Phenomic selection: Predicting quantitative traits in canola with NIRS-Profiles

Lennard Ehrig, Sven Weber, Stefan Abel, Reinhard Hemker, Milka Malenica, Amine Abbadi, Benjamin Wittkop, Rod Snowdon and Andreas Stahl

Model Low N

Limagrain

NPZ INNOVATION
Wir forschen für Qualität.

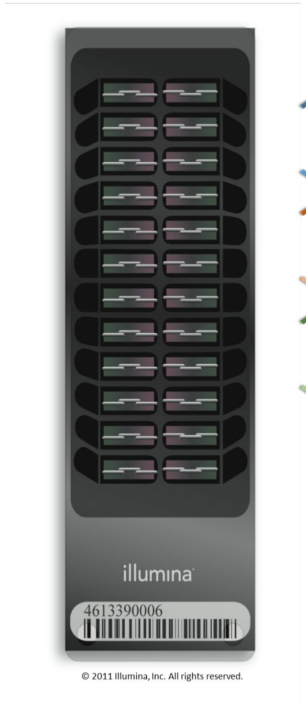
JLU
NEUE WEGE. SEIT 1607.
JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Giessen Plant Breeding

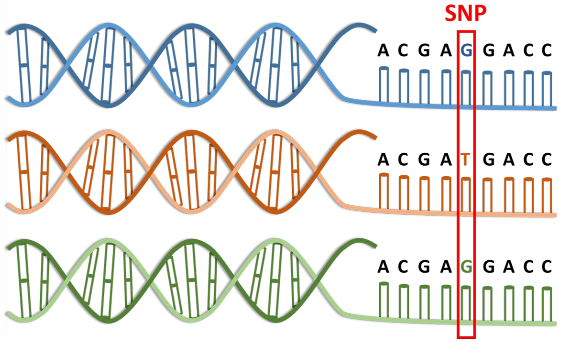
Phenomic selection

- Using NIRS as tool to compute relationship matrices for predicting complex traits as an alternative to genomic selection

Genomic selection

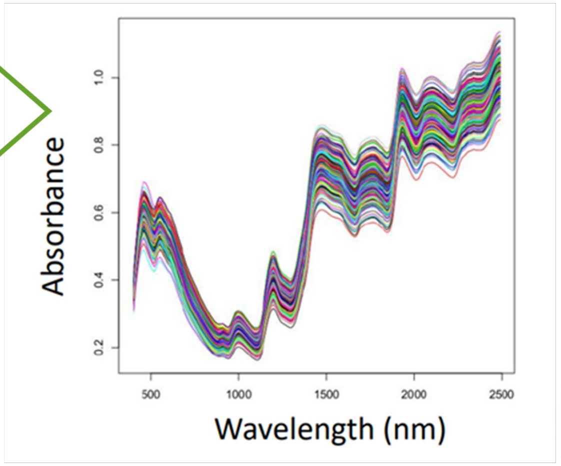


© 2011 Illumina, Inc. All rights reserved.



	SNP1	SNP2	SNP3	SNP4	...
G1	0	1	1	2	...
G2	2	1	0	1	...
G3	2	1	1	0	...
G4	1	1	2	1	...
....	...	...	...	...	...

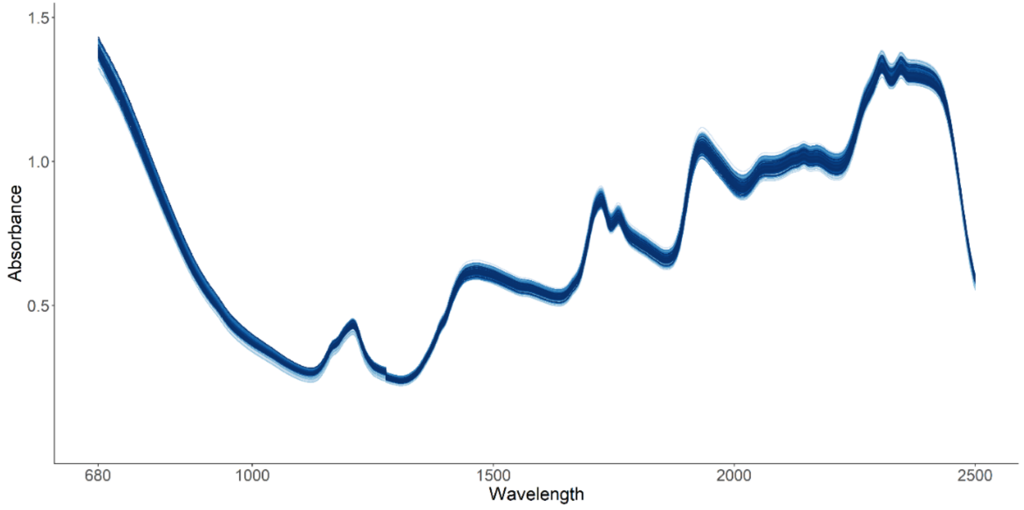
Phenomic selection



	680	681	682	683	...
G1	1,360	1,255	1,353	1,353	...
G2	1,367	1,365	1,363	1,360	...
G3	1,372	1,368	1,367	1,363	...
G4	1,359	1,367	1,366	1,362	...
....	...	...	...	...	...

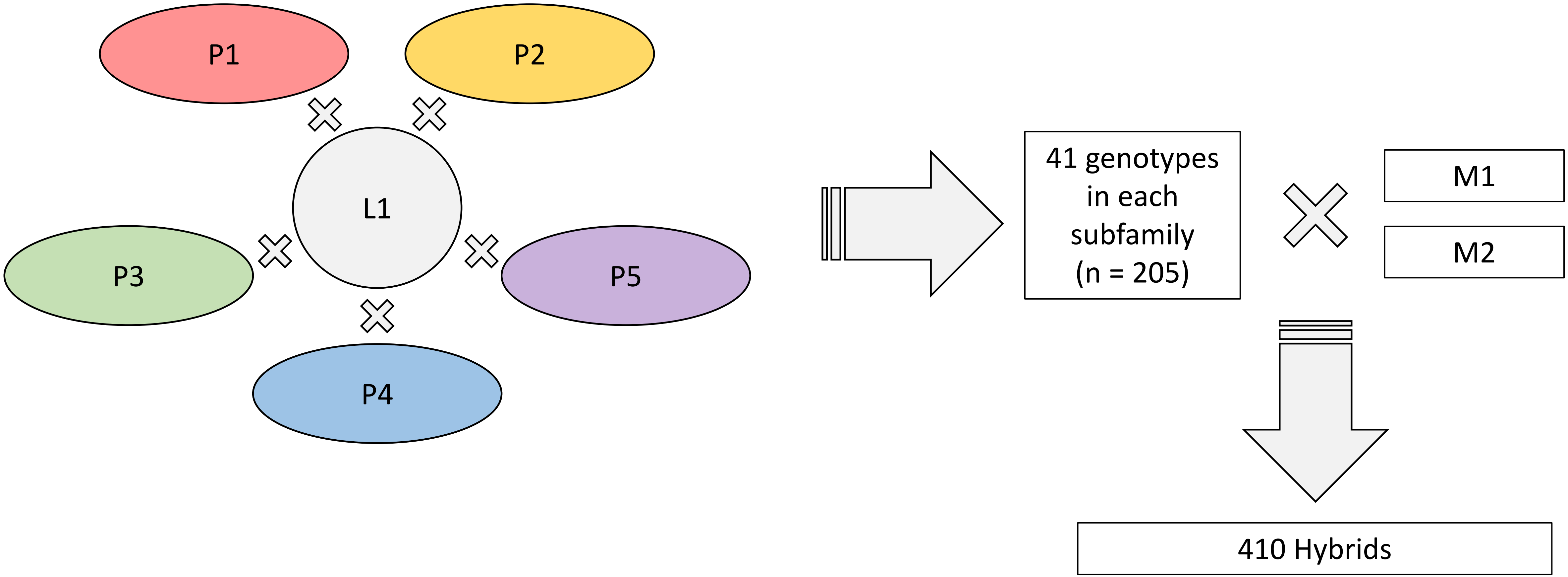
Advantages of phenomic selection

- High-throughput
- Low cost
- (Non-destructive)



