European Cooperative Programme for Plant Genetic Resources (ECPGR)

Minutes of the Fourth EURISCO Advisory Committee meeting

20 September 2024, Tallinn, Estonia

Participants
Külli Annamaa, METK
Theo van Hintum, CGN (Chair)
Lorenzo Maggioni, ECPGR (*ex officio*)
Matija Obreza, Crop Trust
Ludmila Papoušková, CRI
Stephan Weise, IPK (EURISCO Coordinator)

Observers
Marco Marsella, ITPGRFA

Unable to attend
Anne-Françoise Adam-Blondon, INRAE
José María Iriondo Alegría, URJC
Kjell-Åke Lundblad, NordGen
Bettina Müller, Strube Research

The Agenda for this meeting is available online (here).
For a list of acronyms, see Annex 1

1. Welcome and introduction

The Chair of the EURISCO Advisory Committee (AC) welcomed all the participants, including the observers Marco Marsella, ITPGRFA and Külli Annamaa, Vice-Chair of the Documentation and Information Working Group (WG). The agenda was reviewed and adopted. The meeting took place back-to-back of the meeting of the Documentation and Information WG (18-19 September). Relevant discussion and conclusions made by the WG were reported here for consideration of the Advisory Committee.

2. Report on EURISCO activities since the previous AC meeting

S. Weise, EURISCO Coordinator, presented the progress, main activities and developments of EURISCO (PPT available here).

A few points raised by AC members were then discussed/clarified:

- The number of genera and species listed in the EURISCO statistics includes misspellings and synonyms. Therefore, the number of 6,733 genera is overestimated, but the difference compared to the true number is thought to be only a minimal percentage.
- The percentage of accessions including geographic coordinates (ca. 31%) seems to be low, but it should be kept in mind that only a fraction of the EURISCO accessions were collected from the fields. GPS data are important information for the users and deserve attention for improvement of the data coverage.

- Tools are being developed to test and categorize the quality of the data and the main gaps in a database. These tools may be used in the future to analyze data quality in EURISCO.
- The recommended frequency of data updates at country level is once a year. A script can be used for automatic updates (this is the case in the Netherlands).
- Data transfer from EURISCO to Genesys is done manually at least once per year (so far passport data only). It should be possible to automate the process. This requires action from the EURISCO side. It is proposed to look for a technical possibility to introduce this automatism allowing more frequent updates.
- Provision of phenotypic data to EURISCO is still rather infrequent, possibly due to the time-consuming data curation required to fill the existing templates. Simplified forms were developed and might improve the situation.
- There is no dedicated entry point for *in situ* data search and this area of the web interface needs to be developed.
- EURISCO can be advertised as the place where EC project data should be stored.
- An application was submitted for EURISCO to become an ELIXIR Core Data Resource, but this was rejected with no good explanation. It seems evident that EURISCO is the main European core PGR data resource and new attempts should be made to obtain this recognition.
- No objective measurement of the rate of satisfaction of the EURISCO users is available. However, an EURISCO evaluation in the form of a questionnaire was carried out in 2023 within the EURISCO community. The evaluation was rather positive, with some criticism specifically regarding the phenotypic data searches. The results of this evaluation should be circulated to the entire Advisory Committee. It is also recommended to organize direct interviews with users to verify their needs and criticism.
- Physical training remains very important, especially to induct newcomers into the community of focal points. The current 2-year frequency of the in-person training seems to remain the most suitable and the budget is available from ECPGR.

The AC was very positively impressed by the breadth and quality of the activities carried out and the progress made by the EURISCO team and thanked S. Weise for the clear and comprehensive report. The AC was also satisfied with the ongoing collaboration with IPK, which has always maintained a very supportive role.

3. EURISCO's role

The Chair invited the Committee to consider the possibility of an evolution of EURISCO towards becoming mainly a community of data specialists collecting and providing data that are transferred to Genesys, the global catalogue. The task of creating the best functionalities for data searches could be left to Genesys, which is endowed with better resources. This shift of focus would enable EURISCO to better dedicate to strengthening its network of data providers, but also to engage in the provision of other types of data, such as the *in situ* CWR. Phenotyping data gathering could also be reinforced, and linkages with on-farm documentation initiatives could be established.

