Germinate
A Common Platform for Management of PGR Experimental Data

PAG 30 Information flows to harness plant genetic diversity
13th January 2023

Paul Shaw
The International Barley Hub/James Hutton Institute Scotland
The challenge

- Our science is producing larger and more complex datasets.
- We need to provide the infrastructure and tools to store, explore, interact and understand these large datasets.
- We need to develop tools to help provide insight into complex biological data.
- We need to provide easy access to huge volumes of data in digestible chunks.
- Show relationships between data that may not be visible in raw data.
- It’s not about telling people what to think but giving them tools to look for patterns and visual cues.
In a nutshell...

From this...

To this...
“Making experimental data available to breeders, scientific researchers and farmers in approachable and tailored ways will increase the utility and value of the data and knowledge the project will generate ensuring diversity is more accessible for exploitation.”
But how to help try to achieve this?

2. Merge and Wrangle
(Historical and background data)

1. Collect

3. Store

4. Visualize
(and export)

GridScore
Humbug
Germinate Daim / Data templates / BrAPI
Helium
Flapjack
GridScore

Stop writing stuff down!

- Less error-prone data entry.
- Voice feedback using text-to-speech.
- Voice recognition for audio notes.
- Automatic tagging with timestamp, identifier and GPS.

<table>
<thead>
<tr>
<th>Heris2</th>
<th>Heatmap</th>
<th>Trait statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enter data for each trait below. Add comments at the bottom or tag photos with the current location and timestamp.</td>
<td>Please select a trait in the selection box below. The shown graph shows the trait layout where each colored box represents a plot. The color of each plot is shown on a gradient scale. For numerical traits, this will represent the actual value. For categorical traits and text traits, this shows the scoring data as the number of scores per plot. For categorical traits, the color represents the category value.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trait</th>
<th>Plant height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>84.5</td>
</tr>
</tbody>
</table>

2021-03-18_16-19-11-Birka1.jpg
QR-based data sharing

- Share trial setup and data in seconds using QR codes.
- No mucking about – same setup for all!
But how to help try to achieve this?

2. Merge and Wrangle
(Historical and background data)

1. Collect

3. Store

4. Visualize
(and export)

GridScore
Humbug
Germinate Daim / Data templates / BrAPI
Helium
Flapjack
FAO GLIS/Competitors

Genebanks

Pre-breeding Projects

Individual Labs

Germinate

Shallow v Deep Data
Wide and shallow – Narrow and deep

(A) Germplasm Collection Management
- Seed, vegetative tissue and reference samples
- Passport data
- Collection metadata
- Supplementary data
  - Phenotype
  - Genotype

(B) Germplasm Gateways
- Passport data
- Phenotype data

(C) Integration and Pre-Breeding Tools
- Pre-breeding collections
- Phenotype data
- Genotype data
- Add-in visualization tools

Database Solutions for Genebanks and Germplasm Collections 2023
Paul Shaw, Stephan Weise, Matija Obreza, Sebastian Raubach, Susan McCouch, Benjamin Kilian, Peter Werner

(D) Users
- Create and Feedback
  - Phenotypes
  - Genotypes
  - Pedigrees

Researchers

Breeders
Charts
These charts compare the performance of individual genotypes per trait. Each section on the x-axis represents a different genotype. Grouping is based on the chosen color coding option.

Blast score (1-9) [acumetic]

Dead plants [percent]

Resistance [%] [percent]
Germinate species diversity

Barley
Berrybase
Demo
CFC
Grasses
CWR Cowpea
CWR Finger Millet
CWR Grass Pea
CWR Pigeon Pea
CWR Chickpea
CWR Alfalfa
CWR Lentil
CWR Pearl Millet
CWR Barley
CWR Wheat
CWR Sorghum
CWR Eggplant
CWR Rice
CWR Sunflower
CIMMYT Maize
CIMMYT Wheat
Rye
Germinate – RESAS JHI-B1-2 BARGAIN

UAV, ionomics and environmental sensor data.
Tubers cylindrical, about 5.5 x 2.5 x 8.0 cm. Roots very numerous, long, hollow, the tuber slightly swollen between them. Skin black with a reddish sheen. Flesh pale yellow. Yucca ring slightly tinted purple.
Current Projects using Germinate

- Horizon Europe BEST-CROP (2023-2028)
- Crop Trust BOLD (2021-2029)
- RESAS BARGAIN (2022-2026)
- BBSRC Barley Travel Award (2021-2024)
- Horizon 2020 BreedingValue (2021-2024)
- Environmental Indicators (2021-2023)
- Sustainability in Education and Agriculture using MixtureS (2021-2023)
- DivSeek (2020)
- SusCrop ERA-NET BRACE (2021)
- SusCrop ERA-NET WheatSecurity (2023)
- SusCrop ERA-NET ProGrace (2023)
- OzBarley (Australia)
- ICBA Genebank (UAE)
- Quinoa (KAUST)

[Helium](https://helium.hutton.ac.uk) Visualization, exploration and analysis of large scale pedigrees.
Summary

- All our software is free to use, most is open source.
- Germinate is under active development and we welcome contributions, feedback and ideas on how we can improve.
- We are looking to develop better links to tools like Genesys, GrinGlobal, EURISCO, Ensembl & GIGWA.
- BOLD will allow us to take these resources to the next level building on the experiences from the CWR project – this in turn benefits all Germinate users.
- Please just give us a shout if you want to try Germinate or want an extended demo.
- Contact myself or Sebastian Raubach for more information!


Acknowledgements

Hutton Barley, Potato and Soft Fruit Groups

- Sebastian Raubach
- Gaynor McKenzie
- Luke Ramsay
- Robbie Waugh
- Joanne Russell
- Malcolm Macaulay
- Ali Karley
- Niki McCallum
- Glenn Bryan
- Micha Bayer
- Miriam Schreiber
- Pauline Smith
- Derek Stewart
- Iain Milne
- Kelly Houston
- Susan McCallum
- Kaye Smith
- Julie Graham
- Amy Learmonth
- All our funders and partners!
- + Apologies to those who I may have missed.

@GerminateHub
https://ics.hutton.ac.uk/get-germinate
germinate@hutton.ac.uk