A systematic map of cassava farming practices and their agricultural and environmental impacts using new ontologies: Agri-ontologies 1.0

Received: 17 November 2022 | Accepted: 18 April 2023

DOI: 10.1002/2688-8319.12249

Ecological Solutions and Evidence | BRITISH ECOLOGICAL SOCIETY

REGISTERED REPORT STAGE 2

A systematic map of cassava farming practices and their agricultural and environmental impacts using new ontologies: Agri-ontologies 1.0

Amelia S. C. Hood[1,2] | Gorm E. Shackelford[1,3] | Alec P. Christie[1,4] | Hope O. Usieta[5] | Philip A. Martin[1,3,6] | William J. Sutherland[1,3]

[1]Conservation Science Group, Department of Zoology, University of Cambridge, David Attenborough Building, Cambridge, UK

[2]Centre for Agri-Environmental Research, Sustainable Land Management, School of Agriculture, Policy and Development, University of Reading, Reading, UK

[3]Biosecurity Research Initiative at St Catharine's (BioRISC), St Catharine's College, University of Cambridge, Cambridge, UK

[4]Downing College, University of Cambridge, Cambridge, UK

[5]Leventis Foundation (Nigeria), Abuja, Nigeria

[6]Basque Centre for Climate Change (BC3), Leioa, Spain

Correspondence
Amelia S. C. Hood
Email: a.s.hood@reading.ac.uk

Funding information
A G Leventis Foundation; Arcadia Fund; The David and Claudia Harding Foundation

Handling Editor: Marc Cadotte

Abstract

1. Cassava is consumed by 800 million people and is a staple crop in Africa. Its production may increase under climate change due to its high drought tolerance. We produced a systematic map of scientific studies about cassava farming practices, with the aim of identifying knowledge gaps and clusters. Our secondary aim was to develop a classification system for [1] farming interventions and [2] agricultural, economic and environmental outcomes. Standardised classification systems facilitate data reuse, including for evidence synthesis, and promote research efficiency.
2. Following our published protocol, we searched eight publication databases using the search string 'cassava OR mandioca OR manihot OR manioc OR yuca' in December 2017. We screened 36,580 records and included publications that measured the impact of cassava farming practices on agricultural, economic or environmental outcomes, including yield, soil, water, wildlife and labour. We classified the resultant 1599 publications by interventions, outcomes, location, study year and study design. We assessed coding consistency using Kappa scores.
3. We found regional knowledge clusters (Nigeria, Columbia and Brazil accounted for 45.5% of country occurrences) and gaps (e.g. the Democratic Republic of Congo). There were knowledge clusters for interventions testing cultivar type, fertiliser use and diversifying crop rotations and outcomes related to crop production (e.g. yield/biomass). We found knowledge gaps for environmental interventions and outcomes (e.g. 5% of studies measured pollutants or wildlife). In terms of study design, reporting standards were poor (e.g. 24% of studies did not report start dates), average study duration was 2 years, and average publication delays were 4 years. The Kappa scores indicated that we successfully developed

Amelia S. C. Hood and Gorm E. Shackelford contributed equally to this work.

consistent ontologies (named Agri-ontologies 1.0). The map and ontologies are available online: https://www.metadataset.com/.

4. This systematic map of cassava farming practices can direct researchers and funders to knowledge gaps that need addressing, and reviewers to knowledge clusters for synthesis. Better research practices should be promoted within cassava research, as poor reporting standards, short study durations and long publication delays result in an ineffective research environment. This systematic map provides an evidence base for cassava production and the ontologies (Agri-ontologies 1.0) can be applied to other systems to facilitate more efficient and effective synthesis.

KEYWORDS
agricultural taxonomy, cassava *Manihot esculenta*, evidence-based management, interactive evidence map, reporting standards, standardised classification system, subject-wide evidence synthesis, sustainable agriculture, systematic map, terminological ontology agriculture

1 | INTRODUCTION

1.1 | A systematic map of cassava farming practices and their impacts

Cassava (*Manihot esculenta*) is a major global crop, with an increasing annual production that totalled >302 million tonnes in 2020 (FAO, 2019). Cassava (also known as mandioca, manioc and yuca) is a perennial plant primarily grown for its starchy tubers, with the majority used for human consumption—either eaten directly or converted into flour (e.g. tapioca, farinha or garri)—and the remainder used as animal feed or biofuel (FAO, 2000). It is primarily grown via subsistence farming in the tropics and subtropics, with Nigeria and Brazil leading production (see Figure 5 for a global yield and production map). Approximately 800 million people consume it, and in Africa, it is the second most important crop after maize, with >40% of the population using it as a staple crop (FAO, 2019).

The importance of cassava as a food crop may increase further as it has the potential to help to mitigate the negative impacts of climate change on food security (Jarvis et al., 2012). For example, climate change is increasing the frequency and severity of droughts (Malhi et al., 2021), with drought-affected areas projected to increase from 15% to 44% by 2100, and the yield of major crops predicted to decrease by >50% by 2050 (Li et al., 2009). Sub-Saharan Africa, which produces >50% of cassava globally (FAO, 2019), is the most vulnerable region to drought (Li et al., 2009). Furthermore, its population is projected to grow by 1 billion people by 2050, which will further increase the pressure on food systems (United Nations Department of Economic and Social Affairs Population Division, 2019). It is imperative that there is investment in crops, such as cassava, that are more resilient to drought and climate change (Jarvis et al., 2012). Cassava's resilience extends beyond drought tolerance too; it can be grown in soils with low fertility and harvested at any time, hence its nickname: 'the drought, war and famine crop' (Burns et al., 2010). Therefore, cassava has the potential to help to mitigate the impacts of climate change on food security.

Given its current and potential future role, it is important that cassava management is evidence based (Sutherland et al., 2004). This will help to reduce yield gaps, which are currently large (yields of the top-producing cassava countries [18 t/ha] are <50% of potential yields [44 t/ha]; Figure 5). Furthermore, some existing farming practices contribute to environmental degradation, such as depleting soil fertility (Reynolds et al., 2015). If cassava production is to be scaled up, evidence-based farming practices will be crucial for doing this as sustainably as possible whilst maximising yields (Sutherland et al., 2004). Cassava management is varied (e.g. intercropping, fertiliser and rotation practices are diverse), and there are thousands of studies testing different management practices. Several international authorities on cassava farming, such as the International Center for Tropical Agriculture (CIAT), the International Institute of Tropical Agriculture (IITA) and the Food and Agriculture Organisation (FAO) of the United Nations, have made evidence-based agricultural extension manuals and work closely with practitioners to disseminate findings (e.g. FAO, 2013; Howeler & Aye, 2017; International Institute of Tropical Agriculture (IITA), 2017). However, there is no systematic map of cassava research (Shackelford et al., 2018).

Systematic mapping collates, describes and catalogues evidence (e.g. publications or reports) for a specific topic (James et al., 2016). Systematic maps enable researchers, funders, practitioners or policy makers to identify knowledge clusters or gaps and then direct resources accordingly. For example, researchers can conduct syntheses on knowledge clusters and funders can fund field studies on knowledge gaps (James et al., 2016). Systematic maps also direct users to relevant studies, which can improve the efficiency and completeness of evidence gathering (James et al., 2016). Here, we create a global systematic map of cassava farming practices to provide an evidence base for sustainable cassava production.

1.2 | Ontologies for farming interventions and outcomes

The FAIR guiding principles for scientific data management were described in 2016 (Wilkinson et al., 2016) to address the scientific community's growing concern that poor data management was inhibiting data reuse and contributing to a 'significant crisis of reproducibility' (Baker, 2016). These principles outlined four characteristics for data management—that datasets should be 'Findable', 'Accessible', 'Interoperable' and 'Reusable' (Wilkinson et al., 2016). Significant efforts have been made to achieve this, including the development of dataset repositories and citation metrics (Cousijn et al., 2018; Hood & Sutherland, 2021; Konkiel, 2020), journals and funders mandating public data archiving (Mislan et al., 2016) and increased uptake of reporting guidelines or standardised terms for study metadata (Equator Network, 2022; Stevens et al., 2014). Standardised terms can improve reporting standards and address the 'Interoperable' and 'Reusable' aspects of FAIR data management by making it easier to combine datasets (Hopewell et al., 2012; Plint et al., 2006). This facilitates cross-comparison between fields, promotes research efficiency, and enables data reuse for transfer learning, generating synthetic data and synthesis (Todman et al., 2023). Standardised terms can be ordered into multi-level ontologies to show relationships between terms and enable data to be grouped at different levels (Jonquet et al., 2018; Thessen et al., 2015). There are several ontologies relevant to agricultural and ecological research (Jonquet et al., 2018), including ICASA Data Standards for data from agricultural field experiments (Hunt et al., 2006; White et al., 2013), DarwinCore for taxonomic data (Wieczorek et al., 2012) and AgrO for agronomic data (Aubert et al., 2017).

Here, we use this systematic map of cassava to develop hierarchical ontologies for [1] farming interventions and [2] agricultural, economic and environmental outcomes. These ontologies differ from existing ontologies as they have a broad focus and are intended to be comprehensive while also being general enough to apply to other agricultural systems. They are also designed to be facilitate quantitative syntheses, such as meta-analyses, by using 'treatment-treatment' and 'treatment-control' categories, and they apply the use of 'not elsewhere classified (n.e.c.)' categories to identify categories that should be added, i.e. if the 'n.e.c.' category has many publications coded in it. In particular, these ontologies were designed to develop case studies for 'dynamic meta-analysis', which is a method of meta-analysis that enables synthesis at multiple levels via an online interactive interface (Shackelford et al., 2020), with the ultimate aim of increasing the relevance and recency of evidence (e.g. via 'living systematic reviews') and facilitating subject-wide evidence syntheses (Martin et al., 2022; Sutherland & Wordley, 2018). For example, standardised ontologies such as this could eventually contribute to a global map of scientific literature on agriculture.

1.3 | Objectives of the review

The primary objective of this review was to produce a systematic map of scientific studies about cassava farming practices, with the aim of identifying knowledge gaps and clusters. Our secondary objective was to develop a classification system for [1] farming interventions and [2] agricultural, economic and environmental outcomes. These objectives reflect those in our published protocol (Shackelford et al., 2018), with the exception of primary objective five, which relates to the study designs used and was added during the review. Study design is important for interpreting the quality of the mapped studies and the evidence base as a whole. For example, synthesists that have used the following map to identify knowledge clusters will be able to assess whether their chosen topic has a robust study design (e.g. long study duration and randomised treatments) and may choose to select a different topic if not (Christie et al., 2020). We also expanded objective 1; in the protocol, it was limited to cassava fields, and during the review, we expanded it to whole cassava plants (i.e. we included whole plants in glasshouses).

1.3.1 | Primary objectives

The primary objectives were to answer the following questions:

1. Which studies have measured the impacts of cassava farming interventions (e.g. intercropping/tilling) on agricultural, economic and environmental outcomes (e.g. yield/soil)?
2. Which interventions and outcomes have been studied?
3. Which countries have been studied and when?
4. What is the distribution and abundance of studies between different interventions, outcomes, countries and years? In other words, where are the knowledge gaps or knowledge clusters in this map?
5. Which study designs have been used (metrics include reporting completeness, experimental designs, study durations and publication delays)?

