



Table 1. Barley yields (bu/ac) of 2ND25276 and other barley cultivars grown in North Dakota barley breeding program yield trials, 2007-2014.

Entry	Location								Mean
	Carrington	Fargo	Langdon/ Osnabrock	McVille	Minot	Nesson Valley	Sidney	Williston	
Sta. Years	12	19	6	17	11	12	10	16	105
2ND25276	112.5	78.5	107.6	63.4	103.2	115.8	116.0	70.7	91.3
Conlon	101.5	65.2	97.6	59.5	81.9	100.3	105.8	73.1	81.7
Pinnacle	97.8	71.8	108.3	60.0	92.3	115.1	112.8	72.3	86.7
Lacey	112.3	74.8	100.8	58.6	98.6	112.1	116.4	70.2	88.4



Table 2. Malt quality comparisons of 2ND25276 and other barley cultivars grown in North Dakota yield trials, 2007-2014<sup>†</sup>.

Entry	Barley protein (%)	Plump kernels (%)	Malt extract (%)	Wort protein (%)	S/T (%) <sup>†</sup>	Diastatic power (°L)	Alpha- Amylase (20° DU)	Beta- glucan (ppm)
Sta. Yrs.	33	40	26	26	26	26	26	26
2ND25276	11.7	92.3	81.8	5.42	47.8	118	89.8	154
Conlon	13.3	93.9	80.3	5.14	39.5	138	78.0	366
Pinnacle	11.7	90.3	81.6	4.91	43.8	116	70.2	227
Lacey	13.7	85.8	79.4	5.68	42.6	189	77.5	125

<sup>†</sup>Data courtesy of the USDA-ARS Cereal Crops Research Unit, Madison, WI.

<sup>†</sup>Soluble protein to total protein ratio.



Table 3. Disease reaction in the greenhouse to spot-form net blotch disease for 2ND25276 and representative cultivars entered in the Mississippi Valley Barley Nursery (MVBN), 2010-2013<sup>†</sup>.

	Greenhouse Spot-Form Net Blotch (rating) <sup>‡</sup>				
	2010	2011	2012	2013	Mean
2ND25276	3.0	2.0	5.0	2.0	3.0
Morex	7.0	7.0	5.0	3.0	5.5
Robust	3.0	3.0	3.0	3.0	3.0
Legacy	3.0	3.0	3.0	3.0	3.0
Lacey	3.0	3.0	7.0	2.0	3.8
Tradition	5.0	3.0	7.0	3.0	4.5
Pinnacle	7.0	7.0	5.0	7.0	6.5

<sup>†</sup>Data courtesy of Agriculture and Agri-Food Canada – Cereal Research Centre, Winnipeg, Manitoba.

<sup>‡</sup>Reaction categories: 10=VS, 9=S, 7=MS, 5=MR-MS, 3=MR, 1=R

Table 4. Disease reaction to spot blotch disease for 2ND25276, representative cultivars and resistant and susceptible controls in field tests in 2011-2014<sup>†</sup>.

Entry	Row type	Field Spot Blotch (percent) <sup>‡</sup>				Mean
		2011	2012	2013	2014	
2ND25276	2	38.3	18.3	6.3	18.3	20.3
Conlon	2	30.0	53.3	20.0	20.0	30.8
Lacey	6	18.3	8.3	6.7	6.3	9.9
Pinnacle	2	25.0	40.0	12.3	23.3	25.2
Rawson	2	35.0	16.7	9.0	16.7	19.4
Stellar-ND	6	25.0	10.7	5.0	11.3	13.0
Tradition	6	28.0	11.7	5.7	7.3	13.2
NDB112 <sup>§</sup>	6	25.0	11.7	5.0	5.0	11.7
ND5883 <sup>¶</sup>	6	60.0	46.4	28.4	45.0	45.0

<sup>†</sup>Data courtesy of Drs. Stephen Neate and Robert Brueggeman, Department of Plant Pathology, NDSU.

<sup>‡</sup>Numbers indicate percentage of leaf surface covered by lesions.

