Identification and characterization of a *dlx2b cis*-regulatory element both sufficient and necessary for correct transcription during zebrafish tooth development

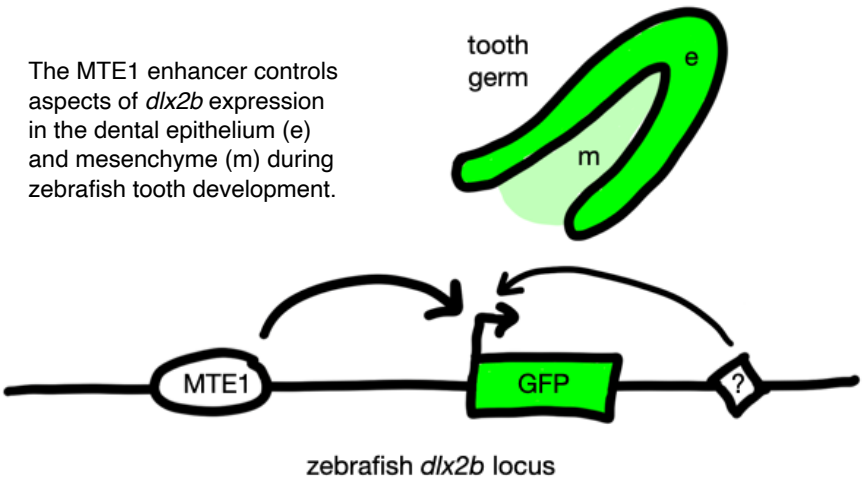
William R. Jackman, Yujin Moon, Drew R. Anderson, Audrey A. DeFusco, Vy M. Nguyen, Sarah Y. Liu, Elisabeth H. Carter, Hana E. Littleford, Elizabeth K. Richards, and Andrea L. Jowdry

Biology Department, Bowdoin College, Brunswick, Maine, wjackman@bowdoin.edu

Abstract

Much remains to be learned about how *cis*-regulatory elements such as enhancers function, especially during vertebrate organ development. To increase knowledge in this area, we have examined the *cis*-regulation of the transcription factor *dlx2b* during zebrafish larval tooth formation. We have created a GFP knock-in line that replicates *dlx2b* expression during tooth development and have also isolated a minimal enhancer of *dlx2b* (MTE1) sufficient for activating most of the tooth germ expression pattern. We have found that four evolutionarily conserved predicted transcription factor binding sites are required for the function of this minimal enhancer in both contexts. When the conserved sequences are mutated in a transgene it eliminates the activity of the enhancer and when they are deleted at the *dlx2b* locus it causes a dramatic alteration in the expression pattern. We hypothesize that disabling this enhancer at the *dlx2b* locus may be enabling other nearby *cis*-regulatory elements to take control of the promoter. These experiments reveal details of how *cis*-regulatory elements are working to control gene expression during organogenesis and highlight how much remains to be learned by empirical studies of gene regulation.

Graphical Abstract



Introduction

Techniques such as mRNA in situ hybridization and reporter transgenics have allowed the visualization of the often-complex spatiotemporal expression patterns of many genes during embryogenesis (e.g. Jensen, 2014; Kvon, 2015). However, the mechanisms behind how the transcription of these genes are regulated to produce complex developmental patterns has been much more difficult to ascertain. Now-classic "promoter-bashing" functional identification and analysis of *cis*-regulatory promoters and enhancers provided a foundation for the field (e.g. Goto et al., 1989), and large-scale bioinformatics work continues to be done at the genomic level to identify *cis*-regulatory elements of developmental genes and to determine their function (e.g. Davis et al., 2018). However, there remains a need for functional examination with individual elements to provide data upon which to test genome-wide predictions and to better understand the nature of *cis*-regulation in general, including enhancer/promoter interactions (Oudelaar et al., 2019) and predictions of sequences with specific functions (Grossman et al., 2017). Thus, new, direct tests of specific parts of developmental *cis*-regulatory elements are beneficial for the field as a whole.

Developing tooth germs are a good system in which to examine *cis*-regulatory function, as the expression of a large number of genes has been characterized in these primordia and the anatomy is relatively simple (Balic and Thesleff, 2015; Catón and Tucker, 2009). Additionally, gene expression and function during tooth development has been investigated in a wide range of vertebrate species ranging from mammals to sharks (Jernvall and Thesleff, 2012; Rasch et al., 2016), and these evolutionary comparisons have shown that, especially in early stages of tooth germ formation, vertebrate teeth develop remarkably similarly (Fraser et al., 2009). A large number of different classes of genes have been examined in tooth formation, including transcription factors like Dlx genes (Borday-Birraux et al., 2006; Zhao et al., 2000), cell signaling molecules like Shh (Seppala et al., 2017), and structural genes involved in cell behavior and differentiation such as Cadherins (Verstraeten et al., 2013). However, despite the relatively abundant information regarding gene expression, very little is yet known about the tooth-related *cis*-regulation of most of these genes.

Only a small number of *cis*-regulatory elements capable of driving specific expression patterns during tooth formation have been previously described. Perhaps the most well characterized are enhancers for mouse *Shh* (Sagai et al., 2009; Seo et al., 2018) and stickleback *bmp6* (Erickson et al., 2015; Cleves et al., 2018; Stepaniak et al., 2021), both of which have been isolated in reporter constructs as well as functionally examined with mutagenesis experiments that have identified potential transcription factor binding sites. Genomic regions near other tooth-related genes have been shown to drive tooth-specific reporter expression, such as from medaka *sp7* (Renn and Winkler, 2009) and human *RUNX2* (in zebrafish; Kague et al., 2012), but the exact position and makeup of enhancers in these regions has not been determined. Similarly, reporters capable of driving tooth-specific expression have been isolated from 4 kb of genomic sequences immediately 5' of mouse *Dlx2* (Thomas et al., 2000) and zebrafish *dlx2b* (Jackman and Stock, 2006), but again, specific enhancers in these regions have not been identified.

Thus, even in the few tooth-related enhancers identified, much remains to be learned about their makeup and function.

Dlx genes are homeodomain transcription factors involved with many aspects of development in vertebrates, especially craniofacial formation and the development of teeth (Jeong et al., 2008). As vertebrate tooth germs begin to form, the most significant tissues involved are the inner dental epithelium, which will eventually differentiate into ameloblasts and secrete the hypermineralized outer part of the tooth (enamel or enameloid, depending on the species), and the dental mesenchyme or papilla, a cranial neural crest derived tissue that will differentiate into odontoblasts and make the bone-like dentin inner layer (Huysseune et al., 1998; Peters and Balling, 1999). Dlx genes are expressed in overlapping spatiotemporal domains of these two developing tissues, with six orthologs activated during both mammalian and zebrafish tooth development (Borday-Birraux et al., 2006; Zhao et al., 2000). Zebrafish *dlx2b* is a particularly approachable gene in the context of tooth development, as it is expressed in a more tooth-specific way relative to neighboring tissues than other Dlx paralogs (Borday-Birraux et al., 2006). Zebrafish develop larval teeth in tissue lining the ventral, posterior part of the pharynx, just anterior to the esophagus (Huysseune et al., 1998). *dlx2b* is expressed in the inner dental epithelium during early morphogenesis stages of tooth development, as well as later in both the epithelium and mesenchyme tissues in differentiating ameloblasts and odontoblasts, respectively, and this transcriptional activity continues until tooth attachment (Borday-Birraux et al., 2006; Jackman et al., 2004). The *cis*-regulation of *dlx2b* has been studied in regard to enhancers located 3' to the gene which control aspects of its expression during brain development (MacDonald et al., 2010) and, as mentioned above, 4 kb of genomic sequence 5' to the gene are capable of driving reporter expression during tooth formation (Jackman and Stock, 2006), but little else is known about the function of this 5' region.

In this study, we have pared down this *dlx2b* 5' genomic region to a minimal functional enhancer of tooth germ cell reporter expression, and performed mutagenesis tests on predicted transcription factor binding sites of interest. Additionally, we have created a *dlx2b* GFP knock-in line that recapitulates the tooth germ expression of the gene, and tested the function of the identified minimal tooth enhancer by deleting its core sequences and observing the resulting change in expression. Together, these experiments provide new empirical data regarding the *cis*-regulatory control of tooth organogenesis.

Results

Identification of a dlx2b minimal tooth enhancer

Approximately two days prior to the appearance of a given mineralized zebrafish pharyngeal tooth (Fig. 1A), the cells of its tooth germ are already well-organized and transcribing a number of tooth-related genes, including *dlx2b* (Fig. 1B). Tooth germ *dlx2b* expression has been observed in at least the first four zebrafish tooth germs that form ($4V^1$, $3V^1$, $5V^1$, and $4V^2$; Borday-Birraux et al., 2006) and thus appears to be a robust marker of odontogenesis.

