A 190 base pair, TGF-β responsive tooth and fin enhancer is required for stickleback *Bmp6* expression

Priscilla A. Erickson[a], Phillip A. Cleves[a], Nicholas A. Ellis[a], Kevin T. Schwalbach[a], James C. Hart[a], and Craig T. Miller[a,*]

[a] Department of Molecular and Cell Biology, University of California, Berkeley, CA, 94720

* Address correspondence to: ctmiller@berkeley.edu

Abstract

The ligands of the Bone Morphogenetic Protein (BMP) family of developmental signaling molecules are often under the control of complex *cis*-regulatory modules and play diverse roles in vertebrate development and evolution. Here, we investigated the *cis*-regulatory control of stickleback *Bmp6*. We identified a 190 bp enhancer ~2.5 kilobases 5' of the *Bmp6* gene that recapitulates expression in developing teeth and fins, with a core 72 bp sequence that is sufficient for both domains. By testing orthologous enhancers with varying degrees of sequence conservation from outgroup teleosts in transgenic reporter gene assays in sticklebacks and zebrafish, we found that the function of this regulatory element appear to have been conserved for over 250 million years of teleost evolution. We show that a predicted binding site for the TGFß effector Smad3 in this enhancer is required for enhancer function and that pharmacological inhibition of TGFß signaling abolishes enhancer activity and severely reduces endogenous *Bmp6* expression. Finally, we used TALENs to disrupt the enhancer *in vivo* and find that *Bmp6* expression is dramatically reduced in teeth and fins, suggesting this enhancer is necessary for expression of the *Bmp6* locus. This work identifies a relatively short regulatory sequence that is required for expression in multiple tissues and, combined with previous work, suggests that shared regulatory networks control limb and tooth development.

Keywords

Bone Morphogenetic Protein, enhancer, tooth development, stickleback, zebrafish, Bmp6, TGFß

Introduction

Bone Morphogenetic Protein (BMP) ligands, the largest subfamily of TGFβ ligands, play multiple essential roles during vertebrate development (Hogan, 1996; Kingsley, 1994; Massagué, 2012), including during craniofacial and tooth development (Nie et al., 2006). Many vertebrate organs develop through reciprocal permissive and instructive signaling between adjacent epithelial and mesenchymal tissues, often involving multiple BMP ligands (Bellusci et al., 1996; Dassule and McMahon, 1998; Dudley et al., 1999; Jung et al., 1998). These pleiotropic functions of BMP ligands are orchestrated by typically large, modular, regulatory regions, which work together to drive complex spatiotemporally restricted expression patterns (Pregizer and Mortlock, 2009).

In humans, regulatory variation in *Bmp* genes has been associated with developmental disorders including brachydactyly and other birth defects (Dathe et al., 2009; Justice et al., 2012), as well as colorectal cancer (Houlston et al., 2008; Lubbe et al., 2012). In other animals, variation in the expression of BMP genes has also been associated with major evolved changes in morphology, including beak shape in Darwin's finches (Abzhanov et al., 2004), jaw size and shape in cichlid fish (Albertson et al., 2005), and tooth number in stickleback fish (Cleves et al 2014).

While the *cis*-regulatory architecture of *Bmp2*, *Bmp4*, *Bmp5*, and *Bmp7* has been studied in mice (Adams et al., 2007; Chandler et al., 2007; Guenther et al., 2008; Jumlongras et al., 2012), less is known about *Bmp6* and *Bmp* gene regulation in other vertebrates. Although not required for viability in the mouse, *Bmp6* is required for axial skeletal patterning (Solloway et al., 1998), kidney function (Dendooven et al., 2011), and physiological iron regulation (Andriopoulos et al., 2009). Non-coding variants in human *Bmp6* have been associated with

human height variation (Gudbjartsson et al., 2008; Wood et al., 2014), as well as orofacial clefting birth defects (Shi et al., 2012). A *cis*-regulatory allele of stickleback *Bmp6* with reduced *Bmp6* expression in developing tooth tissue has recently been shown to be associated with evolved increases in tooth number in derived freshwater sticklebacks, likely adaptive for the shift in diet in freshwater sticklebacks relative to their marine ancestors (Cleves et al., 2014).

BMP signaling plays complex and, in general, poorly understood roles during the development of placodes. During tooth development, multiple BMP genes are expressed dynamically in developing odontogentic epithelia and mesenchyme (Aberg et al., 1997; Vainio et al., 1993). Several lines of evidence reveal BMP signaling plays activating roles during odontogenesis. First, epithelial BMP4 activates *Msx* expression in the mesenchyme, and exogenous BMP from a bead (Bei and Maas, 1998; Chen et al., 1996) or transgene (Zhao et al., 2000) can partially rescue tooth development in *Msx1* mutant mice. Second, in mice, teeth arrest at the bud-to-cap transition in *Bmpr1a* mutants (Andl et al., 2004; Liu et al., 2005). Third, exogenous BMP4 beads can induce molar development in mice (Kavanagh et al., 2007). Fourth, in fish, pharmacological inhibition of BMP signaling can inhibit tooth formation in cichlids (Fraser et al., 2013). In contrast, other evidence supports BMP signaling playing inhibitory effects during the development of teeth and other placodes. In mice, *Pax9* expression marks early dental mesenchyme, and BMP2 and BMP4 inhibit *Pax9* expression (Neubüser et al., 1997). In zebrafish, inhibition of BMP signaling produces supernumerary teeth with altered morphology (Jackman et al., 2013). During development of both feather and hair placodes, BMPs play inhibitory roles (Botchkarev et al., 1999; Jung et al., 1998; Mou et al., 2006, 2011), and suppression of epithelial BMP signaling is required for hair placode induction (reviewed in Biggs and Mikkola, 2014). Together these results suggest that complex positive and negative

interactions between epithelial and mesenchymal BMPs are critical for placode development, yet the regulation of these interactions remains less well understood.

Despite the major role BMP signaling plays during tooth development, little is known about the *cis*-regulatory sequences that drive dynamic *Bmp* expression in early developing odontogenic epithelia and mesenchyme. In mice, a late-stage ameloblast enhancer has been identified for the *Bmp4* gene (Feng et al., 2002); however this enhancer is not reported to be active during embryogenesis, or in dental mesenchyme. A second enhancer of mouse *Bmp4* has been described that is active during embryogenesis and drives expression in dental epithelium but not mesenchyme (Jumlongras et al., 2012). Tooth epithelial and mesenchymal enhancers of the mouse *Bmp2* gene have been localized to a ~150 kb region 3' of *Bmp2* (Chandler et al., 2007), however these enhancers have not yet been further mapped, and in general, cis-regulation of BMPs in dental mesenchyme is poorly understood. Furthermore, since mice are monophyodonts that form one wave of primary teeth and no replacements, less is known about cis-regulatory elements that drive expression in developing and replacement teeth in polyphyodont vertebrates (such as fish) that replace their teeth continuously. Because of the recently identified *cis*-regulatory allele of *Bmp6* associated with evolved changes in stickleback tooth number (Cleves et al., 2014) and to dissect epithelial and mesenchymal *cis*-regulation of vertebrate Bmp signaling, we sought to begin to identify the *cis*-regulatory architecture of the stickleback *Bmp6* gene.

Methods:

Animal statement and fish husbandry:

All animal work was approved by the Institutional Animal Care and Use Committee of the University of California-Berkeley (protocol number R330). Sticklebacks (*Gasterosteus aculeatus*) were raised in ~10% seawater (3.5 g/l Instant Ocean salt, 0.217 ml/l 10% sodium bicarbonate) at 18° C, and crosses were generated by *in vitro* fertilization. Zebrafish (*Danio rerio*) were raised in a recirculating system under standard conditions, and embryos were collected either from natural spawning or *in vitro* fertilization and raised at 28.5 degrees (Westerfield, 2007).

BAC Isolation and Recombineering:

Bacterial Artificial Chromosomes (BACs) from the CHORI-213 and CHORI-215 (Salmon River marine and Paxton benthic freshwater stickleback, respectively) BAC libraries were identified by overgo screening (Ross et al., 1999) using the following overgoes: 5'-TGTGACGTTGACCTCAGCTAGACT-3' and 5'-GAGGATTTAAACCGGGAGTCTAGC -3'. BAC ends were sequenced using Sp6 and T7 primers and mapped to the stickleback genome using the UCSC browser. BAC CHORI-215-29E12 was chosen for reporter analysis because *Bmp6* was relatively centrally located in the BAC. Inverted Tol2 sites were recombineered into the Lox511 site of the pTarbac2.1 backbone according to Suster et al. (2011) using primers PTARBAC_tol2FWD and PTARBAC_tol2REV, and ampicillin resistance was used to select successfully recombineered BAC clones. To place GFP into exon 1 of *Bmp6* as a reporter, a GFP/kanamycin resistant cassette was amplified from pGFP-FRT-Kan-FRT (Suster et al., 2011) using primers GFP_Bmp6_for and GFP_Bmp6_rev (Table S1), which contained 50 bp homology to the beginning and end of the first exon of stickleback *Bmp6,* respectively. This

construct was then recombineered into the BAC containing iTol2 sites to produce the final reporter BAC (see Fig. 6A-C).

Enhancer Constructs:

The vector for the stickleback 2.8 kb enhancer/promoter construct was generated using pENTRbasGFP and pTolDest (Villefranc et al., 2007) using Gateway cloning to produce a construct with the carp *ß-actin* basal promoter (Scheer and Campos-Ortega, 1999) upstream of EGFP, flanked by Tol2 sites (Urasaki et al., 2006). Next, a 2,810 bp sequence upstream of the predicted *Bmp6* transcriptional start site was PCR amplified from BAC CHORI-213-256N24 using primers Gac_3kb_for and Gac_3kb_rev and cloned upstream of the carp *ß-actin* promoter using a ClaI restriction site. Blocks of conserved sequences within the 2.8 kb construct were identified as CS1, CS2, and CS3 from the UCSC 8 species Multiz conservation track (see Fig. 1A). These sequences were cloned into ClaI site of the carp *ß-actin* reporter construct using primers shown in Table S1. CS1 was cloned with Gac_3kb_for and Gac_CS1_rev. CS2 was cloned with Gac_CS2_for and Gac_CS2_rev. CS3 was cloned with Gac_CS3_for and Gac_3kb_rev. CS2+3 was cloned with Gac_CS2_for and Gac_3kb_rev. Because the CS1 fragment drove weak expression with the *ß-actin* promoter, we switched to using a well-characterized zebrafish *hsp70* promoter construct, which we found to drive much brighter expression in transgenic stickleback embryos. CS1 and CS2+3 were also cloned into the *hsp70* promoter construct for additional testing using the same genomic primer sequences but with Nhe and BamHI restriction sites in place of ClaI. The 190 bp and 72 bp enhancer sequences were amplified from the 2.8 kb construct with primers indicated in Table S1 and cloned into the *hsp70* construct.

147 The orthologous enhancer sequences were identified in other teleost genomes using the
148 UCSC genome browser (genome.ucsc.edu) to identify sequence conservation. Zebrafish and
149 medaka (*Oryzias latipes*) wild-type genomic DNA was isolated by standard phenol-chlorofom
150 extraction and enhancers were amplified using primers (Table S1) designed from the respective
151 genome assemblies (zv9/danRer7 and oryLat2) and cloned into the *hsp70* promoter construct.
152 The Atlantic cod (*Gadus morhua*) enhancer DNA sequence was identified by sequence
153 conservation on contig CAEA01327401 of the Atlantic cod genome assembly (UCSC,
154 gadMor1). This short, unassembled contig is flanked by repetitive sequence, but the intervening
155 sequence contains a 94 bp stretch that has 92.4% sequence identity to the stickleback enhancer
156 and is likely the orthologous sequence. We synthesized a 130 bp construct of Atlantic cod
157 sequence by using two primers for amplification (Gmo_for and Gmo_rev, see Table S1) and two
158 additional overlapping oligonucleotides as template (Gmo_temp1 and Gmo_temp2). The
159 template oligonucleotides were added to standard Phusion (NEB) PCR reaction at a
160 concentration of 0.05 μM to amplify the full 130bp sequence, which was then cloned into the
161 Tol2 construct as described above.
162
163 *Sequence Analysis:*
164 Sequence alignments were generated using ClustalW2
165 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (Larkin et al., 2007) and Boxshade
166 (http://www.ch.embnet.org/software/BOX_form.html). Binding sites were predicted with the
167 UniProbe database (http://the_brain.bwh.harvard.edu/uniprobe/) (Newburger and Bulyk, 2009)
168 and PROMO (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3)
169 (Farre et al., 2003; Messeguer et al., 2002).

Imaging and Microscopy:

Transgenic lines were imaged using a Leica DM2500 compound microscope equipped with a Leica DFC500 camera, a Leica M165FC dissecting microscope equipped with a DFC340 FX camera, or a Zeiss 700 confocal microscope. Transgenic fish were fixed for 4 hours at 4°C in either 4% paraformaldehyde in 1X PBS or 10% neutral buffered formalin. For Alizarin red fluorescent counterstaining of GFP lines, 0.01% Alizarin red was added to the fixative. Tooth number was counted on the DM2500 with TX2 filter to visualize Alizarin-stained teeth. Tooth germs with GFP+ epithelia were counted on photographs of GFP fluorescence.

Fish injections and line generation:

Transposase mRNA was produced from the pCS2-TP plasmid (Kawakami et al., 2004) with the mMessage mMachine SP6 *in vitro* transcription kit (Ambion) according to manufacturer's instructions and purified with a Qiagen RNeasy column. Zebrafish injections were performed with 25 ng/μL plasmid DNA and 37.5 ng/μL transposase and 0.05% phenol red as previously described (Fisher et al., 2006). Because stickleback embryos are much larger than zebrafish embryos, the DNA and RNA concentrations were increased to 37.5 and 75 ng/μL respectively. Stable transgenic lines were generated by outcrossing injected fish to non-transgenic fish and visually screening for fluorescent transgenic offspring. At least two stable lines were observed for each construct to ensure fluorescent patterns were due to the transgene and not artifacts of the transgene integration sites.

Site directed mutagenesis:

193 Mutagenesis primers were designed using the online Quickchange tool
194 (http://www.genomics.agilent.com/primerDesignProgram). For constructs containing multiple
195 mutations, the mutagenesis was performed in multiple rounds. Mutagenesis reactions were
196 performed with 125 ng of each primer, 50 ng plasmid template, 200 μM dNTPs, and Pfu Turbo
197 polymerase and buffer. Cycling conditions were 95°C for 30 seconds, followed by 16 cycles of
198 95°C / 30 s, 55°C / 60 s, and 68°C /780 s. Primer sequences can be found in supplementary
199 Table 1; the mutated sequences are shown in Fig. 3A. *DpnI* was added immediately after cycling,
200 and the reaction was incubated for 1 hr at 37°C, then immediately transformed into Top10
201 chemically competent *E. coli* cells.
202
203 *Drug treatments:*
204 SB431542 and XAV939 (Sigma) were dissolved in DMSO to concentrations of 100 μM
205 and 10 μM, respectively. The drug was then diluted into stickleback water or zebrafish system
206 water to working concentrations (25-100 μM for SB431542 and 5-10 μM for XAV939). A
207 DMSO vehicle control was done in parallel with all drug treatments. Drug treatment was
208 performed in 6- or 24-well cell culture dishes. Sticklebacks were treated from 2 dpf to 5 dpf for
209 observation of pectoral and median fin expression, and for 5-7 days post-hatching for
210 observation of tooth GFP. Zebrafish were treated beginning at 10 hpf for observation of median
211 fin and beginning at 24 hpf for pectoral fin and tooth expression. For multiday treatments, fresh
212 solution was applied every 48 hours until the end of the experiment.
213
214 *In situ hybridization (ISH):*

Bmp6 in situ hybridization was performed on embryos and newly-hatched juveniles as previously described (Cleves et al., 2014). For pharyngeal tooth and gill *in situs*, the branchial skeleton was dissected out of the embryo and cut along the dorsal midline prior to the hybridization step.

Mutagenesis using TALENs:

TAL Effector Nucleases (TALENs) were targeted to the predicted Smad3 binding site within the 190 bp enhancer using TAL Effector Nuclear Targeter 2.0 (https://tale-nt.cac.cornell.edu/) using the Cermak architecture (Cermak et al., 2011; Doyle et al., 2012). TALEN plasmids were generated using the RVDs shown in Table S4. TALEN mRNAs were produced with the Mmessage Mmachine kit (Ambion), purified with Qiagen RNeasy columns, and injected into one-cell stickleback embryos at a concentration of 40 ng/µL for each mRNA plus 0.05% phenol red. Embryos and juvenile fish were screened for lesions in the Smad3 site by screening for loss of an XbaI cut site in a 144 bp PCR product amplified with primers Gac_190_for and Gac_72_rev (see Fig. 4G). F1 animals with deletions visible on a 2% agarose gel (~15 bp or larger) were crossed to generate animals used in *in situ* hybridization. Because the F1 parents carried different TALEN-induced lesions, the F2 animals used for ISH were transheterozygotes for two slightly different alleles of the enhancer deletion (see Fig. 6E).

Results:

A *Bmp6* reporter BAC recapitulates endogenous *Bmp6* expression

To begin to identify the *cis*-regulatory architecture of the stickleback *Bmp6* gene, we generated a *Bmp6* reporter line by identifying a bacterial artificial chromosome (CHORI

BAC215-29E12) containing 180 kb of sequence starting ~52 kb upstream of *Bmp6*. Inverted Tol2 sequences were recombineered into the backbone of this BAC and the first exon of *Bmp6* was replaced with GFP coding sequence. This transgenic construct drove GFP reporter expression in a variety of tissues throughout development (Fig. S1), including the embryonic tailbud following somitogenesis (3.5 dpf), the embryonic heart and ventrolateral cells in the pharyngeal region (4 dpf), the distal edge of the developing pectoral fin, and the distal edge of the median fin (5 dpf). After hatching (10-15 dpf), expression was seen in oral and pharyngeal teeth, the pericardium, cells surrounding the opercle and branchiostegal rays, gill buds, and gill rakers.

We compared this transgene expression pattern to the expression pattern of endogenous *Bmp6* mRNA via *in situ* hybridization. We observed *Bmp6* expression in nearly all of the same domains as the reporter BAC (Fig. S2), including the tailbud (at 3.5 dpf), heart, the distal edges of the median and pectoral fins (at 5 dpf), gills, gill rakers, and in the previously described (Cleves et al., 2014) epithelium and mesenchyme of developing teeth (assayed at ~12 dpf). However, several domains observed by *in situ* hybridization were not observed in the BAC transgenic line, including the notochord, the dorsal medial diencephalon, the eyes, and the ears (Fig. S2), suggesting that regulatory elements lying outside of the 180 kb of genomic sequence contained within the BAC control these *Bmp6* expression domain.

A conserved 190 bp enhancer drives tooth, median fin, and pectoral fin expression in both stickleback and zebrafish

To begin to identify regulatory elements contained within this 180 kb genomic interval containing *Bmp6*, we first cloned a construct containing ~2.8 kb of sequence immediately

upstream of stickleback *Bmp6* containing regions of sequence conserved among other teleosts (Fig. 1A). This construct drove GFP expression in a number of tissues that were similar to expression patterns driven by the BAC (Fig. S3, compare to Fig. S1), including the tailbud, the heart, pectoral and median fins, oral and pharyngeal teeth, gills, and the pericardium. Other domains driven by the BAC were not observed in the 5' construct, including gill rakers, opercle, and branchiostegal rays; these domains are likely driven by more distal regulatory elements contained within the BAC but excluded from the 2.8 kb sequence. Combined, these results suggest that much of the regulatory information for *Bmp6* is contained within the 2.8 kb upstream sequence, but that other regulatory elements drive additional expression domains.

