



Strategic Research Center

GenSAP

Genomic Selection in Animals and Plants

Ministeriet for Fødevarer,
Landbrug og Fiskeri
NaturErhvervstyrelsen



gudop



The
Danish Council for
Strategic Research



Genetic improvement programs

genome wide markers



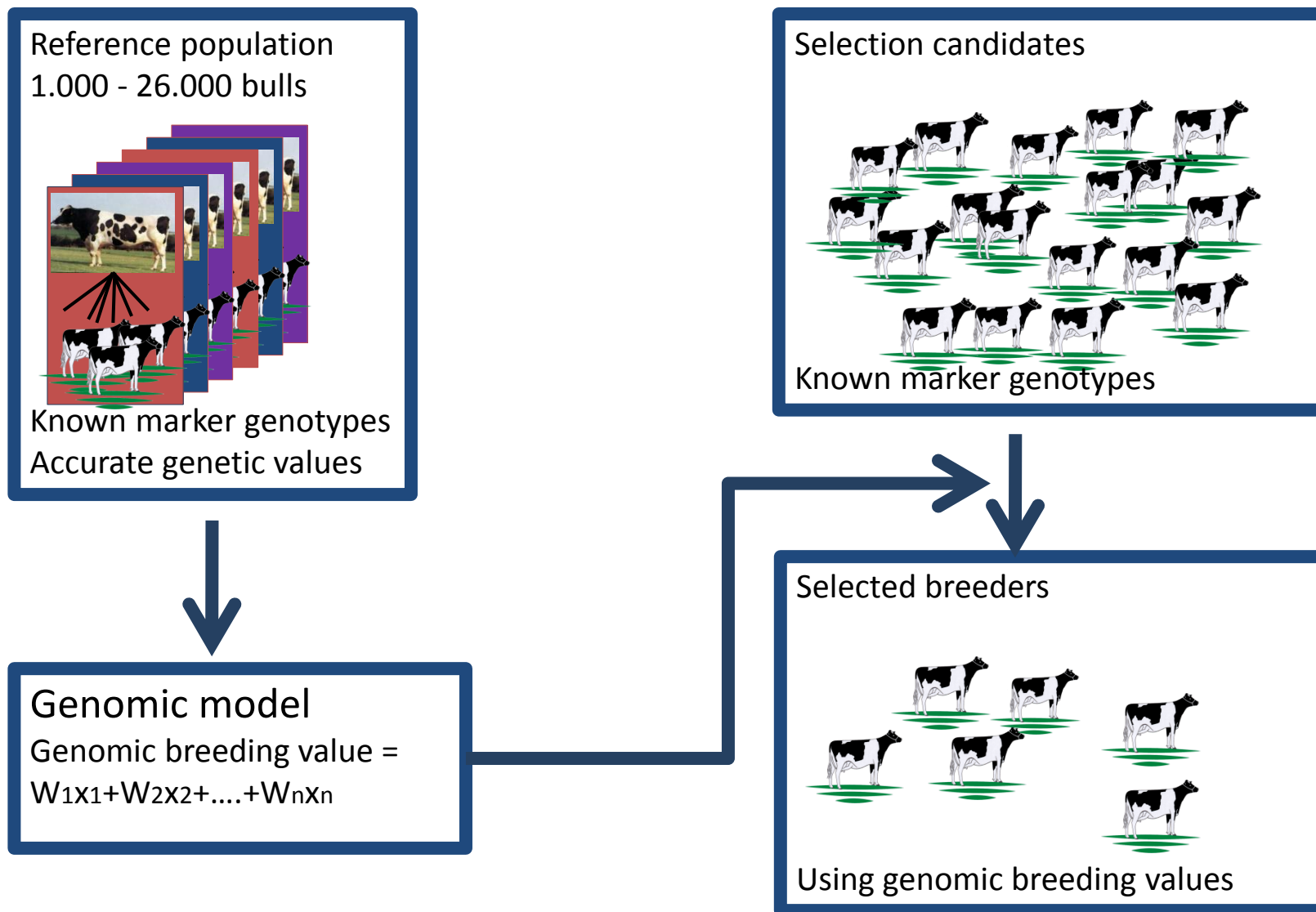
- Breeding objective
- Trait recording
- Genetic evaluation
- Selection and mating

