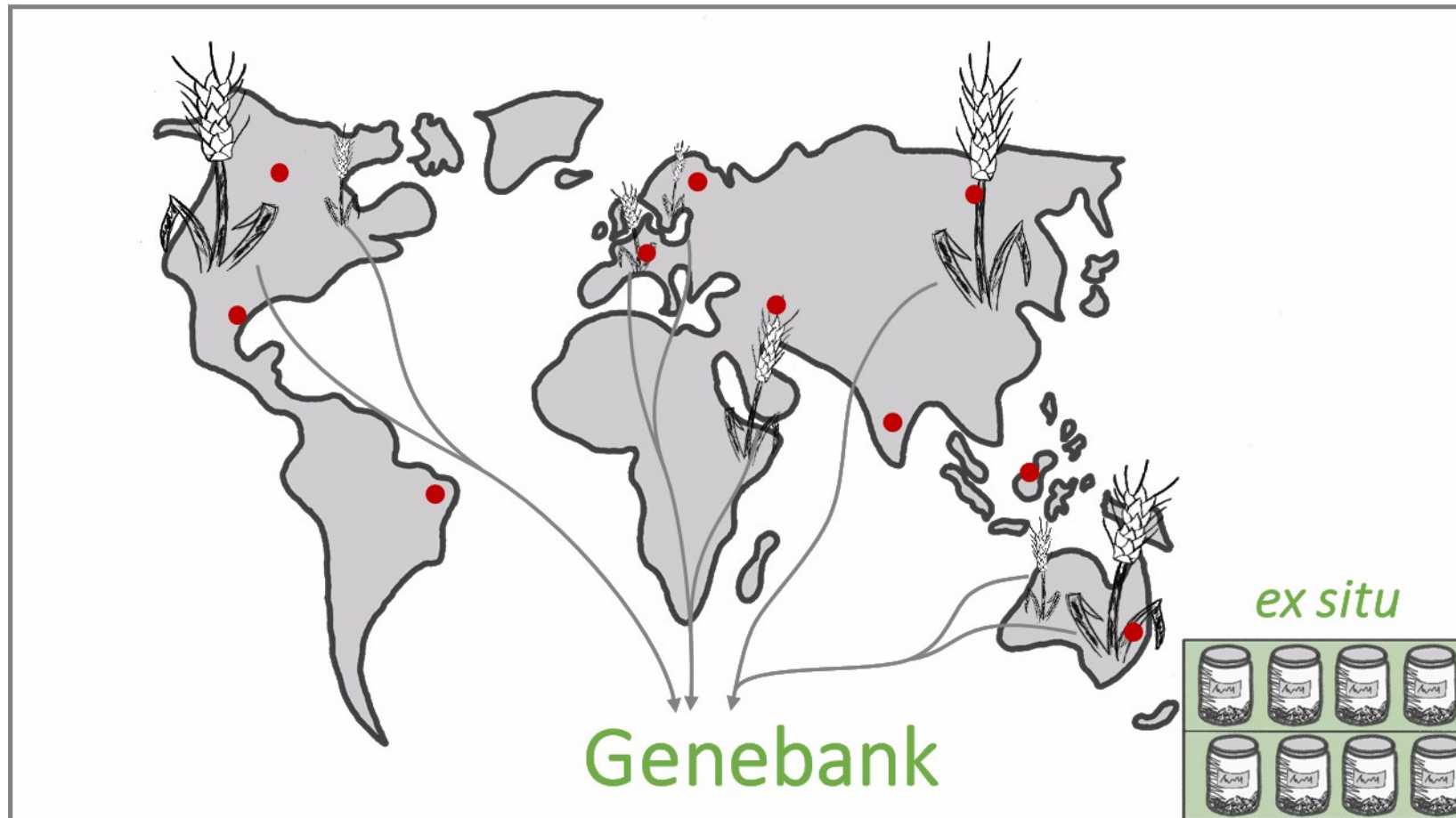


# Report on EURISCO activities since the previous meeting

Third Meeting of the EURISCO Advisory Committee  
7<sup>th</sup> July 2021



# Collections of plant genetic resources



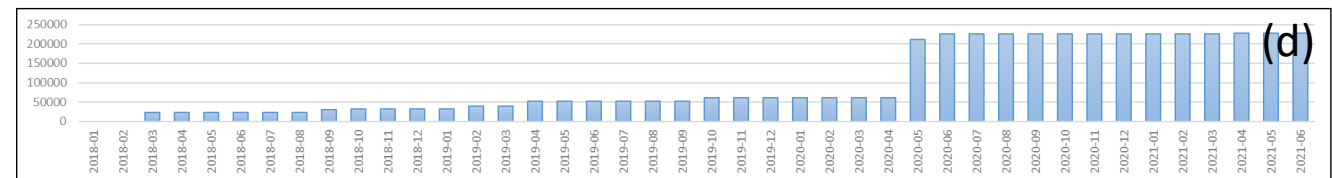
