

Genomics Assisted Breeding and Field-Based High Throughput Phenotyping

Jesse Poland

Wheat Genetics Resource Center
Applied Wheat Genomics Innovation Lab
Kansas State University



**DOE ARPA-E – Advanced Plant Phenotyping Workshop
Chicago, IL
June 18, 2014**

Building a better car... plant breeding style



1. More efficiently select which 'model' will perform best
2. Understand the parts

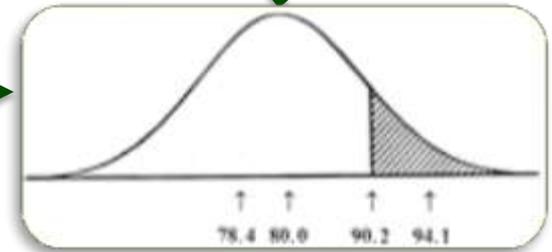
(Accelerating) The Breeding Cycle



Crossing

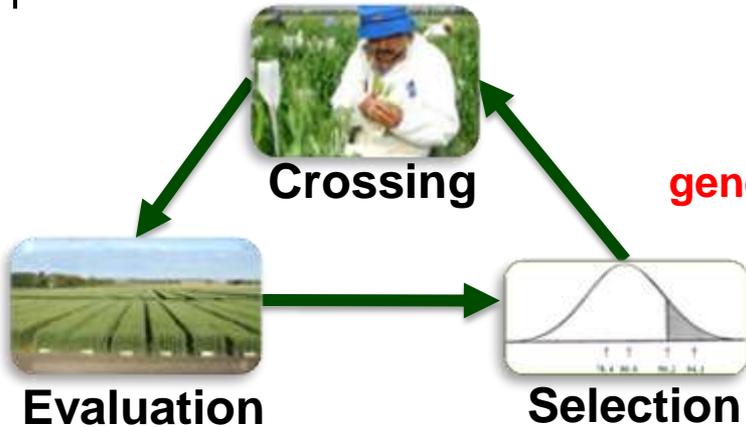


Evaluation



Selection

The breeder's (favorite) equation:



genetic gain over time

$$R_t = \frac{i r S_A}{y}$$

selection intensity i selection accuracy r genetic variance S_A years per cycle y

Selection Intensity

- ✓ Increase (to a limit)
- ✓ Need bigger populations

Selection Accuracy

- ✓ Increase
- ✓ More precise measurements
- ✓ Reduce Errors
- ✓ Correct for environment

Genetic Variance (Diversity)

- ✓ Increase
- ✓ Mixed bag (not all good)
- ✓ A must have

Years per Cycle

- ✓ Decrease!
- ✓ Constant 'rate' of return

Genomic Selection

Prediction of *total genetic value* using dense *genome-wide markers*

- ✓ Estimate Kinship (realized relationship) between breeding with markers



The Breeding Funnel

Years



Crossing



Early Generation Testing



Prelim Yield testing
(thousands)



Replicated Yield testing
(hundreds)



Advanced Yield testing
(tens)



Elite

Un-adapted

Exotic



Varieties (one)

The Breeding Funnel

Years



Crossing



GENOMIC SELECTION



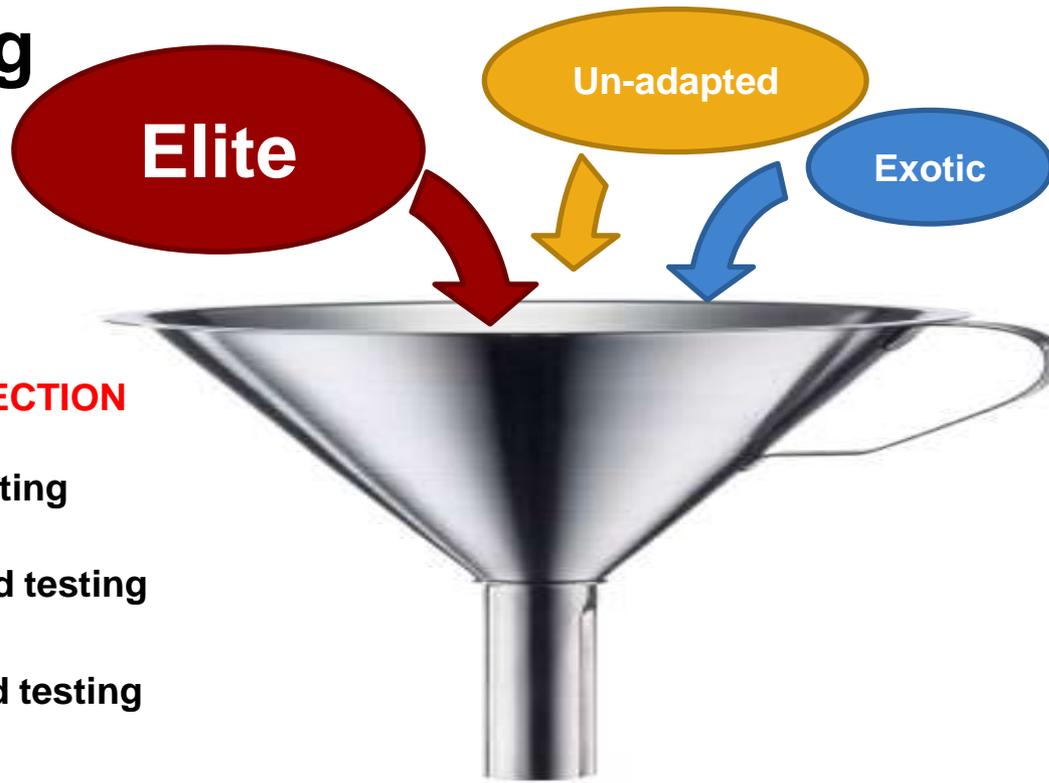
Prelim Yield testing
(thousands)



Replicated Yield testing
(hundreds)



Advanced Yield testing
(tens)



Varieties (LOTS!)

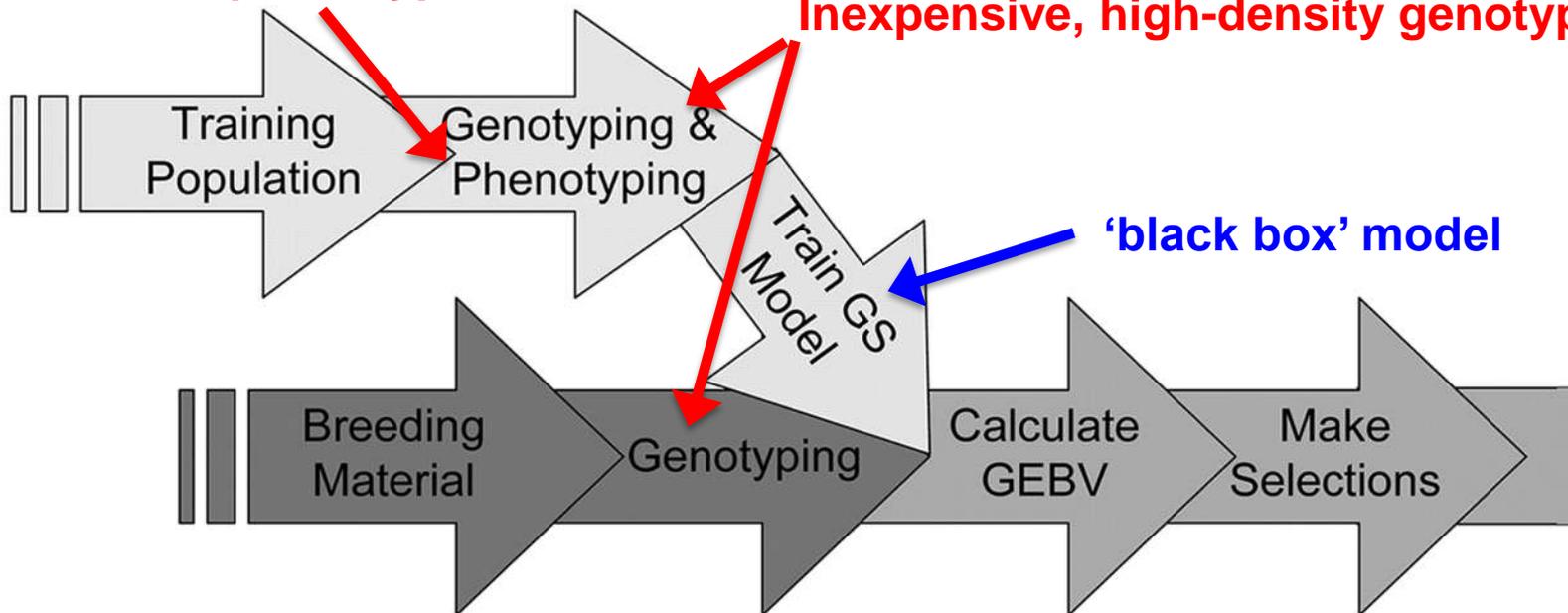
Genomic Selection

Needed:

- 1) Training Population (genotypes + phenotypes)
- 2) Selection Candidates (genotypes)

Accurate phenotypes

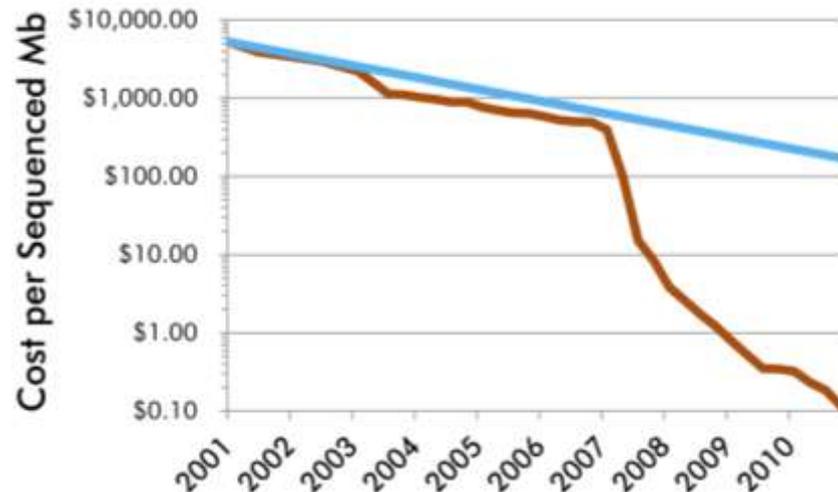
Inexpensive, high-density genotypes



Genotyping-by-sequencing (GBS)

Whole genome profiling by reduced representation sequencing

- + Amazing developments in sequencing output
- + Very good for wheat where polyploidy and duplications cause problems with hybridization/PCR assays
- + Polymorphism discovery simultaneous with genotyping
- + No ascertainment bias
- + Low per sample cost
- Complex bioinformatics
- *Requires paradigm shift in molecular markers*



Whole-genome profile for <\$10 per sample

Genotyping-by-sequencing (GBS)

“...massively parallel sequencing of multiplexed reduced-representation genomic libraries.”