- NIRS is often measured routinely to phenotype for metabolic components
  - Glucosinolate
  - Oil and protein content

Materials



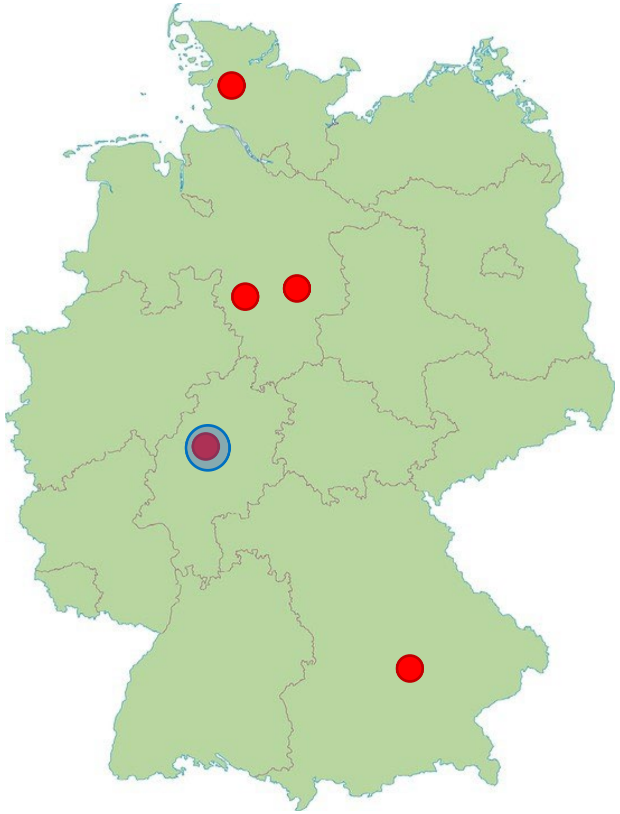
P1
P2
L1
P3
P5
P4

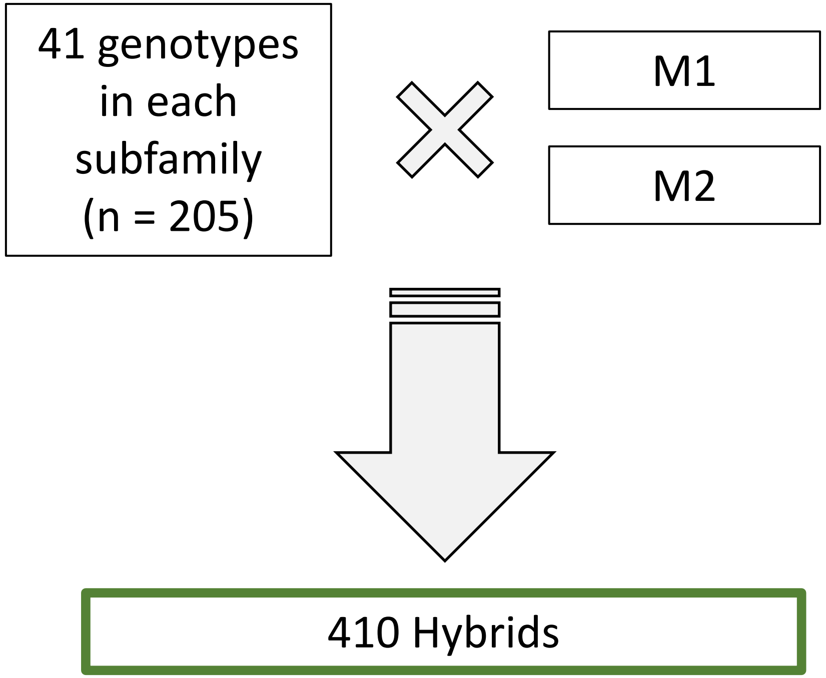
41 genotypes in each subfamily (n = 205)

M1
M2

410 Hybrids

Phenotyping





41 genotypes in each subfamily (n = 205)

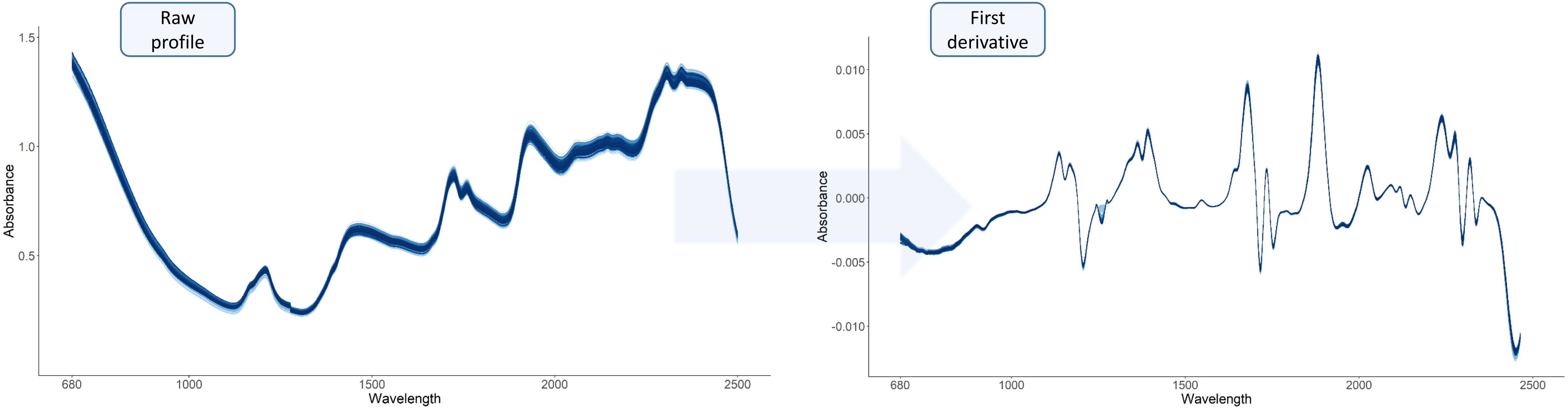
M1

M2

410 Hybrids

- Grown in 2019/2020 and 2020/2021
- P-rep design with 1.2 replications
- Phenotyped for yield, plant height, flowering time, (oil and protein content)

