

Linkage disequilibrium on chromosome 6 in Australian Holstein-Friesian cattle

Mehar S. KHATKAR*, Peter C. THOMSON, Imke TAMMEN,
Julie A.L. CAVANAGH, Frank W. NICHOLAS, Herman W. RAADSMA

Centre for Advanced Technologies in Animal Genetics and Reproduction (ReproGen), and
Co-operative Research Centre for Innovative Dairy Products, Faculty of Veterinary Science,
University of Sydney PMB 3, Camden NSW 2570, Australia

(Received 17 June 2005; accepted 20 April 2006)

Abstract – We analysed linkage disequilibrium (LD) in Australian Holstein-Friesian cattle by genotyping a sample of 45 bulls for 15 closely-spaced microsatellites on two regions of BTA6 reported to carry important QTL for dairy traits. The order and distance of markers were based on the USDA-MARC linkage map. Frequencies of haplotypes were estimated using the E-M approach and a more computationally-intensive Bayesian approach as implemented in PHASE. LD was then estimated using the Hedrick multiallelic extension of Lewontin normalised coefficient D' . Estimates of D' from the two approaches were in close agreement ($r = 0.91$). The mean estimates of D' for marker pairs with an inter-marker distance of less than 5 cM ($n = 13$) are 0.57 and 0.51, and for distances more than 20 cM ($n = 44$) are 0.29 and 0.17, estimated from the E-M and Bayesian approaches, respectively. The Malecot model was fitted for the exponential decline of LD with map distance between markers. The swept radii (the distance at which LD has declined to $1/e$ (~37%) of its initial value) are 11.6 and 13.7 cM for the above two methods, respectively. The Malecot model was also fitted using map distance in Mb from the bovine integrated map (bovine location database, bLDB) in addition to cM from the MARC map. Overall, the results indicate a high level of LD on chromosome 6 in Australian dairy cattle.

linkage disequilibrium / dairy cattle / markers / chromosome 6

1. INTRODUCTION

Linkage disequilibrium, or non-random association between alleles at different loci within a population, has recently become important in the context of gene mapping, especially with the availability of high-density maps and reports of extensive LD in various species [5, 17].

The potential advantage of LD mapping over conventional linkage mapping is in utilising historic meioses which provide higher mapping resolution [36].

* Corresponding author: MeharK@camden.usyd.edu.au

LD analysis has proved powerful for high-resolution mapping of genes for monogenic disorders [5] and complex diseases in humans [5, 18]. In dairy cattle, successful efforts have been made towards positional cloning of causal mutations affecting fat percent and protein percent using combined linkage and LD mapping approaches [4, 11]. However, this type of analysis is limited to specific types of resource populations and is not practicable for novel traits measured only in groups of animals with no family structure. With increasing numbers of single-nucleotide polymorphism (SNP) markers becoming available, it is now possible to conduct genome-wide LD-mapping in populations with no family structure [16].

Recent simulation studies and reviews have suggested that optimum designs for such LD mapping studies depend mainly on the extent and pattern of LD across the genome in the studied population [1, 48]. Marker-marker LD provides a benchmark for evaluating optimal marker density. Hence, a number of studies have attempted to describe the extent of inter-marker LD in various species/populations, and have investigated the relationship between LD and inter-marker map distance. Some of these are described below.

The extent of LD has been studied most thoroughly in humans, where substantial LD extends over physical distances up to 100 kb in several populations [3, 16, 39], and up to several megabases in isolated populations [3, 13, 29, 38]. The extent of LD in livestock is expected to be higher than in humans, since the forces producing LD (admixture, selection and small effective population sizes) are more extreme in livestock populations [33]. The evidence supporting this expectation is accumulating: high levels of LD have been found to extend for many centimorgans in dairy cattle [10], sheep [28], pigs [34] and horses [43]. These studies used family information to infer the most likely phase of the dams' haplotypes. Tenesa *et al.* [42] reported some evidence of LD on BTA6 in a sample of unrelated United Kingdom dairy cattle. Sutter *et al.* [41] used several samples of unrelated animals to estimate LD in a range of dog breeds. Studies in humans, dogs and pigs indicate that the pattern and extent of the LD is often population-specific [19, 34, 37]. This emphasises that estimates of LD in specific populations are required before carrying out LD-based fine-mapping studies in those populations.

Chromosome 6 (BTA6) has been reported to carry important QTL affecting milk production traits. In this study we explored the extent of LD in Australian Holstein-Friesian cattle in the two important regions of chromosome 6 known to carry QTL affecting milk production in dairy cattle. To estimate LD, we used algorithms that do not rely on family information for inferring haplotypes. The animals in the samples are representative of the current Holstein-Friesian

population in Australia and thus provide an indication of the extent of LD in this population for chromosome 6.