In order to better understand the *cis*-regulatory control of *dlx2b* during this process we undertook two approaches. One methodology was to create a reporter line driven by the endogenous *cis*-regulatory sequences at the *dlx2b* locus so we could have a way of assessing the full expression pattern of *dlx2b* more easily and at later stages than feasible using mRNA in situ hybridization. Our approach was to make a knock-in (KI) allele at the locus, termed $dlx2b^{KI}$, where the coding sequence for the green fluorescent protein (GFP) along with an attached promoter was inserted using CRISPR/Cas9 (see Methods, Fig. S1; Ota et al., 2016). In individuals heterozygous for this $dlx2b^{KI}$ allele, tooth-related GFP expression is observed in a pattern that appears to recapitulate those seen both by mRNA in situ hybridization and previous reporter analysis (Fig. 1C; Jackman & Stock 2006) and tooth germs and subsequent tooth development appears morphologically normal. We hypothesize that $dlx2b^{KI}$ is reporting the normal expression pattern of *dlx2b* in developing tooth germs because the overall expression pattern of *dlx2b* appears to be recapitulated (including early CNS expression (MacDonald et al., 2010); not shown), and this method has reproduced accurate expression patterns for other developmental genes (Kimura et al., 2015; Ota et al., 2016). Similarly to *dlx2b* mRNA (Borday-Birraux et al., 2006), GFP expression from the $dlx2b^{KI}$ allele is observed in the inner dental epithelium in early (morphogenesis) stages of tooth germ formation and at later (cytodifferentiation) stages is present in the dental mesenchyme as well. However, differently than reported for *dlx2b* mRNA, we also sometimes observed a small amount of dental mesenchyme expression in morphogenesis stage tooth germs. This pattern appears to be the same for at least the first four teeth to form ($4V^1$, $3V^1$, $5V^1$, and $4V^2$). In the figures presented here, we have chosen to show 3 day post fertilization (dfp, 72-78 hpf) images as representatives, as these often display $3V^1$ in a relatively early/morphogenesis stage in the same focal plane as $4V^1$, which is in a later/cytodifferentiation stage, and thus captures most all of the variation in expression that we observed in tooth germs at various stages. We noticed no tooth-specific differences in expression (e.g. $4V^1$ vs. $3V^1$, $5V^1$, or $4V^2$) for any of the results reported in this study.

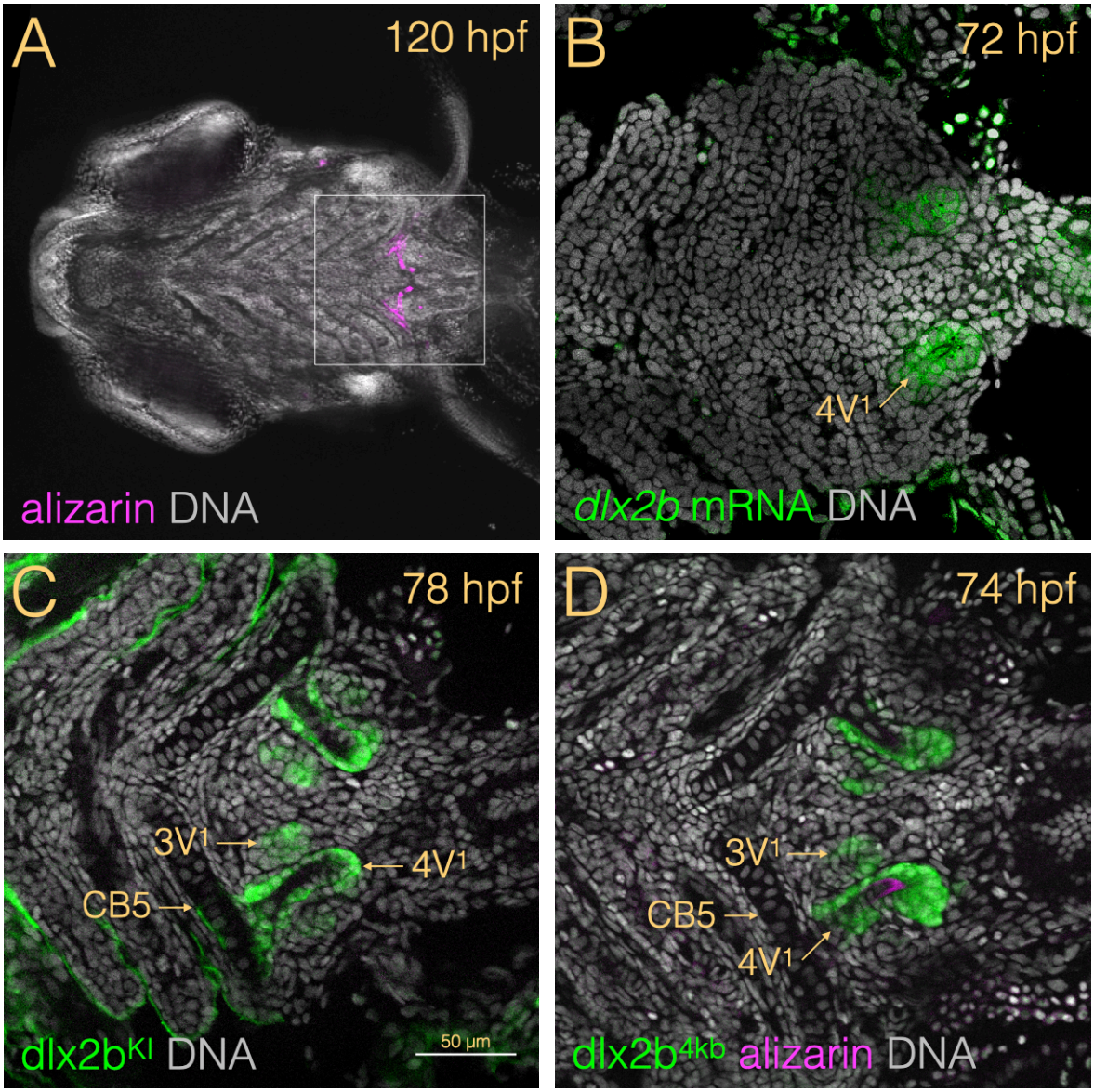


Fig. 1. Endogenous and reporter expression of *dlx2b* in developing zebrafish tooth germs. (A) Ventral view with anterior to the left of a 120 hpf larval head labeled with alizarin red S (magenta) to show mineralizing tissues and DAPI (white) to stain nuclei and thus provide a cellular context. The tooth-forming region in the posterior pharynx is indicated (box), highlighting the approximate zoom and cropping of the subsequent panels, in which the right-side (bottom) tooth germs are labeled. (B) Optical section of *dlx2b* mRNA in situ hybridization (green) at 72 hpf with the $4V^1$ tooth germ indicated. (C) Expression from the $dlx2b^{KI}$ reporter line showing the $3V^1$ tooth germ at a morphogenesis stage and $4V^1$ at cytodifferentiation. (D) GFP expression in the $dlx2b^{4kb}$ reporter, additionally stained with alizarin red S to show developing tooth mineralization. CB5 = fifth ceratobranchial cartilage (onto which all of the teeth eventually attach).

The second approach to understand *dlx2b cis*-regulation involved isolating a minimal enhancer from the *dlx2b* locus capable of driving a normal tooth germ expression pattern. As mentioned above, it was previously determined that the 4kb of genomic sequence immediately 5' of the *dlx2b* transcription start site, when used as a promoter for GFP reporter expression in a transgene allele here referred to as dlx2b^{4kb}, is sufficient to drive expression in a *dlx2b*-like pattern (Fig. 1D; Jackman and Stock 2006). Using a promoter-bashing type approach, we tested a series of reporter constructs, first trimming sequences from the 5' end of the 4kb region, and then, using non-*dlx2b* minimal promoters, from the 3' end (Table S1). The result was the isolation of a 213 bp region located 817-604 bp 5' of the *dlx2b* translation start site which was sufficient to drive GFP expression in developing tooth germs, a sequence which we designate as the *dlx2b* minimal tooth enhancer 1 (MTE1). This region contains sequence homology with that of other vertebrate genomes, including humans, and four regions of very high conservation that may represent conserved transcription factor binding sites (Fig. 2). This homology had been noted previously (Jackman and Stock 2006) and, in retrospect, would have been a good phylogenetic-footprinting type guide to have followed in isolating this enhancer.

Fig. 2. Location and sequence of the *dlx2b* minimal tooth enhancer MTE1. (A) Diagram of the 5' end of the zebrafish *dlx2b* gene, including the 5' UTR and 817 bp of upstream non-coding sequence. The location of the identified 213 bp minimal tooth enhancer MTE1 is indicated. The black bars below the gene schematic represent regions of homology with the *X. tropicalis* and human genomes (UCSC Genome Browser Zv9/danRer7). (B) Sequence of MTE1. Underlines indicate regions with identifiable homology with other vertebrate genomes and bold-case highlights stretches of six or more base pairs that are 100% conserved between zebrafish and tetrapods. Gray boxes indicate the extent of predicted (possible) transcription factor binding sites.

The GFP reporter transgene termed dlx2b^{MTE1} was then established as a stable line. When its expression is compared with the dlx2b^{KI} (Fig. 3A) and dlx2b^{4kb} (Fig. 3B) lines, the GFP expression pattern of dlx2b^{MTE1} appears extremely similar (Fig. 3C), with strong early expression in the inner dental epithelium and later expression in the dental mesenchyme. When a smaller region of MTE1 spanning 744-674 bp upstream of the *dlx2b* translation start site was tested in transient assays, no reporter expression was seen (n=13; Fig. 3D), suggesting that there are sequences important for the function of the enhancer on the edges of the 213 bp MTE1, and not just within the highly conserved central region.

