Stephan Weise

EURISCO – The European Search Catalogue for Plant Genetic Resources

EURISCO training workshop 2023
12–14 September 2023, Plovdiv, Bulgaria

Collections of plant genetic resources

- Approx. 1,800 PGR collections around the world
- Approx. 600 PGR collections in Europe
- 5.8 million *ex situ* accessions globally

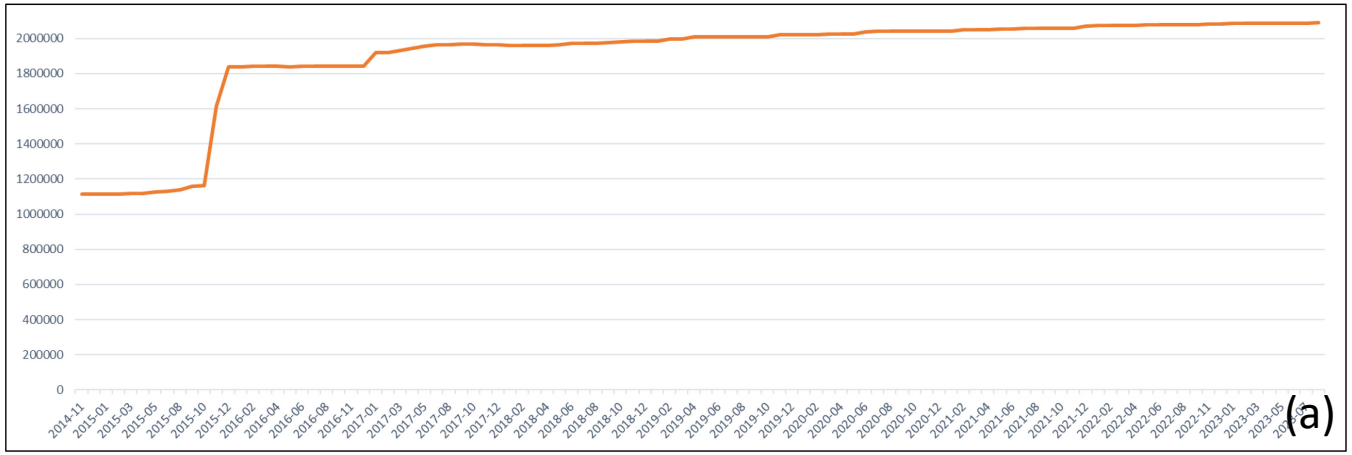
European approach: EURISCO

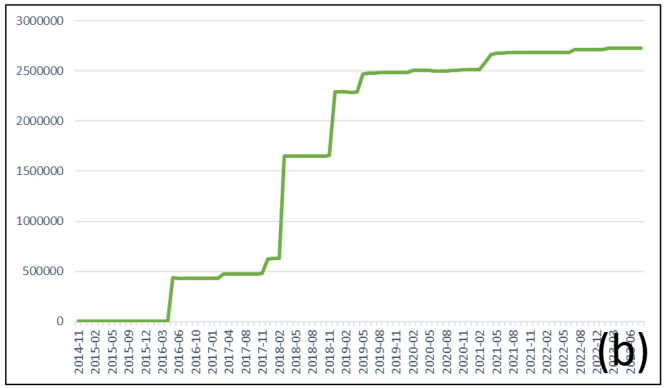
- What is EURISCO?
  - European information system for plant genetic resources
  - Search catalogue for *ex situ* collections
  - Accession-level information system
- Purpose
  - Provides passport data and phenotypic data about plant germplasm accessions maintained in Europe
  - Assists in meeting national obligations
    - Food and Agriculture Organization of the United Nations (FAO)
    - Convention on Biological Diversity (CBD)
    - International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA)

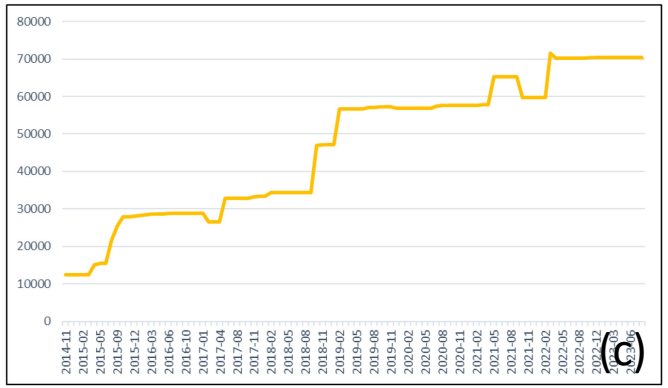
Key facts

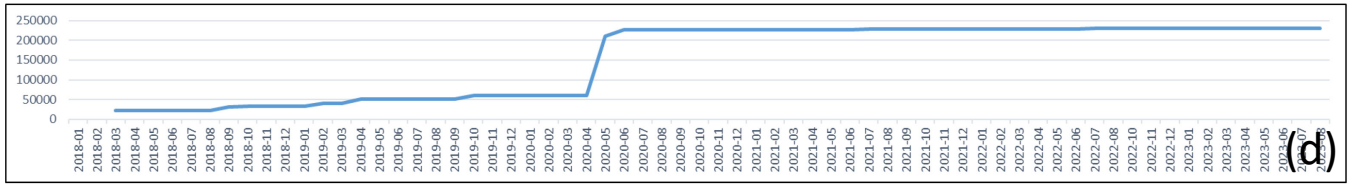
- **2,088,704** accessions (a)
- **2,726,998** phenotypic data records (b)
- **6,731** genera
- **45,115** species
- **430,317** MLS accessions
- **70,427** AEGIS accessions (c)
- **230,758** PUIDs/DOIs (d)

