IDENTIFICATION OF COMMON HAPLOTYPES FOR FINE SCALE MAPPING OF QTL FOR RETAIL BEEF YIELD

Y. Li, B. Harrison, R. Bunch and W. Barendse
CRC for Cattle and Beef Quality
CSIRO Livestock Industries, Queensland Biosciences Precinct, 306 Carmody Rd, St Lucia, QID 4067, Australia

SUMMARY
Haplotype information is an essential ingredient in many analyses of fine-scale mapping of QTL. Various methods are being used for identification of haplotypes. When genotyping methods do not provide phase information, one method that can be used to infer phase is to reconstruct haplotypes by choosing the most probable haplotype assignment, given the genotype data and the estimated population haplotype frequencies. We applied this method to fine scale mapping of QTL for retail beef yield in Beef CRC populations. It proved to be an efficient method to identify the QTL region affecting the quantitative trait.
Keywords: common haplotype, QTL, retail beef yield, Beef CRC population.

INTRODUCTION
Haplotype information is an essential ingredient in many analyses of fine-scale mapping of QTL, such as for resistance to disease (Risch and Merikangas 1996), milk quality (McPartlan *et al.* 2001) and growth traits (Li *et al.* 2002). These information will greatly facilitate the identification and cloning of the causative genes. It is expected some common haplotypes originating from common ancestors may carry on and segregate among individuals of a breeding line, particularly when selection is applied. Our focus here is population data where genotyping methods do not provide phase information due to lack of parental genotypes. One statistical method that can be used to infer phase is to reconstruct haplotypes by choosing the most probable haplotype assignment given the genotype data and the estimated population haplotype frequencies (Excoffier and Slatkin 1995). This paper reports on the value of using that method to identify common haplotypes for retail beef yield using the Beef CRC animals.

MATERIALS AND METHODS
Animals. Two groups of beef cattle consisting of temperate (Angus, Hereford and Shorthorn) and tropically adapted (Brahman, Belmont Red and Santa Gertrudis) breeds were used for the study. They were chosen from the CRC DNA bank. Information stored in the CRC database was used to select animals across a range of 3 purchasing markets and 4 finishing regimes (see Table 1). The first group (260 individuals) comprised animals of extremes (high and low) of retail beef yield (ADJRBY). In essence, the procedure was to select cattle in each cohort which were of extreme phenotypes, ensuring that no sire was represented by a cluster of offspring, that all markets and finishing regimes were included in each extreme, so that extremes were not biased by being representative of a particular market or finishing regime. The second group (528 individuals) comprised the first group (extreme) and an addition of 268 non-extremes animals. The second group was chosen to be

representative of across-the-range population. These two groups form the base for the identification of common haplotypes (Table 1).

Table 1. Part I: Number of sires and contemporary groups within each population. Part II: Number of animals in each breed/finish/market class

	Effect	Extreme	Combined
Total		260	528
Part I	Sires	87	110
	Contemporary Groups[A]	151	183
Part II	Breed		
	Angus	37	78
	Belmont Red	39	92
	Brahman	43	96
	Hereford	50	95
	Santa Gertrudis	51	94
	Shorthorn	31	82
	Finish		
	Pasture South	62	135
	Pasture North	46	107
	Grain South	96	163
	Grain North	56	123
	Market		
	Domestic	106	199
	Korean	102	215
	Japanese	52	114

[A] Contemporary group was defined as the combination of herd of origin, cohort and kill code.

Genotyping and haplotype identification. A single chromosome (anonymous for IP protection) with 9 DNA microsatellite makers was used for this study. The chromosome was chosen for fine-scale mapping because of the presence of a QTL for retail beef yield identified in the CBX experiments (Hetzel *et al.*, 1997) and further confirmed by the Beef CRC marker evaluation project (Li *et al.*, unpublished). Five hundred and twenty-eight animals were genotyped using 9 microsatellite markers. The haplotypes (allele linkage phases) of the animals were established according to the orders of linked markers from public maps i.e. haplotypes were reconstructed by choosing the most probable haplotype assignment, given the genotype data and the estimated population haplotype frequencies. The SAS program (Version 8.2, SAS Inst. Inc., Cary, NC) was used to derive the frequencies of pair-wise marker alleles.

Statistical Analysis. Analyses were performed between the most commonly observed haplotypes and retail beef yield (ADJRBY) using the SAS mixed model procedure, where the difference between animals with and without haplotypes was tested. A complete dominance effect of the haplotype was assumed, in which animals carrying either one or two copies of the haplotype were treated the same. Fixed effects in the model included finish and haplotype type. Contemporary groups were treated as a random effect. The statistical model was Trait = mean + contemporary group + haplotype + finish +

carcass weight within market endpoint. Contemporary group was defined as the combination of herd of origin, cohort and kill code. Carcass weight within market endpoint (Japanese, Korean, domestic) was used as a covariate to adjust for differences in weight and to a lesser extent, age effects. Since breed was confounded with contemporary group, it was not independently fitted in the model. All effects but haplotype were nested within breed. Haplotype type was defined as 1 when the individual had the haplotype or 0 when the individual was without the haplotype. When the most common haplotype could not be determined between two adjacent loci due to similar frequencies of two haplotypes, the haplotype type was then defined as 1, 2 or 0 (i.e. two common haplotypes for the adjacent loci). Summary statistics of the trait are presented in Table 2.