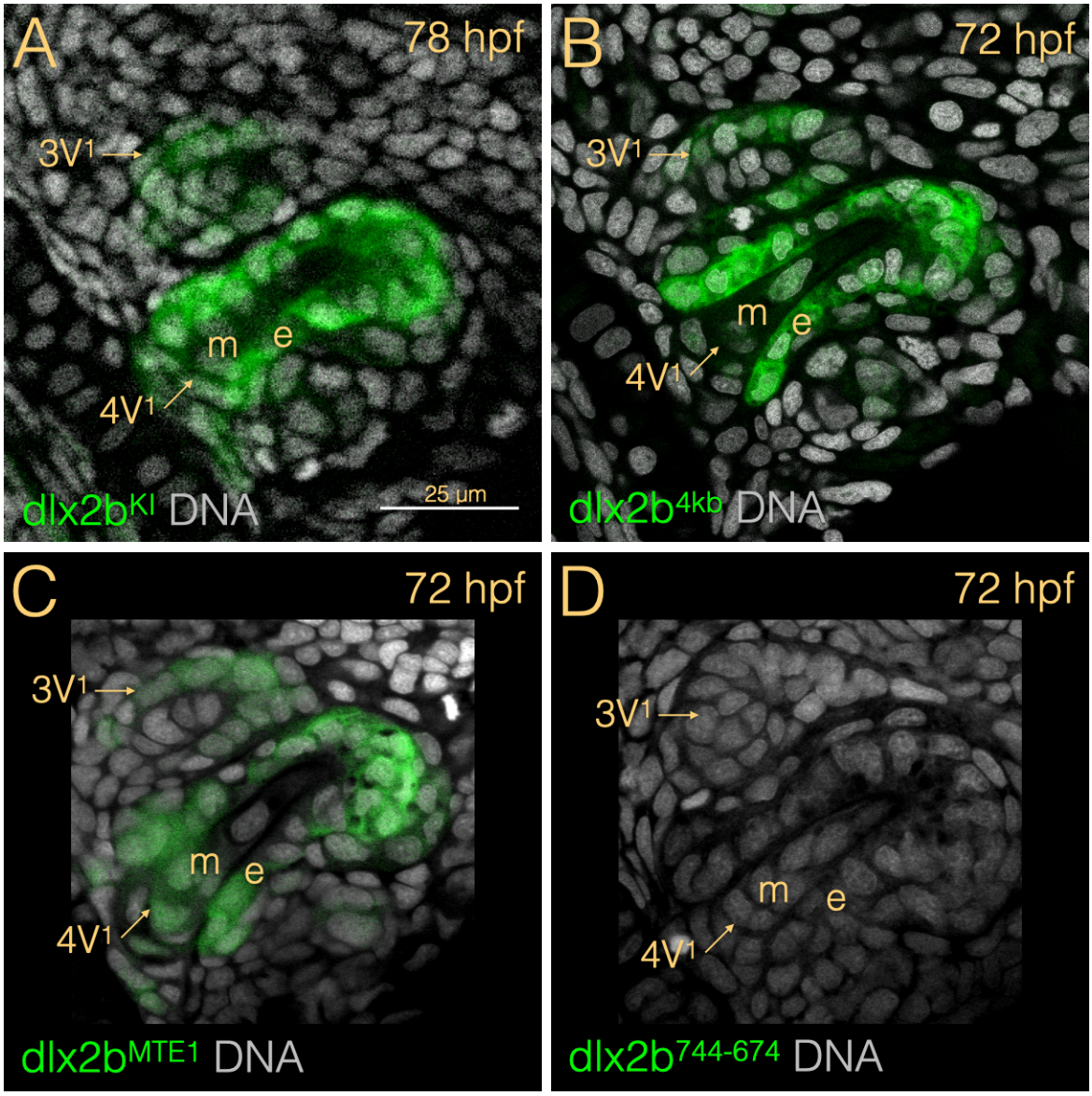


Fig. 3. Tooth germ reporter expression. Ventral view close-up of one side of the tooth forming region with the tooth germs 4V^1 and 3V^1 indicated. The dlx2b^{KI} reporter line (A), the dlx2b^{4kb} line (B), and dlx2b^{MTE1} (C), all exhibit GFP expression primarily in the inner dental epithelium (e) but also somewhat in the dental mesenchyme (m), whereas no GFP expression was observed in the dlx2b$^{744-674}$ truncated reporter (D).

Characterization of the function of MTE1

We next wanted to test if several predicted transcription factor (pTF) binding sites within the MTE1 region were necessary for tooth germ enhancer activity. Our approach was to create a series of reporter constructs with the pTF binding sites altered so as to maximally disrupt their consensus sequences and test these mutated versions in transient injection assays (Fig. S2; Table S1). We found that none of six types of pTF sites tested were required by themselves for tooth germ expression. However, when all four of the most highly conserved pTF sites were mutated (Dlx, FoxA, Cebp, and Ap1) in a construct designated dlx2b^{MTE1mDFCA}, tooth germ reporter expression was eliminated (n=30; Fig. 4). Although we only tested a relatively small number of the many possible combinations of pTF mutations that would be possible to examine, these data suggest that the activity of the MTE1 enhancer is resistant to complete loss of function except by significant sequence alteration.

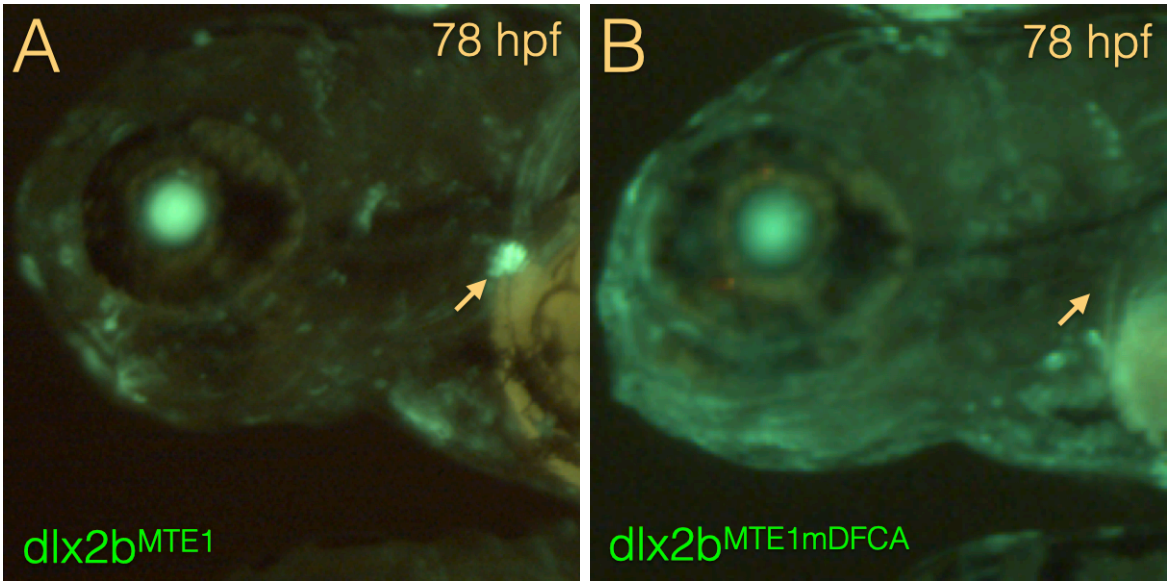


Fig. 4. Conserved transcription factor binding sites are required for MTE1 function. Zebrafish larvae injected transiently with either the dlx2b^{MTE1} GFP reporter construct (A) or the dlx2b^{MTE1mDFCA} construct with the predicted Dlx, FoxA, Cebp, and Ap1 binding sites all mutated (B). Arrows indicate the location of the tooth-forming region.

While none of the four highly conserved pTF sites appeared to be required by themselves for MTE1 enhancer activity, this analysis led us to investigate further a zebrafish Cebp gene, as these proteins have previously been implicated as key activators of amelogenin during ameloblast maturation in developing mammalian tooth germs (Huang et al., 2013; Zhou and Snead, 2000). To that end we examined tooth germ mRNA expression of zebrafish *cebpa*, as well as GFP reporter expression in a cebpa^{KI} line constructed similarly to dlx2b^{KI}. We identified tooth-germ-specific expression with both methods, localized mostly to the inner dental epithelium in a domain near the tooth base, overlapping a part of the *dlx2b* expression pattern (Fig. S3).

Next we wanted to test the necessity of the *dlx2b* MTE1 region for regulating normal tooth gene expression. To do this in a feasible way, rather than targeting an unaltered *dlx2b* locus, we used CRISPR/Cas9 to create a deletion within the MTE1 in the dlx2b^{KI} allele so that we would be able to immediately gauge any resulting changes to expression via GFP. Using two guide RNAs, we deleted 88 bp within the *dlx2b* MTE1, generating an allele termed dlx2b$^{KI\Delta MTE1}$ (Fig. S4). This deletion spans from 760 to 673 5' of the *dlx2b* translation start site, and eliminates all four of the highly conserved pTF biding sites Dlx, FoxA, Cebp, and Ap1. The mutation had a major effect on the tooth germ GFP expression pattern, seemingly reversing the pattern relative to the dental epithelium and mesenchyme (Fig. 5). Whereas normally dlx2b^{KI} GFP expression is observed during morphogenesis and cytodifferentiation stages mostly in the inner dental epithelium with lower-appearing levels of expression in the dental mesenchyme (Fig. 5A), in the dlx2b$^{KI\Delta MTE1}$ allele, GFP is strongly expressed in the dental mesenchyme (Fig. 5B) with weaker, occasionally observed expression in the distal part of the inner dental epithelium (not shown). Thus, it appears likely that in the *cis*-regulatory context of the genomic locus surrounding *dlx2b*, the MTE1 has an important function in directing correct tissue-specific expression within developing tooth germs.

