HAL
open science

Crop wild relatives of the brinjal eggplant ([i]Solanum melongena):[/i] Poorly represented in genebanks and many species at risk of extinction

Mindy M. Syfert, Nora P. Castaneda-Alvarez, Colin K. Khoury, Tiina Särkinen, Chrystian C. Sosa, Harold A. Achicanoy, Vivian Bernau, Jaime Prohens, Marie-Christine Brand-Daunay, Sandra Knapp

To cite this version:

Mindy M. Syfert, Nora P. Castaneda-Alvarez, Colin K. Khoury, Tiina Särkinen, Chrystian C. Sosa, et al.. Crop wild relatives of the brinjal eggplant ([i]Solanum melongena):[/i] Poorly represented in genebanks and many species at risk of extinction. American Journal of Botany, 2016, 103 (4), pp.635-651. 10.3732/ajb.1500539 . hal-01315233

HAL Id: hal-01315233

https://hal.science/hal-01315233v1

Submitted on 12 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AJB Advance Article published on March 29, 2016, as 10.3732/ajb.1500539.
The latest version is at http://www.amjbot.org/cgi/doi/10.3732/ajb.1500539

RESEARCH ARTICLE

AMERICAN JOURNAL OF BOTANY

Crop wild relatives of the brinjal eggplant (*Solanum melongena*): Poorly represented in genebanks and many species at risk of extinction[1]

Mindy M. Syfert[2,9], Nora P. Castañeda-Álvarez[3,4], Colin K. Khoury[3,5], Tiina Särkinen[6], Chrystian C. Sosa[3], Harold A. Achicanoy[3], Vivian Bernau[3], Jaime Prohens[7], Marie-Christine Daunay[8], and Sandra Knapp[2,9]

PREMISE OF THE STUDY: Crop wild relatives (CWR) provide important traits for plant breeding, including pest, pathogen, and abiotic stress resistance. Therefore, their conservation and future availability are essential for food security. Despite this need, the world's genebanks are currently thought to conserve only a small fraction of the total diversity of CWR.

METHODS: We define the eggplant genepool using the results of recent taxonomic and phylogenetic studies. We identify the gaps in germplasm accessions for eggplant (*Solanum melongena* L.) CWR by comparing georeferenced herbarium records and germplasm accessions using a gap analysis methodology implementing species distribution models (SDM). Preliminary conservation assessments using IUCN criteria were done for all species and were combined with the gap analysis to pinpoint where under-collected and threatened CWR species coincide with high human disturbance and occur outside of protected areas.

KEY RESULTS: We show that many eggplant CWR are poorly represented in genebanks compared to their native ranges. Priority areas for future collecting are concentrated in Africa, especially along the Kenya-Tanzania border. Fourteen species of eggplant CWR are assessed as threatened or near-threatened; these are also concentrated in eastern Africa.

CONCLUSIONS: The knowledge base upon which conservation of wild relative germplasm depends must take into account both taxonomic and phylogenetic advances. Beyond traditional research focus on close relatives of crops, we emphasize the benefits of defining a broad CWR genepool, and the importance of assessing threats to wild species when targeting localities for future collection of CWR to improve crop breeding in the face of environmental change.

KEY WORDS agriculture; crop genepools; eggplant; food security; gap analysis; genebanks; Red List assessments; Solanaceae; *Solanum*; species distribution modeling

Domestication and subsequent improvement has generally led to narrowing of the genetic diversity of our commonly cultivated crop varieties (Tanksley and McCouch, 1997). Limited genetic diversity also limits the potential of crop breeding efforts, especially in the face of environmental change. This situation has reinvigorated research into the use of crop wild relatives (CWR), the wild cousins of domesticated crop species, in breeding efforts (Tanksley and McCouch, 1997; Hodgkin and Hajjar, 2008; Dempewolf et al., 2014). Crop wild relatives share a relatively recent common ancestry with domesticated species and, due to that close relationship, are reservoirs of genetic traits that can be useful in crop improvement (Tanksley and McCouch, 1997; Guarino and Lobell, 2011). Interest in CWR has its roots in Nikolai Vavilov's centers of origin of cultivated plants (as compiled in Vavilov, 1992) concepts where regions identified as rich in wild species related to crops were suggested to be the centers of domestication and origin of these crops, and where significant genetic diversity was thought to reside, both in terms of related wild species and in the crop itself. Despite this long history of interest and the recognized value of wild and weedy crop

[1] Manuscript received 30 December 2015; revision accepted 4 February 2016.
[2] Department of Life Sciences, Natural History Museum, Cromwell Road, London, SW7 5BD, UK;
[3] International Center for Tropical Agriculture (CIAT), Km 17 Recta Cali-Palmira, Cali, Colombia;
[4] School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK;
[5] Centre for Crop Systems Analysis, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, Netherlands;
[6] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK;
[7] Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Valencia, Spain; and
[8] INRA, Unité de Genetique & Amélioration des Fruits et Legumes, UR 1052, Domaine St. Maurice, CS 60094 F-84143, Montfavet cedex, France.
[9] Authors for correspondence (e-mail: s.knapp@nhm.ac.uk; m.syfert@nhm.ac.uk)
doi:10.3732/ajb.1500539

Copyright 2016 by the Botanical Society of America

relatives, they remain largely understudied and their natural populations are increasingly at risk due to habitat loss and conversion, climate change, and invasive species (e.g., Ford-Lloyd et al., 2011).

In the past decade, the International Treaty on Plant Genetic Resources for Food and Agriculture (FAO, 2009) has provided momentum for efforts to better conserve and share such wild genetic resources. More recently, international instruments such as the United Nations' Sustainable Development Goals, the updated Global Strategy for Plant Conservation (CBD, 2010b) and the Aichi Biodiversity Targets (CBD, 2010a) have specifically targeted the comprehensive conservation of wild plant genetic resources.

The conservation of CWR is a matter of urgency because many of these species are considered to be threatened in their natural habitats due to the conversion of land to industrial agricultural production, urbanization, mining, the spread of invasive species, pollution, and climate change (Jarvis et al., 2008; Brummitt et al., 2015). A number of initiatives have been enacted to conserve such genetic resources in the wild, but these efforts have treated a limited number of species and geographic regions (Meilleur and Hodgkin, 2004), and much greater efforts are needed to adequately conserve wild genetic resources as a whole (Maxted and Kell, 2009; Iriondo et al., 2012).

Genebanks, also known as seedbanks, have been an important component in the conservation and availability of wild species useful for plant breeding (Tanksley and McCouch, 1997), enabling their maintenance in controlled environments and facilitating their availability to crop breeders and other researchers. Assessing the level of representation of wild species in these genebanks is thus important for evaluating the state of conservation of these species, and for guiding future efforts to fill important gaps in these collections (Maxted et al., 2008; Vimal et al., 2010). Gap analysis methodologies perform such assessments by identifying those wild taxa, geographic locations, and particular traits or adaptations that are un- or under-conserved in genebanks (Maxted et al., 2008) based upon available information resources (i.e., occurrence records from herbaria and genebanks) and using geographic and species distribution modeling methods (Ramírez-Villegas et al., 2010; Parra-Quijano et al., 2011; Castañeda-Álvarez et al., 2015, in press).

Efforts to protect wild plant genetic resources have focused on major cereal crops and legumes, and the CWR of vegetables are still inadequately represented in these collections; only 5% of the accessions in global genebanks correspond to wild relatives of vegetable crops (FAO, 2010). Vegetables are an important source of nutrients, antioxidants, and bioactive compounds, and are essential for a healthy diet and for reducing the risk of some major chronic diseases (Slavin and Lloyd, 2012).

Eggplant, also known as brinjal or aubergine (*Solanum melongena* L., Solanaceae), is an important vegetable crop cultivated worldwide, and is the second most important solanaceous fruit crop after tomato (*S. lycopersicum* L.) (FAOSTAT, 2015). Because of its importance for food security, eggplant is included with 34 other crops in the Annex 1 of the International Treaty on Plant Genetic Resources for Food and Agriculture (Fowler et al., 2003; FAO, 2009), a mechanism established to facilitate access and use of plant genetic resources. Unlike tomato and potato (*S. tuberosum* L.), eggplants have their origins in the Old World (Weese and Bohs, 2010). Three distinct cultivated species of eggplants are known: (1) the Gboma eggplant (*S. macrocarpon* L.); (2) the scarlet eggplant (*S. aethiopicum* L.); and (3) the brinjal eggplant itself (*S. melongena*). While the two former are mainly grown in small scale local agricultural systems in Africa (Daunay and Hazra, 2012), the brinjal eggplant is cultivated worldwide, with a focus in Mediterranean and Asia in particular, where it ranks among the top five most important vegetable crops (Frary et al., 2007). In this article we focus on the brinjal eggplant or aubergine, commonly referred to and referred to hereafter as the eggplant.

Eggplant is an important source of antioxidants, vitamins, and minerals (Raigón et al., 2008; Gramazio et al., 2014). The most important nutritional components of eggplant are phenolic compounds that give the fruit its antioxidant properties (Plazas et al., 2013) which are beneficial for a number of metabolic and cardiovascular ailments. Breeding objectives for eggplant are mostly oriented toward developing high-yielding, early maturing, and high fruit-quality varieties, along with stress resistance and high antioxidants (Singh and Kumar, 2007; Daunay and Hazra, 2012; Gramazio et al., 2014). Breeders have performed research with wild relatives of eggplant for yield increase, fruit quality, disease resistance, and more recently, improved nutritional content (Table 1). However, in contrast to tomato, where most modern cultivars carry genes introgressed from CWR (Díez and Nuez, 2008), at present no modern commercial cultivars of eggplant with traits introgressed from CWR have been released.

All three cultivated eggplant varieties are members of the diverse genus *Solanum* L., which is one of the largest genera of flowering plants and has more than 1000 species distributed worldwide (Bohs, 2005; Särkinen et al., 2013). All eggplants belong to the Leptostemonum clade, a monophyletic group informally called the "spiny" solanums, and within the Leptostemonum clade, a species-rich subclade composed exclusively of Old World taxa (the Old World clade sensu Stern et al., 2011; Vorontsova et al., 2013) from Africa, Australia, and Asia (including Eurasia and the Middle East).

The putative progenitor of *Solanum melongena* is *S. insanum* L., a species widespread in tropical Asia from Madagascar to the Philippines. The eggplant was domesticated in Asia; Vavilov (1951) included eggplant in his Indo-Burman center of diversity but more recently authors have suggested it could have two centers of domestication in China and/or India (e.g., Daunay and Janick, 2007; Wang et al., 2008; Weese and Bohs, 2010), with an additional and independent center of domestication in the Philippines (Meyer et al., 2012a). Eggplant is one of only a handful of crops that are postulated to have multiple centers of domestication (Meyer et al., 2012b). *Solanum insanum* and *S. melongena* share pollinators and freely interbreed where cultivated and wild forms come together (Meyer et al., 2012a; Davidar et al., 2015). Wild populations are distinct from cultivated forms (Meyer et al., 2012a) and in general harbor higher levels of genetic diversity (Mutegi et al., 2015).

The rest of the close wild-relative species of eggplant occur outside the main centers of eggplant domestication, mostly in tropical eastern Africa and the Middle East. Recent taxonomic studies have clarified the species identities and distributions of this complex set of taxa (Knapp et al., 2013; Vorontsova and Knapp, in press), and phylogenetic analyses have clarified their relationships. Previous classification systems (Lester and Hasan, 1991; Daunay and Hazra, 2012) for these taxa used two species names only (*Solanum melongena* and *S. incanum* L.) with a set of groups within each largely defined by geography and delimited by letters (e.g., *S. incanum* group D for what Knapp et al., 2013 defined as *S. lichtensteinii* Willd.). This nomenclature complicated recording of species identities in genebanks with many accessions being identified only to

TABLE 1. Spiny solanums used in eggplant breeding and improvement programs.

