DOI: 10.1534/genetics.108.087114

Canine Polydactyl Mutations With Heterogeneous Origin in the Conserved Intronic Sequence of *LMBR1*

Kiyun Park,* Joohyun Kang,* Krishna Pd. Subedi,* Ji-Hong Ha† and Chankyu Park*,[1]

**Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejon 305-701, Republic of Korea and †Department of Genetic Engineering, College of Natural Sciences, Kyungpook National University, Daegu 702-701, Republic of Korea*

Manuscript received January 15, 2008
Accepted for publication June 9, 2008

ABSTRACT

Canine preaxial polydactyly (PPD) in the hind limb is a developmental trait that restores the first digit lost during canine evolution. Using a linkage analysis, we previously demonstrated that the affected gene in a Korean breed is located on canine chromosome 16. The candidate locus was further limited to a linkage disequilibrium (LD) block of <213 kb composing the single gene, *LMBR1*, by LD mapping with single nucleotide polymorphisms (SNPs) for affected individuals from both Korean and Western breeds. The ZPA regulatory sequence (ZRS) in intron 5 of *LMBR1* was implicated in mammalian polydactyly. An analysis of the LD haplotypes around the ZRS for various dog breeds revealed that only a subset is assigned to Western breeds. Furthermore, two distinct affected haplotypes for Asian and Western breeds were found, each containing different single-base changes in the upstream sequence (pZRS) of the ZRS. Unlike the previously characterized cases of PPD identified in the mouse and human ZRS regions, the canine mutations in pZRS lacked the ectopic expression of sonic hedgehog in the anterior limb bud, distinguishing its role in limb development from that of the ZRS.

CONGENITAL malformations of the vertebrate limb are often observed as abnormal numbers of digits. One class of defect is preaxial polydactyly (PPD), which is caused by alterations in the anteroposterior patterning of limb development (Hill *et al.* 2003; Horikoshi *et al.* 2003). Dogs in general have four digits in the hind limb, presumably due to an evolutionary adaptation (Galis *et al.* 2001). However, in a few breeds, including Great Pyrenees, five or six digits have been intentionally maintained through breeding. In addition, most breeds often display this extra digit as a genetic variation, although the number of digits (five) in the forelimb is essentially unchanged. This type of dominant genetic alteration has been commonly called "dewclaw," but is more accurately described as "hind-limb-specific PPD." Rarely, polydactyl mutations involving both the fore and hind limbs occur, as reported in breeds like the Norwegian Lundehund (Fogle 2000), although their genetic bases are poorly understood. Except for its exclusive appearance in the hind limb, hind-limb-specific canine PPD resembles human PPD type I, which increases one or more biphalangeal thumbs (Temtamy and McKusick 1978; Zguricas *et al.* 1999). Mouse mutations such as *Hemimelic extra toes* (*Hx*) and M100081 share a similarity with human PPD, although the *Sasquatch* (*Ssq*) mutation displays a unique phenotype of semidominant inheritance with no association with other anomalies, so that heterozygotes contain an extra triphalangeal thumb only in the hind limb, whereas homozygotes have defects in both the fore and hind limbs.

Previous genomewide linkage studies have revealed that human and mouse PPDs are associated with the *LMBR1* gene on HSA7q36 and MMU5 (Hing *et al.* 1995; Clark *et al.* 2000). However, no mutations were identified within the coding region of *LMBR1* (Heus *et al.* 1999). The *LMBR1* gene with 17 exons encodes a putative transmembrane protein homologous to a lipocalin receptor (Wojnar *et al.* 2003). The expression of *LMBR1* was not localized in developing mouse limbs (Hill *et al.* 2003), and its overexpression in a developing chick limb did not interfere with normal skeletal development (Maas and Fallon 2004). The targeted deletion of the first exon in a mouse resulted in a loss of distal limb structure (Clark *et al.* 2001), while a deletion in exon 4 in a human caused a distal limb truncation called acheiropodia (Ianakiev *et al.* 2001). These data suggest that the *LMBR1* coding region is not directly involved in the limb patterning.

Sonic hedgehog (*Shh*) is expressed in the mesenchyme along the posterior margin of the limb in a region called the zone of polarizing activity (ZPA). The spatiotemporal expression of *Shh* is driven by a regulatory element, called the ZPA regulatory sequence (ZRS), contained in the intron 5 of the *LMBR1* gene, which is located ~1 Mb from *Shh* (Lettice *et al.* 2003). The proximity of ZRS to *Shh* is preserved in mammals,

[1]*Corresponding author*: Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejon 305-701, Republic of Korea. E-mail: ckpark@kaist.ac.kr

including dogs (located on CFA16, http://genome.ucsc.edu). The location and role of the ZRS sequence was defined by the insertion mutant *Ssq* in mice, near to which the human PPD, as well as the mouse *Hx* and M100081 mutations, revealed substitutions (Lettice *et al.* 2003). The ZRS mutant displays an ectopic expression of *Shh* in the anterior limb bud, while the wild type expresses *Shh* only in the posterior ZPA region (Sharpe *et al.* 1999; Blanc *et al.* 2002). This has led to the hypothesis that ZRS serves as a long-range regulator of *Shh*. A knockout mouse lacking the ZRS region failed to express *Shh* in a developing limb bud, supporting the hypothesis (Sagai *et al.* 2005). This phenotype is similar to that of the *Shh*-knockout mouse. Furthermore, the single-base change in ZRS causing PPD resulted in an ectopic expression of the reporter in the anterior region of a limb bud (Maas and Fallon 2005). The data imply that the anterior expression of *Shh* due to the altered ZRS function is an underlying mechanism of PPD occurring in both fore and hind limbs. The *LMBR1* gene and its surrounding region are rich in multispecies conserved sequence (MCS) elements, including ZRS, which were functionally tested for their tissue-specific expressions using zebrafish (Goode *et al.* 2005). Most of these MCS regions from the Fugu genome, including three in the *LMBR1* gene, exhibited diverse activities during development (Goode *et al.* 2005). So far, there has been no evidence that human and mouse PPDs are caused by mutations in the two other MCSs in *LMBR1* (Sagai *et al.* 2005).

In a previous study, canine PPD in a hind limb was mapped to CFA16 using a pedigree analysis of the Korean breed, Sapsaree (Park *et al.* 2004). Here, we report on the positional cloning and characterization of the affected gene. On the basis of human synteny, the affected region defined previously was further reduced to <213 kb by linkage disequilibrium (LD) mapping. Simultaneously, we identified two affected LD haplotypes, one for Korean breeds and the other for Western breeds. Sequencing the coding and intronic regions of a single candidate gene, *LMBR1*, revealed base substitutions in the intronic MCS region located upstream of ZRS, termed preZRS, or pZRS, which exhibited a limb-specific enhancer activity and altered the expression pattern in the limb bud in a transgenic mouse when mutated. The mutational effect of pZRS was distinct from that of a ZRS reported previously (Maas and Fallon 2004).

MATERIALS AND METHODS

Cosmid and BAC screening: The cosmid library obtained from Beagle is from P. Henthorn at the University of Pennsylvania. Canine BAC clones were obtained by hybridizing with a high-density grid containing an 8.1-fold canine BAC library (Li *et al.* 1999). We used the sequence around the REN85M08 marker as a probe for screening the genomic library. Details of the hybridization conditions and sequencing strategy are described in the supplemental materials and methods.

