Principles for the Inclusion of CWR Data in EURISCO

Prepared by Theo van Hintum and José Iriondo, within the framework of the ECPGR project 'Extension of EURISCO for Crop Wild Relatives (CWR) in situ *data and preparation of pilot countries' data sets' (CWR data in EURISCO), funded by the German Federal Ministry of Food and Agriculture.*

Agreed by project partners and EURISCO Advisory Committee in May 2022

Introduction

Populations of crop wild relatives (CWR)[1] occurring *in situ* are potentially valuable resources for crop science and plant breeding. Therefore, they need to be conserved and made available to users. However, the current conservation of, and access to these CWR populations varies strongly. *In situ* conservation of CWRs is often in the hands of nature conservation organizations, who are sometimes not even aware that they are managing these resources. Other CWR populations occur in farmers' fields, roadsides and other locations, where they are not managed at all. Furthermore, information about the CWR populations, their occurrence and availability, is hardly available.

The issue of CWRs has recently received much attention, e.g. from EU-funded projects such as Farmer's Pride and from the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA). The latter coordinated and led the publication of a descriptor list for CWRs conserved *in situ* (Alercia *et al.*, 2021). For ECPGR and its database EURISCO, the issue of properly handling information about *in situ* CWRs has been on the table for a while, but for various reasons never resulted in substantial improvements.

In Europe, depending on the country, information about CWRs is heterogeneous: sometimes it is scattered over various sources or not available at all, whereas in some other European countries national checklists of CWRs, priority lists, population occurrence records and *ex situ* and *in situ* conservation assessments are available (e.g. Maxted et al., 2007; Smekalova, 2008; Phillips et al., 2014; Landucci et al., 2014; Labokas et al., 2018; Taylor et al., 2017; Rubio Teso et al. 2018; van Treuren et al. 2017). In some cases, specific websites have been created to showcase CWR in a country, providing information about the occurrence, distribution, availability and other data (such as https://www.cwrnl.nl/en/CWRnl-1.htm with information about CWRs occurring in the Netherlands). This heterogeneity of cases is one of the reasons why it is difficult for users (plant breeders and crop scientists) to find out about and access these resources.

To improve the situation this proposal aims at:

1- Supporting the development of CWR National Inventories providing information on the CWR taxa and occurrence of CWR populations, their conservation status and their availability.
2- Feeding EURISCO with information on CWR populations that are – in principle – available.

The approach suggested in this proposal is based on various documents such as the *Concept for a possible extension of EURISCO for* in situ *crop wild relative and on-farm landrace data* (Weise *et al.* 2020), the *Descriptors for Crop Wild Relatives conserved* in situ (Alercia *et al.* 2021), and on discussions held in various platforms. It also relies on previous publications on CWR descriptors, such as those of Thormann et al. (2017) and Bioversity International & University of Birmingham (2017). It presents an approach that will hopefully result in the desired outcome: a) properly organized information about CWRs occurring in a country, arranged in an *in situ* CWR National Inventory, and b) information in EURISCO about CWR populations occurring *in situ* and potentially accessible for use. However, success will depend on the willingness of the European countries to create these National Inventories for CWRs and

---

[1] This document is focused on CWR, however, most of the approaches proposed here can also be applied to wild food plants.

to share data from them. This development would enable them to comply with the recommendations of international plans, treaties and conventions, including the *Plant Genetic Resources Strategy for Europe* (https://ecpgr.cgiar.org/pgrstrategy21; hereafter, PGR Strategy). It will also be necessary that EURISCO accommodates the required changes. The ECPGR project 'CWR data in EURISCO', funded by the German Federal Ministry of Food and Agriculture, will make it possible to prepare EURISCO for data inclusion and to organize the data flow mechanism from National Inventories. A few pilot countries will be enabled to provide their data and test the system.

Background

CWRs in the global context

The study of the effects of global change on our planet's biodiversity has increased the awareness about the vulnerability of plant genetic resources occurring in natural habitats and the need to take action to conserve them. Thus, the *Second Global Plan of Action for Plant Genetic Resources for Food and Agriculture* (FAO, 2011) identified the study and inventory of PGRFA as one of the first priority actions that countries should take to address this problem, as well as the need to promote regional collaboration and coordination, taking into account the cross-border distribution of species. At the European level, the *ECPGR Concept for* in situ *conservation of crop wild relatives in Europe* (Maxted et al. 2015) stressed the importance of identifying the important CWR diversity both at the national and regional levels. In this context, Weise et al. (2020) explored the possible extension of EURISCO for *in situ* CWR and on-farm landrace data and proposed a set of descriptors that could be used for this purpose. In 2021, the Secretariat of the FAO ITPGRFA published an international standard of descriptors for CWRs conserved *in situ* in an attempt to promote the documentation of these genetic resources and "enable countries to compile and exchange data held by different national and international organizations" (Alercia et al., 2021). Presently, the European PGR Strategy, which has been presented in the context of the GenRes Bridge project (http://www.genresbridge.eu/), also identifies the objective of significantly increasing CWR and wild food plant inventories in Europe "to enable a more comprehensive view of available CWRs and wild food plants' (WFP) genetic diversity, to better understand how this diversity is distributed across the region and its neighbouring countries, and to identify which are the priority populations to actively conserve". Among the actions to take in connection with this objective, the PGR Strategy recommends supporting "national authorities to appoint CWR National Focal Points and then provide data – as appropriate – from *in situ* conserved populations of CWRs and WFP (from National Inventories) to a centralized information system, based on agreed data exchange standards".

CWRs conserved ex situ

CWRs do not only occur *in situ* (in their wild habitat). Many populations have been sampled and included in *ex situ* genebanks. These are documented in Europe in EURISCO. EURISCO was set up between 2000 and 2003 in a EU-funded project to provide a complete overview of PGR in Europe. It concentrated purely on *ex situ* genebanks. National Focal Points (NFP) were nominated in all European countries. These NFPs were supported in developing National Inventories by combining the data of the genebanks in the country. The data in these National Inventories were uploaded in a standardized format, based on the Multi-Crop Passport Descriptor (MCPD) list (Alercia et al., 2015), to a central repository: EURISCO. This approach proved successful in the *ex situ* genebanks community. Furthermore, building the capacity of genebank documentation specialists resulted in a substantial improvement in the quality of genebank documentation, and sharing of information resulted in increased transparency of genebank activities. Now EURISCO provides an overview of the content of most genebanks in Europe, allowing PGR users and colleague genebanks to identify materials. There is, however, no assurance that the material in EURISCO is available for use.

Currently, based on data downloaded on 2 December 2021, EURISCO contains data on 1,364,862 accessions in European genebanks (where the UK *Arabidopsis* accessions and the 'historic accessions', which are accessions that no longer exist, are excluded). About 22% of these do not have an indication

about the biological status of the accessions, and of the remaining 1,064,326 accessions, nearly 25% are labelled as 'wild' or 'weedy', accounting for 265,242 accessions. This considerable number of accessions could be considered to be CWRs conserved *ex situ* in European (*s.l.*) genebanks. About half of these *ex situ* conserved wild populations (49.1%) originate from European countries (incl. Russia, excl. Turkey). Closer inspection of these wild species shows, however, that 31.3% of the accessions are maintained in the Millennium Seed Bank of the Royal Botanic Gardens, Kew, the UK (and might thus not be true CWR populations, but wild species that are not related to crops). Together with the Vavilov Institute in Russia and the Lieberman Germplasm Bank in Israel, nearly half (49.1%) of the accessions are accounted for. Also looking at the genera covered by the CWRs listed in EURISCO, a high level of skewness can be observed: only 21 genera account for half (49.8%) of the accessions, with cereals (44%) and grasses (24%) contributing accessions most to these 21 genera. After removing the 4.4% of the accessions without a known origin country, it appears that half of all CWR populations originated in just ten countries, with Israel (12.3%) contributing most, followed by Spain (7.4%) and Russia (5.6%). Surprisingly, Madagascar appears in place nine on the ranking of best-represented countries with 2.6% of the wild accessions, but these were all held by the Millennium Seed Bank. (A more in-depth study of the conservation of CWR in *ex situ* genebanks was not possible in the timeframe of the preparation of this proposal, it would however be very interesting.)

