GGCE → EURISCO → Genesys

Matija Obreza
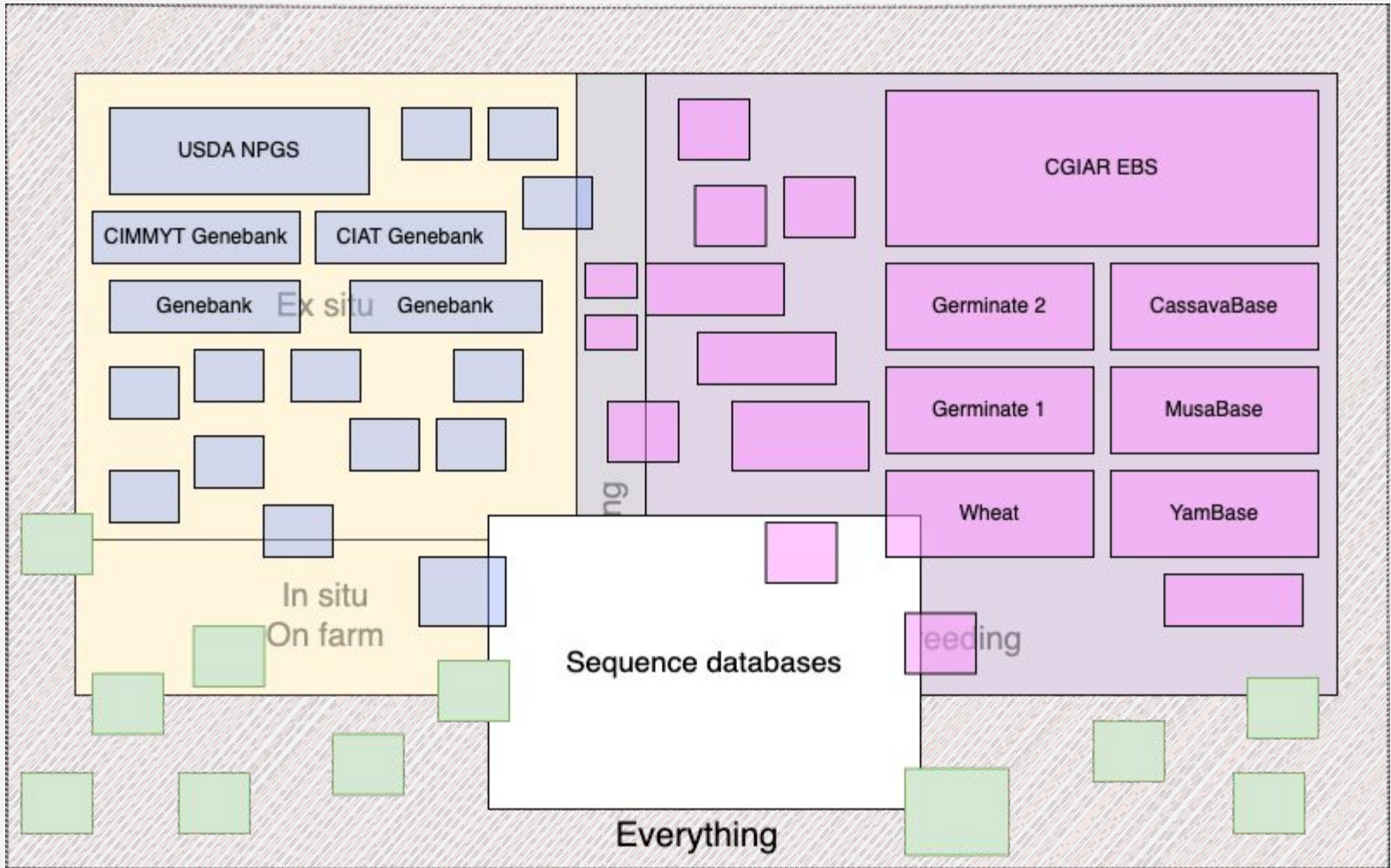