← **Genomic prediction**

← **Optimisation**



Genomic selection





Genomic selection

- Prior to GenSAP
- Implemented in dairy cattle and pig breeding using 50K markers
 - Work very well for large homogeneous breeds
 - Works poorly across breeds and long genetic distances
- Large potential in other species
- New technologies available
 - Whole Genome Sequence (WGS) data
 - Other OMICS data



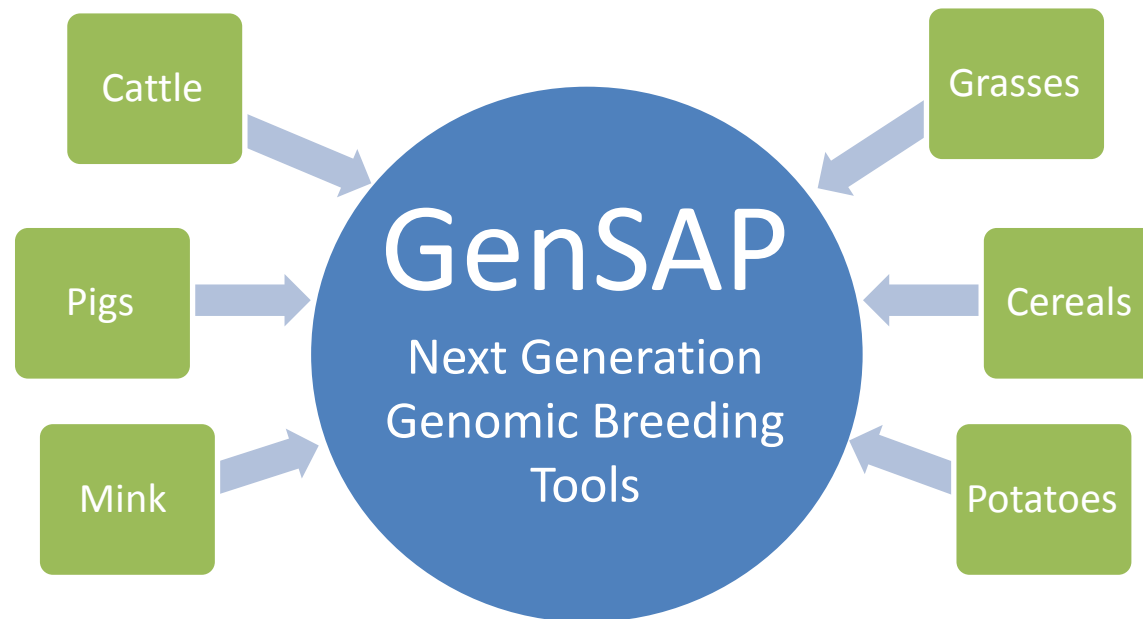
Vision

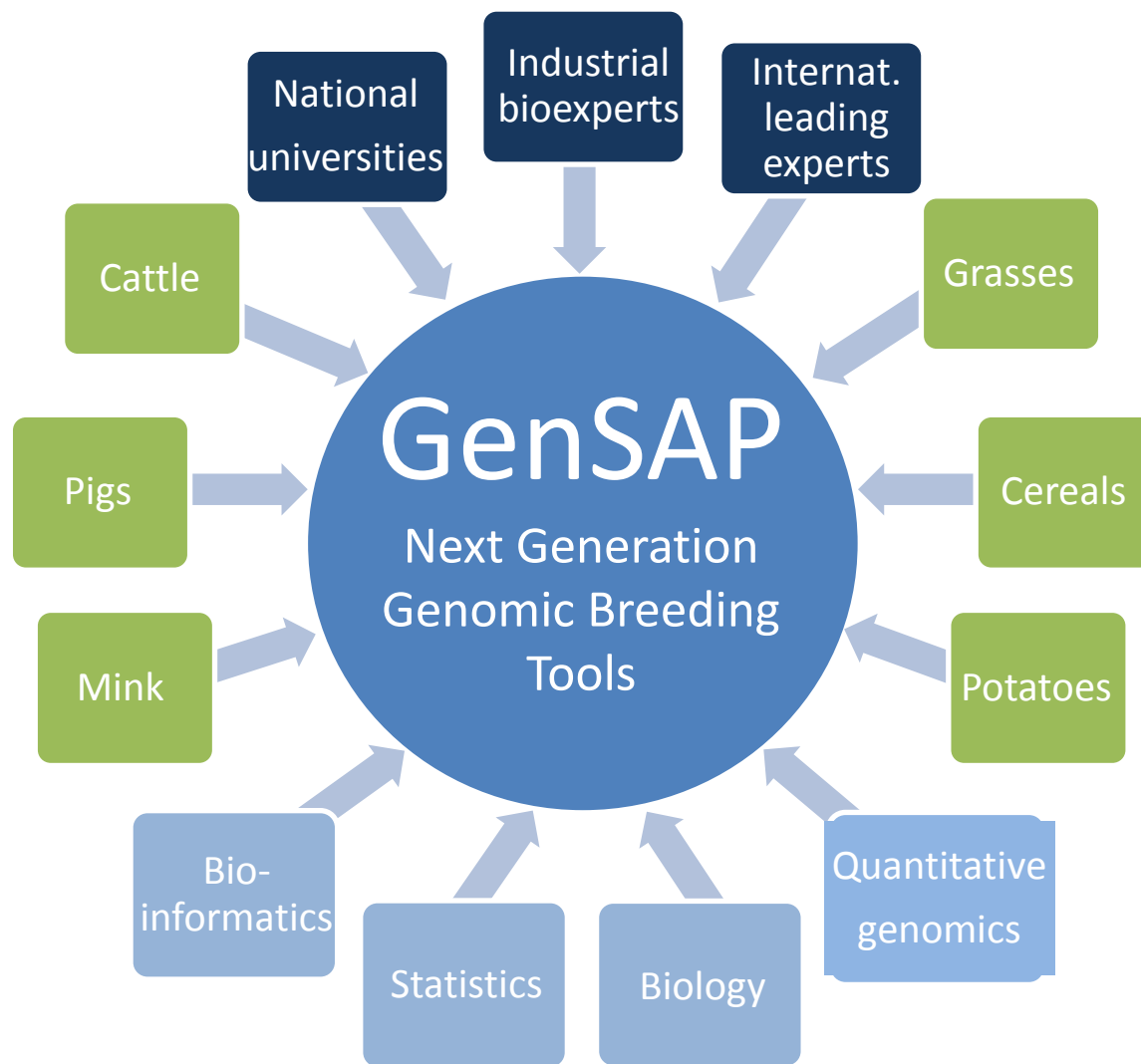