<sup>§</sup>Resistant control

<sup>¶</sup>Susceptible control





**Table 5. Deoxynivalenol (DON) accumulation for 2ND25276, representative cultivars, and Fusarium head blight (FHB) resistant and susceptible controls in field tests in 2009-2013. These trials were irrigated and inoculated and typically gave higher disease and DON than experienced by farmers.**

Entry	Row type	DON (ppm) <sup>†</sup>					Mean
		2009	2010	2011	2012	2013	
Sta. Yrs.		2	2	2	2	2	10
2ND25276	2	47.5	21.4	5.3	7.4	4.6	17.2
Robust	6	59.5	35.6	7.5	11.1	6.2	24.0
Lacey	6	55.0	24.9	10.5	15.5	5.3	22.2
Tradition	6	41.9	30.2	9.6	6.5	5.0	18.6
Stellar-ND	6	35.3	23.8	7.9	8.8	8.8	16.9
Quest <sup>‡</sup>	6	34.0	24.9	4.4	10.4	3.2	15.3
Conlon	2	28.8	20.1	3.1	3.1	3.0	11.6
Pinnacle	2	65.7	25.6	5.3	6.9	3.8	21.4
ND20493 <sup>§</sup>	6	12.5	4.0	5.8	3.5	3.4	5.8

<sup>†</sup>Data courtesy of Dr. Paul Schwarz, NDSU Department of Plant Sciences.

<sup>‡</sup> FHB-resistant cultivar

<sup>§</sup> Resistant control



# NDSU Breeding Scheme

Year	Generation	Location	Process	Number of plants or lines	Number of locations
1	Crossing	Greenhouse (fall)	Crossing block	200 crosses	1
1	F <sub>1</sub>	Greenhouse (winter)	F <sub>1</sub> increase	5,000 plants	1
1	F <sub>2</sub>	North Dakota (summer)	F <sub>2</sub> selection	400,000 plants	1
2	F <sub>3</sub>	New Zealand or Puerto Rico	F <sub>3</sub> advancement	85,000 plants	2
2	F <sub>4</sub>	North Dakota	Progeny rows	25,000 lines	1
3	F <sub>5</sub>	New Zealand or Arizona	Seed increase	5,000 lines	2
3	F <sub>6</sub>	North Dakota	Preliminary yield trials	1,200 lines	2
4	F <sub>7</sub>	North Dakota	Intermediate yield trials	160 lines	7
5	F <sub>8</sub>	North Dakota	Advanced yield trial	40 lines	7
6	F <sub>9</sub>	North Dakota	Variety yield trial (VYT)	10 lines	7
7	F <sub>10</sub>	North Dakota & region	VYT and AMBA Pilot Scale	8 lines	14
8	F <sub>11</sub>	North Dakota & region	VYT and AMBA Pilot Scale	4-8 lines	14
9	F <sub>12</sub>	North Dakota & region	VYT and AMBA Plant Scale	2 lines	20
10	F <sub>13</sub>	North Dakota & region	VYT and AMBA Plant Scale	2 lines	20

# How Long Does it Take to Make These Tables?

- Balanced data vs. unbalanced data.
- How do you make similar tables for all lines in the program that have been tested two or more years?
- $200 \text{ lines} * 30 \text{ minutes/table} = 12.5 \text{ working days}$
- $200 \text{ lines} * 4 \text{ hours/table} = 20 \text{ weeks}$



