

Using genotypes from Australia's numerically smaller sheep breeds to expand the genomic reference population

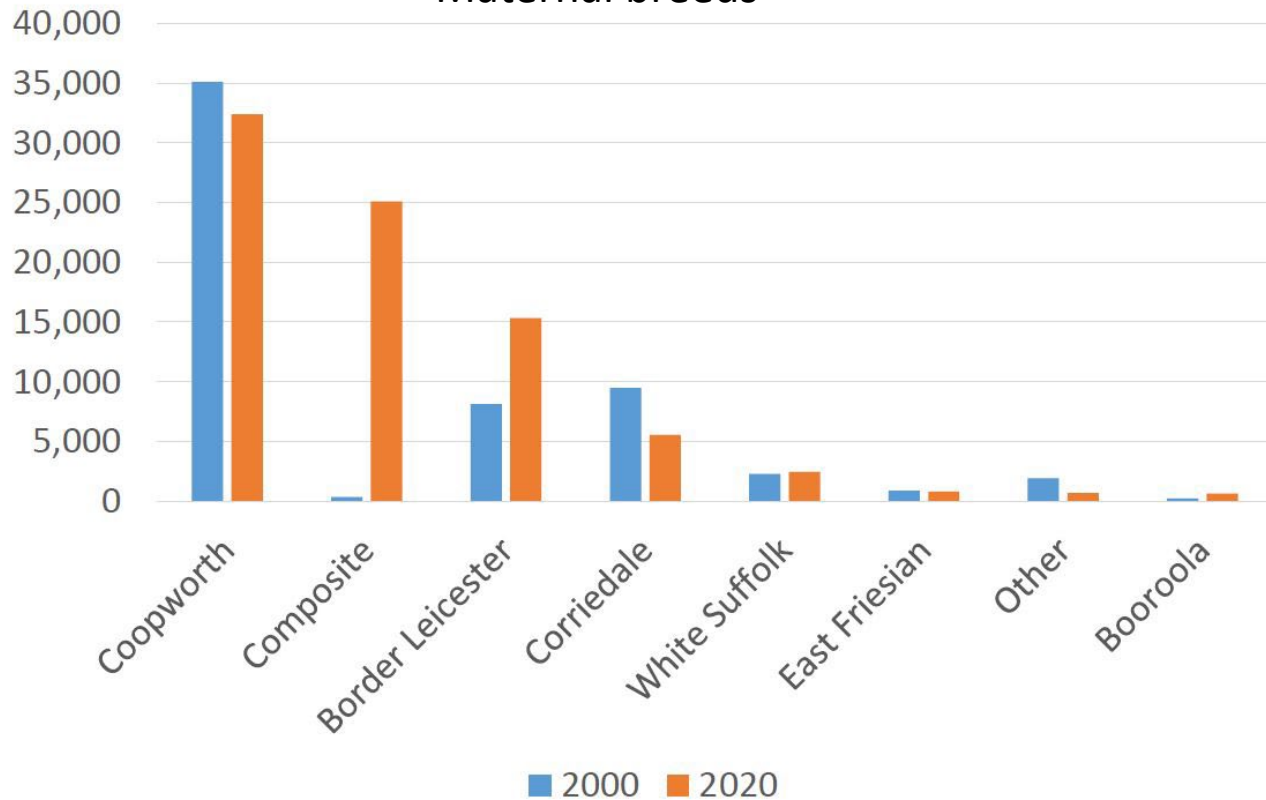
P. Alexandri, P.M. Gurman, K. Gore, R. MacArthur-Onslow, D. J. Brown

