
The European *ex situ* PGR Information Landscape

Theo van Hintum, Frank Begemann, Lorenzo Maggioni
ECPGR Documentation and Information Network Coordinating Group

PGR Information Landscape

■ the presentation

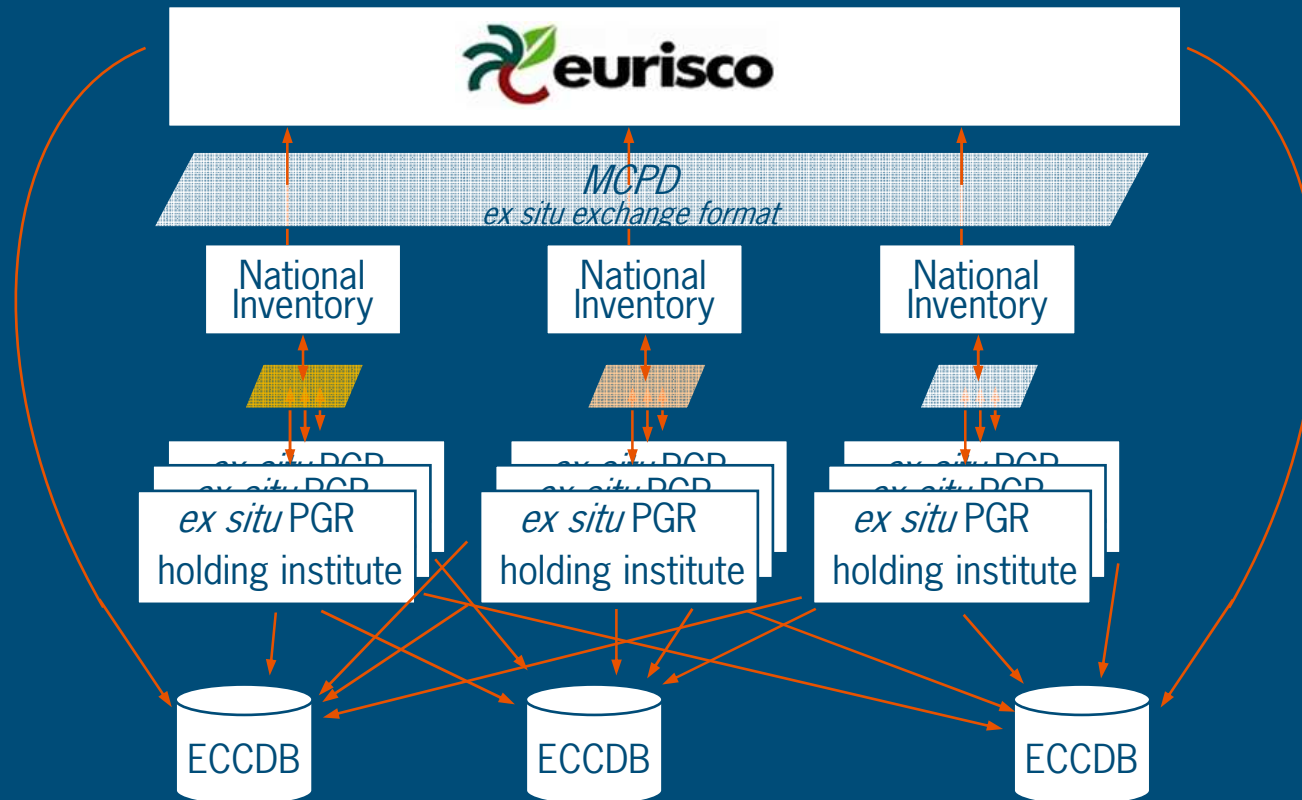
- components of the landscape
 - systems at institutional level
 - systems at national level
 - systems at European level
- changes in the landscape
- technical developments and challenges
- next steps

- conclusions

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- the current landscape



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

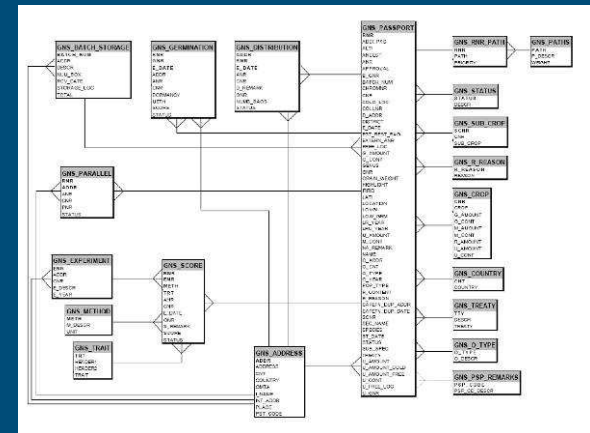
- basic elements of the landscape: local documentation systems of actors conserving PGR
 - paper
 - spreadsheets
 - database management software (DBMS)



