential genes involved in sucker formation, we performed genome scans using F_{ST} outlier analysis. We used the hierarchical method (Excoffier *et al*., 2009), a modified approach of Beaumont and Nichols (1996) implemented in the ARLEQUIN software package version 3.5.1.3 (Excoffier and Lischer, 2010) to detect loci under directional selection. We conducted hierarchical island model simulations on two populations (enset and Entada) with 50,000 simulations to generate the joint distribution of F_{ST} versus heterozygosity. Loci that fall outside the 99% confidence intervals of the distribution were identified as outliers being putatively under selection. The putative function of genes with outlier SNPs was identified using the Gene Ontology (GO) annotation using Blast2GO software tool version 3.0 (Conesa *et al*., 2005).

Results

Phylogenetic and Principal Component Analysis (PCA)

Previously we developed 2,823 high quality SNP markers which were polymorphic in both cultivated enset (226 genotypes), wild enset (10 genotypes) and the 129 Entada landraces. The joint phylogenetic analysis involving all enset genotypes grouped the cultivated enset, wild enset and Entada landraces into three distinct clusters (**Figure 2).** However, 12 genotypes collected as Entada landraces cluster with cultivated enset. These genotypes produce suckers and resemble Entada phenotypically. The PCA analyses confirm that Entada is completely different from other ensets, while cultivated enset, cultivated enset with suckers and wild enset group together (**Figure 3A, B)** with PC1, PC2 and PC3 accounting for 59.12, 10.71, and 17.23% of the variation, respectively. Phylogenetic analysis of the Entada landraces showed several clusters/branches (**Figure 4A, B**). One landrace from Sidama is clearly separated from other clusters. Ten landraces from South Ari and two from North Ari clustered together. Several landraces from South Ari, North Ari and Sidama are clustered together. The phylogenetic analysis didn't show clear differences between the regions. However, the PCA analysis showed that Entada from the Sidama and North Ari regions was completely separated, while Entada from Sidama partly overlapped with Entada from South Ari (**Figure 5A, B**). PC1, PC2 and PC3 explain 62.84, 21, and 16% of the variation, respectively.

Genetic Diversity and Differentiation of Entada Populations

AMOVA analysis of population differentiation was performed between and within the three populations/regions, i.e., the growing regions Sidama, South Ari and North Ari. The AMOVA results show that nearly all variation is present within individuals (99.52%), only very little between populations/regions (0.48%) and none among individuals within populations (**Table 1**). Observed heterozygosity (H_o) was 0.99 in all the three regions, while expected heterozygosity (H_e) was around 0.50 in all three regions (**Table 2**). No significant differentiation was detected between populations/regions

based on pairwise F_{ST} values which ranged from -0.04 to -0.02 (**Table 3)**. The differentiation (F_{ST}) between the three populations/regions was 0.005.

Loci Under Selection

We employed a F_{ST} outlier method; hierarchical structure model (Excoffier *et al*., 2009), using Arlequin to detect true positive loci under selection by comparing cultivated enset with Entada landraces (**Figure 6**). We identified eight candidate loci under positive selection based on F_{ST} values that displayed differentiation higher than the 99% limit of the confidence interval (**Figure 6 and Table 4**). Among the eight loci, four have putative gene functions, i.e., *Lateral suppressor protein, Auxin response factor 2A, Cytokinin dehydrogenase,* and *Scarecrow-like protein 18* (**Table 4).**

Discussion

Genetic Relationship and Diversity

This study shows that the Entada landraces are clearly different from all the other cultivated and wild enset both morphologically and genetically. The genetic differentiation of cultivated and wild enset, can to a large degree be explained by cultivation status and reproduction methods (Birmeta *et al.,* 2004; Gerura *et al.,* 2019; Olango *et al.,* 2015). Wild enset regenerates from botanical seeds and hence lack spontaneous suckers, while cultivated enset and Entada landraces are propagated by suckers. However, formation of suckers is different; suckers are induced in cultivated enset, while in Entada suckers are formed spontaneously (Bekele and Shigeta, 2011; Olango *et al.,* 2015). In addition to these differences in reproduction, various factors like long-term evolutionary history, genetic drift, gene flow and selection (Godoy *et al*., 2018; Schaal *et al*., 1998) can explain the genetic structure evident from the phylogenetic analyses. The results of the phylogenetic analysis show a clear distinction between enset, Entada and wild. However, 12 of the Entada landraces clustered together with cultivated enset. Morphological characters observed in the field confirms that these 12 Entada landraces are similar to the cultivated enset with natural suckers (**Figure 2**).

We found insignificant molecular variation (0.48%) between populations/regions and none among individuals within populations/regions. These results are in agreement with studies on plantain, belonging to the same plant family, using SSR markers (Cyrille *et al*., 2019; Quain *et al*., 2018). However, we detected very high molecular variation (99.5%) within individuals, indicating that complete clonal propagation is frequent, and sexual reproduction is absent. The high level of molecular diversity within individuals may also be partially maintained through clonal determination of clones from the founding population, as they share a common history within a clonal lineage (Balloux *et al*., 2003; Hangelbroek *et al*., 2002). According to some studies, absence of sex will promote divergence between alleles within loci, as the two copies will accumulate different mutations over time (Judson and Normark, 1996; Rousset, 2002). Overall, the pattern of genetic structure and diversity observed in Entada confirms that Entada landraces are clonal plants with the dominance of vegetative offspring. Knowledge about genetic diversity in landraces of Entada can be used in designing conservation strategies and maximizing its use in breeding programs.

Genotypic Diversity in Entada Landraces

Balloux *et al*., (2003) simulated the effect of clonal or partial clonal reproduction on the population genetics of neutral markers in diploid organisms. They found that high rates of clonal reproduction increase heterozygosity, maintain higher genetic diversity at each single locus while the genotypic diversity is reduced, resulting in F_{ST} values being drastically reduced. Population size also increases towards extreme since the polymorphism is protected within individuals due to fixed heterozygosity. Our observations are very much in line with these theoretical predictions. We found that observed heterozygosity (H_o) was extremely high (0.99) in the Entada landraces, expected heterozygosity (H_e) was intermediate (0.50), and F_{ST} was zero. Overall, Entada landraces had very low genotypic diversity, probably due to the origin of Entada landraces from one or a few mutants that has been picked up by farmers and propagated and spread across farms. Thus, our results suggest that Entada has originated from one or a few clones.

Identification of candidate genes responsible for suckers Development

F_{ST} outlier approaches have applied in studies in many crops, i.e., tomato (Sim *et al*., 2011), European beech (Laura *et al*., 2018), soybean (Li *et al*., 2014), banana (Sardos *et al*., 2022), and common bean (Papa *et al*., 2007) for identifying adaptive differentiations. Markers detected in these crops are mapped to the genomic regions with known QTL/genes related to domestication. These loci may be directly under selection. Putative function of the candidate SNP loci detected in the present study revealed that at least four genes with known function are involved in axillary shoot formation. Genes involved in axillary meristem development have been studied in various plant species. Mutations in the *Lateral suppressor* genes in *Arabidopsis* (LAS) (Greb *et al*., 2003) and in tomato (*Ls*) (Schumacher *et al*., 1999) inhibits axillary shoot formation during the vegetative phase. Further, it is well known that the phytohormones, auxin and cytokinin interact to regulate many plants growth and developmental processes (Schaller *et al*., 2015). Functional characterization of some auxin response factors based on the phenotypes of the loss-of-function and gain-of-function mutants showed abnormal abscission of the floral organs (Ellis *et al*., 2005) and impaired hypocotyl elongation and auxin homeostasis (Goetz *et al*., 2006), while the tomato *SlARF3* was found to participate in the formation of trichomes and epidermal cells (Zhang *et al.,* 2015). The cytokinins, which are positive regulators of shoot growth and negative regulators of root growth (Werner *et al*., 2003), are implicated in the control of the shoot architecture (Han *et al*., 2014). Thus, candidate genes detected in this study are involved in axillary growth and they might have important influences on the natural and induced sucker formations in Entada and enset populations.

Conclusion

To the best of our knowledge, this study is the first application of SNP markers to study molecular diversity in landraces of Entada in Ethiopia. The results clearly demonstrate that there is no clonal diversity among Entada landraces from the three regions in Ethiopia. It confirms that Entada is naturally propagated by spontaneous suckers. Low genetic diversity and the little structuring of genetic variation between regions provide an important basis for developing strategies for the conservation of this subspecies in the active germplasm bank in Ethiopia. Furthermore, we identified genes that probably are involved in axillary shoot

growth crucial for sucker formation. Knowledge about the genetic regulation of sucker formation is important for characterization of germplasm and for developing breeding strategies.

Acknowledgements

The authors thank the Norwegian Agency for Development Cooperation (NORAD) for funding this study, which is part of the project 'Controlling Disease in Sweet potato and Enset in South Sudan and Ethiopia to Improve Productivity and Livelihoods under Changing Climatic Conditions using Modern Technologies' under the NORHED program (Agreement no. ETH-13/0017).

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Balloux, F., Lehmann, L., de Meeûs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics 164*(4), 1635-1644. https://doi.org/doi.org/10.1093/genetics/164.4.1635

Beaumont, M. A., Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*(1377), 1619-1626. https://doi.org/doi.org/10.1098/rspb.1996.0237

Bekele, E., Shigeta, M. (2011). Phylogenetic relationships between Ensete and Musa species as revealed by the trnT trnF region of cpDNA. *Genetic Resources and Crop Evolution*, *58*(2), 259-269. https://doi.org/10.1007/s10722-010-9568-2

Birmeta, G., Nybom, H., Bekele, E. (2004). Distinction between wild and cultivated enset (*Ensete ventricosum*) gene pools in Ethiopia using RAPD markers. *Hereditas*, *140*(2), 139-148. https://doi.org/10.1111/j.1601-5223.2004.01792.x

Bolger, A. M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/doi.org/10.1093/bioinformatics/btu170

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633-2635. https://doi.org/doi:10.1093/bioinformatics/btm308

Brandt, S., Spring, A., Hiebsch, C., McCabe, J., Tabogie, E., Diro, M., Wolde-Michael, G., Yntiso, G., Shigeta, M., Tesfaye, S. (1997). *The "Tree Against Hunger": Enset-based agricultural systems in Ethiopia*. (Vol. 56). American Association for the Advancement of Science, Washington, DC, USA.

Brown, A. (1978). Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet*, *52*(4), 145-157. https://doi.org/DOI:10.1007/BF00282571

Campbe, B. (1987). The use of wild fruits in Zimbabwe. *Econ Bo*, *41*(3), 375-385. https://doi.org/doi.org/10.1007/BF02859054

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674-3676. https://doi.org/10.1093/bioinformatics/bti610

Cyrille, K. K. G., Désiré, P., da Sylvère, S., Fulgence, T. D. E. (2019). Genetic diversity and structure of plantain (Musa sapientum L.) landraces from Côte d'Ivoire using SSR markers. *European Journal of Biotechnology and Bioscience*, *7*(1), 36-43.

Ellis, C. M., Nagpal, P., Young, J. C., Hagen, G., Guilfoyle, T. J., Reed, J. W. (2005). AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in Arabidopsis thaliana. https://doi.org/doi.org/10.1242/dev.02012

Excoffier, L., Hofer, T., Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*(4), 285-298. https://doi.org/10.1038/hdy.2009.74

Excoffier, L., Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Gerura, F. N., Meressa, B. H., Martina, K., Tesfaye, A., Olango, T. M., Nasser, Y. (2019). Genetic diversity and population structure of enset (Ensete ventricosum Welw Cheesman) landraces of Gurage zone, Ethiopia. *Genetic Resources and Crop Evolution*, *66*(8), 1813-1824. https://doi.org/10.1007/s10722-019-00825-2

Godoy, F. M. d. R., Lenzi, M., Ferreira, B. H. D. S., Silva, L. V. D., Zanella, C. M., Paggi, G. M. (2018). High genetic diversity and moderate genetic structure in the self-incompatible, clonal Bromelia hieronymi (Bromeliaceae). *Botanical Journal of the Linnean Society*, *187*(4), 672-688. https://doi.org/doi.org/10.1093/botlinnean/boy037

Goetz, M., Vivian-Smith, A., Johnson, S. D., Koltunow, A. M. (2006). AUXIN RESPONSE FACTOR8 is a negative regulator of fruit initiation in Arabidopsis. *The Plant Cell*, *18*(8), 1873-1886. https://doi.org/doi.org/10.1105/tpc.105.037192

Greb, T., Clarenz, O., Schäfer, E., Müller, D., Herrero, R., Schmitz, G., Theres, K. (2003). Molecular analysis of the LATERAL SUPPRESSOR gene in Arabidopsis reveals a conserved control mechanism for axillary meristem formation. *Genes Development*, *17*(9), 1175-1187. https://doi.org/doi/10.1101/gad.260703.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307-321. https://doi.org/10.1093/sysbio/syq010

Han, Y., Yang, H., Jiao, Y. (2014). Regulation of inflorescence architecture by cytokinins. *Frontiers in Plant Science*, *5*, 669. https://doi.org/doi.org/10.3389/fpls.2014.00669

Hangelbroek, H., Ouborg, N., Santamaría, L., Schwenk, K. (2002). Clonal diversity and structure within a population of the pondweed Potamogeton pectinatus foraged by Bewick's swans. *Molecular Ecology*, *11*(10), 2137-2150. https://doi.org/10.1046/j.1365-294X.2002.01598.x

Hawkes, J. G. (2013). *The diversity of crop plants*. Harvard University Press. https://doi.org/doi.org/10.4159/harvard.9780674183551

Hinze, L. L., Hulse-Kemp, A. M., Wilson, I. W., Zhu, Q.-H., Llewellyn, D. J., Taylor, J. M., Spriggs, A., Fang, D. D., Ulloa, M., Burke, J. J. (2017). Diversity analysis of cotton (Gossypium hirsutum L.) germplasm using the CottonSNP63K Array. *BMC Plant Biology*, *17*(1), 37. https://DOI:10.1186/s12870-017-0981-y

Judson, O. P., Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology and Evolution*, *11*(2), 41-46. https://doi.org/https://doi.org/10.1016/0169-5347(96)81040-8

Kölliker, R., Herrmann, D., Boller, B., Widmer, F. (2003). Swiss Mattenklee landraces, a distinct and diverse genetic resource of red clover (Trifolium pratense L.). *Theoretical and Applied Genetics*, *107*(2), 306-315. https://doi.org/10.1007/s00122-003-1248-6

Laura, C.-A., Matthias, A., Markus, M., Christoph, S., Reiner, F., Krutovsky, K. V. (2018). Genetic variation and signatures of natural selection in populations of European beech (Fagus sylvatica L.) along precipitation gradients. https://doi.org/DOI:10.1007/s11295-018-1297-2

Letunic, I., Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*, 256-259. https://doi.org/doi.org/10.1093/nar/gkz239

Li, Y.-H., Reif, J. C., Jackson, S. A., Ma, Y.-S., Chang, R.-Z., Qiu, L.-J. (2014). Detecting SNPs underlying domestication-related traits in soybean. *BMC Plant Biology*, *14*(1), 1-8. https://doi.org/doi:10.1186/s12870-014-0251-1

Magbagbeola, J., Adetoso, J., Owolabi, O. (2010). Neglected and underutilized species (NUS): a panacea for community focused development to poverty alleviation/poverty reduction in Nigeria. *Journal of Economics and International Finance*, *2*(10), 208-211.

