The Value of Haplotyping

S Lloyd, D Bayard, S Lester, J Williamson, R Dawkins

CY O'Connor ERADE Village Foundation, Murdoch University and University of Western Australia.

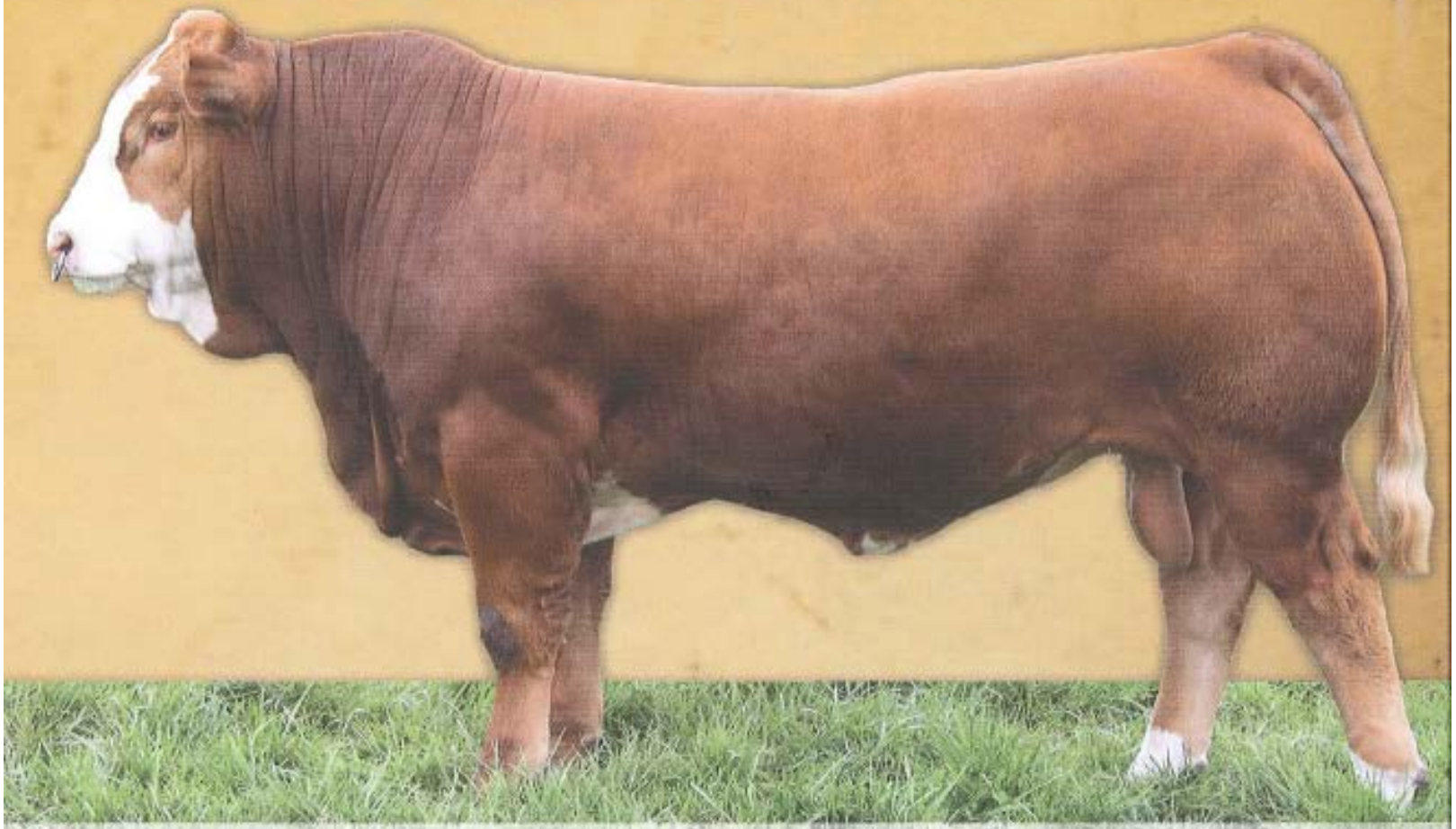
Aim

- Improve the utility of genetic selection for health and commercial returns
- Avoid negative outcomes

How?

GOLDRUSH

THE ONCE IN A LIFETIME BULL.



We are taking expressions of interest for a very limited semen release.

WOONALLEE GOLDRUSH (P)
ID. WEEP6061 DOB. 09/05/11
SIRE. W. BHR TORNADO (P) (ET)
MATERNAL GS. LOCKEIM PRIMAL

	CED	CEM	Birth	200 Wt	400 Wt	600 Wt	Carcase	EMA	Rib Fat	Rump	RBY	IMF
EBV	-5.7	-3.8	+3.4	+27	+48	+50	+34	+2.8	+1.6	+0.6	+0.6	0.0
ACC	41%	36%	77%	69%	69%	69%	56%	43%	50%	53%	41%	34%

ABACUS approach based on NeoDarwinism

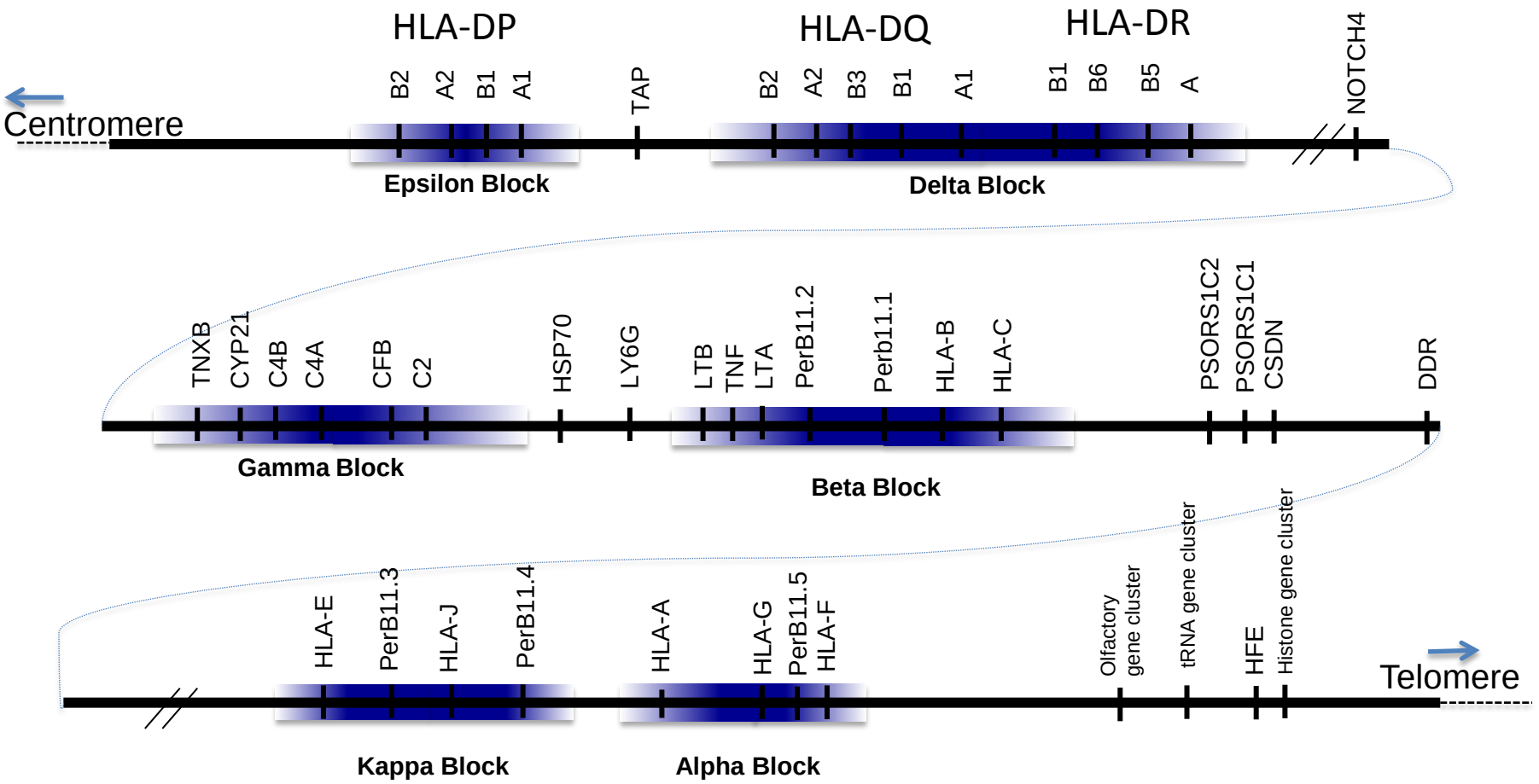
1. Genetic diversity or inherited variation requires on-going mutation.
2. Diversity accumulates through errors in copying of DNA.
3. The unit of inheritance is the allele.
4. At each locus, alleles may be deleterious, beneficial or neutral.
5. Meiotic recombination scrambles maternal and paternal alleles.
6. Linkage disequilibrium between SNPs define haplotypes.

Challenged by Genomic sequencing.