## How NDSU Is Addressing These Data Management Needs?

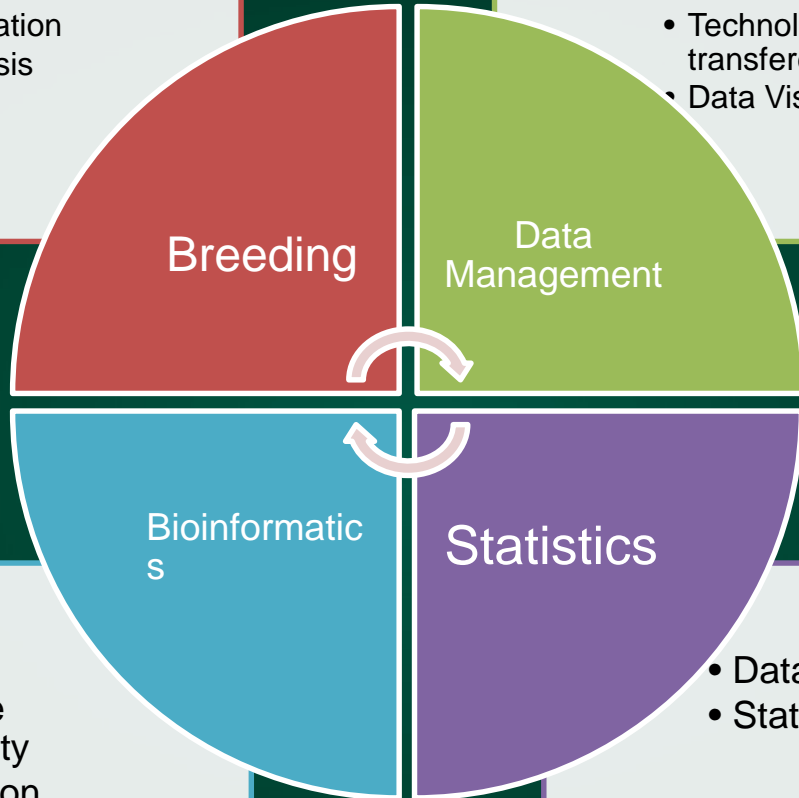


- Department identified bioinformaticists as a priority for the 2013 SBARE & Legislative process.
- Hiring bioinformaticists identified as the top priority for the AES in the 2015 SBARE & Legislative process.
- Two Breeding Pipeline Data Managers hired in summer 2016.

# Pipeline Database Management Team



- Data Compliance
- Standardization
- Data Analysis
- Efficiency



- QAQC
- Technology transference
- Data Visualizations



- Scripting language
- IT Security
- Automation

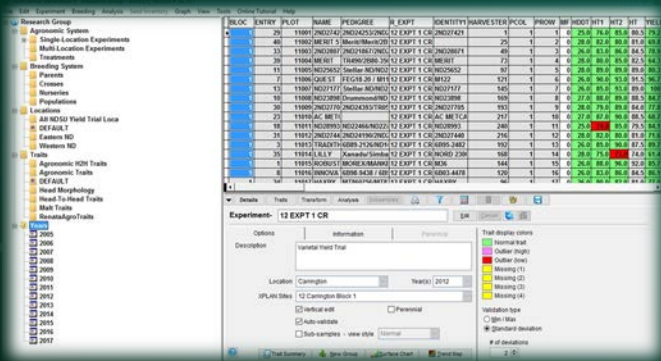
- Data migration
- Statistics





# Requirements for Efficient Data Management

- Single platform for managing
  - phenotype data
  - genotype data
- Single server for housing platforms and data.
- Automating routine processes (queries and calculation of means or BLUPs)



The screenshot shows a software interface with a tree view on the left and a data table on the right. The tree view includes categories like 'Research Group', 'Agronomic System', 'Multi-Location Experiments', 'Breeding System', 'Populations', 'Traits', and 'Years'. The data table has columns for 'BLOC', 'ENTRY', 'PLOT', 'NAME', 'PGENGID', 'R\_EXPT', 'IDENTITY', 'HARVEST', 'POOL', 'SPROV', 'BP', 'HDOT', 'HTY', 'HT2', 'HT', and 'WEELE'. The table contains multiple rows of data with numerical values.



