

Imputation with shellfish genomes : pitfalls

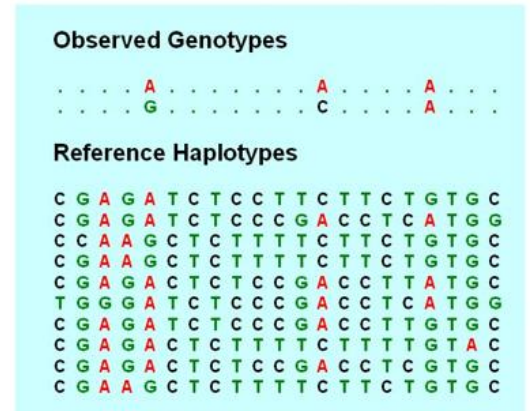
VIVALDI Final Conference
27 November 2019
Brest, France

Binyam Dagnachew & Luqman Aslam



Why WG imputation?

- Use of WG sequence data is expected to increase identification of QTL
 - Expensive to sequence
 - Imputation may help to



Individual data

- Sequenced 67 (10 GP, 20 P & 37 F1)
- Genotyped 1,530 F2

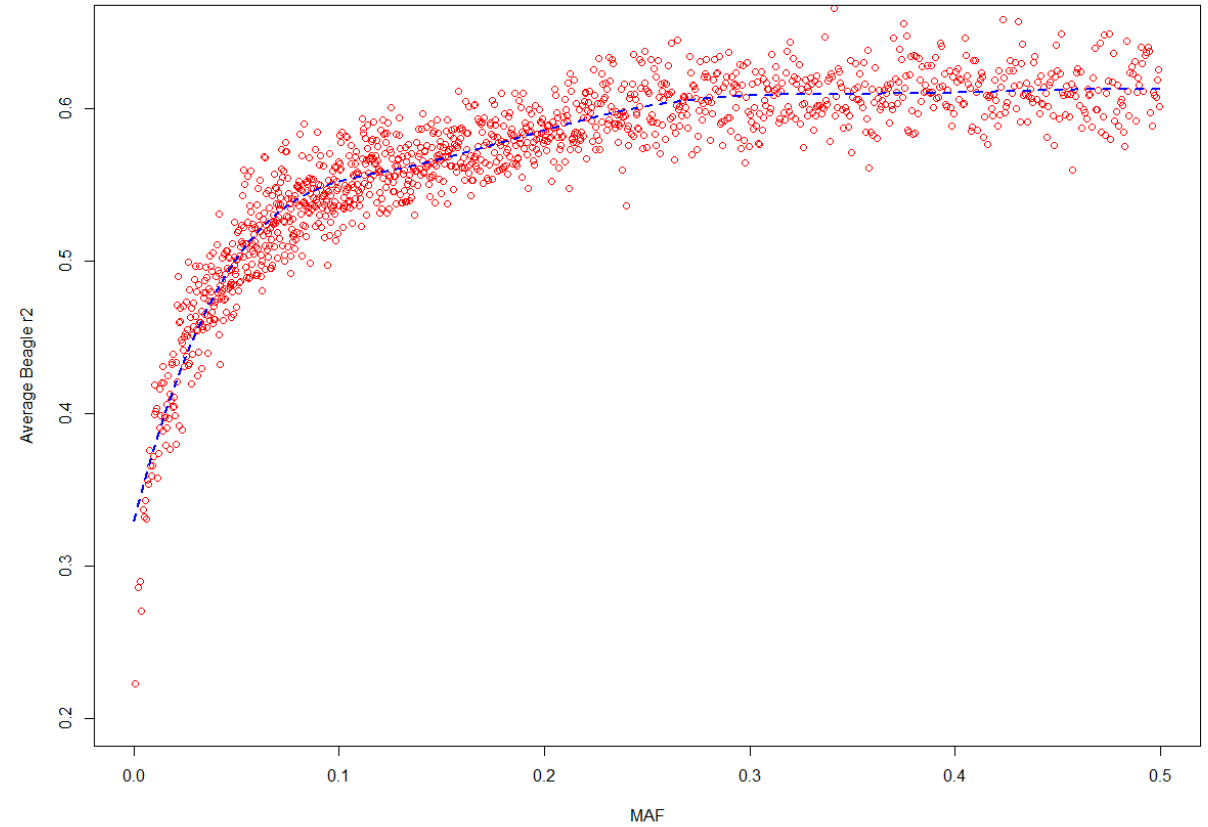
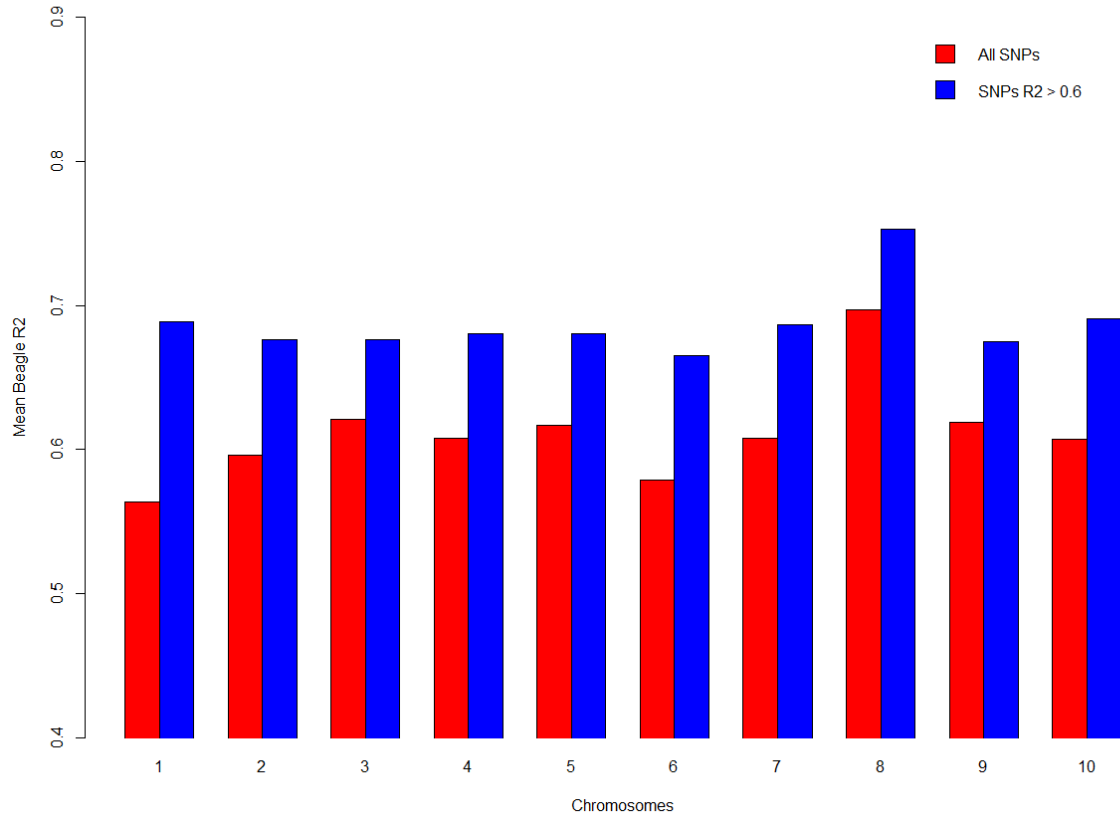
Genomic data