It can be concluded that, although a very large number of wild and weedy populations are conserved *ex situ* in European genebanks, these are not necessarily CWRs or European, and those that are, are highly skewed towards a few cereal and forage crops from a highly restricted geographic origin. The representation of specific crops or crop groups needs to be considered on a crop-by-crop basis.

CWRs in situ *in EURISCO*

As indicated above, CWR in *ex situ* genebank collections are included in EURISCO. And although it is in principle possible to include *in situ* populations in EURISCO, provided that they are managed at the standard of *ex situ* collections, i.e. with a 'holding institute' that can also be approached for access (Maxted and Palme, 2016; Maxted 2021), this approach has never been implemented. However, assuming that EURISCO should provide information about all PGR that are – in principle – available in Europe, the scope regarding CWRs could be expanded to include information about naturally occurring CWR populations, as long as there is an entry point for obtaining them.

In an optimal situation, access to PGR material would be provided by the owner/manager of the PGR population: the genebank managing the accession, the farmer growing the landrace or the manager of the nature reserve where a CWR population grows. However, in many cases that will not work since many owners/managers are not experienced (or interested) in providing access. In these situations, it might be more effective to have a liaison organization acting as an intermediary between the PGR owner/manager and the user: if the user wants to have access to specific PGR material, he/she approaches the liaison organization that mediates between the user and the owner/manager. Certainly, in the case of CWRs, this might improve access considerably. To facilitate the entire process of identifying potentially available *in situ* CWR populations and the respective owners/managers and liaison organizations, an *In situ* CWR National Inventory Focal Point should be established. Depending on the country, this Focal Point could be the genebank that currently provides the EURISCO National Focal Point, or another specific institution dedicated to this challenge. This Focal Point could be tasked to create a national *in situ* CWR database that includes the checklist of CWR taxa that the country wished to actively conserve, plus extended information that allows making this decision. This database would also contain occurrence data of populations corresponding to those prioritized taxa, which would help determine the best locations where active *in situ* conservation can be implemented and the best populations where the material can be made accessible. This institution would then upload the latter to EURISCO.

This proposal intends to support a collaborative process at the European level aiming to organize information on *in situ* CWR occurrence, conservation status and availability. To reach this objective, the establishment of two components is proposed and promoted:

1) Development of *in situ* CWR National Inventories including population data (hereafter, CWR-NIs)
2) Transfer of selected data from the CWR-NIs to a central online available database: EURISCO

The CWR-NIs will document the CWR populations in a country, allowing better management of these populations, and access to selected information about the populations documented in the CWR-NI that are in principle accessible, in EURISCO.

Box 1. Scope of the CWR-NI as opposed to the CWR *in situ* extension of EURISCO

The ***In situ* CWR National Inventory** is created by a designated *In situ* CWR National Inventory Focal Point and includes a list of occurrences of populations belonging to those CWR taxa that the country has identified as a priority for conservation. Descriptors recommended for the generation of this CWR-NI are listed in Annex 1 since these also facilitate the data transfer to EURISCO, but it is up to the individual country to decide whether to extend, reduce or modify this list. The CWR-Ni aims to monitor the extent of CWR diversity in the country and offer the means to identify the most suitable populations for active conservation.

The CWR *in situ* extension of EURISCO is created by including in EURISCO a subsection of the population data of each CWR-NI, corresponding to those populations that are 'actively conserved' (see definition in Box 2) and might potentially be made available to users. The identification of this subsection will be at the discretion of each country/Focal Point and the purpose will be to offer to the public a list of populations for which it is possible to identify an owner/managing institution and a liaison institution (if different from previous) that can be approached to inquire about the possibility and modality to obtain genetic material. In other words, these are lists of populations, a sample of which might be in principle available to users. Descriptors to be used are necessarily those listed in Annex 2, four of which are mandatory.

Approach

To allow access to information about *in situ* conserved PGR in EURISCO, two elements are required:

- The development of CWR-NIs, providing overviews of the CWR populations or species occurring in the country with indication of their location, last observation, management, managing organization, liaison, terms of use, etc. These CWR-NIs should be created by, or on behalf of, the national administration in charge of PGR.
- Organization of a proper flow of information about the CWR populations to be made visible from the CWR-NI to EURISCO. The extent to which the content of the CWR-NIs is to be made visible in EURISCO, both in terms of which populations/species and what data about these items, is a local decision to be made by the administration under which the inventory was created.

In the German-funded ECPGR project, these two elements will be developed. First, in a few countries, CWR-NIs will be created, allowing the development of a practical and generally applicable methodology for doing so. Secondly, the adjustments to EURISCO and the flow of data from the CWR-NIs to EURISCO will be facilitated and initiated with data from the selected countries.

These two steps of the approach will be described in more detail below.

The development of CWR-NIs

The first step in the development of CWR-NIs consists of the identification and selection of the CWRs of interest for the country to obtain a checklist of priority CWRs. There are different approaches to doing this (Maxted et al., 2013), but most of them start with the compilation of a list of the national flora and the identification of the crops whose CWRs will be considered. The identification of these crops should consider the socioeconomical relevance of crops both at the national and the international level or could involve a complete list of crop genera. The relevance of crops at the country level can be obtained from statistical reports about crop production areas and economic revenues of crops in the country, and the country's plant variety list. At the international level, the FAO world primary crop list (FAO, 2022), the crops or taxa included in Annex 1 of the ITPGRFA (ITPGRFA, 2022) or the EU database of registered plant varieties (EU, 2022) may be a good starting point. The data regarding the selection of the crops to be considered and the motivation and/or prioritization might be useful to store. However, as this would increase the complexity of this step, the decision to do so and how to do it is left to the actor(s) responsible for this step.

Once the target crops have been identified, the populations of the wild plant species occurring in the country that belong to the same genera as these crops should be inventoried. In some crops, the scope may be broadened to include species of closely related genera (e.g. the genera *Beta* and *Patellifolia* in beet) that are still considered to be part of the crop genepool. Further selection and/or prioritization is usually made based on the crossability of CWRs with their corresponding crops, the threat status of the species and their endemic or native status. Magos Brehm et al. (2017), Kell et al. (2017), and Nielsen et al. (2017) provide a good synthesis of the criteria and procedures that can be used in this step. Some illustrative examples of the generation of CWR checklists are available in Maxted et al. (2007), Smekalova (2008), Phillips et al. (2014), Landucci et al. (2014), Labokas et al. (2018), Taylor et al. (2017), Rubio Teso et al. (2018) and van Treuren et al. (2017). The taxonomic level of the checklist should also be defined at this stage (i.e. whether it will be composed of just species or also contain taxa at the infraspecific level).

The second step involves the creation of a database structure that will hold the information that is considered of interest for the management and use of the National Inventory of *in situ* CWRs.

The database structure will contain 1) information at the taxon level that was used in the generation of the CWR checklist and 2) information at the population level that will provide the specific details about each population:

1. Information available at the taxon level used for the generation of the checklist:

    - Taxonomy of the CWR (family, genus, species, subtaxon, authority, common name)
    - Crossability of the CWR with the associated crop (genepool)
    - Threat status, legislative protection, endemicity
    - Related crop (scientific name and/or common name)

2. Information at the population level to provide the characteristics of each record:

    - Site descriptors (geographic coordinates, name of the site or municipality, country, site protection, habitat descriptors)
    - Population descriptors (most recent observation date, holding institution, biological status, presence of *ex situ* accessions, herbarium specimen, availability of the material)
    - Population management descriptors (threats, conservation actions in place)

In case additional information about the population is available, such as characterization and evaluation data, or information about the trends of population size across time, etc., appropriate solutions (i.e. specific dataframes) should be created (this was considered outside the scope of this proposal). The

Interactive Tool for Crop Wild Relative Conservation Planning (Magos Brehm et al., 2017) provides a detailed account of the process and information sources that can be followed for the development of the National databases on *in situ* CWR.

