**Fargo Genotyping Laboratory**

Jason Fiedler
Research Plant Geneticist
USDA-ARS Small Grains
Genotyping Laboratory
Fargo, ND

Multi-species 3K genotyping platform

Outline

- Array overview

- Technical validation

- Do the array SNPs match the source SNPs?

- Can we use imputation to increase the # of SNPs (harmonization)?

- Can dual hybridization be used to decrease cost?

- Are the markers suitable for GS?
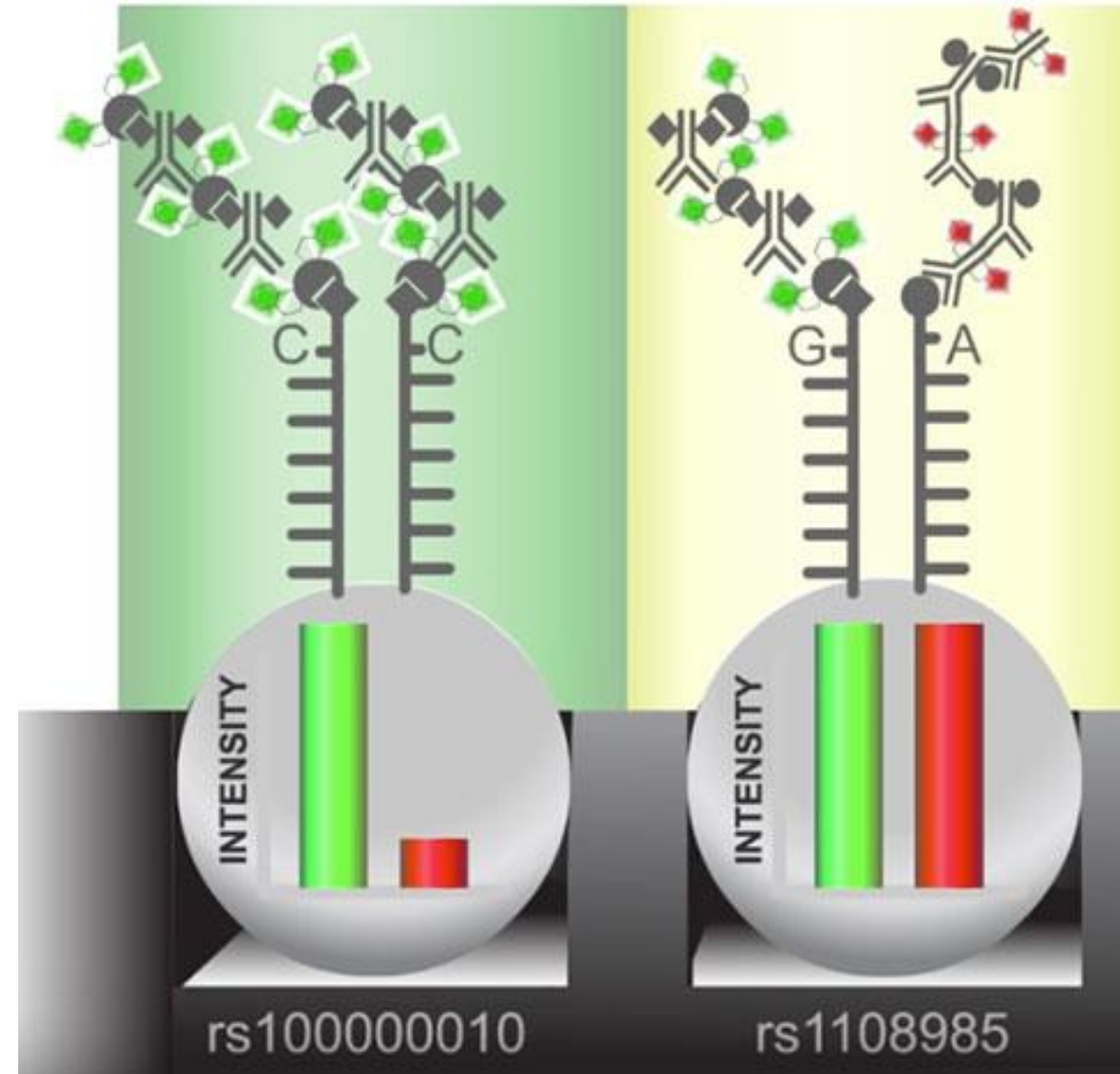
The Illumina multi-species 3K array

- Fargo has significant infrastructure/expertise in Illumina arrays.
  - Iscan – 250K, Tecan Robot – 110K, Misc - 80K

- Robust durable technology, fast protocol, requires no bioinformatics.
    - GUI for clustering

- Less flexible – add-on content or redesigns every 3-4 years

- More expensive

- ~3,000 SNPs each in four crop species (Soy, Wheat, Oat and Barley).

- $14/sample price ($12 for a larger order)

