

Could taxonomic misnaming threaten the ex situ conservation and the usage of plant genetic resources?

Filippo Guzzon¹  · Nicola M. G. Ardenghi¹ 

Received: 4 September 2017 / Revised: 24 November 2017 / Accepted: 28 November 2017 /
Published online: 21 December 2017
© Springer Science+Business Media B.V., part of Springer Nature 2017

Abstract Ex situ conservation of plant germplasm, especially seed banking, is a favourable and widely used method for the conservation of plant genetic resources (PGR). The long-term conservation of these resources is fundamental for food security and plant breeding in order to stem the losses in agrobiodiversity and to meet the global challenges that agriculture is facing. The conservation and accessibility of PGR relies on their correct taxonomic labelling and on the building of a searchable database that links ex situ collections together. In the current study, we analysed the impact of taxonomic misnaming in the two major PGR databases (Genesys PGR, EURISCO), listing accessions conserved worldwide. The aim was to understand if taxonomic misnaming issues prevent PGR conservation. We chose as a case-study seed collections of accessions of the genus *Citrullus* (watermelon genepool), the taxonomy and nomenclature of which have been largely revised in recent times. We observed that taxonomic misnaming greatly limits PGR conservation with only 3% of the accessions listed in the databases correctly named; moreover, 28% were affected by taxonomic errors that prevent the establishment of the accessions' taxonomic identity, with consequences on their conservation and exploitation. The existence of the problem was also confirmed by the experimental propagation of three misnamed accessions. We suggest herein a series of actions that, put in place, could solve the extant misnaming issues in the databases and prevent their reoccurrence, allowing the correct conservation and the usability by the stakeholders of all the accessions.

Communicated by Daniel Sanchez Mata.

This article belongs to the Topical Collection: Ex-situ conservation.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10531-017-1485-7>) contains supplementary material, which is available to authorized users.

✉ Filippo Guzzon
filippo.guzzon01@universitadipavia.it

¹ Department of Earth and Environmental Sciences, University of Pavia, I-27100 Pavia, Italy

Keywords Botanical nomenclature · Database · Genebank · Seed conservation · Taxonomy · Watermelon

Introduction

In the next decades, a growing human population, changing diets, extreme climate conditions, lower availability of natural resources, higher competition for arable soils with non-food crops, soil degradation, and the need to minimize harmful impacts on ecosystems, will pose new challenges to global food security (Godfray et al. 2010; Asseng et al. 2015). To cope with the aforementioned challenges, both agricultural production and sustainability need to increase (Tilman et al. 2002). The selection of new crop cultivars with favourable traits such as increased drought and heat tolerance, and input use efficiency, will be key to this process (Esquinas-Alcazar 2005; McCouch et al. 2013). To meet this goal, the eroded genepool of modern crop plants needs to be broadened and the widest genetic diversity needs to be available and exploitable, in order to select for the improved cultivars of the future (Ford-Llyod et al. 2011; Guarino and Lobell 2011; Vincent et al. 2013; Warschefsky et al. 2014). This is achievable only by conserving and keeping available, for research and breeding programmes, the widest possible genepool for each crop, especially landraces and crop wild relatives (CWR) (McCouch et al. 2013). In many areas of the world, in the last few decades, genetic erosion (the loss of genetic diversity in the form of alleles and genotypes as well as domestic crop accessions) of higher than the 70% has been observed (Hammer et al. 1996; Veteläinen et al. 2009; van de Wouw et al. 2010). Given this scenario, it is of key-importance to conserve agrobiodiversity in the long-term. More specifically, this is the diversity of crop species used in different agro-ecosystems as well as the genetic diversity within and among crop and CWR accessions (Last et al. 2014). Furthermore, the conservation of plant genetic resources for food and agriculture (PGRFA) is fundamental to achieving target 9 of the 2011–2020 Global Strategy for Plant Conservation as well as target 13 of the Aichi biodiversity targets. Thus, the effective conservation of agrobiodiversity and its sustainable use is considered to be of pivotal importance by the Convention of Biological Diversity (CBD) (Convention of Biological Diversity 2011, 2012).

One of the most effective strategies to ensure the conservation and availability of PGRFA is through *ex situ* conservation in genebanks (McCouch et al. 2013; Davies and Allender 2017). In particular, seed banking shows several advantages as a long-term *ex situ* conservation strategy for plant species and therefore is used for the maintenance of most of the PGR (plant genetic resources) collections *ex situ*. Seed material is relatively easy to collect, can be stored in small spaces, can provide a decent sample of the genetic diversity within the species and often remains viable for long periods. Furthermore, seed collections do not require a high level of maintenance and are also economically viable (Li and Pritchard 2009; Riviere and Muller 2017). Worldwide there are more than 1750 genebanks containing over 7.4 million accessions (FAO 2010; Davies and Allender 2017).

All countries are highly dependent upon PGRFA conserved beyond their borders. This global interdependence, and therefore global flows of PGRFA among conservation facilities, are likely to increase in order to cope with future challenges, especially climatic changes (Galluzzi et al. 2016).

The stakeholders of PGRFA indicated as a major constraint in the usage of conserved germplasm the difficulty in accessing it and to obtaining associated information (Kell et al.

2017). The accessibility of PGR accessions is strictly linked with the existence and updating of global information databases, this involves the gathering of data for accessions from many collections into a centralized source, therefore facilitating the transfer of the germplasm among institutions. Consequently, the building and improvement of information systems that link collections together in order to create a single searchable database of genetic resources is a high priority in PGR conservation and utilization (Houry et al. 2010).

The correct and clear taxonomic identification and labelling of accessions is of key importance to making accessions of PGR, conserved ex situ, usable (Dempewolf et al. 2017). Taxonomic issues are indeed demonstrated to highly threaten the effectiveness of conservation programmes (Garnett and Christidis 2017). Moreover, the updated and precise taxonomic labelling of PGR accessions in public databases is also fundamental to performing prioritization studies aimed at finding gaps in ex situ conservation (Dempewolf et al. 2014). This is because a correct taxonomic naming of accessions is essential for the identification of PGR taxa that are currently underrepresented in ex situ conservation facilities worldwide and therefore have a high priority for future collecting missions and urgent conservation measures (Maxted et al. 2010; Castañeda-Álvarez et al. 2016).

