

Towards a flexible definition of Core Sets based on the haplotype diversity of German sheep

J. Geibel^{1,2}, C. Reimer^{1,2}, A. Weigend¹, Y. Shakya³,
H. Melbaum⁴, K. Gerdes⁵ and S. Weigend^{1,2}

¹ Friedrich-Loeffler-Institut, Institute of Farm Animal Genetics

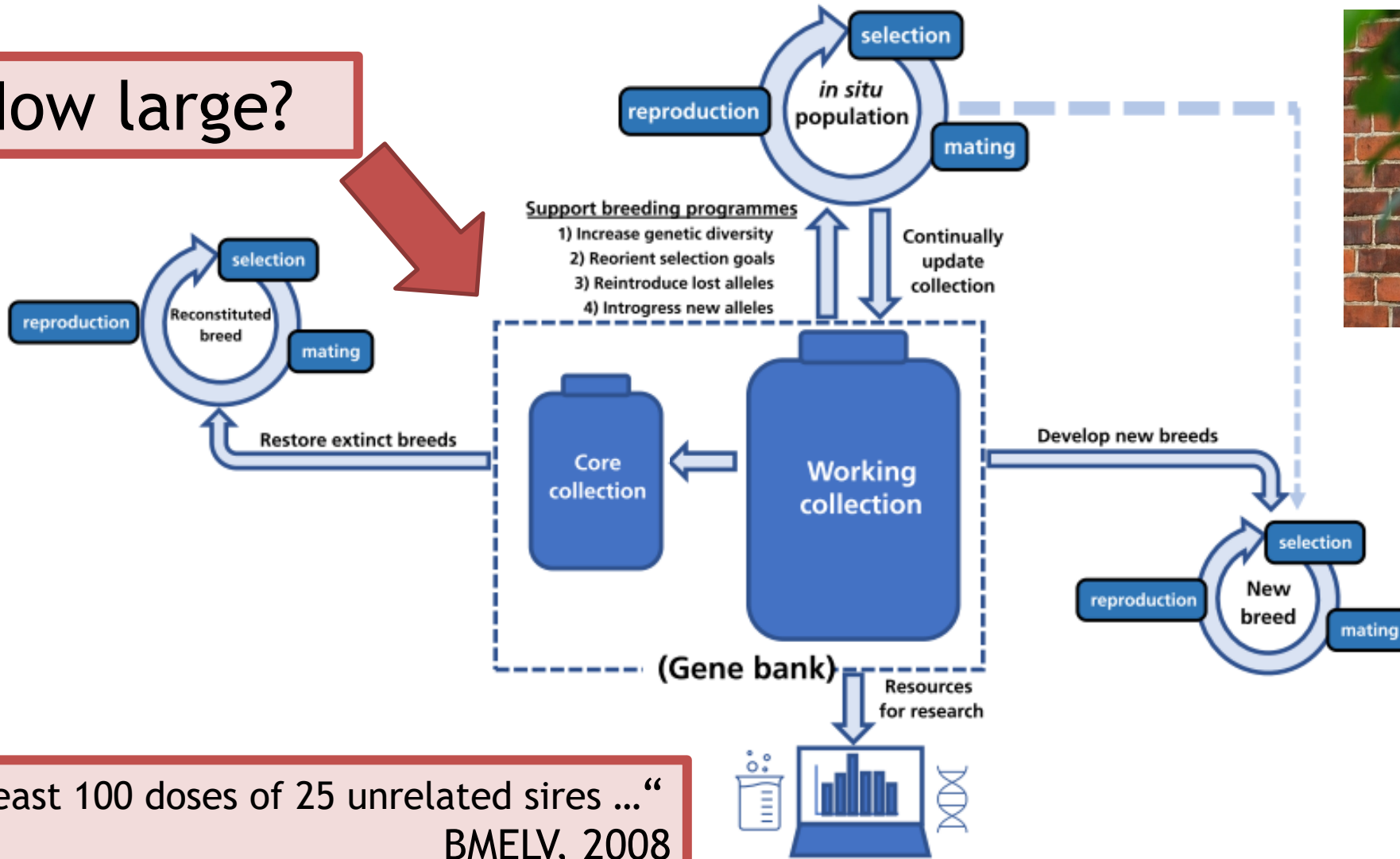
² University of Goettingen, Center for Integrated Breeding Research

³ University of Goettingen, European Master of Animal Breeding and Genetics

⁴ Landes-Schafzuchtverband Weser-Ems e.V.

