FGF4 retrogene on CFA12 is responsible for chondrodystrophy and intervertebral disc disease in dogs

Emily A. Brown1, Peter J. Dickinson1, Tamer Mansour2, Beverly K. Sturges3, Miriam Aguilar4, Amy E. Young5, Courtney Korf6, Jenna Lind7, Cassandra L. Ettinger8, Samuel Varon9, Rachel Pollard9, C. Titus Brown9.a, Terje Raudsepp1, and Danika L. Bannasch2.a.d.1

1Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616; 2Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California, Davis, CA 95616; 3Department of Animal Science, University of California, Davis, CA 95616; 4Gene Center, University of California, Davis, CA 95616; 5Barney and Russum Animal Clinic, Fairfield, CA 94533; and 6Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843

Edited by Harris A. Lewin, University of California, Davis, CA, and approved September 18, 2017 (received for review June 2, 2017)

Chondrodystrophy in dogs is defined by dysplastic, shortened long bones and premature degeneration and calcification of intervertebral discs. Independent genome-wide association analyses for skeletal dysplasia (short limbs) within a single breed (PBonferroni = 0.01) and intervertebral disc disease (IVDD) across breeds (PBonferroni = 4.0 × 10⁻¹⁹) both identified a significant association to the same region on CFA12. Whole genome sequencing identified a highly expressed FGF4 retrogene within this shared region. The FGF4 retrogene segregated with limb length and had an odds ratio of 51.23 (95% CI = 46.69, 56.20) for IVDD. Long bone length in dogs is a unique example of multiple disease-causing retrocopies of the same parental gene in a mammalian species. FGF signaling abnormalities have been associated with skeletal dysplasia in humans, and our findings present opportunities for both selective elimination of a medically and financially devastating disease in dogs and further understanding of the ever-growing complexity of retrogene biology.

GWAS | inherited | genetic | dysplasia | chondrodysplasia

Variation in domestic dog (Canis familiaris, CFA) morphology has long fascinated both scientists and pet owners. Domestication of the dog from the wolf and the subsequent variation in size and shape within purebred dog breeds is a remarkable feat of animal breeding and selection. One of the most extreme examples of dog breed differences is in limb length, as extremely short limbs define many breeds. This morphological feature is present in breeds from all over the world and from all American Kennel Club (AKC) groups, indicating that the underlying genetic causes are likely very old.

Extensive examination of growth plates has been performed on many of these short-legged dog breeds (dachshund, Pekingese, French bulldog, spaniels, beagle), as these breeds are also prone to intervertebral disc disease (IVDD) (1–3). Histopathological analysis of the bones of puppies from these breeds demonstrated that their short stature is due to defects in endochondral ossification, the process whereby cartilage is replaced with bone, in the developing limb. The long bone growth plates show disorganization of the proliferative zone and reduction in the depth of the maturation zone (1–4). In addition to the long bones, similar but more subtle changes exist in endochondral ossification of the vertebral bodies (1, 2).

The intervertebral disc (IVD), which sits between vertebral bodies, is composed of an outer fibrous basket, called the annulus fibrosis, made of 70% collagen and an inner gel-like layer that is a remnant of the embryonic notochord, called the nucleus pulposus (5). Together, these structures and the cartilaginous endplates allow for flexibility of the vertebral column. In chondrodystrophic dogs, the nucleus pulposus is gradually replaced by chondrocyte-like cells in chondroid metaplasia (or metamorphosis) that occurs between birth and 1 y of age (1, 2). Recent studies have shown that in advanced stages of degeneration in nonchondrodystrophic dogs there is also replacement of notochordal cells by chondrocyte-like cells, similar to the changes observed in chondrodystrophic dogs, although this happens at an older age (3, 6–10). The replacement of the nucleus pulposus with chondrocyte-like cells is seen in humans, and chondrodystrophic breeds have been proposed as models for human degenerative disc disease (3, 7, 11, 12).

