

Analysis of Allele Fidelity, Polymorphic Information Content, and Density of Microsatellites in a Genome-Wide Screening for Hip Dysplasia in a Crossbreed Pedigree

R. G. MATEESCU, Z. ZHANG, K. TSAI, J. PHAVAPHUTANON, N. I. BURTON-WURSTER, G. LUST, R. QUAAS, K. MURPHY, G. M. ACLAND, AND R. J. TODHUNTER

From the Baker Institute for Animal Health (Mateescu, Burton-Wurster, Lust, and Acland) and the Department of Clinical Sciences (Todhunter), College of Veterinary Medicine, Cornell University, NY 14853; the Department of Animal Science, College of Agriculture and Life Sciences, Cornell University, NY 14853 (Zhang, Quaas); the Department of Veterinary Pathobiology, College of Veterinary Medicine, Texas A&M University, TX 77845 (Tsai, Murphy); and the Department of Companion Animal, Faculty of Veterinary Medicine, Kasetsart University, Nakhon-Pathom 73140, Thailand (Phavaphutanon).

Address correspondence to Rory Todhunter, Box 32, College of Veterinary Medicine, Cornell University, Ithaca NY 14853, or e-mail: rjt2@cornell.edu.

Abstract

Recent advances in genomics resources and tools are facilitating quantitative trait locus mapping. We developed a crossbreed pedigree for mapping quantitative trait loci for hip dysplasia in dogs by crossing dysplastic Labrador Retrievers and normal Greyhounds. We show that one advantage to using a crossbreed pedigree is the increased marker informativeness in the backcross/ F_2 population relative to the founder populations. We also discuss three factors that affect the detection power in the context of this crossbreed pedigree: being able to detect and correct genotyping errors, increasing marker density for chromosomes with a sparse coverage, and adding individuals to the mapping population as soon as they become available.

Our interest is in the clinical characterization and genetic basis for canine hip dysplasia. To dissect the underlying genetics of this common heritable trait in dogs, an experimental pedigree was established in 1994 by crossing unaffected Greyhounds and dysplastic Labrador Retrievers. This pedigree includes over 150 dogs spanning four generations of backcrosses and intercrosses (Todhunter et al. 1999, 2003). Clinical assessment of hips is determined using multiple diagnostic approaches when dogs are 4 and 8 months of age (Bliss et al. 2002; Lust et al. 2001). A genome-wide screen was undertaken on this crossbreed pedigree in collaboration with the Mammalian Genotyping Service (Marshfield, WA). At the time of this undertaking, Minimal Screening Set 1 (172 markers; Richman et al. 2001) and the canine genetic map of 2001 (Breen et al. 2001) were used to develop a 240-marker set with microsatellite markers to be resolved using a gel-based system.

The identification of genes contributing to variation in canine hip dysplasia requires genetic data of high fidelity. The genotypes on 147 crossbreeds at 247 loci were assessed for correct Mendelian inheritance patterns and repeatability. Genotyping errors occur when the observed genotype does not correspond to the true underlying genetic information as a result of a mistake in data entry or a misinterpretation of the pattern on a gel. Even a small number of genotyping errors can have negative consequences, increasing the estimated recombination fraction (Terwilliger et al. 1990) and reducing the evidence for linkage (Abecasis et al. 2001; Goring and Terwilliger 2000).

In this report, we describe the method used to identify reading errors in marker allelic size and the approach taken to correct such errors. We also present a comparison of the microsatellite informativeness between the pure breeds and the crossbreed dogs. We further show, using one

chromosome as an example, how correcting genotyping errors, increasing the number of markers, and adding more dogs to the analysis affects quantitative trait locus (QTL) mapping resolution.

Materials and Methods

Pedigree

Selected as founders for our pedigree were seven Greyhounds (two males and five females) with excellent hip conformation chosen from racing stock, along with seven Labrador Retrievers (three males and four females) with hip dysplasia and secondary hip osteoarthritis and one female Labrador Retriever with an intermediate phenotype but from a dysplastic lineage (Todhunter et al. 1999). The Greyhound is one of the few breeds in which hip dysplasia is rare, as evidenced by hip joint conformation consistently scored as "perfect" (Beling et al. 1975) or "normal" (PennHIP hip registry, Malvern, PA; Cardinet et al. 1983). The Greyhound founders were assumed to be homozygous for alleles protective against hip dysplasia, and the dysplastic Labrador Retrievers were assumed to be homozygous at the loci contributing to hip dysplasia for statistical mapping purposes. The crossbred pedigree consists of four generations (Greyhound and Labrador founders, F₁, backcrosses to both founders, and F₂ individuals) comprising 159 dogs.