Australian sheep genetic evaluation

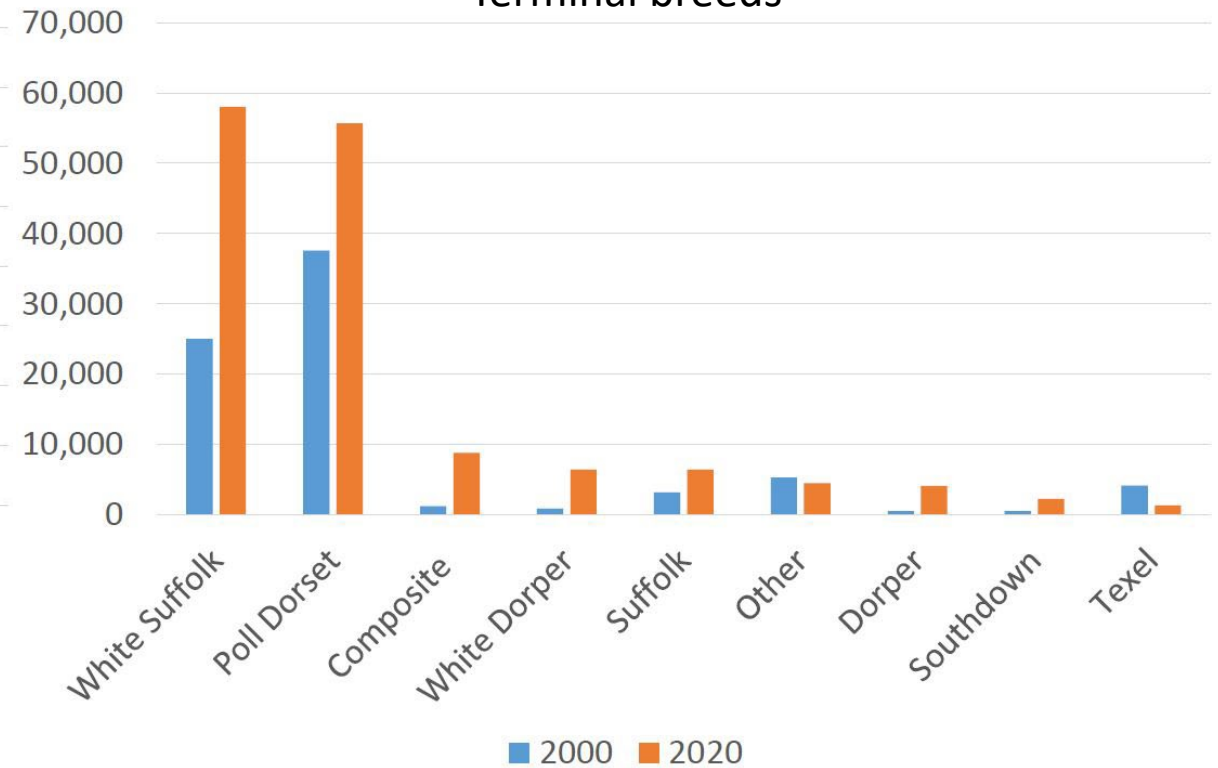
- First national genetic evaluation of Australian sheep implemented in 2005
- Separate large scale analyses
 - Terminal breeds: Poll Dorset, White Suffolk
 - Maternal breeds: Border Leicester, Coopworth
- Structure of analysis data can change over time

