EURISCO, the European Search Portal for Plant Genetic Resources

Status quo & planned developments

EURISCO training workshop, 12th to 14th October 2016, Angers, France
BACKGROUND
Introduction I

• Development of European information system for plant genetic resources
  – Started in 2001 (EU project EPGRIS)
  – EURISCO: Search catalogue for *ex situ* collections; available online since 2003
  – National collections represented by National Inventories (NIs)
  – 43 countries involved (Nordic Countries → NordGen)
  – Network of National Focal Points (NFPs) links NIs ↔ EURISCO
Introduction II

- Development of European information system for plant genetic resources (cont.)
  - Accession-level information system
  - Provides passport information about PGR 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)
Data flow

EURISCO

ex situ PGR data
Contents of EURISCO

• 1,842,760 accessions
• 376 holding institutes
• 43 countries
• 6,233 genera (including synonyms, spelling variants)
• 41,649 species names (unique combinations genus + species, including synonyms)
• 368,446 MLS accessions
• 28,899 AEGIS accessions

as of 2016-09-07
Genera

- Arabidopsis: 36%
- Triticum: 8%
- Hordeum: 6%
- Zea: 3%
- Solanum: 2%
- Phaseolus: 2%
- Pisum: 2%
- Avena: 2%
- Vitis: 1%
- Other: 37%

as of 2015-12-07
Biological status as of 2015-12-07
Regions of origin

as of 2016-02-10
Data quality in EURISCO

• 587,391 accessions with collecting information
  – 91,808 different collecting sites
  – But only 188,456 accessions with coordinates (10% of all accessions; 32% of accessions with collecting information)

• 1,161,631 accessions with donor information

• 960,334 accessions with country of origin
  – 24 different geographic regions

as of 2016-09-07
EURISCO HOSTED BY IPK
IPK became host

• Since 2003:
  – EURISCO hosted by Bioversity International, Rome

• October 2012:
  – Request for tenders for hosting EURISCO

• March 2013:
  – IPK won the bid

• May and October 2013
  – Preparatory meetings in Gatersleben and Rome

• Since 15th April 2014:
  – Sub-contract between Bioversity International and IPK
  – Transfer of the EURISCO responsibilities to IPK (04-09/2014)
Challenges and decisions

• Challenges faced:
  – Outdated systems
  – Insufficient technical documentation
  – High costs for transfer of as-is status

• Decisions:
  → No transfer of existing solutions
  → New development from scratch
Technical tasks

• Analysis of the former web application

• Reengineering of database schema for web application

• Import (and cleansing) of current data set
  – Migration path MySQL → Oracle RDBMS

• New web application for searching EURISCO data
  – PL/SQL for functionality; APEX for rendering
  – First release with basic functionality published 09/2014
  – Continuously improved

• New web application for updating NI data
  – PL/SQL for functionality; APEX for rendering
New architecture: Overview

- EURISCO intranet
  - EURISCO format
  - Staging area
  - Integrity checks
  - EURISCO

- EURISCO web
  - EURISCO
  - Synchronisation
New architecture: DB schema features

• EURISCO intranet
  – 49 tables
  – 460 indexes
  – 90 triggers
  – 9 PL/SQL packages
  – 110 functions and procedures
  – 27 Java classes

• EURISCO web
  – 45 tables
  – 25 materialised views
  – 585 indexes
  – 6 PL/SQL packages
  – 25 functions and procedures
  – In-memory features
EURISCO WEB
New web application
Four standard searches
- Taxonomy
- Accession
- Biological status
- Collecting site
Search results
**Accession details**

<table>
<thead>
<tr>
<th><strong>National inventory</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>National Inventory Code: DEU</td>
</tr>
<tr>
<td>National Inventory: Germany</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Holding institute</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Institute Code: DEU146</td>
</tr>
<tr>
<td>Institute Name: Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Accession</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accession Number: TRI 10454</td>
</tr>
<tr>
<td>Accession Names: White Kandahar (land race)</td>
</tr>
<tr>
<td>Country Of Origin: Afghanistan</td>
</tr>
<tr>
<td>MLS Status: part of the MLS</td>
</tr>
<tr>
<td>AEGIS Status: not part of the AEGIS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Taxonomy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genus: Triticum</td>
</tr>
<tr>
<td>Species: sp.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Acquisition/storage</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition Date: 1973</td>
</tr>
<tr>
<td>Acquisition Source: Other</td>
</tr>
<tr>
<td>Germplasm Storage: Long term</td>
</tr>
</tbody>
</table>

...
Advanced search form

- **Genus**: TRITICUM
- **Species**: 
  - DICOCOIDES
  - MONOCOCCUM
  - SPELTA
  - DURUM
- **Status**: 100 (Wild)
  - 110 (Natural)
  - 120 (Semi-natural/wild)
  - 130 (Semi natural/town)
- **Weedy**: 200 (Weedy)
- **Traditional cultivar/landrace**: 300 (Traditional cultivar/landrace)
- **Breeding/research material**
User specific export – by species
User specific export – by NI
EURISCO INTRANET
EURISCO intranet