DNA and Genotyping

DNA was isolated from peripheral blood by phenol-chloroform extraction. Gel-based electrophoretic separation was used to size microsatellite alleles at the NHLBI Mammalian Genotyping Service, Marshfield Medical Research Foundation (Marshfield, WA; Weber and Broman 2001) and capillary electrophoresis was used at Cornell University at the Bioresource Center on an Applied Biosystems Incorporated 3730 sequencer (Foster City, CA).

Microsatellites

From the integrated canine genetic map (Breen et al. 2001), we selected 240 microsatellite markers, 142 from the linkage map and 98 from the radiation hybrid (RH) map (Lou et al. 2003). Seven additional microsatellite markers were identified for CFA37 (personal communication with E. Kirkness, Institute for Genomic Research, Rockville, MA). The order and spacing of markers on this chromosome were calculated based on meiotic recombination in our pedigree using multipoint analysis available from MULTIMAP (Matise et al. 1993). This marker set covers approximately 80% of the estimated length of the canine genome (2.4 Gb).

Allele Check

The QTL mapping programs require marker genotype data in a pedigree to follow the Mendelian inheritance principles. Therefore, the first step in data editing is to check for Mendelian inheritance errors. We developed a program

(ZZ, <http://www.people.cornell.edu/pages/zz19/research/genoped>) that checks the genotypes in the pedigree for inconsistencies between parents and offspring.

Allele Correction

The second step in data editing is correcting these errors. A program to accomplish this task (RM) identifies all possible alleles for a given marker in the grandparent's generation, assumes that these alleles are measured without errors, follows each allele through the pedigree, and checks for inconsistencies within a narrow interval. Finally, inconsistencies are corrected. The range was defined as ± 2 bp relative to the grandparent allele size for the tetranucleotide and ± 1 bp for di- and trinucleotide repeat microsatellites. These errors were corrected by calling the alleles that differ by 1 or 2 bp as one allele. The errors outside these ranges were left uncorrected, resulting in missing data.

Polymorphic Information Content

To assess the quality of marker genotype data in our crossbred pedigree, the total number of alleles and the mean number of alleles per locus were determined for each marker for the two founder groups, the F₁ breeders, and the backcross/F₂ population. The differences in observed allele frequencies for pairwise comparisons between the Labrador Retriever and Greyhound founders were tested at each locus. Because the sample size is small relative to the size of marker contingency tables, a chi-square test may not be valid; thus, Fisher's exact test is a more appropriate test of no association.

The polymorphism information content (PIC) for each marker was determined separately for the four groups of animals using the following equation:

$$PIC = 1 - \sum_{i=1}^n p_i^2 - 2 \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i^2 p_j^2 \right]$$

where p_i is the frequency of the i^{th} allele, and n is the number of alleles (Botstein et al. 1980).

To analyze the change in the PIC between the four different breed groups, the following statistical model was used:

$$Y_{ij} = \mu + \text{Breed}_i + e_{ij}$$

where Y_{ij} = PIC for marker j in breed i ; μ = overall mean; Breed_i = breed (i = Labrador Retriever, Greyhound, F₁, backcross); e_{ij} = error associated with ij^{th} observation, assumed to be normally distributed $N(0, \sigma_e^2)$.

QTL Mapping

QTL mapping was performed using a regression approach originally described by Haley and Knott (1992). A web-based version of this tool is available (Seaton et al. 2002). The software, QTL Express (<http://latte.cap.ed.ac.uk/>), analyzes data from different mating schemes, including the combined backcross/F₂ design employed for our crossbred pedigree. Hip traits measured and analyzed included the distraction index (DI; a measure of hip laxity), the dorsolateral subluxation (DLS) score, and the Norberg angle. Chromosome-wide

Table 1. Distribution of genotyping errors on initial and corrected data following a genome-wide screen with 247 markers on 159 dogs from a crossbreed pedigree

