Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Identification of quantitative trait loci for canine hip dysplasia by two sequential multipoint linkage analyses

Lan Zhu [a], Su Chen [a], Zhuoxin Jiang [b], Zhiwu Zhang [c], Hung-Chih Ku [a], Xuesong Li [a], Melinda McCann [a], Steve Harris [d], George Lust [e], Pual Jones [d] & Rory Todhunter [f]

[a] Department of Statistics, Oklahoma State University, Stillwater, OK, 74078, USA

[b] Department of Biostatistics, University at Buffalo - The State University of New York, Buffalo, NY, 14214, USA

[c] Institute for Genomic Diversity, 175 Biotechnology Building, Cornell University, Ithaca, NY, 14853, USA

[d] The WALTHAM Centre for Pet Nutrition, Waltham on the Wolds, Leicestershire, UK

[e] Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, 14853, USA

[f] Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, 14853, USA

Version of record first published: 02 Apr 2012

To cite this article: Lan Zhu, Su Chen, Zhuoxin Jiang, Zhiwu Zhang, Hung-Chih Ku, Xuesong Li, Melinda McCann, Steve Harris, George Lust, Pual Jones & Rory Todhunter (2012): Identification of quantitative trait loci for canine hip dysplasia by two sequential multipoint linkage analyses, Journal of Applied Statistics, 39:8, 1719-1731

To link to this article: http://dx.doi.org/10.1080/02664763.2012.673121

Journal of Applied Statistics
Vol. 39, No. 8, August 2012, 1719–1731

Taylor & Francis
Taylor & Francis Group

Identification of quantitative trait loci for canine hip dysplasia by two sequential multipoint linkage analyses

Lan Zhu[a]*, Su Chen[a], Zhuoxin Jiang[b], Zhiwu Zhang[c], Hung-Chih Ku[a], Xuesong Li[a], Melinda McCann[a], Steve Harris[d], George Lust[e], Pual Jones[d] and Rory Todhunter[f]

[a]Department of Statistics, Oklahoma State University, Stillwater, OK 74078, USA; [b]Department of Biostatistics, University at Buffalo – The State University of New York, Buffalo, NY 14214, USA; [c]Institute for Genomic Diversity, 175 Biotechnology Building, Cornell University, Ithaca, NY 14853, USA; [d]The WALTHAM Centre for Pet Nutrition, Waltham on the Wolds, Leicestershire, UK; [e]Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA; [f]Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

(*Received 17 June 2011; final version received 3 March 2012*)

Canine hip dysplasia (CHD) is characterized by hip laxity and subluxation that can lead to hip osteoarthritis. Studies have shown the involvement of multiple genetic regions in the expression of CHD. Although we have associated some variants in the region of fibrillin 2 with CHD in a subset of dogs, no major disease-associated gene has been identified. The focus of this study is to identify quantitative trait loci (QTL) associated with CHD. Two sequential multipoint linkage analyses based on a reversible jump Markov chain Monte Carlo approach were applied on a cross-breed pedigree of 366 dogs. Hip radiographic trait (Norberg Angle, NA) on both hips of each dog was tested for linkage to 21,455 single nucleotide polymorphisms across 39 chromosomes. Putative QTL for the NA was found on 11 chromosomes (1, 2, 3, 4, 7, 14, 19, 21, 32, 36, and 39). Identification of genes in the QTL region(s) can assist in identification of the aberrant genes and biochemical pathways involving hip dysplasia in both dogs and humans.

Keywords: reversible jump Markov chain Monte Carlo; multipoint linkage analysis; single nucleotide polymorphism; quantitative trait loci; canine hip dysplasia

1. Introduction

Canine hip dysplasia (CHD) is a very common disease in dogs, especially in large breed dogs such as Saint Bernards, Labrador Retrievers, German Shepherds, Golden Retrievers, and Rottweilers. The breed occurrence rate estimated by Orthopedic Foundation for Animals ranges from 1% up

*Corresponding author. Email: lan.zhu@okstate.edu

ISSN 0266-4763 print/ISSN 1360-0532 online

http://dx.doi.org/10.1080/02664763.2012.673121
http://www.tandfonline.com

to 75% (http://www.offa.org/hipstatbreed.html.). The symptoms of CHD include lameness in the back legs and swaggering or bunny-hopping gait. CHD develops into osteoarthritis [26,35].

CHD is among the most frustrating diseases in veterinary medicine because of the difficulties and high expenses of treatment. It may cost in the price range of $1000–$6000 if a surgery is needed [38]. For those who own a dog with hip dysplasia, they suffer not only from financial losses, but more importantly, emotional losses. Therefore, investigators seek methods that prevent dogs from developing hip dysplasia. One way to prevent CHD is through the selection of breeding parents who have excellent hip conformation. Studies have shown that dogs whose parents suffer from hip dysplasia have a higher chance of developing CHD [21] indicating that CHD is heritable.

Heritability of CHD varies greatly among studies. Leppanen *et al*. [25] estimated that the heritability of CHD in German Shepherd dogs was moderate (0.31–0.35) while Hamann *et al.* [17] reported that it was lower (0.26 ± 0.03). The range of heritability for CHD (0.1–0.68) is a function of the pedigree studied and the trait measured [6]. Other studies claimed that purebred dogs were more prone to CHD compared to cross-breed dogs [5,24,41]. Four independent studies [22,23,27, 34] and one of the follow up studies [44] showed biometric evidence for a putative quantitative trait locus contributing to CHD. The latter study refined an interval on CFA11 that subsequently allowed identification of a deletion in fibrillin 2 that was associated with a worsening of hip conformation [12] but no investigators have mapped a major locus controlling more than 20% of trait variation [7,28,34,44,45]. These results showed that CHD was a genetically complex disease to which as many as 10–20 quantitative trait loci (QTL) may contribute.