1.3.2 | Secondary objectives

The secondary objective was to develop and test—using this systematic map of cassava as an example—ontologies of [1] farming interventions (e.g. intercropping or tilling) and [2] agricultural, economic and environmental outcomes (e.g. yield or soil). These ontologies are hierarchical classifications of farming practices and agri-environmental outcomes and are intended to be general enough to be reused for other crops.

2 | MATERIALS AND METHODS

2.1 | The systematic map

The following systematic map was produced according to the corresponding published systematic map protocol (Shackelford et al., 2018). This protocol was based on the Collaboration for Environmental

Evidence (CEE) guidelines for systematic mapping (James et al., 2016) and the RepOrting standards for Systematic Evidence Synthesis (ROSES) reporting checklist (Supporting Information 1; Haddaway et al., 2018). The published protocol was developed in consultation with our three project partners: the African Cassava Agronomy Initiative (ACAI; Pypers, 2017), the Conservation Evidence group at the University of Cambridge (Sutherland et al., 2021), and the Leventis Foundation, which runs agricultural schools in West Africa (The Leventis Foundation Nigeria [Internet], n.d.). The Leventis Foundation funded this map and uses it to identify knowledge gaps to study in field trials. The Conservation Evidence group uses the map to identify knowledge clusters and conduct syntheses.

2.1.1 | Searches

We searched for studies using the search string 'cassava OR mandioca OR manihot OR manioc OR yuca'. These are the common synonyms for cassava, but not its products (e.g. tapioca), since we were interested in pre-harvest management of cassava (Gade et al., 2002; Hillocks et al., 2002). As the aim of our systematic map was to map the impacts of multiple interventions on multiple outcomes (i.e. it was an open-framed question; James et al., 2016), we did not include any search terms for interventions (e.g. tilling) or outcomes (e.g. yield).

In late December 2017, we searched for publications from several sources:

1. Publication databases. We searched two generic publication databases (*Scopus* and *Web of Science Core Collection*), two agricultural databases (*AGRICOLA* and *AGRIS*) and one conservation database (*Conservation Evidence*). The generic and conservation databases mostly include peer-reviewed publications, whereas the agricultural databases also include grey literature.
2. Internet searches. We screened the first 500 results sorted by 'relevance' from one search engine (*Google Scholar*; Haddaway et al., 2015).
3. Specialist searches. We screened the first 500 results from two grey literature repositories (*the Document Repository of the Food and Agriculture Organisation* [*FAO*] and *the repository of the Consultative Group on International Agricultural Research* [*CGIAR*] *Centres, in CGSpace*: *A Repository of Agricultural Research Outputs*) sorted by 'relevance'. CGSpace includes publications from many organisations that study cassava: CIAT, IITA, Bioversity International, and the CGIAR Research Program on Roots, Tubers and Bananas.

2.1.2 | Article screening

The records returned from these searches were screened for duplicates in Endnote X8. We screened the remaining records at [1] title and abstract stage and [2] full text stage. The screening process was procedurally independent as the authors that screened the records had not authored articles that would be relevant to include in the map.

2.1.3 | Screening at title/abstract stage

At title and abstract stage, the study was excluded if the title or abstract:

1. was not in English
2. the study had no relevant 'PICO/PECO' components (where 'P' = populations/subjects, 'I/E' = interventions/exposures, 'C' = comparators and 'O' = outcomes; James et al., 2016). The interventions/exposures and outcomes refer to the ontologies, which are described in '2.2 Agri-ontologies 1.0' below, and the populations/subjects and comparators are defined directly below. If the title and abstract were too ambiguous to know whether the study had any or all of the relevant PICO/PECO components (e.g. 'Annual Report 2010'), we included them.

2.1.4 | Populations/subjects

We included studies that measured the impact of cassava farming practices on agricultural or environmental outcomes. We included studies that looked at other crops in rotation with cassava or intercropped with cassava. We included studies on whole cassava plants in pots, but not modelling studies, studies of cassava pests or natural enemies where cassava was absent or studies on cassava cuttings or germplasm (e.g. laboratory studies of cassava in vitro). We included studies of cassava fields before or after cassava was planted (e.g. prepping the field or cover cropping post-harvest) but not of post-harvest management of cassava (e.g. storing). For studies comparing cassava with other habitat types, studies about the conversion of a habitat to cassava agriculture were included, but studies that compared existing cassava fields with other habitat types without cassava were excluded. We did not restrict the timing of studies and the scope was global.

2.1.5 | Comparators

We included studies that compared a population with an intervention with a population without an intervention. This included controlled studies (e.g. comparison of a plot with treatment applied [treatment] and a plot without treatment applied [control]), correlated studies (e.g. comparison of a plot with pre-existing application of a treatment [treatment] and a plot without treatment [control]) and before-and-after studies (e.g. comparison of a plot with itself before [control] and after treatment is applied [treatment]). We included 'treatment-treatment' comparisons (e.g. a comparison of a plot with one fertiliser type applied [control] vs. another plot with a different fertiliser type applied [treatment]). We excluded studies that had no comparators.

2.1.6 | Screening at full-text stage

The publications were included past the title and abstract screening stage, then assessed at the full-text stage and were excluded if

1. A full text of the study could not be found. We searched the publication title in Google, Google Scholar and Web of Science with and without quotations.
2. A full text of the study was behind a paywall according to the journal subscriptions at the University of Cambridge in 2017, where the study was conducted.
3. The full text was not available in English.
4. The study was secondary literature (i.e. studies that cite other studies as the source of numerical results).
5. The study had no relevant PICO components. We only considered outcomes that were reported via numerical results in figures or tables.
6. The study was a duplicate. In cases where there was a larger book with smaller chapters, we included each as smaller chapters, unless the entire book was written by the same authors.
7. There was no full text (e.g. the study was an audio file or poster).
8. The study was retracted.

Note that we assessed these exclusion criteria sequentially, so—for example—if a publication was excluded because the full text was not available in English (criteria 3), we did not assess whether it was secondary literature (criteria 4).

2.1.7 | Consistency in article screening

All publications were screened by one person (GES), but a second person (ASCH) screened 10% of publications, selected at random, to check the consistency of the screening process ($n = 2038$ at title and abstract stage and $n = 502$ at full-text stage). We discussed any disagreements and clarified the eligibility criteria where necessary. We calculated Kappa scores as a measure of agreement (Cohen, 1960). Kappa ranges from −1 to 1, with Kappa >0.6 considered as 'moderate' agreement (Cohen, 1960). We used 0.6 as a threshold to indicate agreement.

2.1.8 | Data coding strategy

We coded the interventions (described in 'Agri-ontologies 1.0' below), outcomes (described below), study location(s) (country), study year(s) and experimental design. Experimental design included the following categories, with the option to select multiple categories: before-and-after, blocked (i.e. a controlled study with treatments next to each other spatially), controlled, correlated (site-comparison), paired (i.e. a correlated study with sites that are selected to have similar characteristics), randomised and replicated. We only coded outcomes that were reported via numerical results in figures or tables. These were coded using an online web app with drop-down menus (www.metadataset.com). We did not critically appraise the validity of these studies as the map can be used for multiple methods of evidence synthesis which may have different criteria for critical appraisal. We did not contact the authors for any missing or unclear information. Many interventions that we coded were confounded (e.g. for a study that varied fertiliser application in tandem with varying cultivar, we would code this as two single interventions related to 'fertiliser' and 'cultivar'). We did not code publications with multiple studies into separate entries, but if the same intervention was tested multiple times in different experiments within one publication, we coded this intervention multiple times within the publication. If the same outcome was tested multiple times within one publication, we coded it once. One author (GES) coded the majority of publications (983/1599). Then we assessed consistency in coding using Kappa scores (described in 'Agri-ontologies 1.0' below). Once agreement was reached, the second author (ASCH) coded the remaining publications (616/1599).

2.1.9 | Deviations from the protocol

This systematic map deviated from the protocol (Shackelford et al., 2018) in three respects:

1. We expanded the scope to include coding 'experimental design'.
2. We expanded the scope to include studies on whole cassava plants (e.g. cassava plants in pots, but not leaves in petri dishes), where the original protocol specified that only studies in fields would be included.
3. We did not code outcomes, years or experimental design for the intervention '10.10.10.TT.20. Planting a different variety/cultivar (e.g. a disease-resistant variety)' (i.e. we only coded the intervention). This was due to time constraints. This intervention was the most common (939 intervention occurrences, 24% of intervention occurrences), and it would be difficult to apply a synthesis to this intervention as there are numerous cultivars, most of which are regional. For publications that tested this intervention and others, we coded the outcomes, years and experimental designs of the other interventions.

2.1.10 | Limitations of the method

The following systematic map of cassava is subject to four limitations:

1. We only included publications written in English, which can bias syntheses (Konno et al., 2020). However, in this case, only 1% of full texts were excluded for this reason. We did

not have the resources to work in other languages, and the stakeholders involved were working in countries where English is an official language (Nigeria, Ghana, Kenya and the United Kingdom).
2. We excluded publications that were behind a paywall according to journal subscriptions at the University of Cambridge where the study was done. This was 6% of full texts.
3. We excluded publications that we could not find as full texts, which was 24% of records. In many cases, we suspect this was because the full texts do not exist (e.g. the record refers to an unpublished conference paper or a magazine article without an online archive).
4. The searches were conducted at the end of December 2017, so the map is already out of date. This delay was partly due to setbacks caused by the coronavirus pandemic. Nevertheless, this map forms a solid foundation for future updates.

2.1.11 | Data mapping method and analysis

This systematic map of cassava farming practices is freely available online via an interactive format (an 'evidence atlas') on https://www.metadataset.com/ (registration required) and in spreadsheet form on a research data archive (Hood et al., 2022). Users can select interventions or outcomes at multiple levels and view the relevant publications and their study location (country level). In this paper, we summarise these findings according to this study's objectives outlined above. Data were wrangled, plotted and analysed visually using R version 4.1.1 (R Core Team, 2020) with R studio version 1.4.1717 (RStudio Team, 2019), using the Tidyverse (Wickham, 2017), rworldmap (South, 2011), Simple Features (Pebesma, 2018) and cowplot (Wilke, 2019) packages. We used local regression smoothers (LOESS) to visually interpret some temporal trends. Several plots were amalgamated in *Inkscape* (Inkscape, 2020).

