Report of the workshop on 'Genebank Metrics' held on 24–25 February 2026 in Prague, Czech Republic

Compiled by Lorenzo Maggioni, Stephan Weise, Filippo Guzzon and Theo van Hintum

Organization

On 24–25 February 2023, a workshop dedicated to the topic of 'Genebank Metrics' was convened at the Hotel Meritum in Prague. The event was organized by ECPGR in collaboration with the Czech Agrifood Research Center (CARC) within the framework of the ECPGR AEGIS Plus project and received financial support from the German Federal Ministry of Agriculture, Food and Regional Identity (BMLEH). Participation comprised 11 trainees, 3 trainers, 2 supporting participants, 4 observers, and 1 technical support staff member (see the list of participants in Annex 1).

The meeting had the following starting objective: "As a next step towards broad adoption of universal genebank metrics, the metrics will be presented, discussed and, where possible, improved by a group of documentation specialists with hands-on experience in genebank documentation. Where possible, steps towards adoption will be made."

Genebank metrics refer to quantitative indicators that characterize the status of a genebank with respect to the size and composition of its collections, the completeness of associated data and documentation, conservation status, and the availability and distribution of material[1]. The established framework encompasses 10 mandatory and 38 additional metrics, including, for example, the total number of accessions, the number of accessions requiring viability testing, and the number of regenerated seed lots. Ideally, such metrics would be generated automatically through integrated data management systems. In practice, however, many genebanks lack documentation structured in a manner that permits straightforward calculation. Moreover, the effective application of these metrics presupposes well-defined and consistently implemented standard operating procedures, enabling, inter alia, the identification of accessions requiring regeneration or viability assessment.

The meeting commenced with welcoming remarks and participants' introductions. Morning sessions were devoted to the presentation and critical discussion of the conceptual framework of genebank metrics, as well as of each individual indicator. These exchanges facilitated the

[1] See the seminal publication: van Hintum T, Bartha B, Niggli C, Avagyan A, Vogl S, Achathaler L, Holubec V, Papouskova L, Ferrari F, Rossi G, Simon A, Horváth jr. L, Kowalik R, Boczkowka M, Weise S, Opperman M, van Zonneveld M, Obreza M, Wijnker E, Chayut N, Chairi F, Axelsson J, Steffensen LL, Guzzon F (2025) A proposal for genebank metrics to enhance collection management. *Plant Genetic Resources: Characterization and Utilization* First View. DOI: 10.1017/S147926212510021X

identification of methodological challenges, ambiguities in definitions, gaps in coverage, and practical obstacles to calculation. Afternoon sessions adopted a hackathon format, during which trainees worked directly with the documentation of their respective genebanks, either brought to the meeting or accessed remotely. This collaborative setting encouraged peer exchange, mutual learning, and the sharing of practical experiences. In parallel, afternoon sessions were devoted to discussions among the remaining participants on strategies to further enhance the adoption and robustness of the metrics. Particular attention was given to aligning them with existing reporting systems and to integrating them into plant genetic resources (PGR) documentation platforms, such as GRIN-Global Community Edition and EURISCO. The final session concluded with a comprehensive recap, reflections from participants, and expressions of appreciation to those who had contributed to the successful organization of the meeting.

Outcomes

The following introductory presentations were given:

- Genebank metrics – a genebank in a few numbers (Theo van Hintum, CGN)
- Application of the genebank metrics to the German Federal *Ex Situ* Genebank for Agricultural and Horticultural Crops (Stephan Weise, IPK)
- Genebank Metrics – Insights in the CGN collection (Laura Reiniers, CGN)
- Metrics Description (Filippo Guzzon, MSB)
- Use of the genebank metrics in aggregator information systems like EURISCO and Genesys (Stephan Weise, IPK)

(The presentations and more can be found on the AEGIS Plus pages on the AEGIS website: https://www.ecpgr.org/aegis/projects/aegis-plus)

These presentations prompted questions for clarification, discussion and proposals for the further development of the metrics. The discussion is summarized below, structured according to the main topics addressed:

1. Size and composition of the collection

- Some genebanks conserve material that are not considered accessions under long-term conservation, for various reasons, and that may enter the collection in the future. It was clarified that the metrics apply only to material that is fully part of the collection under long-term conservation, conserved for future generations of users. A clear line must be drawn to define what is included in this collection and what is not.
- For example, pre-breeding material that is only temporarily conserved and not characterized should not be considered part of the collection (and thus not included in the metrics). It is the responsibility of curators to determine which important lines –

those that are characterized and maintained over the long term – should enter the collection.
- While it is acknowledged that the definition of a landrace can be controversial, the metrics will not attempt to resolve this issue. Whatever a genebank considers to be a landrace will be counted as such.
- It was clarified what constitutes a 'removed accession': an accession that has been removed from the collection because it no longer exists (e.g. it died or seeds were depleted) or was clearly redundant (duplicated in the genebank or elsewhere). All associated data are retained, but the accession is no longer considered part of the collection. This corresponds to changing an accession's status from Active to Historical.

2. Documentation
- The average number of phenotypic data points can be calculated either across the entire collection or only for accessions that contain such data.
- The difficulty of compiling the metric on the number of accessions with genomic data was noted. It was acknowledged that this would become easier if the use of DOIs were more systematically implemented, and if those DOIs were referenced in the metadata of genomic datasets.
- A question was raised about whether there should be separate metrics for characterization and evaluation.
    - Generally, characterization refers to the measurement of traits that are not influenced by the environment – i.e. genetically stable and highly heritable. Traits that are less heritable are the subject of evaluation, requiring specific experimental design (e.g. replicated and multi-year, multi-location trials) and/or equipment to be measured. However, the distinction is not always clear-cut.
    - Some participants suggested that separating characterization and evaluation in the metrics would be useful to highlight the most valuable traits under 'evaluation', which explain how accessions perform under different environmental conditions.
    - Others argued that a differentiation based on trait importance would be difficult to define and standardize, and therefore it might be better not to attempt to distinguish between characterization and evaluation.
    - The number of observations of distinct traits per accession was mentioned as an alternative to the current metric, to filter out repetitive characterization traits observed during each regeneration.
- Currently, it is not possible to distinguish characterization from evaluation in EURISCO; standardization would be required to do so, if at all feasible.
- An ongoing issue is how to link genomic data to a given accession when those data originate from a subset of that accession (often a single plant). A solution to this problem

is being considered within EURISCO, which could also facilitate the calculation of this metric.

- Reporting on the number of accessions with genotypic data is currently difficult because researchers often do not report genotyping results back to the genebanks. If they did, it would be possible to link these data to passport information.
- Another challenge is defining what qualifies as genotypic data. For example, SSRs may not be considered sufficiently advanced. Alternatively, all marker types (SSRs, AFLPs, WGS, etc.) could be included but assigned different value levels (similar to the Passport Data Completeness Index (PDCI)).
- The usefulness of providing genotypic data to users of PGR accessions was also questioned.
- It was concluded that while there are options for improving the metrics related to phenotypic and genotypic information, this may not be the current priority.
- Passport data provided by genebanks evolve and improve over time. It might be useful to record the date of the last update as an indicator of data quality.

3. **Conservation**
   - A question was raised about the possibility of distinguishing between monitoring for medium-term and long-term storage. Many genebanks only monitor viability in their base collection.
   - Other genebanks test viability in their active collection, assuming that viability in the base samples should only be equal or better. This metric was originally intended to monitor viability measurements of base samples, as it is essential that base samples remain viable over the long term. If only the base sample is tested and the active sample is stored under different conditions, we may not know whether the active seed lot is deteriorating, though this would affect the quality of distributed material, not the long-term security of the sample. That said, if standard operating procedures (SOPs) specify that testing is done on active samples as a proxy for the base collection, this is acceptable. There are also genebanks that do not distinguish between active and base collections in terms of storage conditions.
   - A discussion developed around whether it is worthwhile to aim for the standard temperature of -20°C, given the energy costs, compared to storage at 5°C and its impact on seed longevity. Numerous studies address this issue and generally recommend -20°C. The possibility of using stand-alone freezers was also mentioned as a way to avoid the discomfort of entering large -20°C rooms to access material.
   - A question was raised as to whether initial viability testing is included in the number of viability tests considered in the metrics. The metrics might aim to distinguish between initial and subsequent viability monitoring tests.