2. MATERIALS AND METHODS

A panel of 45 progeny-tested Australian Holstein-Friesian bulls born during 1989–1999 was selected for this study. Using pedigree information obtained from the Australian Dairy Herd Improvement Scheme (ADHIS; <http://www.adhis.com.au>), these 45 bulls were selected from all the bulls in these cohorts, on the basis of least relationship, as assessed by the coefficient of co-ancestry. To represent the population better, the bulls with more daughters were also preferentially selected. The bulls were considered unrelated for the purpose of this study. The mean kinship between these bulls is 0.009 with first and third quartile for kinship being 0 and 0.012, respectively. The kinships between these bulls were computed using FORTRAN programs in the PEDIG package of D. Boichard (<http://dga.jouy.inra.fr/sgqa/diffusions/pedig/pedigE.htm>).

2.1. Extraction and amplification of DNA

Semen samples for the bulls were obtained from Genetics Australia (Bacchus Marsh, Vic, Australia). Genomic DNA was extracted from straws of frozen semen samples by a salting-out method adapted from Heyen *et al.* [15]. The quantities of original DNA available were very small. Hence all DNA samples were amplified genome-wide using a GenomiPhiTM DNA Amplification Kit (Amersham Biosciences, Castle Hill, NSW, Australia). Un-amplified and amplified DNA from five of the bulls was genotyped for nine microsatellites and the results were identical. All subsequent genotyping on this panel was done on amplified DNA.

2.2. Markers and genotyping

Fifteen BTA6 microsatellites, mainly spanning two regions of interest, were selected for genotyping (Tab. I). These regions had previously been identified in a meta-analysis of published QTL [22], as being of potential interest. The average inter-marker distance is 4.6 cM. All the distances are based on the MARC-USDA linkage map (<http://www.marc.usda.gov> – accessed 10 January, 2005). Since the genetic distance between close markers may not be precise,

Table I. Marker name, BTA6 chromosomal location in the MARC map and in the integrated bovine map (bLDB), number of alleles, observed heterozygosity and expected heterozygosity under Hardy-Weinberg equilibrium (HWE), and significance level (*P*) of the test for departure from HWE¹.

Marker Name	MARC Position (cM)	bLDB Position (kb)	Alleles	Observed Heterozygosity	Expected Heterozygosity	<i>P</i> -value
BM1329	35.4	15926	8	75.6	76.5	0.551
BMS1242	52.8	40013	7	75.0	79.3	0.225
BM143	53.7	38245	6	24.4	69.4	0.000
JMP36	56.1	44638	5	70.5	70.2	0.573
BMS690	56.4	44787	4	44.4	60.9	0.022
BL1099	58.4	47443	6	29.5	69.2	0.000
BM4322	63.9	61308	7	39.5	38.9	0.830
BM4621	77.6	80060	5	37.8	39.7	0.190
EL03	88.2	96668	6	71.8	66.9	0.227
MB062	89.4	96678	3	44.4	42.7	1.000
CSPS100	91.5	94535	3	41.9	42.4	0.864
BMS2460	93.5	101961	3	66.7	61.4	0.498
AFR227	97.0	103260	5	52.9	77.9	0.046
BM4311	97.7	103260	4	81.8	70.3	0.232
BMS511	99.3	103607	6	88.1	77.4	0.428

¹The web version of the Genepop program (<http://wbiomed.curtin.edu.au/genepop>) was used to conduct the Fisher exact test for Hardy-Weinberg equilibrium (HWE).

hence as an alternative, the distances were also calculated from the bovine Location DataBase (bLDB) (<http://medvet.angis.org.au/ldb/>; [32]) with distance expressed in kb. The bLDB is an integrated map constructed from all the major public-domain physical and genetic maps of cattle, using the LDB software of Morton *et al.* [30], which had previously been used for construction of an integrated human map before completion of the human annotated sequence.

Genotyping was carried out using a LI-COR 4200 sequencer (LI-COR, Lincoln, USA) with fluorescent dye-labeled primers. Two duplicate DNA samples and one blank sample were included for quality control. Genotypic scoring was done using Gene ImagIR software (<http://www.licor.com/bio/GeneIm/GeneIm1.jsp>).

2.3. Testing for Hardy-Weinberg Equilibrium

The number of alleles, observed heterozygosity and expected heterozygosity under Hardy-Weinberg equilibrium (HWE) for each

microsatellite locus were computed using the web version of Genepop (<http://wbiomed.curtin.edu.au/genepop>). The P -values of the Fisher exact test for deviations from HWE were computed as per Guo and Thompson [12] using Genepop (Tab. I). The numbers of alleles in the MARC *versus* our study were compared using loglinear modelling (to cater for low-number counts), taking into account the paired nature of the data, *i.e.* two observations for each marker. The observed heterozygosity in MARC *versus* our study were compared by a paired t -test.