There are a plethora of methods for QTL mapping [10,30,37,40]. Most of the methods do not work computationally efficiently when there are thousands of markers genotyped on a pedigree with more than three generations. Some methods only work for a few specific pedigree structures, such as nuclear families with two generations, backcross families, or F_2 generation only [10,26]. Although some variance-component (VC) based approaches [1,2] can handle more extended pedigrees (pedigrees that have more than three generations), the big challenge of these methods is to estimate VCs accurately where the underlying assumption of the genetic model plays an important role in the estimation of these components. The performance of these component estimates further affects the accuracy of the model coefficient estimation. The assumption of the genetic model may be violated in real data [40]. Violation of these assumptions may lead to a higher false positive or false negative rate compared to results for data following a normal distribution [4,11]. Some researchers proposed an identity-by-decent (IBD) method, which predicts IBD probabilities at putative QTL positions [30]. Two disadvantages of this method are: firstly, the IBD method is not as efficient as the likelihood method for an explicit trait model, for example, the regression model in which the response variables are the traits [39]; secondly, the IBD method is computationally intensive for handling a pedigree with more than three generations at thousands of markers simultaneously [40].

A Markov chain Monte Carlo (MCMC) method [18,29] allows us to perform QTL mapping on data with a large number of markers and pedigrees with more than three generations. A con is that the total number of QTL has to be assumed in advance in the linkage analysis using MCMC. Mis-specification of the number of QTL may bring down the power of detecting the position and effects of the QTL [9,16]. A reversible jump MCMC method proposed by Green [13] provided a solution to this problem. It enabled the model to vary through changing the number of QTL, which would eliminate the disadvantage brought by the mis-specification of the number of QTL [19]. Another advantage of this method was that it could handle a large number of markers (usually thousands of markers) on one or more chromosome(s) in an arbitrary pedigree, which provided more flexibility in the QTL mapping.

The goal of this study was to identify genomic region(s) or QTL that contributed to CHD using the reversible jump MCMC method and also to demonstrate that this method can be applied to the whole genome analysis with an arbitrary pedigree structure, including the complex pedigree

with mating loops across generations. We used a software package 'LOKI' [19,20] to estimate the unknown parameters in the model including the number and locations of the putative QTL, percentage of genetic variation explained due to the detected QTL, and percentage of total variation explained due to all the QTL for CHD. It is assumed that the total variation in the trait is composed of genetic and environmental variation. We scanned through the genome to estimate the degree of association between genomic regions and phenotypes of CHD. If the degree of association is significantly high for a specific region, the genes on this region may be considered as the putative QTL.

2. Methods

2.1 *Data set*

Animals: Dogs used in this study were originated from close breeding colonies at the Baker Institute for Animal Health at Cornell University, the Guiding Eyes for the Blind in Yorktown Height, NY, and those admitted to the Cornell University Hospital for radiographic evaluation from January 1999 through October 2007 [42]. There were 366 dogs genotyped from a canine pedigree, which was constructed by initially crossing eight Labrador Retrievers (four males and four females) with CHD or secondary hip osteoarthritis and seven greyhounds (two males and five females) from racing stock with excellent or good hip conformation [3,33]. Our study was based on these 366 dogs in families with 3–7 generations. Family members range from 3 to more than 15. The distribution of these 366 dogs by breed was summarized in Table 1.

Phenotypes: The Norberg Angle (NA) [14,36] is measured on a ventrodorsal hip-extended pelvic radiograph. The NA is formed by a line connecting the centers of both femoral heads and the line between the center of a femoral head and the craniodorsal acetabular rim on the same hip. It is one of several radiographic measurements to diagnose the CHD. In this study, the NA ranged from 50° (a subluxated hip) to 123° (a normal hip) [45]. NAs were measured on both hips for each dog and denoted by the left Norberg angle (NAL) for the left hip and right Norberg angle (NAR) for the right hip. The distribution of NA measurements was summarized in Table 2.

Genotypes: 21,455 single nucleotide polymorphism (SNP) markers were genotyped on 366 dogs with Illumina CanineSNP20 BeadChip (Illumina Inc., San Diego CA, http://www.illumina.com/documents/products/datasheets/datasheet_canine_snp20.pdf) across 39 chromosomes (38 autosomes and X chromosome). Refer to Zhou *et al*. [43] for detailed marker selection and genotyping information.

Table 1. Distribution of dogs by breed.

Breed	Number of dogs
Labrador Retriever (LR)	188
Greyhound	7
LR × Greyhound (F1)	8
F1 × LR	67
F1 × Greyhound	16
(F1 × LR) × (F1 × LR)	12
Golden Retriever	15
German Shepherd	17
Newfoundland	18
Rottweiler	14
Bernese Mt. Dog	4
Total	366

Table 2. Distribution of NA measurements.

	Number of dogs	
NA (°)	NAL	NAR
~70	1	1
71–80	5	4
81–90	10	9
91–100	42	42
101–110	206	198
111–120	88	99
Total	352[a]	353[a]

Notes: [a]Not all dogs have both NAL and NAR measurement available. Therefore, the number of total dogs of NAL or NAR does not equal to the total number of dogs included in the study.

2.2 *Theoretical model*

Suppose there are k QTLs controlling the quantitative trait of interest. For the ith QTL, its genotype is either AA, Aa, or aa, coded as '0', '1' or '2', respectively. The Bayesian MCMC model is as follows

$$\mathbf{y} = \mu + X\beta + \sum_{i=1}^{k} G_i\alpha_i + e,$$

where $\mathbf{y}$ is an $(n \times 1)$ vector of the quantitative trait (of n individuals); μ is the overall mean; β is an $(m \times 1)$ vector of covariate effects (such as gender and breed in this study); m is the number of covariates; X is an $(n \times m)$ incidence matrix for covariate effects; G_i is an $(n \times 2)$ incidence matrix for the genotypic effects of ith QTL; α_i is the genotypic effects for ith QTL (including additive and dominant effects); e is an $(n \times 1)$ vector of normally distributed residual error.