We hypothesized that the different anatomical sites of expression driven by the 2.8 kb fragment result from multiple anatomically specific enhancers within this sequence. We first tested three non-overlapping subclones, each containing a block of evolutionarily conserved sequence (Fig. 1A). While the most 5' subclone (CS1) drove robust reporter gene expression in most domains of the 2.8 kb fragment, neither the middle (CS2) nor 3' subclone (CS3) drove detectable GFP expression in fins, teeth, or other domains driven by the 2.8 kb fragment at the 3-5 dpf or post-hatching (10-13 dpf) stages. Furthermore, a construct containing CS2 + CS3 also drove no detectable pattern of GFP with either the *ß-actin* or *hsp70* promoter. Next, we focused on the 5'-most region (CS1), and tested a 190 bp fragment highly conserved within teleosts (Fig. 1B). This 190 bp fragment drove robust GFP expression in the distal edges of the pectoral and median fins, and oral and pharyngeal teeth (Fig. 1C-E). Within developing teeth, GFP expression was observed in the inner dental epithelium (IDE) for all constructs (Fig. S4) as well as the interior mesenchyme of mature functional teeth (Fig. 1D), similar to endogenous *Bmp6* expression during tooth development (Cleves et al., 2014). Robust tooth GFP expression was

seen in all teeth at all stages examined including in juveniles and adults, suggesting tooth enhancer activity is present in both primary and replacement teeth (Fig. 1D-E, data not shown). Some domains, including the gills, were lost when CS1 was reduced to the 190bp fragment, suggesting that flanking sequence is required for these domains. When the orientation of the enhancer was flipped with respect to the *hsp70* promoter, 77% (38/49) of injected fish had pectoral and/or median fin expression at 5 dpf, and 69% (27/39) had oral and/or pharyngeal tooth expression at 13 dpf. This result suggests that this enhancer functions regardless of orientation to the promoter. Combined, our results suggest that most domains driven by the 2.8 kb enhancer are driven by the short 190 bp conserved sequence. This 190 bp minimal sequence does not differ between marine and freshwater sticklebacks, though several marine-freshwater sequence differences exist in the surrounding sequences of CS1.

Conservation of *cis* regulatory elements and *trans* machinery in teleosts

Because we used evolutionary sequence conservation to identify the 190 bp minimal enhancer and the sequence was partially conserved to zebrafish, we hypothesized that this 190 bp stickleback enhancer would show similar activity in transgenic zebrafish. Stickleback and zebrafish are ~250 million years divergent (Near et al., 2012) and share only 3 short blocks (totaling 28 bp, Fig. 2A) of perfectly conserved nucleotides in the middle of the enhancer. However, the stickleback enhancer robustly drove a highly similar expression pattern in zebrafish, with expression in the distal edges of the median and pectoral fins, and pharyngeal tooth epithelium and mesenchyme (Fig. 2B-D), suggesting that the *trans* factors activating the enhancer are conserved in distantly related teleosts. We next asked whether the orthologous sequence from the zebrafish genome had similar enhancer activity in both zebrafish and

sticklebacks. A construct containing 477 bp of sequence from the orthologous region of the zebrafish genome drove weak expression in these expression domains (distal edges of median and pectoral fins, and teeth) in a subset of transgenic zebrafish offspring obtained (Fig. 2E-G and Table S2). In sticklebacks, seven stable transgenic lines with the zebrafish sequence driving GFP had no fin expression, although one transgenic line displayed very faint expression in the distal edges of the median and pectoral fins (Fig. 2H-I). None of the eight lines had GFP expression in teeth (Fig. 2J). Therefore, sticklebacks and zebrafish likely share the *trans* machinery sufficient to drive expression from the stickleback sequence, but the *cis* regulatory information present in the zebrafish orthologous sequence is not sufficient to drive tooth expression in the stickleback *trans* environment.

Because the zebrafish enhancer shows much less sequence conservation to sticklebacks relative to other teleosts (Fig. 2A), we hypothesized that the loss of robustness and loss of tooth expression may be unique to the zebrafish *cis*-regulatory element. We generated constructs containing the orthologous enhancer sequences of a beloniform (medaka) and a gadiform (Atlantic cod), which fall between zebrafish and sticklebacks in the teleost phylogeny (Near et al., 2012). We found that sequences from both additional species drove expression in fins and teeth in both stickleback and zebrafish embryos (Fig. S5, Table S2), although the cod enhancer appeared to be slightly less robust (Table S2).

Based on the apparent partial conservation of enhancer function in zebrafish and the conserved activities of the medaka and cod enhancers, we further shortened the stickleback enhancer to contain the sequence most highly conserved among teleosts, a 72 bp sequence near the center of the 190 bp construct, and hypothesized that it would drive the tooth, median fin, and pectoral fin expression domains. In support of this hypothesis, this construct in a stable line of

zebrafish was sufficient to drive strong GFP expression in teeth and median and pectoral fins (Fig. S6). Notably, the heart domain driven by this construct was considerably brighter relative to the 190 bp enhancer, suggesting that this short sequence may have lost additional repressor elements that limit expression in the heart. A similar pattern of brighter heart expression was observed in stickleback injected with this construct compared to the 190 bp larger construct (data not shown). These results suggest that the flanking conserved sequences are not required for the basic enhancer pattern in fins and teeth, but may be important for fine-tuning the transcriptional output.

A predicted Smad3 binding site is required for enhancer function.

To identify candidate transcription factor binding sites within the 190 bp enhancer, we used UniProbe and PROMO (Newburger and Bulyk, 2009; Farre et al., 2003; Messeguer et al., 2002) and found predicted binding sites of transcription factors in several signaling pathways involved in developmental regulation: FGF (PEA3), retinoic acid (RAR-γ), Wnt (TCF/Lef), and TGFβ (Smad3), as well as a predicted homeodomain binding site (Fig. 3A). We were particularly interested in the homeodomain binding site given the known crosstalk between the *Msx1* and *Bmp4* genes during mouse tooth development (Bei and Maas, 1998; Chen et al., 1996; Jumlongras et al., 2012), and the predicted TCF/Lef sites, given the previously described roles of Wnt signaling regulating *Bmp4* dental mesenchyme expression in mice (Fujimori et al., 2010; O'Connell et al., 2012). We quantified the number of stickleback embryos showing pectoral and/or median fin, as well as pharyngeal and/or oral tooth expression, when injected with constructs containing mutated binding sites. The mutation of TCF/Lef and Smad3 binding sites significantly decreased the percentage of fish with median and/or pectoral fin expression

353 domains, whereas the predicted PEA3, RAR-γ, and homeodomain mutations did not (Fig. 3B).
354 Likewise, only the mutations in predicted TCF/Lef and Smad3 sites affected tooth expression,
355 with especially reduced expression when the predicted Smad3 binding sites were mutated (Fig.
356 3C). We made stable zebrafish lines for each of the Smad3 and TCF/Lef mutated enhancers and
357 found that the Smad3-mutated reporter construct did not drive robust expression in zebrafish fins
358 or teeth, while the TCF/Lef mutated construct did drive these domains, albeit at apparently
359 reduced levels (Fig. S7). Since the Smad3-mutated construct did not drive fin or tooth expression
360 in zebrafish, we generated a stable line in sticklebacks and found that this line similarly did not
361 drive detectable median fin, pectoral fin, or tooth expression (Fig. 4J). Therefore, the predicted
362 Smad3 sites are required for normal enhancer output, while TCF/Lef sites may be responsible for
363 expression level but not tissue specificity.
364
365 **A small molecule inhibitor of TGFβ signaling, but not a small molecule inhibitor of Wnt**
366 **signaling, abolishes enhancer function**
367 Since the predicted Smad3 binding site was necessary for enhancer function, we
368 hypothesized that reducing TGFβ signaling (mediated by Smad3) would result in a loss of
369 expression driven by the enhancer. To pharmacologically inhibit TGFβ signaling, we treated
370 transgenic sticklebacks and zebrafish embryos with SB431542, a specific inhibitor of ALK4/5
371 phosphatase activity that abrogates TGF-β signaling in zebrafish (Inman et al., 2002; Sun et al.,
372 2006). After 6 days of treatment in sticklebacks, GFP expression driven by the 190 bp enhancer
373 was reduced in a dose-dependent manner in the epithelium, but not mesenchyme, of developing
374 pharyngeal teeth, with tooth epithelial expression abolished at 50 μM and reduced at 25 μM (Fig.
375 4A-C). Tooth mesenchymal expression was slightly diminished at 50 μM and apparently

unaffected at 25 μM. Similarly, GFP reporter expression was lost in the pharyngeal teeth of newly hatched zebrafish upon treatment with SB431542 from 24 hpf until 5 dpf (Fig. 4D-F). In sticklebacks, we also saw a reduction, but not complete loss, of pectoral and median fin expression driven by the transgene upon treatment with SB431542 (Fig. S8), while the reduction was more severe in the fins of zebrafish. Combined with our site-directed mutagenesis of the Smad3 binding site result, these pharmacological data suggest that TGFβ signaling mediated by ALK4/5 (likely signaling via Smad3 binding) is necessary for tooth epithelium enhancer activity. However other signals likely contribute to the expression in the pectoral and median fins and tooth mesenchyme, as drug treatment did not completely abolish these expression domains in sticklebacks.

Since the mutation of TCF/Lef binding sites appeared to decrease enhancer activity in sticklebacks and zebrafish (Fig. 3, Fig. S7), we hypothesized that Wnt signaling might be an additional input into the 190 bp *Bmp6* enhancer. To test this hypothesis, we treated transgenic fish with SB431542, XAV939 (a specific inhibitor of the Wnt signaling pathway that is active in zebrafish (Huang et al., 2009)), or both drugs in combination at low and high doses. Treatment with a high-dose combination of XAV939 and SB431542 decreased the standard length of fish (data not shown), possibly indicating a slight developmental delay. With XAV939 or SB431542 treatment alone, there was no effect of the drug on tooth number, suggesting that neither drug alone arrests tooth development. However, the two drugs in combination significantly reduced ventral pharyngeal tooth number (Fig. 5H), including at the low dose that did not affect fish standard length, suggesting that XAV939 is bioactive in sticklebacks and that reducing Wnt and TGFβ signaling together disrupts tooth development.

There was no obvious qualitatively detectable effect of XAV939 treatment on the intensity of enhancer expression in the teeth, either alone or in combination with SB431542 (Fig. 5; compare D and E to A, and compare F and G to B and C). However, tooth mesenchymal GFP in the combined drug treatment appeared slightly lower than with SB431542 treatment alone (insets of Fig. 5). Importantly, we never saw a complete loss of mesenchymal GFP with any drug treatment, but frequently saw complete loss of epithelial GFP with SB431542 treatment. To quantify the effect of drug treatment on epithelial GFP expression, we counted the number of GFP^+ tooth epithelia (regardless of fluorescent intensity) in each treatment and expressed it as a ratio to the total number of Alizarin red-stained teeth. XAV939 had no effect on the relative number of GFP^+ epithelia, while SB431542 had a strong, dose-dependent effect (Fig. 5I). In combination with SB431542, there was no additional effect of XAV939 on reporter expression (GFP^+ epithelia in the combination treatments did not differ from treatment with SB431542 alone). Combined, our results suggest that SB431542, but not XAV939, affects enhancer activity and that simultaneous inhibition of Wnt and TGFβ signaling affects tooth development.

The 190 bp enhancer is necessary for *Bmp6* expression

As an additional test of the importance of the predicted Smad3 binding site, we generated a pair of TALENs designed to induce mutations in this region of the enhancer (see Fig. 4G). This pair of TALENs was highly efficient at producing lesions, detected molecularly by loss of an XbaI restriction site, and confirmed by Sanger sequencing in a subset of individuals (Table S3; example deletions shown in Fig. 6E). Upon injection of these TALENs into a stable transgenic line of the 190 bp enhancer driving GFP, 95% of animals (40 of 42) showed partial or full loss of GFP fluorescence in the pectoral fins and median fin expression at 5 dpf. In those same animals,

95% of animals (39 of 41) also showed partial or complete loss of oral and/or pharyngeal tooth expression at 12-13 dpf (see example in Fig. 4I). Thus, the lesions generated by these TALENs are highly effective at disrupting activity driven by this 190bp enhancer.

We next tested whether the sequence targeted by the TALENs was necessary for *Bmp6* expression by injecting the TALENs into a stable transgenic line of the *Bmp6:GFP* BAC reporter. 91% (61/67) of animals had a reduction or complete loss of pectoral and median fin expression, and 89% (8/9) of dissected tooth plates showed severe reductions of GFP expression in the pharyngeal teeth (representative animals shown in Fig. 6 F-K). Notably, GFP expression in the embryonic and juvenile heart was detectable at seemingly unaffected levels in all animals, suggesting that the enhancer is not necessary for this expression domain. Additionally, gill expression appeared to be reduced but not completely eliminated in all animals observed (n=6), and gill raker expression was only slightly reduced. These data suggest the enhancer is required for some (e.g. pectoral fin, median fin, tooth epithelium), but not all domains of *Bmp6* expression.

Finally, we tested the role of the enhancer in driving endogenous *Bmp6* expression by performing *in situ* hybridization for *Bmp6* in fish *trans*-heterozygous for different TALEN-induced mutations in the predicted Smad3 binding site (Fig. 6E). In these *trans*-heterozygous fish, expression of *Bmp6* was dramatically reduced in fins, tooth epithelia and gills, but gill raker expression appeared similar to wild-type controls (Fig. 6L-Q). Despite the severe loss of *Bmp6* expression in tooth epithelia in mutant fish, expression in the mesenchyme of developing teeth was still detectable, although at apparently reduced levels (Fig. 6N-O). Thus, this enhancer is required to maintain normal levels of *Bmp6* expression in developing fins and tooth epithelia.

444 **TGFβ signaling is necessary for normal *Bmp6* expression levels**

445 Since enhancer activity was lost upon treatment with a TGFβ inhibitor, and the enhancer
446 is required for normal *Bmp6* expression, we predicted that endogenous *Bmp6* expression would
447 likewise be reduced upon inhibition of TGFβ signaling. By *in situ* hybridization, pectoral fin and
448 tooth epithelium expression of *Bmp6* were both reduced upon 100 μM SB431542 treatment (Fig.
449 7A-D). SB431542 treatment also reduced GFP expression in reporter BAC animals in fins and
450 teeth (Fig. 7E-H). The effect of the drug on BAC-driven GFP was not robustly observed with a
451 50 μM treatment (data not shown), despite the strong effect that this dose had on enhancer
452 expression (Fig. 4). Together these data support a model in which TGFβ signaling is required for
453 *Bmp6* expression in teeth and fins and exerts its effect through the putative Smad3 binding site
454 that is necessary for enhancer function.
455
456 **Discussion**
457 **A short, conserved enhancer with pleiotropic expression domains required for *Bmp6* tooth**
458 **and fin expression**
459 Here we have identified a 190 base pair enhancer that is highly conserved in teleosts and
460 is both necessary and sufficient for tooth and fin expression of stickleback *Bmp6*. Site-directed
461 mutagenesis of a predicted Smad3 binding site and pharmacological experiments suggest this
462 enhancer is TGFβ-responsive. Though this enhancer drives expression in several of *Bmp6*'s
463 endogenous domains, our results suggest that like other *Bmp* genes, stickleback *Bmp6* contains a
464 complex *cis*-regulatory architecture composed of multiple modules driving expression in
465 different domains. We detected embryonic expression domains of *Bmp6* by *in situ* hybridization,
466 such as the eye, ear, diencephalon, and notochord, that were not observed in the BAC reporter

line, suggesting that the regulatory elements controlling these domains lie outside of the 180 kb of stickleback DNA included in the BAC. Furthermore, while TALEN mutations severely reduced expression in the fins and teeth, every BAC reporter fish injected with TALENs had GFP expression in the heart, suggesting that the enhancer is not required for heart expression. Thus, the short enhancer presented here contributes to a subset of the endogenous *Bmp6* expression domains, with other domains likely driven by other enhancers greater than ~100 kb away. Evidence for long range distant enhancers of stickleback *Bmp6* is expected, given the frequent finding of long distance regulatory elements for developmental regulatory genes, including other vertebrate *Bmp* genes (reviewed in Preziger and Mortlock, 2009). Interestingly, despite the presence of redundant "shadow" enhancers found in many genes (Calle-Mustienes et al., 2005; Marinić et al., 2013; Perry et al., 2010), this enhancer appears to be required for several *Bmp6* expression domains; additional enhancers did not appear to sufficiently compensate in driving *Bmp6* expression when the 5' enhancer was targeted with TALENs.

Another teleost tooth/fin enhancer has been described with overall similar expression patterns observed in this *Bmp6* enhancer. In zebrafish, an FGF-responsive enhancer mediates *Dlx2* expression in teeth and median and pectoral fins (Jackman and Stock, 2006). Additionally, in mice, a *Bmp4* enhancer drives tooth epithelium and limb bud expression by responding to Pitx and Msx homeodomains (Jumlongras et al., 2012). The shared fin/limb and tooth expression domains of these *cis*-regulatory elements and the one described here suggest that fin and tooth development share multiple *cis*-regulatory networks, with at least three signaling pathways (FGF, Pitx/Msx, and TGFß) involved in generating similar gene expression readouts in teeth and fins/limbs. Gene expression patterns of paired fins are thought to be co-opted from median fin expression domains in agnathans (Freitas et al., 2006). The *Bmp6* enhancer presented here

appears to be teleost-specific, as we did not find evidence of this conserved enhancer sequence in the genomes of lamprey, elephant shark, or spotted gar. Thus, our results suggest that teleosts may have secondarily coopted components of a gene regulatory network in developing median and pectoral fins and teeth.

Elucidating the *cis*-regulatory architecture of stickleback *Bmp6* and evolved changes in *Bmp6*'s *cis*-regulatory architecture will help test the hypothesis that evolved changes in *Bmp6* *cis*-regulation underlie the evolved increases in freshwater stickleback tooth number we previously described (Cleves et al., 2014). Although the 190 bp core *Bmp6* enhancer presented here contains no nucleotide differences between low-toothed marine and high-toothed freshwater sticklebacks, several nucleotide differences exist in the sequence flanking the enhancer, which might contribute to the cis-regulatory differences observed between marine and freshwater alleles of *Bmp6.* Future studies will focus on whether these differences result in differential *cis*-regulatory activity between the marine and freshwater alleles of *Bmp6*.

Conservation and turnover of *cis*- and *trans*-regulatory information

It has been proposed that the *cis*-regulatory architecture of developmental control genes often consist of multiple independent modules, each of which drives expression in a particular tissue or cell type (Carroll, 2008; Stern, 2000). Because the *Bmp6* enhancer drives multiple anatomical expression domains and is only partially conserved to zebrafish, we hypothesized that domains may have been sequentially added to the enhancer during teleost evolution, and that the different anatomical domains would be separable. Contrary to these predictions, our site directed mutagenesis and subcloning experiments of the stickleback *Bmp6* enhancer appeared to affect all or none of the different expression domains, suggesting the different anatomical domains might

not be separable and instead reflect ability to respond to a signal or signals present in multiple tissues.

Furthermore, enhancers from all four teleost species tested were sufficient to drive fin and tooth expression in zebrafish. However, the zebrafish enhancer, the most evolutionary divergent enhancer tested in this study, did not function robustly in sticklebacks, suggesting that the *trans* factors driving expression might have changed during the divergence of the two species. Similarly, testing a zebrafish *Dlx2* tooth and fin enhancer in both zebrafish and Mexican tetra revealed that loss of oral *Dlx2* expression in zebrafish is caused by changes in *trans* factors, as the *Dlx2* zebrafish tooth enhancer is active in tetra oral teeth (Jackman and Stock, 2006). In both *C. elegans* and *Drosophila*, transgenic testing of *cis*-regulatory elements from two fly or worm species in both fly or worm species revealed that the greater the evolutionary distance separating two regulatory elements, the more likely upstream *trans* differences are to have evolved (Gordon and Ruvinsky, 2012). But, subtle changes in *trans*-acting factors can maintain similar expression patterns despite *cis* changes in divergent lineages (Barrière et al., 2012). Our results suggest a combination of conservation and divergence of *trans* factors, as stickleback sequence worked robustly in zebrafish, but zebrafish sequence was not functional in stickleback. Additionally, SB431542 treatment affected the stickleback enhancer in zebrafish more severely than in stickleback. Even at a low dose of SB431542 (25 μM), the enhancer was completely shut off in both epithelia and mesenchyme of zebrafish teeth (see Fig. 4E-F). This result supports potential *trans* regulatory divergence between stickleback and zebrafish, because it suggests that the enhancer's expression may be more sensitive to TGFß signaling in zebrafish than in stickleback.