GenSAP

Next Generation
Genomic Breeding
Tools



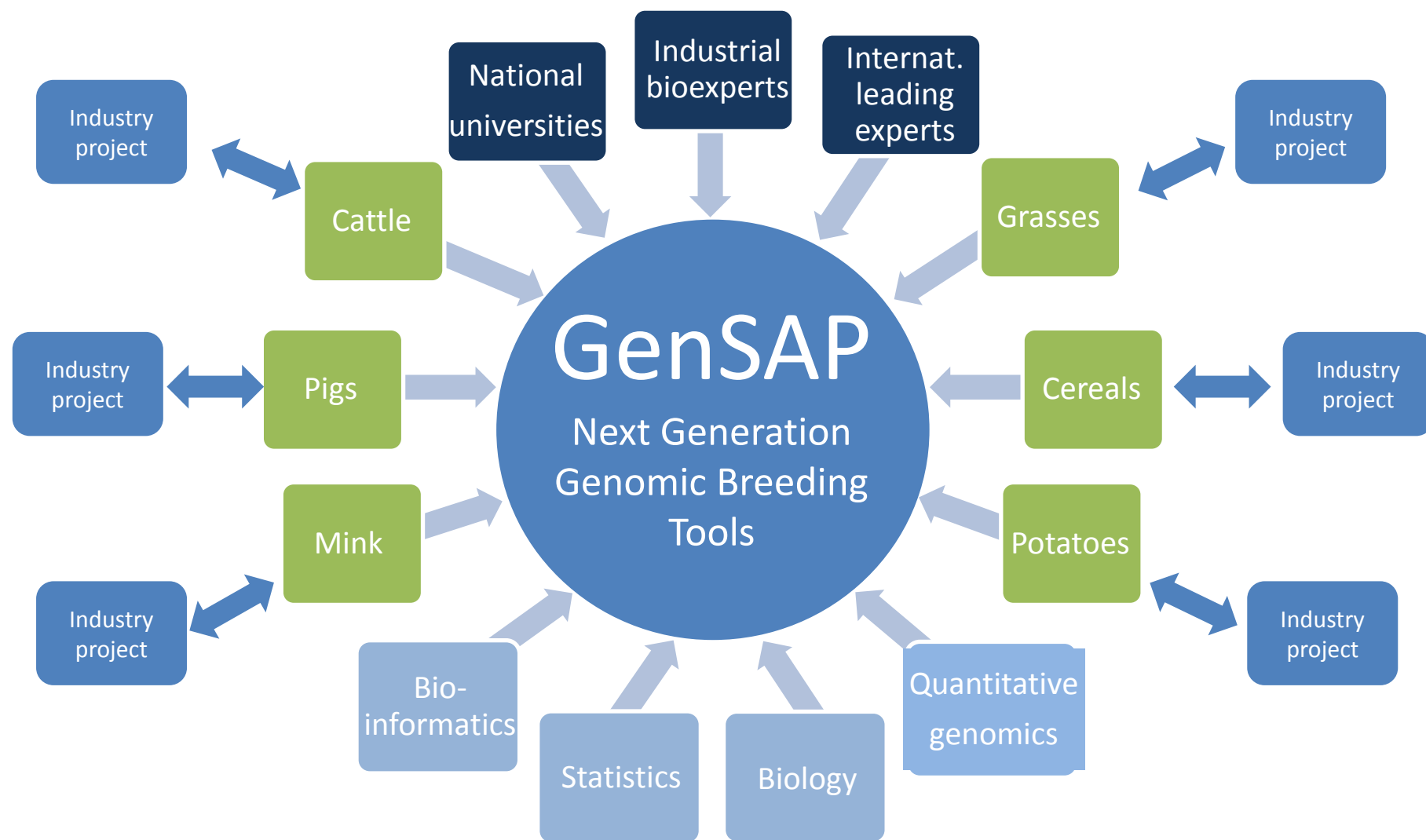
Animal and plant species







Synergy with other projects





GenSAP

**Genetic architecture of
complex traits**

**Genomic prediction
models**

**Assesment and
optimisation tools**

Whole genome sequence

Integration with OMICS data

Structural variants (CNV/inversions/deletions)

Additive, dominance, pleiotropic gene actions

Gene by environment interactions

Gene by gene interactions (depends on genetic background)

Common/rare/private genetic variants

Plant specific models: selfing, ploidity

Proof of concept: mink, potatoe, wheat, barley, ryegrass

- Includes flies as
model organism

- Prediction across breeds
- Scale to huge dimensions

-Genomic information to
handle inbreeding



Example dairy cattle

- **Strategy** to improve of genomic prediction in dairy cattle
- What are the **challenges**
- **Complementarity** and **synergies** between GenSAP and industry project

