A LINE-1 insertion in DLX6 is responsible for cleft palate and mandibular abnormalities in a canine model of Pierre Robin sequence.
A LINE-1 Insertion in DLX6 Is Responsible for Cleft Palate and Mandibular Abnormalities in a Canine Model of Pierre Robin Sequence

Zena T. Wolf1, Elizabeth J. Leslie2, Boaz Arzi3, Kartika Jayashankar1, Nili Karmi1, Zhonglin Jia2,4,5, Douglas J. Rowland6, Amy Young7, Noa Safra1, Saundra Sliskovic1, Jeffrey C. Murray2, Claire M. Wade8, Danika L. Bannasch1,*

1 Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, Davis, California, United States of America, 2 Department of Pediatrics, University of Iowa, Iowa City, Iowa, United States of America, 3 Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California, Davis, Davis, California, United States of America, 4 State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China, 5 Department of Cleft Lip and Palate Surgery, West China Hospital of Stomatology, Sichuan University, Chengdu, China, 6 Center for Molecular and Genomic Imaging, University of California, Davis, Davis, California, United States of America, 7 Department of Animal Science, University of California, Davis, Davis, California, United States of America, 8 Faculty of Veterinary Science, University of Sydney, Sydney, New South Wales, Australia

Abstract

Cleft palate (CP) is one of the most commonly occurring craniofacial birth defects in humans. In order to study cleft palate in a naturally occurring model system, we utilized the Nova Scotia Duck Tolling Retriever (NSDTR) dog breed. Micro-computed tomography analysis of CP NSDTR craniofacial structures revealed that these dogs exhibit defects similar to those observed in a recognizable subgroup of humans with CP: Pierre Robin Sequence (PRS). We refer to this phenotype in NSDTRs as CP1. Individuals with PRS have a triad of birth defects: shortened mandible, posteriorly placed tongue, and cleft palate. A genome-wide association study in 14 CP NSDTRs and 72 unaffected NSDTRs identified a significantly associated region on canine chromosome 14 (24.2 Mb–29.3 Mb; Prho = 4.64 × 10⁻¹⁵). Sequencing of two regional candidate homeobox genes in NSDTRs, distal-less homeobox 5 (DLX5) and distal-less homeobox 6 (DLX6), identified a 2.1 kb LINE-1 insertion within DLX6 in CP1 NSDTRs. The LINE-1 insertion is predicted to insert a premature stop codon within the homeodomain of DLX6. This conserved DLX5 homeobox of a patient with PRS was identified. This suggests the involvement of DLX6 in the development of PRS. These results demonstrate the power of the canine animal model as a genetically tractable approach to understanding naturally occurring craniofacial birth defects in humans.

Citation: Wolf ZT, Leslie EJ, Arzi B, Jayashankar K, Karmi N, et al. (2014) A LINE-1 Insertion in DLX6 Is Responsible for Cleft Palate and Mandibular Abnormalities in a Canine Model of Pierre Robin Sequence. PLoS Genet 10(4): e1004257. doi:10.1371/journal.pgen.1004257

Copyright: © 2014 Wolf et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Editor: Gregory S. Barsh, Stanford University School of Medicine, United States of America

Received November 20, 2013; Accepted February 4, 2014; Published April 3, 2014

Funding: This work was supported by the National Institutes of Health National Institute of Dental and Craniofacial Research [R01DE022532 to DLB and R01DE08559 to JCM]; American Kennel Club- Canine Health Foundation [01424]; Gray Lady Foundation [201017407]; UC Davis Center for Companion Animal Health [C223]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: A DNA test for the DLX6 mutation in Nova Scotia Duck Tolling Retriever is commercially available through BDH Genetics, which is owned by a family member of DLB. This commercial venture had no influence on the interpretation of the results presented within the manuscript.

* E-mail: dlbannasch@ucdavis.edu

Introduction

Cleft palate (CP) is one of the most commonly occurring craniofacial birth defects, affecting approximately 1 in 1500 live human births in the United States [1]. Children born with CP may develop hearing loss, difficulties with speech and eating, and may be at an increased risk for psychiatric disorders and neurological deficits [2–4]. CP occurs when there is a failure in the formation of the secondary palate, which makes up all of the soft palate and majority of the hard palate. Secondary palate development is conserved across mammalian species and proceeds through highly regulated sequential steps: palatal shelf growth, elevation, fusion, and cell differentiation [reviewed in [5]]. Disruptions in any of these pathways may cause a cleft palate and lead to the phenotypic spectrum of CP cases that is observed. CP may occur alone (nonsyndromic) or with other abnormalities (syndromic).

Pierre Robin sequence (PRS, OMIM 261800) is a heterogeneous and phenotypically variable subgroup of CP that affects 1 in 8500 live human births [6]. PRS is characterized by a cleft palate, shortened mandible (micrognathia), and posteriorly placed tongue (glossoptosis). PRS is thought to be the result of a sequence of events caused by the primary defect, micrognathia [7]. The etiology of PRS is still largely unknown and highly variable. PRS may occur alone or as part of a syndrome, such as Stickler syndrome, Velocardiofacial syndrome, and Treacher Collins syndrome [8,9]. A high incidence within families and among twins suggests a genetic etiology. Familial aggregation with an autosomal dominant mode of inheritance has been observed with translocations of 17q24 and a reduction of SOX9 and KCNJ2 expression [10,11]. PRS also occurs at a high incidence among families with cleft lip and palate [12,13]. However, monzygotic twins discordant for PRS are also observed, suggesting that PRS may be a result of the twinning process or of mandibular constraint in utero [13,14].
Results

In an effort to understand the genetic basis of craniofacial birth defects such as PRS, we used a relatively unconventional model organism, the domestic dog. Dogs have naturally occurring birth defects, with inherited orofacial clefts that resemble those observed in humans [15–17]. Domestication and subsequent pedigreed breed creation from a small number of founders has led to a unique genetic background, resulting in a small number of genetic variants being responsible for the broad phenotypic diversity observed [18]. Compared to humans, dogs are amendable to association-based mapping studies with a small number of samples due to their relatively long linkage disequilibrium blocks within breeds [19]. Here, we demonstrate how a naturally occurring model of PRS in the Nova Scotia Duck Tolling Retriever (NSDTR) breed, characterized by CP and relative micrognathia, led to the identification of both the first mutation responsible for cleft palate in dogs and candidate genes for PRS in people.