# Why use a relational database?

Flat File Database	Relational Database
<b>Advantages</b>	
<ul style="list-style-type: none"><li>• Easy to use, quick start</li></ul>	<ul style="list-style-type: none"><li>• High level of security (authentication process)</li></ul>
<ul style="list-style-type: none"><li>• No extra installation required</li></ul>	<ul style="list-style-type: none"><li>• Powerful data organization and storage</li></ul>
<ul style="list-style-type: none"><li>• Analytic &amp; graphing tools built-in</li></ul>	<ul style="list-style-type: none"><li>• Data centralized, standardized &amp; easily shared</li></ul>
	<ul style="list-style-type: none"><li>• Changes universally applied to all tables</li></ul>
	<ul style="list-style-type: none"><li>• Data easily queried, compiled, and analyzed</li></ul>
	<ul style="list-style-type: none"><li>• Intellectual property protection &amp; compliance</li></ul>
<b>Disadvantages</b>	
<ul style="list-style-type: none"><li>• Multiple files difficult to compile &amp; compare</li></ul>	<ul style="list-style-type: none"><li>• Complex set-up</li></ul>
<ul style="list-style-type: none"><li>• Only understood by primary user in most cases</li></ul>	<ul style="list-style-type: none"><li>• Changes may be difficult</li></ul>
<ul style="list-style-type: none"><li>• Prone to corruption</li></ul>	<ul style="list-style-type: none"><li>• Database needs to be well planned</li></ul>
<ul style="list-style-type: none"><li>• More difficult to secure</li></ul>	<ul style="list-style-type: none"><li>• Requires experienced management</li></ul>
<ul style="list-style-type: none"><li>• Data must be applied individually to each table</li></ul>	

# Managing Phenotype Data

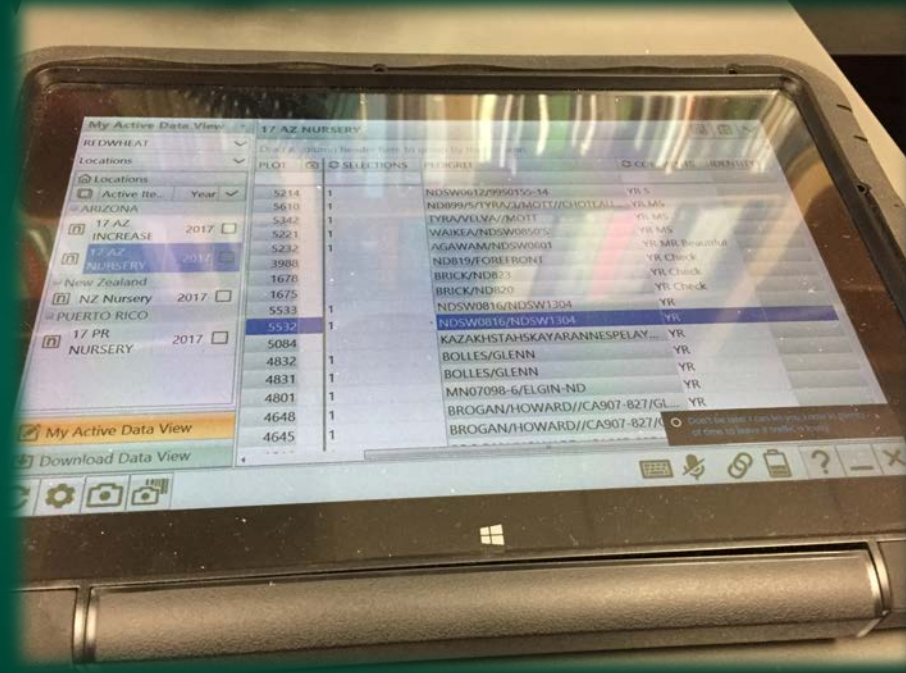
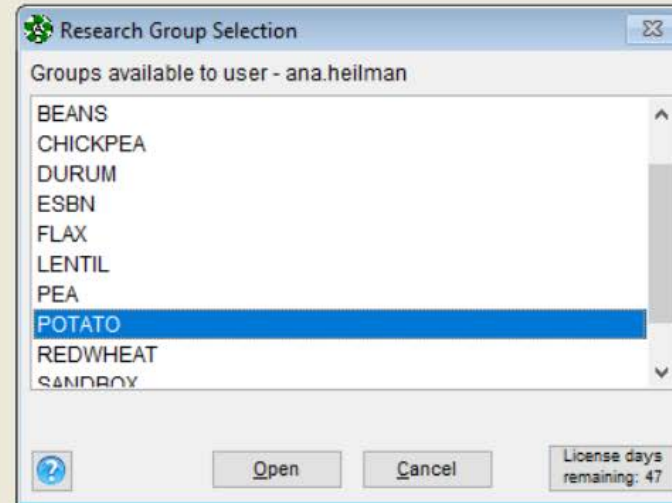
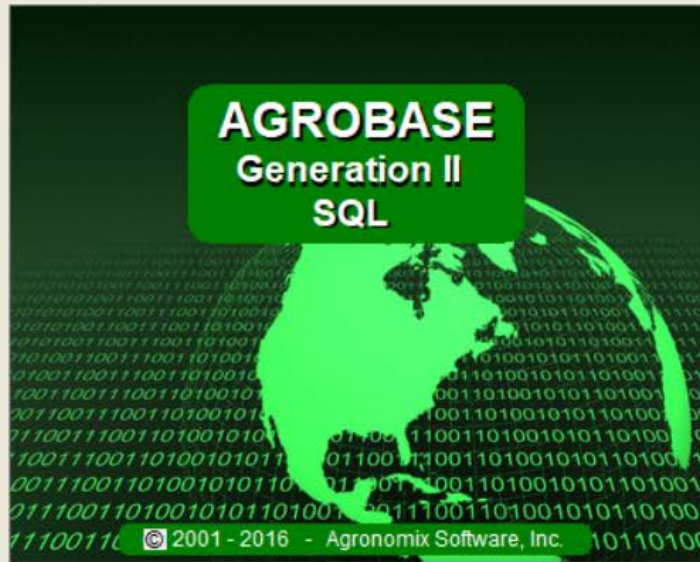


