Review

Received: 13 December 2022 Revised: 25 March 2023 Accepted article published: 24 May 2023 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/jsfa.12710

Connecting data for consumer preferences, food quality, and breeding in support of market-oriented breeding of root, tuber, and banana crops

Elizabeth Arnaud,[a*] Naama Menda,[b*] Thierry Tran,[c,d,e*] Amos Asiimwe,[c,d*] Michael Kanaabi,[f] Karima Meghar,[c,d] Lora Forsythe,[g] Robert Kawuki,[f] Bryan Ellebrock,[b] Ismail Siraj Kayondo,[h] Afolabi Agbona,[h] Xiaofei Zhang,[e] Thiago Mendes,[i] Marie-Angélique Laporte,[a] Mariam Nakitto,[i] Reuben Tendo Ssali,[i] Asrat Asfaw,[h] Brigitte Uwimana,[h] Chukwudi E. Ogbete,[j] Godwill Makunde,[i] Isabelle Maraval,[c,d] Lukas A. Mueller,[b] Alexandre Bouniol,[c,k,d] Eglantine Fauvelle[c,d] and Dominique Dufour[c,d]

Abstract

The 5-year project 'Breeding roots, tubers and banana products for end user preferences' (RTBfoods) focused on collecting consumers' preferences on 12 food products to guide breeding programmes. It involved multidisciplinary teams from Africa, Latin America, and Europe. Diverse data types were generated on preferred qualities of users (farmers, family and entrepreneurial processors, traders or retailers, and consumers). Country-based target product profiles were produced with a comprehensive market analysis, disaggregating gender's role and preferences, providing prioritised lists of traits for the development of new plant varieties. We describe the approach taken to create, in the roots, tubers, and banana breeding databases, a centralised and meaningful open access to sensory information on food products and genotypes. Biochemical, instrumental textural,

* *Correspondence to: E Arnaud, Digital Solutions Team, Digital Inclusion Lever, Bioversity International, Montpellier Office, Montpellier, France. E-mail: e.arnaud@cgiar.org; or N Menda, Boyce Thompson Institute (BTI), Ithaca, NY, USA. E-mail: nm249@cornell.edu; T Tran or A Asiimwe, CIRAD, UMR QualiSud, Montpellier, France. E-mail: thierry.tran@cirad.fr (Tran); E-mail: amos4kasigwa@gmail.com (Asiimwe)*

a *Digital Solutions Team, Digital Inclusion Lever, Bioversity International, Montpellier Office, Montpellier, France*

b *Boyce Thompson Institute (BTI), Ithaca, NY, USA*

c *CIRAD, UMR QualiSud, Montpellier, France*

d *University of Montpellier, Avignon Université, CIRAD, Institut Agro, IRD, Université de La Réunion, Montpellier, France*

e *International Centre for Tropical Agriculture (CIAT), Cali, Colombia*

f *National Crops Resources Research Institute (NaCRRI), Kampala, Uganda*

g *Natural Resources Institute (NRI), Faculty of Engineering & Science, Livelihoods and Institutions Department, University of Greenwich, London, UK*

h *International Institute of Tropical Agriculture (IITA), Ibadan, Uganda*

i *International Potato Center (CIP), Nairobi, Uganda*

j *National Root Crops Research Institute (NRCRI), Umudike, Nigeria*

k *CIRAD, UMR QualiSud, Cotonou, Bénin*

© 2023 The Authors. *Journal of The Science of Food and Agriculture* published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

and sensory analysis data are then directly connected to the specific plant record while user survey data, bearing personal information, were analysed, anonymised, and uploaded in a repository. Names and descriptions of food quality traits were added into the Crop Ontology for labelling data in the databases, along with the various methods of measurement used by the project. The development and application of standard operating procedures, data templates, and adapted trait ontologies improved the data quality and its format, enabling the linking of these to the plant material studied when uploaded in the breeding databases or in repositories. Some modifications to the database model were necessary to accommodate the food sensory traits and sensory panel trials.
© 2023 The Authors. *Journal of The Science of Food and Agriculture* published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

Supporting information may be found in the online version of this article.

Keywords: crop breeding data; consumer preference; sensory information; food quality; roots; tubers; bananas

ACRONYMS

CO	Crop Ontology
GFPP	gendered food product profiles
NIRS	near-infrared spectroscopy
RTB	root, tuber, banana
SOP	standard operating procedure
TPP	target product profile
TD	trait dictionary

INTRODUCTION

Crop breeders' objectives currently focus on agronomic performance of varieties and end-user needs. To secure the adoption of new varieties, a comprehensive market analysis is conducted by multidisciplinary teams to include the food product qualities preferred by the user segments positioned along the value chain,[1] as well as the gender, socioeconomic and cultural drivers for these preferences.[2] A user segment is defined by certain commonalities such as location, role, cultural and socio-economic characteristics. The food quality preferences of the target user segment are analysed to define priority traits summarised in a breeding target product profile (TPP), a data-driven research output guiding the development of new plant varieties. It includes input from growers, agronomists, plant pathologists, end-users, trained panellists, and extension specialists. Results of gender analysis of the preferences for food quality characteristics are compiled in a gendered food product profile (GFPP) that complements the TPP.[3]

This paper describes steps taken to essentially integrate in the crop breeding databases, the sensory information collected through gender-inclusive consumers surveys, sensory panels, biochemical, and instrumental textural characteristics using near-infrared spectroscopy (NIRS) analyses. We included the lessons learned from the 5-year project 'Breeding roots, tubers and banana (RTB) products for end user preferences' (RTBfoods) that involved a network of multidisciplinary teams from international and national partners geographically spread across several African (seven), Latin American (one) and European (two) countries.

The RTBfoods project, the first if its kind for RTB crop breeding, aimed at standardising the data collection for 12 RTB food products and connecting socio-economic survey results and post-harvest and sensory data with agronomic and genetic data, targeting the preferences of diverse user segments (farmers, family and entrepreneurial processors, traders or retailers, and consumers) (Table 1; food products' definitions are in Supporting Information Table S1). We evaluated the preferred qualities of the food products sequentially as follows: (i) recording through surveys, with users freely rating their preferences, (ii) assessment in food sensory tests using hedonic scales or with multi-location participatory processing trials, (iii) measurement by trained sensory panels, (iv) then analysis by biochemical laboratories, and (v) prediction by machines or genetic markers. This favoured over time a

Table 1. List of food products per crop and study countries in the RTBfoods project

RTB crop	Food product	Primary study country	Spillover countries
Cassava	Boiled cassava	Uganda, Colombia	Benin
	Gari-Eba	Nigeria	Cameroon, Benin, Côte d'Ivoire
	Attiéké		
	Fufu	Nigeria	Cameroon
Cooking banana	Boiled plantain	Cameroon	Côte d'Ivoire, Nigeria
	Matooke (East African highland banana)	Uganda	
	Fried plantain aloco	Nigeria	Cameroon
Sweet potato	Boiled sweet potato	Uganda	
	Fried sweet potato	Uganda	Ghana, Côte d'Ivoire
Yam	Boiled yam	Benin	Côte d'Ivoire, Nigeria
	Pounded yam	Nigeria	Côte d'Ivoire, Benin
Potato	Boiled potato	Uganda	Kenya

[a] See Supporting Information Table S1 for the description of the food products.
Abbreviation: RTB: root, tuber, banana.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