as of 2023-08-29



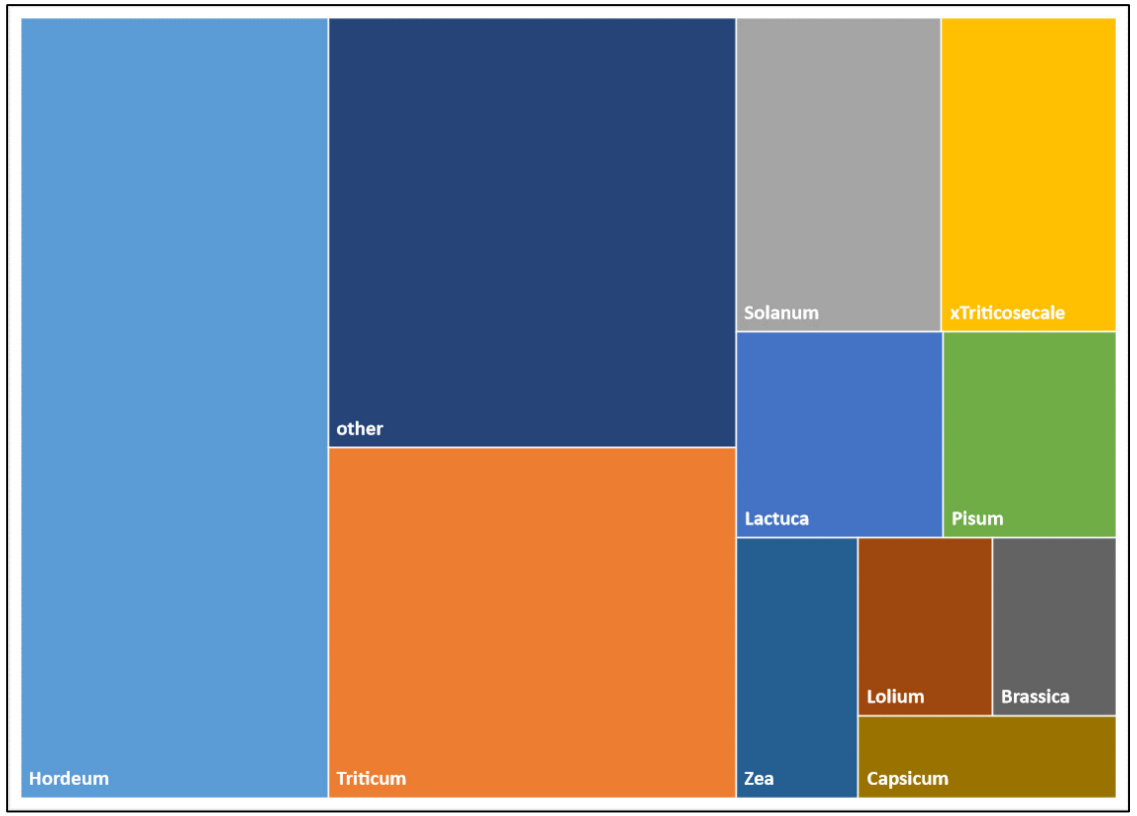






Phenotypic data

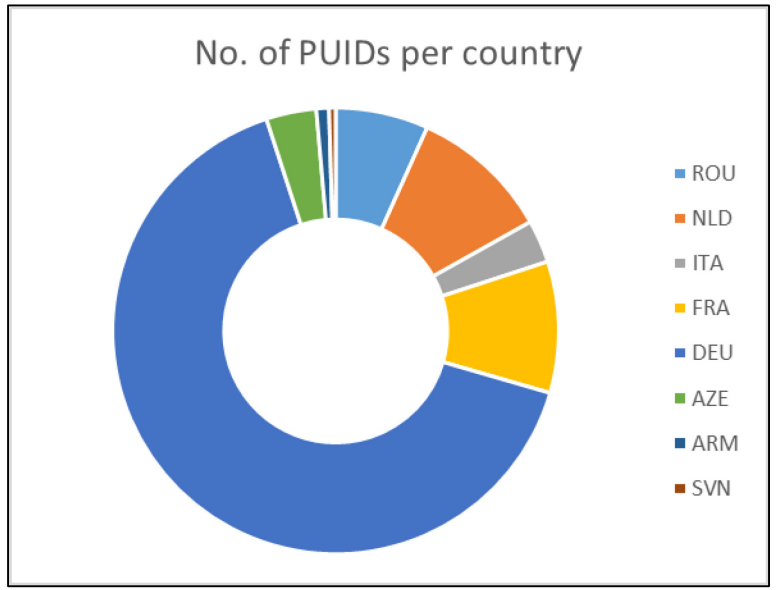
- Extension available since 2016
- Currently, 2,726,998 records of data from 21 countries
- 73 phenotypic datasets with 3,919 experiments
- 91,366 accs. with phenotypic data



as of 2023-08-29

PUIDs/DOIs

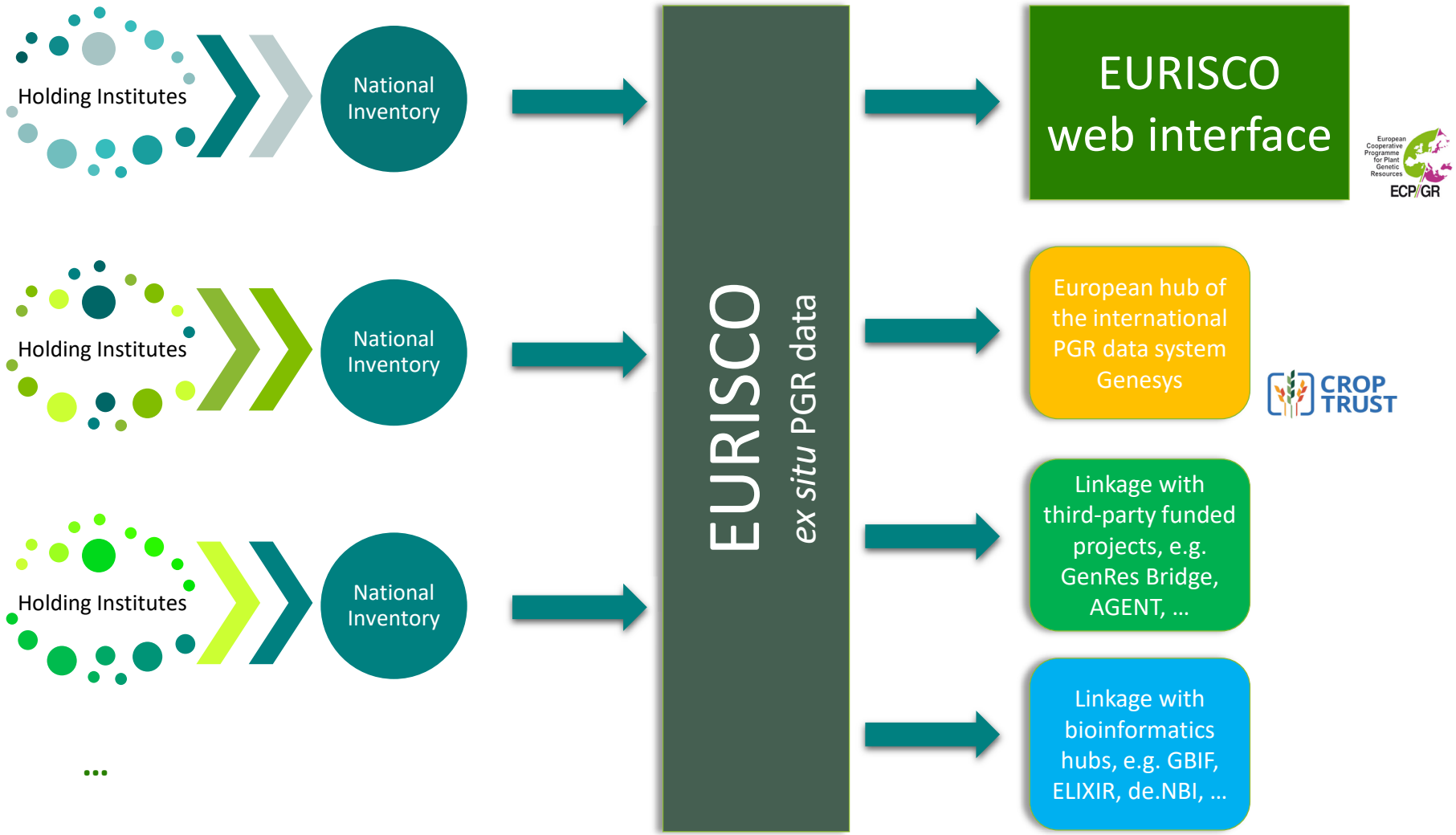
- 230,758 PUIDs
  - DOI assignment is a time consuming process
  - 23 institutes from 8 countries
  - Most of them DOIs (all except Italy)
- Own assignment of DOIs
  - France, Germany
- Use of the Treaty infrastructure
  - Netherlands
- Use of the EURISCO DOI service
  - Armenia, Azerbaijan, Romania, Slovenia
- Other type of PUIDs
  - Italy → about to switch to DOIs