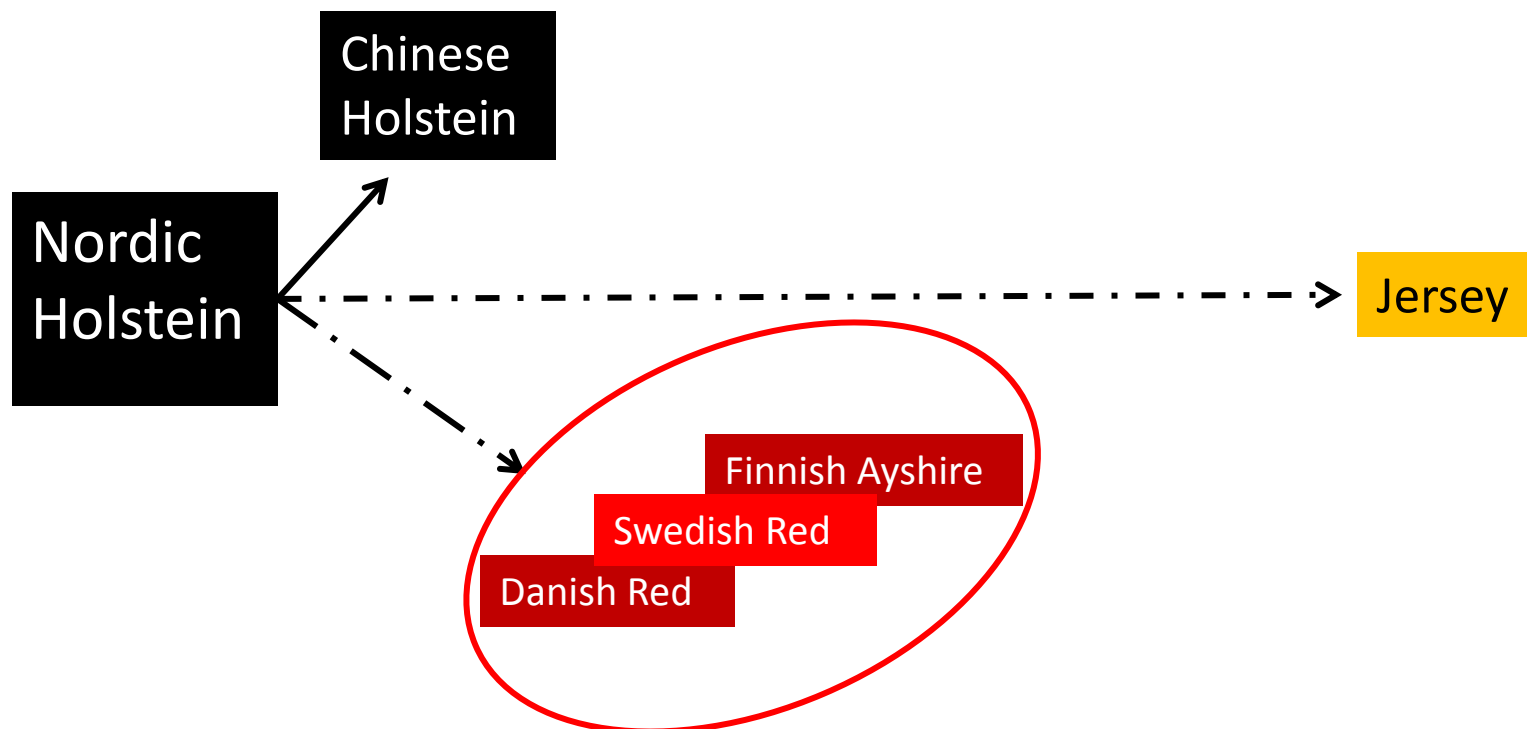


Expand reference population

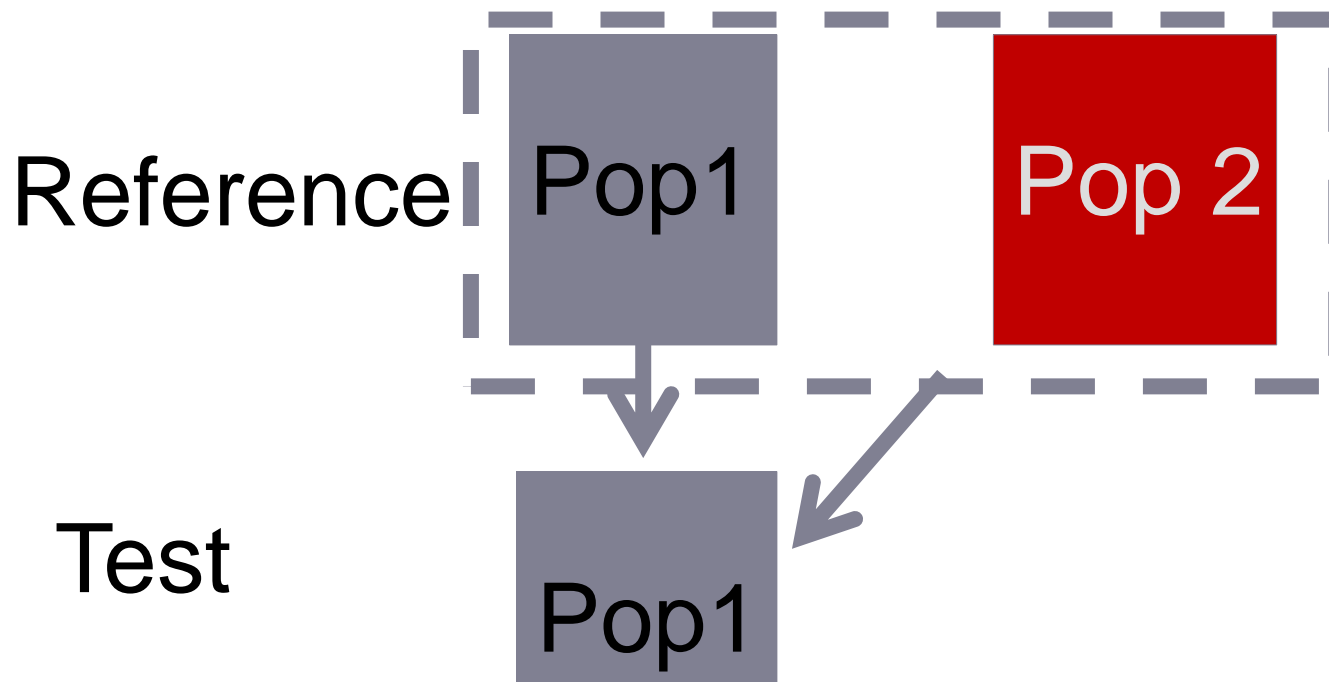
- All progeny tested bulls are already in the reference
- Including ungenotyped individuals using single step GBLUP **increases reliability** and **decreases bias** of predictions (Peipei Ma)
- Genotyping cows for the reference further **increases reliability** of genomic predictions
- Multi breed reference population for breeds without a large bull reference population