A role for TGFß in the regulation of BMPs

To our knowledge, this study is the first to support a role for TGFβ signaling in controlling Bmp signaling via a *cis*-regulatory input. Conditional deletion of *Tgfbr1* (*Alk5*) in mouse neural crest lineages results in reduced expression of *Bmp4* and delayed tooth initiation (Zhao et al., 2008); however, the mechanism of this interaction has not been described. Other studies have shown both positive and negative correlations between *Bmp6* expression and TGFβ levels: *Smad3* -/- chondrocytes have reduced *Bmp6* expression (Li et al., 2006), whereas *Bmp6* expression is increased in *Smad3* -/- tendons undergoing tissue repair (Katzel et al., 2011). Our data suggest that in sticklebacks, TGFß signaling activates *Bmp6* expression in multiple tissues via a predicted Smad3 binding site. In teeth, blocking TGFß signaling using the inhibitor SB431542 caused loss of epithelial reporter expression, but the effect on the mesenchymal expression was less severe (Fig. 4C, Fig. 5). The same pattern was observed in endogenous *Bmp6* expression (Fig. 6O). This result suggests that epithelial and mesenchymal *Bmp6* expression domains respond to partially different signaling pathways, with epithelial expression much more sensitive to TGFß disruption.

We observed that a higher dose of TGFβ inhibitor SB431542 was required to shut off endogenous *Bmp6* expression relative to expression driven solely by the 190bp enhancer. While a 50 μM treatment almost completely eliminated enhancer expression (Fig. 4), at this dose we did not observe a strong difference in GFP expression driven by the reporter BAC. Only at the higher dose of 100 μM did we observe a change in BAC reporter expression and endogenous *Bmp6* expression (Fig. 7). This finding suggests that in its native genomic context, the enhancer may be less sensitive to TGFβ signaling perturbations than when it is isolated in a reporter construct. There may be additional non-TGFβ regulatory elements that drive *Bmp6* expression in

559 the same tooth and fin domains such that a decrease in TGFβ signaling has a less obvious effect
560 at lower doses. Furthermore, the effect of SB431542 treatment on endogenous *Bmp6* expression
561 and BAC reporter expression was not as dramatic as deletion of the Smad3 binding site with
562 TALENs (compare Fig. 6 to Fig. 7). This finding suggests that other non-TGFß factors may bind
563 sequences immediately surrounding the Smad3 binding site to drive enhancer expression.
564 However, the predicted Smad3 site is absolutely required, as loss of this site completely
565 eliminates enhancer activity (Fig. 4J).
566
567 **Combined effects of Wnt and TGFß on tooth development**
568 Although our site-directed mutagenesis experiment indicated that TCF/Lef predicted
569 binding sites might be important for enhancer function (Fig. 3), pharmacological testing with
570 XAV939 did not support the hypothesis that the enhancer requires Wnt signaling inputs for
571 enhancer function. A stable line of zebrafish containing the TCF/Lef mutated reporter also drove
572 robust reporter expression in fins and teeth, providing a second piece of evidence that the
573 enhancer does not require Wnt input. This result was somewhat surprising, as the expression
574 domains driven by the *Bmp6* enhancer are similar to a TCF reporter zebrafish line (Shimizu et
575 al., 2012). The reduction in activity seen from mutating the TCF/Lef sites may have been caused
576 by other unknown binding sites overlapping the mutated base pairs, by inadvertently creating
577 repressive motifs, or by somehow altering the binding of the Smad3 complex. The mutations
578 may have affected the level, but not pattern, of GFP expression, making the construct appear less
579 robust in our transient transgenic assay. We did note that combined treatment with XAV939 and
580 SB431542 caused a slight decrease in mesenchymal tooth GFP expression (see insets of Fig. 5),

however, this effect was less reproducible than the complete loss of epithelial expression seen upon SB431542 treatment alone.

The combination treatment with SB431542 and XAV939 did reduce tooth number in sticklebacks, suggesting that Wnt and TGFβ signaling pathways together are required for maintaining normal tooth development and patterning. In mice, as well as in diphyodont humans and polyphyodonts including snakes and alligators, Wnt signaling is required for tooth formation and replacement (Adaimy et al., 2007; Bohring et al., 2009; Gaete and Tucker, 2013; Genderen et al., 1994; Liu et al., 2008; Wu et al., 2013). In mice, TGFß signaling is also required for tooth development (Ferguson et al., 1998, 2001; Oka et al., 2007). Antisense abrogation of both *TGFB2* and *TGFBRII* in cultured mandibles resulted in accelerated tooth formation (Chai et al., 1994, 1999), however the *TGFB2* knockout mouse has no reported tooth phenotype (Sanford et al., 1997). While the *TGFBRII* knockout dies prior to tooth formation (Oshima et al., 1996), conditional ablation in neural crest cells prevents terminal differentiation of odontoblasts (Oka et al., 2007), while conditional ablation in *Osx*-expressing odontoblasts revealed a necessary role for *TGFBRII* in molar root development (Wang et al., 2013). Furthermore, Wnt and TGFß signaling are required to activate *Eda* and *Edar* in appropriate patterns in the developing tooth germs (Laurikkala et al., 2001). However, to our knowledge, this study is the first to show a partially redundant requirement for TGFß and Wnt during tooth development, as only XAV939 and SB431542 doubly treated fish had reduced tooth numbers. Future studies of this enhancer will further test the hypothesis that this enhancer responds to TGFß signaling to control *Bmp6* expression during tooth and fin development.

Conclusions

We have identified a 190 base pair conserved enhancer required for tooth, fin, and other expression domains of stickleback *Bmp6*. Site directed mutagenesis and pharmacology experiments support the hypothesis that this enhancer responds to TGFß signaling via a Smad3 binding site. Expression driven by this enhancer in tooth epithelial cells appears more sensitive to TGFß levels than expression in tooth mesenchymal cells. To our knowledge, this is the first demonstration of a likely *cis*-regulatory link between TGFß signaling and *Bmp* expression in teeth. *In vivo* deletion of this enhancer using TALENs caused severe disruption of *Bmp6* expression in fins and tooth epithelia, suggesting this enhancer is required for normal expression patterns in a subset of *Bmp6*'s endogenous domains. Finally, we demonstrate that a combination of TGFß signaling and Wnt signaling is required for normal tooth development in sticklebacks.

Acknowledgements

This work was supported by NIH R01 #DE021475. We thank David Kingsley for support and advice on BAC isolation, Tim Howes and David Kingsley for the generous gift of the Tol2/*hsp70* backbone, Daniel Schlenk and Anita Kuepper for providing medaka specimens, Natasha Naidoo for assistance in cloning the medaka reporter construct, and Lisa Kronstad for providing the site-directed mutagenesis protocol.

Figure Legends

Fig. 1. A conserved 190 bp enhancer upstream of *Bmp6* drives gene expression in several domains. (A) The 5' region of stickleback *Bmp6* from the UCSC genome browser (http://genome.ucsc.edu/). The region of genomic DNA used in the 2.8 kb enhancer construct (see Fig. S3) is shown in green, conserved sequences (CS) 1-3 are shown in purple, and the subcloned 190bp enhancer is shown in yellow. The first exon and part of the first intron of *Bmp6* are shown in thick and thin black lines, respectively (bottom). Conservation peaks and alignments (dark blue and grey) are shown from the 8-Species MultiZ track. (B) Zoom in on the middle of CS1, approximately 2.5 kb upstream of the predicted *Bmp6* transcription start site. The 190 bp enhancer, the 72 bp minimal enhancer (see Fig. S6), and a predicted Smad3 binding site (see Fig. 3-4) are shown in yellow, pink, and blue, respectively. The conservation track is shown as dark blue peaks, above green alignments showing conservation to medaka, tetraodon, fugu, and zebrafish, from top to bottom. (C) GFP reporter expression pattern driven by the 190 bp enhancer in a 5 dpf (stage 22, (Swarup, 1958)) stickleback embryo. Strong expression was seen in the distal edge of the developing pectoral fin (arrow), the heart (asterisk), and the distal edge of the median fin (arrowhead). (D) Confocal projection of GFP reporter expression in the ventral pharyngeal tooth plate in a ~10 mm stickleback fry. Expression was observed in the epithelium of developing tooth germs (arrow) and the odontogenic mesenchyme (arrowhead) in the cores of ossified teeth. Bones are fluorescently stained with Alizarin red. (E) GFP reporter expression in the oral teeth (arrow) of a 30 dpf stickleback fry. GFP in the lens is an internal control for the zebrafish *hsp70* promoter used in the transgenic construct. Scale bars = 200 μm.

Fig. 2. Evolutionary functional conservation of the *Bmp6* enhancer in teleosts. (A) Sequence alignments of four teleost sequences relative to the 190 bp stickleback enhancer. The perfectly conserved Smad3 dimer binding site is marked in blue, and purple arrows mark the boundaries of the 72 bp minimal enhancer (see Fig. S6). (B-D) The stickleback sequence reporter construct stably integrated into the zebrafish genome drove expression in the distal edge of the median fin at 24 hpf (arrow in B), the distal edge of the pectoral fin at 48 hpf (arrow in C), and tooth epithelium (arrow) and mesenchyme (arrowhead) at 5 dpf (D). (E-G) A 477 bp construct of zebrafish genomic sequence centered around the conserved sequence of the enhancer drove similar, but weaker expression in the median fin of a 33 hpf zebrafish (arrow in E), pectoral fins of a 48 hpf zebrafish (inset of F), and teeth of a 5 dpf zebrafish (G). (H-I) Although not detected in seven of eight stable lines, in one of eight stable lines, the zebrafish sequence drove faint expression in the distal edges of the median fin (arrow in H) and pectoral fins (arrow in J) of 5 dpf stickleback. However, no expression was detected in tooth germs in newly hatched fry in any line (J). See Table S2 for quantification of expression domains of transgenic lines. Bone is fluorescently stained with Alizarin red in (D, G, J). Scale bars = 200 μm.

Fig. 3. Mutations in predicted Smad3 binding sites severely reduce enhancer function. (A) Binding sites predicted by UniProbe and PROMO are highlighted with a unique color for each signaling pathway. Highlighted sequences represent the "predicted sequence" from PROMO or the "K-mer" from UniProbe. Mutated base pairs are shown with lowercase letters. Nucleotide positions conserved to zebrafish are indicated with an asterisk, and arrows indicate the 72 bp minimal enhancer sequence. (B-C) Sticklebacks were injected with each mutated construct and scored for pectoral fin and/or median fin expression at 5 dpf (B) and oral and/or pharyngeal tooth

668 expression at 12-13 dpf (C). Frequency of expression in these domains is shown as a percentage
669 of the total number of GFP-positive fish (scored as GFP expression driven by the *hsp70*
670 promoter anywhere at 5 dpf or in the lens at 12-13 dpf) on the y-axis.
671
672 **Fig. 4. Pharmacological disruption of TGFβ signaling or TALEN-induced mutations of the**
673 **predicted Smad3 binding site reduce enhancer activity.** (A-C) Treatment of stickleback fry
674 for 7 days in SB431542 (an ALK5 inhibitor) severely reduced GFP expression driven by the 190
675 bp enhancer in a dose-dependent manner. Expression was severely reduced in the epithelia
676 (arrows), but not mesenchyme (asterisks), of pharyngeal teeth at both low (25 μM, B) and high
677 (50 μM, C) doses relative to controls (A). (D-F) SB431542 also eliminated GFP driven by the
678 stickleback enhancer in a zebrafish *trans* environment. (G) The sequence targeted by TALENs
679 contains a predicted Smad3 homodimer binding site (blue). The TALEN binding sites are
680 indicated in purple text and the purple scissors indicate the approximate site of endonuclease
681 activity. The XbaI site used for molecular screening is underlined in green, and the mutagenized
682 sequence of the Smad3 binding site, indicated by orange letters, is shown below. (H-I) Injection
683 of the TALENs into stable transgenic fish carrying the 190 bp reporter construct resulted in near
684 complete loss of GFP expression in 95% of injected animals (I) relative to controls (H). Residual
685 GFP seen in (I) is likely the result of the mosaicism of TALEN-induced lesions. (J). Mutating the
686 predicted Smad3 binding site resulted in a loss of GFP expression in both epithelium and
687 mesenchyme of pharyngeal teeth in 3/3 stickleback lines observed. Bone is fluorescently
688 counterstained with Alizarin red. Scale bars = 200 μm.
689

Fig. 5. Wnt signaling is not required for enhancer function, but Wnt and TGFβ are required for tooth development. Newly hatched stickleback fry were treated with DMSO (control, A), SB431542 (B-C), XAV939 (D-E), or a combination of the two drugs at low (25μM for SB431542 and 5 μM for XAV939, F) or high (50 uM for SB431542 or 10 uM XAV939, G) doses for 5 days. Main panels show Alizarin red and GFP for the ventral tooth plate; insets show GFP only for mesenchyme of a single tooth from the dorsal tooth plate. (B, C) SB431542 reduced GFP in tooth epithelia (arrows) relative to control (A, and see Fig. 3). However, mesenchymal GFP (arrowhead, inset) was less severely reduced. (D, E) XAV939 alone did not affect GFP expression in epithelia (arrows) or mesenchyme (arrowheads) at either dose. (F, G) No strong additional effect on GFP expression was seen when XAV939 and SB431542 were combined, though mesenchymal GFP appeared slightly lower in the combined dose. (H) A combination of SB431542 and XAV939 significantly reduced ventral pharyngeal tooth number. (I) Treatment with SB431542, but not XAV939, decreased the number of green tooth epithelia relative to total ventral teeth (ratio is expressed as a decimal). XAV939 had no additional effect on green epithelia in combination with SB431542. Tukey HSD P-values of relevant comparisons are shown above with asterisks (*=P<0.05, ** =P<0.0005, n.s.=P>0.05). Scale bars = 200 μm

Fig. 6. The 5' 190bp enhancer is necessary for *Bmp6* expression. (A) Schematic of the genomic location of the 180 kb CHORI BAC29E12 with respect to *Bmp6* and nearby genes (coding regions shown in black are *Ipo4*, *Pdcd6*, *Txndc5*, *Muted*, *Eef1e1*, and *Slc35b3* from left to right). (B) Recombineering strategy for introducing GFP into the first exon *of Bmp6;* grey bars indicate exons. (C) Final circular BAC with inverted Tol2 sites for transposition and GFP reporter (not to scale). (D) Strategy for introducing TALEN lesions into the 190 bp 5' enhancer.

The same TALENs were used to target the enhancer in stable transgenic BAC fish and at the endogenous *Bmp6* locus (diagram not to scale). (E) Sequences of stable mutant enhancer alleles. For the endogenous locus targeting, F2 fish trans-heterozygous for two different enhancer mutations were generated. Fish in (M) carried alleles 1 and 2; fish in (O) and (Q) carried alleles 1 and 3. The predicted Smad3 binding site is indicated with blue text in the wild type sequence. (F, G) In the reporter BAC, TALEN injection frequently severely reduced GFP expression from the pectoral fin relative to controls at 5 dpf. A small patch of mosaic, unaffected GFP is indicated with the arrow in (G). (H, I) TALEN injection also eliminated much of the *Bmp6* tooth expression (I). (J, K) GFP expression was also reduced in gills (asterisk) and slightly reduced in the gill rakers (arrowhead). (L-M). Mutations in the enhancer caused a reduction in pectoral fin *Bmp6* expression relative to wild-type siblings. (N, O) *Bmp6* expression was also lost in tooth epithelia (arrows), but was not entirely lost in mesenchyme (arrowheads). (P, Q) Expression was also noticeably reduced in gills (asterisk), though gill raker expression (arrows) appears similar to wild-type sibling controls. Scale bars = 100 μm.

Fig. 7. Treatment with TGFβ inhibitor SB431542 reduces *Bmp6* expression. (A-D) Newly hatched stickleback were treated with 100 μM SB431542 or DMSO vehicle control for 5 days, and *Bmp6* expression was assayed by *in situ* hybridization. Drug treatment severely reduced *Bmp6* expression in fins (A, B) and also reduced *Bmp6* expression in tooth epithelia (C, D). Likewise, GFP driven by the *Bmp6* locus in the reporter BAC was also reduced in fins (arrowheads in E, F) and teeth (G, H) after SB431542 treatment. Scale bars = 100 μm.

Supplementary Figure and Table Legends

Fig. S1. Domains of GFP expression in a stickleback *Bmp6* BAC reporter.

The first exon of *Bmp6* was replaced with GFP in a 180 kb stickleback BAC (see Fig. 6). Stable lines carrying this reporter construct displayed GFP expression in a variety of tissues. Expression was detected in the distal edge of the forming median caudal fin (A) and ventrolateral cells surrounding the heart and pharyngeal region (B) at 3 dpf when viewed laterally. At 5 dpf, expression was observed in cells in the distal edge of the median fin (arrow in C, arrowhead points to autofluorescent pigment cell) and the distal edge of the developing pectoral fins (arrows in D). Soon after hatching (at 11-12 dpf), expression was observed in pharyngeal (E) and oral (F, ventral view) teeth. Additionally, GFP^+ cells were observed surrounding the branchiostegal rays (G), opercle (H), and gill rakers (arrow I). Cells in the soft tissue of the gill buds were also seen (asterisk in I). GFP was also observed in cells surrounding the heart (asterisk in J, ventral view and K, lateral view). Bone is fluorescently counterstained with Alizarin red in E-G. Scale bars = 200 μm (A-D); 100 μm (E-K).

Fig. S2. Expression domains of stickleback *Bmp6*. *Bmp6* expression was assayed by whole mount *in situ* hybridization at 3 dpf (A, B), 5 dpf (C-H), and 12 dpf (I-K). Expression was observed in the forming median fin in the tailbud (A, C), heart (lateral view in B), eyes and ears (asterisk) (D), distal edge of the developing pectoral fins (E), dorsal medial diencephalon (F), notochord and dorsal neural tube (G), hindgut and cloaca (H), gill rakers and gill buds (arrow and asterisk in I), branchiostegal rays (J), and pharyngeal teeth (K). Scale bars = 200 μm (A-E); 50 μm (F-K).

Fig. S3. Expression driven by 2.8 kb of genomic sequence upstream of *Bmp6*. During early development, the 2.8 kb reporter construct drove expression in the forming median fin in the tailbud at 3 dpf (A), cells in the developing heart and pharyngeal pouches at 4 dpf (B), the distal edge of the median (C) and pectoral (D) fins at 5 dpf. After hatching (11-14 dpf), additional expression was observed in pharyngeal teeth (E), pericardial cells (F), the developing gills (G), oral teeth (H), the scapulocoracoid cartilage (I), and the distal edge of the opercle (J). In fry (22-30 dpf), expression was observed in the distal tips of fin rays (K) and the developing pelvic spine (arrow) and kidney (asterisk) (L). Red in E, G-K is Alizarin red counterstaining of bone, and yellow spots in H-J are autofluorescent pigment cells. Scale bars = 200 μm (A-D); 100 μm (E-H, J, K); 500 μm (I, L).