Olango, T. M., Tesfaye, B., Pagnotta, M. A., Pè, M. E., Catellani, M. (2015). Development of SSR markers and genetic diversity analysis in enset (Ensete ventricosum (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. *BMC Genetics*, *16*(1), 98. https://doi.org/DOI 10.1186/s12863-015-0250-8

Papa, R., Bellucci, E., Rossi, M., Leonardi, S., Rau, D., Gepts, P., Nanni, L., Attene, G. . (2007). Tagging the signatures of domestication in common bean (Phaseolus vulgaris) by means of pooled DNA samples. *Annals of Botany*, *100*(5), 1039-1051. https://doi.org/doi.org/10.1093/aob/mcm151

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., Hoekstra, H. E. . (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135. https://doi.org/doi:10.1371/journal.pone.0037135

Quain, M. D., Agyeman, A., Dzomeku, B. M. (2018). Assessment of plantain (Musa sapientum L.) accessions genotypic groups relatedness using simple sequence repeats markers. *African Journal of biotechnology*, *17*(16), 541-551. https://doi.org/DOI:10.5897/AJB2017.16363

Rochette, N. C., Rivera-Colón, A. G., Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737-4754. https://doi.org/doi.org/10.1111/mec.15253

Rogstad, S. H. (1992). Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon*, *41*, 701-708. https://doi.org/10.2307/1222395

Rousset, F. (2002). Inbreeding and relatedness coefficients: what do they measure? *Heredity*, *88*(5), 371-380. https://doi.org/DOI:10.1038/sj/hdy/6800065

Sardos, J., Breton, C., Perrier, X., Van den Houwe, I., Carpentier, S., Paofa, J., Rouard, M., Roux, N. (2022). Hybridization, missing wild ancestors and the domestication of cultivated diploid bananas. *Frontiers in Plant Science*, 3668. https://doi.org/doi.org/10.3389/fpls.2022.969220

Schaal, B., Hayworth, D., Olsen, K. M., Rauscher, J., Smith, W. (1998). Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, *7*(4), 465-474. https://doi.org/doi.org/10.1046/j.1365-294x.1998.00318.x

Schaller, G. E., Bishopp, A., Kieber, J. J. (2015). The yin-yang of hormones: cytokinin and auxin interactions in plant development. *The Plant Cell*, *27*(1), 44-63. https://doi.org/doi.org/10.1105/tpc.114.133595

Schumacher, K., Schmitt, T., Rossberg, M., Schmitz, G., Theres, K. (1999). The Lateral suppressor (Ls) gene of tomato encodes a new member of the VHIID protein family. *Proceedings of the National Academy of Sciences*, *96*(1), 290-295. https://doi.org/doi.org/10.1073/pnas.96.1.290

Shigeta, M. (1990). Folk In-Situ Conservation of Ensete [Ensete ventricosum (Welw.) EE Cheesman]: Toward the Interpretation of Indigenous Agricultural Science of the Ari, Sowthwestern Ethiopia.

Shigeta, M. (1992). *The ethnobotanical study of ensete (Ensete ventricosum) in the southwestern Ethiopia* Kyoto University]. Japan. http://hdl.handle.net/2433/86243

Sim, S., Robbins, M., Van Deynze, A., Michel, A., Francis, D. (2011). Population structure and genetic differentiation associated with breeding history and selection in tomato (Solanum lycopersicum L.). *Heredity*, *106*(6), 927-935. https://doi.org/10.1038/hdy.2010.139

Tsykun, T., Rellstab, C., Dutech, C., Sipos, G., Prospero, S. (2017). Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus Armillaria cepistipes. *Heredity*, *119*(5), 371-380. https://doi.org/10.1038/hdy.2017.48

Weir, B. S., Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358-1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Werner, T. s., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H., Schmülling, T. (2003). Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *The Plant Cell*, *15*(11), 2532-2550. https://doi.org/doi.org/10.1105/tpc.014928

Yemataw, Z., Muzemil, S., Ambachew, D., Tripathi, L., Tesfaye, K., Chala, A., Farbos, A., O'Neill, P., Moore, K., Grant, M. (2018). Genome sequence data from 17 accessions of Ensete ventricosum, a staple food crop for millions in Ethiopia. *Data in brief*, *18*, 285-293. https://doi.org/doi.org/10.1016/j.dib.2018.03.026

Zeven, A. C. (1998). Landraces: A review of definitions and classifications. *Euphytica*, *104*(2), 127-139. https://doi.org/10.1023/a:1018683119237

Zhang, X., Yan, F., Tang, Y., Yuan, Y., Deng, W., Li, Z. (2015). Auxin response gene SlARF3 plays multiple roles in tomato development and is involved in the formation of epidermal cells and trichomes. *Plant and Cell Physiology*, *56*(11), 2110-2124. https://doi.org/doi.org/10.1093/pcp/pcv136

Tables

Table 1. Analysis of Molecular Variance (AMOVA) among and within populations of the three regions in Entada

Source	df	Sum of squares	Variance component	Percentage of variation
Among populations	2	184.15	1.88	0.48
Among individuals within populations	114	1100.41	-383.06	0.00
Within populations	117	8474	777.15	99.52
Total	233	86024.56	395.98	

Table 2. Genetic diversity based on SNP genotyping data between the three regions of Entada

Population	No. of genotypes	Observed (H_o)	Expected (H_e)
South Ari	87	0.99	0.50
North Ari	17	0.99	0.51
Sidama	13	0.99	0.52

Table 3. Average pairwise population differentiation (F_{ST})

Pairwise F_{ST}	South Ari	North Ari
North Ari	-0.02	
Sidama	-0.02	-0.04

Table 4. Candidate genes under directional selection and potentially involved in sucker formation, detected using F_{ST} outlier analysis.

SNP ID	Gene name	Gene function	Reference
E-1971	Lateral suppressor protein	Role in secondary shoot formation	https://www.uniprot.org/uniprot/B5M4A5
E-2117	Auxin response factor 2A	Regulates vegetative growth, lateral root formation and flower organ senescence	https://www.uniprot.org/uniprot/Q2LAJ3
E-1685	Cytokinin dehydrogenase	Play a key role in plant growth and development including maintenance of root and shoot meristems	https://www.uniprot.org/uniprot/A0A1S4ARV5
E-2580	Scarecrow-like protein 18	Transcription factor required for axillary (lateral) shoot meristem formation during vegetative development	https://www.uniprot.org/uniprot/Q9ZWC5

Figures

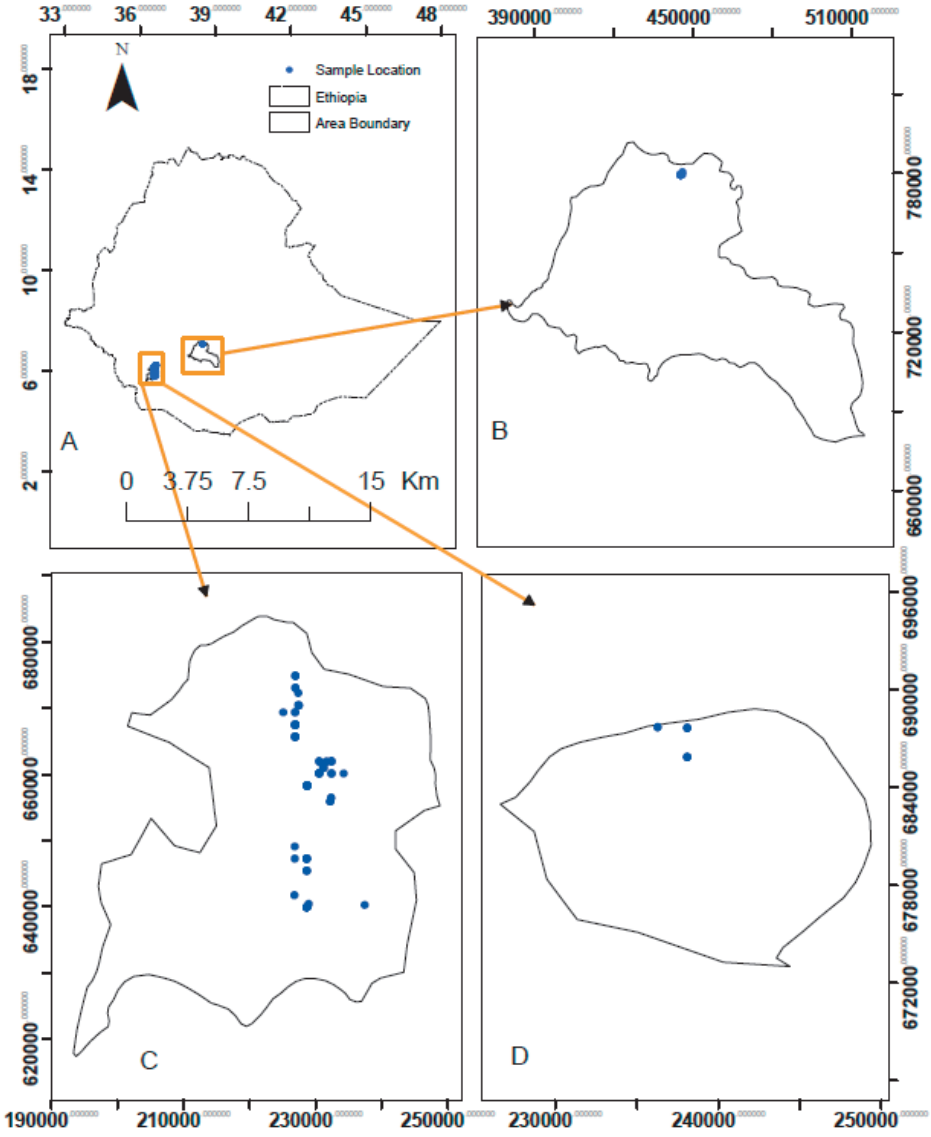


Figure 1. *A: Geographic locations of the collected Entada genotypes. A) An overview of the study areas in Ethiopia; B: Sidama; C: South Ari; D: North Ari*

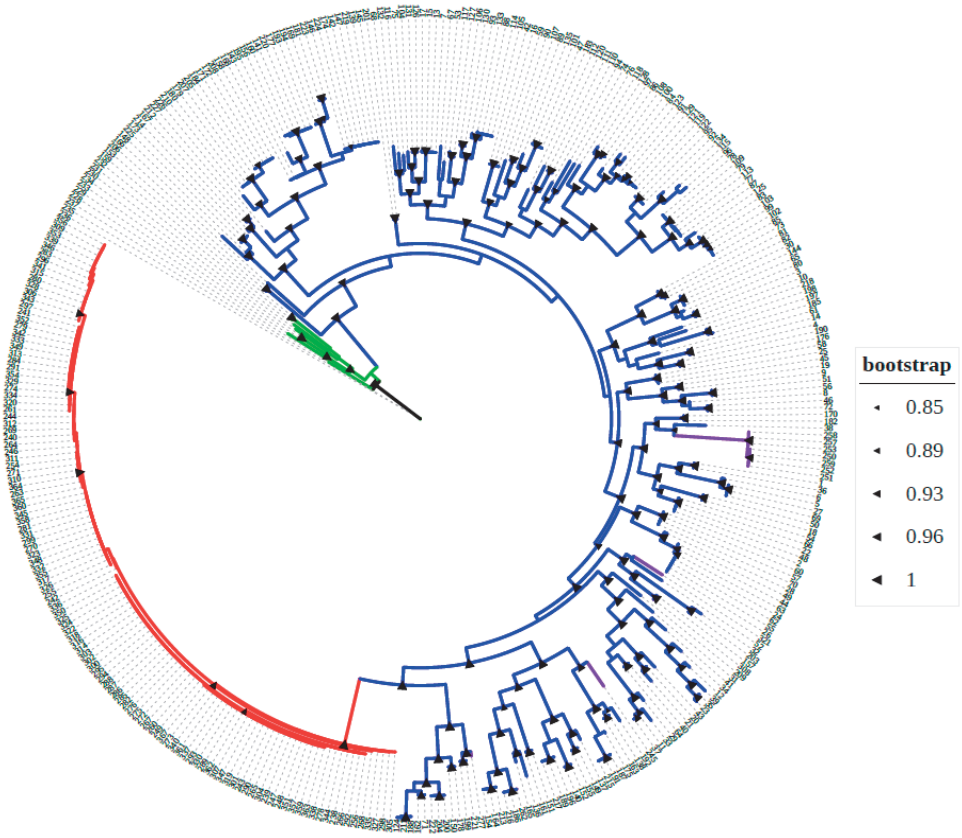


Figure 2. *Maximum likelihood phylogenetic tree with branch length. Colored according to their propagation method and cultivation status, i.e., Cultivated enset: blue; Wild enset: green; Entada landraces: red; Cultivated enset with suckers: purple.*

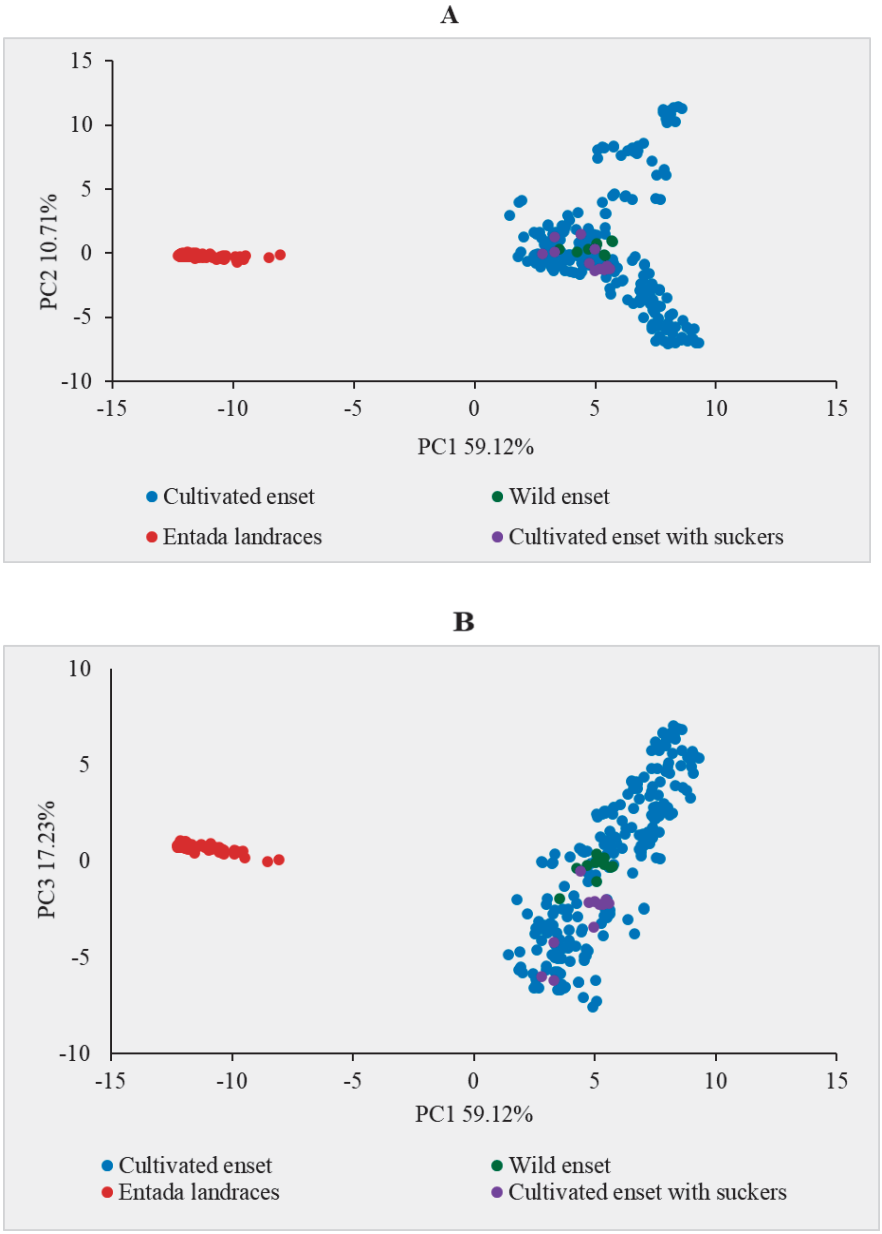
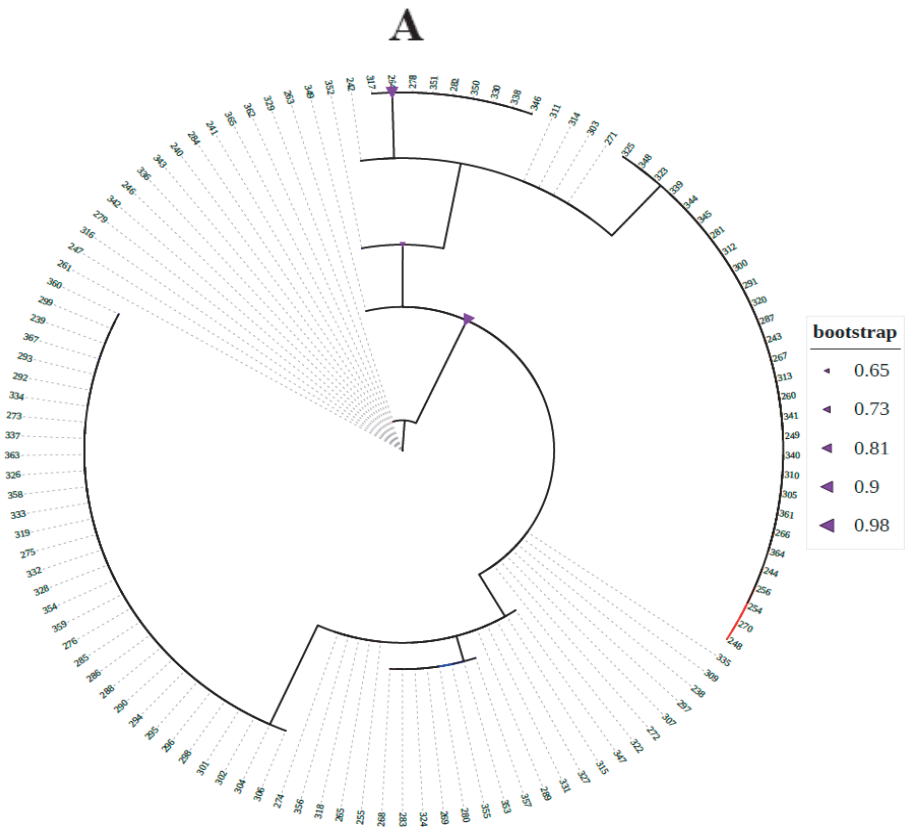