Joining reference populations



Increased reliability from joining populations



$$R^2 = R^2_{\text{REF1+2}} - R^2_{\text{REF1}}$$



Joining reference populations

	Chinese HF
	Nordic HF
	Zhou et al.

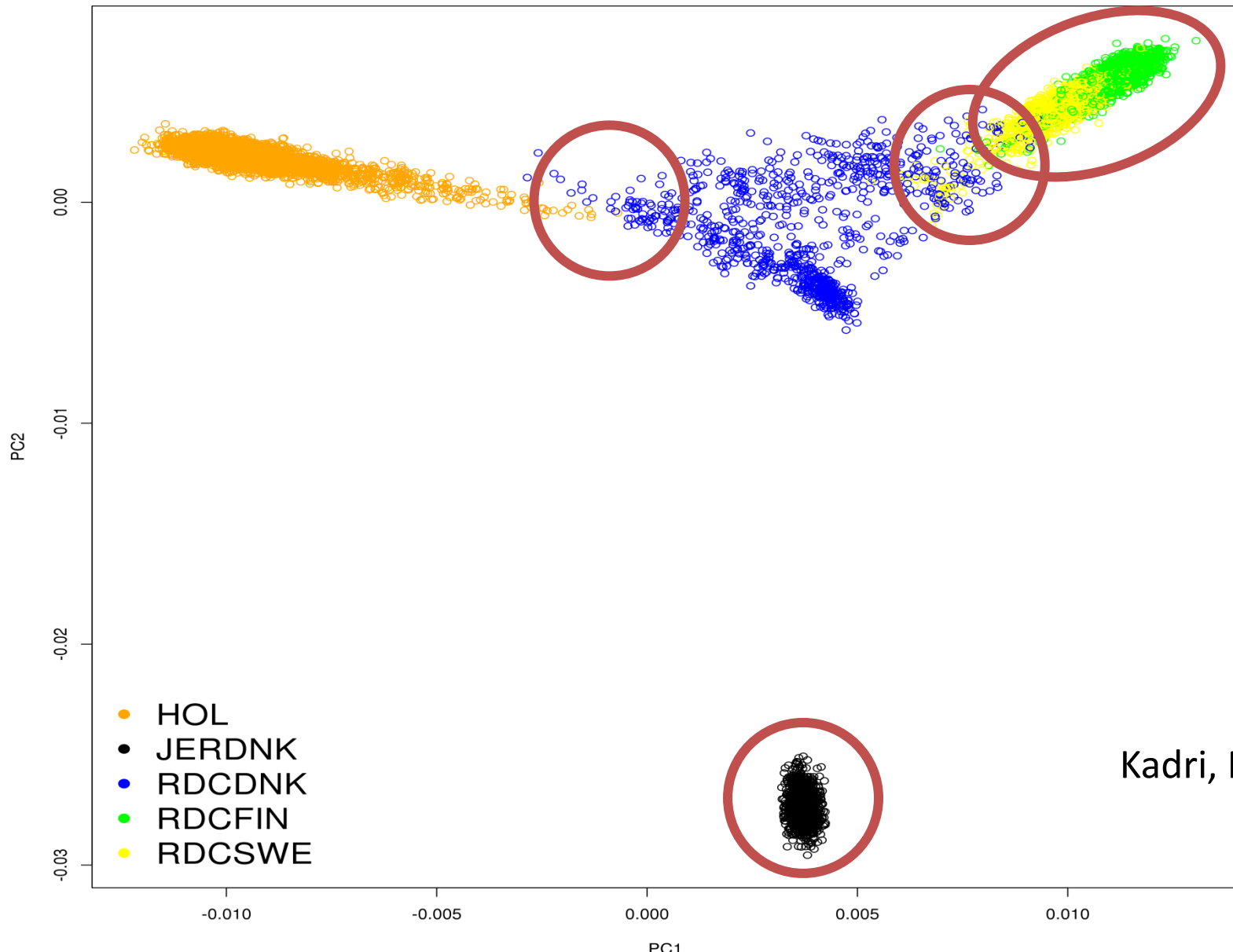
Large improvements when combining highly related populations