ECPGR GRIN-Global Workshop, October 2022
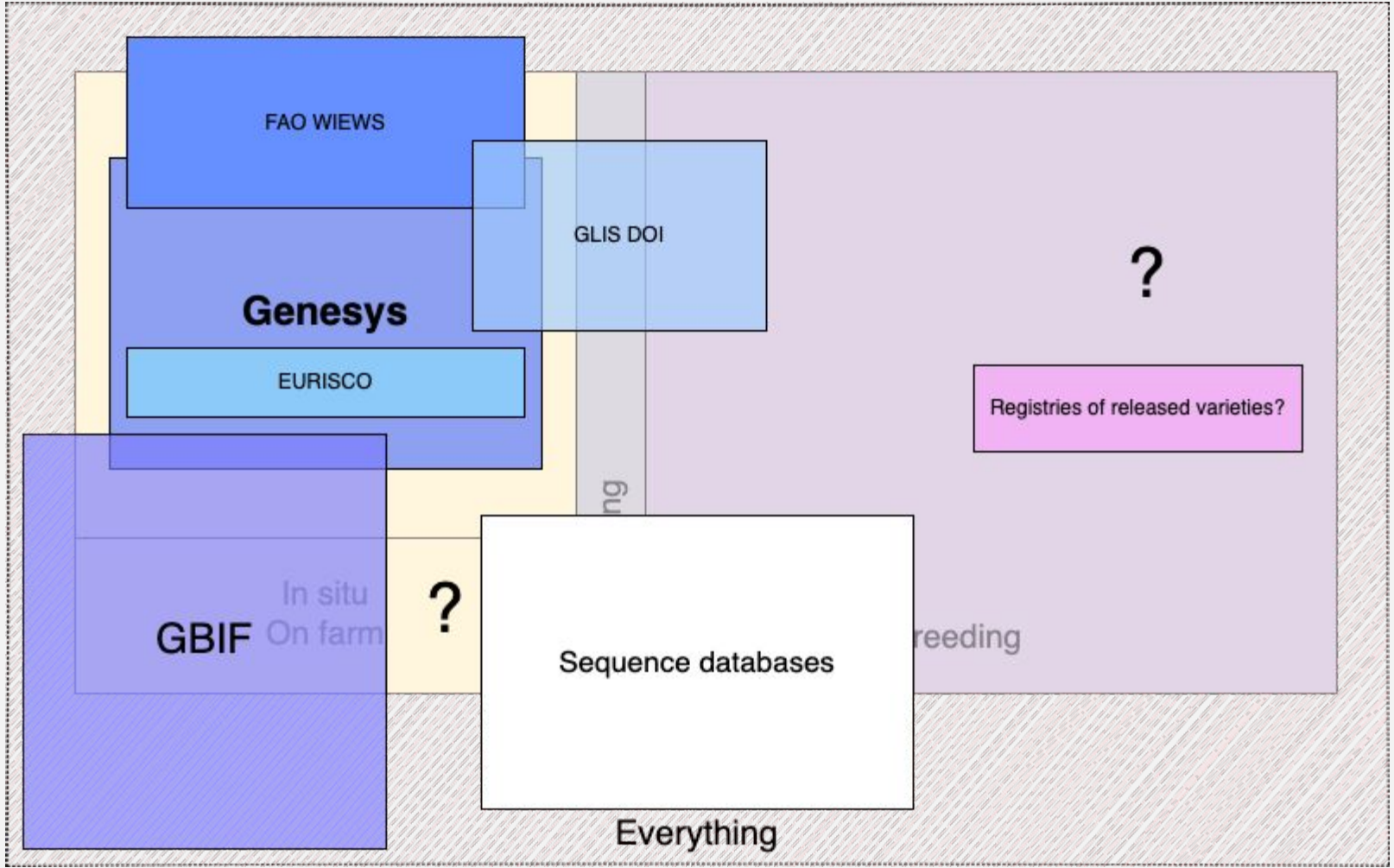
CROP TRUST

