

EURISCO progress report

(April 2014 – April 2018)

15th Steering Committee meeting, 15–17 May 2018, Thessaloniki, Greece

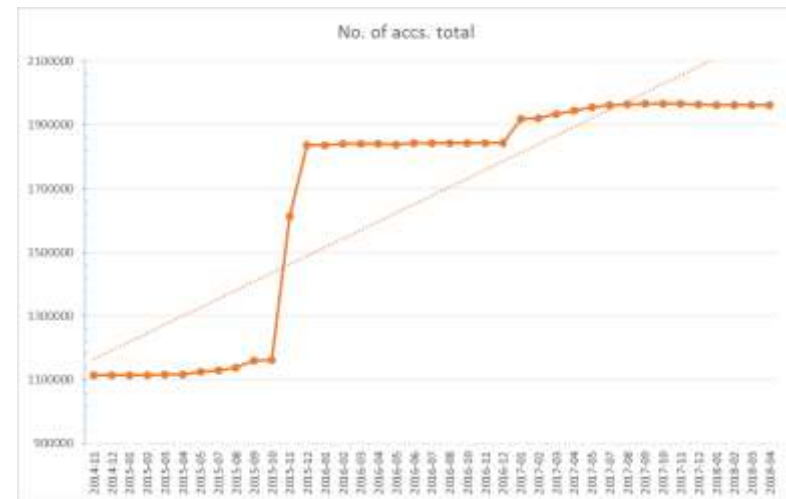
Stephan Weise



CURRENT STATUS OF EURISCO

Contents of EURISCO

- 1,961,985 accessions
- 6,317 genera
(including synonyms, spelling variants)
- 42,974 species names
(unique combinations genus + species, including synonyms)
- 418,679 MLS accessions
- 34,364 AEGIS accessions
- 22,906 DOIs