Low or no improvement when combining distantly related breeds

Increase in R^2	25% (7%)
----------------------	----------

↙ ↘

Low genomic relationships (<0.05)	'High' genomic relationships (>0.05)
$<+1\%$	$+8\%$



Kadri, NK, 2013



Problem with 'distant' phenotypes

- Optimal to use **causal variants** in prediction but they are **many** and **unknown**
- Use of genome wide markers **target most** genetic variation but introduces **noise** due to incomplete linkage disequilibrium
- **Noise increases with distance** from selection candidates to individuals with phenotypes due to recombinations



Genomic relationships at causal loci and markers

(de los Campos et al 2014)

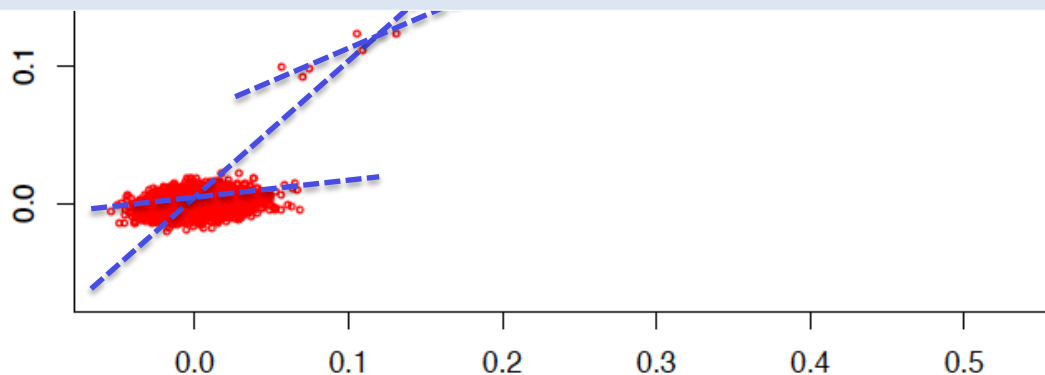
Framingham (LOW)



Across breeds genomic relationships are **low** and **correlate poorly** to relationships at causal loci

Therefore

Use markers **closer to causal variants**



Gij at Causal Loci (5K)



Closer to causal variants

- Increase marker density → whole genome sequence (WGS) data
- WGS data has the potential to **improve** genomic predictions **across breeds** - but **don't** use markers **not** in high LD with QTL (Irene van der Berg)
- Two strategies
 - Use WGS data directly in a **Bayesian** variable selection model
 - Use of selected **QTL markers** from WGS based GWAS improve predictions (Rasmus Brøndum)
- Use **prior knowledge** to improve bayesian model (Mike Goddard)