The Committee agreed on the advantages of establishing stronger linkages with Genesys and the possibility of using its features, software and tools. At the same time, it was noted that EURISCO is a very successful brand and its existence is motivating the European governments and PGR national programmes to collaborate on a common initiative, ensuring a sense of belonging and regional pride. It would therefore be essential for EURISCO not to be diluted into Genesys and disappear as a physical database with its dedicated web interface.

It was clarified that a stronger integration with Genesys could simply mean that EURISCO would maintain its independent database and web interface, but a large part of the functionalities related to database queries by the users would be implemented through Genesys. In this way, EURISCO could maintain a relatively simple and intuitive web interface

based on Genesys APIs and thus duplication of work would be avoided. Features that are not supported by Genesys APIs could be complemented by own ones. This distinction would not be visible to users, as they would only see a single web interface. The users would maintain the possibility to query the EURISCO database looking for just European material, or the Genesys website for global searches.

M. Obreza confirmed the availability of the Genesys team to collaborate towards the above-mentioned stronger integration with EURISCO and explore all the possibilities and necessary technical mechanisms to strengthen such synergy and make it functional.

One aspect to consider is that there are elements in EURISCO that are not in Genesys, and vice versa. In particular, the *in situ* CWR data are at the moment out of the scope of the Crop Trust. It should however be noted that *in situ* data in EURISCO respond to the declared principle of including only those populations that could, in principle, be made available, often through the genebank's intermediation. The strong role of the genebanks in this initiative is something that the Crop Trust may consider as related to its overall mandate, also in the perspective of extending the model to other regions.

Regarding phenotypic data, Genesys facilitates data provision by installing the 'EmbeddedGenesys' feature into local genebank's websites. This system enables powerful and quick searches through the data. The weak element remains the difficulty of the data providers to manually curate the data. A combination of the simplified formats prepared by EURISCO and the searching features offered by Genesys could result in much more efficient data provision and data utilization.

There was consensus by the Committee to start exploring possibilities for creating stronger synergies between EURISCO and Genesys. It was recommended that the dialogue between the EURISCO and Genesys coordinators should intensify to identify and experiment all the appropriate technologies. The EURISCO web interface should be maintained and eventually enable the user to do all searches using the Genesys features within a EURISCO-branded environment. The ECPGR Steering Committee should be kept informed of all the implications related to the proposed steps and be given the opportunity to endorse the process.

The **relationship with GBIF** was discussed. EURISCO is currently not providing data to GBIF. A number of EURISCO data were provided in the past and these are now old and outdated. It was questioned whether there would be a benefit in automatically providing EURISCO data to GBIF, but this does not seem the case for most of the accessions that have no coordinates.

The approach to take would be to provide GBIF with an extracted dataset of the material that has coordinates. This was considered to be a task that Genesys could undertake rather than EURISCO so that all the relevant global data with geographic coordinates could be shared with GBIF.

4. What material to include

In terms of **geographic representation** of EURISCO, there was agreement that there is no intention to extend the catalogue to additional countries. In some cases, only part of the existing collections is reflected in EURISCO. For example, France is including in EURISCO only a small fraction of its collection and it would be desirable to bridge these gaps. On the other hand, it was noted that countries from other regions wishing to share their data, including phenotypic data, are welcome to directly provide these to Genesys.

Regarding the biological status of the material that could be included in EURISCO, the Chair wished to verify the suitability of the British ***Arabidopsis*** collection (ca. 600,000 accessions), which covers a high percentage of what is included in EURISCO, somehow skewing the statistics towards this particular research material. It was reiterated that the EURISCO Data Sharing Agreement refers to plant genetic resources occurring *ex situ*, *in situ*, on farm,

including cultivars, research material and genetic stocks. This guarantees maximum flexibility for the National Inventory Focal Points to decide what they would like to include in EURISCO. From the technical point of view, both S. Weise and M. Obreza confirmed that the *Arabidopsis* data set does not create a problem with the database or the web interface either in EURISCO or in Genesys. No further action was proposed.