- In many institutions, SOPs for viability testing are not standardized. Where no initial viability testing has been conducted, it is difficult to plan monitoring schedules. A protocol is needed to determine when viability should be monitored. While such a protocol may be more or less acceptable to external observers, the essential first step is to define one.
- The possibility of developing metrics to provide information that helps identify gaps in the collection and thereby guide future collecting missions was suggested. Some considered this not feasible, nor desirable, given that no single genebank tries to cover the entire diversity of a crop. However, at the aggregate level (EURISCO), this could be an interesting idea.
- It would be desirable to link data on viability test results obtained through the metrics with research on seed longevity.

4. Availability

- A question was raised about the factors affecting availability. Availability may be restricted if: (1) there is an insufficient quantity of seeds, (2) phytosanitary restrictions apply, or (3) legal issues are involved. Germination percentage is not always considered a restricting factor, as samples can be distributed even when germination falls below the threshold.
- An accession is to be flagged as 'not available' for phytosanitary reasons if national rules prevent the material from being distributed in the country where the genebank is located. Other countries have other rules, and anticipating all countries in the world is not feasible. So, the fact that phytosanitary rules prevent distribution to certain countries does not make the accession unavailable.
- It was asked whether the reasons for unavailability could be more clearly specified in the metrics.
- Some advocated that genebanks should raise the quality threshold for availability by ensuring six criteria are met: (1) sufficient seed quantity, (2) phytosanitary health, (3) legal compliance, (4) sufficiently high-quality data, (5) true-to-type status, and (6) a high germination rate.
- Differing opinions were expressed regarding the need to add these requirements for availability, with some viewing the step as premature and others as unnecessary. In practice, true-to-type verification is often very difficult to carry out, and users typically request material based on limited data, regardless of germination rate. These users are generally very happy to receive the requested material, even if the viability is below the threshold for regeneration.
- It was suggested that unavailability due to insufficient quantity should be distinguished from unavailability due to lack of quality of the seed.

5. **Distribution**
   - Considering that the Standard Material Transfer Agreement (SMTA) requires distinguishing the type of use made of PGR, it would be useful to introduce metrics in this direction, possibly in addition to those on the type of recipient.
   - Many participants confirmed that they record information on the type of use, including categories such as research, breeding, education and others. Hobby growing, repatriation, exchange with other genebanks, direct cultivation, regeneration, and germination are various use categories that genebanks currently track.
   - Despite conflicting opinions about adding metrics for the 'type of use,' it seemed important to capture this information, given the importance of monitoring uses that are permitted under the terms of the SMTA. 'Repatriation' and 'direct use' were also highlighted as potential categories to include.
   - It was suggested to add the number of unique, distinct accessions distributed, in addition to the number of samples distributed, as the latter does not convey the level of diversity being mobilized. This idea was supported, though it was noted that this metric (number of accessions distributed) cannot be summed over periods, the same way as the number of samples, since the same accessions may be distributed in different years.

6. **General concepts**
   - The acronyms used to define the metrics are rather cryptic and do not immediately convey the corresponding metric's title. They could be improved.
   - Considering that PDCI information can evolve and improve over time, obtaining time series of the metrics would be useful to measure the rate of data improvement.
   - Regarding the opportunity to standardize the length of the reporting period (for example, five years) for comparison across genebanks, it was concluded that it is better not to standardize, since each genebank has different reporting obligations. For comparisons across genebanks, a preferential periodicity of five years may still be suggested.
   - The list of mandatory metrics requires stricter definitions, with more precise wording and clearer conceptual boundaries.