- WG Sequence data (~365 k)
- Array (~14k)

Software

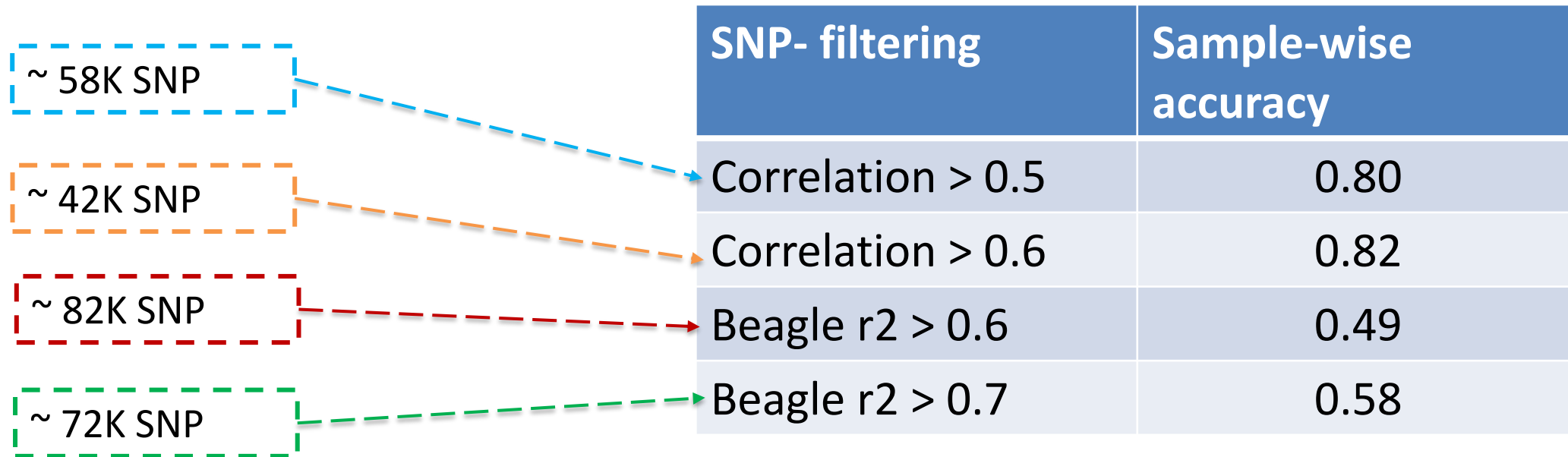
- BEAGLE 5.0
- Alphaimpute 2.0

SNP-wise accuracy



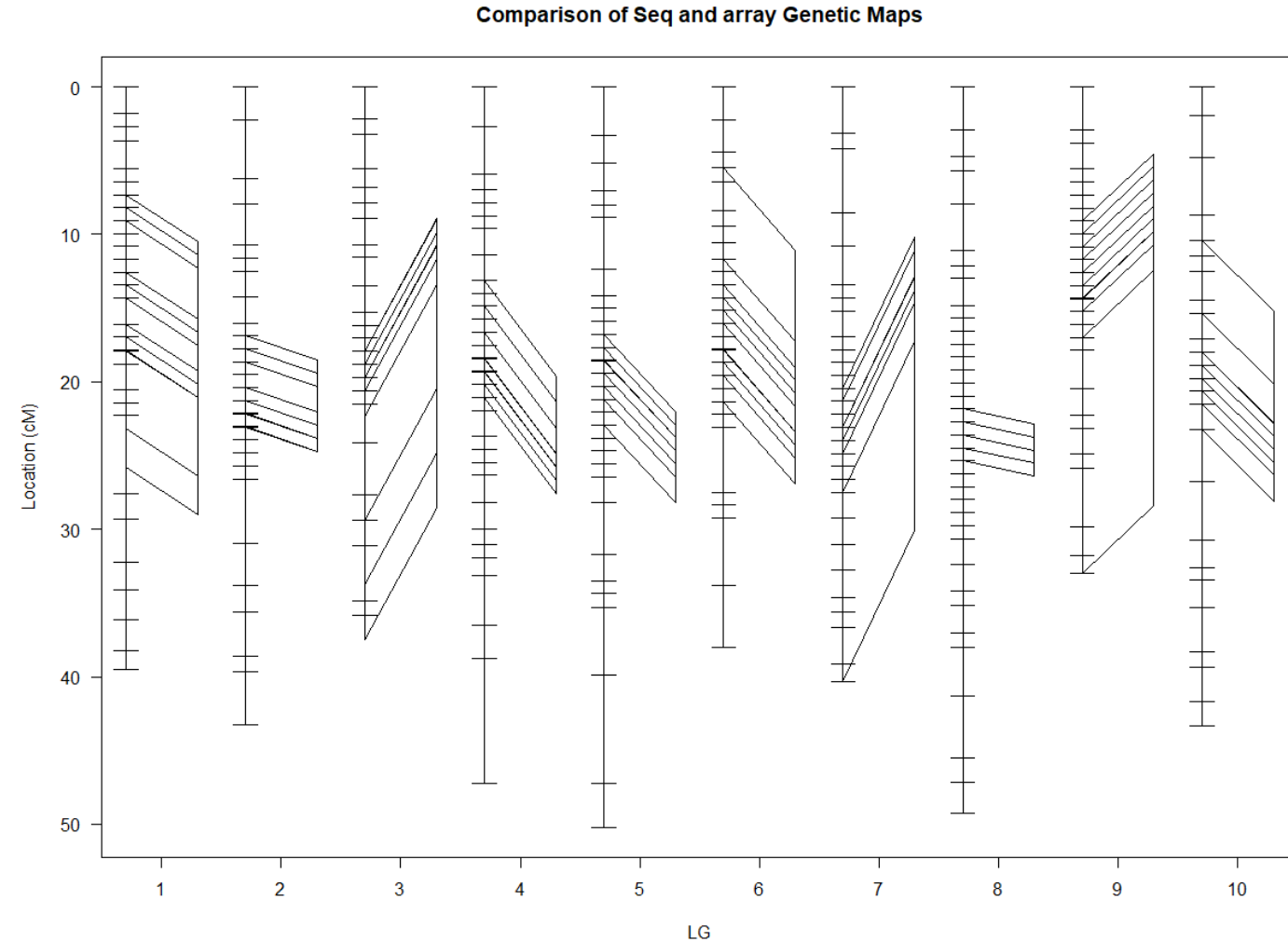
Sample-wise accuracy

	SNP- filtering	Sample-wise accuracy
~ 58K SNP	Correlation > 0.5	0.80
~ 42K SNP	Correlation > 0.6	0.82
~ 82K SNP	Beagle r2 > 0.6	0.49
~ 72K SNP	Beagle r2 > 0.7	0.58