⁵ Chamber of Agriculture in Lower Saxony

How large?



„... at least 100 doses of 25 unrelated sires ...“
BMELV, 2008

Dachin-Burge *et al.*, 2023

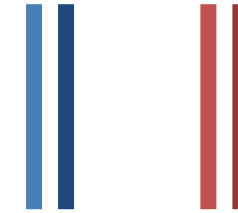
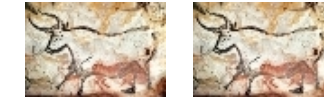
How much „genome of a breed“ do we want to store?

- „Sequential Safe Set“ Analysis (Eding *et al.*, 2002):
 - Maximization of **average** diversity by minimization of mean kinship
 - Condition: N available places

→ How much material represents a breed?

→ Are there differences between breeds?

- By now no easy-to-interpret measures
- Idea: Share of haplotypes of a population



$$\frac{8}{12} = 66\% \rightarrow \frac{10}{12} = 83\%$$

Sheep data

- Genotyping of gene bank and field samples
- Affymetrix IMAGE001 Multispecies Array
~12.000 SNPs for sheep
- Pedigrees (Ovicap)



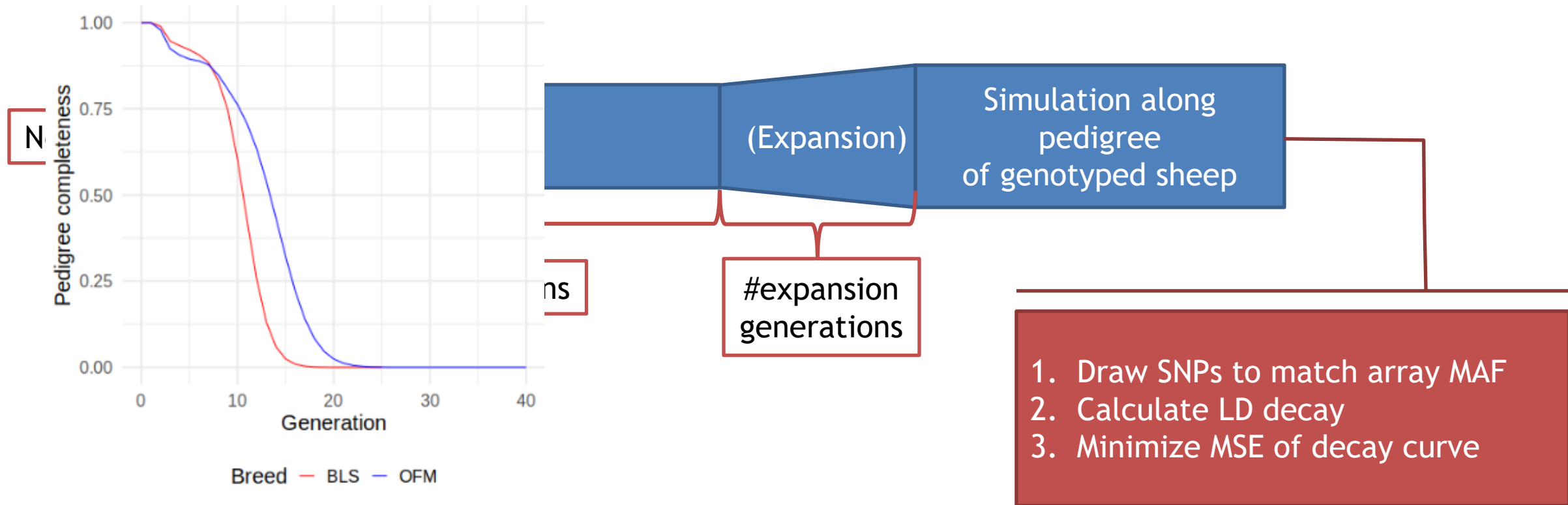
	Ostfriesisches Milchschaaf (OFM)	Bentheimer Landschaaf (BLS)
Gene bank	34	30
Field samples	20	20
Pedigree (2022)	3850	3232