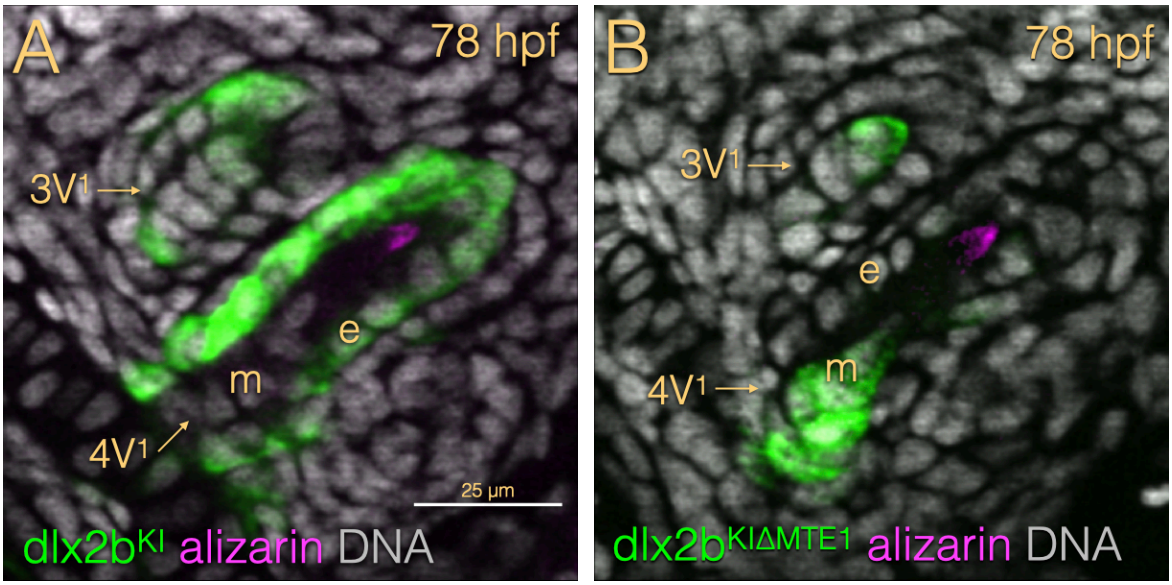


Fig. 5. MTE1 is required for proper tissue-specific expression in developing tooth germs. (A) In dlx2b^{KI} at 78 hpf, GFP expression appears much stronger in the inner dental epithelium (e) than in the dental mesenchyme (m) in both 4V^1 and 3V^1 tooth germs. (B) In dlx2b$^{KI\Delta MTE1}$ at the same stage, the expression pattern appears reversed, with strong GFP signal in the dental mesenchyme and little to none in the epithelium.

Discussion

We have shown that the *dlx2b* minimal enhancer MTE1 is sufficient to recapitulate most or all of the normal *dlx2b* expression pattern in developing tooth germs, especially that of the inner dental epithelium (summarized in Fig. 6). MTE1 contains four regions of conspicuously high sequence conservation that are predicted to represent important transcription factor binding sites (pTFs), but this region by itself, as shown in the dlx2b$^{744\text{-}674}$ reporter experiment, is not capable of driving tooth-specific expression. Additionally, no one of these conserved sites seems necessary for the overall activity of the enhancer, suggesting some redundancy in their function. More work will be required for firm answers regarding what each pTF site might specifically control with regard to *dlx2b* transcriptional activity, but it is nevertheless interesting to consider a few ideas of how they might function. For example, FoxA proteins have been shown to sometimes act as pioneer transcription factors, opening up chromatin regions for further transcription factor binding (Iwafuchi-Doi and Zaret, 2016), and are expressed in the pharyngeal tooth-forming region of zebrafish (Piotrowski and Nüsslein-Volhard, 2000), consistent with the idea that the FoxA pTF site in MTE1 might be important in chromatin modification. Another example is that of Cebp proteins, which have been studied specifically with regard to mouse tooth development and are known to be important in ameloblast differentiation and activation of amelogenin transcription (Zhou and Snead, 2000; Xu et al., 2007). We observed *cebpa* expression in a region of the zebrafish tooth germ inner dental epithelium consistent with these roles, and in a pattern which seems to be a subset of the full *dlx2b* epithelial expression domain (Fig. S3), suggesting that *cebpa* could possibly be be an activator of *dlx2b* in these cells acting through the MTE1 Cebp pTF site. Even though the function of these conserved sequences remains unknown, the activity of MTE1 was eliminated when all four sites were mutated (Fig. 6), emphasizing that these evolutionarily conserved sequences are important and that they may work together, even in isolation from the *dlx2b* locus.

However, it is important also to consider how the MTE1 enhancer might be working in its endogenous locus and how this could be different than how it works in isolation. In this context, we have observed that in both the randomly inserted transgenes dlx2b^{4kb} and dlx2b^{MTE1}, GFP reporter expression appears more restricted to the dental epithelium, especially in early morphogenesis stages (Figs. 3B, 3C), when compared with the GFP knock-in allele dlx2b^{KI} (Figs. 3A, 5A; summarized in Fig. 6). If this observation is accurate, and dlx2b^{KI} indeed does more precisely represent the real expression pattern of *dlx2b* (mRNA in situ hybridization isn't much help in determining what is real, as we find it to be less sensitive than GFP reporters, it is difficult to do at later stages, and the visualization products tend to diffuse), it suggests that there is a small amount of dental mesenchyme transcriptional activity that is not controlled by MTE1, and is instead activated by something else. One explanation for this could be the presence of one or more enhancers located at a different location at the *dlx2b* locus that can drive dental mesenchyme transcriptional activation. This idea is consistent with the results from the dlx2b$^{KI\Delta MTE1}$ allele where the evolutionarily conserved core of MTE1 is deleted and most expression is lost in the dental epithelium but dental mesenchyme expression appears

increased (Fig. 5B). Given these observations, we hypothesize that in a normal situation, MTE1 is interacting with other enhancers and the *dlx2b* promoter, and perhaps due to its proximity to the promoter largely wins, resulting in an expression pattern in the dental epithelium that mirrors what the MTE1 reporter generates in genomic isolation. However, when MTE1 is compromised, other enhancers at the *dlx2b* locus more easily interact with the promoter and their expression patterns are increased (Fig. 6). This hypothesis is inspired by a number of studies on enhancer/promoter interactions (e.g. Bateman et al., 2021; Oudelaar et al., 2019) as well as by the regulation of *bmp6* expression during stickleback tooth development (Cleves et al., 2018). Stickleback *bmp6* contains a mostly epithelial-driving enhancer located immediately 5' to the promoter and a more dental mesenchyme-driving enhancer located in a relatively 3' position in an intron, which is interestingly similar to what we have found thus far with *dlx2b*.

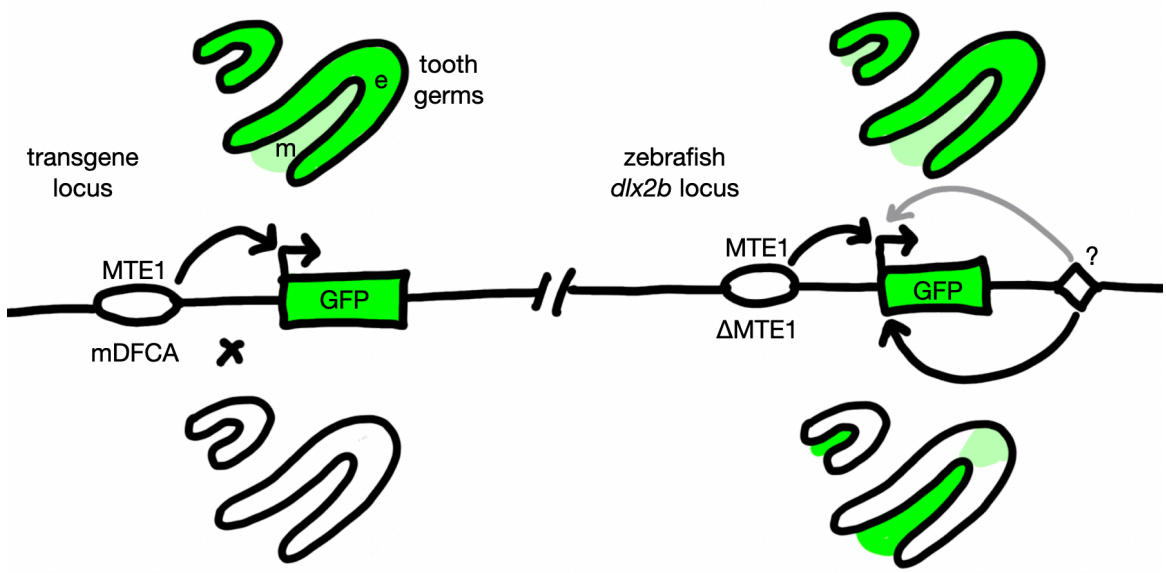


Fig. 6. Model of MTE1 enhancer function. The left side represents the Tol2-based transgene experiments at a remote locus and the right side the knock-in experiments at the *dlx2b* locus. The top of the diagram depicts the normal MTE1 sequence and the bottom the mutated versions. MTE1 is sufficient to drive a mostly-epithelial tooth germ expression pattern even at a remote locus (left, top) but when mutated at a remote locus there is nothing to compensate and all tooth germ expression is lost (left, bottom). In contrast, at the *dlx2b* locus, MTE1 may be a primary driver of epithelial tooth germ expression (right, top), but when mutated, one or more other *cis*-regulatory elements (diamond) maintain expression, but in a more mesenchymal pattern (right, bottom).