- USDA-SoyWheOatBar-3K
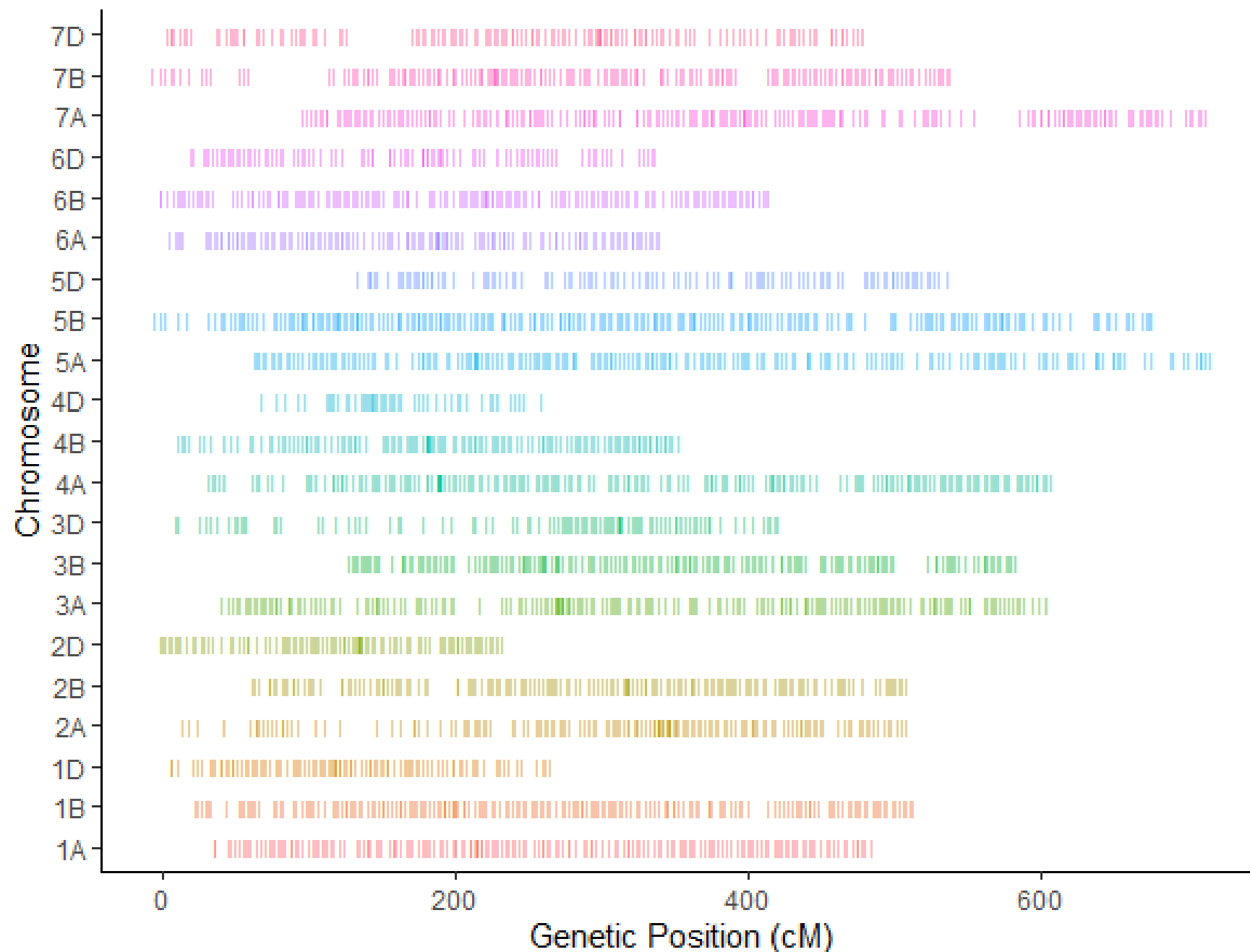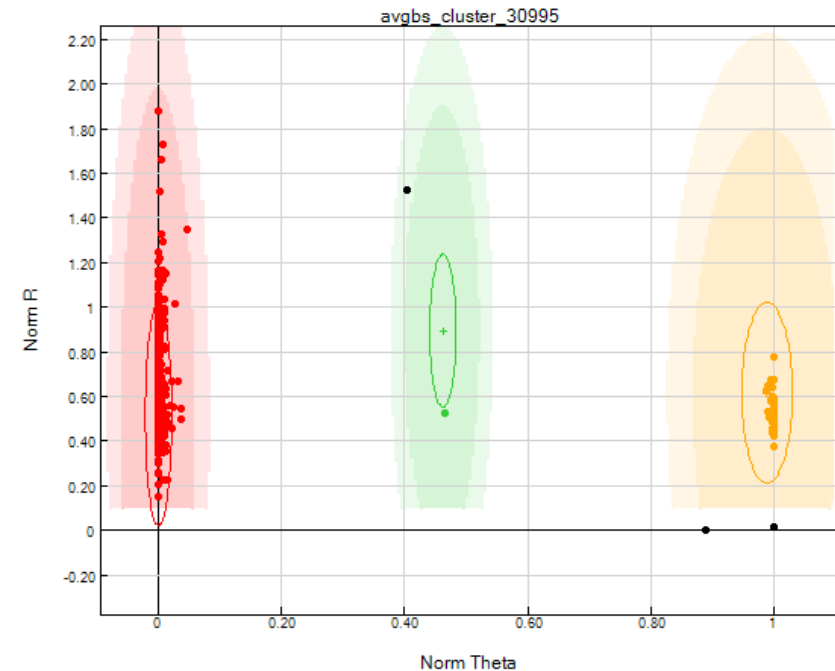
- Been validating since June 2022
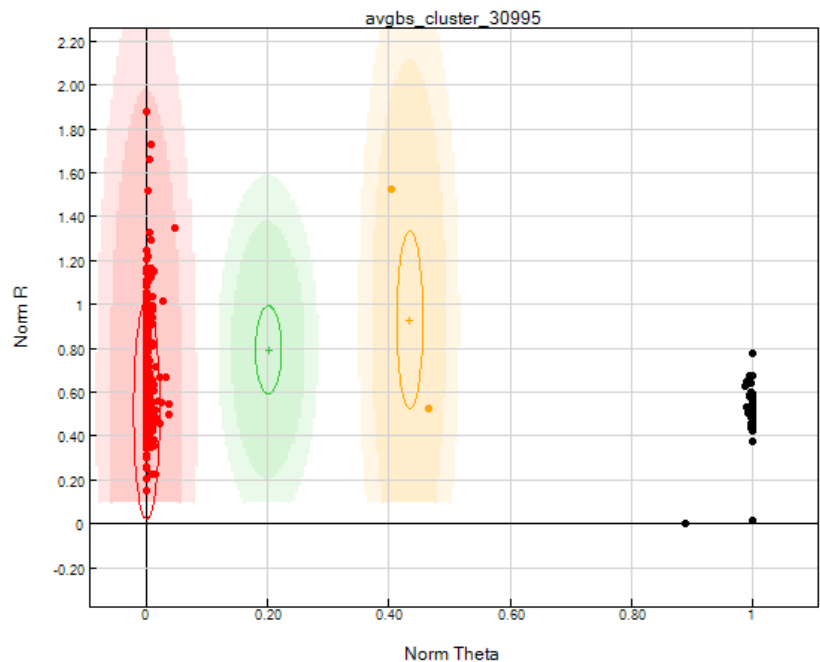
Wheat 3K Positions
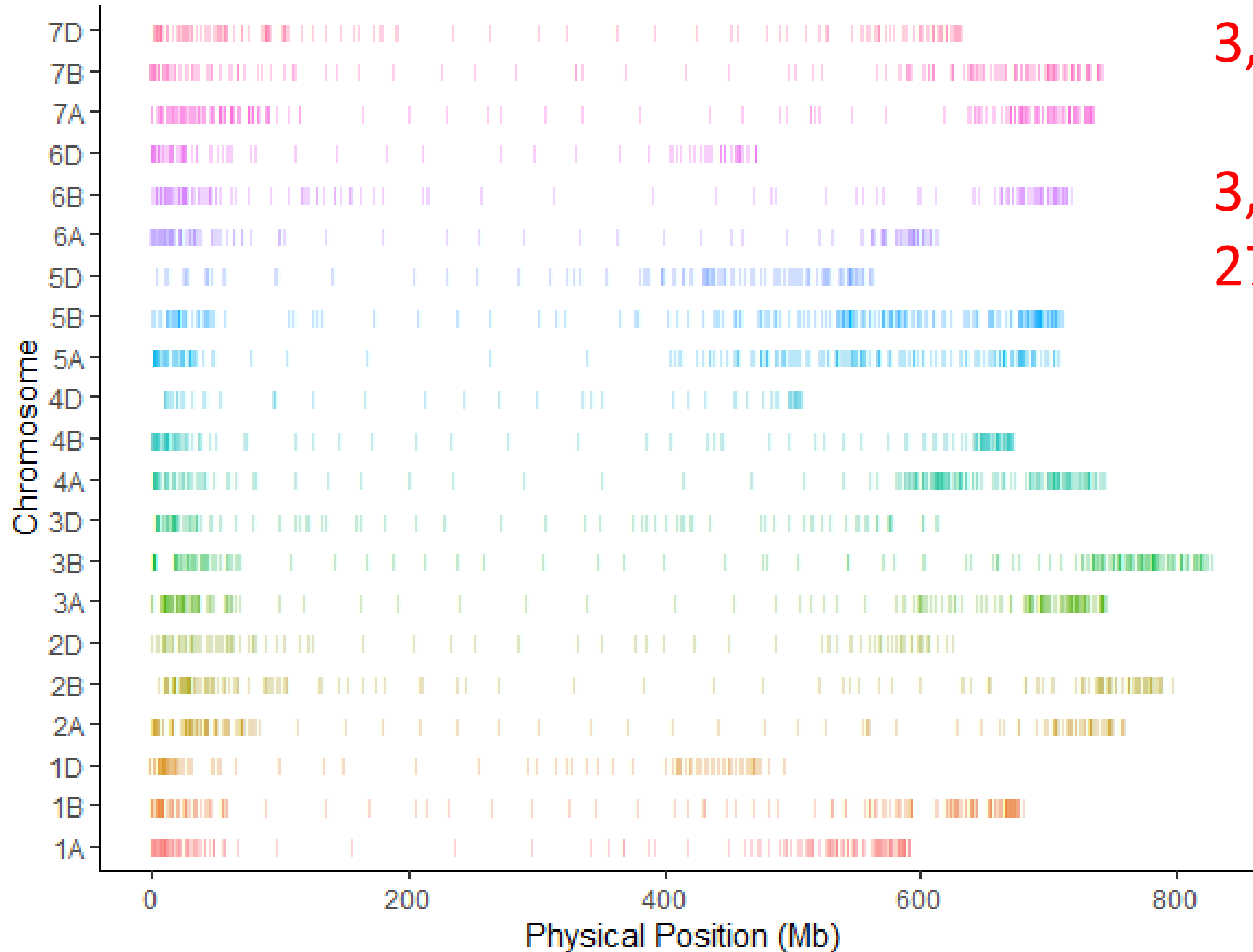