Data changes

Maternal breeds



Terminal breeds



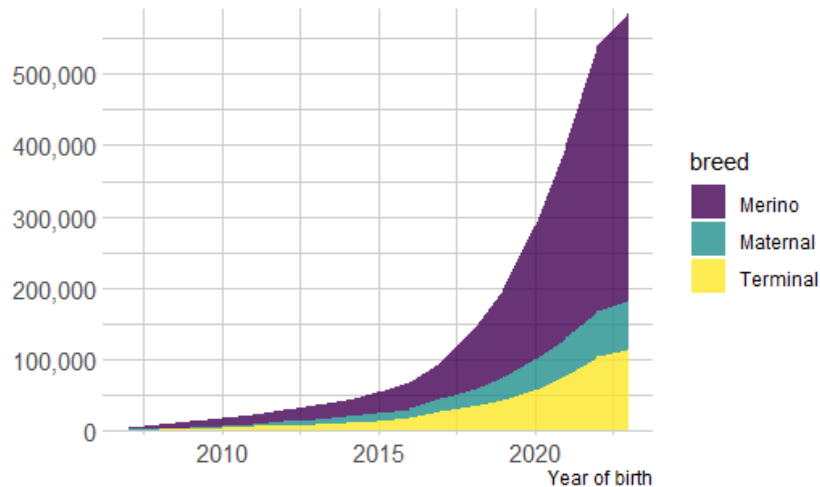
Data submission per year

McMillan et al., 2023

Single step combines all information

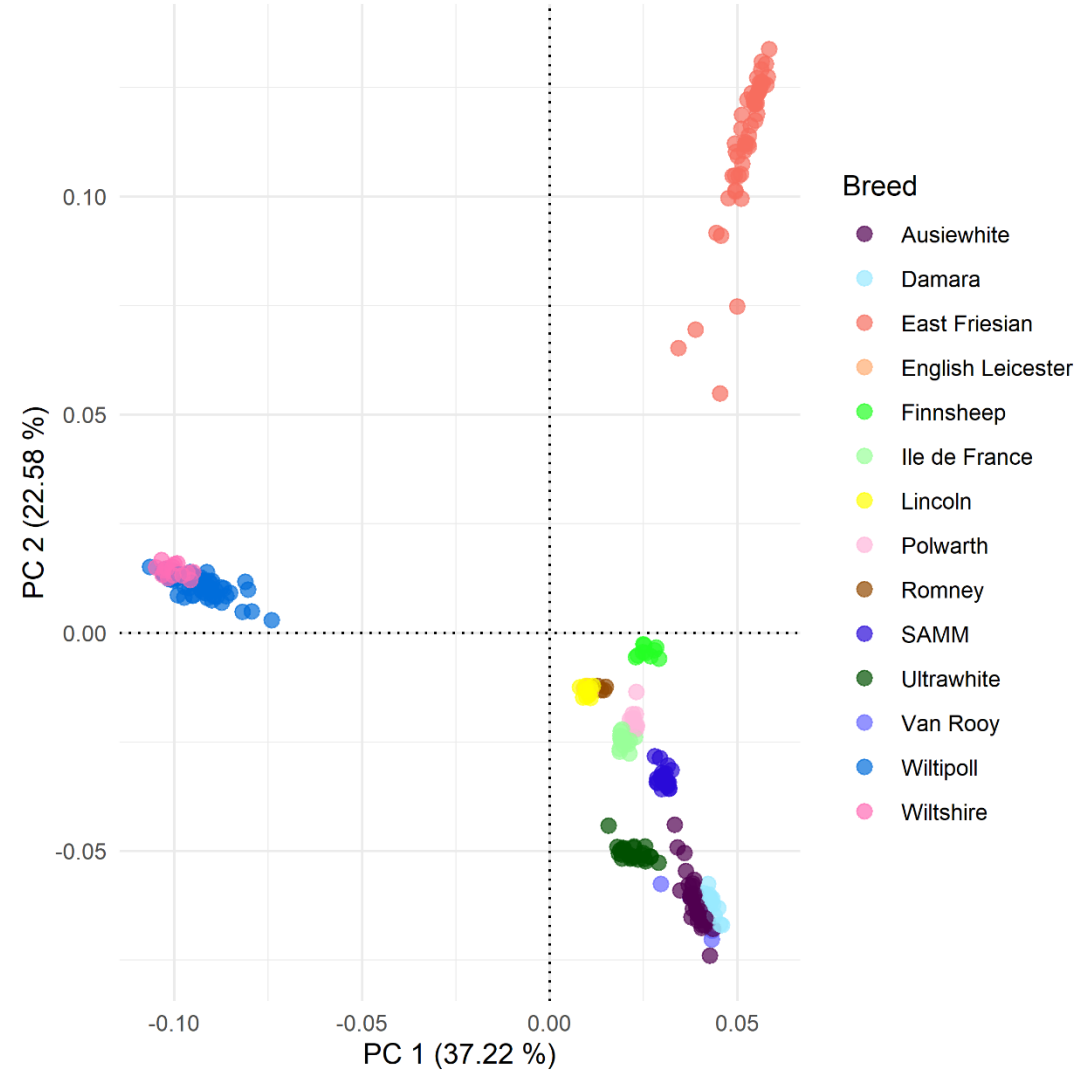
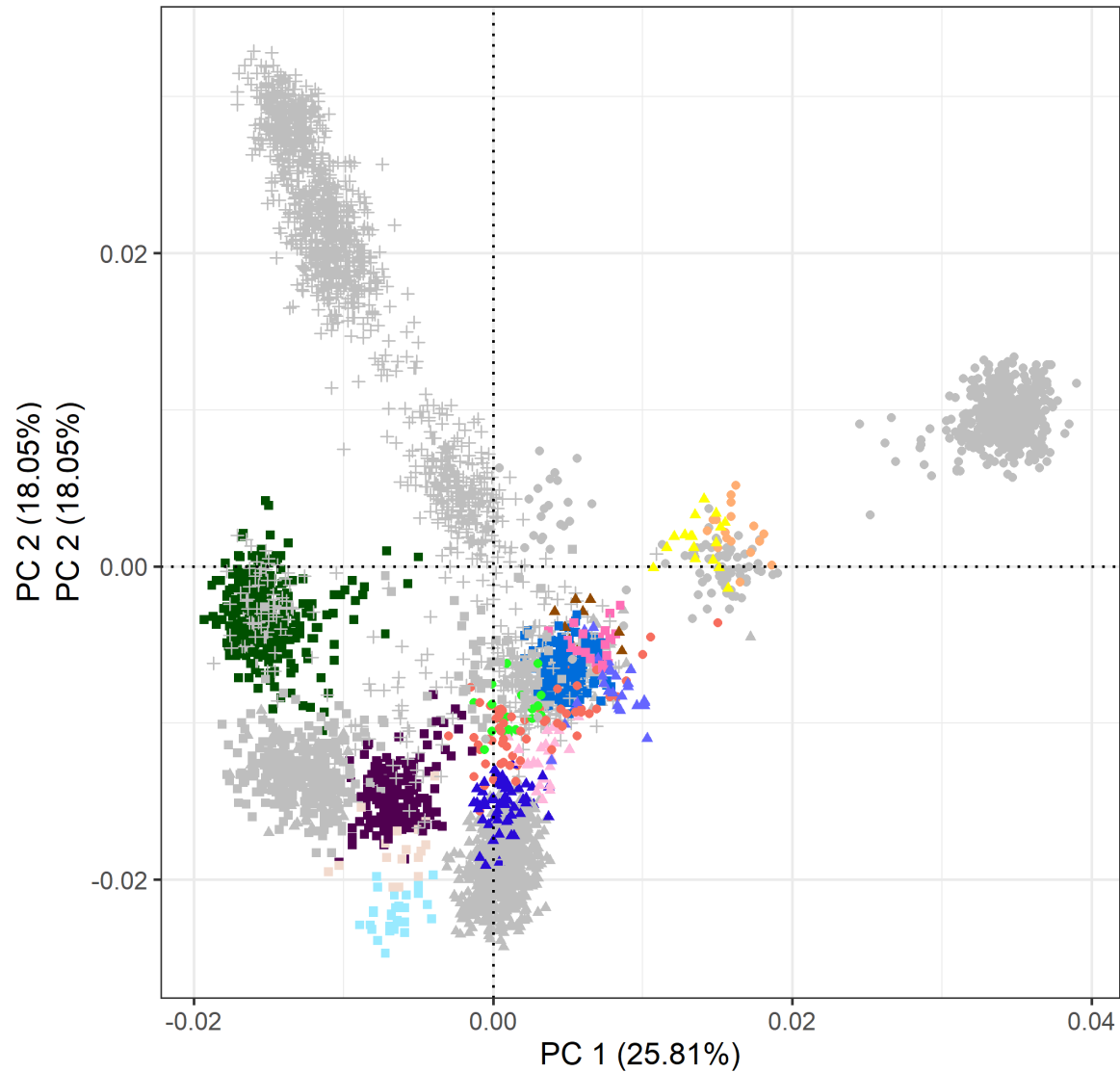


Growth in sheep genotyping



Analysis	Pedigree	Records	Traits	Genetic groups	Genotypes	Added in '23
Merino	4.M	24.8M	94	685	438.7K	165.2K
Maternal	2.7M	7.5M	71	110	74.9K	22.1K
Terminal	3.9M	12.3M	51	93	119.7K	43.5K

What smaller breeds?