Hansen described the two different types of canine IVD prolapse as type I and type II. Type I occurs exclusively in chondrodystrophic breeds and is characterized by premature degeneration of all discs in young dogs. In contrast, type II occurs in older dogs and is usually limited to a single disc with only partial protrusion. In type I disc disease, the calcified nucleus pulposus may undergo an explosive herniation through the annulus fibrosus into the vertebral canal, resulting in inflammation and hemorrhage and causing severe pain and neurological dysfunction (1, 2). In veterinary hospital population studies, breeds with a significant increased risk of IVDV include the beagle, cocker spaniel, dachshund, French bulldog, Lhasa apso, Pekingese,

Significance

Chondrodystrophy, characterized by short limbs and intervertebral disc disease (IVDD), is a common phenotype in many of the most popular dog breeds, including the dachshund, beagle, and French bulldog. Here, we report the identification of a FGF4 retrogene insertion on chromosome 12, the second FGF4 retrogene reported in the dog, as responsible for chondrodystrophy and IVDD. Identification of the causative mutation for IVDD will impact an incredibly large proportion of the dog population and provides a model for IVDD in humans, as FGF-associated mutations are responsible for IVDD and short stature in human achondroplasia. This is a report of a second retrogene copy of the same parental gene, each causing complementary disease phenotypes in a mammalian species.


Conflict of interest statement: The University of California, Davis, has filed a provisional patent entitled: “Methods of Diagnosing Intervertebral Disc Disease and Chondrodystrophy in Canines,” on May 30, 2017.

This article is a PNAS Direct Submission.

This is an open access article distributed under the PNAS license.

Data deposition: The sequence data reported in this paper has been deposited in the Sequence Read Archive (SRA Bioproject no. PRJNA377155) and in the GenBank database (accession nos. MF040221 and MF040222).

To whom correspondence should be addressed. Email: dlbannasch@ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1709082114/-/DCSupplemental.
Pembroke Welsh corgi, and shih tzu (13–15). Pet insurance data suggests a conservative “lifetime prevalence” for IVDD in dogs of 3.5% in the overall population; however, in the chondrodystrophic breed with the highest risk, the dachshund, the lifetime prevalence is between 20–62% with a mortality rate of 24% (9, 10, 16–18). The effect of this disease on dogs and the financial burden to pet owners is enormous.

Skeletal dysplasia (SD), a general term to classify abnormalities of growth and development of cartilage and/or bone resulting in various forms of short stature, occurs in humans and dogs in many forms (19). With advances in molecular genetics, many of the diseases in humans are being reclassified based on the specific underlying causative mutations (20). To a lesser degree, progress has also been made in understanding the molecular nature of SD and the extreme interbreed limb-length variation observed in dogs (21–24). While the mutations causing some subtypes of SD in dogs have been determined, there are still many unexplained types of SD observed within and across dog breeds.

In 2009, the genetic basis for extreme differences in limb length in dogs was investigated by Parker et al. (25) using an across-breed genome-wide association approach. They determined that a FGF4 retrogene insertion on CFA18 ~25 Mb from the parental copy of the FGF4 locus was responsible for the “chondrodysplasia” phenotype in a number of breeds, such as the bassett hound, Pembroke Welsh corgi, and dachshund. However, the FGF4 retrogene insertion on CFA18 failed to explain breeds such as the American cocker spaniel, beagle, and French bulldog, that in addition to dachshunds, were the breeds originally classified as chondrodystrophic based on histopathological and morphological analysis by Hansen (1) and Braund (3). The FGF gene family has similarly been implicated in SD in humans, with mutations in FGF3 found to be responsible for achondroplasia, the most common form of dwarfism, characterized by shortened limbs and abnormal vertebrae and IVDs (20, 26–30). FGF genes are involved in a number of embryological development processes, and specific levels of ligand and receptor are key for appropriate growth and development (31–33).

In this study, genome-wide association analysis in a cohort of Nova Scotia duck tolling retrievers (NSDTRs) with and without severe SD identified a significant association on CFA12 due to a 12-Mb associated haplotype, of which 1.9 Mb was found to be shared in chondrodystrophic breeds. Subsequent genome-wide association analysis of Hansen’s type I IVDD across breeds localized the same 1.9-Mb region on CFA12, suggesting that the locus responsible for SD in the NSDTR is also responsible for type I IVDD and the chondrodystrophic phenotype across dog breeds. A previous genetic investigation of IVDD in dachshunds and limb-length morphology in Portuguese water dogs both identified the same CFA12 locus; however, neither study reported a causative mutation (34, 35). Here, using Illumina paired-end genomic sequencing, we uncover a second FGF4 retrogene insertion (chr12: 33.7 Mb [Canis familiaris (canFam) 3]) in the canine genome and show that it is not only responsible for SD in the NSDTR, but also chondrodystrophy, including the predisposition to Hansen’s type I IVDD, across all dog breeds.