as of 2023-08-29

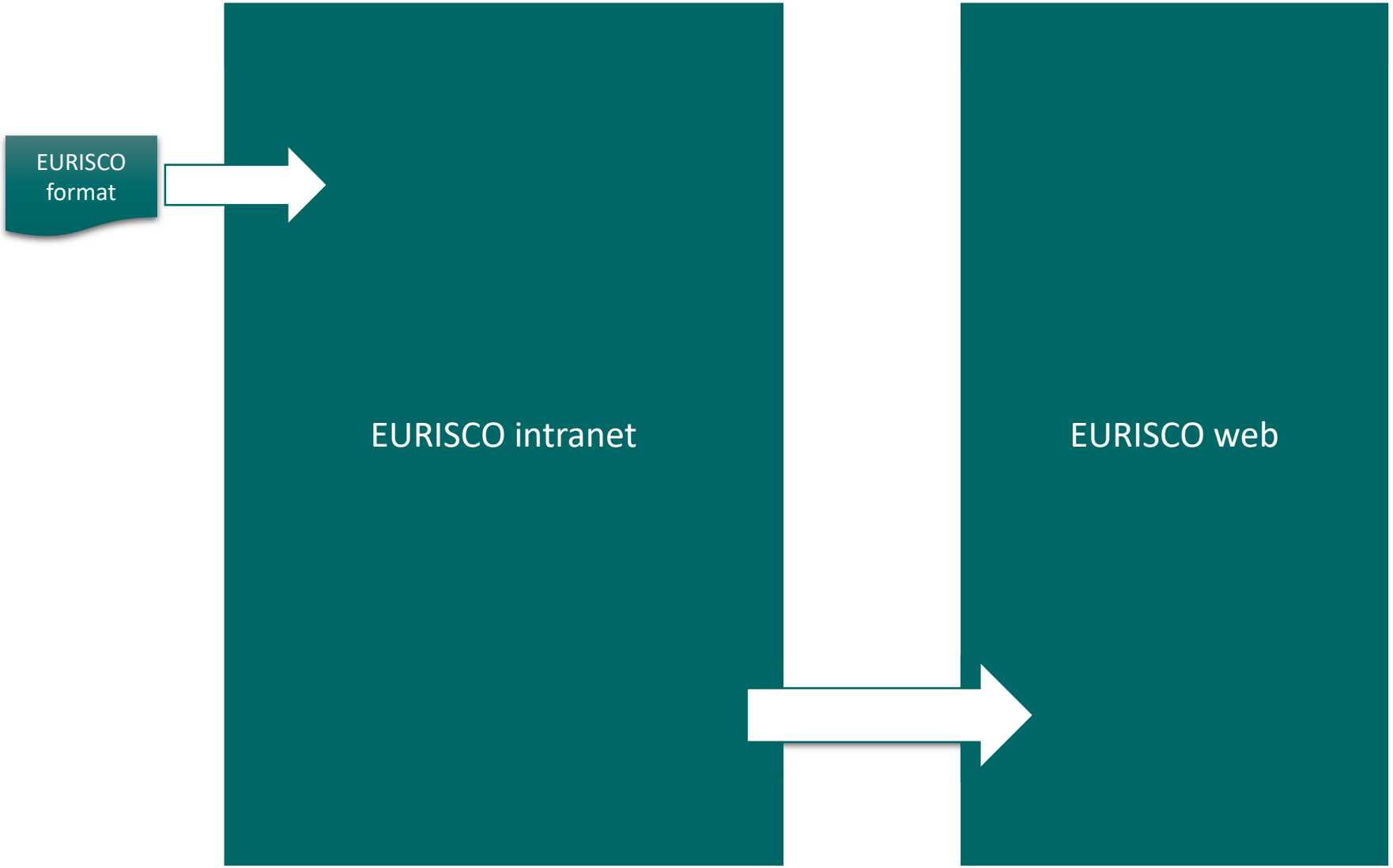
EURISCO development

- Started in 1999 (EU project EPGRIS)
- 43 countries involved
  (Nordic Countries → NordGen)
- National collections represented by
  National Inventories (NIs)
- Network of National Focal Points (NFPs)
  links NIs ↔ EURISCO


Holding Institutes
National Inventory
Holding Institutes
National Inventory
Holding Institutes
National Inventory
...
EURISCO
ex situ PGR data
EURISCO web interface
European hub of the international PGR data system Genesys
Linkage with third-party funded projects, e.g. GenRes Bridge, AGENT, ...
Linkage with bioinformatics hubs, e.g. GBIF, ELIXIR, de.NBI, ...
ECP/GR
CROP TRUST

EURISCO hosting

- On behalf of the European Cooperative Programme for Plant Genetic Resources (ECPGR)
  - Available online since 2003
    - Bioversity International, Rome, Italy
  - Since April 2014:
    - EURISCO hosted by IPK Gatersleben, Germany
    - New development from scratch
    - So far, two 5-year phases (until the end of 2023); further extension endorsed by ECPGR Steering Committee


EURISCO format
EURISCO intranet
EURISCO web

Behind the scenes: database architecture features

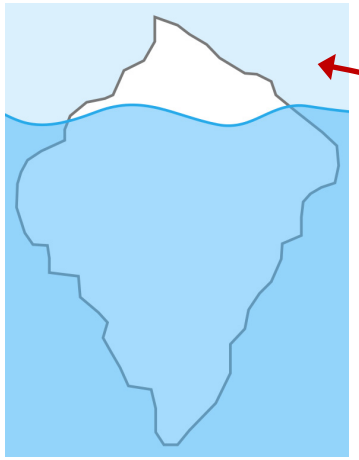
- EURISCO intranet
  - 64 tables
  - 511 indexes
  - 117 triggers
  - 16 PL/SQL packages
  - 202 functions and procedures
    - data upload and import
    - integrity checks
    - updates (passport and phenotypic)
    - taxonomy support
  - 60 sequences
  - 30 Java classes

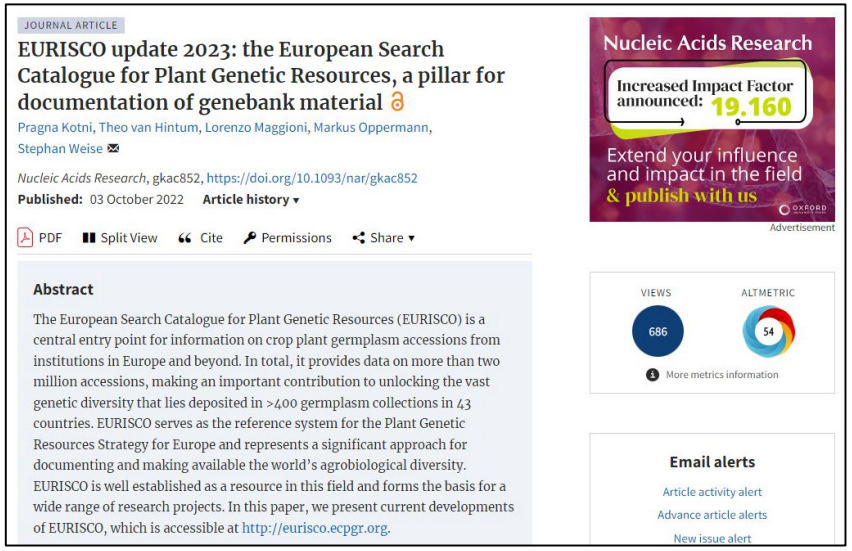
- EURISCO web
  - 56 tables
  - 39 materialised views
  - 771 indexes
  - 11 PL/SQL packages
  - 56 functions and procedures
    - download
    - newsletter
    - statistics
    - phenotypic data visualisation
    - AEGIS status auditing
    - taxonomy support
  - In-memory features
  - 9 Java classes