The descriptors to be used to account for all this information should preferably, as far as possible, follow international standards. A list of proposed descriptors for the CWR-NI is included in Annex 1. This list is based on the principles of efficiency and feasibility and it is, consequently, a selection of descriptors that have been proposed in previous works, such as the FAO *Descriptors for Crop Wild Relatives conserved* in situ (Alercia et al., 2021), the *Concept for a possible extension of EURISCO for* in situ *crop wild relative and on-farm landrace data* (Weise *et al*., 2020), the *Crop wild relative checklist and inventory descriptors v.1.* (Bioversity International and University of Birmingham, 2017), and *CWR checklist and inventory data template v.1* (Thormann *et al*., 2017), as well as the current descriptors for passport data used in EURISCO for *ex situ* accessions.

The third step involves the process for collecting population occurrence data to feed the database. A related concern is to determine the criteria for deciding when a CWR population can be considered as an *in situ* accession that is worth recording into a NI and making it available to potential users. Active *in situ* conservation involves the location, designation, management and monitoring of target taxa in the location where they are found (Maxted, 1997). A strict application of this term would render just a handful of records across Europe and would make pointless the *in situ* extension of EURISCO. On the other hand, a more lenient approach to this concept could include populations that are likely to exist at present, whose location is known, where the land management is compatible with the persistence of the population, and where there is a management institution or person that can be approached that is likely to facilitate access to the material. This approach would provide potential PGR users with a relevant complementary source of material to consider in their breeding programmes. It would also provide *in situ* conservation stakeholders in the country with a set of populations that could be taken into consideration for the establishment of genetic reserves. Although it is up to each country to decide which criteria suit their interests best, some considerations to take in mind are:

1. The most recent observations should be not too old, to ensure that the population currently still exists. Records where current presence of the population is confirmed should be prioritized.

2. Records with precise geolocation data should be prioritized because they will be easier to find if access to them is required. A particular situation concerns widespread species with continuous distributions over a large territory. In these cases, precise geolocation data is not necessary, and their presence may be reported in a more general way, for instance, by adding the coordinates and specifying a radius, or indicating their ubiquity in the country or certain regions.

3. Populations occurring in public land, protected areas, collaborating farms or in long-term ecological research infrastructures (LTER; https://www.lter-europe.net/), where the managers of the land are aware of their existence, may be more likely to be in good status and accessible for use.

4. Populations that, according to the landowners of the site and the competent public authorities, are available for access under the Multilateral System (MLS) of Access and Benefit-sharing of the ITPGRFA should be included.

In essence, the most relevant CWR populations to be considered as *in situ* accessions in the NI will be those whose current presence and precise location are known, are being actively conserved to guarantee their long-term persistence, and are available for access under the MLS.

Site and population occurrence data can be obtained from national and global databases, such as GBIF (https://www.gbif.org/), and biodiversity databases created by national and subnational public administrations and NGOs that may be accessible. Alternatively, a more time-consuming approach is gathering information through searches in public herbaria and chorological bibliographic references.

Special attention must be given to the correct use of taxonomic nomenclature and synonyms in the data query. The information obtained from these sources will often contain a high degree of redundancy and several errors. Many records do not contain accurate geographic coordinates, or the given coordinates are wrong. Therefore, this data collection process requires a significant effort of data curation to solve taxonomic nomenclatural problems and eliminate redundancies and low-quality records. Rubio Teso et al. (2020) provide a detailed account of the data collection and curation processes followed in the preparation of an occurrence database for European priority CWR taxa.

It should be noted that the records obtained this way may refer to population observations made many years back in time. Therefore, it is important to confirm the present occurrence of those populations of interest. To facilitate this and to obtain additional data on *in situ* conserved CWRs, it is advisable to create a national network of stakeholders that may be able to provide direct information on *in situ* CWR populations that are being actively managed and confirm the presence of other targeted CWR populations. This network could include technical staff from subnational and local administrations that are involved in the conservation of legally protected threatened CWRs, and the conservation of protected habitats that include target CWRs as characteristic species (e.g. Annex I habitats in Natura 2000 sites), farmers and farmer associations that conserve CWRs in agricultural margins (CWR population data is likely to be used in the near future in monitoring environmental stewardship schemes in agricultural policies), scientists that manage long-term ecological research (LTER) infrastructures that contain target CWRs, and other people involved in initiatives comprising CWR *in situ* conservation. Ideally, the database containing the CWR-NI should establish a platform where the members of this network could feed their data.

The CWR-NI should be envisioned as a dynamic database that is subject to periodical updates due to changes in the composition of the national CWR checklist, correction of existing data and new data acquisition on population occurrences and active conservation practices. Therefore, each country should periodically update occurrence data from global biodiversity databases and consider the addition of new records to the national *in situ* CWR database after proper assessment.

Once the information about CWRs in a country is compiled, it can be used for multiple purposes, such as developing a national conservation strategy, ensuring the long-term conservation of populations, starting with raising awareness amongst owner/managers and then ensuring commitment and funds, setting up a monitoring system for the vulnerable species, combined with *ex situ* back up. In addition, research and dissemination publications can also be developed, and a website can be made available to raise awareness of these important resources in the country. As such, this CWR-NI is a national tool that can be useful in many instances. It can also be used to identify the CWR populations that can be made available to scientists and plant breeders. Here EURISCO comes into play, as in the next step the flow of selected information from the CWR-NI to EURISCO will be described.

Box 2. Actively conserved populations

EURISCO should receive data only about populations that can (in principle) be made available to users. This restricts the scope to those populations that are 'actively conserved'. As explained in the text, criteria defining 'active conservation' should not be too strict. They may include populations that are likely to exist at the present time, whose location is known, where the land management is compatible with the persistence of the population, and where there is a management institution or person that can be approached that is likely to facilitate the access to the material.

Box 3. Access to the *in situ* material

When *in situ* populations are included in EURISCO and thus considered 'accessible in principle' to users, this does not mean that the user will necessarily be allowed to reach the populations and autonomously collect material of interest (even though some countries may consent to this). It rather means that there is a designated pathway whereby the user can approach a managing (or liaison) institution and receive information on the appropriate way to obtain the material, including terms and conditions, which will vary from country to country (compatibly with the terms of the ITPGRFA Multilateral System or the Convention on Biological Diversity).

Flow of information from the NIs to EURISCO

Once the CWR-NI has been created, selected data can be uploaded to EURISCO to provide a common entry point to enable users to retrieve standardized information on European CWR *in situ* accessions that are in principle accessible for use in scientific research and breeding. This flow of information from the NIs to EURISCO has three aspects. Firstly, it needs to be decided what information can be shared with EURISCO (but also in other public databases). This is a local decision, ECPGR cannot dictate which CWR populations should be made visible in EURISCO, and what information about these populations should be shared. Obviously, in this step, the privacy of the actors involved (as defined in the EU by the General Data Protection Regulation) needs to be respected. From the EURISCO side, the only two requirements are that the information follows the format rules as defined below, and that a potential user should be able, based on the data provided, to approach a contact in the country asking for the possibility to access the population or species. Secondly, the flow of information needs to be facilitated, as it is for the information in the NIs for *ex situ* PGR. An uploading mechanism, very similar to that for *ex situ* data, needs to be created. Obviously, uploading data should only be done with the proper national authorizations, but that is a national issue to be resolved. Thirdly, EURISCO needs to be slightly adapted to be able to accept and present the new information. The proposed changes might also benefit the *ex situ* community, as the principle of presenting a contact for access might be also interesting to them.

(1) Structure of information shared between the CWR-NI and EURISCO

The structure of the information drawn from the CWR-NIs should follow the structure of EURISCO; most descriptors proposed for the CWR-NIs are the same as the ones currently used for uploading information to EURISCO (see Weise, 2021) and are in line with the *Descriptors for Crop Wild Relatives conserved under* in situ *conditions* as published by FAO in 2021. It should be easy to create one file for uploading the selected content of both NIs (*in situ* and *ex situ*). A proposal for a descriptor list for upload of *in situ* CWR records from the CWR-NI is presented in Annex 2.