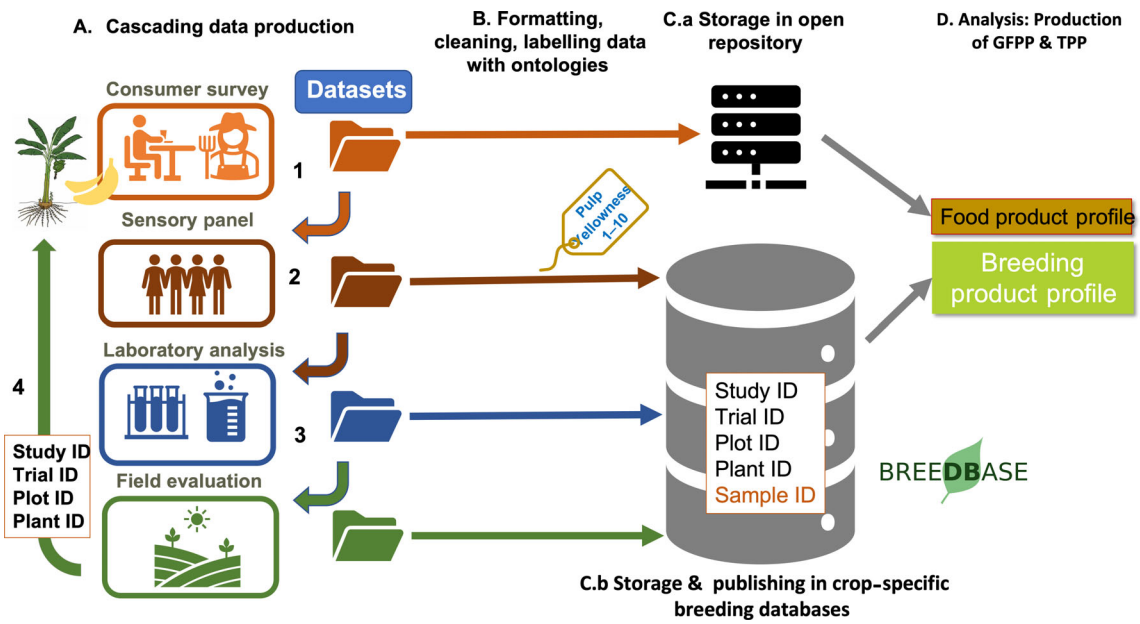

Figure 1. Cascading data flow in RTBfoods project from consumers' preference surveys to trained sensory panels, to biochemical analysis, and selection in the field of the new varieties. GFPP, gendered food product profile; TPP, target product profile.

cascading production of datasets that secured the integration of consumers' prioritised preferences into the subsequent analysis in laboratories or in the field (Fig. 1).

The envisioned data value chain between data generator and end-user has four major objectives: (i) collection of trait data generated by food and social scientists on selected genotypes, for integration into the TPP; (ii) standardisation and publication of datasets in open breeding databases and repositories; (iii) uptake of the results by breeding programmes and food scientists; and (iv) contribution to a final impact on the new variety adoption by targeted user segments.

To develop the evidence supporting the TPP, the resulting project's datasets had to be properly formatted, stored, and interconnected into existing crop-specific breeding databases (CassavaBase, MusaBase, PotatoBase, SweetPotatoBase, YamBase), all using a single crop agnostic open data and ontology-driven model called Breedbase (https://breedbase.org).[4] A consensus on the definitions of traits and variable names is a prerequisite to allow the proper description data to be shared by different breeding programmes. Therefore, Breedbase integrates the Crop Ontology (CO) that unequivocally defines traits and variables for the RTB crops (cropontology.org).[4-7]

MATERIALS AND METHODS

Standard operating procedures for data collection on sensory traits

Evaluation methods for the consumer-preferred traits within the RTB breeding populations were not standardised, hindering the comparison of the results between the different countries and breeding programmes. Using common standard operating procedures (SOPs) to measure users' preferences and sensory traits secures that the measured food qualities are related to the variety acceptability. Therefore, the research teams developed product-specific SOPs along with global guidelines to conduct consumer surveys, participatory processing trials, sensory panels, and biophysical analysis. SOPs developed for sensory panels include lexicons, which are validated lists of food qualities and variables to be measured by trained panellists[8] and compared for consumer adoption of varieties. Lexicons are country specific to align with consumers' preferences that vary according to the cultural and socio-economic context. To secure comparability of the biochemical and NIRS data, the project purchased similar equipment for the laboratories (Colombia, Guadeloupe, and sub-Saharan Africa).

The lists of traits and variables included in the SOPs for sensory panels, food processing trials, and texture analysis were fully described using the CO trait dictionaries for each food product to guide data integration in the breeding databases and storage of the measurement values. All SOPs developed by the project are available with their citation and unique digital object identifier (DOIs) in the Agritrop repository (https://agritrop.cirad.fr/; see DOIs in Supporting Information Tables S2 and S5). Each document has a version number to enable any required revision.

The CO and the trait dictionaries

An 'ontology' is a human and machine-readable collection of classified and uniquely identified concepts with textual definitions and semantic inter-relationships. It thus provides a domain-specific controlled vocabulary and supports the description of data in databases.

In the CO, a 'trait name' is an entity (plant or food product part) combined with a quality (e.g. root colour) and is observed or measured by using a specific method with a defined scale or unit. The combination of the trait name with the method and the scale or unit is called a 'variable'.[6,7] The CO provides a standard framework for composing the name of a trait and a variable that will precisely describe the value of a trait measurement stored in the database. The standard framework is important for the comparison of measurements (Fig. 2). The CO compiles the descriptions and variables of agronomic, morphological, physiological, stress response, and quality traits, enabling the digital capture and aggregation of crop trait data.[5-7]

© 2023 The Authors.
Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

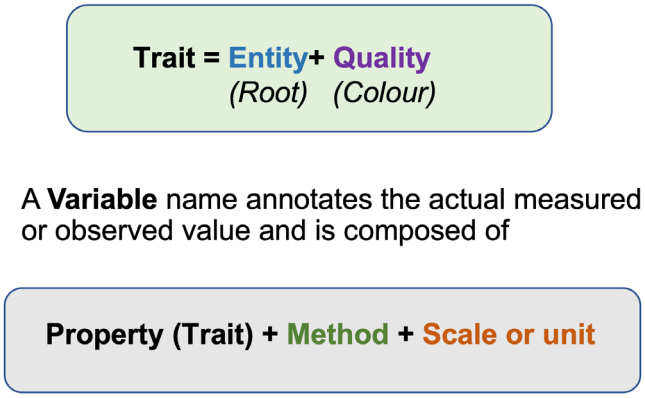

Figure 2. Nomenclature for naming a trait and a variable in the Crop Ontology. Source: Guidelines, version 2.1. Adapted from Pietragalla *et al.*[7]

The CO provides a list of variables with recommended methods and units of measurement to support the harmonisation of the collected data, and thus their comparability. The sensory lexicons vary with the country, so the CO, by recording different evaluation scales for a given trait quality, reflects this variation and enables the description of all possible values in the database. Additionally, new technologies, such as machine learning or high-throughput methods, generate data on crop traits with new methods and scales (prediction, classification, etc.). Therefore, one trait can be linked to several methods of measurement and scales, from visual observation to instrumental measurement methods. This way, the value of a measurement is accurately described with information enabling it to be directly integrated in a comparative analysis or to decide whether to apply some conversion rules before or discard it.

RTB food quality traits were added in the CO with their definitions, methods of measurement, and scales to ensure a harmonised description of data resulting from the assessments or measurements of food products' properties. The CO Trait Dictionary (TD) Template version 5.2 and the user guidelines version 2.1[7] guided the extraction of traits from the lexicons of the SOPs for trained sensory panels.

Publishing food quality data in the open Breedbase model

A decade ago, RTB crop breeding programmes adopted the online Breedbase model, creating one database per crop, namely CassavaBase, MusaBase, PotatoBase, SweetpotatoBase, YamBase. Breedbase uses the ontology-driven Chado schema for storing breeding data[9,10] and records data on field or laboratory crop evaluation trials, trait data on genotypes used, parental selection, crossing design, and experimental design. It enables digital data collection and analyses, and includes decision-making tools.[4]

All phenotypic data, including sensory data, must correspond to a predefined trial record, which includes minimal metadata such as the breeding programme name, location, date, and accessions. The phenotypes measured can be assigned to plots in the field or plots in controlled environments, such as a greenhouse or *in vitro* laboratory, or to individual plants, plant parts, and tissues. Traceable plant collections and genotypes are called accessions. In the case of RTB crops, plants within a plot are typically clonal material. This system allows grouping phenotypic observations by accession across multiple trials. To identify the plant samples and their origin, the indication of the plant accession identifier (ID), along with the trial and plot IDs, are provided by breeders to the food processors and analysis laboratories. This information must remain complete and unaltered in the data templates used by food scientists, and the data collected on the subsamples analysed can be attached to the original plant and trials material in Breedbase. The use of the SOPs and templates improves the quality of data to be uploaded in the breeding databases or in repositories.