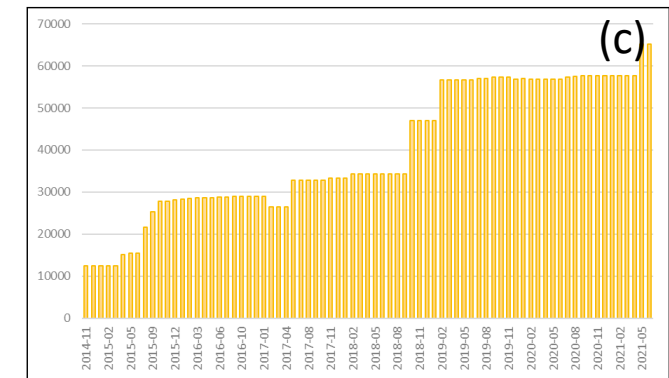
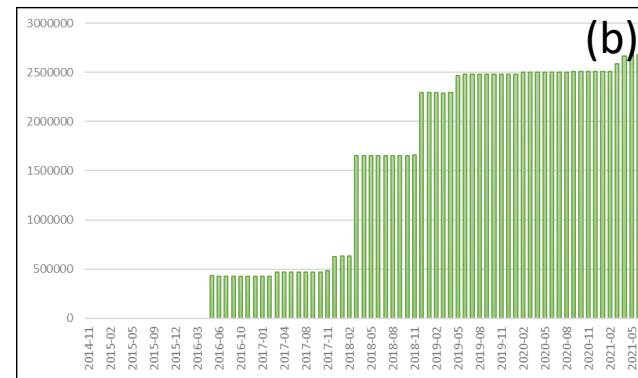
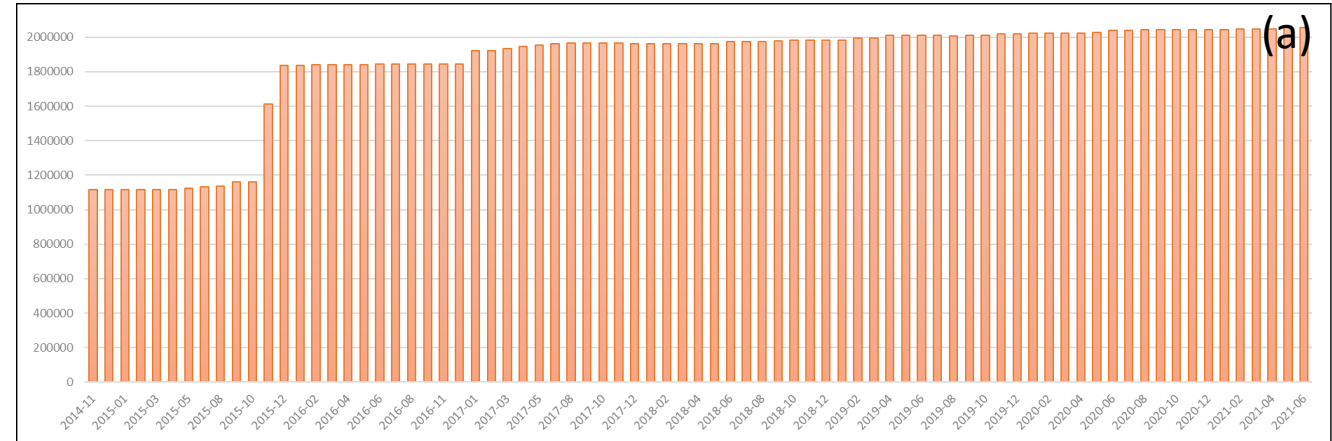
- Approx. 1,800 PGR collections around the world
- Approx. 600 PGR collections in Europe
- 401 collections provide data to EURISCO