2.2 | Agri-ontologies 1.0—Ontologies for farming interventions and outcomes

The intervention and outcome ontologies were developed based on previous work and in consultation with the same three stakeholders that designed the systematic mapping protocol (Shackelford et al., 2018). The ontologies (Figures 1 and 2, listed in full in Supporting Information 3 and 4) are freely available online via an interactive format at https://www.metadataset.com/. These ontologies were designed to be generic enough to apply to other crops and are particularly intended to be used for syntheses, such as systematic maps, reviews or meta-analyses, though they can be used for other purposes too. Both ontologies used 'not elsewhere classified (n.e.c.)' categories to capture interventions or outcomes that did not fit into existing categories. These categories result in a comprehensive map, as all publications are included, and they facilitate updating the ontologies, as having many publications in 'n.e.c.' categories can indicate that a new category should be created if a common

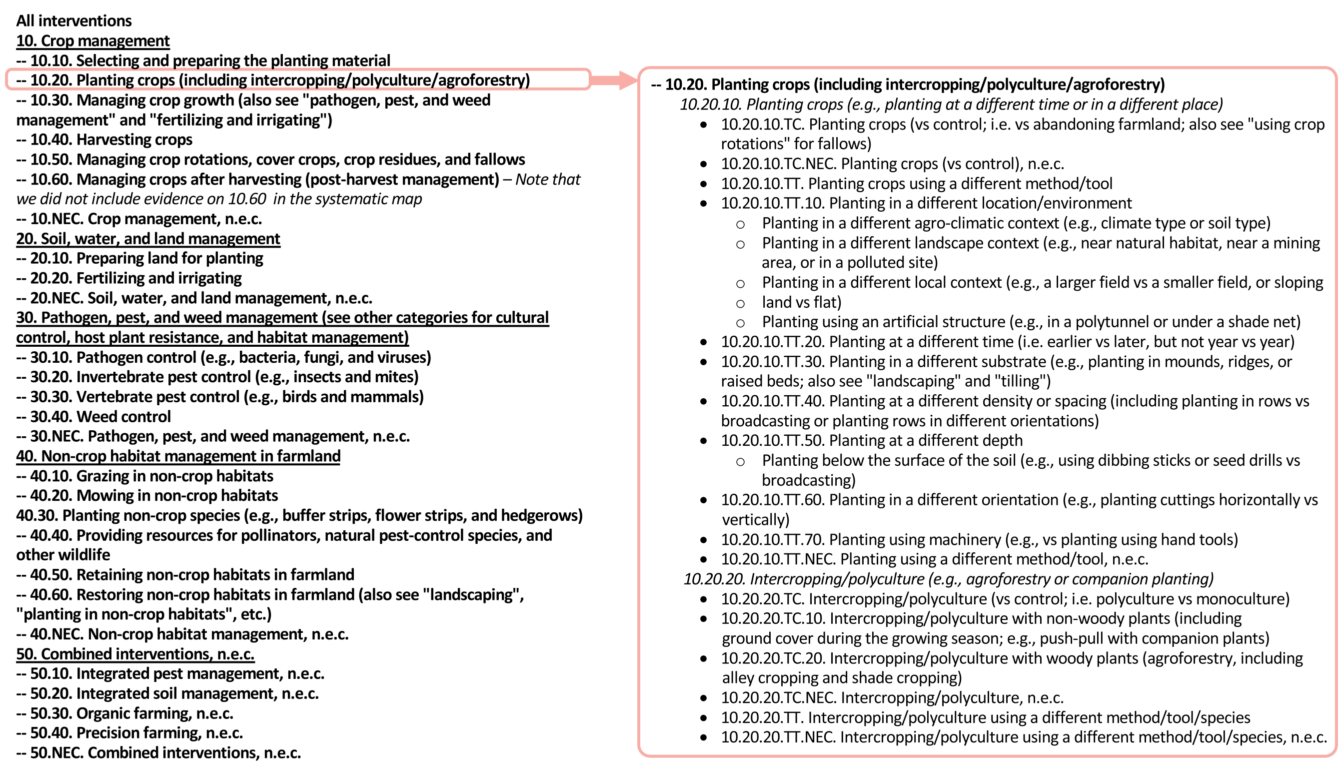

FIGURE 1 The structure of the 'Interventions' ontology, with a sample section expanded to show the lower (more specific) levels of the ontology.

All outcomes

10. Crops
-- 10.10. Crop yield
-- 10.20. Crop quality (including cash crops and cover crops)
-- 10.30. Crop damage, infection, and infestation (also see "pathogens, pests, and weeds")
-- 10.40. Crop growth and survival

20. Soil
-- 20.10. Soil structure and function
-- 20.20. Soil organisms (also see "biodiversity" and "pathogens, pest, and weeds")

30. Water
-- 30.10. Water quantity (use/loss)
-- 30.20. Water quality

40. Pathogens, pests, weeds, and invasive species
-- 40.10. Animals (pests and invasive species)
-- 40.20. Plants (weeds and invasive species)
-- 40.30. Pathogens

50. Pollutants (including greenhouse gases)
-- 50.10. Air pollution
-- 50.20. Greenhouse gases
-- 50.30. Soil pollution

60. Chemicals and energy
-- 60.10. Agrochemicals
-- 60.20. Energy

70. Money, labor, and time
-- 70.10. Money
-- 70.20. Labor
-- 70.30. Time (not including labor measured in person-hours, etc.)

80. Land
-- 80.10. Agricultural land
-- 80.20. Semi-natural land

90. Wildlife (including pollinators and natural enemies)
-- 90.10. Animals (not including livestock or pests)
-- 90.20. Plants (not including crops or weeds)
-- 90.30. Fungi (not including crops or mycorrhizae)

10. Crops
-- 10.10. Crop yield
10.10.10. Crop yield
- Crop yield (including harvest index)
- Crop biomass (including inedible components; e.g., the stems of some crops)
- Crop propagules (including cuttings and seeds not used for food, feed, fuel, etc.)

10.10.NEC. Crop yield, n.e.c.
-- 10.20. Crop quality (including cash crops and cover crops)
10.20.10. Crop appearance (including the size and shape of the edible parts)
10.20.20. Crop cooking quality
10.20.30. Crop storage quality
10.20.40. Crop taste (including acidity, texture, etc.)
10.20.50. Chemical composition of crops
- Dry matter content (i.e. proportion or percentage; also see "Crop yield" for dry matter yield)
- Calories
- ... [Other variables related to composition listed here] ...
- Sugar
- Ash content

10.20.NEC. Crop quality, n.e.c.
10.30. Crop damage, infection, and infestation (also see "pathogens, pests, and weeds")
10.30.10. Crop damage or infection by pathogens (e.g., disease severity)
10.30.20. Crop damage or infestation by pests
10.30.30. Crop damage by weeds
10.30.NEC. Crop damage, n.e.c.
-- 10.40. Crop growth and survival
10.40.10. Crop growth
- Plant growth rate (e.g., photosynthetic rate or radiation use efficiency [RUE])
- Plant size (e.g., canopy cover, ground cover, leaf area index [LAI], or height)

10.40.20. Crop survival (e.g., germination rate)
10.40.NEC. Crop growth and survival, n.e.c.

FIGURE 2 The structure of the 'Outcomes' ontology, with a sample section expanded to show the lower (more specific) levels of the ontology.

intervention or outcome is missing. We revised and clarified these ontologies as we mapped papers, and they should continue to be revised as necessary.

2.2.1 | Agri-ontologies 1.0 intervention ontology

The Agri-ontologies 1.0 intervention ontology included all in-field/on-farm management practices, such as applying amendments, planting in a different agro-ecological zone, irrigating, intercropping, planting crop margins or releasing biological control agents (Figure 1). We included four combined interventions: integrated pest management, integrated soil management, organic farming and precision farming. Many interventions were broken into 'treatment-control' or 'treatment-treatment' comparisons at the lowest level (e.g. a study that compared plots with and without fertiliser would be 'treatment-control', whereas a study that compared two fertiliser types would be 'treatment-treatment'). This distinction was to facilitate reuse of the systematic map for synthesis, as synthesists may wish to focus on one or other type of comparison with conduct a quantitative synthesis.

2.2.2 | Agri-ontologies 1.0 outcome ontology

The Agri-ontologies 1.0 outcome ontology included agricultural or environmental outcomes, such as crop yield, crop quality, soil, water, wildlife, pathogens, pests, pollutants, profits, labour and time (Figure 2). We did not include outcomes on livestock, farmer behaviour or physiology that were not directly relevant to quality or molecular biology (e.g. enzyme activities, molecular markers, gene expression or genetic diversity).

2.2.3 | Consistency in data coding

To check the consistency of coding, two authors (GES and ASCH) coded 100 publications selected randomly. We calculated Kappa scores to test the agreement for coding interventions and outcomes. We discussed any disagreements and revised or clarified the interventions and outcomes ontologies where necessary. As with the article screening, we used a Kappa score of 0.6 as a threshold to signify agreement (Cohen, 1960).

3 | RESULTS AND DISCUSSION

3.1 | Systematic map

3.1.1 | Searches and screening

Following the searches 36,580 records were identified, 20,380 remained after duplicate removal in Endnote X8, 5020 were included after title/abstract screening, 3515 were retrieved at full-text stage, and 1599 were included after full-text screening (Figure 3, see Supporting Information 2 for included and excluded publications).

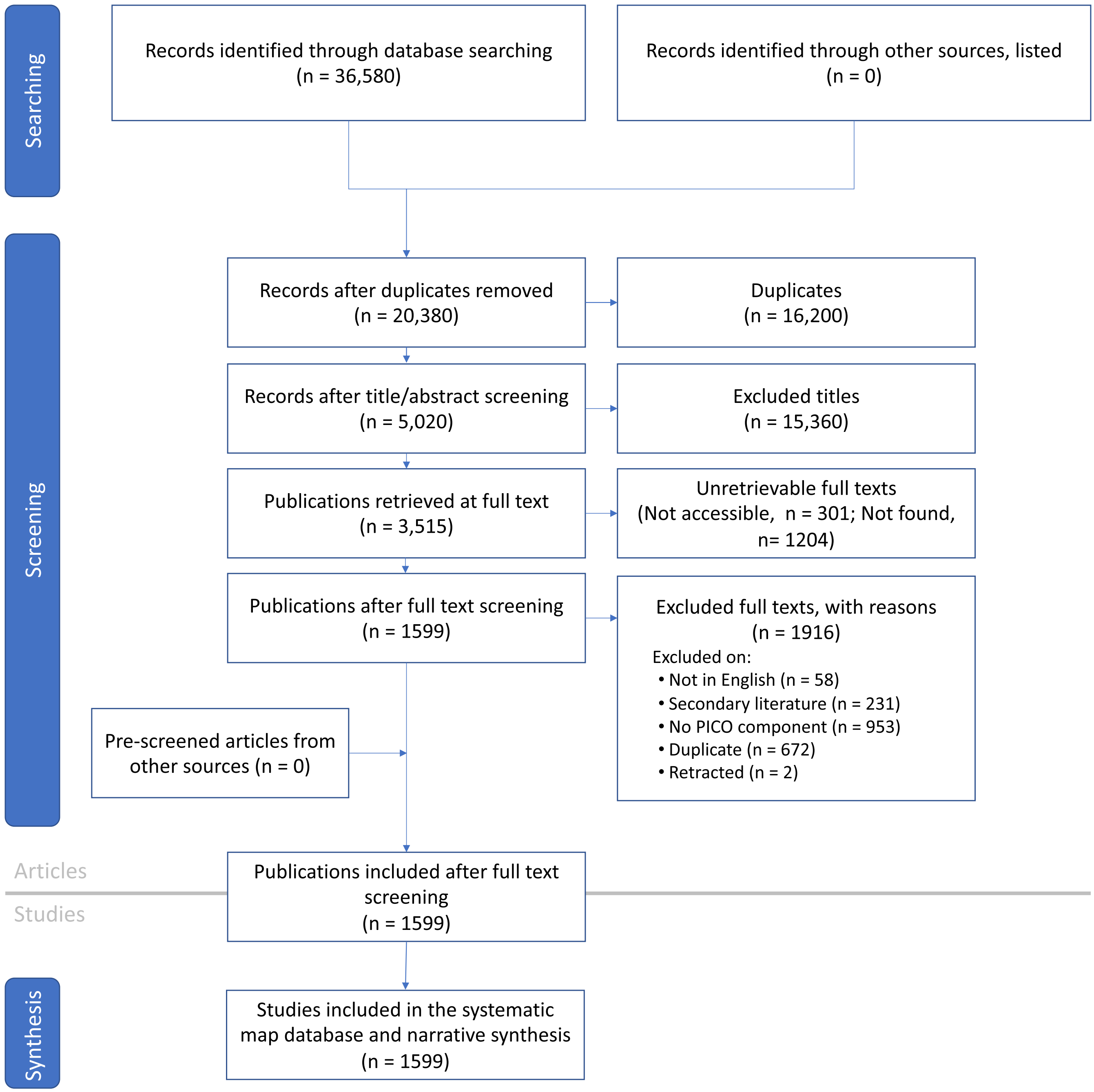

FIGURE 3 A ROSES flow diagram (Haddaway et al., 2017, 2018) showing the searching and screening process.