Alternative Ways of Interpreting the Genome

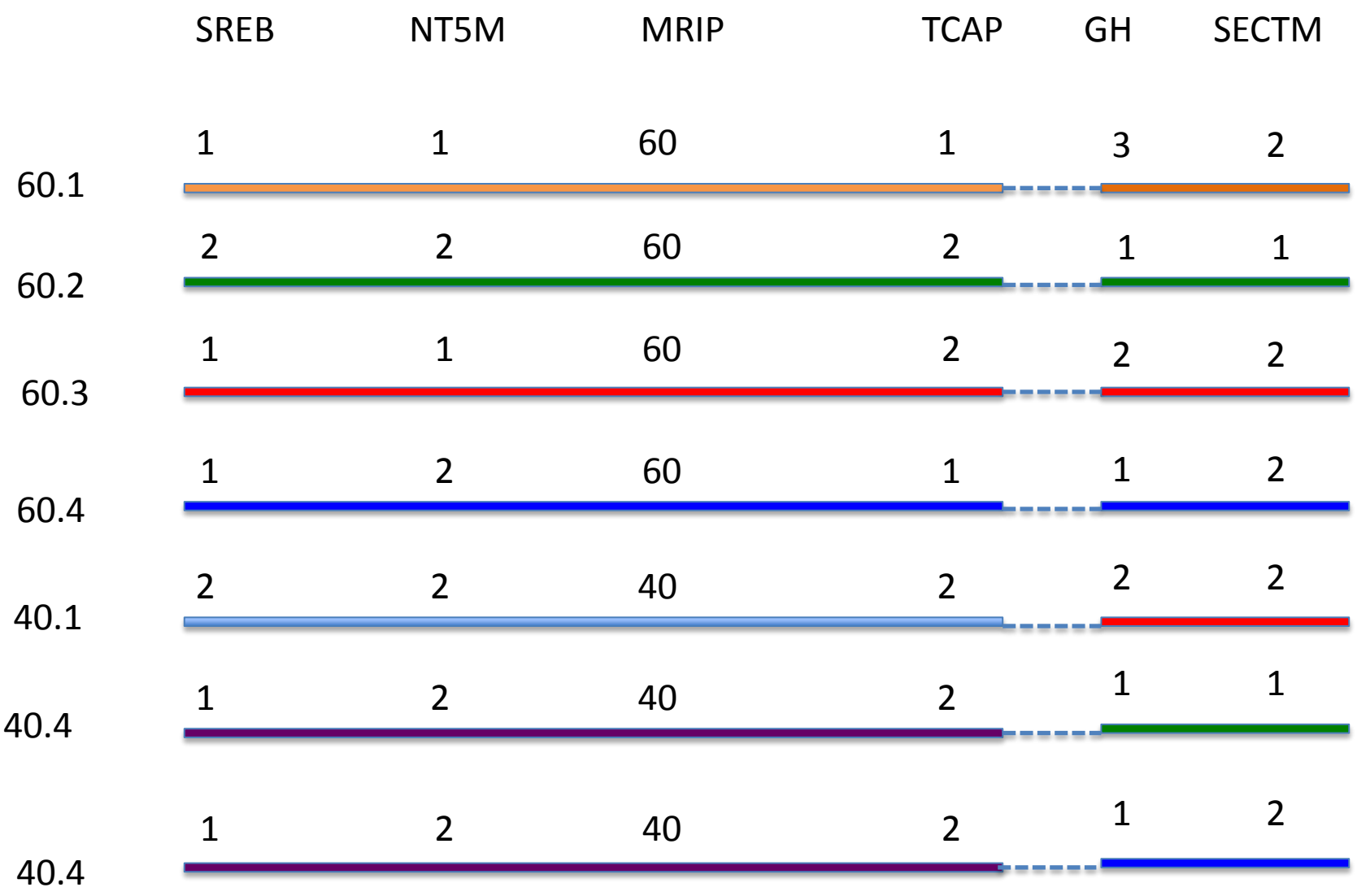
- Diversity is inherited from ancestors rather than created by recent mutation.
- Diversity is regenerated at speciation and maintained by meiotic recombination between the ancestral haplotypes of nearby polymorphic frozen blocks.
- The unit of inheritance is the ancestral haplotype.
- Such sequences carry specific alleles, duplicons, indels and RLEs.

Example of polymorphic blocks in Hosa



Each ancestral haplotype has its own map. Adapted from Dawkins et al (1999) and the MHC Map (http://www.path.cam.ac.uk/~mhc/map/MainMapPage.html). PFBs are shaded. Not all genes are shown. PerB11 is now designated MIC

Bota C19: haplospecific markers.



Relevant genes within a 14Mb region of Bota Chr19

SREBF1 Sterol regulator element binding transcription factor 1
- Cholesterol and FA biosynthesis
- Regulate intracellular lipids

STAT5A transcription factor
- Regulation of mammary tissue development
- Milk-fat percentage
- Fat cell formation and function

GH Growth Hormone
- Breakdown of lipids

Urotensin2 receptor
- glucose metabolism
- insulin resistance
- skeletal muscle fat deposition
- fatty acid metabolism

FASN Fatty acid synthase
- Complex enzyme system involved in FA synthesis

GENOMIC STRATEGY FOR SELECTIVE BREEDING

1. Define polymorphic frozen BLOCKS and their ancestral haplotypes in regions of interest.

2. Develop robust HAPLOSPECIFIC markers.

3. Identify haplotypes which SPECIFIC to breeds with desirable traits whilst excluding those which are common to multiple breeds.

4. Develop minimal requirements for BREED determination and parentage.

5. Compare haplotype frequencies in ELITE versus poor performers within a breed.

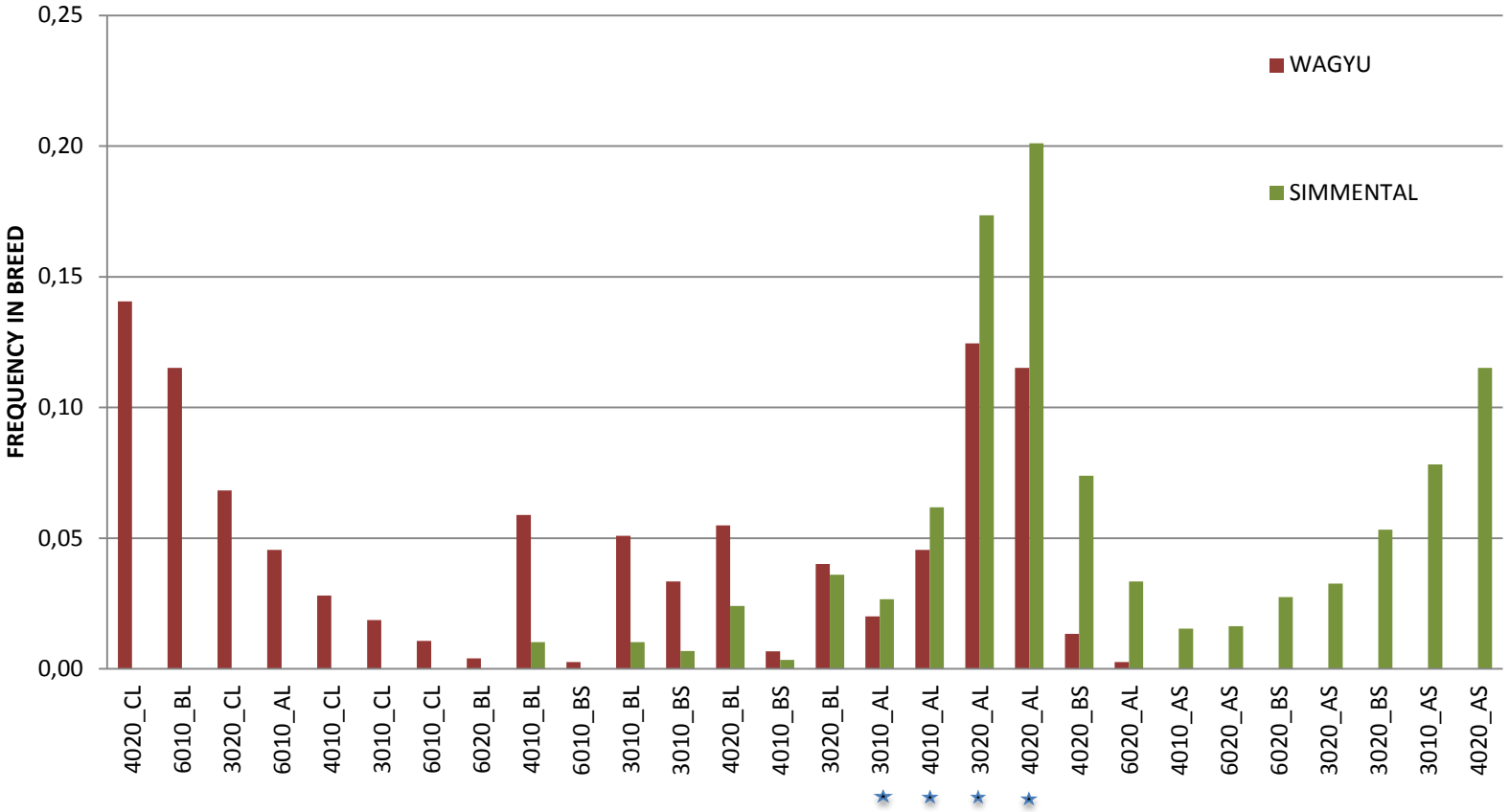
6. Compare haplotypes which arose by RECOMBINATION to localise the operative components.