In order to better understand the extent of taxonomic misnaming in databases of PGR accessions which are conserved ex situ and whether this issue could prevent their exploitation and conservation, we analysed and quantified the occurrence of this problem in seed accessions belonging to the watermelon genepool (*Citrullus*). We decided to focus on this genus as a case study, since it has a relatively small number of taxa (species and subspecies) and, after having been considered for long a critical taxonomic group, its systematics, taxonomy, and nomenclature have been revised and improved in recent years, with the aid of genetic investigation too (see Nesom 2011; Schaefer and Renner 2011; Chomicki and Renner 2014; Renner et al. 2014; Paris 2015). Moreover, *Citrullus* is of significant importance as a vegetable crop (Applequist 2016), with 3.5 million hectares of agricultural land used to cultivate watermelons in 2014 when annual production reached 111 million tons, which was 9.5% of global vegetable production, grown on 6% of the area used globally for the cultivation of vegetables (FAOSTAT 2017). The genus is of great importance in terms of food security in desert and semi-arid areas (Mujaju et al. 2011; Modi and Zulu 2012). *Citrullus* includes eight taxa (seven species, one of which articulated into two subspecies), three are widely cultivated [*C. amarus*, *C. lanatus* subsp. *lanatus*, *C. mucosospermus*], two are only locally cultivated [*C. colocynthis* and *C. lanatus* subsp. *kordophanus*], and three have only wild populations [*C. ecirrhosus*, *C. naudinianus*, *C. rehmi*] (Chomicki and Renner 2014; Paris 2015).

The aims of the current research, adopting *Citrullus* as a case study, are: (a) to understand and define which are the most frequently nomenclature-related issues in databases of PGR accessions, (b) to clarify if these issues could prevent the usage of these accessions, and (c) to identify if and how the discovered issues may be resolved.

Materials and methods

We checked for *Citrullus* accessions in the two major databases of plant genetic resources worldwide: Genesys PGR and EURISCO. Genesys is a global portal that lists more than 3.6 million accessions of plant genetic resources from 482 institutes worldwide. It is a gateway from which germplasm accessions from genebanks can be found and ordered (Genesys 2017). EURISCO is a search catalogue providing information about ex situ plant

collections maintained by institutions located in Europe, and includes data for 1.9 million accessions (EURISCO 2017). Individual databases contributing to EURISCO simultaneously upload data into Genesys (Genesys 2017), but, not knowing the details of updating between the two databases, we decided to consider both separately in our analysis.

We searched as keywords, the existing nomenclatural combinations acquired from literature (e.g., Pangalo 1930; Mansfeld 1959; Fursa 1972; Nesom 2011; Chomicki and Renner 2014; Paris 2015). In order to increase the possibilities of finding accessions, we also searched for the most widespread vernacular names (e.g., colocynth, tsamma, citron, egusi), with reference to geographic provenance and biological status (wild, landrace, modern cultivar) too. The resulting accession names were examined on nomenclatural grounds, by comparing them with the most updated taxonomic and nomenclatural treatments of *Citrullus* (see Nesom 2011; Chomicki and Renner 2014; Renner et al. 2014; Paris 2015); names were further checked in compliance with the rules of the International Code of Nomenclature for algae, fungi, and plants (ICN) (McNeill et al. 2012). A complete list of the accepted names and synonyms we came across in the consulted databases are provided in Online Resource 1.

The encountered misnaming issues were classified into “issue types” and then grouped into two main “issue categories” (Table 1). If an accession was indeed wrongly named, we attributed the misnaming to one or more issue types and then to one or more issue categories. Eventually, when practicable, we proposed, for each accession, the most updated scientific name (Table 1), based on the updated taxonomic and nomenclatural treatment of *Citrullus* reported in Online Resource 1. We then provided a numerical estimation of each misnaming issue type and category found in the two databases, this was intended to evaluate the extent of taxonomic misnaming in PGR databases (Online Resource 2, 3).

Additionally, by analysing the different databases, we found that the name of the same taxon changes among institutions, we referred to this as “variability”. When more scientific names are applied to the same taxon within the same institution we referred to “intra-institution variability”, among different institutions was “inter-institution variability” (Table 2), within the same database “intra-database variability” or between different databases “inter-database variability” (Table 3). We chose *C. amarus* as a target taxon to illustrate this point. In each database, we searched for the current accepted name and its synonyms.

Eventually, in order to verify in practical terms the validity of our doubts, three accessions that appeared to be misnamed in the Genesys and EURISCO databases were propagated at the CREA-ORL institute of Montanaso Lombardo (northern Italy, province of Lodi). The accession names were first revised on nomenclatural grounds and then, after propagation, their taxonomic identity was checked (Table 4; Fig. 1). Cultivation was performed in purity to avoid crossbreeding among the different accessions: female flowers were isolated with paper bags and hand pollination was performed. Herbarium vouchers are stored at the Herbarium of the University of Pavia (PAV).

Results

Misnaming issues: categorization

We identified six types of misnaming issues (Table 1), along with several misprints regarding both the names of the taxa and their authors. Each issue has been classified into two “issue categories”:

Table 1 Misnaming issue types (“Issue types”), with definitions and examples (“Issue types definition”) acquired from the results of our database search

Accession scientific name	Issue types	Issue category	Current scientific name	Issue types definition
<i>Citrullus aedulis</i> Pangalo	Synonym	N	<i>C. lanatus</i> (Thunb.) Matsum. & Nakai subsp. <i>lanatus</i>	The accession name is a synonym of the currently accepted scientific name (see Online Resource 1)
<i>Citrullus lanatus</i> (Thunb.) Matsum. et Nakai subsp. <i>vulgaris</i> var. <i>caffèr</i> (Schrad.) Mansf.	Non-existent name	N	<i>C. amarus</i> Schrad.	The accession name, as written, does not exist. In this case, the combination published by Mansfeld (1959) is <i>C. lanatus</i> var. <i>caffèr</i> (Schrad.) Mansf. (see Online Resource 1), without indication of the subspecies
<i>Citrullus lanatus</i> Mansf.	Authors	N	<i>C. amarus</i> Schrad., <i>C. lanatus</i> (Thunb.) Matsum. & Nakai s.l. or <i>C. mucospermus</i> (Fursa) Fursa	The authors associated with the accession name are incorrect. In this case, the accession name refers to <i>C. lanatus</i> (Thunb.) Mansf., but the combination by Mansfeld (1959) is an isonym (see Art. 6.2, Note 2 of the ICN and Online Resource 1), having been published later than that by Matsumura and Nakai (1916), who are the correct authors of the combination; in addition, the accession name lacks of the basionym’s author, “(Thunb.)”. Since no subspecies is mentioned, this accession name affects the establishment of the taxonomic identity of the accession, thus it represents a “taxonomic error” too, which is properly explained below
<i>Citrullus</i> sp.	No species	T	?	The accession name is represented only by the name of the genus