The threshold to indicate agreement in screening was exceeded following the initial 10% at title and abstract and at full-text stage, so we did not screen additional texts (Title and abstract: Kappa 0.69 and 89% agreement, Full text: Kappa 0.76 and 91% agreement).

Considering the global importance of cassava (e.g. in Africa, it is the second most important crop after maize; FAO, 2019), this is relatively few studies. For example, a Web of Science search for publications prior to December 2017 related to maize (search terms: maize OR 'Zea mays' OR corn) and wheat (search terms: wheat OR triticum) yielded 192,745 and 184,662 records respectively, which is more than nine times the records yielded by our search for cassava (20,380 records), which included multiple search engines and synonyms. Cassava research should be prioritised to increase the evidence base for this globally important crop.

3.1.2 | Temporal distribution

The number of publications has been increasing steadily (Figure 2): mean ± standard error 1967–1977: 3.67 ± 0.91 publications, 1977–1987: 29.60 ± 0.96, 2007–2017: 59.30 ± 1.00. This may indicate a growing interest in researching this globally important crop

(FAO, 2019), or it may follow the wider pattern of increasing publication rates in academia (Fire & Guestrin, 2019; Figure 4).

3.1.3 | Regional distribution

Of the 1599 included publications, we mapped the country-of-study of 998; 76 were excluded as they did not report the country and 525 were excluded as they only tested the intervention 'Planting a different variety/cultivar (e.g. a disease-resistant variety)' (see Section 2). The mapped publications spanned 66 countries, with a particularly high number conducted in Nigeria (*n* = 297), Colombia (*n* = 128) and Brazil (*n* = 82; Figure 5a); these accounted for 45.5% of country occurrences. Country occurrences (*n* = 1113) are greater than mapped publications (*n* = 998) as 47 publications were conducted in multiple countries. In addition to Nigeria, Colombia and Brazil; Thailand (*n* = 55), Indonesia (*n* = 41), Ghana (*n* = 37), India (*n* = 36), Vietnam (*n* = 34), Cameroon (*n* = 31) and Benin (*n* = 30) were the 10 most-studied countries, accounting for 69% of country occurrences (Figure 5a). The remaining countries had under 25 publications each, with the majority (64%) having 5 or fewer.

These results show that research is dominated by a few countries. This may be partly due to the success of large and successful organisations that study cassava, such as CIAT in Colombia and IITA in Nigeria. The number of publications is also partly reflected in production, as six of the seven top producers (exceeding 1000 t/year) are in the top 10 countries for research: Nigeria (5798 t/year), Thailand (3022 t/year), Ghana (1977 t/year), Brazil (1864 t/year), Indonesia (1802 t/year) and Vietnam (1034 t/year; Figure 5). However, half of these countries have under 50 relevant publications, and the remaining top producer (the Democratic Republic of Congo: 3863 t/year) has 11 publications (Figure 5). Therefore, many top-producing countries are understudied, and future research should target these regions (Figure 5). This is particularly important in the context of yield gaps; the average yields of the top seven producing countries (18 t/ha) are less than half of potential yields (44 t/ha; Figure 5). Further research in regions with higher yields may help to inform how to close these gaps and broaden the relevance of cassava research.

3.1.4 | Interventions and outcomes

Many publications tested multiple interventions and outcomes or the same intervention multiple times (*n* = 111), which meant that the number of interventions (3890) and outcomes (4236) tested was greater than the number of mapped publications (1599). The numbers above refer to the lowest (most specific) levels of the intervention and outcome ontologies where each intervention or outcome was represented once. When interpreting these results, please note that outcomes were not coded for the intervention 'Planting a different variety/cultivar (e.g. a disease-resistant variety)' (see Section 2). Additionally, the popularity of interventions and outcomes should be interpreted differently. The authors deliberately measured outcomes, but they may not have deliberately tested an intervention. For example, many studies reported their results broken down by regions with different soil types, which meant that they tested the intervention 'Planting in a different agro-climatic context (e.g. climate type or soil type)'. However, the aim of these studies may have been to test something else (e.g. fertiliser rates), and the agro-climatic context intervention was only tested as an artefact of the way the results were reported (i.e. broken down spatially). This is most likely to have happened with interventions related to space or time (e.g. planting at a different time).

The intervention 'Planting a different variety/cultivar (e.g. a disease-resistant variety)' was far more common than the others (*n* = 939, 24% of intervention occurrences; Figure 6a). The next most common interventions were: 'Planting in a different agro-climatic context (e.g. climate type or soil type)' (*n* = 253); 'Harvesting crops at a different time' (*n* = 217); 'Applying synthetic/mineral/inorganic fertilizer (including N, P, K and other nutrients)' (*n* = 213); 'Intercropping/polyculture with non-woody plants (including ground cover during the growing season; for example, push-pull with companion plants)' (*n* = 204); 'Intercropping/polyculture using a different method/tool/species' (*n* = 195); and 'Using a different amount of amendment/fertilizer (e.g. using the recommended rate of fertilizer)' (*n* = 167; Figure 6a). Together, these seven interventions accounted for 56% of intervention occurrences. The remaining interventions had fewer than 84 publications each (Figure 6a).

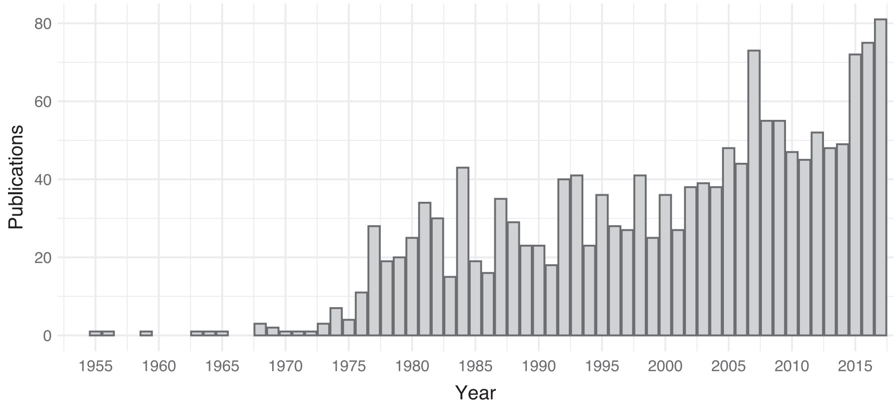

FIGURE 4 A barchart showing the number of mapped publications (*n* = 1599) through time (year of publication).

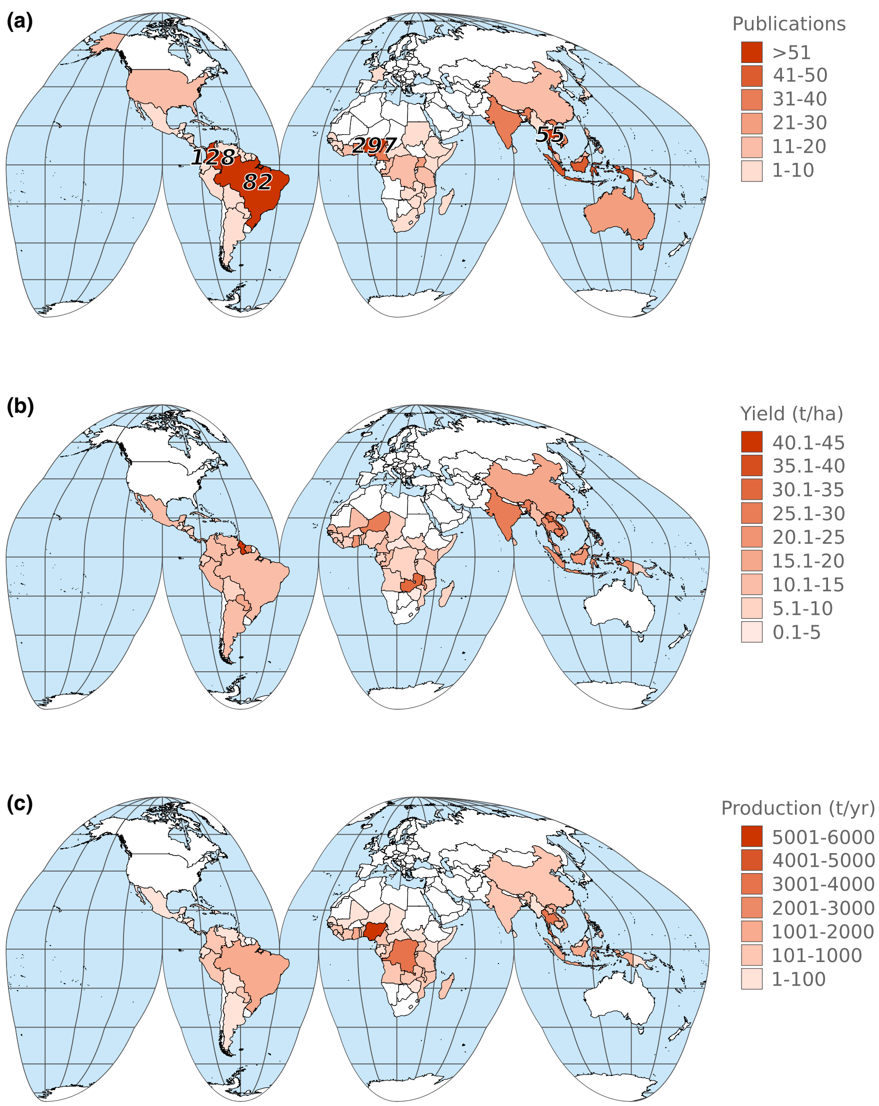

FIGURE 5 World maps using Goode's Homolosine projection, which accurately shows country sizes. Latitude and longitude are shown in grey lines and countries are outlined in black. Brown colour shows (a) the number of publications (*n* = 1004), (b) average cassava yield (t/ha) and (c) total cassava production (t/year). Data from (b) and (c) are country means from 2016 to 2020 according to FAO (FAO, 2019). In (a) there were four countries in the highest bracket, with three outliers; the black numbers show the number of publications for these countries.

One outcome that was far more common than the others: 'Crop yield (per area)' (*n* = 631, 15% of outcome occurrences; Figure 6b). The next most popular outcomes were: 'Plant size (e.g. canopy cover, ground cover, leaf area index [LAI] or height)' (*n* = 252); 'Crop biomass (including inedible components; e.g. the stems of some crops)' (*n* = 204); 'Crop damage or infection by pathogens (e.g. disease severity)' (*n* = 146); 'Dry matter content (i.e. proportion or percentage)' (*n* = 144); 'Starch' content of crops (*n* = 123); and 'Crop yield (per plant)' (*n* = 116; Figure 6b). Together, these seven interventions accounted for 38% of outcome occurrences. These results show that some interventions and outcomes were favoured more than others (i.e. they are knowledge clusters) and that this bias is greater for interventions than outcomes.

