Distributions, *ex situ* conservation priorities, and genetic resource potential of crop wild relatives of sweetpotato [*Ipomoea batatas* (L.) Lam., *I.* series *Batatas*]

Colin K. Khoury[1,2]*, Bettina Heider[3], Nora P. Castañeda-Álvarez[1,4], Harold A. Achicanoy[1], Chrystian C. Sosa[1], Richard E. Miller[5], Robert W. Scotland[6], John R. I. Wood[6], Genoveva Rossel[3], Lauren A. Eserman[7], Robert L. Jarret[8], G. C. Yencho[9], Vivian Bernau[1,10], Henry Juarez[3], Steven Sotelo[1], Stef de Haan[1,3] and Paul C. Struik[2]

[1] International Center for Tropical Agriculture, Cali, Colombia, [2] Centre for Crop Systems Analysis, Wageningen University, Wageningen, Netherlands, [3] International Potato Center, CGIAR Research Program on Roots, Tubers and Bananas, Lima, Peru, [4] School of Biosciences, University of Birmingham, Birmingham, UK, [5] Department of Biological Sciences, Southeastern Louisiana University, Hammond, LA, USA, [6] Department of Plant Sciences, University of Oxford, Oxford, UK, [7] Department of Plant Biology, University of Georgia, Athens, GA, USA, [8] Plant Genetic Resources Conservation Unit, United States Department of Agriculture – Agricultural Research Service, Griffin, GA, USA, [9] Department of Horticultural Science, North Carolina State University, Raleigh, NC, USA, [10] Department of Horticulture and Crop Science, The Ohio State University, Columbus, OH, USA

Crop wild relatives of sweetpotato [*Ipomoea batatas* (L.) Lam., *I.* series *Batatas*] have the potential to contribute to breeding objectives for this important root crop. Uncertainty in regard to species boundaries and their phylogenetic relationships, the limited availability of germplasm with which to perform crosses, and the difficulty of introgression of genes from wild species has constrained their utilization. Here, we compile geographic occurrence data on relevant sweetpotato wild relatives and produce potential distribution models for the species. We then assess the comprehensiveness of *ex situ* germplasm collections, contextualize these results with research and breeding priorities, and use ecogeographic information to identify species with the potential to contribute desirable agronomic traits. The fourteen species that are considered the closest wild relatives of sweetpotato generally occur from the central United States to Argentina, with richness concentrated in Mesoamerica and in the extreme Southeastern United States. Currently designated species differ among themselves and in comparison to the crop in their adaptations to temperature, precipitation, and edaphic characteristics and most species also show considerable intraspecific variation. With 79% of species identified as high priority for further collecting, we find that these crop genetic resources are highly under-represented in *ex situ* conservation systems and thus their availability to breeders and researchers is inadequate. We prioritize taxa and specific geographic locations for further collecting in order to improve the completeness of germplasm collections. In concert with enhanced conservation of sweetpotato wild relatives, further taxonomic research, characterization and evaluation of germplasm, and improving the techniques to overcome barriers to introgression with wild species are needed in order to mobilize these genetic resources for crop breeding.

Keywords: crop diversity, crop improvement, crop wild relatives, food security, gap analysis, plant genetic resources

Introduction

Sweetpotato [*Ipomoea batatas* (L.) Lam.] counts among the world's most important root crops, grown on at least eight million hectares in 114 countries worldwide, with particular significance to food supplies in the tropics and subtropics of East and Southeast Asia, and Sub-Saharan Africa. The success of the crop in these regions is due to its adaptability to a wide range of agro-ecological conditions, ease of propagation from cuttings, low input cultivation requirements, and high productivity as well as nutritive value (Woolfe, 1992). The crop can be cultivated from humid to semi-arid conditions, from sea level to 3000 m.a.s.l. (Huaman, 1987), and can translocate photosynthetic products to the storage roots throughout the growing season, thereby mitigating negative effects of temporary adverse conditions (Kays, 1985). Sweetpotato produces among the highest amounts of edible energy per hectare of all major food crops (De Vries et al., 1967), and is an important source of vitamin A and C, calcium, iron, and a number of essential amino acids (Kays and Kays, 1998; Tumwegamire et al., 2011). In addition to human consumption of the storage roots and young leaves, the crop is utilized for animal feed, fuel, and starch production.

In Sub-Saharan Africa, sweetpotato is predominantly cultivated in small plots characterized by low fertility and drought-prone soils, producing relatively good yields with low inputs and minimal labor costs. The crop has recently become the focus of targeted bio-fortification for enhanced vitamin A. Orange-fleshed varieties have been bred with 50-fold more β-carotene than standard varieties and these newly released varieties rank first among roots and tubers in Sub-Saharan Africa for their nutritional quality (Low et al., 2007; Hotz et al., 2012). Given its adaptability, low-external input requirements, nutritional quality, and improvement potential, it is not surprising that sweetpotato has become a priority in crop based strategies for enhancing food security in the tropics (Pfeiffer and McClafferty, 2007; Bill and Melinda Gates Foundation, 2011; Bouis and Islam, 2012).

The full potential of sweetpotato is far from realized, with particularly large yield gaps (ca. 20 t ha^{-1}) remaining across rain-fed Sub-Saharan Africa due to a range of biotic and abiotic constraints, especially sweetpotato virus disease (SPVD) and sweetpotato weevils (SPW), as well as susceptibility to drought (Sutherland, 1986; Valverde et al., 2007; Ngailo et al., 2013; Kivuva et al., 2015a,b). SPVD is a severe constraint on the continent, caused by the synergistic interaction of two viruses transmitted by whiteflies and aphids, causing yield losses of up to 98% under severe infections (Ngailo et al., 2013). SPW (*Cylas* spp.) are the most devastating insect pests of the crop. *Cylas formicarius elegantulus* Summers, *C. puncticollis* Boheman and *C. brunneus* Fabricius can cause yield losses of between 67 and 100% in Sub-Saharan Africa (Smit, 1997). The concealed feeding behavior, oviposition, and larval development of the weevils in the storage roots make their control very difficult, necessitating the development of improved management options, in particular via enhanced genetic resistance. Drought, and the compounding effect of increasing heat on drought, is a rising concern particularly in regions undergoing significant climatic change, both due

to its direct effect on productivity (Low et al., 2009; Schafleitner et al., 2010) as well as to its association with increased severity of damage from SPW and SPVD (Munyiza et al., 2007; Mwololo et al., 2007). Lack of drought tolerance in high β-carotene sweetpotato varieties has led to constraints in their adoption (Mwanga and Ssemakula, 2011).

Crop wild relatives (CWR) are increasingly being recognized for their potential to contribute valuable traits to breeding programs (Feuillet et al., 2008; Guarino and Lobell, 2011; Dempewolf et al., 2014). CWR have provided breeders with genes for pest and disease resistance, abiotic stress tolerance, and quality traits in an ever increasing number of food crops, such as banana, barley, bean, cassava, chickpea, maize, lettuce, oat, potato, rice, sugarcane, sunflower, tomato, and wheat, among others (Xiao et al., 1996; Hajjar and Hodgkin, 2007; McCouch et al., 2007; Khoury et al., 2010). Yet despite the successful history of contribution to the improvement of major crops, systematic approaches to the use of CWR in the breeding programs of a number of important staples, including sweetpotato, remain underdeveloped.

The morning glory tribe Ipomoeeae contains ~650–900 species and includes the genus *Ipomoea* and nine other related genera (Wilkin, 1999; Mabberley, 2008). Although many genera, subgenera, and sections of the Ipomoeeae are not monophyletic in phylogenetic analyses, *Ipomoea* series *Batatas* (Choisy) D. F. Austin, which contains sweetpotato and 14 closely related CWR (Austin, 1978; McDonald and Austin, 1990), does form a monophyletic lineage (Miller et al., 1999; McDonald et al., 2011; Eserman et al., 2014). These species include wild *I. batatas* (L.) Lam. [including *I. batatas* var. *apiculata* (Martens and Galeotti) McDonald and Austin], *Ipomoea cordatotriloba* Dennstedt, *Ipomoea cynanchifolia* Meisn., *Ipomoea grandiflora* (Dammer) O'Donell, *Ipomoea lacunosa* L., *Ipomoea leucantha* Jacquin, *Ipomoea littoralis* Blume, *Ipomoea ramosissima* (Poir.) Choisy, *Ipomoea splendor-sylvae* House, *Ipomoea tabascana* McDonald and Austin, *Ipomoea tenuissima* Choisy, *Ipomoea tiliacea* (Willd.) Choisy in D. C., *Ipomoea trifida* (H. B. K.) G. Don, and *Ipomoea triloba* L.

Many sweetpotato CWR can be hybridized with the crop through controlled pollinations, somatic cell, and/or ovule culture techniques (Diaz et al., 1996). Crosses involving *I. tabascana*, *I. trifida*, *I. triloba*, *I. littoralis*, *I. grandifolia*, *I. lacunosa*, *I. leucantha*, and wild *I. batatas* in particular have resulted in relatively viable progeny (Nimmakayala et al., 2011). The wild conspecific as well as *I. trifida* have been documented for their contribution to increases in protein and starch content, and nematode and SPW resistance (Iwanaga, 1988; Shiotani et al., 1994), although there is uncertainty for some material as to whether they may actually have been feral forms of the cultivar (Nimmakayala et al., 2011). Species that have been explored for potential traits of use in crop improvement include *I. trifida* and *I. littoralis* for yield and SPW, scab [*Elsinoë batatas* (Saw.) Viegas et Jenkins], and black rot disease (*Ceratocystis fimbriata* Ell. et Halst.) resistance; *I. grandifolia* for sweetpotato stem nematode and SPVD resistance; and *I. triloba* for drought tolerance, root rot resistance, and foliar fungal disease resistances (Iwanaga, 1988; Jarret et al., 1992; Komaki, 2004; Zhang and Liu, 2005; Nimmakayala et al., 2011). Challenges in the creation of viable progeny between the

CWR and the cultivated species are not insignificant, though, due to differences in ploidy and interspecific incompatibility (Martin, 1970, 1982; Teramura, 1979; Lu and Li, 1992; Shiotani et al., 1994; Diaz et al., 1996; Komaki, 2004; Nimmakayala et al., 2011).

A lack of basic knowledge about boundaries between species within *I.* series *Batatas* and a dearth of diagnostic characters enabling differentiation of taxa – to facilitate reliable and accurate species identification – is a fundamental stumbling block constraining the utilization of sweetpotato CWR (Austin, 1978, 1988; Jarret et al., 1992; Diaz et al., 1996; McDonald et al., 2011; Nimmakayala et al., 2011; Eserman et al., 2014). Needed studies have been delayed in part due to the absence of plant materials for research, in particular the availability of specimens with flowers and ripe fruits. Studies that have been performed have often been based upon limited sampling (e.g., single accessions for *I. littoralis* and *I. tabascana*).