| CFA | Markers (#) | Raw data (n = 147) | | Corrected data (n = 159) | |
|-------|-------------|-----------------------|----------|-----------------------------|----------|
| | | Errors | Rate (%) | Errors | Rate (%) |
| 1 | 11 | 54 | 3.5 | 8 | 0.46 |
| 2 | 11 | 101 | 6.38 | 15 | 0.86 |
| 3 | 9 | 85 | 6.6 | 22 | 1.54 |
| 4 | 8 | 118 | 10.17 | 45 | 3.54 |
| 5 | 10 | 40 | 2.83 | 7 | 0.44 |
| 6 | 6 | 80 | 9.43 | 18 | 1.89 |
| 7 | 10 | 180 | 12.57 | 15 | 0.94 |
| 8 | 6 | 52 | 6.06 | 2 | 0.21 |
| 9 | 7 | 37 | 3.65 | 34 | 3.05 |
| 10 | 7 | 46 | 4.65 | 26 | 2.34 |
| 11 | 7 | 82 | 8.18 | 20 | 1.80 |
| 12 | 9 | 63 | 4.89 | 5 | 0.35 |
| 13 | 5 | 1 | 0.14 | 3 | 0.38 |
| 14 | 7 | 70 | 7.09 | 40 | 3.59 |
| 15 | 7 | 27 | 2.69 | 2 | 0.18 |
| 16 | 4 | 12 | 2.09 | 1 | 0.16 |
| 17 | 5 | 29 | 4.18 | 2 | 0.25 |
| 18 | 7 | 49 | 4.97 | 12 | 1.08 |
| 19 | 5 | 9 | 1.29 | 9 | 1.13 |
| 20 | 5 | 55 | 7.74 | 54 | 6.79 |
| 21 | 5 | 51 | 7.31 | 14 | 1.76 |
| 22 | 6 | 58 | 6.84 | 13 | 1.36 |
| 23 | 6 | 21 | 2.45 | 3 | 0.31 |
| 24 | 4 | 5 | 0.88 | 5 | 0.79 |
| 25 | 6 | 95 | 11.14 | 5 | 0.52 |
| 26 | 5 | 14 | 1.96 | 0 | 0 |
| 27 | 6 | 13 | 1.50 | 7 | 0.73 |
| 28 | 6 | 19 | 2.21 | 11 | 1.15 |
| 29 | 4 | 3 | 0.53 | 3 | 0.47 |
| 30 | 7 | 24 | 2.35 | 3 | 0.27 |
| 31 | 5 | 5 | 0.69 | 4 | 0.50 |
| 32 | 4 | 72 | 12.83 | 7 | 1.10 |
| 33 | 5 | 53 | 7.36 | 10 | 1.26 |
| 34 | 4 | 24 | 4.15 | 0 | 0 |
| 35 | 4 | 3 | 0.54 | 0 | 0 |
| 36 | 2 | 3 | 1.05 | 3 | 0.94 |
| 37 | 13 | 114 | 16.08 | 0 | 0 |
| 38 | 4 | 0 | 0 | 0 | 0 |
| X | 5 | 0 | 0 | 0 | 0 |
| Total | 247 | 1,788 | 4.92 | 522 | 1.07 |

significance thresholds for each trait were determined (Churchill and Doerge 1994); the threshold at $P < .05$ and $P < .01$ was obtained from 1,000 permutations. Chromosomes 4, 9, 10, 11 ($P < .01$), 16, 20, 22, 25, 29 ($P < .01$), 30, 35, and 37 were identified to harbor putative QTL for one or more traits (Todhunter et al. in press).

Results

Allelic Size Error Detection

The error rate was 4.92% (out of 36,309 genotypes) for all 247 markers (Table 1). The highest average error rate, 16.08%, was for chromosome 37. Average error rates for chromosome 4, 7, 25, and 32 were more than 10%. Fourteen

Table 2. Descriptive statistics including number of individuals (n); the mean and maximum number of alleles per marker; the median; the 25% and 75% quartiles; and the range for the polymorphism information content (PIC) of 247 microsatellite markers screened on 159 dogs in an experimental canine pedigree representing the Labrador Retriever (L) and Greyhound (G) founders, as well as the F₁ and backcross (BC) generations

| | n | No. of alleles | | PIC | | | |
|----------------|-----|----------------|---------|--------|------|------|--------|
| | | Mean/marker | Maximum | Median | 25% | 75% | Range |
| L | 9 | 3.61 | 8 | 0.54 | 0.37 | 0.65 | 0–0.84 |
| G | 7 | 4.1 | 9 | 0.58 | 0.39 | 0.70 | 0–0.86 |
| F ₁ | 7 | 4.3 | 10 | 0.62 | 0.46 | 0.72 | 0–0.88 |
| BC | 136 | 5.58 | 16 | 0.63 | 0.46 | 0.71 | 0–0.89 |

other chromosomes had one or more markers with error rates more than 30% (data not shown). Only chromosome 38 and the X chromosomes were free of marker error. The marker with the highest error rate was FH2532 on CFA37, with 68% errors, which alone accounted for 89% of all genotyping errors associated with this chromosome (data not shown). Tetranucleotide markers had a disproportionate number of errors compared to the di- and trinucleotides. The error rate for each chromosome after running the allelic size correction program is shown in Table 1. The average error rate for the corrected data was 1.07%, and the total number of errors was reduced by 70%. All chromosomes, except CFA20, have an average error rate less than 4%, and five chromosomes have zero errors.

Marker Informativeness

Summary statistics for PIC are listed in Table 2, including the mean number of alleles per locus; the median; the 25% and 75% quartiles; and the range for the two founder breeds and the F₁ and backcross generations. The marker distribution by number of alleles in the crossbreed pedigree is shown in Figure 1. A decrease in proportion of markers with fewer than three alleles and an increase in those with more than three alleles are observed in the F₁ and backcross generations as compared to the two founder breeds. The increase in the number of alleles per microsatellite locus translates into an increase in the PIC.