Results

Genome-Wide Association Studies for SD and IVDD. A form of SD is common in the NSDTR and is characterized by variable decrease in limb length and associated abnormalities, including long-bone bowing, phyesal widening, and joint incongruity (Fig. 1 A and B). On physical examination, in addition to shorter limbs, SD dogs may also have valgus limb deformities and larger ears (pinnae). While SD is a common phenotype in the breed, the degree of severity is highly variable.

To determine a region of the genome associated with SD in the NSDTR, genome-wide association analysis was performed using 13 NSDTRs with severe SD and 15 NSDTR controls without severe SD. There were 41 single-nucleotide polymorphisms (SNPs) that were genome-wide significant with a P Bonferroni of <0.05, all present between chr12: 35,413,695 and 46,117,273 (top SNP chr12: 36,790,324 P Bonferroni = 0.01) (canFam2) (Fig. 1C and SI Appendix, Fig. S1).

Underlying this strong association for SD in NSDTRs was an ~12-Mb critical interval from chr12: 36–48 Mb (canFam2). Since the NSDTR SD phenotype is not uncommon in different dog breeds, we investigated haplotype sharing across breeds and observed that a portion of this associated haplotype was shared with two breeds of dog considered classically chondrodystrophic: the American cocker spaniel and beagle (1, 3). By plotting the minor allele frequency (MAF) across this interval for 7 American cocker spaniels, 14 beagles, and 13 SD-affected NSDTRs, the critical interval identified via GWAS for SD was shortened to a shared haplotype from chr12: 36.4–38.3 Mb (canFam2) (Fig. 2A).

To test the hypothesis that the same locus was responsible for SD and chondrodystrophy, a second genome-wide association study (GWAS) was performed using IVDD-affected cases (n = 36) and unaffected controls (n = 31) across 26 dog breeds (listed in SI Appendix, Table S6). The most highly associated SNP was located on CFA12 [chr12: 36,909,311 (canFam2)] with a P raw = 3.2 × 10−15, P Bonferroni = 4.0 × 10−10, and odds ratio of 32.67 (Fig. 2B). Observing linkage disequilibrium with the highest associated SNP using r2 values >0.06, the critical interval identified via GWAS for IVDD overlaps with that seen when mapping MAF across breeds and SD in the NSDTR (Fig. 2C).

Identification of FGF4 Insertion on CFA12. To identify a causative variant for SD and IVDD, paired-end whole-genome sequences of two cases, one SD-affected NSDTR and one IVDD-affected dachshund, and 83 unaffected controls were investigated in the associated interval. The average coverage for these samples was 8.7x. There were 9,156 SNP variants and 7,877 insertion/deletion (indel) variants identified from chr12: 33.1–35.5 Mb (canFam3) [chr12: 36.1–38.5 Mb (canFam2)]; however, none segregated with the IVDD phenotype. The same interval was also investigated by visual inspection of BAM files to flag mate
pairs with unusual insert sizes in an effort to identify any large indels. Using the two cases and two controls (one NSDTR and one Saluki) eight large indels (>200 bp) were identified within the interval (SI Appendix, Table S1). Four large indels did not segregate when investigated in additional control genomes, while the remaining four were eliminated after PCR showed lack of segregation between cases and controls.

Visual inspection of the BAM files for read pairs mapping to a different chromosome location identified a region, located at approximately chr12: 33,710,200 (canFam3) that segregated with the two cases and two controls (SI Appendix, Fig. S2). At this location, read mates mapped to chr18: 48.4 Mb (canFam3) and chr7: 68.3 Mb (canFam3) in the NSDTR and dachshund cases, but none of the controls. The reads that mapped to CFA18 aligned to parental FGF4, which was highly suggestive of a FGF4 retrogene insertion at this location. The reads that mapped to CFA7 were investigated by PCR and appear to mark a genome assembly error or a mutation within the dog used for the genome assembly (canFam3).