Many unanswered questions regarding the relationships of CWR to sweetpotato potentially impact the efficiency of breeding strategies for the crop. For example, do species such as *I. tabascana* and *I. tenuissima* represent distinct taxa (hybrid or otherwise) or rather rare variants of the crop? What is the range and genetic diversity present in truly wild forms of the crop conspecific, compared to feral escapees? How accurate are the classifications of species with highly disjunct distributions (e.g., *I. cordatotriloba*)? Are there as yet unrecognized cryptic species within *I.* series *Batatas* [e.g., *I.* 'austinii' (Duncan and Rausher, 2013)]? What are the lineages and genetic resources potential of purported hybrid species (i.e., *I. leucantha* and *I. grandifolia*)? What are the geographic locations of new variation being generated through hybridization among sweetpotato CWR?

The investigation, conservation, and availability of genetic resources of sweetpotato provide a foundation for the crop's long term viability and for its potential for improvement. To contribute to these objectives, we analyzed the comprehensiveness of *ex situ* conservation of sweetpotato CWR through a series of questions: (a) what constitutes a potentially useful wild relative of sweetpotato?, (b) where are these species encountered?, (c) what is the state of conservation and availability of these species to researchers, and what are the highest taxonomic and ecogeographic priorities for further collecting? And finally, (d) what traits do sweetpotato CWR possess that may be valuable to crop improvement?

Materials and Methods

Identification of Target CWR Species and Occurrence Data Compilation

The CWR of sweetpotato analyzed in this study were selected based upon recent and historical taxonomic and molecular phylogenetic research (Austin, 1978, 1997; Austin and Huamán, 1996; Miller et al., 1999; McDonald et al., 2011; Germplasm Resources Information Network [GRIN], 2013; Eserman et al., 2014), identifying those wild species with a relatively close phylogenetic relationship to the crop (i.e., members of *I.* series *Batatas*). We included all 14 wild species comprising the series in the analysis.

Domesticated sweetpotato *I. batatas* (6x) has been proposed as originating from interspecific hybridization involving *I. trifida* (2x, 4x, 6x), *I. littoralis* (2x), and/or *I. leucantha* (2x) (Nishiyama, 1971, 1982; Austin, 1988). The species most closely related to the crop have been posited to be *I. trifida* followed by *I. tabascana* (4x) (Srisuwan et al., 2006; Nimmakayala et al., 2011). Following the genepool concept of Harlan and de Wet (1971) and due to ploidy incompatibility with the cultivated species, the putative closest related taxa to sweetpotato are placed in the secondary genepool: wild forms of *I. batatas* (4x), *I. trifida*, *I. littoralis*, and *I. tabascana* (Jarret et al., 1992; Jarret and Austin, 1994; Rajapakse et al., 2004; Germplasm Resources Information Network [GRIN], 2013). Species classified as tertiary wild relatives include: *I. cordatotriloba* (syn. *Ipomoea trichocarpa* Elliott), *I. cynanchifolia*, *I. grandifolia*, *I. lacunosa*, *I. leucantha*, *I. ramosissima*, *I. splendor-sylvae* (syn. *Ipomoea umbraticola* House), *I. tenuissima*, *I. tiliacea*, and *I. triloba* (Jarret and Austin, 1994; Huang et al., 2002; Rajapakse et al., 2004; Germplasm Resources Information Network [GRIN], 2013).

Occurrence records for these species were acquired from online biodiversity, herbaria, and germplasm databases; through communications with herbaria and genebank managers, and other crop researchers; and via direct recording of provenance data during visits to selected herbaria (Supplementary Table S1). Germplasm data were obtained from repositories that provide straightforward access to genetic resources and associated data to the global research community through online information systems. The occurrence data were then compiled in a standardized format, nomenclature was checked against Germplasm Resources Information Network [GRIN] Taxonomy for Plants (2013) and The Plant List (2010), and duplicate records were eliminated. Existing coordinates were cross-checked to country and being on land (Hijmans et al., 1999), and records with locality information but no coordinates were georeferenced using the Google Maps Geocoder (2013) v.3 application programming interface. Occurrence data were mapped, iteratively evaluated for correctness, and further processed in order to form a final dataset of improved taxonomic and spatial accuracy.

Challenges in using and in improving the large quantities of occurrence data now available from online resources such as the Global Biodiversity Information Facility (GBIF) have been noted (Gaiji et al., 2013), including geographic and nomenclatural data quality and the slow speed with which aggregated datasets are updated (Mesibov, 2013; Otegui et al., 2013; Hjarding et al., 2014). In addition, particular caution must be applied to the occurrence records used in the current paper as ongoing work (unpublished data) indicates that many *Ipomoea* occurrence records in such online databases are identified as synonyms, excluded or invalid names, and that many valid names were applied to specimens well outside of species known ranges. We have identified some of these obvious errors but until all specimen records are correctly identified and checked against an accurate taxonomy these data must be treated with caution.

A total of 5,614 occurrence records for the 14 taxa were included in potential distribution modeling and/or in the conservation analysis, including 749 germplasm records sourced from four genebanks, and 4,865 herbarium and other occurrence reference records sourced from 42 providers. Records per species ranged from eight (*I. tabascana*) to 1,409 (*I. trifida*). Of these, 3,650 records containing unique cross-checked coordinates were used to model species potential distributions and to locate the original collecting site of existing germplasm accessions.

Species Potential Distribution Modeling

A potential distribution model for each species was calculated using the maximum entropy (Maxent) algorithm (Phillips et al., 2006), with a set of ecogeographic variables and unique species presence records as inputs. We chose Maxent due to its extensive application in predicting species distributions (Elith et al., 2006; Phillips and Dudik, 2008; Costa et al., 2010), including those for wild relatives (Ramírez-Villegas et al., 2010; Conolly et al., 2012; Khoury et al., 2015). We performed modeling at a resolution of 2.5 arc-minutes (~5 km^2 cell size at the equator), employing 10,000 background points for model training over the combined distributional range of the sweetpotato CWR. Ecogeographic inputs included altitude and 19 bioclimatic variables from the WorldClim database (Hijmans et al., 2005), and seven major edaphic drivers of plant species distributions with consistent data coverage throughout the range of the sweetpotato CWR species, obtained from ISRIC-World Soil Information (Hengl et al., 2014; Supplementary Table S2). For the edaphic variables we calculated a weighted mean across 0–5, 5–15, 15–30, 30–60, and 60–100 cm soil depth values in order to derive a single data value for 0–100 cm. We then resampled the 30 arc-seconds resolution data to form 2.5 arc-minutes inputs aligned with the WorldClim datasets.

Potential distribution models were produced by calculating the mean of replicates ($k = 5$), and clipped by measuring the shortest distance between the receiver operating characteristic curve (ROC-curve) and the top-left corner of the plot (Liu et al., 2005). Models were constrained per species by a native range defined at the country level as given in GRIN Taxonomy for Plants (Germplasm Resources Information Network [GRIN], 2013), in order to focus prioritization recommendations on those regions with populations exhibiting long-term adaptations to local ecogeographic conditions. We further cross-validated and refined occurrence data based upon our knowledge of native distributions, constraining localities for wild *I. batatas* to Mexico south to Peru, and not in the Caribbean; *I. leucantha* to the USA and Mexico, and *I. littoralis* to points within 100 km of the ocean.

In order to derive robust distribution models for each species, we analyzed Maxent results across three groups of ecogeographic inputs: (a) the full set of 19 bioclimatic variables (Ramírez-Villegas et al., 2010); (b) the bioclimatic variables, altitude, and the additional set of seven edaphic variables, totaling 27 input variables; and (c) a species-specific derivation of the most important drivers of distribution based upon presence data, further refined by removing highly correlated variables. For the ecogeographic variables in the species-specific subset method, we utilized a non-linear iterative partial least squares (NIPALS) algorithm to perform a principal-component analysis (PCA), as NIPALS has the potential to handle data arrays in which the number of observations is less than the number of input variables. We then identified those variables with the greatest contribution (> 0.7 or < -0.7) to the first two principal components per species. Finally, we used a variance inflation factor (VIF) to identify the variables with a low degree of collinearity (see Supplementary Table S3 for a list of variables identified per species). A calibrated area under the ROC curve (cAUC) was obtained to assess the predictive performance of each model (Hijmans, 2012; Khoury et al., 2015).

The three modeling methods were compared against null models, and the species-specific subset variables method showed the least overall spatial sorting bias among methods (Spearman's p for the 19 variables was 0.65, for 27 variables it was 0.72, and for the subset method it was 0.25), although the differences in median AUC distributions across species for each method were not significant ($p = 0.095$) through a Kruskal–Wallis non-parametric analysis of variance test. Maxent models were subjected to a four-fold assessment process including: (a) the fivefold average area under the ROC curve of test data (ATAUC), b) the standard deviation of the test AUC of the five different folds (STAUC), (c) the proportion of the potential distribution coverage with SD above 0.15 (ASD15), and (d) the cAUC value. Models with ATAUC above 0.7, STAUC below 0.15, ASD15 below 10%, and cAUC exceeding 0.40 were considered accurate and stable (Ramírez-Villegas et al., 2010; Khoury et al., 2015).

The potential distribution models of the sweetpotato CWR generally performed well in regard to the modeling assessment process. Four species demonstrated low cAUC values and one of these an elevated ASD15 value, indicating greater uncertainty in the models (Supplementary Table S3). Species-specific subset model outputs for taxa with relatively few distinct occurrence points (< 20; *I. tabascana*, *I. tenuissima*, and *I. cynanchifolia*) lacked sufficient discriminatory power, leading to highly inflated spatial models in comparison to recorded distributions. Potential distribution models for these species were resolved by deriving an ensemble (i.e., overlap) of the three input variation methods, verified by researchers knowledgeable in the distribution of the species as more accurately representing the true known distributions. Potential distribution models based upon the species-specific subset variables method were therefore utilized in subsequent analyses for all species with adequate distinct occurrence points (> 20). The ensemble method was used for the three species with limited occurrence data.

Analysis of Current *Ex Situ* Conservation and Further Collecting Needs for CWR

We adapted a gap analysis methodology proposed by Ramírez-Villegas et al. (2010), combining three metrics used to assess the urgency of further collecting in order to fill gaps in *ex situ* conservation of CWR. The total sample representation of each species in genebank collections was estimated via a sampling representativeness score (SRS), calculated as the number of germplasm

samples (G) divided by the total number of samples [G + herbarium samples (H); i.e., all other records aside from available genebank accessions].