As a caveat, our model relies on an assumption that the dlx2b^{KI} allele with its introduced promoter is behaving like a normal *dlx2b* gene, and thus more work will need to be done both to specifically identify other enhancers at the *dlx2b* locus and to understand how the *cis*-regulatory elements may be together interacting with the endogenous *dlx2b* promoter to generate its complete expression pattern during tooth formation. Some of this information may be acquired along with genome-wide modeling of *cis*-regulatory function based on chromatin structure and other features (e.g. ENCODE) but there is also still much that can be learned from targeted functional studies of *cis*-regulation such as this one.

Methods

Animal husbandry and anatomy

All experimentation with the zebrafish (*Danio rerio*, Hamilton 1822) in this study followed protocols approved by the Institutional Animal Care and Use Committee at Bowdoin College. The zebrafish used were from an in-house strain, derived originally from a mixture of the Tü and AB lines. Embryos were raised in 30% Danieu's embryo medium, with the addition of 0.002% methylene blue to reduce fungal growth, and kept at approximately 28.5˚C. Embryonic stages are reported in hours or days post-fertilization, or using the nomenclature of Kimmel et al. (1995).

Individual teeth are labeled following the convention of Van der heyden & Huysseune 2000, with the first-forming tooth designated $4V^1$, and the subsequently forming teeth as $3V^1$, $5V^1$, and $4V^2$ (the replacement of $4V^1$). These teeth develop in a reliable temporal progression as well as in characteristic locations and were thus possible to unambiguously identify. Names for tooth germ developmental stages are as in Huysseune et al. (1998): initiation (not shown in this study), morphogenesis, cytodifferentiation, and attachment (also not shown).

DNA constructs and transgenic lines

Name	Official Designation	Insertion	Enhancer(s)	Promoter	Reporter
dlx2b^{4kb}	cs1Tg with construct Tg(dlx2b:EGFP)	unknown via Tol2	dlx2b 5' 4 kb	dlx2b endogenous	GFP
dlx2b^{MTE1}	bo1Tg with construct Tg(dlx2b-actb2:EGFP,myl7:GFP)	unknown via Tol2	dlx2b MTE1	ß-actin	GFP
dlx2b^{KI}	bo2Tg with construct Tg3(hsp70l:EGFP)	dlx2b locus	dlx2b endogenous	hsp70	GFP
dlx2b$^{KI\Delta MTE1}$	bo3Tg with construct Tg3(hsp70l:EGFP)	dlx2b locus	dlx2b endogenous	hsp70	GFP
cebpa^{KI}	bo4Tg with construct Tg3(hsp70l:EGFP)	cebpa locus	cebpa 5' endogenous	hsp70	GFP

Table 1. Transgenic lines used in this study.

The GFP knock-in alleles dlx2b^{KI} and cebpa^{KI} were generated via CRISPR/Cas9 using a combination of the methods described in Ota et al., 2016 and Burger et al., 2016. The logic of this method is to insert a plasmid at the target locus containing the Hsp70 promoter situated properly adjacent to the coding sequence of GFP so that a precise integration event is not necessary but nearby enhancers will still be able to interact with the introduced promoter (Ota et al., 2016). Briefly, zebrafish embryos were injected into the blastomeres at early 1-cell or 2-cell stages with about 1 nl of a ribonucleoprotein (RNP) mixture of high-fidelity Cas9 protein (Integrated DNA Technologies, IDT), a guide RNA targeting the desired genomic locus (GCCAAGGCTATCCAGAACAG for

dlx2b, custom synthesized by IDT), a plasmid containing the Hsp70 promoter and GFP coding sequence (Mbait-hs-eGFP; Ota et al., 2016), a second guide RNA to linearize the plasmid, 300 mM KCl to promote RNA solubility (Burger et al., 2016), and a trace amount of phenol red to facilitate visibility of the solution during injection. As they developed, injected (F0) embryos were scored for GFP expression patterns that matched previously reported mRNA expression, and individuals with correct-looking patterns and relatively low mosaicism were raised as potential founders. Approximately 5% of injected F0 fish showed promising GFP expression and of those, about 25% had germline transmission. The dlx2b^{KI} allele is officially designated in the Zebrafish Information Network (ZFIN) as genomic feature bo2Tg with construct Tg3(hsp70l:EGFP). The complete insert region was PCR amplified using Q5 Hi Fidelity DNA polymerase (NEB) and the amplicon sequenced at high coverage using Illumina MiSeq in the Complete Amplicon Next-Generation Sequencing service from the MGH CCIB DNA Core (Fig. S1).

The dlx2b$^{KI\Delta MTE1}$ allele was made by injecting dlx2b^{KI} embryos with two additional sgRNAs targeting either side of the MTE1 highly conserved region (CTGGTTCCGCGCTTTATCCC and CGCCGTCTGTGATTAGTCAG) and scoring F0 fish for reduced levels of GFP expression (Fig. S4). Once a stable line was isolated, the deletion region was PCR amplified with Q5 polymerase and characterized via Sanger sequencing. This dlx2b$^{KI\Delta MTE1}$ allele is designated on ZFIN as genomic feature bo3Tg with construct Tg3(hsp70l:EGFP).

The cebpa^{KI} allele is designated on ZFIN as genomic feature bo4Tg with construct Tg3(hsp70l:EGFP) and was created using GGCGGGTTTTAGATACTCCA as a guide RNA. This allele has not been sequenced, but the plasmid insertion into the *cebpa* locus has been verified by diagnostic PCR. The plasmid is in the reverse orientation (hsp70 promoter and GFP towards the 3' end of the gene) and there is deletion of unknown size 3' of the plasmid insertion.

The allele we refer to in this study as dlx2b^{4kb} is designated as genomic feature cs1Tg with construct Tg(dlx2b:EGFP). It consists of 4 kb of genomic sequence, 5' to the translation start site of the *dlx2b* gene, including the endogenous promoter for *dlx2b*. DNA reporter constructs representing subsets of this original 4 kb region were created during three different time periods of the project, each period using somewhat different methods. The parts of each construct are summarized in Table S1. In the first period, constructs were assembled using the Gateway Tol2kit **(**Kwan et al., 2007), employing the pDestTol2CG2 destination vector, which flanked the insert regions with Tol2 transposon inverted repeats for efficient genomic integration (Kawakami, 2007). The insert region of these plasmids consisted of a fragment of the *dlx2b* 5' genomic *cis*-regulatory region that sometimes included the *dlx2b* endogenous promoter, 500 bp of the ß-actin promoter (Higashijima et al., 1997) if the *dlx2b* promoter was not included, and EGFP coding sequence with a SV40 polyadenylation signal. In the second period of reporter construction, identically arranged insert regions were created (more quickly and reliably) using NEBuilder (New England Biolabs) and PCR amplified using the Q5 proofreading polymerase (New England Biolabs) without cloning into a plasmid. Directly testing PCR products in this manner was rapid but injection of linear DNA produced a high frequency

of deformed embryos that reduced the efficiency of the method. Thus in the most recent period, a final plasmid construct was made using NEBuilder with a pUC19 plasmid backbone, this time using the hsp70 promoter instead of ß-actin, which decreased non-specific GFP expression in injected embryos (Erickson et al., 2015). All constructs, including PCR products, were verified by Sanger sequencing.

To test the function of predicted transcription factor binding sites, mutations were designed to maximally disrupt the predicted transcription factor binding, while at the same time altering a small number of actual nucleotides to make it more likely that we were only disrupting one binding site. To achieve this balance, we examined the binding matrixes for each transcription factor (Cartharius et al., 2005) and changed all nucleotide positions that were invariant or highly probably associated with a functional binding site.

Individual F0 embryos injected with these reporter constructs displayed variable expression in tooth germs, if any expression was seen at all, as is expected from mosaic integration of DNA injected into zebrafish embryos (Ni et al., 2016). Despite this mosaicism, because of the lack of potentially confounding nearby expression driven by the *dlx2b cis*-regulatory regions tested, it was unambiguous when a particular construct was driving tooth-related expression, even if it was present in only a subset of cells in the developing tooth germs. Nevertheless, for certain constructs we wanted to observe in detail the non-mosaic pattern of a stably integrated reporter line, and thus raised these fish to at least the F2 generation (Table S1). Due to housing space limitations, most of these stable lines were discarded after F2 testing, but we have retained the line created from construct #9 (Table S1), which we refer to here as dlx2b^{MTE1}, and is officially designated in ZFIN as genomic feature bo1Tg with construct Tg(dlx2b-actb2:EGFP,myl7:GFP).