Survey data: gendered food product profiles

Gender and social context play important roles in influencing the demand for quality characteristics that follow the gender division of labour in crop cultivation and food processing.[3] Therefore, separate trait scoring tables were produced for men and women, by region, and other important factors according to the context, to identify different preferences in characteristics, and captured this criterion to develop gendered food product profiles (GFPPs). Data on user preferences of food product qualities were collected following a five-step interdisciplinary and participatory methodology. The aim was to integrate GFPPs into the TPP and guide priority setting for mid- to high-throughput protocols, such as biochemical and NIRS.

Step 1 is state of knowledge, which involved a literature review and key informant interviews to establish what was known about the product and the gaps in knowledge, within a specific geographic context, in relation to food science, gender issues and markets, and demand segmentation and location. A list of important characteristics from crop to product were collected with their citations and reasons for their importance among different user groups, and a thematic summary of the findings was provided.

In step 2, food mapping exercises were conducted to identify the various uses of the crop by user segments (e.g. producers, processors, consumers, and local retailers) with effort on capturing gender differences and similarities in preferences, based on the different roles men and women play in the value chain. A list of important characteristics and their description from crop to product was collected, along with the gender-disaggregated scoring, and their prioritisation (by citation or participatory ranking exercise) was developed, along with the names of preferred and non-preferred varieties, and a broader thematic summary of findings.

In step 3, teams conducted a participatory processing diagnosis with experienced processors. Both preferred and non-preferred varieties were included to provide a wide range of technological and physico-chemical characteristics. Processors provided feedback on the varieties before processing, during each processing step, and after processing to identify quality characteristics of the crop and product. Processing parameters were measured at each step. New quality characteristics from this step were added to the GFPP.

In step 4, using the results of step 1 on the user segmentation and demand location of the product, consumer testing was conducted with as many as 300 consumers, with 150 interviews in rural areas and 150 in urban areas. The consumer sampling included an equal number of women and men, from different locations of the city to increase the representation of various socioeconomic and ethnic groups. Groups in rural areas were similar to those visited in step 2.[3] The objective was to better understand the consumer demand and obtain a sensory mapping of the overall liking of each product that could be related to the most- and least-liked characteristics used by each consumer to describe the product. At this stage, new quality characteristics and their prioritisation were eventually added to the GFPP.

 © 2023 The Authors.
Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2. Excerpt of the lexicon for boiled cassava

Type	Attribute	Definition	How to measure?	Scale
Appearance	Yellow	Colour of cassava root surface varies from light yellow to bright yellow	When you receive a cassava root sample, observe the surface and evaluate the intensity of the colour, homogeneity, translucency, and surface smoothness	0: Non-yellow 10: Bright yellow
	White	Colour of cassava root surface varies from cream to bright white		0: Cream 10: Bright white
	Homogeneity of colour	Uniformity of cassava root surface colour		0: Heterogeneous 10: Homogeneous
	Surface smoothness	Absence of roughness, lumps, holes, fibre lines, and ridges along the cassava root		0: Very rough 10: Very smooth
Texture in mouth	Hardness	Mechanical textural attribute relating to the force required to achieve a given deformation, penetration, or breakage of a product	Put a part of boiled sample into your mouth, evaluate during the first bite (between molars) how hard the sample is	0: Soft 5: Firm 10: Hard
	Moisture	Perception of moisture content of food by the tactile receptors in the mouth, and also in relation to the lubricating properties of the product	Put a part of boiled sample into the mouth, chew and evaluate the quantity of water within the sample	0: Dry 10: Moist
	Smoothness	Geometrical textural attribute relating to lack of presence of particles in a product	Put a part of boiled sample into the mouth, chew it and after ten chews evaluate between tongue and palate the number and the size of the particles	0: Lumpy 5: Grainy 10: Smooth

Source: Nuwamanya *et al.*[11]

Then, during step 5, the GFPP is developed and provides a description of a high-quality food product from an evolving list of sensory, processing, and agronomic characteristics, that focuses on a specific region, usually sub-national. Inclusive discussions on the finalised profiles among social scientists, biochemists, food technologists and breeders led to the development of improved selection criteria and methods. DOIs for the guidelines and reports of the 5-step methodology and GFPP are respectively in Supporting Information Tables S2 and S3. Table S4 is an excerpt of the GFPP Template providing definitions of the information captured.

Quality data from trained sensory panel

A sensory descriptive panel is composed of trained assessors who define the sensory attributes that best describe a product being evaluated. Standardisation of a sensory panel through training and use of a lexicon minimises the variation among panellists and produces measurable assessment of the food qualities. The results provide a tool that enables laboratories to evaluate and demonstrate the reliability of the data produced. The lexicon included in an SOP defines specific food attributes corresponding to the preferences identified by the user surveys (GFPPs). It provides a standard template to guide panellists on how to measure the attributes and record answers using categorical scales (Table 2).

Sensory data collected from trained sensory panels should discriminate samples through correct and consistent scoring. Therefore, proper training of the panel must be done using a standardised methodology. The guidelines indicate that a panel must have an equal balance between men and women, as well as good age distribution (from 18 to 60 years). For statistically correct results, a minimum of eight panel members is advisable, but it is highly desirable to have at least ten subjects who are qualified to carry out the test.[12]

Sensory tests with replicated samples allowed checking the repeatability of panellists' scoring skills. Guidelines for curating trained panel sensory data[12,13] recommend that (i) the absolute difference between a panellist's score assigned to sample replicates should not exceed 3 and (ii) each panellist's score for a given sample be compared with the panel mean to check the deviation and identify outliers. Before uploading data in Breedbase, a cleaning exercise was manually performed to eliminate unacceptable data from poorly performing panellists. The process has been extensively used in cleaning descriptive sensory data of RTB food products. This approach ensures reliable data for statistical analysis.

As per the earlier statement, the Breedbase upload process validates the datasets for the presence of information about the plant accession used for the sensory trial, its plots, and the CO ID of each variable in the data spreadsheet column headers.

NIRS data

NIRS was used as a high-throughput phenotyping tool to predict quality traits of food products. Breedbase includes the ability to store and analyse NIRS data thanks to the integration of Waves, an open-source R package, with several cross-validation schemes to assess prediction accuracy.[4,14] RTB breeding programmes used various NIRS devices (Foss XDS, portable handheld SciO, etc.) for data acquisition. However, Breedbase is flexible enough to handle spectral data irrespective of the source.[14] NIRS data were

© 2023 The Authors.
Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3. Template for uploading near-infrared spectroscopy data into Breedbase, developed by Boyce Thompson Institute. The first column is the 'Obersvationunit_name' and subsequent columns store the wavelength value

Observationunit_name	400	402	404	406	408
2019AMDPSerere_813_plant_1_raw_dried11	0.3454392	0.3681009	0.3885303	0.4057705	0.4195259
2019AMDPSerere_968_plant_1_raw_dried11	0.2928936	0.3122137	0.3296063	0.3441868	0.3557591
2019AMDPSerere_1245_plant_1_raw_dried11	0.3078051	0.3282051	0.3464911	0.3617148	0.3736272
2019AMDPSerere_263_plant_1_raw_dried11	0.1898816	0.2010899	0.2109308	0.2188177	0.2246294
2019AMDPSerere_460_plant_1_raw_dried11	0.2673013	0.2840712	0.2989252	0.3110988	0.3204741
2019AMDPSerere_213_plant_1_raw_dried11	0.2724023	0.2910206	0.3076751	0.3214798	0.3322425

reformatted using the Breedbase NIRS template (Table 3). The column 'observationunit_name' must contain an already recorded 'Sample Name' of each sample generated when creating the trial record in the database. The columns' headers of the spectral information always contain the wavelength names (numbers). Storing NIRS data in Breedbase using ontology IDs would require defining a variable for each of the hundreds of wavelengths (typical range 400 or 750 nm up to 2500 nm) that make up an NIRS spectrum. Instead, a separate section was developed specifically to accommodate NIRS data and avoid the need to define variables in the ontology. It is therefore critical to use the 'Sample Name' in the NIRS template to link and track back to the respective trial record. The uploading process requires recording metadata before storing it in the database (see Supporting Information Process S2).