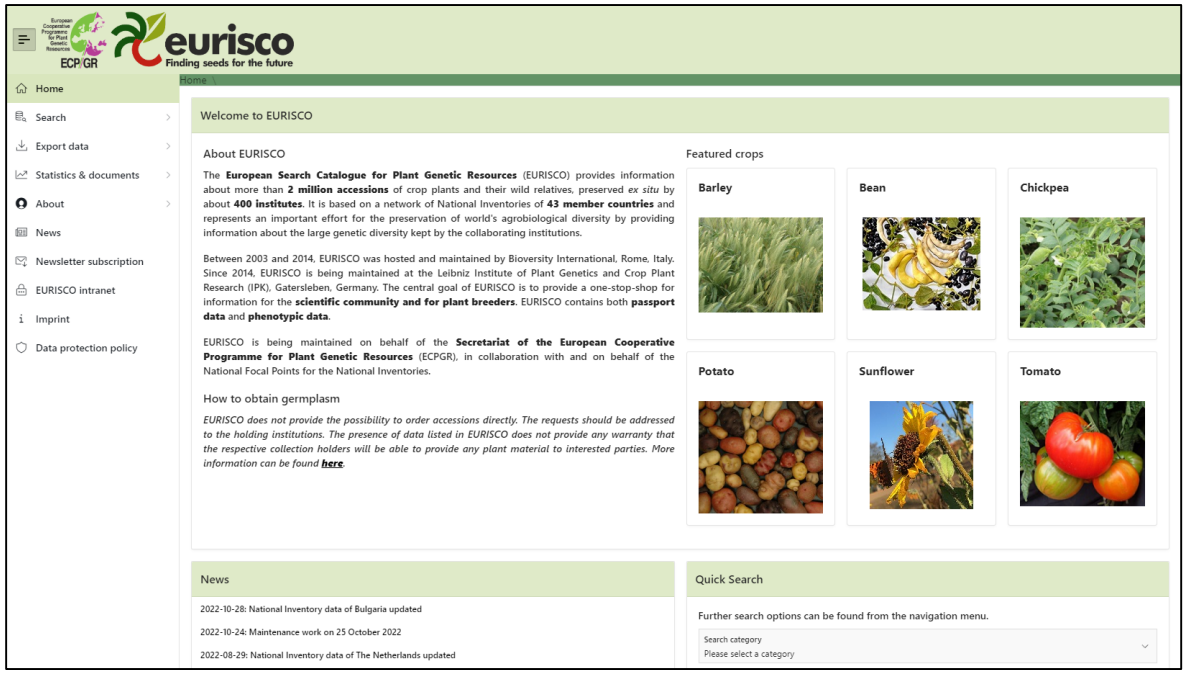
The figures do not include the objects of the *in situ* CWR extension, as it is still under development.

as of 2023-08-29

Web interface







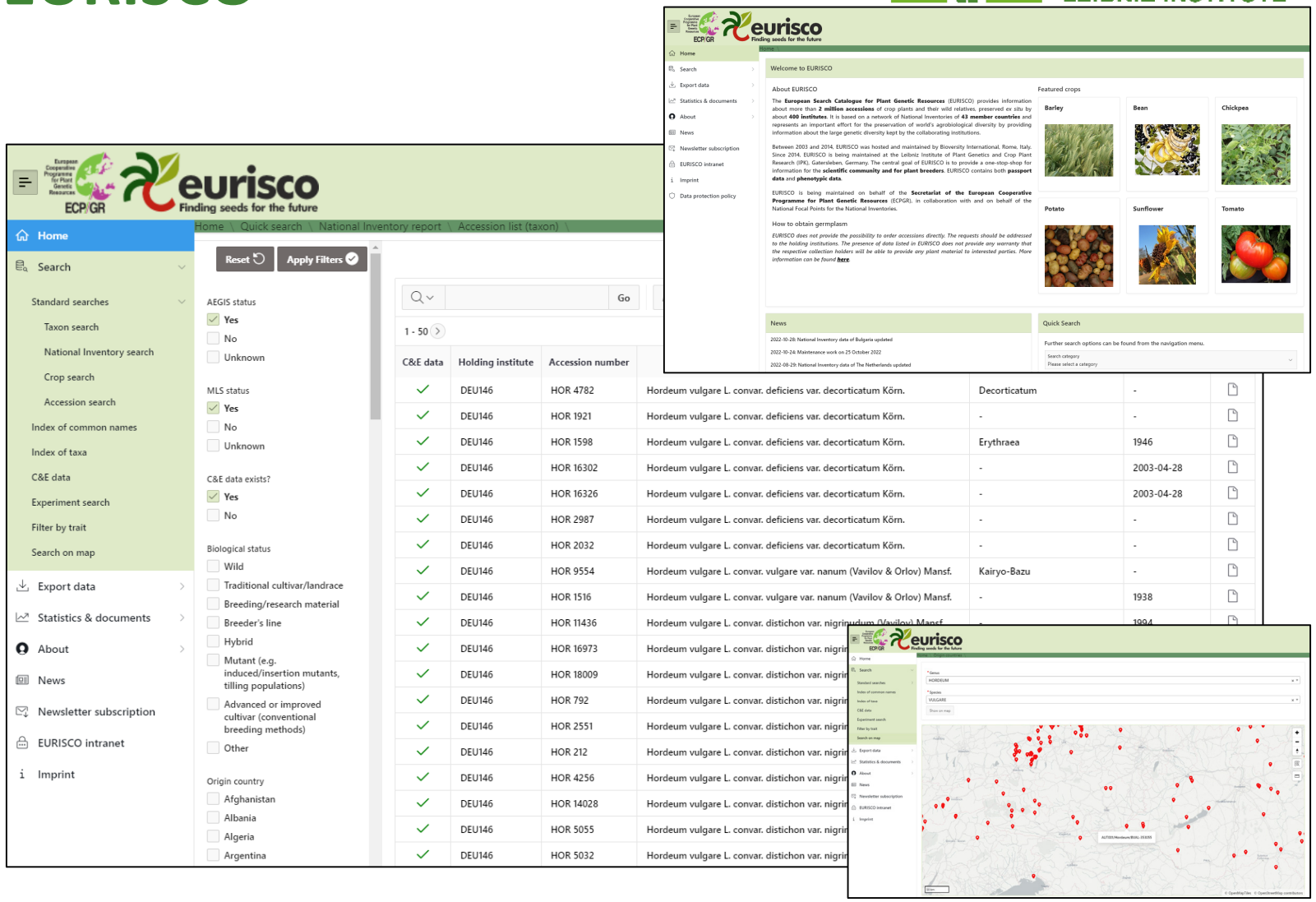
Version history of the public interface

v1.0.0	v1.1.0	v1.2.0	v1.3.0	v1.4.0	v1.5.0	v2.0.0	v2.1.0
• Oct. 2014 • First public version • v1.0.1 – v1.0.6 continuous improvements	• Nov. 2014 • New export functionality + download of full dump • v1.1.1 – v1.1.17 (2015 – 2016) continuous improvements	• Jun. 2016 • C&E data extension; new export functionalities; new advanced search; lots of small improvements • v1.2.1 – v1.2.7 (2016 – 2017) continuous improvements	• Dec. 2017 • Migration to MCPD2; increased usability; lots of small improvements • v1.3.1 – v1.3.5 (2018) continuous improvements	• Jun. 2018 • Taxonomy search simple completely reworked • v1.4.1 – v1.4.9 (2018 – 2019) continuous improvements	• Sep. 2019 • Taxonomy search advanced completely reworked • v1.5.1 – v1.5.4 (2019 – 2020) continuous improvements	• Mar. 2022 • Fully reengineered web interface (new technological basis, additional functionalities) • v2.0.1 – v2.0.8 (2022) contunious improvements	• *Under preparation* • *Improvement of reports & export mechanism; improvement of passport/ phenotypic searches; DOI search*