Sanger Sequencing of DLX5 and DLX6

Located within the interval defined by the genome-wide association study are 21 genes (Table S3). DLX5 and DLX6 (cfa14: 25014704-25033706) were selected for sequencing due to their roles as transcription factors in craniofacial development and the similar phenotypes observed in mutant mice [20–22]. The coding regions and intronic regions with high conservation across species of DLX5 and DLX6 were sequenced in one CP1 case and one unaffected control NSDTR. One intronic nucleotide insertion (25032667-25032667insG) was identified within DLX5 in both the CP1 case and unaffected control NSDTR. A 2056 bp insertion was identified within a highly conserved region of DLX6 intron 2 at cfa14.25016716 in the CP1 case when compared to the unaffected control and reference sequence. The insertion is 82 bp into intron 2 and is flanked by a 13 bp target site duplication. A BLAST search identified 99% similarity with a LINE-1 element (GenBank AC187025.7) [23].

Transcription of the LINE-1 Element

Transcription analysis was performed in cDNA from cerebral cortex for both DLX5 and DLX6. DLX5 and DLX6 transcripts were Sanger sequenced in one neonatal CP1 NSDTR with the LINE-1 insertion and compared to one neonatal unaffected control NSDTR. No polymorphisms were identified within
DLX5. cDNA from the CP1 NSDTR showed both the wildtype DLX6 transcript and a larger transcript, which contained 1281 base pairs of the intronic LINE-1 insertion (Figure 3B).

In order to quantify the amount of both wildtype DLX6 transcript and the larger mutant DLX6 transcript, real-time PCR of DLX6 from cerebral cortex cDNA was performed in 3 neonatal CP1 NSDTRs and 3 neonatal unaffected control NSDTRs. REST analysis indicated that in CP1 NSDTRs, when compared to control NSDTRs, the DLX6 wildtype transcript was downregulated by a mean factor of 0.252 (p = 0.069), while the DLX6 mutant transcript was significantly upregulated by a mean factor of 60.033 (p = 0.012; Figures 3C and 3D) [24]. Real-time PCR was also performed on DLX5 in the same samples. There was no significant change in DLX5 expression levels between CP1 cases and unaffected controls in the tissues examined.

Sequence analysis and translation of the approximately 1.2 kilobase LINE-1 insertion predicts an in frame stop codon after the insertion of a new exon (Figure 4). A premature stop codon is predicted to truncate 17 amino acids from the 60 amino acid functional DNA binding homeodomain.

Segregation Analysis and Allele Frequency

In order to investigate segregation of the DLX6 LINE-1 insertion, PCR genotyping was performed on available DNA from parents, littermates, and the 12 CP1 NSDTRs (Figure 5). All 12 CP1 NSDTRs with the associated haplotype were homozygous for the LINE-1 insertion or mutant allele. Within families of the 12 CP1 NSDTRs, nine parents and 17 littermates were available for genotyping. Nine parents and 14 littermates were heterozygous for the mutant allele, while the 3 remaining littermates were homozygous for the wildtype allele. This indicates that the DLX6 LINE-1 insertion segregates both with the phenotype and with an autosomal recessive mode of inheritance.

Additional dogs were PCR genotyped to determine the allele frequency of the DLX6 LINE-1 insertion (Table 1). Within the NSDTR breed, 96 dogs were genotyped and 80 NSDTRs did not carry the insertion, while the remaining 16 NSDTRs were heterozygous for the insertion. To determine if the insertion was shared among other breeds, 35 affected dogs from 20 other breeds and 284 unaffected dogs from 69 breeds were genotyped. No carriers were identified. This is consistent with a fully penetrant...
autosomal recessive causative mutation that is private to the NSDTR breed.

**Screening of Human Cleft Palate Cases**

To determine if variation in DLX5 or DLX6 contributed to cleft palate in humans, a cohort of patients with a variety of manifestations of cleft palate were sequenced. Sequencing of 59 patients with nonsyndromic CP (NSCP) and 92 patients with nonsyndromic clefting of the lip and palate (NSCLP) identified 7 novel intronic or 3’UTR variants (Table 2). The intronic variant at chr9: 96635849 was present in 1.25% of chromosomes and was absent from 1000 Genomes (p = 0.03) [25]. However, none of the variants was significantly associated when correcting for multiple comparisons. To test for overtransmission of a common allele to affected offspring, a set of 362 case-parent trios from the US were genotyped across a subset of SNPs (rs2272280, rs3801290, and

---

**Figure 2. Phenotype of neonatal CP1 NSDTRs.** A. Neonatal CP1 NSDTR with an extensive cleft of the hard and soft palate. B. Neonatal NSDTR with a normal palate (WT). C. Lateral view of CP1 head exhibiting relative mandibular brachygnathia. D. Lateral view of WT head with a normal jaw relationship. E. Coronal CT image depicting the failure of the palatine processes and nasal septum to fuse in CP1 NSDTRs. F. Coronal CT image depicting midline fusion of palatal process and nasal septum in WT. P – Palatine process, NS – Nasal septum. G. 3D reconstruction of microCT imaging of CP1 and WT skulls with mandibles removed. CP1 skull shows abnormally shaped palatine process and palatine bones. WT skull shows anatomical location of normal palatine sutures and shape of palatine processes and palatine bones. H. 3D reconstruction of mandibles depicting abnormal angulation of the condylar process (*) in CP1 mandibles compared to WT mandibles. doi:10.1371/journal.pgen.1004257.g002
rs3213654) and tested with TDT analysis. These SNPs were not significantly associated with NSCLP or NSCP (p = 0.4, data not shown).