S. Kreide / IPK

# Contents of EURISCO

- 2,055,066 accessions (a)
- 2,680,511 phenotypic data records (b)
- 6,727 genera
- 45,170 species names
- 429,223 MLS accessions
- 65,267 AEGIS accessions (c)
- 226,936 PUIDs (d)

as of 2021-06-22



# Taxonomic composition

Genus	Species	No. accs.	Total
<i>Arabidopsis</i>	<i>thaliana</i>	684,967	685,191
	others	224	
<i>Triticum</i> (wheat)	<i>aestivum</i>	139,892	199,082
	<i>durum</i>	17,439	
	<i>turgidum</i>	15,063	
	<i>monococcum</i>	3,602	
	<i>spelta</i>	3,317	
	<i>aethiopicum</i>	1,024	
	others	18,745	
<i>Hordeum</i> (barley)	<i>vulgare</i>	114,210	124,245
	<i>spontaneum</i>	6,166	
	others	3,869	
<i>Zea</i> (maize)	<i>mays</i>	64,734	64,860
	others	126	
<i>Phaseolus</i> (garden bean)	<i>vulgaris</i>	48,519	54,517
	<i>coccineus</i>	3,234	
	others	2,764	

Genus	Species	No. accs.	Total
<i>Solanum</i> (tomato, potato, eggplant, etc.)	<i>lycopersicum</i>	22,632	54,122
	<i>tuberosum</i>	14,973	
	<i>andigenum</i>	2,814	
	<i>melongena</i>	2,289	
	others	11,414	
<i>Vitis</i> (grape)	<i>vinifera</i>	38,074	47,289
	others	9,215	
<i>Avena</i> (oat)	<i>sativa</i>	33,690	42,244
	<i>sterilis</i>	2,378	
	<i>byzantina</i>	1,986	
	others	4,190	
<i>Pisum</i> (pea)	<i>sativum</i>	35,422	38,481
	others	3,059	
<i>Malus</i> (apple)	<i>domestica</i>	31,720	33,725
	others	2,005	
others			711,310
		Total	2,055,066

as of 2021-06-22

# EURISCO data quality

	Last AC meeting (2018)	2021	Increase
Accs. with collecting information	796,298	841,625	5.69%
No. of different collecting sites	106,301	112,961	6.27%
Accs. with geographic coordinates	166,984 (21% of accessions with collecting information)	215,424 (26% of accessions with collecting information)	29.01%
Accs. with donor information	1,178,522	1,199,810	1.81%
Accs. with country of origin	1,075,327	1,136,236	5.66%
Accs. with AEGIS flag	47,049	65,267	38.72%
Accs. with PUID	32,651	226,936	595.04%

# Passport data updates

Year	No. of updates	Accs. total
2014	1	1,114,995
2015	28	1,837,368
2016	25	1,842,539
2017	55*	1,964,062
2018	36	1,976,608
2019	40	2,019,414
2020	38	2,043,282
2021 (as of 2021-06-22)	28	2,055,066

\* Additional updates after FAO-WIEWS informed that instead of an annual report, the update of the datasets in EURISCO will also be taken into account.

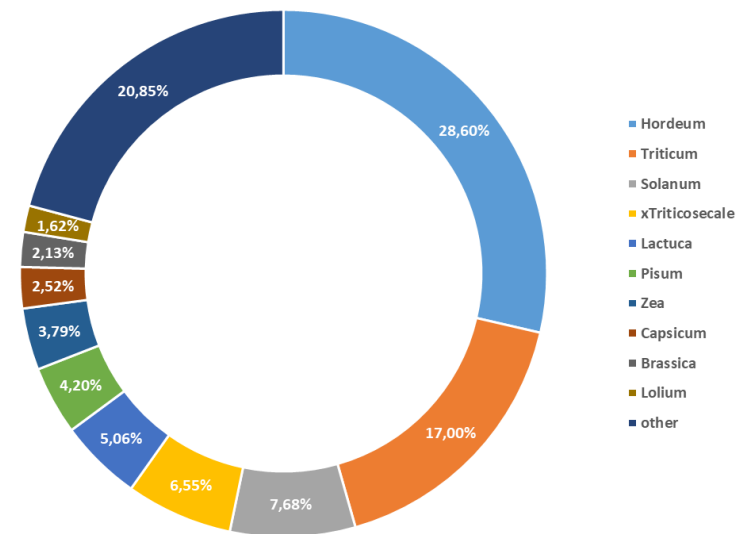
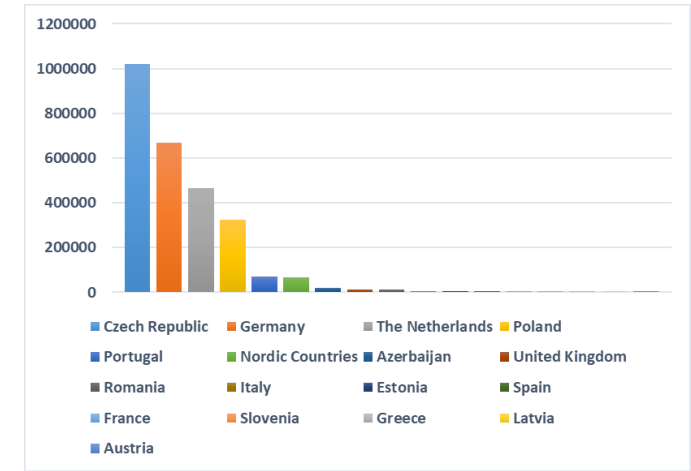
# Phenotypic data updates