The joint posterior distribution of all parameters can be written as: $p(k, G, \beta, \lambda, \delta, \eta, \alpha, \sigma_e^2, \mu|y, M)$, where k is the number of QTL (the maximum number of QTL was assumed to be 10); G are genotypes of the QTL; M are genotypes of all markers; λ is the map position of the QTL; δ indicates which QTL are currently linked; η is the vector of allele frequencies of QTL and markers; σ_e^2 is the variance of the residual error. α (a vector of the genotypic effects of QTL, α_1, $\alpha_2, \ldots, \alpha_k$), μ and y are defined in the above equation. Assume α is normally distributed. All other parameters are independently uniformly distributed in prior.

We assume that the pedigree information is correct and first check for the Mendelian inconsistencies in the genotypes, then a reversible jump MCMC method [13] is used to produce samples from the joint posterior distribution of all unknown parameters in the Bayesian MCMC model. Haldane's mapping function [15] is used to convert the map positions into recombination fractions. The updates of genotypes for all markers (M) and QTL's (G) are performed simultaneously for all individuals according to the 'reverse peeling' algorithm [31] based on the pedigree. The QTL's map position and linkage status (λ and δ) are updated by use of peeling. The changing of a QTL's linkage status affects the dimension of the model, hence reversible jump MCMC steps [13] are used to change the number of QTL's in the model. Each individual parameter generated from the MCMC procedure should follow its marginal posterior distribution. Then we estimate the unknown parameters based on their marginal posterior distributions. In this study, we analyzed our data by a software package named 'LOKI' [19] Bayes Factor (BF) is used as the signal of significant linkage.

BF compares two hypotheses, given the observed data. The equation is as follows:

$$\mathrm{BF} = \frac{\mathrm{Prob(Data|alternative\ hypothesis)}}{\mathrm{Prob(Data|null\ hypothesis)}}.$$

The null hypothesis is that there is no linkage (the degree of association between traits and genotypes at give genomic regions). The alternative hypothesis is that the evidence of linkage exists. BF in our model is a measure of the evidence of linkage. It is computed by the following ratio [40]:

$$\mathrm{BF} = \frac{q_1(1-q_1)}{q_0(1-q_0)},$$

where q_0 is the prior probability of linkage and q_1 is the posterior probability of linkage. According to the general criteria [32], a BF > 20 indicates a strong evidence of linkage, $3 <$ BF ≤ 20 indicates a moderate evidence of linkage, and $1 <$ BF ≤ 3 indicates a weak evidence of linkage. No evidence of linkage if BF ≤ 1. Thus a higher BF indicates a stronger evidence of linkage.

2.3 *Data analysis*

Due to the large number of markers and complicated family structure (inbred loops), the reversible jump MCMC algorithm runs slowly if we analyze the whole genome in the model at once. One possible solution is to break pedigrees down to small families. However, due to inbred loops over multiple generations in this study, it is difficult to find an optimal way to break down pedigrees. An alternative solution is to take two sequential steps in the analysis:

Step 1. *Single chromosome linkage analysis* (*SCLA*): To reduce the computational load and initially locate putative QTL, we started with the SCLA for each trait and proceeded through the entire genome (38 autosomes and X chromosome). For a complex disease, such as CHD, QTL on multiple chromosomes may contribute simultaneously to the trait. The single chromosome analysis cannot handle this effect, so the multi-chromosome analysis was performed afterwards. That is, markers and their qualified neighboring markers (refer to step 2) that were detected harboring putative QTL from SCLA were all included in the multiple chromosome linkage analysis (MCLA) as described in step 2.

Step 2. *MCLA*: In the multi-chromosome linkage analysis, all the chromosome regions that were significant (with BF > 3) in the previous single chromosome analysis were grouped and included in a single model analysis. In other words, all markers that were in the 95% posterior probability interval (PPI) of QTL with BF > 3 in the SCLA were included in the multi-chromosome linkage analysis.

The perfect solution is of course to fit all chromosomes simultaneously, but due to the limitation of the algorithm, we proposed the above sequential analysis. The potential problems associated with this method will be discussed in the discussion section.

3. Results

SCLA: Chromosomes and the 95% PPIs of QTL with BF > 3 were listed in Table 3. As we can see, there are about 11 chromosomes that may harbor QTL for NAL and NAR. Although NAL and NAR are highly correlated, they are not perfectly correlated. Therefore, these 11 chromosomes for NAL are not exactly the same as those for NAR. Our main interest in this study is to find some common regions that may be associated with both NAL and NAR. As we can see from Table 3, some common regions on chromosomes 2, 4, 6, 21, 24, 32, and 39 are detected to be associated with both NAL and NAR. Specifically, 95% PPIs on chromosomes 2, 4, 21, and 24 for NAL

Table 3. BF and 95% PPI in the SCLA.