Figure 3. *Principle Component Analysis (PCA) of the complete collection of different enset genotypes. A: PCA plot of PC1 and PC2; B: PCA plot of PC1 and PC3. The percentages along the axes denote the variances explained by the different PCs.*


A
bootstrap
0.65
0.73
0.81
0.9
0.98

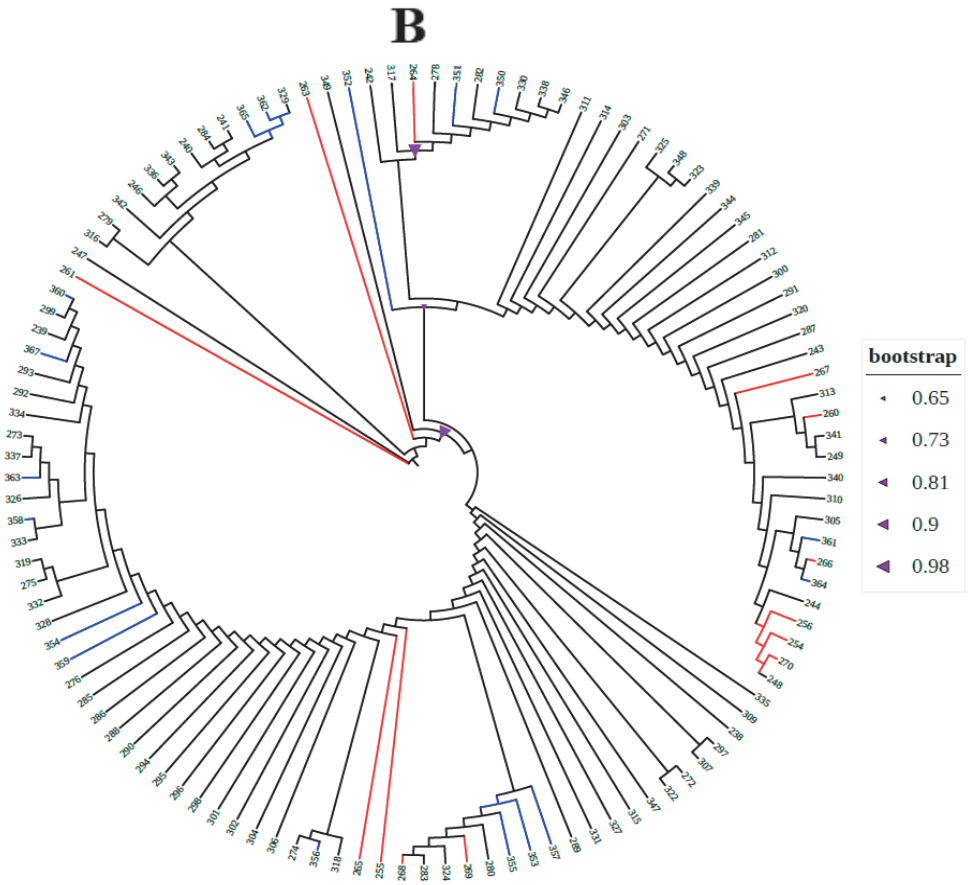
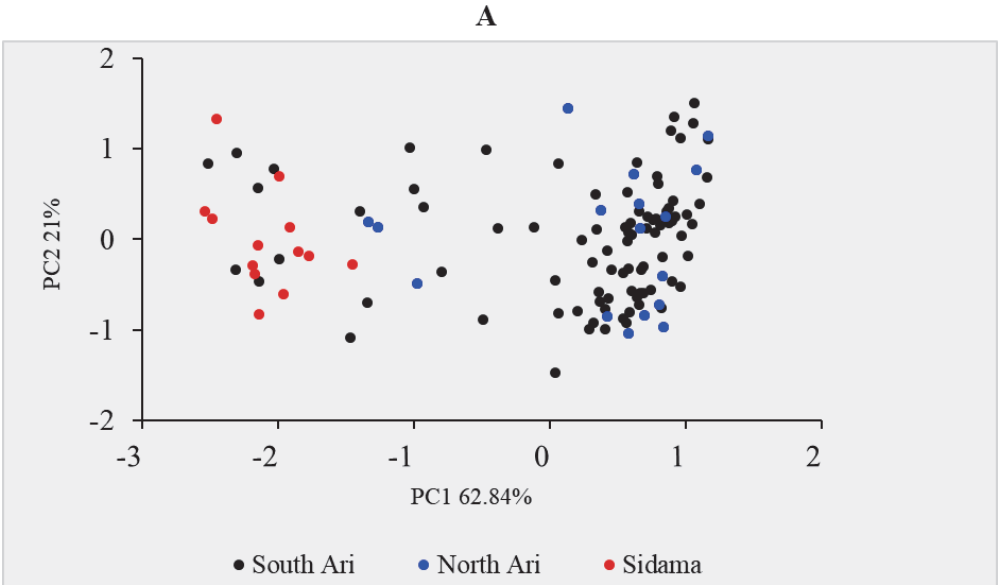


Figure 4. *A: Phylogenic tree maximum-likelihood phylogenetic tree with branch length displayed; B: topological view of the maximum-likelihood phylogenetic tree. The color of the accessions indicates their geographical origin, i.e., South Ari: black; Sidama: red; North Ari: blue.*



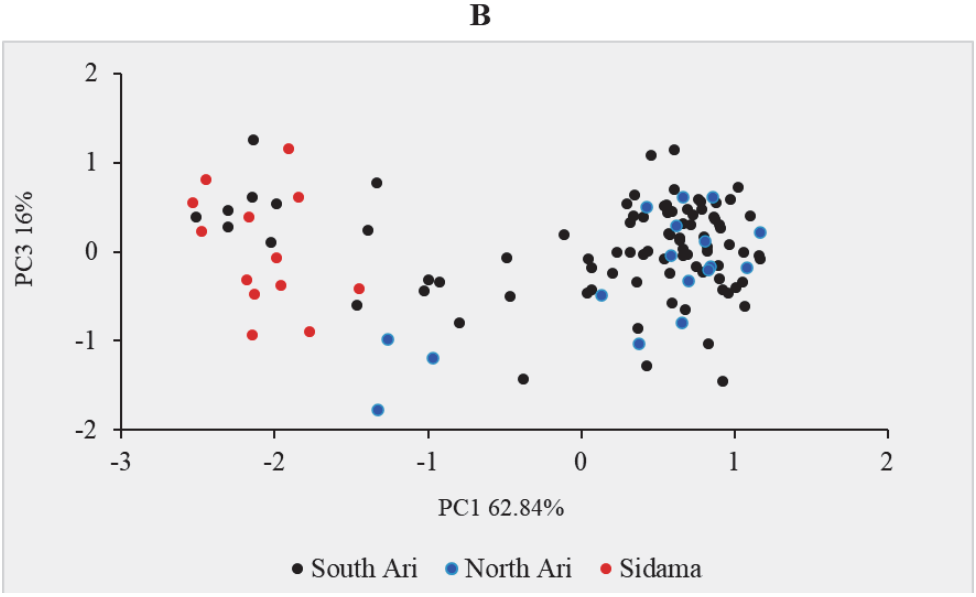


Figure 5. *Principal component analysis (PCA) of the Entada collection. A: PCA plot of PC1 and PC2; B: PCA plot of PC1 and PC3. The percentages along the axes denote the variances explained by the different PCs.*

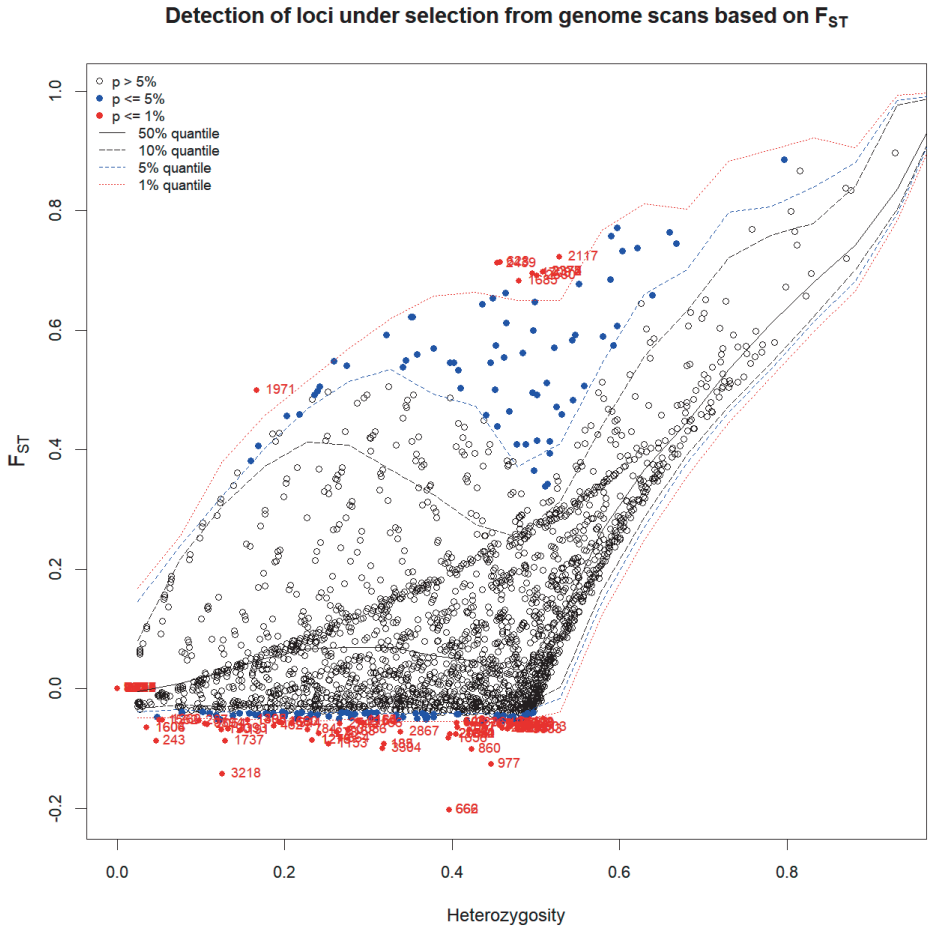


Figure 6. *Candidate loci under selection were identified using F_{ST} based outlier approach (Hierarchical structure model using Arlequin 3.5). F_{ST}: locus–specific genetic divergence among the populations; Heterozygosity: measure of heterozygosity per locus. Loci significant at the 1% level are indicated by red dots.*

Supplementary material

Tables

Supplementary Table 1. Adapter_P1-EcoRI

Name	OligoSequence
GCATG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATG
AACCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCA
CGATC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATC
TCGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAT
TGCAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCAT
CAACC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAACC
GGTTG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTG
AAGGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGA
AGCTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTA
ACACA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACACA
AATTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATTA
ACGGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACGGT
ACTGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGG
ACTTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTC
ATACG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATACG
ATGAG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATGAG
ATTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTAC
CATAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATAT
CGAAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAAT
CGGCT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGCT
CGGTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGGTA
CGTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTAC
CGTCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTCG
CTGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGAT
CTGCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGCG
CTGTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTC
CTTGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGG
GACAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACAC
GAGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGAT
GAGTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTC
GCCGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCGT
GCTGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGA
GGATA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGATA
GGCCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCCA
GGCTC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTC
GTAGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGT
GTCCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCCG
GTCGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCGA
TACCG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCG
TACGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACGT
TAGTA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGTA
TATAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTATAC
TCACG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCACG
TCAGT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCAGT
TCCGG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCCGG
TCTGC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGC
TGGAA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGAA

Supplementary Table 1. Continued …

TTACC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTACC
GCATG_EcoRI_P1.2	/5Phos/AATTCATGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AACCA_EcoRI_P1.2	/5Phos/AATTTGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGATC_EcoRI_P1.2	/5Phos/AATTGATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCGAT_EcoRI_P1.2	/5Phos/AATTATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGCAT_EcoRI_P1.2	/5Phos/AATTATGCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CAACC_EcoRI_P1.2	/5Phos/AATTGGTTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGTTG_EcoRI_P1.2	/5Phos/AATTCAACCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AAGGA_EcoRI_P1.2	/5Phos/AATTTCCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AGCTA_EcoRI_P1.2	/5Phos/AATTTAGCTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACACA_EcoRI_P1.2	/5Phos/AATTTGTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AATTA_EcoRI_P1.2	/5Phos/AATTTAATTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACGGT_EcoRI_P1.2	/5Phos/AATTACCGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTGG_EcoRI_P1.2	/5Phos/AATTCCAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ACTTC_EcoRI_P1.2	/5Phos/AATTGAAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATACG_EcoRI_P1.2	/5Phos/AATTCGTATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATGAG_EcoRI_P1.2	/5Phos/AATTCTCATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATTAC_EcoRI_P1.2	/5Phos/AATTGTAATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CATAT_EcoRI_P1.2	/5Phos/AATTATATGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGAAT_EcoRI_P1.2	/5Phos/AATTATTCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGCT_EcoRI_P1.2	/5Phos/AATTAGCCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGGTA_EcoRI_P1.2	/5Phos/AATTTACCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTAC_EcoRI_P1.2	/5Phos/AATTGTACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGTCG_EcoRI_P1.2	/5Phos/AATTCGACGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGAT_EcoRI_P1.2	/5Phos/AATTATCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGCG_EcoRI_P1.2	/5Phos/AATTCGCAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTGTC_EcoRI_P1.2	/5Phos/AATTGACAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CTTGG_EcoRI_P1.2	/5Phos/AATTCCAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GACAC_EcoRI_P1.2	/5Phos/AATTGTGTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGAT_EcoRI_P1.2	/5Phos/AATTATCTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GAGTC_EcoRI_P1.2	/5Phos/AATTGACTCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCCGT_EcoRI_P1.2	/5Phos/AATTACGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GCTGA_EcoRI_P1.2	/5Phos/AATTTCAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGATA_EcoRI_P1.2	/5Phos/AATTTATCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCCA_EcoRI_P1.2	/5Phos/AATTTGGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GGCTC_EcoRI_P1.2	/5Phos/AATTGAGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTAGT_EcoRI_P1.2	/5Phos/AATTACTACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCCG_EcoRI_P1.2	/5Phos/AATTCGGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
GTCGA_EcoRI_P1.2	/5Phos/AATTTCGACAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACCG_EcoRI_P1.2	/5Phos/AATTCGGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TACGT_EcoRI_P1.2	/5Phos/AATTACGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TAGTA_EcoRI_P1.2	/5Phos/AATTTACTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TATAC_EcoRI_P1.2	/5Phos/AATTGTATAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCACG_EcoRI_P1.2	/5Phos/AATTCGTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCAGT_EcoRI_P1.2	/5Phos/AATTACTGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCCGG_EcoRI_P1.2	/5Phos/AATTCCGGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCTGC_EcoRI_P1.2	/5Phos/AATTGCAGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TGGAA_EcoRI_P1.2	/5Phos/AATTTTCCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TTACC_EcoRI_P1.2	/5Phos/AATTGGTAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

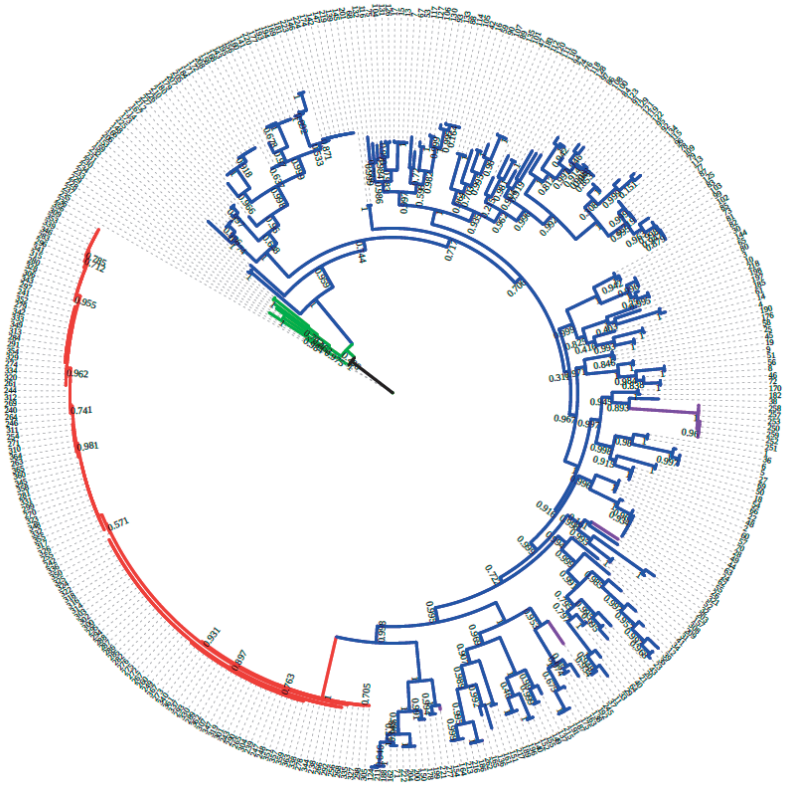
Supplementary Table 2. Adapter_P2-MspI

Name	**OligoSequence**
MspI_P2.1	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
MspI_P2.2	/5Phos/CGAGATCGGAAGAGCGAGAACAA