“massively parallel sequencing”

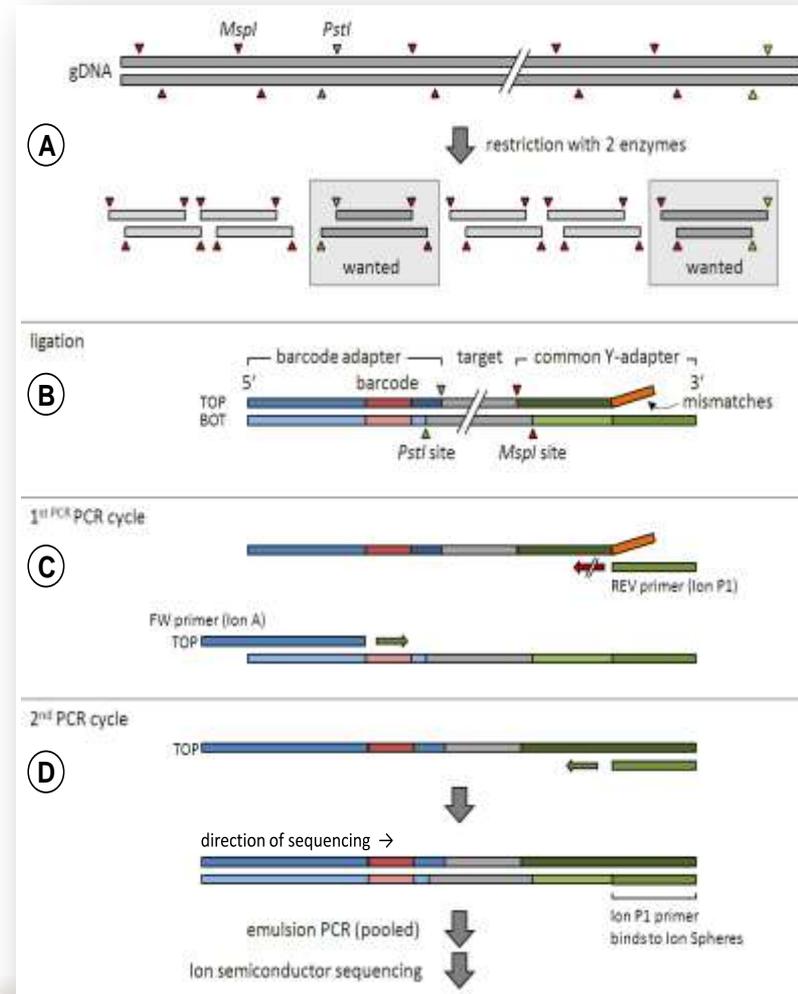
- next-gen sequencing (Illumina)

“multiplex” = using DNA barcode

- unique DNA sequence synthesized on the adapter
- pool 48-384 samples together

“reduced-representation”

- capture only the portion of the genome flanking restriction sites
- methylation-sensitive restriction enzymes
- Target rare, low-copy sites in genome
 - *PstI* (CTGCAG), *MspI* (CCGG)



Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler and S. E. Mitchell (2011). "A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species." *PLoS one* 6(5): e19379.



GS: Prediction of wheat quality

CIMMYT elite breeding lines (n=1,138)

Cycle 45 & 46 International Bread Wheat Screening Nursery (C45IBWSN)

- Genotyping-by-sequencing: 15,330 SNPs (imputed with MVN-EM) (rrBLUP)

Grain quality traits:

- thousand kernel weight
- milling yield
- mix time
- pup loaf volume
(= \$\$\$)

Replicated yield tests

- ✓ 2009 & 2010
- ✓ 6 environments

One replication for quality testing

- ✓ milling
- ✓ dough rheology
- ✓ baking tests

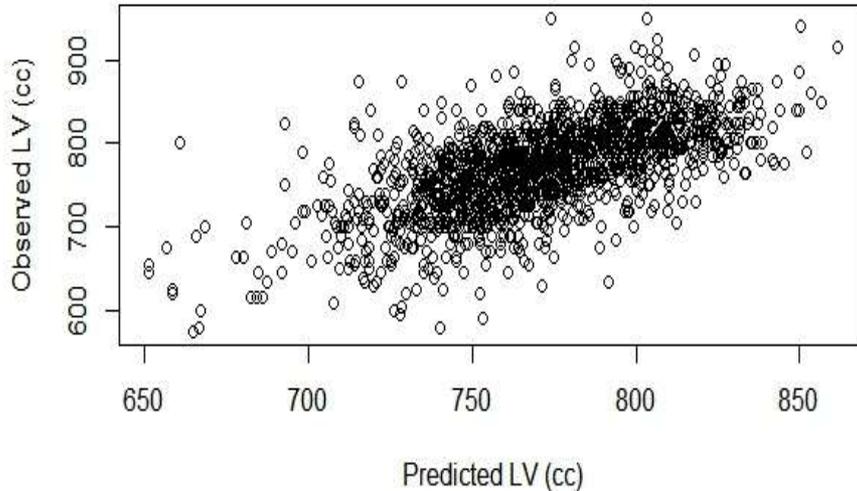
Best Linear Unbiased Estimate (BLUE)

Cross-validation (x100)