as at 2018-04-25

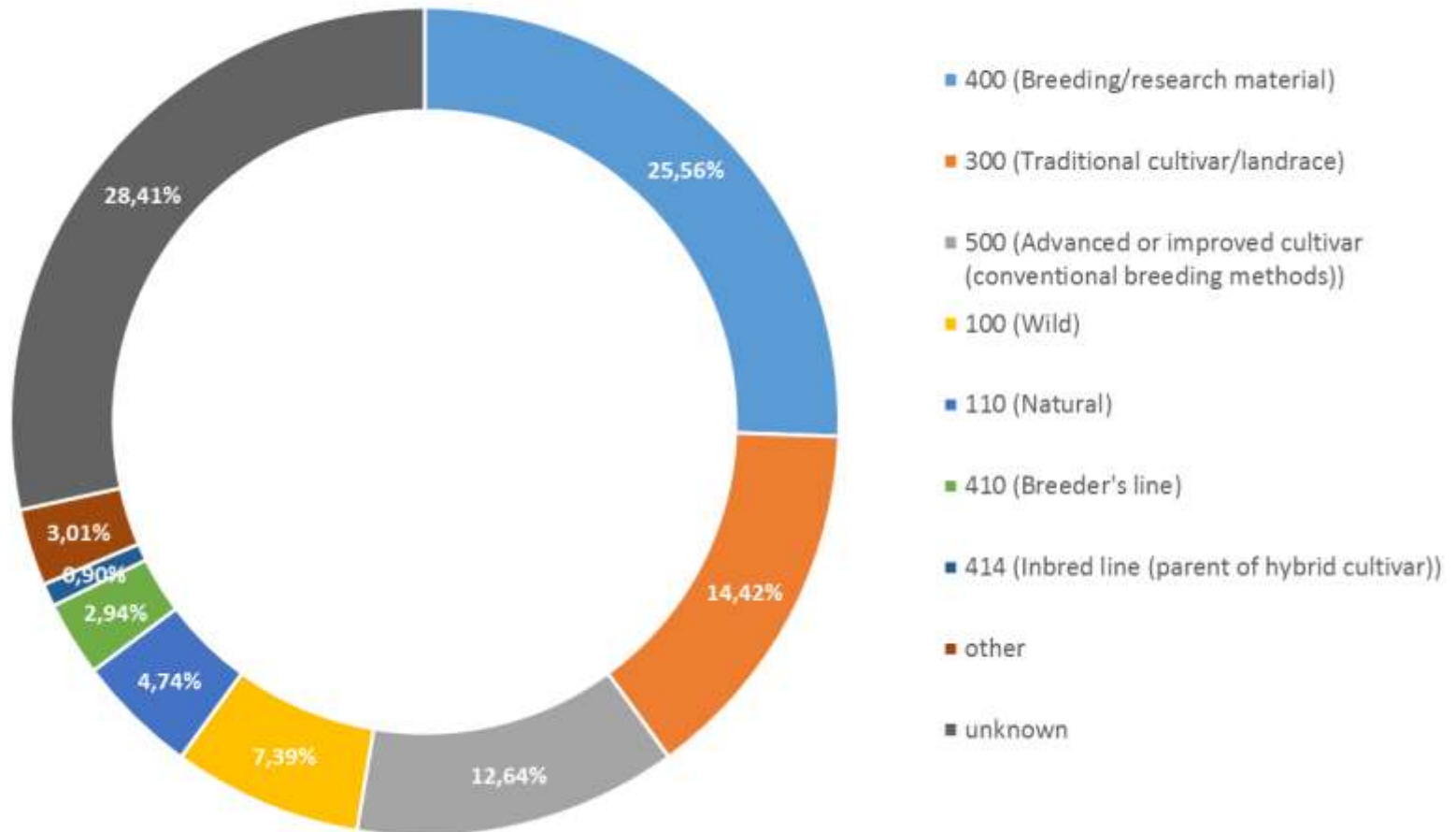
Taxonomic composition

Genus	Species	No. accs.	Total
<i>Arabidopsis</i>	<i>thaliana</i>	681,983	682,191
	others	208	
<i>Triticum</i> (wheat)	<i>aestivum</i>	136,778	189,104
	<i>durum</i>	16,312	
	<i>turgidum</i>	9,601	
	<i>monococcum</i>	3,483	
	<i>spelta</i>	3,202	
	others	19,728	
<i>Hordeum</i> (barley)	<i>vulgare</i>	104,374	121,589
	<i>spontaneum</i>	5,897	
	others	11,318	
<i>Zea</i> (maize)	<i>mays</i>	61,715	61,849
	others	134	
<i>Phaseolus</i> (garden bean)	<i>vulgaris</i>	46,621	52,179
	<i>coccineus</i>	3,064	
	others	2,494	

Genus	Species	No. accs.	Total
<i>Solanum</i> (tomato, potato, eggplant, etc.)	<i>lycopersicum</i>	19,152	51,265
	<i>tuberosum</i>	14,591	
	<i>andigenum</i>	2,814	
	<i>melongena</i>	2,119	
	others	12,589	
<i>Avena</i> (oat)	<i>sativa</i>	33,274	41,299
	<i>sterilis</i>	2,152	
	<i>byzantina</i>	1,968	
	others	3,905	
<i>Malus</i> (apple)	<i>domestica</i>	24,759	31,875
	others	7,116	
<i>Pisum</i> (pea)	<i>sativum</i>	27,417	30,455
	others	3,038	
<i>Vitis</i> (grape)	<i>vinifera</i>	26,096	30,049
	others	3,953	
others			670,130
	Total		1,961,985