The wider knowledge gaps and clusters are more apparent when our intervention and outcome ontologies are viewed as a whole (Figures 7 and 8). Note that higher-level interventions and outcomes (i.e. the higher or less-specific levels of our hierarchical classifications) have fewer occurrences than lower-level ones as we summarised the data to include one of each intervention/outcome per publication to reduce pseudoreplication. In addition, the total number of interventions at the highest level (*n* = 1983, e.g. 'Crop management') and at the middle level (*n* = 2712, e.g. 'Selecting the planting material') and outcomes at the highest level (n = 1571, e.g. 'Crops') and at the middle level (*n* = 2365, e.g. 'Crop yield, biomass, & propagules') is greater than the number of mapped publications (*n* = 1599) because many publications tested multiple interventions/outcomes, or the same intervention over multiple experiments.

At the highest level, the majority of publications tested interventions related to 'Crop management' (*n* = 1395, 70.3% of high-level intervention occurrences), which included mid-level interventions on selecting different types of planting material (*n* = 943, e.g. cultivars), planting in a different context (*n* = 417, e.g. location) and intercropping/polyculture (*n* = 237; Figure 7). Interventions in other categories were rarely studied, with the exception of 'Applying soil amendments/fertiliser' (*n* = 305). Only 139 publications have looked at 'Pathogen, pest & weed management', and there were no publications on vertebrate control, despite there being a number of vertebrate pests in cassava systems (Cudjoe, 1994). Chemical control (*n* = 85) has been studied more than biological (*n* = 39) or physical control methods (*n* = 43, e.g. weeding), and no publications have

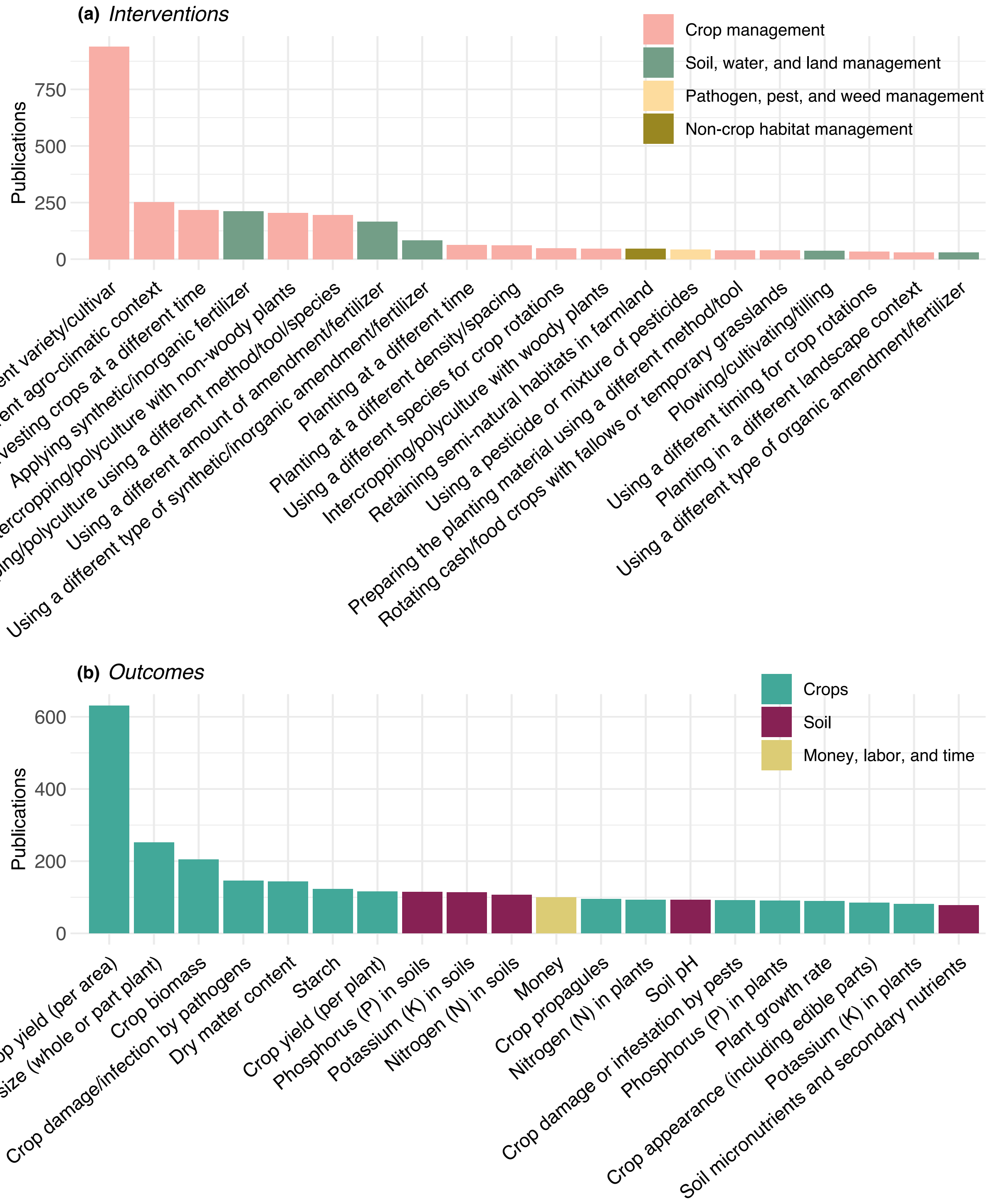

FIGURE 6 Barcharts showing the most studied (a) 20 interventions and (b) 20 outcomes classed at the lowest (most specific) level of the intervention and outcome ontologies. The total number of (a) interventions ($n = 3890$) (b) and outcomes ($n = 4236$) is greater than the number of publications included in the map ($n = 1599$) as many publications tested multiple interventions and outcomes. Colours show the categories of the (a) interventions and (b) outcomes at the highest level in our ontology. Some of these intervention and outcome terms have been abbreviated (see Supporting Information 3 and 4 for the full terms and explanations).

studied biological control of weeds (e.g. undersowing; Figure 7). Few publications have looked at 'Non-crop habitat management' ($n = 67$, e.g. planting in margins; Figure 7).

In terms of outcomes, the majority of publications measured outcomes related to 'Crops' ($n = 913$, 58.3% of high-level outcome occurrences) and 'Soil' ($n = 264$, 16.8% of high-level outcome occurrences; Figure 8). Within the 'Soil' category, 'Soil structure, functions, & chemistry' was measured more than 'Soil organisms' (17.8% of occurrences in the 'Soil' category). Outcomes in other categories were rarely measured, with the exception of 'Money' ($n = 100$, 4.2%

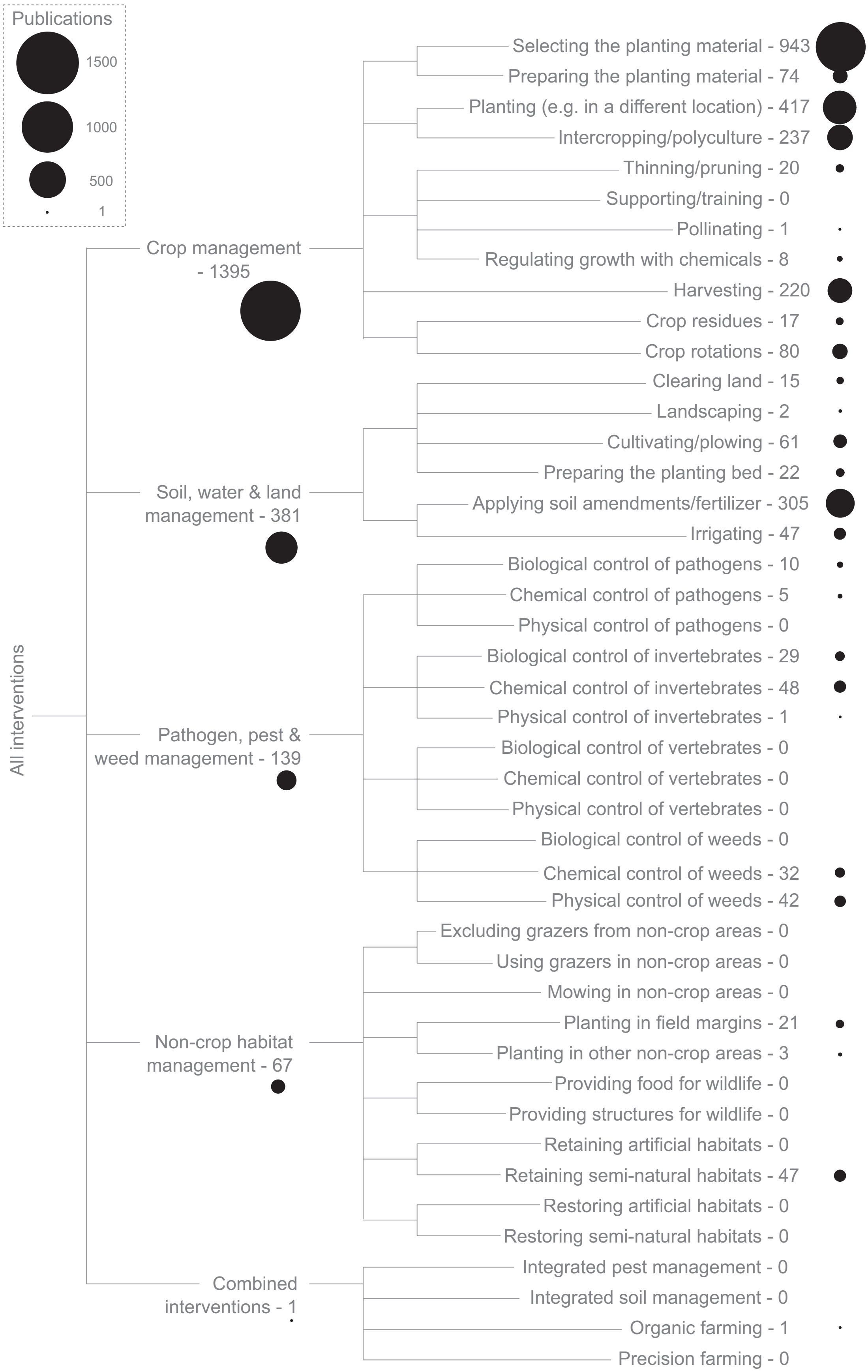

FIGURE 7 A scheme showing the relationship between the highest (least specific) and middle levels of the intervention ontology (see Supporting Information 3 for the full ontology) and the number of publications for each intervention (according to the area of the black points). Some of these intervention terms have been abbreviated (see Supporting Information 3 for the full terms and explanations).