- ✓ Training sets of n=134
- ✓ Validation sets of n=30

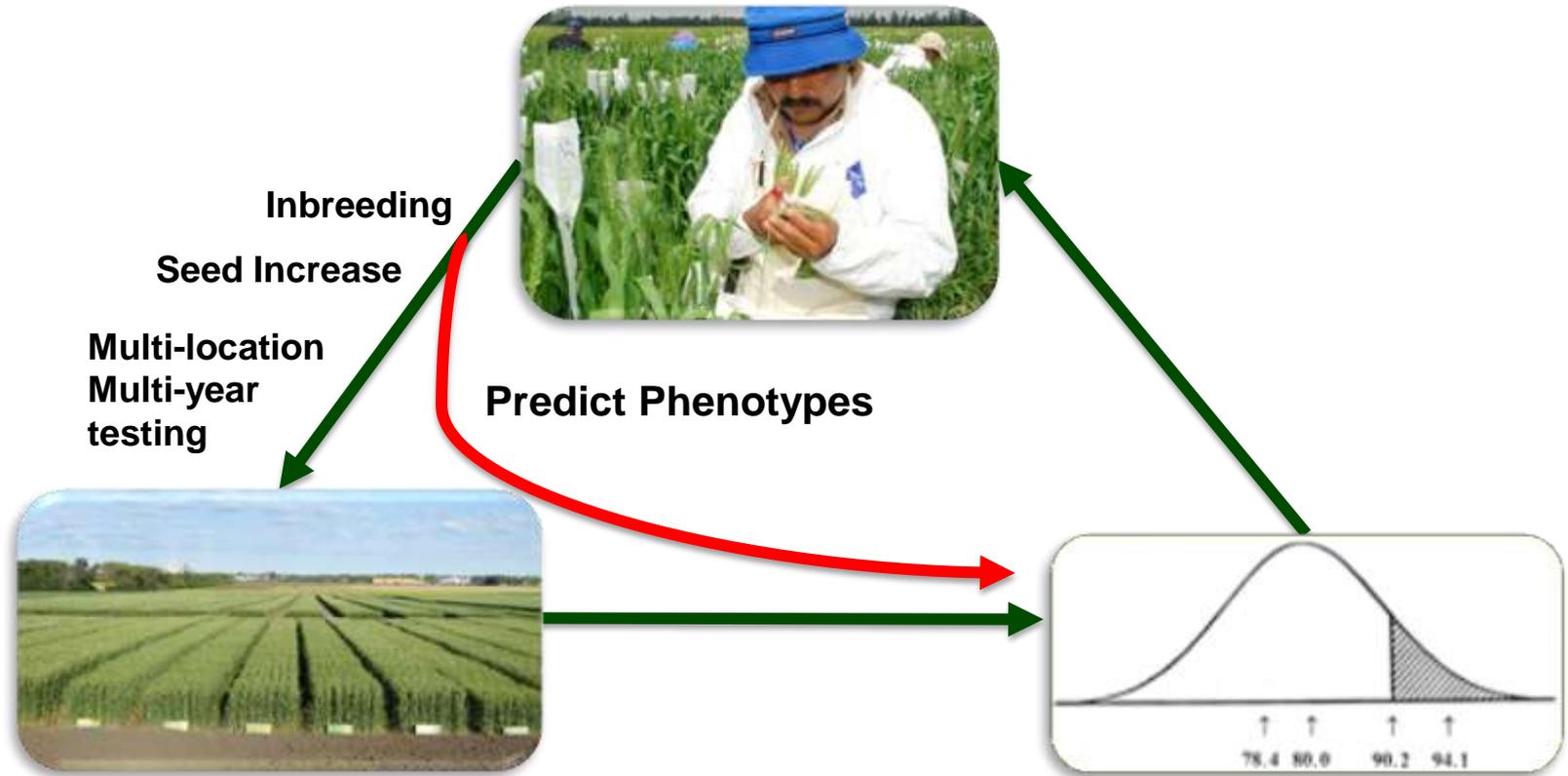


GS: Prediction of wheat quality



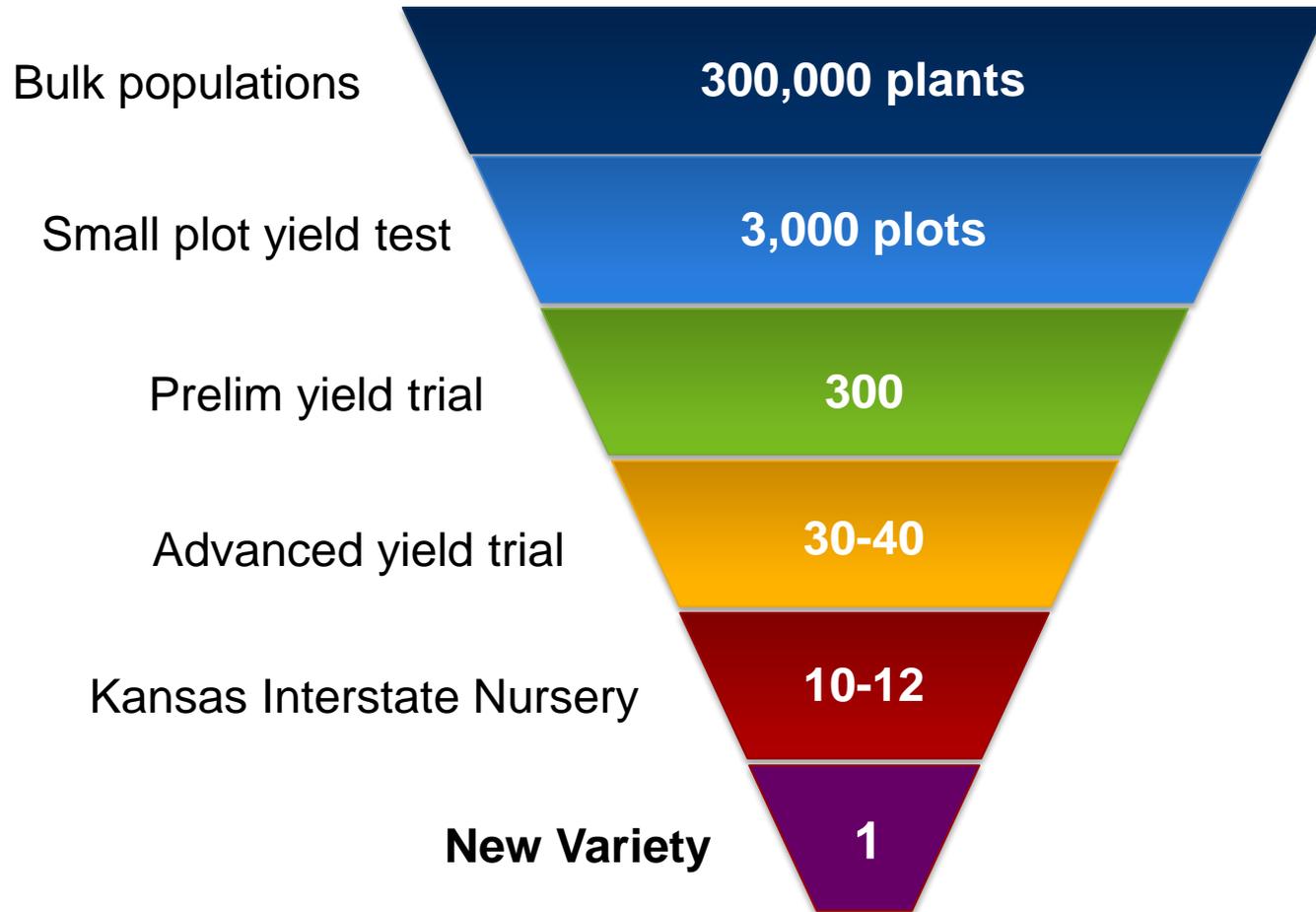
TRAIT	PREDICTION ACCURACY (r)
Test Weight	0.725***
Grain Hardness	0.513***
Grain Protein	0.630***
Flour Protein	0.604***
Flour SDS	0.666***
Mixograph Mix Time	0.718***
Alveograph W	0.697***
Alveograph P/L	0.476***
Loaf Volume	0.638***