7. **Expanding the metrics to other conservation methodologies: *In vitro*, on-farm, cryo- and field conservation**
   - Doubts were expressed about the need and usefulness of developing metrics for on-farm activities, while it was felt that other conservation strategies would deserve attention.

- It was observed that any new metrics would need to be created based on the specificities of each conservation method and that it would not be possible to simply add them to the existing list designed for seed conservation.
- The upcoming peer reviews of a trio of field genebanks, as part of the AEGIS Plus project, could provide a suitable framework for beginning discussions on the development of metrics for field, *in vitro*, and cryoconservation.

The other discussions, some in parallel to the hackathon, yielded important insights, such as:

- The establishment of time series for the defined metrics was considered essential for capturing the developmental trajectory of a genebank. Monitoring changes in metric values over time, i.e. their respective deltas, was therefore recognized as an important indicator in its own right. It was recommended that these data be collected and documented at regular intervals to enable systematic analysis of trends and institutional progress.
- Provided that standard operating procedures are clearly defined and the relevant data are systematically integrated into the database, the implementation of genebank-specific scripts within GRIN-Global Community Edition is entirely feasible. Such scripts would enable the automated calculation of metrics with minimal manual intervention, effectively allowing their generation at the push of a button.
- The metrics should be conceived and communicated primarily as instruments for monitoring and improving internal processes within a genebank. If elevated to the status of Key Performance Indicators (KPIs), they risk being misapplied or distorted, potentially promoting artificial inflation of figures, such as unnecessary distributions or inappropriate data entries, merely to improve numerical outcomes. Such developments would undermine both data integrity and operational quality and should be avoided.
- The collection-related KPIs of the CGIAR Genebank Quality Management System can be readily mapped onto the current set of genebank metrics. However, these correspond to only 3 of the 14 KPIs defined within that system; the remaining indicators address broader institutional dimensions, including human resource management and risk management.
- Several essential aspects of genebank management, including some directly related to collections, are inherently difficult to quantify and therefore not easily translated into metrics. An example is the earlier-mentioned degree to which actions are taken to guarantee that accessions remain true to type. While such elements may resist numerical representation, they can often be effectively safeguarded through robust and consistently applied standard operating procedures.

The hackathon component of the meeting yielded, in addition to valuable exchanges of insight among participants, a preliminary matrix of genebank metrics for the 11 participating institutions. An overview is presented in Annex 2 to this report. It should be emphasized that these data have not yet undergone formal verification and are included primarily to illustrate the analytical value of the metrics in assessing the status of genebanks.

Evaluation

During the final hour of the meeting, both the event itself and the concept of genebank metrics were evaluated collectively by the participants, resulting in an overall highly positive assessment. The following observations were made:

- The metrics contribute not only to the standardization of terminology within the genebank community but also enhance clarity and transparency, thereby supporting improvements in operational quality.
- A genebank that systematically applies the metrics demonstrates effective process control and institutional awareness, even where the resulting values indicate areas requiring further development.
- The process of calculating the metrics reveals the necessity of improved data management. The existence of dispersed datasets, such as spreadsheets stored across multiple computers, was widely recognized as undesirable and as a significant impediment to reliable metric calculation.
- The possibility of comparing metrics over the years, and potentially aggregating them across crops, offers a valuable means of documenting and communicating the development of a genebank over time.

The participants discussed the possibility of writing a follow-up paper on the metrics including: (1) discussed amendments on the already published metrics list, (2) compare the metrics with existing reporting tools (e.g. in the framework of FAO State of the World's Plant Genetic Resources, Crop Trust KPIs, National PGR quantitative reports), to promote the use of genebank metrics for PGR reporting worldwide, and (3) quantitative analyses on the metrics from different genebanks (e.g. participants' institutions and additional key national and international genebanks). The participants expressed their interest in participating in and contributing to this possible follow-up publication.

All participants expressed satisfaction with their participation and conveyed appreciation for the local organization, the logistical coordination provided by ECPGR, and the open and constructive atmosphere that characterizes the European genebank community.