Supplementary Table 3. PCR-primers

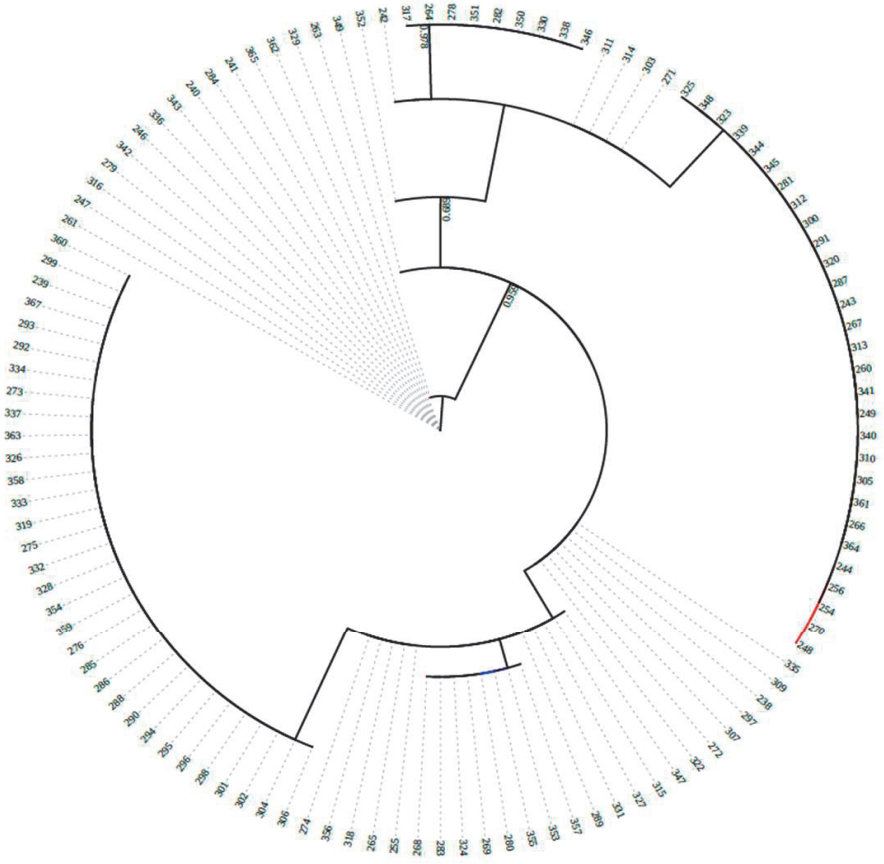
Name	OligoSequence
PCR1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG
PCR2_Idx_6_GCCAAT	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGC
PCR2_Idx_12_CTTGTA	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGC

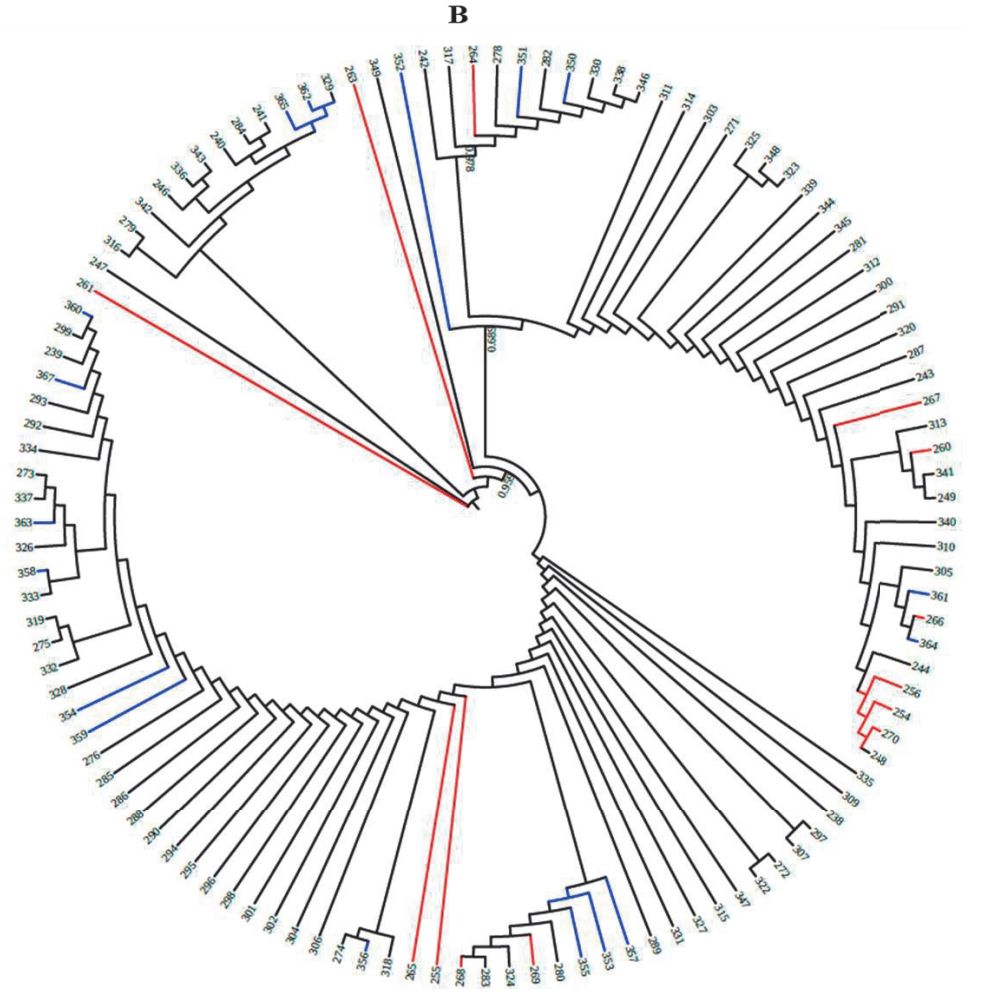
Figures



Supplementary Figure 1. *Maximum likelihood phylogenetic tree with branch length and bootstrap values. Colored according to their propagation method and cultivation status, i.e., Cultivated enset: blue; Wild enset: green; Entada landraces: red; Cultivated enset with suckers: purple.*

A





Supplementary Figure 2. *Phylogenic tree (A) Maximum-likelihood phylogenetic tree with branch length and bootstrap values.; (B) Topological view of the maximum-likelihood phylogenetic tree. Colored according to their geographical origins. Black; South Ari: Blue; North Ari: Red; Sidama.*

Paper IV

Characterization of nutritional composition and bioactive compounds among different forms of enset (*Ensete ventricosum*, (Welw.) Chessman) from Ethiopia

Alye Haile Tefera[1, 2*], Trine Hvoslef-Eide[1], Sylvia Sagen Johnsen[1], Mallikarjuna Rao Kovi[1], Bizuayehu Tesfaye[2], Odd Arne Rognli[1*s]

[1]*Faculty of Bioscience, Department of Plant Sciences, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway*

[2]*School of Plant and Horticultural Science, College of Agriculture, Hawassa University, P.O. Box 05, Awassa, Ethiopia*

*Corresponding Author's E-mail: odd-arne.rognli@nmbu.no

Abstract

Enset (*Ensete ventricosum*) is one of the key endemic starch staple root crops in Ethiopia. It accumulates starch mainly in leaf sheaths and corms. We analysed variation in nutritional composition and bioactive compounds in genotypes of cultivated and wild enset, and Entada. Significant differences were found between genotypes and between the edible enset tissues. The highest amylose content was found in the leaf sheath of the cultivated "Mundraro" genotype and in the corm of wild enset. The average content of amylose in the leaf sheath (45.4%) and corm (26.6%) suggests that starch biosynthesis is more efficient in leaf sheath than in corm tissue. The highest antioxidant capacity (8.4 μmolg^{-1} FW) and total phenolic content (TPC) (100 mg GAE 100 g^{-1} FW) was found in leaf sheaths of the "Kiticho" cultivated enset. However, a wild enset genotype had the highest antioxidant capacity in leaf sheath (9.2 μmolg^{-1} FW) and corm (8.8 μmolg^{-1} FW). Entada landraces had lowest neutral detergent fiber (NDF) and water-soluble carbohydrate (WSC) content compared to cultivated and wild ensets, whereas the wild ensets had highest content of NDF (350.6 g/kg). All ensets had higher contents of dry matter in the corm than in leaf sheaths. Significant variation in mineral content was observed between corms and leaf sheaths except for Mg, P, and Na. Ca and Mg contents were significantly different among cultivated ensets. The content of the minor elements Zn and Cu was highest in the corm, while Fe was highest in the leaf sheath. These results are discussed in relation to the utilization of enset for food and feed. The variation among enset genotypes and edible enset tissues described here are important for selection of accessions for evaluation in multiple environments and for breeding purposes.

Keywords:

Enset (*Ensete ventricosum)*, Corm, Leaf sheath, Starch, Mineral, Total phenolic content, Antioxidant capacity

Introduction

The current trend in food consumption suggests a strong consumer interest in natural and high-quality foods for healthier lifestyles (Grace, 2016). However, malnutrition continues to be a primary public health problem in developing countries, especially in Sub-Saharan countries of Africa (Abebe, 2021). In Ethiopia, undernourishment affects 21.4% of the population and 38.4% of children under the age of five are affected by stunting (von Grebmer *et al*., 2018). Many of the plant species that are cultivated for food are neglected and under-utilized even though they play crucial roles in nutrition and food security, and generate income for rural societies (Magbagbeola *et al*., 2010; Mayes *et al.,* 2012). Several bioactive compounds, such as antioxidants, anti-obesity, immunomodulatory, hypocholesterolemic, antimicrobial, and antidiabetic agents are present in root and tuber crops (Chandrasekara and Josheph Kumar, 2016). Root and tuber crops are the second most important global source of carbohydrates after cereals (Chandrasekara and Josheph Kumar, 2016). They are the primary food source for poor people but they play a minor role in international trade (Oke *et al.,* 1990).

In Ethiopia, root and tuber crops are widely cultivated in the Southern and Southwest parts of the country. Among these crops, enset (*Ensete ventricosum*), a member of the Musaceae family (Borrell *et al*. 2019), is one of the most important starch staple crop (Borrell *et al.,* 2019; Nurfeta *et al.,* 2008; Yemataw *et al.,* 2017). Enset is the only cultivated species in the genus *Ensete* and it is native to Ethiopia (Westphal *et al.,* 1975). Many wild *Ensete* species are found in other countries like in central, eastern and southern parts Africa as well as in Asia (Brandt *et al.,* 1997). Unlike other root and tuber crops such as potato, sweet potato and cassava, enset's nutritional diversity and health benefits have not yet been fully explored (Chandrasekara and Josheph Kumar, 2016). Enset-based farming is a key agricultural system and farmers cultivate various enset landraces under different agroecological conditions (Borrell *et al.,* 2019; Tsegaye, 2002). Enset represents 65% of the total crop production area in the southern regions of Ethiopia and serves as a staple food for 20 to 35% of the population (Borrell *et al.,* 2020; Spring *et al*., 1996; Tuffa, 2019). Enset is also a multipurpose crop used as livestock fodder, for fibre, and traditional medicines in addition to its main role as a food source for local communities (Brandt *et al*., 1997; Mohammed *et al*., 2013; Tamrat *et al.,* 2020). Furthermore, in times of extreme famine, wild enset may be blended with cultivated enset (Tamrat *et al*., 2020) which was done during the great famine in Ethiopia in the years 1888 to 1892 (Tobiaw and Bekele, 2011) and is the reason why enset is called "The Tree Against Hunger" (Brandt *et al.,* 1997). Enset grows best at cooler, higher altitudes between 1200–3100 m a.s.l and can be harvested regularly after 4 to 6 years of transplantation (Borrell *et al.,* 2019; Brandt *et al.,* 1997). In the 2019/20 main crop season survey, 157 million enset plants were harvested, with a total yield of about 9.4 million tonnes per year, making enset one of the largest perennial food crops in Ethiopia (Borrell *et al.,* 2019).

Enset is mainly cultivated for its starch content in the pseudostem (overlapping leaf sheaths), the leaf sheath and the corm (the underground stem), which provide a year-round dietary starch source (Borrell *et al.,* 2019). However, the edible parts of enset vary from place to place, but in general, the pseudostem, young shoots, and the corm are the main edible parts of enset (Atlabachew and Chandravanshi, 2008; Daba and Shigeta, 2016). The main processed food from enset is “Kocho”, a fermented starch obtained from the

mixture of the decorticated leaf sheaths and grated corms; “Bulla”, a white powder produced by drying squeezed sap from scraped leaf sheaths and grated corms; and “Amicho”, boiled enset corm, usually from younger plants (Brandt *et al.,* 1997; Yemataw *et al*., 2014).

Enset is an underexploited starch crop, high in carbohydrates, but low in vitamins, proteins, and essential amino acids (Besrat *et al*., 1979; Tiruha Karssa and Alessio Papini, 2018; Tamrat *et al*., 2020). The major and minor mineral composition of enset products is comparable to sweet potato, taro, and yam (Tsegaye and Struik, 2001). Food from enset contains more calcium and iron than most cereals, tubers, and root crops, and some enset landraces are believed to have medicinal value (Daba and Shigeta, 2016). However, large variations in quality of enset was observed among different genotypes and landraces of enset (Borrell *et al.,* 2019; Tobiaw and Bekele, 2011), due to climatic conditions, genetics, soil conditions (Borrell *et al.,* 2019), and plant age (Tiruha Karssa and Alessio Papini, 2018).

Previous studies on the mineral, protein and carbohydrate content of enset is mostly from processed edible parts of the plant and commercially available food products from local markets (Atlabachew and Chandravanshi, 2008; Bosha *et al*., 2016; Daba and Shigeta, 2016; Nurfeta *et al*., 2008; Tamrat *et al.,* 2020). However, there is lack of studies on the complete nutritional profile including different starch, bioactive compounds, and dietary fibres for important genotypes and landraces. Moreover, it is important to characterize the nutritional diversity in diverse types of enset for selecting future cultivars which are high in essential nutrients and bioactive compounds. Therefore, we selected a set of cultivated and wild enset, and Entada genotypes with different molecular genetic profiles based on our previous genetic diversity study (Haile *et al*. unpublished) for the investigation of nutritional composition (**Figure S1**). We hypothesise that the genotypic variation among enset and Entada genotypes affects the nutritional composition. Therefore, the main objective of this study is to investigate variations in health-promoting major and minor mineral elements, bioactive compounds, starch, and fibre content among diverse types of enset.

Materials and Methods

Sample Collection and Preparation

Fourteen cultivated enset genotypes, 6 originating from Sidama and 8 from Gurage, were collected from the enset maintenance field at the Areka Agricultural Research Centre (AARC) and two wild enset and two Entada landraces, originating from Sidama, were collected at the Hawassa University Research Centre **(Table 1)** based on our previous genetic diversity studies. Five years old enset plants were harvested and the edible parts thoroughly cleaned. The fresh leaf sheaths (LS) and corms (Co) were cut into small pieces, and 2 kg of leaf sheath (LS) and corm (Co) were transferred into individual zipped plastic bags and frozen at -20 ˚C **(Figure 1)**. Leaf sheaths and corms were further sliced, freeze-dried, ground, sieved, and stored at –20 ˚C for later analyses. All analyses were based on dry weight (DW), while the total phenol content and antioxidant capacity were measured based on fresh weight (FW).