Creating the required ontology of food product qualities

The inclusion of food quality traits into the breeding pipeline required a comprehensive list of sensory and biophysical traits, unequivocally named and properly described. The following contains a brief description of the steps taken to produce this list.

Survey data on user preferences

In the GFPP template, to help with the interpretation, descriptive information of the preferred characteristics as expressed by the user segment is provided by the 'Indicator of Characteristic' (Table 4 and Supporting Information Table S4, column B). This information was used just to test the mapping of preferred traits to the CO to label the GFPP.[16]

Table 4. Excerpt of the table 'Indicators of high-quality characteristics of steamed-mashed matooke'

Characteristic	Indicator(s)
Soft texture	On eating – feel in the mouth, smooth on fingers, easy to cut *But what level of softness is desired? Perhaps physicochemical analyses in the laboratory can measure this
Good smell	Inhaling under the nose/by smelling; smells like it has been cooked in banana leaves
Elastic/starchy	Touch; feels elastic (*kunyururuka*)
Homogeneous texture	Visual assessment, feel during eating, no particles (*obukote*), *kutakuterera*, no hard parts after mashing

Source: Marimo *et al.*[15]

Sensory traits extraction from lexicon

The sensory traits were compiled by food scientists into food product and country-specific lexicons based on the preferred food properties revealed by the user surveys (Supporting Information Table S6, for boiled cassava).

Figure 3 illustrates the extraction of traits from the lexicons using CO TD Template version 5.2. The template, designed for measurement of crop variety traits, was adapted to include the food product name (process type and product name) as the main study object and accommodate the format to sensory traits. The format of the first food product TD for steamed cooking banana (matooke) and boiled cassava was validated by food scientists. The adequacy of the TD was then checked with sample datasets from the food science laboratory.

Extraction of food-processing techniques from reports

Qualities of food products are impacted by the processing techniques, so the crop variety must be bred to perform well under the cooking processes. Lack of validated variables and of standard post-harvest processing methods can hinder the interpretation and comparability of processing trials from different food science laboratories. For example, boiling time can be measured either from when the cooking starts or when the water starts to boil and will depend on the size of the root pieces. Therefore, standardisation of the boiling protocol is essential to generate comparable results. Participatory processing trials were conducted by food scientists with local processors (family or small entrepreneurs) applying guidelines to conduct the trial with adequate measurements.[17] To support the description of collected data and its storage in Breedbase, a specific ontology was developed. Some processing techniques with definitions are proposed in the Food Ontology (FOODON, https://foodon.org/), but there is no description of the methods and scales of measurements that make up the variables, and some techniques are missing.

The CO TD Template was again modified with the input of food scientists to accommodate the description of the processing operating units and the format of the variables measured. Traits and variable names were extracted from the reports of the participatory trial surveys. The resulting template captures the name of the processing step in a 'Processing Unit' column (e.g. peeling, washing, boiling), the method used to measure or observe, and the measurement unit (minutes, hours, etc.). The 'Entity' on which 'Qualities' are observed is the piece of food product on which the processing unit is done: 'Raw root' for the 'Peeling' unit and 'Peeled raw root' for the 'Slicing' unit (Table 5 and Supporting Information Table S7). The processing techniques were mapped to FOODON concepts when possible.

 © 2023 The Authors.
Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Food Product	Trait name	Trait class	Trait description
Boiled Cassava	Boiled root surface yellowness	Appearance	Colour of cassava root surface varies from light yellow to bright yellow
Boiled Cassava	Boiled root surface color homogeneity	Appearance	Uniformity of cassava root surface colour

Type	Attributes	Definition	How to measure?	Scale
Appearence	Yellow	Color of the surface of the sample from light yellow to bright yellow	When you receive the sample, observe the surface and evaluate the intensity of the color and its homogeneity	0: very light yellow 10:bright yellow
	Homogeneity of colour	Uniformity of color of the surface of the sample		0 : heterogeneous 10 : homogeneous

Method name	Method description
Appearance method	When you receive a cassava root sample, observe the surface and evaluate the intensity of the yellow colour.
Appearance method	When you receive a cassava root sample, observe the surface and evaluate the colour homogeneity

Scale name	Category 1	Category 2
Yellowness scale 0-10	0=Non-yellow	10=Bright yellow
Colour homogeneity 0-10	0= Heterogeneous	10=Homogeneous

Figure 3. Extraction of sensory traits from the lexicon on boiled cassava to the Crop Ontology template version 5.2.

Processors provided their feedback on the 'good' and 'bad' qualities, which were recorded with their citation frequency. This information was extracted in a separate sheet of the TD, for future labelling with the ontology.

Integration of the food product TDs into the CO and Breedbase

Publishing the resulting TDs in the CO website enables online visualisation of the traits, term search, and download, and applies the Creative Commons license: CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0/). New subclasses were created in the CO 'Quality Trait' class to accommodate the attributes' categorisation of the lexicon: 'Appearance', 'Aroma', 'Taste', 'Texture in hand', 'Texture in mouth'. Figure 4 shows the online display of the variable for measuring the 'Eba sourness in the mouth' trait with its metadata.

Direct composition of new variable names in Breedbase while uploading data, also called 'post-composing', provides necessary flexibility if a project-specific variable was missing in the ontology. The direct variable composition feature is useful for repetitive measurements requiring indication of different points in time, after various treatments, different growing cycles, and/or characterising subsamples (e.g. different parts of the plant). New variables can then be submitted to the CO curation team via a term request form.

RESULTS

Linking all data in Breedbase to the plant material of origin

The survey data and GFPPs

Thirteen GFPPs generated to date by the multidisciplinary teams of the RTBfoods project are available in a Dataverse open repository (Supporting Information Table S3). The qualitative surveys conducted with the end-users could not be processed the same way as the quantitative biophysical data from the field or laboratory evaluations. First, capturing in the ontology the context in which preferences are expressed is complex. Second, a standard process had to be developed to convert unstructured data from surveys into quantifiable data. And finally, strong ethical constraints on data publishing in open access are imposed to anonymise the personal identifiable information before sharing. Therefore, a data workflow was designed from the initial survey forms to the production of the GFPP to be linked to the TPP in Breedbase (Fig. 5).