Table 1 (continued)

Accession scientific name	Issue types	Issue category	Current scientific name	Issue types definition
<i>Citrullus lanatus</i> (Thumb.) Matsum. & Nakai	No subspecies	T	<i>C. lanatus</i> (Thumb.) Matsum. & Nakai s.l., <i>C. amarus</i> Schrad. or <i>C. mucosospermus</i> (Fursa) Fursa	The subspecies is not indicated. As a consequence, in this example, it is impossible to establish whether the accession belong to the nominal subspecies or subsp. <i>kordophanus</i> Ter-Avan. Moreover, since <i>C. amarus</i> Schrad. and <i>C. mucosospermus</i> (Fursa) Fursa have both been treated as subspecies or varieties of <i>C. lanatus</i> until recent times (see Online Resource 1), it cannot be excluded that this name may also refer to one of these two species
<i>Citrullus lanatus</i> (Thumb.) Matsum. & Nakai subsp. <i>lanatus</i> (citroides group)	Conflicting scientific and cultivar names	T	<i>C. amarus</i> Schrad.	The cultivar name is not applied to the appropriate taxon. In this case, taking the cultivar name as correctly indicated (“citroides group”), the correct accession name should be <i>C. amarus</i> Schrad. ‘Citroides’ group

We also listed the accession name (“Accession scientific name”); the issue categories (“Issue category”); and the updated scientific name we proposed (“Current scientific name”)

Author citations of scientific names are included here and also in Tables 2 and 3 as they are in the entries as submitted to the databases

Table 2 Intra- and inter-institution variability of taxon names

Institution	Scientific name	Database(s)	No. of accessions
RUS001	<i>Citrullus lanatus</i> var. <i>citroides</i> (L. H. Bailey) Mansf.	Genesys	41
RUS001	<i>Citrullus lanatus</i> var. <i>capensis</i> (Alef.) Fursa	Genesys	28
RUS001	<i>Citrullus amarus</i> Schrad.	Genesys	0
HUN003	<i>Citrullus lanatus</i> var. <i>caffer</i> (Schrad.) Mansf. ["subsp. <i>vulgaris</i> "]	Genesys EURISCO	25
HUN003	<i>Citrullus lanatus</i> var. <i>citroides</i> (L.H.Bailey) Mansf.	Genesys EURISCO	1
HUN003	<i>Citrullus amarus</i> Schrad.	Genesys EURISCO	0

The scientific names have been transcribed without misprints; although occurring in both Genesys and EURISCO databases, the indication of "subsp. *vulgaris*" within the combination *C. lanatus* var. *caffer* (Schrad.) Mansf. is wrong (see Table 1), thus we inserted it between square brackets. The current accepted taxon name is in bold

Table 3 Intra- and iter-database variability of taxon names

Scientific name	No. of accessions in:	
	Genesys	EURISCO
<i>Citrullus amarus</i> Schrad.	0	0
<i>Citrullus lanatus</i> var. <i>citroides</i> (L. H. Bailey) Mansf.	183	5
<i>Citrullus lanatus</i> var. <i>caffer</i> (Schrad.) Mansf. ["subsp. <i>vulgaris</i> "]	25	26
<i>Citrullus lanatus</i> (Thunb.) Matsum. & Nakai subsp. <i>lanatus</i> ("Citroides" group)	15	18
<i>Citrullus lanatus</i> var. <i>capensis</i> (Alef.) Fursa	28	0

The scientific names have been transcribed without misprints; although occurring in both Genesys and EURISCO databases, the indication of "subsp. *vulgaris*" within the combination *C. lanatus* var. *caffer* (Schrad.) Mansf. is wrong (see Table 1), thus we inserted it between square brackets. The current accepted taxon name is in bold

1. "Nomenclatural inaccuracy" (N): the scientific name associated with the accession does not correspond to the most updated scientific name, but it represents a synonym (homotypic or heterotypic; see Glossary and Art. 14.4 of the ICN) or a spelling variant of the latter; it does not prevent the establishment of the real taxonomic identity of the accession;
2. "Taxonomic error" (T): the scientific name associated with the accession prevents the establishment of the real taxonomic identity of the accession.

Misnaming issues: quantification

By querying the two PGR databases: 8494 single entries of *Citrullus* accessions were obtained, 6631 entries from Genesys and 1863 from EURISCO. The scientific names of 5864 accessions (69.03% of the total) showed nomenclature inaccuracies (N), while the

Table 4 Propagated *Citrullus* accessions

Institution	Accession number	Origin	Common english name(s)	Database	Accession scientific name	Biological status	Accession revised scientific name	Accession revised identity	Issue category
GBR004	83216	Botswana	–	Genesys, EURISCO	<i>Citrullus lanatus</i> (Thunb.) Matsum. & Nakai	Natural	<i>C. amarus/C. lanatus</i> s.l./ <i>C. mucospermus</i>	<i>C. amarus</i>	T
RUS001	4679	Republic of South Africa	–	Genesys	<i>Citrullus lanatus</i> (Thunb.) Matsum. et Nakai var. <i>citroides</i> (Bailey) Mansf.	Founder stock/ base population	<i>C. amarus</i>	<i>C. amarus</i>	N
USA016	PI490380	Mali	Egusi	Genesys	<i>Citrullus lanatus</i> (Thunb.) Matsum. & Nakai var. <i>lanatus</i>	Traditional cultivar/landrace	<i>C. lanatus</i> subsp. <i>lanatus</i>	<i>C. mucospermus</i>	T

We listed: code of the institution where they are stored (“Institution”); accession number (“Accession number”); country where the accessions were originally collected (“Origin”); common English name (“Common English name(s)”; searched database (“Database”); accession scientific name as appearing in the databases (“Accession scientific name”); biological status (“Biological status”); scientific name obtained by revising the original accession name on nomenclatural grounds only (“Accession revised scientific name”); identity of the accession reevaluated after the propagation, in bold (“Accession revised identity”); misnaming issue category obtained from the comparison between the “Accession scientific name” and the “Accession revised identity”