The basic concept in mapping the CWR information on the current EURISCO structure is that the conserved CWR population is considered a PGR similar to an accession in an *ex situ* genebank. The organization managing the population is like the genebank managing an accession, and the identifier (population ID or whatever it is called locally) is like the accession number. The only new concept is the organization that liaises between a potential user and this managing organization. All other EURISCO descriptors can be used for *in situ* managed CWR populations, although sometimes a slightly wider interpretation is required as compared to the existing definition of these descriptors. Also, a new status for the descriptor describing the type of maintenance (STORAGE) needs to be added describing the '*in situ*' status of this PGR.

A complication to including *in situ* CWR information in the current EURISCO structure is that the descriptor of the organization managing the *in situ* CWR population (MNGINSTNAME in the FAO

descriptor list for CWR) is different from the descriptors used for the institute where an *ex situ* sample of this population is maintained (INSTCODE and INSTNAME in the FAO list). In this proposal, in Annex 2, the organization managing the CWR population is considered equivalent to the INSTCODE and INSTNAME in EURISCO, and the population identifier (POPID) equivalent to the accession number (ACCENUMB). Any *ex situ* accession associated with one *in situ* population can be documented in the field for other numbers (OTHERNUMB), documenting the occurrences of these accessions in other institutions.

As not all managing institutions have a FAO-code (and may not be able to obtain one, because the FAO WIEWS code is used for *ex situ* accessions), the addition of INSTNAME was considered necessary – this implies a new descriptor in EURISCO. Other descriptors not yet known to EURISCO are LIAISONCODE, LIAISONNAME, SITEPROT, CONSACTION and POPSRC.

All other descriptors for the upload to EURISCO can remain unchanged – with the exceptions mentioned above.

(2) Uploading CWR-NI to EURISCO

The upload of data from the CWR-NI should be done by a focal point authorized by the ECPGR National Coordinator, one per country. It can be the NFP currently responsible for the upload of the *ex situ* NI, but it could also be someone else, possibly closer to the CWR-NI.

The first step in the upload process is the selection of data from the NI. It needs to be decided what populations and which data can be shared with the world via EURISCO. This can be best documented in the CWR-NI itself (by including descriptors with flags for this purpose); it should be clear for an automatic process which data points can be shared and which cannot.

The decision on which data to share will be based on various considerations, including the reliability of the record, the possibility of providing material of the population and the national policy about disclosure of geographic data of wild plant populations. However, the objective is to share as much information as possible. Sharing information does not necessarily imply accuracy of the information or guaranteed availability of the PGR.

After extraction of the relevant data, they need to be formatted according to the EURISCO *in situ* CWR Upload Format (proposed in Annex 2). The best approach would be to assure that the format in the CWR-NI already follows the Upload Format, but this is not required. If it doesn't follow the format, a recoding step (changing column names, recoding values, etc.) will be required.

Once the upload file is created and properly formatted, an uploading mechanism similar to that available for the *ex situ* NIs should be available. Ideally, it should be the same mechanism that can be used for either *ex situ or in situ* CWR upload files.

(3) Required changes in EURISCO

The required changes in the database structure of EURISCO are limited to the new descriptors:

- The descriptor INSTNAME is new and should be used in case no FAO-WIEWS institution code was available for the institute managing the accession. In these cases, the code 'DUMMY' can be used (see Annex 2).
- The descriptors LIAISONCODE and LIAISONNAME need to be added, documenting the institution that can liaise between the organization managing the CWR population (in INSTCODE and INSTNAME) and the interested user. If these fields are empty, the managing organization can be approached for access (as with *ex situ* genebanks).
- The descriptors SITEPROT, documenting the protection status of the location where the CWR occurs, CONSACTION, describing the type of conservation that occurs, and POPSRC describing the

habitat of the occurrence site of the population, must be added. All three descriptors have a finite set of allowed values (see Annex 2).

- The STORAGE descriptor needs an additional status '60 – *in situ* population'.

The uploading mechanism for the *in situ* CWR data will need to be created accordingly.

The search interface on the web should also be adjusted according to the new descriptors and the new descriptor state (the state '60' to denote *in situ* wild populations in the descriptor 'Type of storage (STORAGE)'). Once a critical mass of *in situ* CWR data is added to EURISCO, a dedicated entry point of the EURISCO website will be added to highlight the occurrence of *in situ* managed CWR in EURISCO. When downloading data, the possibility of selecting or excluding *in situ* data should be added.

Recommendations for next steps

1   To support the creation of CWR-NIs in European countries, specific resources can be created, such as:
   - A web portal with links to the different sources that may be needed to elaborate the national CWR checklist, including the lists of crops that can be used for the selection of CWRs, such as the FAO world primary crop list (FAO, 2022), the crops or taxa included in Annex 1 of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA, 2022) or the EU database of registered plant varieties (EU, 2022), the GRIN-Global CWR dataset, access to GBIF and other global biodiversity databases, and relevant publications about the preparation of CWR checklists.
   - A reference database with information for inclusion in the taxon level (see Annex 1) of the CWR-NIs with information about the CWR taxa in Europe.
   - A customized R script to help countries download CWR occurrence data from GBIF and filter downloaded information according to different quality criteria.
   - An easy prototype database (in Excel or MS Access) to be used for storing the data that might include scripts for extracting the EURISCO upload file.
2   A descriptor list for the upload of data from the CWR-NI should be written and published, possibly based on the proposed descriptors in Annex 2 of this proposal.
3   The EURISCO database structure and search interface should be modified to accommodate the descriptors in the upload format, formulated in the previous point.
4   The mechanism for uploading data should be adjusted according to the descriptor list, formulated in point 2.
5   Once several CWR-NIs have been prepared, data should be uploaded to EURISCO. This feat should be heavily publicized to inspire other countries to join this initiative.
6   To support additional countries interested in creating a CWR-NI, easy access should be provided to various manuals, tools and information to optimally support them in this effort and the subsequent step of uploading the result to EURISCO.

It is to be noted that all changes involving the EURISCO database, search interface and upload mechanism can only be made after prior approval of ECPGR (via the EURISCO Advisory Committee and the ECPGR Steering Committee).

References

Alercia A, Diulgheroff S, Mackay M. (2015) FAO/Bioversity Multi-Crop Passport Descriptors V.2.1. FAO, Bioversity, Rome (Italy), 11 pp.

Alercia A, López F, Marsella M, Cerutti AL (2021) Descriptors for Crop Wild Relatives conserved *in situ* (CWRI v.1). Rome, FAO on behalf of the International Treaty on Plant Genetic Resources for Food and Agriculture. https://doi.org/10.4060/cb3256en

Bioversity International, University of Birmingham. (2017). Crop wild relative checklist and inventory descriptors v.1. Bioversity International, Rome, Italy. pp. 21. https://www.bioversityinternational.org/e-library/publications/detail/crop-wild-relative-checklist-and-inventory-descriptors-v1/

EU (2022) EU Plant variety database https://ec.europa.eu/food/plant/plant_propagation_material/plant_variety_catalogues_databases/search/public/index.cfm, accessed date: April 2022.

FAO (2011) Second Global Plan of Action for Plant Genetic Resources for Food and Agriculture. Rome, 91 pp.

FAO (2022) FAOSTAT. Crops and livestock products: Crops primary https://www.fao.org/faostat/en/#data/QCL, accessed date: April 2022.

ITPGRFA (2022) Annex I – List of crops covered under the Multilateral System https://www.fao.org/plant-treaty/areas-of-work/the-multilateral-system/annex1/en/, accessed date: April 2022

Kell, S.P., Ford-Lloyd, B.V., Magos Brehm, J., Iriondo, J.M., Maxted, N. Broadening the Base, Narrowing the Task: Prioritizing Crop Wild Relative Taxa for Conservation Action. Crop Science (2017) 57:1042–1058.