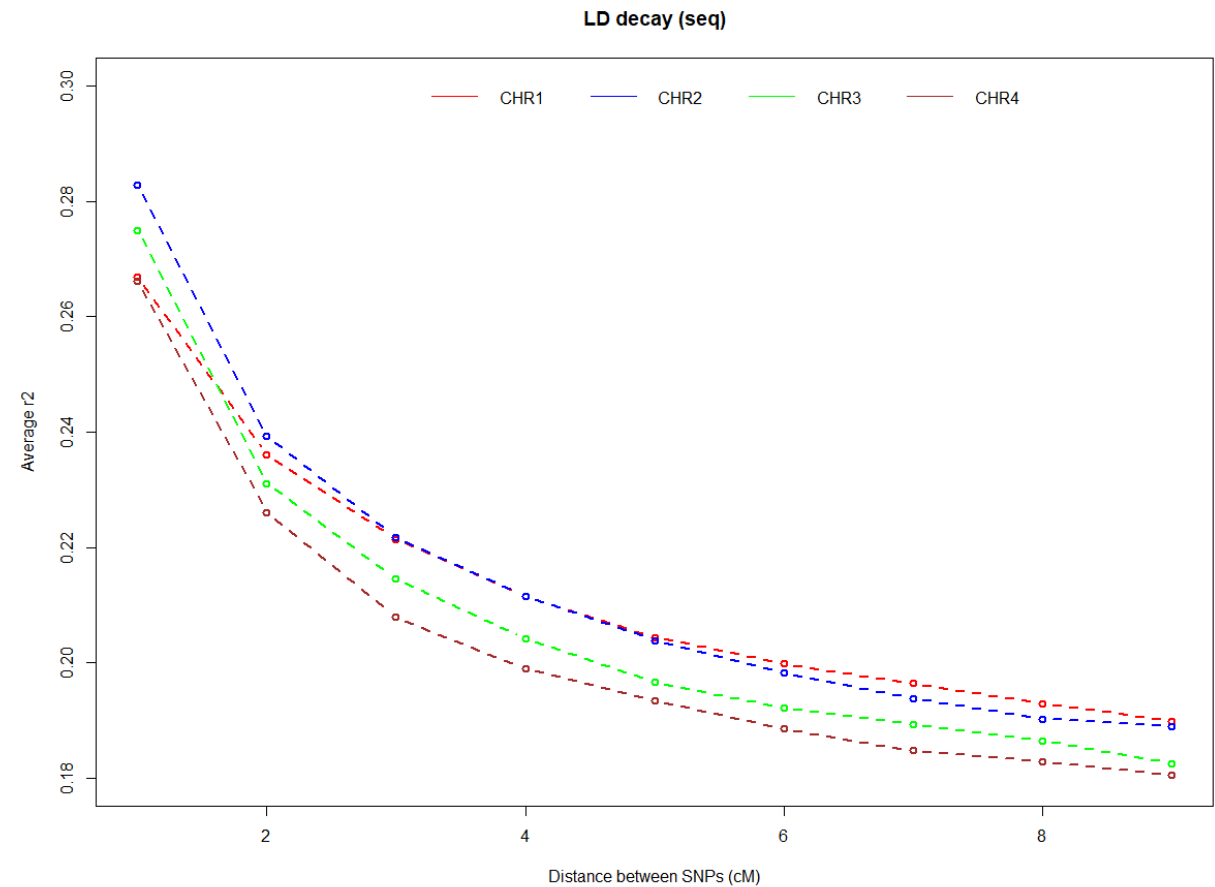
Comparison of genetic map for seq and array data