Species	Reported use	Reference
S. anguivi Lam.	Backcross to obtain cytoplasm substitution male-sterile lines	Khan and Isshiki, 2011
S. incanum L.	Interspecific hybrids as rootstock	Gisbert et al., 2011
	Backcross introgression of high content in bioactive phenolic acids	Prohens et al., 2013
	Resistance to *Fusarium* wilts.	Rao and Kumar, 1980
S. linnaeanum Hepper & P.-M.L.Jaeger	Backcross introgression of resistance to *Verticillium dahliae*.	Sunseri et al., 2003; Liu et al., 2015
	Development of introgression lines	Mennella et al., 2010
S. sisymbriifolium Lam.	Sexual and somatic hybridization for resistance to *Ralstonia solanacearum* and *Verticillium dahliae*. No backcrosses or selfings of hybrids obtained.	Bletsos et al., 1998; Collonnier et al., 2003a
	Resistance to *Ralstonia solanacearum* and *Verticillium dahliae* wilts	Collonnier et al., 2003a
	Resistance to fruit and shoot borers (*Leucinodes orbonalis*)	Chelliah and Srinivasan, 1985
	Resistance to root-knot nematodes (*Meloidogyne* spp.)	Ahuja et al., 1987
	Resistance to spider mite	Schalk et al., 1975
S. violaceum Ortega	Selection of selfings of interspecific hybrids resistant to *Fusarium* wilt	Rao and Kumar, 1980
	Backcross to obtain cytoplasm substitution male-sterile lines	Khan and Isshiki, 2009
S. torvum Sw.	Sexual and somatic hybridization for resistance to *Ralstonia solanacearum*, *Verticillium dahliae* and root-knot-nematodes. No backcrosses or selfings of hybrids obtained	Bletsos et al., 1998; Jarl et al., 1999; Kumchai et al., 2013
	Resistance to soil pathogens *R. solanacearum* and *Verticillium* and as graftstock	Jarl et al., 1999; Collonnier et al., 2003b; Gisbert et al., 2011
	Resistance to root-knot nematodes (*Meloidogyne* spp.)	Ahuja et al., 1987

species level (for a complete explanation of this system see Knapp et al., 2013).

Previous work in analyzing the relationships of eggplant wild relatives used only a small set of taxa that were morphologically similar to the cultivated species (e.g., Lester and Hasan, 1991; Mace et al., 1999). Molecular phylogenetic studies confirmed the close relationship of these species in what was termed the Eggplant clade, but showed that additional species (e.g., *Solanum linnaeanum* Hepper & P.-M.L.Jaeger) were also members of the monophyletic group (Weese and Bohs, 2007). Further studies using a large number of African and Asian species of spiny solanums (Vorontsova et al., 2013; Aubriot et al., in press) confirmed the monophyly of the Eggplant clade, and also identified several poorly resolved groups that were also possible CWR of eggplants. One of these was called the Anguivi grade by Vorontsova et al. (2013) and comprised a large number of African and Asian species whose relationships were poorly resolved, including the other species of cultivated African eggplants *S. aethiopicum* and *S. macrocarpon* (see Fig. 4 in Vorontsova et al., 2013). The re-evaluation of species limits in eggplant wild relatives in the broadest sense coupled with studies further resolving the phylogenetic relationships of Old World spiny solanums in general means these CWR can be set in both a taxonomic and evolutionary framework that allows for progress to be made in conservation analyses.

Although CWR have had some previous use in eggplant breeding, the comprehensiveness of the conservation and availability of these species in genebanks worldwide has not been assessed. Here we use gap analysis and threat assessment methods to identify gaps in genebank collections, to designate those species and geographic areas most threatened and least well- conserved ex situ, and to propose actions to overcome these deficiencies.

MATERIALS AND METHODS

Genepool concept and selection of species— The genepool concept is used in plant breeding to delineate species based on their ability to cross with the domesticated species, and on the relative fertility of the offspring (Harlan and de Wet, 1971). Because crossability relationships have not been established for numerous crops and their associated wild species, surrogate categories have been created based on taxonomic hierarchy (Maxted et al., 2006), and a combination of evolutionary relationships, ploidy level, and reproductive biology (Wiersema et al., 2012). For eggplant, or other crops that are members of large genera such as *Solanum*, the delimitation of biologically meaningful genepools is challenging due to limited crossability data that are systematically reported (Plazas et al., 2016) as well as the extremely large number of potential genepool members. Here we adopt a combined approach to delimit a meaningful genepool concept for the cultivated eggplant taking into account both relatedness as measured by phylogenetic analyses (Vorontsova et al., 2013; Aubriot et al., in press) and, where available, crossability data (Plazas et al., 2016), as outlined in Genepool concept definition in Results.

Occurrence data— We collected occurrence data for the species we defined as wild relatives of the brinjal eggplant (Table 2). Solanaceae Source (http://solanaceaesource.org/) was used as the principal source of occurrence records derived from herbarium specimen data, because it holds information for taxonomically verified specimens used in recent monographic work and recent field collections (Vorontsova and Knapp, in press). All specimens containing sufficient locality data were manually georeferenced to a high standard and those previously georeferenced by others were checked. Additional herbarium specimen data were retrieved from the Global Biodiversity Information Facility (GBIF; http://www.gbif.org) and were similarly checked for nomenclatural and geographic accuracy. Herbarium accessions are mapped in Fig. 1. Genebank accession passport data from multiple genebanks were electronically retrieved from the Global Information Portal for Plant Genetic Resources (Genesys; https://www.genesys-pgr.org), and through direct requests to genebanks holding collections of eggplants and eggplant wild relatives. All data providers of both herbarium and germplasm accessions are listed in Appendix S1 (see Supplemental Data with the online version of this article).

TABLE 2. Eggplant wild relative species with gap analysis results priority categories, and preliminary IUCN Red List status. Species are ordered alphabetically by epithet. Where species have not been included in molecular phylogenetic analyses, relationships are inferred based on morphological similarity (Vorontsova and Knapp, in press). Species where species distribution models (SDMs) were not constructed due to lack of georeferenced records are indicated with an asterisk (*), and species where models failed due to low AUC values (<0.7) are labeled with a cross (+). Samples available for SDMs were based on georeferenced herbarium records and duplicates within a 1 km^2 were removed. Abbreviations in columns 11 to 15 are defined in the text and are as follows: SRS = sampling representativeness score; GRS = geographic representativeness score; ERS = ecological representativeness score; FPS = final priority score; HPS = high priority species; MPS = medium priority species; LPS = low priority species; NFCR = no further collecting required.

Species of *Solanum*	Phylogenetic group (genepool)	Native distribution	Total no. of herbarium records	Total no. of georeferenced herbarium records	Total no. of germplasm records	Total no. of georeferenced germplasm records	No. of samples available for SDMs	Training AUC (complete)	Test AUC (CrossVal)	SRS	GRS	ERS	FPS	*Ex situ* conservation prioritization category
S. adoense Hochst. ex A.Rich.	Anguivi (GP2)	Africa	78	36	0	0	36	0.769	0.708	0.0	0.0	0.0	0.0	HPS
S. agnewiorum Voronts.	Eggplant (GP2)	Africa	12	8	0	0	9	0.89	0.87	0.0	0.0	0.0	0.0	HPS
S. aldabrense C.H.Wright*, 1	Anguivi (GP2)	Aldabra (Seychelles)	30	10	2	2	10	—	—	0.6	0.0	0.0	0.6	HPS
S. anguivi Lam.	Anguivi (GP2)	Africa and Madagascar	717	514	115	46	514	0.747	0.743	1.4	0.6	8.7	3.6	MPS
S. aureitomentosum Bitter+	Eggplant (GP2)	Africa	29	19	0	0	19	—	—	0.0	0.0	0.0	0.0	HPS
S. burchellii Dunal	Anguivi (GP2)	Africa	36	31	5	4	32	0.896	0.867	1.2	0.4	4.2	1.9	HPS
S. campylacanthum Hochst. ex A.Rich.+	Eggplant (GP2)	Africa	1101	715	66	48	715	—	—	0.6	0.2	7.1	2.6	MPS
S. capense L.	Anguivi (GP2)	Africa	58	45	22	12	45	0.795	0.784	2.8	0.7	5.7	3.1	MPS
S. catombelense Peyr.	Anguivi (GP2)	Africa	64	60	6	5	60	0.817	0.770	0.9	0.2	3.7	1.6	HPS
S. cerasiferum Dunal	Eggplant (GP2)	Africa	64	45	5	3	45	0.832	0.801	0.7	0.0	1.3	0.7	HPS
S. cyaneopurpureum De Wild.	Anguivi (GP2)	Africa	36	22	1	1	22	0.777	0.723	0.3	0.2	2.1	0.9	HPS
S. dasyphyllum Schumach. & Thonn. +	Anguivi (GP2)	Africa	209	174	35	19	174	—	—	1.4	0.0	4.0	1.8	HPS
S. deflexicarpum C.Y.Wu & S.C.Huang*	Anguivi (GP2)	China	3	2	0	0	2	—	—	0.0	0.0	0.0	0.0	HPS
S. glabratum Dunal	Anguivi (GP2)	Arabia	132	79	0	0	79	0.799	0.779	0.0	0.0	0.0	0.0	HPS
S. hastifolium Hochst. ex Dunal	Anguivi (GP2)	Africa	130	95	1	0	95	0.843	0.829	0.1	0.0	0.0	0.0	HPS
S. hovei Dunal	Anguivi (GP2)	India	12	9	0	0	8	—	—	0.0	0.0	0.0	0.0	HPS
S. humile Lam.	Anguivi (GP2)	South Africa	44	35	15	12	35	0.889	0.874	2.5	0.4	5.0	2.7	MPS
S. inaequiradians Werderm.	Anguivi (GP2)	Africa	9	7	0	0	7	0.915	0.838	0.0	0.0	0.0	0.0	HPS
S. incanum L.	Eggplant (GP2)	Africa and Middle East	329	184	114	37	184	0.881	0.878	2.6	0.1	5.3	2.7	MPS
S. insanum L.	Eggplant (GP1)	Southeast Asia and Madagascar	173	78	9	9	78	0.871	0.847	0.5	0.1	1.7	0.7	HPS
S. lamprocarpum Bitter	Anguivi (GP2)	Africa	9	9	1	0	9	—	—	1.0	0.0	0.0	0.5	HPS
S. lichtensteinii Willd.	Eggplant (GP2)	South Africa	90	71	15	14	71	0.789	0.772	1.4	0.3	3.8	1.8	HPS
S. lidii Sunding*	Anguivi (GP2)	Canary Islands	7	3	14	4	3	—	—	6.7	10.0	10.0	8.9	NFCR

continued

TABLE 2, *continued*

Species of *Solanum*	Phylogenetic group (genepool)	Native distribution	Total no. of herbarium records	Total no. of georeferenced herbarium records	Total no. of germplasm records	Total no. of georeferenced germplasm records	No. of samples available for SDMs	Training AUC (complete)	Test AUC (CrossVal)	SRS	GRS	ERS	FPS	*Ex situ* conservation prioritization category
S. linnaeanum Hepper & P.-M.L.Jaeger	Eggplant (GP2)	South Africa; Mediterranean	142	110	75	28	110	0.939	0.930	3.5	0.6	2.9	2.3	HPS
S. litoraneum A.E.Gonç.	Anguivi (GP2)	Africa	4	4	0	0	4	0.88	0.85	0.0	0.0	0.0	0.0	HPS
S. macracanthum A.Rich.	Anguivi (GP2)	Africa	29	19	0	0	19	0.881	0.785	0.0	0.0	0.0	0.0	HPS
S. malindiense Voronts.	Anguivi (GP2)	Africa	16	10	0	0	10	0.966	0.964	0.0	0.0	0.0	0.0	HPS
S. mauense Bitter	Anguivi (GP2)	Africa	47	42	1	0	42	0.870	0.854	0.2	0.0	0.0	0.1	HPS
S. multiflorum Roth*	Anguivi (GP2)	India	13	5	0	0	5	—	—	0.0	0.0	0.0	0.0	HPS
S. nigriviolaceum Bitter	Anguivi (GP2)	Africa	33	25	3	2	32	0.892	0.877	0.8	10.0	10.0	6.9	LPS
S. platacanthum Dunal	Anguivi (GP2)	Arabia	18	14	1	1	14	—	—	0.5	7.6	10.0	6.0	LPS
S. polhillii Voronts.	Anguivi (GP2)	Africa	56	30	0	0	25	0.912	0.881	0.0	0.0	0.0	0.0	HPS
S. richardii Dunal	Climbing (GP2)	Africa and Madagascar	119	77	3	2	77	0.725	0.677	0.2	0.0	10.0	3.4	MPS
S. rigidum Lam.[2]	Eggplant (GP2)	Cape Verde Islands	33	32	0	0	25	0.967	0.963	0.0	0.0	0.0	0.0	HPS
S. rubetorum Dunal	Anguivi (GP2)	Africa	56	38	9	1	30	0.851	0.810	1.4	0.2	2.1	1.2	HPS
S. ruvu Voronts.*[3]	Anguivi (GP2)	Africa	1	1	0	0	1	—	—	0.0	0.0	0.0	0.0	HPS
S. setaceum Dammer	Anguivi (GP2)	Africa	39	29	0	0	29	0.79	0.75	0.0	0.0	0.0	0.0	HPS
S. sisymbriifolium Lam.	NW-Sisymbrifolium (GP3)	South America (native); widespread	151	109	68	23	109	0.806	0.778	3.1	0.0	2.8	2.0	HPS
S. sodomeodes Kuntze	Anguivi (GP2)	Africa	21	21	0	0	21	0.841	0.783	0.0	0.0	0.0	0.0	HPS
S. stipitatostellatum Bitter	Climbing (GP2)	Africa	56	34	0	0	38	0.872	0.800	0.0	0.0	0.0	0.0	HPS
S. supinum Dunal	Anguivi (GP2)	Africa	24	19	5	2	19	0.854	0.816	1.7	0.1	0.6	0.8	HPS
S. taitense Vatke	Anguivi (GP2)	Africa	43	33	0	0	33	0.819	0.744	0.0	0.0	0.0	0.0	HPS
S. tomentosum L.+	Anguivi (GP2)	South Africa	56	25	32	17	34	—	—	3.6	1.4	9.2	4.7	MPS
S. torreanum A.E.Gonç.	Anguivi (GP2)	Africa	11	10	0	0	10	—	—	0.0	0.0	NA	0.0	HPS
S. torvum Sw.	NW-Torva (GP3)	Central America (native?); widespread	821	540	125	55	540	0.815	0.816	1.3	0.1	4.5	2.0	HPS
S. umtuma S.Knapp & Voronts.	Eggplant (GP2)	South Africa	22	16	0	0	16	0.789	0.765	0.0	0.0	0.0	0.0	HPS
S. usambarense Bitter & Dammer	Anguivi (GP2)	Africa	36	32	0	0	31	0.825	0.785	0.0	0.0	0.0	0.0	HPS

