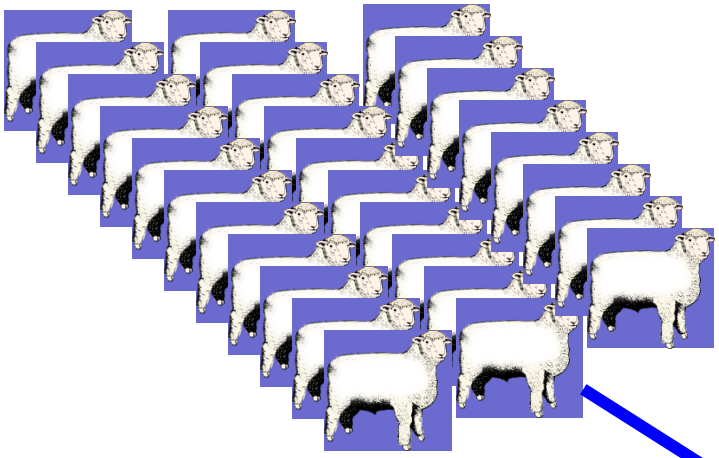
Design of reference populations

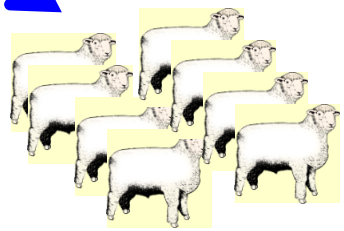
Julius van der Werf

CRC for Sheep Industry Innovation
School of Environmental and Rural Science, UNE, Armidale, NSW

Genomic Prediction: basic idea



1) Somebody (else) measures lots of sheep, and their DNA
→ Reference population



2) A breeder tests DNA on young rams

Prediction from DNA → genomic breeding values - GBV

GBV + Current ASBV → Improved ASBV

Merit depends on trait measurability

Setting up reference populations

Trait is already measured	Early measurement	Late Measurement
YES	No Need	Use industry data (milk, fertility, late wool)
NO	Create Reference population (slaughter)	Create Reference population

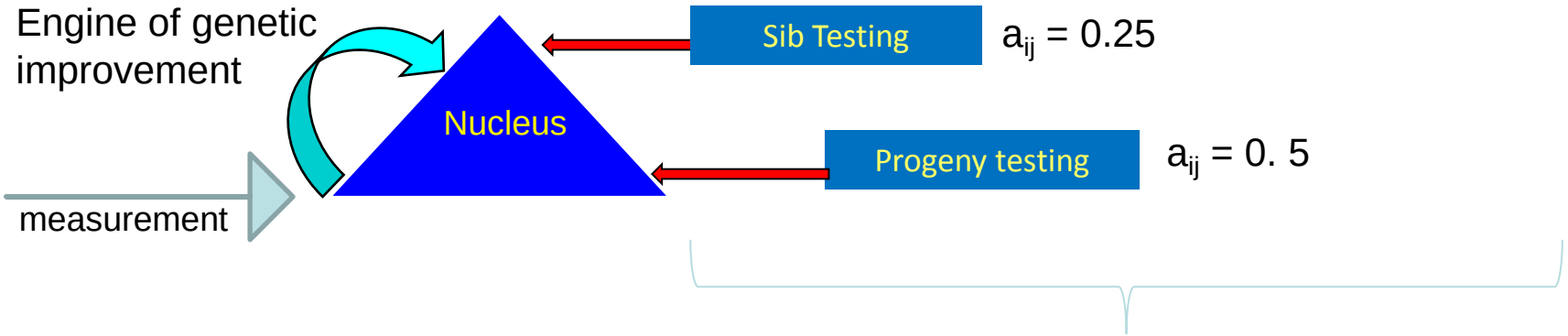
Genomic selection has affected the need for phenotyping !

more…not less

Who pays?