2.4. Estimating linkage disequilibrium between pairs of markers

A number of measures of LD have been suggested in the literature (Devlin and Risch [8]; Hedrick [14]; Jorde [21]; Lewontin [23]). We estimated LD by using the Hedrick [14] multiallelic extension of the Lewontin [23] normalised coefficient D' computed from two-loci haplotype frequencies. For a pair of loci A and B, D' was estimated as

$$D' = D'_{AB} = \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} p_{i \bullet} p_{\bullet j} |D'_{ij}|$$

where $D'_{ij} = D_{ij} / D_{ij, \max}$, $D_{ij} = p_{ij} - p_{i \bullet} p_{\bullet j}$, and

$$D_{ij, \max} = \begin{cases} \min[p_{i \bullet} p_{\bullet j}, (1 - p_{i \bullet})(1 - p_{\bullet j})] & \text{if } D_{ij} < 0 \\ \min[(1 - p_{i \bullet})p_{\bullet j}, p_{i \bullet}(1 - p_{\bullet j})] & \text{if } D_{ij} > 0, \end{cases}$$

and where $p_{i \bullet}$ and $p_{\bullet j}$ are frequencies of alleles i and j at loci A and B, respectively; p_{ij} is the frequency of the pairwise haplotype ij ; and n_A and n_B are the total numbers of the alleles at loci A and B, respectively.

This statistic has been used to measure the extent of LD in human [3] and livestock populations in a range of studies (*e.g.* [10, 28, 34, 42]), thereby facilitating the comparison of the results in this study with other studies. D' allows LD to be measured with highly polymorphic markers and is less sensitive to variation in marker allele frequencies than other measures of LD [14, 46].

2.5. Haplotype frequency estimation

The haplotype frequencies were estimated using the GOLD-ldmax procedure (<http://www.sph.umich.edu/csg/abecasis/GOLD/>; [2]). This procedure

implements a maximum likelihood based expectation-maximisation (E-M) algorithm [9]. Haplotype frequencies were also estimated using a Bayesian algorithm as implemented in PHASE 2.0 [40]. These haplotype frequencies were then used to calculate D' values as detailed above. Chi-square statistics, to test for independence between alleles at two loci, were also computed for each pair of markers using GOLD. PHASE by default assume a stepwise mutation model for microsatellite markers and is appropriate when the length of each allele is known. However, in our data set the actual length of each microsatellite allele was not known. Hence this condition was relaxed to the parent independent model in which each allele has the same chance to mutate to any of the other alleles. The algorithm used in PHASE has been reported as being more reliable than the E-M approach in inferring haplotypes [35, 45].

2.6. Malecot model

There was a general trend of decreasing LD with increasing inter-marker distance, and to quantify this relationship for the current data sets, we fitted the Malecot model [6, 24] to the data. This takes the exponential-asymptotic form,

$$D'_i = (1 - L)M \exp(-\kappa d_i) + L + \varepsilon_i = \rho(d_i) + \varepsilon_i$$

where D'_i is the observed LD for the i th pair of syntenic loci distance d_i apart, L ($0 \leq L \leq 1$) is the residual association at large distances (*i.e.* between unlinked loci), M ($0 \leq M \leq 1$) is the proportion of alleles transmitted from founders (and so is 1 if alleles are monophyletic and less than 1 otherwise), κ (> 0) is the exponential decay rate of D' with physical distance d , and ε_i is the random error deviation for the i th pair of loci, with $\text{var}(\varepsilon_i) = \sigma^2$. $\rho(d)$ represents the expected value of D' at distance d . This declines from $\rho(0) = (1 - L)M + L$ at $d = 0$ to a level $\rho(\infty) = L$ for unlinked loci. Since smaller values of the exponential decay parameter κ are associated with more extensive LD, Morton *et al.* [31] proposed that $1/\kappa$ is a good measure of useful LD for mapping, naming this as the “swept radius”. This is the distance over which LD declines to e^{-1} ($\sim 37\%$) of its initial value. We have also adopted this measure in the current study. The 95 percent of the confidence interval of the swept radius was computed as $1/(\kappa \pm 1.96 \times \text{SE}(\kappa))$. The Malecot model was fitted using the `nls()` function in S-PLUS / R.

3. RESULTS

Marker names, positions on the MARC and bLDB maps, observed and expected heterozygosity, tests of departure from HWE, and the number of alleles in this data set are shown in Table I. The average distance between markers is 4.6 cM (ranged from 0.3 to 17.4 cM). The average number of alleles in this data set is significantly ($P < 0.001$) lower (5.2) as compared to MARC (7.9). The average observed heterozygosity in our study (56.3%) was not significantly different from that in the MARC mapping population (60.5%) ($P = 0.54$). The average observed heterozygosity is lower than the average expected heterozygosity in this study (62.9%) and two markers BM143 and BL1099 show significant deviation from HWE. These two markers were excluded for the further analysis of LD between markers.