The issue of inclusion in EURISCO of **SSD lines data** was thoroughly discussed. These lines are increasingly prepared for genotyping, and phenotyping is also frequently carried out on these same lines to facilitate GWAS. This is the case of the AGENT project and the EVA Networks, which have systematically created, genotyped and phenotyped SSD lines of various crops with the commitment to make the phenotypic data available through EURISCO. As the SSD lines are often not conserved by the genebanks or not maintained with the purpose of free availability, they are not included in EURISCO, or they remain as a subset of the accessions from which they are derived. It is, therefore, a challenge to include in EURISCO phenotypic data of accessions that either do not physically exist anymore or that are not included in EURISCO as distinct entities.

The proposed solution to this problem was discussed in the Documentation and Information WG meeting the day before and proposed as follows: for SSD lines, 'dummy' or 'virtual' entries to EURISCO are created, labelled as research material and historic material, and are linked with the original genebank accessions. In this way, their phenotypic data can be imported, but they would be excluded from passport searches.

This approach was considered feasible, but it would also be important to maintain a reference to the context in which the lines were described, also recording whether the physical SSD samples exist somewhere, in case someone wished to access them. For this, it will be useful that the SSD lines have a DOI assigned.

A solution was agreed in principle of creating virtual entries for SSD lines in order to import their phenotypic data in EURISCO. The EURISCO Coordinator, also in liaison with Genesys, received the task to refine the procedure, verify and test its technical requirements. It will remain necessary that the respective National Focal Points authorize the upload of the phenotypic data of SSD lines derived from national inventory accessions.

5. What data to include

Phenotypic data

During the Documentation and Information WG meeting in the previous days, S. Weise explained the limitations of the currently used template for the transfer of phenotypic data to EURISCO. He proposed a revised version, which should facilitate data compilation, including also more fine-grained metadata information. The Committee welcomed the adoption of the revised template.

The discussion focused on which incentives could be introduced to encourage the provision of phenotypic data to EURISCO. It was concluded that small financial incentives might be helpful, but the possibility of showing the value of the data would be the most effective. This requires building a critical mass of data so that they can be combined to generate illustrative pie charts and histograms showing the diversity of the traits in the collections.

DOIs

It was reiterated that DOIs are essential to extract all possible information related to each accession, including its history, derivation, duplication, past use, genotypic and other associated data. There was agreement that every accession in EURISCO should have a DOI assigned. The service offered by EURISCO, in liaison with the ITPGRFA Secretariat, to assign

DOIs to the accessions upon request of the countries, is working very well and with minimum effort. However, only few countries have decided to take advantage of this service or have autonomously obtained the DOIs for their collection.

The Committee agreed that every occasion should be used to encourage National Inventory Focal Points to assign DOIs to their accessions, including the statement from this report and reminders repeated to the Steering Committee via the ECPGR Bulletin or other means.

CWR descriptors

The Committee was informed that a meeting of the German-funded project on 'Extension of EURISCO for Crop Wild Relatives *in situ* data and preparation of pilot countries' data sets' took place recently (18-19 June 2024, Sadovo, Bulgaria). In this meeting, a revision was proposed for one of the EURISCO descriptors, namely the addition of one state to descriptor CONSACTION (Conservation action in place), to enable adding information about populations managed as part of a genetic reserve. This will be state *5: Managed as part of a CWR genetic reserve*. A new version of the *in situ* CWR EURISCO descriptors will be published on the EURISCO website.

Principles and procedures to assign DOIs to the CWR *in situ* populations were also discussed at the same meeting. Guidelines to inform data providers on suggested options and procedures will be drafted by the EURISCO Coordinator and then publicized.