ANNEX 1

List of Participants

Trainers	
Theo van Hintum	Centre for Genetic Resources, the Netherlands (CGN) - WUR, the Netherlands
Filippo Guzzon	Millennium Seed Bank, Royal Botanic Gardens, Kew, UK
Stephan Weise	Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany
Trainee	
Aridon Ramaj	Institute of Plant Genetic Research of Albania, Albania
Ludmila Papoušková	Czech Agrifood Research Center (CARC), Czech Republic
Domenico De Paola	Institute of Biosciences and BioResources (IBBR CNR), Italy
Francesco Ferrari	University of Pavia, Italy
Graziano Rossi	University of Pavia, Italy
Laura Reiniers	Centre for Genetic Resources, the Netherlands (CGN) – WUR, the Netherlands
Renata Kowalik	Plant Breeding i Acclimatization Institute - National Research Institute (IHAR-PIB), Poland
Silvia Alves	Instituto Nacional de Investigação Agrária e Veterinária (INIAV), Portugal
Iveta Čičová	Research Institute of Plant Production Piešťany (NPPC-VÚRV), Slovakia
Lovro Sinkovič	Agricultural Institute of Slovenia, Slovenia
Beate Schierscher-Viret	Agroscope, Switzerland
Supporting	
Janny van Beem	Global Crop Diversity Trust (Crop Trust)
Matija Obreza	Global Crop Diversity Trust (Crop Trust)
Observers	
Ryoko Machida-Hirano	National Agriculture and Food Research Organization, Japan
Lorenzo Maggioni	ECPGR
Dagmar Janovská	Czech Agrifood Research Center (CARC), Czech Republic
Vojtěch Holubec	Czech Agrifood Research Center (CARC), Czech Republic
Technical support	
Tomáš Čermák	Czech Agrifood Research Center (CARC), Czech Republic

ANNEX 2

Impression of the Genebank Metrics generated by the participants

The participants were asked to ensure that they had access to the documentation of their genebank during the meeting (either by bringing the data or by creating remote access). During the hackathon, they tried to implement the metrics in their systems. In some cases, this proved not possible on-site. If the attempts were successful, these results could not be verified. So the results below should be seen as a first impression of the potential of the metrics, not a current status quo of the genebanks.

Since the reporting periods, the definitions and protocols vary, so not all data can be compared across genebanks.

The totals give the totals of available data, the averages are weighted over the number of accessions in the collections that have data available.

			COMPOSITION														DOCUMENTATION			
Genebank	reporting period (in yrs)	**# accessions**	% wild/weedy	% landrace	% cultivar	% research mat.	% other/unkn.	% domestic origin	% from Asia	% from Africa	% from N. America	% from S. America	% from Europe	% from Oceania	% other/unkn.	time in coll.	PDCI	# pheno/acc	% genotypes	% DOIs
IPK, Germany	6.11	**151140**	15%	38%	11%	28%	9%	15%	21%	8%	3%	6%	44%	0%	17%	39.4	6.45	42.43	na	100%
IPGR, Albania	na	**4121**	na	na	na	na	100%	100%	0%	0%	0%	0%	0%	0%	100%	na	na	na	na	na
CRAC, Czech Republic	5.00	**47224**	11%	7%	15%	60%	8%	19%	8%	0%	10%	1%	72%	1%	7%	34.3	6.9	26.7	8%	5%
IBBR CNR, Italy	na	**38029**	0%	6%	0%	0%	94%	22%	20%	18%	10%	1%	40%	0%	11%	39.6	4.2	na	na	100%
Univ. of Pavia, Italy	na	**4501**	33%	57%	0%	0%	10%	90%	0%	2%	0%	0%	4%	0%	94%	8.3	na	na	na	na
CGN, The Netherlands	5.00	**24023**	16%	30%	9%	35%	10%	16%	26%	7%	6%	7%	43%	1%	11%	31.1	7.92	18.9	10%	100%
IHAR-PIB, Poland	5.00	**81322**	13%	13%	24%	15%	35%	37%	4%	2%	5%	1%	63%	1%	25%	15.8	5.7	14.6	na	0%
INIAV, Portugal	na	**35750**	13%	44%	36%	1%	5%	49%	na	na	na	na	na	na	100%	na	na	na	na	0%
NPPC-VÚRV, Slovakia	5.00	**1397**	1%	8%	33%	122%	-64%	23%	0%	0%	5%	1%	31%	0%	62%	na	na	na	na	na
AIS, Slovenia	na	**6289**	32%	42%	13%	4%	11%	74%	0%	0%	2%	0%	91%	0%	7%	na	na	na	na	na
Agroscope, Swiss	1.00	**17338**	1%	na	na	na	99%	na	na	na	na	na	na	na	100%	na	na	na	na	na
Total		**411134**	12%	26%	17%	24%	20%	26%	15%	7%	5%	3%	52%	1%	17%	32.2				