3,422 SNPs
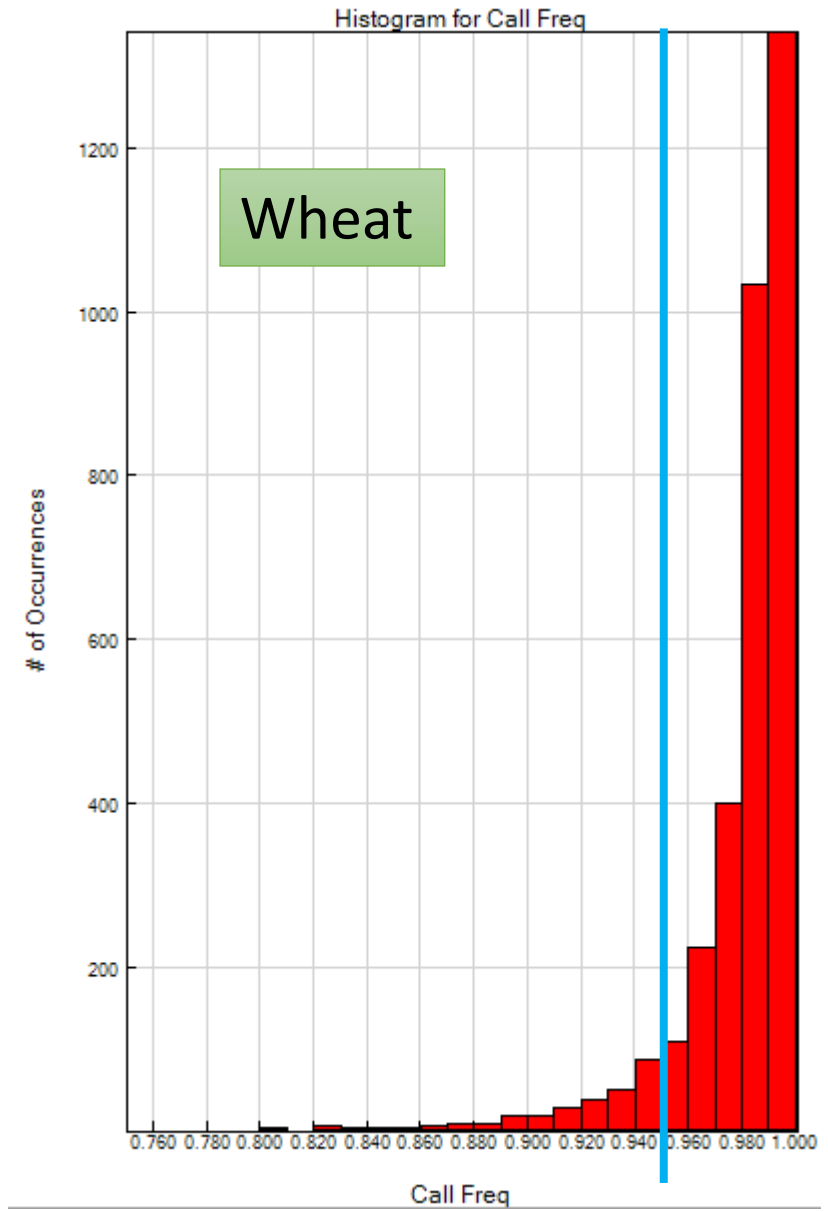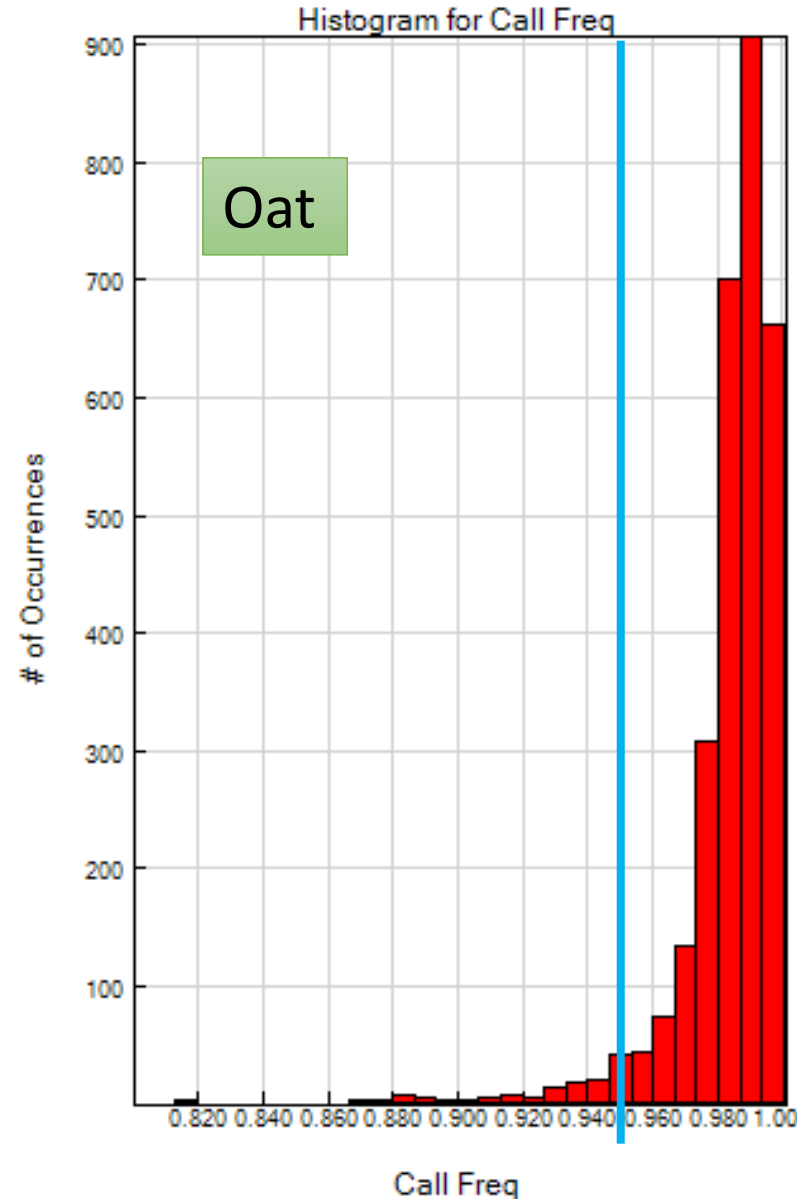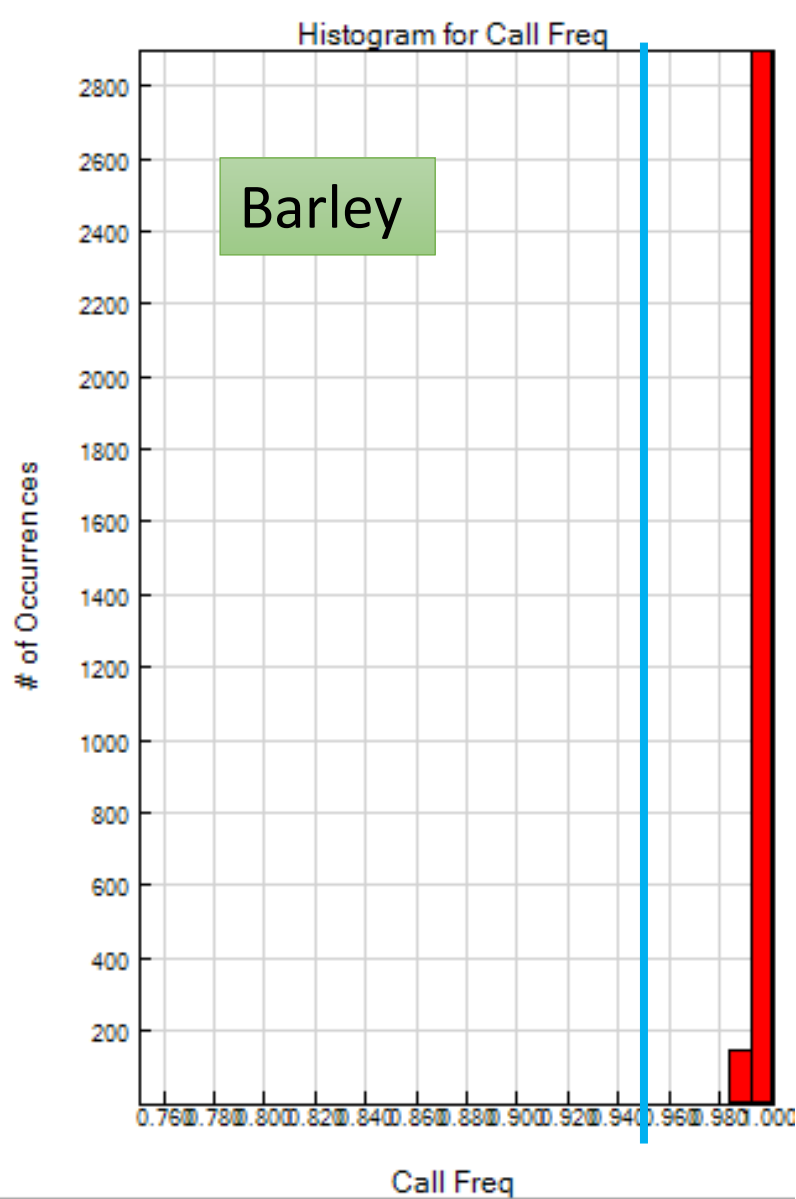
3,149 linkage blocks
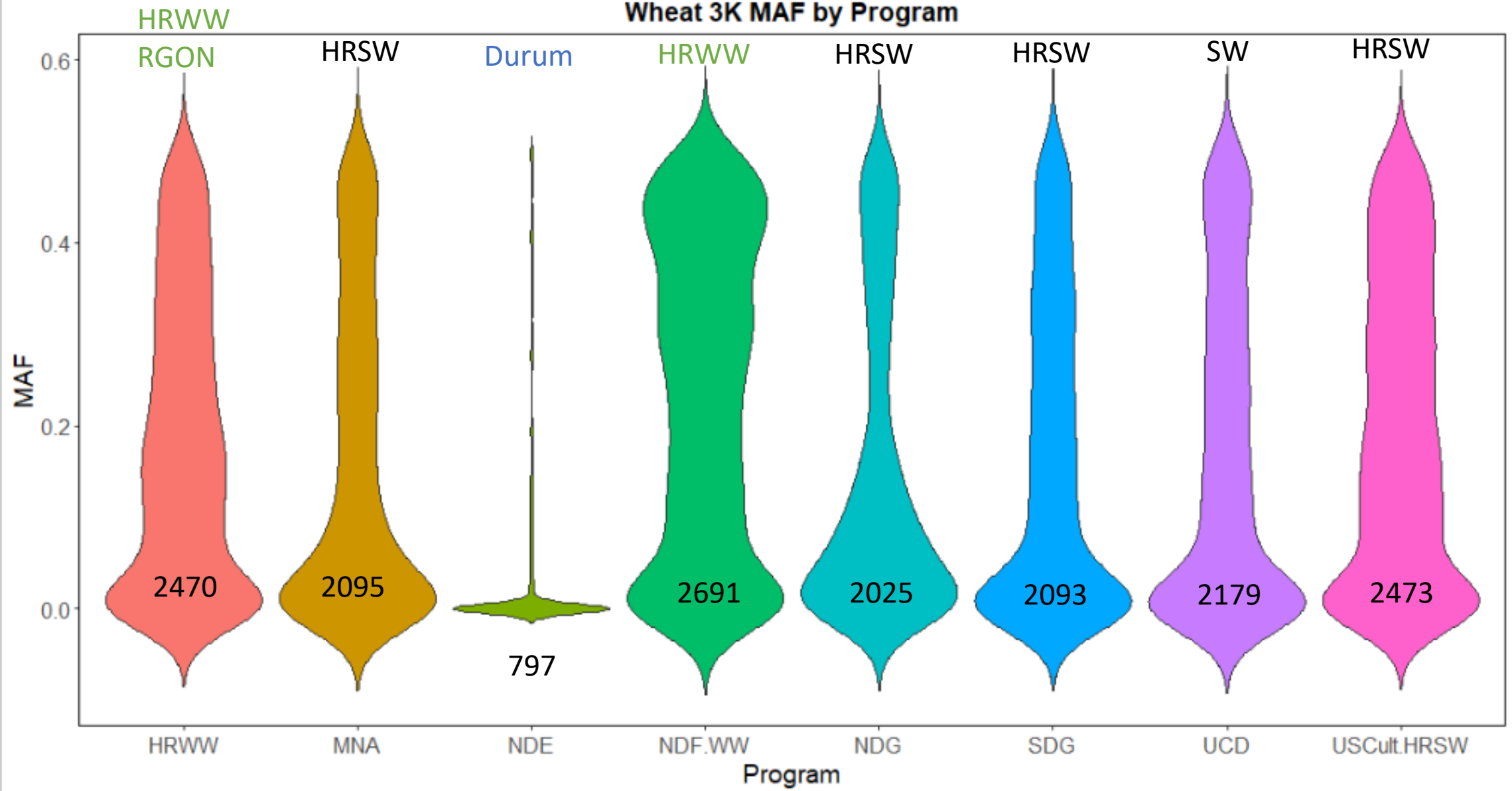273 KIMs