The Breeding Cycle



\$ genotyping < \$ phenotyping

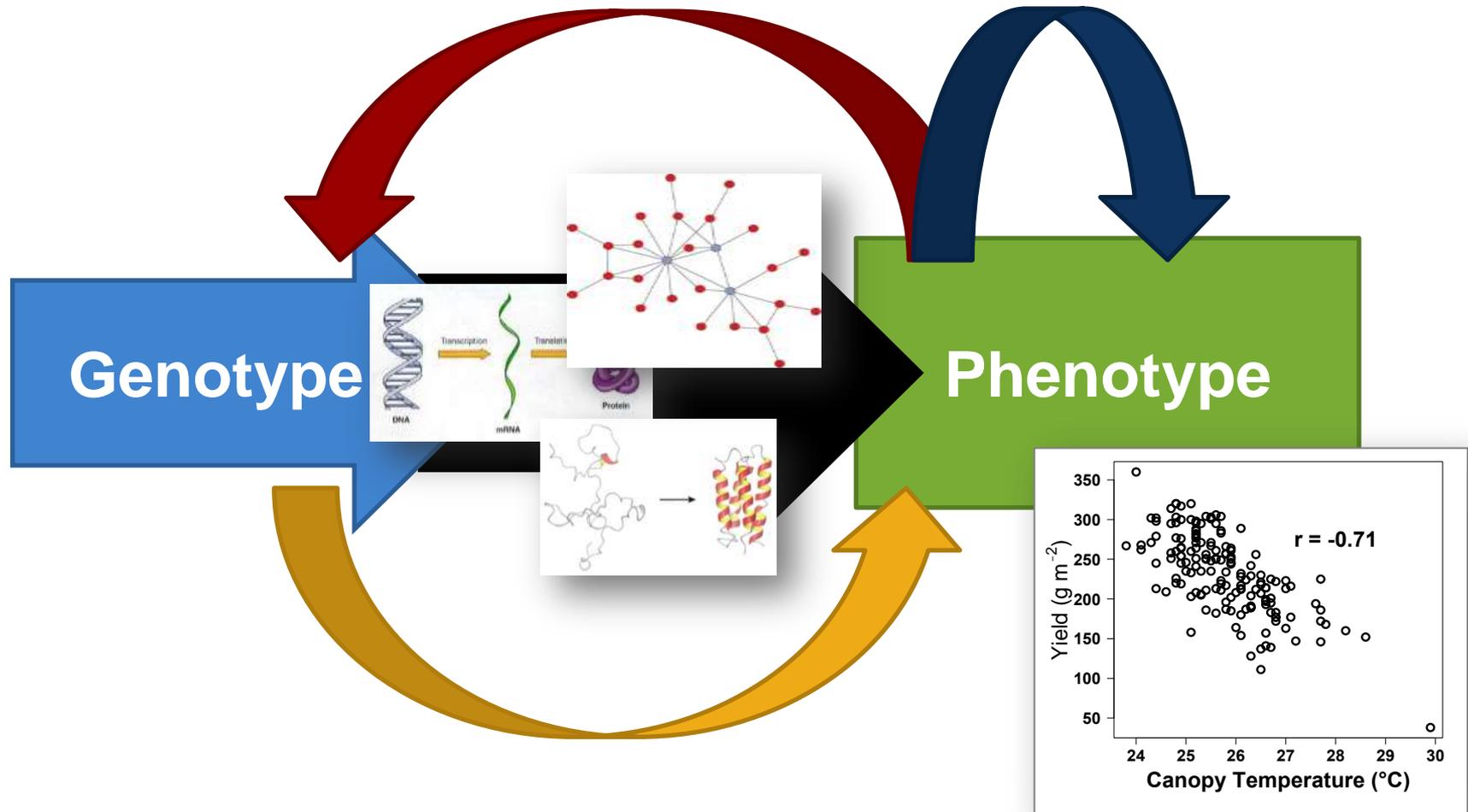
Plant Breeding – It's a numbers game



Increasing selection intensity = more to chose from



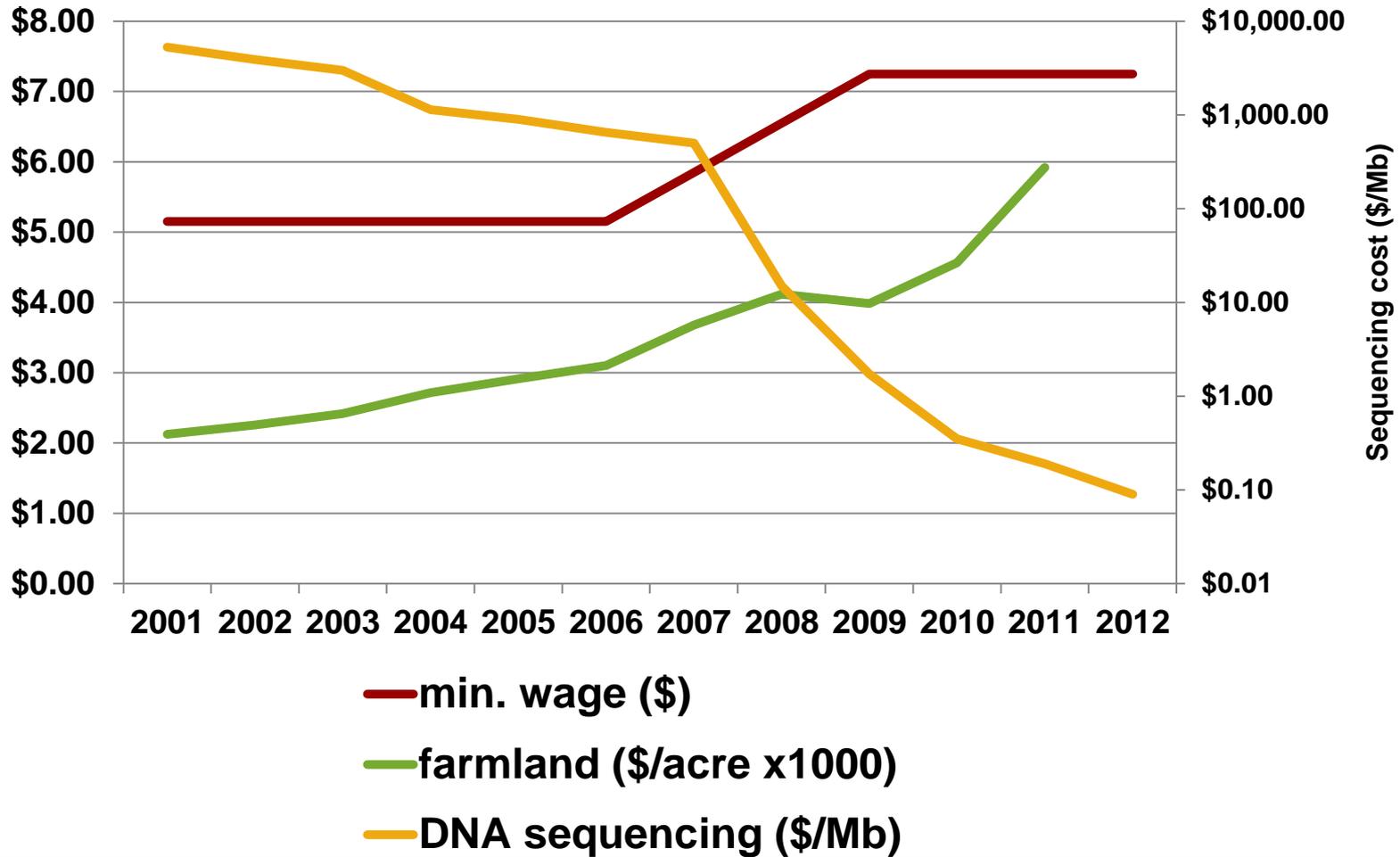
G2P: connecting genotype to phenotype



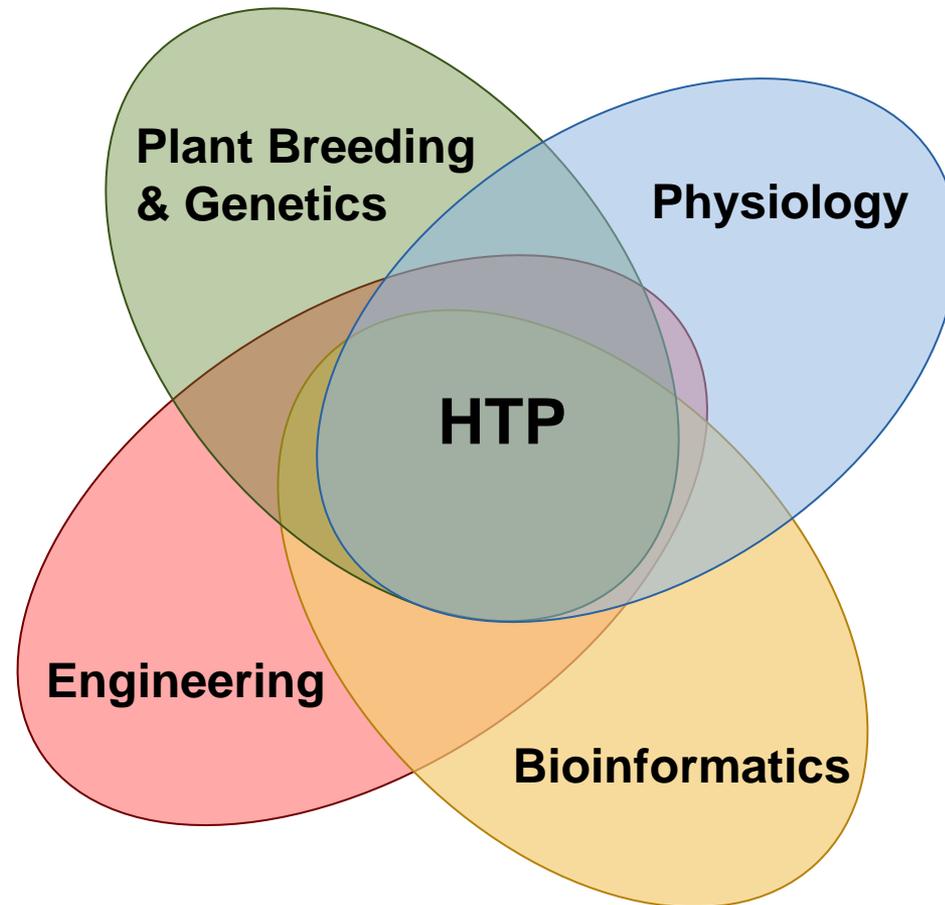
The need for phenotypes:

1. more efficient selection (breeding)
2. understanding the parts (genetics)

Trending: Phenotyping vs Genotyping



A multi-disciplinary approach



Field-based high throughput phenotyping

Defining “field-based high throughput”

- ✓ Fully- (or mostly) automated data collection
- ✓ <1 second per plot (3h for 10,000 plots)
- ✓ Data analysis must be “pipelined”
- ✓ High-resolution \neq high-throughput
- ✓ Field conditions targeting production systems
- ✓ Automated data processing

Phenotyping vehicle



- + Carry lots of equipment
- + flexible deployment
- + easy to operate



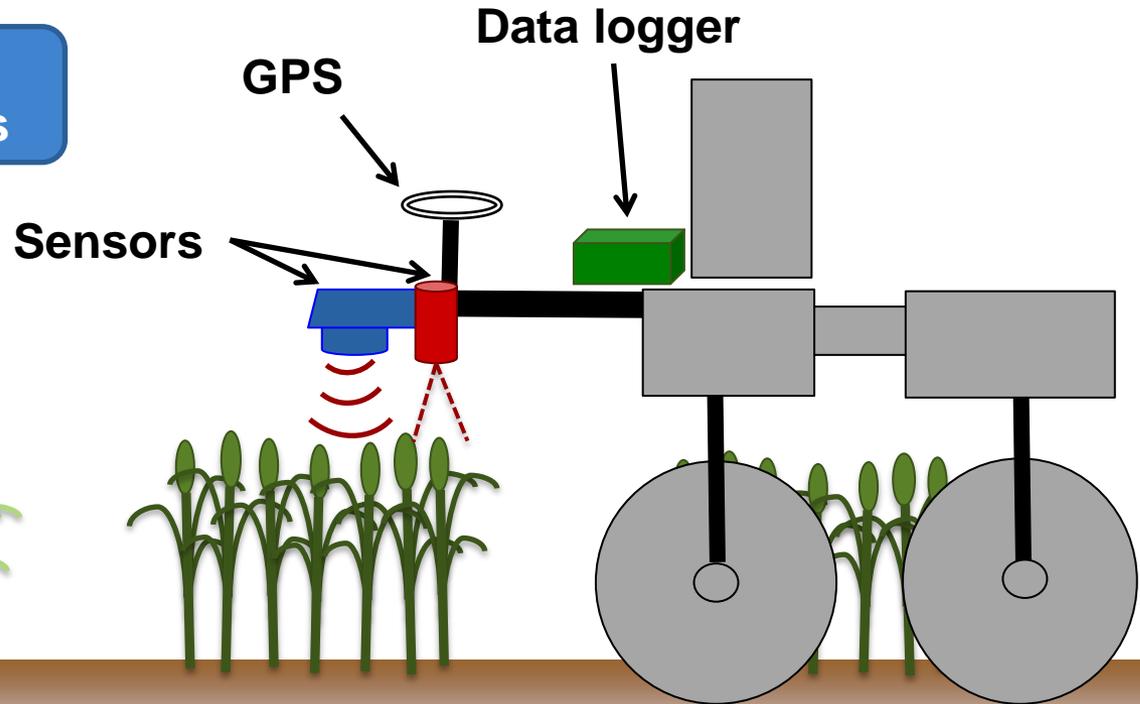
- Can not assay whole field simultaneously
- Not completely automated

Precision Ag meets Plant Genetics



“Geo-referenced proximal sensing”

Physiologically define proximal measurements

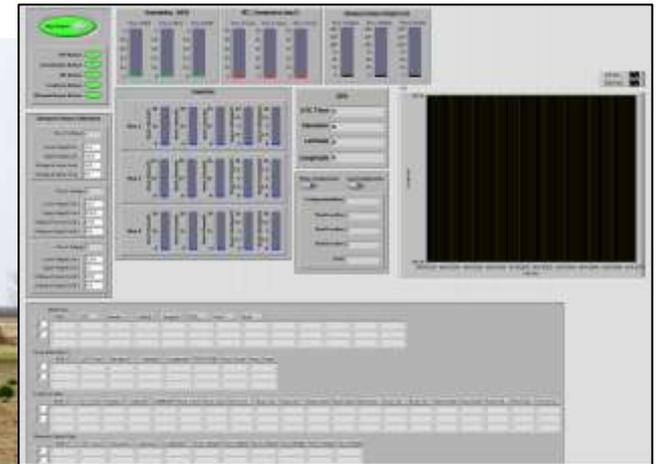


RTK-GPS
(cm level accuracy)

Sensors

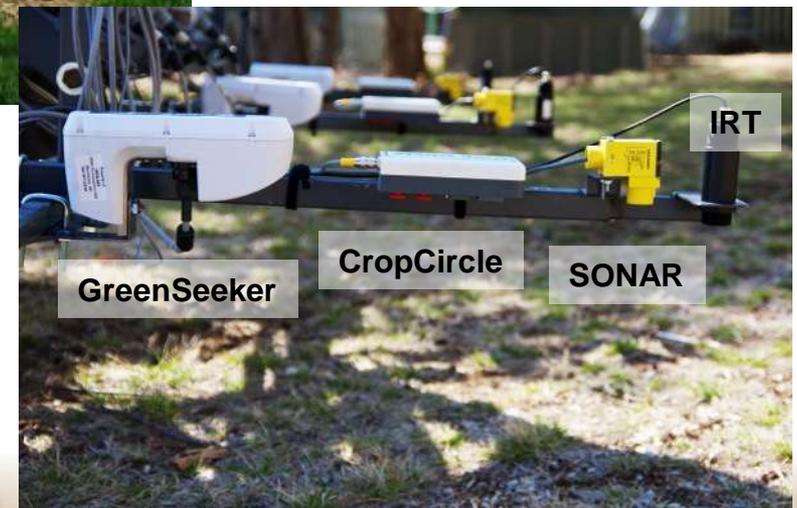
- GreenSeeker = NDVI
- IRT = canopy temperature
- SONAR = plant height