Photo courtesy: Dr. Andrew J Green



- Breeding system management
- Randomizations
- Data collection on tablets
- Analyses of basic experimental designs
- Stores data and means in a relational database

**AGROBASE**  
Generation II  
SQL



# Data Management Work Plan at NDSU

- Educate and support breeders on use of technology platforms (e.g. Agrobase, T.3, JMP, and SAS)
- Integrate legacy data from breeding programs into DB.
- Automate data queries and calculations of:
  - Unbalanced means or BLUPS across years based on
    - Entry
    - Pedigree
    - Parents
  - Balanced means across years
  - Breeder requests
- Expand visualization capabilities for making decisions

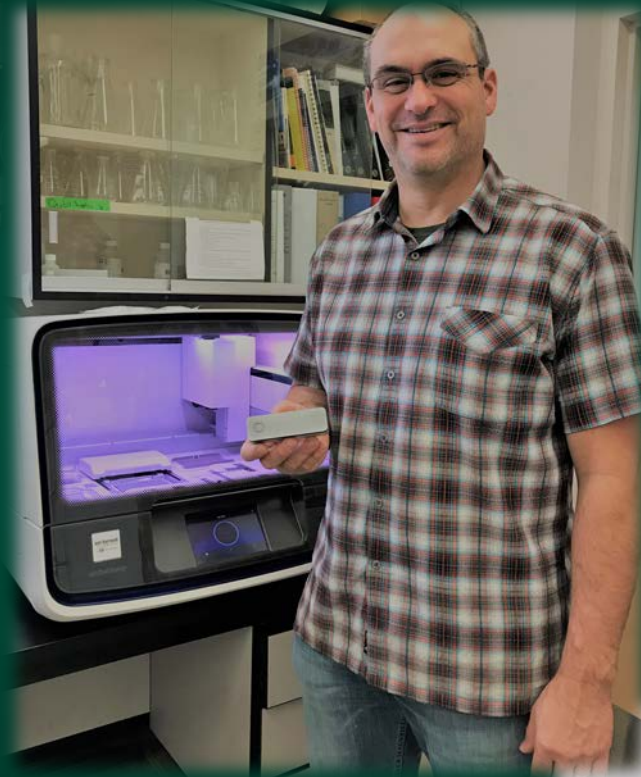
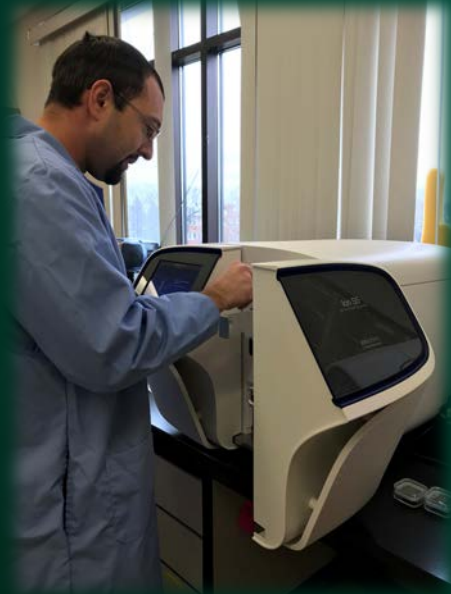
# Managing Genotype Data