continued

TABLE 2, *continued*

Species of *Solanum*	Phylogenetic group (genepool)	Native distribution	Total no. of herbarium records	Total no. of georeferenced herbarium records	Total no. of germplasm records	Total no. of georeferenced germplasm records	No. of samples available for SDMs	Training AUC (complete)	Test AUC (CrossVal)	SRS	GRS	ERS	FPS	*Ex situ* conservation prioritization category
S. usaramense Dammer	Anguivi (GP2)	Africa	23	15	0	0	15	0.891	0.890	0.0	0.0	0.0	0.0	HPS
S. vespertilio Aiton	Anguivi (GP2)	Canary Islands	19	9	18	8	9	0.812	0.751	4.9	10.0	5.0	6.6	LPS
S. viarum Dunal	NW-Acanthophora (GP3)	Brazil (native); widespread	46	34	85	45	34	0.901	0.879	6.5	0.1	4.2	3.6	MPS
S. violaceum Ortega	Anguivi (GP2)	India and Southeast Asia	237	108	84	41	108	0.792	0.776	2.6	0.2	4.6	2.5	HPS
S. zanzibarense Vatke	Climbing (GP2)	Africa	74	58	1	1	58	0.870	0.845	0.1	0.1	1.5	0.6	HPS

Notes: [1] species native to the small Aldabra Islands in the Indian Ocean, for which no global environmental data recorded and hence no SDM calculated; [2] values excluding Caribbean range where species is likely to have been introduced and is now no longer present; [3] species only known from the type.

Species distribution modeling—Georeferenced herbarium records were used as occurrence samples to build species distribution models (SDMs) for species with more than five (5) unique georeferenced records (Wisz et al., 2008). Due to taxonomic uncertainties in genebank records, only verified herbarium specimen data were used to construct SDMs. All available environmental data layers were obtained from the WorldClim database, version 1.4 (http://www.worldclim.org; Hijmans et al., 2005), along with potential evapotranspiration (PET), actual evapotranspiration (AET), and aridity from the CGIAR Consortium for Spatial Information (CGIAR-CSI; http://www.cgiar-csi.org), and enhanced vegetation index (EVI) derived from MODIS imagery from WorldGrids (http://www.worldgrids.org). All the data layers were obtained at a spatial resolution of 30 arc second (~1 km at the equator). Variable selection was based on a combination of cluster, correlation, and principal components analyses, as well as expert judgement based on the ecology of the species. Our final variable selection was driven by choosing variables that were ecologically relevant for eggplant wild relatives, following recommendations for best practice (Elith and Leathwick, 2009). The final selection of six environmental variables were: (1) temperature seasonality; (2) annual mean temperature; (3) the precipitation of the coldest quarter (three-month period); (4) enhanced vegetation index (EVI); (5) actual evapotranspiration (AET); and (6) water deficit (calculated as potential evapotranspiration minus actual evapotranspiration; Stephenson, 1998).

MaxEnt (version 3.3.3; Phillips et al., 2006) was used to build species distribution models as it has been found to be among the best-performing of the various correlative SDM approaches available, and because it is able to perform well with small samples of presence localities (Elith et al., 2006; Pearson et al., 2007). We adopted the default regularization parameters but restricted MaxEnt to using only linear and quadratic functional forms, which constrains models to produce relatively simple models that do not over-fit to the training data (Merow et al., 2013; Syfert et al., 2013). Sampling bias was controlled by including a sampling bias dataset (Phillips et al., 2009) constructed from all georeferenced plant occurrence data from both the GBIF data portal and all georeferenced records from Solanaceae Source (6,200,000 and 47,500 records respectively; downloaded 24 September 2014). The spatial extent of the area modeled also influences model performance (Peterson et al., 2011). In our case, species ranges varied from very small within a single country (e.g., *Solanum deflexicarpum* C.Y.Wu & S.C.Huang) to widespread on multiple continents (e.g., *S. torvum* Sw.). Given the influence of the spatial extent of the area modeled on model performance, we built models with varied extents that fit species ranges to allow for a biologically meaningful fit between a species' occurrence and the associated environmental variables. A 200 km buffer around all presence points was used to delimit the model extent of each species following VanDerWal et al. (2009). All species with >10 georeferenced samples were modeled using a 5-fold cross-validation approach to assess model predictive accuracy. This approach uses 80% of the data to train the model and reserves 20% for model evaluation; this process was repeated until each reserved set had been used to evaluate models (Franklin, 2009). A leave-one-out method was performed for species with sample sizes from 5 to 10 data points (Pearson et al., 2007). SDM performance was evaluated by using the area under the curve (AUC) in a receiver operating characteristic (ROC) plot; an AUC value of 1.0 indicates perfect discrimination ability and a value of 0.5 or less indicates a prediction

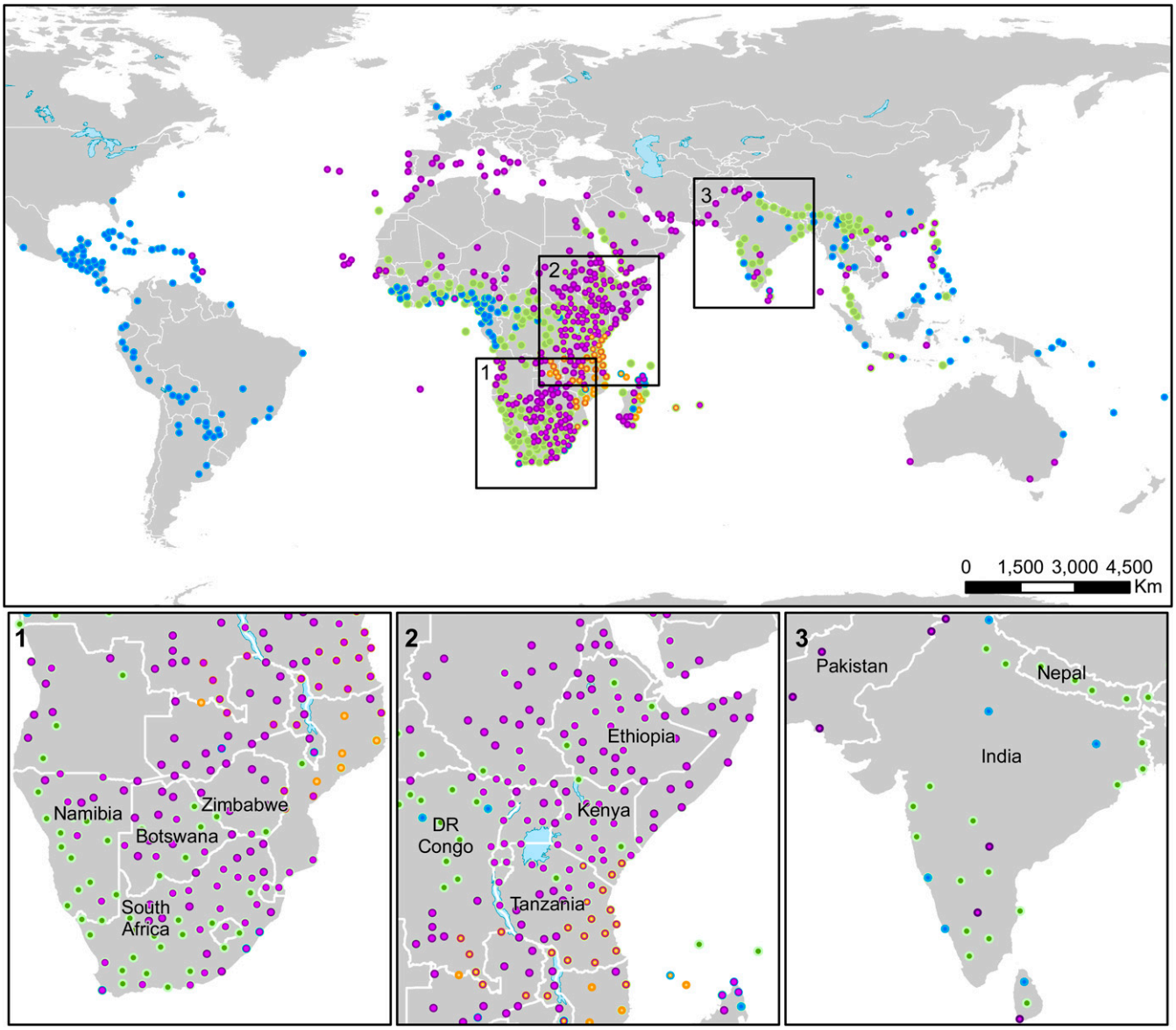

FIGURE 1 Map of herbarium specimens with geographical coordinates of spiny solanums (Leptostemonum Clade) used in this study by phylogenetic group: Purple points = Eggplant clade; green points = Anguivi grade; orange points = Climbing clade; blue points = New World species.

no better than random. Only models with an average test AUC above 0.7 were used for further analyses. In the case that the SDM did not have an average test AUC above 0.7, a convex hull was created and used in place of the SDM for the rest of the ex situ gap analysis. For species with AUC above 0.7, SDMs were then reconstructed using the complete set of occurrences to develop a more robust predictive model (Fielding and Bell, 1997) and these models were used for further analyses. Binary maps of predicted presences and absences were created for each species using a maximum sensitivity and specificity threshold (Liu et al., 2005, 2013).

Ex situ conservation analysis—We used three species-specific gap analysis metrics adapted from Ramírez-Villegas et al. (2010) to estimate the sufficiency of accessions currently conserved in genebanks in representing the full diversity of CWR species based on their potential distributions and the known wild diversity by using data derived from herbarium collections; the sampling representativeness score (SRS), geographic representativeness score (GRS), and ecological representativeness score (ERS). Each metric has a scale from 0 to 10. These metrics measure the representativeness of the germplasm collections in relation to the modeled distribution of each crop wild relative species, and the known wild diversity using data derived from herbarium collections. The SRS is a general measure of how well germplasm records represent the known wild diversity in terms of numbers of records, and is calculated by dividing the number of germplasm records by the total number of records (germplasm + herbarium). An advantage of the SRS metric is that it permits the use of all records, regardless of whether they have geographical coordinates. The GRS focuses on the geographic representativeness of germplasm collections, and is measured by

comparing the current geographic coverage of the germplasm records calculated with a 50 km circular buffer surrounding the original site of collections of each germplasm accession record (Hijmans and Spooner, 2001) against the SDM of each species. The ERS measures ecological representation of germplasm collections, and is calculated by comparing the number of terrestrial ecoregions (Olson et al., 2001) represented in the geographic coverage of the germplasm collections (buffer of 50 km radius around each collection) against the total number of ecoregions encompassed in the SDM of each species. The metrics were averaged to produce a final priority score (FPS), which was then classified into four categories of urgency for further collecting to improve representation in genebank collections: (1) high priority for further collecting (FPS = 0-3); (2) medium priority (FPS = 3.01-5); (3) low priority (FPS = 5.01-7.5; and (4) sufficiently represented (FPS = 7.51-10). Species with no germplasm accessions conserved in genebanks (with or without geographic coordinates) received a final score of 0 and were listed as high priority species.