Simulation set up

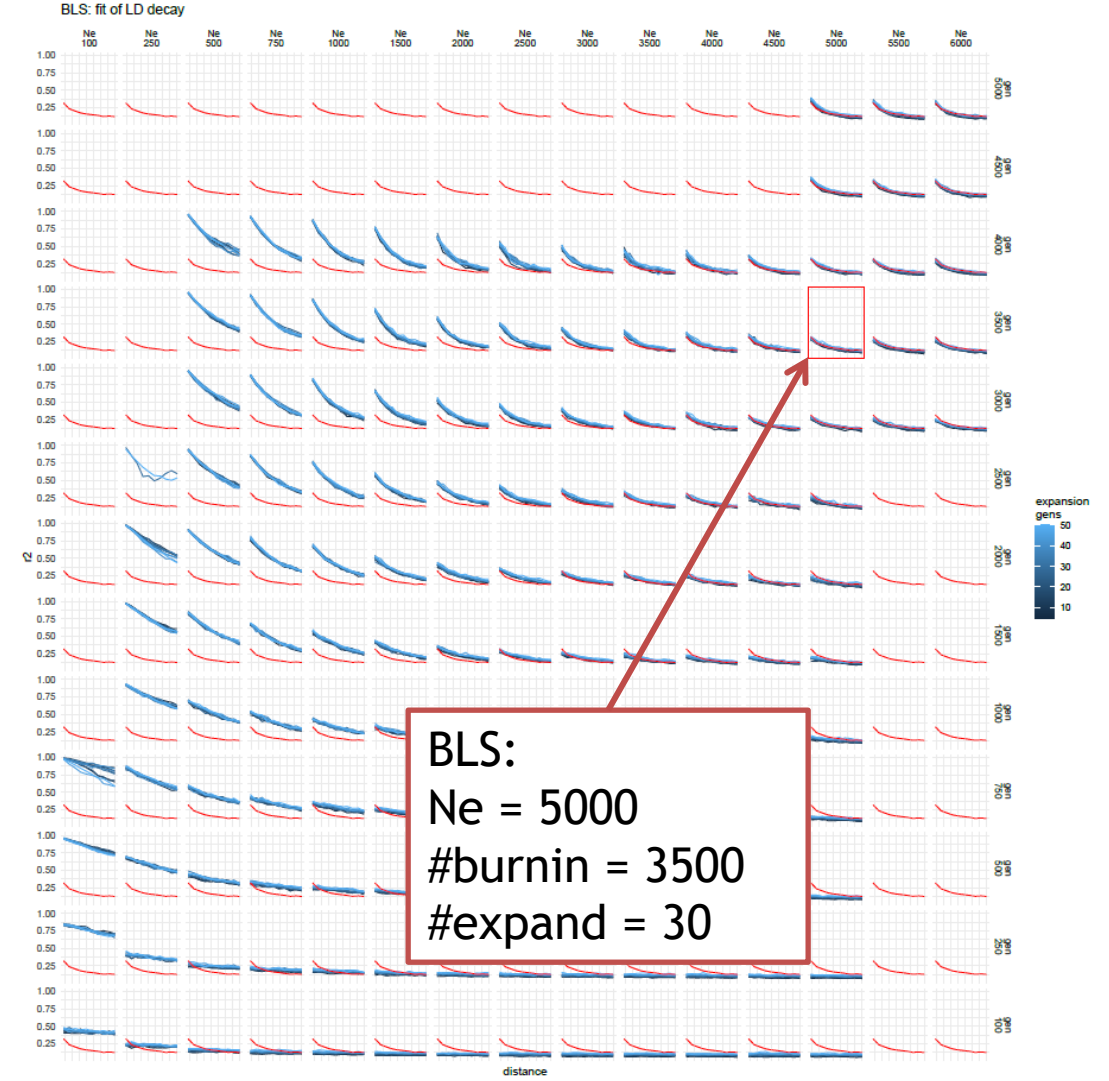