Adaptation of Breedbase to sensory, biophysical, and food-processing datasets

A total of 163 sensory traits were measured through trained panels, extracted in TDs, and uploaded into the crop-specific breeding databases and into the CO website (Table 6). The

Table 5. Excerpt of the trait dictionary for boiled cassava root processing techniques

Parameter	Entity	Property	Trait class	Definition	Method	Variable label
Raw root total weight to peel	Raw root	Total weight to peel	Quality	Total weight of the raw roots to be peeled	Measurement	Total raw roots weight in kg
Peeled raw root washing time	Peeled raw root	Washing time	Quality	Time taken to wash the peeled root	Measurement	Washing time
Raw root steaming time	Steamed cassava root	Steaming time	Quality	Time for steaming cassava roots	Measurement	Steaming time

Note: Concepts extracted from Hamba *et al.*[18]

© 2023 The Authors.
Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

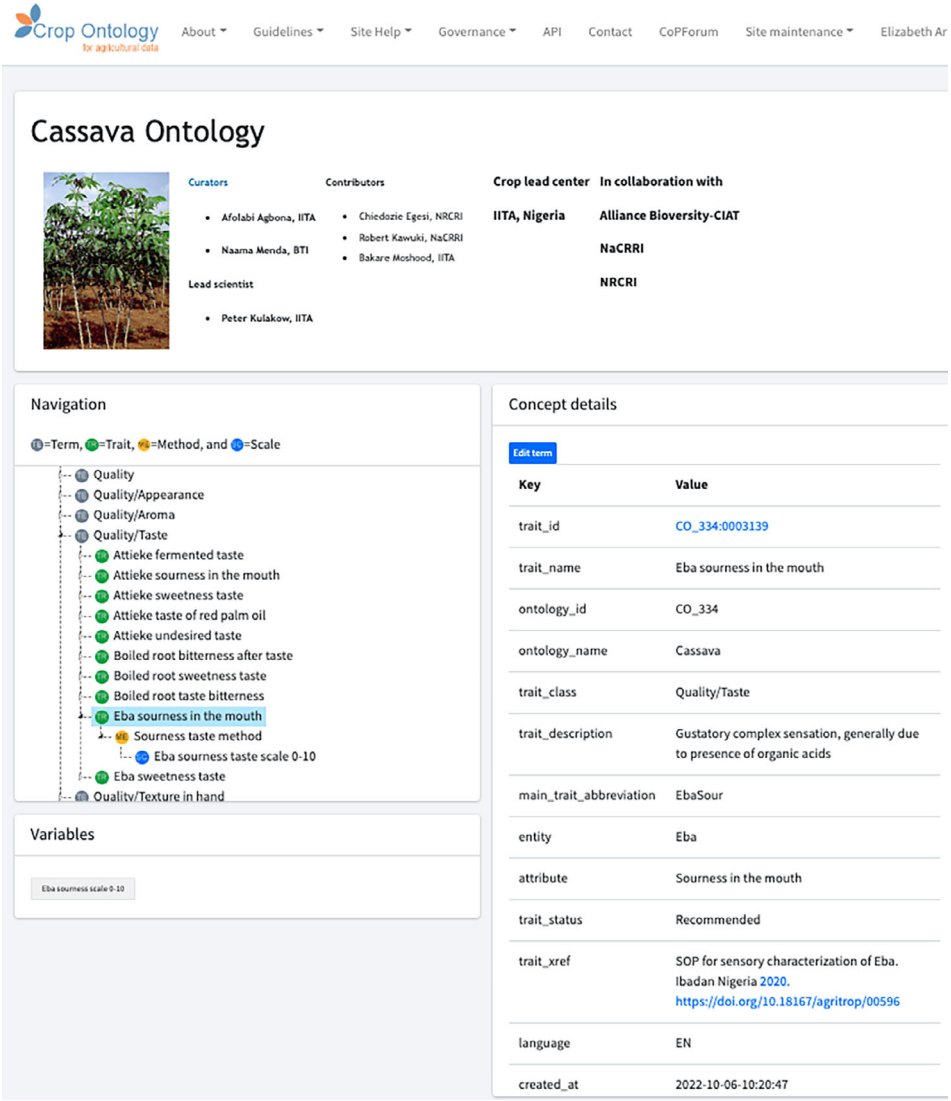

Figure 4. Display in Crop Ontology's 'Quality/Taste' class of the variable name for measuring 'Eba sourness in the mouth' trait using a 1–10 scale, with its descriptive information.

curated data files from the sensory panels were provided back to scientists with a quality report so they could correct or discard some data.

The starting Breedbase upload template is for the creation of the trial design record, followed by the phenotyping values template (xls format). Two non-exclusive approaches were developed for uploading sensory panel data:

(1) Creation of a separate record called 'sensory panel trial' and directly uploading the raw data in this trial (Fig. 6). Breedbase then automatically calculates the means and standard deviations from the raw data and provides intensity scale graphs (Fig. 7). This mainly interests food scientists who may need to check the quality of the raw data, including the repeatability and accuracy of individual panellists.

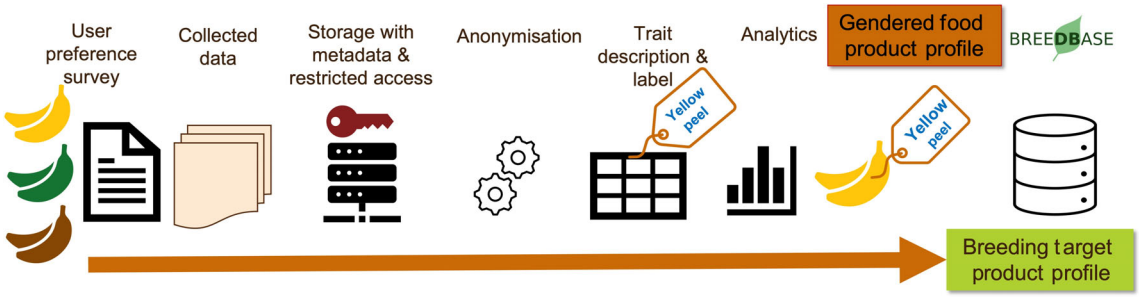

Figure 5. User survey data flow in the RTBfoods project.

© 2023 The Authors.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 6. Number of sensory traits evaluated per crop by sensory panels and extracted into the Crop Ontology (CO)

Food product	Number of sensory traits in CO	CO URI
Attieke (Cote d'Ivoire)	24	https://cropontology.org/term/CO_334:ROOT
Boiled cassava (Benin and Uganda)	33	https://cropontology.org/term/CO_334:ROOT
Boiled plantain (Cameroon)	6	https://cropontology.org/term/CO_325:ROOT
Boiled potato (Uganda)	23	https://cropontology.org/term/CO_330:ROOT
Boiled and steamed sweet potato (Uganda)	26	https://cropontology.org/term/CO_331:ROOT
Boiled yam (Benin and Nigeria)	14	https://cropontology.org/term/CO_343:ROOT
Fufu (Nigeria)	6	https://cropontology.org/term/CO_334:ROOT
Gari/eba (Imo, Osun, Benue States, Nigeria, Cameroun)	11	https://cropontology.org/term/CO_334:ROOT
Matooke (Uganda)	14	https://cropontology.org/term/CO_325:ROOT
Pounded yam	6	https://cropontology.org/term/CO_343:ROOT
Total	**163**	

(2) Directly uploading the means and standard deviations of the sensory panel attached to the field trial record corresponding to the material studied, which makes the data easily accessible and usable by breeders.

The collected user segments' preferences were dissected by teams in charge of the biophysical measurements into measurable biophysical traits, usable for varietal screening. A total of 78 biophysical traits were measured for following products through laboratory-based protocols (e.g. water absorption, texture analysis, dry matter, etc.): boiled cassava (10), fufu (13), eba (13), pounded yam (13), boiled yam (15), boiled plantain (14). The formatted (xls or csv) biophysical files were uploaded in the field trial summary page corresponding to the accessions studied in Breedbase so they were available beside the agronomic datasets and could be analysed with the Breedbase built-in statistical tools.

The upload template only accepts the means of the biophysical traits. Nevertheless, raw data can be uploaded as additional files in the corresponding field trial summary page. The raw data for eight instrumental texture datasets for boiled cassava were uploaded into CassavaBase by the post-harvest quality teams in Uganda and Colombia. Texture-related traits, such as 'Texture characteristic End Force to Max Force Ratio of boiled cassava roots' with the measurement method 'Measurement: Texture-extrusion method', were extracted with the TD Template for future integration in the cassava ontology (https://cropontology.org/term/CO_334:ROOT).