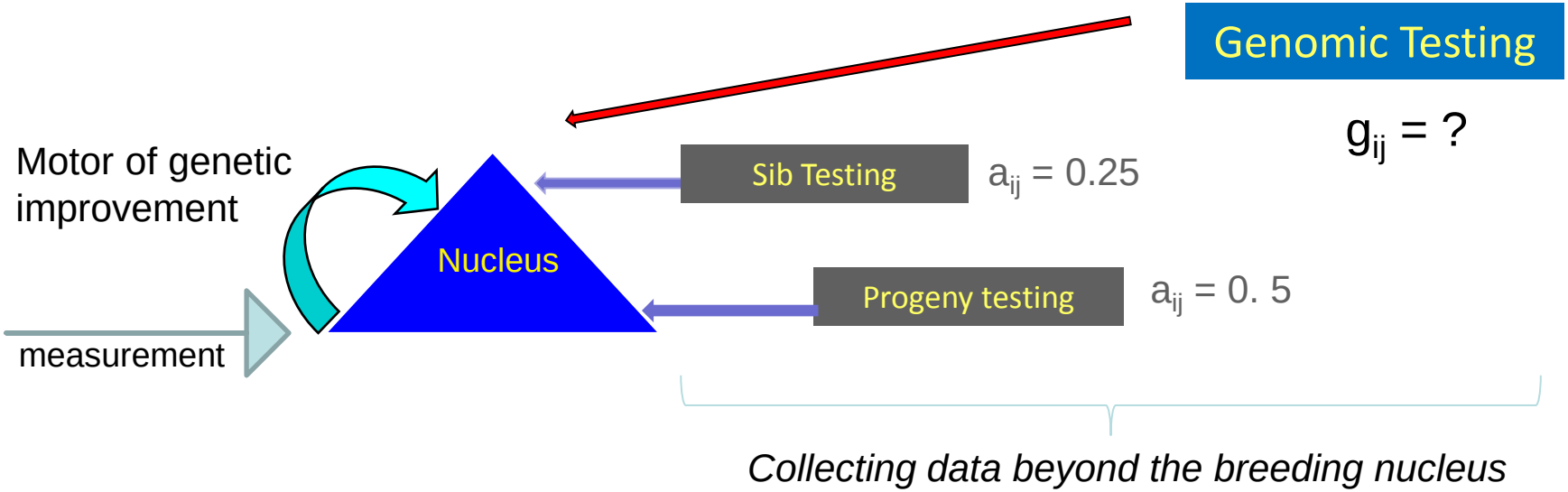
Design of a reference population

Investing in information for genetic improvement pre-genomics

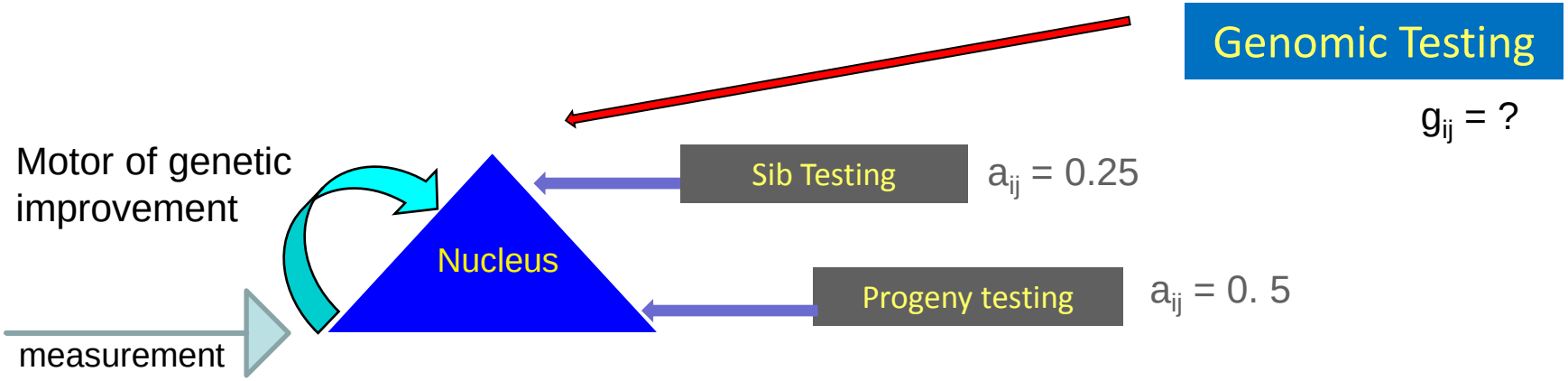


Collecting data beyond the breeding nucleus
- if it is hard to measure within -

Investing in information for genetic improvement



Investing in information for genetic improvement



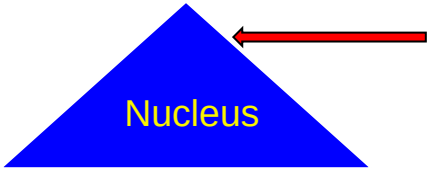
Measure outside nucleus if traits - can not be measured within nucleus
- carcass, eating quality, reproduction

otherwise, reference population can be nucleus

Genomic selection has an advantage over sib or progeny test selection because

1. the information comes earlier
2. can afford to test more distant relatives

Design of Reference Population



Genomic Testing
Relationship paradigm

$a_{ij} > 0.05$

Genomic Testing
LD paradigm

$a_{ij} > 0.0000001$

Relationship paradigm

Need relatives in reference
Need to keep reference 'up to date'
Denser markers maybe of limited benefit
Accuracy limited by relationships and # of relatives
Consider to use IBD inference