D' values were estimated between all pairwise combinations of the thirteen markers, based on two different approaches to infer haplotypes. The mean D' for all pairs is 0.38 ± 0.019 and 0.27 ± 0.019 , with haplotypes inferred by E-M and Bayesian approaches, respectively. The mean estimates of D' for marker pairs with an inter-marker distance of less than 5 cM ($n = 13$) are 0.57 and 0.51, and for distances greater than 20 cM ($n = 44$) are 0.29 and 0.17, estimated from the E-M and Bayesian approaches, respectively. The individual values of D' based on haplotypes inferred by the two approaches were highly correlated ($r = 0.91$, see Fig. 1). The agreement between the two methods appeared greater for the higher values of D' , although in general values of D' calculated using GOLD tend to be higher than those calculated using PHASE.

The distribution on the D' over distance is presented in Figure 2A-D, in each case with the Malecot model fitted. The general decline in the D' over inter-marker distance was evident (Fig. 3). The inverse of the parameter κ of the Malecot model can be used to estimate the swept radius, a distance to which LD may be useful for mapping [26]. We estimated the swept radii as 11.6 (95% confidence interval (CI) 6.6 to 46.9) cM and 13.4 (CI 8.9 to 29.6) cM based on haplotypes inferred from the E-M and Bayesian approaches, respectively (Fig. 2A-B). The Malecot model was also fitted with distance in Mb from the bovine integrated map (bLDB), giving estimates of swept radii of 15.4 (CI 8.9 to 55.7) Mb and 17.9 (CI 12 to 35) Mb, for the two approaches, respectively (Fig. 2C-D). Note that the relatively wide CI were due to the small number of markers and bulls in this study. The parameter estimates from the Malecot model from the two methods and using the two types of map distances are presented in Table II. A steeper decline was observed in the chi-square values when plotted over distance between markers (Fig. 4).

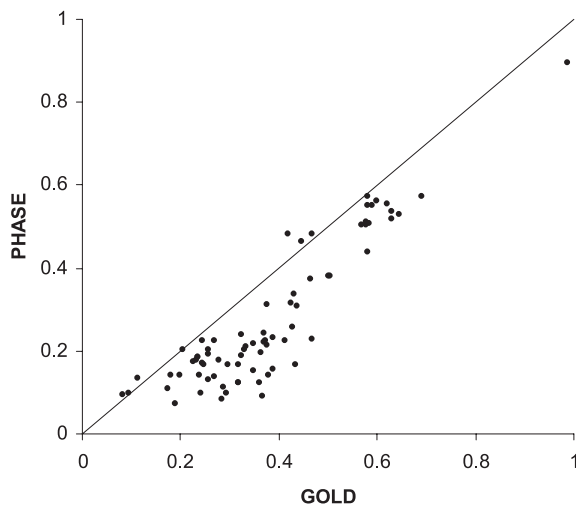


Figure 1. Relationship of D' computed from haplotypes inferred using E-M (GOLD) and Bayesian approaches (PHASE). The 45° line indicates theoretical equality of the two approaches.

4. DISCUSSION

In this study we quantify the extent of LD in Australian Holstein-Friesian dairy cattle in the two regions of BTA6 previously known to carry important QTL [22]. Australian Holstein-Friesian cattle are under intense selection for milk production traits and have a small effective population size of ~95 [25]. Also, there has been a continuous import of germplasm from different countries in addition to an ongoing within-country progeny testing program (<http://www.adhis.com.au>).

The D' values estimated from haplotypes inferred using the E-M and Bayesian approaches were comparable ($r = 0.91$). However, the Malecot model fitted better ($R^2 = 0.63$) with the D' estimated from Bayesian-inferred haplotypes as compared to E-M-inferred haplotypes ($R^2 = 0.43$) (Fig. 2 A-B; Tab. II). This conclusion about the fit of the Malecot model was similar when the distances in Mb units from the integrated map bLDB were used: the R^2 was 0.66 and 0.46 for the two methods of inferring haplotypes, respectively (Fig. 2C-D; Tab. II).