Labokas J, Maxted N, Kell S, Magos Brehm J, Iriondo JM. (2018) Development of national crop wild relative conservation strategies in European countries. *Genetic Resources and Crop Evolution* 65: 1385–1403.

Landucci F, Panella L, Lucarini D, *et al.* (2014). A prioritized inventory of crop wild relatives and wild harvested plants of Italy. *Crop Science* 547: 1628–1644.

Magos Brehm J, Kell S, Thormann I, Gaisberger H, Dulloo ME, Maxted N. (2017). Interactive Toolkit for Crop Wild Relative Conservation Planning version 1.0. University of Birmingham, Birmingham, UK and Bioversity International, Rome, Italy. Available at: www.cropwildrelatives.org/conservation-toolkit/.

Maxted N, Avagyan A, Frese L, *et al.* (2015) ECPGR Concept for in situ conservation of crop wild relatives in Europe. Wild Species Conservation in Genetic Reserves Working Group. European Cooperative Programme for Plant Genetic Resources. Rome, Italy.

Maxted N, Ford-Lloyd BV, Hawkes JG (1997) Complementary conservation strategies. In: Maxted N, Ford-Lloyd BV, Hawkes JG (Eds) *Plant Genetic Conservation: The In Situ Approach.* Chapman & Hall, London, pp.20-55.

Maxted N, Magos Brehm J, Kell SP. (2013) Resource book for preparation of national conservation plans for crop wild relatives and landraces. Commission on Genetic Resources for Food and Agriculture. Food and Agriculture Organization of the United Nations, Rome, Italy. 457 pp. Available online: http://www.fao.org/agriculture/crops/thematic-sitemap/theme/seeds-pgr/resource-book/en/

Maxted N, Scholten M, Codd R, Ford-Lloyd B. (2007) Creation and use of a national inventory of crop wild relatives. *Biological Conservation* 140(1-2):142-159.

Nilsen LB, Maxted N, Mba C, Dulloo ME, Ghosh K, Magos Brehm J, Kell SP, Diulgheroff S, Noorani A, Furman B. (2017) Voluntary guidelines for the conservation and sustainable use of crop wild relatives and wild food plants. Food and Agriculture Organization of the United Nations, Rome, Italy. 93 pp. http://www.fao.org/3/a-i7788e.pdf

Phillips J, Kyratzis A, Christoudoulou C, Kell S, Maxted N. (2014) Development of a national crop wild relative conservation strategy for Cyprus. *Genetic Resources and Crop Evolution* 61: 817–827.

Rubio Teso ML, Torres Lamas E, Parra-Quijano M, de la Rosa L, Fajardo J, Iriondo JM. (2018) National inventory and prioritization of crop wild relatives in Spain. *Genetic Resources and Crop Evolution* 65.

Rubio Teso ML, Álvarez Muñiz C, Gaisberger H, Kell S, Lara-Romero C, Magos Brehm J, Maxted N,

Iriondo JM (2020) In situ plant genetic resources in Europe: crop wild relatives. Farmer's Pride. https://more.bham.ac.uk/farmerspride/wp-content/uploads/sites/19/2020/10/D1.2_In_situ_PGR_in_Europe_crop_wild_relatives.pdf

Smekalova TN. (2008) National Crop Wild Relative In Situ conservation Strategy for Russia In: Maxted N, Ford-Lloyd BV, Kell SP, Iriondo JM, Dulloo ME, Turok J, eds. *Crop Wild Relative Conservation and Use*. Oxfordshire, UK: CAB International, 143–151.

Taylor NG, Kell SP, Holubec V, Parra-Quijano M, Chobot K, Maxted N. (2017) A systematic conservation strategy for crop wild relatives in the Czech Republic. *Diversity and Distributions*: 1–15.

Thormann I, Kell, SP, Magos Brehm J, Dulloo E, Maxted N. (2017) CWR checklist and inventory data template v.1. https://doi.org/10.7910/DVN/B8YOQL, Harvard Dataverse, V4.

van Treuren R, Hoekstra R, van Hintum TJL (2017) Inventory and prioritization for the conservation of crop wild relatives in The Netherlands under climate change. *Biological Conservation* 216: 123-139

Weise, S. (2021). Memo - Extension of EURISCO for *in situ* data. https://www.ecpgr.cgiar.org/fileadmin/templates/ecpgr.org/upload/EURISCO/Advisory_Committee/Memo_in_situ_data_EURISCO_v0.3.pdf

Weise, S., Kreide, S. and Maxted, N. (2020) Concept for a possible extension of EURISCO for *in situ* crop wild relative and on-farm landrace data. https://more.bham.ac.uk/farmerspride/wp-content/uploads/sites/19/2021/09/D2.5_EURISCO_in_situ_extension_concept.pdf

Annex 1. Descriptors recommended for the generation of a National Inventory of *in situ* Crop Wild Relatives.

The descriptors used in the database structure of a National Inventory of *in situ* Crop Wild Relatives (CWR-NI) should, as far as possible, follow international standards. This proposal is based on the FAO *Descriptors for Crop Wild Relatives conserved* in situ (Alercia et al., 2021), the *Concept for a possible extension of EURISCO for* in situ *crop wild relative and on-farm landrace data* (Weise et al., 2020), the report 'Extension of EURISCO for *in situ* data" (Weise, 2021) and the current descriptors for passport data used in EURISCO for *ex situ* accessions.

The text describing the descriptors has been obtained from the FAO *Descriptors for Crop Wild Relatives conserved* in situ (Alercia *et al.*, 2021) when these descriptors occurred in this list. When the descriptor was not available in this list, the definition was obtained from the alternative source where the descriptor had been previously proposed. Slight changes were made and/or some notes were added after the definitions of the descriptors, when necessary, to make recommendations concerning their use and interpretation. The source of the descriptor is indicated in the last column of the lists. It can contain the values 'FAO' referring to Alercia *et al.* (2021), 'Farmer's Pride' referring to Weise *et al.* (2020); 'SADC' referring to Thormann *et al.* (2017); 'Bioversity' referring to Bioversity International and University of Birmingham (2017); 'MCPD' referring to FAO/Bioversity Multi-Crop Passport Descriptors V.2.1 (Alercia *et al.*, 2015), 'Plinian Core' referring to the Plinian Core working group on biological species information of Biodiversity Information Standards (TDWG) (https://www.tdwg.org/community/species/plinian-core/) or 'new' in case the descriptor was not listed in these sources.

The information to be collected by the National Inventory is organized at two levels. (1) The taxon level provides the information about the CWR taxa included in the CWR-NI, whereas (2) the population level contains the information describing each population and its location. Thus, the database structure can be set up in two tables, one at each level. The two tables can then be related through the TAXONID descriptor, unique to each taxon. In the process of selecting the descriptors, the principle of parsimony was applied, aiming at gathering all relevant information with the minimum number of descriptors and in the simplest way.

Ideally, most of the information at the taxon level should be downloadable from a central reference point (possibly maintained by EURISCO) to facilitate the process of creating a CWR-NI and to avoid the use of different taxonomic classification systems.

It should be stressed that these descriptors are merely suggestions, each CWR-NI can have the structure and contain the information considered important to the country compiling it. The only requirement is that at least the mandatory fields for upload to EURISCO can be extracted. These mandatory fields are indicated below.

Bioversity International and University of Birmingham (2017) and Thormann *et al.* (2017) provide a fuller account of descriptors that can be used at the taxon level.