Quantification of Amylose Content

The Megazyme K-Amyl amylose/amylopectin analysis kit was used to quantify the amylose (AM) content following the manufacturer's protocol (Megazyme, Wicklow, Ireland). 20-25 mg of each sample was dissolved and rehydrated in dimethyl sulfoxide (DMSO) by boiling. A maize reference sample of 68% amylose was included with each batch. The starch was precipitated using ethanol alcohol (EtOH) to remove lipids and then re-dissolved in DMSO by boiling. After diluting the samples in concanavalin A (ConA) solvent, they were filtered through Whatman filter paper. Amylopectin was precipitated using ConA, while the amylose was enzymatically degraded to obtain the glucose. An aliquot sample that was not treated with ConA was also degraded to glucose to measure the total starch content. Two aliquots of amylopectin and two aliquots of total starch were treated with glucose determinant reagent (glucose oxidase/peroxidase; GOPOD (D-Glucose assay kit)), and their absorbances were read using a spectrophotometer at 510 nm wavelength. The amylose content (%) was calculated as (Abs510-Amyl/Abs510-total)*66.8. The conversion factor (66.8) originates from the dilutions during the protocol steps.

Quantification of Water-Soluble Carbohydrates (WSC)

Samples were extracted in 0.05 M Na-acetate buffer at room temperature for 18 h and filtered through a filter paper. Sucrose and fructans in the purified extract were hydrolysed using 0.074 M H_2SO_4 at 90 ˚C for 70 minutes. Monosaccharides were further converted to glucose-6-phosphate and fructose-6-phospate by an enzymatic method using the kit K-FRUGL (D-Fructose/D-Glucose Assay Kit D) (Megazyme, Wicklow, Ireland). The absorbance for NADPH at 340 nm before and after the reaction is measured spectrophotometrically. The increase in absorbance is directly proportional to the glucose and fructose concentration. The analysis gives the total sum of monosaccharides, sucrose and fructans as a result. Detailed description of the method used can be found in Randby *et al.*, (2010).

Determination of the Carbon (C) and Nitrogen (N)

Simultaneous CHNS (Carbon, Hydrogen, Sulfur and Nitrogen analyzer) analysis requires high-temperature combustion in an oxygen-rich environment and is based on the classical Pregl-Dumas method. The samples were combusted with oxygen in a combustion chamber at a temperature of 1150 ˚C. Then the combustion products were passed through a hot reduction tube (850 ˚C) having a helium gas. The CHNS analyzer detect CO_2, H_2O and N_2 with a heating wire detector (TCD - Thermal Conductivity Detector). The reference method used to measure CHNS contents was the combustion or Dumas's method (AOAC 992–23 or ISO 16634–1: 2008) ISO 16634: 2008; Cereals, legumes, ground cereals, oilseeds, oilseeds, and animal feed.

Quantification of Neutral Detergent Fibres (NDF)

The samples were heated in a neutral soap solution (neutral detergent) to dissolve the cells, while the cell wall remained undissolved. The undissolved fraction (NDF) mainly contains hemicellulose and cellulose. In contrast, the neutral detergent soluble (NDS) fraction consists of lipids, sugars, organic acids, water-soluble compounds, pectin, starch, non-protein nitrogen and water-soluble proteins fractions. The amount of NDF is determined gravimetrically as aNDF on organic matter basis (aNDFom) after incinerating the samples at 550 ˚C to remove inorganic compounds (Mertens *et al.* 2002). The analysis was performed using

an Ankom200 Fiber Analyzer (Ankom Technology) at the Laboratory for Organic Analytical Chemistry (LabTek) at NMBU.

Quantification of Major and Minor Minerals

The samples were decomposed with ultrapure concentrated HNO_3 (nitric acid) prepared by dissolving 0.25 g of sample with 5 mL of HNO_3 in acid-washed Teflon tubes at 260 °C in a Milestone Ultraclave (260 °C for 20 minutes). The samples were then diluted by adding 50 mL deionized water and analyzed using ICP-OES (Inductively coupled plasma atomic emission spectroscopy) (Agilent 5110 ICP-OES) and ICP-MS (Inductively coupled plasma mass spectrometry) (Agilent 8800 ICP-MS). Reference material and blank samples were decomposed at the same time. LOD (limit of detection) and LOQ (limit of quantification) are calculated from 3*SD (standard deviation) and 10*SD on the blank samples (n = 3). Phosphorus (P), potassium (K), sulfur (S), calcium (Ca), and magnesium (Mg) were quantified using the ICP-OES. Sodium (Na), iron (Fe), copper (Cu) and zinc (Zn) were quantified using the ICP-MS. Dry matter content was determined gravimetrically based on sample weight loss after being heated in an oven at 120 °C for 48 hours.

Bioactive Compounds in Enset Leaf Sheath and Corm

The total phenolic content (TPC) of enset leaf sheaths and corms were determined using the Folin-Ciocalteau method (Singleton *et al.,* 1999). Three grams (g) of fresh homogenate sample was weighed in a centrifuge tube and then extracted using 30 mL methanol. After extraction, the liquid sample was centrifuged, and the supernatant was mixed with Folin Ciocalteu reagent and 7.5% (w/v) sodium carbonate. Finally, the prepared sample was incubated for 15 minutes and measured at 765 nm wavelength. The absorption was equivalent to the sum of the individual contributions of the different classes of phenols present in the sample. Total phenolic content was assessed against a calibration curve of gallic acid, and the results were presented in milligram (mg) gallic acid equivalents (GAE) per 100 g FW. The antioxidant capacity (AO) analysis was carried out by using a FRAP assay on a Konelab 30i (Thermo Electron Corp. Vantaa, Finland) (Benzie and Strain, 1999). Three grams of fresh homogenate sample was weighed in a centrifuge tube then extracted in 30 mL methanol, centrifuged and the supernatant was mixed with acetate buffer, TPTZ (2,4,6-tri-pyridyl-s-triazine) and iron trichloride, incubated for 10 minutes and the absorbance was measured at 595 nm. Standards were prepared using Trolox to which readings were compared, following the Konelab 30i outline and method (Volden *et al*., 2008; Zargar *et al.,* 2011).

Data Analysis

Analysis of variance was performed using the R software (version 3.6.2) and Proc GLM in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) to test differences in contents of various nutritional composition and bioactive compounds among and within different forms of enset. Multiple means comparisons were tested using the Tukey test, and differences between considered statistically significant at $P \leq 0.05$. The correlation coefficients were estimated using Pearson correlation.

Results

Variation in Amylose Content, NDF and WSC

Amylose content in the leaf sheath varied from 22.0% (Astara) to 64.1% (Mundraro), and from 21.0% (Agade) to 62.6% (Wild2 enset) in the corm (**Table 2**). The difference between the average amylose content in leaf sheath (45.36%) and corms (26.60%) was highly significant (P≤0.001, **Table 3**). Wild enset had significantly more amylose in the corm than both cultivated enset and Entada (52.28, 23.39, 24,45, respectively; P≤0.001, **Table 3**). The amylose content in the leaf sheath was not significantly different among the different enset types, nor between the two geographical regions for any of the two tissues.

The WSC content in the leaf sheath varied from 3.7% (Entada2) to 25.5% (Kiticho), and from 4.3% (Ginbowe) to 17.2% (Astara) in the corm (**Table 2**). The difference between the average WSC content in leaf sheaths (9.79%) and corms (7.83%) was significant (P≤0.05, **Table 3**). Cultivated enset had significantly higher WSC content in leaf sheaths than both Entada and wild enset (P≤0.001, **Table 3**), while there were no significant differences between WSC contents of the corms, nor between the two geographical regions for any tissues.

The NDF content of the leaf sheath varied from 66g/kg (Entada2) to 350.6 g/kg (Wild2) (**Table 2**). On average, wild enset had significantly higher NDF content (324.42 g/kg) than both cultivated enset (123.93 g/kg) and Entada (66.62 /kg) (P≤0.001, **Table 3**). Also, ensets from Sidama had significantly higher NDF content than ensets from Gurage (141.84 vs. 110.50 g/kg; P≤0.01, **Table 3**)**.** A heat map of the chemical composition is presented in **Figure 3**, which gives a quick overview of the main differences between the contents in the two tissues.

Variation in Major and Minor Minerals

The concentrations of N and C are presented in **Table 2**, and the major minerals Ca, K, Mg, P, S, and Na, and the minor elements Zn, Cu, Fe are presented in **Table 6**. Heat maps of these elements are presented in **Figure 3** and **Figure 4.** The nitrogen content in leaf sheaths varied from 0.3% (Wild1) to 1.6% (Derea) and from 0.3% (Wild1 and Entada2) to 3.0% (Derea) in the corms (**Table 2**). On average, leaf sheaths contained significantly more N than corms (P≤0.001, **Table 3**), and cultivated enset contained significantly more N than both Entada and wild enset, especially in corms (P≤0.001), but also in leaf sheaths (P≤0.5).

The carbon content in leaf sheaths varied from 37.8% (Wild1) to 42.5% (Mundraro), and from 33.8% (Wild2) to 43.5% (Mundraro) in the corms (**Table 2**). On average, corms contained significantly more C than leaf sheaths (P≤0.05, **Table 3**), and cultivated enset and Entada contained significantly more C than wild enset, both in leaf sheaths and corms (P≤0.001). Both leaf sheaths and corms of ensets from Gurage contained significantly more C than ensets from Sidama (P≤0.01 and P≤0.001, respectively, **Table 3**).

The Ca content in leaf sheaths varied from 1.10 (Badedet) to 13.00 g/kg (Wild2), and from 1.10 (Gena and Astara) to 6.10 g/kg (Wild2) in the corms (**Table 6**). On average, leaf sheaths contained significantly more Ca (3.2 g/kg) than corms (1.3 g/kg) (P≤0.01, **Table 7**). The variation between genotypes was significant (P≤0.01), caused by the major difference between wild enset with average Ca contents of 12.50 and 4.80 g/kg in leaf sheaths and corms, respectively, compared to less than 2 g/kg Ca in leaf sheaths of cultivated

and Entada and less than 1 g/kg in the corms. There were no significant differences between genotypes from the two regions as regards Ca content in leaf sheaths and corms.

The potassium (K) content in leaf sheaths varied from 18.0 (Badedet) to 62.0 g/kg (Wild2), and from 13.0 (Lemat) to 39.0 g/kg (Wild2) in the corms (**Table 6**). On average, leaf sheaths contained significantly more K (31.6 g/kg) than corms (20.9 g/kg) (P≤0.05, **Table 7**), and differences between genotypes and regions were non-significant. The magnesium (Mg) content in leaf sheaths varied from 0.38 (Nichew) to 2.00 g/kg (Wild2), and from 0.32 (Gena) to 3.50 g/kg (Wild2) in the corms (**Table 6**). On average, differences between genotypes were significant (P≤0.01, **Table 7**), and the Mg content of corms was significantly highest among the Sidama genotypes (P≤0.05, **Table 7**), while there was no significant average difference between leaf sheaths and corms.

There were no significant variations for phosphorus (P) nor sodium (Na) contents, although there was a tendency that the content of P was higher in leaf sheaths than in corms on average (P=0.07, **Table 7**). For sulphur (S), iron (Fe), copper (Cu) and zinc (Zn), the differences between tissues were highly significant (P≤0.001, **Table 7**) for all elements. On average the content of S, Cu, and Zn were higher in the corms (1.5 g/kg, 4.4 mg/kg, and 145.5 mg/kg, respectively) than in the leaf sheaths (0.4 g/kg, 1.7 mg/kg, and 9.2 mg/kg, respectively). On the other hand, the Fe content was higher in leaf sheaths (26.9 mg/kg) than in corms (12.1 mg/kg). Differences between genotypes and regions were not significant for S, Fe, Cu and Zn.

Total Phenolic Content

The total phenolic content (TPC) in the leaf sheaths varied from 32.1 mg (Derea) to 108.0 mg GAE/100g FW (Kiticho), and from 39.4. (Astara) to 104.3 mg GAE/100g FW (Agade) in the corms (**Table 4**). The difference between the average TPC contents in leaf sheaths and corms was not significant (**Table 5**). Entada had significantly lower TPC content (44.33 mg GAE/100g FW) in the leaf sheaths than cultivated and wild enset (66.62 and 71.48 mg GAE/100g FW, respectively) (P≤0.05, **Table 5**), and there was a significant difference between the average TPC content in corms of cultivated and Entada (63.88 mg GAE/100g FW) compared with wild enset (80.18 mg GAE/100g FW) (P≤0.05, **Table 5**). There were no significant differences between the two geographical regions for any tissues.

Antioxidant Capacity Content

The antioxidant capacity (AO) in the leaf sheaths varied from 2.4 (Agade) to 9.2 $\mu molg^{-1}$ FW (Wild1), and from 2.4 (Lemate) to 8.8 $\mu molg^{-1}$ FW (Wild2) in the corms (**Table 4**). As with TPC, the difference between the average AO contents in leaf sheaths and corms was not significant (**Table 5**). Wild enset had significantly highest AO content (6.54 $\mu molg^{-1}$ FW) in the leaf sheaths compared with cultivated (5.11) and Entada (3.41 $\mu molg^{-1}$ FW) (P≤0.05, **Table 5**), and there was a highly significant difference between the average AO content in corms of cultivated and Entada (4.68 mg $\mu molg^{-1}$ FW) compared with wild enset (6.81 mg $\mu molg^{-1}$ FW) (P≤0.001, **Table 5**). There were no significant differences between the two geographical regions for any tissues.

Among the cultivated ensets, the highest antioxidant capacity was found in the leaf sheath of Kiticho and in the corm of Agade, while Agade leaf sheath and Lemate corm had the lowest antioxidant capacity.

Dry Matter Content

For all genotypes, except Medasho, Ado, Badedet, and Wild2, the dry matter (DM) content in the corm was much higher than in leaf sheaths (**Figure 2**). The highest DM content in corm was observed in Entada2 (32.7%) and Nichewa (32.3%), and the lowest was observed in wild ensets and Badedet. The highest DM content in leaf sheaths was observed in Badedet (27.3%) followed by Ado (22.7%), while Wild1 (5.6%) and Wild2 (3.3%) had the lowest DM contents (**Figure 2)**.

Pearson Correlation Coefficients

Pearson correlation coefficients among amylose, NDF, WSC, phenolic and antioxidant contents in leaf sheaths and corms among the fourteen cultivated ensets are presented in **Figure 5.** As expected, total phenolic content (TPC) and antioxidant capacity (AO) is highly positively correlated, both in leaf sheaths (r=0,98***) and in corms (r=0.97***) (**Figure 5**). The correlation between the amylose content in leaf sheath (AM_LF) and corm (AM_Corm) was moderate positive (r=0.41), while correlations between AM_LF and NDF content in leaf sheaths (NDF_LF) and WSC in leaf sheaths (WSC_LF) was relatively large a positive (r=0.65** and 0.49, respectively). NDF content in leaf sheaths (NDF_LF) was highly positively correlated with WSC_LF (r=0.78**), while the latter was moderately positively correlated with total phenolic content (TPC_LF) and antioxidant capacity (AO_LF) in leaf sheaths (r=0.50 and 0.49, respectively). All other correlations were weak and there were no strong negative correlations between these traits. Correlations around 0.5 would probably be significant with a larger dataset.

Discussion

Variation of Amylose, NDF and WSC in Different Genotypes

On average, higher amylose content was observed in leaf sheaths than in corm, suggesting that the starch biosynthesis was more efficient in leaf sheaths than in the corm **(Table 2, Figure 3)**. Although not statistically significant, genotypes from Sidama had higher amylose content than Gurage genotypes in both tissues, most pronounced in corms (P=0.07) **(Table 3)**. The Mundraro genotype from Sidama is especially interesting since it has highest amylose content of all cultivated ensets both in leaf sheath and corm (**Table 2**). The differences in amylose content are most certainly due to genotypic variation, at least for variation between the cultivated ensets which had been grown in the same field for many years and were sample at the same age. The differences are also due to variation in the main edible tissues (Gebre-Mariam *et al*., 1996; Moorthy, 2002; Seung, 2020). The amylose content of some of the enset genotypes is very high compared to other root and tuber crops like potato and cassava, which has amylose contents around 20%. Except for two enset genotypes, the content of amylose in the leaf sheath (36.2-64.1%) is higher than in banana (24.4-40.7%) (Fontes *et al*., 2017; Ravi and Mustaffa, 2013; Waliszewski *et al*., 2003). The average amylose content in corms (26.6%) is higher than the 21% previously reported by (Hirose *et al.*, 2010) but lower than 29% reported by (Gebre-Mariam *et al*., 1996). It is interesting that the amylose content we report for enset is higher than found in different potato cultivars (Liang *et al.,* 2019). Genotypes with the highest amylose content can be used as sources for maltodextrin and glucose syrup production. Hirose *et al.* (2010)

characterized the properties of enset starch and found that enset starch had excellent properties compared with potato starch, corn starch, and sago.

Water soluble carbohydrates (WSC) consist of monosaccharides (glucose and fructose), sucrose and fructans. Starch is a polymer of glucose units, either in the linear form amylose or the branched form amylopectin. Thus, there should be a relationship between carbohydrate and starch/amylose contents. The correlation was 0.49 in the leaf sheaths but negative (-0.18) in the corm. A moderate positive correlation in the leaf sheaths is as expected since glucose make up only a part WSC. The negative correlation in the corm is difficult to explain. As regards water soluble carbohydrates (WSC), the leaf sheaths contain significantly more carbohydrates than corms, and cultivated enset has higher content in this tissue than Entada and wild enset. The Kiticho genotype stands out since it has the highest WSC content in leaf sheaths and is also among the highest in WSC of the corm **(Table 2)**.