Fig. S4. Enhancer GFP and *Bmp6* expression are detected in the inner but not outer dental epithelium. (A-C) GFP expression driven by the reporter BAC (A), 2.8 kb reporter construct (B), and 190bp reporter construct (C) was limited to the inner dental epithelium (IDE) as visualized under differential interference contrast optics. (D) *Bmp6* mRNA expression was also restricted to the IDE as previously reported (Cleves et al 2014). The outer dental epithelium (ODE) is indicated with white arrows in A-D. (E-H) Images from A-D with the the outer edge of the ODE traced with white dashed lines and the outer edge of the IDE traced with black dashed lines. Scale bars = 100 μm.

Fig. S5. Atlantic cod and medaka enhancers drive fin and tooth expression in both stickleback and zebrafish. Orthologous *Bmp6* enhancer sequences from two species from clades that evolved between zebrafish and sticklebacks, medaka and Atlantic cod, drove similar

781 expression patterns in stickleback (A-D) and zebrafish (E-H). Expression was observed in the
782 distal edges of the pectoral fins (arrows) at 5 dpf in stickleback (A, C) or 48-56 hpf zebrafish (E,
783 G). Later in development, pharyngeal tooth expression was observed at 20 dpf in stickleback (B,
784 D) or 5 dpf zebrafish (F, H). Bright neural expression in (C) was not seen in other lines and was
785 likely an artifact of the transgene integration site. Scale bars = 200 μm.
786
787 **Fig. S6. 72bp of conserved stickleback genomic sequence is sufficient for enhancer domains**
788 **but increases heart expression.** The minimally sufficient 72 bp construct drove expression in
789 (A) mesenchyme (arrowhead) and epithelium (arrow) of a 5 dpf zebrafish ventral tooth plate, (B)
790 the distal edge of the median fin in a 24 hpf zebrafish and (C) the distal edge of the pectoral fin
791 (arrow) in a 48 hpf zebrafish. The intensity of heart expression was noticeably increased
792 (asterisk, compare to Fig. 2C), suggesting that the shortened sequence had lost some repressor
793 activity. Scale bars = 100 μm (A); 200 μm (B-C).
794
795 **Fig. S7. Mutation of Smad3 but not TCF/Lef predicted binding sites affects reporter**
796 **expression in zebrafish**. Zebrafish stable lines were obtained for two constructs that appeared to
797 show reduced activity in sticklebacks. (A-C) The wild-type 190 bp stickleback enhancer drove
798 expression in the distal edge of the median fin (A), distal edge of the pectoral fin (B) and
799 pharyngeal teeth (C) of zebrafish. Images in A-C are the same as in Fig. 2 for comparison with
800 D-J. (D-F) The TCF/Lef mutated construct showed expression in the median fin at 24 hpf (arrow
801 in D), pectoral fin at 48 hpf (arrow in E), and pharyngeal teeth at 5 dpf (F) in all lines observed.
802 Brain expression in E was not typical and is likely an artifact of the transgene integration site.
803 (G-I) In nearly all (8/9) lines observed, the Smad3 mutated construct lacked expression in the

804 median fin (arrow in G), pectoral fin (arrow in H), and teeth (I). One of 9 lines had very faint
805 expression in these domains. Scale bars = 200 μm.
806
807 **Fig. S8. SB431542 reduces reporter GFP expression in the median and pectoral fins in both**
808 **sticklebacks and zebrafish.** Treatment with 50 μM SB431542 reduced, but did not completely
809 eliminate, GFP reporter expression driven by the 190 bp enhancer relative to vehicle (DMSO)
810 controls in the pectoral fins (A, B, E, F) and median fins (C, D, H, G) of both stickleback (A-D)
811 and zebrafish (E-H) embryos. Scale bars = 400 μm.

812 **References:**

813 Aberg, T., Wozney, J., and Thesleff, I. (1997). Expression patterns of bone morphogenetic
814 proteins (Bmps) in the developing mouse tooth suggest roles in morphogenesis and cell
815 differentiation. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *210*, 383–396.

816 Abzhanov, A., Protas, M., Grant, B.R., Grant, P.R., and Tabin, C.J. (2004). Bmp4 and
817 morphological variation of beaks in Darwin's finches. Science *305*, 1462–1465.

818 Adaimy, L., Chouery, E., Mégarbané, H., Mroueh, S., Delague, V., Nicolas, E., Belguith, H., de
819 Mazancourt, P., and Mégarbané, A. (2007). Mutation in WNT10A Is Associated with an
820 Autosomal Recessive Ectodermal Dysplasia: The Odonto-onycho-dermal Dysplasia. Am. J.
821 Hum. Genet. *81*, 821–828.

822 Adams, D., Karolak, M., Robertson, E., and Oxburgh, L. (2007). Control of kidney, eye and
823 limb expression of Bmp7 by an enhancer element highly conserved between species. Dev.
824 Biol. *311*, 679–690.

825 Albertson, R.C., Streelman, J.T., Kocher, T.D., and Yelick, P.C. (2005). Integration and
826 evolution of the cichlid mandible: The molecular basis of alternate feeding strategies. Proc.
827 Natl. Acad. Sci. U. S. A. *102*, 16287–16292.

828 Andl, T., Ahn, K., Kairo, A., Chu, E.Y., Wine-Lee, L., Reddy, S.T., Croft, N.J., Cebra-Thomas, J.A.,
829 Metzger, D., Chambon, P., et al. (2004). Epithelial Bmpr1a regulates differentiation and
830 proliferation in postnatal hair follicles and is essential for tooth development. Dev. Camb.
831 Engl. *131*, 2257–2268.

832 Andriopoulos, B., Corradini, E., Xia, Y., Faasse, S.A., Chen, S., Grgurevic, L., Knutson, M.D.,
833 Pietrangelo, A., Vukicevic, S., Lin, H.Y., et al. (2009). BMP6 is a key endogenous regulator of
834 hepcidin expression and iron metabolism. Nat. Genet. *41*, 482–487.

835 Barrière, A., Gordon, K.L., and Ruvinsky, I. (2012). Coevolution within and between
836 Regulatory Loci Can Preserve Promoter Function Despite Evolutionary Rate Acceleration.
837 PLoS Genet *8*, e1002961.

838 Bei, M., and Maas, R. (1998). FGFs and BMP4 induce both Msx1-independent and Msx1-
839 dependent signaling pathways in early tooth. Development *125*, 4325–4333.

840 Bellusci, S., Henderson, R., Winnier, G., Oikawa, T., and Hogan, B.L. (1996). Evidence from
841 normal expression and targeted misexpression that bone morphogenetic protein (Bmp-4)
842 plays a role in mouse embryonic lung morphogenesis. Development *122*, 1693–1702.

843 Biggs, L.C., and Mikkola, M.L. (2014). Early inductive events in ectodermal appendage
844 morphogenesis. Semin. Cell Dev. Biol. *25–26*, 11–21.

845 Bohring, A., Stamm, T., Spaich, C., Haase, C., Spree, K., Hehr, U., Hoffmann, M., Ledig, S., Sel, S.,
846 Wieacker, P., et al. (2009). WNT10A Mutations Are a Frequent Cause of a Broad Spectrum

of Ectodermal Dysplasias with Sex-Biased Manifestation Pattern in Heterozygotes. Am. J. Hum. Genet. *85*, 97–105.

Botchkarev, V.A., Botchkareva, N.V., Roth, W., Nakamura, M., Chen, L.H., Herzog, W., Lindner, G., McMahon, J.A., Peters, C., Lauster, R., et al. (1999). Noggin is a mesenchymally derived stimulator of hair-follicle induction. Nat. Cell Biol. *1*, 158–164.

Calle-Mustienes, E. de la, Feijóo, C.G., Manzanares, M., Tena, J.J., Rodríguez-Seguel, E., Letizia, A., Allende, M.L., and Gómez-Skarmeta, J.L. (2005). A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. Genome Res. *15*, 1061–1072.

Carroll, S.B. (2008). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell *134*, 25–36.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. gkr218.

Chai, Y., Mah, A., Crohin, C., Groff, S., Bringas, P., Le, T., Santos, V., and Slavkin, H.C. (1994). Specific transforming growth factor-beta subtypes regulate embryonic mouse Meckel's cartilage and tooth development. Dev. Biol. *162*, 85–103.

Chai, Y., Zhao, J., Mogharei, A., Xu, B., Bringas Jr., P., Shuler, C., and Warburton, D. (1999). Inhibition of transforming growth factor-β type II receptor signaling accelerates tooth formation in mouse first branchial arch explants. Mech. Dev. *86*, 63–74.

Chandler, R.L., Chandler, K.J., McFarland, K.A., and Mortlock, D.P. (2007). Bmp2 Transcription in Osteoblast Progenitors Is Regulated by a Distant 3′ Enhancer Located 156.3 Kilobases from the Promoter. Mol. Cell. Biol. *27*, 2934–2951.

Chen, Y., Bei, M., Woo, I., Satokata, I., and Maas, R. (1996). Msx1 controls inductive signaling in mammalian tooth morphogenesis. Development *122*, 3035–3044.

Cleves, P.A., Ellis, N.A., Jimenez, M.T., Nunez, S.M., Schluter, D., Kingsley, D.M., and Miller, C.T. (2014). Evolved tooth gain in sticklebacks is associated with a cis-regulatory allele of Bmp6. Proc. Natl. Acad. Sci. *111*, 13912–13917.

Dassule, H.R., and McMahon, A.P. (1998). Analysis of Epithelial–Mesenchymal Interactions in the Initial Morphogenesis of the Mammalian Tooth. Dev. Biol. *202*, 215–227.

Dathe, K., Kjaer, K.W., Brehm, A., Meinecke, P., Nürnberg, P., Neto, J.C., Brunoni, D., Tommerup, N., Ott, C.E., Klopocki, E., et al. (2009). Duplications Involving a Conserved Regulatory Element Downstream of BMP2 Are Associated with Brachydactyly Type A2. Am. J. Hum. Genet. *84*, 483–492.

881 Dendooven, A., van Oostrom, O., van der Giezen, D.M., Willem Leeuwis, J., Snijckers, C., Joles,
882 J.A., Robertson, E.J., Verhaar, M.C., Nguyen, T.Q., and Goldschmeding, R. (2011). Loss of
883 Endogenous Bone Morphogenetic Protein-6 Aggravates Renal Fibrosis. Am. J. Pathol. *178*,
884 1069–1079.

885 Doyle, E.L., Booher, N.J., Standage, D.S., Voytas, D.F., Brendel, V.P., VanDyk, J.K., and
886 Bogdanove, A.J. (2012). TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL
887 effector design and target prediction. Nucleic Acids Res. *40*, W117–W122.

888 Dudley, A.T., Godin, R.E., and Robertson, E.J. (1999). Interaction between FGF and BMP
889 signaling pathways regulates development of metanephric mesenchyme. Genes Dev. *13*,
890 1601–1613.

891 Farre, D., Roset, R., Huerta, M., Adsuara, J.E., Rosello, L., Alba, M.M., and Messeguer, X.
892 (2003). Identification of patterns in biological sequences at the ALGGEN server: PROMO
893 and MALGEN. Nucleic Acids Res. *31*, 3651–3653.

894 Feng, J.Q., Zhang, J., Tan, X., Lu, Y., Guo, D., and Harris, S.E. (2002). Identification of Cis-DNA
895 Regions Controlling Bmp4 Expression during Tooth Morphogenesis in vivo. J. Dent. Res. *81*,
896 6–10.

897 Ferguson, C.A., Tucker, A.S., Christensen, L., Lau, A.L., Matzuk, M.M., and Sharpe, P.T. (1998).
898 Activin is an essential early mesenchymal signal in tooth development that is required for
899 patterning of the murine dentition. Genes Dev. *12*, 2636–2649.

900 Ferguson, C.A., Tucker, A.S., Heikinheimo, K., Nomura, M., Oh, P., Li, E., and Sharpe, P.T.
901 (2001). The role of effectors of the activin signalling pathway, activin receptors IIA and IIB,
902 and Smad2, in patterning of tooth. Development *128*, 4605–4613.

903 Fisher, S., Grice, E.A., Vinton, R.M., Bessling, S.L., Urasaki, A., Kawakami, K., and McCallion,
904 A.S. (2006). Evaluating the biological relevance of putative enhancers using Tol2
905 transposon-mediated transgenesis in zebrafish. Nat. Protoc. *1*, 1297–1305.

906 Fraser, G.J., Bloomquist, R.F., and Streelman, J.T. (2013). Common developmental pathways
907 link tooth shape to regeneration. Dev. Biol. *377*, 399–414.

908 Freitas, R., Zhang, G., and Cohn, M.J. (2006). Evidence that mechanisms of fin development
909 evolved in the midline of early vertebrates. Nature *442*, 1033–1037.

910 Fujimori, S., Novak, H., Weissenböck, M., Jussila, M., Gonçalves, A., Zeller, R., Galloway, J.,
911 Thesleff, I., and Hartmann, C. (2010). Wnt/β-catenin signaling in the dental mesenchyme
912 regulates incisor development by regulating Bmp4. Dev. Biol. *348*, 97–106.

913 Gaete, M., and Tucker, A.S. (2013). Organized Emergence of Multiple-Generations of Teeth
914 in Snakes Is Dysregulated by Activation of Wnt/Beta-Catenin Signalling. PLoS ONE *8*,
915 e74484.

916 Genderen, C. van, Okamura, R.M., Fariñas, I., Quo, R.G., Parslow, T.G., Bruhn, L., and
917 Grosschedl, R. (1994). Development of several organs that require inductive epithelial-
918 mesenchymal interactions is impaired in LEF-1-deficient mice. Genes Dev. *8*, 2691–2703.

919 Gordon, K.L., and Ruvinsky, I. (2012). Tempo and Mode in Evolution of Transcriptional
920 Regulation. PLoS Genet *8*, e1002432.

921 Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V.,
922 Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., et al. (2008). Many
923 sequence variants affecting diversity of adult human height. Nat. Genet. *40*, 609–615.

924 Guenther, C., Pantalena-Filho, L., and Kingsley, D.M. (2008). Shaping Skeletal Growth by
925 Modular Regulatory Elements in the Bmp5 Gene. PLoS Genet. *4*.

926 Hogan, B.L. (1996). Bone morphogenetic proteins: multifunctional regulators of vertebrate
927 development. Genes Dev. *10*, 1580–1594.

928 Houlston, R.S., Webb, E., Broderick, P., Pittman, A.M., Di Bernardo, M.C., Lubbe, S., Chandler,
929 I., Vijayakrishnan, J., Sullivan, K., Penegar, S., et al. (2008). Meta-analysis of genome-wide
930 association data identifies four new susceptibility loci for colorectal cancer. Nat. Genet. *40*,
931 1426–1435.

932 Huang, S.-M.A., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O.,
933 Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and
934 antagonizes Wnt signalling. Nature *461*, 614–620.

935 Inman, G.J., Nicolás, F.J., Callahan, J.F., Harling, J.D., Gaster, L.M., Reith, A.D., Laping, N.J., and
936 Hill, C.S. (2002). SB-431542 Is a Potent and Specific Inhibitor of Transforming Growth
937 Factor-β Superfamily Type I Activin Receptor-Like Kinase (ALK) Receptors ALK4, ALK5,
938 and ALK7. Mol. Pharmacol. *62*, 65–74.

939 Jackman, W.R., and Stock, D.W. (2006). Transgenic analysis of Dlx regulation in fish tooth
940 development reveals evolutionary retention of enhancer function despite organ loss. Proc.
941 Natl. Acad. Sci. *103*, 19390–19395.

942 Jackman, W.R., Davies, S.H., Lyons, D.B., Stauder, C.K., Denton-Schneider, B.R., Jowdry, A.,
943 Aigler, S.R., Vogel, S.A., and Stock, D.W. (2013). Manipulation of Fgf and Bmp signaling in
944 teleost fishes suggests potential pathways for the evolutionary origin of multicuspid teeth.
945 Evol. Dev. *15*, 107–118.

946 Jumlongras, D., Lachke, S.A., O'Connell, D.J., Aboukhalil, A., Li, X., Choe, S.E., Ho, J.W.K.,
947 Turbe-Doan, A., Robertson, E.A., Olsen, B.R., et al. (2012). An Evolutionarily Conserved
948 Enhancer Regulates Bmp4 Expression in Developing Incisor and Limb Bud. PLoS ONE *7*,
949 e38568.

950 Jung, H.-S., Francis-West, P.H., Widelitz, R.B., Jiang, T.-X., Ting-Berreth, S., Tickle, C., Wolpert,
951 L., and Chuong, C.-M. (1998). Local Inhibitory Action of BMPs and Their Relationships with

952 Activators in Feather Formation: Implications for Periodic Patterning. Dev. Biol. *196*, 11–
953 23.

954 Justice, C.M., Yagnik, G., Kim, Y., Peter, I., Jabs, E.W., Erazo, M., Ye, X., Ainehsazan, E., Shi, L.,
955 Cunningham, M.L., et al. (2012). A genome-wide association study identifies susceptibility
956 loci for nonsyndromic sagittal craniosynostosis near BMP2 and within BBS9. Nat. Genet. *44*,
957 1360–1364.

958 Katzel, E.B., Wolenski, M., Loiselle, A.E., Basile, P., Flick, L.M., Langstein, H.N., Hilton, M.J.,
959 Awad, H.A., Hammert, W.C., and O'Keefe, R.J. (2011). Impact of Smad3 loss of function on
960 scarring and adhesion formation during tendon healing. J. Orthop. Res. Off. Publ. Orthop.
961 Res. Soc. *29*, 684–693.

962 Kavanagh, K.D., Evans, A.R., and Jernvall, J. (2007). Predicting evolutionary patterns of
963 mammalian teeth from development. Nature *449*, 427–432.

964 Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N., and Mishina, M.
965 (2004). A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated
966 Genes in Zebrafish. Dev. Cell *7*, 133–144.

967 Kingsley, D.M. (1994). What do BMPs do in mammals? Clues from the mouse short-ear
968 mutation. Trends Genet. *10*, 16–21.

969 Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H.,
970 Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version
971 2.0. Bioinformatics *23*, 2947–2948.

972 Laurikkala, J., Mikkola, M., Mustonen, T., Åberg, T., Koppinen, P., Pispa, J., Nieminen, P.,
973 Galceran, J., Grosschedl, R., and Thesleff, I. (2001). TNF Signaling via the Ligand–Receptor
974 Pair Ectodysplasin and Edar Controls the Function of Epithelial Signaling Centers and Is
975 Regulated by Wnt and Activin during Tooth Organogenesis. Dev. Biol. *229*, 443–455.

976 Li, T.-F., Darowish, M., Zuscik, M.J., Chen, D., Schwarz, E.M., Rosier, R.N., Drissi, H., and
977 O'Keefe, R.J. (2006). Smad3-deficient chondrocytes have enhanced BMP signaling and
978 accelerated differentiation. J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. *21*, 4–16.

979 Liu, F., Chu, E.Y., Watt, B., Zhang, Y., Gallant, N.M., Andl, T., Yang, S.H., Lu, M.-M., Piccolo, S.,
980 Schmidt-Ullrich, R., et al. (2008). Wnt/β-catenin signaling directs multiple stages of tooth
981 morphogenesis. Dev. Biol. *313*, 210–224.

982 Liu, W., Sun, X., Braut, A., Mishina, Y., Behringer, R.R., Mina, M., and Martin, J.F. (2005).
983 Distinct functions for Bmp signaling in lip and palate fusion in mice. Development *132*,
984 1453–1461.

985 Lubbe, S.J., Pittman, A.M., Olver, B., Lloyd, A., Vijayakrishnan, J., Naranjo, S., Dobbins, S.,
986 Broderick, P., Gómez-Skarmeta, J.L., and Houlston, R.S. (2012). The 14q22.2 colorectal
987 cancer variant rs4444235 shows cis-acting regulation of BMP4. Oncogene *31*, 3777–3784.