File-DNA-Sequencers\_from\_Flickr\_57080968.jpg



<https://thumbs.dreamstime.com/z/genetic-fingerprinting-as-fingerprint-dna-emerging-out-as-medical-identification-symbol-paternity-test-64869820.jpg>



<https://www.extremetech.com/wp-content/uploads/2013/03/minion-usb-dna-sequencer.jpg>

# How do we Manage Genotype Data?

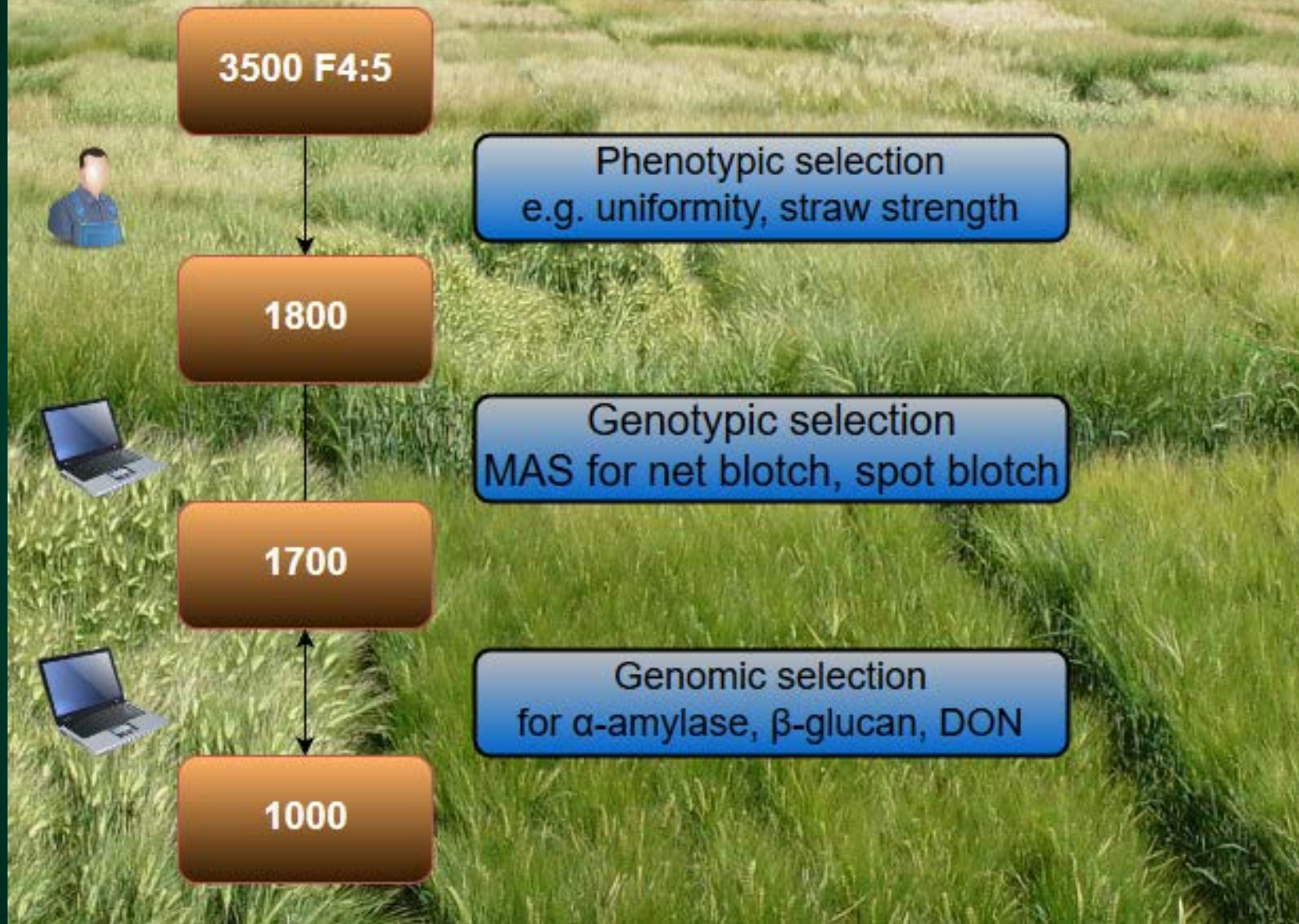
- 525 parents screened with 50,000 SNPs
  - 28.25 million data points
- 3,500 lines screened with 450 SNPs
  - 1.575 million data points
- How do you pair genotype data with phenotype data for GWAS or genomic selection?