Wheat 3K Positions

Processing validation sets

- Barley: 2,273 | Oat: 3,071 | Wheat: 5,397
- Manually checked/manipulated SNPs to better represent the data
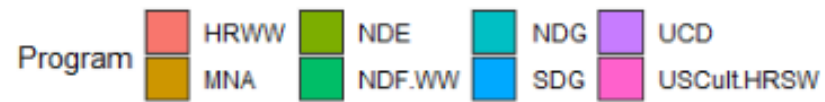 - Optimized positions used on all outgoing projects
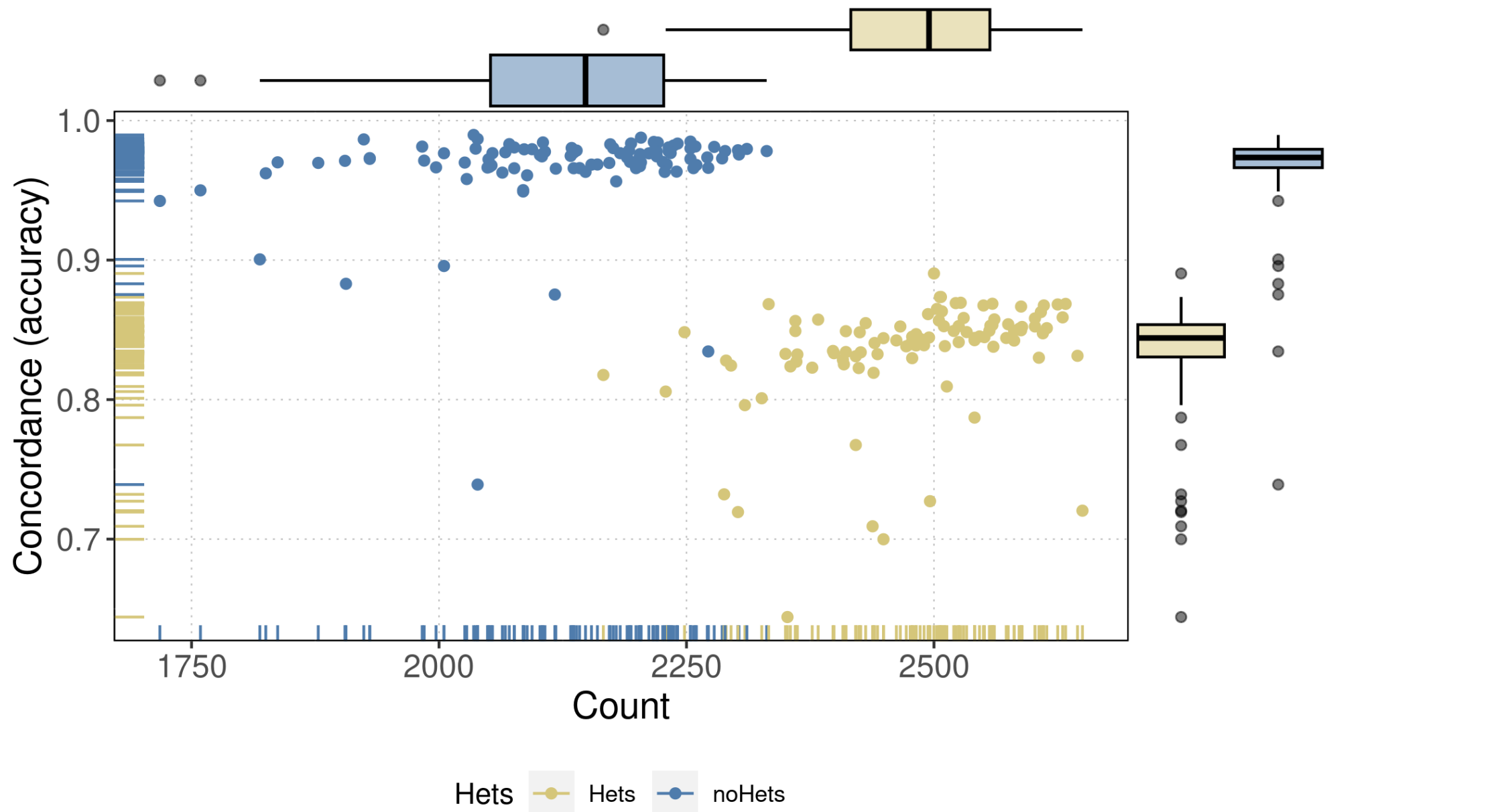- Conversion to VCF/Hapmap