The sufficiency of geographic coverage of germplasm collections of each species was estimated through a geographic representativeness score (GRS), calculated as the share of the combined total area of circular buffers of 50 km (CA50) placed around existing germplasm collection points compared with the overall potential geographic distribution of the species.

The comprehensiveness of ecological coverage of germplasm collections of each species was estimated through an ecological representativeness score (ERS), calculated by estimating the distinct ecosystem classifications (Olson et al., 2001) represented in the CA50 of existing germplasm collection points compared with the diversity of ecosystems in which the overall potential geographic distribution model of the species occurs.

A final priority score (FPS) for further collecting for *ex situ* conservation was assigned to each species by averaging the three gap analysis metrics (SRS, GRS, and ERS). FPS scores were further classified into four categories of urgency for collecting: high priority species (HPS) for taxa whose $0 < FPS \leq 2.5$ or when no germplasm accessions are currently conserved, medium priority species (MPS) when $2.5 < FPS \leq 5$, low priority species (LPS) when $5 < FPS \leq 7.5$, and 'no further collecting recommended' (NFCR) when $7.5 < FPS \leq 10$. We produced collecting priorities maps for all species, displaying the geographic areas that have not yet been collected from within the potential distributions of taxa.

The ecogeographic data preparation, species distribution modeling, and gap analysis were written and performed in R v2.15.1 (R Core Team, 2013), utilizing packages maptools (Bivand and Lewin-Koh, 2014), rgdal (Bivand et al., 2014), SDMTools (van der Wal et al., 2014), raster (Hijmans, 2014), sp (Pebesma and Bivand, 2005; Bivand et al., 2013), dismo (Hijmans et al., 2013), and plsdepot (Sanchez, 2012). Resulting spatial files were mapped in ArcMap v.10 (ESRI, 2011). Collecting priorities spatial files were analyzed using the Zonal Statistics tool in ArcMap to list the countries prioritized for further collecting for *ex situ* conservation.

Expert Evaluation of Conservation Assessment Results

In order to validate and/or expose deficiencies in our findings, we subjected the gap analysis numerical and spatial results to an evaluation performed by seven crop experts with experience in the systematics, distribution, and/or conservation status of CWR of sweetpotato: Richard E. Miller, Southeastern Louisiana University; Robert W. Scotland and John R. I. Wood, University of Oxford; Genoveva Rossel, International Potato Center; Lauren A. Eserman, University of Georgia; Robert L. Jarret, USDA – ARS Griffin; and G. Craig Yencho, North Carolina State University. These experts were first asked to provide an evaluation of the sufficiency of germplasm collections per species based only upon their knowledge of the total number of accessions conserved, and geographic as well as environmental gaps. Such an assessment

[comparable expert priority score (EPS)] was considered directly comparable to the FPS of the gap analysis results.

A second evaluation score (contextual EPS) based on the entirety of expert knowledge, including additional factors such as threats to species *in situ* and prioritization by usefulness in crop breeding, was collected in order to provide additional information to collecting prioritization efforts. In both cases, an EPS between 0 and 10, aligned with the gap analysis results prioritization scale, was requested. After these steps, experts were shown the gap analysis data and asked to comment on assessed quantitative results, occurrence data, potential distribution models, and maps of collecting priorities. Following these contributions by experts, input occurrence data were further refined by eliminating clearly incorrect points and adjusting country-level native areas, and the potential distribution modeling and gap analyses were re-run using the refined datasets in order to improve the quantitative and spatial outputs. Expert evaluation metrics displayed in the results pertain to a final evaluation of improved gap analysis outputs, performed by five of the experts.

A multiple factor analysis was used in order to compare the various forms of expert evaluation inputs (i.e., comparable expert priority score, contextual expert priority score, evaluation of gap analysis results score, evaluation of occurrence data, evaluation of potential species distribution models, and evaluation of collecting priorities map) with the gap analysis results. An expert evaluation index was created, which estimated the degree of accord between all experts and the gap analysis results for each species, with a scale from 0 (disagreement) to 100 (agreement). Analyses were performed using R package FactoMineR (Husson et al., 2009).

Identification of Ecogeographic Characteristics of CWR

In order to evaluate the pairwise degree of geographic overlap between sweetpotato CWR distribution models, we calculated an overlap measure equal to the frequency of shared 2.5 arc-minute geographic cells between taxa divided by the sum of the total number of cells in which either of the species are present (Kernohan et al., 2001; Fieberg and Kochanny, 2005). To assess the pairwise degree of ecogeographic niche overlap between species, we used Schoener's index for niche similarity (D) and the adjusted similarity index (I) as outlined in Warren et al. (2008), using species distribution models and the species-specific subset of the 27 ecogeographic layers/ensemble models as inputs. Overlap indices range from 0 (no overlap) to 1 (complete overlap) and were implemented in the R package Phyloclim (Heibl, 2011).

We utilized ecogeographic information in combination with species presence data in order to identify populations of species with the potential for outstanding adaptations to climatic and/or edaphic conditions of interest to sweetpotato breeding objectives. We assessed the relative importance of the 27 ecogeographic variables (Supplementary Table S2) in explaining the total variation through a PCA after confirming the validity of its use through a Bartlett's test performed in R package psych (Revelle, 2015). We created a hierarchical cluster of principal components (HCPC) in order to identify ecogeographic clusters for the sweetpotato

wild relatives, driven by those variables demonstrating ≥15% difference (±) from the mean and a reduction of ≥20% from the mean standard deviation exhibited across all species, using R package FactoMineR. Boxplots for each of the 27 ecogeographic variables were created based upon CWR species occurrence data points, displaying the median and variance parameters per species per variable. Comparable ecogeographic variable data for the sweetpotato crop was extracted from area of cultivation maps (Monfreda et al., 2008) at a resolution of 5 arc-minutes, with a random sample of 1,000 points weighted by harvested area, taken from the major cultivation areas in Asia, Africa, and Latin America.

Results

Distributions of the Wild Relatives of Sweetpotato

Sweetpotato CWR were modeled to occur from the central USA to northern Argentina in the Americas, including the Caribbean (Supplementary Figure S1). Species richness was greatest in central Mexico through Central America to the northern Andean region, and in the Southeastern USA, with up to nine species potentially overlapping in Mexico from the states of Veracruz through the Yucatan peninsula (**Figure 1**). The Mexican and Central American regions of distribution represent one of the posited centers of origin and primary diversity of cultivated sweetpotato (Austin, 1988; Austin and Huamán, 1996; Zhang et al., 2000; Gichuki et al., 2003; Roullier et al., 2013). Northwestern South America, with archeological remains of cultivated sweetpotato from Peru dating back to 8,000 years BP, which are among the oldest recorded domestication events on the continent (Piperno and Pearsall, 1998; Solis et al., 2001), displayed a considerably lesser but still notable degree of CWR species richness. One Old World species (*I. littoralis*; Austin, 1991) was also modeled to occur in coastal areas of Madagascar, South and Southeast Asia, Australia, and the Pacific region.

Analysis of Current *Ex Situ* Conservation and Further Collecting Needs for CWR

Eleven out of the fourteen CWR species were assigned high priority for further collecting due to insufficient genebank accessions in comparison to the total number of occurrence samples (SRS), and to large geographic (GRS) and ecological (ERS) gaps in *ex situ* germplasm collections in comparison to the full potential distributions of the species (**Figure 2**, **Table 1**, Supplementary Figure S2). Six of these taxa are currently represented by ≤10 accessions conserved *ex situ*, moreover, with few exceptions these accessions lack associated geographic occurrence information (Supplementary Table S3). Total sampling representativeness and geographic coverage of species in germplasm collections were particularly lacking for taxa assessed high priority, while gaps in ecological representativeness were less extreme for some species (e.g., *I. cordatotriloba*, *I. triloba*, and *I. splendor-sylvae*). *Ipomoea grandifolia* and *I. trifida* were assessed as relatively well covered in regard to ecosystems represented *ex situ*, which raised their FPS into the medium priority for further collecting category.

Ipomoea tabascana was assessed as of low priority for further collecting due to existing germplasm collections largely covering its very restricted distribution in southern Mexico, resulting in a high GRS score. The mean FPS across all CWR was 1.75 ± 1.82.

Paralleling the distribution of species richness of sweetpotato CWR, the regions identified for further collecting of the greatest number of species included central and southern Mexico and the Southeastern USA, with up to seven species prioritized for further collecting occurring in the same area (**Figure 3**). Further collecting priorities were recognized in a total of 50 countries throughout the range of the genepool (Supplementary Figure S3, Supplementary Table S4). Occurrence data, potential distribution models, and collecting priorities maps for all modeled species are available in an interactive map format at http://www.cwrdiversity.org/distribution-map/.

Expert Evaluation of Conservation Assessment Results

The average of the directly comparable EPS across experts was 2.65 (±1.10) as a mean across species, varying 1.95 points on average from the FPS, with seven taxa designated by the experts as HPS, and seven as MPS (**Table 1**, **Figure 2**). For most species, this mean was highly influenced by one or two experts' giving species considerably less priority than the other experts. Species with closest accord between the gap analysis results and the comparable expert analysis included *I. splendor-sylvae*, *I. trifida*, *I. tenuissima*, *I. littoralis*, *I. tileacea*, *I. ramosissima*, and *I. grandifolia*. Taxa with the greatest difference between gap analysis and comparable expert values included *I. tabascana* and *I. leucantha*.

Two species were assessed by the experts as higher priority for further collecting than the results of the gap analysis. Although *ex situ* collections for the highly restricted distribution of *I. tabascana* were determined in the gap analysis to be fairly comprehensive (LPS), the experts assigned the species HPS for further collecting due to its very limited overall number of germplasm holdings, and to the difficulty in producing viable seed in *ex situ* conservation. *Ipomoea grandifolia* was assessed in the gap analysis as reasonably comprehensively represented in regard to ecosystem diversity, and thus assigned medium priority for further collecting, while the experts gave moderately higher priority to the species.