SNP finding: Genes were selected for SNP discovery on the basis of human synteny. Two genes, *KCNH2* and *ADRB3*, were located on CFA16, while *En2*, *Shh*, *Hlxb9*, *Vipr2*, *Fgfr1*, *Indo*, *Vdac3*, *Unc5D*, *Wrn*, and *NM024025* were selected from HSA7q and HSA 8p. Intronic or exonic sequences 500–1500 bp in size were amplified using primers designed by comparing human sequences with those of a mouse or rat, depending on availability. To assess the primers, we performed PCR with dog cDNA and confirmed the products by sequencing. Sequence data from genomic DNA were analyzed using the PolyPhred option of the PhredPhrap program. To determine linkage to the canine PPD locus, SNPs were genotyped for the Sapsaree pedigree constructed previously (Park *et al.* 2004). We calculated pairwise LOD scores using the MLINK option of the linkage program (Lathrop and Lalouel 1984). Two-point linkage and haplotype analyses were additionally performed using microsatellite markers previously mapped on CFA16.

Detection of mutations: For an analysis of mutations in the canine ZRS, we used primers designed by comparing human and mouse sequences. Sequences from exon 5 to ZRS of *LMBR1* were obtained by small-scale shotgun sequencing and compared among affected and unaffected individuals to identify mutations. Direct sequencing of both strands was performed using Big Dye Terminator version 2.0 (Applied Biosystems). We performed SNP typing by allele-specific oligonucleotide (ASO) hybridization and allele-specific PCR and by using a SnapShot Multiplex kit (Applied Biosystems). For ASO hybridization, amplified products were blotted onto Hybond-N+ membranes (Amersham Biosciences), which were hybridized for 1 hr at 37° in 5× SSC, 1% SDS, and 50 mg/ml of single-stranded salmon sperm DNA with either normal or mutated oligonucleotides labeled with [γ-^{32}P]ATP. Filters were washed at 37° through a final stringency of 0.3× SSC with 1% SDS.

Association analysis: Marker-disease association was tested using the χ^2 method with a STATA program (http://csg.sph.umich.edu). The *P*-value was obtained using Fisher's exact probability test. The *P*-values of the microsatellites were corrected by multiplying the number of microsatellite alleles observed in each locus (P_c). $P_c < 0.01$ was accepted as statistically significant. The odds ratio for canine PPD was calculated from a 2×2 contingency table. Allele frequency was estimated using the GenePop program (http://wbiomed.curtin.edu.au/genepop/).

Luciferase assay: The 632-bp DNA fragments containing either an A or a G allele of the DC-1 site were amplified from the DNA of a DC-1 homozygous individual. The 703-bp fragments containing the published human mutations were obtained by PCR-based mutagenesis. Fragments were cloned into the *Kpn*I/*Sac*I site of the pGL2-promoter vector (Promega). P19 embryonal carcinoma cells were cultured in media (MEMA, GIBCO, Life Technologies) supplemented with 10% fetal calf serum, 2 mm L-glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin. Transfections were performed using a Lipofectamine reagent. Cells were transfected with 1 μg of the construct and 0.5 μg of the pSV-β-galactosidase vector (Promega) as an internal control for measuring transfection efficiency. After 48 hr, cells were collected and measured for luciferase activity with the luciferase assay reagent (Promega) using a luminometer (EG&G Berthold, Microlumat LB96P). The β-galactosidase activities in the lysis buffer were measured with a spectrophotometer using the appropriate enzyme assay system (Promega).

Transgenic analysis: Transgenic constructs were generated using the β-globin minimal promoter and the bacterial *lacZ* reporter gene (kindly presented by Robb Krumlauf). Constructs

containing pZRS and ZRS were made by cloning the 1.9-kb DNA, generated by PCR from normal and affected dogs, containing 5′-*Kpn*I and *Xho*I restriction sites. The PCR products were sequenced to ensure the absence of errors. A vector fragment was removed by *Kpn*I/*Sac*II digestion. Transgenic mice were generated by pronuclear injection using standard protocols. Embryos for staining were obtained at E10.5–15.5 (assuming the day of plug check as E0.5). Transgenic males were subsequently used as studs for FVB background females. All harvested embryos had their yolk sacs retained to allow for PCR genotyping. Then β-gal staining was performed using standard procedures (Maas and Fallon 2005).

Whole-mount *in situ* hybridization: The experiment was carried out as described previously (Hecksher-Sorensen *et al.* 1988; Hammond *et al.* 1998). RNA probes for *in situ* hybridization of a canine *Shh* sequence were obtained by PCR using primers (GenBank accession no. XM_845357, forward/5′-GGACCCGGCAGGGGATTTGG-3′ and reverse/5′-GGCCGC AGAGTCGTTGAGCG-3′) and labeled using a DIG RNA labeling kit (Roche). The solution for hybridization contained 50% formamide, 5× SSC, 2% blocking powder (Boehringer), 0.15% Triton X-100, 0.5% CHAPS, 5 mm EDTA, 100 μg/ml heparin, and 50 μg/ml of yeast RNA. The chemicals used for hybridization were from Sigma Chemical, unless otherwise stated. Prehybridization was carried out for 5 hr at 65° in a hybridization solution. The DIG-labeled RNA probe (1 μg/ml) was denatured for 3 min at 80° in a 50-μl hybridization mix, which was then used to hybridize with embryos overnight at 65°. Washing was carried out as described (Hammond *et al.* 1998), followed by staining in AP1 (100 mm NaCl, 50 mm $MgCl_2$, 100 mm Tris, pH 9.5) with a 20 μl/ml NBT/BCIP stock solution (Roche, 18.75 mg/ml NBT and 9.4 mg/ml BCIP in 67% dimethylsulfoxide). For developing, the embryos were rinsed in AP1 and mixed with PBST to stop the reaction and then washed several times in PBS with 1% Triton X-100 to adjust the background color, followed by fixing overnight at 4° in 4% paraformaldehyde. Details are described in the supplemental materials and methods.

RESULTS

SNP markers in the affected region: Initially, we screened for canine SNPs located in the region around *LMBR1* on CFA16, which share syntenies to both HSA7q and HSA 8p, since the affected locus lies within the syntenic junction. The linkage of the markers to CFA16 was determined by two-point and haplotype analyses. The deduced syntenic order based on the SNP markers was as follows: REN176D05–8T7R–REN85M08–14SP6R–ZRSi5–*HLXB9*–CXX876–*FGFR1*–*ADRB3*, with linkages to CFA16, as shown in Figure 1 (supplemental Table S1). Since the previous study (Park *et al.* 2004) excluded the CXX876 microsatellite marker located distally from the recombination-free affected region, we suspect that the candidate region is syntenic to HSA7q. Thus, *LMBR1* was highlighted as a likely candidate gene involved in developmental defects of the preaxial limb (Heus *et al.* 1999; Clark *et al.* 2000). On the basis of an earlier study reporting that REN85M08 is the most closely linked marker to the canine PPD locus (Park *et al.* 2004), we attempted to screen additional SNPs from canine BAC clones obtained with the sequence probes flanking the REN85M08 repeat polymorphism, which were further expanded as contigs. The newly identified SNPs were ordered on the basis of the map of BAC contigs (supplemental Figure S1) as follows: 8T7R–*RNF*32–REN85M08–14SP6R–*LMBR*1/SNPs–*NOM*1–*HLXB*9/SNPs (Figure 1A).