Excellent Call Rates
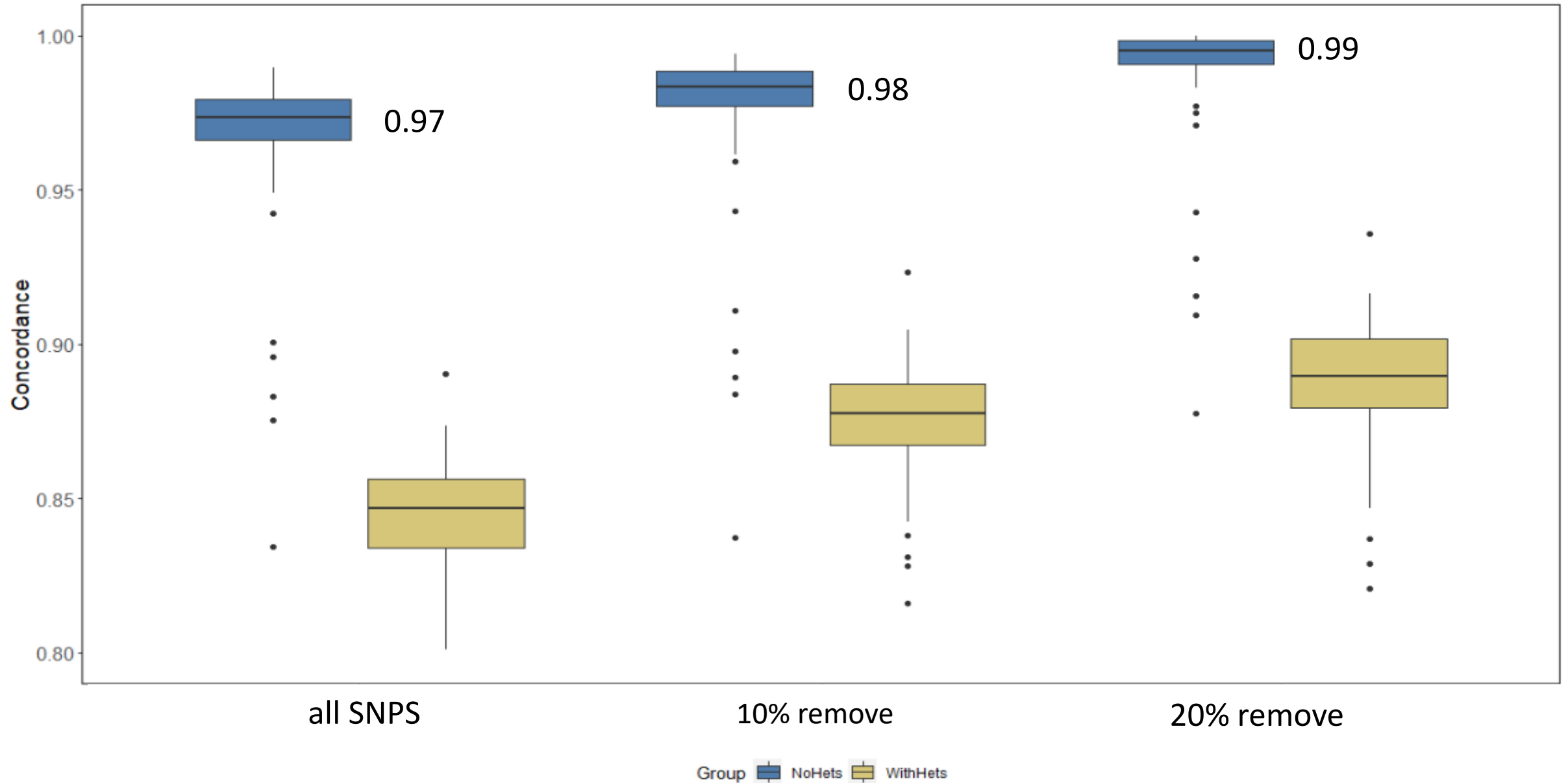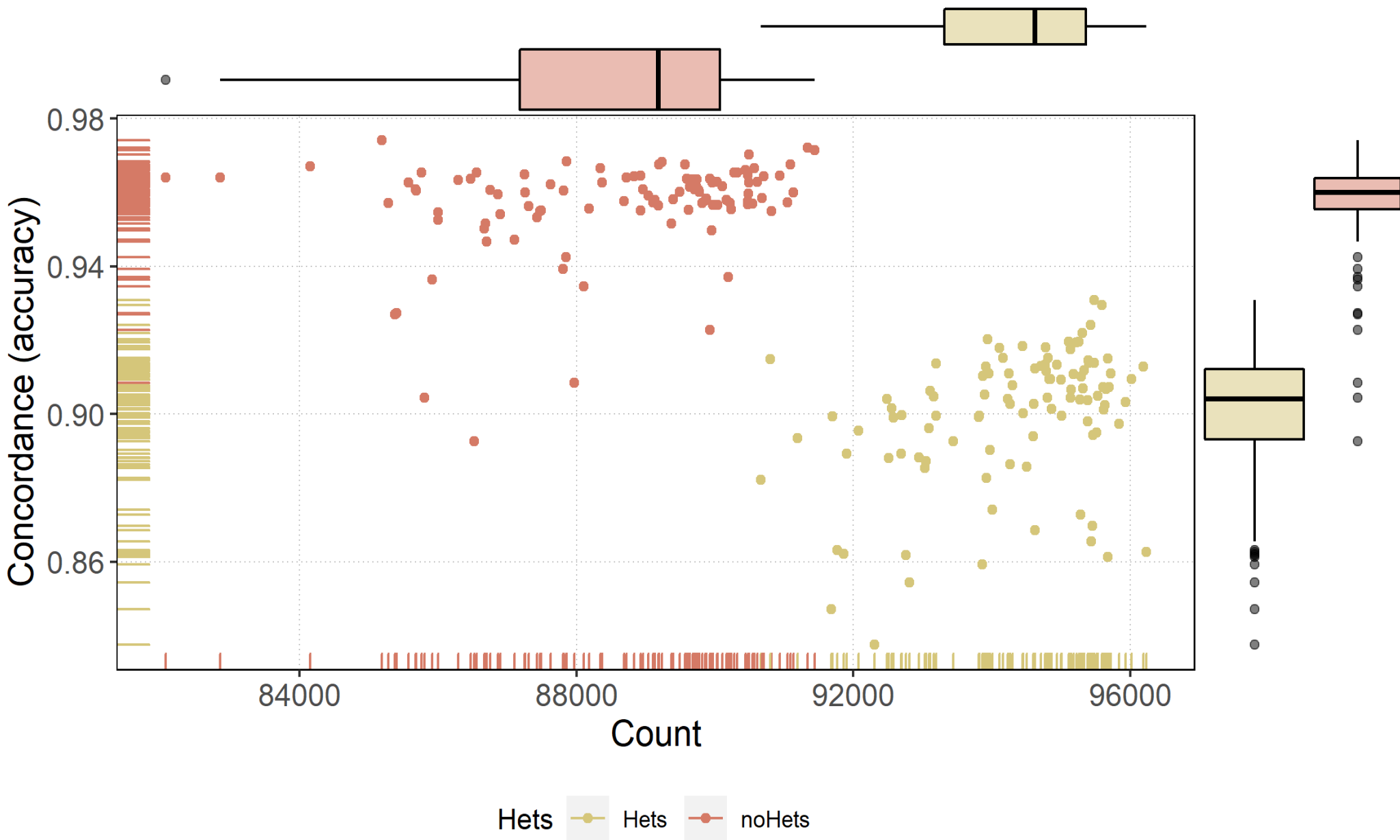
Wheat 3K MAF by Program

