Central Crop Databases: status at the beginning of ECPGR Phase IX

Synthesis of a survey conducted among Database Managers

E. Lipman
ECPGR Secretariat

Tailoring the Documentation of Plant Genetic Resources in Europe to the Needs of the User
Workshop of the Documentation and Information Working Group, 20-22 May 2014, Prague, Czech Republic

Contents

- The survey: background, results obtained and methodology for the analysis
- Presentation of the databases
- Summary / Synthesis of the survey

Background of the survey

- Thorough **check of the availability status of the Central Crop Databases (CCDBs)** listed on the ECPGR website (Sept.-Dec. 2013)
  Numerous corrections and updates in collaboration with DB managers

  → **The existence of a total of 48 CCDBs was confirmed**

- In preparation of this Workshop, **survey** including two complementary approaches:

- **Comparison of the data in EURISCO and the respective CCDBs** *(see presentation by Theo van Hintum)*

- **Mailing of questionnaires to all DB managers** (Dec. 2013-Jan. 2014). The questionnaires were pre-filled with data easily available from the respective websites.

The questionnaire - presentation

Structured in two main parts

1. Table form (“Facts and figures”)
   IDENTIFICATION of the DB
   CONTENTS AND STRUCTURE - *no. of accessions, date of last update, average age of records, contributors, geographical coverage, details on passport data and C&E data (descriptors used, accessibility)*
   OTHER INFORMATION provided on the DB’s website

2. Open questions
   MANAGEMENT
   - How much time (average number of hours per year over the last three years) is invested in the database?
   - Is your mechanism of data gathering and updating working well / sustainable?

   ASSESSMENT OF USE AND VALUE
   - What do the databases add to the data available in EURISCO?
   - Do you have an idea about the use of your database and the appreciation of the users?

   PERSPECTIVES
   - Are you planning to continue maintaining the database in the next five years and, if positive, are sufficient resources available to do that at your institute?

   ANY OTHER COMMENTS

Results

- **Good response level: 41 completed questionnaires received out of 48 (= 85%)**

- For one DB (*Glycine),* a reply was received with an incomplete questionnaire. The DB Manager was unable to participate in the survey and indicated the need to identify another host for the *Glycine* CCDB.

- For six DBs (Chicory, *Cicer*, *Lens*, Pepper, Minor Forage Legumes and *Pisum*) the questionnaires were not returned, despite repeated reminders.

Methodology for the analysis

- **Replies were reviewed** in collaboration with DB Managers (January-April 2014).

  This allowed many further updates and corrections, e.g.
  - changes in URLs
  - changes in DB Managers
  - updates of 4 DBs *(see later, number of accessions)*

- **Revised questionnaires were compiled** in a global file made available online

  Note: some errors or imprecisions were corrected during the review process, but it was not possible to check in depth all information provided. If some inaccuracies remain, responsibility lies with the authors of the replies.

→ **Analytical summary** prepared by compiler

Note: whenever possible (reliable) factual data for "non-answering" DBs were integrated - therefore the total number of DBs for which data are available varies with items.

ECP/GR

Central Crop Databases:
status at the beginning of ECPGR Phase IX

Foreword

A thorough check of the availability status of the Central Crop Databases (CCDBs) listed on the ECPGR Web pages was carried out from September to December 2013. The existence of a total of 48 CCDBs was confirmed.

In preparation of the Documentation and Information Workshop to be held in May 2014, a survey was undertaken to gather detailed data about all CCDBs, including two complementary approaches:

- A comparison of the data in EURISCO and the respective Central Crop Databases (CCDBs)
- The mailing of questionnaires to all DB managers, for them to check and complete (December 2013-January 2014). The questionnaires were pre-filled with easily available data.

The results of the first part will be presented at the Workshop, the present document contains the results of the second part, i.e. the information provided by DB managers in the questionnaires.

A total of 41 completed questionnaires was returned (= 85%).

For one DB, a reply was received with an incomplete questionnaire: Glycine. The DB manager was unable to participate in the survey and indicated the need to identify another host for the DB.

For six DBs the questionnaires were not returned, despite repeated reminders (Chicory, Cicer, Lens, Pepper, Minor Forage Legumes and Pisum).

All replies received were reviewed and queries were sent to DB managers for clarifications/precisions when necessary (January-April 2014). This allowed many updates and corrections, all integrated here (e.g. change in URLs, change in DB manager, update of the DB itself[1]).

The present document integrates the questionnaires for all DBs, including the "pre-filled" versions when no reply had been received.

A synthesis of the information provided will be presented at the Documentation and Information Workshop.

Note: the CCDBs are listed in alphabetical order of crop/species (see next page, contents).

Elinor Lipman, compiler
April 2014

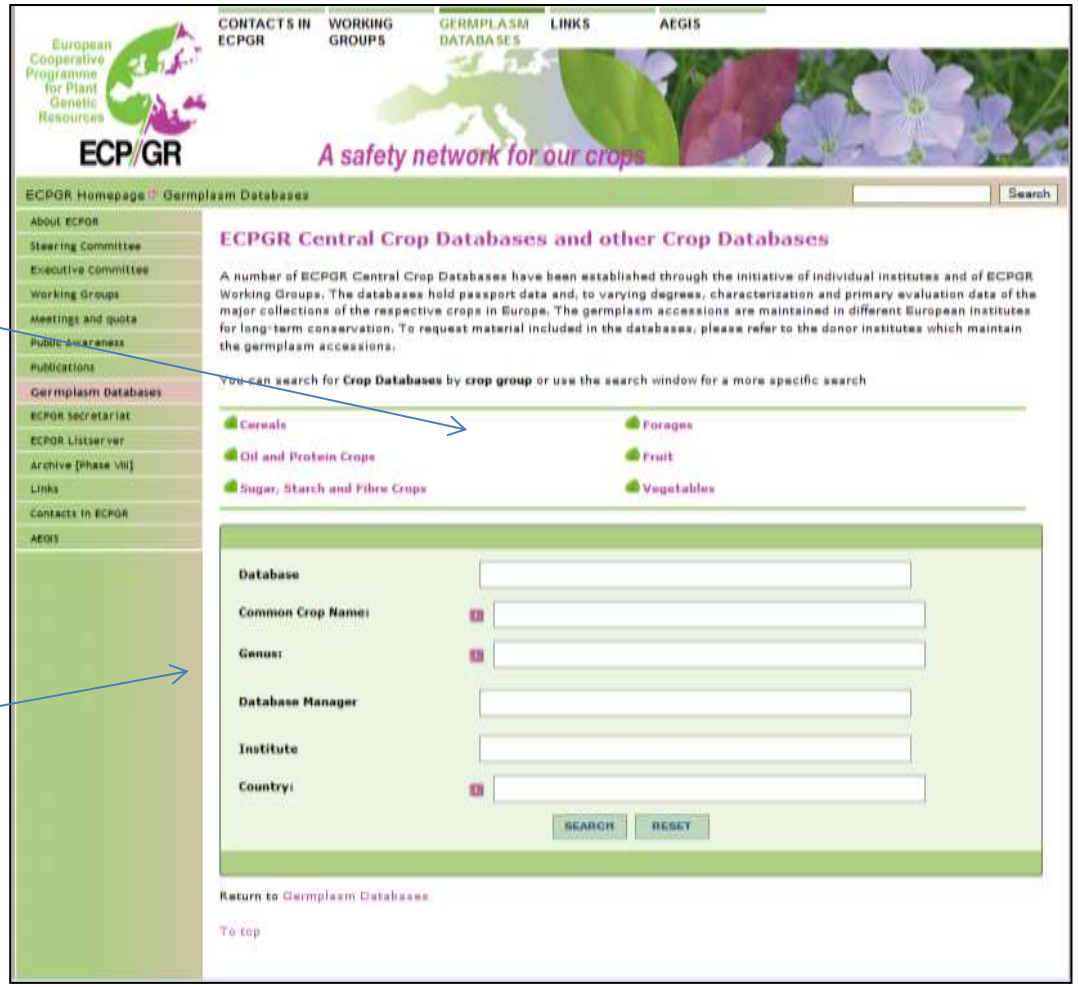
[1] Databases updated during the survey: Prunus (Jan 2014), Triticale (Feb 2014), Vitis (April 2014), Wild Potato (Jan 2014)

Access to CCDBs from ECPGR website

http://www.ecpgr.cgiar.org/germplasm_databases/central_crop_databases.html

Direct links to “crop groups” pages (corresponding to the ex-Networks, but the lists also include “other crop databases”)

More specific searches (crop name, genus, DB manager, institute, country)