The screening of 31 patients with syndromic CP identified 4 novel rare variants, two of which were protein coding (Table 2). These include a synonymous variant at p.Phe282 and a novel missense variant in a patient with PRS (DLX5: p.Ile192Met; NM_005221.5:c.576C> G). The missense variant was predicted to be damaging by SIFT, but was inherited from the affected individual's unaffected mother. However, this variant affects a

![Figure 3. PCR amplification and expression analysis of LINE insertion in CP1 and WT NSDTRs.](image)

**A.** Gel image of long range PCR amplification of the DLX6 intronic LINE-1 insertion from genomic DNA of an unaffected NSDTR (WT; 2500 bp band), CP1 NSDTR (CP1; 4600 bp band), and negative water control (Neg). CP1 has a 2.1 kb LINE insertion. **B.** Gel image of RT-PCR amplification of DLX6 from WT and CP1 cDNA. CP1 expresses both a wildtype (1390 bp) and mutant (2600 bp) transcript. The mutant transcript has a 1.2 kb insertion. GAPDH was used to control for cDNA concentrations. **C.** Relative DLX6 gene expression ratios by transcript of cerebral cortex cDNA from 3 neonatal CP1 NSDTRs compared to 3 neonatal WT NSDTRs. Boxes represent the interquartile range, and the dotted lines within represent median gene expression. Whiskers of the boxplot represent minimum and maximum observations. Relative expression levels were normalized to the housekeeping gene, B2M. Statistical significance is reported as p < 0.05 (*). **D.** Summary of relative expression of wildtype and mutant DLX6 transcripts of CP1 NSDTRs compared to WT NSDTRs. Fold change and p-values were calculated using REST2009 [24].

doi:10.1371/journal.pgen.1004257.g003

![Figure 4. Schematic illustration of genomic and cDNA DLX6 gene structure in unaffected (WT) and CP1 NSDTRs (CP1).](image)

Nucleotides boxed and in bold are the 13 base pair target site duplication identified as part of the DLX6 LINE-1 insertion. Conservation represents the UCSC genome browser comparative genomics conservation track of human, dog, mouse, and rat sequence conservation. The region of conservation represented in red is the region disrupted by the LINE-1 insertion. Image is not drawn to scale.

doi:10.1371/journal.pgen.1004257.g004
highly conserved residue of the DNA-binding domain of DLX5 that is conserved across vertebrates and was absent from both the 1000 Genomes and NHLBI ESP databases [25,26]. An additional 15 patients with Pierre Robin sequence were sequenced, but no additional mutations were identified.

Discussion

A naturally occurring animal model for PRS was discovered in twelve cases of CP1 NSDTRs, which exhibit relative mandibular brachygnathia and cleft palate. A genome-wide association study within NSDTRs with CP identified a shared 5.1 Mb homozygous haplotype among 12 CP1 cases with relative mandibular brachygnathia. From the associated interval, DLX5 and DLX6 were selected as regional candidate genes based on their roles in development. Sequencing of these regional candidate genes in NSDTRs identified an intronic LINE-1 insertion within DLX6 that segregates with the phenotype in the breed. This discovery prompted the sequencing of the same genes in a human cohort of CP cases, which identified a damaging missense mutation within the homeobox of DLX5 in a patient with PRS was identified.

The regional candidate genes, DLX5 and DLX6, make up a pair of convergently transcribed homeobox containing transcription factors [27]. They are involved in patterning of craniofacial structures, inner ear, limb, and brain development with roles in chondrocyte and osteoblast differentiation [21,28,29]. The functions of Dlx5 and Dlx6 were studied by targeted inactivation of murine homologs. Developmental expression of both genes was observed in the first pharyngeal arch, brain, and skeleton [27]. Single gene mutants (Dlx5<sup>−/−</sup> and Dlx6<sup>−/−</sup>) were perinatal lethal and resulted in brain, bone, and inner ear defects, with craniofacial abnormalities including a cleft palate, hypoplastic condylar process, and shortened mandible [20–22]. Double mutants (Dlx5<sup>−/−</sup>;Dlx6<sup>−/−</sup>) exhibited more severe craniofacial, inner ear, and bone defects, and were a phenocopy of split hand/foot malformation [28]. They also exhibited homeotic transformation of the mandible into a maxilla indicating that these genes were responsible for normal patterning of the mandible [21,28,29].

Phenotypic similarity between CP1 NSDTRs, the PRS patient, and mutant mice are observed with the changes in the condylar process, cleft palate, and relatively shortened mandibles. There may be associated condylar hypoplasia in PRS, but this was not investigated in the PRS patient [30,31]. Shortened mandibles have

Table 1. Summary of allele frequencies of genotyping results.

| Sample            | N   | Allele Frequency* |
|-------------------|-----|-------------------|
| NSDTRs            | 96  | 17%               |
| Non NSDTRs with cleft | 35  | 0%                |
| Non NSDTRs        | 284 | 0%                |

*Allele frequencies indicate the number of individuals with at least one copy of the mutant allele.

doi:10.1371/journal.pgen.1004257.t001
Table 2. Minor allele frequency of DLX5 and DLX6 variants among human cleft cohort and control cohort.