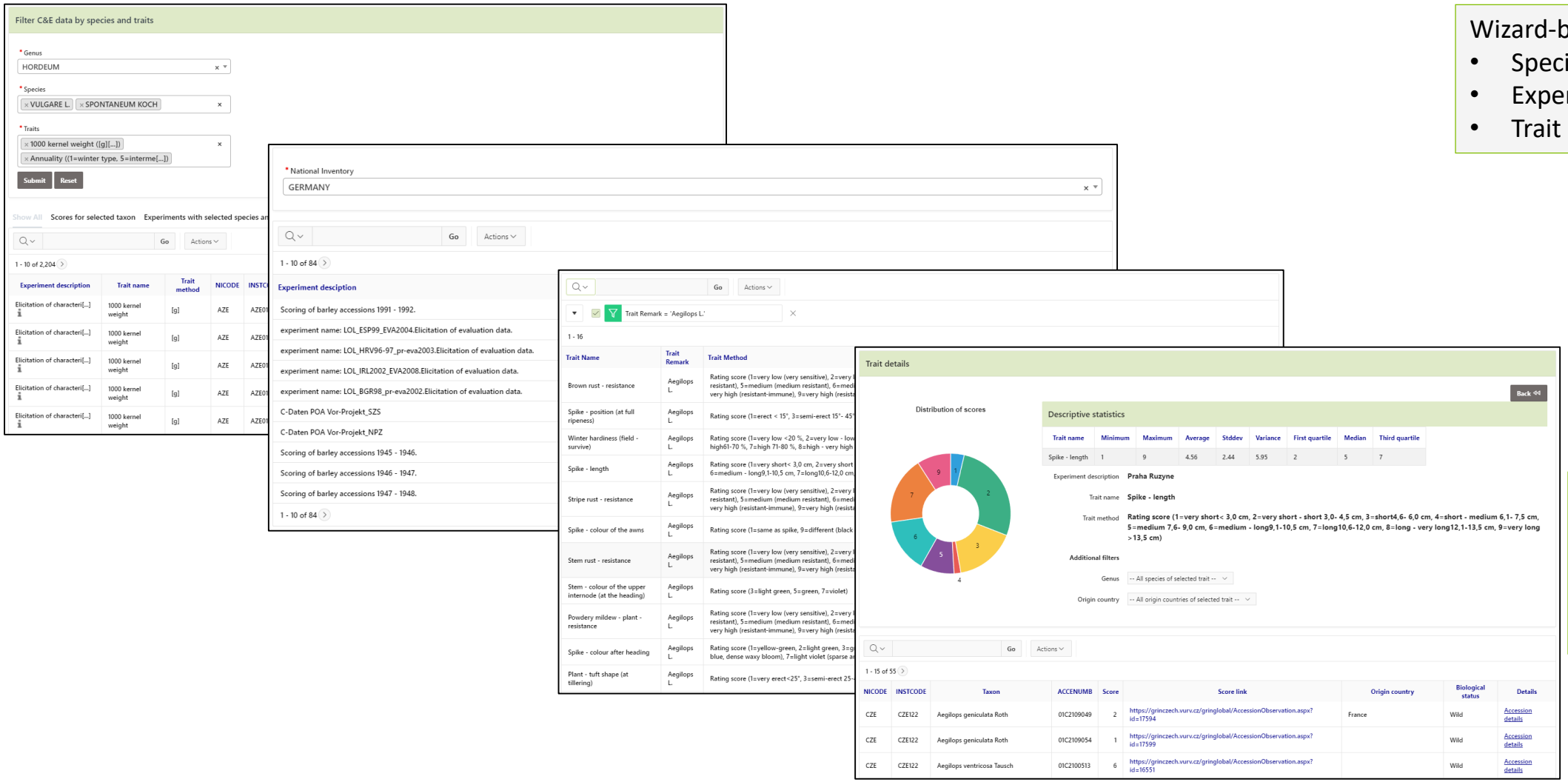
A total of 70 versions and sub-versions of the public EURISCO web interface have been completed since 2014, 26 of which since 2019.

Passport data search in EURISCO

- Four standard searches:
  - Taxon search (incl. synonyms)
  - National Inventory search
  - Crop search
  - Accession search
- Index of common names
- Index of taxa
- Map-based search
- Various data export features



Phenotypic data search in EURISCO



Wizard-based searches for

- Species and trait
- Experiment
- Trait

Refine result

- Sort
- Filter
- Download
- Chart

AEGIS

- AEGIS - A European Genebank Integrated System
- ECPGR initiative for improving the coordination of
  - Conservation and management of PGRFA
  - Access to PGRFA
- Aims:
  - Conservation of genetically unique and important accessions
  - Making materials available for breeding and research
  - Safe long-term conservation (with common agreed standards)
  - Reduction of redundancy
  - Clarification of responsibilities for conservation

AEGIS data in EURISCO

- No physical collection → “virtual” genebank
- AEGIS accessions labelled in EURISCO
  - Including tracking of AEGIS status
- 70,427 AEGIS accessions
- 35 member countries

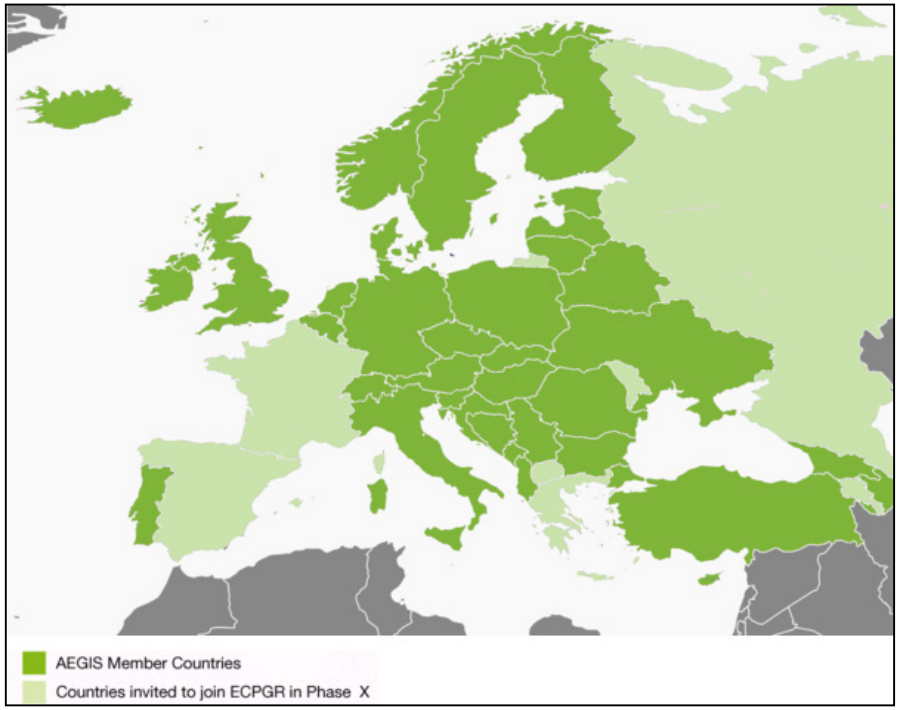
AEGIS status auditing

Go | Actions

1 - 15

NICODE	INSTCODE	AEGISSTAT old	AEGISSTAT current	Change timestamp	DML operation	No. of accessions
ITA	ITA380	1		2023-06-05	Existing AEGIS accession completely removed from EURISCO	44 (click for details)
ITA	ITA380		1	2023-06-05	Existing EURISCO accession labelled as part of AEGIS	44 (click for details)
CZE	CZE050		1	2023-01-04	Existing EURISCO accession labelled as part of AEGIS	1 (click for details)
CZE	CZE122		1	2023-01-04	Existing EURISCO accession labelled as part of AEGIS	12 (click for details)
ITA	ITA380	0	1	2022-11-25	Existing EURISCO accession labelled as part of AEGIS	14 (click for details)
ITA	ITA380	1	0	2022-11-25	Existing AEGIS accession de-flagged	4 (click for details)
ITA	ITA380	1		2022-11-25	Existing AEGIS accession completely removed from EURISCO	5 (click for details)
ITA	ITA381	0	1	2022-11-25	Existing EURISCO accession labelled as part of AEGIS	1 (click for details)
ITA	ITA381	1	0	2022-11-25	Existing AEGIS accession de-flagged	1 (click for details)
ITA	ITA378	1		2022-11-18	Existing AEGIS accession completely removed from EURISCO	1 (click for details)
ITA	ITA380	0	1	2022-11-18	Existing EURISCO accession labelled as part of AEGIS	33 (click for details)
ITA	ITA380	1	0	2022-11-18	Existing AEGIS accession de-flagged	2 (click for details)
ITA	ITA380	1		2022-11-18	Existing AEGIS accession completely removed from EURISCO	2 (click for details)
ITA	ITA380		1	2022-11-18	Existing EURISCO accession labelled as part of AEGIS	6 (click for details)
ITA	ITA381	0	1	2022-11-18	Existing EURISCO accession labelled as part of AEGIS	1 (click for details)