List of the CCDBs in relation to the ECPGR Crop Working Groups

Next slides: quick overview of CCDBs classified by “crop groups” (= ex-Networks, Phase VIII) and within each, by Working Group

The 48 existing CCDBs are distributed unevenly across WGs

- The taxonomical coverage of WGs and of CCDBs is variable (one or more genus/genera/species..) → One WG may have one or several DBs
- There are “satellite” DBs, not linked to a WG
- Only one WG does not have a CCDB (Medicinal and Aromatic Plants)

Cereals: 3 WGs, 6 DBs

Crop groups (=ex-Networks Phase VIII)	Working Groups Phase IX	Databases	DB managers	Hosting institute
Cereals	***Avena***	Avena	Christoph Germeier	JKI, Quedlinburg, Germany
	Barley	Barley	Helmut Knüpffer and Markus Oppermann	IPK, Gatersleben, Germany
	Wheat	Wheat	Iva Faberová	Crop Research Institute (CRI), Prague, Czech Republic
	-	Maize	Violeta Anđelković	Maize Research Institute "Zemun Polje", Belgrade, Serbia
	-	Secale	Marcin Zaczyński	Plant Breeding and Acclimatization Institute – National Research Institute (IHAR-PIB), Radzikow, Poland
	-	Triticale	Beate Schierscher-Viret	Agroscope, Nyon, Switzerland

Forages: 1 WG, 10 DBs

Crop groups (=ex-Networks Phase VIII)	Working Groups Phase IX	Databases	DB managers	Hosting institute
Forages	**Forages**	*Dactylis*	Bartosz Tomaszewski	Botanical Garden of IHAR, Poland
		Festuca	Bartosz Tomaszewski	Botanical Garden of IHAR, Poland
		Lolium & Trifolium	Ian Thomas	IBERS, Aberystwyth University, UK
		Medicago Annual	Valentín Maya Blanco	CICYTEX, Badajoz, Spain
		Medicago Perennial	Stéphane Fourtier	CRG INRA, Lusignan, France
		Minor Forage Grasses	Anna Palmé	NordGen, Alnarp, Sweden
		Minor Forage Legumes	Lajos Horváth	RCAT, Tápiószele, Hungary
		Phleum	Anna Palmé	NordGen, Alnarp, Sweden
		Poa	Evelin Willner	IPK Genbank Teilsammlungen Nord Malchow/Poel, Germany
		Trifolium subterraneum	Valentín Maya Blanco	CICYTEX, Badajoz, Spain

Fruit: 3 WGs, 6 DBs

Crop groups (=ex-Networks Phase VIII)	Working Groups Phase IX	Databases	DB managers	Hosting institute
Fruit	***Malus/Pyrus***	*Malus*	Matthew Ordidge	National Fruit Collections, UK
	Malus/Pyrus	*Pyrus*	Marc Lateur and Patrick Houben	CRA-W (Walloon Agricultural Research Centre), Gembloux, Belgium
	Prunus	*Prunus*	Marine Blouin	INRA, Bordeaux, France
	Vitis	*Vitis*	Erika Maul	JKI, Siebeldingen, Germany
	-	Minor Fruit Trees	Edgardo Giordani	Department of Agri-Food and Environmental Science, University of Florence, Italy
	-	*Ribes/Rubus*	Darius Ryliskis	Vilnius University, Vilnius, Lithuania

Oil and Protein Crops: 1 WG, 8 DBs

Crop groups (=ex-Networks Phase VIII)	Working Groups Phase IX	Databases	DB managers	Hosting institute
Oil and protein Crops	**Grain Legumes**	*Cicer*	Maria da Graça Mendonça Pereira	INIAV, Portugal
		Glycine	Margarita A. Vishnyakova	VIR, St. Petersburg, Russian Federation
		Lens	Necla Tas (acting for Lerzan Gul Aykas)	AARI, Izmir, Turkey
		Lupinus	Paweł Barzyk	Poznań Plant Breeders, Ltd., Poland
		Phaseolus	Wolfgang Kainz	AGES, Linz, Austria
		Pisum	Mike Ambrose	JIC, Norwich, UK
		Vicia faba	Gérard Duc	INRA Dijon, France
		Vigna	Wolfgang Kainz	AGES, Linz, Austria

Sugar, Starch and Fibre crops: 4 WGs, 4 DBs

Crop groups (=ex-Networks Phase VIII)	Working Groups Phase IX	Databases	DB managers	Hosting institute
Sugar, Starch and Fibre Crops	***Beta***	*Beta*	Lothar Frese and Christoph Germeier	JKI, Quedlinburg, Germany
	Fibre Crops (Flax and Hemp**)**	Flax	Martin Pavelek	AGRITEC Ltd., Šumperk, Czech Republic
	Medicinal and Aromatic Plants	-		
	Potato	Cultivated Potato	Heather Campbell	SASA, Edinburgh, UK
	Potato	Wild Potato	Roel Hoekstra	PRI, CGN, Wageningen, The Netherlands

Note: the Hemp CCDB is listed on the website but still under construction

Vegetables: 6 WGs, 14 DBs

Crop groups (=ex-Networks Phase VIII)	Working Groups Phase IX	Databases	DB managers	Hosting institute
Vegetables	***Allium***	*Allium*	Joachim Keller	IPK Gatersleben, Germany
	Brassica	*Brassica*	Frank Menting and Noor Bas	CGN, Wageningen, The Netherlands
	Cucurbits	Cucurbits	María José Díez Niclos and José Miguel Blanca	COMAV, Valencia, Spain
	Leafy Vegetables	Chicory	Pascal Coquin	DB maintained by GEVES, France and hosted by CGN, The Netherlands
	Leafy Vegetables	*Lactuca*	Frank Menting and Rob van Treuren	CGN, Wageningen, The Netherlands
	Leafy Vegetables	Minor Leafy Vegetables	Ulrike Lohwasser	DB maintained by IPK Gatersleben, Germany and hosted by CGN, The Netherlands
	Leafy Vegetables	Spinach	Frank Menting and Rob van Treuren	CGN, Wageningen, The Netherlands
	Umbellifer Crops	Umbellifer	Charlotte Allender	Genetic Resources Unit, University of Warwick, Wellesbourne, United Kingdom
	Solanaceae	Eggplant	Gerard M. van der Weerden	Experimental Garden and Genebank, Radboud University Nijmegen, The Netherlands
	Solanaceae	Groundcherry	José Vicente Valcarcel and José Miguel Blanca	COMAV, Valencia, Spain
	Solanaceae	Pepino	José Vicente Valcarcel and José Miguel Blanca	COMAV, Valencia, Spain
	Solanaceae	Pepper	Necla Tas (acting for Lerzan Gul Aykas)	AARI, Izmir, Turkey
	Solanaceae	Tomato	Frank Menting and Willem van Dooijeweert	CGN, Wageningen, The Netherlands
	Solanaceae	Tree Tomato	José Vicente Valcarcel and José Miguel Blanca	COMAV, Valencia, Spain

Homepages

Three types of access

- **ECPGR homepage (11 DBs)**

(linking or not to an institutional site)