Trait	Chromosome #	BF	95% PPI(cM)
NAL	2	5.17	(78.47, 83.97)
NAL	4	3.27	(4.1, 7.2) (89.7, 91.7*[a])
NAL	6	4.09	(75.25, 79.05)
NAL	7	6.57	(66.05, 82.15)
NAL	19	3.34	(46.8, 50.8)
NAL	21	7.99	(48.55, 51.05)
NAL	24	5.88	(20.45, 33.05)
NAL	32	3.48	(4.85, 6.75)
NAL	36	3.55	(12.87, 16.27)
NAL	37	5.15	(10.67, 10.97) (28.27, 34.27*)
NAL	39	4.02	(34.35, 41.45) (62.25, 81.85*)
NAR	1	7.95	(57.36, 67.06) (91.36, 92.66*)
NAR	2	15.57	(77.47, 86.87)
NAR	3	3.85	(3.85, 5.05*) (14.45, 18.65) (49.15, 49.55)
NAR	4	3.73	(4.7, 10.0) (40.7, 41.8) (88.9, 91.7*)
NAR	6	9.88	(18.05, 24.25*) (69.95, 71.25)
NAR	14	9.34	(3.45, 4.25*) (45.45, 57.85)
NAR	15	3.5	(58.95, 67.65)
NAR	21	3.24	(37.45, 38.35) (50.55, 54.35*)
NAR	24	3.32	(29.45, 33.05) (48.75, 50.75*)
NAR	32	4.97	(33.55, 35.55)
NAR	39	12.48	(17.65, 18.25) (38.45, 41.55) (61.35, 66.25*)

Note: [a]The highest BF falls into the interval with the sign of asterisk (*).

overlap those regions for NAR. For example, the 95% PPI on chromosome 2 for NAL is 78.47–83.97 cM, which is the subset of the 95% PPI 77.47–86.87 cM on the same chromosome for NAR. Figure 1 shows the evidence of linkage (BF) along chromosomes 2, 4, 21, and 24, respectively. The percentage of total variation in the trait accounted by QTL on chromosomes 2, 3, 4, 21, 24, and 39 are listed in Table 4, respectively. As the table shows, each QTL on chromosomes 2, 4, 21, and 24 explained about 7.6–11% of the total variation and about 1.9–2.4% of the genetic variation for NA.

MCLA: MCLA certainly reduced the number and the range of significant QTL as we can see from Table 5. Compared to the SCLA, the evidence of linkage (BFs) on all chromosomes was reduced, except on chromosome 19 for NAL and chromosome 21 for NAR. Supporting evidence for some QTL even dropped below or close to the weakest significant level (BF = 1), such as for chromosomes 6, 24, and 37 for NAL and chromosomes 6, 15, and 24 for NAR. It was interesting to see that signals became stronger on chromosome 19 (BF = 4.78 in MCLA compared to 3.34 in SCLA) for NAL and on chromosome 21 (BF = 5.71 in MCLA compared to 3.24 in SCLA) for NAR. The strongest signals in SCLA were BF = 7.99 on chromosome 21 for NAL and BF = 15.57 on chromosome 2 for NAR; the strongest signals in MCLA were BF = 7.62 on chromosome 21 for NAL (which is about the same as in SCLA) and BF = 5.71 on the same chromosome 21 for NAR. The peak signal regions for NAL and NAR on chromosome 21 did not overlap perfectly, but they were in the range of 42–51 cM.

Although the BFs were reduced, chromosome 2 showed weak evidence of linkage for both NAL and NAR on about the same genomic region 79.47–79.67 cM for NAL and 77.9–83.1 cM for NAR.

The left NA-specific QTL (not significant for NAR) were detected on chromosome 19 between 50.55 and 50.75 cM with moderate evidence. Right NA-specific QTL (not significant for NAL) were detected on chromosomes 3 between 19.13 and 22.33 cM and on chromosome 4 between 54.5 and 61.7 cM.

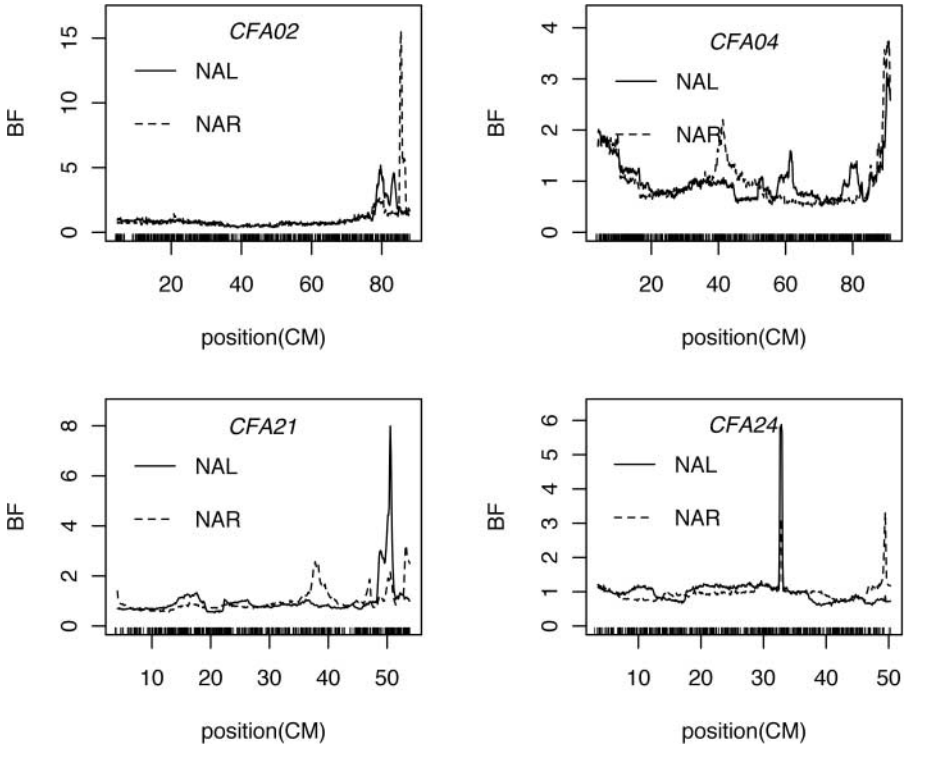

Figure 1. BF (signal of significance) along the chromosome (CFA02, CFA04, CFA21, and CFA24) in the SCLA for trait NAL (solid line) and NAR (dashed line).