The TDs describing the food product processing techniques and their variables have not yet been uploaded into the ontology. The teams still need to agree on the display of these concepts in different Cos, since some techniques (e.g. slicing, boiling) are common to all food products. Therefore, to date, participatory processing datasets are not yet uploaded into Breedbase.

NIRS spectral datasets

A total of 21 844 NIRS spectra of fresh grated cassava roots and fresh or boiled yam tubers have been uploaded to date into CassavaBase and YamBase respectively. The number of uploaded spectra by institute, crop, product, and trial is shown in Supporting Information Table S8. As detailed earlier herein, NIRS datasets were uploaded in a dedicated NIRS section of Breedbase; plot names stored with the NIRS data enabling linking each dataset to its respective field trial record. A reverse link to the available NIRS datasets should automatically be inserted in the field trial summary page, so users are aware of their existence.

DISCUSSION

The benefit of using the CO model

In the user preference survey results, the informative value of the 'Indicator of the Characteristic', providing a textual definition of preferred qualities, varied considerably. The final GFPP only includes indicators that were validated with socioeconomists and breeders. A test was performed on the correspondence of the main categories of the Matooke GFPP with the ontology classes and trait mapping (Table 7), but it still needs to be expanded to all GFPPs.

The RTBfoods project has, through an iterative process, generated best practices for developing sensory lexicons to generate comparable data. An attribute in a lexicon must be a single element with a clear definition and a well-defined 'how to measure' method that relies on an agreed categorical scale. The systematic extraction of variables for sensory traits into the format of the CO TD identified some inconsistencies in the first versions of the lexicons, particularly in the way attributes were named or classified, and in the use of different scales (0–10, 0–100). When different categorical scales were used by research teams for the same sensory trait and method, several variable names had to be created to enable proper labelling of the collected data for future comparative analysis. When measurement methods of sensory traits were similar across various food products, the definitions were harmonised by selecting the clearest and better expressed method (in English).

To upload sensory datasets, the original TD Template was modified so the column 'Growth Stage' was replaced by 'Food Product' and a new context of use (Trained sensory panel) was added. In the TD template, a 'Method type' must be allocated. Food scientists recommended 'Measurement' for results of a trained sensory panel and 'Estimation' for hedonic sensory assessment. The variable name, following the nomenclature recommended by the CO, was then created (Table 8). The abbreviated variable name was then used as a column header in the dataset template file.

This adaptation of the CO TD template worked well for connecting traits, methods, and scales, as well as for the definitions. The

© 2023 The Authors.

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

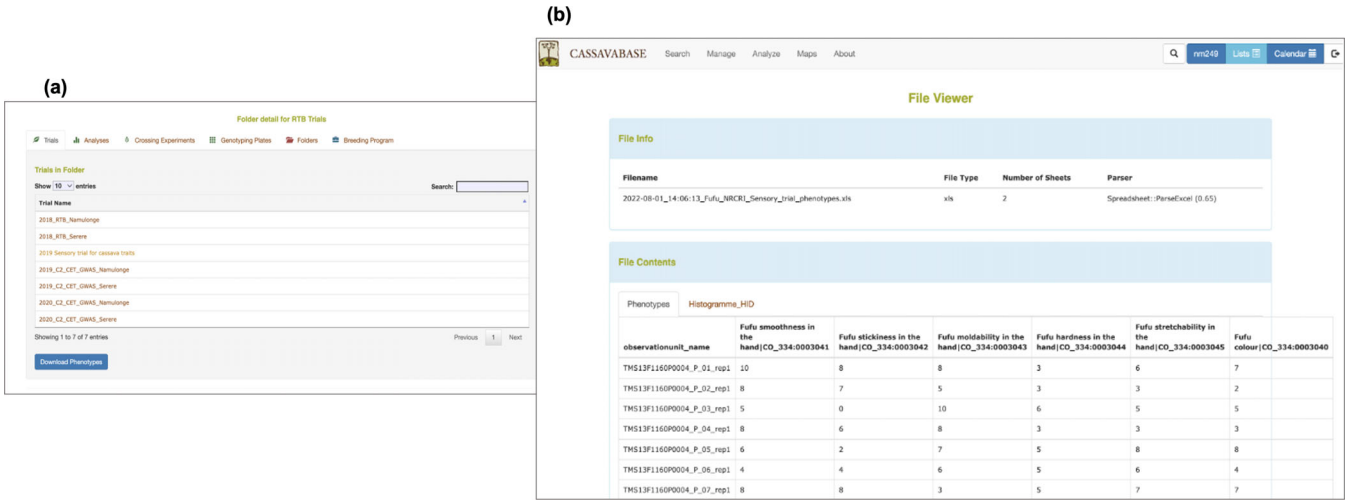

Figure 6. (a) Sensory trial for boiled cassava created in CassavaBase. (b) Visualisation of the uploaded raw sensory data file with the trait name and the Crop Ontology variable identifier in the column headers – accessed 16 November 2022.

formatted sensory traits were uploaded in their respective crop-specific ontology for boiled cassava, attieke, eba, matooke, boiled sweet potato, boiled and pounded yam.

The trait classification in the CO was modified to accommodate the specific categorisation of attributes in the sensory lexicon. Sub-classes, like 'Quality Traits/Aroma' were created in the 'Quality Traits' class.

Sometimes, breeders and food scientists use different terminologies for similar plant parts or qualities. The ontology allows these terms to be added as synonyms, creating the necessary correspondence, and avoiding losing a domain-specific vocabulary. For example, breeders use in their database the term 'Boiled storage root', whereas food scientists use 'Boiled potato'. Additionally, many terms in the ontologies have acronyms that are commonly used by breeders, but those may not be self-explanatory. Those acronyms are designated as synonyms to the full trait name. For example, the term 'Dry matter content by table top NIRS' in percentage has a synonym of 'dmNIRS' (https://cassavabase.org/cvterm/77879/view).

Impact of the RTBfoods project on cassava and yam breeding

The percentage of positive change in number of quality traits recorded into CassavaBase and YamBase depicts the increase in number of quality traits assayed by a partner (Table 9). A comparison of the number of quality traits assayed by breeding programmes before and after the initiation of the RTBfoods project (2018) using data uploaded to CassavaBase (www.cassavabase.org) and YamBase (www.yambase.org) was performed.