The contextual expert priority score per species, which also included the expert's knowledge of threats *in situ* as well as usefulness for crop improvement, did not vary widely from the comparable score [mean across all experts and species was 2.76 (±0.85); mean difference between comparable and contextual expert scores across all species and experts was 0.11]. Due largely to differences between the perceptions of relative sufficiency in regard to the overall number of germplasm accessions by the experts versus gap analysis results (e.g., for *I. tabascana*, *I. triloba*, and *I. leucantha*), the comparable and contextual assessments did not correlate well with the gap analysis results for the genepool as a whole (Supplementary Figures S4A,B).
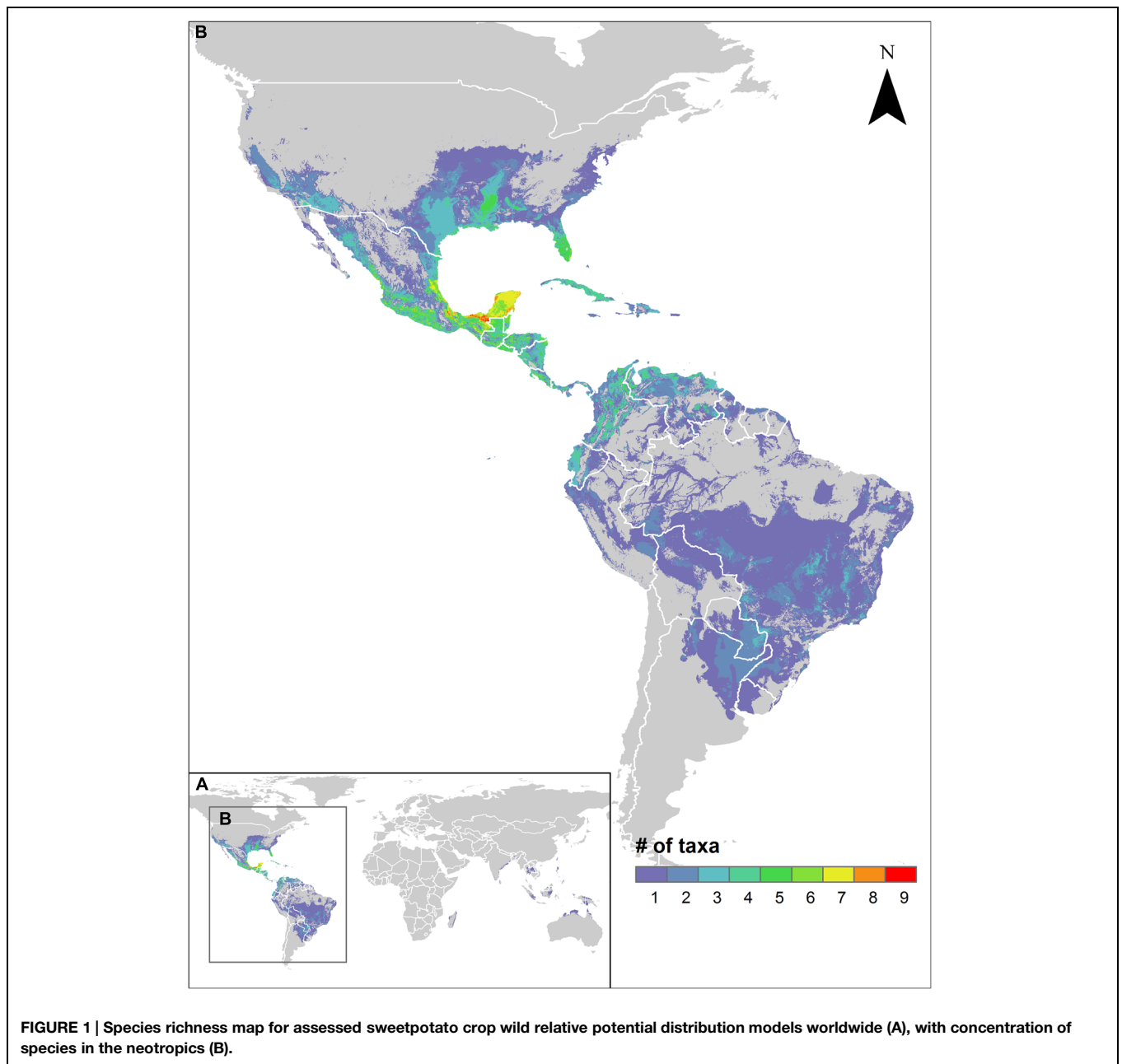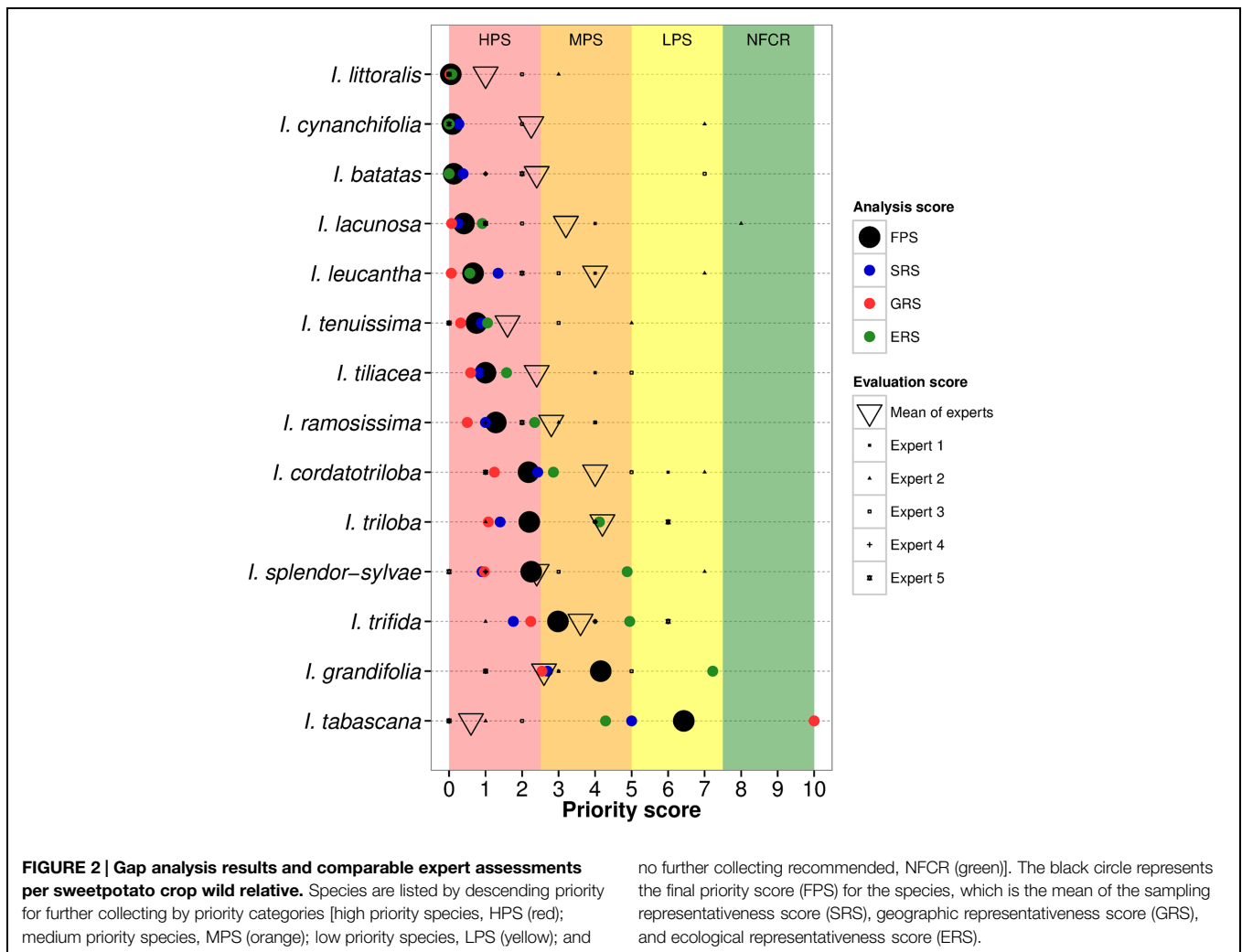
FIGURE 1 | Species richness map for assessed sweetpotato crop wild relative potential distribution models worldwide (A), with concentration of species in the neotropics (B).

The multiple factor analysis combining the comparable expert priority score, contextual expert priority score, evaluation of gap analysis results score, evaluation of occurrence data, evaluation of potential species distribution models, and evaluation of collecting priorities map, revealed sufficient agreement among experts and variables to justify the use of the expert evaluation index, although variation in expert option was notable for many species. Data inputs and resulting distribution models were generally assessed positively as a mean across experts, with eight species receiving very positive index values. Those species with the highest accord between all variables and experts and the gap analysis results included *I. littoralis*, *I. spendor-sylvae*, *I.*

ramosissima, *I. tenuissima*, and *I. tiliacea*. Those species with the lowest accord included *I. tabascana*, *I. triloba*, and *I. grandifolia* (Supplementary Figure S4). FPS results were particularly influenced by spotty occurrence records for the majority of species, with gaps recognized by the experts.

Identification of Ecogeographic Characteristics of CWR

The analysis of geographic and ecogeographic accord between pairwise potential species distribution models segregated species well into temperate North American (e.g., *I. lacunosa*, *I. tenuissima*), Mesoamerican (e.g., *I.*

FIGURE 2 | Gap analysis results and comparable expert assessments per sweetpotato crop wild relative. Species are listed by descending priority for further collecting by priority categories [high priority species, HPS (red); medium priority species, MPS (orange); low priority species, LPS (yellow); and no further collecting recommended, NFCR (green)]. The black circle represents the final priority score (FPS) for the species, which is the mean of the sampling representativeness score (SRS), geographic representativeness score (GRS), and ecological representativeness score (ERS).

splendor-sylvae, *I. tabascana*), widely distributed tropical (e.g., *I. triloba*, *I. trifida*), and South American (e.g., *I. grandifolia*, *I. cynanchifolia*) areas (Supplementary Figure S5).

Strong linear relationships were found between bioclimatic variables within the study area, justifying the application of the PCA, with 75.6% of variance explained through four principal components (Supplementary Figure S6). The first principal component (32% of variation) was correlated with temperature extremes and fluctuation. The second component (21.6% of variation) was most occupied by precipitation variables related to drought. The third component (13.9% of variation), was related with altitude, and the final component (8.1%) with soil texture characteristics.

Species occurrence data segregated into temperate and tropical ecogeographic clusters, with the great majority of species' distributions strongly occurring within a single cluster. Ecogeographic variables most strongly influencing the definition of the temperate cluster (1) included those associated with temperature variation and relatively low precipitation, elevation,

and soil organic matter. The most determinant variables in the tropical cluster (2) were related to relatively high and consistent temperatures. Those species displaying a notable proportion of occurrences within both clusters included *I. cynanchifolia* and *I. triloba*, and to a lesser degree *I. leucantha* and *I. ramosissima* (**Table 1**, Supplementary Figure S6).