We produced maps to identify areas prioritized for further collecting of eggplant CWR seeds by excluding the original sites of collection of existing germplasm accessions with coordinates (50 km buffer around each record) from the geographic coverage derived from SDMs. When an appropriate SDM was not produced, 50 km circular buffers around each herbarium record were used as proxies for geographic coverage. Remaining geographic areas represent areas where a species is potentially present but has not yet been collected (i.e., collecting gaps). Individual maps of species from the high and medium priority categories were added together to produce a priority collecting hotspots map. Maps illustrating collecting hotspots were created for the following groups of species: (1) Eggplant clade; (2) Anguivi grade; (3) Climbing clade; and (4) the New World species.

All gap analyses were performed in R version 2.15 using maptools (Bivand and Lewin-Koh, 2014), rgdal (Bivand et al., 2014), SDMTools (VanDerWal et al., 2014), raster (Hijmans, 2014), sp (Pebesma and Bivand, 2005; Bivand et al., 2013), and dismo (Hijmans et al., 2014) packages. All raster files were mapped in ESRI ArcGIS 10 (ESRI, 2013).

In situ conservation assessment—We evaluated the preliminary conservation status of the wild relatives of eggplant using the IUCN Red List Categories and Criteria (IUCN, 2012), one of the most widely used systems for assessing the conservation status of species in their natural habitats (Rodrigues et al., 2006; Mace et al., 2008). Threatened species were placed within three categories (critically endangered, endangered, and vulnerable) based on five criteria (A-E; IUCN, 2012). A fourth category, near threatened, was used when species partially met the IUCN threat criteria and were present in areas with documented rapid change in land use. Because population data were lacking for the eggplant wild relatives, several of the Red List criteria were inapplicable and we focused on using estimates of geographic range as the key determinant of extinction risk (criterion B). The extent of occurrence (EOO), which measures the overall geographic spread of the localities in which a species is found (Gaston and Fuller, 2009), was calculated by applying a minimum convex polygon (MCP; also known as a convex hull) around the known species locations (IUCN, 2012). Area of occupancy (AOO), the other geographical range element to criterion B, measures the area with in the EOO that is occupied by a taxon. We focused on EOO rather than AOO because not enough information was available from herbarium specimen records to know whether populations were well-sampled within the species range. AOO calculations were only performed for those species with a preliminary status indicating some degree of threat (Table 3). The effect of land use changes for each species was assessed in Google Earth (http://www.google.co.uk/earth/) using current and historical land use images to visually determine fragmentation within the EOO due to human activities such as mining, agriculture, and urban habitation. Following IUCN (2012) species were considered threatened if the EOO was < 20,000 km^2 (< 100 km^2 = critically endangered, < 2000 km^2 = endangered, and < 20,000 km^2 = vulnerable) and if a minimum of two of the following subcategories were met: (1) severely fragmented habitat; (2) continuing decline of habitat; and (3) extreme range fluctuations (see Appendix S2 for details of IUCN criteria).

We also examined the intersection of threatened species of eggplant CWR and protected areas to further refine our prioritization of collecting and conservation action. First, we focused on regions with priority collecting hotspots (i.e., those in which more than seven under-collected species were estimated to occur in the same pixel) and evaluated regions within and outside of protected areas as defined from the World Database on Protected Areas (WDPA, 2010). We also prepared a map of hotspots for further collecting of those species preliminarily assessed at risk for extinction by overlaying the collecting priority map for each species identified in one of the threatened categories (including those in the near threatened category). To further prioritize regions for conservation actions, the extent of area in collecting hotspots for species at risk for extinction was combined with the overall area for general priority collecting hotspots to identify regions with multiple species poorly represented in genebanks that coincide with regions where there are also significant conservation threats.

TABLE 3. IUCN threat assessments for eggplant wild relatives at risk for extinction. Extent of Occurrence (EOO) and Area of Occupancy (AOO) calculations described in the text. All listings are new here with the exception of *Solanum lidii* (formally assessed by Rodríguez Delgado et al., 2011). EOO and AOO values for all species are in Appendix S3 (see Supplemental Data with the online version of this article) and all occurrence data used in these calculations are archived on the NHM Data Portal (http://dx.doi.org/10.5519/0006757).

Species of *Solanum*	EOO (km^2)	AOO (km^2)	Prelim IUCN category
S. agnewiorum Voronts.	7946	32	VU B1ab(iii)
S. aldabrense C.H.Wright	3437	32	EN B1ab(iii)
S. deflexicarpum C.Y.Wu & S.C.Huang[1]	0	8	CR B1ab(iii)
S. inaequiradians Werderm.	2101	28	EN B1ab(iii)
S. lidii Sunding[2]	<5	5	CR B2ab(ii,iii) +C2a(i)
S. litoraneum A.E.Gonç.	6196	16	VU B1ab(iii)
S. malindiense Voronts.	2381	40	EN B1ab(iii)
S. nigriviolaceum Bitter	43 749	16	NT
S. platacanthum Dunal	45 123	56	NT
S. rigidum Lam.	135 509 (35 310)[3]	124 (108)[3]	LC (NT)[3]
S. ruvu Voronts.	0	4	EX [?]
S. setaceum Dammer	13 564	112	VU B1ab(iii)
S. torreanum A.E.Gonç.	16 804	40	VU B1ab(iii)
S. vespertilio Aiton	2336	36	EN B1ab(iii)

Notes: [1] EOO could not be calculated because there are only two distinct localities; [2] AOO value is as reported from Rodríguez Delgado et al. (2011); [3] values excluding Caribbean range where species is likely to have been introduced and is now no longer present.

RESULTS

***Genepool concept definition*—**Based on the existing phylogenetic hypotheses and information on crosses between the wild relatives and the cultivated eggplant, we included 52 eggplant CWR species in our eggplant wild relative genepool concept (Table 2). Species were assigned to clades, grades, and genepools based on morphological evidence when they had not been included in a previous phylogenetic analysis (Table 2; Vorontsova and Knapp, in press). Genepool 1 (GP1) was defined as the cultivated eggplant and its widely accepted wild progenitor *Solanum insanum* (Knapp et al., 2013), following Maxted et al.'s (2006) concept. Genepool 2 (GP2) was defined based on phylogenetic evidence from Vorontsova et al. (2013), where the eggplant and its closest relatives are resolved in a large monophyletic clade, that includes three elements: (1) the monophyletic Eggplant clade that includes the most closely related species (excluding *S. insanum* and *S. melongena* itself); (2) the small climbing clade; and (3) the set of species from the paraphyletic Anguivi grade where relationships among the component species are not well-resolved. Lastly, genepool 3 (GP3) is defined to include phylogenetically more distantly related species native to the New World which have already been used in eggplant breeding due their resistance traits (e.g., *S. sisymbriifolium* Lam., *S. torvum*). The Eggplant clade included 10 species, the Climbing clade three species, the Anguivi grade 36 species, and the three remaining species were included in a nonphylogenetically defined group we here call the New World group (Table 2). *Solanum anguivi* Lam. (wild ancestor of cultivated *S. aethiopicum*) and *S. dasyphyllum* Schumach. & Thonn. (wild ancestor of cultivated *S. macrocarpon*) are included in our analysis as members of the Anguivi grade.

***Occurrence data*—**A total of 5618 occurrence records from herbarium specimens and 936 germplasm accession records were compiled. Records with geographic coordinates ranged from 1 to 1100 per species (Table 2). All herbarium specimen records for the eggplant genepools defined above are mapped in Fig. 1. All occurrence data are available for download from the Natural History Museum Data Portal (http://dx.doi.org/10.5519/0006757).

***Species distribution models*—**SDMs with adequate predictive performance to use in the gap analysis were produced for 37 CWR species [median test AUC was 0.83 ± 0.06 SD (from 5-fold cross-validated models)]. SDMs could not be built for five species with fewer than five georeferenced records (*Solanum aldabrense* C.H.Wright, *S. deflexicarpum, S. lidii* Sunding, *S. litoraneum* A.E.Gonç., and *S. ruvu* Voronts.). A further seven species were not included even though they had slightly higher record numbers (5-25) because AUC values were below our targeted threshold of 0.7 (Table 2). While it is expected that sample sizes of less than 25 would yield SDMs with poor predictive performance, we also found that three species with very large numbers of records (*S. campylacanthum* Hochst. ex A.Rich. [n = 715], *S. dasyphyllum* [n = 174] and *S. richardii* Dunal [n = 77]) also yielded SDMs with poor predictive performance. *Solanum campylacanthum* and *S. dasyphyllum* are both very widely distributed in many different habitats in continental Africa, and it is likely that climatic factors are not of primary importance in determining their distributions. It is possible that species with poor SDM performance have distributions that are constrained to factors (e.g., dispersion or biotic interactions) beyond the variables included to build the SDMs. For instance, a recent field study in the Kenyan savannah found that *S. campylacanthum* abundance greatly decreased in plots where several mammals browsed on different parts of the plants (Pringle et al., 2014), suggesting biotic interactions could be important factors influencing the distribution of this species.

***Ex situ conservation analysis*—**The wild relatives of eggplant are currently exceedingly poorly represented in global germplasm collections (Table 2). The final priority scores (FPS) for the gap analysis classified 40 species as high priority for further collecting to improve the state of their conservation in genebanks, eight (8) as medium priority for further collecting to improve the state of their conservation in genebanks, three as low priority for further collecting, and only one species as currently sufficiently represented in genebanks (Table 2). A total of 34 of the 52 species (65%) have five or fewer accessions publicly available in genebanks.

The sampling representativeness score (SRS) influenced the assessment of the low level of representation of eggplant CWR species in genebanks, with many species represented by very few genebank accessions (Appendix S4a). Island species with narrow distributions (i.e., *Solanum lidii* and *S. vespertilio* Aiton) were an exception, with high SRS (Table 2). Species like *S. campylacanthum* had the largest number of herbarium specimen records, but also one of the largest gaps with respect to the number of germplasm collections (1101 herbarium records as opposed to 66 germplasm accessions). This mismatch results from the use of revised species names and delimitations for verified herbarium material (Vorontsova and Knapp, in press), and the outdated naming of accessions in genebanks with a previous set of species names (based on Lester and Hasan, 1991).

The low level of representation of the wild relatives of eggplant in genebanks was also strongly influenced by the inadequacy of geographic coverage of samples (i.e., geographic locations where germplasm has already been collected as compared to the total modeled distributions of eggplant CWR) (Appendix S4b). Forty seven of the 52 species used in our analyses had a geographic representativeness score (GRS) of less than 2 (Table 2). Only species with relatively restricted distributions had a GRS above 7.5 (*Solanum lidii*, *S. vespertilio*, *S. platacanthum* Dunal, and *S. nigriviolaceum* Bitter*);* these species were also assessed as relatively well-represented with respect to ecological regions (ERS; ERS ≥ 5).

Ecological representation (ERS scores) of eggplant wild relatives is somewhat higher than either the SRS or GRS, but still falls below the ideal (see Appendix S4c). Twenty four of the 52 species analyzed had an ERS of < 2. Some widespread species such as *Solanum anguivi*, *S. richardii*, and *S. tomentosum* L. displayed a relatively comprehensive ecological representativeness (ERS >8). Caution should be taken with this result however, the terrestrial ecoregions defined by Olson et al. (2001) are relatively crude, and more tightly defined vegetation types that more accurately represent the diversity of terrestrial biomes (e.g., Särkinen et al., 2011, as done for tropical dry forests in South America) could significantly change this representation.

Regions identified as hotspots for further collecting of multiple medium and high priority species included eastern and southern Africa and southeast Asia (Fig. 2). Several areas along the Tanzania-Kenya border have as many as 10 species occurring in the same area and were identified as high or medium priority collecting areas (Fig. 2, inset 2). Further hotspots include western South Africa (Fig. 2, inset 1) and the region of the Western Ghats in India (Fig. 2, inset 3).

