

The Triticeae Toolbox: Small Grains Breeding Database



David Waring
Jean-Luc Jannink
Clay Birkett



Outline



1. Where are we going?
 - T3 is FAIR
 - T3 minimizes data problems
2. How will we get there?
 - Breeders contribute phenotypes
 - Phenotypes are inter-related by alleles
3. What are the perks?
 - The latest data is accessible
 - Tools facilitate breeding tasks
 - Analyses best done centrally can be

Where are we going?

Taxpayers pay for WheatCAP data



T3 is FAIR:

Findable

Metadata and data should be easy to find

Accessible

Once a user finds the data, it should be easily accessible

Interoperable

Data should be integrated with other data

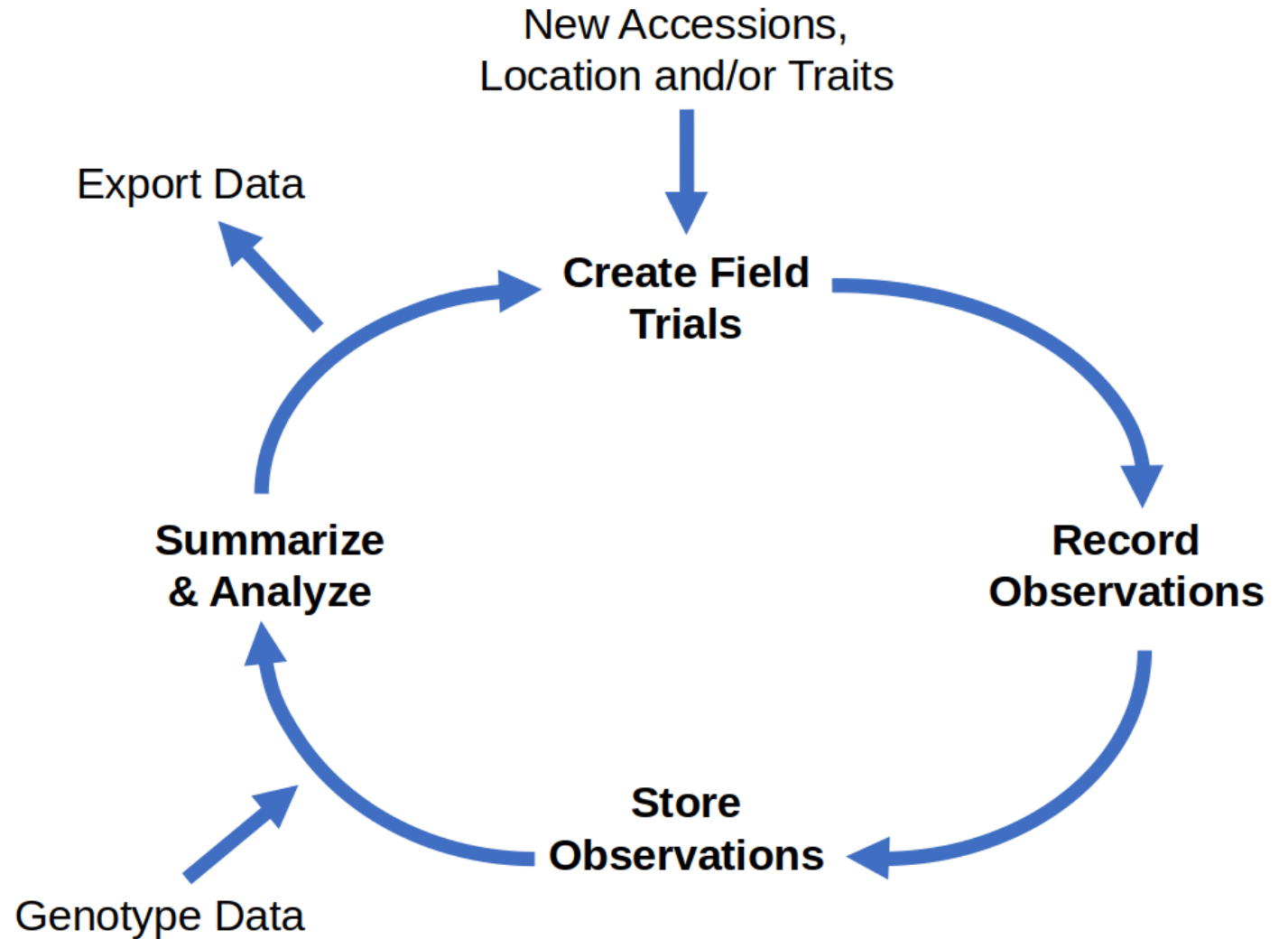
Reusable

Users should be able to combine data for new uses

T3 minimizes sources of error



- From start to finish no source of data corruption
- Tools for each step:
 - Crossing
 - Trial Design
 - Data Collection
 - Summarize



Everyone has access



<https://wheat.triticeaetoolbox.org>

(<https://wheatcap.triticeaetoolbox.org>)

<https://barley.triticeaetoolbox.org>

<https://oat.triticeaetoolbox.org>

How will we get there?

Uploading Phenotypes



General Workflow:

1. Create Upload Templates
2. Optional: Upload Templates to Sandbox
 - Test the format of the files
 - See how data is presented
3. Submit the Templates
 - Email directly to djw64@cornell.edu
 - Phenotype Submission Form

wheat-sandbox.triticeaetoolbox.org
barley-sandbox.triticeaetoolbox.org
oat-sandbox.triticeaetoolbox.org

Uploading Phenotypes



Example Templates

Download Excel templates from homepage

Blank Templates

- Have correct column headers
- Contain no data

Example Templates

- Have correct column headers
- Contain sample data from workshop