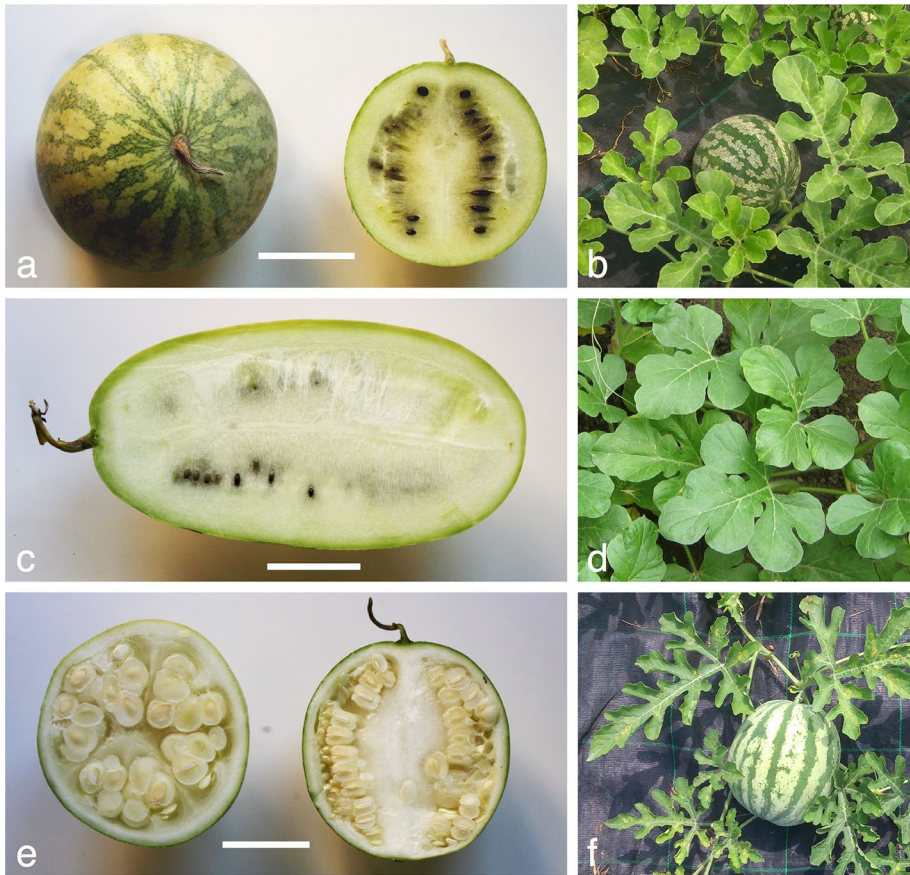


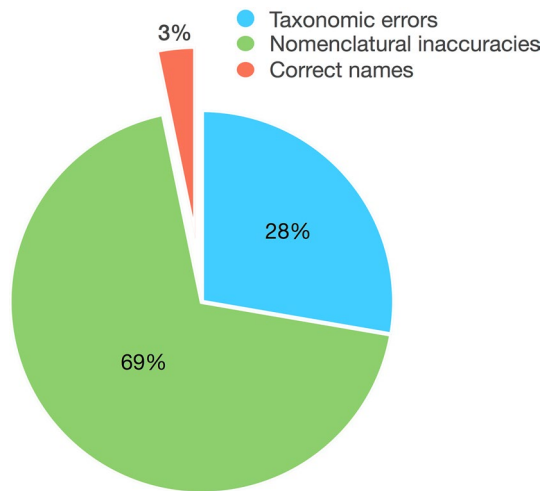
Fig. 1 Propagated *Citrullus* accessions (see Table 1): **a, c, e** fruits, scale bar = 5 cm; **b, d, f** foliage. **a** GBR004-83216: *C. amarus*, **b** RUS001-4679: *C. amarus*, **c** USA016-PI490380: *C. mucospermus*. Photos by T. Abeli (**a, c, e**), V. Ottobriano (**b, f**), and N. M. G. Ardenghi (**d**)

scientific names of 2355 accessions (27.72% of the total) showed taxonomic errors (T) (Online Resource 2, 3; Fig. 2).

In detail, 9.85% of the misnaming issues encountered in both the databases fell within the issue type “authors” (N); 66.79% fell within “synonym” (N); 1.40% fell within “non-existent name” (N); 0.94% fell within “no species” (T); 26.60% fell within “no subspecies” (T); and 0.38% fell within “conflicting scientific and cultivar names” (T) (the sum of the aforementioned percentages exceeds 100% because some of the accessions showed more than one issue type). Misprints regarding the taxon names or their authors affected 12.54% of the accessions (Online Resource 2, 3).

Only 275 accessions (3.23% of the total) were correctly and unambiguously named according the most recent and updated taxonomic and nomenclatural treatment of *Citrullus* (Online Resource 1, 2, 3). Considering each database individually the percentages are similar: 2.88% (190 accessions) in Genesys and 4.59% (85 accessions) in EURISCO (Online Resource 2, 3).

Fig. 2 Percentages of the nomenclature inaccuracies (green), taxonomic errors (light blue), and correct names (red) pooling together the data of *Citrullus* accessions extracted from both Genesys and EURISCO databases



Variability

Taking the name *Citrullus amarus* as an example, we discovered the following variability concerning the accessions' names:

1. “Intra- and inter-istitution variability”: in Table 2, it can be seen that within two institutions (RUS001 and HUN003) different accessions belonging to the same species are named using different nomenclatural combinations, which are synonyms of the current accepted name. Moreover, each of the two institutions uses a name not employed by the other institution (*C. lanatus* var. *capensis* for RUS001, *C. lanatus* var. *caffer* for HUN003). None of the accessions have been updated with the currently accepted scientific name, *C. amarus*.
2. “Intra- and inter-database variability”: Table 3 reveals that both databases employ more nomenclatural combinations (four in Genesys and three in EURISCO) to name accessions belonging to the same species. Moreover, one database (Genesys) uses a name (*C. lanatus* var. *capensis*) not adopted by the other (EURISCO). Similarly to what found in “intra- and inter-istitution variability”, there are no accessions bearing the name *C. amarus* in the two databases.