LD paradigm

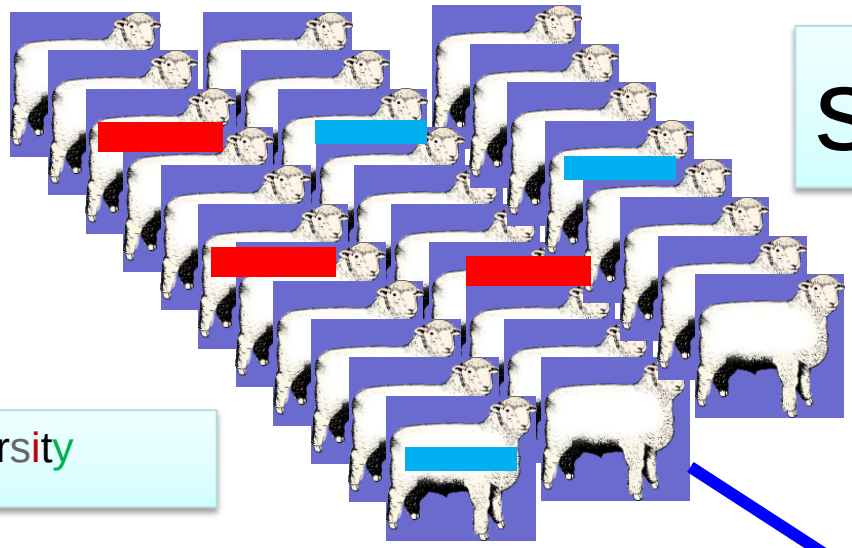
May achieve prediction across breeds
Reference population of long lasting benefit
Accuracy limited by marker density and size of reference
Requires detectible average effects
across wide range of genetic background

Summarizing Genomic Prediction

- What information is used?

- Based on very many small – genomic- relationships
- Does not require ‘direct relatives’ to be tested
- Can be based on distant relatives ‘some generations away’
- …..but the number of small relatives needs to be large (thousands)
- Can not predict across breed