LD mapping with microsatellites and SNPs: Using several SNPs and microsatellites covering a few megabases of a previously suggested region of the canine PPD locus, we performed LD mapping. Since the linkage analysis was conducted with selected pedigrees from a much larger population (>1000 dogs), we selected 77 other Sapsarees with PPD for high-resolution LD mapping. Genotyping with these markers along with a haplotype analysis led to the exclusion of REN176D05-8T7R and *NOM1* markers located at the 3′ and 5′ sides of *LMBR1*, respectively, from the LD region, including the canine PPD locus (Figure 1A). The REN85M08 marker was initially positioned between 8T7R and 14SP6R SNPs on the basis of PCR analyses with BAC contigs (supplemental Figure S1) and accurately confirmed using a recent dog genome map (http://genome.ucsc.edu). Thus, we limited the locus of the canine PPD to <1 Mb (Figure 1A) within which a few genes, including *LMBR1*, are located.

Association analysis of the *LMBR1* region in other breeds with PPDs: To further map the canine PPD locus with better resolution, we collected 47 normal and PPD dogs from other breeds (supplemental Table S3). An association analysis was performed for REN85M08 and ShotR2 markers closely linked to the *LMBR1* gene, using a distantly located REN124F19 marker as a control (Table 1). Using χ^2 analysis, a significant *P*-value was obtained for REN85M08 and ShotR2 markers. We additionally tested other SNPs (supplemental Table S2) identified in the introns of *LMBR1*. Our data further support an association with the *LMBR1* region, *e.g.*, $\chi^2 = 12.35$, $P = 4.41 \times 10^{-4}$ for ZRSi5. The most significantly associated allele of REN85M08 was absent in normal individuals, which may seem unique, perhaps due to its recent origin and tight linkage to a rare microsatellite allele. For minimizing allelic bias, we attempted a balanced inclusion of each breed for case and control samples taken from nine different breeds (supplemental Table S3).

Assessing LD haplotypes around *LMBR1* and their phenotypic association: Since the PPD loci in other breeds are associated with the *LMBR1* region, we extended the LD mapping to their genomes (supplemental Table S3). Genotyping with most SNP and microsatellite markers of the *LMBR1* region revealed recombination breaks between REN85M08 and 14SP6R SNP in three affected individuals of the Shih Tzu, Yorkshire terrier, and standard poodle breeds and between Int3 and *NOM1* SNPs in the Shih Tzu (supplemental Figure S1). Thus, the boundary of the canine PPD locus was marked by the REN85M08 and *NOM*1 markers, suggesting a candidate region of <213 kb (Figure 1A). LD mapping and association analysis confirm that the 140-kb region of *LMBR1* constitutes the core of the affected locus. Accordingly, we

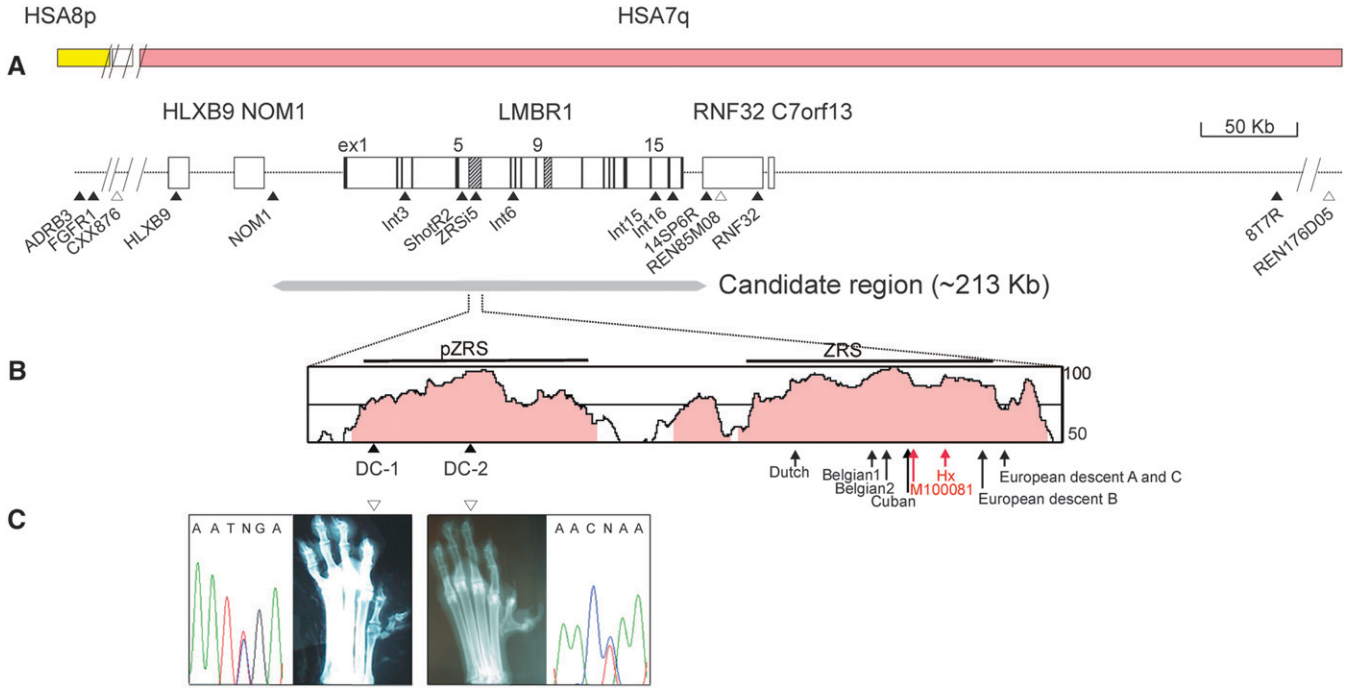

FIGURE 1.—Genomic analysis of the PPD locus on CFA16. (A) The genomic structure around the MCS regions of the *LMBR1* gene. The SNPs (solid triangles) and microsatellites (open triangles) are indicated below. The black bar and shaded regions within the *LMBR1* gene represent exon and MCS regions, respectively. The yellow and pink bars represent human syntenic regions on HSA8p and HSA7q. The deduced candidate region (213 kb) around the *LMBR1* gene for PPD is designated below. (B) The highly conserved region between human and dog, analyzed using VISTA (http://genome.lbl.gov/vista), is presented. The region with >50% identity analyzed with windows of 100 bp appears on the plot, while those >75% are shaded in pink. The identified PPD mutations, DC-1 and DC-2, are presented (solid triangles). The red and black arrows indicate the sites of two mouse mutations, Hx-G/A and M100081-A/G, and six human PPD mutations: Dutch-C/G (105 bp from 5′-ZRS), Belgian1-A/T (305), Belgian2-T/C (323), Cuban-G/A (404), family B-C/G (621), and families A and C-A/G (740) from northern European descent (LETTICE *et al.* 2003; GURNETT *et al.* 2007). The phenotypic patterns of preaxial polydactyly are indicated below. (C) A sequence electropherogram of DC-1 (Sapsaree) and DC-2 (Great Pyrenees) mutations read in the reverse direction are shown with hind foot radiographs.

systematically analyzed this region for the presence of an LD block by typing genetically independent individuals for various breeds. For the Sapsaree, deducing LD haplotypes was fairly straightforward due to the availability of pedigree information. In fact, the genetic history of the whole Sapsaree population is derived from less than eight founders (PARK *et al.* 2004). From the Sapsaree and other breeds, 14 SNP markers (supplemental Table S4) of the candidate region were typed, which were initially assessed with the HaploXT/GOLD and SIMWALK2