| Gene | Position (Hg19) | rs ID | Location | Alleles   | NAF (%) | p-value | NAF (%) | p-value | NAF (%) | p-value | NAF (%) | p-value | NAF (%) | p-value | NAF (%) | p-value |
|------|-----------------|-------|----------|-----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|      | DLX5            |       |          |           |         |         |         |         |         |         |         |         |         |         |         |         |
|      | 96651698        |       | intronic | C/T       | 0.6%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         |
|      | 96651406        | rs1207733 | intronic | G/A       | 7.2%    |         | 7.0%    |         | 9.1%    |         | 1.1%    |         | 1.9%    |         | 6.1%    |         | 1.9%    |         |
|      | 96651391        | rs193302639 | intronic | G/A       | 0.6%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         |
|      | 96650408        |       | intronic | C/T       | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 1.1%    |         | 0.0%    |         | 0.0%    |         |
|      | 96650342        | I192M | C/G      | 0.0%      |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 1.1%    |         | 1.9%    |         | 0.0%    |         |
|      | 96650216        | rs35273378 | S234R    | C/A       | 1.2%    |         | 1.8%    |         | 2.4%    |         | 1.1%    |         | 1.9%    |         | 0.0%    |         | 1.2%    |         |
|      | 96650098        | rs19967826 | P274A    | C/G       | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         |
|      | 96649884        | rs11277247 | 3’ UTR   | -/C       | 29.3%   |         | 30.2%   |         | 30.3%   |         | 29.5%   |         | 32.7%   |         | 43.0%   |         | 43.0%   |         |
|      |                 |       |          |           |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
|      | DLX6            |       | intronic | C/T       | 0.6%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         |
|      | 96635741        | rs185935123 | intronic | C/T       | 1.3%    |         | 0.0%    |         | 0.0%    |         | 1.8%    |         | 3.8%    |         | 0.4%    |         | –       |         |
|      | 9663849         |       | intronic | G/A       | 1.25%, p = 0.03 | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         |
|      | 96636865        | rs1207727 | intronic | A/C       | 32.2%   |         | 3.2%    |         | 11.0%   |         | 6.4%    |         | 6.3%    |         | 5.8%    |         | –       |         |
|      | 96637047        | rs141424124 | Q178Q   | G/A       | 0.6%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 2.4%    |         |
|      | 96637185        |       | intronic | C/T       | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | –       |         |
|      | 96637654        |       | intronic | A/G       | 0.6%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | –       |         |
|      | 96638021        | rs1207728 | intronic | G/C       | 20.1%   |         | 25.0%   |         | 25.6%   |         | 20.7%   |         | 21.7%   |         | 19.7%   |         | –       |         |
|      | 96638045        | rs1004278 | intronic | G/C       | 37.4%   |         | 26.8%   |         | 27.9%   |         | 26.3%   |         | 26.1%   |         | 30.5%   |         | –       |         |
|      | 96638237        | rs3801290 | intronic | A/G       | 36.8%   |         | 26.8%   |         | 27.9%   |         | 26.3%   |         | 26.1%   |         | 30.5%   |         | –       |         |
|      | 96638267        | rs3801289 | intronic | A/C       | 29.1%   |         | 37.5%   |         | 36.0%   |         | 41.3%   |         | 34.8%   |         | 36.7%   |         | –       |         |
|      | 96638763        |       | intronic | T/C       | 0.0%    |         | 0.0%    |         | 0.0%    |         | 1.1%    |         | 1.9%    |         | –       |         | –       |         |
|      | 96639323        |       | P282P    | A/G       | 0.6%    |         | 1.0%    |         | 1.3%    |         | 1.2%    |         | 1.9%    |         | 0.2%    |         | –       |         |
|      | 96639368        | rs3213654a | 3’ UTR   | A/G       | 1.15%, p = 0.04 | 2.9%    |         | 1.3%    |         | 2.4%    |         | 1.9%    |         | 0.0%    |         | 0.0%    |         |
|      | 96639505        | rs147798666 | 3’ UTR   | -/CTCT   | 20.1%   |         | 23.5%   |         | 23.7%   |         | 19.0%   |         | 21.2%   |         | –       |         | –       |         |
|      | 96639654        |       | 3’ UTR   | C/A       | 0.6%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | 0.0%    |         | –       |         | –       |         |
|      | 96639683        |       | 3’ UTR   | C/T       | 0.0%    |         | 1.0%    |         | 1.4%    |         | 0.0%    |         | 0.0%    |         | –       |         | –       |         |
|      | 96639753        | rs34606666 | 3’ UTR   | C/T       | 37.9%   |         | 27.1%   |         | 29.2%   |         | 25.6%   |         | 26.0%   |         | 30.5%   |         | –       |         |
|      | 96640130        | rs2272280b | 3’ UTR   | C/T       | 1.4%    |         | 2.7%    |         | 1.16%, p = 0.01 | 3.3%    |         | 1.9%    |         | 0.0%    |         | –       |         |
|      | 96640311        | rs191071332 | 3’ UTR   | T/C       | 2.7%    |         | 0.9%    |         | 1.2%    |         | 1.1%    |         | 0.0%    |         | 1.2%    |         | –       |         |

N: Number of individuals sequenced; CLP: Cleft palate; NS: Nonsyndromic; CP: Cleft lip and palate.

- Genotypes count not available.
- a,b These SNPs have minor allele frequencies from 7–11% in non-European populations.
- SNPs with MAF listed in bold are statistically significant when compared to MAF of 1000 genomes and ESP databases [25,26].
been observed to cause cleft palate in mice studies [7,32–34]. Palatal development is a highly regulated sequence of events involving the repositioning of the palatal shelves from a lateral to horizontal orientation over the tongue. During this process, the tongue repositions by elongation of the mandible, allowing for reorientation of the palatal shelves (reviewed in [3]). If the mandible does not elongate, the tongue cannot reposition, which in turn obstructs palatal shelf elevation and leads to a cleft palate [7]. This supports the role of DLX5 and DLX6 in the phenotype observed in the PRS patient and CP1 NSDTRs. This also suggests that the CP1 NSDTRs are naturally occurring animal models for PRS because, although other disorders may be associated with cleft palate and micrognathia, these dogs lack additional abnormalities [35].

DLX5 and DLX6 contain homeoboxes that regulate transcription by binding to specific DNA sequences. Homeoboxes are highly conserved nucleotide sequences of 180 base pairs. Point mutations located within the homeobox give rise to disease at a higher frequency than mutations located within the rest of the gene and often show a dominant effect due to haploinsufficiency [36]. DLX5 is sensitive to haploinsufficiency because, although Dlx5+/− mice appear normal, closer investigation indicates that they have a decreased bone mineral density [20,37]. Mutations within the homeobox may also affect proper structural folding of the protein and DNA binding specificity leading to a mutant phenotype [38]. This is observed in a homozygous nonsense mutation in the homeobox of DLX5 in two affected family members in a consanguineous pedigree with split-hand/foot malformation (SHFM) [39]. The affected individuals were noted to have clefts of the hands, but no cleft palate or craniofacial abnormalities with the exception of a mildly pronounced forehead. The homozygous nonsense mutation identified in the PRS patient (NM_005221.5:c.576C>T) is located within the homeobox sequence of DLX5, which supports its contribution to the observed phenotype.