Table 4. Percentage of total variation (genetic variation) due to QTL in the SCLA.

	Trait	
Chromosome	NAL	NAR
CFA02	7.6% (2.3%)	9.1% (2.4%)
CFA03	7.3% (2.3%)	8.5% (2.5%)
CFA04	8.0% (2.4%)	9.8% (2.3%)
CFA21	7.8% (2.2%)	9.2% (2.3%)
CFA24	7.6% (1.9%)	11.0% (1.9%)
CFA39	6.8% (2.6%)	6.7% (2.5%)

Table 6 lists the percentage of the total variation and genetic variation explained by QTL in the MCLA for NAL and NAR. About 3.12–10.3% of the total phenotypic variation was explained by all the detected QTL on the chromosomes listed in Table 5.

The estimated posterior probability of the QTL number is shown in Figure 2. The maximum posterior probability in Table 5 indicates that 9 QTL in the dog genome might be associated with trait NAL and 8 QTL might be associated with trait NAR. Although the posterior densities for both traits show unimodal distribution, the probability of 6–10 QTL for both traits is about 80–85% (the sum of the last five categories of k in Figure 2).

Table 5. BF and 95% PPI in the MCLA.

Trait	Chromosome #	BF	95% PPI (cM)
NAL	2	2.51	(79.47, 79.67)
NAL	4	1.44	(82.25, 91.75)
NAL	6	0.70	–[a]
NAL	7	2.84	(66.55, 67.35)
NAL	19	4.78	(50.55, 50.75)
NAL	21	7.62	(50.15, 50.25)
NAL	24	0.55	–
NAL	32	2.39	(6.25, 6.45)
NAL	36	2.32	(15.55, 15.65)
NAL	37	0.98	–
NAL	39	1.82	(61.98, 63.38*)[b] (72.48, 72.58)
NAR	1	1.53	(57.85, 62.15)
NAR	2	1.23	(77.9, 83.1)
NAR	3	3.06	(19.13, 22.33)
NAR	4	3.14	(54.5, 61.7)
NAR	6	1.05	–
NAR	14	1.27	(3.65, 4.05)
NAR	15	0.62	–
NAR	21	5.71	(42.65, 44.35)
NAR	24	1.05	–
NAR	32	2.46	(35.65, 35.85)
NAR	39	1.81	(39.35, 40.15)

Notes: [a]The hyphen indicates that no significant QTL was detected on this chromosome.
[b]The highest BF falls into the interval with the sign of asterisk (*).

Table 6. Percentage of total variation (genetic variation) due to QTL in MCLA.

Trait	NAL	NAR
Multiple CFAs	10.30% (4.68%)	3.12 % (1.58%)

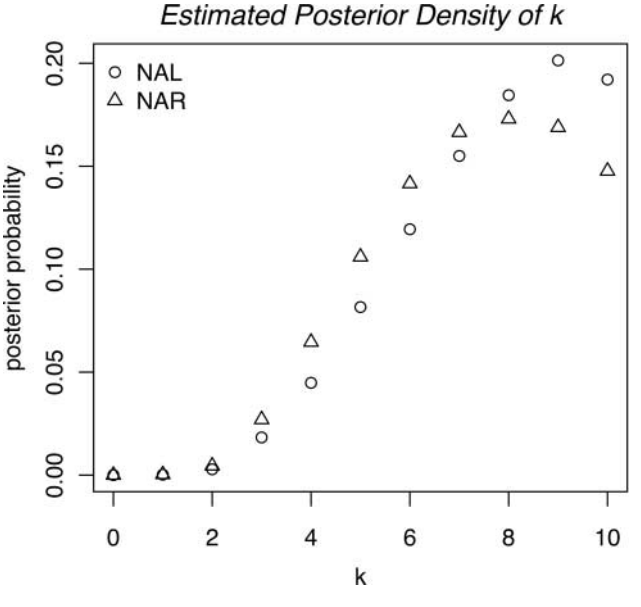

Figure 2. Estimated posterior density of k, the number of QTL, in the MCLA for NAL and NAR.

Table 7. Summary of QTL ranges.

Chromosome #	95% PPI (cM)
1	57.85–62.15
2	77.9–83.10
3	19.13–2.33
4	54.5–61.70, 82.25–91.75
7	66.55–67.35
14	3.65–4.05
19	50.55–50.75
21	42.65–44.35, 50.15–50.25
32	6.25–6.45, 35.65–35.85
36	15.55–15.65
39	39.35–40.15, 61.98–63.38, 72.48–72.58

In summary, MCLA by a reversible jump MCMC method found evidence of the presence of putative QTL for trait NA on 11 chromosomes, including 1, 2, 3, 4, 7, 14, 19, 21, 32, 36, and 39 (Table 7).

4. Discussion

While MCLA ideally should take the whole genome into account, it is computationally too intensive on a large number of markers, which increases significantly the number of iterations in the procedure of estimating the genotypes of QTL between adjacent markers. The computational load is even heavier when the sample size is large with a complicated pedigree structure. In this study, we have 21,455 SNP markers genotyped across 39 chromosomes from 366 dogs across 3–7 generations with inbred loops. Currently, there is no efficient method available to take all markers into account in a single model. So we start from a single chromosome multipoint linkage analysis in a Bayesian framework with an equal prior probability of linkage at each segment of the genome. In the SCLA, we allow multiple QTL on the same chromosome with additive effects on the trait under study to reduce the chance of missing true QTL. Of course, the trade-off is that it might increase the false positive rate. To reduce the number of false positive QTLs computationally and increase the resolutions of the QTL regions, one way is to refine significance levels of BF, so they can be more specific to studies of QTL mapping. Alternatively, this problem can be partially addressed in the MCLA.