One of the advantages of using a crossbreed pedigree is the increased informativeness of the markers for linkage mapping, which is associated with an increase in proportion of moderately and highly informative markers in F₁ and backcross generations relative to the founders (Table 3). Approximately 60% of the markers in the backcross generation demonstrated high information content, compared to only 39% in the Labrador Retriever founders and 48% in the Greyhound founders.

The differences in observed allelic frequencies for pairwise comparisons between the Labrador Retriever and Greyhound founders were tested at each locus. In all, 4 out of 247 markers (REN193A22, REN166C13, AHTH134Ren, RENo2C20) were monomorphic in both founder breeds. For 162 of the remaining 243 markers, the allelic frequencies

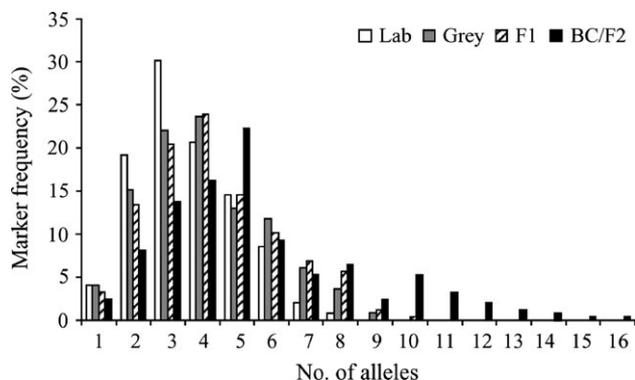


Figure 1. Marker distribution versus number of alleles in the Labrador Retriever (Lab) and Greyhound (Grey) founders and the F₁ and backcross/F₂ (BC/F₂) populations.

were significantly different between the two founder populations (the Fisher test for 127 markers had $P < .01$ and for 35 markers had $.01 < P < .05$). The most prominent markers (with highest Fisher test statistic) were REN41D20, REN150M24, REN130F03, LEI002, and FH2060. We also evaluated the difference in numbers of alleles between the two founder breeds for each marker. In the Greyhounds, 24.4% of the markers had up to three fewer alleles; 28% had the same number; and 47.6% had up to six more alleles relative to the Labrador Retriever founders.

The least squares means and standard deviation for PIC in the two founder breeds, F₁, and backcross generations are shown in Table 4. Pairwise comparison of PIC least squares means shows that the two founder breeds were not different ($P = .1$). However, PIC was significantly higher in F₁ and backcross individuals ($P < .01$) relative to the Labrador Retriever founders and in backcross individuals ($P < .01$) relative to the Greyhound founders.

Factors Affecting Mapping Power on CFA06

CFA06 is used as an example to illustrate the importance of three factors in detecting QTL: correction of genotype errors, marker density, and number of observations. Six markers

Table 3. Descriptive statistics for 247 microsatellite markers in the Labrador Retriever (L), Greyhound (G), and F₁ and backcross/F₂ (BC/F₂) populations

| | Range of PIC values | | |
|-------------------|-------------------------|------------------------------------|------------------------------|
| | < 0.30 Uninformative | 0.3–0.59 Moderately informative | > 0.60 Highly informative |
| L | 37 (15.04%) | 114 (46.34%) | 95 (38.62%) |
| G | 34 (13.82%) | 94 (38.21%) | 118 (47.97%) |
| F ₁ | 27 (10.98%) | 90 (36.58%) | 129 (52.44%) |
| BC/F ₂ | 22 (8.94%) | 76 (30.89%) | 148 (60.16%) |

PIC = polymorphism information content.

Table 4. Least square means (LSM) and standard deviation (SD) for polymorphism information content (PIC) in the Labrador Retriever founders (L), Greyhound founders (G), and F₁ and backcross/F₂ (BC/F₂) populations

| | PIC | | Het | |
|-------------------|------|------|------|------|
| | LSM | SD | LSM | SD |
| L | 0.50 | 0.19 | 0.55 | 0.20 |
| G | 0.53 | 0.21 | 0.58 | 0.21 |
| F ₁ | 0.57 | 0.20 | 0.62 | 0.20 |
| BC/F ₂ | 0.58 | 0.20 | 0.63 | 0.19 |

covering 45.5 cM of CFA06 were initially genotyped on 147 dogs from the crossbreed pedigree. In checking the data for Mendelian inheritance errors, 87 errors were identified for markers FH2164 and FH2561 (33 and 54, respectively). To analyze the data using QTL Express, all errors were changed to missing values (QTL Express requires genotypes at each marker to be consistent with Mendelian inheritance). An interval mapping analysis was performed, and the values of the *F*-test for three traits (DI on the left side and DLS score on both the right and the left side) at each location along CFA06 are plotted in Figure 2A. From these results, we conclude that no QTL on CFA06 affects any of these traits.