To investigate the potential FGF4 insert on CFA12, the region was PCR amplified using primers flanking the insertion site from genomic DNA of an IVDD-affected beagle. Wild-type dogs without the insert had a single 615-bp band, while dogs homozygous for the CFA12 FGF4 insertion had an ~4-kb product. Sanger sequencing showed the insertion on CFA12 is 3,209 bp long (GenBank accession no. MF040221) and includes parental FGF4 cDNA (i.e., FGF4 exons spliced without introns), as shown in the insert schematic comparing parental FGF4 to the CFA12 insert (Fig. 3). The insert also contains a majority of the predicted 5′-untranslated region (UTR), which includes the transcription start site (TSS) as only PCR primers FGF4 TSSF1 and FGF4 R1 yielded a product in RT-PCR using cDNA from neonatal beagle IVD (SI Appendix, Table S8). The insertion location is intergenic between the 3′-UTR of OGFRL1 ~9.5 kb on the proximal side and ~350 kb to the RIMS1 gene on the distal side. It is also intronic to an unvalidated antisense transcript (NASEQ_As_0027291).

To compare the CFA12 FGF4 retrogene to the previously identified CFA18 FGF4 retrogene, it was necessary to obtain the full-length sequence of the CFA18 insertion (25). The cloned product was sequenced using the flanking and common internal primers (SI Appendix, Table S8), yielding a 2,665 bp insert (GenBank accession no. MF040222). While it contained the same length 5′-UTR and FGF4 cDNA as that seen in the CFA12 FGF4 insert, the 3′-UTR was shortened in comparison. The 3′-UTR of the CFA18 FGF4 insert was followed by a sequence containing 30 adenine and one guanine residues and a different target site duplication (TSD) sequence (AAG TCA GAC AGA G). The 5′-UTR shared between the two canine FGF4 retrogenes was compared with the human FGF4 gene and was 295 bp long and 91% identical at the nucleotide level. Within the highly conserved 5′-UTR that was transposed, there are 79 conserved transcription factor binding sites between dog and human (https://ecrbrowser.dcode.org). There are also multiple DNase I hypersensitivity sites as well as H3K4me3 and H3K9ac marks (www.genome.ucsc.edu) within the human sequence.

**Association of FGF4 Retrogene with SD, Height, and IVDD.** To assay the insertions in additional dogs, insertion- and allele-specific PCR-based genotyping assays were developed for both the CFA12 FGF4 insertion and the previously identified CFA18 FGF4 insertion (Fig. 4). Twelve SD NSDTR cases from the GWAS were genotyped and were homozygous for the CFA12 FGF4 insertion, while all controls were heterozygous or wild type. Additionally, IVDD cases (n = 7) from the NSDTR breed were collected and were either homozygous mutant or heterozygous for the CFA12 FGF4 insertion (SI Appendix, Table S3). All NSDTRs tested for the CFA18 FGF4 insertion (n = 31) were wild type, including SD and IVDD cases. NSDTRs with known height (n = 20 males) at the withers were genotyped for the CFA12 FGF4 insertion to investigate the association of height with genotype status. Height and genotype were significantly associated in a dose-dependent manner (all P < 0.04) when comparing wild-type, heterozygous, and homozygous dogs (Fig. 4B).

To assess the significance of association of the CFA12 FGF4 insertion with IVDD across breeds, dogs used in the IVDD GWAS were genotyped for both insertions. All dogs’ genotypes were concordant with phenotype except for one case, a rottweiler (SI Appendix, Table S2). When associated with IVDD, the CFA12 FGF4 association was more highly associated than both the most highly associated SNPs from the GWAS, as well as the CFA18 FGF4 insertion (Fig. 4C). To further investigate the association of the CFA12 4.5 insertion with IVDD, 33 additional cases were genotyped for
Table S4. Comparison of breeds with high (11479), medium (20.2), and low (18.94) height at the withers (in).

<table>
<thead>
<tr>
<th>Breed</th>
<th>Frequency</th>
<th>Height at the withers (in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Highland white terrier</td>
<td>11</td>
<td>20.2</td>
</tr>
<tr>
<td>Coton de Tulear</td>
<td>9</td>
<td>19.5</td>
</tr>
<tr>
<td>Beagle</td>
<td>4</td>
<td>19.2</td>
</tr>
<tr>
<td>Boston terrier</td>
<td>10</td>
<td>19.1</td>
</tr>
<tr>
<td>Afghan hound</td>
<td>8</td>
<td>18.9</td>
</tr>
</tbody>
</table>