The screenshot shows the T3/Wheat website interface. The browser tab is labeled 'T3/Wheat' and the URL is 'https://wheat.triticeaetoolbox.org'. The navigation bar includes links for 'T3/Wheat', 'Search', 'Manage', 'Analyze', 'Maps', 'About', 'Contact Us', 'Register', and 'Login'. The main content area is divided into two columns. The left column, titled 'Templates', contains two sections: 'Blank Upload Templates' and 'Example Templates'. Both sections list the same set of available templates: 'Locations (.xls)', 'Accessions (.xls)', 'Pedigrees (.txt)', 'Trials (.xls)', and 'Phenotype Observations (.xls)'. The right column, titled 'Submit Phenotype Data', provides instructions on how to submit data, including a link to the 'Phenotype Submission Form' and a button to 'Visit the T3/Wheat Sandbox'. At the bottom, there is a 'Submit Genotype Data' section and a 'User Guides' section with links to 'Read Documentation' and 'Watch Tutorials'.

Template Submission



- Accessions
- Trials
- Observations

Phenotype Submission Form

- Submit templates
 - We curate (errors / issues / duplicates)
 - We upload
-
- Get help from your student 😊

The screenshot shows the T3/Wheat website interface. The top navigation bar includes links for Search, Manage, Analyze, Maps, About, Contact Us, and user options. The left sidebar contains a 'v Watkins language collection' section with links like 'Exome capture Watkins', 'WheatCAP Regulatory', and 'GBS KSU 2020'. Below this is a 'Site update' section dated November 1, 2020, with links for 'Analysis', 'Population Structure (PCA)', 'Clustering (k-means)', 'Manage (contact us to use these features)', 'Mixed Model analysis', 'Seed Lots', and 'Crosses'. The main content area features a 'Submit Phenotype Data' section with instructions on how to submit data, a 'Submit Genotype Data' section, and a 'User Guides' section. A blue button labeled 'Phenotype Submission Form' is prominently displayed in the main content area, with an orange arrow pointing to it from the left sidebar. The bottom of the page includes a 'Read Documentation' button.