HTP: Platform configuration

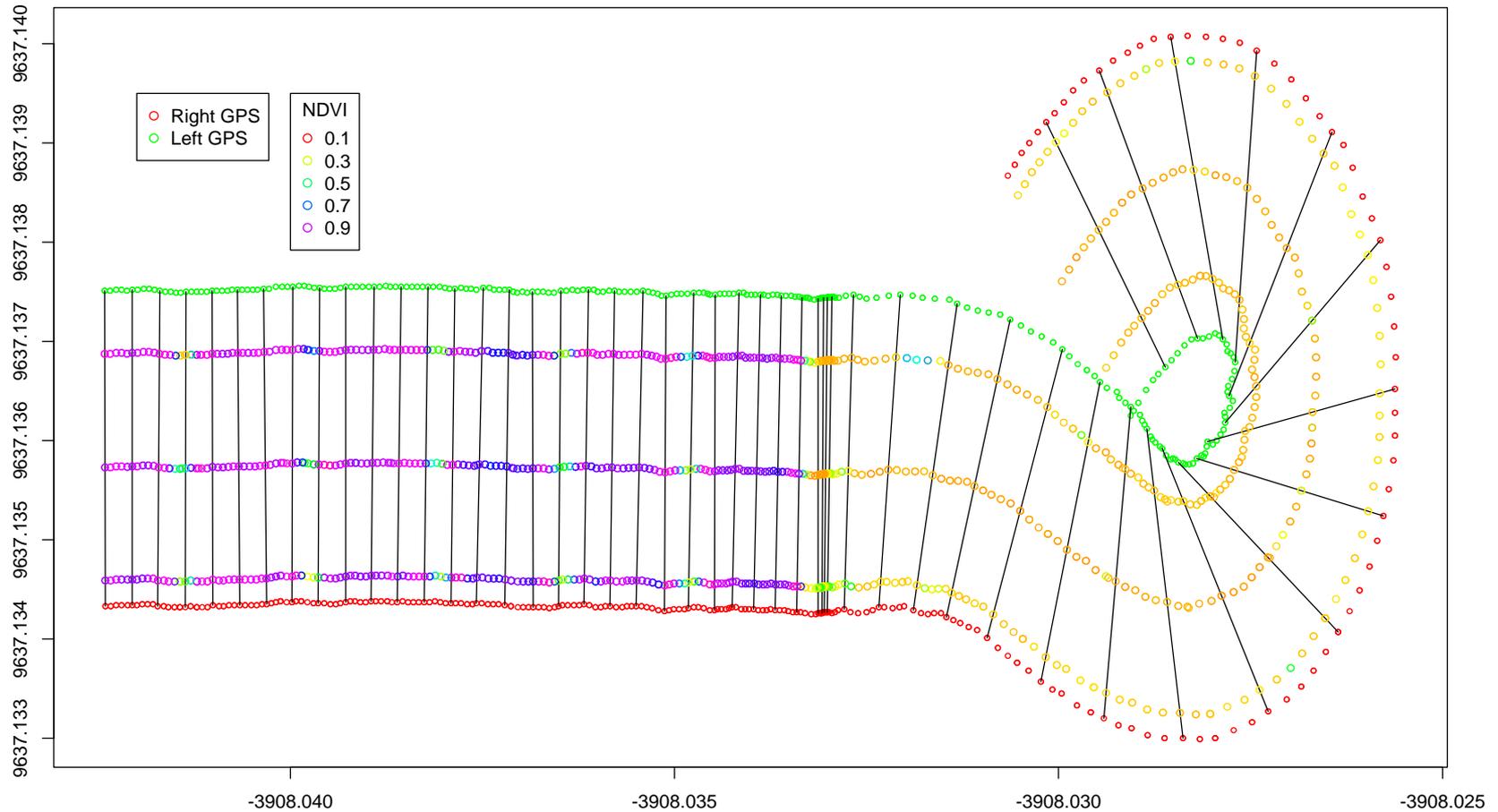


LabView program

- ✓ 10 Hz sampling
- ✓ Real-time feedback
- ✓ Flat file output

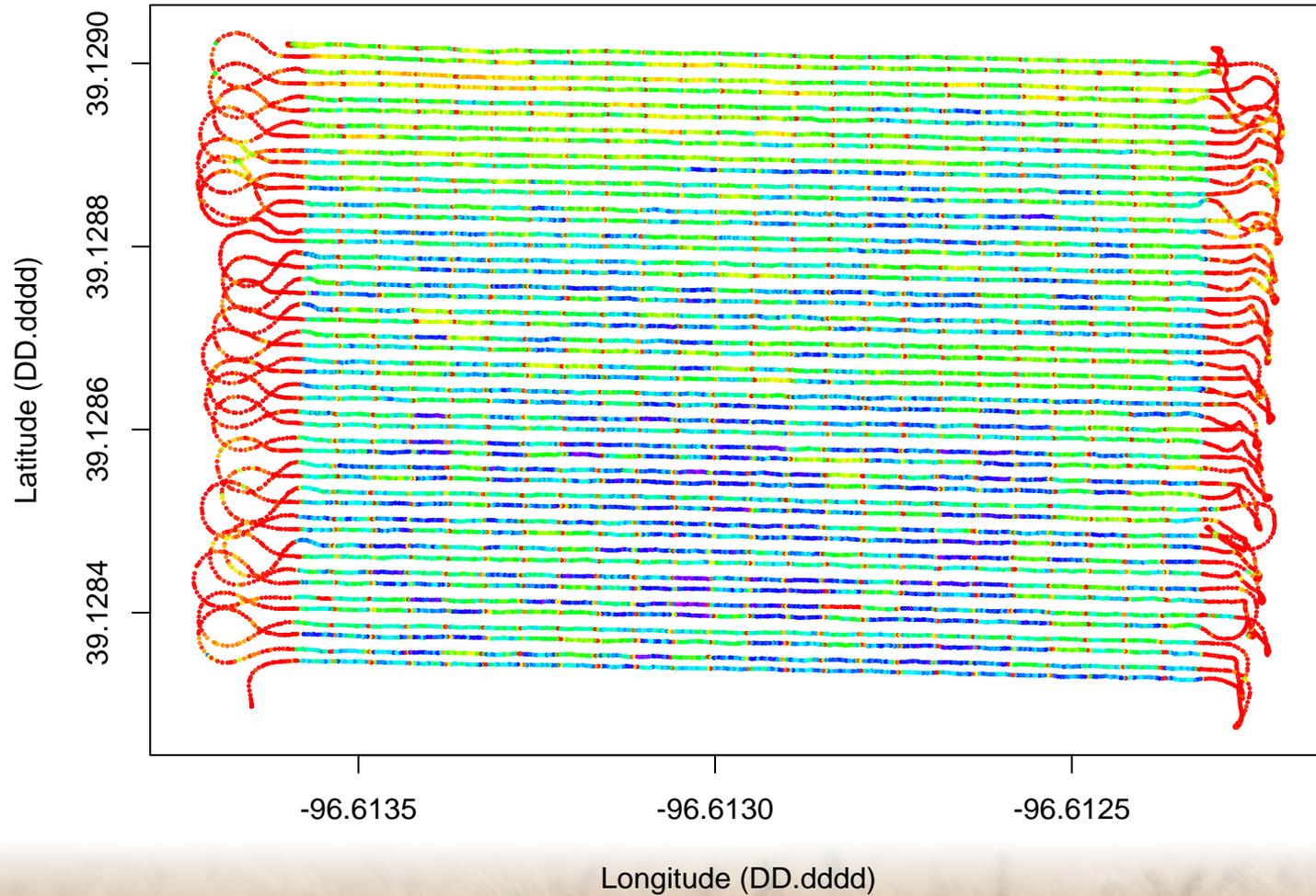


HTP: Multiple sensor orientation



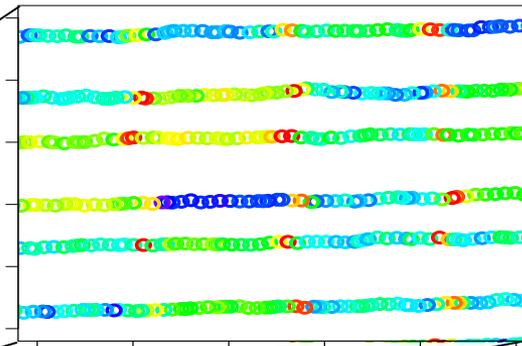
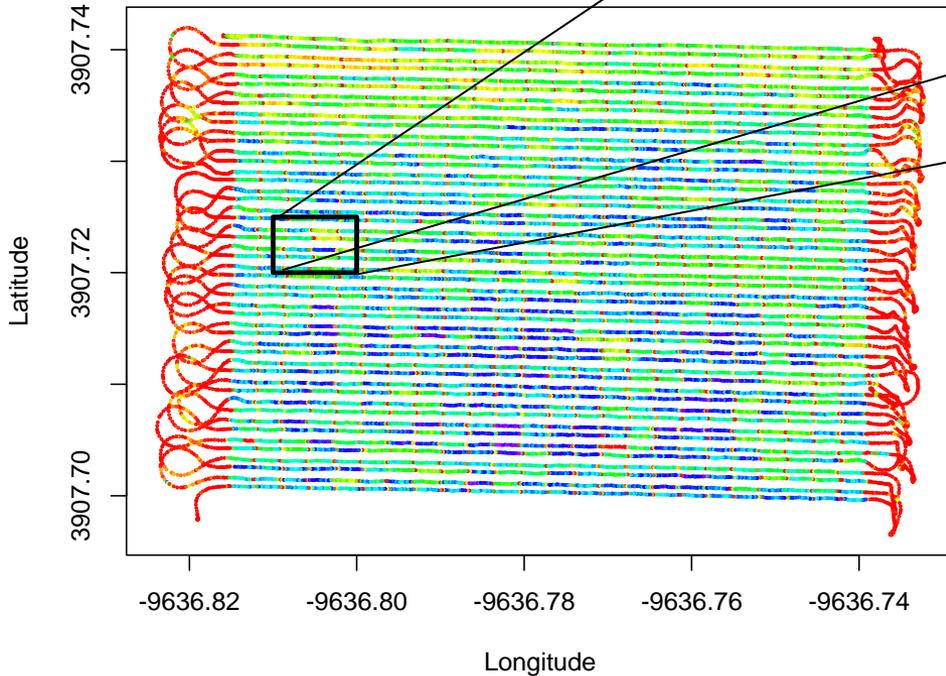
NDVI – raw data