In the next step, errors for markers FH2164 and FH2561 were corrected using our allelic size correction program. The 33 errors for marker FH2164 and 53 of the 54 errors for marker FH2561 were successfully resolved. The remaining error for marker FH2561 is likely a mutation; therefore, it was set as a missing value. Interval mapping analysis was performed again, using the corrected data set, and the plot of the *F*-test statistic is shown in Figure 2B. Note the rise of the *F*-test value for all three traits at the right telomeric end of the chromosome, with the *F*-test for the left DI trait approaching significance. These results suggest that a QTL around the 45 cM position might reside on this chromosome. Note that most of the corrected errors (53) were for FH2561 marker, which is located at 45 cM.

The approximate length of CFA06 is 87 cM, and the aforementioned six markers provided coverage for only half of the chromosome. Moreover, the distribution of the six markers was not optimal, because of a large interval of 22 cM between markers REN149M14 and FH2561.

The most recent version of the canine map contains 3,300 markers at 1 Mb resolution (Guyon et al. 2003). That work included a description of Minimal Screening Set 2, which comprised 327 markers. To expedite linkage studies and positional cloning efforts, the second set was multiplexed into chromosome-specific panels (Clark et al. 2004). To better characterize this chromosome and to be able to confirm or reject the possibility of a QTL, seven more markers from this set were genotyped on the same 147 individuals. The coverage of CFA06 was extended to 74.4 cM by the addition of three markers at the right telomeric end of the chromosome. The coverage was also improved by adding two highly informative