as at 2018-04-25

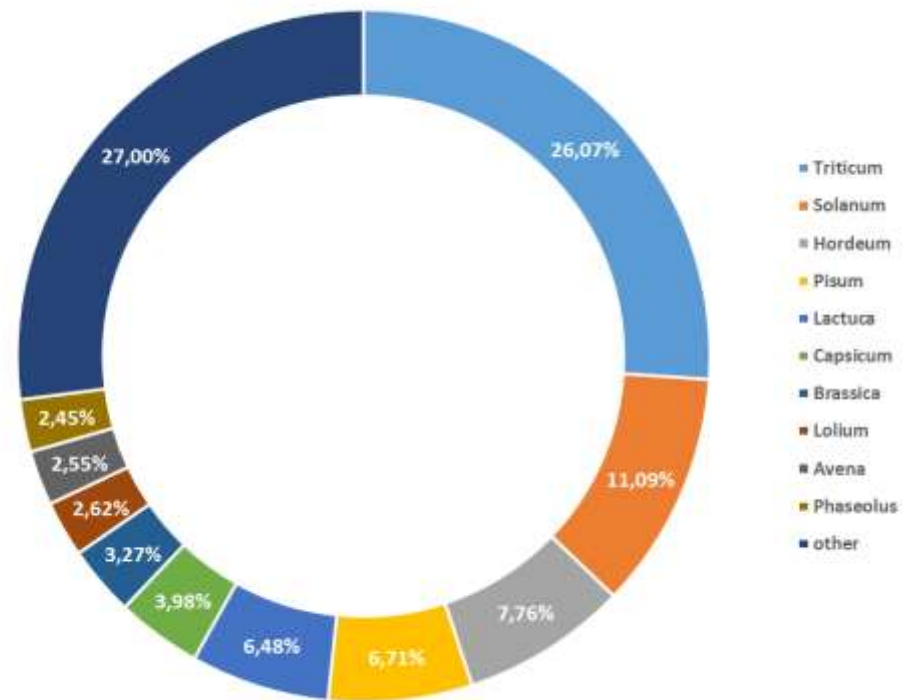
Biological status



as at 2018-04-25

C&E data

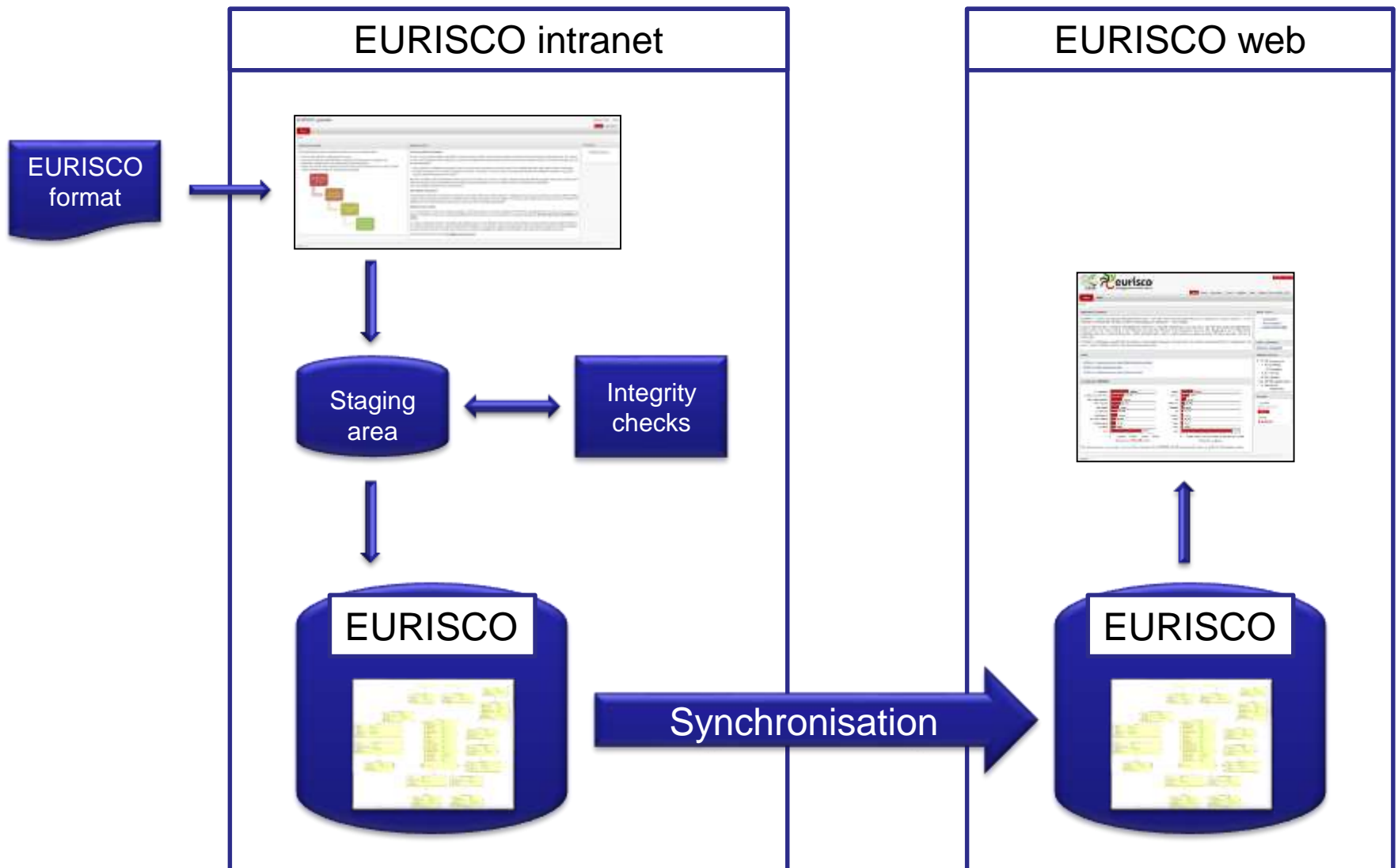
- Currently, 1,652,895 records of data from seven countries
 - Czech Republic
 - Germany
 - Latvia
 - The Netherlands
 - Poland
 - Romania
 - United Kingdom