NDVI - 2012.05.10

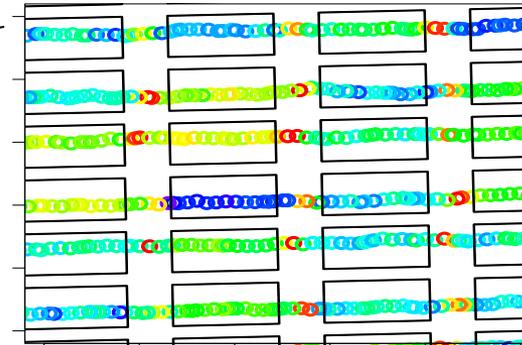


Assigning data to field entries

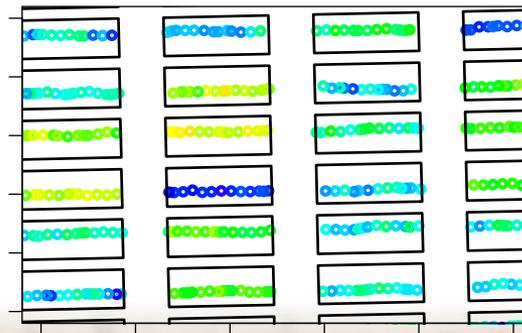
NDVI - 2012.05.10



Raw data

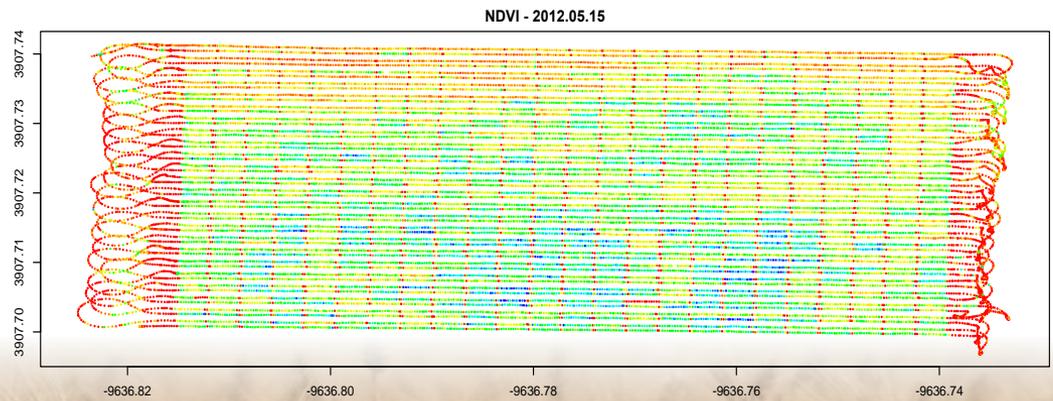
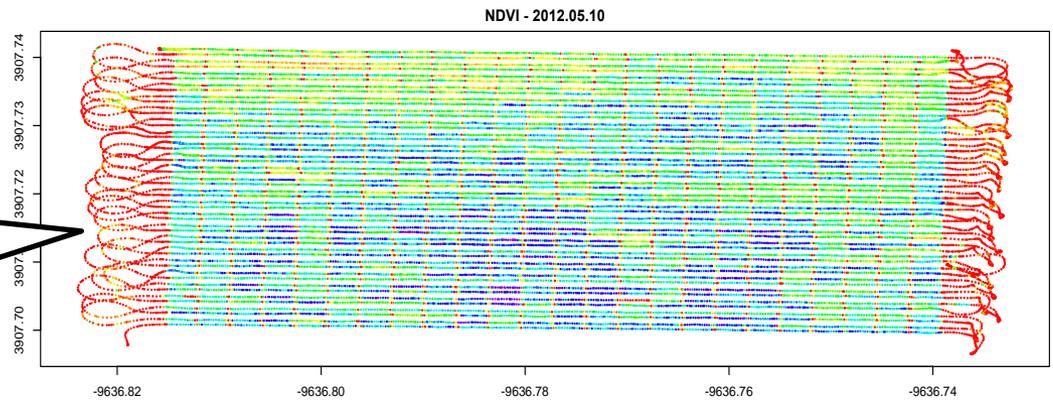
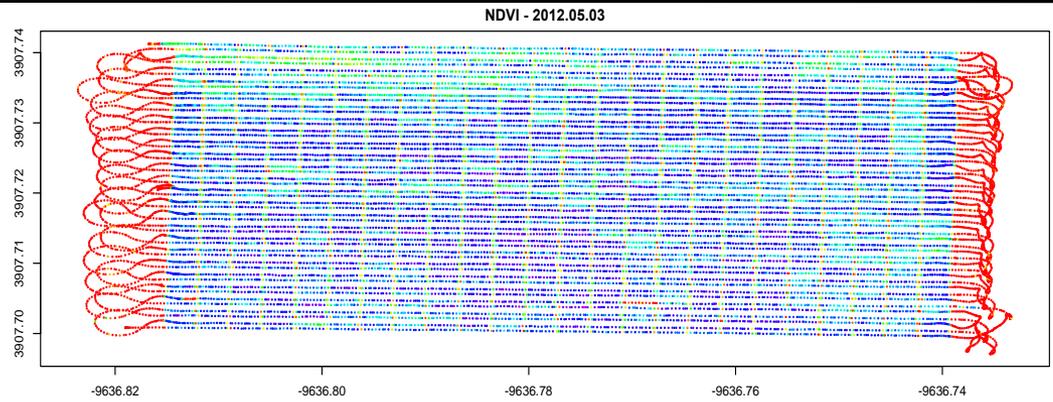