Marinić, M., Aktas, T., Ruf, S., and Spitz, F. (2013). An Integrated Holo-Enhancer Unit Defines Tissue and Gene Specificity of the Fgf8 Regulatory Landscape. Dev. Cell *24*, 530–542.

Massagué, J. (2012). TGFβ signalling in context. Nat. Rev. Mol. Cell Biol. *13*, 616–630.

Messeguer, X., Escudero, R., Farré, D., Núñez, O., Martínez, J., and Albà, M.M. (2002). PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics *18*, 333–334.

Mou, C., Jackson, B., Schneider, P., Overbeek, P.A., and Headon, D.J. (2006). Generation of the primary hair follicle pattern. Proc. Natl. Acad. Sci. *103*, 9075–9080.

Mou, C., Pitel, F., Gourichon, D., Vignoles, F., Tzika, A., Tato, P., Yu, L., Burt, D.W., Bed'hom, B., Tixier-Boichard, M., et al. (2011). Cryptic Patterning of Avian Skin Confers a Developmental Facility for Loss of Neck Feathering. PLoS Biol *9*, e1001028.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M., and Smith, W.L. (2012). Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. *109*, 13698–13703.

Neubüser, A., Peters, H., Balling, R., and Martin, G.R. (1997). Antagonistic Interactions between FGF and BMP Signaling Pathways: A Mechanism for Positioning the Sites of Tooth Formation. Cell *90*, 247–255.

Newburger, D.E., and Bulyk, M.L. (2009). UniPROBE: an online database of protein binding microarray data on protein–DNA interactions. Nucleic Acids Res. *37*, D77–D82.

Nie, X., Luukko, K., and Kettunen, P. (2006). BMP signalling in craniofacial development. Int. J. Dev. Biol. *50*.

O'Connell, D.J., Ho, J.W.K., Mammoto, T., Turbe-Doan, A., O'Connell, J.T., Haseley, P.S., Koo, S., Kamiya, N., Ingber, D.E., Park, P.J., et al. (2012). A Wnt-bmp feedback circuit controls intertissue signaling dynamics in tooth organogenesis. Sci. Signal. *5*, ra4.

Oka, S., Oka, K., Xu, X., Sasaki, T., Bringas Jr., P., and Chai, Y. (2007). Cell autonomous requirement for TGF-β signaling during odontoblast differentiation and dentin matrix formation. Mech. Dev. *124*, 409–415.

Oshima, M., Oshima, H., and Taketo, M.M. (1996). TGF-beta receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis. Dev. Biol. *179*, 297–302.

Perry, M.W., Boettiger, A.N., Bothma, J.P., and Levine, M. (2010). Shadow enhancers foster robustness of Drosophila gastrulation. Curr. Biol. CB *20*, 1562–1567.

Pregizer, S., and Mortlock, D.P. (2009). Control of BMP gene expression by long-range regulatory elements. Cytokine Growth Factor Rev. *20*, 509–515.

1021 Ross, M.T., LaBrie, S., McPherson, J., and Stanton, V.P. (1999). Screening large-insert
1022 libraries by hybridization. In Current Protocols in Human Genetics, N.C. Dracopoli, J.L.
1023 Haines, and B.. Korf, eds. (New York: John Wiley and Sons), pp. 5.6.1–5.6.52.

1024 Sanford, L.P., Ormsby, I., Groot, A.C.G., Sariola, H., Friedman, R., Boivin, G.P., Cardell, E.L., and
1025 Doetschman, T. (1997). TGFbeta2 knockout mice have multiple developmental defects that
1026 are non-overlapping with other TGFbeta knockout phenotypes. Development *124*, 2659–
1027 2670.

1028 Scheer, N., and Campos-Ortega, J.A. (1999). Use of the Gal4-UAS technique for targeted gene
1029 expression in the zebrafish. Mech. Dev. *80*, 153–158.

1030 Shi, M., Murray, J.C., Marazita, M.L., Munger, R.G., Ruczinski, I., Hetmanski, J.B., Wu, T.,
1031 Murray, T., Redett, R.J., Wilcox, A.J., et al. (2012). Genome wide study of maternal and
1032 parent-of-origin effects on the etiology of orofacial clefts. Am. J. Med. Genet. A. *158A*, 784–
1033 794.

1034 Shimizu, N., Kawakami, K., and Ishitani, T. (2012). Visualization and exploration of Tcf/Lef
1035 function using a highly responsive Wnt/β-catenin signaling-reporter transgenic zebrafish.
1036 Dev. Biol. *370*, 71–85.

1037 Solloway, M.J., Dudley, A.T., Bikoff, E.K., Lyons, K.M., Hogan, B.L., and Robertson, E.J. (1998).
1038 Mice lacking Bmp6 function. Dev. Genet. *22*, 321–339.

1039 Stern, D.L. (2000). Perspective: Evolutionary Developmental Biology and the Problem of
1040 Variation. Evolution *54*, 1079–1091.

1041 Sun, Z., Jin, P., Tian, T., Gu, Y., Chen, Y.-G., and Meng, A. (2006). Activation and roles of
1042 ALK4/ALK7-mediated maternal TGFβ signals in zebrafish embryo. Biochem. Biophys. Res.
1043 Commun. *345*, 694–703.

1044 Suster, M.L., Abe, G., Schouw, A., and Kawakami, K. (2011). Transposon-mediated BAC
1045 transgenesis in zebrafish. Nat. Protoc. *6*, 1998–2021.

1046 Swarup, H. (1958). Stages in the Development of the Stickleback Gasterosteus aculeatus
1047 (L.). J. Embryol. Exp. Morphol. *6*, 373–383.

1048 Urasaki, A., Morvan, G., and Kawakami, K. (2006). Functional Dissection of the Tol2
1049 Transposable Element Identified the Minimal cis-Sequence and a Highly Repetitive
1050 Sequence in the Subterminal Region Essential for Transposition. Genetics *174*, 639–649.

1051 Vainio, S., Karavanova, I., Jowett, A., and Thesleff, I. (1993). Identification of BMP-4 as a
1052 signal mediating secondary induction between epithelial and mesenchymal tissues during
1053 early tooth development. Cell *75*, 45–58.

1054 Villefranc, J.A., Amigo, J., and Lawson, N.D. (2007). Gateway compatible vectors for analysis
1055 of gene function in the zebrafish. Dev. Dyn. *236*, 3077–3087.

Wang, Y., Cox, M.K., Coricor, G., MacDougall, M., and Serra, R. (2013). Inactivation of Tgfbr2 in Osterix-Cre expressing dental mesenchyme disrupts molar root formation. Dev. Biol. *382*, 27–37.

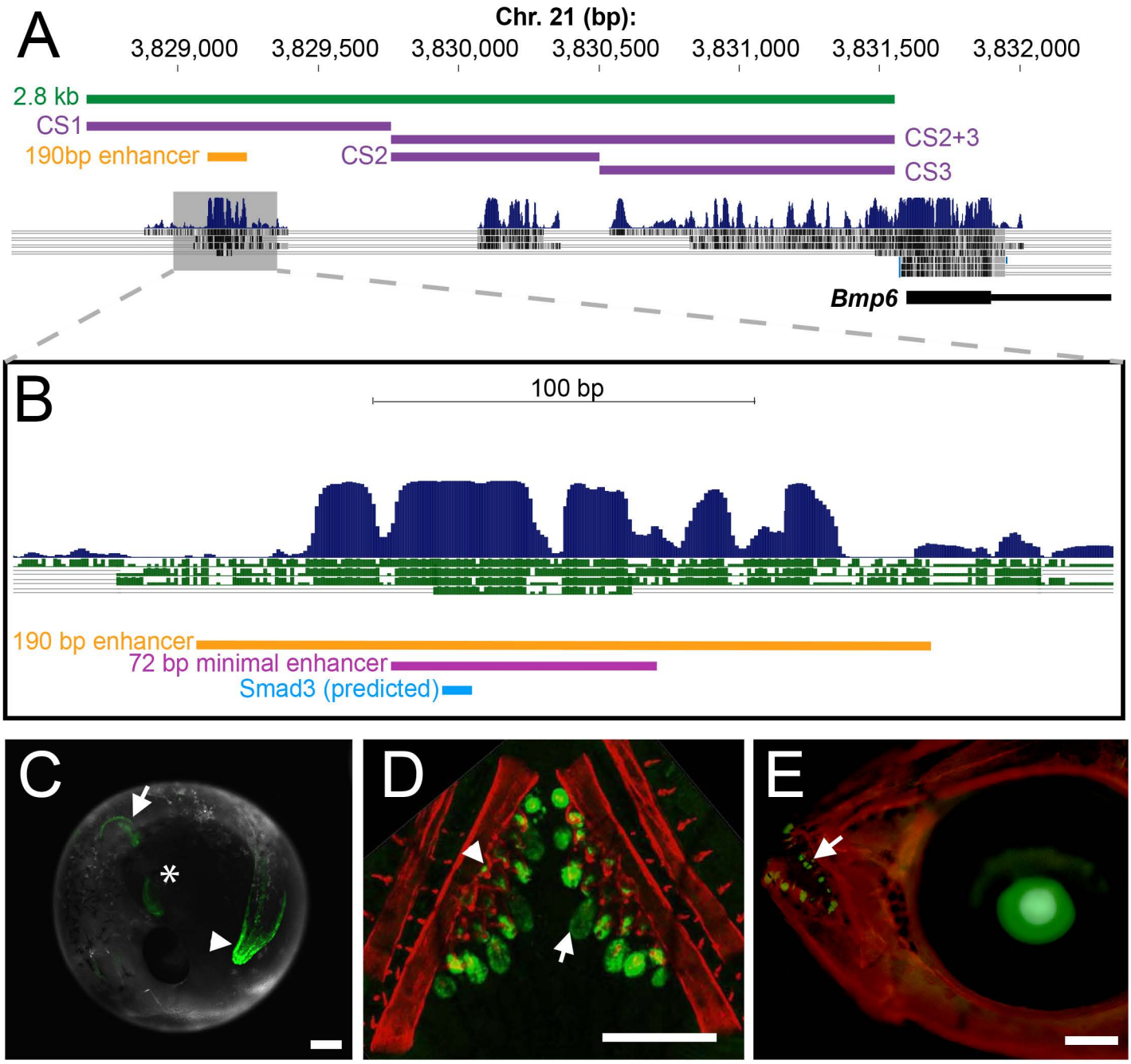
Westerfield, M. (2007). The Zebrafish Book: A guide for the Laboratory Use of Zebrafish (Danio rerio), 5th Edition (Eugene, OR: University of Oregon Press).

Wu, P., Wu, X., Jiang, T.-X., Elsey, R.M., Temple, B.L., Divers, S.J., Glenn, T.C., Yuan, K., Chen, M.-H., Widelitz, R.B., et al. (2013). Specialized stem cell niche enables repetitive renewal of alligator teeth. Proc. Natl. Acad. Sci. U. S. A. *110*, E2009–E2018.

Zhao, H., Oka, K., Bringas, P., Kaartinen, V., and Chai, Y. (2008). TGF-β type I receptor Alk5 regulates tooth initiation and mandible patterning in a type II receptor-independent manner. Dev. Biol. *320*, 19–29.

Zhao, X., Zhang, Z., Song, Y., Zhang, X., Zhang, Y., Hu, Y., Fromm, S.H., and Chen, Y. (2000). Transgenically ectopic expression of Bmp4 to the Msx1 mutant dental mesenchyme restores downstream gene expression but represses Shh and Bmp2 in the enamel knot of wild type tooth germ. Mech. Dev. *99*, 29–38.

Figure 1


A
Chr. 21 (bp):
3,829,000
3,829,500
3,830,000
3,830,500
3,831,000
3,831,500
3,832,000
2.8 kb
CS1
190bp enhancer
CS2
CS2+3
CS3
Bmp6
B
100 bp
190 bp enhancer
72 bp minimal enhancer
Smad3 (predicted)

C

D

E

Figure 2

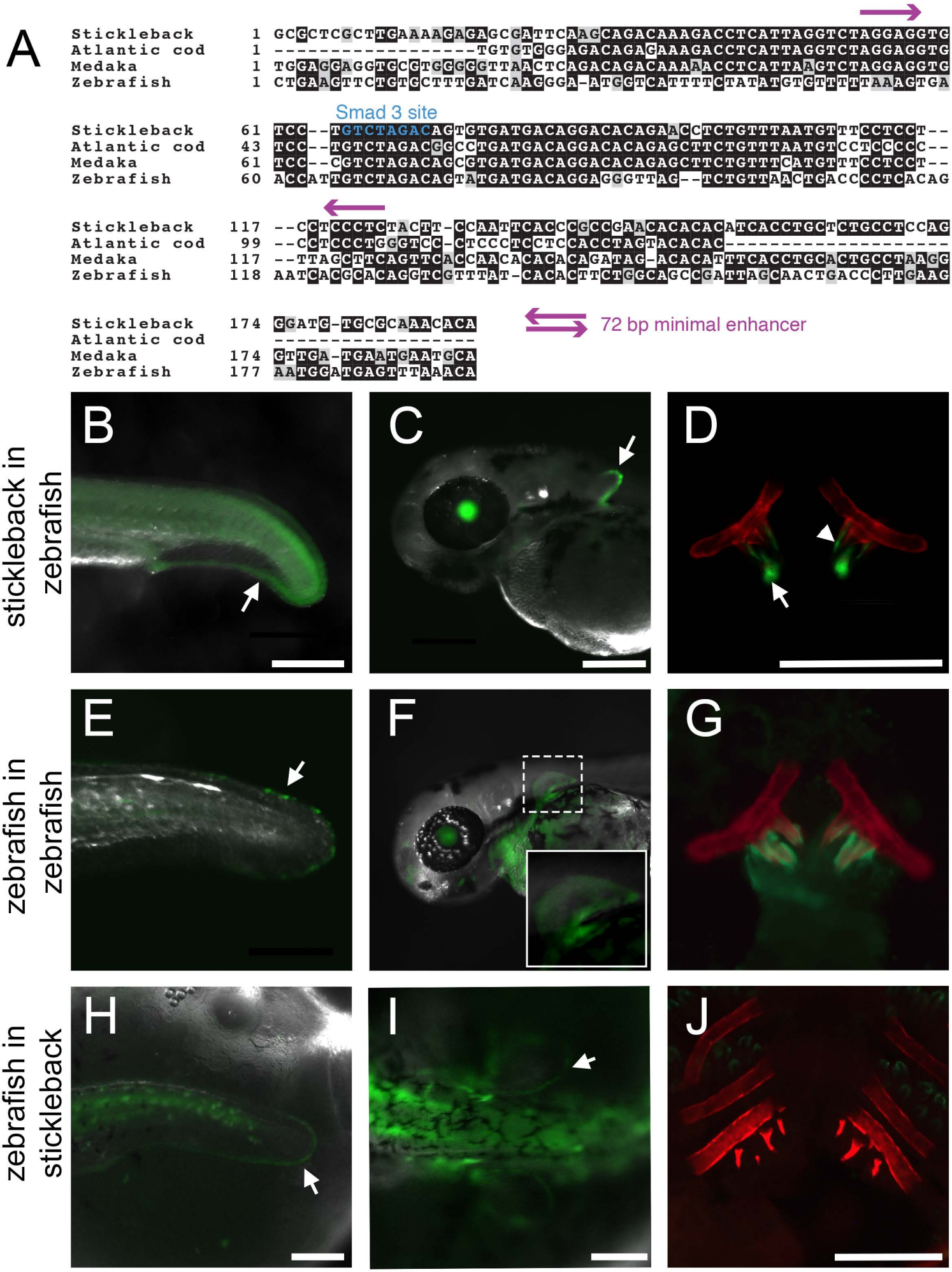

A
Stickleback 1 GCGCTCGCTTGAAAAGAGAGCGATTCAAGCAGACAAAGACCTCATTAGGTCTAGGAGGTG
Atlantic cod 1 ----------TGTGTGGGAGACAGAGAAAGACCTCATTAGGTCTAGGAGGTG
Medaka 1 TGGAGGAGGTGCGTGGGGGTTAACTCAGACAGACAAAAACCTCATTAAGTCTAGGAGGTG
Zebrafish 1 CTGAAGTTCTGTGCTTTGATCAAGGGA-ATGGTCATTTTCTATATGTGTTTTTAAAGTGA
Smad 3 site
Stickleback 61 TCC--TGTCTAGACAGTGTGATGACAGGACACAGAACCTCTGTTTAATGTTTCCTCCT--
Atlantic cod 43 TCC--TGTCTAGACGGCCTGATGACAGGACACAGAGCTTCTGTTTAATGTCCTCCCC--
Medaka 61 TCC--CGTCTAGACAGCGTGATGACAGGACACAGAGCTTCTGTTTCATGTTTCCTCCT--
Zebrafish 60 ACCATTGTCTAGACAGTATGATGACAGGAGGGTTAG--TCTGTTAACTGACCCTCACAG
Stickleback 117 --CCTCCCTCTACTT-CCAATTCACCCGCCGAACACACACATCACCTGCTCTGCCTCCAG
Atlantic cod 99 --CCTCCCCTGGGTCC-CTCCCTCCTCCACCTAGTACACAC----------
Medaka 117 --TTAGCTTCAGTTCACCAACACACACAGATAG-ACACATTTCACCTGCACTGCCTAAGG
Zebrafish 118 AATCACGCACAGGTCGTTTAT-CACACTTCTGGCAGCCGATTAGCAACTGACCCTTGAAG
Stickleback 174 GGATG-TGCGCAAACACA
Atlantic cod ----------
Medaka 174 GTTGA-TGAATGAATGCA
Zebrafish 177 AATGGATGAGTTTAAACA
72 bp minimal enhancer
stickleback in zebrafish
zebrafish in zebrafish
zebrafish in stickleback
B
C
D
E
F
G
H
I
J