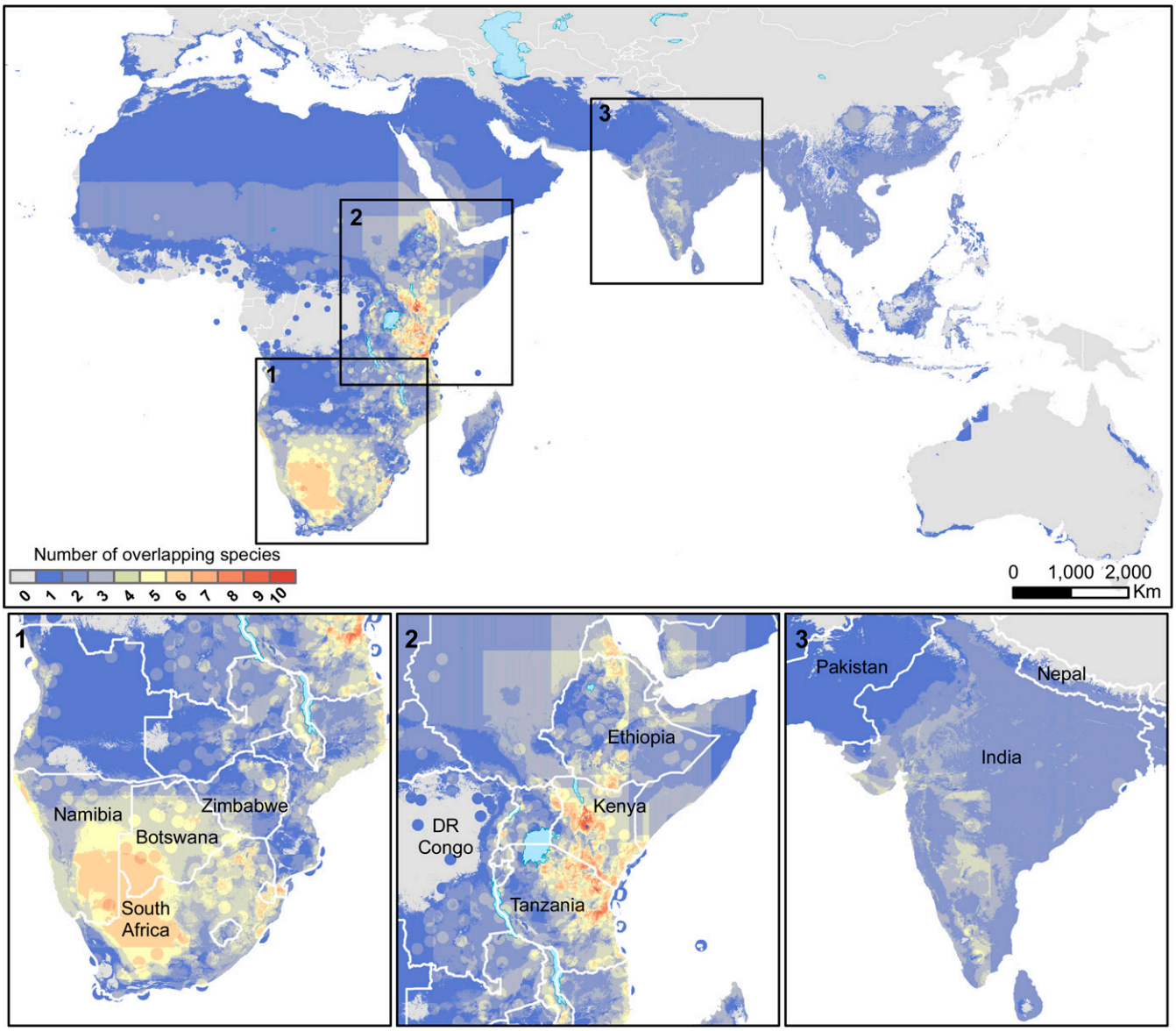

FIGURE 2 Map of future collecting hotspots for 48 species of cultivated eggplant wild relatives classified as medium or high priority based on the gap analysis. Areas with orange-red color are highest priority collecting hotspots; these include more than seven species. Insets zoom in on regions identified as 1 (southwestern Africa), 2 (eastern Africa), and 3 (India) on the larger scale map.

These "secondary" hotspots are even clearer when the data are analyzed by phylogenetic grouping. The Eggplant clade, which is comprised of the 10 species most closely related to the cultivated eggplant, has priority areas with the greatest number of overlapping species for further germplasm collecting in Kenya and Tanzania, and with a small area to the south in Zimbabwe (Appendix S5a). Western South Africa and Namibia is identified as an additional priority area for further collecting for the Anguivi grade (Appendix S5b), and areas in coastal eastern Africa from Kenya south to Mozambique are priorities for the species of the Climbing clade (Appendix S5c).

In situ conservation assessment—*Solanum lidii* is the only species of the 52 eggplant wild relatives that has been previously formally assessed and included on the IUCN Red List (critically endangered; Rodríguez Delgado et al., 2011). We have assessed an additional nine species as threatened (EOO < 20,000 km^2) according to preliminary Red List criteria due to small geographic ranges (Fig. 3) and occurrence in areas with high habitat fragmentation and degradation (IUCN criterion B1ab (iii); Table 3). Our analysis shows that one additional species is likely to be critically endangered (*S. deflexicarpum*), four species endangered (*S. aldabrense, S. inaequiradians* Bitter, *S. malindiense* Voronts., and *S. vespertilio*), and four species vulnerable (*S. agnewiorum* Voronts., *S. litoraneum*, *S. setaceum* Bitter, and *S. torreanum* A.E.Gonç.; Table 3). In addition to the 10 eggplant wild relatives assessed as threatened, two species with EOOs greater than 20,000 km^2 are likely to be near threatened (*S. platacanthum* and *S. nigriviolaceum)* based on their occurrence in regions with high land degradation. A third species, *S. rigidum* Lam., could also be considered to be near threatened because it has

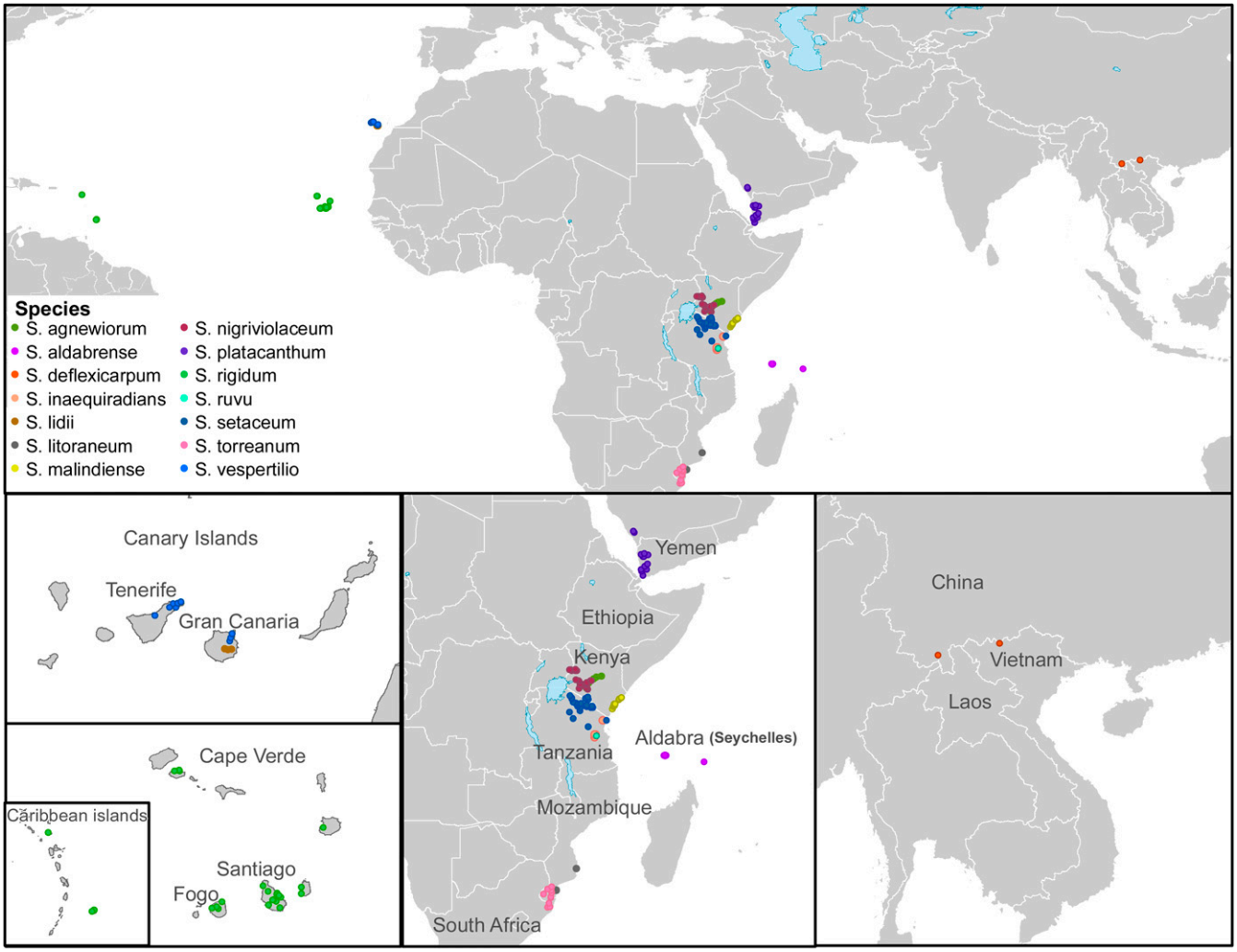

FIGURE 3 Map of georeferenced specimens of eggplant wild relatives identified as at risk for extinction. One species (*Solanum lidii*) has been formally assessed as critically endangered, nine species (*S. agnewiorum*, *S. aldabrense*, *S. deflexicarpum*, *S. inaequiradians*, *S. litoraneum*, *S. malindiense*, *S. setaceum*, *S. torreanum*, *S. vespertilio*) were identified in this study as threatened , three (*S. nigriviolaceum*, *S. platacanthum*, *S. rigidum*) as near threatened according to the preliminary IUCN Red List assessment criteria, and one (*S. ruvu*) has been previously suggested to be extinct in the wild.

been recently shown to be endemic to the Cape Verde Islands although there are a few old specimens from the Caribbean, possibly dispersed by humans (Knapp and Vorontsova, 2013). The most poorly known species of the eggplant wild relatives is *S. ruvu* that is only known from the type and has been suggested to be extinct in the wild; recent collecting trips to the type locality and similar habitats have failed to encounter extant populations (Vorontsova and Mbago, 2010). In total, one quarter of the species of eggplant CWR (14 spp.) are assessed to have some level of threat to their wild populations (see Table 3), while the rest (38 spp.) are assessed here as of least concern, meaning their populations are widespread and not of immediate conservation concern (see Table 3 for EOO and AOO values for all threatened species and Appendix S3 for EOO values for all species).

Five of the species identified at risk for extinction are restricted to eastern Africa, primarily in Kenya, Tanzania, and Uganda (Fig. 3). This is also the region identified as containing multiple species under-represented in ex situ collections, particularly along the border between Kenya and Tanzania (Fig. 4A). Only 15% (ca. 10,000 km^2 of a total 65,000 km^2) of hotspot areas identified as high priority for further collecting (i.e., seven or more eggplant wild relatives occurring in one pixel) and also containing Red List species are protected (Fig. 4B). More than half of these target collecting priorities occur in the hot semiarid habitats generally known as the southern *Acacia-Commiphora* bushlands and thickets ecoregion (Olson et al., 2001) and are the southern edge of White's (1983) Somali-Masai *Acacia-Commiphora* deciduous bushland and thicket phytochorion. More recent satellite images show that a significant proportion of habitats in this region have been converted to croplands (European Space Agency, 2009) and that vegetation today is a complex and fragmented mosaic.

