Enset (*Ensete ventricosum*) Landrace Database

Muzemil, S; Olango, T.M, Yemataw, Z, et al.

Note: This is a draft enset landrace database and a manuscript describing the database methodology, curation process, and comprehensive analysis is in preparation. Please check back for updates or contact the corresponding authors if you're interested to contribute.

DOI: 10.5281/zenodo.8287638 **Version:** 0 (Draft) **License:** CC BY-SA 4.0 **Last Updated:** 2025-12-23

This database represents an effort to create a publicly available, open-access record of enset landrace lists with associated attributes that can be freely utilized by researchers, breeders, conservationists, and policymakers worldwide.

Background: The Crop and Its Significance

Enset (*Ensete ventricosum* (Welw.) Cheesman), commonly known as the "false banana" is a giant herbaceous monocarpic perennial belonging to the family Musaceae—the same family as bananas (*Musa* spp.). Unlike its fruit-bearing relatives, enset is cultivated for its starchy pseudostem and underground corm, which are harvested, fermented, and processed into storable food products. This unique crop serves as the primary staple food for approximately 20 million people across the densely populated highlands of southwestern Ethiopia, making it one of Africa's most important indigenous food security crops.

Despite its critical importance to Ethiopian food security and livelihoods, enset remains one of Africa's most understudied domesticated crops. It hasn't been cultivated outside Ethiopia, even though wild *E. ventricosum* populations occur naturally across the highlands of East and Southern Africa—from Ethiopia through Kenya, Uganda, and Tanzania to Zimbabwe and South Africa. This paradox of a crop with proven food security value yet extremely limited geographic adoption underscores the need for comprehensive research and documentation.

Landrace Diversity: A Genetic Treasure Under Threat

Domesticated enset exhibits extraordinary phenotypic diversity, with +1,200 named landraces documented across Ethiopia's enset-growing regions. These landraces—developed and maintained through centuries of farmer selection and vegetative (clonal) propagation—span multiple ethnic groups and at least 27 distinct linguistic communities. This rich agrobiodiversity encompasses extensive variation in:

- **Morphological traits**: Pseudostem and leaf coloration (green, red, purple, black), plant size and architecture, corm dimensions
- **Food and nutritional quality**: Kocho (fermented bread) texture and taste, bulla (starch) yield, amicho (boiled corm) palatability
- **Agronomic performance**: Maturity period, yield potential, drought tolerance, cold tolerance
- **Disease resistance**: Tolerance to bacterial wilt (*Xanthomonas vasicola* pv. *musacearum*), corm rot (*Erwinia* spp.), and other pathogens

- **Fiber quality**: Strength and durability of leaf sheath fibers used for rope, construction, and handicrafts
- **Medicinal properties**: Traditional uses for bone healing, placenta expulsion, liver ailments, and other therapeutic applications

Farmers actively recognize, select, and maintain landraces with specific trait combinations suited to their local environments, cultural preferences, and economic needs. This indigenous knowledge represents an invaluable but increasingly threatened resource.

Conservation Challenge: Why This Database Matters

Enset landrace diversity faces mounting threats from multiple sources: the spread of bacterial wilt disease (now present on up to 80% of farms in some regions), climate change and environmental degradation, agricultural modernization pressures, rural-urban migration of knowledgeable farmers, and the erosion of traditional cultivation practices. The loss of valuable enset genotypes has already been documented, with some landraces disappearing before they could be properly characterized or conserved.

Current germplasm conservation efforts remain critically insufficient. As a vegetatively propagated crop with recalcitrant seeds, enset cannot be easily stored in conventional seed banks. Field gene banks—the primary conservation method—are costly to establish and maintain, requiring continuous cultivation and management. Only a handful of Ethiopian institutions currently maintain living collections, and these represent only a fraction of the total diversity present in farmers' fields.

Compounding these challenges, comprehensive landrace information has historically remained fragmented and inaccessible. Data are scattered across peer-reviewed publications, unpublished theses and dissertations, institutional survey reports, government records, and undocumented indigenous knowledge systems. This fragmentation has severely limited the ability of researchers, breeders, and policymakers to understand the full scope of enset diversity, identify priority landraces for conservation, or develop evidence-based improvement strategies.

Purpose and Scope: Building an Open Resource

This database aims to addresses the critical information gap by consolidating enset landrace records from 32 source references spanning the period 1991–2017 into a single, standardized, and openly licensed resource. The current draft represents an initial effort to systematically organize landrace names, geographic distributions, trait information, and source documentation into an accessible format.

Primary objectives of this database include:

- **Research acceleration**: Providing consolidated baseline data on landrace diversity, nomenclature, and geographic distribution to support genetic, agronomic, and ethnobotanical studies.
- **Conservation prioritization**: Enabling identification of rare, threatened, or undersampled landraces and geographic regions requiring urgent collection efforts
- **Breeding support**: Linking landrace names to trait information (disease resistance, drought tolerance, food quality) to facilitate parent selection and improvement programs.
- **Knowledge preservation**: Documenting traditional nomenclature and associated indigenous knowledge before it is lost to cultural erosion.
- **Comparative analysis**: Facilitating cross-regional and cross-ethnic studies of landrace diversity, naming patterns, and trait distributions.
- **Policy guidance**: Informing conservation strategies, agricultural development programs, and food security initiatives with empirical data.

Current Status and Future Development

This database is currently in draft form. The present version organizes landrace records and their associated traits extracted from published literature, grey literature (theses, reports), and institutional survey data. We acknowledge that this compilation is incomplete and may contain errors in transcription, standardization, or interpretation.