The Entada landraces had lowest NDF content in the leaf sheaths, wild enset highest and of the cultivated ensets, Mundraro and Kiticho had significantly higher NDF content than the other genotypes (**Table 2**). However, the NDF content in this study was lower than the values reported by Nurfeta *et al*., (2009); they found NDF content of 562 g/kg DM for leaves compared to the average and maximum NDF content of 139.8 and 350.6 g/kg DM, respectively, in the present study. These values are difficult to compare since there is no information about the genotype used by Nurfeta *et al.,* (2009) and factors like age, maturity and preparation of samples are different. NDF is important for the feed value of enset leaves, and low NDF content means that the leaves have high energy concentration and provide a good forage for ruminants. Thus, the Entada genotypes should be very promising as feed sources providing that the yield is satisfactory. The NDF content of wild enset in the present study is comparable with that of banana pseudostem (Carmo *et al*., 2018).

The large positive correlations between NDF content on the one hand, and WSC and amylose in the leaf sheaths (r=0.78 and 0.65, respectively, **Figure 5**) is somewhat surprising. NDF is composed of hemicellulose and lignin, and it is expected that there would be a negative correlation between WSC and NDF as found in a study of the inheritance of WSC in the forage grass cocksfoot (*Dactylis glomerata*) by Sanada *et al.* (2007). Leaves of grasses do not accumulate starch, this is a big difference compared with leaf sheaths of enset. Thus, it is likely that the explanation for the positive correlation is that high WSC content creates a good source for starch deposition. WSC is also providing the building blocks (glucose and other sugars) of cellulose and hemicellulose, which are main components of NDF. The positive correlation can only be explained by the composition of NDF, i.e., it must be made up of a low proportion of lignin and mostly hemicellulose. Also, the NDF content among the cultivated enset genotypes was low compared to other studies, which could affect these relationships. This show that the lower NDF content in leaf sheath might be related to the high amylose content in leaf sheaths of cultivated and wild enset as found by Tuffa (2019).

Compositional Variation in Minerals

Significant variation of mineral content was observed between corm and leaf sheath of enset except for contents of Mg, P and Na. Mineral contents did not vary according to the geographical region, except for Mg in corms which was significantly highest among Sidama genotypes (**Table 7**). Significant variation in Ca and Mg contents is interesting since some enset genotypes are believed to have medicinal value and are used for that purpose by enset farmers. This may be because enset contains large amounts of Ca and P (Daba and Shigeta 2016). The wild enset had much larger content of Ca than the cultivated ensets in both tissues, which could be a genotypic effect but also an environmental effect since the wild ensets were not grown in the same experimental field as the cultivated ensets. However, another study (only on cultivated enset) showed lower Ca contents in the leaf sheath and corm than the values reported in this study (Nurfeta *et al*., 2008).

The content of the minor minerals S, Fe, Zn, and Cu was significantly higher in corms than in leaf sheaths, except for Fe (**Table 7, Figure 4**). In this context, we found that enset is a good source of zinc and copper, and that enset corm contains higher levels of zinc and copper than leaf sheath. The highest Fe content was found in leaf sheaths of wild enset and the lowest in the corms of the Entada landraces. All accessions had higher Cu content in the corm than the leaf sheath. These results are in agreement with previous reports (Debebe *et al.,* 2012).

Variation in Bioactive Compounds - Total Phenolic Content (TPC) and Antioxidant Capacity (AO)

Consumption of food rich in phenolic content has been shown to reduce the risk of heart disease by acting as antioxidants (Kaur and Kapoor, 2002). The antioxidant activity of phenolics is mainly because of their redox properties (Rice-Evans *et al*., 1997). As expected, very strong positive correlations were observed between TPC and AO contents both in leaf sheaths and corms. Antioxidant capacity is the most important in this context, therefore we focus most of the discussion on this component. The average TPC and AO contents in leaf sheaths and corms were not significantly different, but within tissues, there were significant differences both within and between different types of enset (Table 4 and 5). Of the cultivated enset, Kiticho had highest TPC and AO contents in leaf sheaths, while Agade had highest contents in corms, but lowest contents in the leaf sheath. This demonstrates very different characteristics of these two commonly grown genotypes. However, genotype Wild1 had significantly highest antioxidant capacity in leaf sheaths (9.2 $\mu molg^{-1}$ FW) and Wild2 in corms (8.8 $\mu molg^{-1}$ FW) of all enset genotypes (**Table 4, Figure 3**). This demonstrates that there are valuable untapped genetic resources among wild ensets that can be utilized for improving food quality of enset products. In general, the content of bioactive compounds in tissues of Entada genotypes is average, and Entada does not seem to be promising sources of antioxidants.

Our study indicates that, except for two of the cultivated ensets, the others have higher TPC contents than reported for enset by Forsido *et al.,* (2013). The quantity of TPC content will vary with genotype and tissue part (Chung *et al.,* 2008; Patthamakanokporn *et al.,* 2008). Like banana, the enset plant has different content of total phenol in different edible parts (Kandasamy and Aradhya, 2014). All edible enset tissue parts have high amounts of total phenolic content and thus enset can serve as an equally good source of dietary polyphenols as other Ethiopian staple carbohydrate foods, as also shown by Forsido *et al*., (2013). Further,

the total phenolic content of most of the enset accessions are very high compared with other tuber vegetables reported in the literature (Cornago *et al.*, 2011; Shan *et al*., 2005). However, it is difficult to compare our results with the results of other studies because of the differences in the range of genotypes and tissues studied, the sampling, storage and preparation methods, and the methods used to estimate total phenolic content and antioxidant capacity (Chung *et al.,* 2008; Patthamakanokporn *et al.,* 2008).

Variation in Dry Matter Content

The enset genotypes differed considerably in dry matter content, for most of them, DM content was much higher in corms than in leaf sheaths. This has also been found in other studies (Debebe *et al.,* 2012; Negash, 2002; Nurfeta *et al*., 2008). There are, however, some genotypes, i.e., Medasho, Ado, Badedet, and one of the wild ensets (Wild2), that has very similar DM content in the two tissues. Wild2 has very low DM content, and this type would an interesting genotype for feeding ruminants during the dry season in areas where the water supply is scarce, as pointed out by Nurfeta *et al*. (2008).

Conclusion

This study demonstrates that there is significant variation in nutritional content and bioactive compounds among well-known cultivated enset genotypes, Entada and wild enset, and between these types of enset. Leaf sheaths and corms also differ in contents of most of the components. Since the cultivated enset plants were grown under similar condition and harvested at same age of maturity, the variation between genotypes and tissues is, to large extend, due to genetic differences among genotypes. Samples of the Entada and wild genotypes were taken from clones grown at another location, and therefore differences between cultivated enset, Entada and wild enset would be influenced by environmental factors. The amylose content was highest in leaf sheaths of the Mundraro genotype and corms of the wild enset genotype Wild2. Most enset accessions are good sources of antioxidants, the best sources are from leaf sheaths of wild enset, followed by the cultivated ensets Kiticho, Mundiraro and Separa, and from corms of Wild2 enset, Agade and Mundraro. The variation among enset genotypes and edible enset tissues described are important for selection of accessions for evaluation in multiple environments and for breeding purposes.

Acknowledgement

The authors thank the Norwegian Agency for Development Cooperation (NORAD) for funding this study, which is part of the project 'Controlling Disease in Sweet potato and Enset in South Sudan and Ethiopia to Improve Productivity and Livelihoods under Changing Climatic Conditions using Modern Technologies' under the NORHED program (Agreement no. ETH-13/0017). The authors would like to thank Kari Grønnerød for conducting the analyses of total phenol content and antioxidant capacity.

Conflict of interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Abebe, A. (2021). Statistical Analysis of Correlates of Malnutrition in Ethiopia: Children under Age Five. *J Food Microbiol Saf Hyg*, *6*, 147.

Atlabachew, M., Chandravanshi, B. S. (2008). Levels of major, minor and trace elements in commercially available enset (Ensete ventricosum (Welw.), Cheesman) food products (Kocho and Bulla) in Ethiopia. *Journal of Food Composition and Analysis*, *21*(7), 545-552. https://doi.org/10.1016/j.jfca.2008.05.001

Benzie, I. F., Strain, J. (1999). [2] Ferric reducing/antioxidant power assay: direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. *Methods in Enzymology*, *299*, 15-27. https://doi.org/10.1016/S0076-6879(99)99005-5

Besrat, A., Mehansho, H., Bezuneh, T. (1979). Effect of varietal difference and fermentation on protein quantity and quality of ensete.[kocho, Ethiopia]. *Nutrition Reports International (USA)*.

Borrell, J. S., Biswas, M. K., Goodwin, M., Blomme, G., Schwarzacher, T., Heslop-Harrison, J. S. P., Wendawek, A. M., Berhanu, A., Kallow, S., Janssens, S., Molla, E. L., Davis, A. P., Woldeyes, F., Willis, K., Demissew, S., Wilkin, P. (2019). Enset in Ethiopia: a poorly characterized but resilient starch staple. *Annals of Botany*, *123*(5), 747-766. https://doi.org/10.1093/aob/mcy214

Borrell, J. S., Goodwin, M., Blomme, G., Jacobsen, K., Wendawek, A. M., Gashu, D., Lulekal, E., Asfaw, Z., Demissew, S., Wilkin, P. (2020). Enset-based agricultural systems in Ethiopia: A systematic review of production trends, agronomy, processing and the wider food security applications of a neglected banana relative. *Plants, People, Planet*, *2*(3), 212-228. https://doi.org/DOI:10.1002/ppp3.10084

Bosha, A., Dalbato, A. L., Tana, T., Mohammed, W., Tesfaye, B., Karlsson, L. M. (2016). Nutritional and chemical properties of fermented food of wild and cultivated genotypes of enset (Ensete ventricosum). *Food Research International*, *89*, 806-811. https://doi.org/10.1016/j.foodres.2016.10.016

Brandt, S., Spring, A., Hiebsch, C., McCabe, J., Tabogie, E., Diro, M., Wolde-Michael, G., Yntiso, G., Shigeta, M., Tesfaye, S. (1997). *The "Tree Against Hunger": Enset-based agricultural systems in Ethiopia*. (Vol. 56). American Association for the Advancement of Science, Washington, DC, USA.

Carmo, T. D. d., Barbosa, P. M., Geraseev, L. C., Costa, D. S., Seles, G. M., Duarte, E. R. (2018). Intake and digestibility of lamb fed diets containing banana crop residues. *Pesquisa Agropecuária Brasileira*, *53*, 197-205. https://doi.org/doi.org/10.1590/S0100-204X2018000200008

Chandrasekara, A., Josheph Kumar, T. (2016). Roots and tuber crops as functional foods: a review on phytochemical constituents and their potential health benefits. *International journal of food science 2016*, 15. https://doi.org/10.1155/2016/3631647

Chung, Y. C., Chiang, B. H., Wei, J. H., Wang, C. K., Chen, P. C., Hsu, C. K. (2008). Effects of blanching, drying and extraction processes on the antioxidant activity of yam (Dioscorea alata). *International Journal of Food Science and Technology*, *43*(5), 859-864. https://doi.org/10.1111/j.1365-2621.2007.01528.x

Cornago, D. F., Rumbaoa, R. G. O., Geronimo, I. M. (2011). Philippine Yam (Dioscorea spp.) tubers phenolic content and antioxidant capacity. *Philippine Journal of Science*, *140*(2), 145-152.

Daba, T., Shigeta, M. (2016). Enset (Ensete ventricosum) production in Ethiopia: its nutritional and socio-cultural values. *Agriculture Food Sciences Research*, *3*(2), 66-74. https://doi.org/10.20448/journal.512/2016.3.2/512.2.66.74

Debebe, A., Chandravanshi, B., Wondimu, T. (2012). Metallic nutrients in enset (Ensete ventricosum) corm cultivated in Wolliso and Wolkite Towns in Ethiopia. *Ethiopian Journal of Science*, *35*(2), 71-80.

Dhingra, D., Michael, M., Rajput, H., Patil, R. (2012). Dietary fibre in foods: a review. *Journal of Food Science and Technology*, *49*(3), 255-266. https://doi.org/DOI 10.1007/s13197-011-0365-5

Fontes, S. d. M., Cavalcanti, M. T., Candeia, R. A., Almeida, E. L. (2017). Characterization and study of functional properties of banana starch green variety of Mysore (Musa AAB-Mysore). *Food Science and Technology*, *37*, 224-231. https://doi.org/doi.org/10.1590/1678-457X.18916

Forsido, S. F., Rupasinghe, H. V., Astatkie, T. (2013). Antioxidant capacity, total phenolics and nutritional content in selected Ethiopian staple food ingredients. *Tropical Resource Mangement Papers*, *64*(8), 915-920. https://doi.org/10.3109/09637486.2013.806448

Gebre-Mariam, T., Abeba, A., Schmidt, P. (1996). Isolation and physico-chemical properties of enset starch. *Starch-Stärke*, *48*(6), 208-214. https://doi.org/10.1002/star.19960480603

Grace, D. (2016). Influencing food environments for healthy diets through food safety. In *FAO, Influencing food environments for healthy diets* (pp. 43-75). FAO.

Hirose, R., Tezuka, Y., Kondo, T., Hirao, K., Hatta, T., Nemoto, S., Saio, K., Takahashi, S., Kainuma, K. (2010). Characteristic physico-chemical properties and potential uses of Enset (Ensete ventricosum) starch: comparative studies with starches of potato, sago and corn. *Journal of Applied Glycoscience*, *57*(3), 185-192. https://doi.org/doi.org/10.5458/jag.57.185

Kandasamy, S., Aradhya, S. M. (2014). Polyphenolic profile and antioxidant properties of rhizome of commercial banana cultivars grown in India. *Food Bioscience*, *8*, 22-32. https://doi.org/10.1016/j.fbio.2014.10.001

Karssa, T., Papini, A. (2018). Effect of clonal variation on quality of Kocho, traditional fermented food from enset (Ensete ventricosum), Musaceae. *International Journal of Food Science and Nutrition Engineering*, *8*, 79-85. https://doi.org/10.5923/j.food.20180803.04

Karssa, T., Papini, A. (2018). Effect of clonal variation on quality of Kocho, traditional fermented food from enset (Ensete ventricosum), Musaceae. *Int. J. Food Sci. Nutr. Eng*, *8*, 79-85. https://doi.org/DOI:10.5923/j.food.20180803.04

Kaur, C., Kapoor, H. C. (2002). Anti-oxidant activity and total phenolic content of some Asian vegetables. *International Journal of Food Science & Technology*, *37*(2), 153-161. https://doi.org/doi.org/10.1046/j.1365-2621.2002.00552.x

Liang, Z., MU, T.-h., ZHANG, R.-f., SUN, Q.-h., XU, Y.-w. (2019). Nutritional evaluation of different cultivars of potatoes (Solanum tuberosum L.) from China by grey relational analysis (GRA) and its application in potato steamed bread making. *Journal of integrative agriculture*, *18*(1), 231-245. https://doi.org/10.1016/S2095-3119(18)62137-9

Magbagbeola, J., Adetoso, J., Owolabi, O. (2010). Neglected and underutilized species (NUS): a panacea for community focused development to poverty alleviation/poverty reduction in Nigeria. *Journal of Economics and International Finance*, *2*(10), 208-211.

Mayes, S., Massawe, F., Alderson, P., Roberts, J., Azam-Ali, S., Hermann, M. (2012). The potential for underutilized crops to improve security of food production. *Journal of Experimental Botany*, *63*(3), 1075-1079. https://doi.org/10.1093/jxb/err396

Mohammed, B., Gabel, M., Karlsson, L. M. (2013). Nutritive values of the drought tolerant food and fodder crop enset. *African Journal of Agricultural Research*, *8*(20), 2326-2333. https://doi.org/10.5897/AJAR12.1296

Moorthy, S. N. (2002). Physicochemical and functional properties of tropical tuber starches: a review. *Starch-Stärke*, *54*(12), 559-592. https://doi.org/10.1002/1521-379X(200212)54:12<559::AID-STAR2222559>3.0.CO;2-F

Negash, A., Tsegaye, A., van Treuren, R., Visser, B. (2002). AFLP analysis of enset clonal diversity in south and southwestern Ethiopia for conservation. *Crop Science*, *42*(4), 1105-1111. https://doi.org/10.2135/cropsci2002.1105

Nurfeta, A., Tolera, A., Eik, L., Sundstøl, F. (2009). Feeding value of enset (Ensete ventricosum), Desmodium intortum hay and untreated or urea and calcium oxide treated wheat straw for sheep. *Journal of Animal Physiology and Animal Nutrition*, *93*(1), 94-104. https://doi.org/10.1111/j.1439-0396.2007.00784.x

Nurfeta, A., Tolera, A., Eik, L. O., Sundstøl, F. (2008). The supplementary value of different parts of enset (Ensete ventricosum) to sheep fed wheat straw and Desmodium intortum hay. *Livestock Science*, *119*(1-3), 22-30. https://doi.org/10.1016/j.livsci.2008.02.010

Oke, O. L., Redhead, J., Hussain, M. A. (1990). Roots, tubers, plantains and bananas in human nutrition. *FAO food and nutrition series*, 24.