Cultivation

The cultivation in purity confirmed that the names of two out of the three misnamed propagated accessions did not correspond to their actual taxonomic identity (taxonomic error). While the identity of RUS001-4679, although stored with a misprinted heterotypic synonym [*Citrullus lanatus* (Thunb.) Matsum. et Nakai var. *citroides* (Bailey) Mansf.], thus a nomenclatural inaccuracy; see Online Resource 1 and Table 1], was confirmed (*C. amarus*; Table 4; Fig. 1), the identity of GBR004-83216 and USA016-PI490380 turned out to be incorrect or unclear. Specifically, the identity of GBR004-83216 appeared to be doubtful already at the stage of the nomenclatural revision, since the institution did not provide any

intraspecific rank for *C. lanatus*, preventing any safe choice among *C. amarus*, the subspecies of *C. lanatus*, and *C. mucosospermus* (see Table 1, third example). The cultivation allowed us to resolve this issue, revealing that the correct identity of the accession is *C. amarus* (Table 4; Fig. 1). In USA016-PI490380, the scientific name of the accession (*C. lanatus* var. *lanatus*, a homotypic synonym of *C. lanatus* spp. *lanatus*, see Online Resource 1) appears to be in conflict with the vernacular name (Egusi), which applies to another species (*C. mucosospermus*, see Online Resource 1). In this case, propagation also proved to be decisive in resolving the issue, being *C. mucosospermus* the plant having emerged from the seeds constituting the accession (Table 4; Fig. 1).

Discussion

The results of our database search and subsequent revision of the accessions' scientific names, highlight the fact that taxonomic misnaming issues actually greatly limit the conservation and usage of *Citrullus* seed material conserved ex situ; only 3% of the material is correctly and unambiguously named, in conformity to the most updated taxonomic and nomenclatural treatment. Some sort of nomenclatural inaccuracies has been found for 69% percent of the material (Fig. 2; Online Resource 2, 3), which could be merely solved by updating the databases through the application of the current taxonomic and nomenclatural treatment. On the other hand, 28% of the accessions showed taxonomic errors (Fig. 2, Online Resource 2, 3) and therefore cannot be unequivocally attributed to any existing taxon; as a consequence, they are prevented from being employed in any research, breeding, cultivation, reintroduction or conservation project. Their usage can be recovered only by means of re-determination (if a herbarium voucher of the seed accession is available) or re-propagation, followed by re-determination, which enables the revision of the accession's taxonomic identity and its re-accessioning under the correct name. If this re-propagation and re-determination process is not undertaken, research and breeding activities on 28% of the *Citrullus* accessions currently stored worldwide will be corrupted by an erroneous taxonomic identification, causing their priceless genetic variability and the potential of their useful traits to be improperly exploited. Moreover, it is possible, considering that species barriers to crossing in *Citrullus* are weak (Assis et al. 2000), that, during the re-propagation phase, introgressants or other intermediates might be found. Specifically, introduction and cropping of dessert watermelons, *Citrullus lanatus*, in parts of Africa in which other *Citrullus* species are indigenous, might have resulted in introgression of dessert watermelon genes into indigenous germplasm, thus complicating taxonomic identification. We suggest classifying these introgressants as *Citrullus* sp. in the databases, with their possible intermediate origin in their passport data. The results of our propagation experiment (Fig. 1; Table 4) show unequivocally in practical terms the existence of the problem of taxonomic identity, since the revised taxonomic identity of two out of the three propagated accessions did not correspond to the originally adopted accession name. On the other hand, the propagation experiment demonstrates the feasibility and effectiveness of a growth and re-determination phase in recovering the usability of accessions affected by taxonomic issues.

Our results clearly show great variability in the taxonomic and nomenclatural treatments adopted by the different institutions and even by the same institution (Tables 2, 3). As shown by Table 2, different accessions belonging to the same taxon are named using different synonym combinations, leading to an apparent overestimation of the taxa conserved ex situ by each institution. Specifically, in our example, two taxa (*C. lanatus* var. *citroides* and *C. lanatus* var. *capensis* in RUS001, and *C. lanatus* var. *caffer* and *C. lanatus* var. *citroides*

in HUN003) appear to be conserved rather than one (*C. amarus*) (Table 2). On the other hand, it is even worse to notice that the variability in accession naming, along with the aforementioned misnaming issues, can lead to a great underestimation of the overall number of accessions conserved per taxon. As shown in Table 3, for instance, it appears that no seed accession of *C. amarus* is conserved worldwide. This would make this species, which is widespread in southern Africa (Paris 2015) and frequently cultivated in all the tropical and sub-tropical areas of the world (Laghetta and Hammer 2007), of extremely high priority for ex situ seed conservation measures, when in fact accessions stored under four different synonyms actually exist (49 in EURISCO and 251 in Genesys). The apparent and erroneous underestimation of the number of stored accessions of a certain taxon is of particular relevance since it could diminish the importance of prioritization studies intended to unveil which taxa are currently underrepresented in long-term conservation facilities and therefore which of them should be the target of collecting missions. Prioritization studies are indeed based on the quantification of the accessions conserved ex situ of a particular target species or group of species (Maxted et al. 2010; Castañeda-Álvarez et al. 2016).

Achieving the goal of having the biggest number of accessions correctly named is fundamental to standardizing the naming process of accessions in the PGR databases, using as a reference the most updated taxonomic and nomenclatural treatments. Special attention should be paid to the choice of taxonomic ranks, without neglecting the infraspecific ones (e.g., subspecies), both for their employment in distinguishing between closely related wild and domesticated taxa (see e.g., Galasso et al. 2017), and for their possible raising to higher ranks (e.g., species) as a result of the advances in taxonomic knowledge. This latter case is illustrated by *C. lanatus*, which until recent times has been divided into various subspecies and varieties, that have been subsequently regarded as independent species (Online Resource 1). Thus, when accessions are stored under names without mention of subspecies or variety, they are almost impossible to interpret on taxonomic grounds. In our case, a simple "*C. lanatus*" entry may refer to *C. amarus*, *C. mucosospermus* or one of the two subspecies of *C. lanatus* (Tables 1, 2; Online Resource 2, 3).