DISCUSSION

Defining genepools in large genera such as *Solanum* is challenging and we have approached this with a compromise solution that takes into account phylogenetic relationships and crossability (Vorontsova

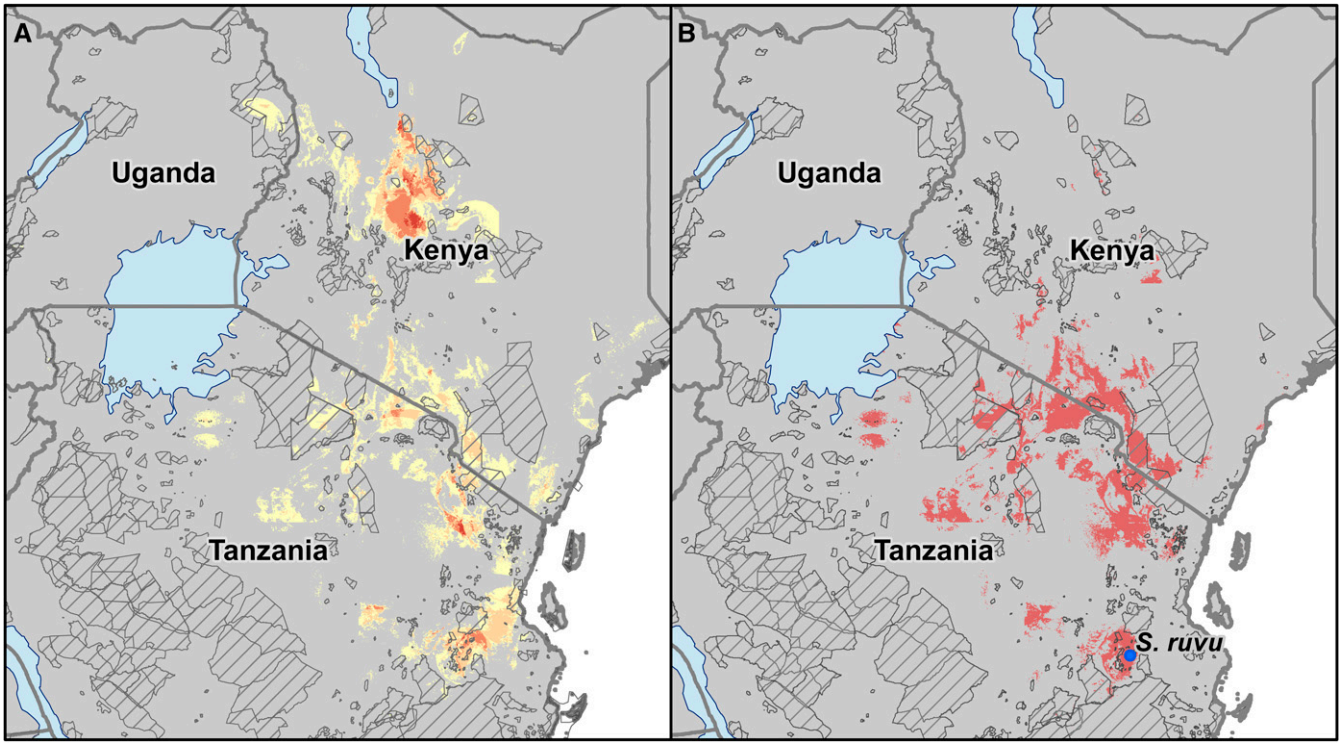

FIGURE 4 Collecting priority hotspots in relation to protected areas in eastern Africa. (A) Collecting priority hotspots (7 or more species per pixel); and (B) collecting priority hotspots with at least one species that has a threatened preliminary status using IUCN Red List criteria. The georeferenced type specimen for *Solanum ruvu* (possibly extinct in the wild) is included. Hatched polygons represent protected areas (WDPA, 2010).

et al., 2013; Plazas et al., 2016). The concept presented here uses primarily phylogenetic relationships to define GP2, and GP3 includes less closely related but useful species which have previously been used in eggplant breeding but that thus far cannot be easily crossed with *S. melongena* (Plazas et al., 2016). Our newly suggested genepool concept for eggplant includes one wild species (the wild progenitor of the eggplant, *S. insanum*) in GP1, 48 species in GP2, and three more widespread weedy species with native distributions in the New World in GP3 (see Table 2). Our revised genepool based on phylogenetic and taxonomic criteria better represents the true wild diversity of eggplant CWR, and provides an initial framework for current and future breeding efforts. A more systematic approach to record results of trials of interspecific crosses between the cultivated eggplant with its wild relatives can help in producing a refined genepool concept more aligned with the Harlan and de Wet (1971) approach, that would be widely useful in the plant breeding community. Our results here focus on the wild species, but in eggplants and many other crops the dividing line between wild and cultivated can be difficult to determine (see Davidar et al., 2015).

The use of a broad genepool concept is potentially important to best capture traits desirable for crop improvement that may be present in species not closely related to or directly crossable with the crop. In this case, the selection of those distantly related, but potentially useful taxa might best look at weedy, invasive species, whose resistance to both biotic and abiotic stresses is likely to be well-developed (Daunay and Hazra, 2012). Our use of three New World species of spiny solanums that are distantly related to the eggplant is an illustration of this concept. We deliberately chose species that had been used in eggplant breeding, and that are also widespread outside their putative native ranges. Widespread species may hold a wealth of genetic variation and adaptive traits and thus may be of great significance to breeders in the long run, especially in light of new techniques for hybridization and gene editing.

Despite the recent taxonomic revision (Knapp et al., 2013) of species-level identities and names for members of the Eggplant clade, taxonomic challenges persist within wild eggplants. A systematic specimen identification of eggplant relatives has been carried out and it is reflected in the occurrence records of herbarium samples derived from Solanaceae Source. Germplasm collections, however, clearly require re-identification, in part due to the revised species circumscriptions of Knapp et al. (2013) but also because this information is not frequently revised or updated. For instance, *Solanum campylacanthum* accessions are often named as *S. incanum* (groups A and B) in genebanks, but we found many genebanks did not include the letter designation that would allow us to directly compare identifications (see also Meyer et al., 2012a, where all close wild relatives were labeled as *S. incanum*). Consequently it is likely that germplasm collections are underestimated for *S. campylacanthum* and overestimated for *S. incanum*.

Overall, the gap analysis reveals significant gaps in genebank holdings of eggplant CWR with 40 out of 52 species ranked as medium or high priority for future collecting for ex situ conservation. Most of the species currently considered of high value for eggplant breeding (Daunay and Hazra, 2012; Rotino et al., 2014) are currently poorly represented in genebank collections. Improving the current representativeness of eggplant CWR in genebanks is vital

for conserving the genetic diversity in these species, while at the same time widening source options for mining of valuable traits by eggplant breeders.

With large geographic regions identified as priorities for further collecting, our additional analyses to refine the collecting priorities based on IUCN Red List threat assessment categories can help to prioritize the most threatened species (see final column in Table 2). Geographic collecting gaps often coincide with areas known to have degraded habitats due to conversion to agriculture and urbanization (Vorontsova, 2010), so geographic and species collecting gaps can also directly inform strategies for the management of threatened species in situ. Of all prioritized areas, the Tanzania-Kenya border is remarkable in possessing as many as 10 species considerably under-represented in genebanks, of which four are among the closest relatives of the cultivated eggplant and five are at potential risk of extinction.

Three important facets for future collecting of eggplant CWR for conservation are of importance in prioritization of germplasm acquisition. Firstly, those CWR species that are threatened in their native ranges; secondly, those whose populations harbor significant genetic diversity; and thirdly, those species whose biology and distribution suggests they might harbor traits of significance for eggplant breeding in the future.

Of the several eggplant wild relatives identified here as threatened with extinction, the only one formally listed as critically endangered on the Red List is the Canary Island endemic *Solanum lidii* (Rodríguez Delgado et al., 2011; see Fig. 3). The population size and genetic diversity of extant populations of *S. lidii* is limited, which may compromise their medium term viability (Prohens et al., 2007). Fortunately 14 accessions of the species are conserved in genebanks (see Table 2), and wild populations occur in the Reserva Natural Especial de Los Marteles on the island of Gran Canaria (Rodríguez Delgado et al., 2011), where it may be able to continue to evolve in its natural habitat. The design and implementation of management plans for its in situ conservation will be vital to secure the long-term conservation of this species. Conversely, both *S. deflexicarpum* and *S. agnewiorum* have small range sizes, occur in areas of high habitat degradation, but have no germplasm accessions in genebanks. *Solanum deflexicarpum* is a member of the Anguivi grade closely related to *S. violaceum* (Aubriot et al., in press), a species previously used in eggplant breeding (see Table 1) and is here assessed as endangered. With a very small range size and a fragmented distribution (see Fig. 3) *S. deflexicarpum*, despite being less threatened as assessed using the IUCN criteria, is clearly important for future collecting and use in crop improvement. The same holds true for *S. agnewiorum*, only recently described from eastern Africa. This forest species is a surprising member of the Eggplant clade (Vorontsova et al., 2013; Aubriot et al., in press) and is thus a prime candidate for crossing experiments with the cultivated eggplant.

Widespread species are also important reservoirs of genetic variation for potential use in crop improvement. *Solanum insanum*, the wild progenitor of eggplant, is extremely widely distributed from Madagascar throughout south and southeastern Asia (as far east as the Philippines). Meyer et al. (2012a) working the Asian range of both *S. melongena* and *S. insanum* (identified in their study as *S. incanum*) found considerable genetic diversity within both species that in part paralleled geography, but not completely. Mutegi et al. (2015) worked with populations of *S. insanum* in southern India and showed that populations of the wild species (*S. insanum*) were genetically distinct from the cultivated plants, and were highly structured geographically. They separated *S. insanum* into three major genetic groups and suggested that geographic isolation and limited gene flow have structured the genetics of this species in the study area. This high genetic divergence among populations suggests that further population level sampling across the range of *S. insanum* will uncover genetic novelty worth conserving in genebanks, especially given the lack of reproductive barriers between *S. insanum* and *S. melongena* (Davidar et al., 2015; Plazas et al., 2016).

Solanum campylacanthum is equally widespread in Africa, and occurs in a huge variety of mostly disturbed and human-altered habitats (Vorontsova and Knapp, in press). Leaf shape is particularly variable in this species (see Fig. 2 in Knapp et al., 2013), leading to its having more than 70 synonyms and having been classified as *S. incanum* group A (wider leaves) and *S. incanum* group B (narrower leaves) in previous classification systems (Lester and Hasan, 1991). Morphological variability of this magnitude is likely to be mirrored in genetic variability, although this is not always the case (Kardos et al., 2015). This again suggests that for conservation in genebanks, populations across the range of *S. campylacanthum* should be collected. Our analyses of hotspots of collecting gaps for all species identify an efficient way in which to begin such a challenging collecting program.

Weeds are plants that are successful at colonizing new areas and often are those found in disturbed areas. Many of the eggplant CWR analyzed here could be classified as weeds: *Solanum campylacanthum*, *S. insanum*, and *S. anguivi* all have very large distributions in many different sorts of disturbed and semidisturbed habitats (Vorontsova and Knapp, in press). Weediness facilitates commensalism with humans, ease of establishment, and possibly also crossability with domesticates, but not necessarily other traits thought to be important for domestication, such as loss of secondary metabolites or changes to fruit and seed morphology (see Meyer et al., 2012b for a discussion and critique of domestication syndrome traits). Weedy species are also often resistant to many pests and diseases; biotic and abiotic resistance may help to account for their broad ranges. The species of CWR that have to date provided useful pest and disease resistance have all been weedy (Table 1), but most of these are only partially interfertile with *S. melongena* (Daunay and Hazra, 2012) or require more advanced techniques (e.g., somatic hybrids and tissue culture) to interbreed (e.g., Daunay et al., 1993; Kumchai et al., 2013) or, alternatively, to be used as rootstocks (Rahman et al., 2002; Prohens et al., 2013). Richardson and Rejmánek (2011) included the eggplant relatives *S. torvum* and *S. sisymbriifolium* in a recent global analysis of invasive species, and it could be argued that invasives are the ultimate weeds. We included three distantly related, weedy (possibly invasive) species in our GP3 for eggplant and suggest that in including such taxa in a gap analysis for germplasm collecting both native and nonnative ranges should be part of the analysis. Populations of weedy (invasive) species outside their native range may have escaped from local pests and diseases and be exhibiting ecological release, or conversely they could be very resistant to whatever biotic challenges they confront. Distinguishing between these two possibilities is important in the context of collecting CWR for conservation in genebanks, especially if weedy CWR are resistant to a great variety of challenges. In this context *S. viarum* Dunal and *S. torvum* (both included in our GP3) may be particularly interesting because their nonnative ranges include regions in southeastern Asia where

eggplant is thought to have been domesticated (see Fig. 1) and presumably also where pests and disease that have coevolved with *S. melongena* occur.

Our study using herbarium specimen data identified and georeferenced to a high standard and compared to data taken from genebank collections has shown that eggplant wild relatives are extremely poorly represented, both taxonomically and geographically, in genebanks worldwide. Our high resolution species distribution models can be of use in targeting areas for further collecting to improve the conservation status of these species ex situ. Combining the ex situ gap analysis with an assessment of conservation threat in situ has revealed hotspots for conservation action; many of the species under the most threat exist only outside of current protected areas. Collection of these species will also be important for biodiversity conservation ex situ. Earlier unpublished versions of our analyses have already been useful to prioritize collecting efforts of eggplant CWR species in eastern Africa (Dempewolf et al., 2014). Working with traditional genepool concepts is challenging in large genera such as *Solanum*, but using a combined approach and incorporating recent taxonomic and phylogenetic advances we have identified both species and regions for further germplasm collecting to both increase understanding of eggplant wild relatives and strengthen eggplant breeding in the future.