The integrated bLDB map uses as its framework the best available radiation-hybrid mapping information, which provides the best physical map available at present in cattle. Genetic maps may be inaccurate over small intervals where one is most interested in precisely quantifying the relationship between LD and

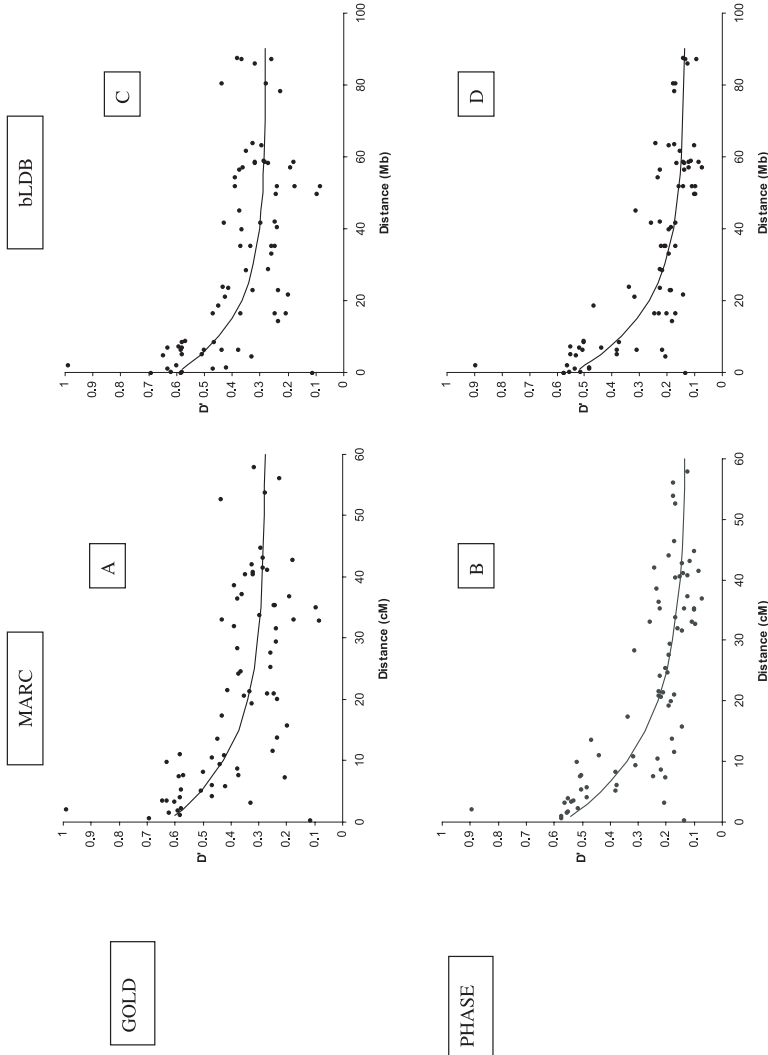


Figure 2. Distribution of D' values between syntenic marker pairs as a function of map distance. The lines indicate the fitted Malecot models. **A:** D' estimated from haplotypes inferred by E-M (GOLD), with MARC map distances; **B:** D' estimated from haplotypes inferred by the Bayesian approach (PHASE), with MARC map distances; **C:** D' estimated from haplotypes inferred by E-M (GOLD), with bLDB map distances; and **D:** D' estimated from haplotypes inferred by the Bayesian approach (PHASE), with bLDB map distances.

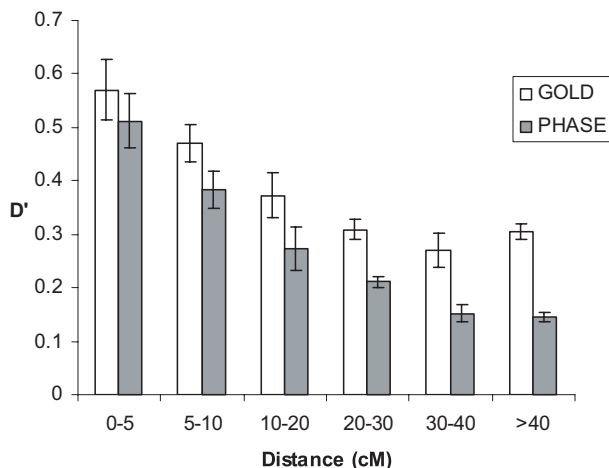


Figure 3. Average D' over different distances by two methods of inferring haplotypes, E-M (GOLD) and Bayesian approach (PHASE). The error bars indicate the standard errors of means.

distance. Inaccurate estimates of distances in the genetic maps are mainly because of the limited recombination events available in the resource population used for linkage mapping. This fact can be noted in the substantial changes recently in the genetic distances in MARC maps after incorporation of new map information [20]. From the mapping perspective, it is important to know the extent of LD over physical distances, so that marker density on the physical map can be determined.

Both the approaches for estimating the frequency of haplotypes used in this study have been reported to be satisfactory when compared with empirically estimated haplotypes [44], the Bayesian approach is somewhat more accurate [35, 45] but requires considerably more computation. Consequently, the E-M algorithm has been implemented in a number of analytical tools to survey LD based on population data *e.g.* [26]. The estimated swept radii from the two approaches are comparable. Based on the estimates of κ in the Malecot model, the results indicate that useful LD extends around 15 cM in Australian Holstein-Friesian dairy cattle (Tab. II). The extent of LD observed in this study in Australian dairy cattle is comparable to that reported for a Dutch Holstein population [10] in which LD between syntenic markers extended over “several tens of centimorgans”. However, there were differences in the sizes of the sample and the marker density used in these two studies.