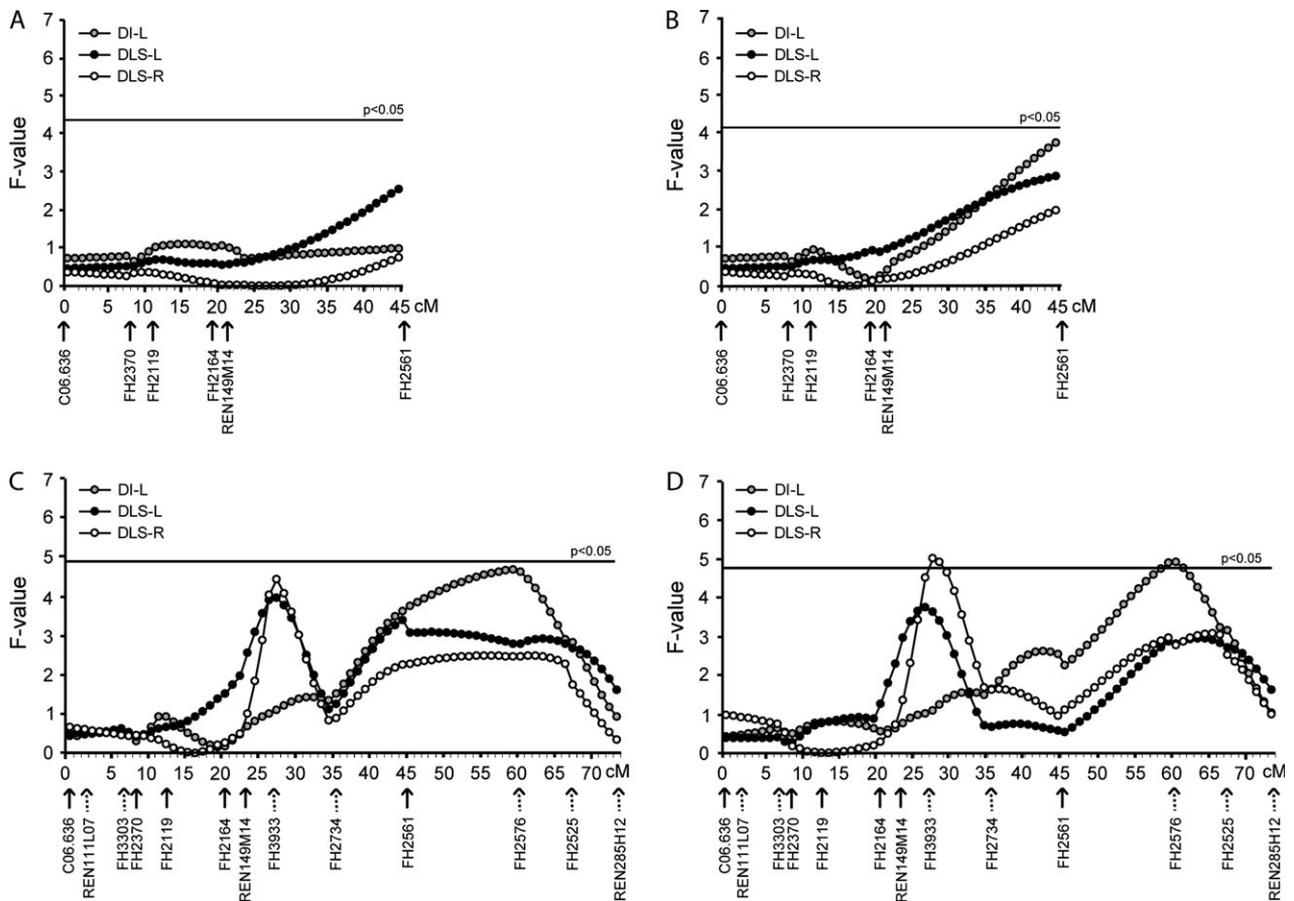


Figure 2. *F*-value plots for left distraction index (DI) and left (L) and right (R) dorsolateral subluxation (DLS) scores from interval mapping on CFA06. The position of the initial six markers is indicated by the solid arrows, and the position of the additional seven markers, by interrupted arrows. **(A)** Mendelian-inheritance errors on markers FH2164 (33) and FH2561 (54) were transformed into missing values. No *F* statistic reached the chromosome-wide threshold indicated by the horizontal bar. **(B)** Mendelian-inheritance errors on markers FH2164 (33) and FH2561 (54) were corrected. No *F* statistic reached the chromosome-wide threshold indicated by the horizontal bar, but there was an increase in *F*-test values at the far right end of the chromosome. **(C)** Interval mapping after correction of Mendelian inheritance errors on markers FH2164 and FH2561 and addition of seven markers. The *F* statistic for the three traits are approaching the chromosome-wide threshold. **(D)** Interval mapping after correction of Mendelian inheritance errors on markers FH2164 and FH2561 and addition of 7 markers and 12 backcross individuals. The *F* statistics for the left DI and right DLS scores exceed the threshold of significance at $\alpha = .05$.

markers in the 22 cM interval between markers REN149M14 and FH2561. A complete description of the 13 markers is provided in Table 5. Interval mapping analysis was performed using the new data set, and the plot of the *F*-test statistic is shown in Figure 2C. The addition of the three markers at the telomeric end of the chromosome resulted in a peak definition for the left DI at 60 cM with an *F*-test value close to the chromosome-wide significance threshold of .05. This result supports the presence of a QTL for this trait suggested by the previous analysis using the initial six markers.

Interestingly, this analysis suggests a QTL at 28 cM on CFA06 for both left and right DLS scores. A closer examination of the markers in this chromosomal region shows that the only marker available in this area from the initial set of six markers is not very informative (REN149M14), having only

three alleles segregating in the population and a PIC of 0.54. This explains why the analysis with the first six markers showed no QTL present in this region. The addition of two highly informative markers at position 27 and 35 cM (FH3933 with a PIC = 0.73 and FH2734 with a PIC = 0.66) revealed the presence of a QTL in this region.

This work was expanded with the addition of 12 new backcross individuals that were genotyped for all 13 markers. To determine whether this addition strengthened the data, interval mapping analysis was performed using all 159 individuals, and the plot of the *F*-test is shown in Figure 2D. The addition of 12 individuals increased the peak *F* value for both the left DI and the right DLS score above the threshold of significance at $\alpha = .05$ and refined the definition of the peak for the left DI.

Table 5. Location (cM), polymorphism information content (PIC), and number of alleles for 13 markers on CFA06 on the first 147 individuals and on the final 159 individuals from the crossbreed pedigree

| Marker | cM | <i>n</i> = 147 dogs | | <i>n</i> = 159 dogs | |
|-----------|------|---------------------|----------|---------------------|----------|
| | | PIC | # allele | PIC | # allele |
| C06.636 | 0 | 0.13 | 3 | 0.13 | 3 |
| REN111L07 | 3.6 | 0 | 1 | 0 | 1 |
| FH3303 | 7.4 | 0.86 | 12 | 0.86 | 12 |
| FH2370 | 8.8 | 0.80 | 10 | 0.85 | 11 |
| FH2119 | 12.3 | 0.75 | 5 | 0.76 | 7 |
| FH2164 | 20.4 | 0.71 | 6 | 0.72 | 8 |
| REN149M14 | 23.4 | 0.54 | 3 | 0.54 | 3 |
| FH3933 | 27.5 | 0.73 | 8 | 0.73 | 8 |
| FH2734 | 35.2 | 0.66 | 4 | 0.66 | 4 |
| FH2561 | 45.5 | 0.81 | 9 | 0.80 | 11 |
| FH2576 | 60.6 | 0.81 | 13 | 0.81 | 13 |
| FH2525 | 67.1 | 0.62 | 7 | 0.62 | 7 |
| REN285H12 | 74.4 | 0.18 | 2 | 0.18 | 2 |

These results suggest that a QTL is segregating on CFA06 for both the DI and the DLS score. This example is a clear illustration of the increase in power to detect QTL obtained by carefully correcting genotyping errors, by strategically adding markers to provide a better coverage of the chromosome, and by adding individuals in the analysis.