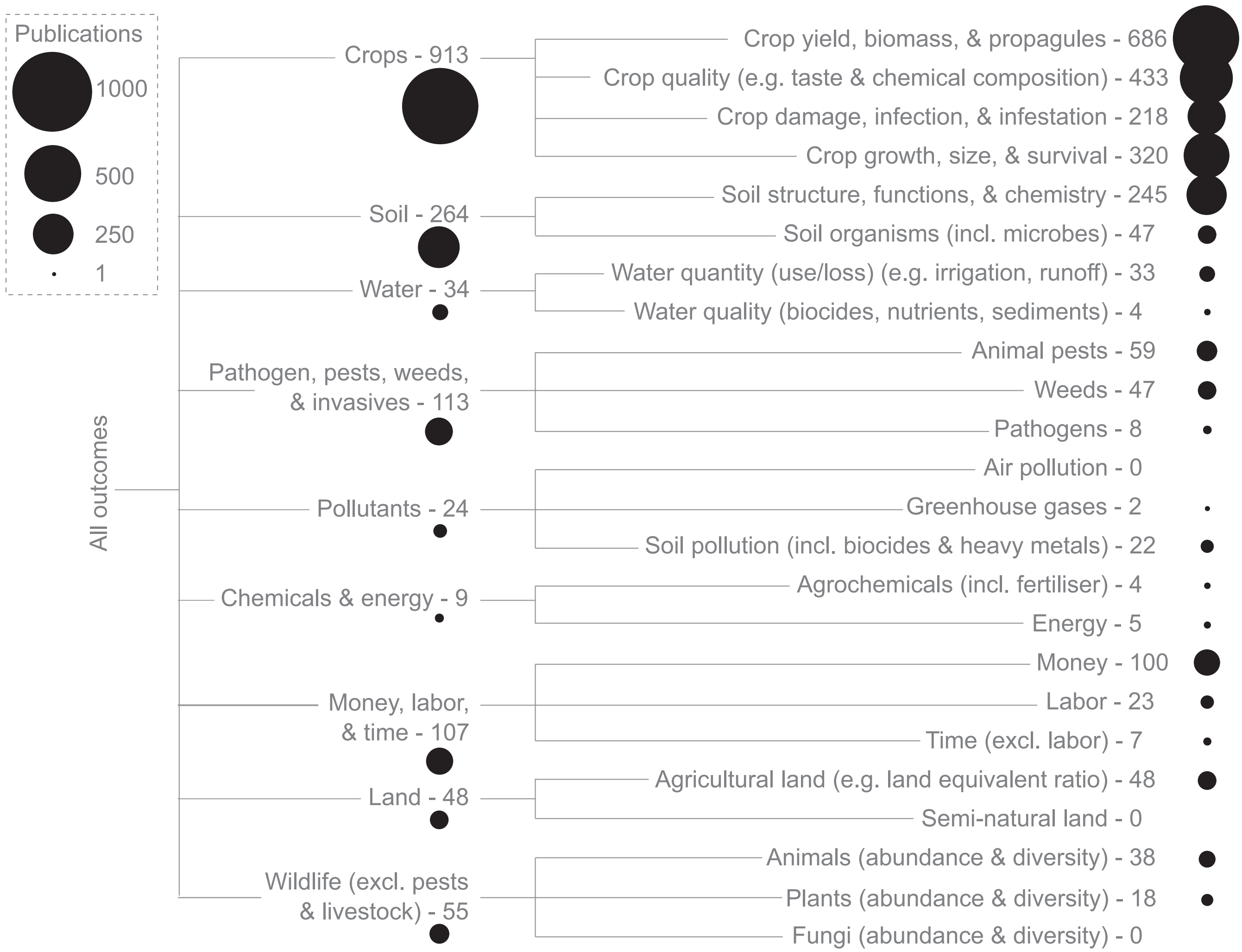

FIGURE 8 A scheme showing the relationship between the highest (least specific) and middle levels of the outcome ontology (see Supporting Information 4 for the full ontology) and the number of publications for each outcome (according to the area of the black points). Some of these outcome terms have been abbreviated (see Supporting Information 4 for the full terms and explanations).

of mid-level outcome occurrences). Few publications measured environmental outcomes, with only 24 publications that studied 'Pollutants' and 55 that studied 'Wildlife' (combined 5% of higher-level outcome occurrences) (Figure 8).

Taken together, these results show that there are major knowledge gaps, particularly related to environmental management and outcomes. The majority of research has focussed on production, with minimal research into non-crop habitat, wildlife or pollutants. It is imperative that this knowledge gap is addressed so cassava production and expansion can be managed as sustainably as possible, as evidence-based management can reduce negative impacts on biodiversity (Sutherland et al., 2004; Walsh et al., 2015). This is likely to be important for maintaining yield stability in cassava too by providing resilience to yield fluctuations, as publications in other systems have shown that biodiversity can achieve this (Dardonville et al., 2022; Haughey et al., 2018). We also identify knowledge clusters, such as interventions related to cultivar selection, harvesting, fertiliser application, intercropping and crop rotations, which can be used for syntheses. Therefore, synthesists should conduct syntheses in knowledge clusters, and funders, stakeholders and researchers should target resources towards improving the knowledge base around environmental management and outcomes in cassava systems.

3.1.5 | Study design

We included four aspects of study design to assess the quality of mapped studies and evidence base as a whole: reporting standards (reporting completeness), experimental design, study duration and publication delay. We tracked reporting standards for country and experimental dates; 24% ($n=384$) of publications did not report experimental start dates and 7% ($n=112$) did not report country-of-study. These are basic experimental details that should be reported in every study so that study context can be properly understood and to facilitate reproducibility and data reuse. We did not track reporting standards for experimental design because having a study comparator was a criterion for publication inclusion during mapping.

Some experimental designs are more biased than others (Christie et al., 2020). Generally, designs that incorporate more of the features listed in Table 1 are considered more robust than simpler designs, and experimental randomised controlled studies and observational before-after controlled or paired/blocked controlled studies are considered to be more robust than simpler controlled or before-after designs, as well as correlational studies (Christie et al., 2020). However, there are other aspects of experimental design (e.g. statistical matching, sample size and plot size) that affect study bias and robustness (Christie et al., 2020). Furthermore, some designs cannot be applied to certain questions (e.g. a study comparing primary forest to another ecosystem cannot be controlled because you cannot create a new primary forest). For publications included in our systematic map, the most common design was a randomised blocked design ($n=1069$, 36.23% of experimental design occurrences; Table 1), which is very robust (Christie et al., 2020). Other robust designs included before-after-control-impact (BACI) experiments ($n=61$, 2.07%) and randomised controlled trials ($n=68$, 2.3%) (Christie et al., 2020). However, the proportion of simpler designs (e.g. before-after designs, unreplicated controlled designs, correlative designs and non-randomised controlled designs without pairing or blocking) was also large ($n=1753$, 59.4%).

Study duration and publication delay was recorded for 821 publications, as 253 did not report start dates and 525 were not tracked as they only tested cassava cultivars (see Section 2). The study duration was 2.0 ± 0.1 years (mean ± standard error). It has also been decreasing, with visual inspection using LOESS smoothers showing estimated values of 2.7 years in 2000 and 1.3 years in 2017 (Figure 9a). Short study durations are partly due to funding landscapes, where projects are often funded over short periods and researchers are required to demonstrate impact within this time (Grove, 2018). This hinders research, as long-term studies contribute disproportionately both to scientific understanding and to policy in ecology (Hughes et al., 2017; Mills et al., 2015). Long publication delays are also problematic, as they reduce researcher, practitioner and funder access to the latest and most relevant research, which creates an inefficient research environment (Christie et al., 2021). Here, we show that publication delay is 4.0 ± 0.1 years (mean ± standard error), which is longer than delays in the related field of conservation science (3.2 ± 0.1) (Christie et al., 2021). Visual inspection of publication delay through time using LOESS smoothers showed that this has been relatively constant, with some reduction in the last 15 years (approximately −0.7-year delay; Figure 9b).

Based on our findings, we recommend that researchers and other stakeholders (e.g. journals and funders) promote the use of reporting checklists or guidelines by authors, reviewers and editors, to improve reporting standards (Hopewell et al., 2012; Plint et al., 2006; Stevens et al., 2014; Turner et al., 2012). Long-term projects should also be prioritised to increase study duration, as should more robust experimental designs. Publication delays have multiple causes, and changes are needed throughout the research process to address these causes (e.g. simplifying formatting requirements, improving

TABLE 1 All possible experimental designs for studies with comparators: controlled, correlated and before-after. Designs that are not shown (e.g. randomised correlated studies) are not possible. Bars and numbers indicate the number of interventions that were tested using each design ($n=2951$). Some study designs have commonly used names, which are referenced in the 'design' column.

Comparator	Before-after	Replicated	Randomised	Blocked/paired	*Design*	*n*
Controlled						429
	✓					10
		✓				740
			✓			1
	✓	✓			*BACI*[a]	5
	✓	✓		✓	*BACI*[a]	6
	✓	✓	✓	✓	*R-BACI*[b]	50
		✓		✓		114
		✓	✓		*RCT*[c]	68
		✓	✓	✓	*RCB*[d]	1069
			✓	✓		1
Correlated						182
	✓					0
		✓				245
				✓		4
	✓	✓				3
		✓		✓		4
	✓	✓		✓		0
Before-after	✓					11
	✓	✓				9

[a]Before-after-control-impact.

[b]Randomised before-after-control-impact.

[c]Randomised controlled trial or randomised control-impact.

[d]Randomised blocked design, including randomised complete blocked design.