Histology

For mRNA in situ hybridization, riboprobes were created from custom-synthesized DNA templates matching parts of the genes of interest (gBlocks from IDT). For *cebpa*, the sequence used to make the riboprobe started 685 bp from the translation start site and extended 1000 bp into the 3' UTR (GenBank:BC056548). This was used as template for PCR, in which a T3 RNA polymerase binding site for transcription of the antisense probe was added as part of the reverse primer. For *dlx2b*, the gene-specific sequence consisted of the first 474 bp of the 3' UTR (GenBank:BC134899) and the T3 site was synthesized directly into the gBlock template. Synthesis and purification of probes was performed as previously described (Thisse and Thisse, 2008). Hybridization and developing steps were performed as described in Talbot et al. (2010), except that digoxigenin-labeled riboprobe was detected with an anti-digoxigenin antibody conjugated to alkaline phosphatase and visualized with a Fast Red reaction (Lauter et al., 2011), and that DAPI or Sytox Green (depending on the microscope to be used later) was added to the overnight antibody incubation in order to simultaneously observe cell nuclei for orientation purposes.

GFP antibody labeling and alizarin red S staining of mineralized teeth was performed as in Yu et al. (2015). We found that we could combine these two methods in a single

specimen (e.g. Fig. 1D; Fig. 5) if the alizarin staining was performed after the antibody label and visualized relatively quickly (within 1-2 days).

Microscopy & Image Processing

Photographs of head and tooth GFP expression in living zebrafish embryos were taken with a Leica MZ16F stereoscope with a DCF300FX camera. For fixed tissue, to get a clear view of the tooth-forming region, the heart and yolk of embryos or larvae were manually removed with insect pins and the specimens were mounted ventral side up in glycerol under a small coverslip, as perviously described (Yu et al., 2015). Z-stacks for 3D image analysis were taken using either a Zeiss 510 Meta laser scanning confocal microscope, a Leica TCS SP8 confocal microscope, or a Zeiss Axio Imager M2 with an Apotome 2 structured illumination attachment. Brightness and contrast levels were adjusted uniformly across each image using FIJI/ImageJ (Schindelin et al., 2012). For most images, a single Z-slice is shown, but for certain specimens (Fig. 1A, Fig. 5) it was desirable to show a thicker Z representation. For these, the stacks were rendered in 3D and visualized with FluoRender (Wan et al., 2017). Colors for fluorescence images were selected to facilitate visibility for diverse vision types (Wong 2011). Final figures were assembled using Keynote (Apple Inc.).

Bioinformatics

Visualization of evolutionary sequence conservation at the *dlx2b* genomic locus was done using the Jul. 2010 (Zv9/danRer7) zebrafish genome assembly in the UCSC Genome Browser (Kent et al., 2002) and guide RNA targets were chosen with help from the CRISPR/Cas9 Sp. Pyog. target sites tracks. Predicted transcription factor binding sites were located using a combination of PROMO (Messeguer et al., 2002) and MatInspector (Cartharius et al., 2005). DNA sequence analysis and Figs. S1 and S4 were done with Geneious R11 and Geneious Prime 2022.0.1.

End Matter

Author Contributions & Notes

W.R.J. designed research; W.R.J., Y.M., D.R.A., A.A.DF., V.M.N., S.Y.L., E.H.C., H.E.L., E.K.R., and A.L.J. performed research, collected, and analyzed data; W.R.J. wrote the paper. The authors declare no conflict of interest.

Acknowledgements

We thank Atsuo Kawahara for graciously sending us the GFP knock-in plasmid.

Funding

The work was supported by the National Institute of Dental and Craniofacial Research (NIH R15DE023667) and the Maine IDeA Network of Biomedical Research Excellence (NIH P20GM0103423).

References

Balic, A., Thesleff, I., 2015. Tissue Interactions Regulating Tooth Development and Renewal, in: Current Topics in Developmental Biology. Elsevier, pp. 157–186. https://doi.org/10.1016/bs.ctdb.2015.07.006

Bateman, J.R., Johnson, J.E., Locke, M.N., 2012. Comparing Enhancer Action in *Cis* and in *Trans*. Genetics 191, 1143–1155. https://doi.org/10.1534/genetics.112.140954

Borday-Birraux, V., heyden, C., Debiais-Thibaud, M., Verreijdt, L., Stock, D.W., Huysseune, A., Sire, J.-Y., 2006. Expression of Dlx genes during the development of the zebrafish pharyngeal dentition: evolutionary implications. Evol Dev 8, 130–141. https://doi.org/10.1111/j.1525-142X.2006.00084.x

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L.M., Catena, R., Jinek, M., Robinson, M.D., Mosimann, C., 2016. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. Development dev.134809. https://doi.org/10.1242/dev.134809

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., Werner, T., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933–2942. https://doi.org/10.1093/bioinformatics/bti473

Catón, J., Tucker, A.S., 2009. Current knowledge of tooth development: patterning and mineralization of the murine dentition. Journal of Anatomy 214, 502–515. https://doi.org/10.1111/j.1469-7580.2008.01014.x

Cleves, P.A., Hart, J.C., Agoglia, R.M., Jimenez, M.T., Erickson, P.A., Gai, L., Miller, C.T., 2018. An intronic enhancer of Bmp6 underlies evolved tooth gain in sticklebacks. PLoS Genet 14, e1007449. https://doi.org/10.1371/journal.pgen.1007449

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., Onate, K.C., Graham, K., Miyasato, S.R., Dreszer, T.R., Strattan, J.S., Jolanki, O., Tanaka, F.Y., Cherry, J.M., 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Research 46, D794–D801. https://doi.org/10.1093/nar/gkx1081

Erickson, P.A., Cleves, P.A., Ellis, N.A., Schwalbach, K.T., Hart, J.C., Miller, C.T., 2015. A 190 base pair, TGF-β responsive tooth and fin enhancer is required for stickleback Bmp6 expression. Developmental Biology 401, 310–323. https://doi.org/10.1016/j.ydbio.2015.02.006

Fraser, G.J., Hulsey, C.D., Bloomquist, R.F., Uyesugi, K., Manley, N.R., Streelman, J.T., 2009. An Ancient Gene Network Is Co-opted for Teeth on Old and New Jaws. PLoS Biol 7, e1000031. https://doi.org/10.1371/journal.pbio.1000031

Goto, T., Macdonald, P., Maniatis, T., 1989. Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell 57, 413–422. https://doi.org/10.1016/0092-8674(89)90916-1

Grossman, S.R., Zhang, X., Wang, L., Engreitz, J., Melnikov, A., Rogov, P., Tewhey, R., Isakova, A., Deplancke, B., Bernstein, B.E., Mikkelsen, T.S., Lander, E.S., 2017. Systematic dissection of genomic features determining transcription factor binding and enhancer function. Proc Natl Acad Sci USA 114, E1291–E1300. https://doi.org/10.1073/pnas.1621150114

Higashijima, S., Okamoto, H., Ueno, N., Hotta, Y., Eguchi, G., 1997. High-Frequency Generation of Transgenic Zebrafish Which Reliably Express GFP in Whole Muscles or the Whole Body by Using Promoters of Zebrafish Origin. Developmental Biology 192, 289–299. https://doi.org/10.1006/dbio.1997.8779

Huang, Z., Newcomb, C.J., Zhou, Y., Lei, Y.P., Bringas, P., Stupp, S.I., Snead, M.L., 2013. The role of bioactive nanofibers in enamel regeneration mediated through integrin signals acting upon C/EBPα and c-Jun. Biomaterials 34, 3303–3314. https://doi.org/10.1016/j.biomaterials.2013.01.054

Huysseune, A., Van der heyden, C., Sire, J.-Y., 1998. Early development of the zebrafish (Danio rerio) pharyngeal dentition (Teleostei, Cyprinidae). Anatomy and Embryology 198, 289–305. https://doi.org/10.1007/s004290050185

Iwafuchi-Doi, M., Zaret, K.S., 2016. Cell fate control by pioneer transcription factors. Development 143, 1833–1837. https://doi.org/10.1242/dev.133900

Jackman, W.R., Draper, B.W., Stock, D.W., 2004. Fgf signaling is required for zebrafish tooth development. Developmental Biology 274, 139–157. https://doi.org/10.1016/j.ydbio.2004.07.003

Jackman, W.R., Stock, D.W., 2006. Transgenic analysis of Dlx regulation in fish tooth development reveals evolutionary retention of enhancer function despite organ loss. Proceedings of the National Academy of Sciences 103, 19390–19395. https://doi.org/10.1073/pnas.0609575103

Jensen, E., 2014. Technical Review: In Situ Hybridization: AR Insights. Anat. Rec. 297, 1349–1353. https://doi.org/10.1002/ar.22944

Jeong, J., Li, X., McEvilly, R.J., Rosenfeld, M.G., Lufkin, T., Rubenstein, J.L.R., 2008. Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. Development 135, 2905–2916. https://doi.org/10.1242/dev.019778

Jernvall, J., Thesleff, I., 2012. Tooth shape formation and tooth renewal: evolving with the same signals. Development 139, 3487–3497. https://doi.org/10.1242/dev.085084

Kague, E., Gallagher, M., Burke, S., Parsons, M., Franz-Odendaal, T., Fisher, S., 2012. Skeletogenic Fate of Zebrafish Cranial and Trunk Neural Crest. PLoS ONE 7, e47394. https://doi.org/10.1371/journal.pone.0047394

Kawakami, K., 2007. Tol2: a versatile gene transfer vector in vertebrates. Genome Biol 8, S7. https://doi.org/10.1186/gb-2007-8-s1-s7