Year	No. of updates	No. of records total
2016	2	427,602
2017	9	624,963
2018	3	2,293,141
2019	5	2,482,274
2020	6	2,513,267
2021 (as of 2021-06-22)	8	2,680,511



# Phenotypic data updates (cont.)

- Extension available since 2016
- Currently, 2,680,511 records of data from seventeen countries
- 68 phenotypic datasets with 3,862 experiments
- 90,569 accs. with phenotypic data



as of 2021-06-22



# Outcomes 2018

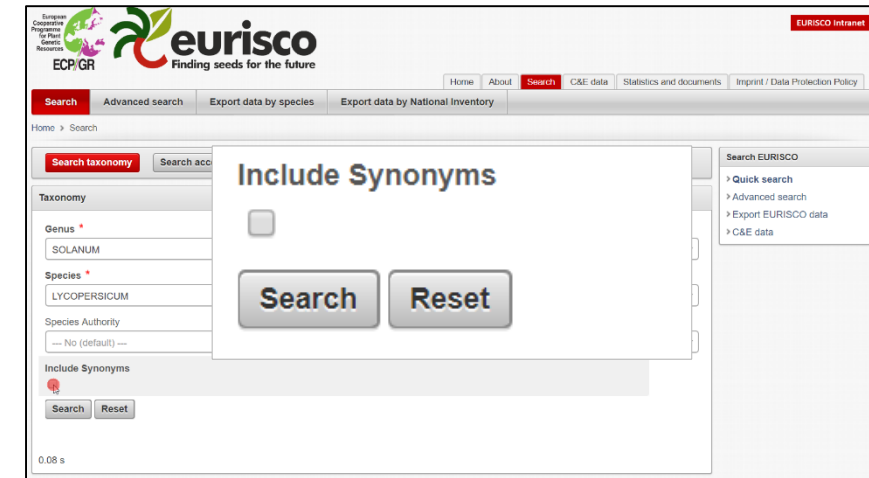
- EURISCO intranet
  - Improvement of taxonomy support (started in 2017)
- Public EURISCO application
  - Rework of simple search functionalities (taxonomy, accession, status, collecting site) + improvement of reports
  - Improved taxonomy feature for simple search
  - Rework of search functionality for phenotypic data + improvement of query performance + additional download features
  - Lots of small improvements under the bonnet
- Support for crop portals
  - Development of the European Poa Database into a crop portal with automatic updates of passport and phenotypic data from EURISCO → blueprint for other Crop WGs

# Outcomes 2018 – taxonomy example

- Challenges
  - Data from >400 institutes
    - Different taxonomic schools, opinions, traditions
    - No uniform scientific names, e.g. different author abbreviations
    - Misspellings
  - Knowledge on data background needed for searches
- Improvement of taxonomic search
  - Automatic mapping against public repositories
  - Identification of synonym candidates
    - GRIN, Mansfeld  
+ taxonomic terms accepted in EURISCO (based on user feedback)
  - Obtaining more complete search results
  - However, limited to available data!