NIRS profile

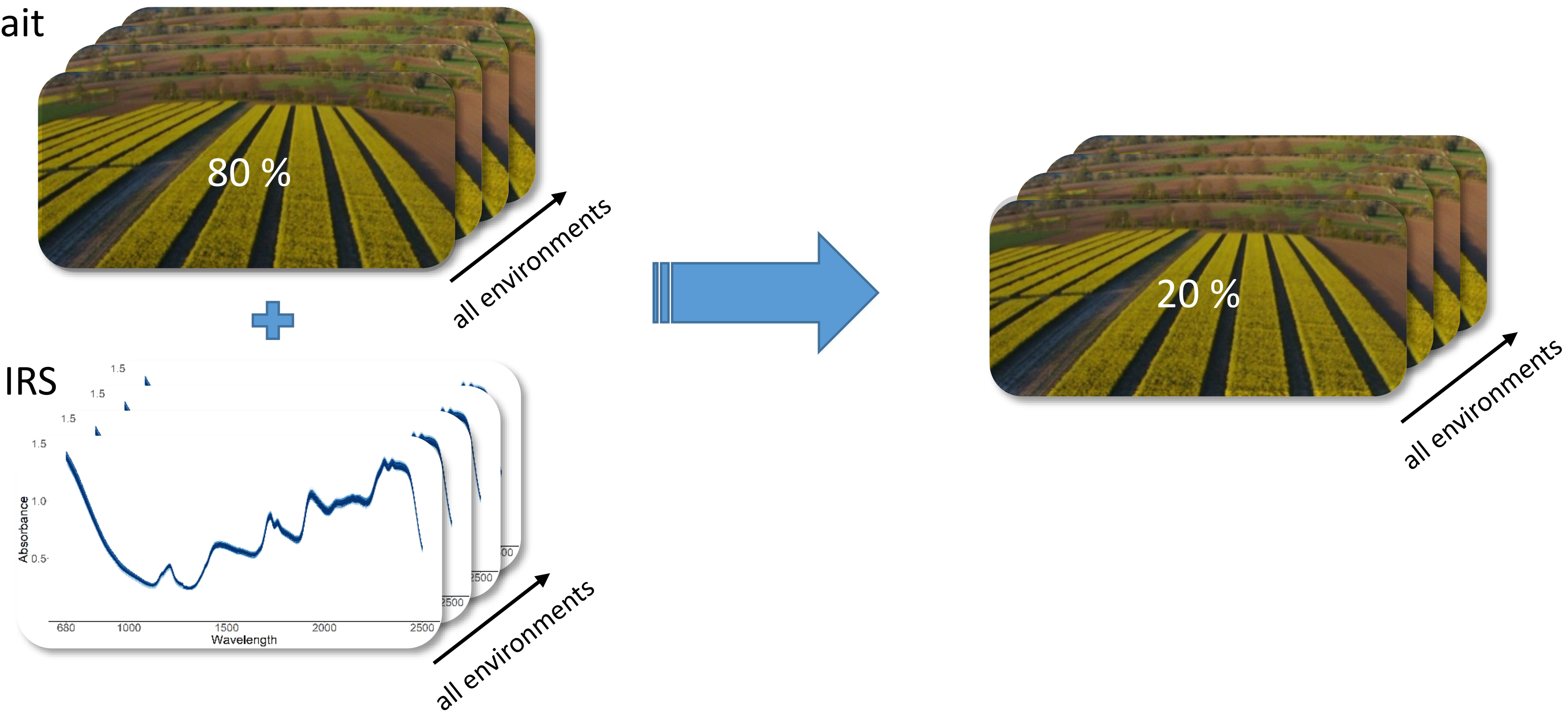


- 680 – 2500 in 1 nm steps, measured on F2 seeds
- Each Wavelength adjusted for environmental and field design effects
- First derivative built by using Savitzky–Golay filter with a window size of 37 then data was scaled and centred