Ecogeographic niche assessments of sweetpotato CWR based upon occurrence data points revealed large differences in potential adaptation to temperature, precipitation, and edaphic characteristics (**Table 1**, **Figure 4**, Supplementary Figure S7). Such adaptation for many species fell well outside the modeled niche of the cultivated species, particularly for high temperatures both in wet and dry conditions, as well as high precipitation. Species of notable adaptation to high mean annual, monthly, and quarterly temperatures included *I. littoralis*, *I. tabascana*, *I. trifida*, *I. leucantha*, *I. tiliacea*, *I. tenuissima*, *I. triloba*, *I. splendor-sylvae*, and wild *I. batatas*. USA species *I. lacunosa* stood out for adaptation to low temperatures, with *I. grandifolia*, *I. cordatotriloba*, *I. tenuissima*, and *I. ramossissima* also demonstrating cold tolerance. These same species were among those displaying the widest adaptation to temperature fluctuation.

TABLE 1 | Utilization characteristics, number of germplasm accessions conserved *ex situ*, further collecting priorities, and potential adaptive traits associated with ecogeographic niches of sweetpotato crop wild relatives.

Taxon	Gene pool	Ploidy	Germplasm accessions	Gap analysis priority	Mean expert priority	Eco geographic cluster	Potential adaptation to
Ipomoea batatas	2	4x = 60	4 (0)	HPS	HPS	2	Heat, high precipitation, drought, precipitation seasonality, clay soils
I. cordatotriloba	3	2x, 4x	103 (67)	HPS	MPS	1	Cold, temperature variation, clay soils, sandy soils
I. cynanchifolia	3	2x = 30	1 (0)	HPS	HPS	1,2	Drought, precipitation seasonality, sandy soils
I. grandifolia	3	2x = 30	124 (83)	MPS	MPS	1	Cold, temperature variation, clay soils, sandy soils
I. lacunosa	3	2x = 30	10 (1)	HPS	MPS	1	Cold, temperature variation, drought
I. leucantha	3	2x = 30	18 (15)	HPS	MPS	1,2	Heat, drought, precipitation seasonality, sandy soils
I. littoralis	2	2x = 30	2 (2)	HPS	HPS	2	Heat, high precipitation, drought, precipitation seasonality, sandy soils
I. ramosissima	3	2x = 30	34 (30)	HPS	MPS	2,1	Cold, high precipitation, clay soils
I. splendor-sylvae	3	2x = 30	16 (9)	HPS	HPS	2	Heat, high precipitation, drought, precipitation seasonality, clay soils
I. tabascana	2	4x = 60	4 (2)	LPS	HPS	2	Heat, high precipitation, clay soils
I. tenuissima	3	2x = 30	3 (1)	HPS	HPS	1	Heat, cold, temperature variation, sandy soils
I. tiliacea	3	4x = 60	61 (44)	HPS	HPS	2	Heat, high precipitation, clay soils
I. trifida	2	2x,3x,4x,6x	248 (159)	MPS	MPS	2	Heat, high precipitation, drought, precipitation seasonality
I. triloba	3	2x = 30	121 (74)	HPS	MPS	2,1	Heat, drought

Ploidy data adapted from Nimmakayala et al. (2011). Germplasm accessions displays both the total number of accessions recorded in genebanks, as well as the number of accessions with unique geographic coordinates (i.e., unique populations) in parenthesis. HPS = high, MPS = medium, and LPS = low priority species for further collecting.

Crop wild relative of sweetpotato occurring in areas of high precipitation included *I. ramosissima, I. littoralis, I. splendor-sylvae, I. tabascana, I. tiliacea, I. trifida,* and wild *I. batatas. Ipomoea littoralis, I. trifida, I. splendor-sylvae, I. leucantha, I. cynanchifolia, I. triloba, I. lacunosa,* and wild *I. batatas* were distributed in regions with low precipitation. These species were also among those displaying the widest adaptation to precipitation seasonality. Sweetpotato CWR also displayed variation in adaptation to edaphic characteristics. *Ipomoea tabascana, I. grandifolia, I. tiliacea, I. splendor-sylvae, I. ramosissima, I. cordatotriloba,* and wild *I. batatas* occurred in clay soils, while *I. tenuissima, I. littoralis, I. cynanchifolia, I. grandifolia, I. leucantha,* and *I. cordatotriloba* were distributed in sandy soils.
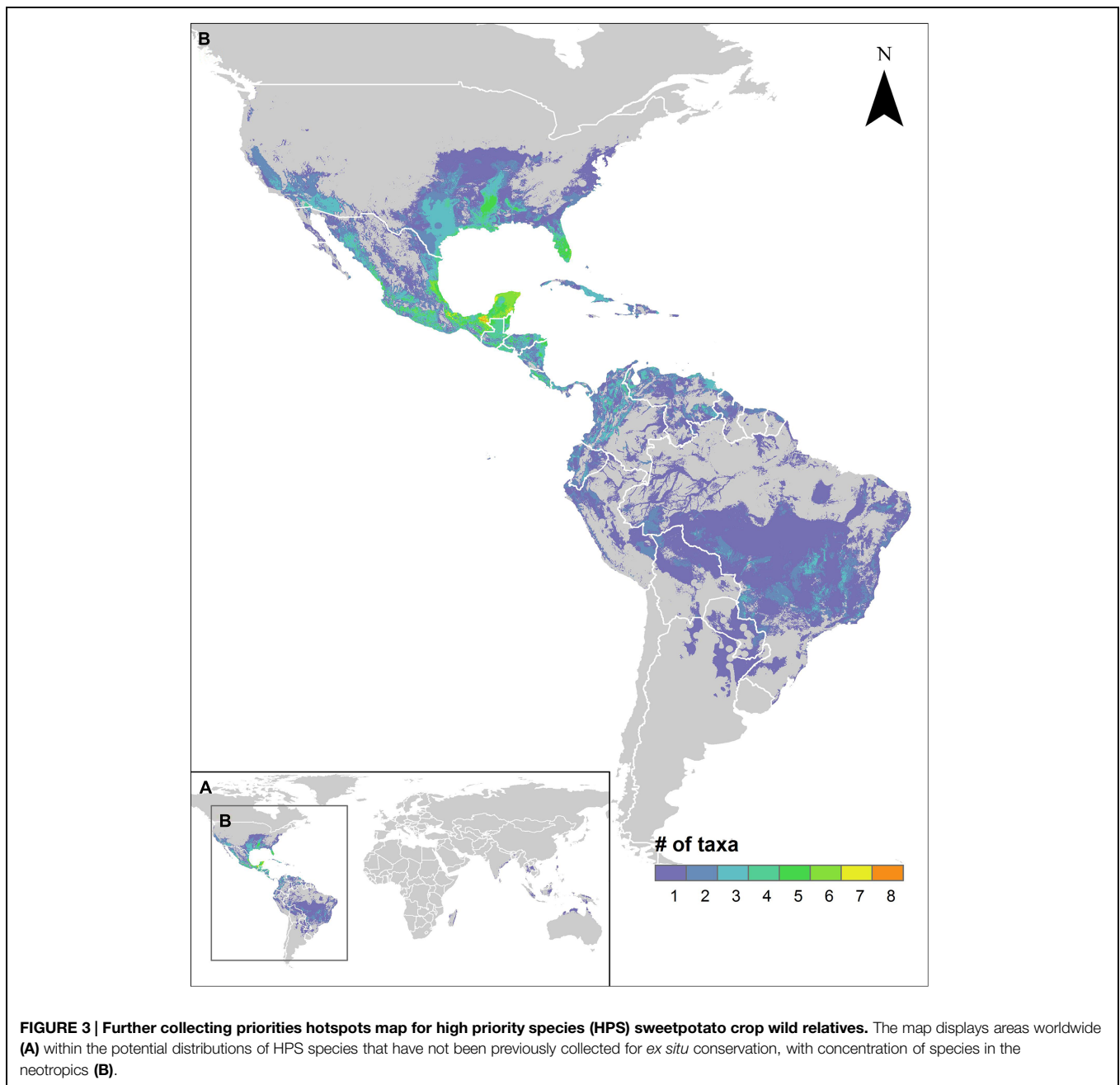
Discussion

This article utilizes the most current knowledge on species concepts within *I. series Batatas*. Due to taxonomic uncertainties and to the notable dearth of study material for sweetpotato CWR, the results represent a preliminary understanding of the geography and conservation status of the series, to be further refined following increased levels of collecting sufficient to support the needed taxonomic and crossability research. Further collecting of germplasm thus serves two purposes: (a) conserves genetic resources for the long-term and makes these resources available to breeders; and (b) provides the basic materials needed by

researchers to understand the diversity present in the CWR of sweetpotato.