PGRFA domain

Ex situ

In situ
On farm

Prebreeding

Breeding

Everything
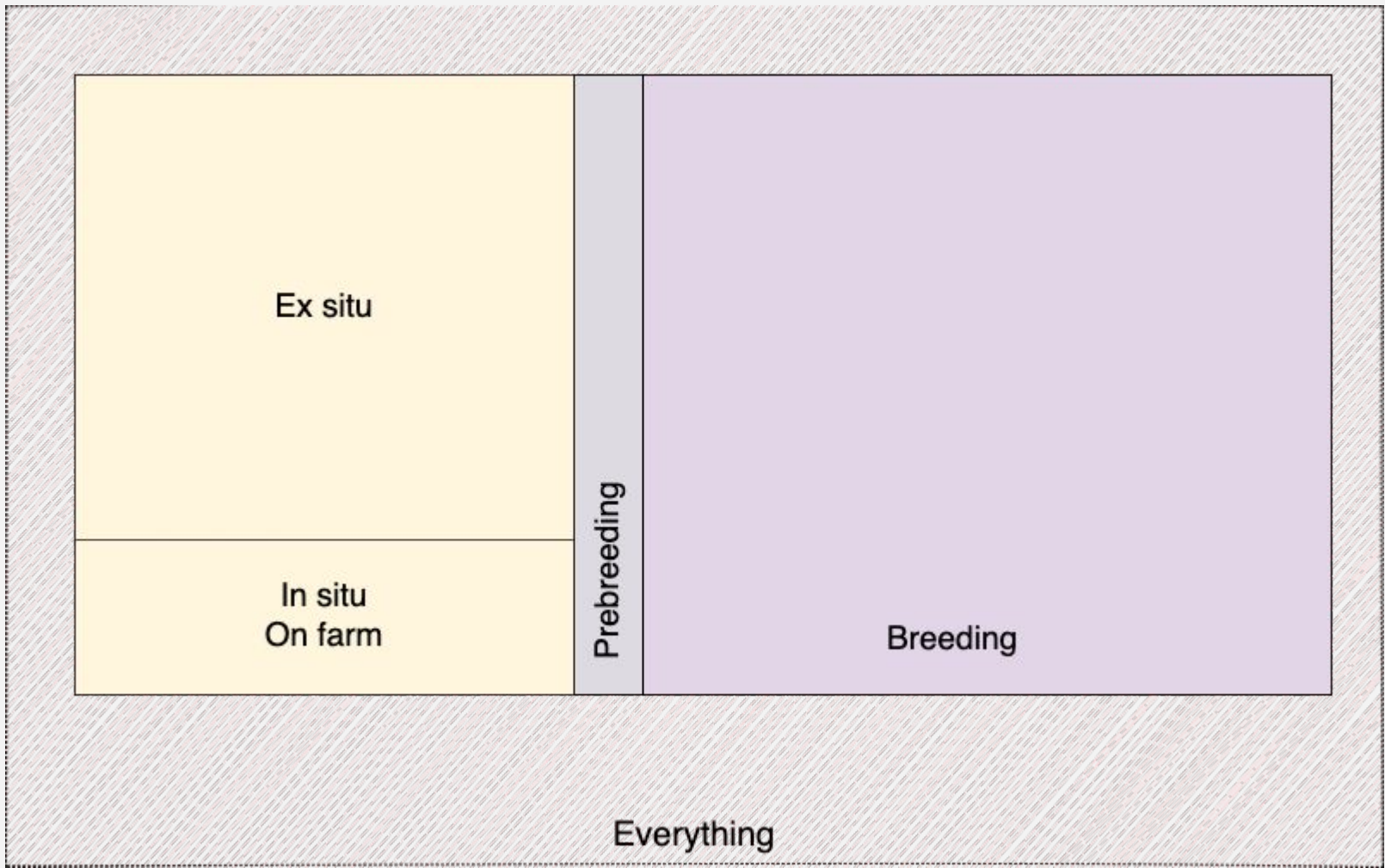
Database hell

Organized chaos

GGCE / EURISCO / Genesys

Internal genebank data management

Every genebank needs an information system to effectively maintain their collections.

- Requires **a lot of resources**
- People **retire**
- Technologies **change**

Adopt (not adapt) an existing system!

- Learn from the community
- Share with the community

Publicly accessible databases

Standardization: Data from different genebanks is made available to users in a standard format.

Safety and **security**: Only selected data is exposed to the public. A copy of the data is hosted in a separate database. Master data is behind a firewall.

Convenience: More data means better utility for users of genebanks and easier discovery of available material. Economy of scale.
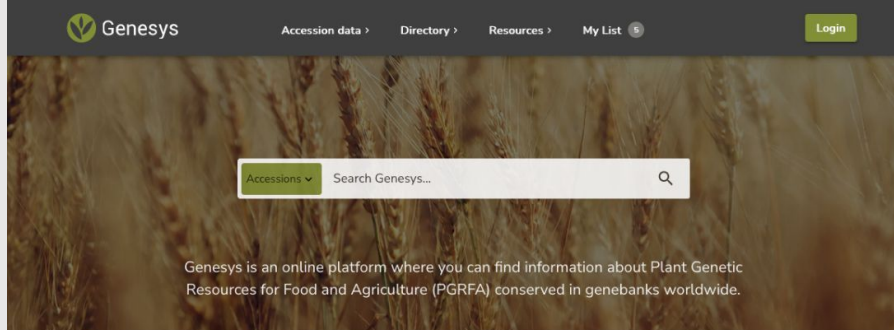
CROP TRUST

Genesys

~ Est. 2011 ~

www.genesys-pgr.org

Genesys is an online platform where you can find information about Plant Genetic Resources for Food and Agriculture (**PGRFA**) in crop genebanks worldwide.