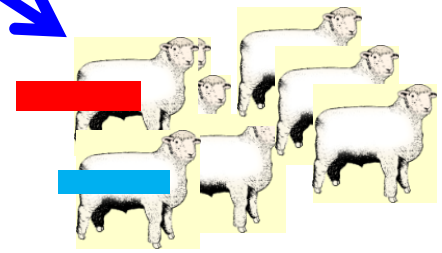
Design of reference populations



size

diversity

relatedness



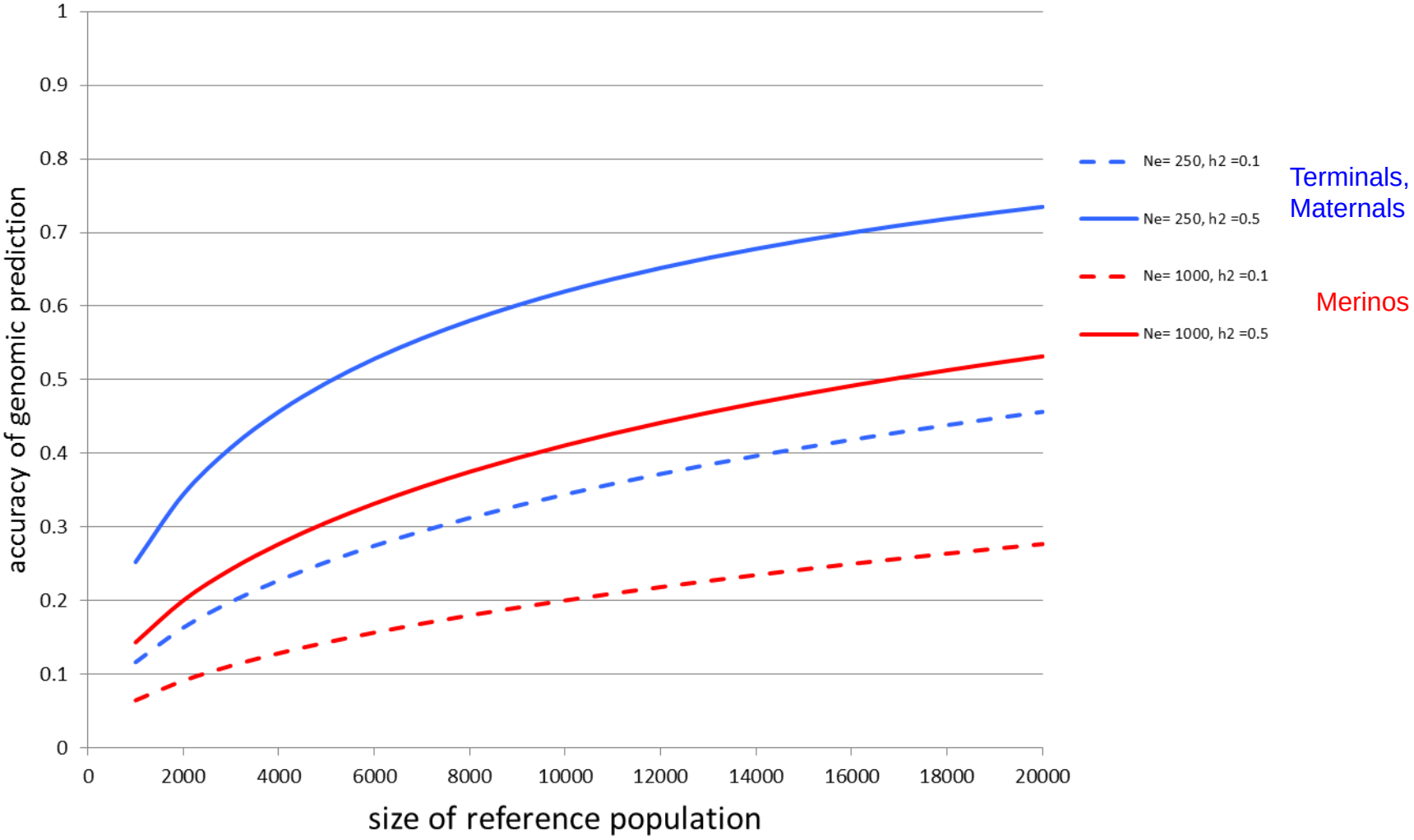
Select on impact → merit

→ diversity

Multi-breed
Across breed?
Longevity of RefPop?

Accuracy of genomic prediction depending on size of reference population

Goddard 2009



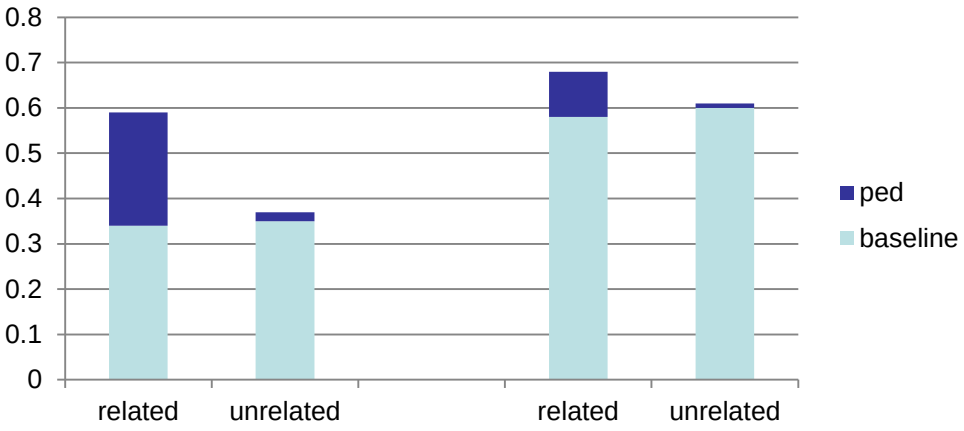
design of reference population

- Relatedness between reference population and selection candidates
- Across breeds or lines?
- Number of sires, nr of progeny per sire, which dams?

Sources of information contributing to GBV accuracy

	BLUP	GBLUP	half life
1. Variation between families	++	++	1 gen
2. Variation within families	0	+	1 gen
3. Markers tracking effects of genome segments/LD	0	++.+	several gen's

Info on 'unrelated'



smaller ref pop larger ref pop

Depending on size of reference population