# Outcomes 2018 – taxonomy example (cont.)

- Rework of the search interface
  - Allows to include synonyms
  - Shows the distribution of matching types
  - Significant performance improvement



The screenshot shows the EURISCO search interface. A modal dialog box titled "Include Synonyms" is displayed, featuring a checkbox and two buttons: "Search" and "Reset". The background interface shows search filters for Genus (SOLANUM) and Species (LYCOPERSICUM).

© NIAB 2019  
ISSN 1479-2621

Plant Genetic Resources: Characterization and Utilization (2019) 17(6): 559-561  
doi:10.1017/S1479262119000339

### Short Communication

#### Advancement of taxonomic searches in the European search catalogue for plant genetic resources

Stefanie Kreide\* , Markus Oppermann  and Stephan Weise 

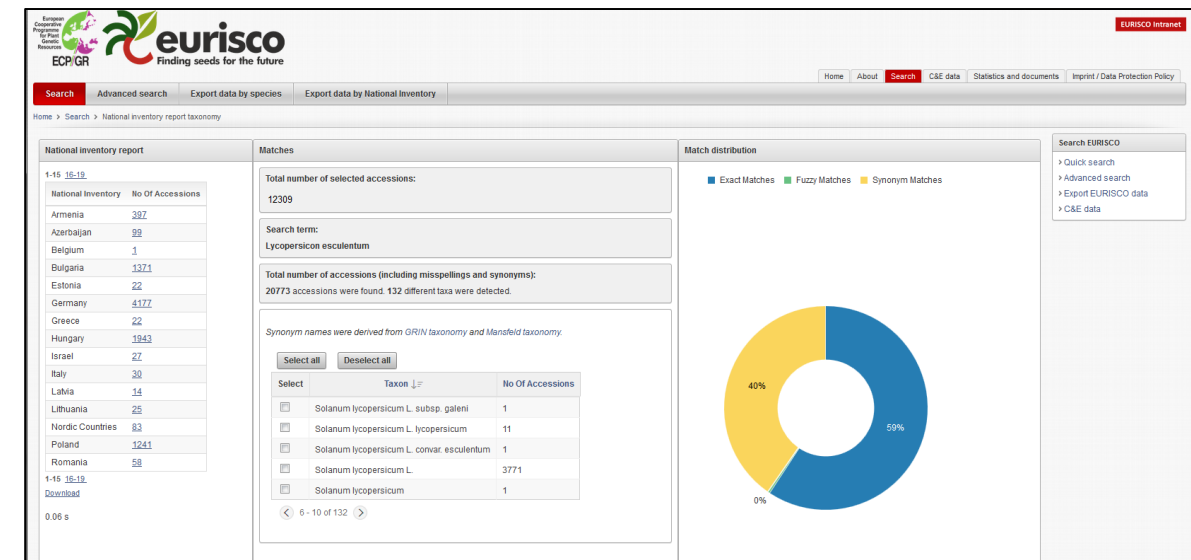
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstraße 3, 06465 Seeland, Germany

Received 6 September 2019; Accepted 4 October 2019 – First published online 31 October 2019

**Abstract**  
Genebanks play an important role in the conservation of global plant biodiversity. The European Search Catalogue for Plant Genetic Resources (EURISCO) was created as a central entry point to provide information on these collections. However, a major challenge lies in the heterogeneity of scientific plant names. This makes the selection of suitable plant material, e.g. for research or breeding purposes, significantly more difficult. For this reason, the taxonomic backbone of EURISCO has been completely revised. Search terms entered by users are now automatically checked against taxonomic reference repositories, allowing a variety of synonyms to be identified. In addition, a fuzzy search has been implemented, which makes the search function tolerant of erroneous data (e.g. caused by typing errors). Besides improvements of the search interface, more support will be given to EURISCO's data providers. The new developments provide a tool that makes it easier to identify problem cases within the data, such as accepted/non-accepted taxonomic names, and will successively improve the quality of taxonomic information in EURISCO.

**Keywords:** data integration, EURISCO, genebank community, plant genetic resources, taxonomy

**Introduction**  
Crop plants are a major source of human and animal nutrition (Grossnick and Dhillon-Brown, 2020), whereas a key role in conserving PGREFA around the world, of which 625 collections comprising more than two million accessions are maintained in Europe (Engels and Maggioni, 2012). In the context of the European Search Catalogue for Plant Genetic Resources (EURISCO),



The screenshot shows the EURISCO search results page for the search term "Lycopersicon esculentum". It displays a table of national inventory reports, a list of matches with checkboxes, and a donut chart showing the distribution of match types: Exact Matches (59%), Fuzzy Matches (40%), and Synonym Matches (0%).