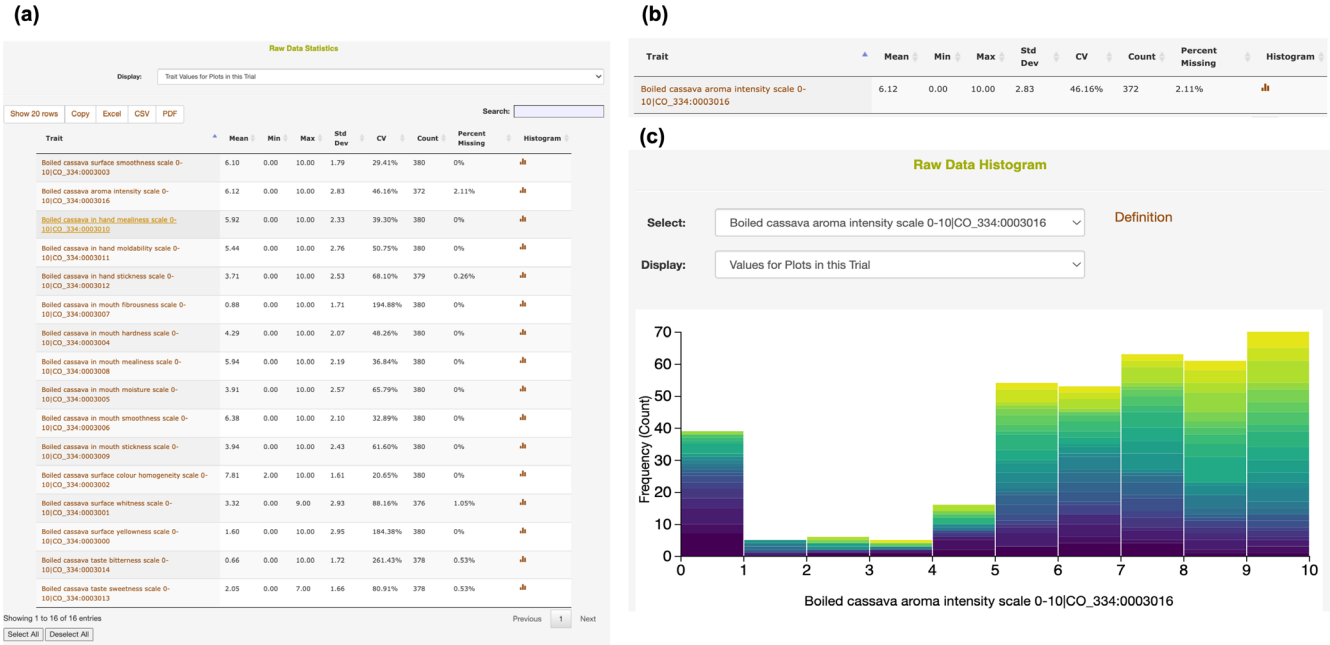

Figure 7. (a) Mean values and standard deviations of traits measured by sensory panel. (b) Blow up of the aroma intensity mean and standard deviation. (c) Breedbase graph on boiled cassava aroma intensity scale 0–10, CO_334:0003016.

 © 2023 The Authors.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 7. Example of a preference 'Matooke Smooth feel' recorded in the matooke food product profile template mapped to a trait and variable in Crop Ontology

Columns of food product profile template				Crop Ontology		
		Definition	Example	Mapping	Definition	Annotation
A	Category or family	Class to which the quality belongs	Sensory quality	Trait class	Agronomic, biotic stress, abiotic stress, quality, etc.	Quality- Sensory in Mouth
B	Characteristic	Quality preferred or disliked as expressed by the informant	Matooke Smooth feel	Variable name and label	Variable name composed by Trait name, Method name and Scale or Unit, all abbreviated	Matooke_Texture homogeneity_assessment by mouth_cat1-5-CO_325:0002033
C	Indicator of the characteristic	Complementary information indicating the way the informant estimates the quality	Soft feel in mouth during eating	Trait Method	*Matooke Smoothness* Method used to measure or assess the trait and provide a value	CO_325:0002046 Texture in mouth
		n/a		Scale/unit	Categorical scale with scoring classes or unit	10: Smooth

Source: Arnaud *et al.*, 2022.[6]

Table 8. Example of variable names for matooke and the label that will be used in the database

Variable name	Variable label
Matooke yellowness measurement scale 0–19	MatYell_Meas_1to10
Matooke homogeneity of colour measurement scale 0–10	MatHomCol_Meas_1to10

The development of specific data templates and the creation of ontologies for food product qualities have increased the quality and clarity of datasets generated. Publishing the RTBfoods project data in the crop-specific Breedbase and open repositories guarantees the desired open access, while the use of common ontology and consistent plant material IDs supports data interoperability.

The cascading data generation concept of the project, where each research module had to deliver validated data to the next module, has slowed down the progress of uploading data in breeding databases but has secured the provision of quality data.

Recording the IDs of the plant accessions studied, the field trial name, and plot number in all subsequent data files, beside the subsampling coding, is a crucial practice that needs to be promoted from the inception of similar projects to enable connecting all datasets back to the plant material agronomic data. Applying the SOPs and using the recommended ontologies from the beginning reduce the data reformatting effort.

The priority datasets to be uploaded were the ones generated from plant accessions provided by breeders. Material used as checks can be recorded when designing the trial. The material collected in fresh markets with its local name is not currently added in Breedbase, and data will be stored in the project open repository. Decisions can be made to create specific market trials in the Breedbase where this type of material and related data can be recorded. However, a variety can be attributed to diverse names in markets, which confuses the connection to the original variety name and complicates data interpretation in the context of a breeding programme.

The difference between users and scientists regarding 'trait naming' adds another challenge to be addressed by (i) the addition of an indicator of characteristics clearly describing the preference so a breeder can securely interpret to which phenotype it relates, and (ii) the mapping of biophysical and chemical traits to sensory traits.

During their participatory consumer trials, breeding teams should use the SOPs, lexicons, and templates developed by food scientists to support the quality of data collected and its interoperability with food science data.

A product profile includes traits with their threshold values defined during the project and required to meet the breeding targets. For example, 'Biofortified cassava for enhanced nutrition' is one of the four main target product profiles. Traits include β-carotene (>15 ppm fresh weight), dry matter content (0.3 g kg^{-1} of fresh weight) and fresh root yield (25 t ha^{-1}). This list of target traits enables the selection of genotypes showing values closest to the thresholds. As an additional feature, Breedbase should enable the scoring of clones against threshold values to highlight which clones have the most promising profiles for adoption. To this end, all TPPs need to be integrated into the crop-specific Breedbase and connected to GFPPs stored in the project open repository.

© 2023 The Authors.
Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 9. Number of recorded agronomic and quality traits evaluated over years in breeding programs and percentage increase in number of quality traits assayed from 2018 to 2022

Institution	Crop	Duration/years	No. field trials	No. traits phenotyped	No. quality traits	Change in no. quality traits (2018–2022) (%)
IITA	Cassava	1974–2017	1975	195	25	87.5
		2018–2022	153	342	34	
NaCRRI	Cassava	2012–2017	71	63	9	14.3
		2018–2022	53	86	17	
NRCRI	Cassava	2006–2017	47	95	9	26.5
		2018–2021	78	77	7	
CIAT	Cassava	1979–2017	1225	51	9	47.1
		2018–2022	164	117	63	
IITA	Yam	2001–2017	61	132	14	33.3
		2018–2022	154	145	21	
NRCRI	Yam	2016–2017	10	120	10	16.7
		2018–2020	8	121	12	
CNRA	Yam	2016–2017	5	53	3	40

Note: Analysis performed by M. Kanaabi, IITA. Source: CassavaBase and YamBase, accessed in September 2022.
Abbreviation: IITA, International Institute of Tropical Agriculture; NaCRRI, National Crops Resources Research Institute; NRCRI, National Root Crops Research Institute; CIAT, International Centre for Tropical Agriculture; CNRA, Centre National de Recherche Agronomique, Côte d'Ivoire.

LESSONS LEARNED FROM CONNECTING DATA ON FOOD QUALITY AND BREEDING

- The RTBfoods project has increased the number of quality traits assayed by partner breeding programmes, and using the data templates and ontologies for food product traits improved the quality of the datasets generated.
- Gender, culture, and socioeconomic factors have an important role when collecting food product profile data. It is important to collect inclusive input reflecting the local agriculture and food system preferences.
- Developing SOPs and data templates that integrate the protocols with a defined list of standards and contextual variables to be measured in multi-country and multi-partner projects is a prerequisite to proper data standardisation.
- Each SOP for sensory panels is specific to a food product and a country. Traits, variables, and scales vary as the consumers' preferences change with the cultural and socioeconomic context. In the CO, one trait can have several methods and scales to reflect the difference.
- The CO uses standardised vocabularies for sensory traits included in the SOPs. It facilitates collecting and describing comparable data for consumer preferences, food quality, and breeding. By integrating defined and structured vocabularies of the multidisciplinary teams (e.g. food scientists, socioeconomists, breeders), the ontology supports the interpretation of results across domains.
- Recording identifiers of plant accessions, field trial names, and plot numbers in all subsequent data files is crucial for connecting all datasets back to the plant material agronomic data.

ACKNOWLEDGEMENTS

We are grateful to the CGIAR Roots, Tubers and Banana Research Programme funded by CGIAR Fund council and the grant opportunity INV-008567 (formerly OPP1178942): Breeding RTB Products for End User Preferences (RTBfoods), to the French Agricultural Research Centre for International Development (CIRAD), Montpellier, France, by the Bill & Melinda Gates Foundation (BMGF): https://rtbfoods.cirad.fr.