Define plot boundaries



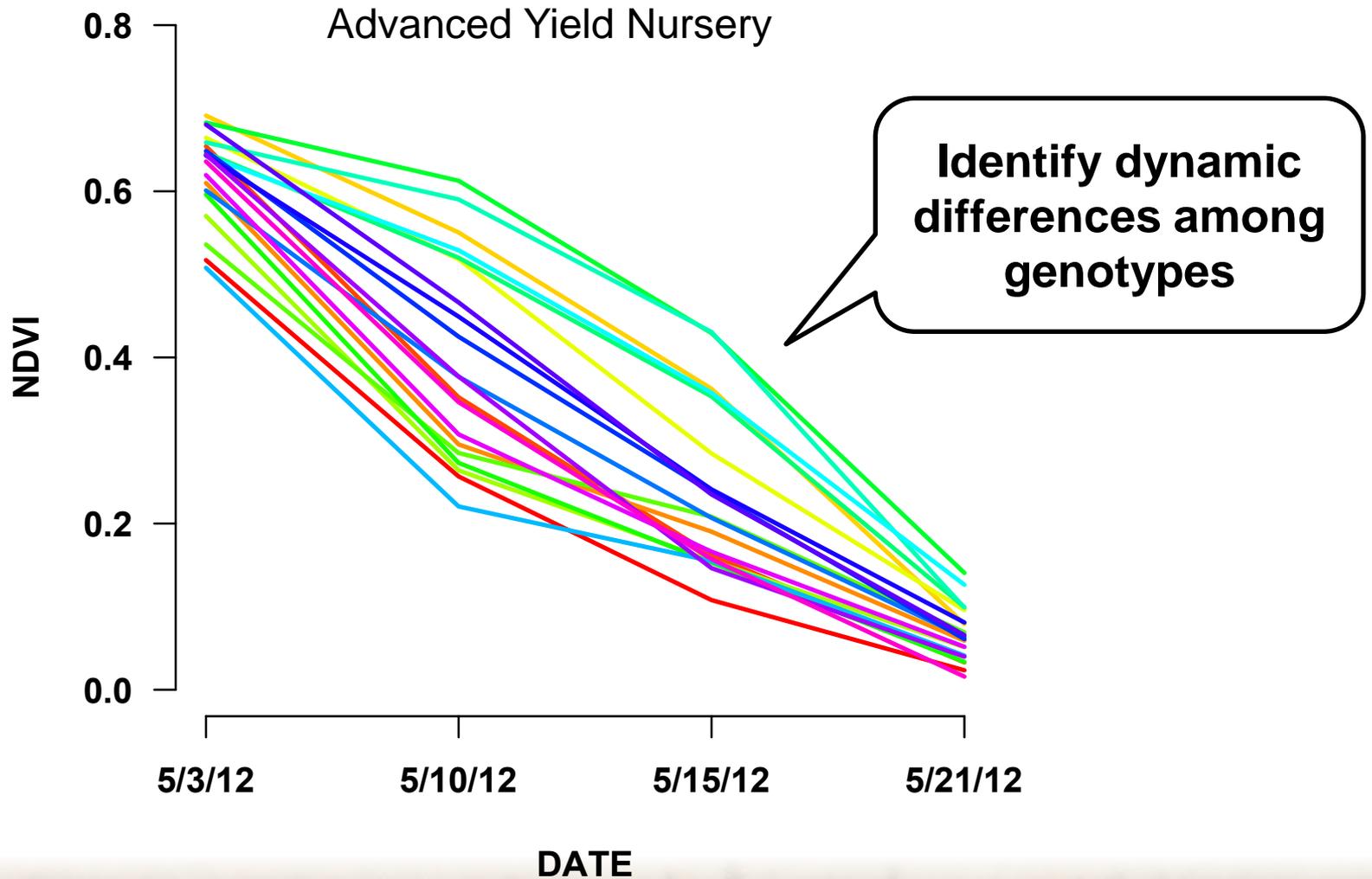
Trim data
Assign to plots

NDVI: Multi-temporal measurements



Rapid assessment
enables repeated
measurements over
time

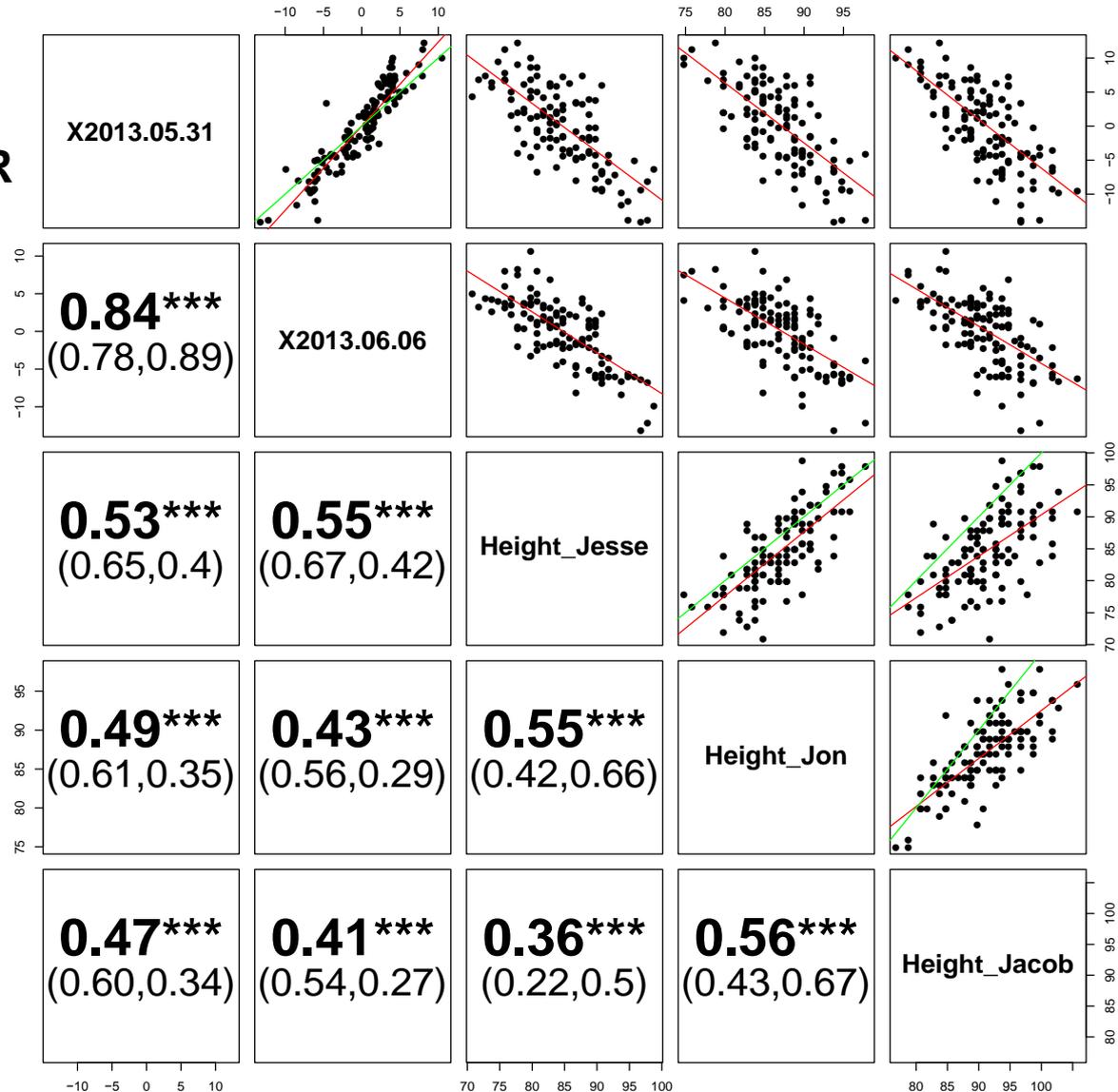
NDVI: Multi-temporal measurements



Phenotyper: Increased accuracy

Plant Height w/ SONAR

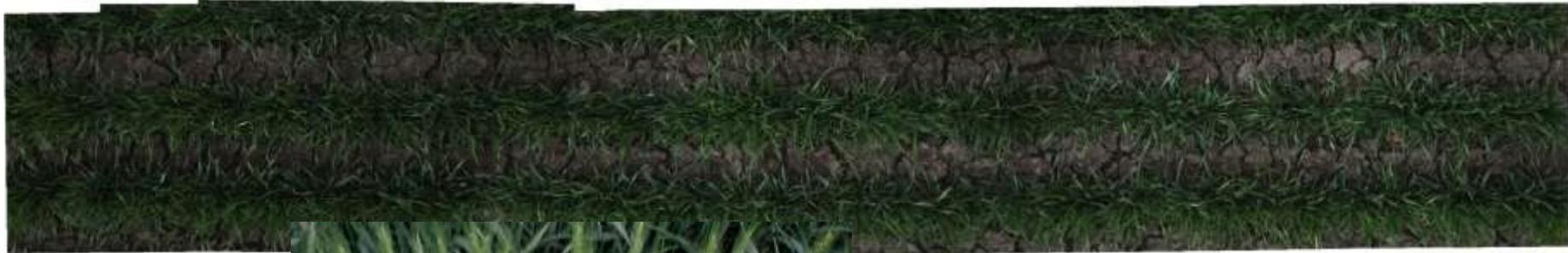
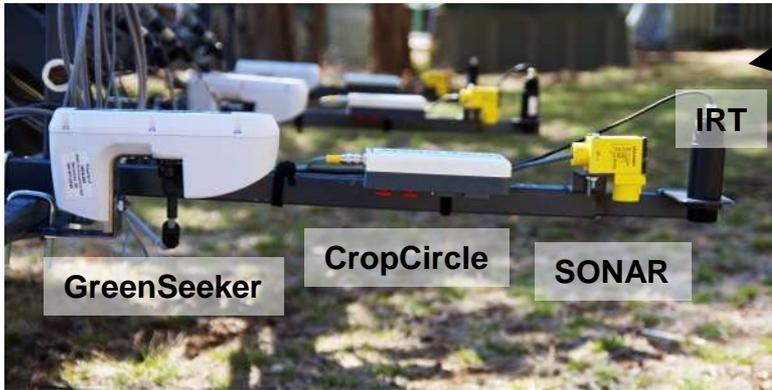
- 40 varieties
- 3 reps
- 1.3m x 3m plots



HTP platforms of all shapes and sizes...



HTP: Imaging



HTP: small Unmanned Aerial Systems (sUAS)



DJI S800 Hexacopter



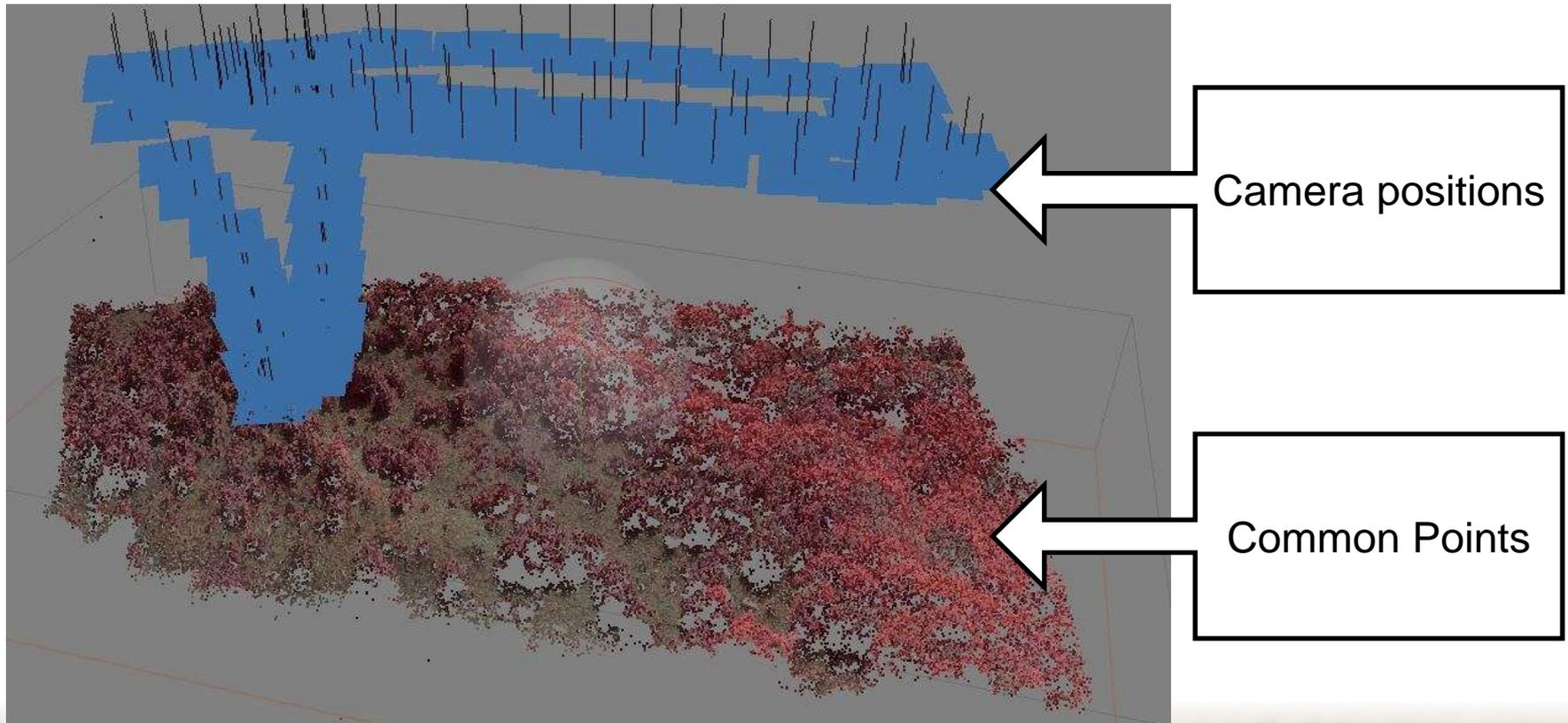
- + Not too expensive
- + flexible deployment
- + Image whole field