- Array did **not** cover the whole genomic region

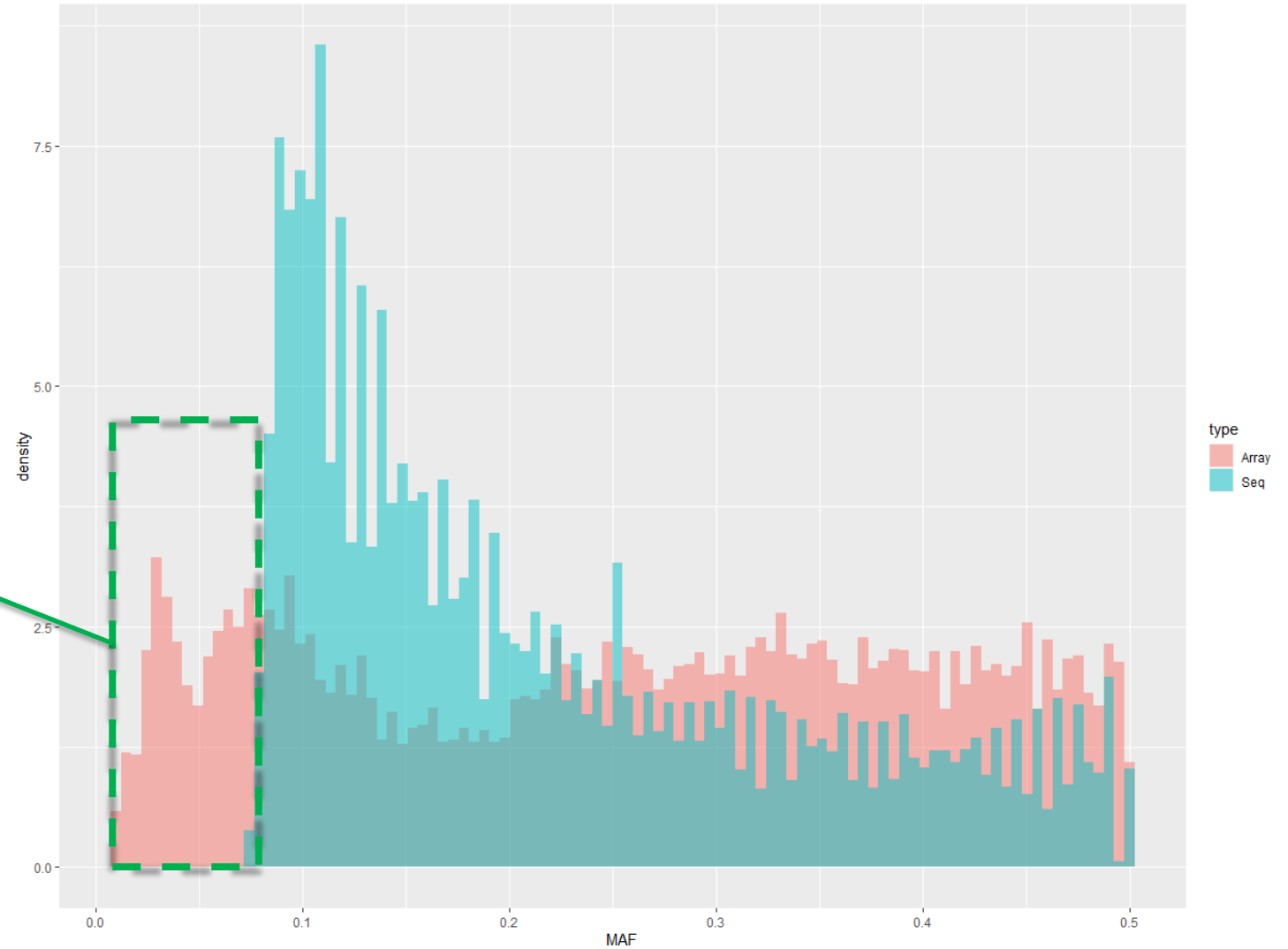
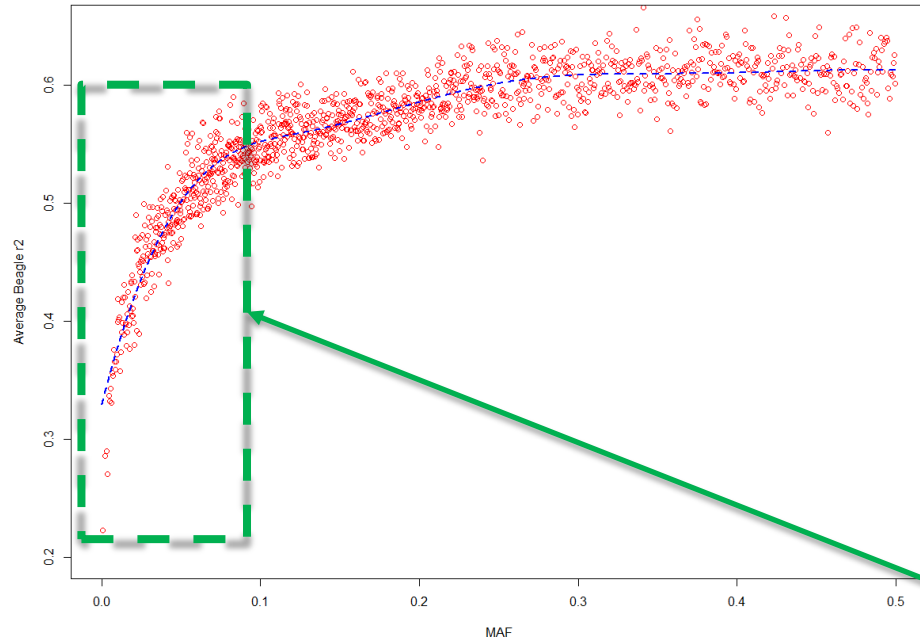