ACKNOWLEDGEMENTS

The authors thank the managers curating Solanaceae in both herbaria and genebanks for contributing data for these analyses. Helpful comments by two anonymous reviewers greatly improved the manuscript. This work was undertaken as part of "Adapting Agriculture to Climate Change: Conserving, Protecting and Preparing Crop Wild Relatives" (http://www.cwrdiversity.org/) supported by the Government of Norway and the Natural History Museum's Natural Resources Initiative. Taxonomic and systematic work on *Solanum* and eggplant CWR was supported by National Science Foundation (NSF) grant DEB-0316614 'PBI *Solanum* – a world treatment' to SK. MMS was funded by the Crop and Pest Wild Relatives project as part of the NHM Natural Resources Initiative.

LITERATURE CITED

Ahuja, S., M. C. Mukhopadhyaya, A. Singh, and S. P. Ahuja. 1987. Effects of infestation of eggplant (*Solanum melongena*) with root knot nematode (*Meloidogyne incognita*) on the oxidative enzymes and cell wall constituents in their roots. *Capsicum and Eggplant Newsletter* 6: 98–99.

Aubriot, X., P. Singh, and S. Knapp. In press. Tropical Asian species show the Old World clade of "spiny solanums" (subgenus *Leptostemonum* Bitter, the Leptostemonum Clade: Solanaceae) is not monophyletic. *Botanical Journal of the Linnean Society* 181(2).

Bivand, R., T. Keitt, and B. Rowlingson. 2014. rgdal: Bindings for the geospatial data abstraction library [online]. Website http://cran.r-project.org/package=rgdal [Accessed 12 December 2014].

Bivand, R., and N. Lewin-Koh. 2014. maptools: Tools for reading and handling spatial objects [online]. Website http://cran.r-project.org/package=maptools [Accessed 12 December 2014].

Bivand, R. S., E. Pebesma, and V. Gomez-Rubio. 2013. Applied spatial data analysis with R, 2nd edition. Springer Verlag, New York, New York, USA.

Bletsos, F., D. Roupakias, M. L. Tsaktsira, A. B. Scaltsoyjannes, and C. C. Thanassoulopoulos. 1998. Interspecific hybrids between three eggplant (*Solanum melongena* L.) cultivars and two wild species (*Solanum torvum* Sw. and *Solanum sisymbriifolium* Lam.). *Plant Breeding* 117: 159–164.

Bohs, L. 2005. Major clades in *Solanum* based on ndhF sequence data. *In* R. C. Keating, V. C. Hollowell, and T. B. Croat [eds.], A festschrift for William G. D'Arcy: The legacy of a taxonomist, Monographs in Systematic Botany, vol. 104, 27-49. Missouri Botanical Garden Press, St. Louis, Missouri, USA.

Brummitt, N., S. P. Bachman, E. Aletrari, H. Chadburn, J. Griffiths-Lee, M. Lutz, J. Moat, et al. 2015. The sampled Red List Index for plants, phase II : Ground-truthing specimen-based conservation assessments. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 370: 20140015.

Castañeda-Álvarez, N. P., S. de Haan, H. Juárez, C. K. Khoury, H. A. Achicanoy, C. C. Sosa, V. Bernau, et al. 2015. Ex situ conservation priorities for the wild relatives of potato (*Solanum* L. Section *Petota*). *PLoS One* 10: e0122599.

Castañeda-Álvarez, N. P., C. K. Khoury, H. A. Achicanoy, V. Bernau, H. Dempewolf, R. J. Eastwood, L. Guarino, et al. In press. Global conservation priorities for crop wild relatives. *Nature Plants* doi:10.1038/nplants.2016

CBD. 2010a. Convention on biological diversity, decision X/2. The strategic plan for biodiversity 2011–2020 and the Aichi biodiversity targets [online]. Website http://www.biodiv.be/convention/strategic-plan-2011-2020/ [Accessed 30 November 2014].

CBD. 2010b. Convention on biological diversity. Global strategy for plant conservation [online]. Website http://www.cbd.int/undb/media/factsheets/undb-factsheet-gspc-en.pdf [Accessed 5 January 2015].

Chelliah, S., and K. Srinivasan. 1985. Resistance in bhendi, brinjal and tomato to major insect and mites pests. *In* Proceedings of National Seminar in Breeding Crop Plants for Resistance to Pests and Diseases (May 25-27, 1983), 43-44. School of Genetics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India.

Collonnier, C., I. Fock, M.-C. Daunay, A. Servaes, F. Vedel, S. Siljak-Yakolev, V. Souvannavong, and D. Sihachakr. 2003a. Somatic hybrids between *Solanum melongena* and *S. sisymbrifolium*, as a useful source of resistance against bacterial and fungal wilts. *Plant Science* 164: 849–861.

Collonnier, C., I. Fock, I. Mariska, A. Servaes, F. Vedel, S. Siljak-Yakolev, V. Souvannavong, and D. Sihachakr. 2003b. GISH confirmation of somatic hybrids between *Solanum melongena* and *S. torvum*: Assessment of resistance to both fungal and bacterial wilts. *Plant Physiology and Biochemistry* 41: 459–470.

Daunay, M.-C., M. H. Chaput, D. Sihachakr, M. Allot, F. Vedel, and G. Ducreux. 1993. Production and characterization of fertile somatic hybrids of eggplant (*Solanum melongena* L.) with *Solanum aethiopicum* L. *Theoretical and Applied Genetics* 85: 841–850.

Daunay, M.-C., and P. Hazra. 2012. Eggplant. *In* K. Peter, and P. Hazra [eds.], Handbook of vegetables, 257–322. Studium Press, Houston, Texas, USA.

Daunay, M.-C., and J. Janick. 2007. History and iconography of eggplant. *Chronica Horticulturae* 47: 16–22.

Davidar, P., A. A. Snow, M. Rajkumar, R. Pasquet, M.-C. Daunay, and E. Mutegi. 2015. The potential for crop to wild hybridization in eggplant (*Solanum melongena*; Solanaceae) in Southern India. *American Journal of Botany* 102: 129–139.

Dempewolf, H., R. J. Eastwood, L. Guarino, C. K. Khoury, J. V. Müller, and J. Toll. 2014. Adapting agriculture to climate change: A global initiative to collect, conserve, and use crop wild relatives. *Agroecology and Sustainable Food Systems* 38: 369–377.

Díez, M., and F. Nuez. 2008. Tomato. *In* J. Prohens, and F. Nuez [eds.], Vegetables II SE - 7, Handbook of plant breeding, 249–323. Springer Verlag, New York, New York, USA.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, et al. 2006. Novel methods improve prediction of species distributions from occurrence data. *Ecography* 29: 129–151.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697.

ESRI. 2013. ArcGIS for desktop. Computer program and documentation available from ESRI, website http://www.esri.com/software/arcgis/arcgis-for-desktop [Accessed 2 January 2013].

European Space Agency. 2009. GlobCover 2009 (Global land cover map) [online]. Website http://due.esrin.esa.int/page_globcover.php [Accessed 23 February 2015].

FAO. 2009. International treaty on plant genetic resources for food and agriculture. Food and Agricultural Organization, Rome, Italy.

FAO. 2010. The second report on the state of the world's plant genetic resources for food and agriculture. Food and Agricultural Organization, Rome, Italy.

FAOSTAT. 2015. FAOSTAT [online]. Website http://www.faostat.fao.org [Accessed 30 November 2015].

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49.

Ford-Lloyd, B. V., M. Schmidt, S. J. Armstrong, O. Barazani, J. Engels, R. Hadas, K. Hammer, et al. 2011. Crop wild relatives—undervalued, underutilized and under threat? *Bioscience* 61: 559–565.

Fowler, C., G. Moore, and G. Hawtin. 2003. The international treaty on plant genetic resources for food and agriculture: A primer for the future harvest centres of CGIAR. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.

Franklin, J. 2009. Mapping species distributions: Spatial inference and prediction. Cambridge University Press, Cambridge, UK.

Frary, A., S. Doganlar, and M.-C. Daunay. 2007. Eggplant. *In* C. Kole [ed.], Vegetables SE - 9, Genome mapping and molecular breeding in plants, 287–313. Springer Verlag, Berlin, Germany.

Gaston, K. J., and R. A. Fuller. 2009. The sizes of species geographic ranges. *Journal of Applied Ecology* 46: 1–9.

Gisbert, C., J. Prohens, M. D. Raigón, J. R. Stommel, and F. Nuez. 2011. Eggplant relatives as sources of variation for developing new rootstocks: Effects of grafting on eggplant yield and fruit apparent quality and composition. *Scientia Horticulturae* 128: 14–22.

Gramazio, P., J. Prohens, M. Plazas, I. Andújar, F. J. Herraiz, E. Castillo, S. Knapp, et al. 2014. Location of chlorogenic acid biosynthesis pathway and polyphenol oxidase genes in a new interspecific anchored linkage map of eggplant. *BMC Plant Biology* 14: 350.

Guarino, L., and D. B. Lobell. 2011. A walk on the wild side. *Nature Climate Change* 1: 374–375.

Harlan, J., and J. de Wet. 1971. Toward a rational classification of cultivated plants. *Taxon* 20: 509–517.

Hijmans, R. J. 2014. raster: Geographic data analysis and modeling [online]. Website http://cran.r-project.org/package=raster [Accessed 12 December 2014].

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.

Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2014. dismo: Species distribution modeling [online]. Website http://cran.r-project.org/package=dismo [Accessed 12 December 2014].

Hijmans, R. J., and D. M. Spooner. 2001. Geographic distribution of wild potato species. *American Journal of Botany* 88: 2101–2112.

Hodgkin, T., and R. Hajjar. 2008. Using crop wild relatives for crop improvement: Trends and perspectives. *In* N. Maxted, B. V. Ford-Lloyd, S. P. Kell, J. Iriondo, E. Dulloo, and J. Turok [eds.], Crop wild relatives: Conservation and use, 535–548. CABI, Wallingford, UK.

Iriondo, J. M., N. Maxted, S. P. Kell, B. V. Ford-Lloyd, C. Lara-Romero, J. Labokas, and J. Magos Brehm. 2012. Quality standards for genetic reserve conservation of crop wild relatives. *In* N. Maxted, M. E. Dulloo, B. V. Ford-Lloyd, L. Frese, J. Iriondo, and M. A. A. Pinheiro de Carvalho [eds.], Agrobiodiversity conservation: Securing the diversity of crop wild relatives and landraces, 72–77. CABI, Wallingford, UK.

IUCN. 2012. IUCN Red List Categories: Version 3.1, 2nd edition. International union for the conservation of nature, Gland, Switzerland.

Jarl, C. I., E. M. Rietveld, and J. M. de Haas. 1999. Transfer of fungal tolerance through interspecific somatic hybridisation between *Solanum melongena* and *S. torvum*. *Plant Cell Reports* 18: 791–796.

Jarvis, A., A. Lane, and R. J. Hijmans. 2008. The effect of climate change on crop wild relatives. *Agriculture, Ecosystems & Environment* 126: 13–23.

Kardos, M., A. Husby, S. E. McFarlane, A. Qvarnström, and H. Ellegren. 2015. Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. *Molecular Ecology Resources*, published online in advance of print, 24 December 2015.

Khan, M. M. R., and S. Isshiki. 2009. Functional male-sterility expressed in eggplant (*Solanum melongena* L.) containing the cytoplasm of *S. kurzii* Brace & Prain. *Journal of Horticultural Science & Biotechnology* 84: 92–96.

Khan, M. M. R., and S. Isshiki. 2011. Development of a cytoplasmic male-sterile line of eggplant (*Solanum melongena* L.) with the cytoplasm of *Solanum anguivi*. *Plant Breeding* 130: 256–260.

Knapp, S., and M. S. Vorontsova. 2013. From introduced American weed to Cape Verde Islands endemic: The case of *Solanum rigidum* Lam. (Solanaceae, *Solanum* subgenus *Leptostemonum*). *PhytoKeys* 25: 35–46.

Knapp, S., M. S. Vorontsova, and J. Prohens. 2013. Wild relatives of the eggplant (*Solanum melongena* L.: Solanaceae): New understanding of species names in a complex group. *PLoS One* 8: e57039.

Kumchai, J., Y.-C. Wei, C.-Y. Lee, F.-C. Chen, and S. W. Chin. 2013. Production of interspecific hybrids between commercial cultivars of eggplant (*Solanum melongena* L.) and its wild relative *S. torvum*. *Genetics and Molecular Research* 12: 755–764.