Figure 3

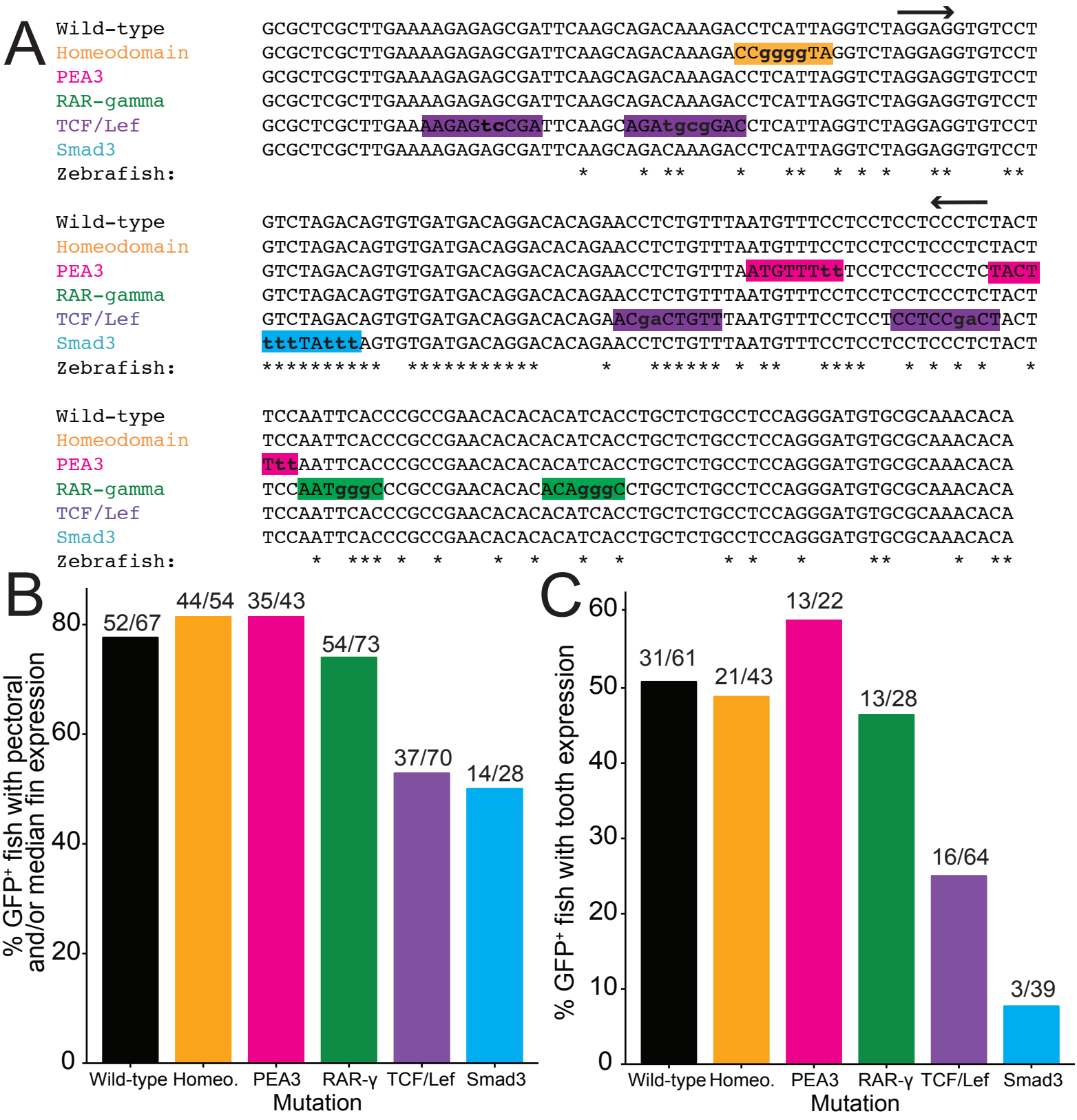

A
Wild-type
Homeodomain
PEA3
RAR-gamma
TCF/Lef
Smad3
Zebrafish:
GCGCTCGCTTGAAAAGAGAGCGATTCAAGCAGACAAAGACCTCATTAGGTCTAGGAGGTGTCCT
GCGCTCGCTTGAAAAGAGAGCGATTCAAGCAGACAAAGACCggggTAGGTCTAGGAGGTGTCCT
GCGCTCGCTTGAAAAGAGAGCGATTCAAGCAGACAAAGACCTCATTAGGTCTAGGAGGTGTCCT
GCGCTCGCTTGAAAAGAGAGCGATTCAAGCAGACAAAGACCTCATTAGGTCTAGGAGGTGTCCT
GCGCTCGCTTGAAAAGAGtcCGATTCAAGCAGAtgcgGACCTCATTAGGTCTAGGAGGTGTCCT
GCGCTCGCTTGAAAAGAGAGCGATTCAAGCAGACAAAGACCTCATTAGGTCTAGGAGGTGTCCT
GTCTAGACAGTGTGATGACAGGACACAGAACCTCTGTTTAATGTTTCCTCCTCCTCCCTCTACT
GTCTAGACAGTGTGATGACAGGACACAGAACCTCTGTTTAATGTTTCCTCCTCCTCCCTCTACT
GTCTAGACAGTGTGATGACAGGACACAGAACCTCTGTTTAATGTTTttTCCTCCTCCCTCTACT
GTCTAGACAGTGTGATGACAGGACACAGAACCTCTGTTTAATGTTTCCTCCTCCTCCCTCTACT
GTCTAGACAGTGTGATGACAGGACACAGAACgaCTGTTTAATGTTTCCTCCTCCTCCgaCTACT
tttTAtttAGTGTGATGACAGGACACAGAACCTCTGTTTAATGTTTCCTCCTCCTCCCTCTACT
TCCAATTCACCCGCCGAACACACACATCACCTGCTCTGCCTCCAGGGATGTGCGCAAACACA
TCCAATTCACCCGCCGAACACACACATCACCTGCTCTGCCTCCAGGGATGTGCGCAAACACA
TttAATTCACCCGCCGAACACACACATCACCTGCTCTGCCTCCAGGGATGTGCGCAAACACA
TCCAATgggCCGCCGAACACACACAgggCCTGCTCTGCCTCCAGGGATGTGCGCAAACACA
TCCAATTCACCCGCCGAACACACACATCACCTGCTCTGCCTCCAGGGATGTGCGCAAACACA
TCCAATTCACCCGCCGAACACACACATCACCTGCTCTGCCTCCAGGGATGTGCGCAAACACA
B
% GFP+ fish with pectoral and/or median fin expression
0
20
40
60
80
52/67
44/54
35/43
54/73
37/70
14/28
Wild-type
Homeo.
PEA3
RAR-γ
TCF/Lef
Smad3
Mutation
C
% GFP+ fish with tooth expression
0
10
20
30
40
50
60
31/61
21/43
13/22
13/28
16/64
3/39
Wild-type
Homeo.
PEA3
RAR-γ
TCF/Lef
Smad3
Mutation

Figure 4

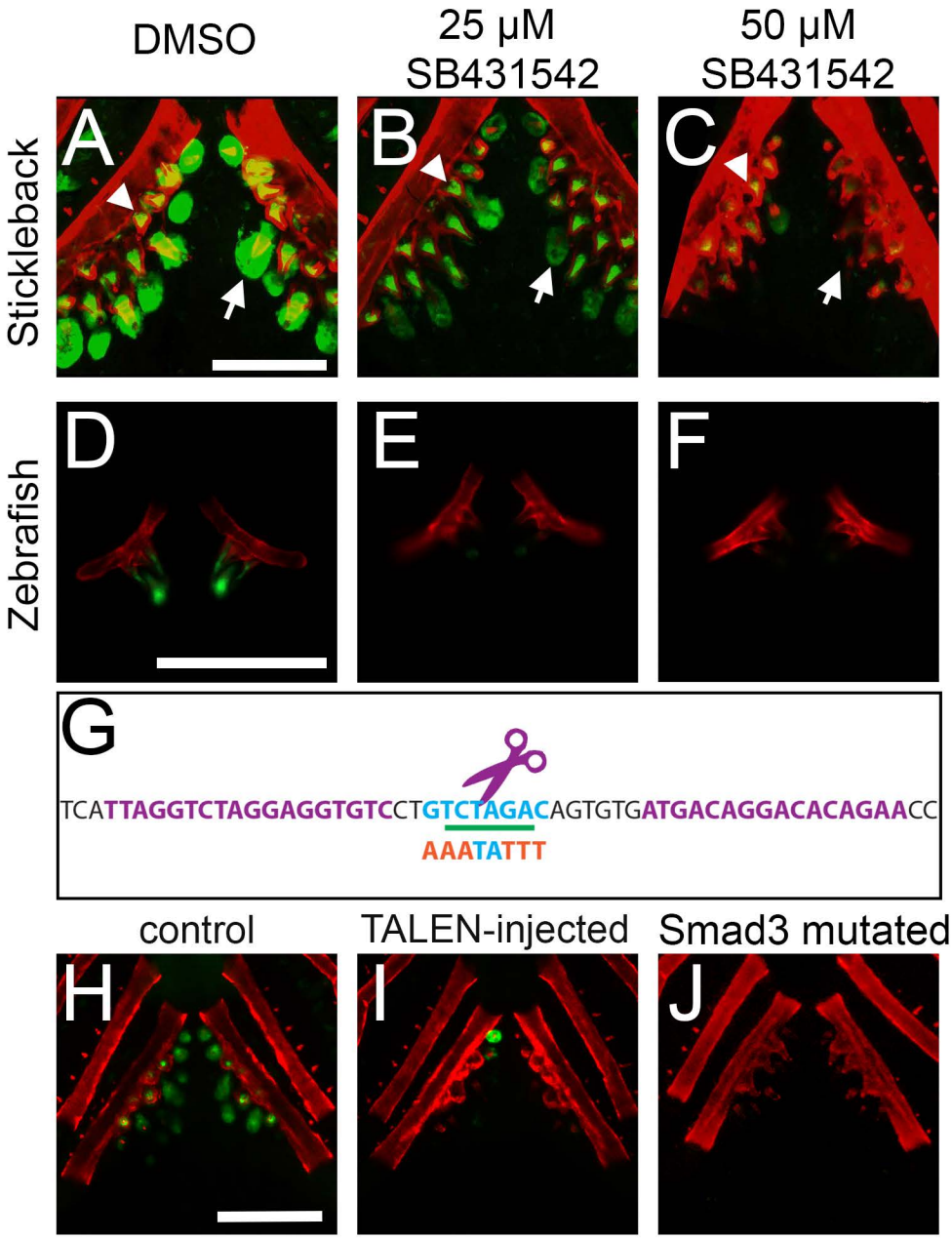

DMSO
25 μM SB431542
50 μM SB431542
Stickleback
Zebrafish
A
B
C
D
E
F
G
TCATTAGGTCTAGGAGGTGTCCTGTCTAGACAGTGTGATGACAGGACACAGAACC
AAATATTT
control
TALEN-injected
Smad3 mutated
H
I
J

Figure 5

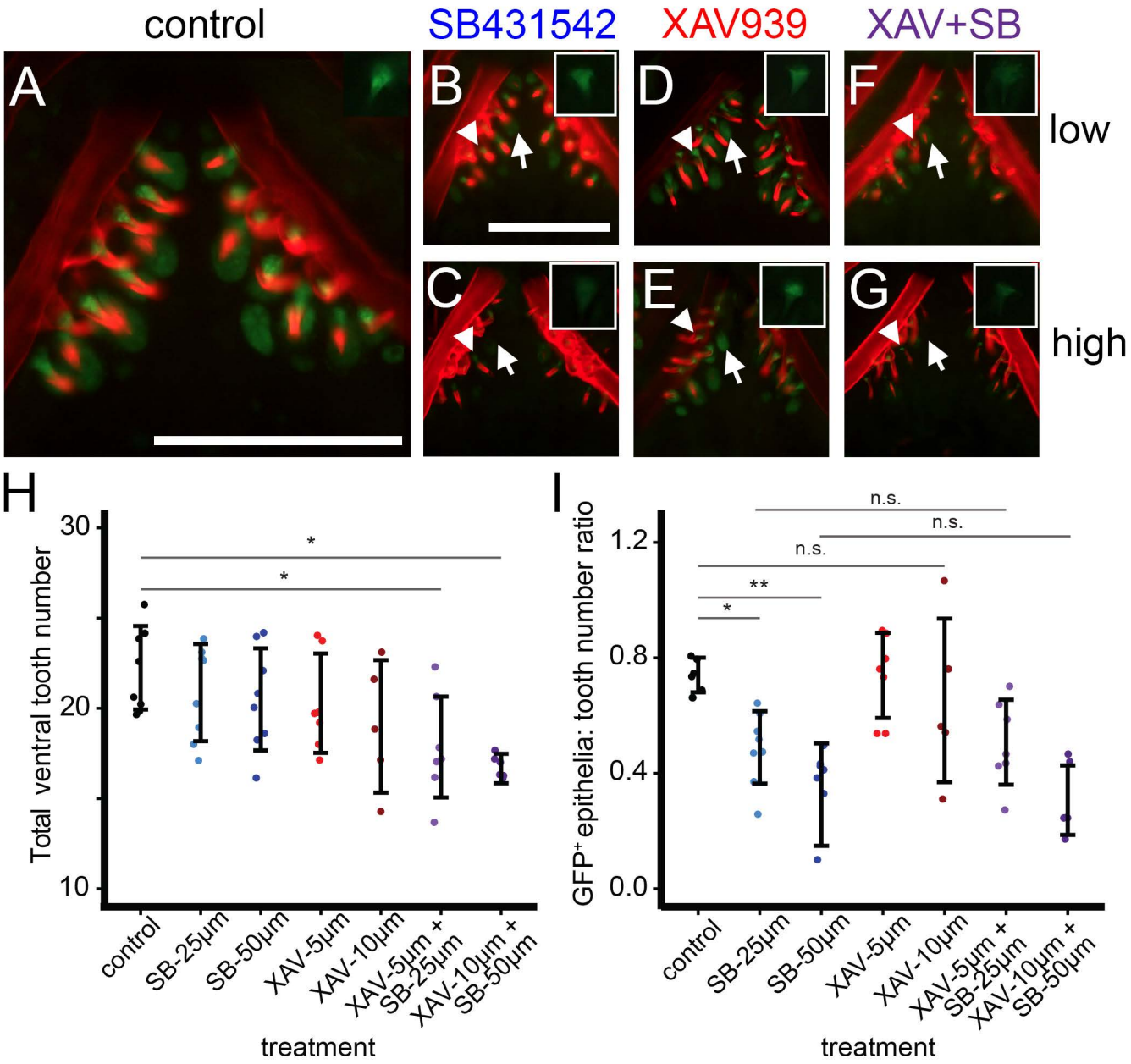

control
SB431542
XAV939
XAV+SB
A
B
C
D
E
F
G
low
high
H
Total ventral tooth number
30
20
10
*
*
control
SB-25µm
SB-50µm
XAV-5µm
XAV-10µm
XAV-5µm + SB-25µm
XAV-10µm + SB-50µm
treatment
I
GFP+ epithelia: tooth number ratio
1.2
0.8
0.4
0.0
n.s.
n.s.
n.s.
**
*
control
SB-25µm
SB-50µm
XAV-5µm
XAV-10µm
XAV-5µm + SB-25µm
XAV-10µm + SB-50µm
treatment

Figure 6

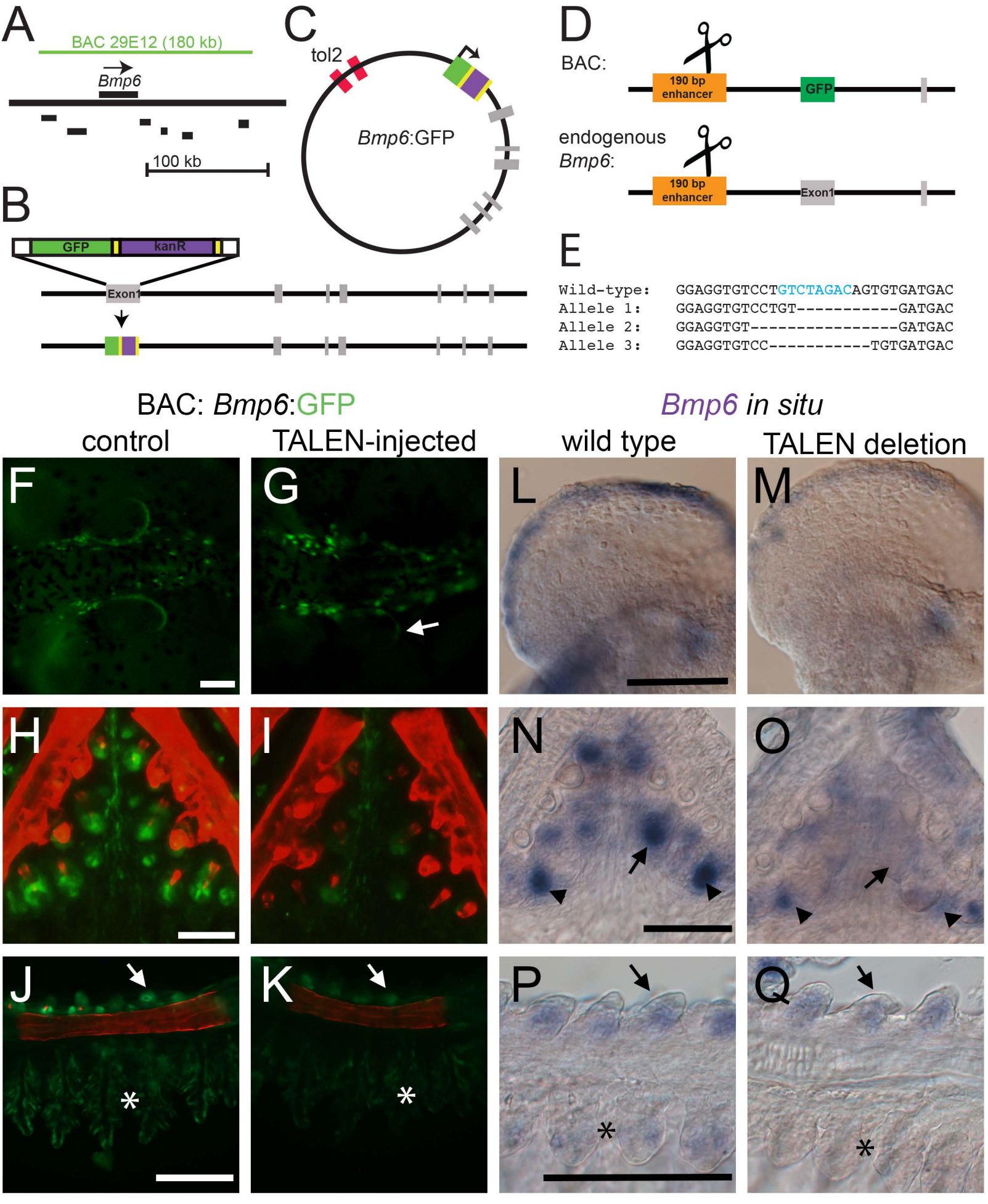

A
BAC 29E12 (180 kb)
Bmp6
100 kb
B
GFP
kanR
Exon1
C
tol2
Bmp6:GFP
D
BAC:
190 bp enhancer
GFP
endogenous Bmp6:
190 bp enhancer
Exon1
E
Wild-type: GGAGGTGTCCTGTCTAGACAGTGTGATGAC
Allele 1: GGAGGTGTCCTGT----------GATGAC
Allele 2: GGAGGTGT----------GATGAC
Allele 3: GGAGGTGTCC----------TGTGATGAC
BAC: Bmp6:GFP
control
TALEN-injected
Bmp6 in situ
wild type
TALEN deletion
F
G
L
M
H
I
N
O
J
K
P
Q