To avoid the re-occurrence of all the aforementioned problems, a process of taxonomic and nomenclatural peer review is urgently needed before making each new accession public, in order to guarantee that each single accession is usable by the stakeholders. This great number of misnaming issues and variabilities in the application of taxon names also compromises the database analyses that, nowadays, are fundamental to planning and performing ex situ conservation programmes intended to find out taxa and locations underrepresented in current seed collections (Castañeda-Álvarez et al. 2016; Guzzon and Müller 2016).

It is known that the efficiency of species conservation measures greatly depends on the reliability of the taxonomic information available (Bortolus 2008). The conservation and utilization of PGR depends on the building and update of global searchable databases (Khoury et al. 2010). Our investigation demonstrates that taxonomic misnaming threatens the ex situ conservation efforts of the genus *Citrullus*. Our results confirm that taxonomic issues are a major problem in aggregator databases and that those issues have serious implications for the uncritical use of specimen data from botanical collections (Goodwin et al. 2015). On the basis of our experience, we here propose a series of actions useful for understanding the extent of the phenomenon and for solving its detrimental effects across the gene pools of plant genetic resources for food and agriculture (PGRFA):

1. Perform studies similar to the present one on further PGRFA gene-pools to unveil the extent of similar problems also in other taxa;
2. Establish taxonomic authorities in order to provide an updated and standardized taxonomic and nomenclatural treatment that should be followed in the application of the accessions' scientific names for a certain gene-pool within and the among the PGRFA databases;
3. Update the nomenclature inaccuracies following the most recent nomenclatural treatments;
4. Propagate, identify, and correct the scientific names of accessions affected by taxonomic errors;
5. Always link a herbarium voucher to each seed accession, in order to allow a quick taxon re-determination and to avoid the lengthy and costly process of re-propagation. The collection and preservation of herbarium specimens is often not considered by institutions involved in germplasm conservations;
6. Perform a peer review of the accessions' scientific names before their publication in databases.

This series of actions appears to us to be the only way to solve and prevent the occurrence of large numbers of taxonomic and nomenclatural issues with detrimental effects on PGR conservation and usage. Such procedures require investment in personnel and resources, and therefore would require adequate recognition and funding from government-supported sources.

The current study has been undertaken on the genus *Citrullus*, a relatively small genus whose taxonomy has been resolved in recent times. Nevertheless, a remarkable number of issues emerged, with consequences for the interpretation of the data stored in the databases and the practical usage of some accessions. In light of this observation, how serious may be the situation for accessions belonging to critical taxonomic groups with a great importance for food security (e.g., *Triticum*, *Musa*; see Hammer et al. 2011; Čížková et al. 2015) as well as other agriculture-related activities (e.g., *Festuca* s.l.; see Ardenghi et al. 2017), still characterized by controversial and unsolved taxonomy? This highlights once again the key-importance of taxonomic studies, often neglected in modern biology, for species conservation (Garnett and Christidis 2017). Moreover, we suggest that studies similar to the current one should be performed on more crop gene-pools that significantly contribute to global food security in order to get an idea of the scale of the issues highlighted in this study and for them to be corrected where possible. Gene-pools of different sizes will have to be considered. Of particular importance are gene-pools that had a recent taxonomic revision, for instance the pea (*Lathyrus oleraceus* subsp. *oleraceus*), lentil (*Vicia lens* subsp. *lens*), broad bean (*Vicia faba*) and wheat, which are considered some of the founder crops of Neolithic agriculture in the Fertile Crescent and hold a fundamental importance in food security (van Slageren 1994; Kilian et al. 2011; Schaefer et al. 2012).

While the current paper focuses more on PGRFA for their role in food security, the extent of taxonomic misnaming could also affect the conservation of seed accessions of endangered plants and of wild species of interest for reintroduction and translocation programmes.

Acknowledgements The authors are grateful to Robert Jarrett (USDA, U.S.A), Irina Gashkova (Vavilov Institute of Plant Industry, Russian Federation), and the Millennium Seed Bank of the Royal Botanic Garden (U.K.) for donating the seed accessions. We are also thankful to Massimo Schiavi (CREA-ORL, Italy) for performing the cultivation in purity. We also thank Thomas Abeli and Valentino Ottobriano (University of

Pavia, Italy) for taking pictures of the fruits and cultivated plants. We are also indebted to Giulia Ardenghi (Stradella, Italy) for the digital image preparation, and Jonas V. Müller (Millennium Seed Bank, U.K.) for the help in finding essential literature sources. We are thankful to Sarah Hanson (Brighton) for the linguistic check. Finally, we thank two anonymous reviewers for their extremely useful comments on an earlier version of the manuscript.

Compliance with ethical standard

Conflict of interest The authors declare that they have no conflict of interest.