	A	B	C	D		
1	January	Product 1	Product 2	Product 3		
2	North	12.1	12.2	12.3		
3	South	A	B	C	D	
4	East	February	Product 1	Product 2	Product 3	
5	West	1	North	12.1	12.2	
	2	South	A	B	C	D
	3	East	March	Product 1	Product 2	Product 3
	4	West	1	North	12.6	13.2
		2	South	9	14.5	
		3	East	12.2	16.8	
		4	West	12.7	13.1	

↓

	A	B	C	D
1	First Quarter	Product 1	Product 2	Product 3
2	North	38.3	41	
3	South	39.6	45.1	
4	East	37.1	43.2	
5	West	41.6	43.6	



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- database management software (DBMS)
 - popular brands: MySQL, MS-Access, Oracle
 - allows for proper database management
 - data integrity
 - data security
 - data processing
 - requires investments
 - license, installation and maintenance
 - requires application based on data-model
 - define structure of the data (what data?)
 - define and implement the functionalities



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ data categories

- genebank management data
 - internal use only (data models vary)
- passport data
 - broad external use (institute, national, European, incl. breeders)
 - fairly standardized (MCPD-list)
- characterization & evaluation data
 - broad external use (breeders, researchers, others incl. institutes)
 - range of models varying level of detail
- data about distribution and use of germplasm
 - institute (internal), national, international administration (e.g. Treaty)

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- coding systems for *ex situ* PGR documentation
 - hardly any standard coding systems, controlled vocabularies or ontologies are available
 - MCPD contains or refers to coding systems for a/o countries, origin types, population types
 - institute codes remains a problem – attempt of World Information and Early Warning System on Plant Genetic Resources (WIEWS)
 - trait names – several standards are available (Bioversity, UPOV) – rather low acceptance, incomplete crop coverage, inconsistency amongst standards
 - no systems for taxonomy, or other descriptors (user type, coordinate type, etc.)

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- systems at national level
 - most prominent: National Inventory
 - initiated by EPGRIS project as part of the establishment of EURISCO
 - established in most European countries – 1.1 million accessions in 38 countries
 - standardized data (MCPD)
 - act as national interface to systems at regional or global level
 - not always optimally accessible individually - all accessible via EURISCO
 - content at discretion of National Focal Point (NFP)
 - restricted to passport data
 - good entry point for coordination and capacity building

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ systems at regional level

● European Central Crop Databases (ECCDBs)

- long history – role of ECPGR
- role in collaboration (GENRES projects)
- very many established: 62 ECCDBs with 0.75 million accessions
- mainly passport data, some (12 ECCDBs) with C&E data
- data come directly from local systems plus EURISCO
- wide variety with regard to completeness, data quality, age of datasets, inclusion of C&E data, and possibility to search or download them via the web



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ systems at regional level

● EURISCO

- established by EU project EPGRIS (2000-2003)
- now under responsibility of ECPGR
- managed by Bioversity in collaboration with National Focal Points of National Inventories
- 1.1 million accessions from >240 holding institutions in 38 participating countries
- new interface recently introduced
- data from National Inventories – some are old !



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- changes in the landscape
 - requirements International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and the Convention on Biological Diversity (CBD)
 - registration of MLS material – EURISCO can easily accommodate
 - reporting of transactions - EURISCO might play a role
 - new global Accession Level Information System (ALIS) is being established
 - EURISCO can act as data source
 - AEGIS requires a level of data management
 - EURISCO can easily accommodate

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- changes in the landscape
 - role of ECCDBs is changing from passport data gathering points to crop specific PGR entry points
 - C&E data and research results (markers etc.) are more relevant
 - new roles might arise (catalyze crop groups activities: improving data quality at data source, supporting activities related to AEGIS, creating crop portals)
 - new role requires EURISCO to be the one-stop-shop for passport data

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ technological changes

- quality of data is getting higher
 - experts are 'closer to the data'
 - higher exposure of data
- other types of data are becoming available and required
 - the user requires C&E data
 - making these accessible proves to be a major challenge
 - molecular and other types of data are being generated
- more services can be provided
 - on-line access allows on-line ordering and handling MTAs

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ technological changes

● establishment of virtual genebanks

- data sources and data providers are separated – web-services allow direct access for computers to databases
- user can search and order material from a combination of genebanks without needing to know where the data / material comes from
- technology is available (ref GBIF), agreement on the policy level and an upgrade of most local systems is required

● crop portals

- changing the focus from the data provider (genebank) to specific user groups (breeders / scientists / policy makers)
- data from a variety of sources, incl. genebanks, scientific literature