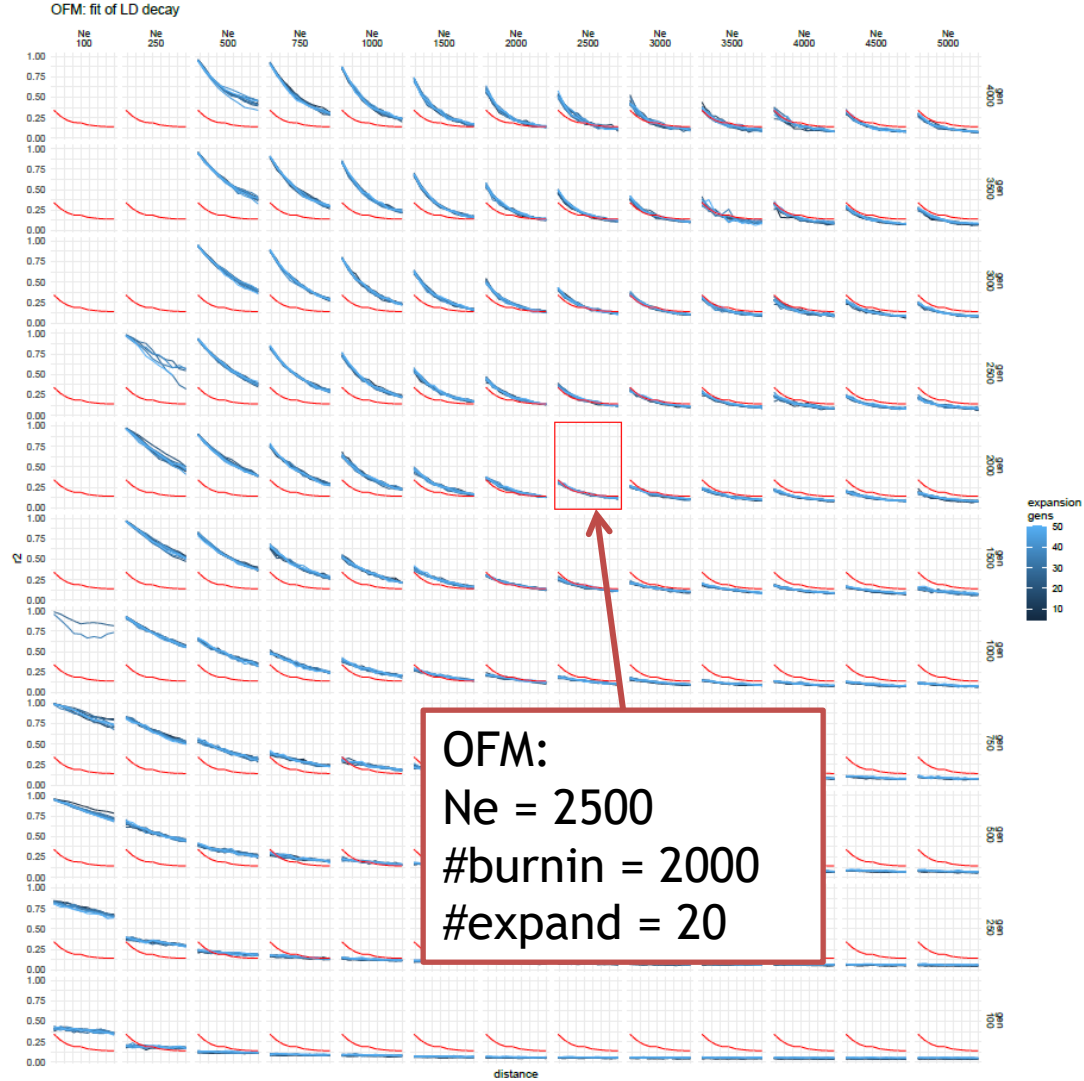