Figure 4. The association between CFA12 and IVDD in dogs. (A) Association of CFA12 with IVDD in dogs. (B) Association of CFA18 with IVDD in dogs. (C) Association of CFA18 and CFA12 with IVDD in dogs. The association was determined by quantitative RT-PCR (si-qPCR) performed on samples from dogs with and without IVDD. The association was also confirmed by genotyping and sequencing of the CFA12 FGF4 retrogene insertion. The results are shown in Table S4.
for appropriate embryonic axial growth and segmentation, and FGF4/FGF8 murine hypomorphs are characterized by altered vertebral morphology and smaller limb buds (39, 40). Additionally, FGF8 hypomorphs are observed to have either hypoplastic or nonexistent external ear structures (41). In mice, creation of a gain-of-function FGF4 copy to replace an inactive FGF8 gene was able to rescue limb development; however, it also caused abnormal tissue deposition and postaxial polydactyly, highlighting that levels of FGF throughout embryonic development must be properly controlled for normal limb formation (31). While the specific embryonic expression pattern of FGF4 in dogs with four to six copies of the gene is unknown, we hypothesize that the insertion site milieu on CFA12 versus CFA18 is contributing to differences in expression between the retrogenes, leading to the differences in phenotype.

A survey of retrogenes in the canine reference genome reported ~70 functional retrogenes in the dog; however, only the previous CFA18 FGF4 retrogene insertion has been reported to be associated with a disease-causing phenotype (25, 42). Similarly in humans, the formation of processed pseudogenes in general, as well as those that retain their intended function and cause disease, is rare (43–46).

Both copies of the canine FGF4 retrogenes have signatures of having arisen from RNA retrotransposed by LINE-1 integrate and reverse transcriptase including flanking TSUs and polyA tracts (class 1 templated sequence insertion polymorphism) (47). The insertion sites of the two FGF4 retrogenes are very different. The CFA18 site is within a LINE element and the CFA12 insertion site is intergenic between OGFRL1 and RIMS. The CFA18 FGF4 retrogene insertion was predicted to be expressed due to insertion near sequences with promoter properties (25). While the CFA12 FGF4 insertion is placed near a potential TATA box and RNA Pol II promoter, it is more likely that the CpG island included in the retrogene is driving expression (48–50). This hypothesis is supported by the finding that a majority of retrogene expression is actually due to genomic context and contribution of CpG islands, not through the use of nearby promoters (51). Human FGF4 shares the large CpG island observed in dogs and other species. Within the highly conserved 5′-UTR that was transposed, there are many conserved transcription factor binding sites between dog and human as well as multiple methylation marks further supporting that both CFA12 and CFA18 FGF4 retrogenes contain the necessary components for transcription. To our knowledge, this is a unique documentation of a second retrogene insertion of the same parental gene resulting in a disease phenotype in a mammalian species. Due to the lack of resources available to identify these types of mutations, it is likely that there are other phenotype inducing retrocopies present in the canine genome that have yet to be discovered.

Chondrodystrophy-associated mutation events occurred a very long time ago, as there are descriptions of short-legged dogs dating back over 4,000 y (52). In addition, both mutations occur concurrently in very unrelated dog breeds from diverse breed groupings and geographical locations. The fact that FGF4 has been retrotransposed twice in dogs in the last 3–4,000 y makes it likely that this has happened at other times. The large CpG island in the 5′-end of the parental FGF4 gene may enable phenotypic consequences more readily than for other retrogenes. Once the FGF4 retrogene appeared and produced an obvious phenotype, strong selection was likely applied to retain it, aided by the semidominant nature of the mutation.

The NSDTR is the smallest of the retriever dog breeds, and based on the association of the CFA12 FGF4 insertion with height, we hypothesize that the heterozygous phenotype is aesthetically desirable and that selection is maintaining the insertion at a relatively high allele frequency. Investigation of the CFA12 FGF4 insertion in additional breeds also showed high allele frequency in multiple small- and medium-sized dog breeds. In breeds also containing the CFA18 FGF4 insertion, there is an even more dramatic decrease in height (e.g., bassett hound, Cardigan Welsh corgi, dachshund, etc.), supporting that both FGF4 retrogenes affect long-bone length.