Discussion

Genotyping errors occur when observed genotypes do not reflect the true alleles. These errors arise from mistakes in data entry, sample mishandling, or errors introduced by the genotyping process itself (Ewen et al. 2000). Genotyping errors are detrimental to linkage analyses (Cherny et al. 2001; Goldstein et al. 1997); therefore, the identification of errors is necessary for accurate analysis of data. Although large genotyping data sets will likely contain errors, and the inclusion of incorrect data can result in erroneous conclusions (Abecasis et al. 2001; Terwilliger et al. 1990), little attention is generally given to correcting the identified errors. The cleaning of genotype data should be an integral and important component of a successful genome scan for QTL detection. Too often, data editing is limited to identifying Mendelian inheritance errors and to changing the markers with errors into missing values in order to run the analysis. Described herein is the importance of detecting genotyping errors, even those with modest error rates. In our study, an error rate of 4.87% was detected following a whole-genome screen with 247 microsatellite markers on 147 dogs from a crossbreed pedigree and was subsequently reduced to 1.07% with the use of an allele-correction program. Correction of genotyping errors increased the detection power of linkage analysis and allowed detection of an underlying QTL on CFA06 that did not show initial evidence for harboring QTL.

A major advantage in using a crossbreed pedigree is the increased informativeness of the markers for linkage analysis of complex traits. Evidence for this in our pedigree is the increase in proportion of markers with higher numbers of

alleles in the F₁ and backcross/F₂ generations as compared to those of the two founder breeds.

Fine mapping of putative QTL is time-consuming and costly. Therefore, it is useful to exclude, with a high level of confidence, certain chromosomes and chromosomal regions. The marker set used in this study was chosen to provide optimal coverage of the canine genome. Even so, intervals between some markers exceeded 10 cM. Addition of more markers in these regions will improve the mapping resolution or result in a definite exclusion of these chromosomal regions for finer mapping. Also, the addition of highly informative markers in a 22 cM interval on CFA06 allowed us to map a QTL at a chromosome-wide significance level of 0.05.

One of the major issues when envisaging positional cloning of the mapped QTL is the poor mapping resolution that is typically achieved after the initial genome-wide screen. Confidence intervals for the location of the QTL are of the order of 20 to 30 cM and typically contain as many as 500 to 1,000 genes. This is due to the limited number of recombination events in the available pedigree and the fact that many QTL effects are likely to reflect the combined action of multiple linked genes. The resolution of QTL mapping is limited by the information gained from observing the genotypic states of the markers (Darvasi et al. 1993). The observed recombinants can be limited by small sample size and missing genotypic data. In one example, the addition of 12 backcross individuals increased the detection power of linkage analysis and helped to narrow the QTL region on CFA06.

In conclusion, by maximizing the information for QTL mapping, even in a less-than-optimum genome-wide screen, it is possible to detect putative QTL in dog pedigrees. This can be achieved by creating a crossbreed pedigree, which will result in increased marker informativeness. Detection and correction of genotyping errors is a crucial step in ensuring that maximum information is extracted from available data. Adding markers in regions with poor coverage, as well as adding new individuals as they become available, would also increase the power of the analysis.

Acknowledgments

We thank Elizabeth Corey, Margaret Vernier-Singer, and Alma J. Williams for excellent technical assistance. Support was provided by the Morris Animal Foundation; Collaborative Grant Program, College of Veterinary Medicine and Biotechnology Center for Advanced Technology, Cornell University; Pfizer; American Border Collie Association; Van Sloun Foundation; and Mammalian Genotyping Service. This article was presented at the second international conference on the "Advances in Canine and Feline Genomics: Comparative Genome Anatomy and Genetic Disease," Universiteit Utrecht, Utrecht, Netherlands, October 14–16, 2004.

References

- Abecasis GR, Cherny SS, and Cardon LR, 2001. The impact of genotyping error on family-based analysis of quantitative traits. *Eur J Hum Genet* 9: 130–134.