The editorial comments by Hernán Ceballos and Dominique Dufour, as well as the final proofreading of the manuscript by Clair Hershey, greatly improved the quality of this paper.

AUTHOR CONTRIBUTION

Elizabeth Arnaud, Naama Menda: writing, review, and Editing. Thierry Tran, Amos Asiimwe, Michael Kanaabi, Karima Meghar, Lora Forsythe, Ismail Siraj Kayondo, Robert Kawuki, Afola Agbona, Xiaofei Zhang, Bryan Ellebrock, Isabelle Maraval: writing (supporting). Lora Forsythe, Alexandre Bouniol, Marie-Angélique Laporte, Isabelle Maraval, Thierry Tran, Xiaofei Zhang, Afola Agbona, Godwill Makunde, Karima Meghar, Reuben Tendo Ssali, Chukwudi E. Ogebete, Miriam Nakitto, Robert Kawuki: investigation, methodology, validation. Amos Asiimwe, Elizabeth Arnaud, Marie-Angélique Laporte, Naama Menda, Michael Kanaabi, Karima Meghar, Afola Agbona: data curation, standard, methodology. Naama Menda, Bryan Ellebrock, Lukas A. Mueller: software, methodology, visualization. Lora Forsythe, Karima Meghar, Alexandre Bouniol, Isabelle Maraval, Thierry Tran, Xiaofei Zhang, Reuben Tendo Ssali, Chukwudi E. Ogebete, Godwill Makunde, Ismail Siraj Kayondo, Asrat Asfaw, Thiago Mendes, Brigitte Uwimana, Miriam Nakitto: formal analysis. Lora Forsythe, Alexandre Bouniol, Elizabeth Arnaud, Xiaofei Zhang, Lukas A. Mueller, Asrat Asfaw: supervision (lead). Dominique Dufour: funding acquisition (lead), conceptualization (lead), project administration (lead), supervision (supporting). Eglantine Fauvelle: funding acquisition (supporting), conceptualization (supporting), project administration (supporting), supervision (supporting).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Agritrop at https://agritrop.cirad.fr/.

© 2023 The Authors.

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Dufour D, Hershey C, Hamaker BR and Lorenzen J, Integrating end-user preferences into breeding programmes for roots, tubers and bananas. *Int J Food Sci Technol* **56**:1071–1075 (2021). https://doi.org/10.1111/ijfs.14911.

2 Thiele G, Dufour D, Vernier P, Mwanga ROM, Parker ML, Schulte Geldermann E *et al*., A review of varietal change in roots, tubers and bananas: consumer preferences and other drivers of adoption and implications for breeding. *Int J Food Sci Technol* **56**:1076–1092 (2021). https://doi.org/10.1111/ijfs.14684.

3 Forsythe L, Tufan H, Bouniol A, Kleih U and Fliedel G, An interdisciplinary and participatory methodology to improve user acceptability of root, tuber and banana varieties. *Int J Food Sci Technol* **56**:1115–1123 (2021). https://doi.org/10.1111/ijfs.14680.

4 Morales N, Ogbonna AC, Ellerbrock BJ, Bauchet BJ *et al*., Breedbase: a digital ecosystem for modern plant breeding. *G3 (Bethesda)* **12**: jkac078 (2022). https://doi.org/10.1093/g3journal/jkac078.

5 Shrestha R, Matteis L, Skofic M, Portugal A, McLaren G, Hyman G *et al*., Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the Crop Ontology developed by the crop communities of practice. *Front Physiol* **3**:326 (2012). https://doi.org/10.3389/fphys.2012.00326.

6 Arnaud E, Laporte M-A, Kim S, Aubert C, Leonelli S, Miro B *et al*., The ontologies community of practice: a CGIAR initiative for big data in agrifood systems. *Patterns* **1**:100105 (2020). https://doi.org/10.1016/j.patter.2020.100105.

7 Pietragalla J, Valette L, Shrestha R, Laporte M-A, Hazekamp T and Arnaud E, *Guidelines for Creating Crop-Specific Ontology to Annotate Phenotypic Data: Version 2.1*. Alliance Bioversity International–CIAT, Rome, Italy, p. 38 (2022) https://hdl.handle.net/10568/110906.

8 Suwonsichon S, The importance of sensory lexicons for research and development of food products. *Foods* **8**:27 (2019). https://doi.org/10.3390/foods8010027.

9 Mungall CJ, Emmert DB and The FlyBase Consortium, A Chado case study: an ontology-based modular schema for representing genome-associated biological information. *Bioinformatics* **23**:i337–i346 (2007). https://doi.org/10.1093/bioinformatics/btm189.

10 Jung S, Menda N, Redmond S, Buels RM, Friesen M, Bendana Y *et al*., The chado natural diversity module: a new generic database schema for large-scale phenotyping and genotyping data. *Database* **2011**:bar051 (2011). https://doi.org/10.1093/database/bar051.

11 Nuwamanya E, Iragaba P, Kawuki R, Nanyonjo AR, Kanaabi M, Khakasa E *et al*., *Sensory characterization of boiled cassava. Biophysical Characterization of quality traits, WP2*. RTBfoods, Kampala, Uganda, p. 18 (2021). https://doi.org/10.18167/agritrop/00599.

12 Maraval I, Forestier-Chiron N and Bugaud C, *RTBfoods manual – part 1 – sensory analysis. Training a panel in sensory analysis and implementing descriptive tests, tutorial: how to process data in sensory analysis*. CIRAD–RTBfoods, Montpellier, France, p. 54 (2018). https://doi.org/10.18167/agritrop/00573.

13 Bugaud C, Maraval I and Forestier-Chiron N, *RTBfoods manual – part 2 – tutorial. Monitoring panel performance and cleaning data from descriptive sensory panels for statistical analysis. Biophysical characterization of quality traits, WP2*. RTBfoods, Montpellier, France, p. 14 (2021). https://doi.org/10.18167/agritrop/00582.

14 Hershberger J, Morales N, Simoes CC, Ellerbrock B, Bauchet G, Mueller LA *et al*., Making waves in Breedbase: an integrated spectral data storage and analysis pipeline for plant breeding programs. *Plant Phenome J* **4**:e20012 (2021). https://doi.org/10.1002/ppj2.20012.

15 Marimo P, Kenneth K, Khamila S, Tinyiro SE, Bouniol A, Ndagire L *et al*., *Participatory processing diagnosis of matooke in Uganda. Understanding the drivers of trait preferences and the development of multi-user RTB product profiles, WP1, Step 3*. RTBfoods, Kampala, Uganda, p. 38 (2022). https://doi.org/10.18167/agritrop/00615.

16 Arnaud E, Khamila S, Kibooga C, Ndagire LY and Marimo P, *Guidelines for Mapping the Preferences in Gender Sensitive Product Profiles to Crop Ontology and Creating a Consumer Segment Ontology. Version 1.0*. Alliance Bioversity–CIAT, Rome, Italy (2022) https://hdl.handle.net/10568/118232.

17 Fliedel G, Bouniol A, Kleih U, Tufan H and Forsythe L, *RTBfoods Step 3: Participatory Processing Diagnosis and Quality Characteristics*. CIRAD–RTBfoods, Montpellier, France (2018). https://doi.org/10.18167/agritrop/00570.

18 Hamba S, Nanyonjo AR, Kanaabi M, Kawuki RS and Bouniol A, Participatory processing diagnosis of boiled cassava in Uganda, in *Understanding the Drivers of Trait Preferences and the Development of Multi-User RTB Product Profiles, WP1*. RTBfoods, Kampala, Uganda, p. 26 (2021). https://doi.org/10.18167/agritrop/00625.

View publication stats

10970010, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jsfa.12710 by Nigeria Hinari NPL, Wiley Online Library on [08/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License