A comprehensive analysis of landrace diversity patterns, trait distributions, geographic clustering, and nomenclature harmonization is currently underway. The findings will be presented in a forthcoming peer-reviewed manuscript that will provide detailed methodology, validation procedures, and analytical results.

We actively welcome contributions from the research community. If you have unpublished landrace data, corrections to existing records, additional trait information, or suggestions for database improvement, please contact the corresponding author. Collaborative efforts will strengthen this resource and maximize its value for enset research and conservation.

Database Summary

Metric	Value
Total records	64,747
Unique curated landrace names	1,724
Administrative zones covered	25
Linguistic groups	27
Survey years span	1991–2017
Number of source references	32

Data Fields

The database contains 24 columns organized into the following categories:

Identification

Column	Description
Curated_name	Standardized landrace name
Reported_name	Original name as reported in source

Source Information

Column	Description
Institution	Contributing research institution
References	Literature citation or source
Survey_Year	Year(s) of data collection

Geographic Information

Column	Description
region	Administrative region (SNNPR, Oromia)
zone	Administrative zone
Wereda	District
Kebeles_landrace_recorded	Local administrative units surveyed
Geographic_proximity	Location details
Altitude	Elevation range (m.a.s.l.)
HHD	Number of households surveyed

Linguistic & Cultural Classification

Column	Description
lingrp	Linguistic group of farming community
lanfam	Language family (Omotic, Cushitic, Semitic)
famsys	Farming system classification

Agronomic Traits

Column	Description
Gender	Plant sex (Male/Female)
Food_value	Food uses (Kocho, Bulla, Amicho)
Xvm_tolerant	Xanthomonas wilt resistance
Medicinal	Medicinal uses
Drouhgt_stress	Drought tolerance

Morphological Descriptors

Column	Description
Pseudostem color	Color of pseudostem
midrib color	Leaf midrib color
colour petiole	Petiole color

Unsorted Administrative Regions or Zones

	Unique Landraces	Studies Reporting	Total Records
Wolaita	319	12	521
Gurage	383	11	478
Hadiya	319	9	464
Sidama	231	9	423
Kembata-Tembaro	314	8	421
Dawro	207	6	317
Gamo Goffa	230	7	276
Yem SPW	223	4	275
Kaffa	104	4	148
Sheka	98	3	129
Hadiya and Kembata-Tembaro	96	1	102
South Omo	78	2	78
Gurage and Silte	59	1	61
Sidama and Gedeo	52	1	55
Gedeo	35	4	40
W/SW Shewa	36	2	36
Oromo	35	3	35
Jimma	29	1	29
East Shewa	20	1	21
NA	14	1	15
Silte	3	2	3
Gedeo, Gurage	1	1	1
Keffa	1	1	1
Wolaita, Gamo Goffa	1	1	1
Yem	1	1	1

Language Families

Family	Unique Landraces	Studies Reporting	Total Records
Omotic	1,154	19	1,701
Cushitic	997	18	1,631
Semitic	383	12	473
Semetic	68	1	69
NA	14	1	15

Farming Systems

	Unique Landraces	Studies Reporting	Total Records
Ensetcost (Enset-dominant mixed)	1,331	21	2,185
Ensetdom (Enset-dominated)	664	16	993
Ensetminor (Enset as minor crop)	191	5	273
Ensetsec (Enset secondary)	71	2	73
NA	14	1	15
Cushitic	1	1	1

Contributing Institutions

Institution	Records
SARI (Southern Agricultural Research Institute)	1,932
Hawassa University	264
Addis Ababa University (AAU)	187
Jimma University	93

Survey Years

Year	Records
1991	171
1995	94
1996	155
1998-2000	166
1999	98
2002	142
2003	34
2006	41
2008	303
2008-2009	274
2009	68
2010	105
2011	37
2012	142
2013	627
2014	151
2015	126
2016	65
2017	131

References (Publications, dissertations and other reports)

Reference	Records
SARI (Unpublished)	633
Zerihun et al. 2016	572
Zerihun, 2010	274
Kefale and Sandford, 1991	171
Admasu, 2002	166
Woldemariam, 2008	152
Almaz, 2002	142
Yem TPC, 2015	131
Zippel and Kefale, 1995	119
Melese, 2012	119
CCRP-SARI 2013	118
Sabora, 2010	105
Bizuayehu, 2002	98
Ambachew et al 2018	93
SARI (unpublished)	88
Asnaketch, 1997	82
Fikre and Alemar, 2017	81
Temesgen, 2008	78
Gizachew, 2000	77
Spring 1996	73
Feta, 2008	73
Belachew et al. 2017	68
Hailu, 2009	68
Tesfaye and Zerihun et al. 2018	65
Chekora et al. 2017	55
Nida, 1996	52
Yemane and Fassil, 2006	41
Adanech Jarso, 2017	38
Eyasu, 2003	34
CCRP-SARI, 2013	24
Hunduma et al., 2015	20
Habtamu and Zemede 2011	17

Usage Notes

1. **Multiple records per landrace**: The database contains multiple entries for many landraces, reflecting documentation from different surveys, sources, and time periods. Users should aggregate by Curated_name when counting unique landraces.

2. **Data completeness**: Not all fields are populated for every record. Trait information (disease resistance, drought tolerance, food value) is available for a subset of entries.

3. **Name standardization**: The Curated_name column provides standardized names to reconcile spelling variations in Reported_name.

4. **Citation**: When using this database, please cite: > Muzemil, S., Olango, T.M., Yematawa, Z., et al. (2023). Enset (Ensete ventricosum) landrace database (Version 0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.8287638

Interested to Contribute?

We welcome contributions from researchers, breeders, and conservationists working on enset. Please contact the corresponding authors, to discuss data submission, collaboration opportunities, or suggestions for database improvement.