# NDSU Breeding Scheme

Year	Generation	Location	Process	Number of plants or lines	Number of locations
1	Crossing	Greenhouse (fall)	Crossing block	200 crosses	1
1	F <sub>1</sub>	Greenhouse (winter)	F <sub>1</sub> increase	5,000 plants	1
1	F <sub>2</sub>	North Dakota (summer)	F <sub>2</sub> selection	400,000 plants	1
2	F <sub>3</sub>	New Zealand or Puerto Rico	F <sub>3</sub> advancement	85,000 plants	2
2	F <sub>4</sub>	North Dakota	Progeny rows	25,000 lines	1
3	F <sub>5</sub>	New Zealand or Arizona	Seed increase	5,000 lines	2
3	F <sub>6</sub>	North Dakota	Preliminary yield trials	1,200 lines	2
4	F <sub>7</sub>	North Dakota	Intermediate yield trials	160 lines	7
5	F <sub>8</sub>	North Dakota	Advanced yield trial	40 lines	7
6	F <sub>9</sub>	North Dakota	Variety yield trial (VYT)	10 lines	7
7	F <sub>10</sub>	North Dakota & region	VYT and AMBA Pilot Scale	8 lines	14
8	F <sub>11</sub>	North Dakota & region	VYT and AMBA Pilot Scale	4-8 lines	14
9	F <sub>12</sub>	North Dakota & region	VYT and AMBA Plant Scale	2 lines	20
10	F <sub>13</sub>	North Dakota & region	VYT and AMBA Plant Scale	2 lines	20

# Application of Genomic Selection at NDSU



# Favorable Phenotypes in the Arizona Winter Nursery



# Weak Strawed 2-rowed Phenotypes in the Arizona Winter Nursery



# The Triticeae Toolbox

Home

<https://triticeaetoolbox.org/>

## Home

- T3 Barley ▾
- Choose database
- T3 Wheat
- T3 Barley**
- T3/Oat
- Breeders Datafarm, Barley
- 
- T3 Wheat sandbox
- T3 Barley sandbox
- T3/Oat sandbox

### Triticeae Toolbox (T3)

generated by the Triticeae Coordinated Agricultural Project (CAP), funded by the Department of Agriculture (USDA). T3 contains SNP, phenotypic, and pedigree data to integrate rapidly expanding DNA marker and sequence data with traditional phenotype and enabling breeders to select on marker data alone. T3 will also link (e.g., to the Triticeae Genebank, PLEXdb, and GRIN) for functional analyses to identify causal polymorphisms.

The software and data structure for T3 were developed as part of the Barley CAP project for its database, The H

# T3/Barley

<https://triticeaetoolbox.org/barley/>

Home Select ▾ Analyze ▾ Download ▾ Reports ▾ Manage ▾ Resources ▾

## Quick Links

[Login/Register](#)

**Current selections:**

[Lines: 0](#)

[Markers: All](#)

[Traits: 0](#)

[Phenotype Trials](#)

[Genotype Experiments](#)

Quick search...

## What's New

**Data Selection**

[Physical Map](#) added for IBSC 2016 assembly.