The Dlx5 homozygous nonsense mutation within the PRS patient does not segregate with a Mendelian mode of inheritance, but this is not expected since PRS is a complex trait. PRS was not apparent in the mother, but she may have exhibited subtle craniofacial abnormalities that went unnoticed. This likely complex inheritance suggests a role for additional, not yet identified loci or environmental factors. Support for this is observed when mice with a homozygous expression of a Dlk5 null allele and targeted inactivation of a transcription factor responsible for clefting in people, Msx1, exhibit cleft palate (Dlk5−/−; Msx1−/−) [40]. However, when expression of both genes is disrupted (Dlk5−/−; Msx1−/−), mice have no cleft palate.

Nine noncoding sequence variants were identified within DLX5 and DLX6 of human patients with orofacial clefts. Functional analysis of the nine variants may provide further insight into the possible contribution of these genes to a cleft palate phenotype. Noncoding regions have important regulatory functions as they have been observed to disrupt gene expression. Reduced DLX5 and DLX6 expression was observed in an autistic patient with a SNP in the DLX5 and DLX6 intergenic region [41]. Hearing loss and craniofacial defects including cleft palate and micrognathia have been observed in the deletion of a DLX5 and DLX6 enhancer region [42]. The LINE-1 insertion within DLX6 of the CP1 NSDTRs is also located within a noncoding region that is highly conserved. According to the UCSC genome browser, the HMR conserved transcription factor binding site regulation track indicates that the LINE-1 insertion identified within DLX6 CP1 NSDTRs disrupts a binding domain for SUZ12. SUZ12 is a long noncoding RNA that encodes a core of the polycomb repressive complex2 that has regulatory functions in the developing embryo [43]. The exact interaction of SUZ12 and DLX6 is not yet known, indicating that the LINE insertion may do more than disrupt transcription.

LINE elements are transposable elements that comprise 21% of the human genome, 16% of the dog genome, and often insert into intronic regions [44,45]. Intronic LINE element insertions are observed to cause disease in Duchenne-like muscular dystrophy and chronic granulomatous disease through the introduction of a new exon [46–48]. cDNA sequence from CP1 NSDTRs homozygous for the LINE-1 insertion indicate that 1 kb of the LINE element is spliced into the DLX6 transcript. Premature protein truncation of DLX6 is predicted due to an in frame stop codon after the formation of a new exon. Both the Dlx6 mutant mice (Dlx6−/−) and CP1 NSDTRs phenotypes are the result of truncating the same 3’ sequence of the homeodomain [22].

The LINE insertion disrupts transcription of DLX6 within CP1 NSDTRs and leads to downregulation of wildtype DLX6 transcript when compared to unaffected NSDTRs. As a result, only 25% of the normal expression levels are produced. The reduced expression of wildtype DLX6 transcript is not enough to prevent CP and the mandibular abnormalities. DLX5 and DLX6 expression levels have been observed to vary based on the timepoint and tissue examined with complex transcriptional regulation from tissue specific enhancers and noncoding RNAs [20,21,27,28,42,49–51]. It is possible that the LINE insertion works to disrupt appropriate timing of tissue specific expression since it is inserted into a highly conserved region within intron 2. Although not statistically significant, expression analysis from additional biological replicates from the correct tissue during the correct embryonic timepoint would likely yield significant values in the CP1 NSDTRs.

This discovery provides a genetic tool for the NSDTR breeder who wishes to avoid producing cleft palate affected puppies. The LINE insertion identified within the CP1 NSDTRs is unique to a subset of cases with CP within the breed and cannot be used as a tool to prevent against all genetic causes of CP. CP disease heterogeneity is observed in the NSDTR breed as 2 of the 14 CP cases did not possess relatively shortened mandibles and the associated LINE-1 insertion. This is unexpected in a purebred dog breed with few founders where the inbreeding coefficient is 0.26 [52]. This indicates that even in relatively genetically homogenous samples, heterogeneity and a complex etiology that mimics human cleft cases such as PRS is observed. This highlights the promise of the dog as an animal model for birth defects to identify multiple genes and/or pathways involved in craniofacial development.

In summary, identifying a mutation in an animal model with naturally occurring birth defects has enabled the identification of new candidate genes for PRS in people. This supports the continued use of the naturally occurring birth defects found within dogs and their unique genetic background to further our understanding of commonly occurring birth defects in people.

Materials and Methods

Ethics Statement

The use of samples involving human participants was approved by the institutional review board at the University of Iowa (approval #s 1998040800 and 1998040811). Informed consent was obtained and all clinical investigation was conducted according to the principles expressed in the Declaration of Helsinki. The collection of canine samples used in this study was approved by the University of California, Davis Animal Care and Use Committee (protocol #16092).
Canine Samples and DNA Extraction

Blood and tissue samples from NSDTRs with cleft palate (n = 14), healthy littermates (n = 24), parents (n = 11), unaffected NSDTRs (n = 153), and dogs with cleft palate across 20 breeds (n = 35) were collected from privately owned animals. When available, tissue was collected at post mortem examination and flash frozen. The evaluations of orofacial clefts were performed by visual inspection of affected dogs. Blood samples from unaffected dogs (n = 284) across 69 other breeds were collected from the William R. Pritchard Veterinary Medical Teaching Hospital. Genomic DNA was extracted from EDTA whole blood and tissue samples using Gentra Puregene DNA purification extraction kit (Qiagen, Valencia, CA).