* SNPs maf >= 0.05

3K arrays SNPs match the original exome capture call

# Removing poor performing SNPs increases taxa accuracy

HRSW BEAGLE imputed exome capture taxa concordance

Average concordance of imputed wheat data sets

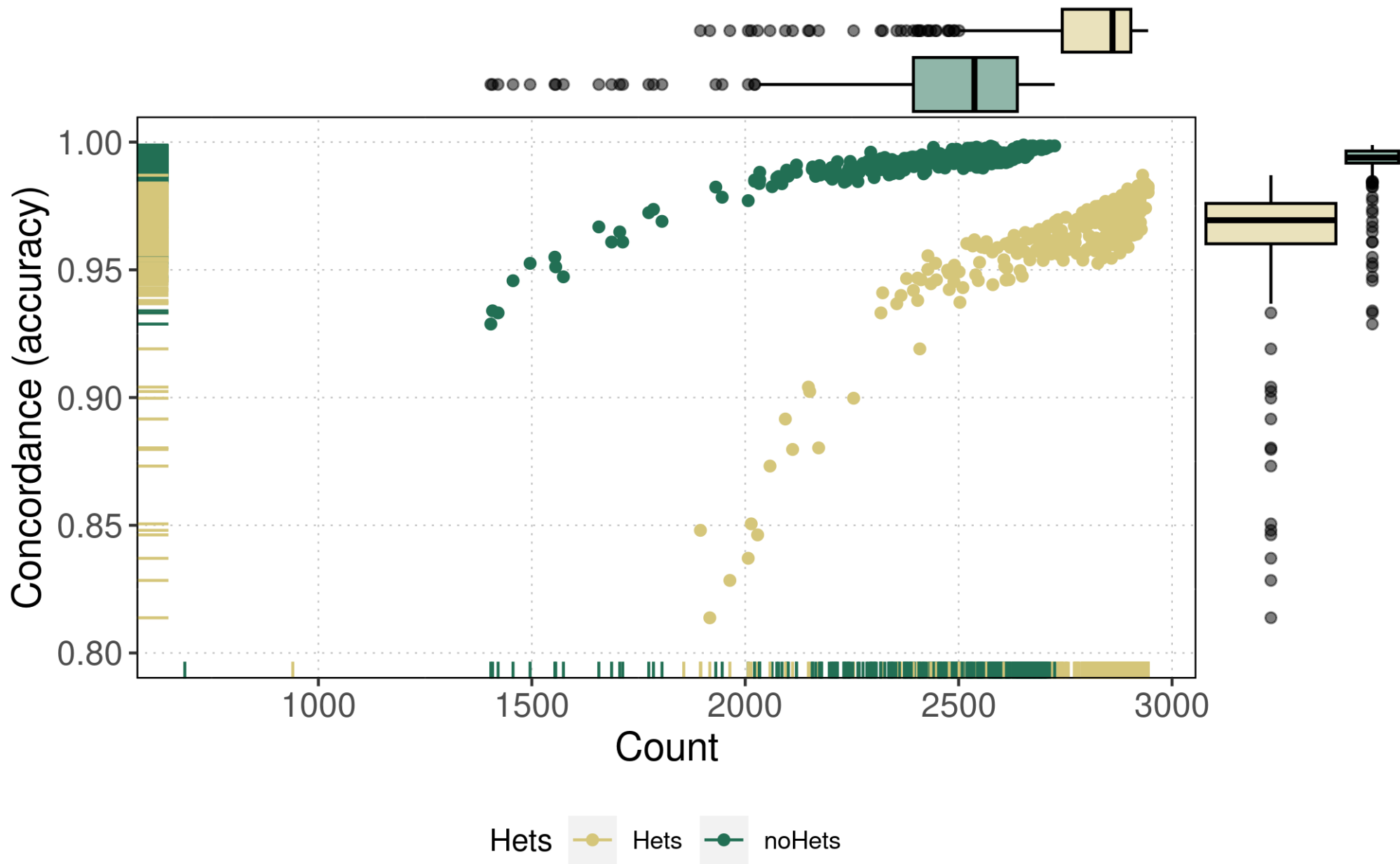| Pop	Imputed platform	Accuracy No hets	Count No hets	Accuracy With hets	Count With hets
HRSW cultivars/ breeding	exome capture (maf1)	0.957	88,569	0.900	94,293
HRSW cultivars/ breeding	exome capture (top 90%)	0.980	80,032	0.925	84,855
MSU breeding	GBS (maf1)	0.954	14,603	0.836	16,725
UMN breeding	GBS (maf1)	0.960	3,040	0.718	4,064
HRSW cultivars/ NDSU breeding	90K (maf1)	0.931	22,593	0.625	33,776
HRSW cultivars/ NDSU breeding	90K (top 80%)	0.981	18,976	0.695	27,017