- 68,821 accs. with C&E data

as at 2018-04-25

Architecture: Overview



Architecture: DB schema features

- EURISCO intranet
 - 59 tables
 - 502 indexes
 - 108 triggers
 - 14 PL/SQL packages
 - 177 function and procedures
 - data upload and import
 - integrity checks
 - updates (passport and C&E)
 - taxonomy support
 - 27 Java classes
- EURISCO web
 - 46 tables
 - 29 materialised views
 - 698 indexes
 - 9 PL/SQL packages
 - 38 function and procedures
 - download
 - newsletter
 - statistics
 - C&E data visualisation
 - AEGIS status auditing
 - taxonomy support
 - In-memory features
 - 4 Java classes

Web interface



Published online 31 August 2016

Nucleic Acids Research, 2017, Vol. 45, Database issue D1003–D1008
doi: 10.1093/nar/gkx755

EURISCO: The European search catalogue for plant genetic resources

Stephan Weise^{1,*}, Markus Oppermann¹, Lorenzo Maggioni², Theo van Hintum³ and Helmut Knüpffer¹

¹Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstr. 3, 06466 Stadt Seeland, Germany, ²Biodiversity International, Via del Tre Denari 472/a, 00057 Maccarese (Fiumicino), Rome, Italy and ³Centre 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 675 collections are maintained in Europe

data of The Netherlands updated
data of Switzerland updated - New AEGIS accessions
data of the Czech Republic updated - New C&E data



Get an overview of the composition of the EURISCO data. More detailed information can be found at the [statistics section](#)

Search EURISCO

- Quick search
- Advanced search
- Export EURISCO data
- C&E data

EURISCO newsletter

Subscribe / unsubscribe

Regions of origin

Statistical overview

1,961,965 Accessions
372 Institutes
43 Countries
6,317 Genera
42,974 Species
418,679 MLS accessions
34,364 AEGIS accessions
22,906 DOIs

Site rating

Your rating:

Submit

Average:

★★★★★

39 (sub)versions since 2014

Passport data in EURISCO

- Four standard searches:
 - Taxonomy
 - Accession
 - Biological status
 - Collecting site
- Advanced search
- Different user-specific export features