GRIN-Global and **Genesys** development was initiated in **2008**, in collaboration with USDA and Bioversity, respectively.

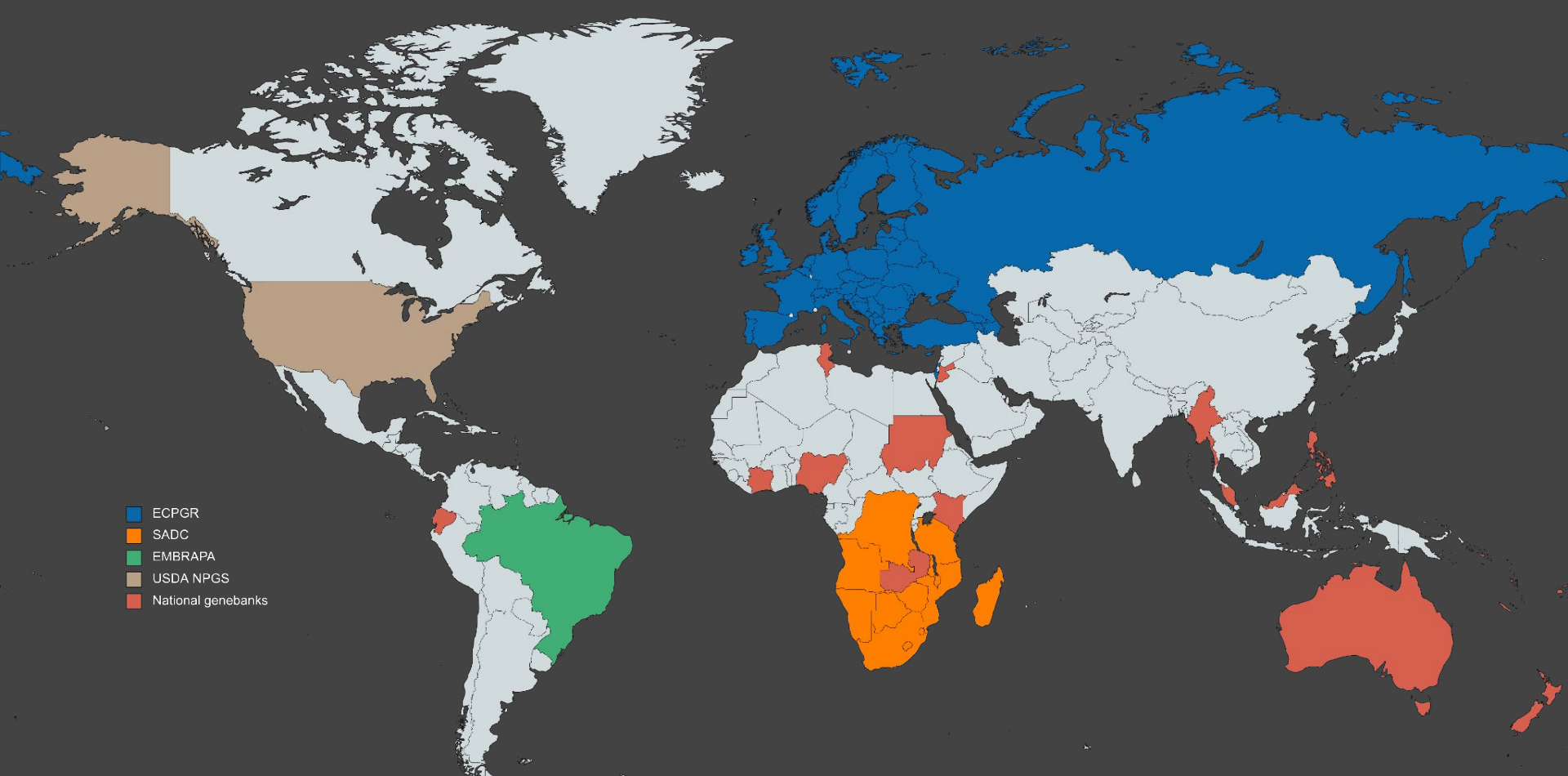