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ technological changes

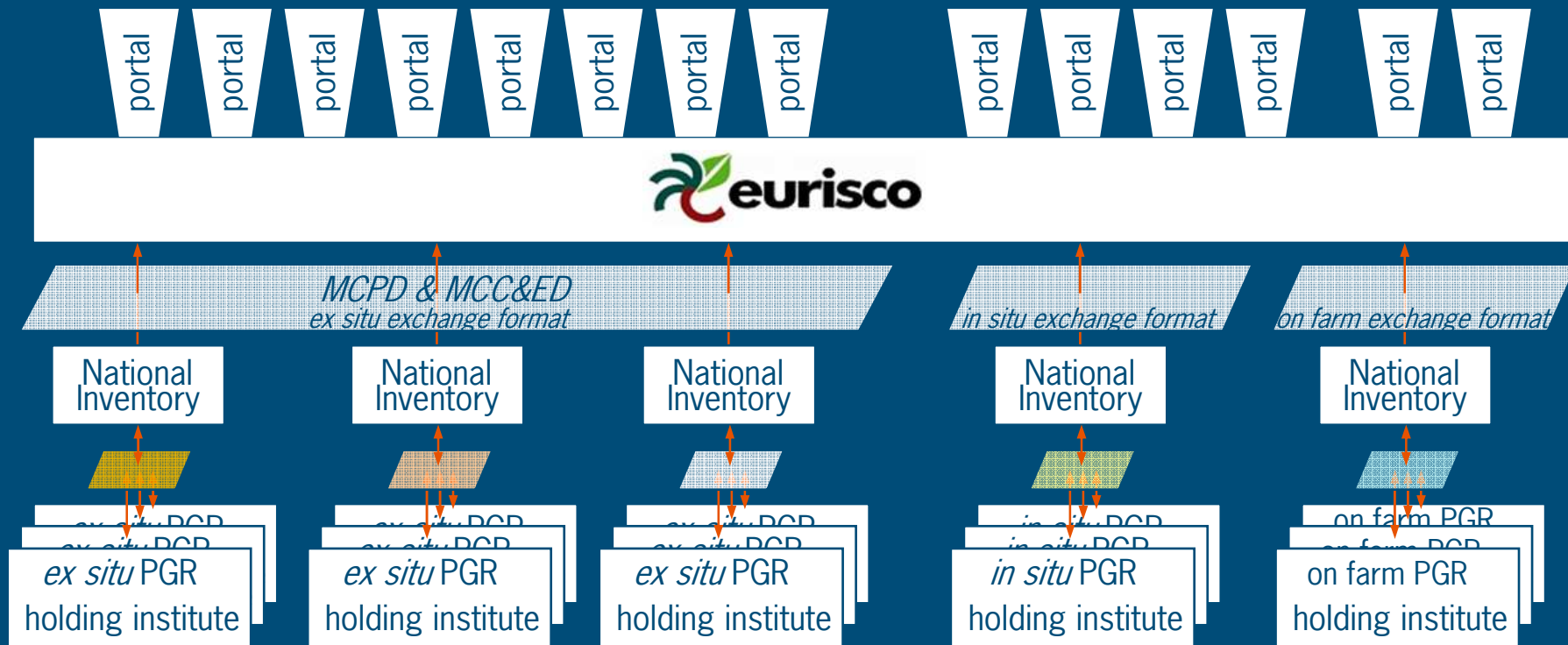
● relationship EURISCO – ECCDBs

- EURISCO should be the 'source' for passport data, giving web-service access - but: National Inventories need to be completed
- ECCDBs should develop into crop portals
 - more interfacing
 - more services
 - less databasing

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ the dreamscape



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ next steps

● standards

- develop and adopt more and better standards: e.g. C&E data
- expansion of MCPD to accommodate Treaty and AEGIS requirements
- compliance to Access to Biological Collection Data (ABCD)
- introduction and use of life science identifiers (LSID)

● technology

- adopt existing technology to PGR community - establish few testing sites
- invest in open source genebank documentation system (possibly GRIN-Global provides a starting point)

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ next steps

● capacity building

- many collaborating institutions still lack technical and personal capacity: need for teaching and training materials, teaching workshops, staff exchanges and other capacity building activities
- for a start: NFPs and ECCDB managers provide a good audience

● improve data quality

- garbage in – garbage out
- institutions should concentrate on improving their own data quality

Centre for Genetic Resources, the Netherlands

PGR Information Landscape

■ coordination

● current actors

- ECPGR Documentation and Information Network Coordinating Group
 - EPGRIS3 (a self-funded initiative)
 - Global Crop Diversity Trust
 - Generation Challenge Programme (GCP)
 - Global Public Goods Programme (GPG2)
- Europe could benefit more from these programs if more priority would be given, and more capacity would be made available to PGR documentation at all levels: institutional, national and European



Centre for Genetic Resources, the Netherlands

PGR Information Landscape

- concluding remarks
 - PGR documentation is crucial for PGR use and coordination of PGR activities
 - the technology is available
 - obstacles are
 - low data quality
 - low standardization
 - lack of technical knowledge
 - increasing priority of PGR documentation makes complete sense

Centre for Genetic Resources, the Netherlands