• Upload component for National Inventory Focal Points

• Standardised data exchange formats
  – Passport data
  – Phenotypic (characterisation and evaluation) data

• Details will follow during the hands-on sessions
CHARACTERISATION & EVALUATION DATA
C&E data: Extending EURISCO

• EURISCO backend completely extended for C&E data
  – Exchange format
  – Database schema
  – Upload tools
  – Data integrity checks

• Extension of EURISCO web interface for C&E data – first version available
  – Wizard-based searches for
    • Genus
    • Species and trait
    • Experiment
    • Trait
C&E data: Genus
C&E data: Species and trait
C&E data: Experiment

Filter C&E data by experiment

The report below lists all experiments, which contain characterisation & evaluation (C&E) data. Please use the search bar below to define filters.

Experiment Description

<table>
<thead>
<tr>
<th>Sowing date</th>
<th>Planting date</th>
<th>IVT glasshouse</th>
<th>Soil culture</th>
<th>Sowing, 5 plants per field, collection no. 567-869, experiment 1</th>
<th>Test data CON</th>
<th>Data remark</th>
<th>Uploaded</th>
<th>Uploaded AI</th>
<th>Experiment description</th>
<th>Details</th>
<th>Experiment Start Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>February 2</td>
<td>April 17</td>
<td>XII, heated,</td>
<td>soil culture, 2 stems, 5 plants per field, collection no. 765-869,</td>
<td>experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Test data CON</td>
<td>Data remark</td>
<td>Uploaded</td>
<td>Uploaded AI</td>
<td>Sowing date = February 24, planting date = April 17, IVT glasshouse no. XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 765-869, experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Details</td>
<td>Experiment Start Year</td>
</tr>
<tr>
<td>February 16</td>
<td>April 6</td>
<td>XII, heated,</td>
<td>soil culture, 2 stems, 5 plants per field, collection no. 444-443,</td>
<td>experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Test data CON</td>
<td>Data remark</td>
<td>Uploaded</td>
<td>Uploaded AI</td>
<td>Sowing date = February 16, planting date = April 6, IVT glasshouse XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 444-443, experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Details</td>
<td>Experiment Start Year</td>
</tr>
<tr>
<td>March 15</td>
<td>April 6</td>
<td>XII, heated,</td>
<td>soil culture, 2 stems, 5 plants per field, collection no. 660-762,</td>
<td>experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Test data CON</td>
<td>Data remark</td>
<td>Uploaded</td>
<td>Uploaded AI</td>
<td>Sowing date = March 15, planting date = April 6, IVT glasshouse XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 660-762, experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Details</td>
<td>Experiment Start Year</td>
</tr>
<tr>
<td>February 28</td>
<td>April 13</td>
<td>VIII, IX,</td>
<td>soil culture, 2 stems, 5 plants per field, collection no. 765-869,</td>
<td>experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Test data CON</td>
<td>Data remark</td>
<td>Uploaded</td>
<td>Uploaded AI</td>
<td>Sowing date = February 28, planting date = April 13, IVT glasshouse VIII, IX, heated, soil culture, 2 stems, 5 plants per field, collection no. 765-869, experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Details</td>
<td>Experiment Start Year</td>
</tr>
<tr>
<td>March 11</td>
<td>April 6</td>
<td>XII, heated,</td>
<td>soil culture, 2 stems, 5 plants per field, collection no. 871-984,</td>
<td>experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Test data CON</td>
<td>Data remark</td>
<td>Uploaded</td>
<td>Uploaded AI</td>
<td>Sowing date = March 11, planting date = April 6, IVT glasshouse XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 871-984, experimentals L. de Groot and G. Pet, standard = Brunna</td>
<td>Details</td>
<td>Experiment Start Year</td>
</tr>
</tbody>
</table>

Traits in selected experiment

<table>
<thead>
<tr>
<th>Trait Name</th>
<th>Trait Remark</th>
<th>Trait Method</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fruit corona</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit diameter</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit growth</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Flower growth</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Flower height</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Flower number</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Flower size</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Flower color</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit color</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit cracking</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit maturity</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit number</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit size</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit thickness</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit weight</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Fruit yield</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Leaf number</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Leaf size</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Leaf shape</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Leaf texture</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Leaf color</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Plant height</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Plant number</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Plant size</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Plant shape</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Plant texture</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Plant color</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Root number</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Root size</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Root thickness</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Root weight</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Root yield</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Stem number</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Stem size</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Stem shape</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Stem texture</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
<tr>
<td>Stem color</td>
<td></td>
<td></td>
<td>scores</td>
</tr>
</tbody>
</table>
## C&E data: Trait

The report below lists the definitions of all phenotypic traits, which are currently available in EURISCO. Please use the search bar below to define filters.