On the one hand, high LD makes the application of LD mapping feasible for genome-wide QTL mapping, but on the other hand it can limit the resolution of

Table II. Parameter estimates of fitted Malecot models. The Malecot model takes the form $D'_i = (1 - L)M \exp(-\kappa d_i) + L + \varepsilon_i = \rho(d_i) + \varepsilon_i$ with $\text{var}(\varepsilon_i) = \sigma_s^2$.

Parameter	GOLD (E-M)		PHASE (Bayesian)	
	MARC	bLDB	MARC	bLDB
L	0.276±0.031	0.279±0.027	0.127±0.030	0.135±0.026
M	0.487±0.067	0.445±0.053	0.509±0.043	0.467±0.035
κ	0.086±0.033	0.065±0.024	0.073±0.020	0.056±0.014
σ_s	0.1223	0.1195	0.1033	0.1001
R^2	0.434	0.460	0.635	0.657

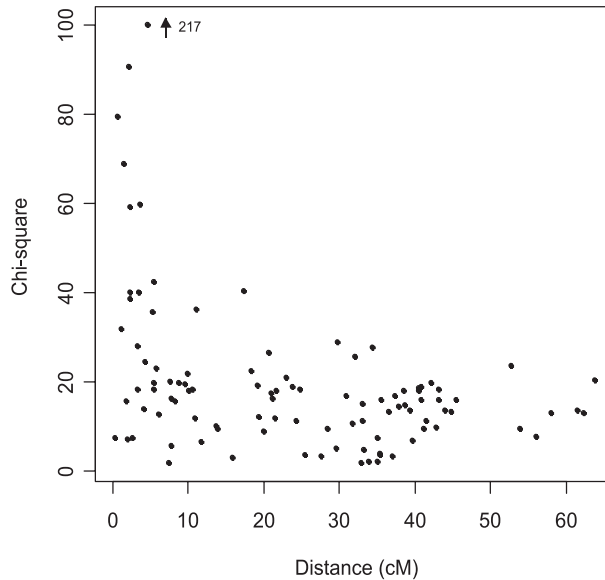


Figure 4. Distribution of chi-square values between marker pairs as a function of genetic distance in centimorgan (cM). One extreme value (chi-square =217) has not been plotted to scale.

fine mapping and may hinder positional cloning of the actual mutation. However, the results obtained from population-wide LD of markers with economic traits can be readily applicable in improving breeding stock through MAS [7] without the need to find the causative mutation.

In the human genome, the extent of LD is much smaller (5–100 kb) than reported in this study and in the Dutch Holstein population [10]. Recently Hinds *et al.* [16] reported the analysis of 1.58 million human SNP and suggested that the same power in association mapping of complex traits can be achieved by

using around 300 000 SNP. In contrast, in dairy cattle, the far higher levels of LD mean that far fewer markers will be required for similar power.

The standard measure of LD provides a starting point only, and may not fully capture the complexity of the occurrence of LD. By fitting the Malecot model in windows of small segments of the genome with very close markers, it is possible to assess how patterns of LD change not only between chromosomes, but also within chromosomes. Using such approaches, recent studies in humans have suggested that the genome is organised into the blocks of haplotypes with high LD, separated by regions of low LD [16, 26]. Recently, metric LD maps scaled in LD units (LDU) have been developed in humans, and have been suggested as a tool for targeted and optimal marker placement, which should lead to more powerful gene mapping and population studies [27, 47]. Similar efforts of creating LDU maps to understand the pattern of LD over the cattle genome are required to design optimum experiments for LD-based localisation of genetic variants affecting economic traits in cattle. The advent of high density SNP maps in bovine and high throughput genotyping platforms [17] now provide scope to undertake such studies in dairy cattle.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Genetics Australia for semen samples, ADHIS for pedigree data and the technical support provided by M. Jones and Y. Husaini in the laboratory. M.S.K. was a recipient of an F. H. Loxton Bequest scholarship, and a CRC-IDP support allowance. The work reported here was supported with funds from the Co-operative Research Centre for Innovative Dairy Products.