Challenges for multibreed evaluation

- Changes in breed structure
 - Smaller proportion of pure bred animals
 - Rise in composite animals
- Accurate modelling of breed and heterosis
- Breed composition reporting
- Multibreed genetic evaluation for diverse industry needs

Why re-assess reference population

- Genomic reference population first developed in 2017
 - Allele frequencies used to construct breed-adjusted genomic relationship matrixes in terminal and maternal analyses
- Not all breeds were included
 - Include new breeds/composites and numerically smaller breeds
- Number of animals has increased exponentially
 - Need for a more representative genomic reference population

Overall approach

- Use genetic information to detect the presence of population structure
- Assign individuals to clusters based on their genotypes
- Simultaneously estimate cluster allele frequencies
- STRUCTURE 2.3.4, assuming independent loci and admixture between populations

More details on filtering

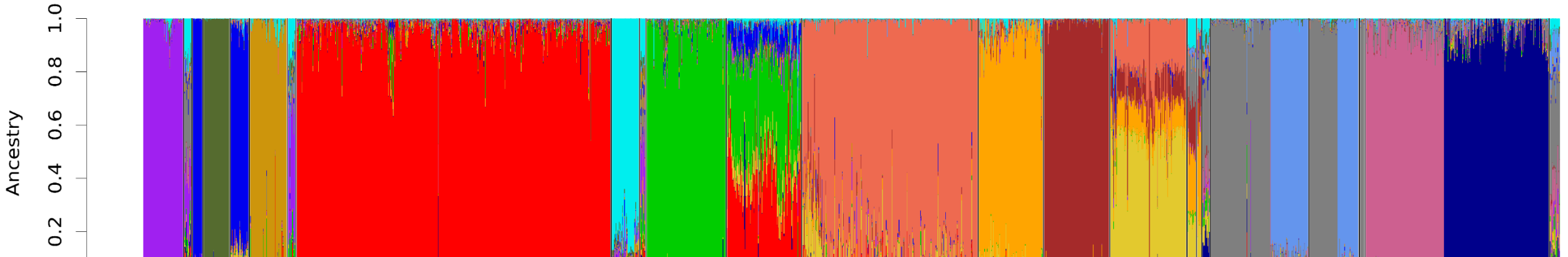
- Clustering algorithms tend to pick up family structures
- Detect unrelated animals from the same breed
 - Pedigree based breed composition proportion > 0.9
 - GRM filter < 0.2 off diagonal within breed