We include all significant genomic regions from the SCLA in the second stage analysis – MCLA. An equal prior probability of linkage was assumed for each region under the null hypothesis of no linkage. In this model with multiple chromosomes, we consider the additive effects from all previously significant genomic regions simultaneously. With the presence of other putative QTL, some previously significant signals (BFs on chromosome 6, 24, and 37 for NAL; chromosome 15 for NAR) from the single chromosome analysis have faded away and become insignificant, which might be the false positive ones in the first stage analysis.

In general, all signals (BFs) of significance were weaker in the MCLA than in the single chromosome analysis except those on chromosome 19 and 21. This makes intuitive sense. Since CHD is a complex disease, under the assumption of QTL on one chromosome only in the SCLA, the effect of genomic regions might be overestimated, leading to a higher BF in the first stage analysis. On the other hand, signals on chromosome 19 and 21 get stronger in the MCLA than in the single chromosome analysis. Although we do not have biological evidence at this stage to explain why these signals get stronger, one of the possible reasons might be some interactions among QTL that result in amplifying the effects of QTL on these chromosomes. Further analysis on significant regions on chromosome 19 and 21 will be needed to refine the QTL location since Marschall and

Distl [28] also found significant evidence of association between QTL on chromosomes 19 and 21 and CHD in German Shepherd dogs. In addition, from our study, the signals at regions on chromosome 21 remain the highest for both NAL and NAR, although the QTL positions are more than 4.5 cM away from what Marschall and Distl [28] detected in German Shepherd dogs. We find a significant QTL region on chromosome 19 between 50.55 and 50.75 cM, which is consistent with Marschall and Distl's findings (50.5 Mb) on the same chromosome for German Shepherd dogs [28]. Whether a QTL in this region is segregating across more than one breed needs further investigation.

The whole-genome screen in previous studies showed multiple QTL contributing to the different phenotypes of CHD [7,8,28,36,43,44]. The number of chromosomes with evidence of putative QTL was 11 in our study, whereas Chase *et al*. [7,8] reported only 2 on a single chromosome (CFA01) reaching genome-wide significance, Todhunter *et al*. [36] found 12, each with chromosome-wide significance, Marschall and Distl [28] found 19, Zhu *et al*. [44] found 2 on CFA11 but only analyzed linkage on 2 chromosomes (CFA11 and 29), and Zhou *et al*. [43] found 3 chromosomes with putative QTL for traits of CHD. The inconsistent results among different studies are likely due to differences in dog populations (different number of dogs, different dog breeds, and different pedigree structures), number and density of genomic markers, allele frequencies of the QTL, distribution of CHD traits under the study, and statistical analysis methods. Compared to the study of Zhu *et al*. [44], more dogs and more dense markers on more than two chromosomes were included in the current study and the CHD traits used in the present study were also different. Although the individual dogs, phenotypes, and markers included in this study were the same as one of the two data sets (Illumina CanineSNP20 Bead arrays) in Zhou *et al*. [43], the statistical findings were not as close as one would expect. Ideally, if the density of markers was high enough and if both statistical approaches were powerful enough, one would expect to see some common QTL detected by both methods. Multipoint linkage analysis and single marker analysis for detecting QTL both rely on linkage disequilibrium between markers and QTL. When we do not have extremely high-resolution (density) markers, it is very rare or impossible to detect a QTL by linkage analysis right on the marker position. However, the advantage of multipoint linkage analysis methodology compared with single marker analysis is that multiple markers were included in a single model while the pedigree structure and kinship between individuals were simultaneously considered. This model allows for the effect of multiple QTL, which are in linkage disequilibrium (LD) with markers, to be revealed as a functional unit for the phenotype under study. Therefore, multipoint linkage has increased power for small data sets. For single marker association, unless every SNP in linkage disequilibrium with the QTL is informative (segregating), the association will be missed. Finally, significantly associated SNPs detected by Zhou *et al*. [43] were mostly from a second data set that was not included in this current study. However these significant SNPs may in strong LD with true QTL in regions detected by our approach as markers physically on different chromosomes may show significant LD simply due to sampling.

In our study, if we could include the whole genome markers in a single model, no multiple comparison issue would be involved in this Bayesian framework. However, due to the extremely intensive computational load, we currently can only include some markers that are highly likely to be linked to a single trait and only allow up to 10 QTL in the model. This is the limitation of this approach. The effect of a prior number of QTL on the results of our analysis needs to be further explored. Moreover, additive QTL effects only as described by the model would not be enough to explain the true underlying biological mechanism. Another limitation of the current model is that only one trait is considered in the model and we analyzed NAL and NAR separately, while left- and right-side hip traits are actually highly correlated. Advanced statistical models, which take multiple correlated traits into account and/or include interactions between QTL, are in high demand. Moreover, regions detected in this study are still in megabases. The positions of QTL have to be refined by increasing the density of informative markers in these regions, collecting

more individuals within breed and across breeds [45], and developing more powerful statistical tools, such as the multiple-trait model, which takes the correlation between traits into account. The analysis of multiple traits simultaneously should have more power than single-trait analysis, especially when we are interested in the identification of common QTL for multiple correlated traits. Genes located in QTL regions identified in these studies and involved in the development of osteoarthritis in man and mouse may be preferentially chosen for SNP development and for testing of linkage with CHD.

Acknowledgements

The authors gratefully acknowledge the technical assistance of Dr Dana Brunson in running programs for data analysis in this study. The computational work was performed at the OSU High Performance Computing Center (OSUHPCC) at Oklahoma State University (OSU). We also thank the anonymous reviewers for their invaluable critical comments.

References

[1] L. Almasy and J. Blangero, *Multipoint quantitative-trait linkage analysis in general pedigrees*, Am. J. Hum. Genet. 62 (1998), pp. 1198–1211.