1 - 15



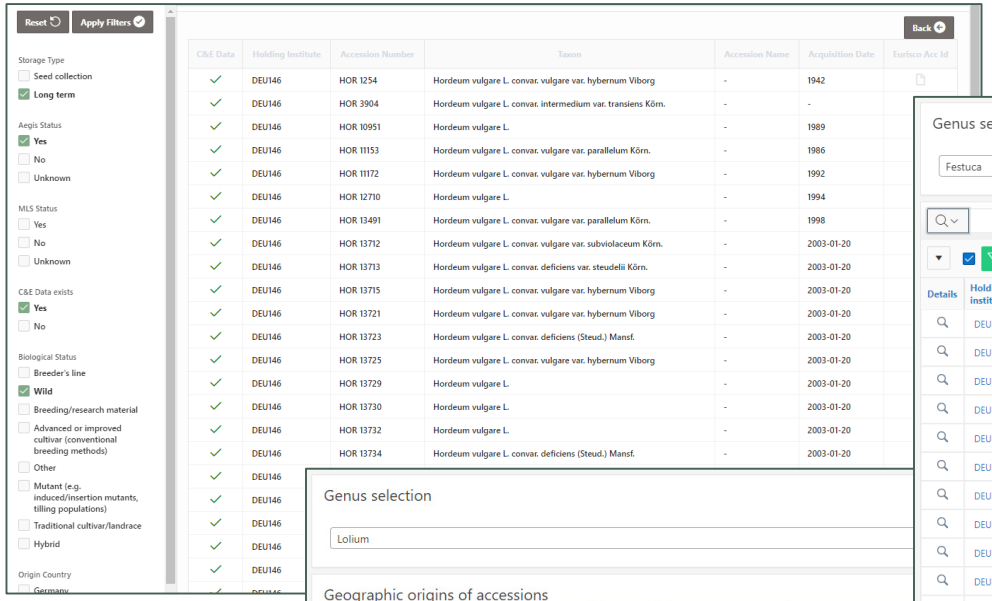
as of 2023-08-29

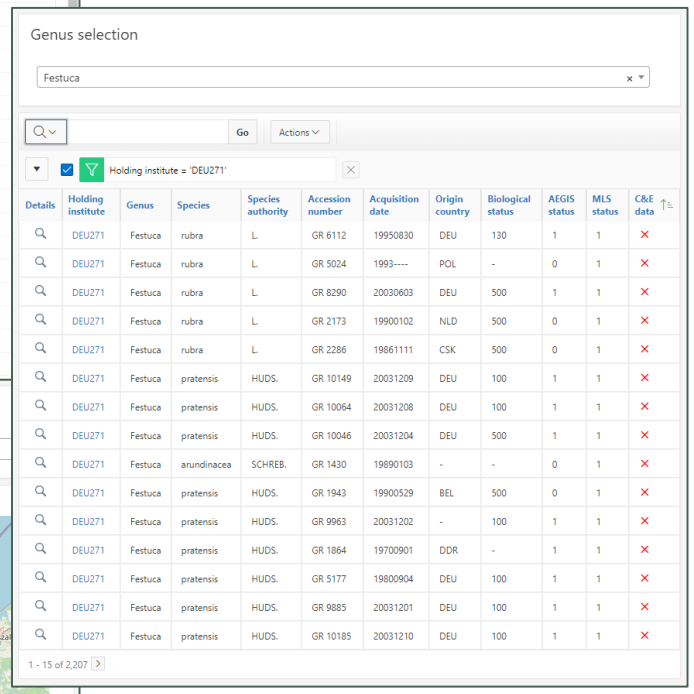
International integration and data provision of EURISCO

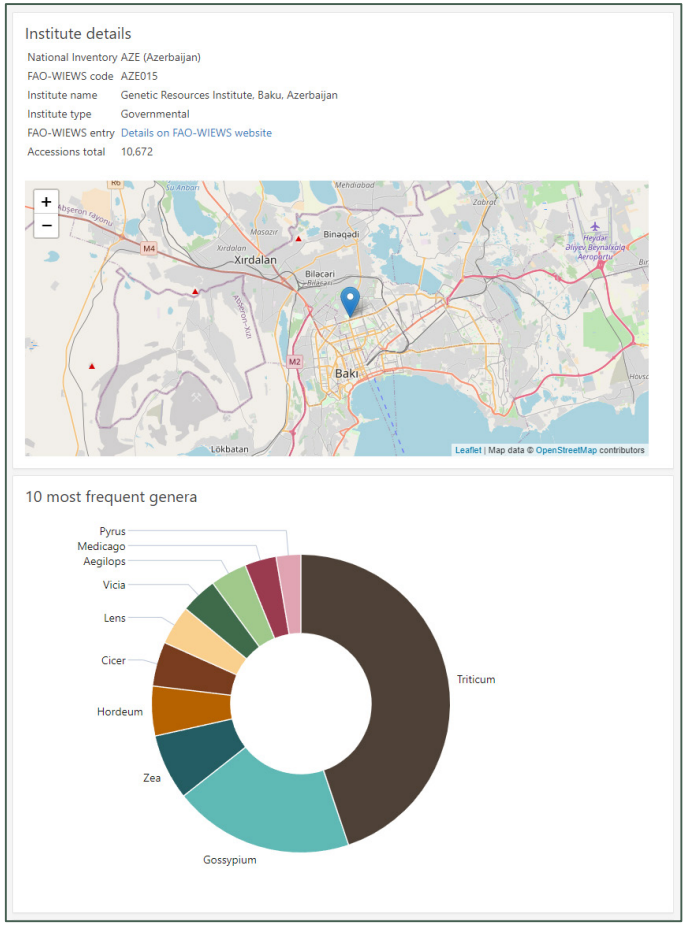
- Genesys (Global Crop Diversity Trust)
- FAO-WIEWS
- Germinate (JHI)
- Vital part of the global information system (GLIS)
- Data provider for ECPGR crop working groups