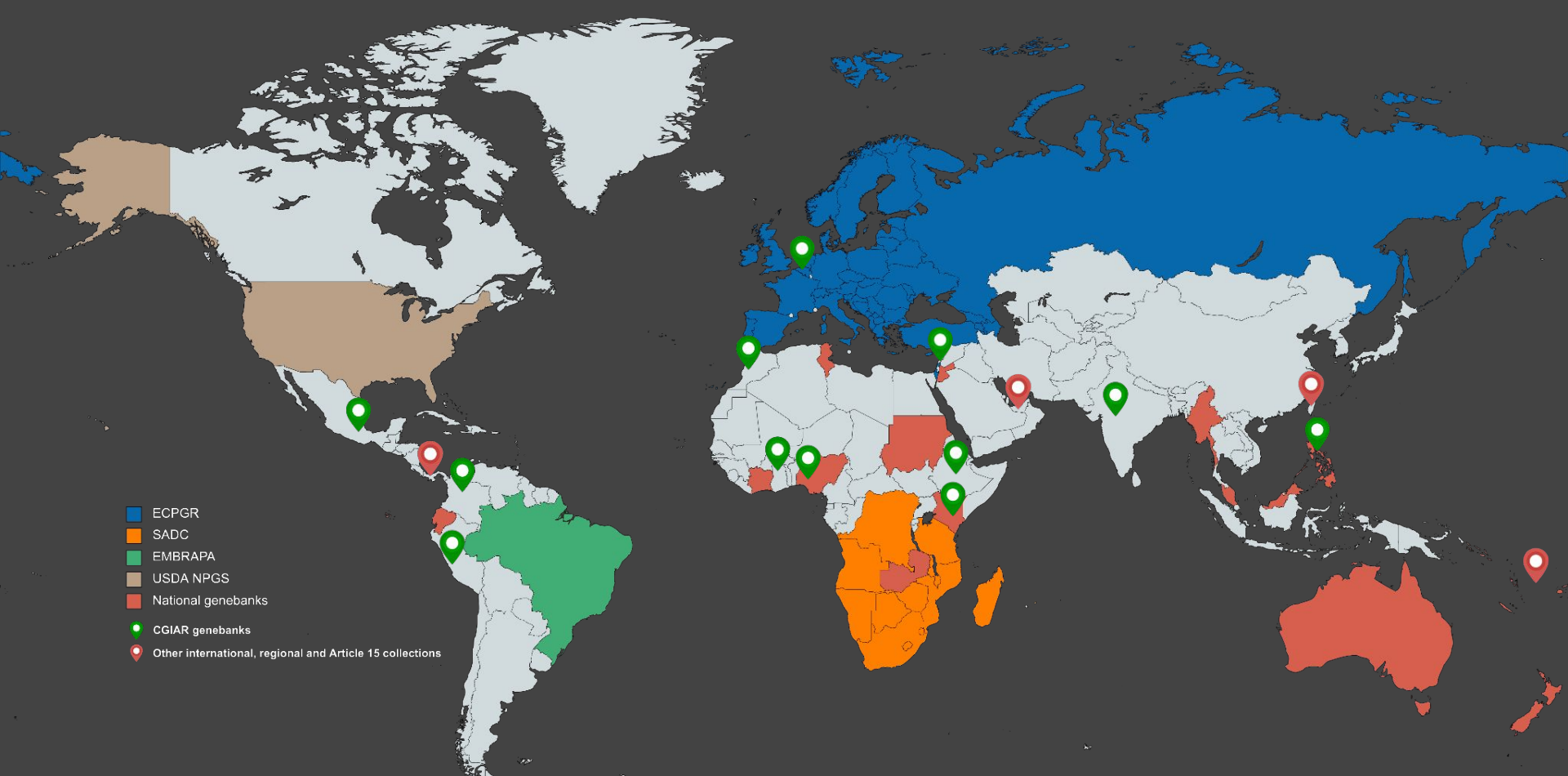CROP TRUST

Scope of Genesys

Genesys focuses on **accession-level** data about the material in long-term *ex situ* conservation.