- **Institutional homepage (35 DBs)** - *preferable*

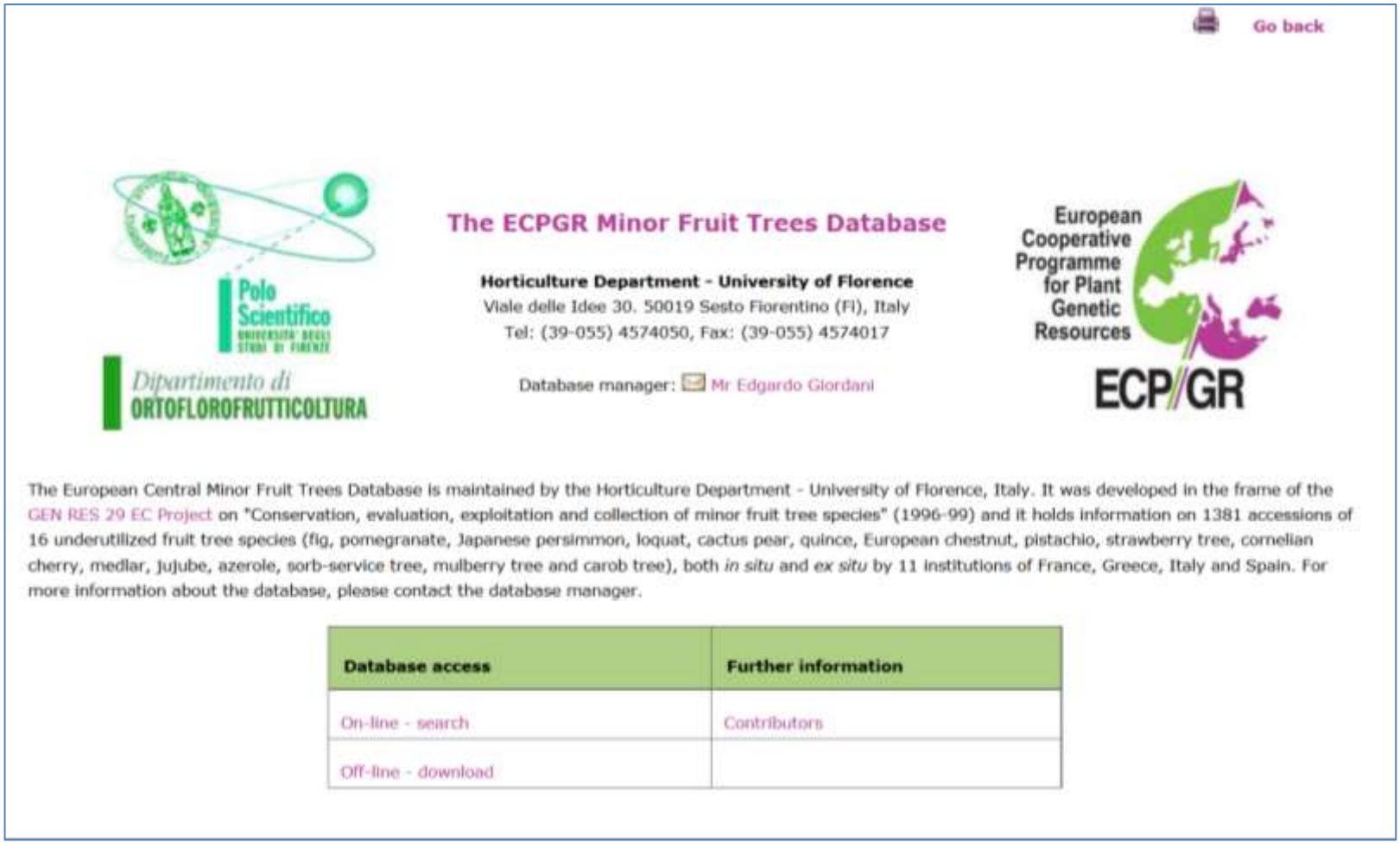
- **Google Fusion Tables**

Specificity of the **Forages WG**, created to facilitate the identification of **AEGIS candidates**

2 DBs provide as their only homepage the URL of a Google Fusion Table (*Dactylis* and *Festuca*)

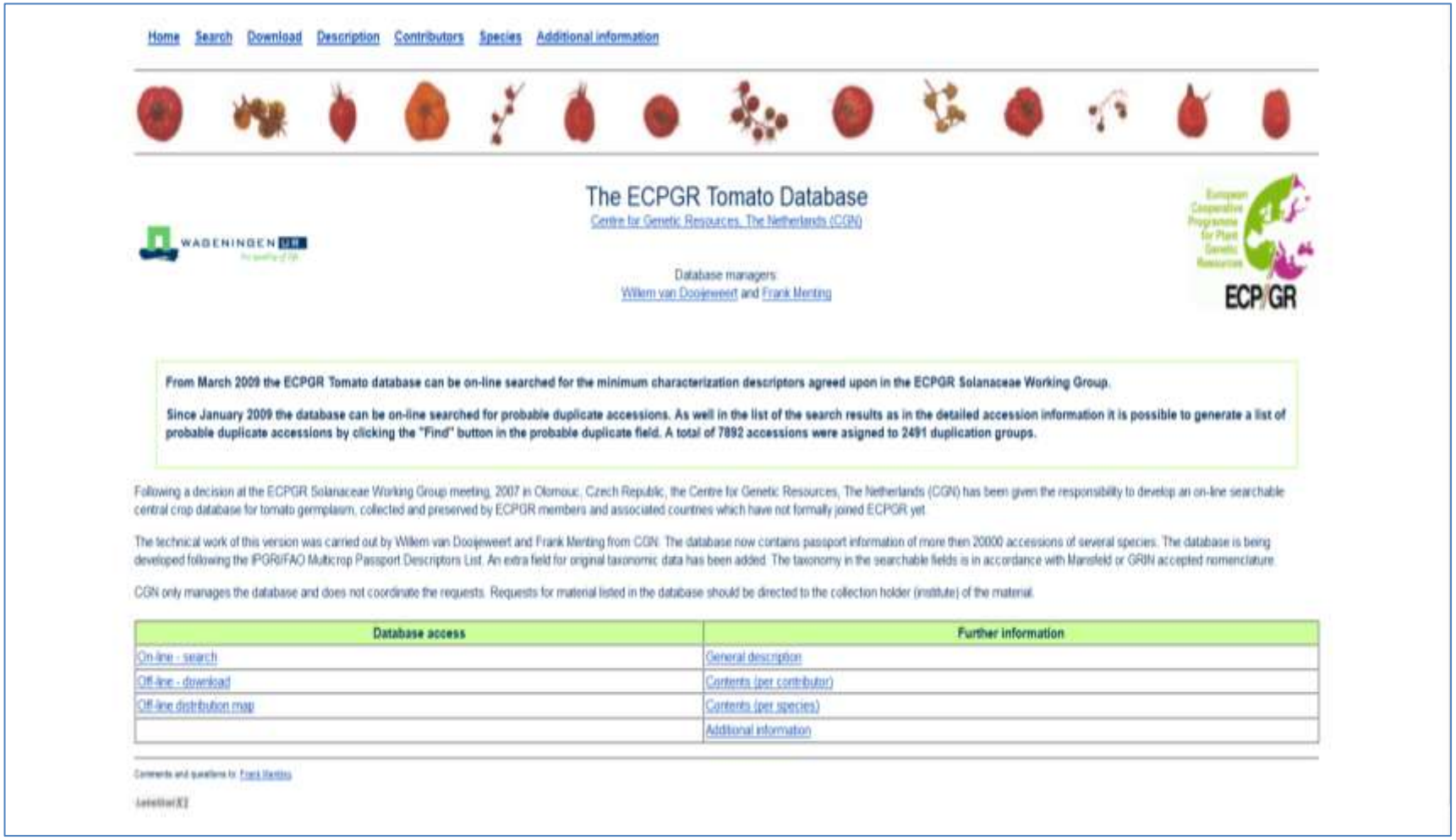
Homepage example 1 – ECPGR page (Minor Fruit Trees)

http://www.ecpgr.cgiar.org/germplasm_databases/list_of_germplasm_databases/crop_databases/crop_database_windows/minor_fruit_trees.html



Homepage example 2 – institutional page (Tomato)

http://documents.plant.wur.nl/cgn/pgr/tomato//



Home Search Download Description Contributors Species Additional information

The ECPGR Tomato Database

Centre for Genetic Resources, The Netherlands (CGN)

Database managers:
Willem van Dooijeweert and Frank Menting

From March 2009 the ECPGR Tomato database can be on-line searched for the minimum characterization descriptors agreed upon in the ECPGR Solanaceae Working Group.

Since January 2009 the database can be on-line searched for probable duplicate accessions. As well in the list of the search results as in the detailed accession information it is possible to generate a list of probable duplicate accessions by clicking the "Find" button in the probable duplicate field. A total of 7892 accessions were asigned to 2491 duplication groups.

Following a decision at the ECPGR Solanaceae Working Group meeting, 2007 in Olomouc, Czech Republic, the Centre for Genetic Resources, The Netherlands (CGN) has been given the responsibility to develop an on-line searchable central crop database for tomato germplasm, collected and preserved by ECPGR members and associated countries which have not formally joined ECPGR yet.