REFERENCES

- [1] Abdallah J.M., Goffinet B., Cierco-Ayrolles C., Perez-Enciso M., Linkage disequilibrium fine mapping of quantitative trait loci: a simulation study, *Genet. Sel. Evol.* 35 (2003) 513–532.
- [2] Abecasis G.R., Cookson W.O., GOLD—graphical overview of linkage disequilibrium, *Bioinformatics* 16 (2000) 182–183.
- [3] Ardlie K.G., Kruglyak L., Seielstad M., Patterns of linkage disequilibrium in the human genome, *Nat. Rev. Genet.* 3 (2002) 299–309.
- [4] Blott S., Kim J.J., Moisisio S., Schmidt-Kuntzel A., Cornet A., Berzi P., Cambisano N., Ford C., Grisart B., Johnson D., Karim L., Simon P., Snell R., Spelman R., Wong J., Vilkki J., Georges M., Farnir F., Coppeters W., Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution

- in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition, *Genetics* 163 (2003) 253–266.
- [5] Botstein D., Risch N., Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease, *Nat. Genet.* 33 Suppl. (2003) 228–237.
 - [6] Collins A., Morton N.E., Mapping a disease locus by allelic association, *Proc. Natl. Acad. Sci. USA* 95 (1998) 1741–1745.
 - [7] Dekkers J.C., Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons, *J. Anim. Sci.* 82 Suppl. (2004) E313–E328.
 - [8] Devlin B., Risch N., A comparison of linkage disequilibrium measures for fine-scale mapping, *Genomics* 29 (1995) 311–322.
 - [9] Excoffier L., Slatkin M., Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population, *Mol. Biol. Evol.* 12 (1995) 921–927.
 - [10] Farnir F., Coppeters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D., Georges M., Extensive genome-wide linkage disequilibrium in cattle, *Genome Res.* 10 (2000) 220–227.
 - [11] Grisart B., Coppeters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M., Snell R., Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine *DGAT1* gene with major effect on milk yield and composition, *Genome Res.* 12 (2002) 222–231.
 - [12] Guo S.W., Thompson E.A., Performing the exact test of Hardy-Weinberg proportion for multiple alleles, *Biometrics* 48 (1992) 361–372.
 - [13] Hall D., Wijsman E.M., Roos J.L., Gogos J.A., Karayiorgou M., Extended intermarker linkage disequilibrium in the Afrikaners, *Genome Res.* 12 (2002) 956–961.
 - [14] Hedrick P.W., Gametic disequilibrium measures: proceed with caution, *Genetics* 117 (1987) 331–341.
 - [15] Heyen D.W., Beever J.E., Da Y., Evert R.E., Green C., Bates S.R., Ziegler J.S., Lewin H.A., Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semiautomated parentage testing, *Anim. Genet.* 28 (1997) 21–27.
 - [16] Hinds D.A., Stuve L.L., Nilsen G.B., Halperin E., Eskin E., Ballinger D.G., Frazer K.A., Cox D.R., Whole-genome patterns of common DNA variation in three human populations, *Science* 307 (2005) 1072–1079.
 - [17] Hirschhorn J.N., Daly M.J., Genome-wide association studies for common diseases and complex traits, *Nat. Rev. Genet.* 6 (2005) 95–108.
 - [18] Hugot J.P., Chamaillard M., Zouali H. *et al.*, Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease, *Nature* 411 (2001) 599–603.
 - [19] Hyun C., Filippich L.J., Lea R.A., Shepherd G., Hughes I.P., Griffiths L.R., Prospects for whole genome linkage disequilibrium mapping in domestic dog breeds, *Mamm. Genome* 14 (2003) 640–649.