7. Consider functional explanations ENCODED by haplotypes.

8. Test haplotype interactions using AI, ET and Cloning experiments.

9. Use MAS to BLEND preferred haplotypes into already successful herds, or to increase the frequency of haplotypes already present.

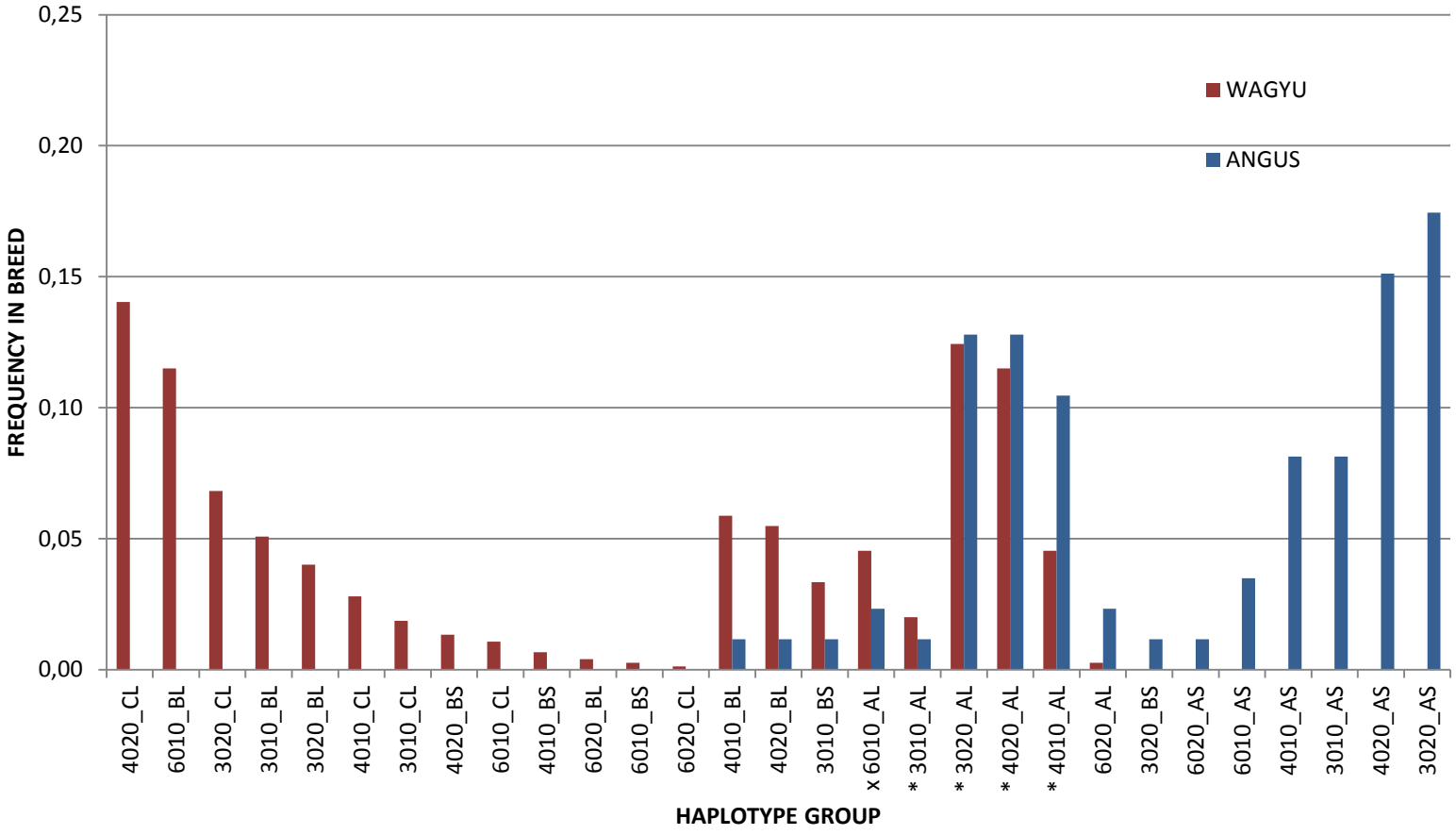
MRIP_TCAP_GH_FASN markers show haplotypes with different frequencies in Wagyu and Simmental breeds.



★ These haplotypes are present in all cattle breeds tested

Bota Chr19 markers measured in 190 Wagyu and 290 Simmental cattle.

MRIP_TCAP_GH_FASN defined haplotypes also have different frequencies in Wagyu and Angus.



* These haplotypes are present in all cattle breeds tested

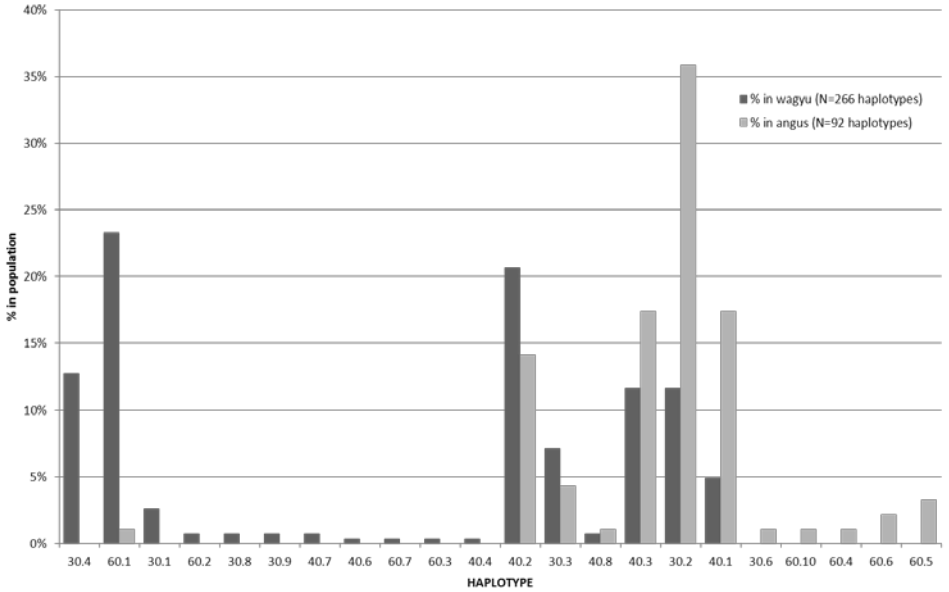
X 6010_AL is shared between Wagyu and Angus, but is not found in Simmental.

Chr19 markers measured in 190 Wagyu and 43 Angus cattle.