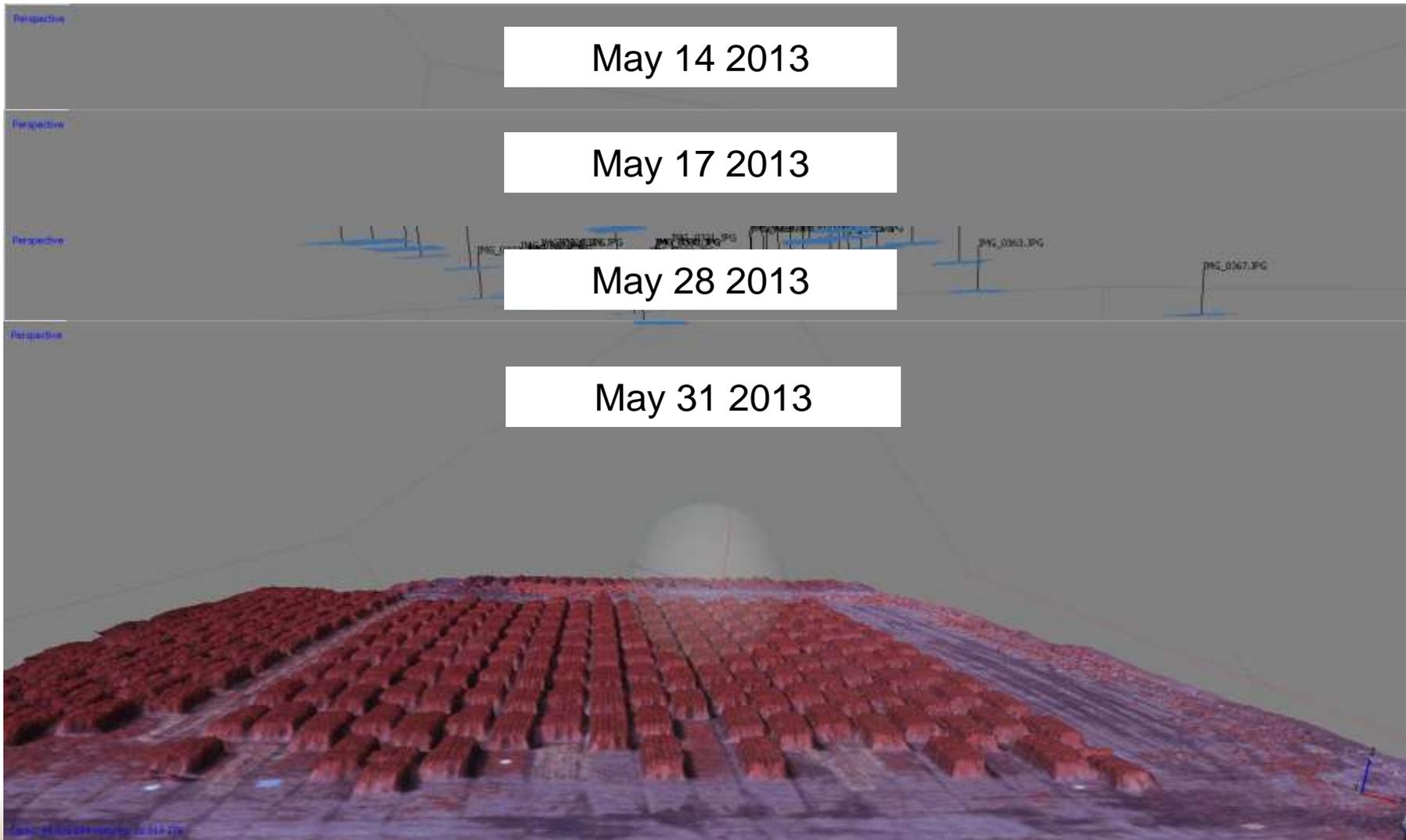


- Need trained pilot
- FAA restrictions?
- Limited payload (<1kg)
- Crashes

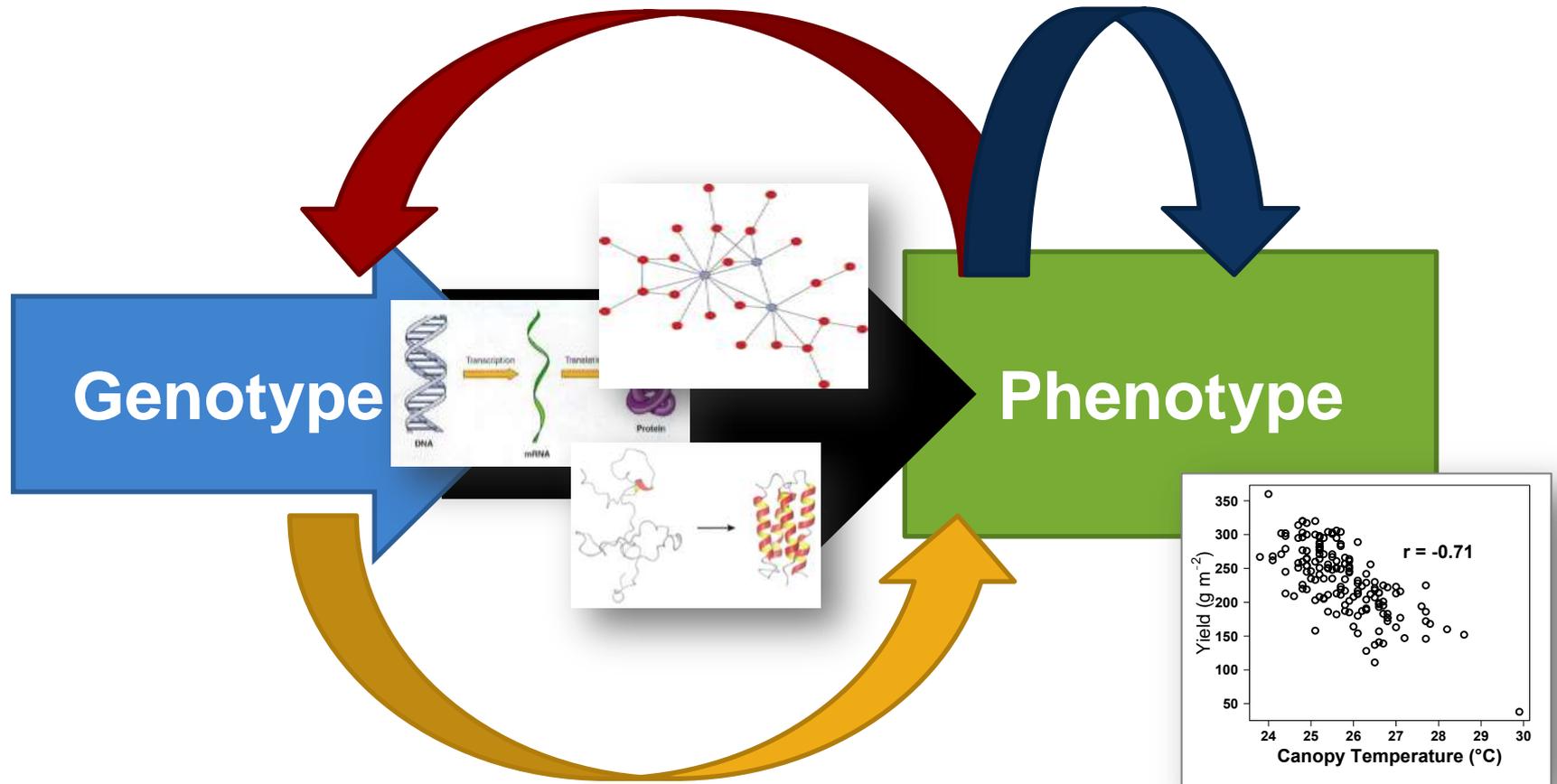
HTP: sUAS platform and 3D modeling

- Ortho mosaic from multiple images





G2P: connecting genotype to phenotype



To make it work....

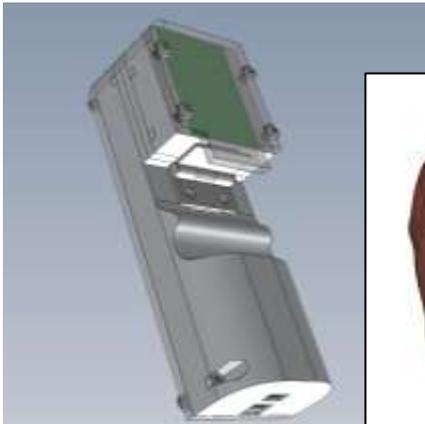
1. Start with the breeding program
 - many failures in 'genomic assisted breeding'
2. USER FRIENDLY!
 - pragmatic triumphs
 - short learning curve
3. Timely assessment under any conditions



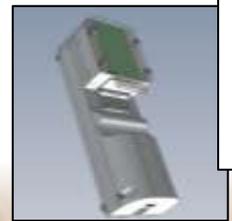
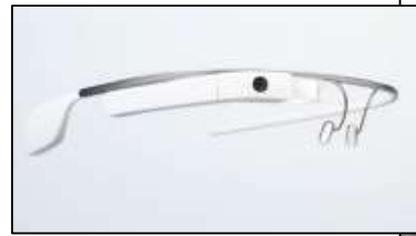
HTP: The future is here...

Implementation of existing technology

- ✓ Commercial and existing sensors
- ✓ Imaging
- ✓ Low-cost, modular 'nodes'



Interactive data collection and analysis





Shuangye Wu
 Josh Sharon ★
 Ryan Steeves ★
 Jared Crain ★
 Sandra Dunckel
 Trevor Rife
 Daljit Singh
 Narinder Singh
 Traci Viinanen
 Xu (Kevin) Wang
 Lisa Borello
 Erena Edae

Allan Fritz
 Andy Auld
 Shaun Winnie

Naiqian Zhang
 Jed Barker ★
 Spencer Kepley
 Yong (Ike) Wei
 Randy Price
 Kevin Price

Steve Welch
 Nan An★
 Dale Schinstock



Pedro Andrade-Sanchez ★
 John Heun



Jeffery White ★
 Kelly Thorp ★
 Andrew French ★
 Mike Salvucci ★

Michael Gore ★

www.fieldphenomics.org

www.wheatgenetics.org

*If we knew what it was we were doing, it would not be called research, would it?
 - Albert Einstein*



Ravi Singh
 Susanne Dreisigacker
 Matthew Reynolds
 David Bonnett
 Rick Ward