TABLE 1

Allelic associations between the CFA16 microsatellite and SNP markers

Marker	Distance (Mb)[a]	No. of alleles	Frequency of associated allele[b] (no. of individuals)		Odds (C.I. of 95%)[c]	χ^2 (1)[d]	*P*-value	P_c value[e]
			Normal	PPD				
REN85M08	21.1	7	0/48 (24)	17/42 (21)	—	23.95	9.87×10^{-7}	6.91×10^{-6}
ShotR2	21.2	2	21/48 (24)	41/44 (22)	10.63 (2.51 45.06)	16.24	5.58×10^{-5}	
REN124F19	49.6	6	13/26 (23)	17/42 (19)	2.05 (0.81 5.18)	2.46	0.12	0.72

[a] Distances were based on the canine genome database (May 2005 for Boxer, http://genome.ucsc.edu).
[b] Frequencies of highly associated alleles are presented as the number of a particular allele among the number of all alleles. In microsatellite analysis, *P*-values were obtained for all combinations, each with a particular allele versus the other alleles combined. Only the *P*-value for the most highly associated allele is shown. Individuals were selected from nine breeds (details in supplemental Table 3S).
[c] C.I. denotes the confidence interval. The odds ratio for REN85M08 could not be estimated due to the lack of PPD-associated alleles in normal individuals.
[d] χ^2 values with degrees of freedom in parentheses were obtained using Pearson's method.
[e] *P*-values were corrected by multiplying the number of microsatellite alleles observed for each locus.

TABLE 2

Haplotype-based association analysis in Western breeds

Haplotype[a]	Haplotype frequency[b]		Odds ratio (C.I. of 95%)[c]	χ^2 (1)	*P*-value[d]	P_c value[e]
	Normal	PPD				
1-CCCAGC	0.250	0.727	8.00 (2.76 23.22)	20.95	4.60×10^{-6}	2.30×10^{-5}
2-CCCAAC	0.125	0.000		5.88	0.015	0.075
3-TCCCAC	0.312	0.068	6.21 (1.54 25.11)	8.71	0.0032	0.016
4-TCCCGC	0.250	0.137	2.11 (0.74 6.33)	1.88	0.17	
5-CTTAGT	0.063	0.068	1.09 (0.21 5.80)	0.01	0.91	

[a] The SNPs used to construct haplotypes are Int16, Int15, Int6, ZRSi5, ShotR2, and Int3, whose positions on the canine genome are designated in supplemental Table 6S.

[b] Individuals were selected from nine breeds, including 24 normal (10 homozygous and 14 heterozygous) and 22 PPD dogs (10 homozygous and 12 heterozygous).

[c] C.I. denotes the confidence interval.

[d] The *P*-values for the genotype level association with haplotype 1 are 6.27×10^{-6} ($\chi^2 = 20.40$) and 1.31×10^{-2} ($\chi^2 = 6.15$) for dominant and recessive modes of inheritance, respectively.

[e] The *P*-values were corrected by multiplying the number of haplotypes.

programs. Pairwise disequilibrium measurements based on D' values indicated that the associated region with significant *P*-values (<0.05) lies in an LD block extending ~113 kb, which includes Int3 and Int16 SNP markers (see Figure 1A; supplemental Table S5). Consequently, a more extensive LD analysis was performed with the same six SNPs. The five deduced haplotypes shown in Table 2 were repeatedly obtained with various parameter settings, *e.g.*, with D' score ranges of 0.7–0.9, confirmed by HAPBLOCK and PHASE programs (Stephens *et al.* 2001; Zhang *et al.* 2005). The reliability of haplotype prediction is high due to the frequent presence (0.49) of homozygous individuals. For ambiguous cases of predicted haplotypes, allele-specific PCR was performed.

Intriguingly, the deduced haplotypes of the Sapsaree were fairly diverse, consisting of six different species (Figure 2). Only two (no. 3 and no. 5) overlapped with the haplotypes of Western breeds. According to previous phylogenetic analyses with mitochondrial (Leonard *et al.* 2002; Savolainen *et al.* 2002) and microsatellite (Parker *et al.* 2004) markers, Western breeds share a fairly recent ancestor. A neighbor-joining tree showing a genealogical relationship was drawn (Figure 2) on the basis of previous reports (Parker *et al.* 2004) with an addition of the Sapsaree and other native Asian dogs (Kim *et al.* 2001). When this tree was juxtaposed with haplotype patterns found in each breed, only a subset of 10 haplotypes was identified in the breeds of Western origin. This is consistent with the current theory that Western breed diversity is derived from a small number of founders through a population bottleneck (Sutter and Ostrander 2004). Haplotype 10 is unique in wolves, while haplotypes 5–9 were distributed over dogs of Asian origin, including the Shiba Inu, Akida, Jindo, Sapsaree, and Korean Tosa, which is a hybrid of a Korean mongrel and Japanese Tosa. A variant of haplotype 3 in the Sapsaree is associated with the polydactyl phenotype, while haplotype 1 is associated with the PPD in other breeds. Thus, we inferred two affected haplotypes of different origin, specifically, one for Asian breeds and the other for Western breeds. Tracing the genealogical roots of the two affected haplotypes, 1* and 3*, through the neighbor-joining tree suggests that haplotype 3* found in Asian breeds is likely to be more ancient than haplotype 1* in Western breeds.

Since the haplotypes were based on the affected LD region, we confirmed the results in Table 1 by attempting a haplotype-based association analysis for 47 individuals with normal and canine PPD phenotypes. A highly significant *P*-value was obtained for the association of haplotype 1 to the PPD phenotype (Table 2), thus verifying the PPD association of the *LMBR1* gene in Western breeds. When the same analysis was performed at a genotype level by assuming a dominant inheritance, we obtained an even higher significance ($\chi^2 = 20.40$, $P = 6.27 \times 10^{-6}$) compared to that at the haplotype level ($\chi^2 = 20.95$, $P_c = 2.30 \times 10^{-5}$).

Mutational heterogeneity for canine PPD in the MCS region of *LMBR1* intron 5: A comparison of the complete cDNA sequences of *LMBR1* from wild-type and affected individuals of the Sapsaree breed revealed no differences, with the exception of a silent mutation in exon 6 (data not shown). In mice and humans, no polydactyl mutations were identified in the coding region of *LMBR1* (Heus *et al.* 1999; Clark *et al.* 2000), but they were present in the ZRS sequence (700 bp, starting from 4514 bp of intron 5), which is part of the MCS region (~1133 bp). We sequenced a 10.4-kb region from a Sapsaree, including the ZRS, encompassing the whole 5′ region of intron 5 with 104-bp exon 5 and identified four single nucleotide changes. Three of these mutations, designated as ShotF1-2, ShotR2, and ZRSi5, were variable in affected and unaffected individuals, which is characteristic of SNP (supplemental Table S4). However, the remaining mutant (DC-1), found in the 1-kb region upstream of the 5′-end of the ZRS