National Inventory	No Of Accessions
Armenia	397
Azerbaijan	98
Belgium	1
Bulgaria	1371
Estonia	22
Germany	4177
Greece	22
Hungary	1943
Israel	27
Italy	30
Latvia	14
Lithuania	25
Nordic Countries	83
Poland	1241
Romania	58

Select	Taxon	No Of Accessions
<input type="checkbox"/>	Solanum lycopersicum L. subsp. galeni	1
<input type="checkbox"/>	Solanum lycopersicum L. lycopersicum	11
<input type="checkbox"/>	Solanum lycopersicum L. convar. esculentum	1
<input type="checkbox"/>	Solanum lycopersicum L.	3771
<input type="checkbox"/>	Solanum lycopersicum	1

Match distribution: Exact Matches (59%), Fuzzy Matches (40%), Synonym Matches (0%)

# Outcomes 2019

- EURISCO intranet
  - Extension for crop-specific passport data to handle additional information not included in MCPD
    - Frequently requested by various ECPGR Crop WGs
    - Simple and general extension following EAV approach → never used :-(
  - Implementation of DOI registration service in close collaboration with ITPGRFA
- Public EURISCO application
  - Implementation of general improvements
  - API for linking to EURISCO passport data from external information systems (e.g. Commonwealth Potato Collection, JHI)
  - Improved taxonomy feature for advanced search
  - Thorough performance tuning for phenotypic data searches
  - Implementation of download option for complete phenotypic experiments in MS Excel format
  - Crop portal for forages in the frame of the ECPGR Grant Scheme Activity “ImprovLoliumCol”

# Outcomes 2020

- No EURISCO developer between February and November
- EURISCO intranet
  - Rework of update mechanism for National Focal Points started (native PL/SQL implementation)
- ECPGR-EVA
  - Development of two exchange formats for phenotypic data (a simplified one, a more fine-grained one)
  - Recording of basic database requirements of the EVA infrastructure

## Outcomes 2021 (ongoing)

- EURISCO intranet
  - Final tests of reworked update mechanism for National Focal Points
- Public EURISCO application
  - Full reengineering of the web interface (ongoing)
    - Regular technical revision
    - Responsive design with clearer layout
    - Support of mobile platforms
    - Critical review of functionalities
  - RESTful services (ongoing) as additional means of access to EURISCO data
- ECPGR-EVA
  - EVA database infrastructure + web interface under development

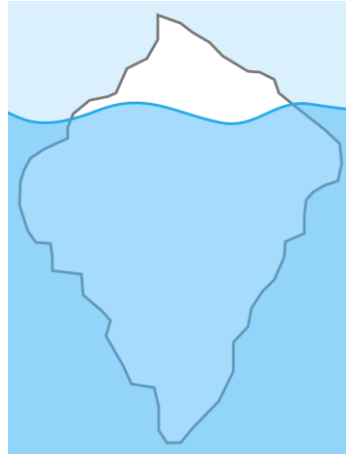
# Architecture: DB schema features

- EURISCO intranet
  - 64 tables
  - 510 indexes
  - 113 triggers
  - 14 PL/SQL packages
  - 181 functions and procedures
    - data upload and import
    - integrity checks
    - updates (passport and phenotypic)
    - taxonomy support
  - 58 sequences
  - 30 Java classes
- EURISCO web
  - 48 tables
  - 32 materialised views
  - 673 indexes
  - 9 PL/SQL packages
  - 42 functions and procedures
    - download
    - newsletter
    - statistics
    - phenotypic data visualisation
    - AEGIS status auditing
    - taxonomy support
  - In-memory features
  - 7 Java classes

as of 2021-06-22



# Web interface



The screenshot shows the EURISCO website with the following elements:

- Header:** EURISCO logo with tagline "Finding seeds for the future", ECP-GR logo, and "EURISCO Intranet" link.
- Navigation:** Home, News, Search, C&E data, Statistics and documents, Imprint / Data Protection Policy.
- Main Content:**
  - Welcome to EURISCO: Introduction text about the European Search Catalogue for Plant Genetic Resources.
  - How to obtain germplasm: Information on ordering accessions.
  - News: Links to national inventory data updates for France, Albania, and Spain.
  - Accessions in EURISCO: Two bar charts showing the number of accessions per national inventory and per genus.
- Right Sidebar:**
  - Search EURISCO: Quick search, Advanced search, Export EURISCO data, C&E data.
  - EURISCO newsletter: Subscribe / unsubscribe.
  - Regions of origin: World map.
  - Statistical overview: 2,055,066 Accessions, 401 Institutes, 43 Countries, 6,727 Genera, 45,170 Species, 429,223 MLS accessions, 65,267 AEGIS accessions, 226,936 DOIs.
  - Site rating: Your rating and Average.