Linkage Disequilibrium (LD)

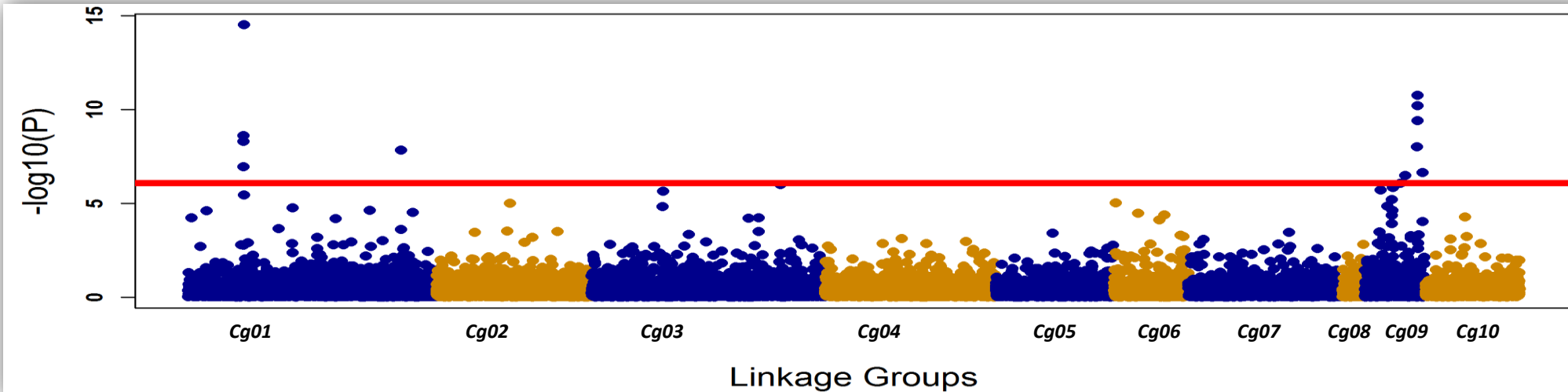
- The LD decaying
 - Low number of shared haplotypes
- Low pairwise LD R^2 values across the Genome
 - Recombination
 - Mapping errors