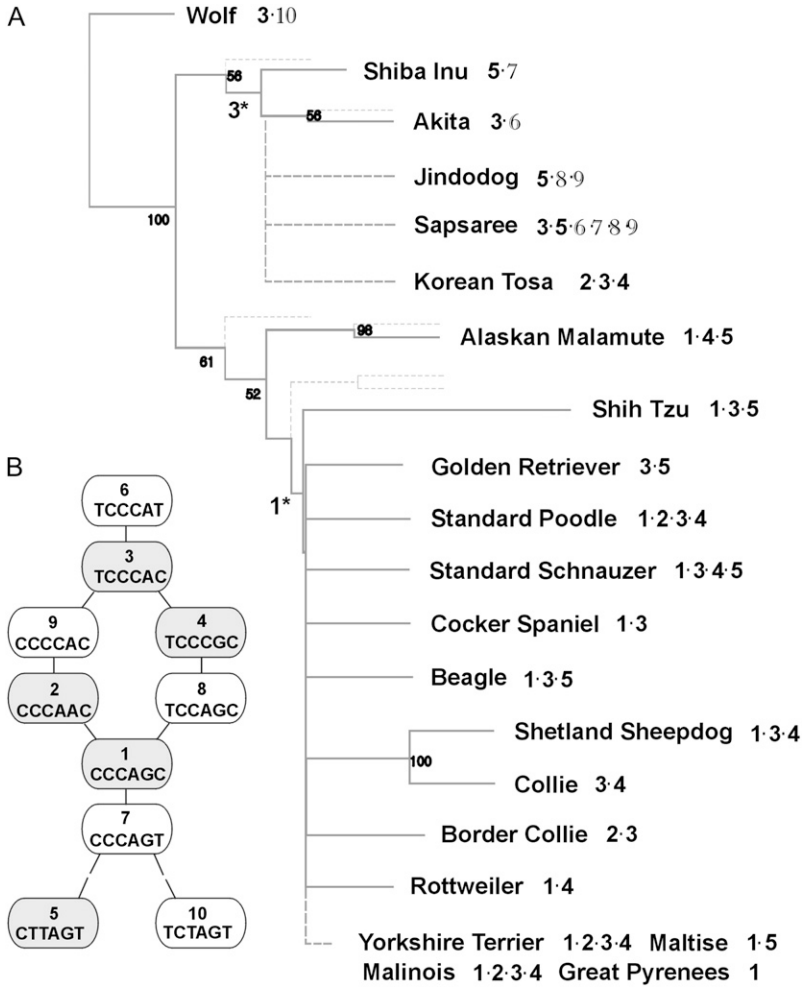

FIGURE 2.—The genealogical relationships between different breeds and their haplotypes. (A) The backbone of the phylogenetic tree (including dotted lines) and bootstrap numbers are adopted from a previous report (PARKER *et al.* 2004). The dashed line includes breeds analyzed previously (KIM *et al.* 2001) or in this study. The numbers with asterisks represent the affected haplotypes from Asian and Western breeds. (B) The SNP markers used to construct the haplotypes are Int16, Int15, Int6, ZRSi5, ShotR2, and Int3. The shaded circles represent haplotypes found in Western breeds, and the open circles (also for numbers in A) signify the unique haplotypes in Asian breeds. The relationships among different haplotypes were analyzed using the NETWORK program (http://www.fluxus-engineering.com) and are shown schematically. The double dashes specify two mismatches.

(Figure 1B), differs in terms of dependence on the affected status. For example, 144 heterozygous and 6 homozygous affected individuals harbored G/A and A/A, respectively, whereas the 65 unaffected individuals displayed homozygous G/G (supplemental Table S4). The same DNA alteration pattern in DC-1 was observed in six Korean Tosas, whereas DC-1 was invariable (G/G) in all individuals (42) of the Western breeds. A search for DNA alterations specific for canine PPD individuals of Western breeds led to the identification of a single nucleotide change, G/A (denoted DC-2), located 246 bp downstream of DC-1 (Figure 1, B and C). The DC-2 mutation was identified in 26 affected individuals of Western breeds with either heterozygosity (G/A in 14) or homozygosity (A/A in 12), whereas the 23 unaffected individuals were homozygous (G/G). On the other hand, there were seven more MCS in the candidate region, three in the intergenic regions between the *NOM1* and *LMBR1* genes, and four in the intronic regions (SHARPE *et al.* 1999; LETTICE *et al.* 2003; ROBERT and LALLEMAND 2006) in *LMBR1* (Figure 1A and supplemental Figure S2). A comparison of the sequences in this region did not disclose any differences between normal and PPD animals (data not shown).

A sequence with mutational changes, designated pZRS (632-bp region located 585 bp upstream of ZRS) and identified on the basis of human–dog similarities, is less well conserved in other species. However, the nucleotide residues corresponding to the DC-2 and DC-1 positions are well conserved in mammalian species (supplemental Figure S3). While none of the previous polydactyl mutations are localized in this region, it is possible that future searches will reveal changes in pZRS.

The MCS region serves as a transcriptional enhancer modulated by polydactyl mutations: We investigated the role of the ZRS in transcriptional regulation by generating a luciferase reporter construct. Interestingly, the ZRS exhibited transcription-enhancing activity in the P19 embryonal carcinoma cell line, which is more than sevenfold relative to that of the vector control (Figure 3A). The magnitude of the enhancer activity was not considerably altered when the insert size was reduced to 131 bp (1787–1917 bp in Figure 3A). However, the core ZRS region displaying almost 100% identity in mammalian sequences exhibited the highest enhancer activity (data not shown). When we introduced DC-1 or DC-2 mutations in the wild-type ZRS region, >50% reduction in enhancer activity was evident (Figure 3A). Similar reductions were observed

