

Canine Health Foundation  
Progress Report  
January 15, 2007

CHF Grant No. 600: Linkage Analysis of Craniomandibular Osteopathy  
(CMO) in Terrier Breeds

Principal Investigator: Patrick J. Venta, Ph.D.  
Microbiology & Molecular Genetics  
College of Veterinary Medicine  
Michigan State University  
East Lansing, MI 48824-4320

### **Introduction**

CMO (craniomandibular osteopathy) is an inherited non-cancerous overgrowth of bone primarily occurring in the jaw region of affected dogs, most of which belong to particular terrier breeds. The disease generally disappears about a year after the initial onset at several months of age. It is rarely fatal, but can cause pain to the puppy and a good deal of distress to the owner. We are continuing our efforts to identify this gene, so that breeders can identify carriers by a diagnostic test and then make informed breeding decisions to avoid producing affected offspring. We have excluded many candidate genes in previous work (including one early in 2005, *COL1A1*, a gene that produces a very similar phenotype in human patients [Gensure et al., 2005]), and random samplings of the genome. Part of this work has been previously published (Housley et al., 2004). The current work is to exclude the remaining regions of the genome using a systematic whole genome linkage approach.

### **Production of SINE markers**

Approximately 10,000 variable canine SINES were identified by Wang and Kirkness (2005) by comparison of the sequenced Boxer and Poodle genomes. We have designed PCR primers to amplify a subset of these SINES to cover two thirds of the canine genome at about 10 to 20 cM resolution. Originally the primers were designed manually, but we have now developed a semi-automated method that extracts the SINE and flanking DNA sequences, masks nearby repetitive DNA that would interfere with primer design, and allows the sequence to be imported into the Primer3 program for automated design. Because part of our strategy for automated SINE genotyping is dependent upon the presence of a BamHI restriction enzyme site in the allele containing the SINE, we also search for BamHI sites in the region flanking the SINE insertion point that might interfere with the assay. As expected, very few SINES have flanking BamHI sites, but those that do have been excluded from the marker set. As we mentioned in our grant proposal, we have been able to use supplementary internal funding to synthesize the PCR primer sets for these markers. So far, 164 SINE primer sets covering CFA1-18 have been

completely designed and purchased (68% of the genome), and SINEs for the remaining chromosomes have been selected but the design has not yet been completed.

### **Testing for marker variation in CMO parents**

Pedigree-based linkage analysis requires heterozygosity at both loci (i.e., CMO gene and marker) to be informative. Because pedigrees are collected (ascertained) because of the presence of CMO-affected offspring, the parents are obligate heterozygotes, and informativeness is thus dependent solely upon heterozygosity for the marker. To test for heterozygosity of the SINE markers, we look for the presence of both alleles in DNA pooled from the carrier parents. We have used pools from each of the three breeds separately. The SINEs from the work of Wang and Kirkness (2005) were identified using only two chromosomes; the one consensus from the sequence of the two chromosomes of the one Boxer dog, and the sequence traces from the Poodle genome project. Theoretical predictions based upon using only two chromosomes to identify polymorphisms suggest that half of such markers will have minor allele frequencies (MAFs) of greater than 0.25 and half will have MAFs less than 0.25 (Eberle and Kruglyak, 2000); this result, incidentally, is independent of marker type for biallelic markers). Because of drift in the relatively small effective population sizes of dog breeds, slightly fewer markers would be expected to have MAFs greater than 0.25 (e.g., Brouillette and Venta, 2002). From this it would be expected that somewhat less than half of the markers would be variable enough to provide good linkage information, and this is very close to what we have found to this point. We were, however, a bit surprised to find more variation in the MAF of the markers across the three breeds, particularly with respect to the Cairn and West Highland White terrier breeds, because these breeds have been reported to be closely related (that is, a marker might have good variability in one breed, but be monomorphic in another breed). This result may be due to very small effective population sizes during some point in the history of the breeds. Despite the slightly lower amount of variability than we had anticipated, it has still been sufficient given the number of pedigrees that we have to allow the exclusion of large regions of chromosomes for containing the CMO gene.

### **Linkage data on canine chromosomes 1, 2, and 3**

We are systematically scanning chromosomes in numerical order (that is, from largest to smallest). In order to minimize the cost of primers, our strategy is to use two successive rounds of primer design – the first to exclude most of a given chromosome, and the second to fill any gaps in coverage caused by uninformative (i.e., low or non-variable) SINE markers. With chromosome 1, all of the chromosome was excluded with the exception of two small gaps that will need to be filled. Most of the markers that we tested for this chromosome turned out to be useful. However, the markers for the first round of chromosomes 2 and 3 were not as informative as those used for chromosome 1 and, although good stretches have been excluded, a much greater amount of these two chromosomes will require the development of additional SINE markers. Although it is possible that these chromosomes might be less variable, overall, for biological reasons (e.g., selection for desirable characteristics that have caused loss of variability by selective sweeps) it seems more likely that the result of lower marker variability was simply random chance. We plan to fill the gaps on these chromosomes with a second

round of SINEs, and if variation is again low (which, again, we view as unlikely), we will use the SNPs that we have identified in other work to cover these chromosomes. In summary, the gene is not on CFA1 (although two small gaps remain to be filled), and it is not on roughly half of CFA2 or CFA3. We expect to exclude a large portion of the genome in the next two months.