[2] C.I. Amos, *Robust variance-components approach for assessing genetic linkage in pedigrees*, Am. J. Hum. Genet. 54 (1994), pp. 535–543.

[3] C. Beling, P. Gustafsson, and H. Kasstrom, *Metabolism of estradiol in greyhounds and German Shepherd dogs. In investigation with special reference to hip dysplasia*, Acta Radiol. Suppl. 344 (1975), pp. 109–120.

[4] J. Bethony, J. Williams, L. Almasy, R. Correa-Oliveira, J. Blangero, and S. Williams-Blangero, *Genetic analysis of quantitative traits in highly ascertained samples: Total serum IgE in families with asthma*, Genet. Epidemiol. 21(Suppl. 1) (2001), pp. S174–S179.

[5] S. Bliss, R. Todhunter, R. Quaas, G. Casella, R. Wu, G. Lust, A. Williams, S. Hamilton, N. Dykes, A. Yeager, R. Gilbert, N. Burton-Wurster, and G. Acland, *Quantitative genetics of traits associated with hip dysplasia in a canine pedigree constructed by mating dysplastic Labrador Retrievers with unaffected Greyhounds*, Am. J. Vet. Res. 63(7) (2002), pp. 1029–1035.

[6] G. Breur, G. Lust, and R. Todhunter, *Genetics of hip dysplasia and other orthopedic traits*, in *Genetics of the Dog*, A. Ruvinsky and J. Sampson, eds., CAB International, Wallingford, Oxon, UK, Chapter 9, 2002, pp. 407–438.

[7] K. Chase, D.F. Lawler, F.R. Adler, E.A. Ostrander, and K.G. Lark, *Bilaterally asymmetric effects of quantitative trait loci (QTLs): QTLs that affect laxity in the right versus left coxofemoral (hip) joints of the dog (Canis familiaris)*, Am. J. Med. Genet. 124 (2004), pp. 239–247.

[8] K. Chase, D.F. Lawler, D.R. Carrier, and K.G. Lark, *Genetic regulation of osteoarthritis: A QTL regulating cranial and caudal acetabular osteophyte formation in the hip joint of the dog (Canis familiaris)*, Am. J. Med. Genet. 135 (2005), pp. 334–335.

[9] M.-H. Dizier, C. Bonafti-Pellie, and F. Clerget-Darpoux, *Conclusions of segregation analysis for family data generated under two-locus models*, Am. J. Hum. Genet. 53 (1993), pp. 1338–1346.

[10] R. Elston, S. Buxbaum, K. Jacobs, and J. Olson, *Haseman and Elston revisited*, Genet. Epidemiol. 19 (2000), pp. 1–17.

[11] W. Forrest and E. Feingold, *Composite statistics for QTL mapping with moderately discordant sibling pairs*, Am. J. Hum. Genet. 66 (2000), pp. 1642–1660.

[12] S. Friedenberg, L. Zhu, Z. Zhang, W. Foels, P. Schweitzer, W. Wang, P. Fisher, N. Dykes, E. Corey, M. Singer, S. Jung, X. Sheng, L. Hunter, S. McDonough, G. Lust, S. Bliss, U. Krotscheck, T. Gunn, and R. Todhunter, *Evaluation of a fibrillin 2 gene haplotype associated with hip dysplasia and incipient osteoarthritis in dogs*, AJVR 72 (2011), pp. 530–540.

[13] P.J. Green, *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika 82 (1995), pp. 711–732.

[14] P.O. Gustafsson, S.-E., Olsson, H. Kasstrom, and B. Wennman, *Skeletal development of Greyhounds, German Shepherd dogs andtheir crossbreed offspring*, Acta Radiol. 344 (1975), pp. 81–108.

[15] J.B.S. Haldane, *The combination of linkage values, and the calculation of distance between the loci of linked factors*, J. Genet. 8 (1919), pp. 299–309.

[16] C. Haley and S. Knott, *A simple regression method for mapping quantitative trait loci in line crosses using flanking markers*, Heredity 69 (1992), pp. 315–324.