A total of 78.6% of the wild relatives of sweetpotato considered in this study were assessed as high priority for further collecting for *ex situ* conservation. With general agreement from expert evaluators of medium to high importance for all species, it is clear that much remains to be done to safeguard the wild genetic resources of this critically important root crop. Included in this list of priorities are species with very few germplasm accessions accessible to the global community in genebank information systems, including *I. cynanchifolia, I. littoralis, I. tenuissima, I. tabascana, I. lacunosa, I. leucantha, I. splendor-sylvae,* and clearly designated wild forms of the crop conspecific *I. batatas.* Such species represent the highest level of priority for further collecting for use in systematic analyses as well as genetic resources conservation. As the species diversity gaps in genebank collections largely align with the geographic distribution of species richness of sweetpotato CWR, hotspots in central Mexico to Central America, and in the extreme Southeastern USA, represent particularly high priority regions for efficient collecting of the sweetpotato genepool (**Figure 3**). Additional unrepresented populations of HPS such as *I. littoralis* and *I. cynanchifolia* occur outside those regions, thus targeted collecting throughout the geographic distribution of the genepool is necessary in order to form comprehensive germplasm collections.
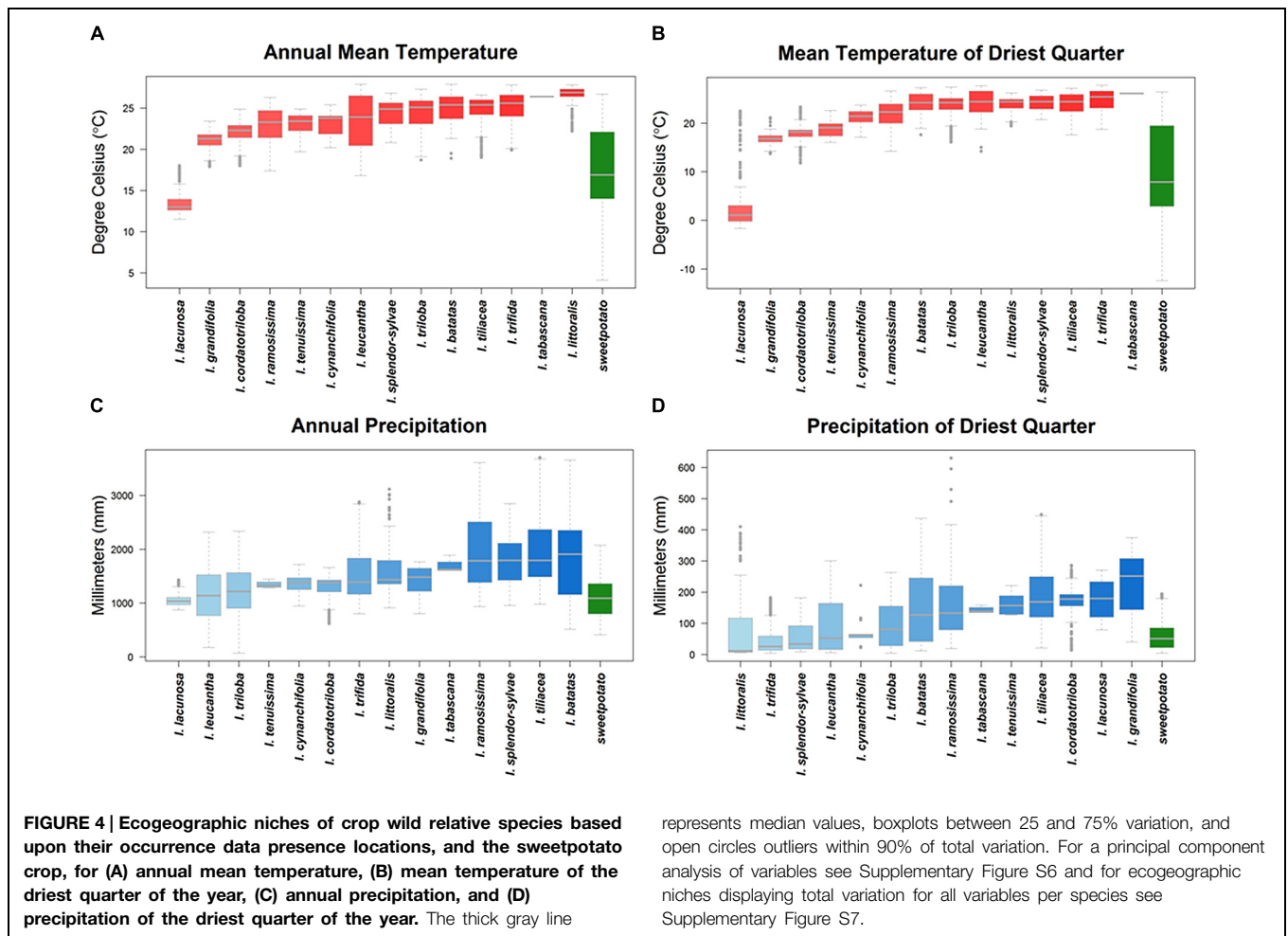
Due to having relatively large potential distributions occupying a diversity of ecosystems, species such as *I. triloba, I.*

FIGURE 3 | Further collecting priorities hotspots map for high priority species (HPS) sweetpotato crop wild relatives. The map displays areas worldwide **(A)** within the potential distributions of HPS species that have not been previously collected for *ex situ* conservation, with concentration of species in the neotropics **(B)**.

cordatotriloba, and *I. tiliacea* were categorized as high priority, and *I. trifida* and *I. grandifolia* as medium priority for further collecting despite having sizable currently existing germplasm collections. As the cost of conserving and investigating germplasm *ex situ* is significant, a further assessment of what constitutes sufficient representation of these species in germplasm collections is warranted. Given adequate resources, further collecting may be worthwhile, as extremely valuable traits sourced from CWR of some crops have been found in only a few populations despite screening of a large number of accessions (Brar and Khush, 1997), and accessions of individual CWR species such as *I. triloba* have been shown to possess notable variation in traits such as tolerance

to precipitation (Martin and Jones, 1973; Nimmakayala et al., 2011).

As Maxent models are based upon known presence points for species and are thus subject to sampling bias, they may not fully capture the possibility of occurrence of populations of CWR species in unique climates (Araújo and Guisan, 2006; Loiselle et al., 2007; Kramer-Schadt et al., 2013). Further field exploration of climatic extremes beyond the edges of the distributions created through these methods may therefore lead to the discovery of new populations with particularly valuable adaptations to abiotic stress (Williams et al., 2009). Investigation of non-native populations (e.g., *I. trifida* in Asia) may also yield novel adaptations

FIGURE 4 | Ecogeographic niches of crop wild relative species based upon their occurrence data presence locations, and the sweetpotato crop, for (A) annual mean temperature, (B) mean temperature of the driest quarter of the year, (C) annual precipitation, and (D) precipitation of the driest quarter of the year. The thick gray line represents median values, boxplots between 25 and 75% variation, and open circles outliers within 90% of total variation. For a principal component analysis of variables see Supplementary Figure S6 and for ecogeographic niches displaying total variation for all variables per species see Supplementary Figure S7.

of use in crop improvement. As techniques for the utilization of distantly related germplasm improve, the exploration of other *Ipomoea* species outside of *I.* series *Batatas* may also result in the identification of beneficial traits [e.g., *Ipomoea purpurea* (L.) Roth, for stem nematode and SPVD resistance (Cao et al., 2009)].

Analysis of geographic overlap and ecogeographic similarities between species, as well as ecogeographic clusters among all species, can supplement morphological and genetic analyses in differentiating useful genetic resources, and can serve as a point of departure for identifying taxonomically problematic populations for further investigation. These analyses may also indicate geographic areas of particular interest in regard to high rates of hybridization, as in the case of *I. cordatotriloba* and *I. lacunosa* (Duncan and Rausher, 2013), which indeed were identified as sharing similar geographies and ecogeographic niches. The purported hybrid descendent of these species, *I. leucantha*, was modeled as containing a much more extended latitudinal gradient in the northern hemisphere than its parents, as well as a differing ecogeographic niche, including potential adaptation to high heat and to drought.

Genetic resistance is essential to efforts to overcome major biotic and abiotic constraints in sweetpotato production. As these constraints are often interrelated, e.g., drought stress with SPW and SPVD damage, enhancement of broad resistance for traits such as drought may improve yield across relatively large geographic areas, without the need to breed for resistance to localized viral strains (Ngailo et al., 2013). Such broad tolerance may also improve adoption rates for sweetpotato varieties with other desirable characteristics, such as high β-carotene content.

Reliable funding for germplasm collections is paramount in order to safeguard sweetpotato CWR genetic resources in the long-term and to continue to make *ex situ* collections available to the global community. Further investment in genebank information systems, *ex situ* conservation technologies (i.e., storage, testing, and regeneration), safety duplication of unique germplasm, and ensuring facilitated access to genetic diversity is equally essential (Food and Agriculture Organization of the United Nations [FAO], 2002, 2010; Khoury et al., 2010). In order to maximize the usefulness of conserved germplasm, characterization and evaluation for traits of interest, and increased breeding research, which have been limited for CWR of sweetpotato, are also needed. Further research combining morphological studies, trait evaluations, and genetic diversity analyses is likewise critically needed for elucidating species boundaries and highlighting accessions of particular value for use in breeding. Recent focused research has produced quick gains, including

the identification of new species (Duncan and Rausher, 2013; Wood et al., in press). Through these actions the crop research community will contribute to ensuring the long term viability of this important root crop.

Author Contributions

CK, BH, NC-A, HA, CS, HJ, and SdH conceived and designed the study. CK, NC-A, HA, CS, and VB gathered data and performed the analyses. CK, BH, NC-A, HA, CS, RM, RS, JW, GR, LE, RJ, CY, SdH, and PS analyzed the results. CK and BH wrote the manuscript. CK, BH, NC-A, HA, CS, RM, RS, JW, GR, LE, RJ, CY, VB, HJ, SdH, and PS edited the manuscript.

Acknowledgments

We dedicate this research to the late Dan F. Austin, who provided the foundation for our understanding of the taxonomy of *Ipomoea* series *Batatas*. We thank the taxonomic and

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00251/abstract

References

Araújo, M. B., and Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *J. Biogeogr.* 33, 1677–1688. doi: 10.1111/j.1365-2699.2006.01584.x

Austin, D. F. (1978). The *Ipomoea batatas* complex – 1. Taxonomy. *Bull. Torrey Bot. Club* 105, 114–129. doi: 10.2307/2484429

Austin, D. F. (1988). "The taxonomy, evolution and genetic diversity of sweetpotatoes and related wild species," in *Exploration, Maintenance, and Utilization of Sweetpotato Genetic Resources*, ed. P. Gregory (Lima: International Potato Center), 27–60.

Austin, D. F. (1991). *Ipomoea littoralis* (Convolvulaceae) - Taxonomy, distribution, and ethnobotany. *Econ. Bot.* 45, 251–256. doi: 10.1007/BF02862052

Austin, D. F. (1997). Dissolution of *Ipomoea* series Anisomerae (Convolvulaceae). *Taxon* 28, 359–361.

Austin, D. F., and Huamán, Z. (1996). A synopsis of *Ipomoea* (Convolvulaceae) in the Americas. *Taxon* 45, 3–38. doi: 10.2307/1222581

Bill and Melinda Gates Foundation. (2011). *Agricultural Development Strategy Overview*. Available at: https://docs.gatesfoundation.org/Documents/agricultural-development-strategy-overview.pdf

Bivand, R., and Lewin-Koh, N. (2014). *Maptools: Tools For Reading and Handling Spatial Objects. R Package version 0.8-30.* Available at: http://CRAN.R-project.org/package=maptools

Bivand, R., Keitt, T., and Rowlingson, B. (2014). *Rgdal: Bindings for the Geospatial Data Abstraction Library. R Package version 0.8-16.* Available at: http://CRAN.R-project.org/package=rgdal

Bivand, R., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied Spatial Data Analysis With R*, 2nd Edn. New York, NY: Springer. Available at: http://www.asdar-book.org/

Bouis, H., and Islam, Y. (2012). *Delivering Nutrients Widely through Biofortification: Building on Orange Sweet Potato - Scaling Up in Agriculture, Rural Development and Nutrition. Focus 19, Brief 11.* Washington, DC: International Food Policy Research Institute.