Molecular Breeding
Program Simulator
(MoBPS; Pook et al. 2020)

- Background Variation:
 - Chromosome 10 and 20
 - Beta-distributed allele frequency spectrum, WGS density



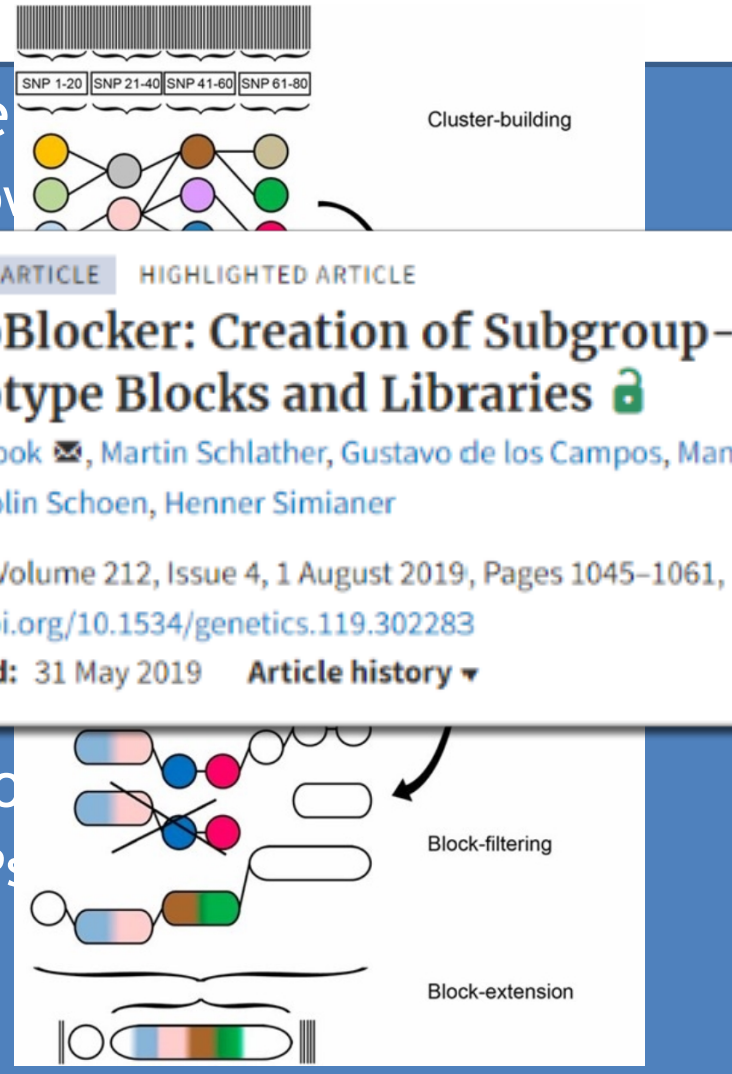
Simulation optimization



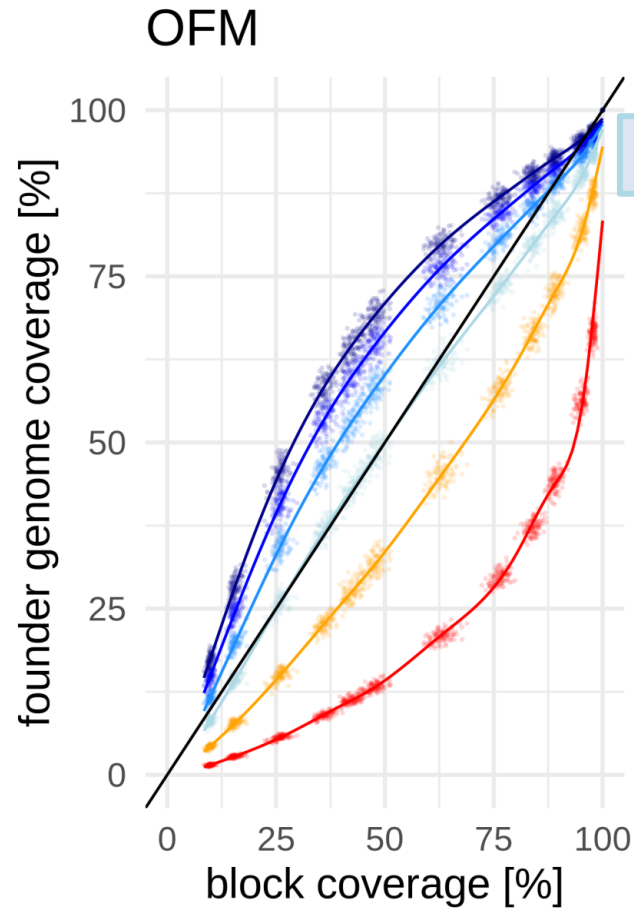
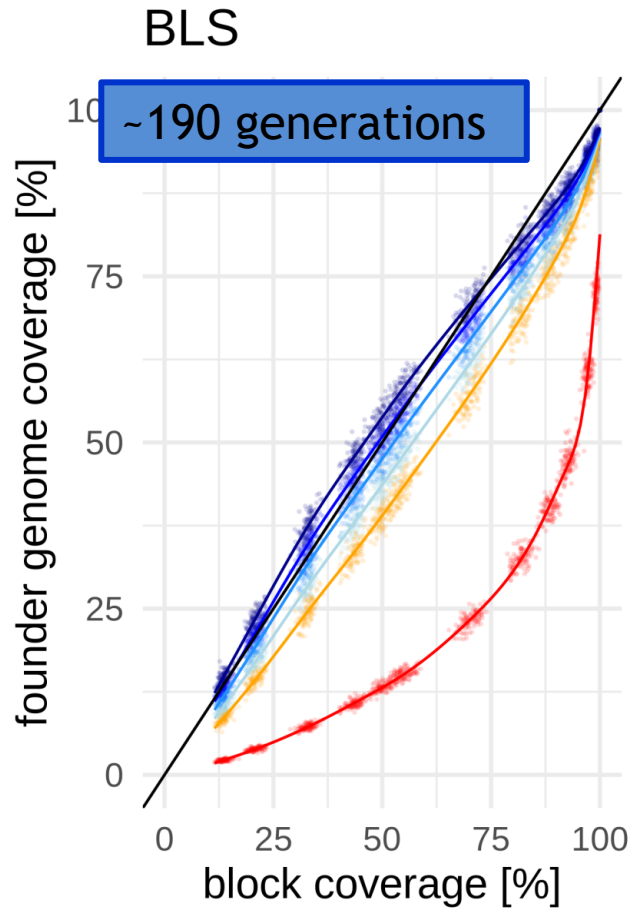
Identification of haplotypes using HaploBlocker

- Parameters influence

- Longer starting window
→ longer and larger blocks
 - „target_coverage“
covered in block
→ shorter and smaller blocks
 - Influence of SNP density
with larger starting window
- Currently no filter for blocks
- Stick to setting for no filter
- window_size = 10 SNPs
 - Target_coverage = 1
 - SNP density = WGS