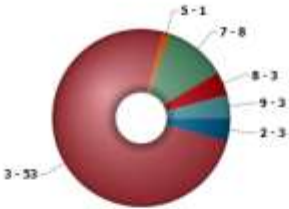
The screenshot displays the EURISCO website interface. At the top, the EURISCO logo is visible with the tagline "Finding seeds for the future". Below the logo, there are navigation tabs for "Home", "About", "Search", "Data sets", "Statistics and Downloads", "Support", and "Data Protection Policy". A search bar is located at the top left, and a "Search" button is next to it. The main content area is divided into several sections:

- National inventory:** A section for national inventory data.
- Holding entities:** A section for holding entities, including a note about accessions and contact information for the EURISCO database.
- Accession:** A section for accession data, showing the Institute Code (FRT091) and Institute Name (Portuguese Bank of Plant Germplasm, Braga, Portugal).
- Taxonomy:** A section for taxonomy data, showing the Genus (*Brassica*), Species (*oleracea*), Species Authority (*L.*), Subtaxa (*var. capitata*), and Subtaxa Authority (*DC.*).
- Collecting site:** A section for collecting site data, showing the Collecting Number (350814 A), Collecting Institute Code (FRT091), Collecting Date (2014-03-25), Collecting Latitude (46.330611), Collecting Longitude (-7.130556), Collecting Elevation (872), and Collecting Site (Portugal, Guarda).
- Map:** A map showing the location of the collecting site in Portugal, Guarda.

On the right side of the page, there is a sidebar with a "Search EURISCO" section, including options for "Quick search", "Advanced search", "Export full record data", "CSD data", "Download passport", and "CSD data". Below this, there is a "Last update" section showing the last update of accession records as 2017-05-29.

C&E data in EURISCO

Trait details



Descriptive statistics

Trait Name	Minimum	Maximum	Average	Stdev	Variance	First Quartile	Median	Third Quartile
Fruit altitude	2	8	3.8	1.93	3.72	3	3	3

Experiment description: Sowing date = February 2, Planting date = April 17, I/VT glasshouse XII, heated, soil culture, 2 stems, 4 plants per field, collection no. 567-658, experimental H. Roolofsen and G. Pet, standard = Bruinsma Wonder

Trait name: Fruit altitude

Additional filters

Genus:

Origin Country:

Accession scores for selected trait

Rows 10

1 - 10 of 71 >

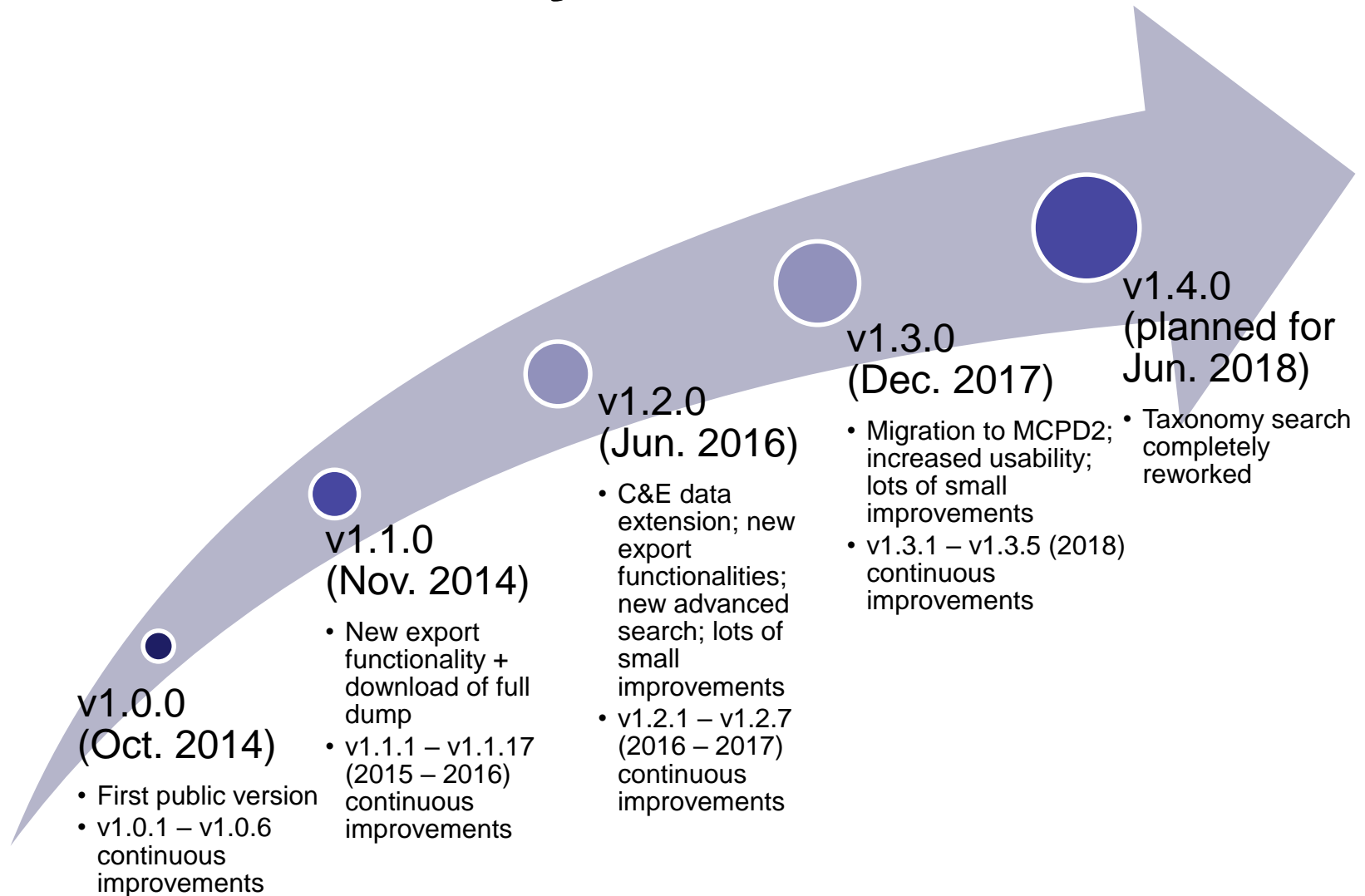
NICODE	INSTCODE	GENUS	ACCENUMB	Score	Score Link	Origin Country	Biological Status	Details
NLD	NLD037	Capsicum	CGN16913	3	-	Germany	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16914	3	-	Netherlands	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16916	8	-	Israel	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16917	7	-	Israel	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16918	3	-	Hungary	Traditional cultivar/landrace	Accession details
NLD	NLD037	Capsicum	CGN16919	3	-	Hungary	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16920	7	-	-	-	Accession details
NLD	NLD037	Capsicum	CGN16904	9	-	-	-	Accession details
NLD	NLD037	Capsicum	CGN16905	3	-	-	Traditional cultivar/landrace	Accession details
NLD	NLD037	Capsicum	CGN16906	3	-	-	Advanced or improved cultivar (conventional breeding methods)	Accession details

1 - 10 of 71 >

0.02 s

- Wizard-based searches for
- Genus
 - Species and trait
 - Experiment
 - Trait

Version history of the web interface



NEW HOSTING OF EURISCO

EURISCO transfer

- Preparation of transfer (2013)
 - Preparatory meetings
 - High costs for transfer of as-is status
→ New development from scratch
 - Analysis of the former web application
- System reengineering and data migration
 - Frontend reengineering (public EURISCO application)
 - Database structures and public web application
 - PL/SQL for functionality; APEX for rendering
 - Backend reengineering (EURISCO intranet)
 - Database structures and intranet web application
 - PL/SQL for functionality; APEX for rendering
 - Import (and cleansing) of current data set
 - Migration path MySQL → Oracle RDBMS