Genesys **does not** host information on *in situ* germplasm, breeding materials or released varieties, unless formally part of genebank collections.

Institute code	Holding insti...	Accession number	Accession name	Taxonomy	Crop name	Crop	Biological status of accession	Provena
CIV033	Africa Rice Ce...	WAB0001126	TOG 5284	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Ivory Coa
CIV033	Africa Rice Ce...	WAB0001255	TOG 5552	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Nigeria
CIV033	Africa Rice Ce...	WAB0001407	TOG 6196	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Guinea
CIV033	Africa Rice Ce...	WAB0001413	TOG 6203	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Guinea
CIV033	Africa Rice Ce...	WAB0001421	TOG 6211	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Nigeria
CIV033	Africa Rice Ce...	WAB0001430	TOG 6221	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Nigeria
CIV033	Africa Rice Ce...	WAB0001443	TOG 6235	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001491	TOG 6285	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001544	TOG 6339	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001547	TOG 6342	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001551	TOG 6346	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001553	TOG 6349	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001563	TOG 6359	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001642	TOG 6441	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Mali
CIV033	Africa Rice Ce...	WAB0001647	TOG 6449	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Mali
CIV033	Africa Rice Ce...	WAB0001649	TOG 6456	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Mali
CIV033	Africa Rice Ce...	WAB0001664	TOG 6472	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001681	TOG 6489	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001689	TOG 6497	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001698	TOG 6506	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001712	TOG 6520	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001734	TOG 6542	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001748	TOG 6556	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001749	TOG 6557	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001765	TOG 6573	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001781	TOG 6589	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia
CIV033	Africa Rice Ce...	WAB0001819	TOG 6629	*Oryza glaberrima* Steud.	Rice	Rice	Traditional cultivar/Landrace	Liberia

Networks and national genebanks

Legend:
- ECPGR
- SADC
- EMBRAPA
- USDA NPGS
- National genebanks

CROP TRUST

Created with mapchart.net

International, regional and ITPGRFA Article 15 collections

Legend:
- ECPGR
- SADC
- EMBRAPA
- USDA NPGS
- National genebanks
- CGIAR genebanks
- Other international, regional and Article 15 collections

CROP TRUST

Created with mapchart.net

Tools for genebanks

https://validator.genesys-pgr.org

Spelling of taxonomic names based on GRIN Taxonomy

GIS test of coordinates, country of provenance

GENUS	GENUS_check	SPECIES	SPECIES_check	GRINTAX_speciesId	GRINTAX_species
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolo	bicolor		
Sorghum	OK	bicolor	OK	35092	true
Sorgho	Sorghum;Sorgum	bicolor	(Sorghum) bicolor		
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true
Sorghum	OK	bicolor	OK	35092	true

CROP TRUST

Exploring PGRFA data

- Full-text search and filtering

- Maps

- Data summaries (Overviews)

- Finding accessions by climate

- Finding similar/duplicate accessions

- History of updates to passport data

- Initiating requests for material

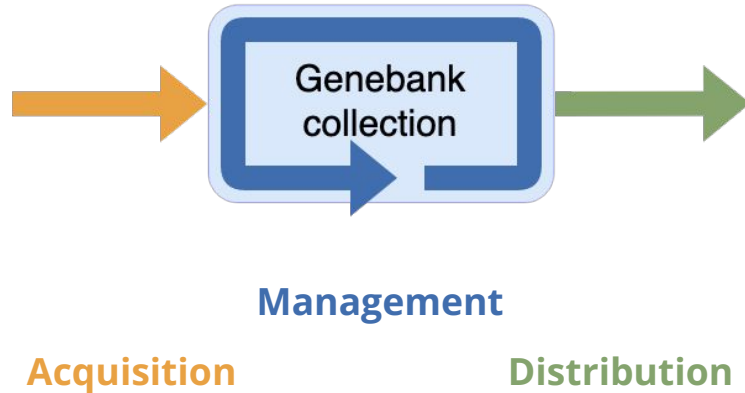- *Coming soon:* Subsetting Tool

CROP TRUST

Accession map

About 44,673 accessions   Excluding Historical ⊗   Genus: Zea ⊗

OVERVIEW   ACCESSIONS   MAP   IMAGES   DOWNLOAD KML

Map of collecting sites of *Zea* spp.