Patthamakanokporn, O., Puwastien, P., Nitithamyong, A., Sirichakwal, P. P. (2008). Changes of antioxidant activity and total phenolic compounds during storage of selected fruits. *Journal of Food Composition Analysis*, *21*(3), 241-248. https://doi.org/10.1016/j.jfca.2007.10.002

Ravi, I., Mustaffa, M. M. (2013). Starch and amylose variability in banana cultivars. *Indian Journal of Plant Physiology*, *18*(1), 83-87. https://doi.org/DOI:10.1007/s40502-013-0014-2

Rice-Evans, C., Miller, N., Paganga, G. (1997). Antioxidant properties of phenolic compounds. *Trends in Plant Science*, *2*(4), 152-159. https://doi.org/doi.org/10.1016/S1360-1385(97)01018-2

Seung, D. (2020). Amylose in starch: towards an understanding of biosynthesis, structure and function. *New Phytologist*, *228*(5), 1490-1504. https://doi.org/doi.org/10.1111/nph.16858

Shan, B., Cai, Y. Z., Sun, M., Corke, H. (2005). Antioxidant capacity of 26 spice extracts and characterization of their phenolic constituents. *Journal of agricultural food chemistry*, *53*(20), 7749-7759. https://doi.org/10.1021/jf051513y

Singleton, V. L., Orthofer, R., Lamuela-Raventós, R. M. (1999). [14] Analysis of total phenols and other oxidation substrates and antioxidants by means of folin-ciocalteu reagent. In *Methods in Enzymology* (Vol. 299, pp. 152-178). Elsevier. https://doi.org/10.1016/S0076-6879(99)99017-1

Spring, A., Haile, B., Tesfaye, S., Abebe, Y., Amaldegen, A., Gizachew, W., Tadesse, T. (1996). Enset farming systems in southern region, Ethiopia: Report on a rapid rural appraisal in Gurage, Hadiya, and Sidama zones. Addis Ababa: GTZ.

Tamrat, S., Borrell, J. S., Biswas, M. K., Gashu, D., Wondimu, T., Vásquez-Londoño, C. A., Heslop-Harrison, P. J., Demissew, S., Wilkin, P., Howes, M.-J. R. (2020). Micronutrient composition and microbial community analysis across diverse landraces of the Ethiopian orphan crop enset. *Food Research International*, *137*, 109636. https://doi.org/10.1016/j.foodres.2020.109636

Tobiaw, D. C., Bekele, E. (2011). Analysis of genetic diversity among cultivated enset (Ensete ventricosum) populations from Essera and Kefficho, southwestern part of Ethiopia using inter simple sequence repeats (ISSRs) marker. *African Journal of Biotechnology*, *10*(70), 15697-15709. https://doi.org/10.5897/AJB11.885

Tsegaye, A. (2002). *On indigenous production, genetic diversity and crop ecology of enset (Ensete ventricosum (Welw.) Cheesman).* Wageningen University and Research.

Tsegaye, A., Struik, P. C. (2001). Enset (Ensete ventricosum (Welw.) Cheesman) kocho yield under different crop establishment methods as compared to yields of other carbohydrate-rich food crops. *Netherlands Journal of Agricultural Science*, *49*(1), 81-94. https://doi.org/10.1016/s1573-5214(01)80017-8

Tuffa, A. C. (2019). Value chain and nutritional analyses of Warqe food products in relation to post-harvest losses. *Swedish University of Agricultural Sciences, Uppsala, Sweden.*

Volden, J., Borge, G. I. A., Bengtsson, G. B., Hansen, M., Thygesen, I. E., Wicklund, T. J. F. C. (2008). Effect of thermal treatment on glucosinolates and antioxidant-related parameters in red cabbage (Brassica oleracea L. ssp. capitata f. rubra). *Food Chemistry*, *109*(3), 595-605. https://doi.org/10.1016/j.foodchem.2008.01.010

Von Grebmer, K., Bernstein, J., Hammond, L., Patterson, F., Sonntag, A., Klaus, L., Fahlbusch, J., Towey, O., Foley, C., Gitter, S. (2018). Global hunger index: Forced migration and hunger. *International food policy institute*, 68.

Waliszewski, K. N., Aparicio, M. A., Bello, L. s. A., Monroy, J. A. (2003). Changes of banana starch by chemical and physical modification. *Carbohydrate polymers*, *52*(3), 237-242. https://doi.org/10.1016/S0144-8617(02)00270-9

Westphal, E., Stevels, J. M. C., Stevels, J. M. C. (1975). *Agricultural systems in Ethiopia* (Vol. 826). Wageningen: Centre for Agricultural Publishing and Documentation.

Yemataw, Z., Chala, A., Ambachew, D., Studholme, D. J., Grant, M. R., Tesfaye, K. (2017). Morphological variation and inter-relationships of quantitative traits in enset (Ensete ventricosum (Welw.) Cheesman) germplasm from south and south-western Ethiopia. *Plants*, *6*(4), 56. https://doi.org/10.3390/plants6040056

Yemataw, Z., Mohamed, H., Diro, M., Addis, T., Blomme, G. (2014). Ethnic-based diversity and distribution of enset (Ensete ventricosum) clones in southern Ethiopia. *Journal of Ecology and the Natural Environment*, *6*(7), 244-251. https://doi.org/DOI:10.5897/JENE2014.0450

Zargar, M., Azizah, A. H., Roheeyati, A. M., Fatimah, A. B., Jahanshiri, F., Pak-Dek, M. S. (2011). Bioactive compounds and antioxidant activity of different extracts from Vitex negundo leaf. *Journal of Medicinal Plants Research*, *5*(12), 2525-2532. https://doi.org/10.5897/JMPR.9000293

Tables

Table 1. Enset (*Ensete ventricosum* (Welw.) Cheesman) genotypes selected for nutritional analyses.

No	Vernacular name	Sequence code	Biological status	Geographical origin
1	Gena	68	Cultivated	Sidama
2	Mundraro	48	Cultivated	Sidama
3	Kiticho	2	Cultivated	Sidama
4	Ado	58	Cultivated	Sidama
5	Sediso	65	Cultivated	Sidama
6	Medasho	1	Cultivated	Sidama
7	Ginbowe	114	Cultivated	Gurage
8	Agade	108	Cultivated	Gurage
9	Nichew	109	Cultivated	Gurage
10	Badedet	115	Cultivated	Gurage
11	Derea	111	Cultivated	Gurage
12	Lemat	112	Cultivated	Gurage
13	Separa	105	Cultivated	Gurage
14	Astara	129	Cultivated	Gurage
15	Entada1	255	Cultivated	Sidama
16	Entada2	263	Cultivated	Sidama
17	Wild enset1	235	Wild	Sidama
18	Wild enset2	229	Wild	Sidama

Table 2. Chemical composition of cultivated enset, wild enset and Entada landraces in the leaf sheath (LS) and corm (Co) (based on dry weight (DW).

	Amylose[1]		WSC[2]		NDF[2]	Carbon[2]		Nitrogen[2]	
Genotypes	LS	Co	LS	Co	LS	LS	Co	LS	Co
Mundraro	64.1^{a}	29.6^{c}	14.6^{b}	7.6^{ef}	209.0^{c}	42.5^{a}	43.5^{a}	0.5$^{f–i}$	1.6^{b}
Sediso	55.2^{c}	22.9^{efg}	13.4^{c}	5.0^{i}	134.9^{de}	42.3^{a}	43.2^{ab}	0.5^{ijk}	1.4^{c}
Kiticho	54.1^{cd}	22.2^{efg}	25.5^{a}	10.1^{c}	208.2^{c}	40.3^{d}	42.8$^{a–d}$	0.5^{jkl}	1.1^{ef}
Gena	45.4^{e}	22.8^{efg}	7.1^{g}	9.9^{cd}	104.6$^{e–h}$	40.3^{d}	42.1^{de}	0.7^{cd}	1.1^{ef}
Ado	36.3^{h}	21.6^{fg}	6.4^{gh}	5.2^{i}	97.0^{ghi}	41.1^{bcd}	42.6^{bcd}	0.5$^{h–k}$	1.2^{e}
Medasho	29.4^{i}	25.1^{de}	7.0^{g}	6.0^{gh}	97.3^{ghi}	41.7^{ab}	42.5^{bcd}	0.6^{efg}	0.9^{i}
Lemat	52.1^{cd}	21.6^{efg}	14.5^{b}	8.1^{e}	131.0^{def}	42.1^{ab}	42.1^{cde}	0.6^{fgh}	1.3^{d}
Badedet	46.5^{e}	24.8^{def}	5.7^{hi}	11.1^{b}	83.1^{hij}	41.9^{ab}	42.9^{abc}	0.4^{l}	1.1^{fg}
Separa	44.8^{ef}	21.6^{fg}	8.4^{f}	7.2^{f}	108.3$^{e–h}$	42.0^{ab}	42.2^{cde}	0.7^{cde}	0.9^{i}
Derea	41.7^{fg}	23.0^{efg}	11.3^{d}	7.5^{ef}	147.2^{d}	41.6^{abc}	43.1^{ab}	1.6^{a}	3.0^{a}
Ginbowe	39.4^{gh}	26.5^{cd}	8.3^{f}	4.3^{j}	101.2^{fgh}	41.8^{ab}	43.1^{ab}	1.0^{b}	0.9^{i}
Agade	37.0^{h}	21.0^{g}	12.3^{d}	5.2^{i}	97.3^{ghi}	41.6^{abc}	43.1^{ab}	0.7^{cd}	0.9^{hi}
Nichew	37.0^{h}	21.6^{fg}	13.9^{bc}	7.8^{ef}	94.9$^{g–j}$	41.8^{ab}	42.8$^{a–d}$	0.6^{def}	0.6^{j}
Astara	22.0^{j}	21.4^{fg}	9.8^{e}	17.2^{a}	121.0$^{d–g}$	41.8^{ab}	42.7$^{a–d}$	0.7^{c}	1.0^{gh}
Entada1	54.8^{c}	23.8^{def}	4.8^{i}	7.6^{ef}	67.2^{ij}	41.7^{ab}	42.8$^{a–d}$	0.6^{efg}	0.4^{k}
Entada2	51.0^{d}	24.7^{def}	3.7^{j}	6.4^{g}	66.0^{j}	41.5^{abc}	42.6^{bcd}	0.5$^{g–j}$	0.3^{k}
Wild1	59.4^{b}	42.0^{b}	4.8^{i}	5.6^{hi}	298.3^{b}	37.8^{e}	41.7^{e}	0.3^{m}	0.3^{k}
Wild2	46.5^{e}	62.6^{a}	5.1^{i}	9.3^{d}	350.6^{a}	40.6^{cd}	33.8^{f}	0.4^{kl}	1.0^{h}
Mean	45.4	26.6	9.8	7.8	139.8	41.4	42.2	0.6	1.1

LS: leaf sheath; Co: corm; NDF: neutral detergent fiber; WSC: water-soluble carbohydrates; [1]based on two replications; [2]based on three replications. Means within a column followed by the same letter(s) are not significantly different according to Tukey's test ($P \leq 0.05$).

Table 3. ANOVA analysis of chemical composition across tissues (corm and leaf sheath), cultivation status, genotypes and geographical origin of genotype.

Factors	Compared fraction	df	AM[1]	NDF[2]	WSC[2]	N[2]	C[2]
	Corm		26.60	-	7.83	0.62	42.19
Tissue (corm and leaf sheath)	Leaf sheath		45.36	-	9.79	1.05	41.35
	P-value	1	0.001***	-	0.02*	0.001***	0.01**
Cultivation status							
Cultivated vs. Wild (leaf sheath)	Cultivated		52.94	116.77	10.40	0.65	39.21
	Wild		44.41	324.41	4.93	0.36	41.62
	P-value	1	0.13 ns	0.001***	0.01**	0.01**	0.001***
Cultivated vs wild (corm)	Cultivated		23.39	-	7.88	1.10	42.75
	Wild		52.28	-	7.44	0.64	37.75
	P-value	1	0.001***	-	0.73 ns	0.08 ns	0.001***
Genotypes							
Enset vs. Entada (leaf sheath)	Enset		43.20	123.93	11.29	0.67	41.63
	Entada		52.89	66.62	4.22	0.54	41.62
	P-value	1	0.09 ns	0.001***	0.001***	0.30 ns	0.96 ns
Enset vs. Entada (corm)	Enset		23.26	-	8.01	1.21	42.76
	Entada		24.25	-	7.01	0.32	42.69
	P-value	1	0.44 ns	-	0.47 ns	0.001***	0.76 ns
Enset vs. Wild vs. Entada (leaf sheath)	Enset		43.20	123.93[c]	11.28a	0.67[a]	41.62[a]
	Entada		52.89	66.62[b]	4.22b	0.54[ab]	41.63[a]
	Wild		52.94	324.42[a]	4.93b	0.36[b]	39.21[b]
	P-value	2	0.07 ns	0.001***	0.00***	0.03*	0.001***
Enset vs. wild vs. Entada (corm)	Enset		23.26[b]	-	8.01	1.21[a]	42.76[a]
	Entada		24.25[b]	-	7.02	0.32[b]	42.69[a]
	Wild		52.28[a]	-	7.43	0.64[b]	37.75[b]
	P-value	2	0.001***	-	0.71 ns	0.001***	0.001***
Geographical origin	Sidama		24.02	-	7.30	1.03	42.43
corm	Gurage		22.69	-	8.54	1.35	43.42
	P-value	1	0.16 ns	-	0.23 ns	0.07 ns	0.001***
leaf sheath	Sidama		47.02	141.84	12.33	0.57	41.26
	Gurage		40.05	110.50	10.50	0.74	41.88
	P-value	1	0.07ns	0.01**	0.25 ns	0.06 ns	0.004**

Table 4. Antioxidant capacity (μmolg^{-1} FW) and total phenolic (mg GAE 100 g^{-1} FW) content in different parts of cultivated enset, wild enset and Entada landraces (based on fresh weight (FW)).

	Antioxidant capacity		Total phenol content	
Genotypes	LS	Co	LS	Co
Kiticho	8.4^{b}	5.0^{ef}	108.0^{a}	67.6^{cd}
Gena	5.5^{ef}	4.3^{g}	72.1^{d}	61.9^{de}
Medasho	5.3^{fg}	5.2^{de}	72.2^{d}	68.5^{cd}
Mundraro	5.2^{fg}	7.3^{c}	59.4^{f}	94.2^{b}
Sediso	4.8^{h}	4.5^{g}	71.5^{d}	62.3^{de}
Ado	3.7^{j}	3.3^{j}	51.7^{g}	48.8^{hi}
Nichewa	7.6^{c}	3.6^{i}	98.9^{b}	61.6^{def}
Separa	7.4^{d}	4.8$_{f}$	85.9^{c}	64.0^{de}
Lemate	5.7^{e}	2.4^{k}	68.7^{de}	45.6^{ij}
Astara	5.1^{gh}	2.5^{k}	67.7^{de}	39.4^{j}
Gimbowe	5.1^{gh}	3.6^{i}	63.4^{ef}	63.4^{de}
Badedet	3.4^{k}	5.4^{d}	45.7^{h}	73.1^{c}
Derea	2.5^{m}	4.0^{h}	32.1^{i}	51.7^{ghi}
Agade	2.4^{l}	8.2^{b}	35.4^{i}	104.3^{a}
Entada1	4.3^{i}	4.9^{ef}	51.3^{g}	62.0^{de}
Entada2	2.6^{l}	4.2^{gh}	37.4^{i}	53.7^{fgh}
Wild1	9.2^{a}	4.8^{f}	98.3^{b}	58.5^{efg}
Wild2	3.9^{j}	8.8^{a}	44.7^{h}	101.8^{ab}
Mean	5.1	4.8	64.7	65.7

Based on three replications. Means within a column followed by the same letter (s) are not significantly different according to Tukey's test (P ≤ 05).

Table 5. ANOVA analysis of bioactive compounds across tissues (corm and leaf sheath), cultivation status, genotypes, and geographical origin of genotypes.