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The Human Genome Browser at UCSC. Genome Research 12, 996–1006.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., Schilling, T.F., 1995. Stages of embryonic development of the zebrafish. Dev. Dyn. 203, 253–310. https://doi.org/10.1002/aja.1002030302

Kimura, Y., Hisano, Y., Kawahara, A., Higashijima, S., 2015. Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. Sci Rep 4, 6545. https://doi.org/10.1038/srep06545

Kvon, E.Z., 2015. Using transgenic reporter assays to functionally characterize enhancers in animals. Genomics 106, 185–192. https://doi.org/10.1016/j.ygeno.2015.06.007

Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant, J.M., Yost, H.J., Kanki, J.P., Chien, C.-B., 2007. The Tol2kit: A multisite gateway-based construction kit forTol2 transposon transgenesis constructs. Dev. Dyn. 236, 3088–3099. https://doi.org/10.1002/dvdy.21343

Lauter, G., Söll, I., Hauptmann, G., 2011. Two-color fluorescent in situ hybridization in the embryonic zebrafish brain using differential detection systems. BMC Dev Biol 11, 43. https://doi.org/10.1186/1471-213X-11-43

MacDonald, R.B., Debiais-Thibaud, M., Talbot, J.C., Ekker, M., 2010. The relationship between dlx and gad1 expression indicates highly conserved genetic pathways in the zebrafish forebrain. Dev. Dyn. 239, 2298–2306. https://doi.org/10.1002/dvdy.22365

Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J., Alba, M.M., 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics 18, 333–334. https://doi.org/10.1093/bioinformatics/18.2.333

Ni, J., Wangensteen, K.J., Nelsen, D., Balciunas, D., Skuster, K.J., Urban, M.D., Ekker, S.C., 2016. Active recombinant Tol2 transposase for gene transfer and gene discovery applications. Mobile DNA 7, 6. https://doi.org/10.1186/s13100-016-0062-z

Ota, S., Taimatsu, K., Yanagi, K., Namiki, T., Ohga, R., Higashijima, S., Kawahara, A., 2016. Functional visualization and disruption of targeted genes using CRISPR/Cas9-mediated eGFP reporter integration in zebrafish. Sci Rep 6, 34991. https://doi.org/10.1038/srep34991

Oudelaar, A.M., Harrold, C.L., Hanssen, L.L.P., Telenius, J.M., Higgs, D.R., Hughes, J.R., 2019. A revised model for promoter competition based on multi-way chromatin interactions at the α-globin locus. Nat Commun 10, 5412. https://doi.org/10.1038/s41467-019-13404-x

Peters, H., Balling, R., 1999. Teeth: where and how to make them. Trends in Genetics 15, 59–65. https://doi.org/10.1016/S0168-9525(98)01662-X

Piotrowski, T., Nüsslein-Volhard, C., 2000. The Endoderm Plays an Important Role in Patterning the Segmented Pharyngeal Region in Zebrafish (Danio rerio). Developmental Biology 225, 339–356. https://doi.org/10.1006/dbio.2000.9842

Rasch, L.J., Martin, K.J., Cooper, R.L., Metscher, B.D., Underwood, C.J., Fraser, G.J., 2016. An ancient dental gene set governs development and continuous regeneration of teeth in sharks. Developmental Biology 415, 347–370. https://doi.org/10.1016/j.ydbio.2016.01.038

Renn, J., Winkler, C., 2009. Osterix-mCherry transgenic medaka for in vivo imaging of bone formation. Dev. Dyn. 238, 241–248. https://doi.org/10.1002/dvdy.21836

Sagai, T., Amano, T., Tamura, M., Mizushina, Y., Sumiyama, K., Shiroishi, T., 2009. A cluster of three long-range enhancers directs regional *Shh* expression in the epithelial linings. Development 136, 1665–1674. https://doi.org/10.1242/dev.032714

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Seo, H., Amano, T., Seki, R., Sagai, T., Kim, J., Cho, S.W., Shiroishi, T., 2018. Upstream Enhancer Elements of *Shh* Regulate Oral and Dental Patterning. J Dent Res 97, 1055–1063. https://doi.org/10.1177/0022034518758642

Seppala, M., Fraser, G., Birjandi, A., Xavier, G., Cobourne, M., 2017. Sonic Hedgehog Signaling and Development of the Dentition. JDB 5, 6. https://doi.org/10.3390/jdb5020006

Stepaniak, M.D., Square, T.A., Miller, C.T., 2021. Evolved Bmp6 enhancer alleles drive spatial shifts in gene expression during tooth development in sticklebacks. Genetics iyab151.

Talbot, J.C., Johnson, S.L., Kimmel, C.B., 2010. *hand2* and Dlx genes specify dorsal, intermediate and ventral domains within zebrafish pharyngeal arches. Development 137, 2507–2517. https://doi.org/10.1242/dev.049700

Thisse, C., Thisse, B., 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc 3, 59–69. https://doi.org/10.1038/nprot.2007.514

Thomas, B.L., Liu, J.K., Rubenstein, J.L.R., Sharpe, P.T., 2000. Independent regulation of *Dlx2* expression in the epithelium and mesenchyme of the first branchial arch. Development 127, 217–224.

Van der heyden, C., Huysseune, A., 2000. Dynamics of tooth formation and replacement in the zebrafish (Danio rerio) (Teleostei, Cyprinidae). Dev. Dyn. 219, 486–496. https://doi.org/10.1002/1097-0177(2000)9999:9999<::AID-DVDY1069>3.0.CO;2-Z

Verstraeten, B., van Hengel, J., Sanders, E., Van Roy, F., Huysseune, A., 2013. N-cadherin is Required for Cytodifferentiation during Zebrafish Odontogenesis. J Dent Res 92, 365–370. https://doi.org/10.1177/0022034513477424

Wan, Y., Otsuna, H., Chien, Chi-Bin, Hansen, C., 2009. An interactive visualization tool for multi-channel confocal microscopy data in neurobiology research. IEEE Trans. Visual. Comput. Graphics 15, 1489–1496. https://doi.org/10.1109/TVCG.2009.118

Wong, B., 2011. Points of View: Color blindness. Nat Methods 8, 441–441. https://doi.org/10.1038/nmeth.1618

Xu, Y., Zhou, Y.L., Gonzalez, F.J., Snead, M.L., 2007. CCAAT/Enhancer-binding Protein δ (C/EBPδ) Maintains Amelogenin Expression in the Absence of C/EBPα in Vivo. Journal of Biological Chemistry 282, 29882–29889. https://doi.org/10.1074/jbc.M702097200

Yu, J.C., Fox, Z.D., Crimp, J.L., Littleford, H.E., Jowdry, A.L., Jackman, W.R., 2015. Hedgehog signaling regulates dental papilla formation and tooth size during zebrafish odontogenesis. Dev. Dyn. 244, 577–590. https://doi.org/10.1002/dvdy.24258

Zhao, Z., Stock, D.W., Buchanan, A.V., Weiss, K.M., 2000. Expression of Dlx genes during the development of the murine dentition. Dev Gene Evol 210, 270. https://doi.org/10.1007/s004270050314

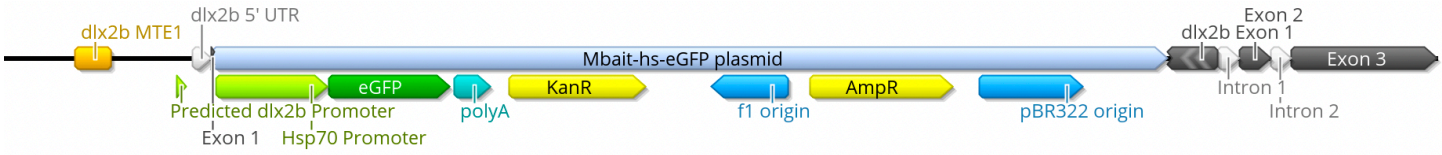
Zhou, Y.L., Snead, M.L., 2000. Identification of CCAAT/Enhancer-binding Protein alpha as a Transactivator of the Mouse Amelogenin Gene. The Journal of Biological Chemistry 275, 12273–12280.