Link

Genesys and genebank operations

Management

Acquisition   **Distribution**

Acquisition (improving collections):

- Which material is already conserved?
- What is unique? Gap analysis
- Any updates relevant to material received from other genebanks?

Management (updating Genesys):

- Availability
- Type of germplasm maintenance

Distribution (for use):

- Finding and requesting material

Beyond passport data

- Accession **images** and scanned collecting forms

- Collection **subsets** (core- and mini-core collections, reference sets, etc.)

- **Datasets** of trait observations annotated with trait descriptors and other metadata

Aegilops tauschii mini-core collection

The Wheat Genetics Resource Center at Kansas State University has genotyped a collection of 600 *Ae. tauschii* accessions and selected a set of 40 lines that are the most genetically diverse and capturing more than 95% of the allelic diversity. This mini-core includes 28 accessions from Lineage 1 [strangulata type] and 12 from Lineage 2 [eusquarrosa type]. The mini-core represents 13 countries spanning the distribution of *Ae. tauschii*; three accession are of unknown origin; 27.5% are from Iran and 17.5% are from Afghanistan, countries with the largest density of this species.

Single nucleotide polymorphism (SNP) discovery and genotyping was performed in single step with Tassel 5 GBSv2 pipeline using the *Ae. tauschii* genome assembly (Aet v4.0; NCBI BioProject

PRJNA341983) as the reference. Tassel was run with bowtie2 aligner for tags mapping in Linux HPC environment via shell script. Population level SNP filtering was performed and SNPs with a minor allele frequency (MAF) less than 0.01 and missing data more than 20% were removed. SNPs with heterozygosity greater than 5% were removed because *Ae. tauschii* accessions are highly inbred. Individual samples with more than 80% missing SNP calls and more than 5% heterozygosity were also removed.

All SNPs were used to select a representative core-set from the *Ae. tauschii* collection. The core-set was selected in two steps. First, the software package PowerCore was used with default settings, which selects the lines to retain most diverse alleles by

implementing advanced M (maximization) strategy. Then the number of selected accessions was further reduced with a phenotypically guided selection using the available phenotypic data for a leaf rust composite culture, stem rust race TTKSK, and Hessian fly biotype D resistance. The diversity captured by the MiniCore was assessed by the percent segregating SNPs present in the selected accessions relative to the whole collection.

Reference:
Genomic analysis confirms population structure and identifies inter-lineage hybrids in *Aegilops tauschii*. 2019. Singh N, Wu S, Tiwari V, Sehgal S, Raupp J, Wilson D, Abbasov M, Gill B, and Poland J. Frontiers in Plant Science. https://doi.org/10.3389/fpls.2019.00009

Type of subset	Selective
Method of selection	Genoyping-by-sequencing
Crop	Wheat
Number of accessions	40
Institute	Wheat Genetics Resource Center
Creation date	2015
Source	☑ https://doi.org/10.3389/fpls.2019.00009

Subset metadata

Subset creators

Data curator	Eduard A
Data curator	Bikram S
Data curator	Jesse Po
Data manager	W. John Raupp Kansas State University
Data curator	Sunish Sehgal South Dakota State University [current affiliation]

CROP TRUST

Public genebank websites

Embedded Genesys

Instead of hosting and maintaining a separate database, paying for a web server and setting up a new domain name, genebanks can use Embedded Genesys to integrate their data from Genesys directly into existing institutional websites with a few lines of Javascript.

The latest feature added to Embedded Genesys is the Similarity Search. When an accession is marked as unavailable, users can search for similar accessions and request one of the provided alternatives instead.

Summary

Genebank + GGCE

- Manage all data related to PGRFA in your collection and effectively operate your genebank
- Community knowledge exchange, support, technical solutions

ECPGR + Eurisco

- Regional
- Knowledge exchange, standards, support, technical solutions

Genesys

- Cross-regional and global
- Knowledge exchange, standards, support, technical solutions

CROP TRUST

www.genesys-pgr.org

What's new?
https://www.genesys-pgr.org/content/news

Webinar: How to use Genesys?
https://www.genesys-pgr.org/content/news/129/recording-how-to-use-genesys

Webinar: Subsets and trait data in Genesys
https://www.genesys-pgr.org/content/news/136/recording-subsets-and-trait-data-in-genesys

Searching for similar accessions
https://www.genesys-pgr.org/content/news/104/searching-for-similar-accessions

Get in touch!
helpdesk@genesys-pgr.org

CROP TRUST

www.croptrust.org