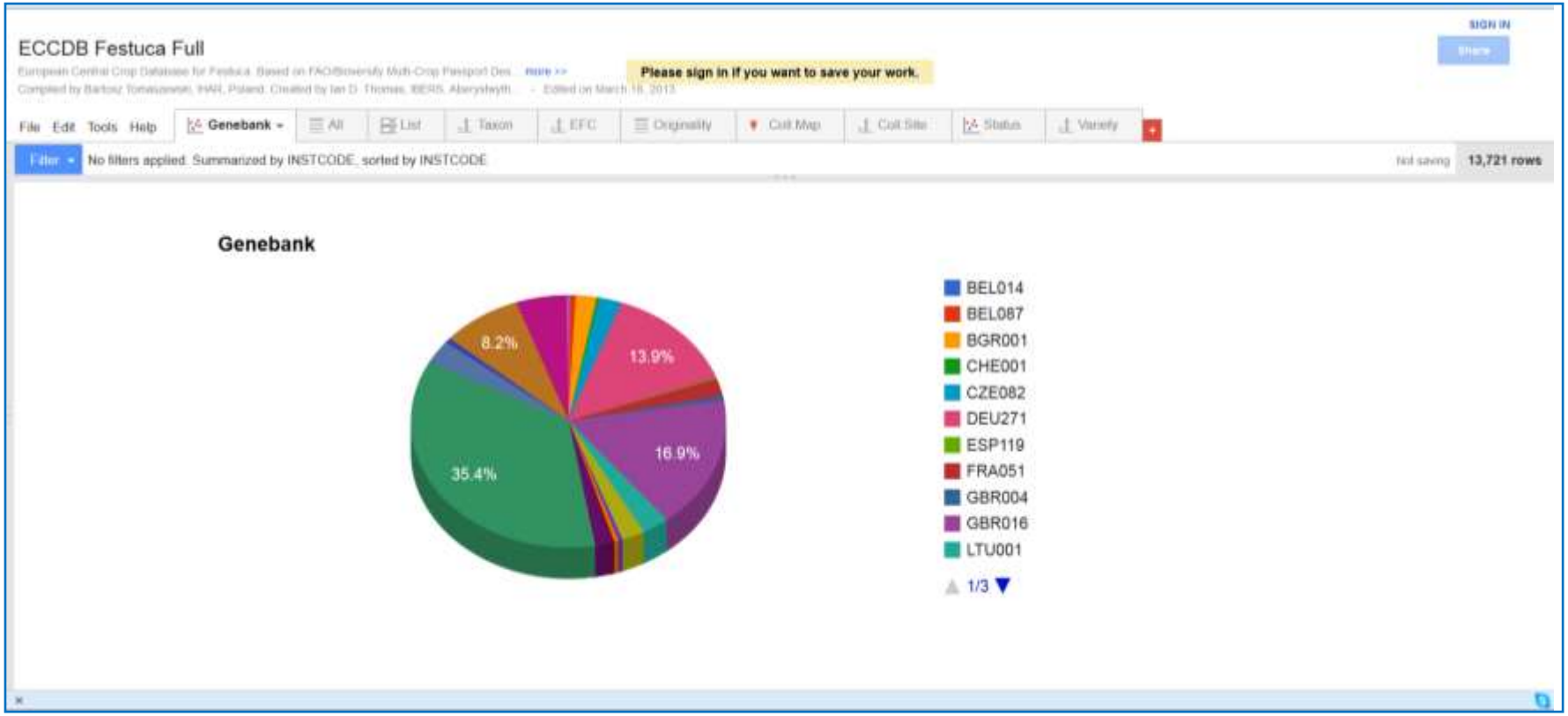
The technical work of this version was carried out by Willem van Dooijeweert and Frank Menting from CGN. The database now contains passport information of more then 20000 accessions of several species. The database is being developed following the IPGRI/FAO Multicrop Passport Descriptors List. An extra field for original taxonomic data has been added. The taxonomy in the searchable fields is in accordance with Mansfeld or GRIN accepted nomenclature.

CGN only manages the database and does not coordinate the requests. Requests for material listed in the database should be directed to the collection holder (institute) of the material.

Database access	Further information
On-line - search	General description
Off-line - download	Contents (per contributor)
Off-line distribution map	Contents (per species)
	Additional information

Homepage example 3 – Google Fusion Table (*Festuca*)

https://www.google.com/fusiontables/DataSource?docid=1Lcsf7_nAJbViL14dnp1PR4Eu9mY8EMmoC1n1kZY#chartnew:id=10



Date of creation

Range:
from 1980 (Minor Forage Grasses) to 2011 (*Ribes/Rubus*)

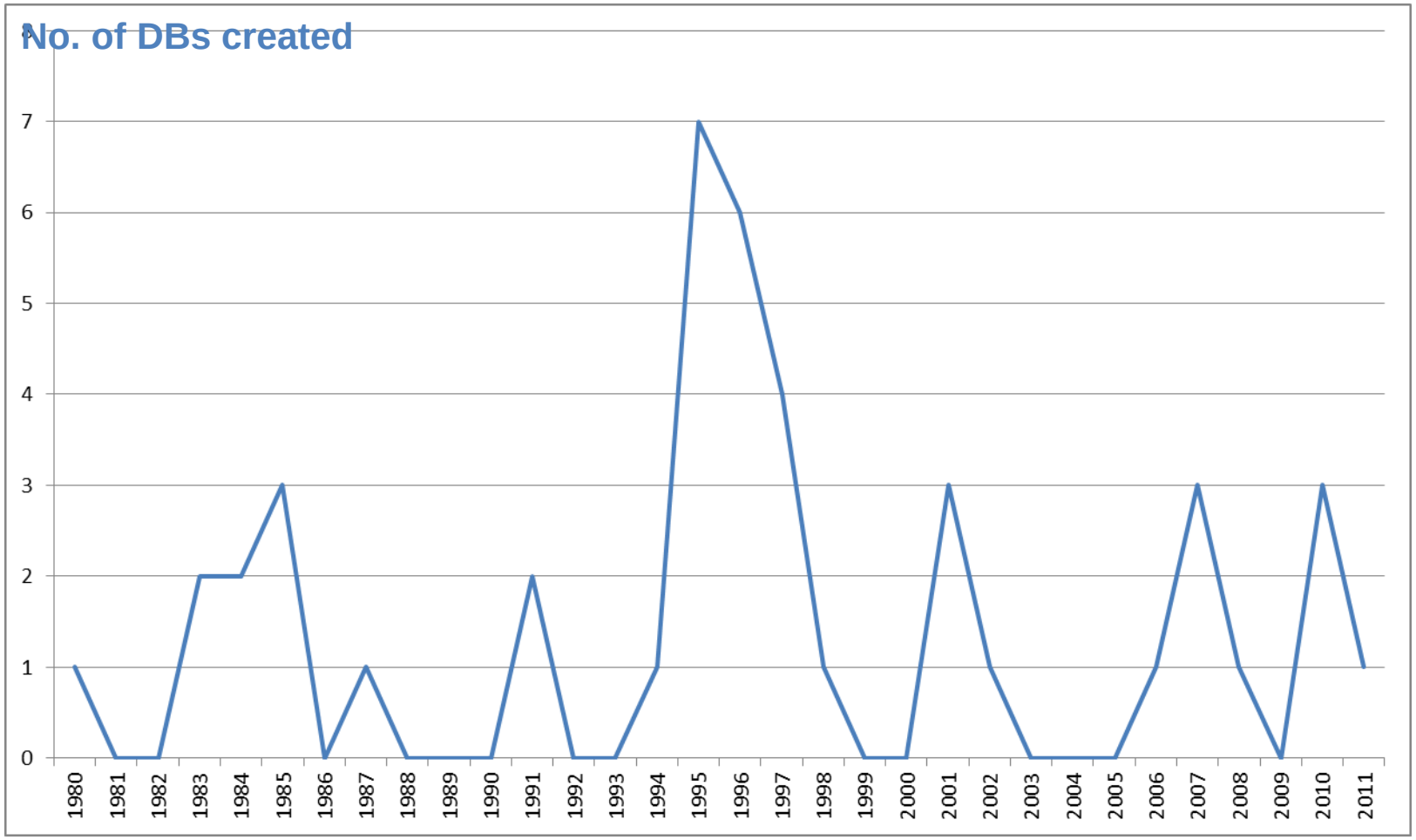
Notes:

Sometimes complete renewal at long intervals – esp. **Forages** DBs reviewed 2010
e.g. Minor Forage Grasses: components 1980, current structure 2010
Perennial *Medicago*: first version 1985, restructured and updated 2013

Or transfer from one hosting institute to another
e.g. *Prunus*: created 1983 at Nordgen, since 1994 at INRA-Bordeaux
Allium: created 1984 at HRI-Wellesbourne, since end December 2011 at IPK

→ *date considered for the graph (next slide) = earlier one*

Chronology of CCDB creation
"DB boom" in 1995-1996 = Budapest effect!