References

- Appelquist WL (2016) Report of the nomenclature committee for vascular plants: 67. *Taxon* 75:169–182. <https://doi.org/10.12705/651.15>
- Ardenghi NMG, Foggi B, Orsenigo S, Maggioni L, Cauzzi P, Rossi G (2017) *Festuca* and allied genera (Poaceae) as crop wild relatives: checklists and red lists are urgently required. *Plant Biosyst* 151:6–8. <https://doi.org/10.1080/11263504.2016.1165753>
- Asseng S, Ewert F, Martre P, Rötter RP, Lobell DB, Cammarano D, Kimball BA, Ottman MJ, Wall GW, White JW, Reynolds MP, Alderman PD, Prasad PVV, Aggarwal PK, Anothai J, Basso B, Biernath C, Challinor AJ, De Sanctis G, Doltra J, Fereres E, Garcia-Vila M, Gayler S, Hoogenboom G, Hunt LA, Izaurralde RC, Jabloun M, Jones CD, Kersebaum KC, Koehler A-K, Müller C, Naresh Kumar S, Nendel C, O’Leary G, Olesen JE, Palosuo T, Priesack E, Eyshi Rezaei E, Ruane AC, Semenov MA, Shcherbak I, Stöckle C, Stratonovitch P, Streck T, Supit I, Tao F, Thorburn PJ, Waha K, Wang E, Wallach D, Wolf J, Zhao Z, Zhu Y (2015) Rising temperatures reduce global wheat production. *Nat Clim Change* 5:143–147. <https://doi.org/10.1038/nclimate2470>
- Bortolus A (2008) Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *Ambio* 37:114–118. [https://doi.org/10.1579/0044-7447\(2008\)37\[114:ECITBS\]2.0.CO;2](https://doi.org/10.1579/0044-7447(2008)37[114:ECITBS]2.0.CO;2)
- Castañeda-Álvarez NP, Khoury CK, Achicanoy HA, Bernau V, Dempewolf H, Eastwood RJ, Guarino L, Harker RH, Jarvis A, Maxted N, Müller JV, Ramírez-Villegas J, Sosa CC, Struik PC, Vincent H, Toll J (2016) Global conservation priorities for crop wild relatives. *Nat Plants* 2:16022. <https://doi.org/10.1038/nplants.2016.22>
- Chomicki G, Renner SS (2014) Watermelon origin solved with molecular phylogenetics including Linnaean material: another example of museomics. *New Phytol* 205:526–532. <https://doi.org/10.1111/nph.13163>
- Čížková J, Hřibová E, Christelová P, Van den Houwe I, Häkkinen M, Roux N et al (2015) Molecular and cytogenetic characterization of wild *Musa* species. *PLoS ONE* 10(8):e0134096. <https://doi.org/10.1371/journal.pone.0134096>
- Convention on Biological Diversity (2011) Aichi BiodiversityTargets. <https://www.cbd.int/sp/targets>. Accessed 2 Sep 2017
- Convention on Biological Diversity (2012) 2011–2020 Global Strategy for Plant Conservation Botanic Gardens Conservation International. Richmond, UK
- Davies LR, Allender CJ (2017) Who is sowing our seeds? A systematic review of the use of plant genetic resources in research. *Genet Res Crop Evol* 6:1–10. <https://doi.org/10.1007/s10722-017-0491-7>
- de Assis JGA, Queiroz M, de Araujo SMC, Bandel G, Martins PS (2000) Implications of the introgression between *Citrullus colocynthis* and *C. lanatus* characters in the taxonomy, evolutionary dynamics and breeding of watermelon. *Plant Genet Resour Newsl* 121:15–19
- Dempewolf H, Eastwood RJ, Guarino L, Khoury CK, Müller JV, Toll J (2014) Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecol Sust Food* 38:369–377
- Dempewolf H, Baute G, Anderson J, Kilian B, Smith C, Guarino L (2017) Past and future use of wild relatives in crop breeding. *Crop Sci* 57:1070–1082
- Esquinas-Alcázar J (2005) Protecting crop genetic diversity for food security: political, ethical and technical challenges. *Nature* 6:946–953. <https://doi.org/10.1038/nrg1729>
- EURISCO (2017) EURISCO: finding seeds for the future. <http://eurisco.ipk-gatersleben.de>. Accessed 29 June 2017
- FAO (2010) The second report on the state of the world’s plant genetic resources for food and agriculture. Rome. <http://www.fao.org/docrep/013/i1500e/i1500e.pdf>. Accessed 20 July 2017

- FAOSTAT (2017) Food and Agriculture Organization of the United Nations. <http://www.fao.org/faostat/en/#data>. Accessed 20 July 2017
- Ford-Lloyd BV, Schmidt M, Armstrong SJ, Barazani O, Engels J, Hadas R, Hammer K, Kell SP, Kang D, Khoshbakht K, Li Y, Long C, Lu B-R, Ma K, Nguyen VT, Qiu L, Ge S, Wei W, Zhang Z, Maxted N (2011) crop wild relatives—undervalued, underutilized and under threat? *Bioscience* 61:559–565. <https://doi.org/10.1525/bio.2011.61.7.10>
- Fursa TB (1972) К систематике рода *Citrullus* Schrad. [On the taxonomy of genus *Citrullus* Schrad.]. *Bot Zhurn (Moscow & Leningrad)* 57:31–41
- Galasso G, Banfi E, Bartolucci F, Ardenghi NMG (2017) Notulae to the Italian native vascular flora: 3. *Ital Botanist* 3:35–36. <https://doi.org/10.3897/italianbotanist.3.13200>
- Galluzzi G, Halewood M, Noriega IL, Vernooy R (2016) Twenty-five years of international exchanges of plant genetic resources facilitated by the CGIAR genebanks: a case study on global interdependence. *Biodiver Conserv* 25:1421–1446. <https://doi.org/10.1007/s10531-016-1109-7>
- Garnett ST, Christidis L (2017) Taxonomy anarchy hampers conservation. *Nature* 546:25–27. <https://doi.org/10.1038/546025a>
- Genesys (2017) Gateway to genetic resources. <https://www.genesys-pgr.org/welcome>. Accessed 20 May 2017
- Godfray H CJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C (2010) Food Security: the challenge of feeding 9 billion people. *Science* 327:812–818. <https://doi.org/10.1126/science.1185383>
- Goodwin ZA, Harris DJ, Filer D, Wood JRI, Scotland RW (2015) Widespread mistaken identity in tropical plant collections. *Curr Biol* 25:R1057–R1069
- Guarino L, Lobell DB (2011) A walk on the wild side. *Nat Clim Change* 1:374–375. <https://doi.org/10.1038/nclimate1272>
- Guzzon F, Müller JV (2016) Current availability of seed material of onset (*Ensete ventricosum*, Musaceae) and its Sub-Saharan wild relatives. *Genet Res Crop Evol* 63:185–191
- Hammer K, Knupffer H, Xhuveli L, Perrino P (1996) Estimating genetic erosion in landraces—two case studies. *Genet Resour Crop Evol* 43:329–336
- Hammer K, Filatenko AA, Pistrick K (2011) Taxonomic remarks on *Triticum* L. and x *Triticosecale* Wittm. *Genet Res Crop Evol* 58:3–10
- Kell S, Marino M, Maxted N (2017) Bottlenecks in the PGRFA use system: stakeholders’ perspectives. *Euphytica* 230:170. <https://doi.org/10.1007/s10681-017-1935-z>
- Khoury C, Laliberté B, Guarino L (2010) Trends in ex situ conservation of plant genetic resources: a review of global crop and regional conservation strategies. *Genet Res Crop Evol* 57:625–639. <https://doi.org/10.1007/s10722-010-9534-z>
- Kilian B, Mammen K, Millet E, Sharma R, Graner A, Salamini F, Özkan H (2011) *Aegilops*. In: Kole C (ed) Wild crop relatives genomic and breeding resources, cereals. Springer, Berlin, pp 1–76
- Laghetti G, Hammer K (2007) The Corsican citron melon (*Citrullus lanatus* (Thunb.) Matsum. et Nakai subsp. *lanatus* var. *citroides* (Bailey) Mansf. ex Greb.) a traditional and neglected crop. *Genet Res Crop Evol* 54:913–916. <https://doi.org/10.1007/s10722-007-9220-y>
- Last L, Arndorfer M, Balázs K, Dennis P, Dyman T, Fjellstad W, Friedel JK, Herzog F, Jeanneret P, Lüscher G, Moreno G, Kwikiriza N, Gomiero T, Paoletti MG, Pointereau MG, Sarthou J-P, Stoyanova S, Wolfrum S, Kölliker R (2014) Indicators for the on-farm assessment of crop cultivar and livestock breed diversity: a survey-based participatory approach. *Biodivers Conserv* 23:3051–3071
- Li DZ, Pritchard CW (2009) The science and economics of ex situ plant conservation. *Trends Plant Sci* 14:614–621. <https://doi.org/10.1016/j.tplants.2009.09.005>
- Mansfeld R (1959) Vorläufiges Verzeichnis landwirtschaftlich oder gärtnerisch kultivierter pflanzenarten. *Kulturpflanze* 2:1–659
- Matsumura J, Nakai T (1916) *Catalogus seminum et sporarum in Horto Botanico Universitatis Imperialis Tokyoensis per annos 1915 et 1916 lectorum*. Tokyo
- Maxted N, Kell SP, Toledo Á, Dulloo E, Heywood V, Hodgkin T, Hunter D, Guarino L, Jarvis A, Ford-Lloyd B (2010) A global approach to crop wild relative conservation: securing the gene pool for food and agriculture. *Kew Bull* 65:561–576. <https://doi.org/10.1007/s12225-011-9253-4>
- McCouch S, Baute GJ, Bradeen J, Bramel P, Bretting PK, Buckler E, Burke JM, Charest D, Cloutier S, Cole G, Dempewolf H, Dingkuhn M, Feuillet C, Gepts P, Grattapaglia D, Guarino L, Jackson S, Knapp S, Langridge P, Lawton-Rauh A, Lijua Q, Lusty C, Michael T, Myles S, Naito K, Nelson RL, Pontarollo R, Richards CM, Rieseberg L, Ross-Ibarra J, Rounsley S, Hamilton RS, Schurr U, Stein N, Tomooka N, van der Knaap E, van Tassel D, Toll J, Valls J, Varshney RK, Ward J, Waugh R, Wenzl P, Zamir D (2013) Feeding the future. *Nature* 499:23–24. <https://doi.org/10.1038/499023a>