Genome-Wide Association Study

Genome-wide SNP genotyping was performed with 14 cases and 72 controls using the Illumina CanineHD BeadChip (Illumina, San Diego, CA) with 173,662 markers. Samples had a genotyping call rate of ≥90%. 63,195 SNPs were excluded due to a minor allele frequency of ≤0.05 and 2,994 SNPs were excluded for a missing call rate of ≤10%. Chi-square analysis was performed using Plink [53]. 100,000 permutations were performed to correct for multiple tests. To determine if population stratification was present, quantile-quantile plots were generated in R using Plink output data [34]. The genomic inflation factor was also calculated using Plink.

CT Imaging

High-resolution micro-computed tomography (micro-CT) was used to evaluate craniofacial structures in 4 CP1 NSDTRs that were homozygous for the LINE-1 insertion, and in 3 normal NSDTRs homozygous for the wildtype allele. Samples were imaged at the Center for Molecular and Genomic Imaging (UC Davis). The skulls were kept in zip-lock bags and allowed to warm to the CT scanner temperature (29°C) inside of a custom plastic holder. CT images were obtained on the Centers MicroXCT-200 specimen CT scanner (Carl Zeiss X-ray Microscopy). The CT scanner has a variable x-ray source capable of a voltage range of 20–90 kV with 1–8W of power. Samples were mounted on the scanners sample stage, which has submicron level of position adjustments. Scan parameters were adjusted based on the manufacturer’s recommended guidelines: source and detector distances were 108 mm and 33 mm, respectively; the manufacturers LE4 custom filter was used for beam filtration; the voltage and power were set to 70 kV and 8W, respectively; and 1600 projections were acquired over 360-degrees with an exposure time of 1.5 s. Images were reconstructed on an isotropic voxel grid with 51.1507 microns per edge. Digital TIFF images were imported into Amira 5.4.5 (Visualization Sciences Group, FEI). For all specimens, tridimensional reconstructive (3D) images were generated to assess the spatial relationship of the bones. 3D length measurements were performed using the 3D length tool. Visual inspection of the micro-CT images was performed to identify any abnormalities.

PCR Amplification of Microsatellites and Linkage Analysis

Four regional microsatellites were mined from the UCSC genome browser (CanFam 2.0) “Variation and Repeats” database within cfa14: 24.2 Mb–29.3 Mb. Fluorescently labeled microsatellite primers were designed (Table S4) using Primer3 [55]. Microsatellites were PCR amplified using the following protocol: 94°C for 12 minutes, 7 cycles of 95°C for 20 seconds, 65°C for 30 seconds, 72°C for 2 minutes, 5 cycles of 93°C for 20 seconds, 55°C for 30 seconds, 72°C for 2 minutes, followed by a final annealing of 72°C for 2 minutes. Genotyping was performed on an ABI 3100 Genetic Analyzer (Applied Biosystems, CA) in 4 CP1 NSDTRs and 32 control NSDTRs. All genotypes were analyzed using STRand software [56]. LOD score analysis of microsatellite data was performed using Mendel software’s LOCATION_SCORES option in 59 individuals from three pedigrees [57]. The disease trait was coded as autosomal recessive and assumed to be fully penetrant.

Sequence of Candidate Genes

Primers to amplify the exons, intron/exon boundaries, and regions of high conservation within the intragenic and intergenic regions of DLX5 and DLX6 were designed in Primer3 (Table S3) [55]. PCR products were amplified in one CP1 NSDTR and one unaffected NSDTR. Areas with high GC content were amplified using Invitrogen AccuPrime GC-Rich DNA Polymerase protocols (Life Technologies, Grand Island, NY). PCR products were cleaned using ExoSAP-IT and sequenced using the Big Dye Terminator mix on an ABI 3100 Genetic Analyzer (Applied Biosystems, CA). Sequences were analyzed using Chromas (Technelysium, Tewantin, QLD, Australia) and Vector NTI (Informax, Frederick, MD). Sequences were aligned to each other and the Boxer reference sequence (CanFam 2.0) to identify any polymorphisms [19]. The DLX6 LINE-1 insertion was amplified using LongAmp Taq PCR Kit (New England Biolabs Ipswich, MA) and cloned using the Invitrogen TOPO TA Cloning kit (pCR2.1-TOPO vector) with One Shot TOP10 Chemically Competent E. coli. Products were isolated with the Qiaprep Spin Miniprep kit (Qiagen, Valencia, CA) and sequenced using an ABI 3100 Genetic Analyzer (Applied Biosystems, CA). LINE element sequence was identified using BLAST 2.2.28 [25].

RNA Extraction/cDNA Sequencing

Expression of DLX5 and DLX6 was evaluated in NSDTR lymphocytes, unaffected adult beagle tissue from cerebellum, cerebral cortex, heart, kidney, liver, skeletal muscle, skin, spinal cord, spleen, testis, and thymus. Total RNA was isolated from tissue samples using Qiagen QIAamp Blood Mini Kit (Valencia, CA) tissue protocols. Adult beagle total RNA was obtained from Zyagen (San Diego, CA). RNA was synthesized into cDNA using Invitrogen Superscript III First Strand Synthesis System to RT PCR product (Life Technologies, Grand Island, NY). GAPDH was amplified (forward primer 5’ AAGATTGTCCAGGATGCTGCC 3’ and reverse primer 5’ CCAGGAAATGAGCTTGACAAA 3’) in these tissues to ensure that equivalent amounts of cDNA were added. Invitrogen 5’ prime RACE system for Rapid Amplification of cDNA ends kit (Life Technologies, Grand Island, NY) was used to sequence the 5’ prime end of DLX6 from embryo cDNA. RACE primers were designed using Primer3 (Table S3) [55]. RACE PCR products were cloned using the Invitrogen TOPO TA Cloning kit (pCR2.1-TOPO vector) with One Shot TOP10 Chemically Competent E. coli. Products were isolated with the Qiaprep Spin Miniprep kit (Qiagen, Valencia, CA) and sequenced using an ABI 3100 Genetic Analyzer (Applied Biosystems, CA).