Taxon level information

Description	Descriptor name	Source
Taxon ID Unique taxon identifier. Example: TaxonID 21356 **NOTE**: Use preferably existing identifiers, e.g. those used by GBIF (taxonKey) or by other national or international taxonomical backbone reference systems such as GRIN-Tax (TaxonID). If identifiers from different sources are used, specify the source by means of key-value pairs, e.g. GBIF:2706054 or GRIN:102504.	**TAXONID**	SADC
Family Taxon family, in Latin. Initial uppercase required.	**FAMILY**	Bioversity
Genus Genus name for taxon. **NOTE**: This descriptor should be considered mandatory.	**GENUS**	FAO
Species Species epithet portion of the scientific name. **NOTE**: This descriptor is considered mandatory in the FAO list.	**SPECIES**	FAO
Species authority Provide the authority for the species name. It is recommended to use the Catalogue of Life (https://www.catalogueoflife.org/).	**SPAUTHOR**	FAO
Subtaxon Subtaxon can be used to store any additional infraspecific epithet. The following abbreviations are allowed: 'subsp.' (for subspecies); 'var.' (for botanical variety); 'f.' (for form).	**SUBTAXA**	FAO
Subtaxon authority Subtaxon authority at the most detailed taxonomic level. **NOTE**: Subtaxon authority associated to the infraspecific epithet specified in SUBTAXA.	**SUBTAUTH**	FAO

Description	Descriptor name	Source
Use-Value What is the plant used for. CWR: crop wild relative; WFP: wild food plant; OTHER: other ecosystem services. Multiple entries are separated by a semicolon (;) without space. N**OTE:** This descriptor should be used when NIs contain records of wild plants other than crop wild relatives.	**USE_VALUE**	Plinian Core
Related crop species The scientific name(s) of the crops to which the taxon is related. Multiple entries are separated by a semicolon (;) without space. Example: *Brassica oleracea* L.;*Brassica juncea* (L.) Czern. **NOTE**: Depending on the interests of each country, alternative crop classification systems may be used. When possible, use the list of crop names used by GRIN.	**RELATEDCROP**	Bioversity
Genepool Genepool category assigned to the closest crop according to the GRIN-Global[2] database. GP1: primary; GP2: secondary; GP3: tertiary.	**GENEPOOL**	new
National category The Red List category[3] according to national criteria. CR: Critically endangered; EN: Endangered; VU: Vulnerable; NT: Near threatened; LC: Least concern; DD: Data deficient; NE: Not evaluated. **NOTE:** Use the most recent assessment.	**NATIONAL_CAT**	Bioversity
Legislative protection Informs whether the taxon is legally protected at the European, national or subnational levels. HD: Habitats Directive; NAT: National level; SUB: Subnational level.	**LEGSTATUS**	new

---

[2] https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomysearchcwr
[3] https://www.iucnredlist.org/resources/categories-and-criteria

Description	Descriptor name	Source
Distribution status The distribution status of the taxon within the geographic area of the checklist or inventory, indicating whether it is a National endemic, Regional endemic, Cosmopolitan, or Unknown. **NOTE**: 'Regional' is defined here as a geographic area comprising different countries rather than a sub-unit within a country.	**DIST_STATUS**	Bioversity

Accession/population-level information

These descriptors provide information relevant to the CWR populations.

Description	Descriptor	Source
Population identifier The identifier (sequential number or code) that the National Inventory uses to identify each population. Each distinct population should be given a population unique identifier. **NOTE:** This descriptor is mandatory for upload to EURISCO.	**POPID**	FAO
Persistent unique identifier Any persistent, unique identifier (preferably a DOI) assigned to the population so it can be unambiguously referenced at the global level and the information associated with it harvested through automated means. Report only one PUID for each population. The Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture (PGRFA) is facilitating the assignment of a persistent unique identifier (PUID), in the form of a DOI, to PGRFA at the accession level.[4] **NOTE**: This descriptor should be assigned only to those CWR populations that the National Focal Point considers as long-term available sources of germplasm (e.g. the population is being monitored and is potentially available under the terms of the MLS).	**PUID**	MCPD
Observation date [YYYYMMDD] The most recent date the population was observed. YYYY: year; MM: month; DD: day. Missing data (MM or DD) should be indicated with hyphens or '00' [double zero]. **NOTE**: This descriptor is considered mandatory in the FAO list.	**OBSDATE**	FAO
Biological status of the population The coding scheme proposed can be applied by using the general codes or the more specific codes. 100: Wild (110: Natural, 120: Semi-natural/wild, 130: Semi-natural/sown) 200: Weedy 999: Other (elaborate in REMARKS field) **NOTE**: Use 100 for wild populations, unless there is specific information indicating its semi-natural status.	**SAMPSTAT**	FAO

---

[4] https://www.fao.org/plant-treaty/areas-of-work/global-information-system/doi/en/

Description	Descriptor	Source
Managing institute code FAO WIEWS code of the institution responsible for the population (e.g. protected area authority, nature reserve manager, national park manager, private landowner, etc.). **NOTE**: This descriptor is new and mandatory for upload to EURISCO.	**MNGINSTCODE**	new
Managing institute name Name and short address of the institution responsible for the population, to be used if the previous file is empty. **NOTE**: the description was changed from the FAO list.	**MNGINSTNAME**	FAO
Liaison institute code FAO WIEWS code of the institution that can liaise between the organization managing the CWR population and the interested user to facilitate access to the material. **NOTE**: If this and the next field are empty, the managing organization can be approached for access.	**LIAISONCODE**	new
Liaison institute name Name, and brief address, of the institution that can liaise between the organization managing the CWR population and the interested user. **NOTE**: If this and the previous field are empty, the managing institute can be approached for access.	**LIAISONNAME**	new
Other identifiers associated with the population Any other identifiers known to exist in *ex situ* collections for this population. INSTCODE and identifier are separated by a colon without space. Pairs of INSTCODE and identifier are separated by a semicolon without space. Example: INSTCODE:identifier;INSTCODE:identifier;… When the institute is not known, the identifier should be preceded by a colon. Example::identifier;:identifier;… **NOTE**: This descriptor is the most efficient way of linking this record to an *ex situ* record of the same population. This way, there is no need to repeat the information concerning the *ex situ* accession.	**OTHERNUMB**	MCPD

Description	Descriptor	Source
Index Herbariorum code[5] Index Herbariorum code of the herbarium that holds a herbarium specimen of the population.	**HERBCODE**	FAO
Name of the herbarium holding herbarium specimens Name of the herbarium that holds herbarium specimens of the population. This descriptor should only be used if HERBCODE is empty.	**HERBNAME**	new
Herbarium specimen number This is the unique identifier for herbarium specimens collected and is assigned when a specimen is entered into the collection.	**SPECNUMB**	FAO
Conservation actions in place Indication whether conservation actions related to the population are in place. Use the IUCN classification scheme for conservation actions in place.[6] 0: No conservation actions 1: Monitoring and Planning 2: Land/Water Protection and Management 3: Species Management 4: Education and Legislation 99: Other (elaborate in REMARKS field) **NOTE**: There can be more than one answer. For example: 1; 3.	**CONSACTION**	FAO
MLS status of the material The status of the *in situ* accession of the CWR population with regards to the Multilateral System of Access and Benefit-sharing of the International Treaty, if available. 0: Not available under the MLS 1: Available under the MLS **NOTE**: The definition of this descriptor in Alercia et al. (2021) refers to the corresponding *ex situ* accession, but in the CWR-NI it should be applied to the *in situ* accession. The *ex situ* accession will have its own MLSSTAT descriptor in the genebank inventories.	**MLSSTAT**	FAO

---

[5] http://sweetgum.nybg.org/science/ih/
[6] Available from:
https://nc.iucnredlist.org/redlist/content/attachment_files/dec_2012_guidance_conservation_actions_in_place_classification_scheme.pdf (adapted).