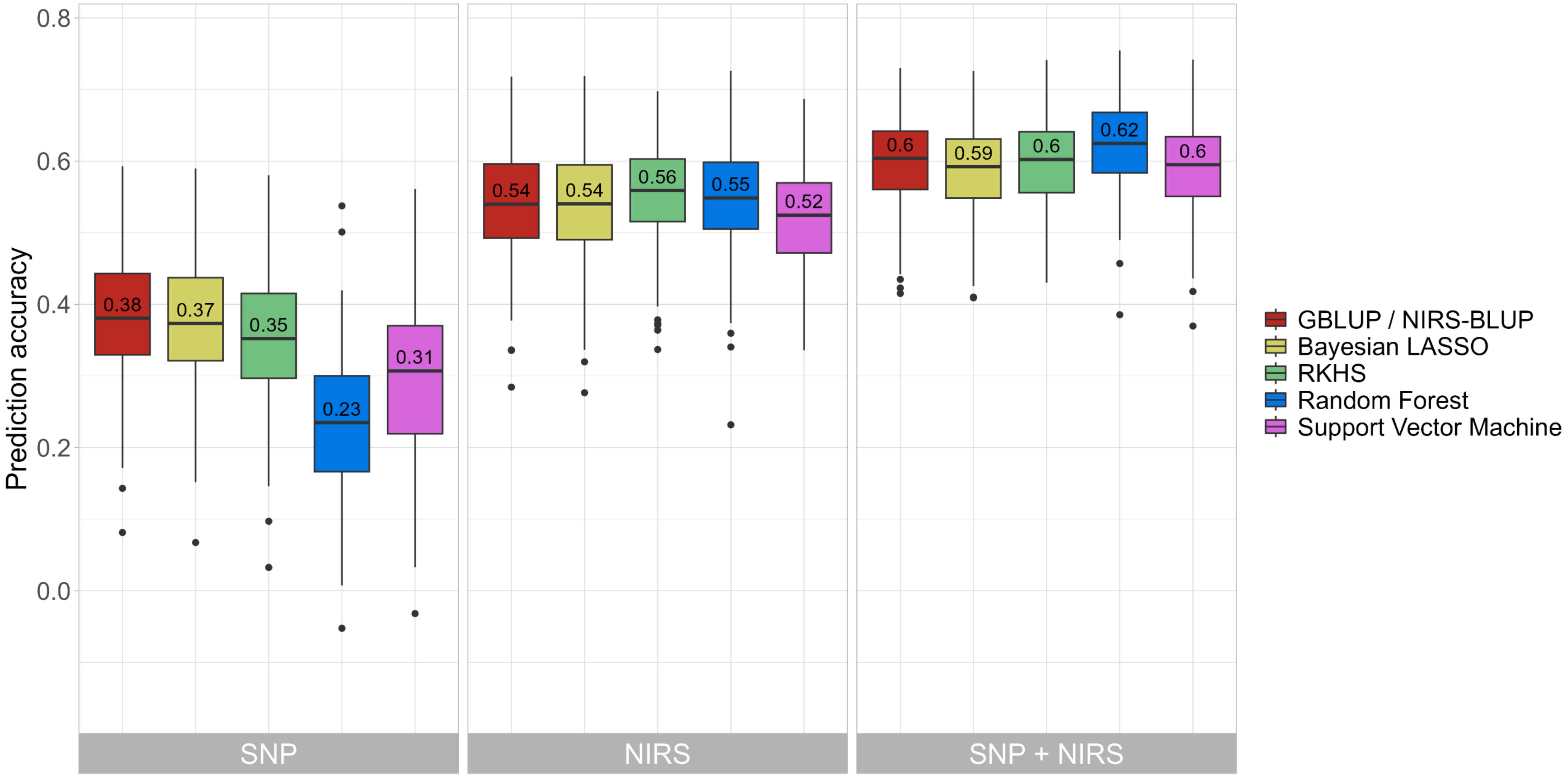
Scenario 1

Training

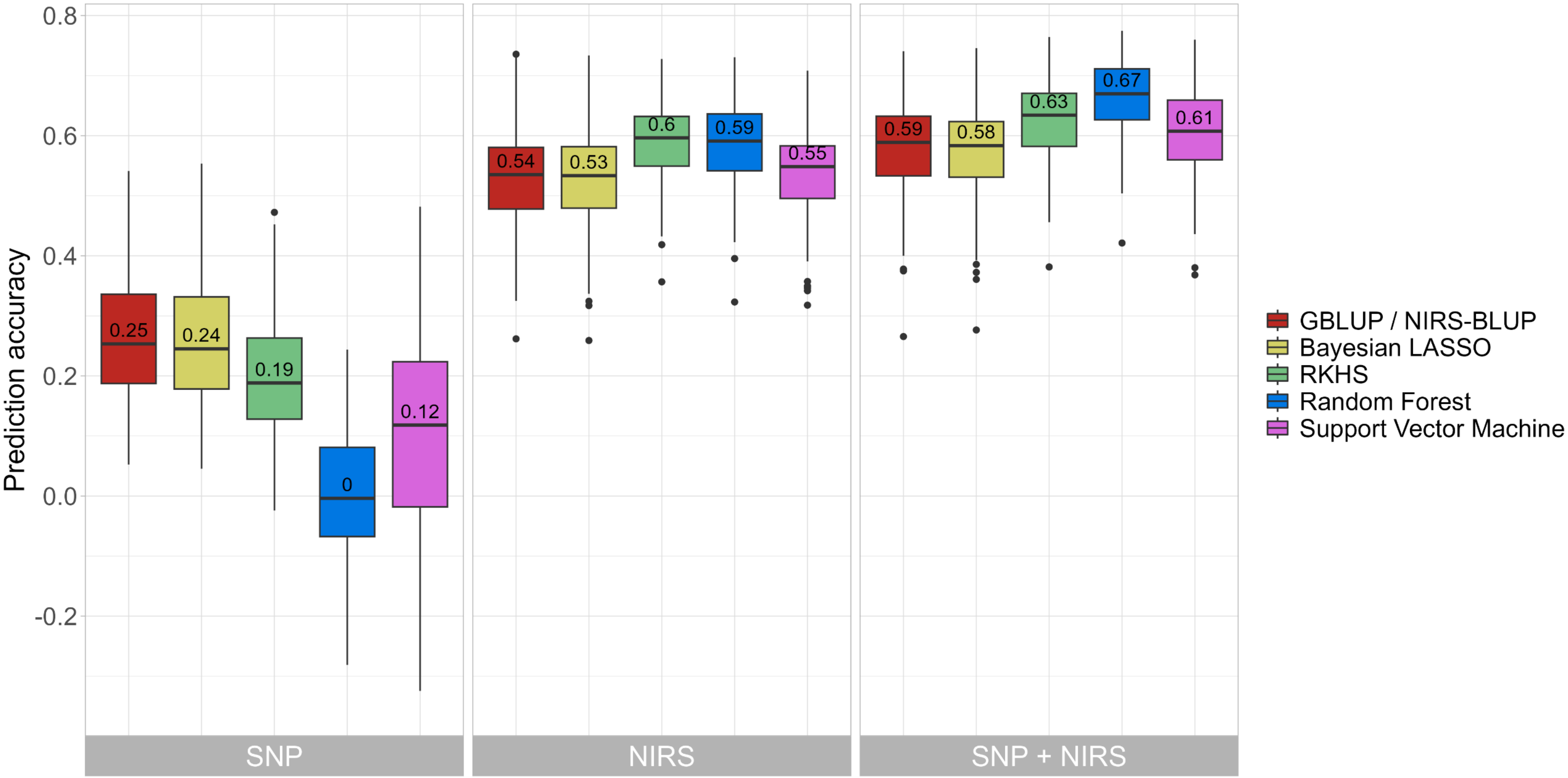
Prediction/Test



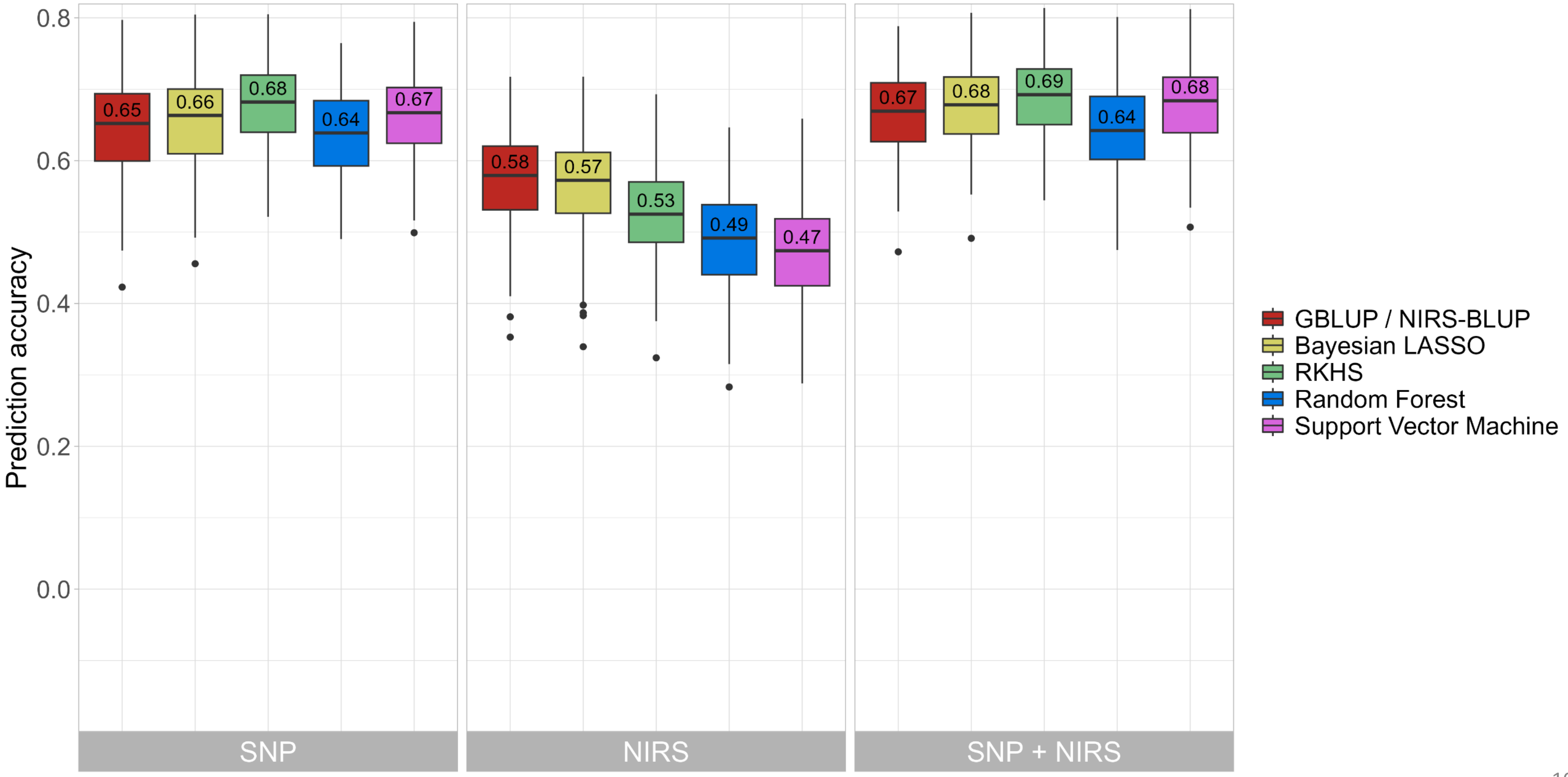
S1: Yield prediction



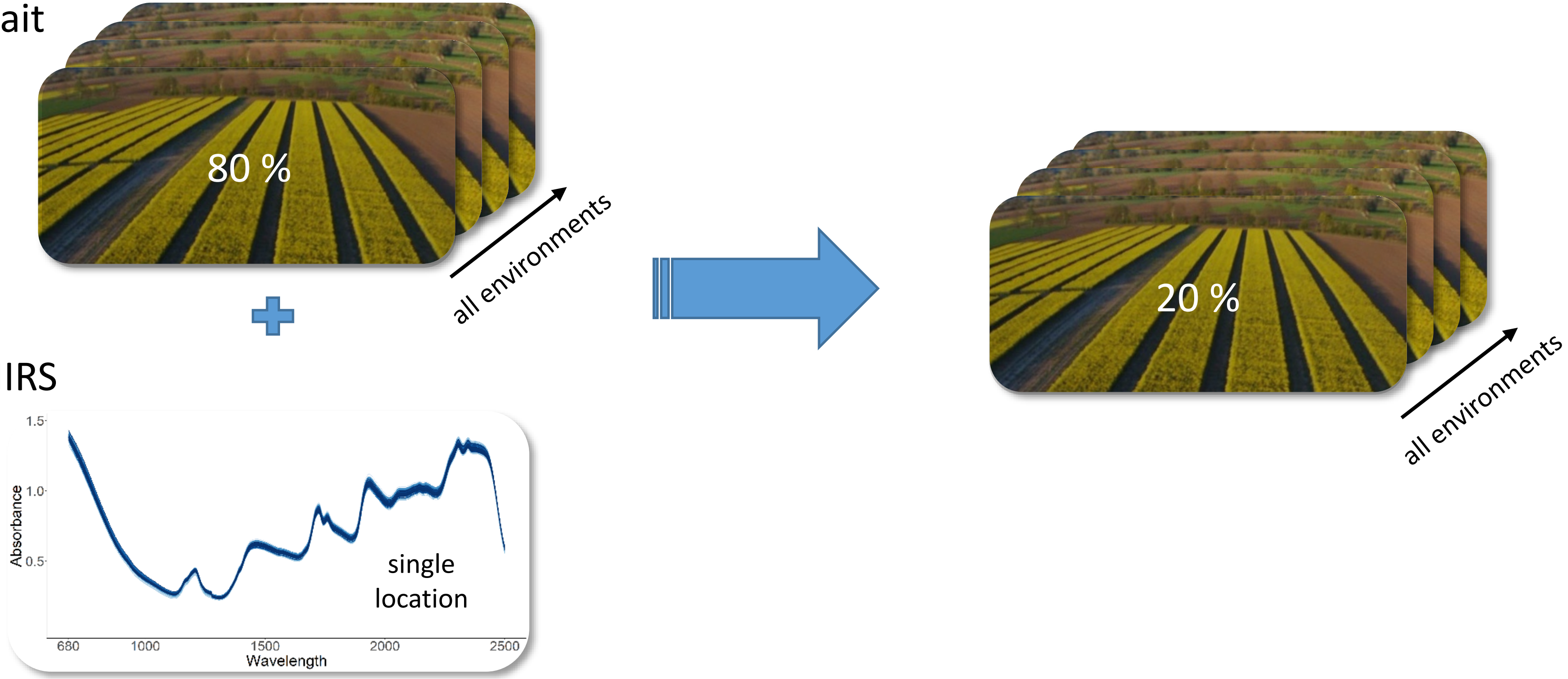
S1: Plant height



S1: Flowering time



Scenario 2



Training

Trait

80 %

all environments

NIRS