Number of accessions

Range:

from 87 (Tree Tomato) to 180 639 (Wheat)

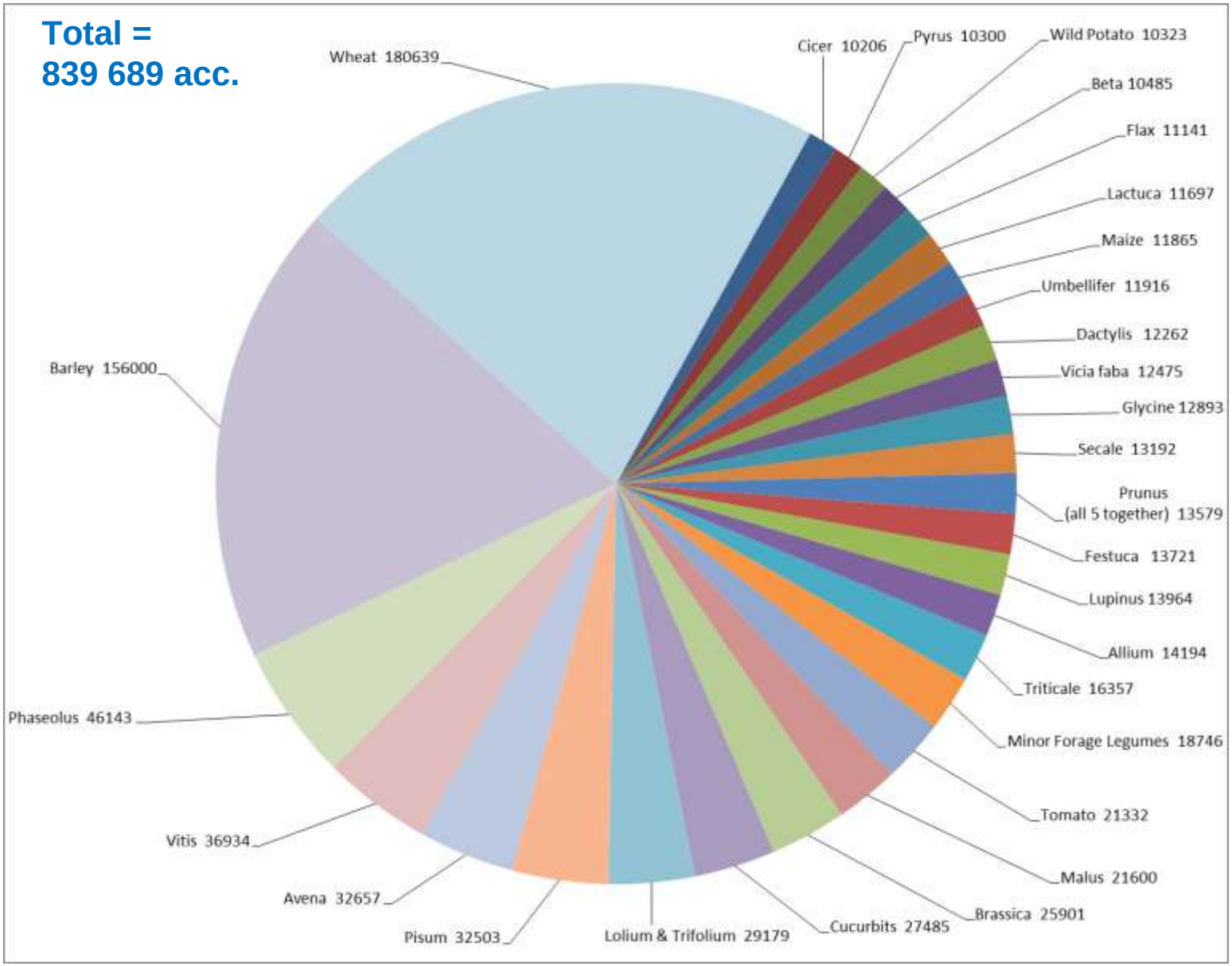
Total = 917 927 records

- **29 DBs with > 10 000 acc. = total 839 689 acc. (91.5%)** *(graph, next slide)*
  Out of which 2 clearly stand out, with more then 100 000 acc.: **Wheat and Barley,** together making up for ca. 37% of grand total
- **19 DBs with < 10 000 acc. = total 78 238 acc. (8.5%)**

Notes:

- *Generally number at December 2013*
- *Some exceptions for 4 DBs updated during survey process: Prunus and Wild Potato - Jan 2014, Triticale - Feb 2014, Vitis - April 2014*
- *Total numbers given here include non-European holdings + these recent updates, therefore will be slightly different from the numbers in data file analysis (comparison with EURISCO)*

29 DBs with more than 10 000 acc.



Date of last update

Range: from 1998 (*Malus*) to 2014 (4 DBs updated during survey)

Notes:
Need to differentiate between "technical update" and updating of records

Special cases worth mentioning:

- "**Old" DBs linked to specific project,** not updated after end of project (e.g. Minor Fruit Trees 1999, Barley 2001)

- **Update performed permanently/directly online** by contributors: Cultivated Potato, *Vitis*

Average age of records

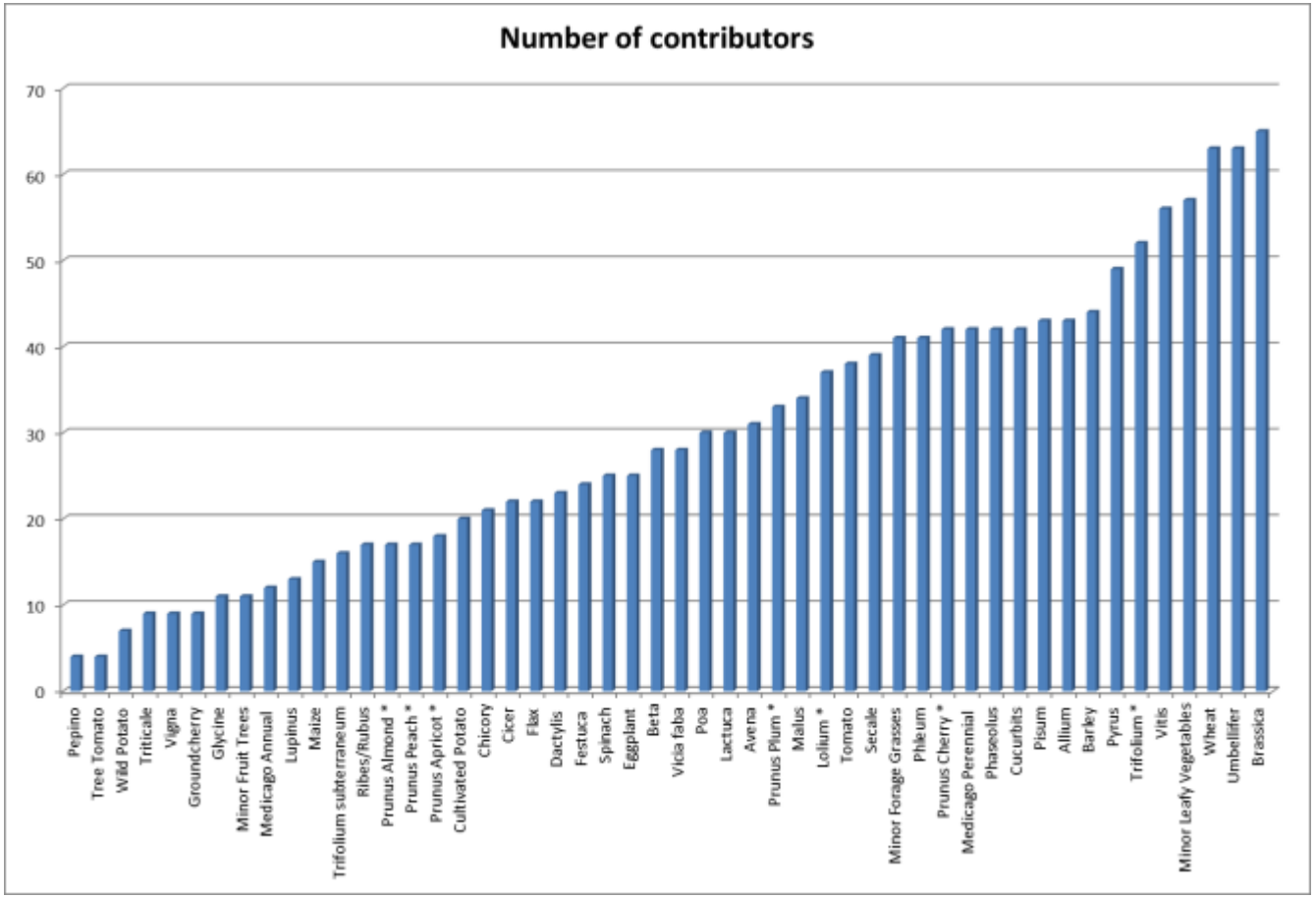
Range: from 0.75 year to 26 years

Number of contributors

Range:
from 4 (Pepino and Tree Tomato) to 65 (*Brassica*)