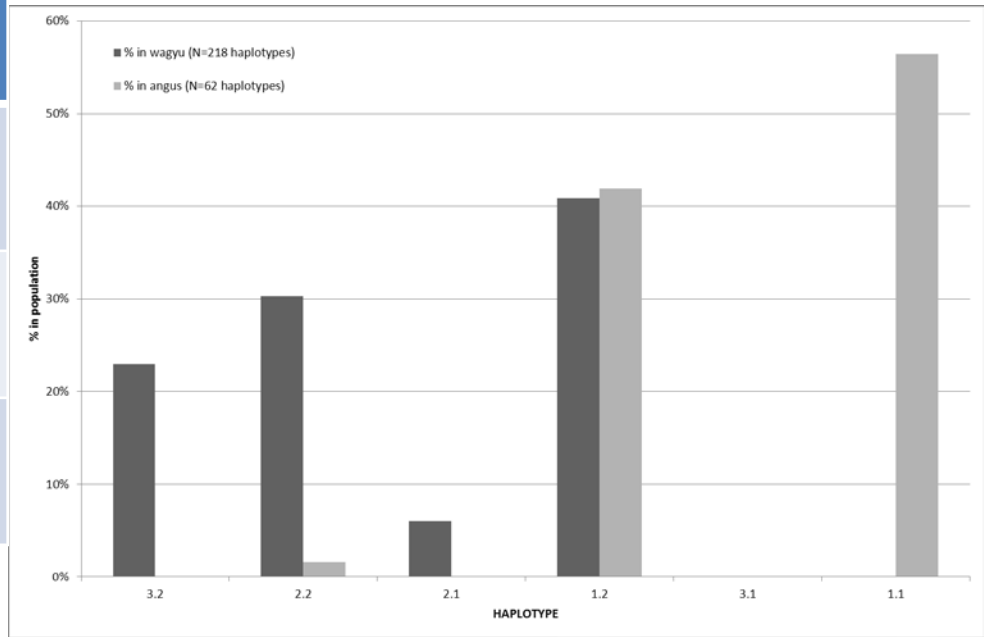
Test for Wagyu *v* Angus

At least one of:
MRIP 60.1, 30.4, 60.2, 30.8, 30.9,
SECTM 3.2, 2.2, 2.1

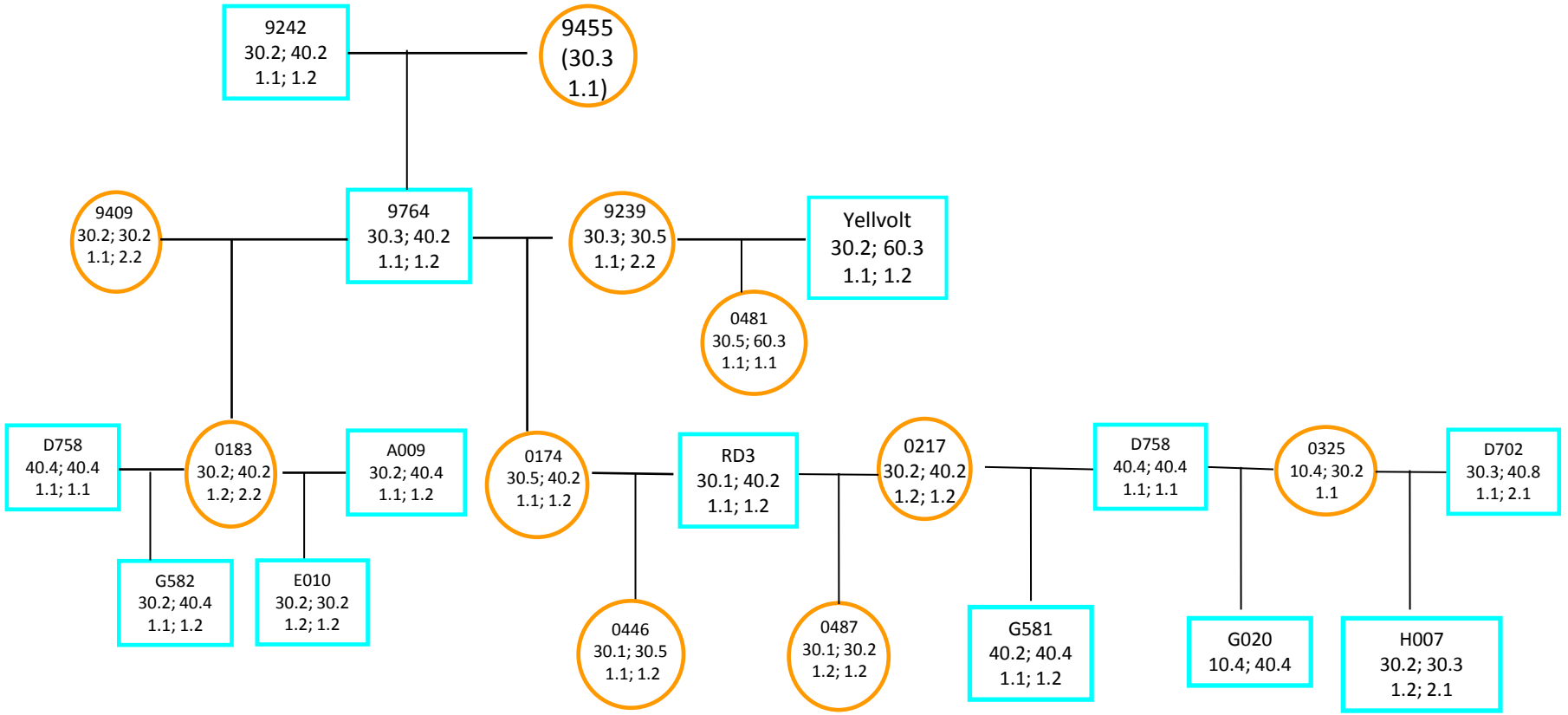
None of:
MRIP 60.5,60.6, 60.4, 60.10,30.6
SECTM 1.1



	WAGYU	ANGUS	Predictive value
Positive test	120	2	PPV=98%
Negative test	15	41	NPV=73%
TOTAL ANIMALS	135	43	



Confirm segregation in multi-generation families



Confirm Dam and assign Sire from mated bulls.

		SREB	NT5M	MRIP	TCAP	GH	FASN
J008	phenotype	LL	20 22	40 40	10 20	AB	SL
DAM A535	haplotype 1	L	20	30	10	A	S
	haplotype 2	L	20	40	20	B	S
SIRE 1 D702	haplotype 1	L	20	30	10	B	L
	haplotype 2	L	22	40	10	A	L
SIRE 2 D758	haplotype 1	S	20	40	20	A	S
	haplotype 2	S	20	40	20	A	S
J008	haplotype 1	L	20	40	20	B	S
	haplotype 2	L	22	40	10	A	L

Conclude D702 sire rather than D758 to match SREB and FASN

from haplotype 2 of A535

from haplotype 2 of D702

Testing at birth and point of sale can detect meat substitution or incorrect labelling

ID	SREB	NT5M	MRIP	TCAP	GH	FASN
Ear sample G593	SL	2022	4040	1020	AB	SL
Meat sample 171	SL	2022	4040	1020	AB	SL
Ear sample G572	LL	2020	3040	2020	AA	SL
Meat sample 172	LL	2020	3040	2020	AA	SL
Ear sample G529	LL	2020	3030	2020	AA	LL
Meat sample 173	LL	2020	3030	2020	AA	LL
Ear sample G596	LL	2020	3030	2020	AA	LL
Meat sample 174	LL	2020	3040	2020	AA	SL
Ear sample G518	LL	1022	3040	1020	AA	LL
Meat sample 175	LL	1022	3040	1020	AA	LL

Meat Sample 174 results do not match with original animal G596 indicating substitution or mislabelling has occurred.

What is Meant by HAPLOTYPING?

- Comparing and contrasting alternative sequences at polymorphic regions of the genome

Comparison of Genetic Markers

	MARKER		
	SNP	Sequencing	Haplotyping
Coverage	1 bp	500 bp	14 Mb
Inherited Polymorphism	Y/N	Y	Y
Reproducibility	N	Y/N	Y
Success in Humans	N	Y	Y
Cost	High	High	Low
Independence	N	N	Y

Lessons from genomic sequencing

- **Age and sex related penetrance** implying gene regulation of multiple pathways.
- Traits are carried by **extensive ancestral haplotypes**; a few monogenic diseases are exceptions.
- Important candidates include **indels, retro-elements** and **duplications** in addition to sequence polymorphism.
- Genetic **redundancy** permits advantages.
- **Four** Unlinked paralogous regions contribute to complex functions and survival.
- Expression, penetrance and **regulation** remain poorly understood but will be important biologically and commercially.

AN ALTERNATIVE STRATEGY FOR SELECTIVE BREEDING

1. Define polymorphic frozen blocks and their ancestral haplotypes in regions of interest.

2. Develop robust haplospecific markers.

3. Identify those haplotypes which are relatively frequent in breeds with desirable (or interesting) traits whilst excluding those which are common to multiple breeds.

4. Develop minimal requirements for reliable determination of breed and parentage.

5. Compare haplotype frequencies in elite versus poor performers within a breed, and in cross breeds, allowing for vagaries of penetrance.

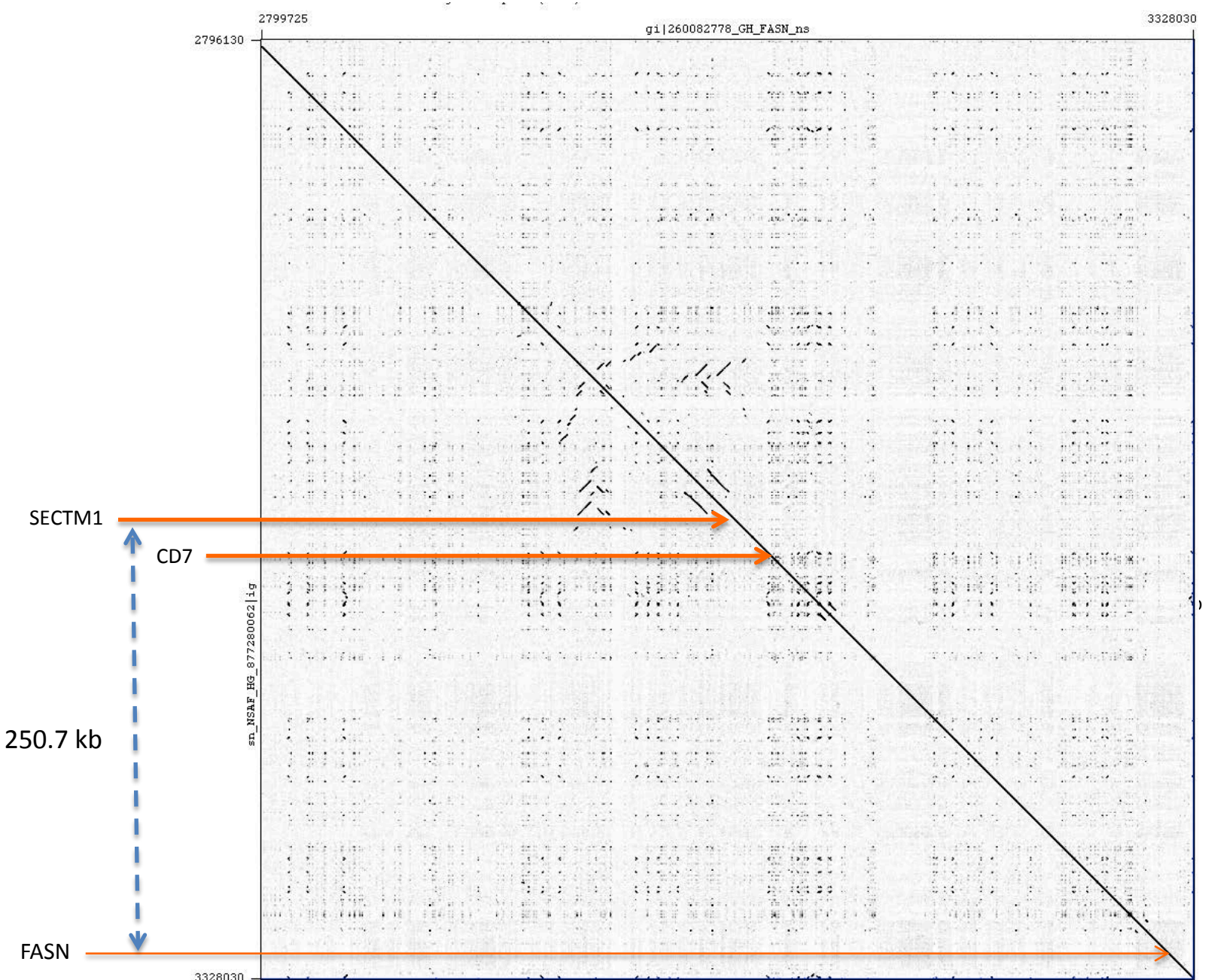
6. Compare candidate haplotypes to determine whether these have arisen by ancestral recombination events which can help to localise the operative components.