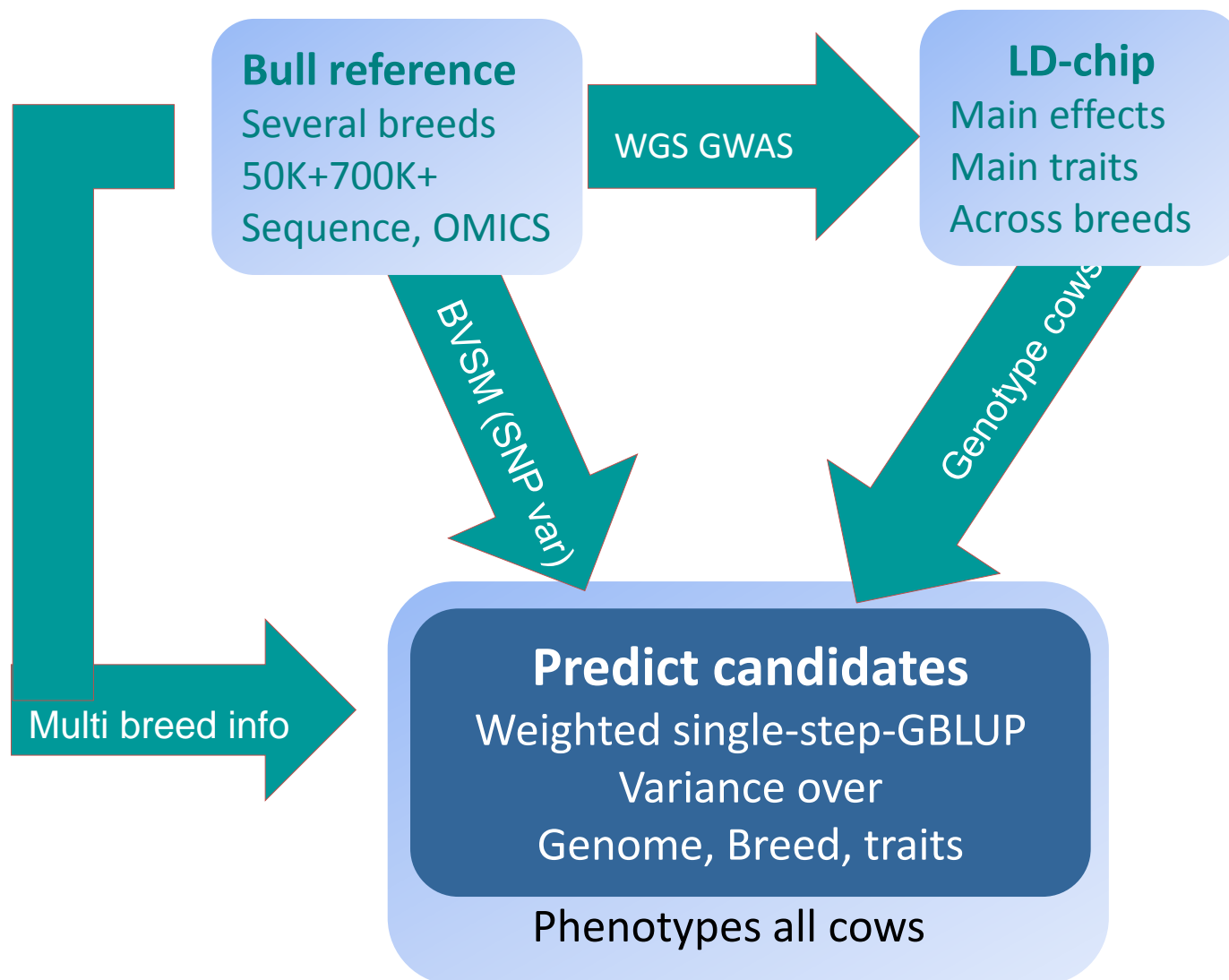


Requirements/challenges for the next generation genomic breeding tool

- Scale to huge number of individuals
 - Millions **ungenotyped** and hundreds of thousands **genotyped** individuals
- Use information from advanced models
 - Bayesian **mixture** models, **haplotypes**, **QTL** etc.
- Information from WGS data across several breeds
- Potential integration with other 'OMICS' data



Genomic strategy





GenSAP and industry projects

Industry project

- Data generation
- Validation of models
- Implementation

GenSAP

- Develop methods
- Proof of concept

Include all genotyped
and ungenotyped cows

cow genotypes
V&I



SS-GBLUP method

Improved
prediction models
(closer to causal genes)

BVS models
Haplotype models
LD chip
V&I



WGS data (AU and 1000 genomes)
Sequence based models
Weighted GBLUP
Integrative models

Across breeds prediction

Nordic breeds
Validation



Nordic and French breeds
sequence based



Synergy with other projects