[Reports](#)

## Home: T3/Barley

### Welcome to The Triticeae Toolbox (T3)

T3 is the web portal for the data generated by the [Triticeae Coordinated Agricultural Food and Agriculture \(NIFA\)](#) of the United States Department of Agriculture ([USDA](#)) wheat and barley germplasm in the Triticeae CAP. [More...](#)

**T3 Barley** is an extension of The Hordeum Toolbox (THT), which was created by the THT data.

**Participants:** The templates and instructions for data submission are [here](#). If you are interested in the [Usage Policy](#).

[Data submission](#)

data analysis - Bing image x T3/Barley x Large Data Managers Joi x

Secure <https://triticeaetoolbox.org/barley/>

Apps ServiceNow Home - JMP User Co J.P. Morgan Commer Join the Meeting | Go Agronomix Software MobileDemand Microsoft Office Hom T

# T3/Barley

<https://triticeaetoolbox.org/barley/>

Home Select Analyze Download Reports Manage Resources

Cluster

- Cluster Lines by Genotype
- Cluster Lines 3D (pam)
- Cluster Lines 3D (hclust)

Filter outliers Cluster

Optimize training set

Selection Index

Traits and Trials statistics

Canopy Spectral Reflectance

**Genomic Association and Prediction**

Compare Trials

Track Alleles through Pedigree

Parse Purdy Pedigrees

Allele Data Conflicts

BLAST

Haplotype Data

## Quick Links

[Login/Register](#)

**Current selections**

Lines: 0

Markers: All

Traits: 0

[Phenotype Trials](#)

[Genotype Experiments](#)

Quick search...

## What's New

**Data Selection**

[Physical Map](#) added for IBSC 2016 assembly.

## Triticeae Toolbox (T3)

...for the data generated by the [Triticeae Coordinated Agricultural Project](#) (NIFA) of the United States Department of Agriculture (USDA). ...mplasm in the Triticeae CAP. [More...](#)

...sion of The Hordeum Toolbox (THT), which was created by the

**Participants:** The templates and instructions for data submission are [here](#). If your data

[Usage Policy](#).

[Data submission](#)

# The Triticeae Toolbox

Home Select ▾ Analyze ▾ Download ▾ Manage ▾ About T3 ▾

## Quick Links

[Login/Register](#)

**Current selections:**

Lines: 0

Markers: All

Traits: 0

[Phenotype Trials](#)

[Genotype Experiments](#)

## What's New

### Data Analysis Tools

Outlier detection on traits and trials using Bonferroni–Holm with Median Absolute Deviation (MAD).

Selection of optimized [training sets](#) for use in genomic prediction, using an R package called STPGA .

### Download

Historical [weather data](#) using

## Select Phenotypes

Select multiple options by holding down the Ctrl(PC) Command(Mac) key while clicking.  
Selecting traits and trials will NOT affect currently selected lines

Category	Traits	Trials
Agronomic	FFB Severity	15GHNB_2R
Disease	leaf rust (0-9)	15GHNB_6R
Malting quality	leaf rust (LR30) (0-4)	14GHNB_2R
Morphological	leaf rust (LR8) (0-4)	14GHNB_6R
Quality for food and other uses	leaf rust seedling	13GHNB_2R
Winter growth habit	net blotch (1-10)	13GHNB_6R
	net blotch net form reaction type (0-9)	12GHNB_2R
	powdery mildew (0-4)	12GHNB_6R
	powdery mildew (0-9)	ABAR_2009_Blacksburg
	scald (0-8)	HSADV_2009_Blacksburg

Save Phenotype Selection

Traits	Trials	Lines
net blotch (1-10)	15GHNB_2R	265
net blotch (1-10)	14GHNB_2R	356
net blotch (1-10)	13GHNB_2R	272
net blotch (1-10)	12GHNB_2R	247



# Having Data in Databases Allows for Optimization of:

- Markers to include for genomic selection.
- Lines to use in training population.
- Statistical models to use for genomic selection of specific traits.
- Predictions of progeny performance from planned crosses.

# Next Steps for NDSU Data Management Team

- Determine needs for hosting T3 locally and updating system to handle more crops.
- Develop Addins in JMP for automating queries and analyses.
- Make applications available on tablet and mobile devices for NDSU breeding programs.

# Any Questions?