- [20] Ihara N., Takasuga A., Mizoshita K., Takeda H., Sugimoto M., Mizoguchi Y., Hirano T., Itoh T., Watanabe T., Reed K.M., Snelling W.M., Kappes S.M., Beattie C.W., Bennett G.L., Sugimoto Y., A comprehensive genetic map of the cattle genome based on 3802 microsatellites, *Genome Res.* 14 (2004) 1987–1998.
- [21] Jorde L.B., Linkage disequilibrium and the search for complex disease genes, *Genome Res.* 10 (2000) 1435–1444.
- [22] Khatkar M.S., Thomson P.C., Tammien I., Raadsma H.W., Quantitative trait loci mapping in dairy cattle: review and meta-analysis, *Genet. Sel. Evol.* 36 (2004) 163–190.
- [23] Lewontin R.C., The interaction of selection and linkage. I. General considerations: heterotic models, *Genetics* 49 (1964) 49–67.
- [24] Malécot G., *Les Mathématiques de l’Hérédité*, Masson et Cie, Paris, 1948.
- [25] Man W.Y.N., Pedigree analysis of Holstein Friesians in Australia, PhD thesis University of Sydney (2004).
- [26] Maniatis N., Collins A., Xu C.F., McCarthy L.C., Hewett D.R., Tapper W., Ennis S., Ke X., Morton N.E., The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis, *Proc. Natl. Acad. Sci. USA* 99 (2002) 2228–2233.
- [27] Maniatis N., Morton N.E., Gibson J., Xu C.F., Hosking L.K., Collins A., The optimal measure of linkage disequilibrium reduces error in association mapping of affection status, *Hum. Mol. Genet.* 14 (2005) 145–153.
- [28] McRae A.F., McEwan J.C., Dodds K.G., Wilson T., Crawford A.M., Slate J., Linkage disequilibrium in domestic sheep, *Genetics* 160 (2002) 1113–1122.
- [29] Mohlke K.L., Lange E.M., Valle T.T., Ghosh S., Magnuson V.L., Silander K., Watanabe R.M., Chines P.S., Bergman R.N., Tuomilehto J., Collins F.S., Boehnke M., Linkage disequilibrium between microsatellite markers extends beyond 1 cM on chromosome 20 in Finns, *Genome Res.* 11 (2001) 1221–1226.
- [30] Morton N.E., Collins A., Lawrence S., Shields D.C., Algorithms for a location database, *Ann. Hum. Genet.* 56 (Pt 3) (1992) 223–232.
- [31] Morton N.E., Zhang W., Taillon-Miller P., Ennis S., Kwok P.Y., Collins A., The optimal measure of allelic association, *Proc. Natl. Acad. Sci. USA* 98 (2001) 5217–5221.
- [32] Nicholas F.W., Barendse W., Collins A., Dalrymple B.P., Edwards J.H., Gregory S., Hobbs M., Khatkar M.S., Liao W., Maddox J.F., Raadsma H.W., Zenger K.R., Integrated maps and Oxford grids: maximising the power of comparative mapping, 29th International Conference on Animal Genetics, September 11–16, 2004, Tokyo, Japan (2004).
- [33] Nsengimana J., Baret P.V., Linkage disequilibrium and the genetic distance in livestock populations: the impact of inbreeding, *Genet. Sel. Evol.* 36 (2004) 281–296.
- [34] Nsengimana J., Baret P., Haley C.S., Visscher P.M., Linkage disequilibrium in the domesticated pig, *Genetics* 166 (2004) 1395–1404.
- [35] Qin Z.S., Niu T., Liu J.S., Partition-ligation-expectation-maximization algorithm for haplotype inference with single-nucleotide polymorphisms, *Am. J. Hum. Genet.* 71 (2002) 1242–1247.

- [36] Risch N., Merikangas K., The future of genetic studies of complex human diseases, *Science* 273 (1996) 1516–1517.
- [37] Sawyer S.L., Mukherjee N., Pakstis A.J., Feuk L., Kidd J.R., Brookes A.J., Kidd K.K., Linkage disequilibrium patterns vary substantially among populations, *Eur. J. Hum. Genet.* 13 (2005) 677–686.
- [38] Service S.K., Ophoff R.A., Freimer N.B., The genome-wide distribution of background linkage disequilibrium in a population isolate, *Hum. Mol. Genet.* 10 (2001) 545–551.
- [39] Shifman S., Kuypers J., Kokoris M., Yakir B., Darvasi A., Linkage disequilibrium patterns of the human genome across populations, *Hum. Mol. Genet.* 12 (2003) 771–776.
- [40] Stephens M., Donnelly P., A comparison of Bayesian methods for haplotype reconstruction from population genotype data, *Am. J. Hum. Genet.* 73 (2003) 1162–1169.
- [41] Sutter N.B., Eberle M.A., Parker H.G., Pullar B.J., Kirkness E.F., Kruglyak L., Ostrander E.A., Extensive and breed-specific linkage disequilibrium in *Canis familiaris*, *Genome Res.* 14 (2004) 2388–2396.
- [42] Tenesa A., Knott S.A., Ward D., Smith D., Williams J.L., Visscher P.M., Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes, *J. Anim. Sci.* 81 (2003) 617–623.
- [43] Tozaki T., Hirota K.I., Hasegawa T., Tomita M., Kurosawa M., Prospects for whole genome linkage disequilibrium mapping in thoroughbreds, *Gene* 346 (2005) 127–132.
- [44] Xu C.F., Lewis K., Cantone K.L., Khan P., Donnelly C., White N., Crocker N., Boyd P.R., Zaykin D.V., Purvis I.J., Effectiveness of computational methods in haplotype prediction, *Hum. Genet.* 110 (2002) 148–156.
- [45] Xu H., Wu X., Spitz M.R., Shete S., Comparison of haplotype inference methods using genotypic data from unrelated individuals, *Hum. Hered.* 58 (2004) 63–68.
- [46] Zapata C., The D' measure of overall gametic disequilibrium between pairs of multiallelic loci, *Evolution* 54 (2000) 1809–1812.
- [47] Zhang W., Collins A., Gibson J., Tapper W.J., Hunt S., Deloukas P., Bentley D.R., Morton N.E., Impact of population structure, effective bottleneck time, and allele frequency on linkage disequilibrium maps, *Proc. Natl. Acad. Sci. USA* 101 (2004) 18075–18080.
- [48] Zondervan K.T., Cardon L.R., The complex interplay among factors that influence allelic association, *Nat. Rev. Genet.* 5 (2004) 89–100.