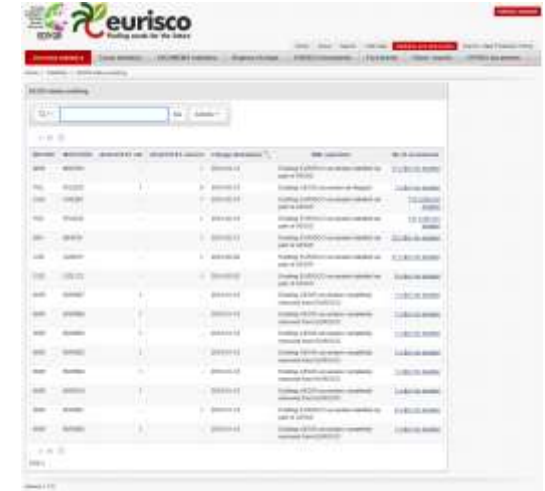


Outcomes 2014

- EURISCO intranet
 - EURISCO transfer (main task)
 - Switch to new upload mechanism (incremental updates)
- Public EURISCO application
 - EURISCO transfer (main task)
 - New export feature with customised export files in CSV format

Outcomes 2015

- EURISCO intranet
 - Java-based passport importer developed
 - Data quality/integrity checks improved
 - AEGIS status tracking
- Public EURISCO application
 - Statistics section with various reports
 - Improved search functionality (additional fields/filters)
 - Adjustments “under the hood”
- Extension for C&E data
 - Data exchange format
 - Database schema
 - Upload and check mechanism
 - Public web interface (prototype; continued in 2016)



The screenshot displays the EURISCO public application interface. At the top, the EURISCO logo is visible with the tagline "Finding seeds for the future". Below the logo, there is a navigation menu and a search bar. The main content area features a table with multiple columns, likely representing different data points or records. The table has a header row and several data rows, each with a unique identifier and associated information. The interface is clean and professional, typical of a scientific data management system.

Outcomes 2016

- EURISCO intranet
 - Interface extended for importing C&E data
 - AEGIS support as well as quality/integrity checks improved
- Public EURISCO application
 - Additional export functionalities implemented
 - Advanced search completely reworked
 - Extended performance tuning
- Extension for C&E data
 - Public web interface extended for C&E and tested
 - Additional features implemented (based on users' feedback)
 - Productive data imported
 - Official release of C&E extension