Genotyping

Genotyping primers were designed using Primer3 [55]. PCR genotyping was performed using a shared FAM labeled forward primer (5’ ACCATCGCTTGCAGGAAAAT 3’) and wildtype (5’ CTAGGCCCAATTCGCTGCC 3’) were designed. The PCR program...
was as follows: 94°C for 12 minutes, 35 cycles of 94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 45 seconds, followed by 72°C for 20 minutes. The wildtype product produced a 171 base pair product and the mutant product produced a 184 base pair product. GeneScan 500 ROX size standards were used and the reaction was analyzed on an ABI 3100 Genetic Analyzer (Applied Biosystems, CA). 96 unrelated unaffected NSDTRs and 284 unaffected dogs across 69 breeds were genotyped for the insertion. Nonsyndromic cleft palate cases (n = 35) across 20 breeds were genotyped. All genotypes were analyzed using STRand software [56].

Real Time Quantitative PCR

Primer sequences were generated using Primer3Plus (http://primer3plus.com/). A shared forward primer (5’aacctcag-tacctggcccttc 3’) was designed with reverse primers unique to wildtype (5’catacctacgtggttgtg 3’) and mutant (5’aacctcag-tacctggcccttc 3’). Semi quantitative PCR using AmpliTaq Gold DNA Polymerase was performed to test the quality of cDNA and primers, to confirm product size and to check for the presence of genomic DNA contamination. Real-time PCR was performed using the Rotor-Gene SYBR Green PCR Kit (QIAGEN, Valencia, CA) using a 2-step cycle protocol (45 cycles; Initial denaturation-5 minutes at 95°C; Annealing- 15 seconds at 95°C; Extension- 90 seconds at 95°C; Final Melt curve) on the Rotor Gene Q real-time PCR system. Cerebral cortex from 3 neonatal unaffected NSDTRs and 3 neonatal CP1 NSDTRs were run in triplicates with each replicate containing 0.2 ng template cDNA. cDNA was prepared as described above. All data was normalized to the housekeeping gene B2M5 [58]. Amplification and takeoff values were analyzed and graphed by REST2009 to determine any significant expression differences in DLX5 and DLX6 transcript levels between control and affected cDNA samples [24].

Human Samples

The case cohort consisted of 92 samples from individuals with nonsyndromic cleft lip with cleft palate (NSCLP) from the US, 59 samples from individuals with nonsyndromic cleft palate (NSCP) from the US and the Philippines, and 46 samples from individuals with cleft palate syndromes. The 46 syndromic samples consisted of 30 samples from individuals with PRS and 16 samples from individuals with nonsyndromic cleft palate (NSCP) from the US, 59 unrelated CEPH samples (CEU) were used as controls. Genotypes from European or Asian populations available from the 1000 Genomes Project and the NHLBI Exome Sequencing Project were used as controls [25,26]. Fisher’s exact test, implemented in STATA (v12.1), was used to test for differences in allele frequencies between NSCLP and NSCP cases and controls [60]. For insertions or deletions where allele frequencies were unavailable, we sequenced a set of 60 unrelated CEPH samples.

Genotyping of Human Samples

Three SNPs (rs2272280, rs3801290, and rs3213654) were genotyped in 362 case-parent trios (1090 individuals) with nonsyndromic cleft lip with or without cleft palate (NSCL/P) from the US using TaqMan SNP Genotyping Assays (Life Technologies, Grand Island, NY) on the ABI Prism 7900HT, and were analyzed with SDS 2.3 software (Applied Biosystems, Foster City, CA). The Family Based Association Test (FBAT) in STATA (v12.1) was used to test for association these NSCL/P case-parent trios [60].

Supporting Information

Table S1 Location of microsatellite markers on CFA14, calculated LOD scores, and recombination fraction (θ). Genomic location based on Can Fam 2.0 assembly. (DOCX)

Table S2 Average lengths of neonatal NSDTRs mandibles. Length measurements of the mandible are taken from the angular process to rostral tip of the mandibular body. (DOCX)

Table S3 Genes located within associated interval on chromosome 14. Genomic locations are based on the Can Fam 2.0 assembly and refer to chromosome 14 base pair locations. Genes listed were identified using the UCSC genome browser RefSeq gene annotation track. (DOCX)

Table S4 Primers and annealing temperatures for sequencing of canine samples. Genomic locations are based on the Can Fam 2.0 assembly and refer to chromosome 14 base pair locations. * LINE insert begins at cfa14.25016704. S – primers used for sequencing. (DOCX)

Table S5 Primers and annealing temperatures for sequencing of canine samples. Genomic locations are based on the hg19 assembly and refer to chromosome 7 base pair locations. S – primers used for sequencing. (DOCX)

Acknowledgments

The authors thank all of the people, dog owners, and dogs that contributed to this study.

Author Contributions

Conceived and designed the experiments: ZTW BA EJL NK JCM CMW DLB. Performed the experiments: ZTW EJL DJR KJ NK ZJ AY SS. Analyzed the data: ZTW EJL BA NK NS KJ JCM CMW DLB. Contributed reagents/materials/analysis tools: JCM DLB. Wrote the paper: ZTW EJL BA KJ DLB.

References

1. Parker SE, Mai CT, Canfield MA, Rickard R, Wang Y, et al. (2010) Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004–2006. Birth Defects Res A Clin Mol Teratol 88: 1008–1016.

2. Persson C, Elander A, Lohmander-Agerskov A, E S (2002) Speech outcomes in isolated cleft palate: impact of cleft extent and additional malformations. Cleft Palate Craniofac J 39: 397–400.
3. Christensen K, PB M (2002) Facial clefing and psychiatric diseases: a follow-up of the Danish 1936–1986 Facial Cleft cohort. Cleft Palate Craniofac J 39: 392–396.

4. Conrad AL, Canady J, Richman L, P N (2008) Incidence of neurological soft signs in children with isolated clef lip of the lip or palate. Percept Mot Skills 106: 197–206.

5. Bush JO, RJ (2012) Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development. Development 139: 231–243.

6. Bush PG, Williams AJ (1993) Incidence of the Robin Anomaloid (Pie Robin syndrome). Br J Plast Surg 36: 434–437.

7. Ricks JE, Ryder VM, Bridgewater LC, Schaua B, RE S (2002) Altered mandibular development precedes the time of palate closure in mice homozygous for disproportionate microtum: an oral clefing model supporting the 2Pe Pierre Robin sequence. Teratology 65: 116–120.