Published online 31 August 2016

Nucleic Acids Research, 2017, Vol. 45, Database issue D1003–D1008  
doi: 10.1093/nar/gkw755

## EURISCO: The European search catalogue for plant genetic resources

Stephan Weise<sup>1,\*</sup>, Markus Oppermann<sup>1</sup>, Lorenzo Maggioni<sup>2</sup>, Theo van Hintum<sup>3</sup> and Helmut Knüpffer<sup>1</sup>

<sup>1</sup>Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstr. 3, 06466 Stadt Seeland, Germany, <sup>2</sup>Bioversity International, Via del Tre Denari 472/a, 00057 Maccarese (Fiumicino), Rome, Italy and <sup>3</sup>Centre for Genetic Resources, The Netherlands (CGN), Wageningen University and Research Centre, P.O. Box 16, 6700 AA Wageningen, The Netherlands

Received July 28, 2016; Revised August 19, 2016; Accepted August 21, 2016

### ABSTRACT

The European Search Catalogue for Plant Genetic Resources, EURISCO, provides information about 1.8 million crop plant accessions preserved by almost 400 institutes in Europe and beyond. EURISCO is being maintained on behalf of the European Cooperative Programme for Plant Genetic Resources. It is based on a network of National Inventories of 43

typic characterisation of genebank accessions, i.e. collecting information about traits such as disease resistance, drought tolerance and yield components. These data are usually generated on selected material, resulting in non-orthogonal, highly incomplete data sets. Nevertheless, the analysis of these data allows meaningful results, e.g. the identification of promising new alleles (5). Around the world, there are about 1800 genebank collections conserving PGRFA. Thereof about 625 collections are maintained in Europe

54 (sub)versions since 2014

# Version history of the web interface

v1.0.0	v1.1.0	v1.2.0	v1.3.0	v1.4.0	v1.5.0	v2.0.0
<ul style="list-style-type: none"> <li>• Oct. 2014</li> <li>• First public version</li> <li>• v1.0.1 – v1.0.6 continuous improvements</li> </ul>	<ul style="list-style-type: none"> <li>• Nov. 2014</li> <li>• New export functionality + download of full dump</li> <li>• v1.1.1 – v1.1.17 (2015 – 2016) continuous improvements</li> </ul>	<ul style="list-style-type: none"> <li>• Jun. 2016</li> <li>• C&amp;E data extension; new export functionalities; new advanced search; lots of small improvements</li> <li>• v1.2.1 – v1.2.7 (2016 – 2017) continuous improvements</li> </ul>	<ul style="list-style-type: none"> <li>• Dec. 2017</li> <li>• Migration to MCPD2; increased usability; lots of small improvements</li> <li>• v1.3.1 – v1.3.5 (2018) continuous improvements</li> </ul>	<ul style="list-style-type: none"> <li>• Jun. 2018</li> <li>• Taxonomy search simple completely reworked</li> <li>• v1.4.1 – v1.4.9 (2018 – 2019) continuous improvements</li> </ul>	<ul style="list-style-type: none"> <li>• Sep. 2019</li> <li>• Taxonomy search advanced completely reworked</li> <li>• v1.5.1 – v1.5.4 (2019 – 2020) continuous improvements</li> </ul>	<ul style="list-style-type: none"> <li>• Expected in 2021</li> <li>• Fully reengineered web interface</li> </ul>

# Network maintenance + development

- Contact with EURISCO stakeholders
- Definition of new services, e.g. with regard to DOIs
- Advancement of current and definition of new standards, e.g. with regards to phenotypic data
- Coordination with initiatives such as Genesys and GLIS
- Bilateral communication with regard to the coverage of EURISCO
- Cooperation with ECPGR Working Groups
- Preparation of work plans and reports
- Helpdesk behind the scenes (should not be underestimated ;-)