- *EURISCO is used as a data provider for various projects*

Example of data provision





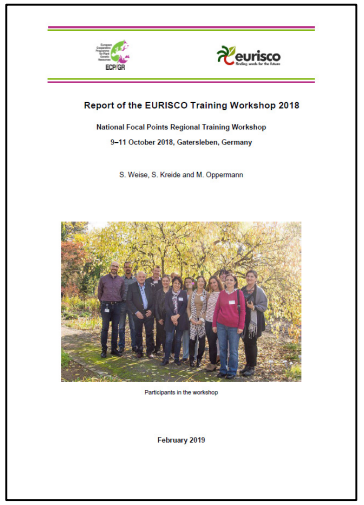


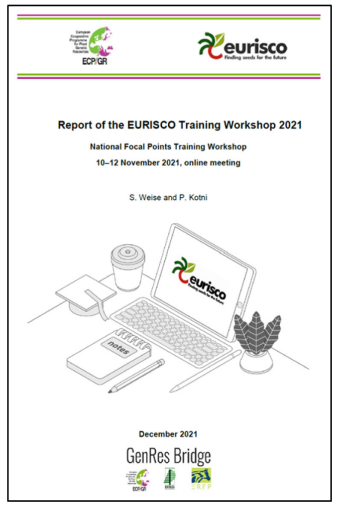
Network maintenance + development

- 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 ;-))**

EURISCO training workshops

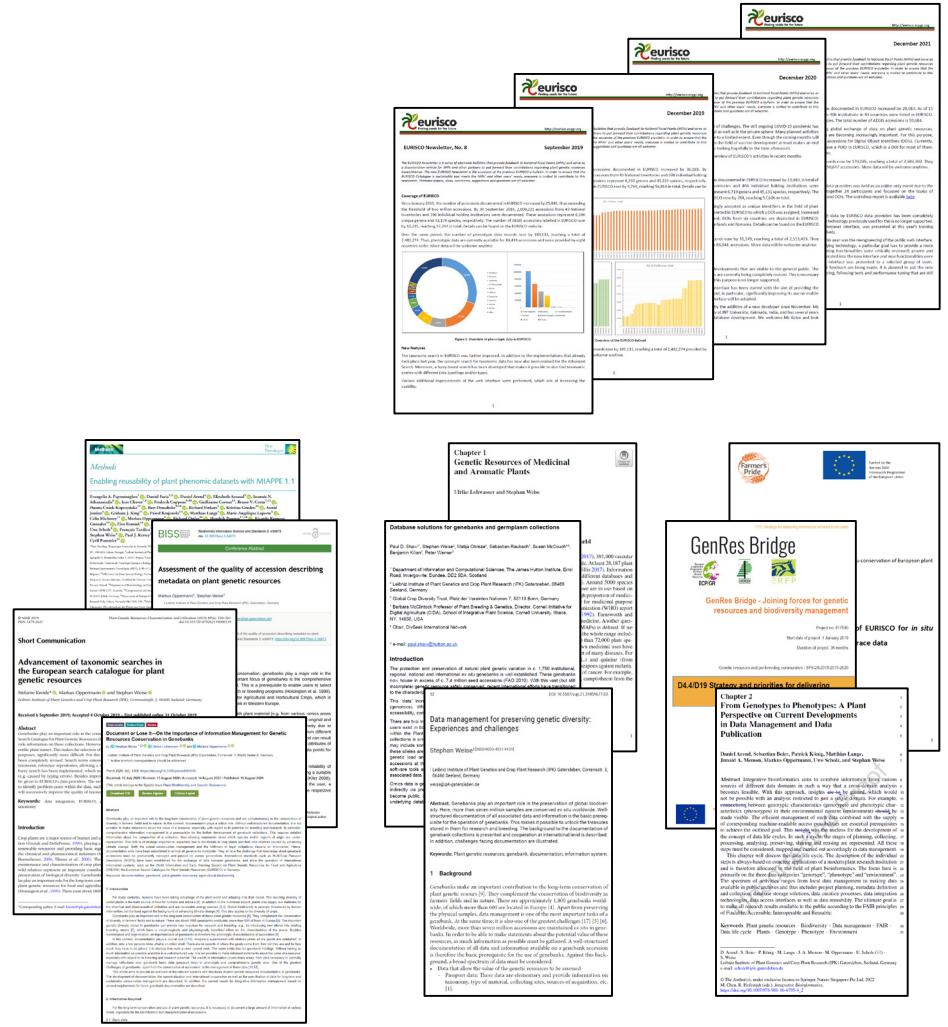
- Indispensable
  - Refreshing knowledge on data preparation and provision
  - Stay in touch with data providers
  - Discussion of changes and extensions
  - Continuous increase of data quality
- Switch to biennal trainings in 2018
- Additional online training on specific topics on request





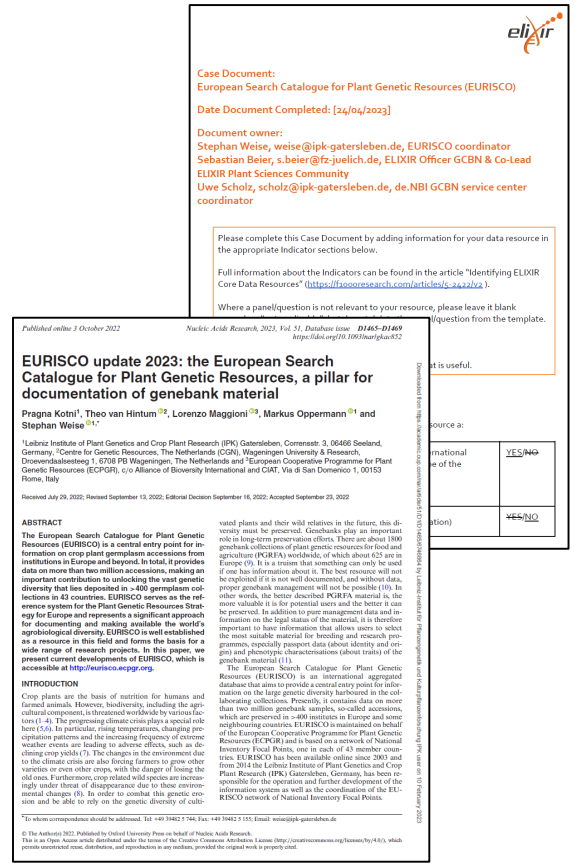
Dissemination in ECPGR context

- Regularly short information in ECPGR bulletin
- EURISCO newsletter
- Various reports
- Presentations on several ECPGR workshops
- ECPGR Grant Scheme activities



Dissemination beyond ECPGR

- Journal articles
- EURISCO talks and posters on several conferences
- Involvement in various committees
- Application as ELIXIR Core Data Resource



elixir

Case Document:
European Search Catalogue for Plant Genetic Resources (EURISCO)

Date Document Completed: [24/04/2023]

Document owner:
Stephan Weise, weise@ipk-gatersleben.de, EURISCO coordinator
Sebastian Beier, s.beier@fz-juelich.de, ELIXIR Officer GCBN & Co-Lead ELIXIR Plant Sciences Community
Uwe Scholz, scholz@ipk-gatersleben.de, de.NBI GCBN service center coordinator

Please complete this Case Document by adding information for your data resource in the appropriate Indicator sections below.

Full information about the Indicators can be found in the article "Identifying ELIXIR Core Data Resources" (https://f1000research.com/articles/5-2422/v2).

Where a panel/question is not relevant to your resource, please leave it blank

t is useful.

ource a:

rnational e of the	YES/~~NO~~
tion)	~~YES~~/NO

Published online 3 October 2022 *Nucleic Acids Research, 2023, Vol. 51, Database issue* ***D1465–D1469***
https://doi.org/10.1093/nar/gkac852

EURISCO update 2023: the European Search Catalogue for Plant Genetic Resources, a pillar for documentation of genebank material

Pragna Kotni[1], Theo van Hintum[2], Lorenzo Maggioni[3], Markus Oppermann[1] and Stephan Weise[1,*]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstr. 3, 06466 Seeland, Germany, [2]Centre for Genetic Resources, The Netherlands (CGN), Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands and [3]European Cooperative Programme for Plant Genetic Resources (ECPGR), c/o Alliance of Bioversity International and CIAT, Via di San Domenico 1, 00153 Rome, Italy

Received July 29, 2022; Revised September 13, 2022; Editorial Decision September 16, 2022; Accepted September 23, 2022

ABSTRACT

The European Search Catalogue for Plant Genetic Resources (EURISCO) is a central entry point for information on crop plant germplasm accessions from institutions in Europe and beyond. In total, it provides data on more than two million accessions, making an important contribution to unlocking the vast genetic diversity that lies deposited in ∼400 germplasm collections in 43 countries. EURISCO serves as the reference system for the Plant Genetic Resources Strategy for Europe and represents a significant approach for documenting and making available the world's agrobiological diversity. EURISCO is well established as a resource in this field and forms the basis for a wide range of research projects. In this paper, we present current developments of EURISCO, which is accessible at http://eurisco.ecpgr.org.