7. Consider functional explanations for observed haplotype associations.

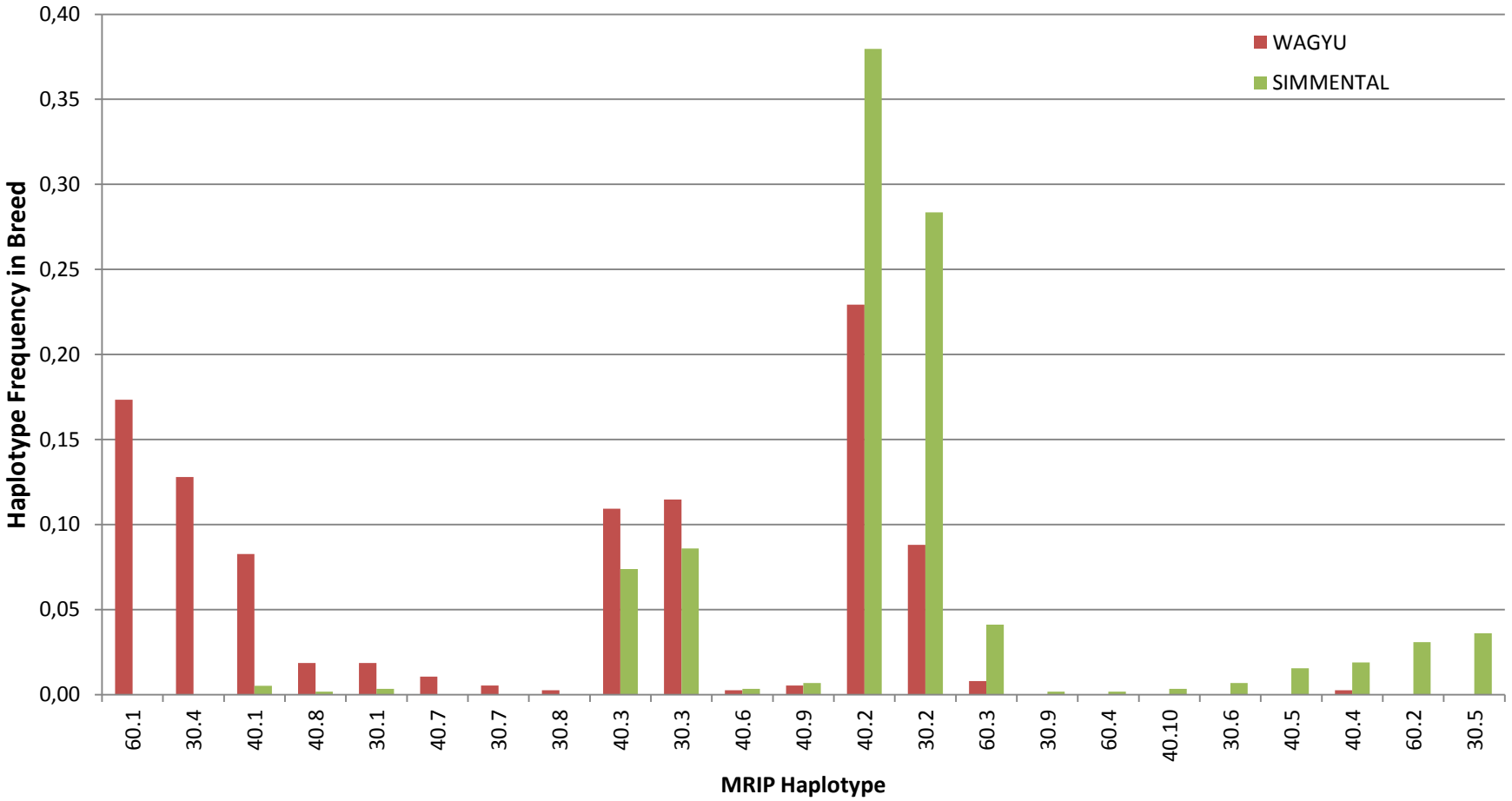
8. Test the hypotheses of specific haplotype associations and interactions using AI, ET and Cloning experiments.

9. Use Marker Assisted Selection to blend preferred haplotypes into already successful herds, or to increase the frequency of haplotypes already present.

A duplicated region containing SECTM1 is close to FASN



Haplotypes common in Wagyu but not found in Simmental are likely candidates for associations with high marbling and low melting point fat.



- Conclusions
  - initially based on MHC – Chr6
  - most polymorphic region of the genome
- Principles apply to
  - RCA – Chr1
  - Duplotyping* (21 chromosomes)
    - Polymorphism >>>>> Expected

Assumption – Same Rules will Apply to Cattle

* McLure *et al* Genomics 102(2013):15-26

Duplotyping
Degree of Polymorphism greater than expected

Name (CYO_)	Region	Total # Products
3_2	3p21-cen to 3q11.2	58
6_5	6p22	57
1_3	1p36.22	55
6_1	6p11.2	52
6_6	6p22	52
22_3	22q11.23	45
8_3	8p23.1	43
10_7	10q22.3	43
1_13	1p36	42
19_1	19q13.2	40
Y_9	Yq11.23	39
10_6	10q22.3	38
1_6	1p21	35
4_3	4q28-q31	35
X_13	Xq28	34
1_11	1q21-q23	33
9_4	9q34	32
10_2	10p11.2	32
22_4	22q13.1	32
Y_11	Yq11.23	31
X_5	Xp11.22	30

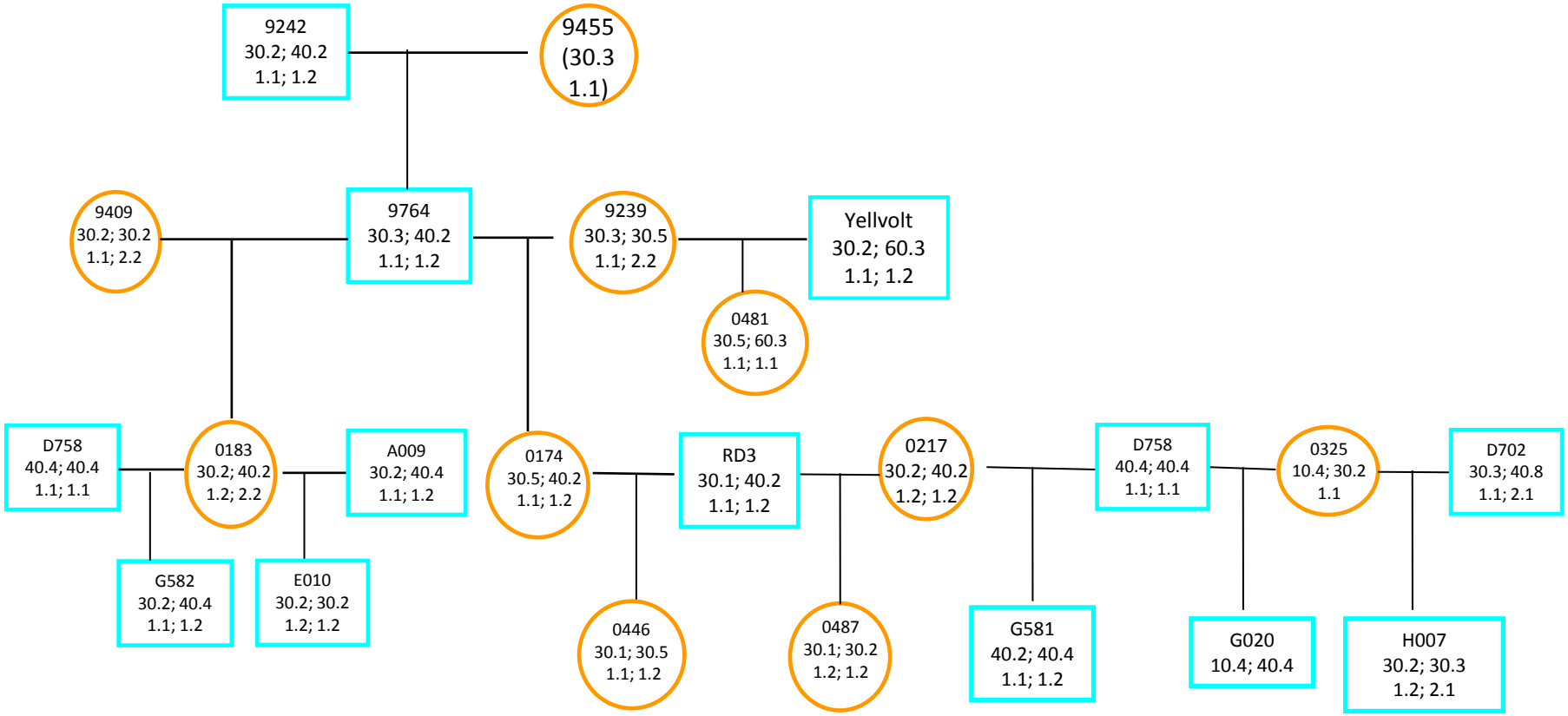
* McLure *et al* Genomics 102(2013):15-26