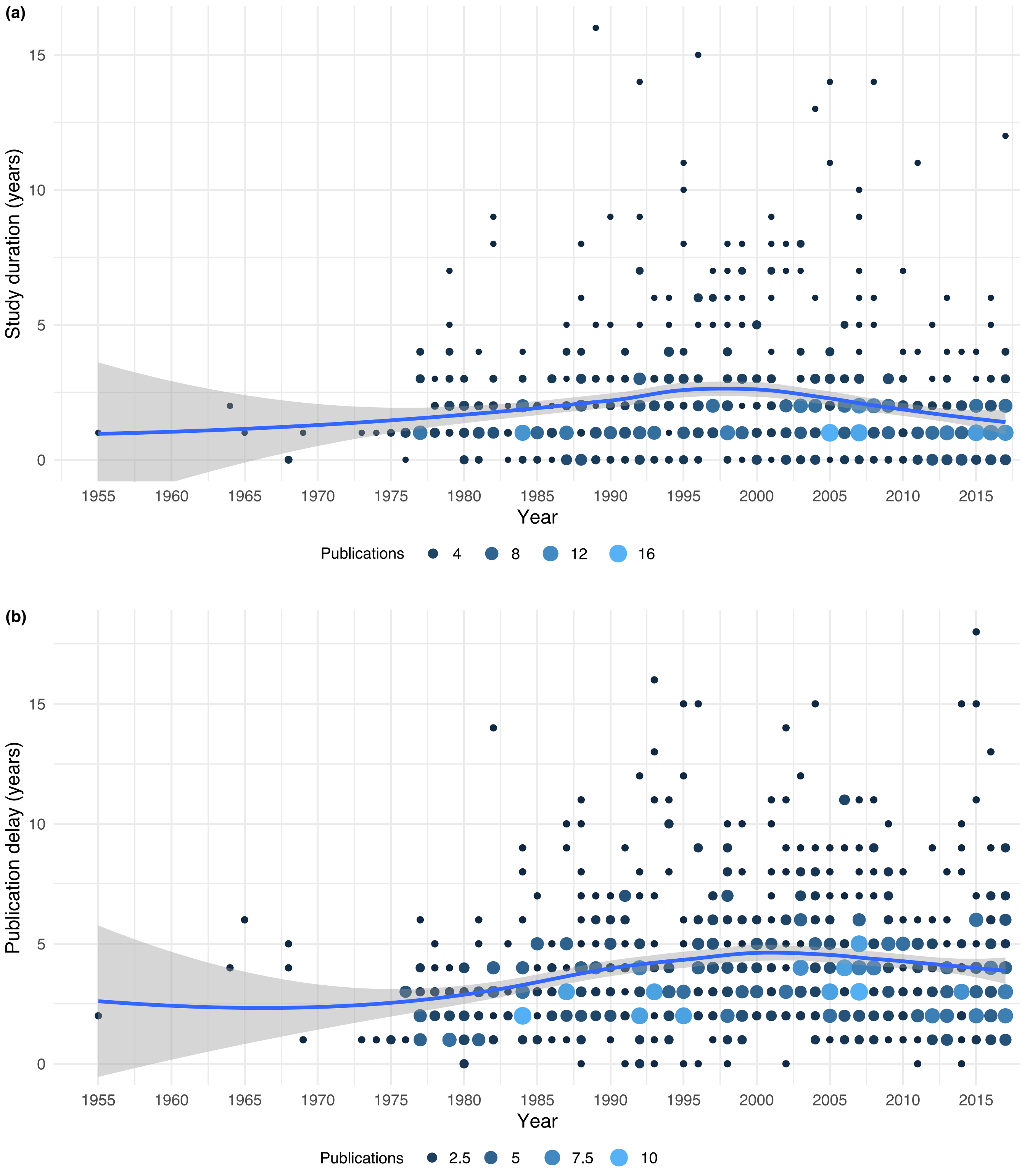

FIGURE 9 Scatterplots showing the (a) study duration and (b) publication delay through time (year of publication). The colour and point area shows the number of publications with that value (maximum: [a] 16, [b] 10). Blue lines show local regression smoothers (LOESS) with a span of 1 and shaded areas show 95% pointwise confidence intervals. Zero values show durations <1 year.

author guidelines, providing assistance to authors writing in a second language, incentivising peer review, allowing submission to multiple journals, or using pre-print servers; Christie et al., 2021). These results can also be used to direct future resources. For example, after synthesists have identified knowledge clusters using the interventions and outcomes maps, they can interrogate these data online (www.metadataset.com) to see whether their chosen topic has a robust study design. If not, they may wish to choose a different topic. Similarly, funders and field researchers can use these results to identify topics that require further research, either because they have been overlooked or because they have been tested with biased study designs.

3.2 | Agri-ontologies 1.0: Ontologies of farming interventions and outcomes

Our secondary objective was to develop and test our hierarchical classification system for [1] farming interventions and [2] agricultural, economic and environmental outcomes. The agreement threshold (Kappa value of 0.6) for coding these interventions and outcomes was not reached after the two authors coded the first ~10% of studies, so we clarified the classifications based on discussions and then added another 1% of studies and recalculated the Kappa scores. After this, the threshold was exceeded (Interventions: Kappa 0.68 and 98% agreement, Outcomes: Kappa 0.67 and 98% agreement). Based on these results, we can conclude that we reliably classified the publications into interventions and outcomes and that our ontologies can be coded consistently (see Figures 1 and 2 for examples, Supporting Information 3 and 4 for the full ontologies, or www.metadataset.com for an interactive online format).

Like other ontologies, Agri-ontologies 1.0 can improve reporting standards and increase the interoperability of data (Hopewell et al., 2012; Plint et al., 2006), therefore promoting data reuse for uses such as transfer learning, generating synthetic data and syntheses (e.g. meta-analysis; Todman et al., 2023). Hierarchical classification systems such as this enable the automatic lumping and splitting of evidence, using methods such as 'dynamic meta-analysis' (Martin et al., 2022; Shackelford et al., 2021). Agri-ontologies 1.0 complements existing ontologies, such as DarwinCore (Wieczorek et al., 2012) and AgrO (Aubert et al., 2017). It is unique by having a broad focus, and being comprehensive whilst also being general enough to apply to other agricultural systems. It applies the use of 'not elsewhere classified' categories to identify common categories that are missing and should be added. In addition, it facilitates quantitative syntheses, such as meta-analyses, by using 'treatment-treatment' and 'treatment-control' categories.

Agri-ontologies 1.0 represents an important step towards 'subject-wide evidence synthesis', as developing subject-wide classification systems facilitates comparison between subjects (Sutherland & Wordley, 2018). For example, if Agri-ontologies 1.0 was used to systematically map the literature on other crops, then these maps would be interoperable, and this could eventually lead to a global map of scientific literature on agriculture.

4 | CONCLUSIONS

Here, we provide two advancements in the field of evidence synthesis. We provide the first example of an evidence atlas for an entire crop and we develop hierarchical classifications of farming interventions and outcomes tailored to evidence synthesis. The results of our searches indicated that there has been little research on cassava compared with other major global crops: maize and wheat. Furthermore, we found regional knowledge gaps and clusters. Three countries (Nigeria, Colombia and Brazil) accounted for nearly half of the country occurrences, and many top-producing countries are under-researched (e.g. the Democratic Republic of Congo produces 3863 t/year and has 11 publications). In Africa, cassava is the second most important crop after maize (FAO, 2019) and its importance may increase further as it is better adapted to the impacts of climate change than other crops (Jarvis et al., 2012). Therefore, increasing the evidence base around cassava in under-researched regions should be prioritised to broaden the relevance of cassava research and reduce existing yield gaps there.

We also found knowledge gaps and clusters in the interventions and outcomes that were tested. The majority of research has focussed on production, with minimal research into non-crop habitats, wildlife or pollutants. Improving the knowledge base around environmental interventions and outcomes will facilitate evidence-based management of cassava, and therefore, production and expansion that is more sustainable and reduces negative impacts on biodiversity (Sutherland et al., 2004; Walsh et al., 2015). Such environmental benefits will likely also promote yield stability in cassava, as biodiversity provides resilience to yield fluctuations in other systems (Dardonville et al., 2022; Haughey et al., 2018). This may help to address existing yield gaps (Figure 5). In terms of knowledge clusters, synthesists can use this map to identify well-researched topics; for example, we found many publications for interventions related to fertiliser use and crop rotations. Stakeholders should aim to increase the knowledge base in environmental topics in particular to facilitate sustainable management of cassava.

There were several concerning trends in terms of study design. Reporting standards were poor, and we recommend that reporting checklists/guidelines are promoted within this field to improve this (Hopewell et al., 2012; Plint et al., 2006; Stevens et al., 2014; Turner et al., 2012). Long-term projects and more robust experimental designs should also be prioritised, and systemic changes are needed across academia to address the problem of long publication delays (Christie et al., 2021). Given the importance of cassava as a staple crop, we urge researchers, funders, policymakers and other stakeholders to use our systematic map to guide syntheses and fill knowledge gaps with primary studies. Doing this could help to close existing yield gaps, increase food security and reduce the impacts of cassava production on the wider environment.

AUTHOR CONTRIBUTIONS

Amelia S. C. Hood: writing—original draft preparation (lead), visualisation (equal), validation (equal), methodology (supporting), investigation (equal), formal analysis (equal), data curation (equal), writing—review and editing (equal). Gorm E. Shackelford: visualisation (equal), validation (equal), supervision (co-lead), software (lead), methodology (lead), investigation (equal), funding acquisition (supporting), formal analysis (equal), data curation (equal),

conceptualisation (lead), writing—review and editing (equal). Hope O. Usieta: methodology (supporting), funding acquisition (supporting), conceptualisation (supporting), writing—review and editing (equal). Alec P. Christie: methodology (supporting), data curation (equal), writing—review and editing (equal). Philip A. Martin: supervision (supporting), methodology (supporting), formal analysis (equal). William J. Sutherland: supervision (co-lead), methodology (supporting), funding acquisition (lead), conceptualisation (supporting), writing—review and editing (equal).

ACKNOWLEDGEMENTS

We are grateful to the A. G. Leventis Foundation, Arcadia, and the David and Claudia Harding Foundation for funding. We thank Dr Neal Haddaway, Dr Silviu Petrovan and Dr Pieter Pypers for their help in writing the protocol that this systematic map was based.

FUNDING INFORMATION

Arcadia Fund; The David and Claudia Harding Foundation; A G Leventis Foundation.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/2688-8319.12249.

DATA AVAILABILITY STATEMENT

The datasets generated and analysed during the current study are available via an interactive map on the Metadataset website (www.metadataset.com). They are also available in spreadsheet form: https://doi.org/10.17864/1947.000429 (Hood et al., 2022).

ORCID

Amelia S. C. Hood https://orcid.org/0000-0003-3803-0603
Gorm E. Shackelford https://orcid.org/0000-0003-0949-0934
Alec P. Christie https://orcid.org/0000-0002-8465-8410
Philip A. Martin https://orcid.org/0000-0002-5346-8868
William J. Sutherland https://orcid.org/0000-0002-6498-0437

REFERENCES

Aubert, C., Buttigieg, P. L., Laporte, M.-A., Devare, M., & Devare, M. (2017). CGIAR agronomy ontology. http://purl.obolibrary.org/obo/agro.owl

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*, 452–454.

Burns, A., Gleadow, R., Cliff, J., Zacarias, A., & Cavagnaro, T. (2010). Cassava: The drought, war and famine crop in a changing world. *Sustainability*, *2*(11), 3572–3607.

Christie, A. P., Abecasis, D., Adjeroud, M., Alonso, J. C., Amano, T., Anton, A., Baldigo, B. P., Barrientos, R., Bicknell, J. E., Buhl, D. A., Cebrian, J., Ceia, R. S., Cibils-Martina, L., Clarke, S., Claudet, J., Craig, M. D., Davoult, D., de Backer, A., Donovan, M. K., … Sutherland, W. J. (2020). Quantifying and addressing the prevalence and bias of study designs in the environmental and social sciences. *Nature Communications*, *11*(1), 6377.

Christie, A. P., White, T. B., Martin, P. A., Petrovan, S. O., Bladon, A. J., Bowkett, A. E., Littlewood, N. A., Mupepele, A. C., Rocha, R., Sainsbury, K. A., Smith, R. K., Taylor, N. G., & Sutherland, W. J. (2021). Reducing publication delay to improve the efficiency and impact of conservation science. *PeerJ*, *9*(October), e12245.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cousijn, H., Kenall, A., Ganley, E., Harrison, M., Kernohan, D., Lemberger, T., Murphy, F., Polischuk, P., Taylor, S., Martone, M., & Clark, T. (2018). A data citation roadmap for scientific publishers. *Scientific Data*, *5*, 1–11.

Cudjoe, A. R. (1994). Vertebrate pests of cassava in Africa and their control. *African Crop Science Journal*, *2*(4), 497–503.