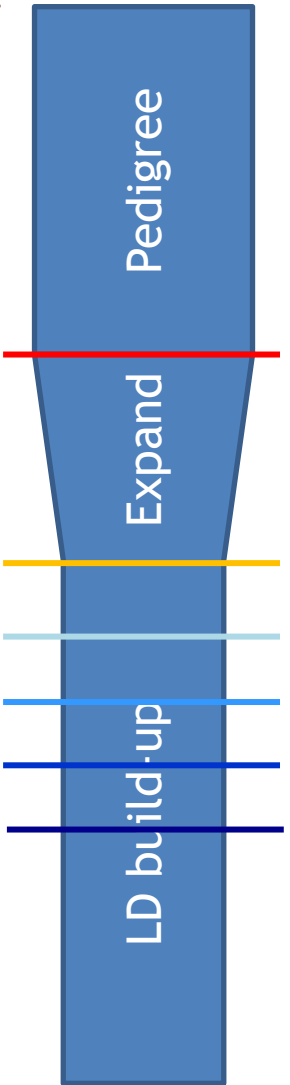
Haplo blocks vs. founder genome content



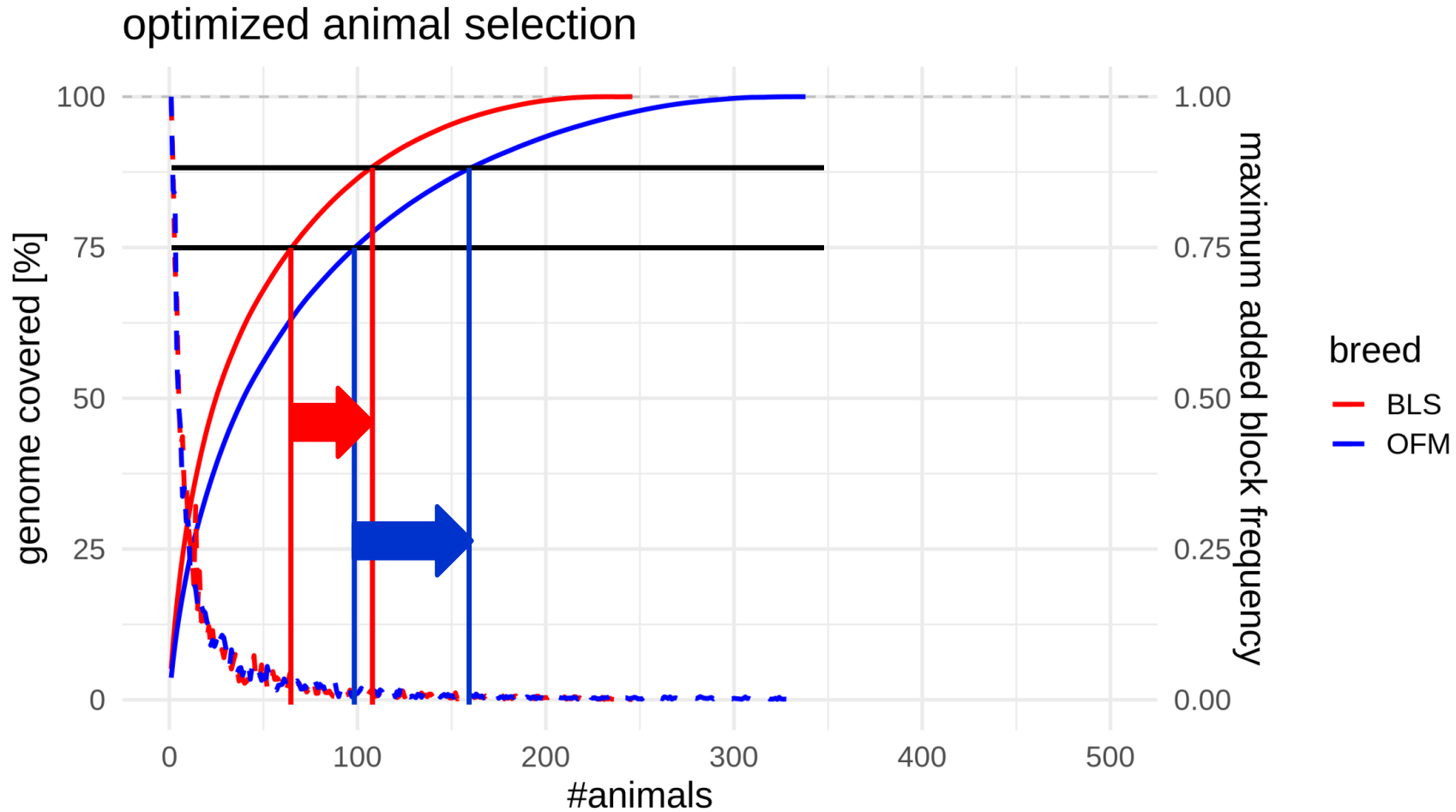
basis

- founder_genome_founder_pedigree
- founder_genome_founder_expand
- founder_genome_burnin_50
- founder_genome_burnin_100
- founder_genome_burnin_150
- founder_genome_burnin_200

Blocks →

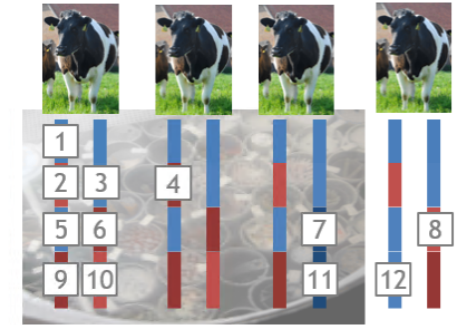


Genome coverage with haplo blocks

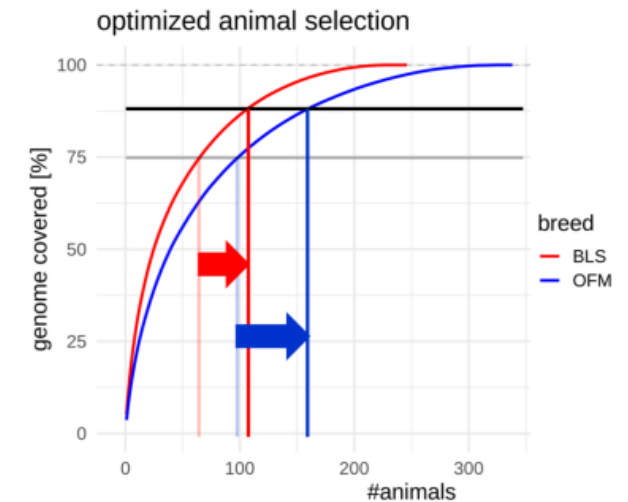


Conclusions

- Usage of haplo blocks to generate an easy-to-interpret and flexible goal for the size of a core reserve
- Shows that a strict size of a core reserve does not deliver appropriate results for all populations
- Best setup of block structure is still under evaluation
- Allows to add further restrictions like minimum block length, or block frequency



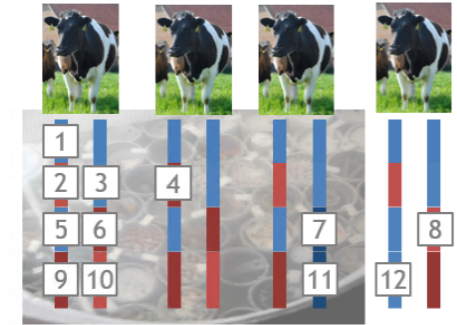
$$\frac{8}{12} = 66\% \rightarrow \frac{10}{12} = 83\%$$



Thank you for your attention!

Conclusions

- Usage of haplo blocks to generate an easy-to-interpret and flexible goal for the size of a core reserve
- Shows that a strict size of a core reserve does not deliver appropriate results for all populations
- Best setup of block structure is still under evaluation
- Allows to add further restrictions like minimum block length, or block frequency



$$\frac{8}{12} = 66\% \rightarrow \frac{10}{12} = 83\%$$