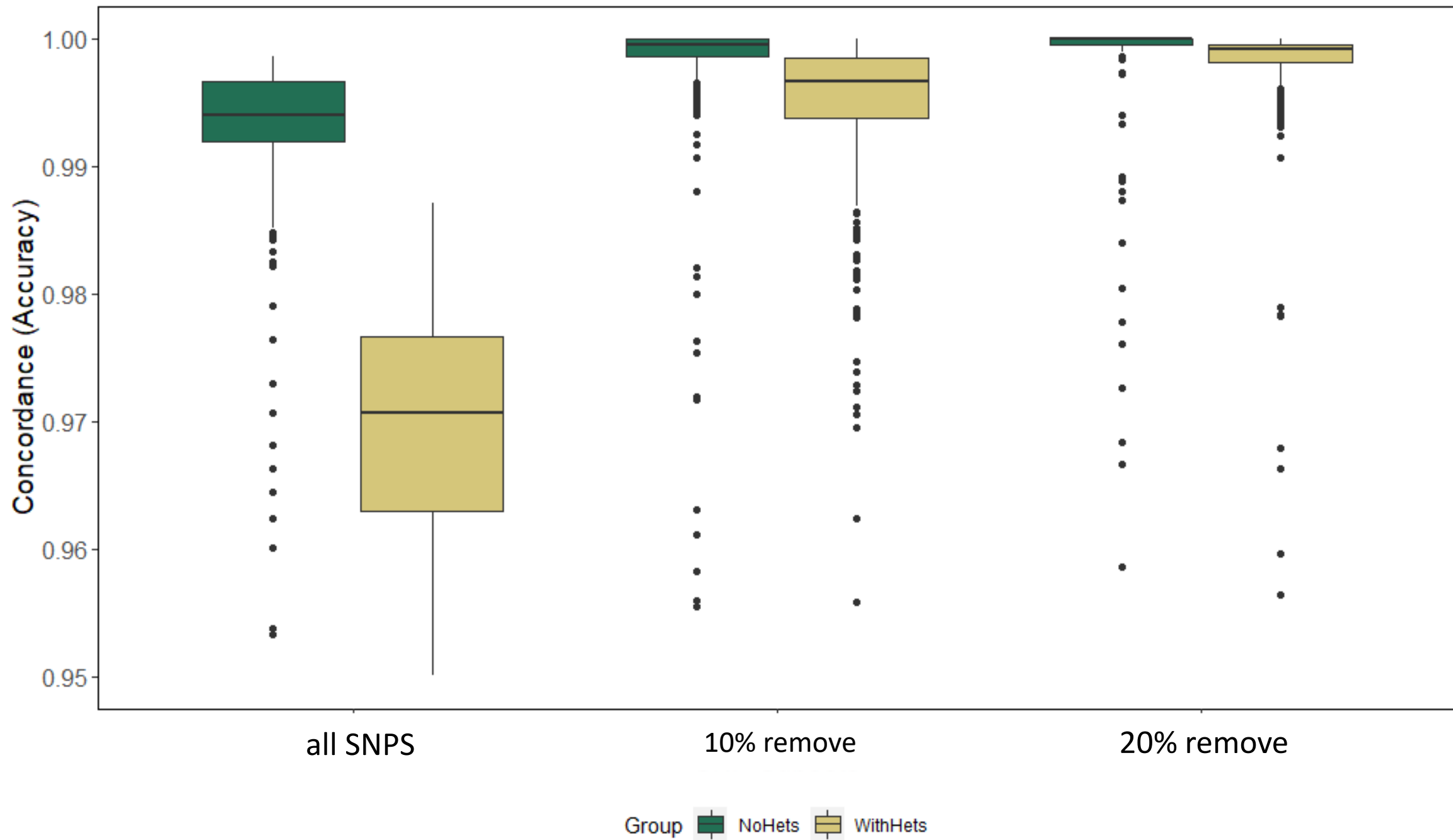# HRSW PHG imputation exome capture concordance (maf1)
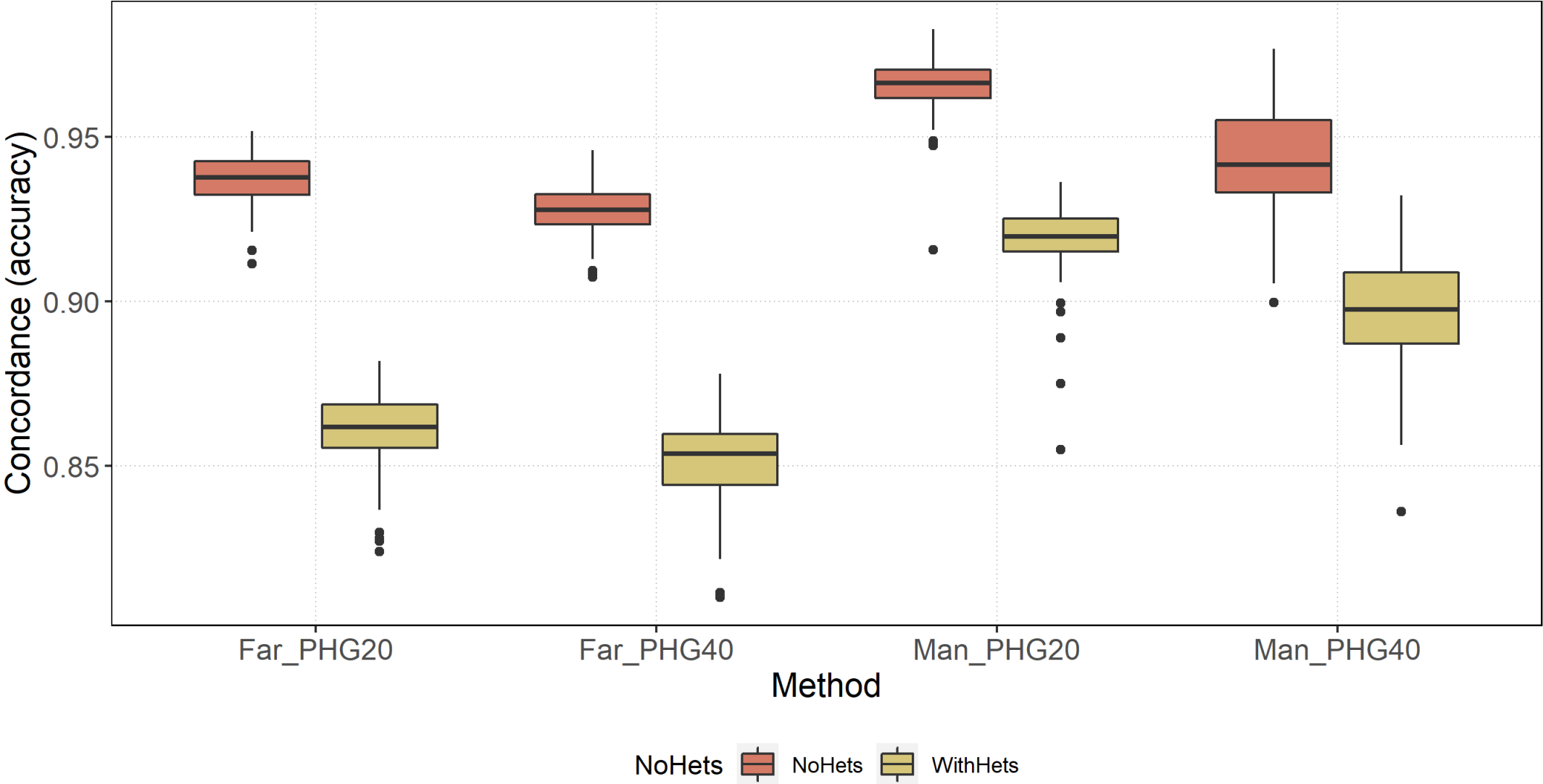
Dual hybridization works well - Barley
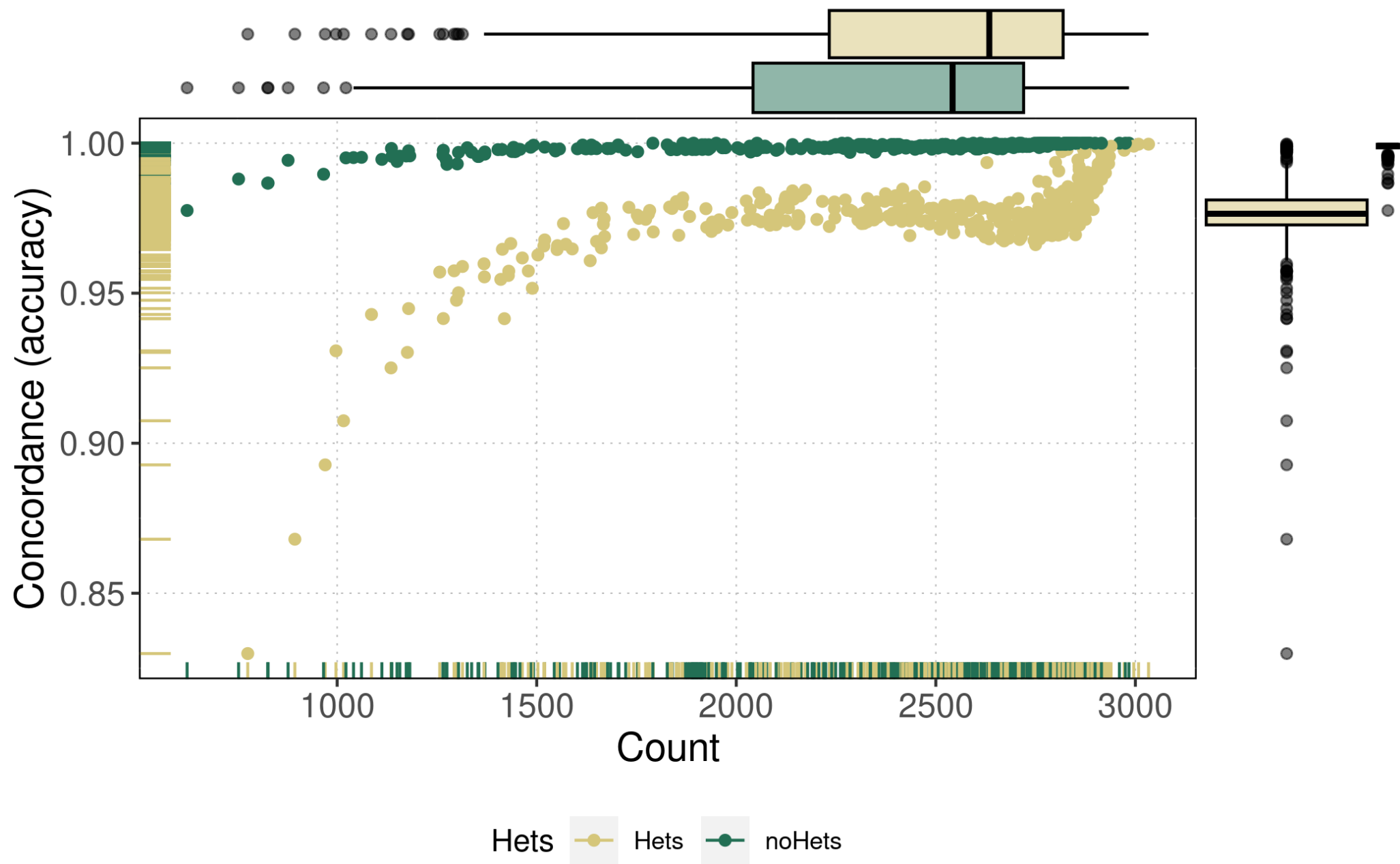
Oat concordance – missing data predicts accuracy

Concordance (accuracy)