Table 2. Summary statistics of retail beef yield in both populations

Population and trait	Mean	Range	SD
Extreme population			
ADJRBY (%)	66.87	56 – 77.16	4.79
Combined population			
ADJRBY (%)	66.72	55 – 77.16	4.47

RESULTS AND DISCUSSION

Identification of common haplotypes. The number of alleles in each marker is shown in Table 3. On average, 12.3 alleles were detected for each locus of the chromosome in the extreme population, with a range of 5 to 25 alleles per locus. The average was slightly higher in the combined population (12.9). In both populations, the phase of the most common haplotypes was readily determined for the first six markers based on the frequencies of alleles at adjacent loci along the chromosome. However, there were difficulties with the last three loci due to the similar frequencies of two haplotypes. The extra two rare alleles in markers 6 and 7 were not the cause of the difficulty. Therefore, two common haplotypes were assigned as 1 and 2 for analysis in these three loci (see Table 4).

Association between a haplotype and ADJRBY. Associations between the most common haplotypes, which were exclusively haplotypes of adjacent loci, and ADJRBY are shown in Table 4. In both populations, a consistent significant effect was identified for the common haplotype 121-153 within markers 1 and 2 ($P<0.05$). The haplotype had a very high frequency of 40% in both populations relative to the other haplotypes (ranging from 9% to 40%). It explained 33% of trait variation in terms of standard deviation. There was no significant effect due to the other haplotypes. Despite of the lack of parental genotype information, the identification of the common haplotype has further confirmed the existence of a QTL for ADJRBY on the chromosome identified by the CBX QTL experiment (Hetzel *et al.*, 1997) and the CRC marker evaluation project. It will provide useful information for further characterization of the gene(s) of interest in the region. Haplotypes can generally be identified using the identity by descent method with the genotype information available at least from the sires of animals. However, with commercial populations, where sires or dams may not be genotyped due to the cost, these results demonstrate that common haplytypes can be used to detect potential QTL.

Table 3. Number of alleles from each microsatellite marker in both populations

Marker	Extreme Population	Combined Population
M1	9	9
M2	8	9
M3	9	9
M4	25	25
M5	19	19
M6	5	7
M7	9	11
M8	13	13
M9	14	14
Average	12.3	12.9

Table 4. Association between haplotypes and ADJRBY in two populations. Two common haplotypes were assigned for markers 7, 8 and 9

	Extreme			Combined		
Haplotype[A]	Freq.	P-value	Effect ± S.D	Freq.	P-value	Effect ± S.D
M1-121, M2-153	0.43	0.046*	1.56 ± 0.73	0.40	0.024*	1.49± 0.629
M2-153, M3-114	0.48	0.068	1.31± 0.68	0.40	0.19	0.798± 0.595
M3-114, M4-193	0.25	0.57	-0.43 ± 0.75	0.21	0.67	-0.272 ± 0.629
M4-193, M5-187	0.11	0.39	-0.92 ± 1.03	0.093	0.48	0.619 ± 0.864
M5-187, M6-163	0.27	0.088	-1.27 ± 0.70	0.21	0.40	0.544 ± 0.643
M6-163, M7-100	0.20	0.67	0.635 ± 0.827	0.17	0.64	0.569 ± 0.708
M6-163, M7-104	0.26		0.50 ± 0.753	0.22		0.458 ± 0.664
M7-100, M8-155	0.15	0.33	0.0462± 0.958	0.13	0.31	-0.137± 0.802
M7-104, M8-141	0.18		1.25± 0.816	0.15		1.12± 0.739
M8-155, M9-181	0.16	0.67	-0.765± 0.932	0.18	0.64	-0.709± 0.779
M8-141, M9-189	0.23		-0.443± 0.784	0.21		-0.334± 0.691

[A] The haplotypes were named by two alleles of a pair of loci. For example M1-121, M2-153 represents a segment of chromosome having allele 121 of M1 and allele 153 of M2.
* $P < 0.05$

REFERENCES

Excoffier, L. and Slatkin, M. (1995) *Mol. Biol. Evol.* **12**:921.

Hetzel, D.J.S., Davis, G.P., Corbet, N.J., Shorthose, W.R., Stark, J., Kuypers, R., Scacheri, S., Mayne, C., Stevenson, R., Moore, S.S. and Byrne, K. (1997) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **12**:442.

Li, C., Basarab, J., Snelling, W.M., Benkel, B., Murdoch, B. and Moore, S.S. (2002) *J. Anim. Sci.* **80**:187.

McPartlan, H., Chamberlain, A., Balasingham, T., Robinson, N., Carrick, M., Bowman, P. and Goddard, M.E. (2001) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **14:**95.

Risch, N. and Merikangas, K. (1996) *Science* **273**:1516.