- McNeill J, Barrie FR, Buck WR, Demoulin V, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Marhold K, Prado J, Prud'homme Van Reine WF, Smith GF, Wiersema JH, Turland NJ (eds) (2012) International Code of Nomenclature for algae, fungi, and plants (Melbourne Code). Regnum Vegetabile 154. Königstein: Koeltz Scientific Books. [online ed., <http://www.iapt-taxon.org/nomen/main.php>]
- Modi AT, Zulu NS (2012) Watermelon landrace seedling establishment and field performance in response to differing water regimes. *Afr J of Agric Res* 7:6016–6021. <https://doi.org/10.5897/ajar11.2087>
- Mujaju C, Zborowska A, Werlemark G, Garkava-Gustavsson L, Andersen SB, Nybom H (2011) Genetic diversity among and within watermelon (*Citrullus lanatus*) landraces in Southern Africa. *J Hortic Sci Biotechnol* 86:353–358. <https://doi.org/10.5897/AJAR11.2087>
- Nesom GL (2011) Toward consistency of taxonomic rank in wild/domesticated Cucurbitaceae. *Phytoneuron* 13:1–33
- Pangalo KI (1930) Арбузы Северного полушария. [Watermelons of the northern hemisphere]. *Trudy po prikladnoi botanike, genetike i selektsii* 23:41–84
- Paris HS (2015) Origin and emergence of the sweet dessert watermelon, *Citrullus lanatus*. *Ann Bot* 116:133–148. <https://doi.org/10.1093/aob/mcv077>
- Renner SS, Chomicki G, Greuter W (2014) (2313) Proposal to conserve the name *Momordica lanata* (*Citrullus lanatus*) (watermelon, Cucurbitaceae), with a conserved type, against *Citrullus battich*. *Taxon* 63:941–942. <https://doi.org/10.12705/634.29>
- Rivière S, Müller JV (2017) Contribution of seed banks across Europe towards the 2020 Global Strategy for plant conservation targets, assessed through the ENSCONET database. *Oryx*. <https://doi.org/10.1017/S0030605316001496>
- Schaefer H, Renner SS (2011) Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). *Taxon* 60:122–138
- Schaefer H, Hechenleitner P, Santos-Guerra A, Menezes de Sequeira M, Pennington RT, Kenicer G, Carine MA (2012) Systematics, biogeography, and character evolution of the legume tribe Fabaeae with special focus on the middle-Atlantic island lineages. *BMC Evol Biol* 12:250
- Tilman D, Cassman KG, Matson PA, Naylor R, Polasky S (2002) Agricultural sustainability and intensive production practices. *Nature* 418:671–677. <https://doi.org/10.1038/nature01014>
- van de Wouw M, Kik C, van Hintum T, van Treuren R, Visser B (2010) Genetic erosion in crops: concept, research results and challenges. *Plant Genet Resour* 8:1–15
- van Slageren MW (1994) Wild wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub. & Spach) Eig (*Poaceae*). Agricultural University Papers, Wageningen
- Veteläinen M, Negri V, Maxted N (2009) European landraces: on-farm conservation management and use. *Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome
- Vincent H, Wiersema J, Kell SP, Dobbie S, Fielder H, Castañeda-Álvarez NP, Guarino L, Eastwood R, Leon B, Maxted N (2013) A prioritised crop wild relative inventory as a first step to help underpin global food security. *Biol Conserv* 167:265–275. <https://doi.org/10.1016/j.biocon.2013.08.011>
- Warschewsky E, Varma Penmetsa R, Cook DR, von Wetttemberg EJB (2014) Back to the wild: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *Am J Bot* 101(10):1791–1800. <https://doi.org/10.3732/ajb.1400116>