Description	Descriptor	Source
Country of occurrence Country where the CWR population was observed or inventoried. Use the Three-letter ISO 3166-1 code of the country where the site is located. **NOTE**: This descriptor is considered mandatory in the FAO list.	**ORIGCTY**	FAO
Location of the occurrence site Location information below the country level that describes the site where the population sample was observed, inventoried or collected, preferably in English. This might include the distance in km and direction from the nearest town, village or map grid reference point, (e.g. 7km south of Curitiba in the state of Parana). **NOTE**: This descriptor may be used to indicate the extension of the population in the case of ubiquitous widespread taxa.	**OCCURSITE**	FAO
Latitude of occurrence site Latitude of the site expressed in decimal degrees. Positive values are north of the Equator; negative values are south of the Equator (e.g. -44.6975). **NOTE**: Latitude and longitude should be provided using the WGS84 datum, as this is the standard used by GBIF.	**DECLATITUDE**	FAO
Longitude of occurrence site Longitude of the site expressed in decimal degrees. Positive values are east of Greenwich Meridian; negative values are west of Greenwich Meridian (e.g. -120.9123). **NOTE**: Latitude and longitude should be provided using the WGS84 datum, as this is the standard used by GBIF.	**DECLONGITUDE**	FAO
Coordinate uncertainty [m] Uncertainty associated with the coordinates in metres. Leave the value empty if the uncertainty is unknown. **NOTE**: This descriptor may also be used to indicate the extension of the population in the case of ubiquitous widespread taxa.	**COORDUNCERT**	MCPD
Elevation of site Elevation of site expressed in metres above sea level. Negative values are allowed.	**ELEVATION**	FAO

Description	Descriptor	Source
Status of occurrence site Habitat of the occurrence site of the population(s). The coding scheme can be applied either by using the general codes or the more specific codes. Multiple values are separated by a semicolon without space. 10: Wild (11: Forest or woodland, 12: Shrubland, 13: Grassland, 14: Desert or tundra, 15: Aquatic habitat) 20: Farm or cultivated area (21: Field, 22: Orchard, 23: Backyard, kitchen or home garden, 24: Fallow land, 25: Pasture, 28: Park) 60: Weedy, disturbed or ruderal habitat (61: Roadside, 62: Field margin) 99: Other (elaborate in REMARKS field)	**POPSRC**	FAO
Site protection Indicate whether the site is protected under any legal or official protection.[7] 0: not protected; 1: strict nature reserve; 2: wilderness area; 3: national park; 4: natural monument or feature; 5: habitat/species management area; 6: protected landscape/seascape; 7: protected area with sustainable use of natural resources. 8: other effective conservation measures (OECM) **NOTE**: OECM includes other areas achieving effective *in situ* conservation outside of protected areas, such as on-farm, long-term ecological monitoring infrastructures, etc.	**SITEPROT**	FAO

---

[7] Follow IUCN Guidelines available at https://portals.iucn.org/library/node/30018

Description	Descriptor	Source
Links to associated information One or more URLs where further information about the CWR can be found. Multiple values are separated by a semicolon without space. **NOTE**: Only use one URL where the associated information remains updated.	**LINKS**	FAO
Remarks The Remarks field is used to add notes or to elaborate on descriptors with value 99 or 999 (= Other). Prefix remarks with the field name they refer to and a colon (:) without space (e.g. SITESTAT:riverside). Distinct remarks referring to different fields are separated by semicolons without space.	**REMARKS**	FAO

Annex 2. Descriptors for uploading passport data of *in situ* CWR to EURISCO

Introduction

This descriptor list provides the descriptors to be used by EURISCO when including records for *in situ* CWR accessions and describes the proposed data exchange format for uploading passport data from the National Inventories for *in situ* CWR to EURISCO.[8] The list of descriptors that can be uploaded to EURISCO is a subset of the list of descriptors recommended for the creation of a National Inventory of *in situ* Crop Wild Relatives in Annex 1, with the addition of the National Inventory Code (NICODE) and Type of germplasm storage (STORAGE).

General formatting rules

The general formatting rules that apply to the *ex situ* data also apply to the *in situ* data:

- If a field allows multiple values, these values should be separated by a semicolon (;) without space (e.g. Accession name: Symphony;Emma;Songino).
- A field for which no value is available should be left empty (e.g. Elevation). If data are exchanged in ASCII format, a field with a missing numeric value should be left empty. If data are exchanged in a database format, missing numeric values should be represented by generic NULL values.
- Dates are recorded as YYYYMMDD. If the month or day are missing, this should be indicated with hyphens or '00' [double zero]. If both (month and day) are missing, two double zeros are needed (e.g. 19750000; 197506--, 19750600).
- Country names: Three-letter ISO codes are used for countries. The ISO 3166-1 standard country or area codes are available online at: https://unstats.un.org/unsd/methodology/m49/. Note: The list of obsolete codes can be found at: http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3#Reserved_code_elements.
- For institutes, the codes from FAO WIEWS should be used. The current set of institute codes is available from the FAO WIEWS site (http://www.fao.org/wiews).
  - If new institute codes are required, they can be generated online by FAO National Focal Points (http://www.fao.org/agriculture/crops/thematic-sitemap/theme/seeds-pgr/gpa/national-focal-points/en/) or they can be requested from: WIEWS@fao.org.
  - In case no FAO WIEWS code of the institution responsible for, and/or organization that manages the CWR population is available and cannot be generated, the code ('DUMMY') can be used.

For institutes that no longer exist, or that were not assigned a FAO WIEWS institute code, please provide full details in the descriptor added for that purpose. In the case of the proposed upload format, this concerns INSTCODE and LIAISONCODE.

Descriptors

The descriptors in this list are a selection from those of the *ex situ* format for upload, with the addition of a few. In case the descriptor name or description is deviating from the *ex situ* upload format, this is indicated in the description.

In case the descriptor name or description is different from the corresponding descriptor in Annex 1 of this proposal, this is also indicated in the description.

Be aware that a significant departure from the proposed list of descriptors for CWR-NI (in Annex 1) is the concept of an *in situ* CWR population being an accession. As a result, the POPID in the CWR-NI

---

[8] Inclusion of the descriptors SITEPROT and CONSACTION in this list was not undisputed amongst the compilers of the list and needs to be (re-)considered.

becomes the ACCENUMB in EURISCO, and the managing institute code and name, MNGINSTCODE/MNGINSTNAME, the INSTCODE/INSTNAME, respectively.

The mandatory fields are, similarly to the *ex situ* upload format, NICODE, INSTCODE, ACCENUMB and GENUS. The combination of these fields has to be unique.

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)
National Inventory code **(NICODE)** Code identifying the National Inventory; the Three-letter ISO 3166-1 code of the country preparing the National Inventory. Exceptions are possible if agreed with EURISCO, such as NGB. Example: NLD	not in the CWR-NI
Persistent unique identifier **(PUID)** Any persistent, unique identifier assigned to the accession so it can be unambiguously referenced at the global level and the information associated with it harvested through automated means. Report one PUID for each accession. The Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture (PGRFA) is facilitating the assignment of a persistent unique identifier (PUID), in the form of a DOI, to PGRFA at the accession level (http://www.fao.org/plant-treaty/areas-of-work/global-information-system/doi/en/). Note: This descriptor should be assigned only to those CWR populations that the National Focal Point considers as long-term available sources of germplasm (e.g. the population is being monitored and potentially available under the terms of the MLS).	Persistent unique identifier (PUID)
Institute code **(INSTCODE)** FAO WIEWS code of the institution responsible for, and/or organization that manages the CWR population (e.g. protected area authority, nature reserve manager, national park manager, private landowner, etc.). The codes consist of the Three-letter ISO 3166 country code of the country where the institute is located plus a number (e.g. COL001). The current set of institute codes is available from http://www.fao.org/wiews. For those institutes not yet having a FAO Code, or for those with 'obsolete' codes, see '*General formatting rules*'. **NOTE**: This description deviates from the *ex situ* upload format	Managing Institute code (MNGINSTCODE)