Supplemental Information

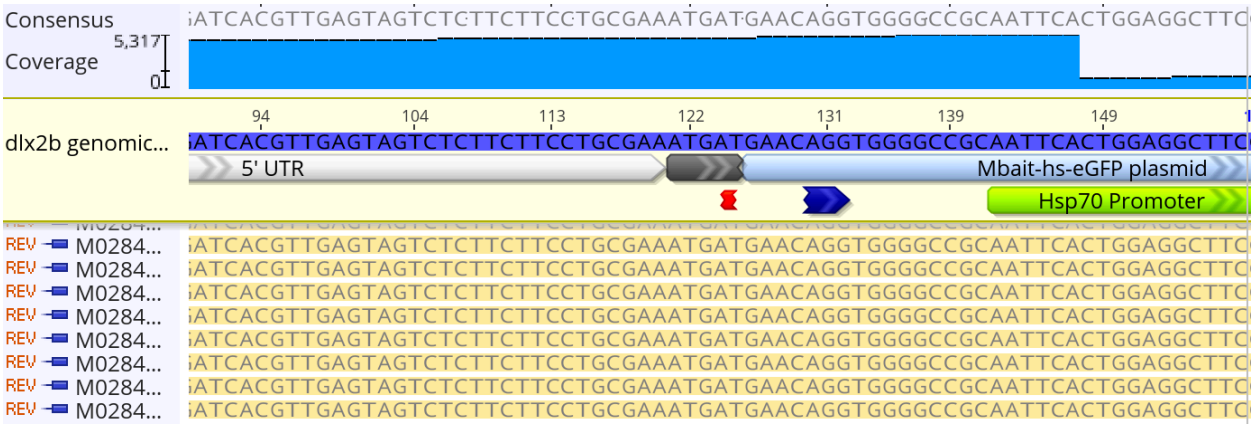
#	dlx2b 5' end	dlx2b 3' end	Mutation	Type of construct	Promoter	Injected molecule	Scored as	Tooth GFP+ / total examined	Tooth activity
1	1895	0	none	Tol2kit (Gateway)	dlx2b	plasmid	line	>100/>100	+
2	1555	0	none	Tol2kit (Gateway)	dlx2b	plasmid	F0	4/7	+
3	930	0	none	Tol2kit (Gateway)	dlx2b	plasmid	F0	3/4	+
4	856	0	none	Tol2kit (Gateway)	dlx2b	plasmid	F0	3/5	+
5	817	0	none	Tol2kit (Gateway)	dlx2b	plasmid	line	>100/>100	+
6	767	0	none	Tol2kit (Gateway)	dlx2b	plasmid	F0	0/19	-
7	622	0	none	Tol2kit (Gateway)	dlx2b	plasmid	F0	0/6	-
8	817	444	none	Tol2kit (Gateway)	ß-actin	plasmid	line	>100/>100	+
9	817	604	none	Tol2kit (Gateway)	ß-actin	plasmid	line	>100/>100	+
10	744	674	none	Tol2kit (Gateway)	ß-actin	plasmid	line	0/13	-
11	0	0	none	Tol2kit (Gateway)	ß-actin	plasmid	line	0/>100	-
12	817	604	Pea3	Tol2kit (Gateway)	ß-actin	plasmid	F0	12/35	+
13	817	604	Cebp	NEBuilder	ß-actin	amplicon	F0	3/5	+
14	817	604	Dlx	NEBuilder	ß-actin	amplicon	F0	3/7	+
15	817	604	FoxA	NEBuilder	ß-actin	amplicon	F0	6/14	+
16	817	604	Ap1	NEBuilder	ß-actin	amplicon	F0	7/11	+
17	817	604	HPM	NEBuilder	ß-actin	amplicon	F0	12/14	+
18	817	604	DFCA	NEBuilder	ß-actin	amplicon	F0	0/30	-
19	817	604	none	NEBuilder	hsp70	plasmid	F0	5/6	+

Table S1. Summary of reporter constructs used in this study. The dlx2b^{MTE1} line is from construct #9. Mutation sequences are shown in Fig. S2.

A



B



C

Fig. S1. The dlx2b^{KI} allele. (A) Schematic diagram of the Mbait-hs-eGFP plasmid insertion into the *dlx2b* locus. Sequences at the 5' (B) and 3' (C) end of the insertion.

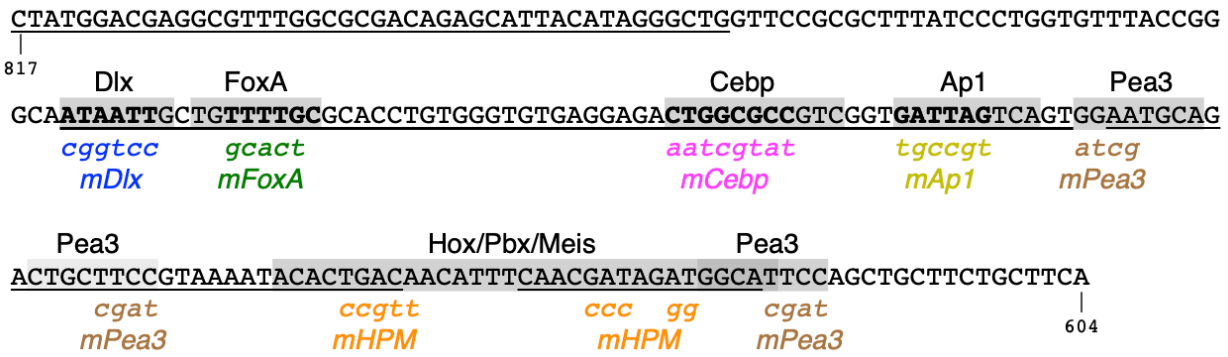


Figure S2. Sequences mutated to test the necessity of various predicted/possible transcription factor binding sites.

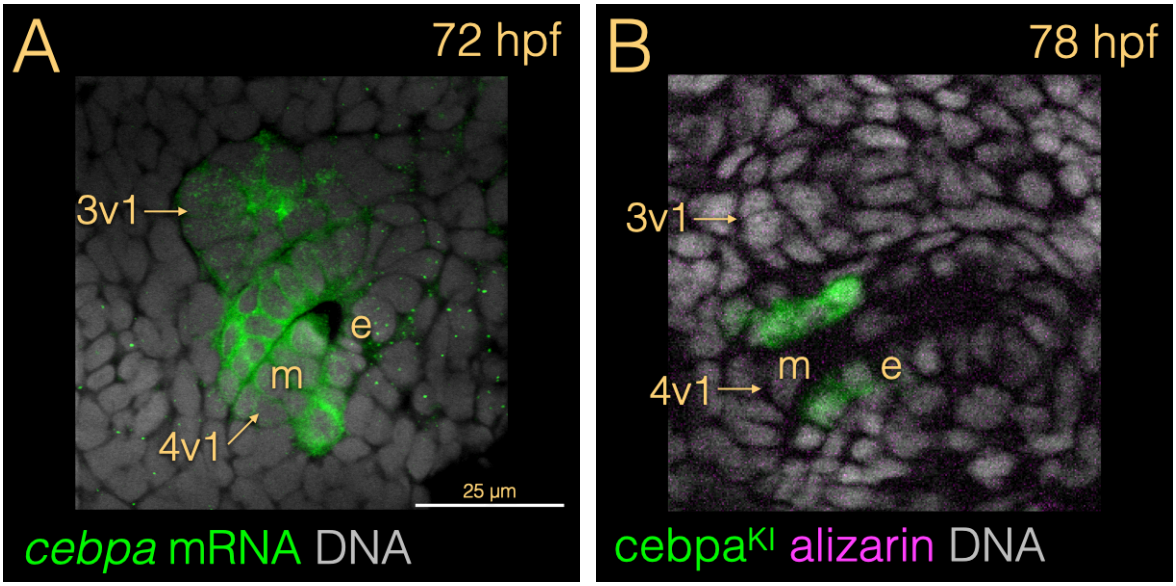


Figure S3. Expression of *cebpa*. (A) mRNA in situ hybridization. (B) GFP expression from an individual heterozygous for the cebpa^{KI} allele. For both methods, expression is strongest in the basal part of the inner dental epithelium.

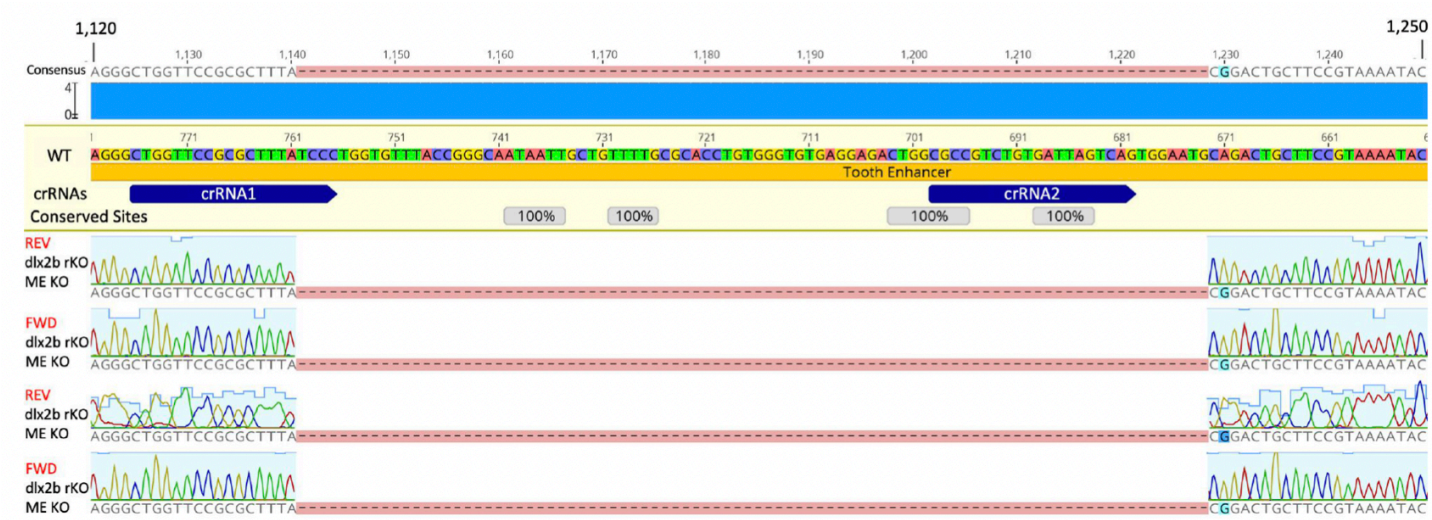


Figure S4: Sequence of the deletion in the dlx2b$^{KI\Delta MTEI}$ allele.