Duplotyping
Variable polymorphism seen in other species

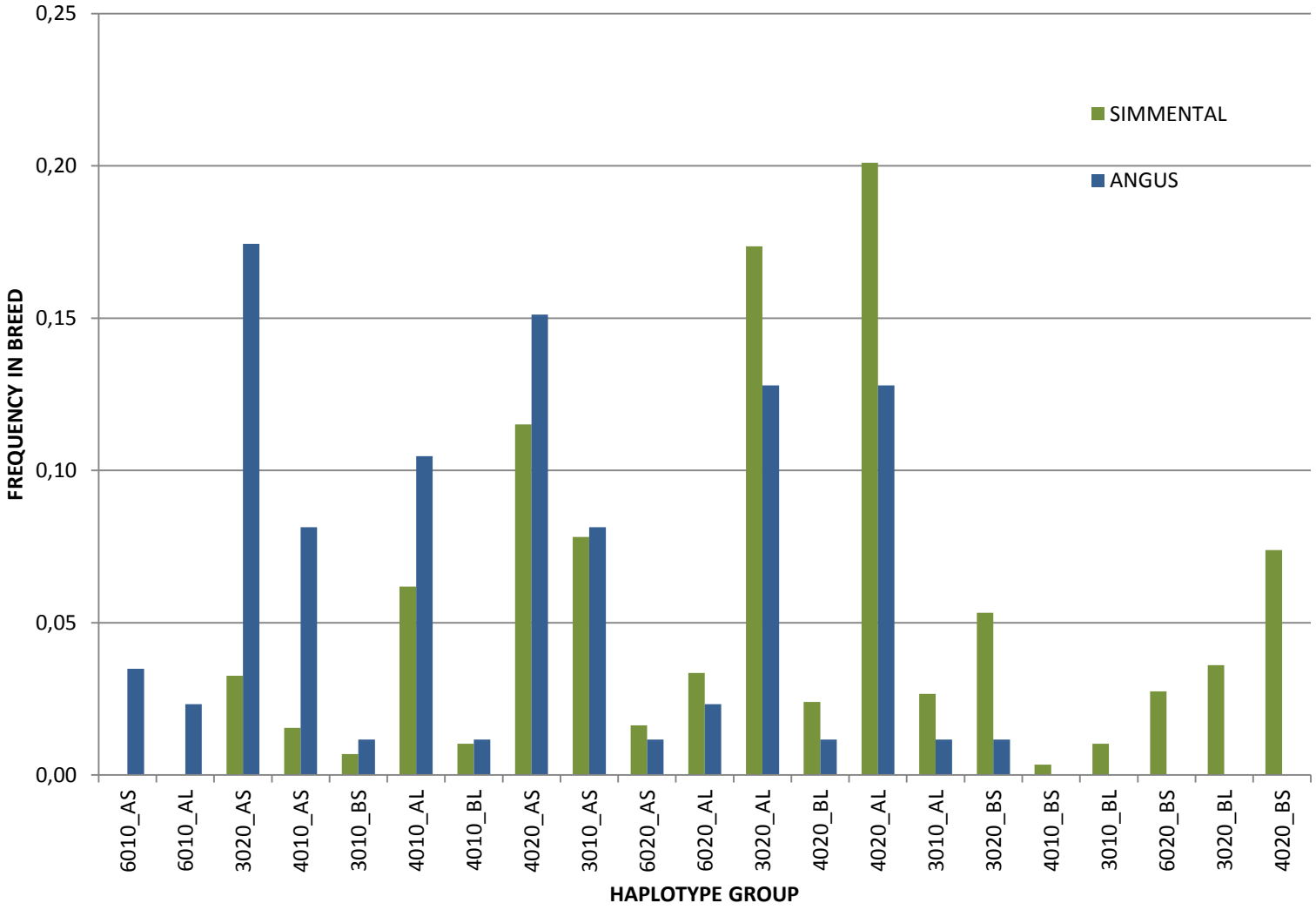
Species	Common Ancestor (MYA)	CYO_8_1	CYO_10_4	CYO_9_4	CYO_9_3	CYO_Y_8	CYO_Y_4	CYO_17_5	CYO_X_14	CYO_6_3	CYO_6_4	CYO_5_1	CYO_Y_10	CYO_12_2	CYO_3_3	CYO_19_1	CYO_Y_9	CYO_Y_11	CYO_10_3	CYO_10_7	CYO_22_4	CYO_5_2	CYO_12_1	CYO_22_3	CYO_2_8	CYO_15_3	CYO_3_2	CYO_Y_7	CYO_10_6	CYO_X_15	CYO_2_5	CYO_Y_5
Human	0	8	9	**>10**	**>10**	5	6	5	5	2	**>10**	6	7	7	6	9	10	**<10**	**>10**	**>10**	6	5	5	**>10**	7	3	4	5	9	7	8	4
Chimp	5	5	8	**>10**	**>10**	6	3	5	8	3	7	3	7	4	4	9	7	8	**>10**	**>10**	4	5	2	8	3	8	4	3	4	4	3	4
Orang	20	6	7	**>10**	**>10**	1	2	3	4		6	2	1	4	4	4	9	1	3	**>10**	1	1	1		2	3	2	1	3	2	2	
Macaque	30-40	1	3	**>10**	**>10**	5	6	10	8	1	**>10**	3	1	5	3	3	5	6	8	6	1	3	3	3	2	5	2	2	5	6	1	
Mouse	90	4	5	8	5	1	5	7	**>10**	**>10**	8			3	1	5	5	3								5						
Rat	90	5	2	**<10**	3	2	2	6	3	1	2	1	1																			
Dog	140	4	**>10**	**>10**	**>10**	4	3	4	2					5			5	2	7	3	2	2	3	1	1	1						2
Horse	140	8	**>10**	**>10**	**>10**	5	5	4	8	4	4	3	1	3	2	3	**>10**		**>10**	5	3	1	1	5	1	3	1	1				1
Cow	140	8	8	**>10**	**>10**	3	2	5	8	3	4	2		1		1	6	1	**>10**	5	3		3	1					1	1		
Sheep	140	7	8	**>10**	**>10**	4	3	3	4	6	6	3		2	1		2	1	9	3	1	1	1						1			
Chicken	220	**>10**	9	7	**>10**	7	4	4	4	2	4	1	1	1	1	1	5		3	2		2										
Budgie	220	**>10**	8	9	**>10**	5	1	2	6	4	3			4	5	4	6	1	4			1										
Snake	220	7	6	8	4	1				1		1	1					3			1											
Axolotil	350	**>10**	3				8																									
Danio	365	2		1														1														
Bee	630	2	5	2															3													
Marron	630	10	3		1																											