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)
Institute name **(INSTNAME)** Name and short address of the organization that manages the CWR population (e.g. protected area authority, nature reserve manager, national park manager, private landowner, etc.). This descriptor should be used only if INSTCODE has the value ('DUMMY') because the FAO WIEWS code for this institute is not available. **NOTE**: This descriptor is new and did not occur in the EURISCO format yet.	Managing institute name (MNGINSTNAME)
Accession number **(ACCENUMB)** This is the unique identifier for CWR populations maintained *in situ*, and is assigned by the organization managing the population. **NOTE**: This description deviates from the *ex situ* upload format.	Population identifier (POPID)
Liaison institute code **(LIAISONCODE)** FAO WIEWS code of Name of the institution that can liaise between the organization managing the CWR population and the interested user. **NOTE**: This descriptor is new and did not occur in the EURISCO format yet.	Liaison institute code (LIAISONCODE)
Liaison institute name **(LIAISONNAME)** Name, and brief address, of the institution that can liaise between the organization managing the CWR population and the interested user. **NOTE**: This descriptor is new and did not occur in the EURISCO format yet.	Liaison institute name (LIAISONNAME)
Genus **(GENUS)** Genus name for taxon. Initial uppercase letter required.	Genus (GENUS)
Species **(SPECIES)** Specific epithet portion of the scientific name in lowercase letters. Only the following abbreviation is allowed: 'sp.'	Species (SPECIES)
Species authority **(SPAUTHOR)** Provide the authority for the species name.	Species authority (SPAUTHOR)
Subtaxon **(SUBTAXA)** Subtaxon can be used to store any additional taxonomic identifier. The following abbreviations are allowed: 'subsp.' (for subspecies); 'var.' (for variety); 'f.' (for form). **NOTE**: The description was slightly modified as cultivar groups cannot occur in CWR populations.	Subtaxon (SUBTAXA)

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)
Subtaxon authority **(SUBTAUTHOR)** Subtaxon authority at the most detailed taxonomic level.	Subtaxon authority (SUBTAUTHOR)
Observation date [YYYYMMDD] **(ACQDATE)** The most recent date the population was observed, where YYYY is the year, MM is the month and DD is the day. Missing data (MM or DD) should be indicated with hyphens or '00' [double zero]. **NOTE**: The name and description of this descriptor have been changed to apply to CWR *in situ*.	Observation date (OBSDATE)
Country of occurrence **(ORIGCTY)** Three-letter ISO 3166-1 code of the country where the CWR population was observed or inventoried. **NOTE**: The name and description of this descriptor have been changed to apply to CWR *in situ*.	Country of occurrence (ORIGCTY)
Location of occurrence site **(COLLSITE)** Location information below the country level where the population sample was observed. This might include the distance in km and direction from the nearest town, village or map grid reference point (e.g. 7km east of Wageningen in the province of Gelderland). **NOTE**: The name and description of this descriptor have been changed to apply to CWR *in situ*.	Location of occurrence site (OCCURSITE)
Latitude of occurrence site **(DECLATITUDE)** Latitude expressed in decimal degrees. Positive values are north of the Equator; negative values are south of the Equator (e.g. -44.6975). **NOTE**: The name of this descriptor has been changed to apply to CWR *in situ*. The accuracy of this information that is going to be disseminated may be adjusted as considered appropriate by each country.	Latitude of occurrence site (DECLATITUDE)
Longitude of occurrence site **(DECLONGITUDE)** Longitude expressed in decimal degrees. Positive values are east of the Greenwich Meridian; negative values are west of the Greenwich Meridian (e.g. +120.9123). **NOTE**: The name of this descriptor has been changed to apply to CWR *in situ*. The accuracy of this information that is going to be disseminated may be adjusted as considered appropriate by each country.	Longitude of occurrence site (DECLONGITUDE)

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)
Coordinate uncertainty [m] (COORDUNCERT) Uncertainty associated with the coordinates in metres. Leave the value empty if the uncertainty is unknown. Can also be used to indicate the size of the distribution area of the CWR. **NOTE**: The description of this descriptor has been changed to apply to CWR *in situ*. The coordinate uncertainty should be adjusted if the accuracy of the geographic coordinates is reduced.	Coordinate uncertainty [m] (COORDUNCERT)
Elevation of site [masl] (ELEVATION) Elevation of site expressed in metres above sea level. Negative values are allowed.	Elevation of site [masl] (ELEVATION)
Status of occurrence site **(POPSRC)** Habitat of the occurrence site of the population(s). The coding scheme can be applied either by using the general codes or the more specific codes. Multiple values are separated by a semicolon without space. 10: Wild (11: Forest or woodland, 12: Shrubland, 13: Grassland, 14: Desert or tundra, 15: Aquatic habitat) 20: Farm or cultivated area (21: Field, 22: Orchard, 23: Backyard, kitchen or home garden, 24: Fallow land, 25: Pasture, 28: Park) 60: Weedy, disturbed or ruderal habitat (61: Roadside, 62: Field margin) 99: Other (elaborate in REMARKS field) **NOTE:** This descriptor is new and did not occur in the EURISCO format yet.	**Status of occurrence site (POPSRC)**

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)
Site protection (SITEPROT) Indicate whether the site is protected under any legal or official protection 0: not protected 1: strict nature reserve 2: wilderness area 3: national park 4: natural monument or feature 5: habitat/species management area 6: protected landscape/seascape 7: protected area with sustainable use of natural resources. 8: other effective conservation measures (OECM) **NOTE:** This descriptor is new and did not occur in the EURISCO format yet.	Site protection (SITEPROT)
Conservation actions in place (CONSACTION) Indication whether conservation actions related to the population are in place. Use the IUCN classification scheme for conservation actions in place. 0: No conservation actions 1: Monitoring and Planning 2: Land/Water Protection and Management 3: Species Management 4: Education and Legislation 99: Other (elaborate in REMARKS field) **NOTE:** This descriptor is new and did not occur in the EURISCO format yet.	Conservation actions in place (CONSACTION)

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)

Biological status of accession   **(SAMPSTAT)**
 	
The coding scheme proposed can be used at two different levels of detail: either by using the general codes (in boldface) such as 100, 200, 300, 400, or by using the more specific codes such as 110, 120, etc.
 	
  100) Wild
 	
   110) Natural
 	
   120) Semi-natural/wild
 	
   130) Semi-natural/sown
 	
  200) Weedy
 	
  999) Other (Elaborate in REMARKS field)
 	
 NOTE: The description of this descriptor has changed (less allowed values).
	Biological status of accession (SAMPSTAT)

Other identifiers associated with the accession **(OTHERNUMB)** The identifier(s) of any sample of this population in an *ex situ* collection. Use the following format: INSTCODE:ACCENUMB;INSTCODE:identifier;… INSTCODE and identifier are separated by a colon without space. Pairs of INSTCODE and identifier are separated by a semicolon without space. When the institute is not known, the identifier should be preceded by a colon.	Other identifiers associated with the population (OTHERNUMB)
Type of germplasm storage **(STORAGE)** For *in situ* CWR populations this descriptor should always have the value 60. 60) *in situ* wild population **NOTE**: Status 60 is a new status! The description of this descriptor has been changed to apply to CWR *in situ*.	not in the CWR-NI

MLS status of the accession   **(MLSSTAT)**
 	
The status of the *in situ* accession of the CWR population with regards to the Multilateral System of Access and Benefit-Sharing (MLS) of the International Treaty, if available.
 	
   0   Not available under the MLS
 	
   1   Available under the MLS
 	
NOTE: The definition of this descriptor in Alercia et al. (2021) refers to the *ex situ* accession, but in the CWR-NI and in EURISCO is applied to the *in situ* accession.
	MLS status of the material (MLSSTAT)

Descriptor and description to be used in EURISCO	Corresponding descriptor from CWR *in situ* NIs (Annex 1)
Remarks (REMARKS) The remarks field is used to add notes or to elaborate on descriptors with value 99 or 999 (= Other). Prefix remarks with the field name they refer to and a colon (:) without space (e.g. COLLSRC:riverside). Distinct remarks referring to different fields are separated by semicolons without space.	Remarks (REMARKS)
Accession URL (ACCEURL) URL linking to additional data about the population. Example: http://gbis.ipk-gatersleben.de/gbis_i/detail.jsf?akzessionId=31805 **NOTE**: This description deviates from the *ex situ* upload format.	Links to associated information (LINKS)