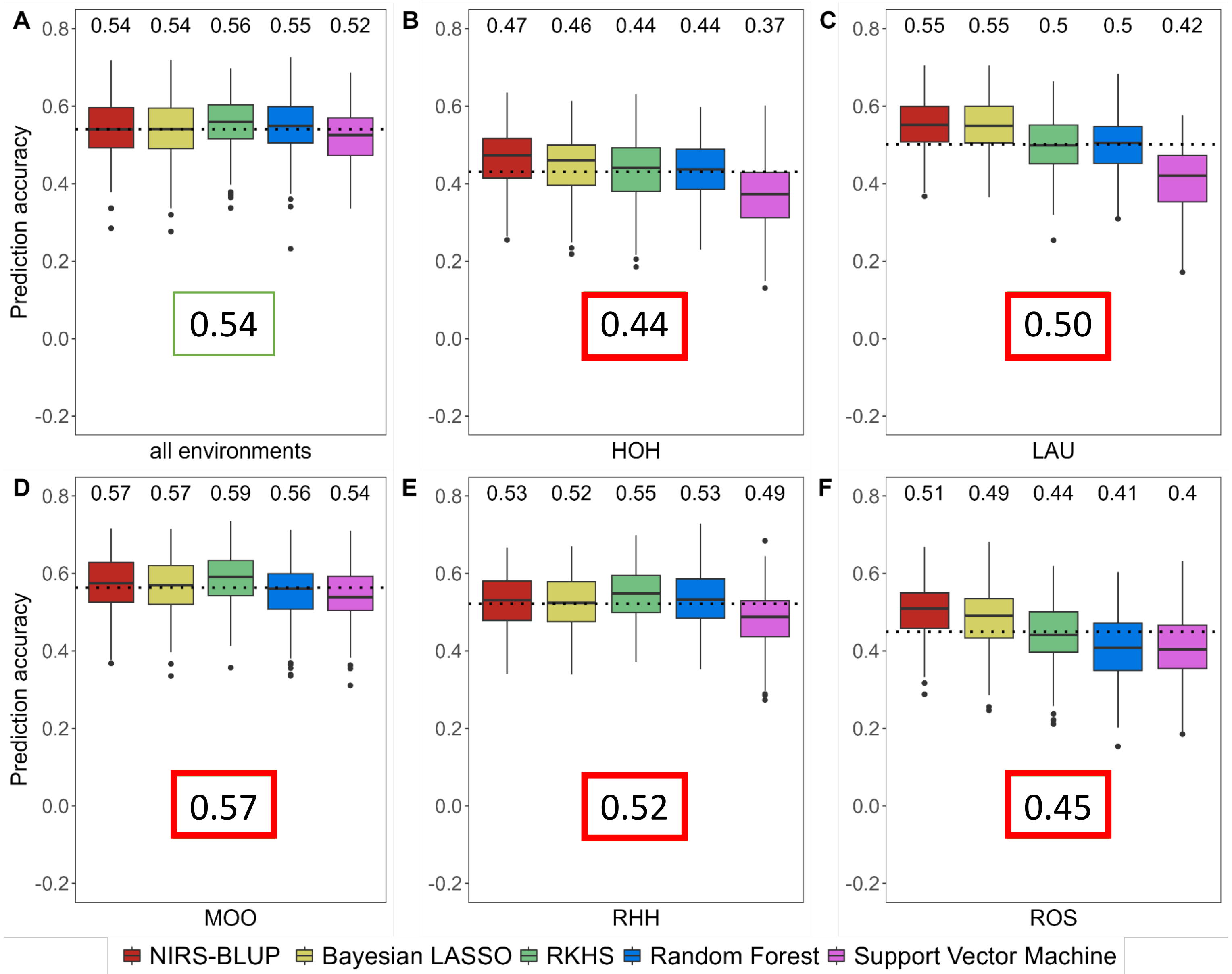
single location

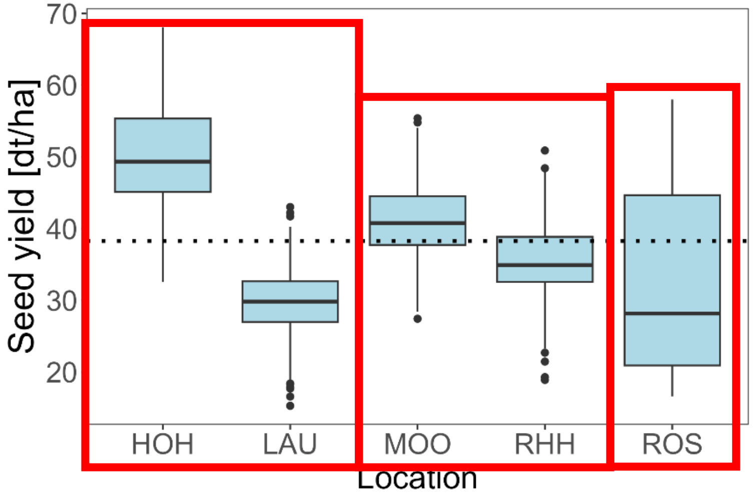
Prediction/Test

20 %

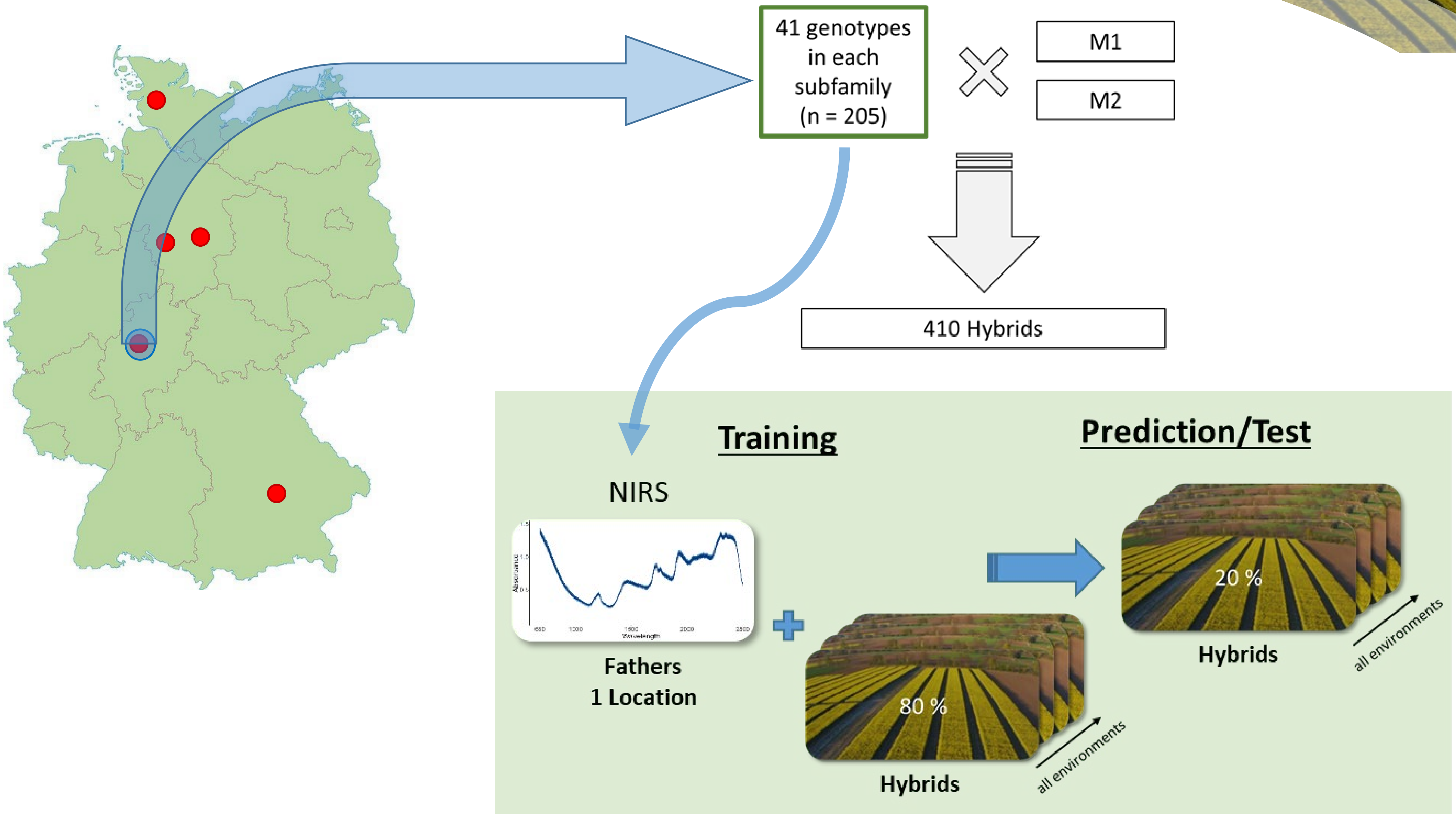
all environments

Seed yield


A
B
C
D
E
F
Prediction accuracy
0.8
0.6
0.4
0.2
0.0
-0.2
0.54 0.54 0.56 0.55 0.52
0.54
all environments
0.47 0.46 0.44 0.44 0.37
0.44
HOH
0.55 0.55 0.5 0.5 0.42
0.50
LAU
0.57 0.57 0.59 0.56 0.54
0.57
MOO
0.53 0.52 0.55 0.53 0.49
0.52
RHH
0.51 0.49 0.44 0.41 0.4
0.45
ROS
🟥 NIRS-BLUP
🟨 Bayesian LASSO
🟩 RKHS
🟦 Random Forest
🟪 Support Vector Machine


Seed yield [dt/ha]
70
60
50
40
30
20
HOH
LAU
MOO
RHH
ROS
Location

Parental prediction



Fathers
1 Location

41 genotypes
in each
subfamily
(n = 205)