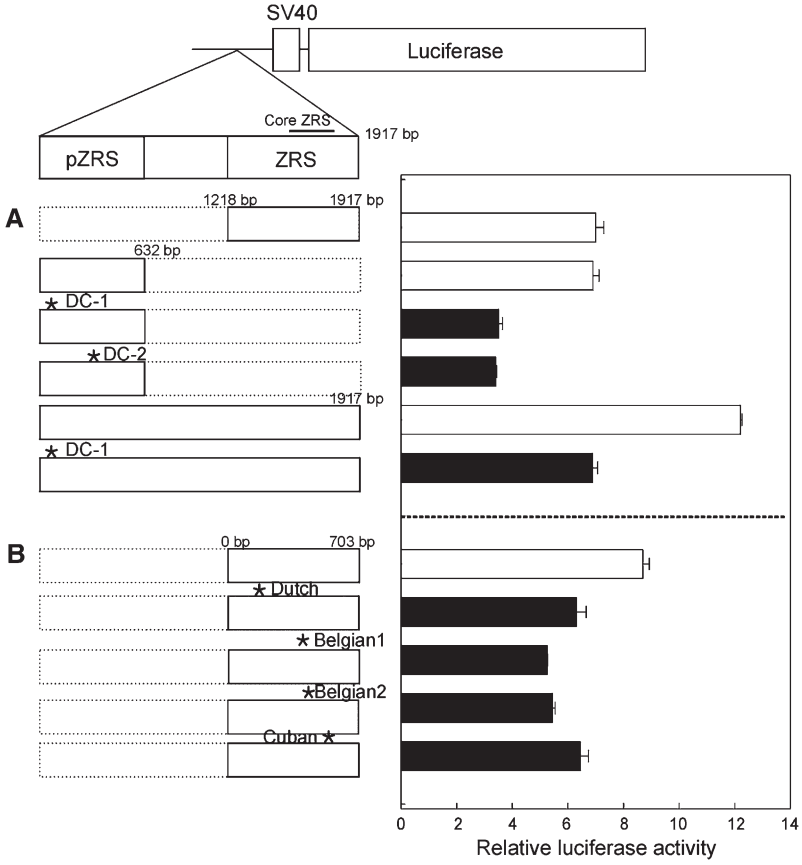

FIGURE 3.—Transcriptional enhancer activities of the ZRS and pZRS regions. (A) The canine ZRS or pZRS sequence with or without the PPD mutation was inserted 269 bp upstream of the luciferase gene under the control of an SV40 promoter. The full-length insert includes a 1917-bp region containing pZRS and ZRS. The asterisks represent the mutational positions. The ZRS region (1218–1917 bp) is a highly conserved region among humans, dogs, and mice in which the core ZRS (1571–1750 bp) is located. The boxes with a solid line (left) represent the sizes of inserts with or without mutations (*). The broken line is used as a reference to designate the size of the full-length insert. After transfection of P19 embryonal carcinoma cells with the appropriate constructs, luciferase activities were determined, standardized with β-galactosidase, and shown as relative values calculated by setting the activity of an empty vector (pGL2, promoter only) as one. The solid bars (right) represent expressions from the constructs with mutations. Data were obtained in triplicate. (B) Transcriptional activity of constructs containing a human ZRS with or without mutations. The constructs with mutations contain nucleotide changes for Dutch-C/G, Belgian1-A/T, Belgian2-T/C, and Cuban-G/A. Data were obtained in triplicate.

in constructs containing human ZRS with four mutations (asterisks in Figure 3B and arrows in Figure 1B), when compared to those of a wild-type ZRS (Figure 3B).

The polydactyly caused by mutated pZRS is not due to an ectopic expression of the anterior *Shh*: To examine whether the canine PPD mutation in pZRS causes an ectopic expression on the anterior limb bud like other ZRS mutations, we designed two transgenic constructs, one containing a 1.9-kb fragment with both pZRS and ZRS regions positioned upstream of the β-globin minimal promoter with *lacZ* and the other with a DC-1 PPD mutation. For the wild-type control, four independent transgenic mice were obtained. At embryonic day 10.5, these lines displayed a β-gal expression within a small patch in both limbs (not shown), which expanded to the posterior region of ZPA (Figure 4, E and F) and spread to the anterior direction of the developing digits (Figure 4H), resulting in a V-shaped staining pattern. From the DC-1 construct, we obtained six independent lines. The stained region was broader in the DC-1 embryos, compared to the wild type at E11.5 (Figure 4G), and extended to the position of the second digit at E13.5 (Figure 4I). In contrast to that of a hind limb, the β-gal staining pattern of the DC-1 forelimb was essentially the same as that of the wild type (Figure 4, J and K). The patterns in the homozygotes were indistinguishable from those in the heterozygotes (data not shown). In addition to the expression in the limb bud, the pZRS-containing constructs created additional β-gal staining in the olfactory bulb region (Figure 4D, asterisk).

Since the previous PPD mutations mapped in the ZRS caused not only the reporter but also an *Shh* expression in the anterior limb bud (MAAS and FALLON 2005), the lack of anterior limb bud expression observed in the canine ZRS with PPD was a bit of a surprise. Thus, we speculate that the canine PPD may be unique in that the mutation phenotypically mimics the normal limb development in which the ZRS regulation, and also *Shh*, is restricted to the ZPA region. To test this possibility, we examined whether the *Shh* expression in the canine embryo with PPD is limited to ZPA by *in situ* analysis. The embryonic period of dogs is defined as 13–34 days after fertilization, in which the limb buds are visible at 24–26 days of embryo development (RUVINSKY 2001). This period corresponds to 10–12 embryonic days for mice, during which the *Shh* expression in the ZPA of a limb bud is apparent (ROBERT and LALLEMAND 2006). We obtained 13 dog embryos of days 23–26 from four Sapsarees in a natural estrus cycle, which were generated after crossing heterozygous Sapsarees with PPD. The embryos were genotyped for DC-1 mutation and a REN85M08 marker to obtain normal or PPD dogs. The *Shh* expressions were observed at day 25 in the ZPA region of the hind limb and were the same in both normal and homo- or heterozygous PPD embryos

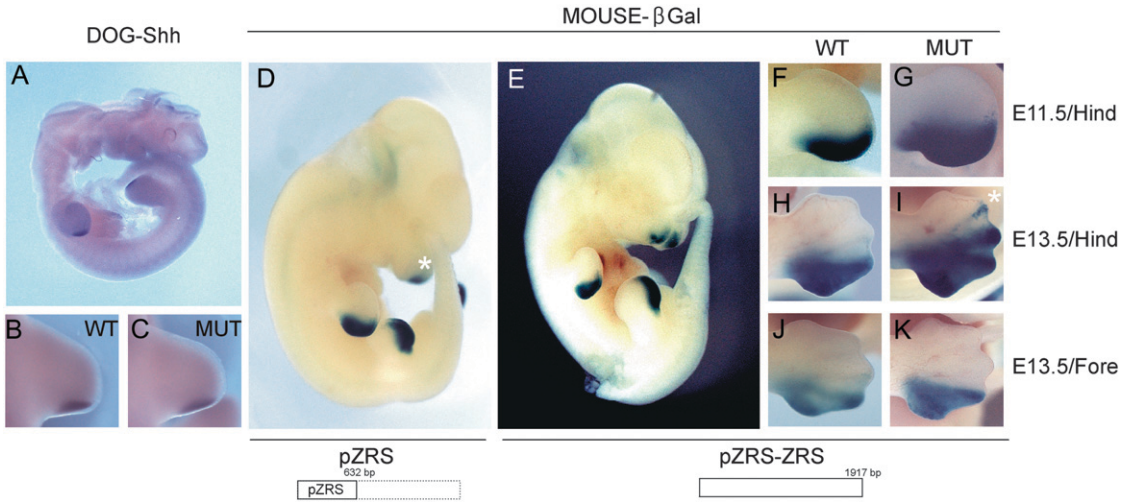

FIGURE 4.—Spatio-temporal patterns of gene expression in a developing embryo. (A, D, and E) Whole mounts, lateral views, and ventral at the left. (B, C, and F–K) Dorsal views and anterior at the top. (A) *Shh* expression in dog embryos at day 25, equivalent to mouse E11.0. (B and C) *Shh* expression in the ZPA region of the hind limb bud in wild-type (B, WT) and PPD (C, MUT) dogs. (D and E) β-Gal staining patterns of transgenic mouse embryos at E11.5 with the wild-type pZRS (D) and the wild-type pZRS plus ZRS (E). The extents of the inserts in the transgenic constructs are shown with boxes (solid lines) at the bottom. The asterisk in D indicates β-gal staining in the olfactory region. (F–I) β-Gal staining in the hind limb buds with wild-type pZRS-ZRS at E11.5 (F) and E13.5 (H) with mutant pZRS-ZRS at E11.5 (G) and E13.5 (I). The asterisk in I indicates the position of a developing second digit. (J and K) β-Gal staining of forelimb buds in wild-type (J) and mutant (K) embryos at E13.5.

(Figure 4, A–C). The expression patterns stayed the same until day 25.5 (data not shown). Unlike the human or mouse mutations in ZRSs, an ectopic expression of *Shh* in the anterior region was not observed in the canine limb bud with PPD. It is likely that the role of pZRS is different in driving an *Shh* expression there.