Lester, R., and S. Hasan. 1991. Origin and domestication of the brinjal eggplant, *Solanum melongena*, from *Solanum incanum*, in Africa and Asia. *In* J. G. Hawkes, R. N. Lester, M. Nee, and N. Estrada [eds.], Solanaceae III: Taxonomy, chemistry and evolution, 369–387. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.

Liu, C., P. M. Berry, T. P. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28: 385–393.

Liu, C., M. White, and G. Newell. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography* 40: 778–789.

Liu, J., Z. Zheng, X. Zhou, C. Feng, and Y. Zhuang. 2015. Improving the resistance of eggplant (*Solanum melongena*) to Verticillium wilt using wild species *Solanum linnaeanum*. *Euphytica* 201: 463–469.

Mace, E. S., R. N. Lester, and C. G. Gebhardt. 1999. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). *Theoretical and Applied Genetics* 99: 626–633.

Mace, G. M., N. J. Collar, K. J. Gaston, C. Hilton-Taylor, H. R. Akçakaya, N. Leader-Williams, and S. N. Stuart. 2008. Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology* 22: 1424–1442.

Maxted, N., E. Dulloo, B. V. Ford-Lloyd, J. M. Iriondo, and A. Jarvis. 2008. Gap analysis: A tool for complementary genetic conservation assessment. *Diversity & Distributions* 14: 1018–1030.

Maxted, N., B. V. Ford-Lloyd, S. Jury, S. Kell, and M. Scholten. 2006. Towards a definition of a crop wild relative. *Biodiversity and Conservation* 15: 2673–2685.

Maxted, N., and S. Kell. 2009. Establishment of a global network for the in situ conservation of crop wild relatives: Status and needs. *FAO Background Study Paper* 39: 1–112. Food and Agriculture Organization, Rome, Italy.

Meilleur, B. A., and T. Hodgkin. 2004. In situ conservation of crop wild relatives: Status and trends. *Biodiversity and Conservation* 13: 663–684.

Mennella, G., G. L. Rotino, M. Fibiani, A. D'Alessandro, G. Francese, L. Toppino, F. Cavallanti, et al. 2010. Characterization of health-related compounds in eggplant (*Solanum melongena* L.) lines derived from introgression of allied species. *Journal of Agricultural and Food Chemistry* 58: 7597–7603.

Merow, C., M. J. Smith, and J. A. Silander. 2013. A practical guide to MaxEnt for modeling species distributions: What it does, and why inputs and settings matter. *Ecography* 36: 1058–1069.

Meyer, R. S., A. E. DuVal, and H. R. Jensen. 2012b. Patterns and processes in crop domestication: An historical review and quantitative analysis of 203 global food crops. *New Phytologist* 196: 29–48.

Meyer, R. S., K. G. Karol, D. P. Little, M. H. Nee, and A. Litt. 2012a. Phylogeographic relationships among Asian eggplants and new perspectives on eggplant domestication. *Molecular Phylogenetics and Evolution* 63: 685–701.

Mutegi, E., A. A. Snow, M. Rajkumar, R. Pasquet, H. Ponniah, M.-C. Daunay, and P. Davidar. 2015. Genetic diversity and population structure of wild/weedy eggplant (*Solanum insanum*, Solanaceae) in southern India: Implications for conservation. *American Journal of Botany* 102: 140–148.

Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. N. Powell, E. C. Underwood, J. A. D'Amico, et al. 2001. Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* 51: 933–938.

Parra-Quijano, M., J. M. Iriondo, and E. Torres. 2011. Improving representativeness of genebank collections through species distribution models, gap analysis and ecogeographical maps. *Biodiversity and Conservation* 21: 79–96.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34: 102–117.

Pebesma, E. J., and R. S. Bivand. 2005. Classes and methods for spatial data in R. R News 5(2) [online]. Website http://cran.r-project.org/doc/Rnews/ [Accessed 12 August 2014].

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araújo. 2011. Ecological niches and geographic distributions. Princeton University Press, Princeton, New Jersey, USA.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259.

Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrer. 2009. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications* 19: 181–197.

Plazas, M., M. P. López-Gresa, S. Vilanova, C. Torres, M. Hurtado, P. Gramazio, I. Andújar, et al. 2013. Diversity and relationships in key traits for functional and apparent quality in a collection of eggplant: Fruit phenolic content, antioxidant activity, polyphenol oxidase activity, and browning. *Journal of Agricultural and Food Chemistry* 61: 8871–8879.

Plazas, M., P. Vilanova, S. Gramazio, A. Rodríguez-Burruezo, A. Fita, F. J. Herraiz, R. Ranil, et al. 2016. Interspecific hybridization between eggplant and wild relatives from different genepools. *Journal of the American Society for Horticultural Science* 141: 34–44.

Pringle, R. M., J. R. Goheen, T. M. Palmer, G. K. Charles, E. DeFranco, R. Hohbein, A. T. Ford, and C. E. Tarnita. 2014. Low functional redundancy among mammalian browsers in regulating an encroaching shrub (*Solanum campylacanthum*) in African savannah. *Proceedings of the Royal Society of London, B, Biological Sciences* 281: 20140390.

Prohens, J., G. J. Anderson, F. J. Herraiz, G. Bernardello, A. Santos-Guerra, D. Crawford, and F. Nuez. 2007. Genetic diversity and conservation of two endangered eggplant relatives (*Solanum vespertilio* Aiton and *Solanum lidii* Sunding) endemic to the Canary Islands. *Genetic Resources and Crop Evolution* 54: 451–464.

Prohens, J., B. D. Whitaker, M. Plazas, S. Vilanova, M. Hurtado, M. Blasco, P. Gramazio, and J. R. Stommel. 2013. Genetic diversity in morphological characters and phenolic acids content resulting from an interspecific cross between eggplant, *Solanum melongena*, and its wild ancestor (*S. incanum*). *Annals of Applied Biology* 162: 242–257.

Rahman, M. A., M. A. Rashid, M. M. Hossain, M. A. Salam, and A. Masum. 2002. Grafting compatibility of cultivated eggplant varieties with wild *Solanum* species. *Pakistan Journal of Biological Sciences* 5: 755–757.

Raigón, M. D., J. Prohens, J. E. Muñoz-Falcón, and F. Nuez. 2008. Comparison of eggplant landraces and commercial varieties for fruit content of phenolics, minerals, dry matter and protein. *Journal of Food Composition and Analysis* 21: 370–376.

Ramírez-Villegas, J., C. Khoury, A. Jarvis, D. G. Debouck, and L. Guarino. 2010. A gap analysis methodology for collecting crop genepools: A case study with *Phaseolus* beans. *PLoS One* 5: e13497.

Rao, G. R., and A. Kumar. 1980. Some observations on interspecific hybrids of Solanum melongena L. *Proceedings of the Indian Academy of Sciences (Plant Sciences)* 89: 117–121.

Richardson, D. M., and M. Rejmánek. 2011. Trees and shrubs as invasive alien species—a global review. *Diversity & Distributions* 17: 788–809.

Rodrigues, A. S. L., J. D. Pilgrim, J. F. Lamoreux, M. Hoffman, and T. M. Brooks. 2006. The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution* 21: 71–76.

Rodríguez Delgado, O., A. García Gallo, and G. Cruz Trujillo. 2011. *Solanum lidii. In:* IUCN 2013. IUCN Red List of threatened species, version 2013.2. Website http://www.iucnredlist.org [Accessed 27 January 2014].

Rotino, G. L., T. Sala, and L. Toppino. 2014. Eggplant. *In* A. Pratap, and J. Kumar [eds.], Alien gene transfer in crop plants, vol. 2, SE - 16, 381–409. Springer Verlag, New York, New York, USA.

Särkinen, T., L. Bohs, R. G. Olmstead, and S. Knapp. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): A dated 1000-tip tree. *BMC Evolutionary Biology* 13: 214.

Särkinen, T., J. R. Iganci, R. Linares-Palomino, M. F. Simon, and D. E. Prado. 2011. Forgotten forests—issues and prospects in biome mapping using Seasonally Dry Tropical Forests as a case study. *BMC Ecology* 11: 27.

Schalk, J. M., A. K. Stoner, R. E. Webb, and H. F. Winters. 1975. Resistance in eggplant (*Solanum melongena* L.) and non tuber-bearing *Solanum* species to carmine spider mite. *Journal of the American Society for Horticultural Science* 100: 479–481.

Singh, M., and R. Kumar. 2007. Eggplant (*Solanum melongena* L.). *In* R. J. Singh [ed.], Genetic resources, chromosome engineering, and crop improvement. Vegetable crops, vol. 3, 473–495. CRC Press, Boca Raton, Florida, USA.

Slavin, J. L., and B. Lloyd. 2012. Health benefits of fruits and vegetables. *Advances in Nutrition* 3: 506–516.

Stephenson, N. L. 1998. Actual evapotranspiration and deficit: Biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography* 25: 855–870.

Stern, S., M. D. F. Agra, and L. Bohs. 2011. Molecular delimitation of clades within New World species of the "spiny solanums" (*Solanum* subg. *Leptostemonum*). *Taxon* 60: 1429–1441.

Sunseri, F., A. Sciancalepore, G. Martelli, N. Acciarri, G. L. Rotino, D. Valentina, and G. Tamietti. 2003. Development of a RAPD-AFLP map of eggplant and improvement of tolerance to *Verticillium* wilt. *Acta Horticulturae* 625: 197–198.

Syfert, M. M., M. J. Smith, and D. A. Coomes. 2013. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS One* 8: e55158.

Tanksley, S. D., and S. R. McCouch. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* 277: 1063–1066.

VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo, and C. Storlie. 2014. SDMTools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises [online]. Website http://cran.r-project.org/package=SDMTools [Accessed 12 December 2014].

VanDerWal, J., L. P. Shoo, C. Graham, and S. E. Williams. 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling* 220: 589–594.

Vavilov, N. I. 1951. The origin, variation, immunity and breeding of cultivated plants (translated from Russian by K. Starr Chester). *Chronica Botanica* 13: 1–366.

Vavilov, N. I. 1992. Origin and geography of cultivated plants (a series of papers translated from Russian by D. Löve, first published in Russian in 1987). Cambridge University Press, London, UK.

Vimal, R., A. S. L. Rodrigues, R. Mathevet, and J. D. Thompson. 2010. The sensitivity of gap analysis to conservation targets. *Biodiversity and Conservation* 20: 531–543.

Vorontsova, M. S. 2010. Overlooked diversity in African *Solanum* (Solanaceae). *Phytotaxa* 37: 31–37.

Vorontsova, M. S., and S. Knapp. In press. A revision of the spiny solanums, *Solanum* subgenus *Leptostemonum* (Solanaceae) in Africa and Madagascar. *Systematic Botany Monographs 100*.

Vorontsova, M. S., and F. M. Mbago. 2010. Species from Tanzanian coastal forests may already be extinct. *Journal of East African Natural History* 99: 227–234.

Vorontsova, M. S., S. Stern, L. Bohs, and S. Knapp. 2013. African spiny *Solanum* (subgenus *Leptostemonum*, Solanaceae): A thorny phylogenetic tangle. *Botanical Journal of the Linnean Society. Linnean Society of London* 173: 176–193.

Wang, J.-X., T.-G. Gao, and S. Knapp. 2008. Ancient Chinese literature reveals pathways of eggplant domestication. *Annals of Botany* 102: 891–897.

WDPA. 2010. World database on protected areas. World Database on Protected Areas Consortium, United Nations Environment Program-World Conservation Monitoring Center (UNEP-WCMC), Cambridge, UK.

Weese, T. L., and L. Bohs. 2007. A three-gene phylogeny of the genus *Solanum* (Solanaceae). *Systematic Botany* 32: 445–463.

Weese, T. L., and L. Bohs. 2010. Eggplant origins: Out of Africa, into the Orient. *Taxon* 59: 49–56.

White, F. 1983. The vegetation of Africa, a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. UNESCO, Paris, France.

Wiersema, J. H., B. León, and E. J. Garvey. 2012. Identifying wild relatives of subtropical and temperate fruit and nut crops. *In* M. K. Aradhya, and D. A. Kluepfel [eds.], I. International symposium on wild relatives of subtropical and temperate fruit and nut crops. *Acta Horticulturae* 948: 285–288.

Wisz, M. S., R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, and A. Guisan, and NCEAS predicting species distributions working group. 2008. Effects of sample size on the performance of species distribution models. *Diversity & Distributions* 14: 763–773.