×

M1

M2

410 Hybrids

Training

NIRS

Fathers
1 Location

\+

80 %

Hybrids

all environments

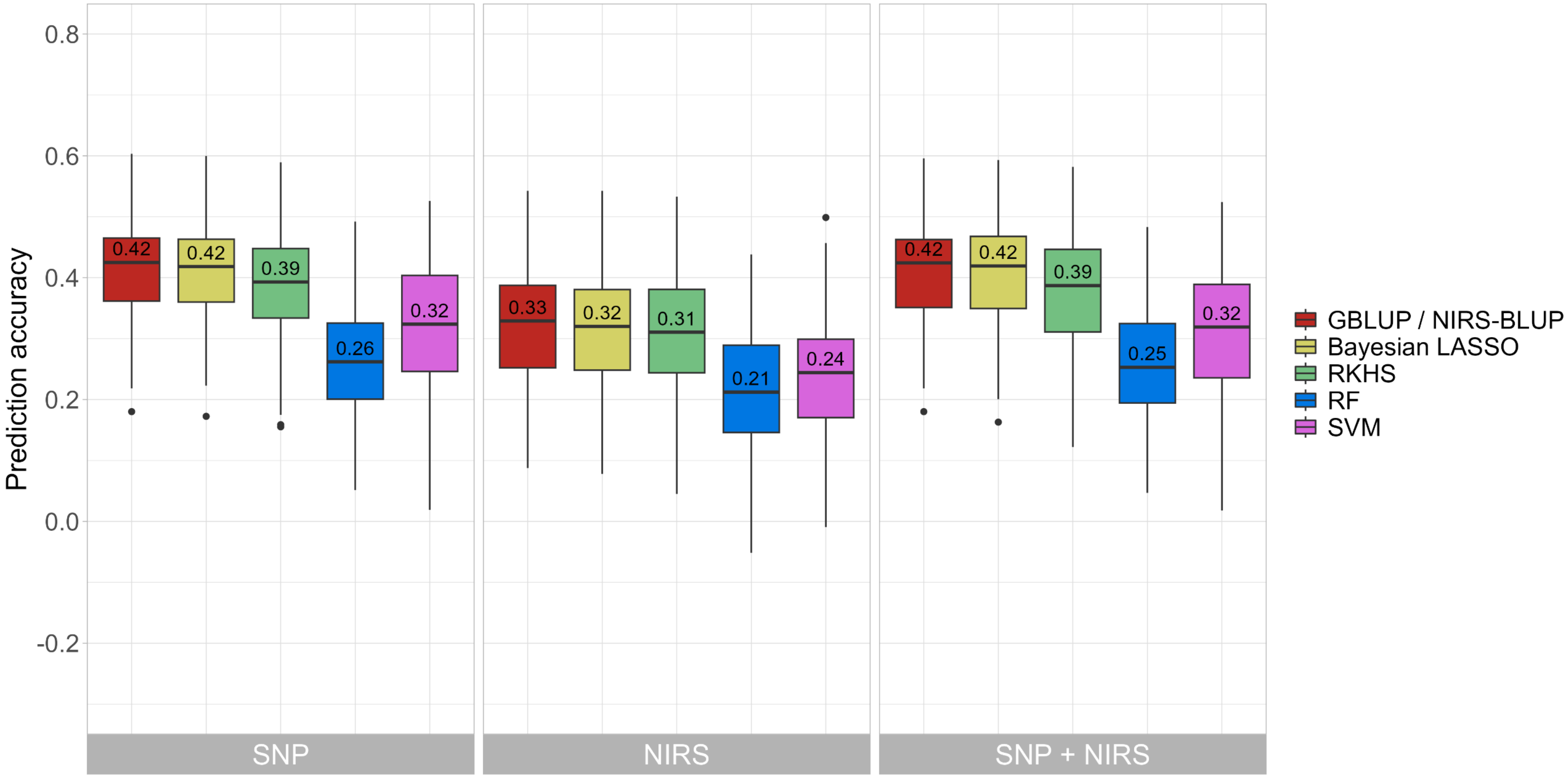
Prediction/Test

20 %

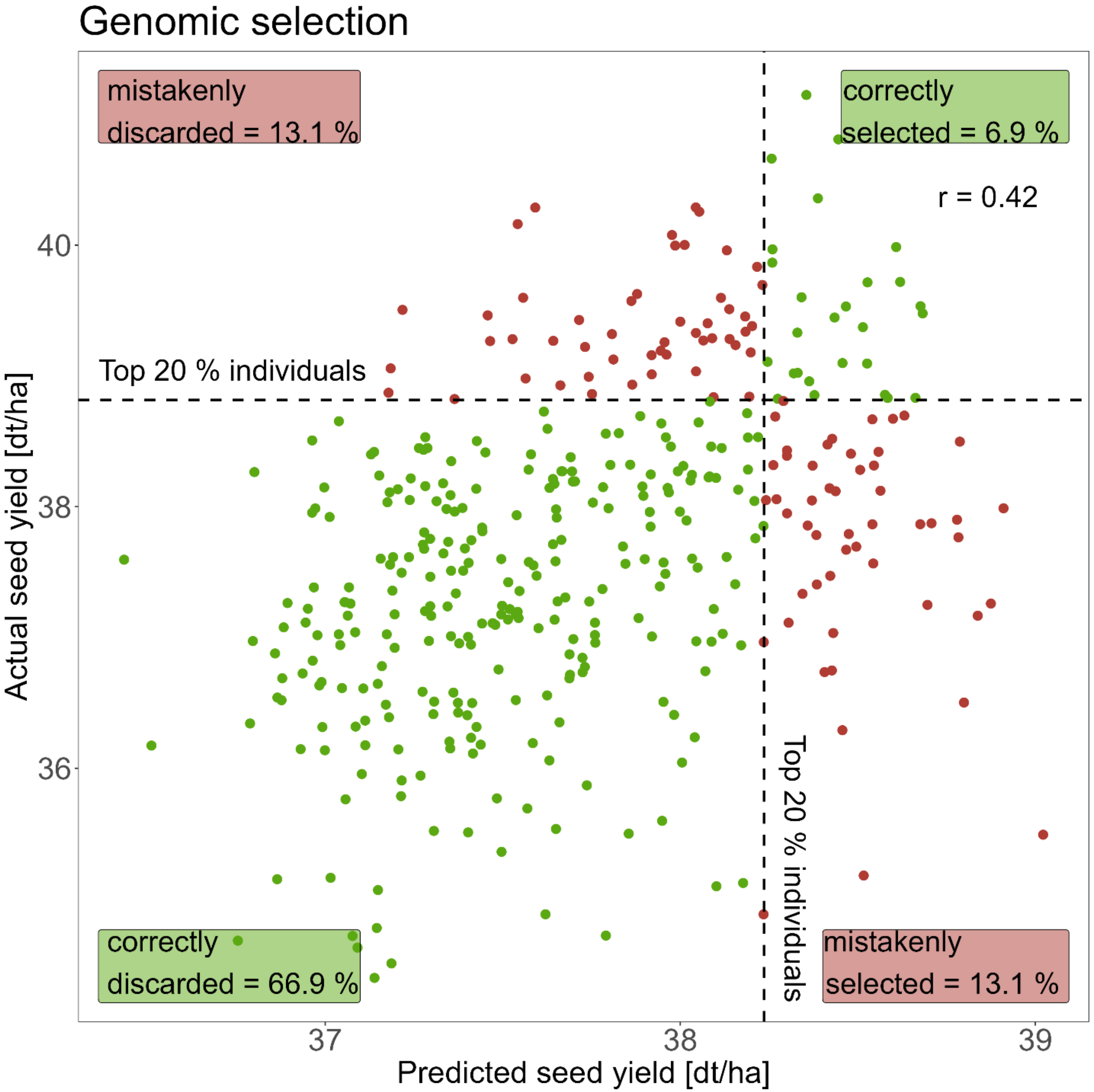
Hybrids

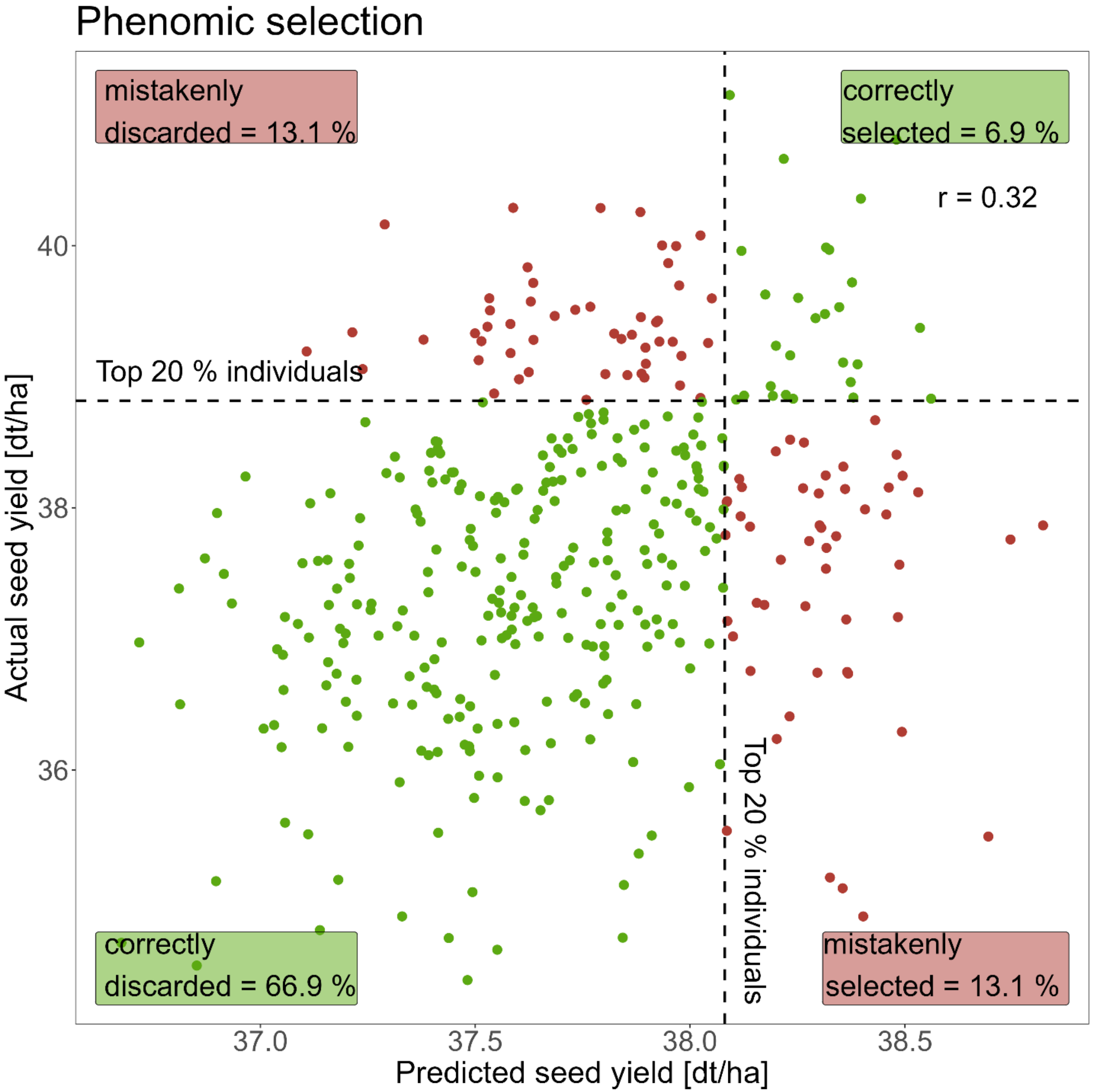
all environments

Parental prediction: Yield



Parent based selection





Why does it work?

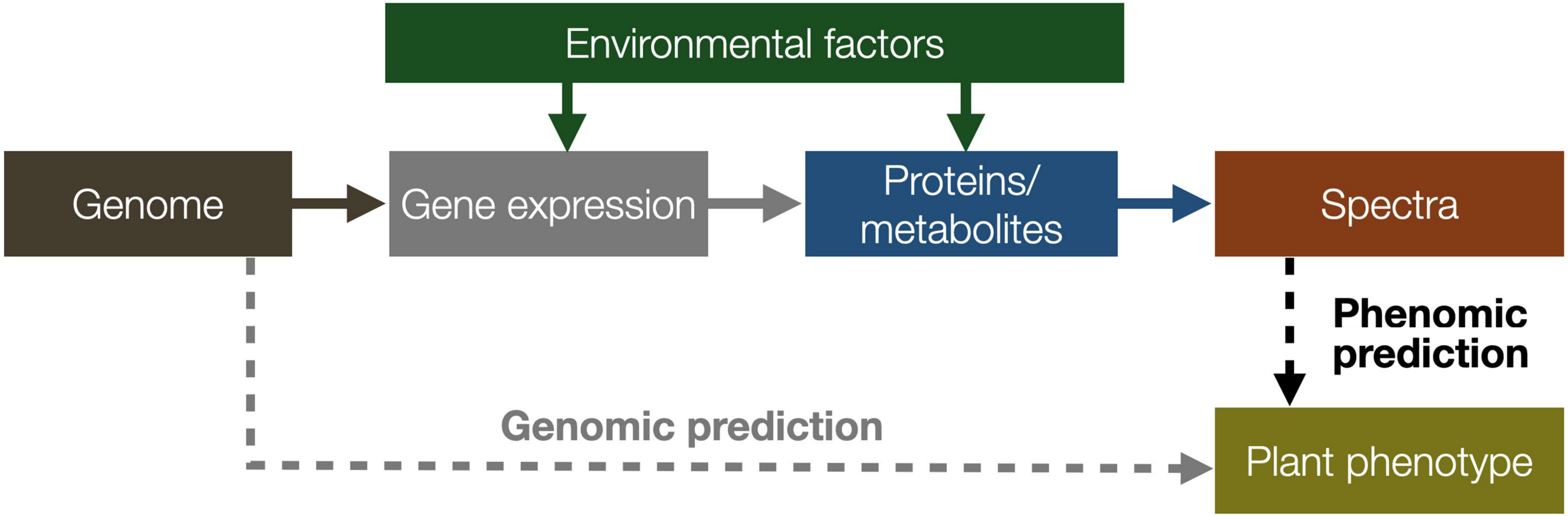


Figure adapted from Charlotte Brault

Hypothesis:

- Spectra indirectly capture genomic variation
- Spectra capture GxE

Conclusions

- Prediction accuracy of **PS outperformed GS** when using NIRS obtained from F2-seeds
- Ongoing **GS** programs can be **improved by including spectral data**
- NIRS data can come from a **single reference location**
- Using **parental NIRS** profiles to select hybrids with high yield works **as good as GS** and is **cheaper** and **faster**

Thank you!

Limagrain

- Stefan Abel
- Reinhard Hemker

NPZ INNOVATION
Wir forschen für Qualität.

- Amine Abbadi
- Dirk Stulgies
- Milka Malenica

Y=G x E x M
CO_2
N_2O
Model Low N

JLU
NEUE WEGE. SEIT 1607.

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Sven Weber
- Benjamin Wittkop
- Rod Snowdon
- Andreas Stahl

Gefördert durch
Bundesministerium für Ernährung und Landwirtschaft
aufgrund eines Beschlusses des Deutschen Bundestages

Projektträger
Bundesanstalt für Landwirtschaft und Ernährung

Lennard Ehrig