* McLure *et al* Genomics 102(2013):15-26

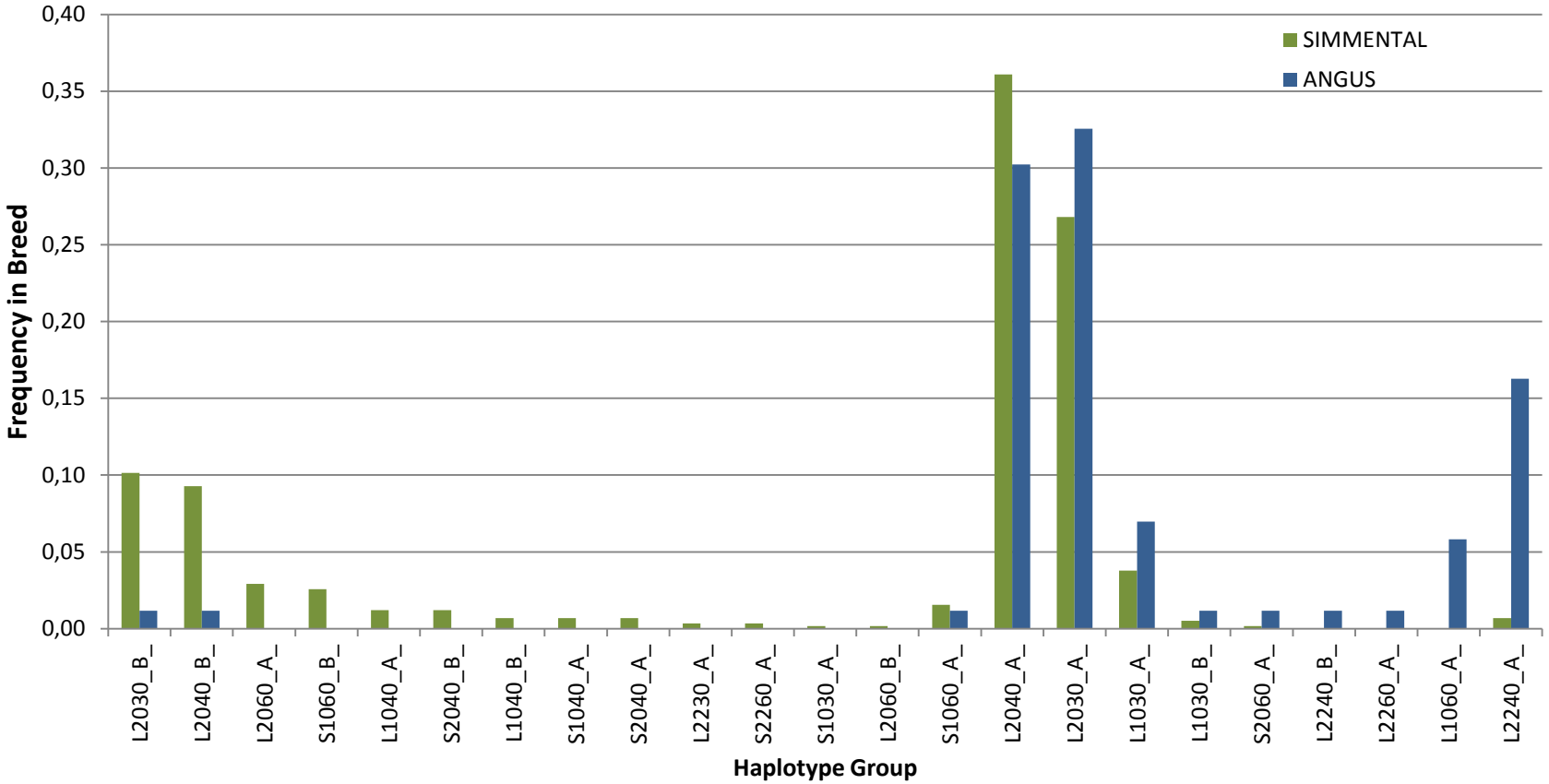
Segregation of MRIP and SECTM haplotypes in 3-generation Bovine families



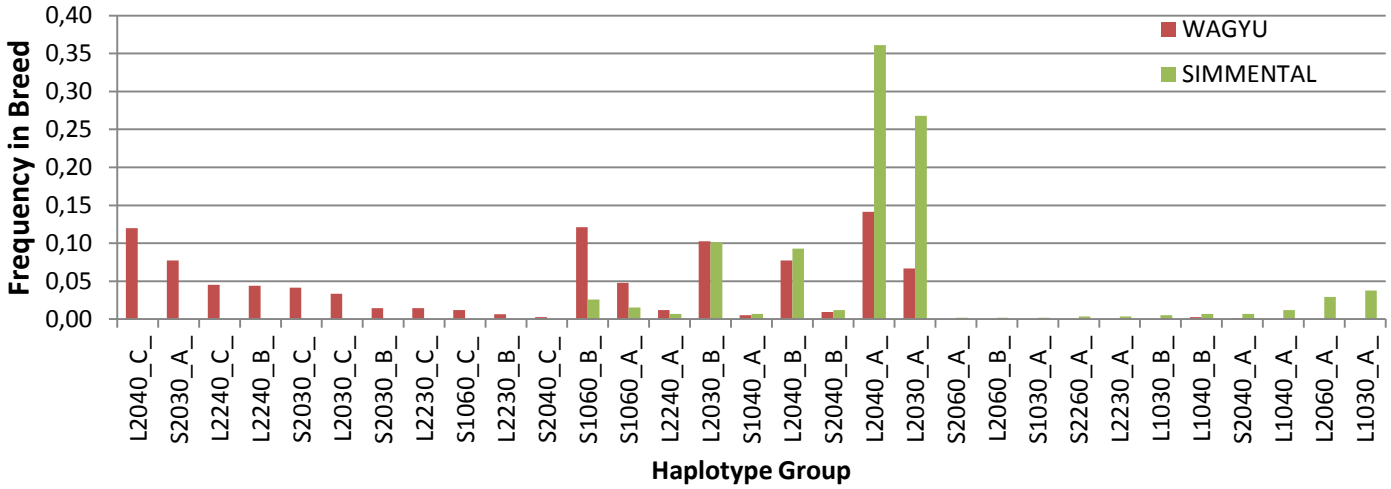
There is much more overlap between haplotypes frequencies in Simmental and Angus.

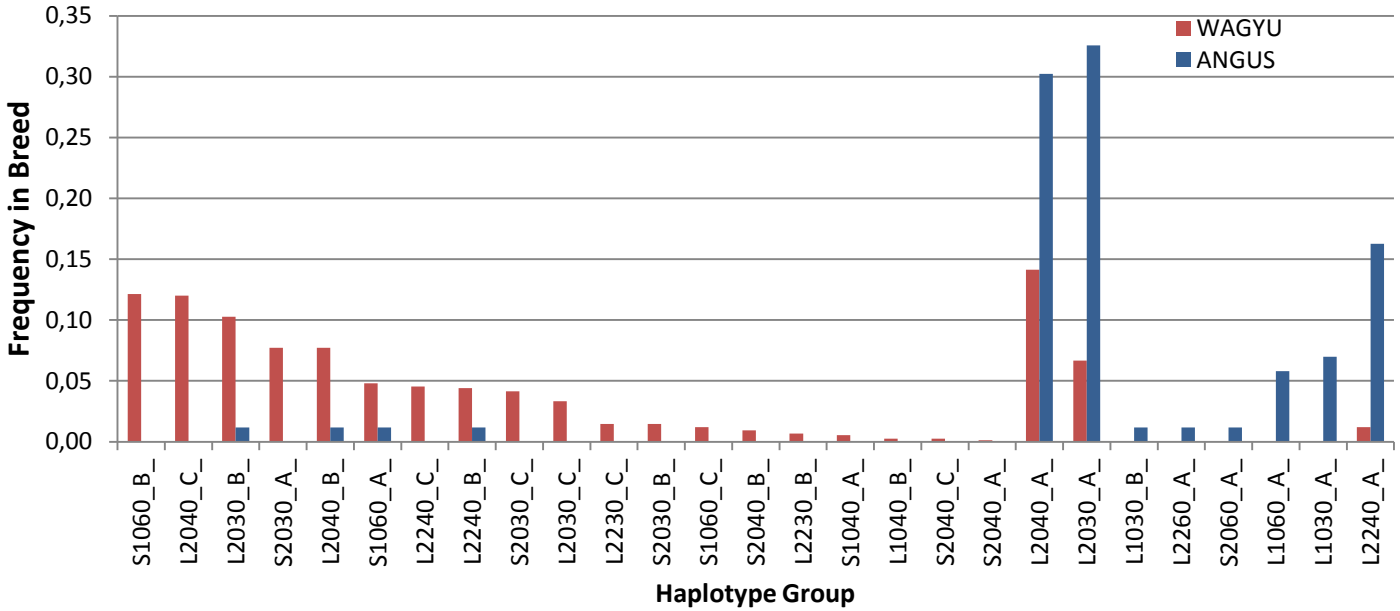


Including SREB and NT5M markers improve the ability to distinguish Simmental from Angus



But SREB and NT5M markers provide less additional information for distinguishing Wagyu from Angus and Simmental





SUMMARY

1. Polymorphic frozen blocks are reservoirs of clinically important genes.
2. Often sex dependent and appear during adult life.
3. Relate to degrees and subsets.
4. Different mechanisms must be encoded within haplotypes.
5. Polymorphic blocks contains regulatory genes.
6. Changes during vertebrate evolution relevant.
7. Gene copy number will be important.
8. Indels are very important candidates.
9. SNPs unsuccessful in explaining function.
10. Relevant genes in paralogous regions.
11. Sequencing multiple haplotypes will reveal mechanisms.
12. Individuality generated through meiotic cross-over not mutation.
13. Unit of inheritance is a haplotype

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes.