Genebank	reporting period (in yrs)	**# accessions**	CONSERVATION age base (yrs)	base germination in rep. period		regeneration in rep. period		base germination needs		base regeneration needs		low germonation	low seed amount	safety duplicated		collegue genebank	Svalbard
IPK, Germany	6.11	**151140**	23.44	88341	58%	24936	16%	na	na	na	na	6856	na	70154	46%	274	69947
IPGR, Albania	na	**4121**	na	na	na	1900	46%	na	na	na	na	na	na	1900	46%	na	na
CRAC, Czech Republic	5.00	**47224**	23.3	12372	26%	5889	12%	1542	3%	3058	6%	5772	118	4340	9%	3396	2117
IBBR CNR, Italy	na	**38029**	37.1	3153	8%	2193	6%	na	na	na	na	na	8030	na	na	na	592
Univ. of Pavia, Italy	na	**4501**	8.3	1359	30%	497	11%	3329	74%	2708	60%	562	2708	48	1%	0	9
CGN, The Netherlands	5.00	**24023**	29.6	10210	43%	1405	6%	2086	9%	844	4%	667	214	23710	99%	23429	20542
IHAR-PIB, Poland	5.00	**81322**	7.3	21222	26%	8350	10%	35813	44%	2441	3%	13406	na	14257	18%	996	13261
INIAV, Portugal	na	**35750**	na	7695	22%	na	na	na	na	na	na	3792	na	na	na	2936	1053
NPPC-VÚRV, Slovakia	5.00	**1397**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
AIS, Slovenia	na	**6289**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
Agroscope, Swiss	1.00	**17338**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
Total		**411134**															

Genebank	reporting period (in yrs)	**# accessions**	AVAILABILITY readily available		legal reasons	phyto reasons	seed reasons	DISTRIBUTION # ever distributed	# ever distr. out	# distr./yr in coll	# out rep. per.	#out domestic		% to commercial	% to public	% to NGO's	% to persons
IPK, Germany	6.11	**151140**	139825	93%	0	na	na	1200000	1200000	na	127758	55581	44%	13%	61%	2%	7%
IPGR, Albania	na	**4121**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CRAC, Czech Republic	5.00	**47224**	46463	98%	0	1951	349	25458	17633	0.02	9869	6198	63%	na	na	na	na
IBBR CNR, Italy	na	**38029**	26924	71%	na	na	2193	na	na	na	1686	986	58%	1%	96%	3%	0%
Univ. of Pavia, Italy	na	**4501**	38	1%	0	0	2708	397	0	0.00	359	0	0%	24%	64%	0%	4%
CGN, The Netherlands	5.00	**24023**	23973	100%	0	0	50	198867	151468	0.22	16997	6064	36%	49%	45%	6%	1%
IHAR-PIB, Poland	5.00	**81322**	42158	52%	17342	na	na	24735	16628	0.02	9271	7223	78%	5%	21%	3%	36%
INIAV, Portugal	na	**35750**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
NPPC-VÚRV, Slovakia	5.00	**1397**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
AIS, Slovenia	na	**6289**	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
Agroscope, Swiss	1.00	**17338**	na	na	na	na	na	na	na	na	675	na	na	na	na	na	na
Total		**411134**															