# International integration of EURISCO

- Genesys (Global Crop Diversity Trust)
- FAO-WIEWS
- Germinate (JHI)
- Vital part of the global information system (GLIS)
- Data provider for ECPGR crop working groups



Food and Agriculture  
Organization of the  
United Nations



**The International Treaty**  
ON PLANT GENETIC RESOURCES  
FOR FOOD AND AGRICULTURE



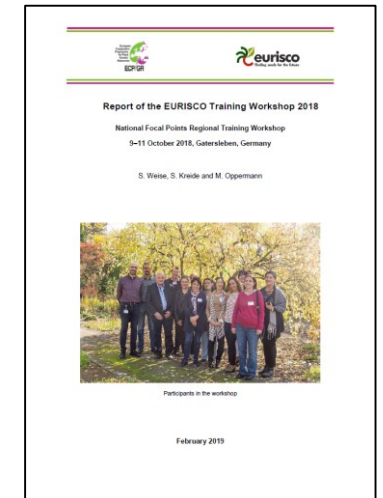
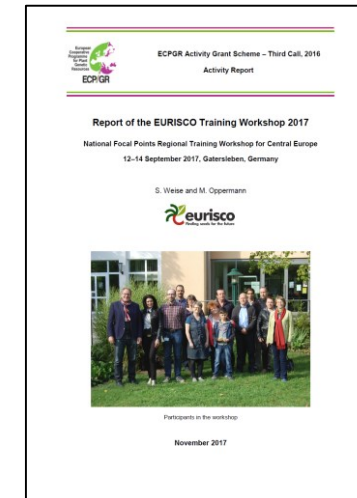
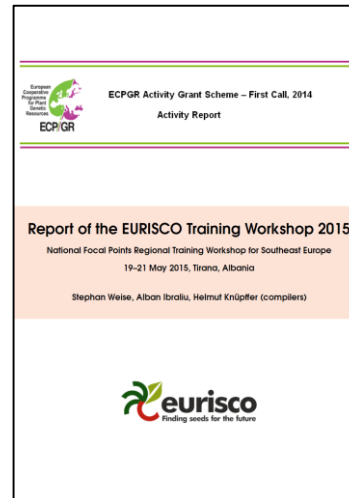
# Participation in project consortia

- Various ECPGR Grant Scheme Activities
- EURISCO Taxonomy (ECPGR, BLE), 2017–2018
  - Improvement of taxonomic backbone of EURISCO
- EUCLEG (Horizon 2020), 2017–2021
  - Coordination: INRA, France
  - Aim: reduction of protein import dependencies of both European and Chinese partners
  - Leader of work package for data management
    - Find data gaps in EURISCO (and try to close them)
    - Manage project data (passport, phenotypic, genetic)
- Farmer's Pride (Horizon 2020), 2018–2020
  - Coordination: University of Birmingham, U.K.
  - Aim: Development of network of *in situ* sites and stakeholders
  - Task leader: Preparation of a concept to extend EURISCO for *in situ* data
- GenRes Bridge (Horizon 2020), (2019–2021)
  - Coordination: European Forest Institute
  - Aim: Join forces of plant, forest and animal genetic resources
  - User training + information system linking
- ECPGR European Evaluation Network (initial funding BLE), 2019–2022
  - Coordination: ECPGR
  - Aim: Implementation of the evaluation network on wheat/barley and vegetable crops
  - Development of infrastructure
- AGENT (Horizon 2020), (2020–2025)
  - Coordination: IPK Gatersleben
  - Aim: Activate genebanks and facilitate access to plant genetic resources
  - Task leader: Data exchange and representation; infrastructure for managing and analysing genotypic and phenotypic data about genetic resources



# EURISCO training workshops

- Most active data providers trained
- Switch to bi-annual trainings
- Next focus:
  - Southeast Europe
  - Northern Europe
- 2020 training should take place in Bulgaria
  - Had to be postponed due to COVID-19 pandemic
  - According to current planning, end of 2021 as an online event
- 2021 individual training for UK representatives



# Dissemination

- Regularly short information in ECPGR bulletin
- EURISCO newsletter twice a year
- Journal articles
- EURISCO talks and posters on several conferences
- Presentations on several ECPGR workshops
- ECPGR Grant Scheme activities





M. Grau / IPK

**Thank you for your attention**