Figure 7

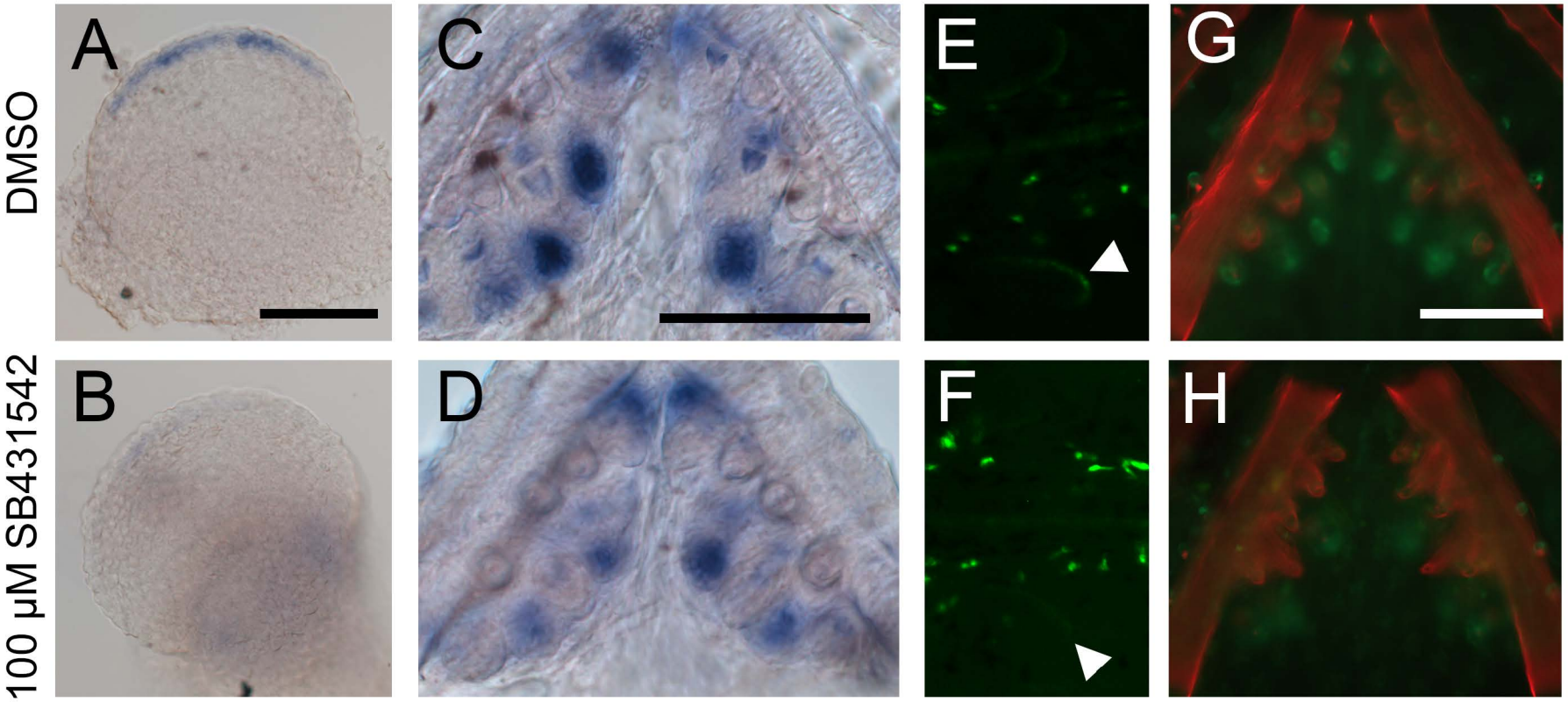


Figure S1

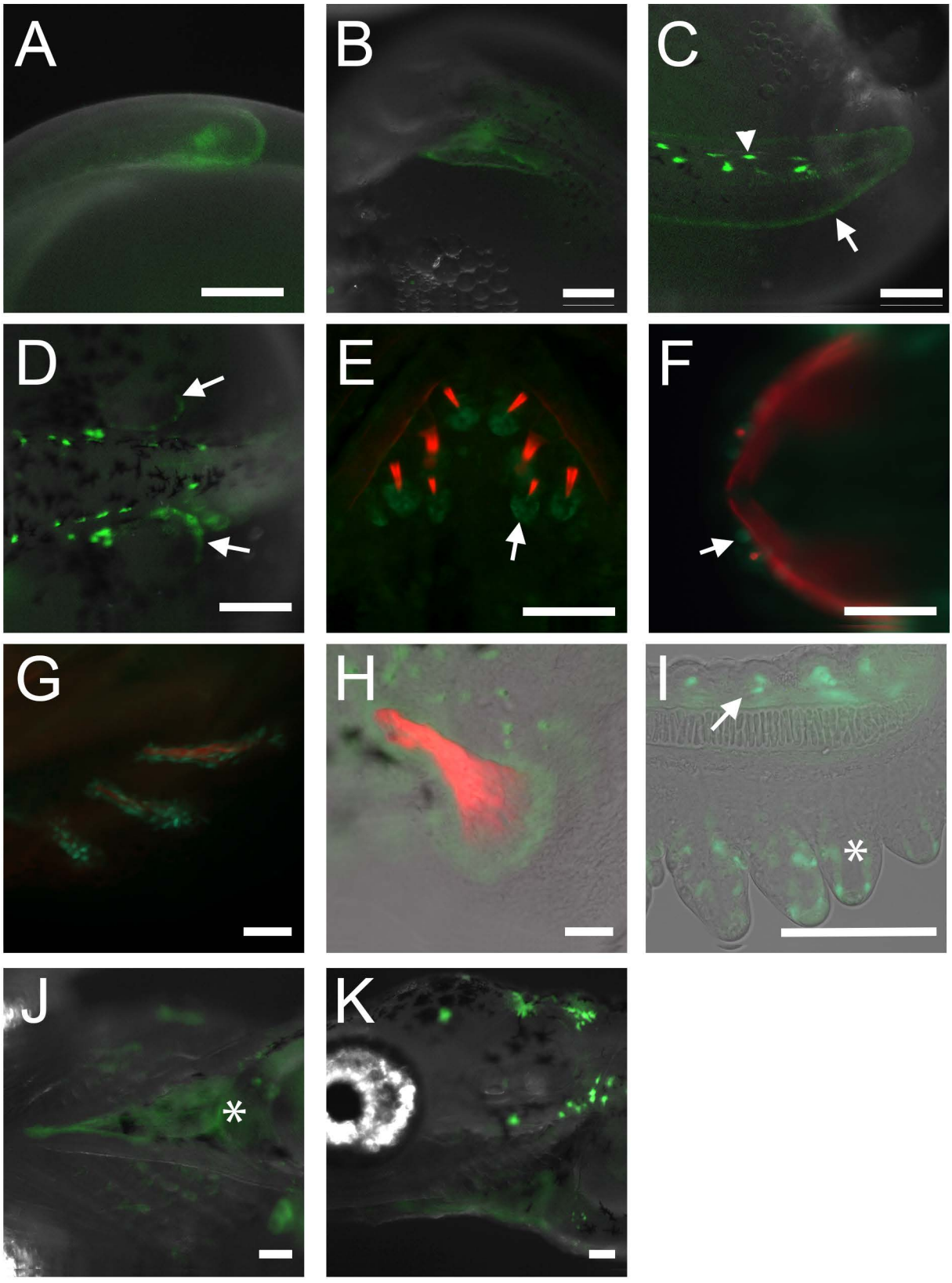


Figure S2

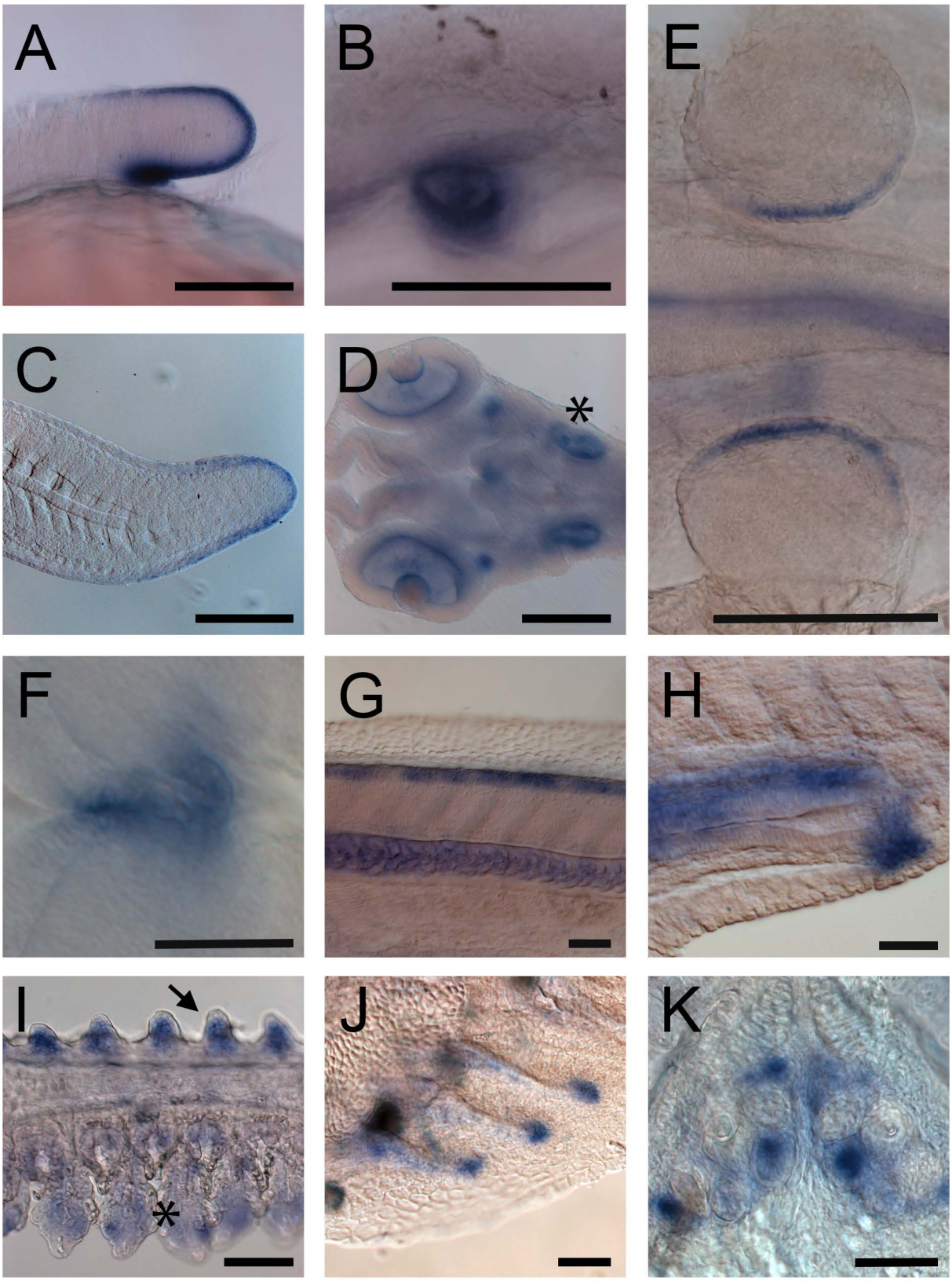


Figure S3

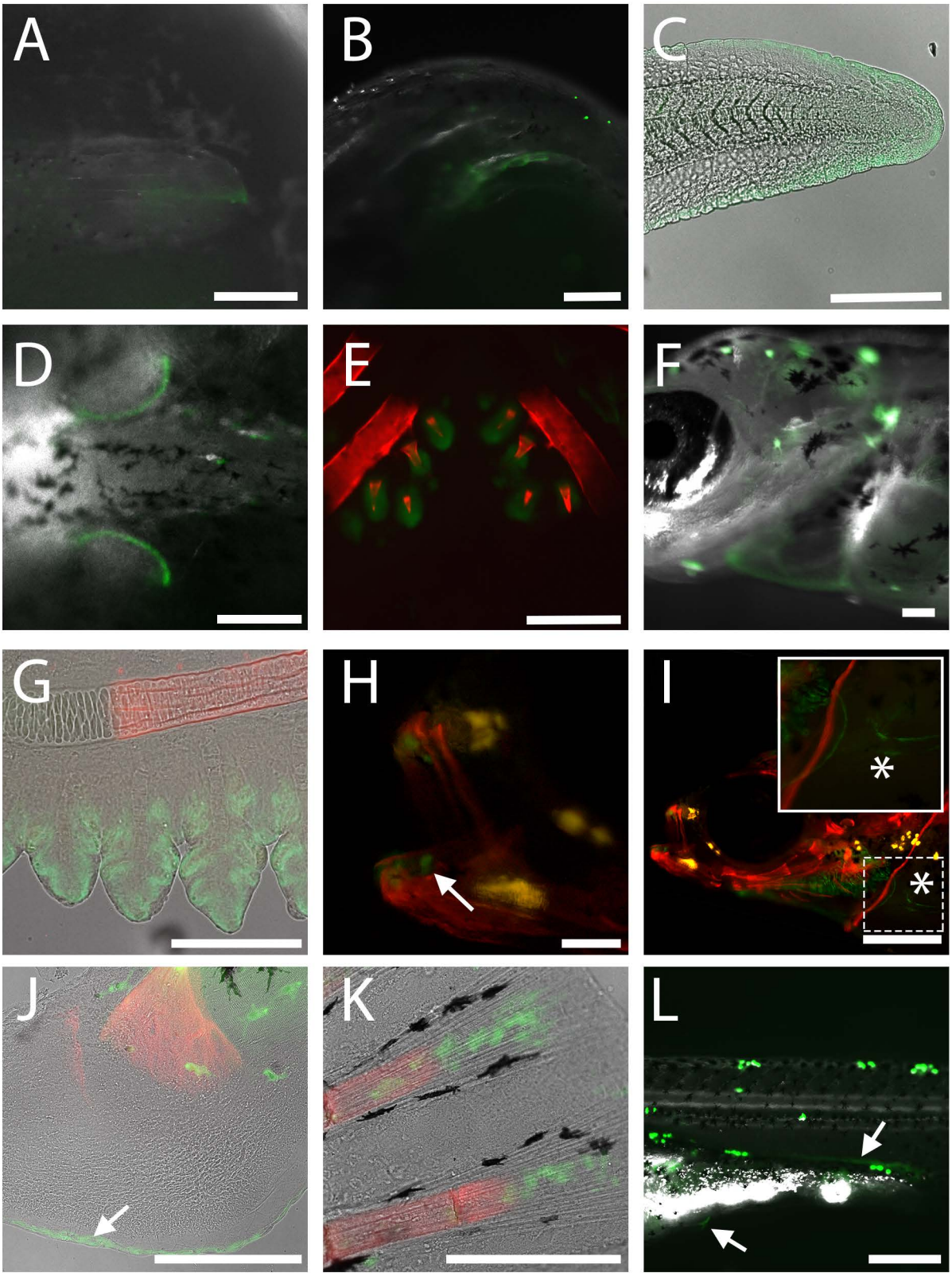


Figure S4

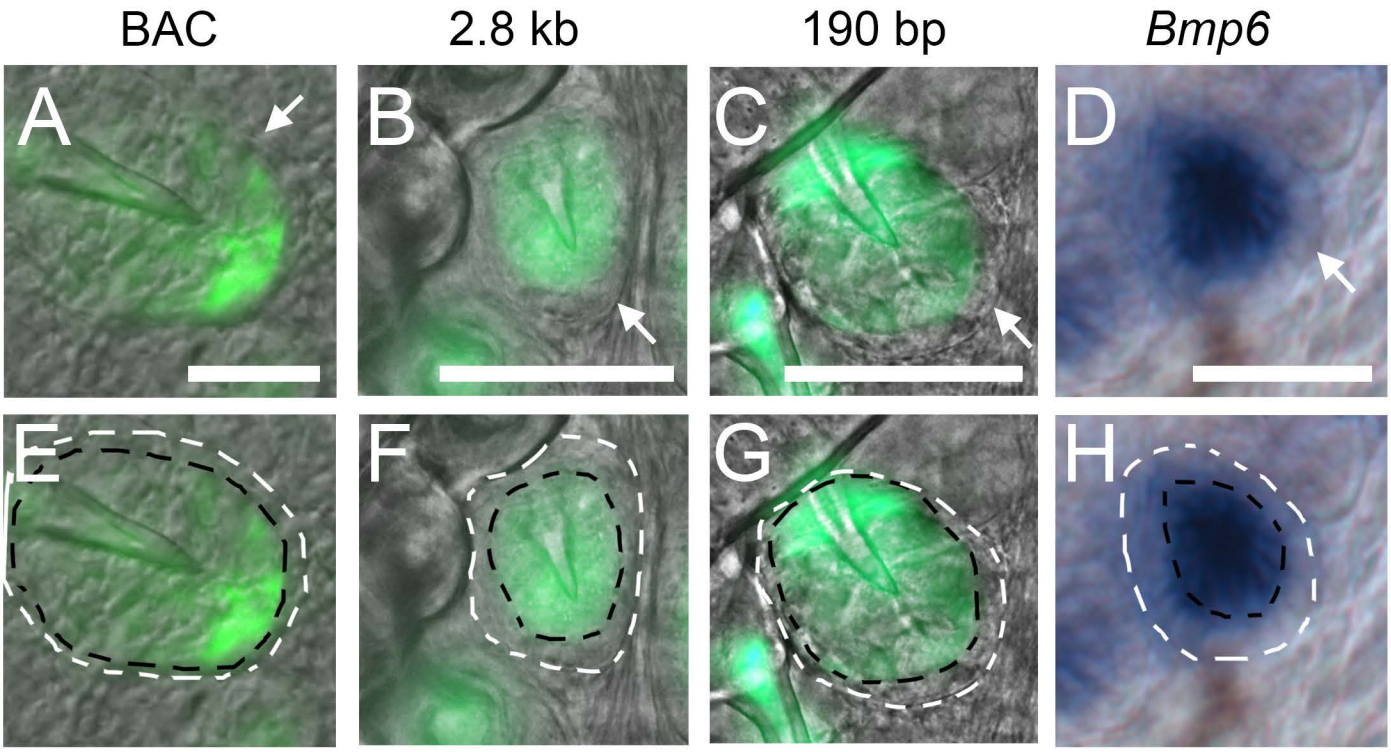


Figure S5

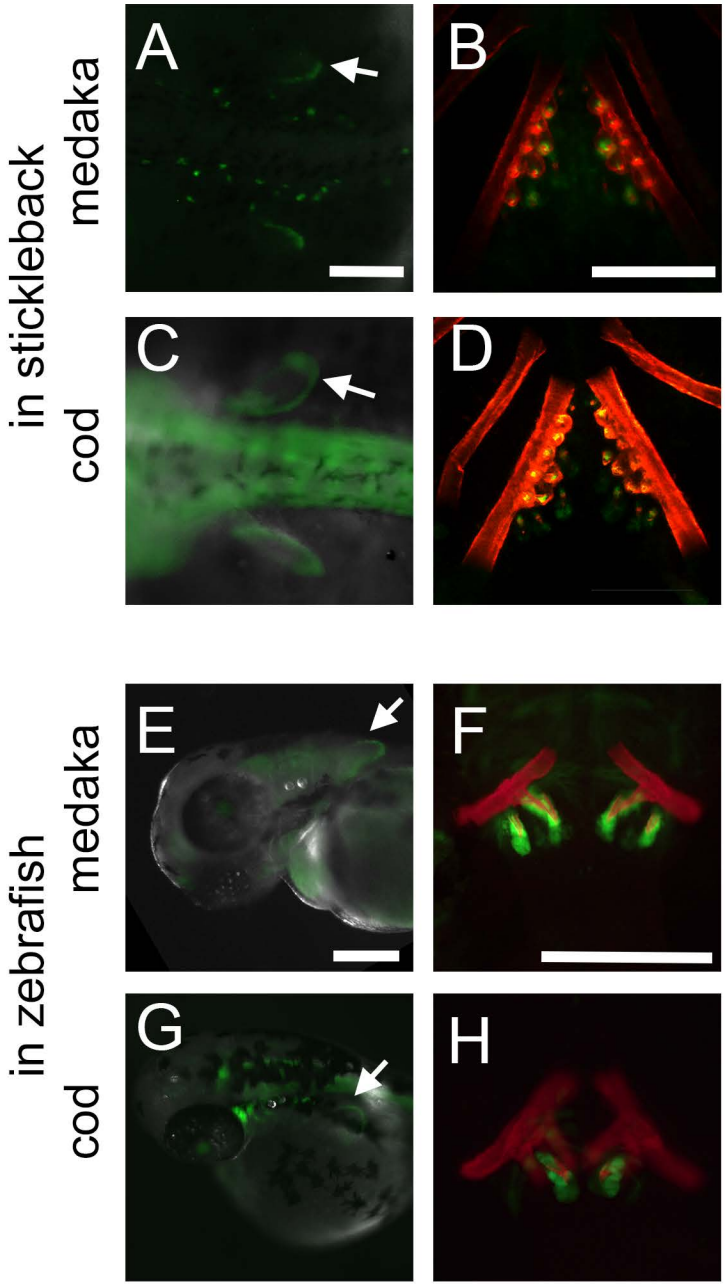

in stickleback
medaka
cod
in zebrafish
medaka
cod
A
B
C
D
E
F
G
H

Figure S6

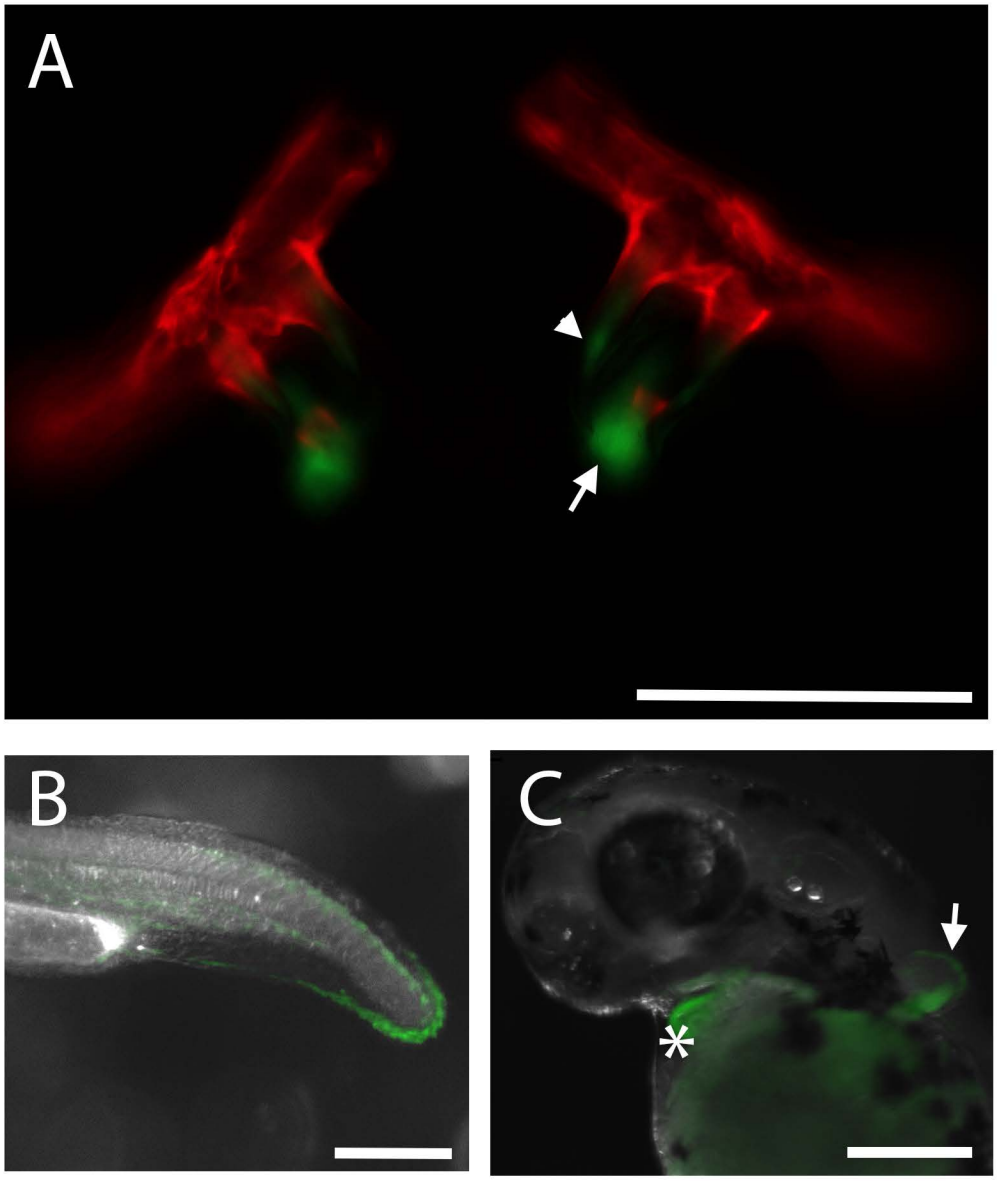


Figure S7

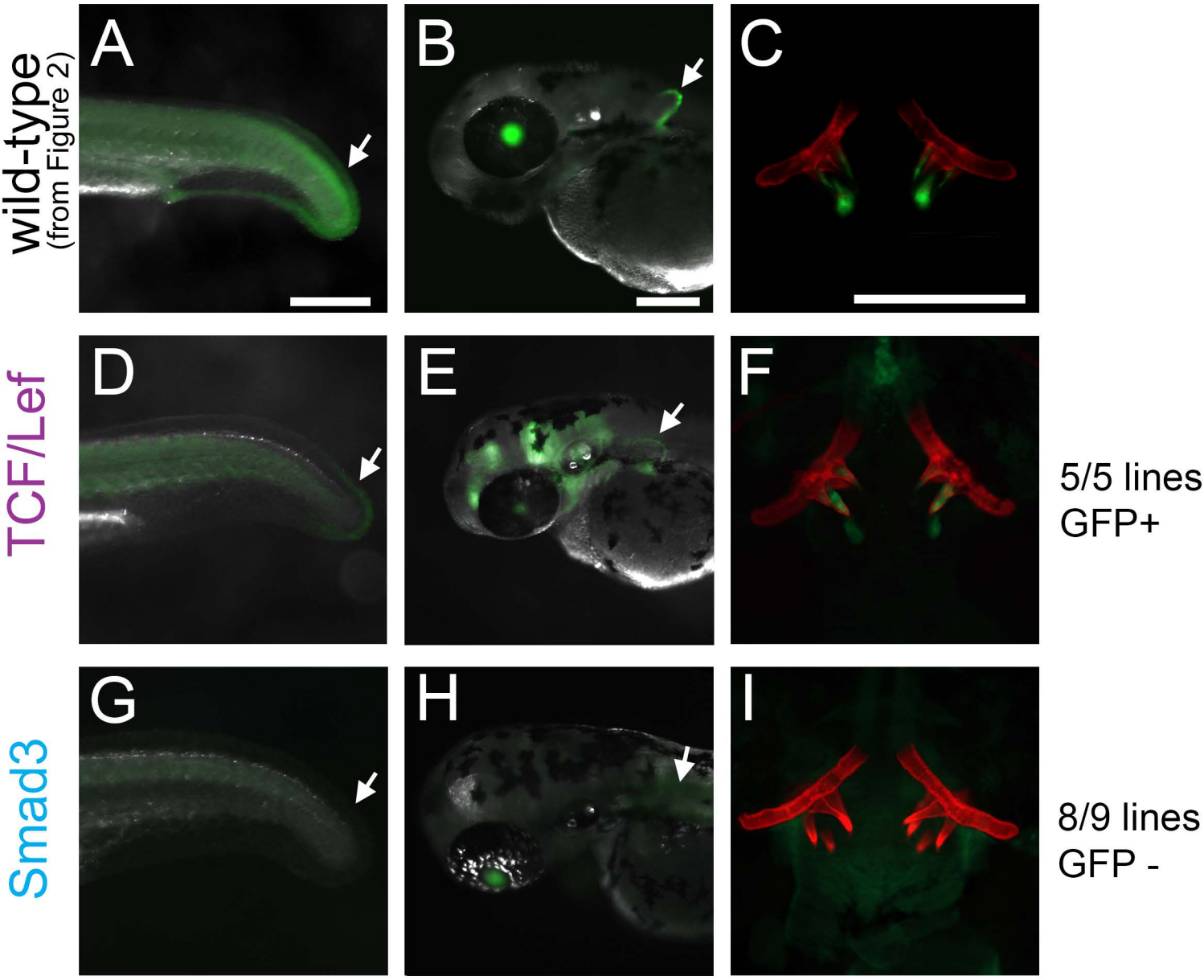

wild-type
(from Figure 2)
A
B
C
TCF/Lef
D
E
F
5/5 lines
GFP+
Smad3
G
H
I
8/9 lines
GFP -

Figure S8

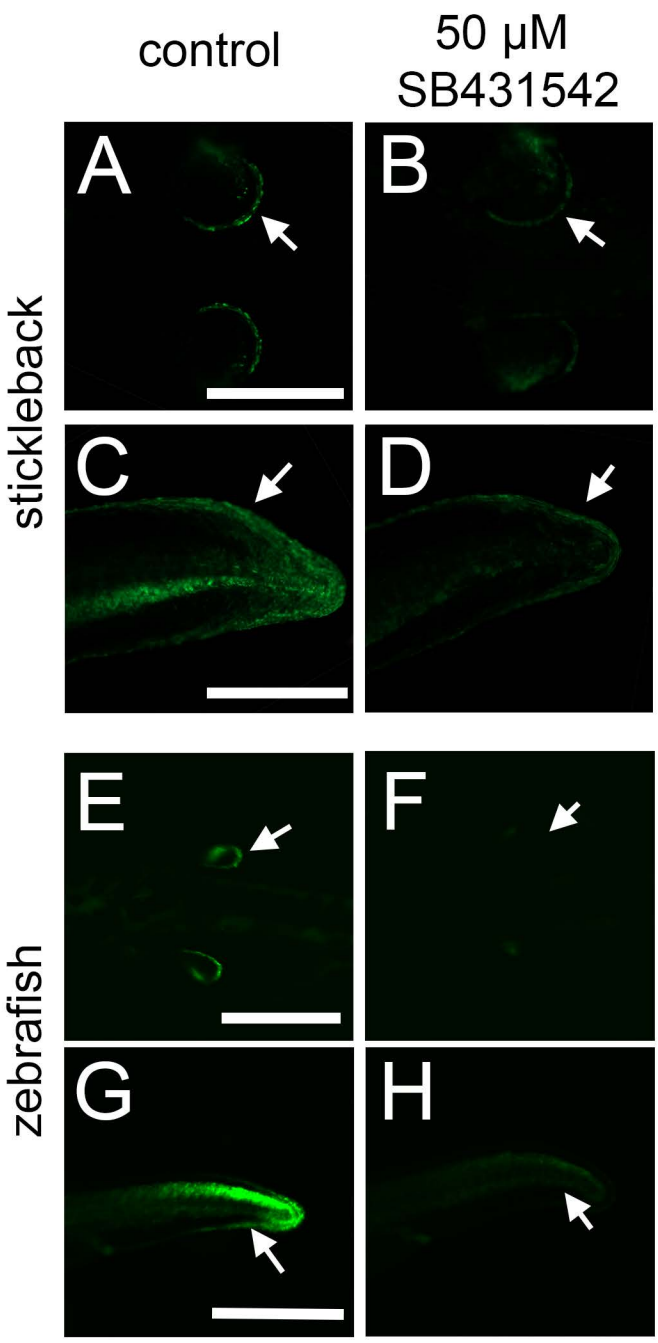


Table S1. Primers used to clone reporter constructs, perform site directed mutagenesis, and recombineer BACs.

Enhancer constructs

Primer Name	Sequence	Purpose
Gac_3kb_for	GCCGATCGATATAGGAAGGCTGGACAACGA	stickleback 3kb forward
Gac_3kb_rev	GCCGATCGATAGAACACAGCGGGGAAACACC	stickleback 3kb reverse
Gac_CS1_rev	GCCGATCGATAGTATGGTGCGTGTGTGCAT	stickleback CS1 reverse
Gac_CS2_for	GCCGATCGATATGCACACACGCACCATACT	stickleback CS2 forward
Gac_CS2_rev	GCCGATCGATGAAACAGCAAGCAATGACGA	stickleback CS2 reverse
Gac_CS3_for	GCCGATCGATTCGTCATTGCTTGCTGTTTC	stickleback CS3 forward
Gac_190_for	GCCGGCTAGCGCGCTCGCTTGAAAAGAGAGC	stickleback 190bp forward
Gac_190_rev	GCCGGGATCCTGTGTGTTTGCGCACATCCC	stickleback 190bp reverse
Gac_72_for	GCCGGCTAGCAGGAGGTGTCCTGTCTAGACA	stickleback 72bp forward
Gac_72_rev	GCCGGGATCCGAGGGAGGAGGAGGAAACATTAAA	stickleback 72bp_rev
Dre_for	GCCGGCTAGCCCCTGAAGTTCTGTGCTTTGATCA	zebrafish forward
Dre_rev	GCCGGGATCCAAGCTGGACATTCCTCTGCAAATG	zebrafish reverse
Gmo_for	GCCGGCTAGCTGTGTACTAGGTGGAGGAGGGAGGGACCCAGGGAGGGGG GGAGGACATT	cod forward
Gmo_temp1	GACGGCCTGATGACAGGACACAGAGCTTCTGTTTAATGTCCTCCCCCCC	cod template 1
Gmo_temp2	CTGTCATCAGGCCGTCTAGACAGGACACCTCCTAGACCTAATGAGGTC	cod template 2
Gmo_rev	GCCGGGATCCGTGTGGGAGACAGAGAAAGACCTCATTAGGTCTAGGAGG	cod reverse
Ola_for	AGTCGCTAGCAATGGAAGCAGTGTGGAGGAGG	medaka forward
Ola_rev	AGCTGGATCCGGCCCTAATCAGTTGTGTTCTGCA	medaka reverse

Mutagenesis constructs

Primer Name	Sequence	Purpose
Smad3_mut1_for	ATTAGGTCTAGGAGGTGTCCTAAATAGACAGTGTGATGACAGGAC	SMAD3 mut. first round forward
Smad3_mut1_rev	GTCCTGTCATCACACTGTCTATTTAGGACACCTCCTAGACCTAAT	SMAD3 mut. first round reverse
Smad3_mut2_for	GTCCTGTCATCACACTAAATATTTAGGACACCTCCTAGACCTAATGAGGT	SMAD3 mut. second round forward
Smad3_mut2_rev	ACCTCATTAGGTCTAGGAGGTGTCCTAAATATTTAGTGTGATGACAGGAC	SMAD3 mut. second round reverse
Pea3_mut1_for	CTCCTCCTCCCTCTACTTTTAATTCACCCGCCGAACAC	PEA3 mut. first round forward
Pea3_mut1_rev	GTGTTCGGCGGGTGAATTAAAAGTAGAGGGAGGAGGAG	PEA3 mut. first round reverse
Pea3_mut2_for	AGGACACAGAACCTCTGTTTAATGTTTGGCCTCCTCCCTCTAC	PEA3 mut. second round forward
Pea3_mut2_rev	GTAGAGGGAGGAGGCCAAACATTAAACAGAGGTTCTGTGTCCT	PEA3 mut. second round reverse
RAR_mut1_for	CTCCTCCTCCCTCTACTTCCAATGGGCCCGCCGAACAC	RAR mut. first round forward

RAR_mut1_rev	GTGTTCGGCGGGCCCATTGGAAGTAGAGGGAGGAGGAG	RAR mut. first round reverse
RAR_mut2_for	TTCACCCGCCGAACACACACAGGGCCTGCTCTGCC	RAR mut. second round forward
RAR_mut2_rev	GGCAGAGCAGGCCCTGTGTGTGTTCGGCGGGTGAA	RAR mut. second round reverse
TCF_mut1_for	GCGCTCGCTTGAAAAGAGTCCGATTCAAGCAGACAAAG	TCF mut. 1st round forward
TCF_mut1_rev	CTTTGTCTGCTTGAATCGGACTCTTTTCAAGCGAGCGC	TCF mut. 1st round reverse
TCF_mut2_for	GTGATGACAGGACACAGAACGACTGTTTAATGTTTCCTCCTC	TCF mut. 2nd round forward
TCF_mut2_rev	GAGGAGGAAACATTAAACAGTCGTTCTGTGTCCTGTCATCAC	TCF mut. 2nd round reverse
TCF_mut3_for	AATGTTTCCTCCTCCTCCGACTACTTCCAATTCACCCG	TCF mut. 3rd round forward