- Beling CG, Gustafsson PO, and Kasstrom H, 1975. Metabolism of estradiol in greyhounds and German shepherd dogs. An investigation with special reference to hip dysplasia. *Acta Radiol Suppl* 344:109–120.
- Bliss S, Todhunter RJ, Quaas R, Casella G, Wu R, Lust G, Williams AJ, Hamilton S, Dykes NL, Yeager A, Gilbert RO, Burton-Wurster NI, and Acland GM, 2002. 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* 63:1029–1035.
- Botstein D, White RL, Skolnick M, and Davis RW, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331.
- Breen M, Jouquand S, Renier C, Mellersh CS, Hitte C, Holmes NG, Cheron A, Suter N, Vignaux F, Bristow AE, Priat C, McCann E, Andre C, Boundy S, Gitscham P, Thomas R, Bridge WL, Spriggs HF, Ryder EJ, Curson A, Sampson J, Ostrander EA, Binns MM, and Galibert F, 2001. Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. *Genome Res* 11:1784–1795.
- Cardinet GH III, Guffy MM, Wallace LJ, and Laben RC, 1983. Canine hip dysplasia in German shepherd dog-greyhound crossbreeds. *J Am Vet Med Assoc* 182:393–395.
- Cherny SS, Abecasis GR, Cookson WO, Sham PC, and Cardon LR, 2001. The effect of genotype and pedigree error on linkage analysis: analysis of three asthma genome scans. *Genet Epidemiol* 21 (suppl. 1):S117–S122.
- Churchill GA and Doerge RW, 1994. Empirical threshold values for quantitative trait mapping. *Genetics* 138:963–971.
- Clark LA, Tsai KL, Steiner JM, Williams DA, Guerra T, Ostrander EA, Galibert F, and Murphy KE, 2004. Chromosome-specific microsatellite multiplex sets for linkage studies in the domestic dog. *Genomics* 84:550–554.
- Darvasi A, Weinreb A, Minke V, Weller JI, and Soller M, 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* 134:943–951.
- Ewen KR, Bahlo M, Treloar SA, Levinson DF, Mowry B, Barlow JW, and Foote SJ, 2000. Identification and analysis of error types in high-throughput genotyping. *Am J Hum Genet* 67:727–736.
- Goldstein DR, Zhao H, and Speed TP, 1997. The effects of genotyping errors and interference on estimation of genetic distance. *Hum Hered* 47:86–100.
- Goring HH and Terwilliger JD, 2000. Linkage analysis in the presence of errors II: marker-locus genotyping errors modeled with hypercomplex recombination fractions. *Am J Hum Genet* 66:1107–1118.
- Guyon R, Lorentzen TD, Hitte C, Kim L, Cadieu E, Parker HG, Quignon P, Lowe JK, Renier C, Gelfenbeyn B, Vignaux F, DeFrance HB, Gloux S, Mahairas GG, Andre C, Galibert F, and Ostrander EA, 2003. A 1-Mb resolution radiation hybrid map of the canine genome. *Proc Natl Acad Sci USA* 100:5296–5301.
- Haley CS and Knott SA, 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315–324.
- Lou XY, Todhunter RJ, Lin M, Lu Q, Liu T, Wang Z, Bliss SP, Casella G, Acland GM, Lust G, and Wu R, 2003. The extent and distribution of linkage disequilibrium in a multi-hierarchical outbred canine pedigree. *Mamm Genome* 14:555–564.
- Lust G, Todhunter RJ, Erb HN, Dykes NL, Williams AJ, Burton-Wurster NI, and Farese JP, 2001. Comparison of three radiographic methods for diagnosis of hip dysplasia in eight-month-old dogs. *J Am Vet Med Assoc* 219:1242–1246.
- Matisse TC, Perlin M, and Chakravarti A, 1993. MultiMap: an expert system for automated genetic linkage mapping. *Proc Int Conf Intell Syst Mol Biol* 1:260–265.
- Richman M, Mellersh CS, Andre C, Galibert F, and Ostrander EA, 2001. Characterization of a minimal screening set of 172 microsatellite markers for genome-wide screens of the canine genome. *J Biochem Biophys Methods* 47:137–149.
- Seaton G, Haley CS, Knott SA, Kearsley M, and Visscher PM, 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* 18:339–340.
- Terwilliger JD, Weeks DE, and Ott J, 1990. Laboratory errors in the reading of marker alleles cause massive reductions in lod score and lead to gross overestimates of the recombination fraction. *Am J Hum Genet* 47(suppl.):A201.
- Todhunter RJ, Acland GM, Olivier M, Williams AJ, Vernier-Singer M, Burton-Wurster N, Farese JP, Grohn YT, Gilbert RO, Dykes NL, and Lust G, 1999. An outcrossed canine pedigree for linkage analysis of hip dysplasia. *J Hered* 90:83–92.
- Todhunter RJ, Casella G, Bliss SP, Lust G, Williams AJ, Hamilton S, Dykes NL, Yeager AE, Gilbert RO, Burton-Wurster NI, Mellersh CC, and Acland GM, 2003. Power of a Labrador Retriever-Greyhound pedigree for linkage analysis of hip dysplasia and osteoarthritis. *Am J Vet Res* 64:418–424.
- Todhunter RJ, Mateescu RG, Lust G, Burton-Wurster NI, Dykes NL, Bliss SP, Williams AJ, Vernier-Singer M, Corey E, Harjes C, Quaas RL, Zhang Z, Gilbert RO, Casella G, Wu R, and Acland GM, in press. Quantitative trait loci for hip dysplasia in a crossbreed canine pedigree. *Mamm Genome*.
- Weber JL and Broman KW, 2001. Genotyping for human whole-genome scans: past, present, and future. *Adv Genet* 42:77–96.

Corresponding Editor: Francis Galibert