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant
7. Gene copy number will be important

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant
7. Gene copy number will be important
8. Indels are very important candidates

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant
7. Gene copy number will be important
8. Indels are very important candidates
9. SNPs unsuccessful in explaining function

Lessons from genetic sequencing in humans are:

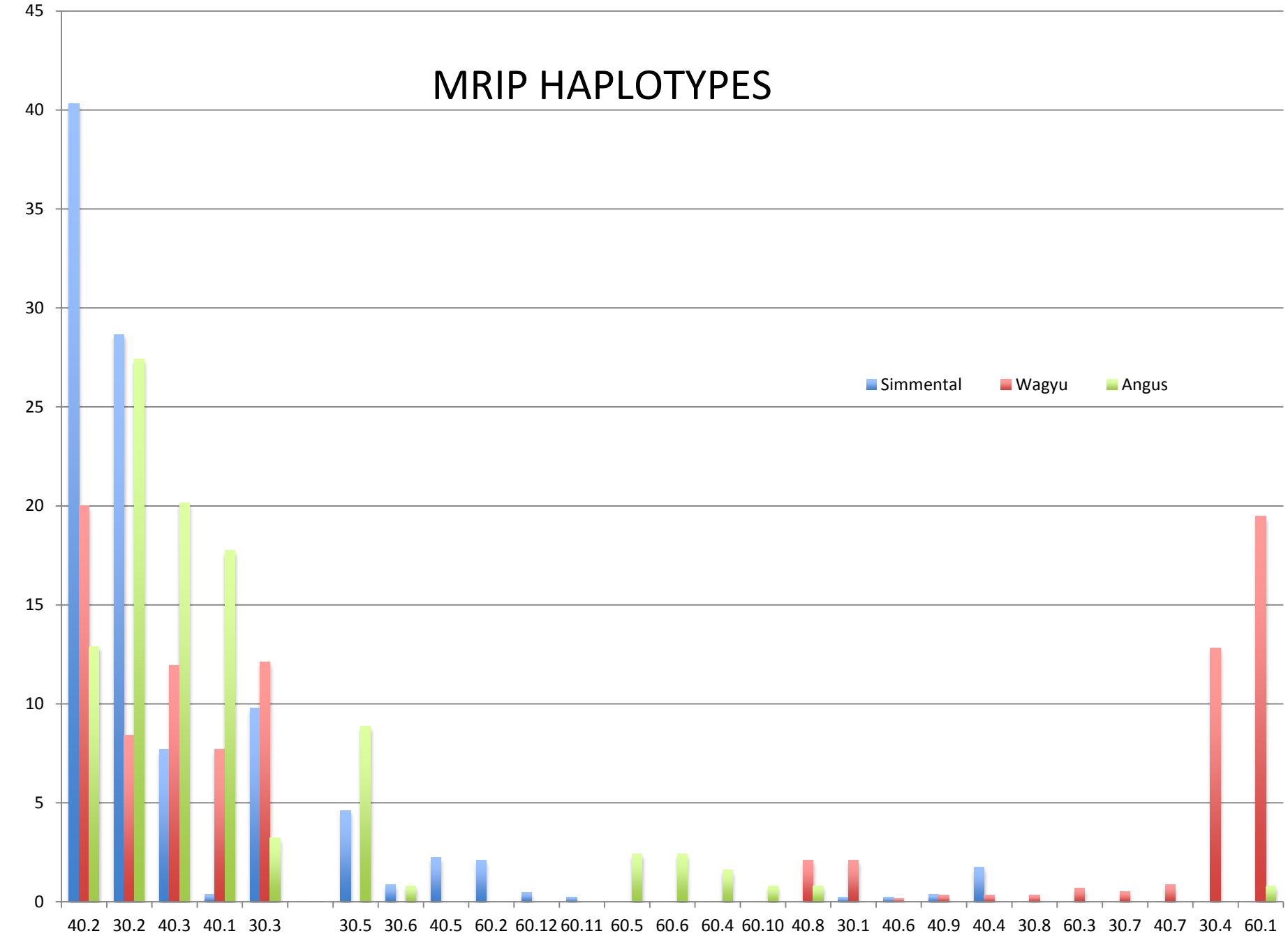
1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant
7. Gene copy number will be important
8. Indels are very important candidates
9. SNPs unsuccessful in explaining function
10. Relevant genes in paralogous regions

Lessons from genetic sequencing in humans are:

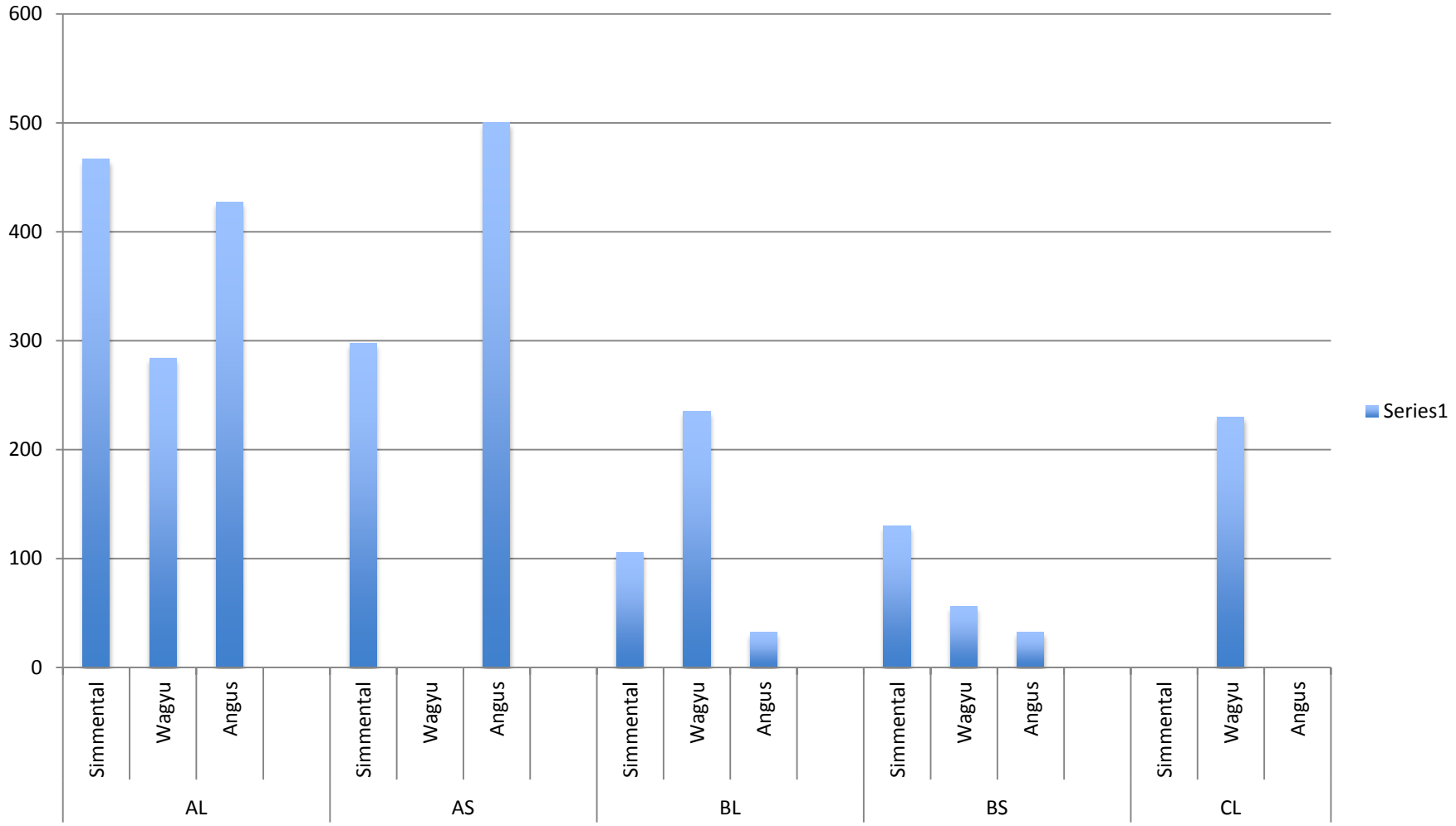
1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant
7. Gene copy number will be important
8. Indels are very important candidates
9. SNPs unsuccessful in explaining function
10. Relevant genes in paralogous regions
11. Sequencing multiple haplotypes will reveal mechanisms

Lessons from genetic sequencing in humans are:

1. Polymorphic frozen blocks are reservoirs of clinically important genes
2. Often sex dependent and appear during adult life
3. Relate to degrees and subsets
4. Different mechanisms must be encoded within haplotypes
5. The region contains regulatory genes
6. Changes during vertebrate evolution relevant
7. Gene copy number will be important
8. Indels are very important candidates
9. SNPs unsuccessful in explaining function
10. Relevant genes in paralogous regions
11. Sequencing multiple haplotypes will reveal mechanisms
12. Individuality generated through meiotic cross-over not mutation


MRIP HAPLOTYPES
Simmental
Wagyu
Angus
45
40
35
30
25
20
15
10
5
0
40.2
30.2
40.3
40.1
30.3
30.5
30.6
40.5
60.2
60.12
60.11
60.5
60.6
60.4
60.10
40.8
30.1
40.6
40.9
40.4
30.8
60.3
30.7
40.7
30.4
60.1

SECTM HAPLOTYPE FREQUENCIES



A single pair of SECTM primers bind in duplicated segments