TCF_mut3_rev	CGGGTGAATTGGAAGTAGTCGGAGGAGGAGGAAACATT	TCF mut. 3rd round reverse
TCF_mut4_for	GAAAAGAGTCCGATTCAAGCAGATGCGGACCTCATTAGGTCTAGGAGGTG	TCF mut. 4th round forward
TCF_mut4_rev	CACCTCCTAGACCTAATGAGGTCCGCATCTGCTTGAATCGGACTCTTTTC	TCF mut. 4th round reverse
Homeo_mut1_for	GCGATTCAAGCAGACAAAGACCGGGGTAGGTCTAGGAGGTGTCCTGTC	Homeodomain mut. forward
Homeo_mut1_rev	GACAGGACACCTCCTAGACCTACCCCGGTCTTTGTCTGCTTGAATCGC	Homeodomain mut. reverse

BAC recombineering

Primer Name	Sequence	Purpose
GFP_Bmp6_for	CTGCAGCTCCAAGAGAGACCCACTTGGGACAGCGGAGAACACAGCGGGG AGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTC	GFP>Bmp6 recombineering
GFP_Bmp6_rev	CCAAGGTTAACGAAGCTCATGACCATGTCTGCGTCATTTAGAAAGGCACTC CGCGTGTAGGCTGGAGCTGCTTC	GFP>Bmp6 recombineering
PTARBAC_tol2FWD	GCGTAAGCGGGGCACATTTCATTACCTCTTTCTCCGCACCCGACATAGATCC CTGCTCGAGCCGGGCCCAAGTG	iTol2 recombineering
PTARBAC_tol2REV	CGCGGGGCATGACTATTGGCGCGCCGGATCGATCCTTAATTAAGTCTACTAA TTATGATCCTCTAGATCAGATCT	iTol2 recombineering

All primers were designed from genomic sequences obtained from UCSC. Gac=*Gasterosteus aculeatus* (stickleback), Dre=*Danio rerio* (zebrafish), Gmo=*Gadus morhua* (Atlantic cod), Ola=*Oryzias latipes* (medaka). For constructs with multiple mutations, the order in which the mutations were introduced is indicated.

Table S2. Enhancer activity of *cis*-regulatory sequences from four species in stickleback and zebrafish *trans* environments.

	***cis*-regulatory element**	**# lines with median fin expression**	**# lines with pectoral fin expression**	**# lines with tooth expression**
Stickleback *trans*	zebrafish	1/8	1/8	0/8
	cod	5/7	5/7	4/7
	medaka	4/5	4/5	4/5
	stickleback	6/6	6/6	6/6
Zebrafish *trans*	zebrafish	4/7	5/7	3/7
	cod	2/2	2/2	2/2
	medaka	5/5	5/5	5/5
	stickleback	2/2	2/2	2/2

Fish injected with each construct were outcrossed to wild-type fish, and offspring were scored for GFP fluorescence in the distal edge of the median fin, distal edge of the pectoral fin, and the pharyngeal teeth for each independent line. For stickleback, median and pectoral fins were scored at 5 dpf and teeth were scored post-hatching (12-20 dpf). For zebrafish, median fins were scored at 24 hpf, pectoral fins were scored at 48 hpf, and teeth were scored at 5 dpf.

Table S3. Efficiency of molecular lesions produced by TALENs.

clutch number	generation	% molecular lesions
1	F0 injected	17/17 (100%)
2	F0 injected	19/19 (100%)
3	F0 injected	9/10 (90%)
Average	*F0 injected*	*98%*
4	F1 outcross	2/10 (20%)
5	F1 outcross	5/10 (50%)
6	F1 outcross	7/10 (70%)
7	F1 outcross	6/9 (67%)
8	F1 outcross	9/10 (90%)
Average	*F1 outcross*	*59%*

A subset of each TALEN clutch was screened at 2 dpf for TALEN-induced lesions. Molecular lesions were identified by PCR amplification with Gac_190_for and Gac_72_rev and digestion with XbaI (see Fig. 4G for illustration). An undigested band indicated the presence of a TALEN-induced lesion. Lesions were confirmed by Sanger sequencing for a subset of F1 animals, including parents of animals used for *in situ* hybridization (see Figure 6E).

Table S4. RVDs used for TALEN construction.

	PfusA RVDs	PfusB RVDs	pLR RVD
5':	NG1	NI1	HD
	NI2	NN2	
	NN3	NN3	
	NN4	NG4	
	NG5	NN5	
	HD6	NG6	
	NG7		
	NI8		
	NN9		
	NN10		
3':	NG1	NN1	NG
	HD2	NG2	
	NG3	HD3	
	NN4	NI4	
	NG5		
	NN6		
	NG7		
	HD8		
	HD9		
	NG10		

Individual RVD monomers were cloned into pFUS_A and the appropriate pFUS_B plasmid. The completed pFUS_A and pFUS_B plasmids were then combined into pTal3-DD (5') and pTal3-RR (3') with the appropriate pLR and sequence-verified by Sanger sequencing (Cermak et al., 2011).