INTRODUCTION

Crop plants are the basis of nutrition for humans and farmed animals. However, biodiversity, including the agricultural component, is threatened worldwide by various factors (1–4). The progressing climate crisis plays a special role here (5,6). In particular, rising temperatures, changing precipitation patterns and the increasing frequency of extreme weather events are leading to adverse effects, such as declining crop yields (7). The changes in the environment due to the climate crisis are also forcing farmers to grow other varieties or even other crops, with the danger of losing the old ones. Furthermore, crop related wild species are increasingly under threat of disappearance due to these environmental changes (8). In order to combat this genetic erosion and be able to rely on the genetic diversity of cultivated plants and their wild relatives in the future, this diversity must be preserved. Genebanks play an important role in long-term preservation efforts. There are about 1800 genebank collections of plant genetic resources for food and agriculture (PGRFA) worldwide, of which about 625 are in Europe (9). It is a truism that something can only be used if one has information about it. The best resource will not be exploited if it is not well documented, and without data, proper genebank management will not be possible (10). In other words, the better described PGRFA material is, the more valuable it is for potential users and the better it can be preserved. In addition to pure management data and information on the legal status of the material, it is therefore important to have information that allows users to select the most suitable material for breeding and research programmes, especially passport data (about identity and origin) and phenotypic characterisations (about traits) of the genebank material (11).

The European Search Catalogue for Plant Genetic Resources (EURISCO) is an international aggregated database that aims to provide a central entry point for information on the large genetic diversity harboured in the collaborating collections. Presently, it contains data on more than two million genebank samples, so-called accessions, which are preserved in ∼400 institutes in Europe and some neighbouring countries. EURISCO is maintained on behalf of the European Cooperative Programme for Plant Genetic Resources (ECPGR) and is based on a network of National Inventory Focal Points, one in each of 43 member countries. EURISCO has been available online since 2003 and from 2014 the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany, has been responsible for the operation and further development of the information system as well as the coordination of the EURISCO network of National Inventory Focal Points.

*To whom correspondence should be addressed. Tel: +49 39482 5 744; Fax: +49 39482 5 155; Email: weise@ipk-gatersleben.de

© The Author(s) 2022. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Access statistics

Year	Unique users	# of hits	# of pages	# of accesses
2019	18,461	49,789	910,173	1,190,423
2020	18,045	59,882	1,315,558	1,579,331
2021	23,787	93,223	1,175,809	1,461,246
2022	31,809	168,541	4,441,965	4,809,474
	92,102	**371,435**	**7,843,505**	**9,040,474**

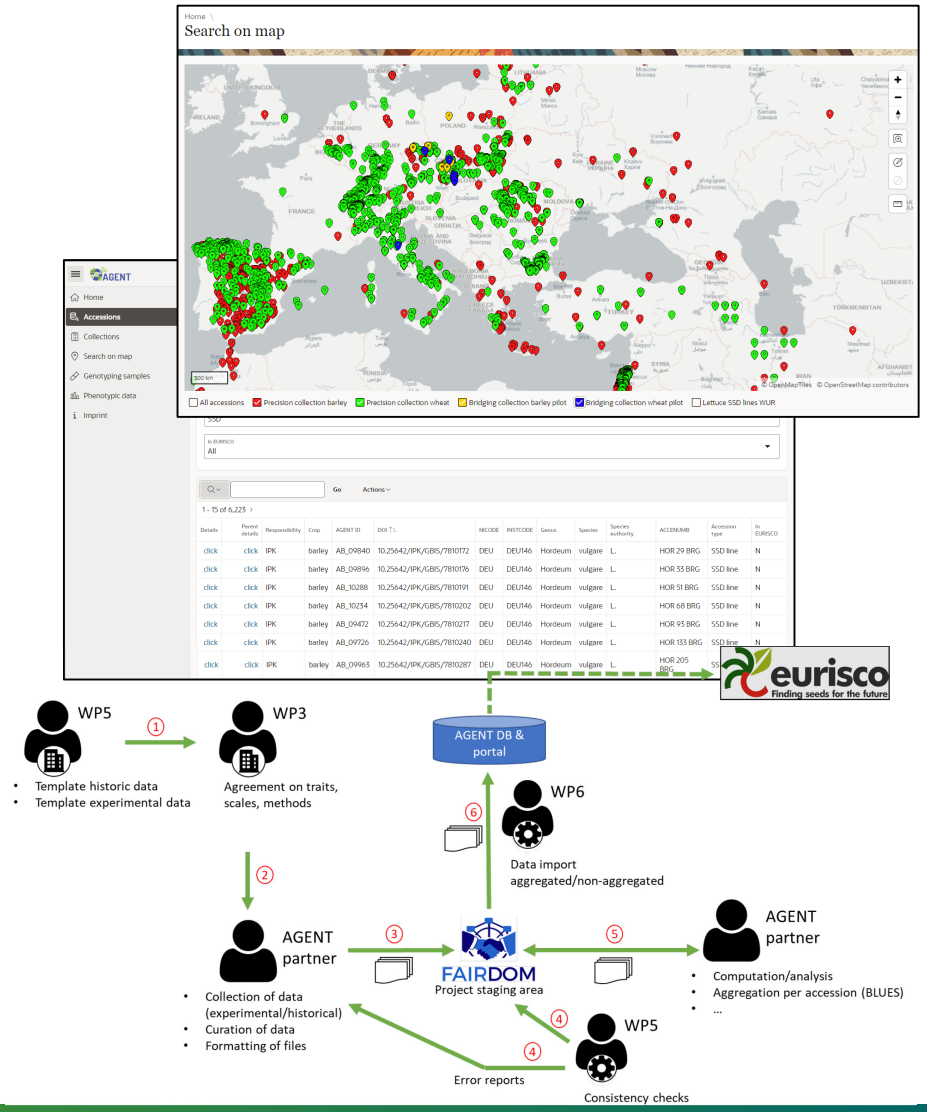
- Unique users:
  Number of distinct IP addresses (=physical users)
- # of hits:
  Every new user who accessed a page and did not access any page of the website within the last 60 minutes
- # of pages:
  Total number of accessed pages (sum of all accesses of all users; HTML pages only)
- # of accesses:
  Total number of all shown content (HTML pages, images, files, …)

Participation in project consortia

- Various ECPGR Grant Scheme Activities
- EUCLEG (Horizon 2020), 2017–2021
- Farmer’s Pride (Horizon 2020), 2018–2021
- GenRes Bridge (Horizon 2020), (2019–2021)
- ECPGR European Evaluation Network (initial funding BLE), 2019–2022

Current project focus – AGENT

- Activated GEnebank NeTwork (AGENT)
  - 2020 – 2025, funding: Horizon 2020 (Sustainable Food Security – Genetic resources and pre-breeding communities, H2020-SFS-2019-2)
  - Involvement in various work packages
    - Proposal of bridging collections
    - Guidelines for project data flow
    - Data exchange formats allowing MIAPPE compliance
    - Data management infrastructure and tools for discovery and visualisation
    - Responsibility for PGR data management



Current project focus – PRO-GRACE

PROMOTING A PLANT GENETIC RESOURCE COMMUNITY FOR EUROPE

- Promoting a Plant Genetic Resource Community for Europe (PRO-GRACE)
  - 2023 – 2025, funding: Horizon Europe (Research infrastructure concept development, HORIZON-INFRA-2022-DEV-01-01)
  - Lead of WP1 – Inventory and information system

- PRO-GRACE aims at developing
  - the concept,
  - regulatory framework and
  - governance

  for a functional and efficient European Research Infrastructure for plant genetic resources

Current/planned developments

- ECPGR phase XI (2024–2028)
  - Continuous improvement of functions and services
  - Specific focus on:
    - *In situ* CWR data
      - Extension of EURISCO started in Autumn 2022
    - Phenotypic data
      - Extend for more fine-grained metadata
      - Take up the idea of datasets (additionally)
      - Strengthen role as repository
    - Data quality (e.g. completeness, reliability) → continuous task
  - Further hosting of EVA
  - Participation in project consortia related to EURISCO (AGENT, PRO-GRACE, …)

  - All further development in close collaboration with ECPGR bodies