Brar, D. S., and Khush, G. S. (1997). Alien introgression in rice. *Plant Mol. Biol.* 35, 35–47. doi: 10.1023/A:1005825519998

Cao, Q., Zhang, A., Ma, D., Li, H., Li, Q., and Li, P. (2009). Novel interspecific hybridization between sweetpotato (*Ipomoea batatas* (L.) Lam.) and its two diploid wild relatives. *Euphytica* 169, 345–352. doi: 10.1007/s10681-009-9967-7

Conolly, J., Manning, K., Colledge, S., Dobney, K., and Shennan, S. (2012). Species distribution modelling of ancient cattle from early neolithic sites in SW Asia and Europe. *Holocene* 22, 997–1010. doi: 10.1177/0959683612437871

Costa, G. C., Nogueira, C., Machado, R. B., and Colli, G. R. (2010). Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodivers. Conserv.* 19, 883–899. doi: 10.1007/s10531-009-9746-8

De Vries, C. A., Ferwerda, J. D., and Flach, M. (1967). Choice of food crops in relation to actual and potential production in the tropics. *Neth. J. Agr. Sci.* 15, 241–248.

Dempewolf, H., Eastwood, R. J., Guarino, L., Khoury, C. K., Müller, J. V., and Toll, J. (2014). Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecol. Sust. Food* 38, 369–377. doi: 10.1080/21683565.2013.870629

Diaz, J., Schmiediche, P., and Austin, D. F. (1996). Polygon of crossability between eleven species of *Ipomoea*: section batatas (Convolvulaceae). *Euphytica* 88, 189–200. doi: 10.1007/BF00023890

Duncan, T. M., and Rausher, M. D. (2013). Evolution of the selfing syndrome in *Ipomoea*. *Front. Plant Sci.* 4:301. doi: 10.3389/fpls.2013.00301

Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151. doi: 10.1111/j.2006.0906-7590.04596.x

Eserman, L. A., Tiley, G. P., Jarret, R. L., Leebens-Mack, J. H., and Miller, R. E. (2014). Phylogenetics and diversification of morning glories (tribe Ipomoeeae, Convolvulaceae) based on whole plastome sequences. *Am. J. Bot.* 101, 92–103. doi: 10.3732/ajb.1300207

ESRI. (2011). *ArcGIS Desktop: Release 10.1.* Available at: http://www.esri.com/software/arcgis/arcgis-for-desktop

Food and Agriculture Organization of the United Nations [FAO]. (2002). *International Treaty on Plant Genetic Resources for Food and Agriculture.* Rome: FAO.

Food and Agriculture Organization of the United Nations [FAO]. (2010). *The Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture.* Rome: FAO.

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. *Trends Genet.* 24, 24–32. doi: 10.1016/j.tig.2007.11.001

Fieberg, J., and Kochanny, C. O. (2005). Quantifying home-range overlap: the importance of the utilization distribution. *J. Wildl. Manag.* 69, 1346–1359. doi: 10.2193/0022-541X(2005)69[1346:QHOTIO]2.0.CO;2

Gaiji, S., Chavan, V., Ariño, A. H., Otegui, J., Hobern, D., Sood, R., et al. (2013). Content assessment of the primary biodiversity data published through GBIF network: status, challenges and potentials. *Biodivers. Inform.* 8, 94–172.

Germplasm Resources Information Network [GRIN]. (2013). *USDA, ARS, National Genetic Resources Program.* Available at: http://www.ars-grin.gov/cgi-bin/npgs/html/wep.pl

Gichuki, S. T., Berenyi, M., Zhang, D., Hermann, M., Schmidt, J., Glössl, J., et al. (2003). Genetic diversity in sweetpotato [*Ipomoea batatas* (L.) Lam.] in relationship to geographic sources as assessed with RAPD markers. *Genet. Resour. Crop Evol.* 50, 429–437. doi: 10.1023/A:1023998522845

Google Maps Geocoder. (2013). *Version 3.* Available at: https://developers.google.com/maps/documentation/geocoding/?hl=pl

Guarino, L., and Lobell, D. B. (2011). A walk on the wild side. *Nat. Clim. Change* 1, 374–375. doi: 10.1038/nclimate1272

Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives for crop improvement: a survey of developments over the past 20 years. *Euphytica* 156, 1–13. doi: 10.1007/s10681-007-9363-0

Harlan, J. R., and de Wet, J. M. J. (1971). Toward a rational classification of cultivated plants. *Taxon* 20, 509–517. doi: 10.2307/1218252

Heibl, C. (2011). *Phyloclim: Integrating Phylogenetics and Climatic Niche Modeling. v.0.9-4.* Available at: http://cran.r-project.org/web/packages/phyloclim/index.html

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., et al. (2014). SoilGrids1km – global soil information based on automated mapping. *PLoS ONE* 9:e105992. doi: 10.1371/journal.pone.0105992

Hijmans, R. J. (2012). Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93, 679–688. doi: 10.1890/11-0826.1

Hijmans, R. J. (2014). *Raster: Geographic Data Analysis and Modeling. R package version 2.2-31.* Available at: http://CRAN.R-project.org/package=raster

Hijmans, R. J., Cameron, S., Parra, J., Jones, P. G., and Jarvis, A. (2005). *WorldClim, version 1.3.* University of California, Berkeley.

Hijmans, R. J., Phillips, S., Leathwick, J., and Elith, J. (2013). *Dismo: Species Distribution Modeling. R package version 0.9-3.* Available at: http://CRAN.R-project.org/package=dismo

Hijmans, R. J., Schreuder, M., De La Cruz, J., and Guarino, L. (1999). Using GIS to check co-ordinates of genebank accessions. *Genet. Resour. Crop Evol.* 46, 291–296. doi: 10.1023/A:1008628005016

Hjarding, A., Tolley, K. A., and Burgess, N. D. (2014). Red List assessments of East African chameleons: a case study of why we need experts. *Oryx* doi: 10.1017/s0030605313001427

Hotz, C., Loechl, C., Lubowa, A., Tumwine, J. K., Ndeezi, G., Nandutu Masawi, A., et al. (2012). Introduction of β-carotene-rich orange sweet potato in rural Uganda resulted in increased vitamin a intakes among children and women and improved vitamin a status among children. *J. Nutr.* 142, 1871–1880. doi: 10.3945/jn.111.151829

Huaman, Z. (1987). "Current status on maintenance of sweetpotato genetic resources at CIP," in *Proceedings of the Exploration, Maintenance and Utilization of Sweetpotato Genetic Resources. Report of the First Sweetpotato Planning Conference 1987.* (Lima, Peru: CIP), 101–120.

Huang, J., Corke, H., and Sun, M. (2002). Highly polymorphic AFLP markers as a complementary tool to ITS sequences in assessing genetic diversity and phylogenetic relationships of sweetpotato (*Ipomoea batatas* (L.) Lam.) and its wild relatives. *Genet. Resour. Crop Evol.* 49, 541–550. doi: 10.1023/A:1021290927362

Husson, F., Josse, J., Le, S., and Mazet, J. (2009). *FactoMineR: Multivariate Exploratory Data Analysis and Dmining with R. V.1.26.* Available at: http://factominer.free.fr

Iwanaga, M. (1988). "Use of wild germplasm for sweetpotato breeding," in *Exploration, Maintenance, and Utilization of Sweetpotato Genetic Resources*, ed. P. Gregory (Lima: International Potato Center), 199–210.

Jarret, R. L., and Austin, D. F. (1994). Genetic diversity and systematic relationships in sweetpotato (*Ipomoea batatas* (L.) Lam.) and related species as revealed by RAPD analysis. *Genet. Resour. Crop Evol.* 41, 165–173. doi: 10.1007/BF00051633

Jarret, R. L., Gawel, N., and Whittemore, A. (1992). Relationships of the sweetpotato [*Ipomoea batatas* (L.) Lam.]. *J. Am. Soc. Hortic. Sci.* 117, 633–637.

Kays, S. J. (1985). "The physiology of yield in sweet potato," in *Sweet Potato Products: A Natural Resource of the Tropics*, ed. J. Bouwkamp (Boca Raton, FL: CRC Press), 19–133.

Kays, S. J., and Kays, S. E. (1998). "Sweet potato chemistry in relation to health," in *Proceedings of the Sweet Potato Production System Towards the 21st Century*, Miyakonojo: Kyushu National Agricultural Experimental Station, 231–272.

Kernohan, B. J., Girzen, R. A., and Millspaugh, J. J. (2001). "Analysis of animal space use and movements," in *Radio Tracking Animal Populations*. eds J. J. Millspaugh and J. M. Marzluff (San Diego: Academic Press), 125–166. doi: 10.1016/B978-012497781-5/50006-2

Khoury, C., Laliberté, B., and Guarino, L. (2010). Trends in ex situ conservation of plant genetic resources: a review of global crop and regional conservation strategies. *Genet. Resour. Crop Evol.* 57, 625–639. doi: 10.1007/s10722-010-9534-z

Khoury, C. K., Castañeda Álvarez, N. P., Achicanoy, H., Sosa, C. C., Bernau, V., Kassa, M. T., et al. (2015). Crop wild relatives of pigeonpea [*Cajanus cajan* (L.) Millsp.]: distributions, ex situ conservation status, and potential genetic resources for abiotic stress tolerance. *Biol. Conserv.* 184, 259–270. doi: 10.1016/j.biocon.2015.01.032

Kivuva, B. M., Githiri, S. M., Yencho, C., and Sibiya, J. (2015a). Combining ability and heterosis for yield and drought tolerance traits under managed drought stress in sweetpotato. *Euphytica* 201, 423–440. doi: 10.1007/s10681-014-1230-1

Kivuva, B. M., Githiri, S. M., Yencho, C., and Sibiya, J. (2015b). Screening sweetpotato genotypes for tolerance to drought stress. *Field Crop Res.* 171, 11–22. doi: 10.1016/j.fcr.2014.10.018

Komaki, K. (2004). "Breeding value of wild species closely related to sweetpotato," in *Proceedings of international workshop on production, utilization and development of sweetpotato*, Korea, 164–172.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schroder, B., Lindenborn, J., Reinfelder, V., et al. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Divers. Distrib.* 19, 1366–1379. doi: 10.1111/ddi.12096

Liu, C., Berry, P. M., Dawson, T. P., and Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 3, 385–393. doi: 10.1111/j.0906-7590.2005.03957.x

Loiselle, B., Jørgensen, P. M., Consiglio, T., Jiménez, I., Blake, J. G., Lohmann, L. G., et al. (2007). Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *J. Biogeogr.* 35, 105–116. doi: 10.1111/j.1365-2699.2007.01779.x

Low, J. W., Arimond, M., Osman, N., Cunguara, B., Zano, F., and Tschirley, D. (2007). A food-based approach introducing orange-fleshed sweet potatoes increased vitamin a intake and serum retinol concentrations in young children in rural Mozambique. *J. Nutr.* 137, 1320–1327.

Low, J., Lynam, J., Lemaga, B., Crissman, C., Barker, I., Thiele, G., et al. (2009). "Sweetpotato in Sub-Saharan Africa," in *The Sweet Potato*, eds G. Loebenstein and G. Thottappilly (Amsterdam: Springer), 359–390.