8. Cohen MM. (1991) Robin sequences and complexes: causal heterogeneity and pathogenetic/phenotypic variability. Am J Med Genet 48: 311–315.

9. Printzlaus A, MA A (2004) Pierre Robin sequence in Denmark: a retrospective simple genetic-based epidemiological study. Cleft Palate Craniofac J 41: 47–52.

10. Jakobsen LP, Ullmann R, Christensen SB, Jensen KE, et al (2007) Pierre Robin sequence may be caused by dysregulation of SOX9 and KCNJ2. J Med Genet 44: 381–386.

11. Benko S, Fantas JA, Amiel J, Kiënian DJ, Thomas S, et al (2009) Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat Genet 41: 359–364.

12. Jakobsen LP, Knudsen MA, Lespinasse J, Garcia Ayuso C, Ramos C, et al. (2009) Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat Genet 41: 359–364.

13. Depew MJ, Liu JK, Long JE, Presley R, Meneses JJ, et al. (1999) Dlx5 regulates gene expression during limb development in the 7q21 SHFM1 locus. Development 126: 3831–3846.

14. Kemp C, Thiele H, Dankof A, Schmidt G, Lauster C, et al. (2009) Cleft lip and/or palate with autosomal recessive transmission in Brittany spaniels. Cleft Palate Craniofac J 46: 81–88.

15. Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, et al. (2010) A simple genetic architecture underlies morphological variation in dogs. PLoS Biol 8: e1000451.

16. Moura E, Cirio SM, CT P (2011) Non-Syndromic Cleft Lip and Palate in Boxer dog. Journal of Dental Research 90: 588–590.

17. Simeone A, Acampora D, Pannese M, D’Esposito M, Stornaiuolo A, et al. (2001) Conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat Genet 41: 359–364.

18. Damante G, Pellizzari L, Esposito G, Fogelari F, Vighino P, et al. (1996) A canine model of Pierre Robin Syndrome. Br J Plast Surg 49: 16–20.

19. Cohen MM Jr. (1998) Syndrome with cleft lip and palate. Cleft Palate J 35: 377–405.

20. Damante G, Pellizzari L, Esposito G, Fogelari F, Vighino P, et al. (1996) A molecular code dictates sequence-specific DNA recognition by homeodomains. EMBO J 15: 4992–5000.

21. Depew MJ, Liu JK, Long JE, Presley R, Meneses JJ, et al. (1999) Dlx5 regulates gene expression during limb development in the 7q21 SHFM1 locus. Development 126: 3831–3846.

22. Damante G, Pellizzari L, Esposito G, Fogelari F, Vighino P, et al. (1996) A canine model of Pierre Robin Syndrome. Br J Plast Surg 49: 16–20.

23. Damage L, Plath K, Zeilinger J, Brambrink T, Meierhofer CA, et al. (2006) Polychromat complexes repress developmental regulators in murine embryonic stem cells. Nature 441: 349–353.

24. Kirkness EF, Bafina V, Halpern AL, Levy S, Remington K, et al. (2003) The dog genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 423: 597–615.

25. Kouwenhoven EN, van Heeringen SJ, Tena JJ, Oti M, Dutilh BE, et al. (2010) Molecular archeology of L1 insertions in the human genome. Genome Biol 3: research0092.

26. Smith BF, Yue Y, Woods BR, Kornegay JN, Shin JH, et al. (2011) An intronic LIN1 element inserted in the dystrophin gene aborts dystrophin expression and results in Duchenne-like muscular dystrophy in the corgi breed. Lab Invest 91: 216–231.

27. Meischl C, Boer M, Ahln A, DR R (2000) A new exon created by intronic insertion of a rearranged LIN-1 element as the cause of chronic granulomatous disease. Eur J Hum Genet 8: 697–703.

28. Landers ES, Linton LM, Birren B, Nussbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.

29. Kroonenhoven EN, van Heeringen SJ, Tena JJ, Oti M, Dutilh BE, et al. (2010) Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 SHFM1 locus. PLoS Genet 6: e1001065.

30. Zerucha T, Stuhmer T, Hatch G, Park BK, Long Q, et al. (2000) A highly conserved enhancer in the Dlx5/Dlx6 region in a family with hearing loss, craniofacial defects, and an inv(7)(q21.3q35). Hum Genet 123: 171–181.

31. Kaufman CB, Mackowick D, Bostock MS, Landman D, et al. (2002) Molecular archeology of L1 insertions in the human genome. Genome Biol 3: research0003.

32. Seegmiller RE, FC F (1977) Mandibular growth retardation as a cause of cleft palate in Moxie-nail mice. Development 136: 4223–4225.

33. Jakobsen LP, Knudsen MA, Lespinasse J, Garcia Ayuso C, Ramos C, et al. (2009) Pierre Robin sequence: a series of 117 consecutive cases. J Pediatr 139: 43: 155–159.

34. Brown KK, Reiss JA, Crow K, Ferguson HL, Kelly C, et al. (2010) Deletion of an enhancer near Dlx5 and Dlx6 in a family with hearing loss, craniofacial defects, and an inv(7)(q21.3q35). Hum Genet 127: 19–31.

35. Pouiras L, Yu M, Lesage-Pelletier C, Macdonald BB, Gagne JP, et al. (2010) An SNP in an ultracpensive regulatory element affects Dlx5/Dlx6 regulation in the forebrain. Development 137: 3089–3097.

36. Poupras P, M Le, J, P, H, C (2001) Pierre Robin sequence: a series of 117 consecutive cases. J Pediatr 139: 43: 155–159.

37. Samee N, Geoffroy V, Marty C, Schiltz C, Vieux-Rochas M, et al. (2009) The genetic basis of the Pierre Robin Sequence. Cleft Palate Craniofac J 46: 81–88.

38. Cohen MM, Jr. (1978) Syndromes with cleft lip and cleft palate. Cleft Palate J 15: 230–237.