Note: some DBs have several components (data files) with different contributors, e.g. *Prunus* (5 files, each species-specific); "*Lolium & Trifolium*" (2 files, one for each genus)

→ They appear separately here



Non-European contributors

9 DBs including Chicory / 9 non-European countries
Total = 50 605 acc. (5.5% of total)

Database	Non-European holdings									
	Australia	Canada	China	India	Iran	Israel	Japan	Syria (ICARDA)	USA	Totals
Avena	0	14	0	0	0	0	0	0	45	59
Barley	9947	0	0	0	0	0	4016	24372	0	38335
Beta	0	0	78	132	229	50	32	0	2255	2776
Chicory	0	0	0	0	0	0	0	0	82	82
Flax	0	0	0	0	0	0	0	0	369	369
Lactuca	0	0	0	0	28	0	0	0	1162	1190
Lupinus	4665	0	0	0	0	0	0	0	1293	5958
Minor Leafy Vegetables	0	0	0	0	0	0	0	0	1475	1475
Spinach	0	0	0	0	0	0	0	0	361	361
Totals	14612	14	78	132	257	50	4048	24372	7042	**50605**

Passport data

- **Which descriptors are used?**

  MCPDs and/or others →

  Remark: no question was asked about data completeness, therefore the quality of contents cannot be evaluated

- **Accessibility of passport data: are they searchable and/or downloadable? →**

Passport descriptors used

(42 DBS considered here)

- **Multi-crop Passport Descriptors (MCPDs): 40 DBs (95%)**
  Including Maize, using "earlier version of MCPDs"
  All others use either V1 (2001), V2 (2012), or combination of both (sometimes subsets)

- **"Other descriptors": 29 DBs (69%)**
  Including the 2 DBs not using MCPDs: Cultivated Potato and Minor Fruit Trees

MCPDs				Other passport descriptors	
V1 or subset*	**V2**	**V1 + V2**	**None**	**Yes****	**No**
29 (69%)	**7** (~16.5%)	**4** (9.5%)	**2** (~5%)	**29** (69%)	**13** (31%)

* including Maize
** including Cultivated Potato and Minor Fruit Trees

What are the “other passport descriptors”?

- **EURISCO** (Allium, Eggplant, Phaseolus, Prunus, Vigna, Wheat)
  or EURISCO modified (descriptors added/removed)
  (Brassica, Lactuca, Minor Leafy Vegetables, Spinach, Tomato)

- **WG or crop-specific**
  (Allium, Forages, Malus, Prunus, Pyrus, Vitis, Wheat)

- **Other additional descriptors (incl. AEGIS, MLS)**
  Avena, Barley, Maize, Secale, Triticale
  Beta, Wild Potato
  Brassica, Eggplant
  Medicago Perennial, Minor Forage Grasses, Phleum, Poa

- **Specific to the DB / to project**
  Cultivated Potato, Minor Fruit Trees

Accessibility of passport data

Are they searchable/downloadable?

Some are searchable but not downloadable, or the opposite, or both

No. of DBs with searchable passport data = 33 out of 44 DBs (75%)

*The **no. of searchable descriptors varies greatly (from 2, Maize, to 43, Beta)***
Full list not provided for all
Combinability..

	Passport data				
	Searchable	**Downloadable**	**Searchable AND downloadable**	**Only searchable**	**Only downloadable**
No. of DBs	**Yes = 33 (75%)** No = 11 (25%)	**Yes = 35 (80%)** (but some only after search) No = 9 (20%)	**24** (55%)	**9** (20%)	**11** (25%)

Characterization and evaluation (C&E) data

21 DBs* out of 48 (~44%) include C&E data

** including Chicory but no details from DBM, so not included in detailed analysis*

Note: In some cases, only for subsets of accessions, e.g.

Avena: AVEQ project database still to be integrated into the EADB

Barley: 406 accessions not included

Cucurbits: only *Citrullus lanatus* and *Cucumis sativus*

Databases with C&E data
1. Avena
2. Barley
3. Beta
4. Brassica
5. Chicory
6. Cucurbits
7. Cultivated Potato
8. Eggplant
9. Flax
10. Lactuca
11. Maize
12. Minor Fruit Trees
13. Minor Leafy Vegetables
14. Phaseolus
15. Ribes/Rubus
16. Secale
17. Spinach
18. Tomato
19. Vitis
20. Wheat
21. Wild Potato

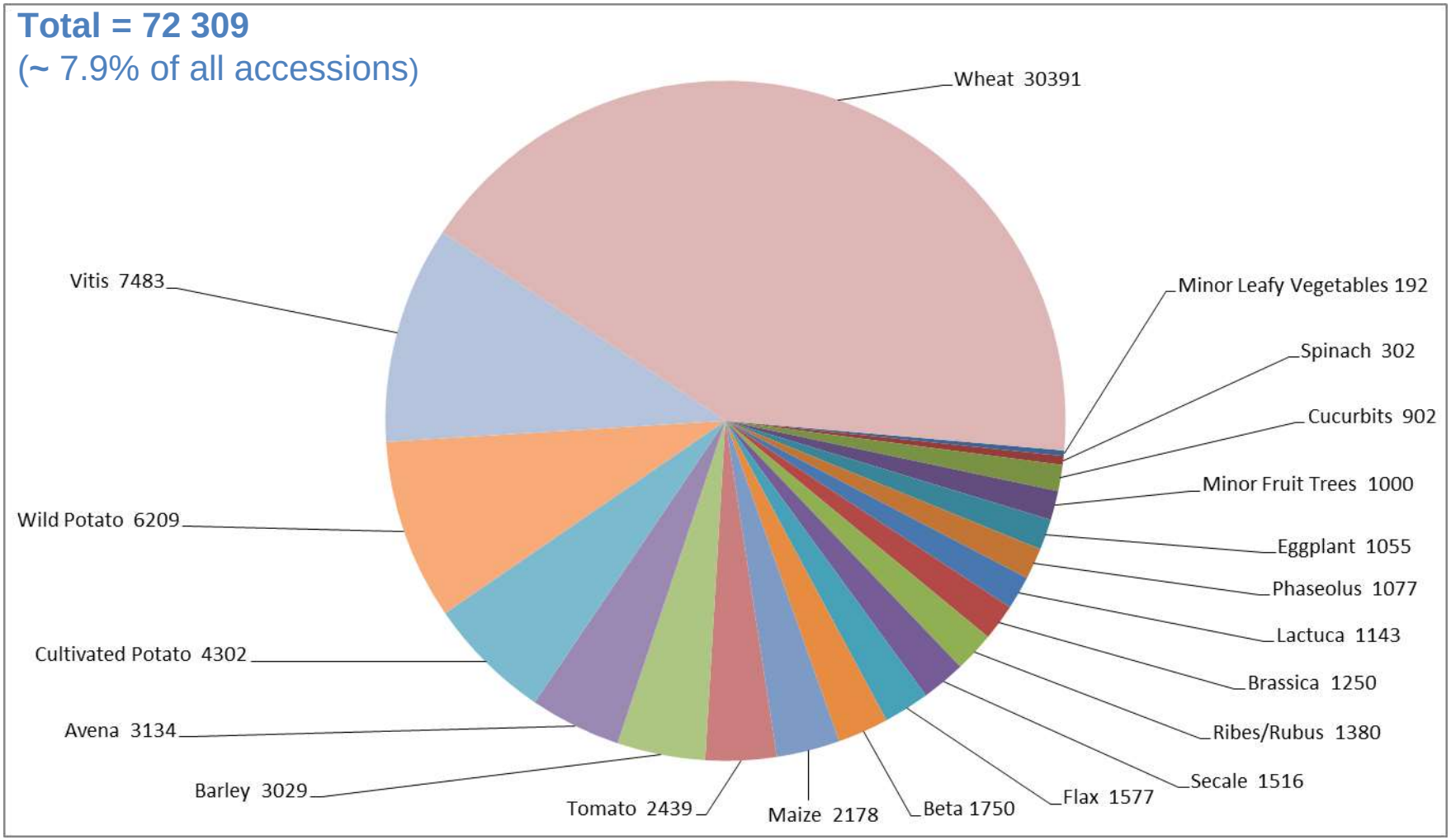
C&E data

Some DBs do not include them yet, but plan to do so and/or have structure ready

Malus	*The database structure does hold the capacity, if not the data*
Poa	*No, not yet in db, should follow in this winter, but we have a lot of C&E data for our accessions (DEU 271)*
Prunus	*Not yet (but structure is ready)*
Pyrus	*Apart from fruit pictures, not yet but it is well planned: the last WG meeting has already defined priority C&E descriptors*
Triticale	*No, but could be integrated if requested and if the organisations transmit us data. (…) no problem to integrate characterization and evaluation data in the future.*
Vicia faba	*During 2013-15, the database will be updated with verified passport data of old and new collections, as well as addition of phenotypic data when available. The database will be prepared for adding molecular markers data. It will need a host website (…)*

Number of accessions with C&E data (20 DBs)

Total = 72 309
(~ 7.9% of all accessions)



C&E data

Number of data points

Few usable replies, listed here to acknowledge the efforts!