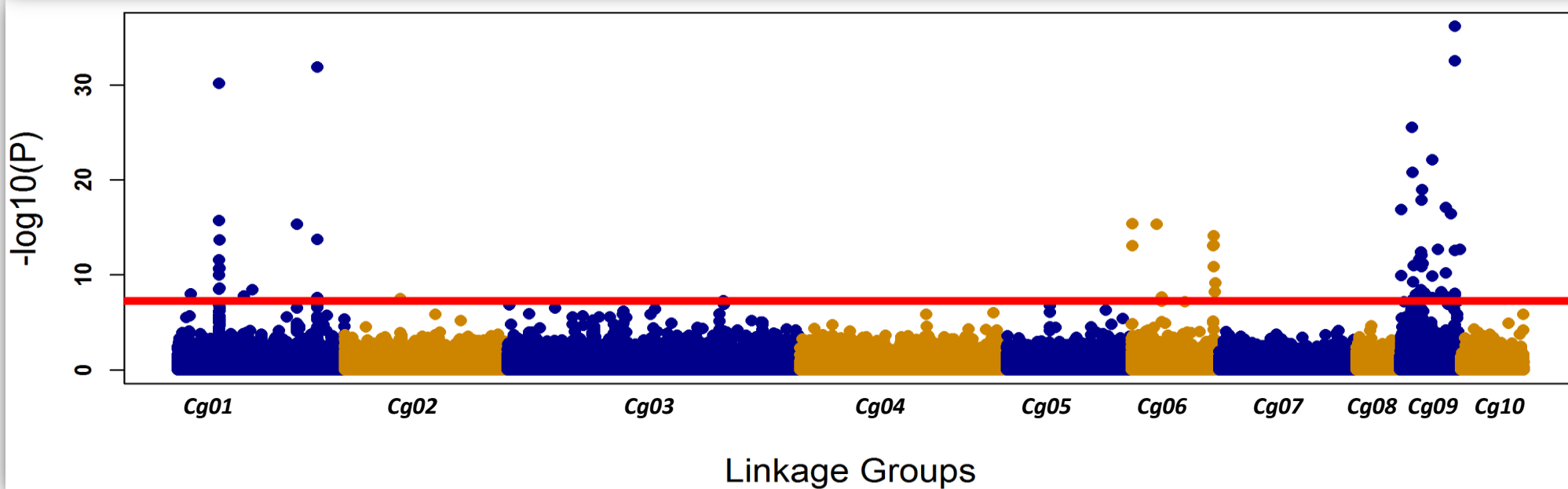
MAF of the array



Manhattan Plot for OsHV-1



Array data



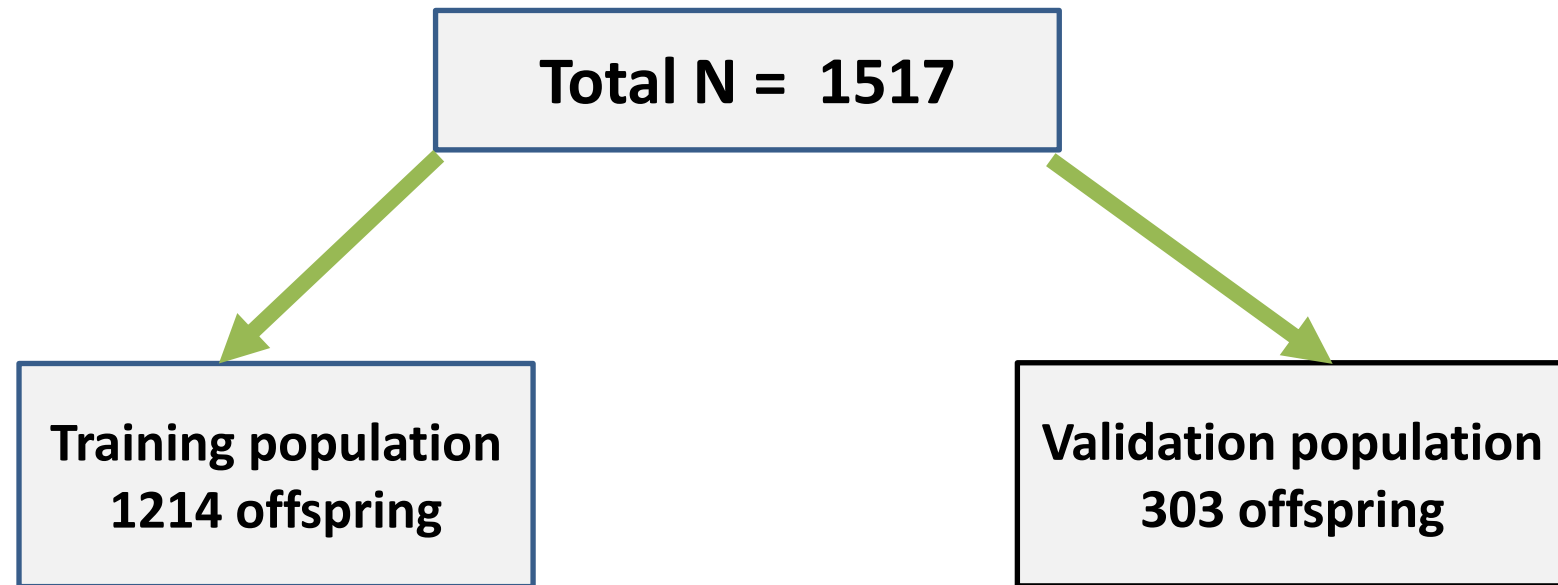
Imputed data

Genomic model and prediction accuracy for

OsHV-1

$$y = \mu + Zu + e$$

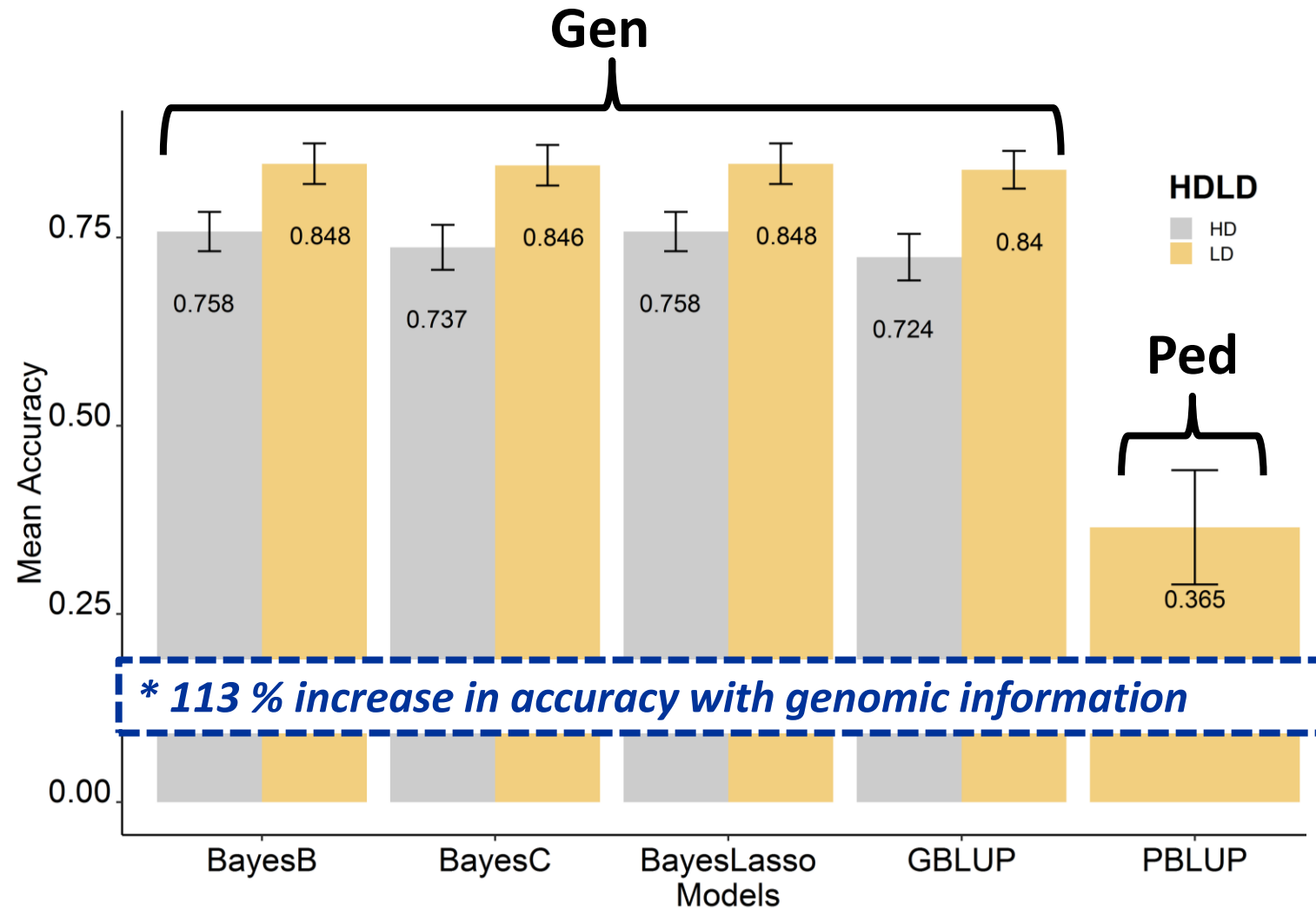
- Random masking of ~ 20% of the offspring in the population
- Repeat the random sampling 20 times



- Correlation between estimated breeding value (PEBV/GEBV) and phenotype

$$r_{corr} = \frac{\rho(G[P]EBV, y)}{\text{sqrt}(h^2)}$$

Genomic prediction accuracy for OsHV-1

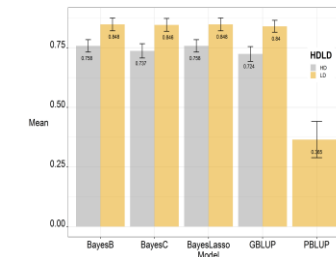
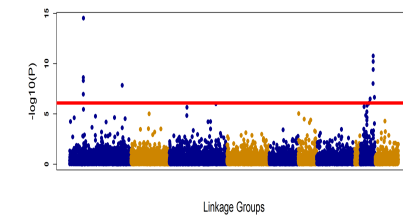
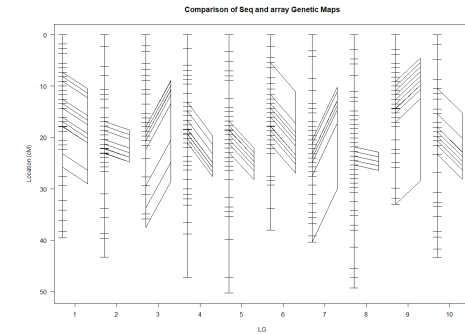


Summing-up

- ❑ Poor imputation accuracy observed
 - Genome assembly (improve??)
 - Optimal selection of SNPs for the array
 - Low pairwise LD and faster decay

- ❑ Amplified a putative QTL region resistance against OsHV-1
 - ❑ High resolution

- ❑ Genomic prediction using genotypic information generated 113% improvement in accuracy of prediction
 - Imputation did not improved prediction accuracy



Thank you for the attention



This project has received funding from
the European Union's Horizon 2020
Research and innovation programme
under grant agreement N° 678589



Luqman Aslam

CONTACT

Binyam.dagnahcew@nofima.no
Luqman.aslam@nofima.no