<table>
<thead>
<tr>
<th>Trait Name</th>
<th>Trait Remark</th>
<th>Trait Method</th>
<th>Trait Group</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petiole and/or midvein enlargement</td>
<td>3=narrow, 5=intermediate, 7=enlarged (See IPGRI descriptors Brassica and Raphanus 1990 4.2.27)</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Silique angle</td>
<td>1=0°; 2=0-22.5°; 3=22.5°; 4=22.5-45°; 5=45°; 6=45-67.5°; 7=67.5°</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Leaf anthocyanin content</td>
<td>1 = very weak, 3 = weak, 5 = medium, 7 = strong, 9 = very strong</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Timo period until marketable maturity</td>
<td>1! = very early, ..., 9! = very late</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Bremia lactucae</td>
<td>Resistance to B. lactucae race BI 5, scale:1 = susceptible, 9 = resistant</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Bremia lactucae</td>
<td>Resistance to B. lactucae race BI 11, scale:1 = susceptible, 9 = resistant</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Bremia lactucae</td>
<td>Resistance to B. lactucae race BI 12, scale:1 = susceptible, 9 = resistant</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Bremia lactucae</td>
<td>Resistance to B. lactucae race BI 14, scale:1 = susceptible, 9 = resistant</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Leaf anthocyanin distribution</td>
<td>1 = localised, 2 = entire</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
<tr>
<td>Bremia lactucae</td>
<td>Resistance to B. lactucae race BI 25, scale:1 = susceptible, 9 = resistant</td>
<td></td>
<td>C&amp;E data (not further specified)</td>
<td>used by experiment(s)</td>
</tr>
</tbody>
</table>
C&E data: C&E scores
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
- 28,899 AEGIS accessions
- 34 member countries

as of 2016-09-08
FUTURE OF EURISCO
More data

• Data quantity
  – Inclusion of additional passport data
    • ~250,000 accessions (van Hintum, 2014)

• Data quality
  – Increase frequency of updates
    • Avg. age 1.16 years
    • Oldest 10% avg. 4.96 years
    (van Hintum, 2014)
Better data

• Data quality
  – Improve taxonomic backbone of EURISCO
    • Management of taxon synonyms
    • Improvement of checks during import
      – GRIN, Catalogue of Life webservice
  – Increase completeness of information
    • Often limited information about certain accessions
    • Some descriptors only sparsely populated
      – SPECIES: 97%
      – BREDCODE: 40%
      – ELEVATION: 11%
Challenges: Sea food?

- **Genera**: Medicago
- **Species**: murex

**Acquisition/storage**
- **Acquisition Source**: Roadside

**Collection**
- **Collecting Institute Code**: [field]
- **Collecting Date**: 1987
- **Collecting Latitude**: 41.4254
- **Collecting Longitude**: -9.0831
- **Collecting Elevation**: 635

Map with a location marker at coordinates 41.4254, -9.0831.
Challenges: Inconsistencies
Challenges: Babylonian aspects
Increase functionality

• Improve import mechanism
  – Nested user accounts
    • Data upload also by holding institutes
    • Check and approval by NFPs

• Improve public web interface
  – Additional reports and download facilities
  – Extend filter possibilities by additional fields
  – Enable case insensitive search for taxa (for standard searches)
  – Provide a full-text search
Extend for *in situ* data

- Will be documented in EURISCO
- Sufficient specification needed
- Data exchange formats need to be agreed
Additional means of access

• Web services
  – Additional means of access
    • Data exchange with Genesys, GBIF etc.
  – Improvement of upload mechanism
    • Selective updates of certain accessions

• Mobile version
Dissemination

- Presentations on several workshops
- Journal article
  - Nucleic Acids Research (database issue), DOI: 10.1093/nar/gkw755
- EURISCO posters on several conferences
- Regularly short information in ECPGR bulletin
- EURISCO newsletter twice a year
- Preparation of project proposals (H2020 and others)
- In preparation: Update of fact sheets
Acknowledgements

• Bioversity International
  – Sónia Dias
  – Simone Mori
  – Milko Škofič
  – Dario Valori

• IPK Gatersleben
  – Andreas Graner
  – Helmut Knüpffer
  – Markus Oppermann
  – Uwe Scholz

• Global Crop Diversity Trust
  – Matija Obreza

• ECPGR
  – Theo van Hintum (Doc&Info group)
  – Lorenzo Maggioni, Elinor Lipman, Lidwina Koop, Jan Engels (ECPGR Secretariat)
  – ECPGR Executive Committee
THANK YOU FOR YOUR ATTENTION