Database	Number of C&E data points
Minor Leafy Vegetables	For characterization 1059 for *Eruca*, 322 for *Valerianella* and 190 for *Atriplex* For evaluation 1002 for *Eruca*
Ribes/Rubus	ca. 55 000
Secale	7015
Flax	9366
Beta	36 513
Maize	34 848
Tomato	15 674
Barley	61 974
Avena	169 799
Wild Potato	33 210
Wheat	20 7058

Accessibility of C&E data

Are they searchable/downloadable?

Same as for passport data: some are searchable but not downloadable, or the opposite, or both

No. of DBs with searchable C&E data = 11 (55%)

*Again the **no. of searchable descriptors varies greatly (from 2, Maize, to 87, Cultivated Potato)***
Combinability mentioned by Avena and Beta: "all descriptors are searchable, combined with each other and with passport descriptors"

	C&E data				
	Searchable	**Downloadable**	**Searchable AND downloadable**	**Only searchable**	**Only downloadable**
No. of DBs	**Yes = 11 (55%)** **No = 9**	**Yes = 13 (65%)** **No = 7**	**6**	**5**	**7**

Accessibility of C&E data (20 DBs)

Database	C&E data searchable		C&E data downloadable	
	Yes	No	Yes	No
Avena	X		X	
Barley	X			X
Beta	X		X	
Brassica		X	X	
Cucurbits	X		X	
Cultivated Potato	X			X
Eggplant	X			X
Flax		X	X	
Lactuca		X	X	
Maize	X		X	
Minor Fruit Trees		X		X
Minor Leafy Vegetables		X	X	
Phaseolus		X		X
Ribes/Rubus		X	X	
Secale	X		X	
Spinach		X	X	
Tomato	X		X	
Vitis	X			X
Wheat	X			X
Wild Potato		X	X	
Total number of DBs (%)	**11 (55%)**	**9 (45%)**	**13 (65%)**	**7 (35%)**

Other information available on the DB websites

Database	Broad categories				
	Links (projects, literature, other sites)	Search forms / Display of results	Photos	Help or reference documents	Other (special collections, taxonomic/gene info)
Allium	x				
Avena	x	x	x	x	
Barley	x				x
Beta	x	x	x	x	
Brassica	x	x			
Cucurbits					x
Cultivated Potato		X		x	
Eggplant	x		x		
Lactuca		X			
Minor Fruit Trees			x		
Minor Leafy Vegetables		x			
Phaseolus			x	x	
Prunus		x	x	x	
Pyrus		x			
Ribes/Rubus			x		
Secale		x		x	
Spinach		x			
Tomato	x	x			
Tree Tomato		x			
Vigna				x	
Vitis	X	x		x	
Wheat		x			x

Management (1)

How much time (average number of hours per year over the last three years) is invested in the database?

Range:
from 0 (*Lactuca* and Spinach) to 700 (*Avena*)

Replies not fully exploitable
Several DB managers mention difficulty in replying – work time *"intermingled with other tasks/responsibilities"*, management of several DBs, impossible to precise exact time...

→ To simplify:

- **No time or very little time invested (from 0 to 10 hours)**
  - Leafy Vegetables DBs (*Lactuca*, Spinach, Minor Leafy Vegetables): decision of the WG *"no longer to invest in the database"*
  - Small DBs requiring only *"few hours":* Groundcherry, Pepino, Tree Tomato
  - Minor Fruit Trees (project terminated)
  - Wild Potato (but much more time in the past, creation of macros for updates)
- **Majority of replies indicate between 10 and 100 hours/year**
- **Most time-consuming: 100 h or more (13 DBs)** →

Management: Most time-consuming DBs (≥100 hours/year)

Database	How much time (average number of hours per year over the last three years) is invested in the database?
Wheat	40-160 hours depending on the data quality and delivery – the most time-consuming was renumbering German (Braunschweig/Gatersleben) and French (GEVES-INRA) collections in 2011 and continuous C&E data conversion from Excel to EWDB format
Poa	approx. 80-100 h per year
Secale	100 hours
Lolium & Trifolium	100 hrs per year.
Allium	~ 100 h
Dactylis	About 125 hours.
Festuca	About 130 hours.
Cucurbits	The database has been continuously revised and compared with EURISCO, with several contributors' databases and with online contributors' websites in order to have a more complete dataset. Approximately 480 hours have been invested in these tasks over the last three years.
Medicago Perennial	During the three last years I restarted a work interrupted since more than 10 years, I spent a lot of time reorganizing files and managing data under database software: I think I spent more than 3 months on the European Perennial Medicago database.
Flax	Approximately 176 hours per year
Lupinus	24 days/year
Vitis	At least 600 hours per year Besides the technical aspects of DB maintenance and improvement (e.g. upload and search functionalities), a lot of time was invested in data checking and assistance to partners.
Avena	This is difficult to estimate, because it is intermingled with work in documentation work packages for research projects, but it is at least 50% of my total working time (ca. 700h)

Management (2)

Is your mechanism of data gathering and updating working well / sustainable?

Replies are not always clear-cut, but **majority of DBs mention difficulties or answer negatively**

- **Positive replies, no problems mentioned**

  Cucurbits, Cultivated Potato*, *Dactylis*, Eggplant, *Festuca*, Groundcherry, *Lupinus*, Pepino, Tree Tomato, *Vicia faba* , *Vitis**, Wheat

  (*) = update performed directly by contributors themselves

- **Working well, but time-consuming**

  Brassica – due to harmonisation of taxonomy and need to check additional fields
  Poa - updating procedure is too heavy

Is your mechanism of data gathering and updating working well / sustainable?

- **Not working so well / negative replies**

  - **Difficulties in data gathering: heavy process** of requesting data to individual contributors, **low response** rate, low quality of data provided (format)

    Some "ambivalent" perception of EURISCO
    + positive as a source of data usable for updates (e.g. *Secale, Lolium&Trifolium*)
    - "competitor" for provision of data by contributors (e.g. *Phaseolus, Vigna*, Barley)

  - **Lack of resources (time/technical)** to carry data acquisition and uploading / software development or technical work needed (e.g. *Avena, Beta, Prunus*)

+ Special cases

Pyrus: awaiting development of *Prunus* DB
(Minor Fruit Trees: no new project)

Assessment of use and value (1)

What do you think your database adds to the data available in EURISCO?

Here again, info grouped into broad categories

- **Data on collections not (yet) included in EURISCO →**
- **Data on type of material not included in EURISCO →**
- **Taxonomic / gene info / historical data →**

- **Use of other descriptors →**
- **C&E data** *(see above, that section)*
- **Photos** *(see above, "other information")*
- **Search and update procedures →**

- **Support to WG activities →**

Data on collections not (yet) included in EURISCO

- DBs containing data from **non-European collections**
  e.g. Lactuca, Spinach…

- **European collections present in CCDBs but not in EURISCO**
  e.g. Barley, *Brassica* (specifically BLR and FRA), *Prunus*, Umbellifer*

  (*) comment from Umbelllifer:
  *It provides **more up-to-date information** in situations where institutional updates to EURISCO are delayed by issues concerning National Focal Points or National Coordinators. Therefore some material is listed in the EUDB and not in EURISCO.*

- **Passport data more complete in CCDB**
  e.g. *Brassica*, Cucurbits *(see specific presentation by M.J. Díez)*

What do you think your database adds to the data available in EURISCO?

Data on type of material not included in EURISCO

- **Material which is not distributable as genetic resource to users**
  but for which there is either living reference material or other documents (herbarium, images etc.) (Allium)

- **(Historical records – see next section)**

What do you think your database adds to the data available in EURISCO?