Animals and genotypes

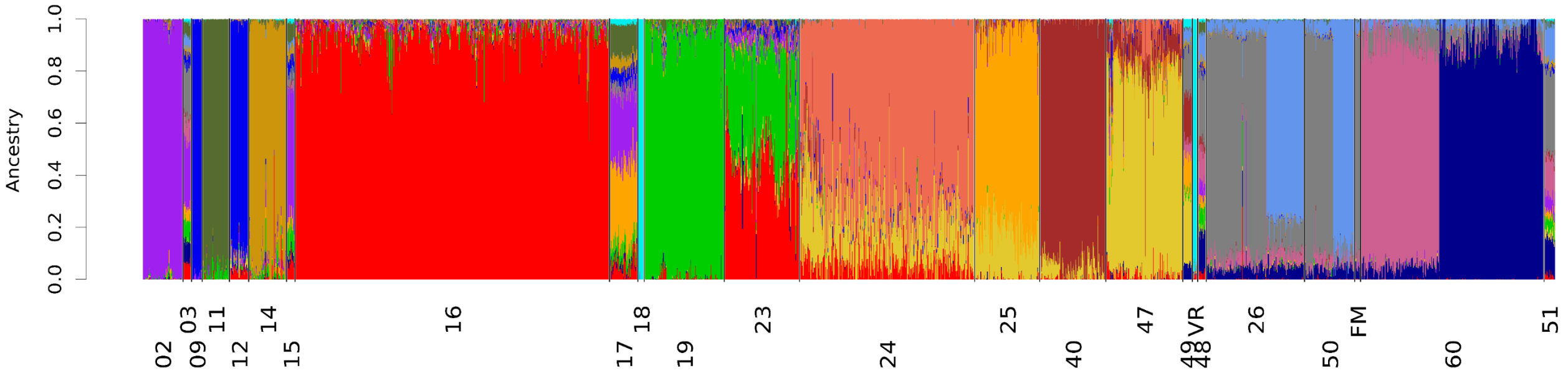
- 4329 animals from different breeds genotyped with different arrays
- Different data sets were used:
 - Common SNPs between different arrays (~3,000 SNPs)
 - Imputed SNPs as used in LAMBPLAN analysis (~62,000 SNPs)

Repeat with raw and imputed genotypes

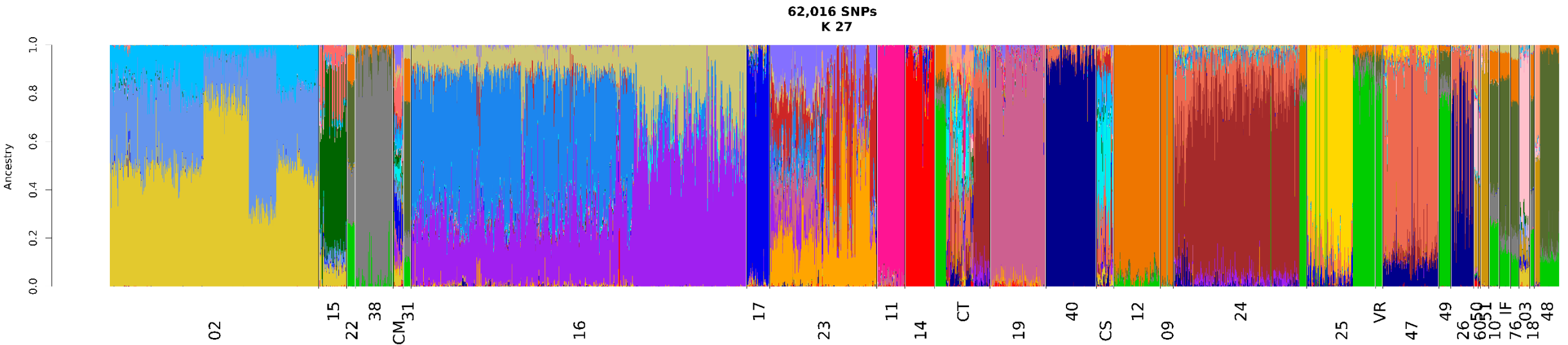
**3K Structure
K 15**



**Ovis 62K Structure
K 15**



What do results look like



	CL1	CL2	CL3	CL4	CL5
ID1	0.168	0.338	0.06	0.362	0.072
ID2	0.173	0.376	0.041	0.352	0.058
ID3	0.214	0.311	0.053	0.349	0.073

Population membership



	CL1	CL2	CL3	CL4	CL5
Allele 1	0.276	0.677	0.408	0.483	0.286
Allele 2	0.724	0.323	0.592	0.517	0.714

Allele Frequency

Conclusions

- Several homogeneous breeds but also breeds including more than one clusters
- Updated allele frequencies representative of Australian sheep genetic variation
 - More accurate definitions of breed structure for all animals

Industry implementation

- Breed component analysis based on estimated allele frequencies
- Implemented in May 2024: new genotypes included in the evaluation
 - 29,000 Terminal
 - 13,000 Maternal

Acknowledgements