Count

Hets: Hets | noHets

Oat Dual Hybridization Taxa Concordance 3K SNPs

Concordance (Accuracy)

all SNPS   10% remove   20% remove

Group ■ NoHets ■ WithHets

Seed Size Predictive Ability MSC vs GBS

Emily Conley & Jim Anderson, UMN
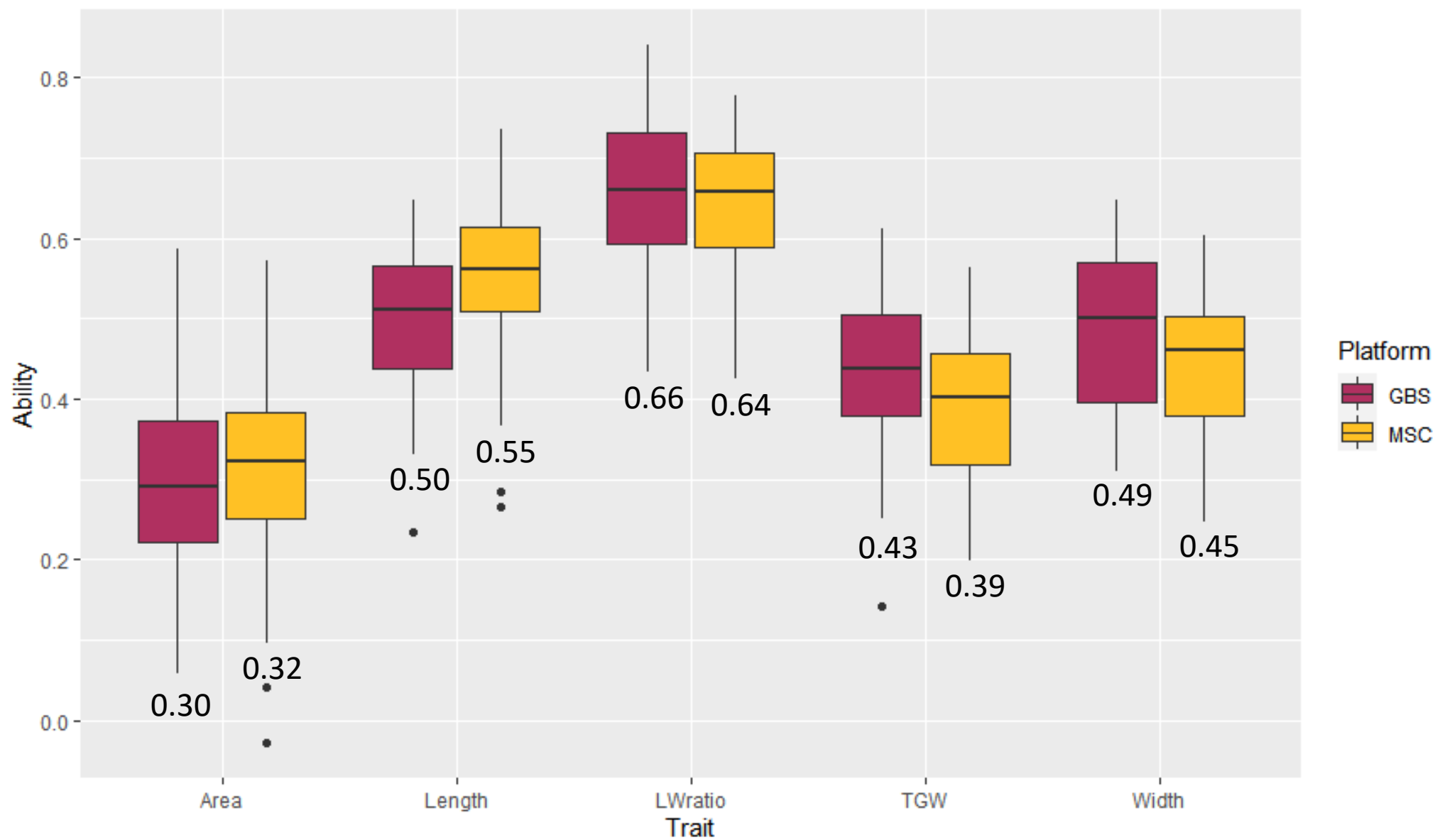
Conclusion

- The 3K array works quite well.
  - XT protocol preferred - 96 samples/chip, 1152 min order
  - HTS format 24 samples/chip, 288 min order
- Imputation can increase the # of SNPs and avoid re-genotyping all training germplasm.
- Dual-mode can immediately lower the price to $7.
  - Verifying that dual mode works with wheat
  - $4.67 in tri-mode
- GS models are similar to higher-density counterparts.
  - Oat and barley as well

- Poster 481 – Monday @3PM
- IWGSC Webinar Jan 26th

Low marker redundancy suggested by LD analysis
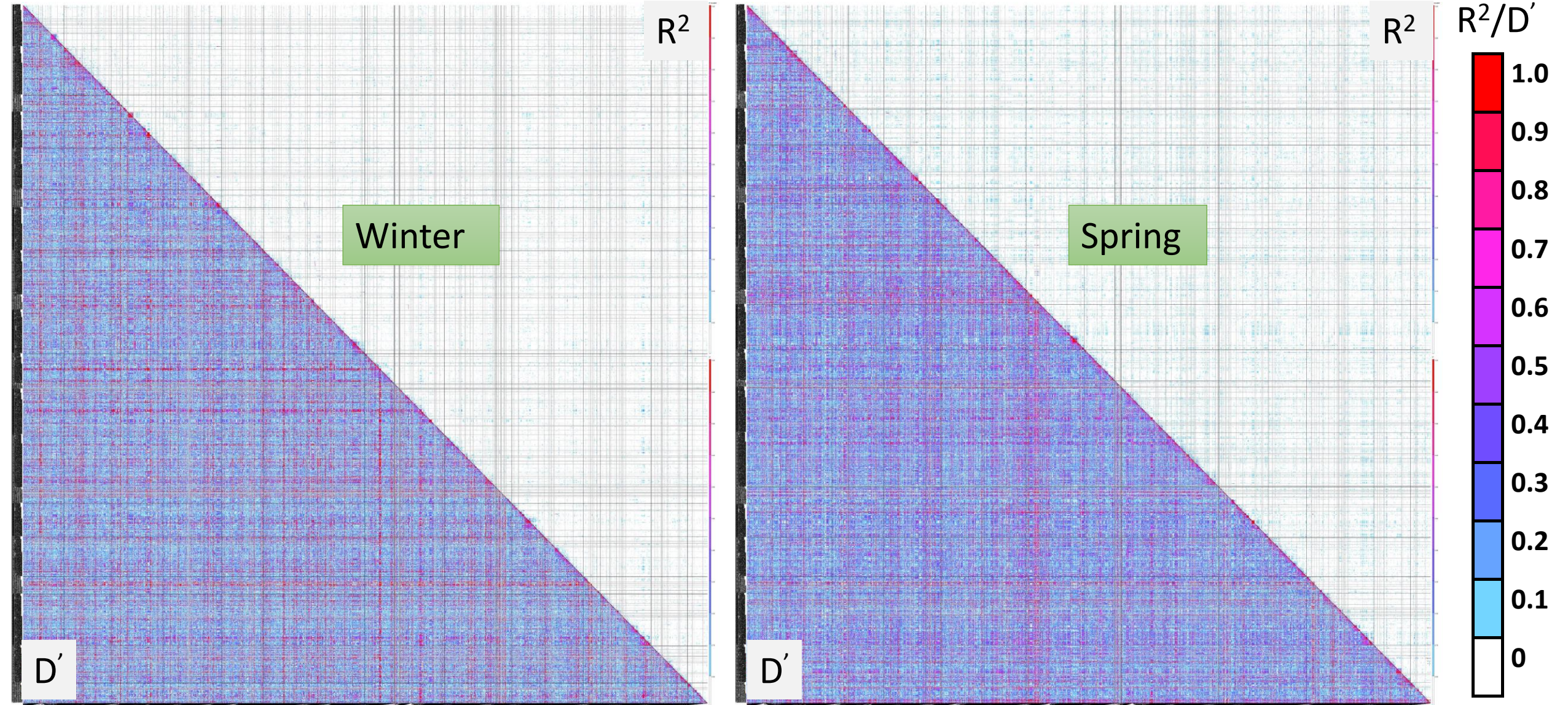
LD distributions – HRSW subgenome