[17] H. Hamann, T. Kirchhoff, and O. Distl, *Bayesian analysis of heritability of canine hip dysplasia in German Shepherd dogs*, J. Anim. Breed. Genet. 120 (2003), pp. 258–268(11).
[18] W.K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika 57 (1970), pp. 97–109.
[19] S. Heath, *Markov chain Monte Carlo segregation and linkage analysis for oligogenic models*, Am. J. Hum. Genet. 61 (1997), pp. 748–760.
[20] S. Heath, G. Snow, C. Thompson, and E. Wijsman, *MCMC segregation and linkage analysis*, Genet. Epidemiol. 14 (1997), pp. 1011–1016.
[21] A. Hedhammar, S. Olsson, S. Andersson, L. Persson, L. Pettersson, A. Olausson, and P. Sundgren, *Canine hip dysplasia: Study of heritability in 401 litters of German Shepherd dogs*, J. Am. Vet. Med. Assoc. 174(9) (1979), pp. 1012–1016.
[22] V. Janutta, H. Hamann, and O. Distl, *Complex segregation analysis of canine hip dysplasia in German Shepherd dogs*, J. Hered. 97 (2006), pp. 13–20.
[23] E. Leighton, *Genetics of canine hip dysplasia*, J. Am. Vet. Med. Assoc. 210 (1997), pp. 1474–1479.
[24] E. Leighton, J. Linn, and R. Willham, *A genetic study of canine hip dysplasia*, Am. J. Vet. Res. 38 (1977), pp. 241–244.
[25] M. Leppanen, K. Maki, J. Juga, and H. Saloniemi, *Estimation of heritability for hip dysplasia in German Shepherd dogs in Finland*, J. Anim. Breed Genet. 117 (2000), pp. 97–103.
[26] G. Lust, A.J. Williams, N. Burton-Wurster, G.J. Pijanowski, K.A. Beck, G. Rubin, and G.K. Smith, *Joint laxity and its association with hip dysplasia in Labrador Retrievers*, Am. J. Vet. Res. 54 (1993), pp. 1990–1999.
[27] K. Maki, L. Janss, and A. Groen, *An indication of major genes affecting hip and elbow dysplasia in four Finnish dog populations*, Heredity 92 (2004), pp. 402–408.
[28] Y. Marschall and O. Distl, *Mapping quantitative trait loci for canine hip dysplasia in German Shepherd dogs*, Mamm. Genome 18 (2007), pp. 861–870.
[29] N. Metropolis, A.W. Rosenbluth, M.N., Rosenbluth, A.H. Teller, and E. Teller, *Equations of state calculations by fast computing machines*, J. Chem. Phys. 21 (1953), pp. 1087–1092.
[30] T.H.E. Meuwissen and M.E. Goddard, *Multipoint identity-by-descent prediction using dense markers to map quantitative trait loci and estimate effective population size*, Genetics 176 (2007), pp. 2551–2560.
[31] J. Ott, *Computer-simulation methods in human linkage analysis*, Proc. Natl. Acad. Sci. 86 (1989), pp. 4175–4178.
[32] A. Raftery, *Approximate Bayes factors and accounting for model uncertainty in generalized linear models*, Bometrika 83 (1996), pp. 251–266.
[33] R.J. Todhunter, G.M. Acland, M. Olivier, A. Williams, M. Vernier-Singer, N. Burton-Wurster, J. Farese, Y. Grön, R. Gilbert, N. Dykes, and G. Lust, *An outcrossed canine pedigree for linkage analysis of hip dysplasia*, J. Hered. 90 (1999), pp. 83–92.
[34] R.J. Todhunter, G. Casella, S.P. Bliss, G. Lust, A.J. Williams, S. Hamilton, N.L. Dykes, A.E. Yeager, R.O. Gilbert, N.I. Burton-Wurster, C.C. Mellersh, and G.M. Acland, *Power of a Labrador retriever-greyhound pedigree for linkage analysis of hip dysplasia and osteoarthritis*, Am. J. Vet. Res. 64 (2003), pp. 418–424.
[35] R.J. Todhunter and G. Lust, *Canine hip dysplasia: Pathogenesis*, in *Textbook of Small Animal Surgery*, D. Slatter, ed., W.B. Saunders, Philadelphia, PA, 2003, pp. 2009–2019.
[36] R.J. Todhunter, R. Mateescu, G. Lust, N. Burton-Wurster, N. Dykes, S. Bliss, A. Williams, M. Vernier-Singer, E. Corey, H. Carlos, R. Quaas, Z. Zhang, R. Gilbert, D. Volkman, G. Casella, R. Wu, and G. Acland, *Quantitative trait loci for hip dysplasia in a cross breed pedigree*, Mamm. Genome 16 (2005), pp. 720–730.
[37] S. Wang, C.J. Basten, and Z.-B. Zeng, *Windows QTL Cartographer 2.5*, Department of Statistics, North Carolina State University, Raleigh, NC, 2007. Available at http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.
[38] W. Whicher, *A dog hip problem*, 2002. Available at http://www.essortment.com/all/doghipproblem_ryoe.htm.
[39] E.M. Wijsman and C. Amos, *Genetic analysis of simulated oligogenic traits in nuclear and extended pedigrees: Summary of GAW10 contributions*, Genet. Epidemiol. 14 (1997), pp. 719–735.
[40] E.M. Wijsman and D. Yu, *Joint oligogenic segregation and linkage analysis using Bayesian Markov Chain Monte Carlo methods*, Mol. Biotechnol. 28 (2004), pp. 205–223.
[41] M. Willis, *Hip dysplasia*, in *Genetics of the Dog*, M.B.Willis, ed., Howell Book House, New York, NY, 1989, pp. 144–179.
[42] Z. Zhang, L. Zhu, J. Sandler, S. Friedenberg, J. Egelhoff, A. Williams, N. Dykes, W. Hornbuckle, U. Krotscheck, N. Moise, G. Lust, and R. Todhunter, *Estimation of heritabilities, genetic correlations, and breeding values of four traits that collectively define hip dysplasia in dogs*, Am. J. Vet. Res. 70 (2009), pp. 483–491
[43] Z. Zhou, X. Sheng, Z. Zhang, K. Zhao, L. Zhu, G. Guo, S.G. Friedenberg, L.S. Hunter, W.S. Vanderberg-Foels, W.E. Hornbuckle, U. Krotscheck, E. Corey, N.S. Moise, N.L. Dykes, J. Li, S. Xu, L. Du, Y. Wang, J. Sandler, G.M. Acland, G. Lust, and R.J. Todhunter, *Differential genetic regulation of canine hip dysplasia and osteoarthritis*, PLoS ONE 5(10) (2010), p. e13219. doi: 10.1371/journal.pone.0013219.

[44] L. Zhu, Z. Zhang, F. Feng, P. Schweitzer, J. Phavaphutanon, M. Vernier-Singer, E. Corey, S. Friedenberg, R. Mateescu, A. Williams, G. Lust, G. Acland, and R. Todhunter, *Single nucleotide polymorphisms refine QTL intervals for hip joint laxity in dogs*, Anim. Genet. 39 (2008), pp. 141–146

[45] L. Zhu, Z. Zhang, S. Friedenberg, S. Jung, J. Phavaphutanon, M. Vernier-Singer, E. Corey, R. Mateescu, N. Dykes, J. Sandler, G. Acland, G. Lust, and R. Todhunter, *The long (and winding) road to gene discovery for canine hip dysplasia*, Vet. J. 181 (2009), pp. 97–100.