Outcomes 2017

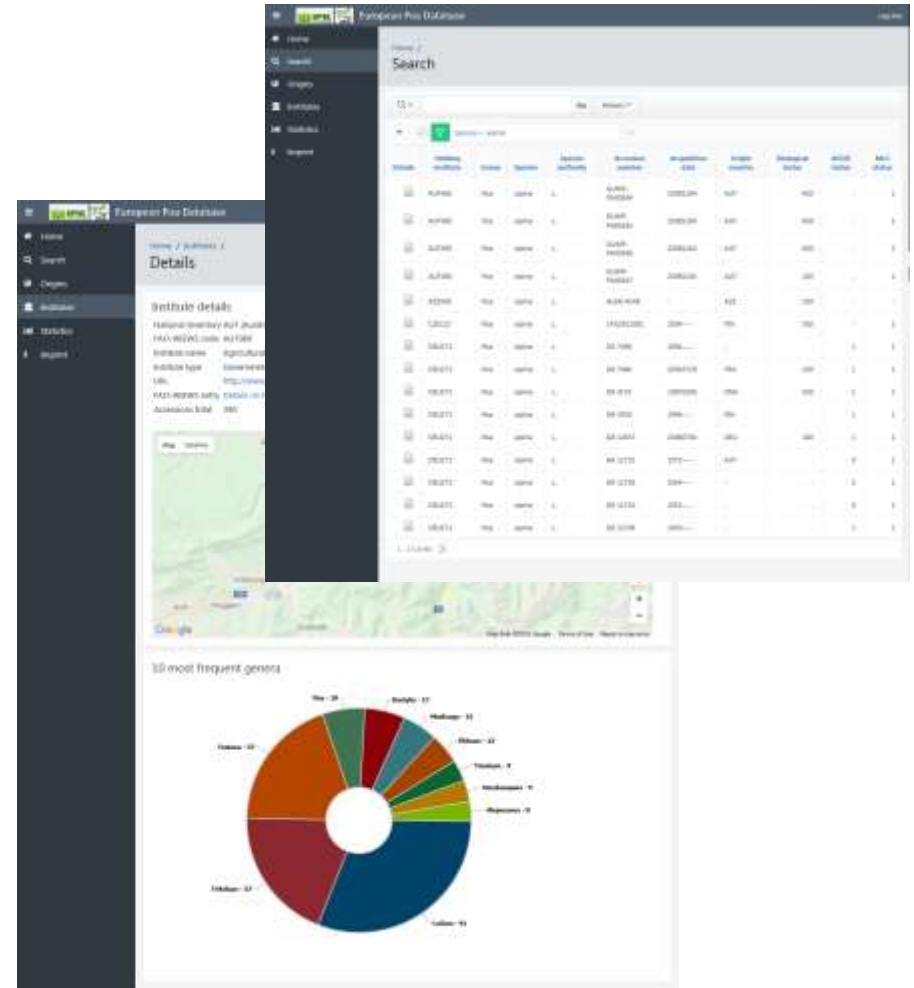
- EURISCO intranet
 - AEGIS support improved further
 - Improvement of taxonomy support started (continued in 2018)
 - Database structures and mapping procedures
 - How-to-documents (handbooks) compiled
 - Implementation of MCPD2 for EURISCO
 - Database structures, upload and check procedures
- Public EURISCO application
 - User survey conducted and requirements for improvement collected
→ addressed all NCs/NFPs; received useful suggestions
- Extension for C&E data
 - Additional flexible procedure for uploading C&E data compiled and implemented → delegation and final approval by NFPs

Outcomes 2018 (ongoing)

- EURISCO intranet
 - Improvement of taxonomy support (started in 2017)
- Public EURISCO application
 - Implementation of requirements from user survey
 - Extension for SSR data (under evaluation)
- Extension for C&E data
 - Improvement of search and download features

Support of ECPGR Central Crop Databases

- Cooperation with various crop WGs
- Example: European *Poa* Database (Forage WG)
- poa.ipk-gatersleben.de (hosted at IPK since the 1990s)
- Blueprint for a crop portal automatically updated from EURISCO



EURISCO COORDINATION

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
- 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 ;-)

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 (as at 2018-04-25)	12	1,961,985

EURISCO training workshops

- Regular training for data providers (esp. for C&E data and in case of staff changes)
- Training materials + test environment developed + meeting reports
- Annual workshops:
 - 2015 Tirana, Albania (Agricultural University of Tirana)
 - 20 participants
 - 2016 Angers, France (INRA)
 - 13 participants
 - Meeting of the EURISCO Advisory Committee
 - 2017 Gatersleben, Germany (IPK)
 - 14 participants
- 2018 training under preparation

Participation in project proposals

- Involved in 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
- Participation in the proposal “GenRes Bridge – Joining forces for genetic resources and biodiversity management”

Dissemination

- Regularly short information in ECPGR bulletin
- EURISCO newsletter twice a year
- Journal article
- EURISCO talks and posters at several conferences
- Presentations at several ECPGR workshops
- ECPGR Grant Scheme Activities

FUTURE OF EURISCO

EURISCO: Plans & challenges

- ECPGR Phase X (2019–2023)
 - Further development in close collaboration with ECPGR bodies
 - Specific focus on:
 - Phenotypic data
 - Harmonisation of traits
 - Implementation of additional use-cases
 - *In situ* data
 - Extension of EURISCO
 - Data quality
 - Completeness
 - Reliability, e.g. geographic coordinates
 - Continuous improvement of functions and services

**THANK YOU FOR YOUR
ATTENTION**