WheatCAP Trials



Breeding program	Trials total	2021 pheno trials	2022 pheno trials
Colorado State University	15	13	0
Cornell University	1	0	1
Kansas State University	5	0	0
Michigan State University	59	0	0
Montana State University	1	0	0
Purdue University	3	0	3
South Dakota State University- Winter Wheat	6	0	6
TAMU	3	0	0
UC Davis	20	0	18
University of Idaho	1	0	1
University of Illinois	57	27	26
University of Nebraska	9	0	3
USDA-ARS Kansas	8	0	8
Utah State University	7	0	3

Some perks

Search Wizard

Seedlots

Barcodes

Data Collection Apps

Imputed Genotypes

Search Wizard



Menu: Search > Wizard

- Filter / Subset
- Combine
- Download phenotypes
- Download genotypes

Search Wizard

Don't see your data? [Refresh Lists](#) [Update Wizard](#)

Traits

fhb

Select All 1/228 Clear

+ FHB incidence - 0-9 percent

+ FHB ISK index - %|CO_321:00

+ FHB plant response - 1-9 res

+ FHB severity - %|CO_321:00

+ FHB severity - Greenhouse -

×

FHB incidence - %|CO_321:00

Match ANY ALL

Add to List... Add

Create New List Create

Locations

IL

Select All 3/42 Clear

+ Clarksville, MD

+ Evansville, IN

+ Fayetteville, AR

+ Merrill, MI

×

Champaign, IL

×

Highland, IL

×

Urbana, IL

Match ANY ALL

Add to List... Add

Create New List Create

Trials

Search

Select All 63/63 Clear

×

5STADV_2001_Urbana

×

5STADV_2002_Urbana

×

5STADV_2003_Urbana

×

5STADV_2004_Urbana

×

5STADV_2005_Urbana

Match ANY ALL

Add to List... Add

Create New List Create

Accessions

Search

Select All 0/1529 Clear

+ 011007A1-14

+ 011007A1-14-2

+ 011007A1-14-44

+ 011010A1-15

+ 011034A1-3

Match ANY ALL

Add to List... Add

Create New List Create

Seedlots



Seedlot AJAX-2002-SL1

Details

[\[Edit Seedlot Details\]](#)

Breeding Program	Cornell University
Seedlot Name	AJAX-2002-SL1
Seedlot Description	Collected on June 1, 2022
Organization	Cornell University
Location Code	Room 105
Box Name	Box A
Quality issues	
Contents	AJAX (accession)
Current count	500
Current weight (g)	NA
Submitters	dwaring87

Transactions

Transactions Table

[\[Add New Transaction\]](#)

Show 10 entries

Search:

Transaction Id	Transaction Date	From	To	Transaction Num Seeds	Transaction Weight (g)	Operator	Transaction Description	Options
1466620	Thu Dec 1 14:47:14 2022	AJAX (accession)	AJAX-2002-SL1 (seedlot)	+500	NA	dwaring87		[Edit]

Showing 1 to 1 of 1 entries

Previous 1 Next

Seedlot: a single packet of seed

Properties:

- Unique Name
- Location
- Box Name
- Accession
- Contents
 - Count
 - Weight

Transactions:

- Each time seed is added or removed
- Linked to Plots from the Seedlot
- Linked to Seedlots when splitting / combining

Barcodes



Barcodes can be generated for:

- Accessions
- Trials
- Plots
- Genotyping Plates

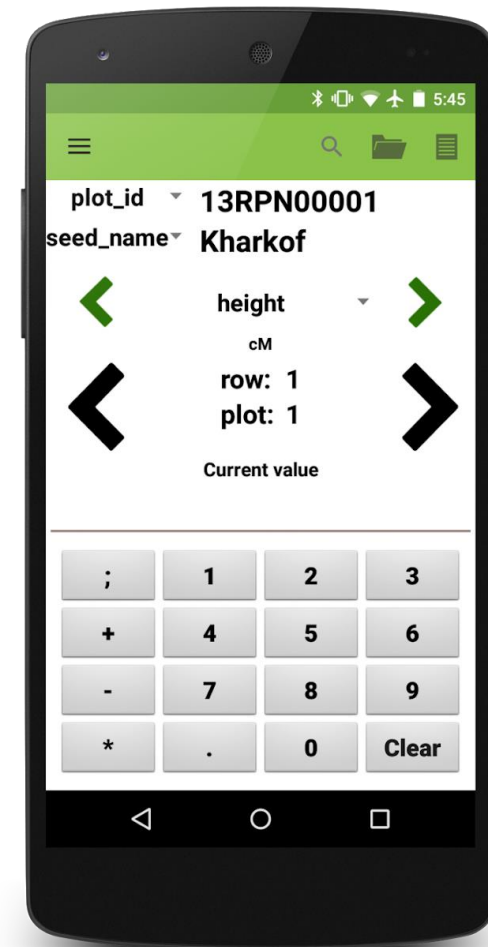
Use Cases:

- Barcoded Plots: Observations recorded for the correct plot
 - Unique barcode for each Plot
 - Android Field Book scans before recording observations
- Seedlots: No genotype mislabeling
 - Unique barcode for each Seedlot
 - Android Coordinate scans when collecting samples for Genotyping



Android App for Data Collection

- Load Fields & Traits from DB (via BrAPI or files)
- Record data for each plot in the app
 - Numeric
 - Counter
 - Categorical
 - Boolean
 - Text
 - Photos
- Integrated barcode scanner
- Export data to store in DB (via BrAPI or files)



Accessing imputed genotypes



files.triticeaetoolbox.org

Imputed protocols

Method: Practical Haplotype Graph - PHG V2

Description: [Imputation pipeline and accuracy](#)

Protocol	Genotype project	description
Infinium 90K v2.1	TCAP90K_HWWAMP	2.9M markers
	TCAP90K_SpringAM_panel	
	TCAP90K_YQV14	
	TCAP90K_CSRVAL14	
	VRND4_UCD_2015	
	TCAP90K_NAMparents_panel	
	TCAP90K_HWWAMP_SRPN	
Infinium 9K v2.1	NSGCwheat9K_4X	
	WorldwideDiversityPanel_9K	
GMS	2021_GMS	

Accessing imputed genotypes



T3/Wheat

Search

Manage

Analyze

Maps

About

Genotyping Protocols

Search

Select All 1/26 Clear

+ Genotyping by multiplexed sequencing (G)

+ Infinium 90K

+ Infinium 90K PHG

+ Infinium 9K

+ Infinium 9K v2.1

×

Infinium 90K v2.1

Genotyping Projects

Search

Select All 1/9 Clear

+ TCAP90K_CSRVAL14_V2

+ TCAP90K_HWWAMP_SRPN

+ TCAP90K_LeafRustPanel_V

+ TCAP90K_NAMparents_V

+ TCAP90K_SpringAM_panel

×

TCAP90K_HWWAMP_V

Related Genotype Data

Imputed Genotype Data available

The genotype projects have been imputed using the Practical Haplotype Graph (PHG). The imputed marker density is 1.3M markers for assembly RefSeq_v1 and 2.9M markers for assembly RefSeq v2.1. For quick access the imputed genotypes can be downloaded from a prepared file. A description of the imputation process can be found in [About PHG imputation](#).

1 protocol, selected protocol1 project, TCAP90K_HWWAMP_V2

Retrieve Genotypes

Download file

all accessions in project

Related Trial Metadata

Related Trial Phenotypes

Imputation roadmap



- The PHG is complicated software
- We are improving in our ability to use it to impute accurately
- The frontend to access imputed data will stay consistent
- Backend improvements will mean periodically new imputations

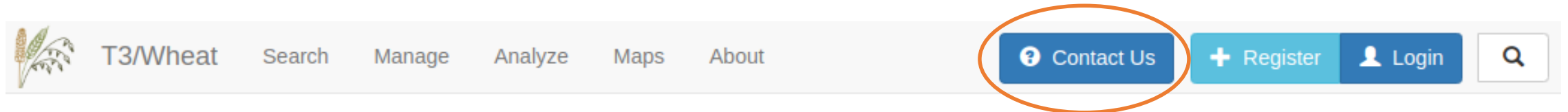
Questions?



Get In Touch:

Email: djw64@cornell.edu; clb343@cornell.edu; jeanluc.work@gmail.com

Website: "Contact Us" form



Survey

Interested in a more in-depth hands-on workshop?

Fill out this survey:



tinyurl.com/t3workshops

Acknowledgements

The Triticeae Toolbox

Jean-Luc Jannink

Clay Birkett

David Waring



Breedbase

Lukas Mueller's Lab

Boyce Thompson Institute