DISCUSSION

The canine PPD model provides a unique opportunity to study limb development due to its hind-limb-specific and atavistic phenotypes. The polydactylies that are generally reported in dogs or cats show additional digits (one or two) in the forelimb or hind limb, along with other symptoms. But the phenotype of dewclaws is specifically confined only to the hind limb. Most human or mice phenotypes with preaxial digit anomalies are generated as a result of abnormalities from the canonical form of five digits. However, the canine PPD studied here appears to be a reversion from the evolutionary loss of the first digit. As described previously (PARK *et al.* 2004), the restored phenotype is slightly overmanifested, generating one or more preaxial digits. As a matter of fact, the external morphology of hind-limb-specific canine PPD is slightly different from other polydactylies in that only the claw is observed as having an extra digit with fibrous tissue connecting the region between the tarsal and the first metatarsal bones, which either is reduced in size or is incomplete with a lesser deposit of bony material (MILLER 1979; PARK *et al.* 2004). Thus, one might expect that the genetic changes associated with the hind-limb-specific canine PPD differ from those of polydactyly in other species. Intriguingly, mutational changes for the canine PPD were found in the region upstream of ZRS, in which no PPD mutations have previously been mapped.

In a developing limb, *Shh* is required for proper anterior/posterior patterning. The previous animal models for PPD indicated that there were *Shh* expressions in both the posterior and anterior regions of the developing limb (MASUYA *et al.* 1995, 1997; SHARPE *et al.* 1999). The *Shh* misexpression in the anterior limb bud leads directly to the formation of extra digits (RIDDLE *et al.* 1993; LOPEZ-MARTINEZ *et al.* 1995), which is caused by a single base change in the ZRS region (MAAS and FALLON 2005). However, the canine PPD mutation found in pZRS did not create an ectopic expression in the anterior limb bud (Figure 4G). Instead, the enhancing activity was extended toward the anterior side by the PPD mutation, which is clearly visible in the later stage of the embryo (Figure 4I). Although it is difficult to conclude whether such differences correlate with the *Shh* expressions in the dog embryo (Figure 4, B and C) due to a technical limit of *in situ* hybridization, the mutational effect on the enhancer activity found only in the hind limb is clearly suggestive. This may reflect the unique genetic alteration restoring the five-digit pattern of the canine hind limb. On the other hand, the pZRS sequence characterized here itself displayed the enhancer function in the ZPA region of the limb bud, as in the ZRS (Figure 4, D and E), with additional enhancer activity in the olfactory pit. Recent findings (CAPELLINI *et al.* 2006) indicate that the ZRS interacts with Pbx, which is required for proximal/distal patterning, as well as with HOX proteins such as Hoxd10 and Hoxd13.

Similarly, the pZRS element may be associated with other developmental regulators. Further investigation with the pZRS element should clarify the mechanism underlying the unique hind-limb-specific digit patterning in canine PPD.

A haplotype analysis of the LD block around *LMBR1* revealed a fairly extended genomic region of >200 kb with no recombination interruption among different breeds. The considerable LD size is attributed to the fact that the breeds have undergone tight population bottlenecks (Sutter and Ostrander 2004), which is also consistent with the sharing of identical affected haplotypes for canine PPD in different breeds. For the LD block around *LMBR1*, we found a total of five haplotypes in Western breeds, consistent with the previous finding of relatively low haplotype diversity (average of 4.5 haplotypes across 80% of the genome at a locus) among various breeds (Sutter and Ostrander 2004; Sutter *et al.* 2004) with a high level of haplotype sharing. We added five more haplotypes from Asian origin and wolf breeds (Figure 2), supporting the current hypothesis that Asian dogs have a more ancient origin. The representation of such diverse haplotypes in our dog colony appears to be due to an initial founder collection made several decades ago from diverse regions of Korea. The two independent origins of canine PPD mutation may reflect its rather frequent incidence, perhaps due to a developmental constraint.

To elucidate the morphological variations in dog breeds, gene-associated repeat expansions/contractions were cataloged for developmental genes, including *Alx4* (Aristaless-like 4), *Runx2*, and *Hox* (Fondon and Garner 2004). Interestingly, homozygous *Alx4* (Δ51 bp) alleles were identified in four Great Pyrenees with bilateral double PPD, while a single PPD Great Pyrenees harbored homozygous full-length alleles. On the other hand, Great Pyrenees hind-limb polydactyly was not a simple Mendelian trait (Fondon and Garner 2004), which was suggested to be consistent with the genetic background dependence of the *Alx4*-null phenotype in mice exhibiting variable PPD phenotypes (Qu *et al.* 1998). Examination of *Alx4* contraction in Great Pyrenees revealed three homozygous ($Alx4^{\Delta 51}$) and three heterozygous genotypes among six bilateral double PPDs. In addition, *Alx4* contraction was not observed in the PPD individuals of two Shetland sheepdogs, two Shih Tzus, three standard poodles, one Yorkshire terrier, one rottweiler, and Sapsarees (data not shown). Thus, it is unlikely that the canine PPD mutation is associated with *Alx4* contraction in dog breeds other than Great Pyrenees. In fact, even in Great Pyrenees, the association with $Alx4^{\Delta 51}$ alleles is genetically ambiguous, whereas DC-1 or DC-2 associations are clearly Mendelian with an autosomal-dominant mode of inheritance.

The authors thank P. Henthorn for providing needed materials. This work was supported in part by the Korea Research Foundation and by Korea Science and Engineering Foundation.

LITERATURE CITED

Blanc, I., A. Bach and B. Robert, 2002 Unusual pattern of Sonic hedgehog expression in the polydactylous mouse mutant Hemimelic extra-toes. Int. J. Dev. Biol. **46:** 969–974.

Capellini, T. D., G. Di Giacomo, V. Salsi, A. Brendolan, E. Ferretti *et al.*, 2006 Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and shh expression. Development **133:** 2263–2273.

Clark, R. M., P. C. Marker and D. M. Kingsley, 2000 A novel candidate gene for mouse and human preaxial polydactyly with altered expression in limbs of Hemimelic extra-toes mutant mice. Genomics **67:** 19–27.

Clark, R. M., P. C. Marker, E. Roessler, A. Dutra, T. C. Schimenti *et al.*, 2001 Reciprocal mouse and human limb phenotypes caused by gain- and loss-of-function mutations affecting Lmbr1. Genetics **159:** 715–726.

Fondon, J. W., III, and H. R. Garner, 2004 Molecular origins of rapid and continuous morphological evolution. Proc. Natl. Acad. Sci. USA **101:** 18058–18063.

Fogle, B., 2000 *Domestic Dog Breeds: The New Encyclopedia of the Dog.* Dorling Kindersley, London.

Galis, F., J. J. M. van Alphen and J. A. J. Metz, 2001 Why five fingers? Evolutionary constraints on digit numbers. Trends Ecol. Evol. **16:** 637–646.

Goode, D. K., P. Snell, S. F. Smith, J. E. Cooke and G. Elgar, 2005 Highly conserved regulatory elements around the Shh gene may contribute to the maintenance of conserved synteny across human chromosome 7q36.3. Genomics **86:** 172–181.

Gurnett, C. A., A. M. Bowcock, F. R. Dietz, J. A. Morcuende, J. C. Murray *et al.*, 2007 Two novel point mutations in the long-range SHH enhancer in three families with triphalangeal thumb and preaxial polydactyly. Am. J. Med. Genet. A **143:** 27–32.