Taxonomic / gene info / historical data

- **Parallel taxonomic system mentioning synonyms and also historical taxonomic** determinations of an accession *(Avena, Beta)*
- **Harmonisation of taxonomy** *and therefore easier to search for crop types (Brassica)*
- **Link to taxonomic DB / more precise taxonomy than in EURISCO**
  Case of Wheat:
  Records linked to the Database Wheat Pedigree and Identified Alleles of Genes
  Uses taxonomy by Dorofeev (27 species in the genus Triticum) / 6 in EURISCO (GRIN)
- Some **older information of named genes** described in literature *(Avena)*
- **Historical records** *(belonging to lost or removed material) that could be used as a source of additional information (Wheat)*

What do you think your database adds to the data available in EURISCO?

Use of other descriptors

As already mentioned, used by 69% of DBs

Some highlights / examples:

- **Homonyms and synonyms: important for Fruit DBs** (*Malus, Pyrus, Vitis*)
- **Descriptors used for the identification of AEGIS: Forages DB** (SEEDAVAIL, EFC, ORIGINALITY, PRIMCOLL, RECDATE, PLOIDY)
- **Descriptors defined in the framework of a project** e.g. Eggplant (EGGNET project)

For complete lists, refer to the compilation of questionnaires

What do you think your database adds to the data available in EURISCO?

Search and update procedures

Highlights
(In addition to information already provided in sections on descriptors)

- **Combined search for characterisation and evaluation data** *(Avena, Beta)*

- **Identification and display of duplicates** (*Avena, Beta, Brassica*, Wild Potato)
  Examples
  - Avena, Beta: groups of duplicated accessions listed together with harmonised passport information for the whole duplicate group.
  - Brassica: For B. oleracea and B. rapa collections probable duplicate groups have been indicated. Probable duplicates are being displayed in the search results and can be found in the download.

- **Update procedure** (*Avena*)
  *.. **keeps historical records and thus maintains consistency and integrity** between passport-, C+E-, and other data, and which also makes transparent, which accessions have been lost by the genebanks from one update to another. It further updates the duplicate search and keeps track on all changes ("corrections") made by the genebanks during the various updates, although this semi-automated update procedure is quite time consuming*

What do you think your database adds to the data available in EURISCO?

Support to WG activities

- **Role in European projects**
  e.g. *Avena, Beta, Brassica*
  Note: also true for many others..

- **Important role for the identification of AEGIS candidates**
  DBs of the Forages WG - mentioned by Minor Forage Grasses, *Phleum, Poa*
  But should be the case of all Forages DBs esp. through Google Fusion Tables

- **Key tool for the management of a collection**
  and helpful in promoting the utilization of genetic resources and the regional coordination of conservation activities (*Dactylis, Festuca)*

What do you think your database adds to the data available in EURISCO?

General considerations on CCDBs vs. EURISCO - Differences and complementarity

- **Developed with different aims → content is not equal**
- **Different data collecting channels (crop curators / National Focal Points)**
  Less distance between CCDB managers and data providers
- **Different responsibilities of data donors**
  CCDB manager: voluntary work
  NFPs: formal agreements

Assessment of use and value (2)

Do you have an idea about the use of your database and the appreciation of the users?

- 14 negative replies
- 9 replies mentioning the existence of a “visitor counter” (either specific to the DB or general for institute website) – no qualitative feedback
- **18 positive replies** – most relevant and concrete examples, *next slide*

Do you have an idea about the use of your database and the appreciation of the users?

Yes - examples

- Requests for material (*Poa*: seed, *Ribes/Rubus*: cuttings)
- Positive feedback and/or requests regarding characterized accessions (Barley, Maize)
- DBs can be sources of potential breeding material (*Dactylis, Festuca*)
- Importance for European projects (*Avena, Beta*) and development of Global Crop Strategy (*Avena*)
- Geographically wide sample of accessions; used to compose the very large working collection of the French AKER project in 2011/2012 (*Beta*)
- DB triggered organization of a collecting mission (*Poa*)
- Impact on public awareness (Minor Fruit Trees)
- Data not available elsewhere on the Internet (Wild Potato)
- **Importance of crop-specific information for users** (breeders, researchers) and WG members as opposed to use of EURISCO by genebank managers (Umbellifers, Wheat)

Perspectives

Are you planning to continue maintaining the database in the next five years and, if positive, are sufficient resources available to do that at your institute?

Majority of positive replies, but many expressing conditions or constraints

Yes, without restrictions = 16 (39%)

Yes, with restrictions = 19 (>46%)

No = 6 (>14%)

Reasons given obviously overlap with replies to question on updating mechanism and sustainability
(Details next slides)

Are you planning to continue maintaining the database in the next five years and, if positive, are sufficient resources available to do that at your institute?

Yes – no restrictions

- Maize
- Secale
- Triticale
- Wheat

- Dactylis
- Festuca
- Lolium & Trifolium

- Vitis

- Lupinus
- Vicia faba

- Cultivated Potato
- Wild Potato

- Cucurbits
- Pepino
- Groundcherry
- Tree Tomato

Are you planning to continue maintaining the database in the next five years and, if positive, are sufficient resources available to do that at your institute?

Yes, with restrictions / constraints

- **Plan to continue, but limited activity**
  - *Beta*: *only maintenance – additional resources needed for upgrading/updating*
  - *Brassica : In a low profile way, only including information/C&E data from ECPGR Brassica WG projects*
  - *Minor Forage Grasses, Phleum - Minor updates are planned but we do not have the time and resources to spend a lot of effort on these data bases*

- **Continuation will depend on EURISCO development**
  - *Phaseolus* - If C&E data cannot be included in EURISCO
  - Tomato – awaiting results of this workshop. *It is preferred to transfer all data to EURISCO. This needs an upgrade of EURISCO*

Are you planning to continue maintaining the database in the next five years and, if positive, are sufficient resources available to do that at your institute?

Yes, but restrictions / constraints

- **Wish to continue, but limiting factors**

  - **Resources in funding / staff**

    Avena
    Malus, Prunus, Ribes/Rubus
    Medicago Annual, Trifolium subterraneum
    To be confirmed after retirement of current DB manager: Allium, Eggplant

  - **- Technical issues**

    Expertise/support needed: Poa, Flax
    Pyrus: awaiting transfer of new version of Prunus DB

Are you planning to continue maintaining the database in the next five years and, if positive, are sufficient resources available to do that at your institute?

No

- **Awaiting developments in EURISCO (and/or outcomes of this Workshop)**

  ➢ **Decision at institutional (hosting institute) level: IPK**

  intends to integrate all DBs maintained (Barley, Allium and Poa) into EURISCO (but keeping them separated for the users). Maybe, depending on the recommendations or decisions of the upcoming Doc&Info meeting, we will look for another solution.

  ➢ **Decision at WG level**

  - Leafy Vegetables (Lactuca, Spinach, Minor Leafy Vegetables)
  - Umbellifer

- **Lack of resources -** can only rely on EC calls (Minor Fruit Trees)

Concluding remarks

What did this survey bring us?

- ✓ Allowed many **updates and corrections** – much needed!
- ✓ **Great amount of info collected** – never done to this extent (last overview Budapest 1996)

- ✓ Pointed out **need for more in-depth analysis of quality aspects** (*data accuracy, completeness, uniqueness, timeliness, consistency..)*

- ✓ **Global picture: great heterogeneity of the CCDBs →**

Concluding remarks - Heterogeneity of the CCDBs

- **In aspects linked to life history and management conditions (impact on sustainability)**

  Cf. **wide ranges for most parameters** investigated

  - Date of creation (changes in hosting and/or coverage in some cases)
  - Contents (size, no. of contributors..)
  - Accessibility of data…

- **In views expressed regarding the continued maintenance of the CCDBs and EURISCO in parallel -** 2 positions:
  - Wish to **discontinue CCDB** (however subject to inclusion of C&E data when applicable) - awaiting outcomes and decisions of this Workshop
  - Wish to **continue CCDB**, judged as complementary to, but not replaceable by EURISCO →

Concluding remarks

Arguments for continuation of CCDBs

- They are at the core of WG activities (incl. AEGIS)
- They "*offer a valuable addition to EURISCO, in that they hold the ability to address crop-specific issues*"
- They (can) offer a wealth of additional information not available in EURISCO (*C&E data, molecular data, pictures, links, texts..)*

→ feasibility to integrate all this info into unique source?

BUT

Consensus on need to avoid duplication of efforts!

→ *see presentation by Theo van Hintum on comparison of data sets…*

Thank you again to all contributors!

And to all Workshop participants,

wishing you success in implementing its ambitious “tailoring” goal…

with a smile!