Lu, S., and Li, T. (1992). Study on the characterization of intra and inter-specific incompatibility of Batatas section. *Acta Agron. Sin.* 18, 161–168.

Mabberley, D. J. (2008). *Mabberley's Plant Book: A Portable Dictionary of Plants, Their Classifications, and Uses*, 3rd Edn. Cambridge: Cambridge University Press.

Martin, F. W. (1970). Self- and interspecific incompatibility in the Convolvulaceae. *Bot. Gaz.* 131, 139–144. doi: 10.1086/336524

Martin, F. W. (1982). "Analysis of the incompatibility and sterility of sweetpotato," in *Proceedings of 1st International Symposium on Sweetpotato*, eds R. L. Villareal and T. D. Griggs (Tainan: AVRDC), 275–283.

Martin, F. W., and Jones, A. (1973). The species of *Ipomoea* closely related to the sweetpotato. *Econ. Bot.* 26, 201–215. doi: 10.1007/BF02861033

McCouch, S. R., Sweeney, M., Li, J., Jiang, H., Thomson, M., Septingsinsih, E., et al. (2007). Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. *Euphytica* 154, 317–339. doi: 10.1007/s10681-006-9210-8

McDonald, J. A., and Austin, D. F. (1990). Changes and additions in *Ipomoea* section Batatas (Convolvulaceae). *Brittonia* 42, 116–120. doi: 10.2307/2807625

McDonald, J. A., Hansen, D. R., McDill, J. R., and Simpson, B. B. (2011). A phylogenetic assessment of breeding systems and floral morphology of North American *Ipomoea* (Convolvulaceae). *J. Bot. Res. Inst. Texas* 5, 159–177. doi: 10.1007/s00606-011-0550-z.1.15

Mesibov, R. (2013). A specialist's audit of aggregated occurrence records. *ZooKeys* 293, 1–18. doi: 10.3897/zookeys.293.5111

Miller, R. E., Rausher, M. D., and Manos, P. S. (1999). Phylogenetic systematics of *Ipomoea* (Convolvulaceae) based on ITS and waxy sequences. *Syst. Bot.* 24, 209–227. doi: 10.2307/2419549

Monfreda, C., Ramankutty, N., and Foley, J. A. (2008). *Farming the planet*: 2. *Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global biogeochemical cycles 22*: GB1022. Available at http://www.geog.mcgill.ca/landuse/pub/Data/175crops2000/ doi: 10.1029/2007GB002947

Munyiza, H., Stevenson, P., Mwanga, R., Talwana, H., Murumu, J., and Odongo, B. (2007). The relationship between stem base and root damage by *Cylas* spp. on sweet potato. *Afr. Crop Sci. J.* 8, 955–957.

Mwanga, R., and Ssemakula, G. (2011). Orange-fleshed sweet potatoes for food, health and wealth in Uganda. *IJSA* 9, 42–49. doi: 10.3763/ijas.2010.0546

Mwololo, J. K., Mburu, M. W. K., Njeru, R. W., Ateka, E. M., Kiarie, N., Munyua, J. K., et al. (2007). Resistance of sweet potato genotypes to sweet potato virus disease in Coastal Kenya. *Proc. ACSS Conf.* 8, 2083–2086.

Ngailo, S., Shimelis, H., Sibiya, J., and Mtunda, K. (2013). Sweet potato breeding for resistance to sweet potato virus disease and improved yield: progress and challenges. *Afr. J. Agric. Res.* 8, 3202–3215.

Nimmakayala, P., Vajja, G., and Reddy, U. K. (2011). "*Ipomoea*," in *Wild Crop Relatives: Genomic and Breeding Resources*, ed. C. Kole C (Berlin: Springer), 123–132.

Nishiyama, I. (1971). Evaluation and domestication of sweetpotato. *Bot. Magazine Tokyo* 84, 377–387. doi: 10.15281/jplantres1887.84.377

Nishiyama, I. (1982). "Autohexaploid evolution of the sweetpotato," in *Proceedings of 1st International Symposium on Sweetpotato*, eds R. L. Villareal and T. D. Griggs (Shanhua: Asian Vegetable Research and Development Center), 263–274.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* 51, 933–938. doi: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Otegui, J., Ariño, A. H., Encinas, M. A., and Pando, F. (2013). Assessing the primary data hosted by the Spanish node of the global biodiversity information facility (GBIF). *PLoS ONE* 8:e55144. doi: 10.1371/journal.pone.0055144

Pebesma, E. J., and Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News 5* (2). Available at: http://cran.r-project.org/doc/Rnews/

Pfeiffer, W. H., and McClafferty, B. (2007). HarvestPlus: breeding crops for better nutrition. *Crop Sci.* 47, 88–105. doi: 10.2135/cropsci2007.09.0020IPBS

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., and Dudik, M. (2008). Modeling of species distributions with maxent: new extensions and comprehensive evaluation. *Ecography* 31, 161–175. doi: 10.1111/j.0906-7590.2008.5203.x

Piperno, D. R., and Pearsall, D. M. (1998). *The Origins of Agriculture in the Lowland Neotropics*. San Diego: Academic Press.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Available at: http://www.r-project.org/

Rajapakse, S., Nilmalgoda, S. D., Molnar, M., Ballard, R. E., Austin, D. F., and Bohac, J. R. (2004). Phylogenetic relationships of the sweetpotato in *Ipomoea* series Batatas (Convolvulaceae) based on nuclear β-amylase gene sequences. *Mol. Phylogenet. Evol.* 30, 623–632. doi: 10.1016/S1055-7903(03)00249-5

Ramírez-Villegas, J., Khoury, C., Jarvis, A., Debouck, D. G., and Guarino, L. (2010). A gap analysis methodology for collecting crop genepools: a case study with Phaseolus beans. *PLoS ONE* 5:e13497. doi: 10.1371/journal.pone.0013497

Revelle, W. (2015). *Psych: Procedures for Phychological, Psychometric, and Personality Research*, v. 1.5.1. Available at: http://cran.r-project.org/web/packages/psych/index.html

Roullier, C., Duputié, A., Wennekes, P., Benoit, L., Fernández Bringas, V. M., Rossel, G., et al. (2013). Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). *PLoS ONE* 8:e62707. doi: 10.1371/journal.pone.0062707

Sanchez. G. (2012). *Package 'Plsdepot'. Partial Least Squares (PLS) Data Analysis Methods*, v. 0.1.17. Available at: http://cran.r-project.org/web/packages/plsdepot/index.html

Schafleitner, R., Tincopa, L. R., Palomino, O., Rossel, G., Robles, R. F., Alagon, R., et al. (2010). A sweet potato gene index established by de novo assembly of pyrosequencing and Sanger sequences and mining for gene-based microsatellite markers. *BMC Genomics* 11:604. doi: 10.1186/1471-2164-11-604

Solis, R. S., Haas, J., and Creamer, W. (2001). Dating caral, a preceramic site in the supe valley on the central coast of Peru. *Science* 292, 723–726. doi: 10.1126/science.1059519

Shiotani, I., Huang, Z. Z., Sakamoto, S., and Miyazaki, T. (1994). The role of the wild *Ipomoea* trifida germplasm in sweet potato breeding. *Acta Hort.* (*ISHS*) 380, 388–398.

Smit, N. E. J. M. (1997). *Integrated Pest Management for Sweetpotato in Eastern Africa*. Ph.D. thesis, Wageningen University, Wageningen.

Srisuwan, S., Sihachakr, D., and Siljak-Yakovlev, S. (2006). The origin and evolution of sweet potato (*Ipomoea batatas* Lam.) and its wild relatives through the cytogenetic approaches. *Plant Sci.* 171, 424–433. doi: 10.1016/j.plantsci.2006.05.007

Sutherland, J. A. (1986). A review of the biology and control of the sweetpotato weevil, cylas formicarius elegaluntus (F). *Trop. Pest Manag.* 32, 304–315. doi: 10.1080/09670878609371084

Teramura, T. (1979). Phylogenetic study of *Ipomoea* species in the section Batatas. *Mem. Coll. Agric. Kyoto Univ.* 114, 29–48.

The Plant List. (2010). *Version 1*. Available at: http://www.theplantlist.org/

Tumwegamire, S., Kapinga, R., Rubaihayo, P. R., LaBonte, D. R., Grüneberg, W. J., Burgos, G., et al. (2011). Evaluation of dry matter: protein, starch, β-carotene, iron, zinc, calcium, and magnesium in East African sweetpotato [*Ipomoea batatas* (L.) *Lam*] germplasm. *Hort. Science* 46, 348–357.

Valverde, R. A., Clark, C. A., and Valkonen, J. P. T. (2007). Viruses and virus disease complexes of sweetpotato. *Plant Viruses* 1, 116–126.

van der Wal, J., Falconi, L., Januchowski, S., Shoo, L., and Storlie, C. (2014). *SDMTools: Species Distribution Modelling Tools: Tools for Processing Data Associated with Species Distribution Modelling Exercises. R Package version 1.1-20*. http://CRAN.R-project.org/package=SDMTools

Warren, D. L., Glor, R. E., and Turelli, M. (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62, 2868–2883. doi: 10.1111/j.1558-5646.2008.00482.x

Wilkin, P. (1999). A morphological cladistic analysis of the Ipomoeeae (Convolvulaceae). *Kew Bull.* 54, 853–876. doi: 10.2307/4111164

Williams, J. N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brian, J. M., et al. (2009). Using species distribution models to predict new occurrences for rare plants. *Divers. Distrib.* 15, 565–576. doi: 10.1111/j.1472-4642.2009.00567.x

Wood, J. R. I., Carine, M. A., Harris, D., Wilkin, P., Williams, B., and Scotland, R. W. (in press). *Ipomoea* (Convolvulaceae) in Bolivia. *Kew Bull.*

Woolfe, J. A. (1992). Sweetpotato, an untapped food source. New York, NY: Cambridge University Press.

Xiao, J., Grandillo, S., Ahn, S. N., McCouch, S. R., and Tanksley, S. D. (1996). Genes from wild rice improve yield. *Nature* 384, 223–224. doi: 10.1038/384223a0

Zhang, D. P., Cervantes, J., Huamán, Z., Carey, E., and Ghislain, M. (2000). Assessing the genetic diversity of sweetpotato (*Ipomoea batatas* (L.) La.) cultivars from tropical America using AFLP. *Genet. Resour. Crop Evol.* 47, 659–665. doi: 10.1023/A:1026520507223

Zhang, S. S., and Liu, L. F. (2005). "Utilization of the wild relatives in sweetpotato breeding," in *Sweetpotato Breeding and Industrialization in China*, ed. Q. C. Liu (Beijing: China Agricultural University Press), 29–32.