Dardonville, M., Bockstaller, C., Villerd, J., & Therond, O. (2022). Resilience of agricultural systems: Biodiversity-based systems are stable, while intensified ones are resistant and high-yielding. *Agricultural Systems*, *197*(June 2021), 103365. https://doi.org/10.1016/j.agsy.2022.103365

Equator Network. (2022). https://www.equator-network.org/

FAO. (2000). The world cassava economy: Facts, trends, and outlook. 1–64. https://www.fao.org/publications/card/en/c/92ab39ea-841c-5851-84fc-8df9590e260f/

FAO. (2013). Cassava, a 21st century crop. Save and grow: Cassava, a guide to sustainable production intensification. 1–18 p. http://www.fao.org/ag/save-and-grow/cassava/en/1/index.html

FAO. (2019). http://faostat.fao.org. http://www.fao.org/faostat/en/#data/QC

Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics: Observing Goodhart's Law in action. *Gigascience*, *8*(6), 1–20.

Gade, D. W., Journal, S., Geography, A., & Gade, D. W. (2002). Names for *Manihot esculenta*: Geographical variation and lexical clarification. *Journal of Latin American Geography*, *1*(1), 43–57.

Grove, L. (2018). *The effects of funding policies on academic research*. University College London.

Haddaway, N., Macura, B., Whaley, P., & Pullin, A. (2017). ROSES flow diagram for systematic maps.

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of google scholar in evidence reviews and its applicability to grey literature searching. *PLoS ONE*, *10*(9), 1–17.

Haddaway, N. R., Macura, B., Whaley, P., & Pullin, A. S. (2018). ROSES reporting standards for systematic evidence syntheses: Pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environmental Evidence*, *7*(1), 4–11. https://doi.org/10.1186/s13750-018-0121-7

Haughey, E., Suter, M., Hofer, D., Hoekstra, N. J., McElwain, J. C., Lüscher, A., & Finn, J. A. (2018). Higher species richness enhances yield stability in intensively managed grasslands with experimental disturbance. *Scientific Reports*, *8*(1), 1–10.

Hillocks, R. J., Thresh, J. M., & Bellotti, A. (2002). *Cassava: Biology, production and utilization*. CABI.

Hood, A. S. C., Christie, A. P., Sutherland, W., & Shackelford, G. E. (2022). Dataset supporting: "A systematic map of cassava farming practices and their agricultural and environmental impacts using new ontologies: Agri-ontologies 1.0". University Reading. Dataset. https://doi.org/10.17864/1947.000429

Hood, A. S. C., & Sutherland, W. J. (2021). The data-index: An author-level metric that values impactful data and incentivizes data sharing. *Ecology and Evolution*, *11*(21), 14344–14350.

Hopewell, S., Ravaud, P., Baron, G., & Boutron, I. (2012). Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: Interrupted time series analysis. *BMJ*, *345*(7864), 1–7.

Howeler, R., & Aye, T. M. (2017). Sustainable management of cassava in Asia: From research to practice. https://cgspace.cgiar.org/bitstream/handle/10568/91531/U17ManAtserWeedNothomDev.pdf?sequence=1&isAllowed=y

Hughes, B. B., Beas-Luna, R., Barner, A. K., Brewitt, K., Brumbaugh, D. R., Cerny-Chipman, E. B., Close, S. L., Coblentz, K. E., de Nesnera, K. L., Drobnitch, S. T., Figurski, J. D., Focht, B., Friedman, M., Freiwald, J., Heady, K. K., Heady, W. N., Hettinger, A., Johnson, A., Karr, K. A., … Carr, M. H. (2017). Long-term studies contribute disproportionately to ecology and policy. *Bioscience*, *67*(3), 271–278.

Hunt, L. A., Hoogenboom, G., Jones, J. W., & White, J. W. (2006, August). ICASA version 1.0 data standards for agricultural research and decision support. 1–37.

Inkscape. (2020). Inkscape project. https://inkscape.org

International Institute of Tropical Agriculture (IITA). (2017). The ABC of weed management in cassava production in Nigeria: A training manual. https://cgspace.cgiar.org/bitstream/handle/10568/91531/U17ManAtserWeedNothomDev.pdf?sequence=1&isAllowed=y

James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence*, *5*(1), 1–13.

Jarvis, A., Ramirez-Villegas, J., Campo, B. V. H., & Navarro-Racines, C. (2012). Is cassava the answer to African climate change adaptation? *Tropical Plant Biology*, *5*, 9–29.

Jonquet, C., Toulet, A., Dutta, B., & Emonet, V. (2018). Harnessing the power of unified metadata in an ontology repository: The case of AgroPortal. *Journal on Data Semantics*, *7*(4), 191–221. https://doi.org/10.1007/s13740-018-0091-5

Konkiel, S. (2020). Assessing the impact and quality of research data using altmetrics and other indicators. *Scholarly Assessment Reports*, *2*(1), 1–16.

Konno, K., Akasaka, M., Koshida, C., Katayama, N., Osada, N., Spake, R., & Amano, T. (2020). Ignoring non-English-language studies may bias ecological meta-analyses. *Ecology and Evolution*, *10*(13), 6373–6384.

Li, Y., Ye, W., Wang, M., & Yan, X. (2009). Climate change and drought: A risk assessment of crop-yield impacts. *Climate Research*, *39*(May), 31–46.

Malhi, G. S., Kaur, M., & Kaushik, P. (2021). Impact of climate change on agriculture and its mitigation strategies: A review. *Sustainability*, *13*(3), 1–21.

Martin, A. P. A., Christie, A. P., Shackelford, G. E., Amelia, S. C., Wang, S., Li, B., Morgan, W., Lee, M., Aldridge, D. C., & Sutherland, W. J. (2022). *Flexible synthesis to deliver relevant evidence for decision-making*. OSF.

Mills, J. A., Teplitsky, C., Arroyo, B., Charmantier, A., Becker, P. H., Birkhead, T. R., Bize, P., Blumstein, D. T., Bonenfant, C., Boutin, S., Bushuev, A., Cam, E., Cockburn, A., Côté, S. D., Coulson, J. C., Daunt, F., Dingemanse, N. J., Doligez, B., Drummond, H., … Zedrosser, A. (2015). Archiving primary data: Solutions for long-term studies. *Trends in Ecology & Evolution*, *30*(10), 581–589.

Mislan, K. A. S., Heer, J. M., & White, E. P. (2016). Elevating the status of code in ecology. *Trends Ecology & Evolution*, *31*(1), 4–7. https://doi.org/10.1016/j.tree.2015.11.006

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*(1), 439–446.

Plint, A. C., Moher, D., Morrison, A., Schulz, K., Altman, D. G., Hill, C., & Gaboury, I. (2006). Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *The Medical Journal of Australia*, *185*(5), 263–267.

Pypers, P. (2017). The African cassava agronomy initiative (ACAI) project model. https://cgspace.cgiar.org/handle/10568/88115

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reynolds, T. W., Waddington, S. R., Anderson, C. L., Chew, A., True, Z., & Cullen, A. (2015). Environmental impacts and constraints associated with the production of major food crops in sub-Saharan Africa and South Asia. *Food Security*, *7*(4), 795–822.

RStudio Team. (2019). *RStudio: Integrated development for R*. RStudio, Inc. http://www.rstudio.com/

Shackelford, G., Martin, P., Hood, A., Christie, A., Kulinskaya, E., & Sutherland, W. (2020). Dynamic meta-analysis: A method of using global evidence for local decision making. *BMC Biology*, *19*, 1–23.

Shackelford, G. E., Haddaway, N. R., Usieta, H. O., Pypers, P., Petrovan, S. O., & Sutherland, W. J. (2018). Cassava farming practices and their agricultural and environmental impacts: A systematic map protocol. *Environmental Evidence*, *7*(1), 1–7. https://doi.org/10.1186/s13750-018-0142-2

Shackelford, G. E., Martin, P. A., Hood, A. S. C., Christie, A. P., Kulinskaya, E., & Sutherland, W. J. (2021). Dynamic meta-analysis: A method of using global evidence for local decision making. *BMC Biology*, *19*(1), 33.

South, A. (2011). rworldmap: A new R package for mapping global data. *The R Journal*, *3*(1), 35–43.

Stevens, A., Shamseer, L., Weinstein, E., Yazdi, F., Turner, L., Thielman, J., Altman, D. G., Hirst, A., Hoey, J., Palepu, A., Schulz, K. F., & Moher, D. (2014). Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: Systematic review. *BMJ*, *348*(June), 1–29. https://doi.org/10.1136/bmj.g3804

Sutherland, W. J., Dicks, L. V., Petrovan, S. O., & Smith, R. K. (2021). What works in conservation.

Sutherland, W. J., Pullin, A. S., Dolman, P. M., & Knight, T. M. (2004). The need for evidence-based conservation. *Trends in Ecology & Evolution*, *19*(6), 305–308.

Sutherland, W. J., & Wordley, C. F. R. (2018). A fresh approach to evidence synthesis comment. *Nature*, *558*(7710), 364–366.

The Leventis Foundation Nigeria [Internet]. https://leventisfoundation.org.ng/

Thessen, A. E., Bunker, D. E., Buttigieg, P. L., Cooper, L. D., Dahdul, W. M., Domisch, S., Franz, N. M., Jaiswal, P., Lawrence-Dill, C. J., Midford, P. E., Mungall, C. J., Ramírez, M. J., Specht, C. D., Vogt, L., Vos, R. A., Walls, R. L., White, J. W., Zhang, G., Deans, A. R., … Mabee, P. M. (2015). Emerging semantics to link phenotype and environment. *PeerJ*, *2015*(12), 1–39.

Todman, L. C., Bush, A., & Hood, A. S. C. (2023). Ecology & Evolution 'small data' for big insights in ecology. *Trends Ecology & Evolution*, *3117*, 1–8. https://doi.org/10.1016/j.tree.2023.01.015

Turner, L., Shamseer, L., Altman, D. G., Schulz, K. F., & Moher, D. (2012). Does use of the CONSORT statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Systematic Reviews*, *1*(1), 60.

United Nations Department of Economic and Social Affairs Population Division. (2019). World population prospects 2019: Highlights. http://www.ncbi.nlm.nih.gov/pubmed/12283219

Walsh, J. C., Dicks, L. V., & Sutherland, W. J. (2015). The effect of scientific evidence on conservation practitioners' management decisions. *Conservation Biology*, *29*(1), 88–98.

White, J. W., Hunt, L. A., Boote, K. J., Jones, J. W., Koo, J., Kim, S., Porter, C. H., Wilkens, P. W., & Hoogenboom, G. (2013). Integrated description of agricultural field experiments and production: The ICASA version 2.0 data standards. *Computers and Electronics in Agriculture*, *96*, 1–12. https://doi.org/10.1016/j.compag.2013.04.003

Wickham, H. (2017). Tidyverse: Easily install and load the "Tidyverse". R package version 1.2.1. https://cran.r-project.org/package=tidyverse

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin core: An evolving community-developed biodiversity data standard. *PLoS ONE*, *7*(1), e29715.

Wilke, C. O. (2019). cowplot: Streamlined plot theme and plot annotations for "ggplot2". https://cran.r-project.org/package=cowplot

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 1–9.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting Information 1. Roses checklist.

Supporting Information 2. FullTexts.

Supporting Information 3. Interventions.

Supporting Information 4. Outcomes.

How to cite this article: Hood, A. S. C., Shackelford, G. E., Christie, A. P., Usieta, H. O., Martin, P. A., & Sutherland, W. J. (2023). A systematic map of cassava farming practices and their agricultural and environmental impacts using new ontologies: Agri-ontologies 1.0. *Ecological Solutions and Evidence*, *4*, e12249. https://doi.org/10.1002/2688-8319.12249