In addition to segregating with height, the CFA12 FGF4 insertion also segregates with Hansen’s type I IVDD susceptibility. Of the IVDD cases genotyped for the CFA12 FGF4 insertion, all were homozygous mutant or heterozygous, except for one, suggesting that one additional copy of FGF4 on CFA12 is sufficient to cause type I IVDD. The single discordant case was a rotweiller, a breed that does not fit the chondrodystrophic phenotype. It is possible that there is another cause of IVDD in non-chondrodystrophic dog breeds occurring without endochondral ossification defects (10). IVDD-affected NSDTRs were also all either homozygous or heterozygous for the CFA12 FGF4 insertion. This supports the idea that while the CFA12 FGF4 insertion is semidominant with respect to height, it is dominant for altered IVDs. Given that the CFA18 FGF4 insertion is not found in the NSDTR and was inconsistently present in the IVDD cases tested, this further supports the idea that the identified insertion on CFA12 is causing both short limbs and Hansen’s type I IVDD in both the NSDTR and across dog breeds.

The breeds with a higher frequency of the CFA12 FGF4 insertion are the same breeds identified in the last 50 y as being predisposed to IVDD. Presence of the CFA18 FGF4 insertion is common in many breeds with IVDD, and it is possible that it may contribute to the disease; however, previous mapping within dachshunds, which are reported “fixed” for the CFA18 FGF4 insertion, show segregation of the associated haplotype on chromosome 12 with IVDD, supporting the idea that the CFA12 FGF4 insertion is the critical factor determining disease status (25, 34). Of particular interest is the lack of reports of IVDD cases in breeds such as the cairn terrier and West Highland white terrier, both of which have the CFA18 FGF4 insertion, but not the CFA12 FGF4 insertion. Similarly, the high incidence of IVDD in breeds such as the American cocker spaniel, beagle, and French bulldog that do not have the CFA18 FGF4 insertion but a high frequency of the CFA12 FGF4 insertion supports the idea that FGF4 specifically from CFA12 is contributing to the IVDD phenotype.

The segregation of the CFA12 FGF4 insertion within dog breeds presents an opportunity for improvement of animal health, as identification of the causative mutation(s) for IVDD in these breeds may aid in the development of genetic testing for the elimination of type I IVDD. Based on the ever-growing popularity of some breeds, the number of animals with this intervertebral disc disease mutation across the globe is in the millions. Myelopathy secondary to IVDD herniation is the most commonly presenting neurological disorder of the spinal cord in dogs (53). The overall health and financial consequences across the spectrum of presentations in companion dogs is immense. Prevention of disease through breeding and eradication has the potential for far-reaching benefits beyond those achievable through advances in surgical or medical therapy. Additionally, the dog may serve as a valuable human–animal model for IVDD. Administration of a tyrosine kinase inhibitor in a mouse model with a gain-of-function mutation in FGFFR3 has been shown to overcome growth defects associated with altered FGF signaling (32). Based on the phenotype and molecular etiology of chondrodystrophy and IVDD in dogs, it has the potential to serve as a bridge between mouse and human studies evaluating the efficacy of targeted pharmacological treatment of FGF-based genetic disorders.

Given the high mortality rate of IVDD and the high cost of surgery, identification of this susceptibility locus could provide a valuable tool for owners, breeders, and veterinarians for mitigating risk of intervertebral disc herniation and resulting myelopathy (9). This could be especially useful in breeds that have both the CFA12 and CFA18 FGF4 retrogene, as they could breed away from the CFA12 FGF4 retrogene, while still maintaining the aesthetically desirable shortness in stature contributed
by the CFA18 FGFR4 retrogene. In breeds with only the CFA12 FGFR4 retrogene, breeders will ultimately decide if prevention of Hansen's type I IVDD outweighs any potential loss of shortness (or gain in height).

**Methods**

Full detailed methods are available in **SI Appendix, SI Methods**. Collection of canine samples was approved by the University of California, Davis, Animal Care and Use Committee (protocol 18561). Genome-wide SNP genotyping was performed using the Illumina Canine HD 174,000 SNP array (Illumina). Sequencing was performed on the Illumina HiSeq. 2500 with 150-bp paired-end reads. Both CFA12 and CFA18 retrogene insertions were cloned and sequenced to determine full-length insertions. The FGF4 insertions on CFA12 and 18 were assayed using a PCR-based genotyping test. qPCR and semi-qPCR were used to assay expression of genes within the interval and FGF4.

**ACKNOWLEDGMENTS.** Funding for this work was provided by the Maxine Adler Endowed Chair Funds, Center for Companion Animal Health, School of Veterinary Medicine, University of California, Davis (UC Davis); NIH, National Institute of Dental and Craniofacial Research R01DE22532; UC Davis Signature Research in Genomics: ACTG NIH, T35 OD010956 and NIH T32 OD010931.