Hammond, K. L., I. M. Hanson, A. G. Brown, L. A. Lettice and R. E. Hill, 1998 Mammalian and Drosophila dachshund genes are related to the Ski proto-oncogene and are expressed in eye and limb. Mech. Dev. **74:** 121–131.

Hecksher-Sorensen, J., R. E. Hill and L. A. Lettice, 1988 Double labeling for whole-mount in situ hybridization in mouse. Biotechniques **24:** 914.

Heus, H. C., A. Hing, M. J. van Baren, M. Joosse, G. J. Breedveld *et al.*, 1999 A physical and transcriptional map of the preaxial polydactyly locus on chromosome 7q36. Genomics **57:** 342–351.

Hill, R. E., S. J. H. Heaney and L. A. Lettice, 2003 Sonic hedgehog: restricted expression and limb dysmorphologies. J. Anat. **202:** 13–20.

Hing, A. V., C. Helms, R. Slaugh, A. Burgess, J. C. Wang *et al.*, 1995 Linkage of preaxial polydactyly type 2 to 7q36. Am. J. Med. Genet. **58:** 128–135.

Horikoshi, T., N. Endo, M. Shibata, P. Heutink, R. E. Hill *et al.*, 2003 Disruption of the C7orf2/Lmbr1 genic region is associated with preaxial polydactyly in humans and mice. J. Bone Miner. Metab. **21:** 1–4.

Ianakiev, P., M. J. van Baren, M. J. Daly, S. P. Toledo, M. G. Cavalcanti *et al.*, 2001 Acheiropodia is caused by a genomic deletion in C7orf2, the human orthologue of the Lmbr1 gene. Am. J. Hum. Genet. **68:** 38–45.

Kim, K. S., Y. Tanabe, C. K. Park and J. H. Ha, 2001 Genetic variability in East Asian dogs using microsatellite loci analysis. J. Hered. **92:** 398–493.

Lathrop, G. M., and J. M. Lalouel, 1984 Easy calculations of lod scores and genetic risks on small computers. Am. J. Hum. Genet. **36:** 460–465.

Leonard, J. A., R. K. Wayne, J. Wheeler, R. Valadez, S. Guillen *et al.*, 2002 Ancient DNA evidence for Old World origin of New World dogs. Science **298:** 1613–1616.

Lettice, L. A., S. J. Heaney, L. A. Purdie, L. Li, P. de Beer *et al.*, 2003 A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum. Mol. Genet. **12:** 1725–1735.

Li, R., E. Mignot, J. Faraco, H. Kadotani, J. Cantanese *et al.*, 1999 Construction and characterization of an eightfold redundant dog genomic bacterial artificial chromosome library. Genomics **58:** 9–17.

Lopez-Martinez, A., D. T. Chang, C. Chiang, J. A. Porter, M. A. Ros *et al.*, 1995 Limb-patterning activity and restricted posterior

localization of the amino-terminal product of Sonic hedgehog cleavage. Curr. Biol. **5:** 791–796.

Maas, S. A., and J. F. Fallon, 2004 Isolation of the chicken Lmbr1 coding sequence and characterization of its role during chick limb development. Dev. Dyn. **229:** 520–528.

Maas, S. A., and J. F. Fallon, 2005 Single base pair change in the long-range Sonic hedgehog limb-specific enhancer is a genetic basis for preaxial polydactyly. Dev. Dyn. **232:** 345–348.

Masuya, H., T. Sagai, S. Wakana, K. Moriwaki and T. Shiroishi, 1995 A duplicated zone of polarizing activity in polydactylous mouse mutants. Genes Dev. **9:** 1645–1653.

Masuya, H., T. Sagai, K. Moriwaki and T. Shiroishi, 1997 Multigenic control of the localization of the zone of polarizing activity in limb morphogenesis in the mouse. Dev. Biol. **182:** 42–51.

Miller, M. E., 1979 *Miller's Anatomy of the Dog*, Ed. 2. Saunders, Philadelphia.

Park, K., J. Kang, S. Park, J. Ha and C. Park, 2004 Linkage of the locus for canine dewclaw to chromosome 16. Genomics **83:** 216–224.

Parker, H. G., L. V. Kim, N. B. Sutter, S. Carlson, T. D. Lorentzen *et al.*, 2004 Genetic structure of the purebred domestic dog. Science **304:** 1160–1164.

Qu, S., S. C. Tucker, J. S. Ehrlich, J. M. Levorse, L. A. Flaherty *et al.*, 1998 Mutations in mouse Aristaless-like4 cause Strong's luxoid polydactyly. Development **125:** 2711–2721.

Riddle, R. D., R. L. Johnson, E. Laufer and C. Tabin, 1993 Sonic hedgehog mediates the polarizing activity of the ZPA. Cell **75:** 1401–1416.

Robert, B., and Y. Lallemand, 2006 Anteroposterior patterning in the limb and digit specification: contribution of mouse genetics. Dev. Dyn. **235:** 2337–2352.

Ruvinsky, A., 2001 Developmental genetics, pp. 431–512 in *The Genetics of the Dog*, edited by A. Ruvinsky and J. Sampson. CABI, Oxon, UK.

Sagai, T., M. Hosoya, Y. Mizushina, M. Tamura and T. Shiroishi, 2005 Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. Development **132:** 797–803.

Savolainen, P., Y. P. Zhang, J. Luo, J. Lundeberg and T. Leitner, 2002 Genetic evidence for an East Asian origin of domestic dogs. Science **298:** 1610–1613.

Sharpe, J., L. Lettice, J. Hecksher-Sorensen, M. Fox, R. Hill *et al.*, 1999 Identification of sonic hedgehog as a candidate gene responsible for the polydactylous mouse mutant Sasquatch. Curr. Biol. **9:** 97–100.

Stephens, M., N. J. Smith and P. Donnelly, 2001 A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. **68:** 978–989.

Sutter, N. B., and E. A. Ostrander, 2004 Dog star rising: the canine genetic system. Nat. Rev. Genet. **5:** 900–910.

Sutter, N. B., M. A. Eberle, H. G. Parker, B. J. Pullar, E. F. Kirkness *et al.*, 2004 Extensive and breed-specific linkage disequilibrium in Canis familiaris. Genome Res. **14:** 2388–2396.

Temtamy, S. A., and V. A. McKusick, 1978 The genetics of hand malformations. Birth Defects Orig. Artic. Ser. **14(3):** i–xviii, 1–619.

Wojnar, P., M. Lechner and B. Redl, 2003 Antisense down-regulation of lipocalin-interacting membrane receptor expression inhibits cellular internalization of lipocalin-1 in human NT2 cells. J. Biol. Chem. **278:** 16209–16215.

Zguricas, J., H. Heus, E. Morales-Peralta, G. Breedveld, B. Kuyt *et al.*, 1999 Clinical and genetic studies on 12 preaxial polydactyly families and refinement of the localisation of the gene responsible to a 1.9 cM region on chromosome 7q36. J. Med. Genet. **36:** 32–40.

Zhang, K., Z. Qin, T. Chen, J. S. Liu, M. S. Waterman *et al.*, 2005 HapBlock: haplotype block partitioning and tag SNP selection software using a set of dynamic programming algorithms. Bioinformatics **21:** 131–134.

Communicating editor: L. Siracusa