Factors	Compared fraction	df	TPC	AO
Tissue	Corm		65.69	4.82
	Leaf sheath		64.68	5.08
	P-value	1	0.79 ^{ns}	0.47 ^{ns}
Cultivation status	Cultivated		63.83	4.89
Leaf sheath	Wild		71.48	6.54
	P-value	1	0.43 ^{ns}	0.06 ^{ns}
Corm	Cultivated		63.88	4.68
	Wild		80.17	6.81
	P-value	1	0.03*	0.001***
Genotypes	Enset		66.62	5.11
Enset vs. Entada	Entada		44.33	3.41
(leaf sheath)	P-value	2	0.01**	0.03*
Enset vs. Entada	Enset		64.74	4.57
(corm)	Entada		57.86	4.58
	P-value	2	0.34 ^{ns}	0.99 ^{ns}
Enset vs. wild vs. Entada	Enset		66.62^{a}	5.11^{ab}
(leaf steath)	Entada		44.33^{b}	3.41^{b}
	Wild		71.48^{a}	6.54^{a}
	P-value	2	0.05*	0.02*
Enset vs. wild vs. Entada	Enset		64.74	4.57
(corm)	Entada		57.86	4.58
	Wild		80.17	6.81
	P-value	2	0.07 ^{ns}	0.09 ^{ns}
Geographical origin	Sidama		67.22	4.91
Corm	Gurage		62.88	4.31
	P-value	1	0.42 ^{ns}	0.23 ^{ns}
Leaf sheath	Sidama		72.48	5.48
	Gurage		62.23	4.83
	P-value	1	0.12 ^{ns}	0.26 ^{ns}

Table 6. Mineral content in leaf sheath and corm tissues of cultivated enset, wild enset and Entada landraces (based on dry weight (DW)).

Genotype	Part used	[1]Ca	[1]K	[1]Mg	[1]P	[1]S	[1] Na	[2]Fe	[2]Cu	[2]Zn
Gena	**Leaf sheath**	**1.20**	**29**	**0.72**	**1.4**	**0.49**	**0.02**	**23**	**2.00**	**9**
	Corm	1.10	14	0.32	0.81	0.75	0.02	7.20	1.70	47
Mundraro	**Leaf sheath**	**2.80**	**43**	**0.87**	**1.50**	**0.56**	**0.02**	**36**	**3.80**	**13**
	Corm	1	20	0.77	1.10	2.20	0.02	18	7.50	480
Kiticho	**Leaf sheath**	**4**	**41**	**0.83**	**1.20**	**0.33**	**0.01**	**23**	**1.70**	**24**
	Corm	1.40	15	0.89	1.10	2.10	0.02	9.10	7.50	240
Ado	**Leaf sheath**	**1.50**	**25**	**0.99**	**1**	**0.34**	**0.01**	**21**	**1.30**	**8.50**
	Corm	0.71	15	0.56	1.20	2	0.01	10	5.10	270
Sediso	**Leaf sheath**	**3.20**	**37**	**0.92**	**1.10**	**0.41**	**0.02**	**20**	**2.00**	**7.60**
	Corm	0.57	14	0.79	1.40	2.30	0.01	10	3.70	210
Medasho	**Leaf sheath**	**2.50**	**23**	**0.58**	**1.30**	**0.46**	**0.01**	**18**	**2.40**	**12**
	Corm	1.20	17	0.75	1.10	1.90	0.01	8.80	6.60	75
Ginbowe	**Leaf sheath**	**1.90**	**28**	**0.71**	**1.10**	**0.39**	**0.01**	**13**	**1.10**	**7.70**
	Corm	0.64	14	0.54	1.10	0.97	0.01	9.90	3.40	76
Agade	**Leaf sheath**	**1.60**	**23**	**0.5**	**0.98**	**0.31**	**0.01**	**15**	**1.10**	**5.60**
	Corm	0.47	9.80	0.35	0.86	1.90	0.01	27	5.80	240
Nichew	**Leaf sheath**	**1.20**	**20**	**0.38**	**1.10**	**0.36**	**0.01**	**24**	**1.60**	**5.10**
	Corm	0.68	9.20	0.43	1.20	1.60	0.01	15	5.10	190
Badedet	**Leaf sheath**	**1.10**	**18**	**0.68**	**0.96**	**0.40**	**0.04**	**9.5**	**1.80**	**6.60**
	Corm	0.73	15	0.35	0.65	1.20	0.04	7.70	4.30	150
Derea	**Leaf sheath**	**1.70**	**31**	**0.90**	**2**	**0.55**	**0.01**	**20**	**2.70**	**21**
	Corm	1.10	14	0.40	1.20	1.90	0.01	6.80	2.50	87
Lemat	**Leaf sheath**	**1.80**	**33**	**0.80**	**0.77**	**0.32**	**0.01**	**14**	**1.50**	**7.30**
	Corm	0.58	13	0.37	0.97	1.10	0.02	7.70	1.70	92
Separa	**Leaf sheath**	**1.90**	**26**	**0.41**	**0.96**	**0.28**	**0.02**	**38**	**1.00**	**6.10**
	Corm	0.76	20	0.43	1.40	2.20	0.01	22	3.50	150
Astara	**Leaf sheath**	**2.40**	**25**	**1.50**	**0.54**	**0.54**	**0.01**	**23**	**2.70**	**5.20**
	Corm	1.10	5.4	0.66	0.49	0.70	0.01	14	8.30	78
Entada1	**Leaf sheath**	**1.80**	**34**	**0.94**	**3.50**	**0.43**	**0.03**	**7.30**	**1.50**	**7.20**
	Corm	0.63	14	0.84	2.10	1.50	0.03	3.30	0.29	15
Entada 2	**Leaf sheath**	**1.50**	**31**	**0.79**	**3.30**	**0.34**	**0.01**	**11**	**1.50**	**4.40**
	Corm	0.87	18	0.85	1.70	1.10	0.07	4.20	0.22	46
Wild1	**Leaf sheath**	**13**	**62**	**1.30**	**5.60**	**0.21**	**0.03**	**81**	**0.69**	**11**
	Corm	3.50	38	2.40	1.30	1.60	0.50	17	7.30	99
Wild2	**Leaf sheath**	**12**	**39**	**2**	**1**	**0.33**	**0.06**	**87**	**0.86**	**3.90**
	Corm	6.10	110	3.50	0.68	0.34	0.07	20	5.20	74

[1] Data are expressed on g/kg and [2] data are expressed on mg/kg.

Table 7. ANOVA analysis of major and minor elements content across tissues (corm and leaf sheath), genotypes and geographical origin of genotypes.

Factors	Compared fraction	df	Ca	K	Mg	P	S	Na	Fe	Cu	Zn
Tissue	Corm		1.3	20.9	0.8	1.1	1.5	0.1	12.1	4.4	145.5
	Leaf sheath		3.2	31.6	0.9	1.6	0.4	0.0	26.9	1.7	9.2
	P-value	1	0.01^{**}	0.05^{*}	0.8^{ns}	0.07^{ns}	0.001^{***}	0.23^{ns}	0.001^{***}	0.001^{***}	0.001^{***}
Genotype	Genotype		2.23	26.21	0.86	1.37	0.96	0.03	19.49	3.08	77.34
	P-value	17	0.01^{**}	0.12^{ns}	0.01^{**}	0.09^{ns}	0.52^{ns}	0.38^{ns}	0.12^{ns}	0.3^{ns}	0.46^{ns}
Region	Sidama		1	15.8	0.7	1.1	1.9	0.02	10.5	5.4	220.3
Corm	Gurage		0.8	12.6	0.4	1	1.4	0.02	13.8	4.3	132.9
	P-value	1	0.12^{ns}	0.12^{ns}	0.02^{*}	0.36^{ns}	0.17^{ns}	1.00^{ns}	0.35^{ns}	0.40^{ns}	0.17^{ns}
Region	Sidama		2.5	33	0.8	1.3	0.43	0.02	23.5	2.2	8.1
Leaf sheath	Gurage		1.7	25.5	0.7	1.1	0.4	0.01	19.56	1.7	12.4
	P-value	1	0.06^{ns}	0.06^{ns}	0.61^{ns}	0.31^{ns}	0.48^{ns}	0.11^{ns}	0.38^{ns}	0.23^{ns}	0.19^{ns}

Figures

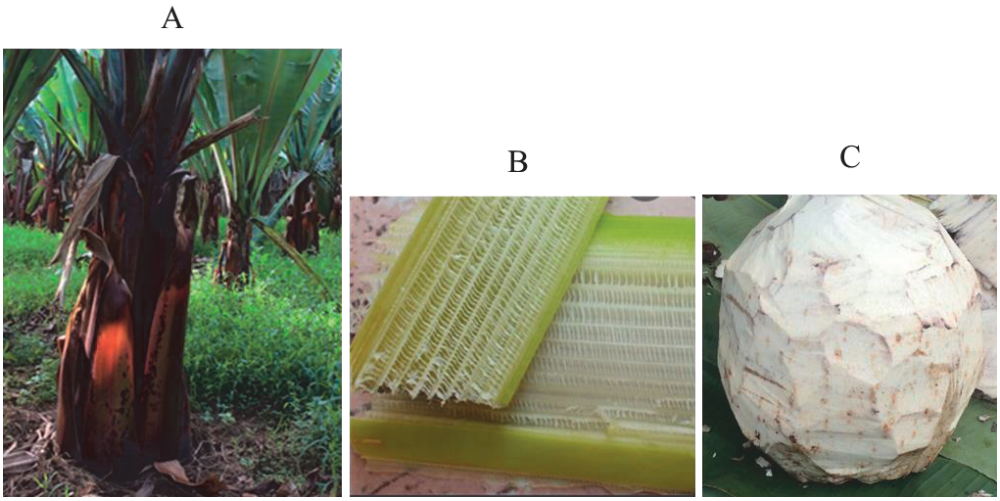


Figure 1. *A: Enset plant morphology; B: fresh leaf sheath cross section; C: fresh corm from underground part.*

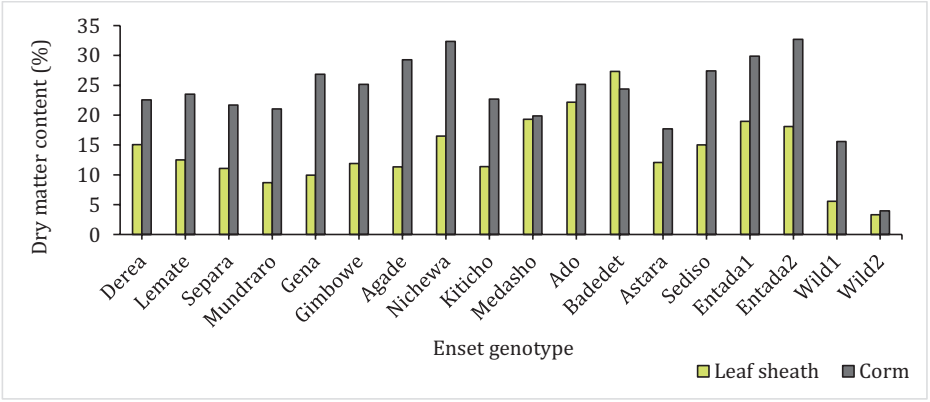


Figure 2. *Dry matter content in leaf sheaths and corms of cultivated enset, wild enset and Entada landraces.*

	AM		NDF	WSC		TPC		AO		C		N	
Genotypes	LS	Co	LS	LS	Co	LS	Co	LS	Co	LS	Co	LS	Co
Gena													
Mundraro													
Kiticho													
Ado													
Sediso													
Medasho													
Ginbowe													
Agade													
Nichew													
Badedet													
Derea													
Lemat													
Separa													
Astara													
Entada1													
Entada 2													
Wild1													
Wild2													

Figure 3. *Heat map showing the chemical contents of leaf sheaths and corms of cultivated enset, wild enset and Entada landraces. LS: leaf sheath; Co: corm; AM: amylose (%); WSC: water soluble carbohydrates (%); NDF: neutral detergent fiber (g/kg); TPC: Total phenolic contents (mg GAE100 g^{-1} FW); AO: antioxidant capacity (µmolg^{-1} FW); C: carbon (%); N: nitrogen (%).*

	Ca		K		Mg		P		S		Na		Fe		Cu		Zn	
Genotypes	LS	Co	LS	Co	LS	Co	LS	Co	LS	Co	LS	Co	LS	Co	LS	Co	LS	Co
Gena																		
Mundraro																		
Kiticho																		
Ado																		
Sediso																		
Medasho																		
Ginbowe																		
Agade																		
Nichew																		
Badadet																		
Derea																		
Lemat																		
Separa																		
Astara																		
Entada1																		
Entada 2																		
Wild1																		
Wild2																		

Min. Max.

Figure 4. *Heat map showing major and minor minerals of leaf sheaths and corms of cultivated enset, wild enset and Entada landraces LS: leaf sheath; Co: corm. Major elements are expressed in g/kg and minor elements mg/kg*

	AM_LF	AM_Corm	NDF_LF	WSC _LF	WSC_corm	TPC_LF	TPC_corm	AO_LF
AM_LF								
AM_Corm	0.41							
NDF_LF	0.65**	0.35						
WSC _LF	0.49	-0.09	0.78**					
WSC_corm	-0.26	-0.18	0.12	0.07				
TPC_LF	0.13	-0.16	0.25	0.5	0.17			
TPC_corm	0.34	0.39	0.16	0.15	-0.36	-0.2		
AO_LF	0.2	-0.07	0.29	0.49	0.17	0.98***	-0.16	
AO_corm	0.33	0.35	0.22	0.13	-0.3	-0.26	0.97***	-0.23

[1] Data obtained from leaf sheaths, [2] data obtained from corm.

Figure 5. *Heat-map displaying the extent and direction of correlations (r) between all pairs of variables. Positive correlations are displayed in red and negative correlations in blue. The intensity of the color is proportional to the magnitude of the correlation coefficient. AM: amylose; WSC: water-soluble carbohydrates; NDF: neutral detergent fiber; AO: antioxidant capacity; TPC: total phenolic content.*

Supplementary Figure

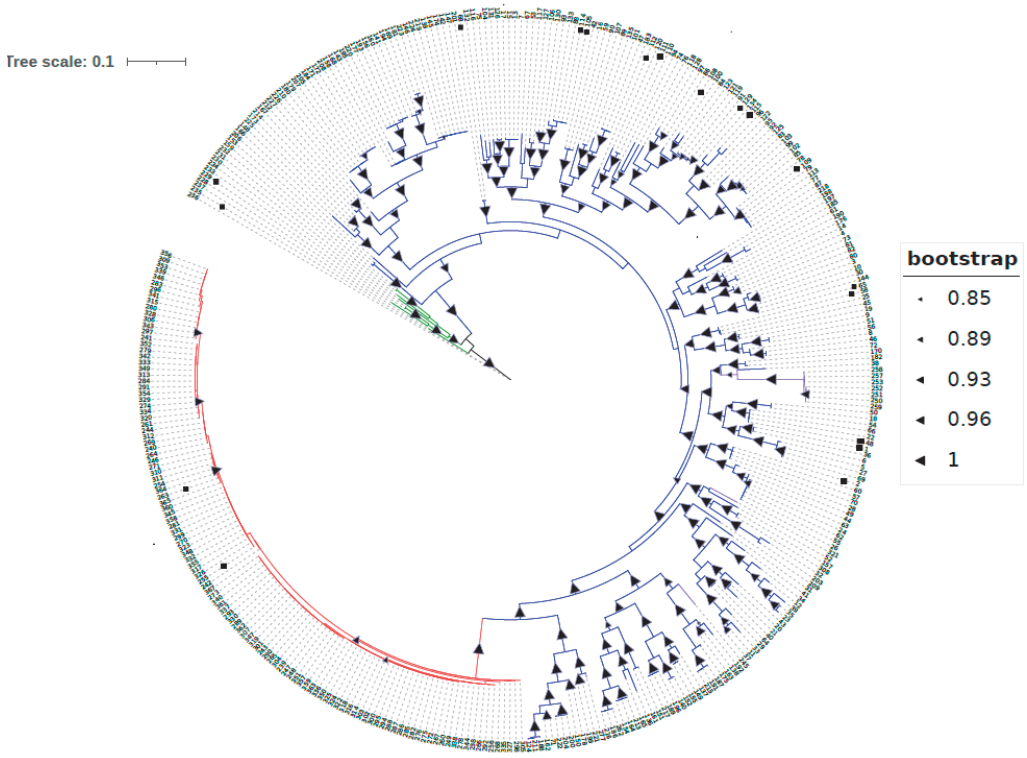


Figure S1. *Phylogenetic tree showing diversity across the wild enset, cultivated enset and Entada. Different colours represent their propagation method and cultivation status, i.e., cultivated enset: blue; wild enset: green; Entada landraces: red; cultivated enset with suckers: purple; selected accessions for nutritional analyses: black dots on the branch length.*

ISBN: 978-82-575-2079-3
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

116763 / ANDVORDGRAFISK.NO