#### **Automated SINE assay**

We have not been using the automated system up to this point because we wanted to have accurate, manually annotated genotyping as control material to verify that the automated system makes accurate calls. Although we could continue with the manual method, we believe that the automated system produces good genotyping data, and we will verify this in the next few weeks. Our original plan was to use a real-time PCR machine in the M.S.U. Research Technology Support Facility, located a few hundred yards from our lab, to scan our microtiter plates for the genotyping calls. However, our department (Microbiology & Molecular Genetics) has recently acquired a Typhoon phosphoimager unit (Amersham Biosciences) that is capable of scanning our plates and putting the data into an appropriate electronic spreadsheet format for import into our linkage analysis program. We will use this instrument because it is housed directly across the hall from our main lab, and because the use of the instrument is free to us, thus saving us time and money.

#### **Decision not to request approval for a change in experimental protocol**

There have been reports on the development of canine SNP genotyping microarrays during the past year. These arrays have the potential to allow rapid whole genome scanning for disease genes such as CMO. We had considered using this system in place of the SINE marker system that we originally proposed for the linkage analysis work for CMO. However, two factors make it seem more reasonable to continue with the SINE marker system. The first is that the microarrays (produced by the Affymetrix and potentially by Illumina corporations) require high quality DNA for accurate genotyping. The samples that we have are buccal (cheek) swabs, and are unlikely to permit direct, accurate genotyping with the SNP arrays. We had hoped that whole genome amplification (WGA) from the swabs would produce higher quality DNA. Unfortunately, we have been unable to produce any WGA DNA from the CMO swabs using commercially available WGA kits, even after several attempts. The second reason is cost. We calculate that the cost for the SNP chips would be somewhat over budget, and although we were prepared to find and use supplementary funding, any failures that might be caused by low quality DNA might put the project far over budget. For these reasons, we have decided to continue with the SINE markers which, despite the greater amount of effort required, have been consistently reliable for genotyping the CMO buccal swab DNA samples. It is also worth mentioning that although SNP chips may very well become the mainstream mapping method for canine genomics, SINEs may be useful markers for narrowing map intervals. This is because SNP chips must have fixed markers for each release (a fairly rare occurrence), whereas SINEs can be easily chosen in a given interval, and easily genotyped using the automated method that we have developed and are using for our linkage scan.

CMO Jan 2007

### **Change in personnel**

Dr. Donna Housley has been responsible for the genotyping assays. However, as a result of several failed grant proposals, she has left the lab (the CHF funding by itself is insufficient to support Dr. Housley full time). Dr. Venta has directly taken over her assigned tasks for the project although, of course, no salary support for him will be derived from the CHF funding. One and possibly two undergraduates will continue to help Dr. Venta with data collection. No change in budget categories is currently requested because it may be possible to hire additional undergraduate help for some of the work. If the salary budget is not used for personnel, we may at some point request permission to use the funds for additional supplies support.

### **Updated Timeframe**

Although we are on track or somewhat ahead of schedule for developing the SINE markers, we are behind in our schedule for testing the markers for linkage to the CMO gene. However, now that our plan is to use primarily the automated SINE genotyping system, we believe we should nearly be able to catch up to our schedule for coverage of the whole genome. It is worth noting that the timing of finding the linked marker is dependent upon random chance, given no a priori reason for suspecting any particular region of the genome for containing the CMO gene. Thus, there is a 50/50 chance of finding the gene after screening half of the genome, and in the case of an early discovery of a linked marker it would not be necessary to screen the complete genome. However, we do want to maintain the schedule in case we are extremely unfortunate and do not find the linkage until that the last one percent of the genome is scanned. Given that we may still have gap filling to do by the end of the funded year, it may be necessary to request a six-month extension to complete this phase of the work.

### **References**

Brouillette, J.A. and P.J. Venta. Within-breed heterozygosity of canine single nucleotide polymorphisms identified by across-breed comparison. *Anim. Genet.* 33:464-467 (2002).

Eberle, M.A. and L. Kruglyak. An analysis of strategies for discovery of single-nucleotide polymorphisms. *Genet. Epidemiol.* 19 (suppl 1):S29-35 (2000).

Gensure R.C., O. Makitie, C. Barclay, C. Chan, S.R. Depalma, M. Bastepe, H. Abuzahra, R. Couper, S. Mundlos, D. Silience, L. Ala Kokko, J.G. Seidman, W.G. Cole, and H. Juppner A novel COL1A1 mutation in infantile cortical hyperostosis (Caffey disease) expands the spectrum of collagen-related disorders. *J. Clin. Invest.* 115:1142-1144 (2005).

Housley, D.J.E., R. Ritzert, and P.J. Venta. Comparative radiation hybrid map of canine chromosome 1 incorporating SNP and indel polymorphisms. *Genomics* 84:248-264 (2004).

Wang, W. and E.F. Kirkness. Short interspersed elements (SINEs) are a major source of canine genomic diversity. *Genome Res.* 15:1798-1808 (2005).

CMO Jan 2007