Crop names

The opportunity to improve the search functionality by crop names was discussed. The pragmatic approach of using the GRIN standard for crop names is currently implemented in EURISCO and this enables mapping the majority of the accessions to an established crop name. It is acknowledged that this standard has many limitations, since different users use different crop names or attribute different sets of taxa to a given crop name. For example, the crop names 'Brusselsprouts' is currently linked to all *Brassica oleracea* accessions and not only to the correct subtaxon. However, a perfect solution is not available at the moment and no specific action was suggested by the Committee.

6. Ordering system

The usefulness of introducing an ordering system to enable the user to directly create an 'ordering basket' while browsing EURISCO was reiterated. A possible procedure had been proposed in the previous days during the Documentation and Information WG meeting. The MCPDs do not include a descriptor related to the availability of the material. This was never introduced upon consideration that maintaining it up to date would probably be unrealistic for many national inventories. The risk to place many orders that would not be successfully processed would be very high. On the other hand, a descriptor which indirectly indicates 'availability' does exist and it is the AEGIS descriptor. In fact, all the accessions that are part of AEGIS should by default be available under SMTA. It was therefore suggested to start implementing the ordering system for the AEGIS accessions, which could be associated to an ordering button.

EURISCO should act as a brokering facilitator enabling the user to order at the same time from several genebanks located in different countries and organized with different procedures for processing requests. Therefore, EURISCO should simply interconnect the requestor with the recipient genebanks, without storing any data related to the requestor or the ordered accessions. This could be possible by providing a specific API to the genebank information systems and thus redirecting specific orders, which would be received by the genebank in exactly the same way as any other order submitted from outside of EURISCO, and enter the regular ordering system (possibly including a click-wrap SMTA approval etc.)

For the case of genebanks that are not organized with an information system that could install the API, it was proposed that EURISCO generates an email that would be sent to the suitable

address where the genebank wishes to receive this type of orders. The latter mechanism would have the drawback that EURISCO should maintain an updated list of emails of genebanks and this requires manual maintenance and is prone to errors. Also in this case the genebank would need to implement a specification in its own protocol.

It was noted that so far Genesys only uses the email part of the mechanism, since most genebanks do not have a suitable information system that can handle orders. In the Genesys mechanism there is also a PDF that is generated with the SMTA, but only after negotiation of what is actually being sent (not including what has been requested). Genesys can be a mechanism for click-wrap acceptance of the SMTA.

The Committee recommended the EURISCO team start implementing the ordering system with IPK and CGN, and seek the interest of other genebanks to join (i.e. Nordgen, others).

7. Operation of the Advisory Committee

The ECPGR Secretary reminded the procedures of operation of the Committee, which are described in the ToRs of the EURISCO Advisory Committee, revised by the Steering Committee in June 2023. Th. van Hintum expressed the wish to step down from the position of Chair of the Advisory Committee, which he has been maintaining since 2003. Names of possible candidates as new members of the Committee and/or new Chair were proposed, and these people will be contacted by the Chair to verify their availability. The subsequent step would be to propose their nomination to the ECPGR Executive Committee.

The meeting was closed with an agreement to try to hold the next meeting in 2025.

Annex 1 - List of Acronyms

AEGIS, A European Genebank Integrated System
CGN, Center for Genetic Resources, The Netherlands
CRI, Crop Research Institute, Czech Republic
CWR, Crop Wild Relatives
DOI, Digital Object Identifier
GBIF, Global Biodiversity Information Facility
GRIN, Germplasm Resources Information Network of the US
GWAS, Genome-Wide Association Studies
INRAE, National Research Institute for Agriculture, Food and the Environment, France
IPK, Leibniz Institute of Plant Genetics and Crop Plant Research, Germany
ITPGRFA, International Treaty on Plant Genetic Resources for Food and Agriculture
MCPDs, Multi-Crop Passport Descriptors
METK, Centre of Estonian Rural Research and Knowledge
SMTA, Standard Material Transfer Agreement
SSD, Single Seed Descent
URJC, Universidad Rey Juan Carlos, Spain