Title
Developmental Genetic Basis of Tooth Number Evolution in Stickleback Fish

Author
Cleves, Phillip Alfonso

Publication Date
2015

Peer reviewed|Thesis/dissertation

Developmental Genetic Basis of Tooth Number Evolution in Stickleback Fish

By

Phillip Alfonso Cleves

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Craig Miller, Chair

Professor Gian Garriga

Professor Nipam Patel

Professor Leslea Hlusko

Spring 2015

Abstract

Developmental Genetic Basis of Tooth Number Evolution in Stickleback Fish

by

Phillip Alfonso Cleves

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Assistant Professor Craig Miller, Chair

Teeth are a classic model for studying vertebrate organogenesis and evolution. Despite the incredible phenotypic diversification of dentition in vertebrates, our understanding of the molecular and developmental basis behind this variation is limited. A derived benthic freshwater stickleback population has evolved a nearly two-fold increase in ventral pharyngeal tooth number compared to their ancestral marine counterparts. This evolved tooth gain provides an excellent system to study the developmental and molecular genetic basis of evolved dental variation. To ask when during development evolved tooth gain appears, we generated lab-reared developmental time courses of a low-toothed marine population and this high-toothed freshwater population. Early in development, no differences in dental patterning are observed. However, at late larval stages, an increase in tooth number, an increase in tooth plate area, and a decrease in tooth spacing arise. We identified genomic regions controlling these evolved patterning changes by mapping quantitative trait loci (QTL) controlling tooth number, area, and spacing in a marine by freshwater F2 cross. One large effect QTL controlling tooth number fine-maps to a genomic region containing an excellent candidate gene, *Bone morphogenetic protein 6 (Bmp6)*. Stickleback *Bmp6* is expressed in developing teeth, but no coding changes are found between the two populations. However, by quantitatively comparing allele specific expression of *Bmp6*, we find *cis*-regulatory changes have down-regulated the relative expression level of the freshwater *Bmp6* allele at late, but not early, stages of development. To functionally test the role of *Bmp6* in controlling tooth patterning, we generated predicted loss-of-function alleles of *Bmp6* in freshwater sticklebacks. We found that *Bmp6* is required for tooth formation and tooth plate area mirroring aspects of the evolved changes. Next, to discover enhancers that contain marine/freshwater sequence differences, we compared the chromosome 21 genomic sequences from fish with the tooth QTL to fish without the QTL. We identified a partially conserved region of the fourth intron of *Bmp6* containing QTL-associated variants. This region is a tooth and fin enhancer that drives partially distinct expression patterns during tooth development compared to a 5' *Bmp6* tooth and fin enhancer we previously discovered that lacks consistent sequence differences associated with the tooth QTL. Future genetic and transgenic approaches will functionally test this intron 4 enhancer of *Bmp6* as a candidate for underlying evolved tooth gain in sticklebacks.

Table of Contents

List of Tables

List of Figures

For my grandfather, Phil Warren Adams, a man who loved life and exploring the wonders of nature

Acknowledgments

None of this would have been possible without the continuous support and encouragement of my mentor Dr. Craig Miller. Over the past 6.5 years, he has consistently gone above and beyond to support me both personally and scientifically. He provided a strong scientific environment in which I could flourish and be productive. Through the ups and downs, he has been there as a mentor and as a friend.

The Miller lab has been a stimulating and fun place to work. I am particularly grateful for my intelligent and generous co-workers: Nick Ellis, Priscilla Erickson, James Hart, and Andrew Glazer. I especially like to acknowledge Andrew Glazer, who has been a great friend from the beginning.

I would also like to thank the wonderful students that I have had the pleasure to work with: Monica Jimenez, Rachel Agoglia, Jessica Grindheim, Meera Garriga and Sara Carsanaro. Without the work of these careful and talented students, this research would have been much more difficult. In particular, Monica Jimenez and Rachel Agoglia have thought me many lessons in how to be an effective mentor, and for that I am grateful.

I am lucky to have a great family that has supported me my entire life. My dad, Mario Cleves, always makes time to help me bounce around ideas or teach me statistics, and my mom, Patti Cleves, has been there to listen to my scientific ideas and provide loving support as needed. I also want to thank Alice Cleves, for her support and encouragement through the years. I thank all my family for affording me the self-confidence and security to follow my dreams regardless of where they might take me. I am very lucky and proud to have them in my life. I also want to thank my brother, Andrew Cleves, for teaching me how to live a happy fulfilling life and for always being my best friend.

Throughout the course of graduate school, I have made many life-long friends. From long climbing trips in Bishop to weekly TV club, I have been extremely lucky to be surround by such great people. I will never forget all the excellent backpacking trips with the crew and seeing concerts with Debbie Thurtle, Aaron Cantor, and Mason Appel. The best memories in graduate school were often rock climbing with many friends, including but not limited to, Alec Sexton, Damian Trujillo, and Dan Richter. I am positive that the relationships I made in graduate school will last a lifetime.

Finally, I would not have had such an amazing past few years of graduate school without my partner, Tara Anderson. She has been a major source of encouragement and support. Thank you very much for being a great partner and I am lucky to have you in my life.

Chapter 1

Introduction

1.1 The molecular substrate of morphological evolution

Understanding the genetic basis of evolved differences in morphology remains a fundamental goal in biology. In the last few decades, evolutionary geneticists have begun to make progress identifying the genes and mutations driving the morphological diversity in animals. This work is aimed at understanding and generalizing how evolution acts to generate diversity. Some of the important questions in the field are: How many loci control morphological diversity in animals? Are certain types of genes more likely to be used in evolution? What types of mutations are used? Are the answers to these questions generalizable across different types of evolution and organisms?

In some systems for some traits, the genetic architecture of morphological evolution is beginning to be understood. In 1930, Ronald Fisher proposed the Infinitesimal Model of evolution postulating that adaptive evolution occurs and can be effectively modeled through many very small effect mutations (1). In 1983, Motoo Kimura elaborated on Fisher's infinitesimal model by suggesting that small effect mutations may not be able to escape being lost by genetic drift and concluded that evolution occurs through mutations of intermediate effect (2). In 1998, Allen Orr extended this theory by simulating the expected effect sizes of adaptive evolution and concluded that an exponential function, where most mutations are of small effect with fewer of large effect fit the data best (3). Modern genetic studies using techniques like Quantitative Trait Locus (QTL) mapping have demonstrated that morphological evolution can occur through large effect loci with much of the variance in a trait controlled by a relatively small number of these loci, with many additional loci of small effect (4–7). These studies suggest that although the genetic architecture of morphological evolution is complex, in many cases it is not controlled by an infinitesimal number of loci, but most closely follows theoretical effect size distribution predicted by Orr.

Fisher's infinitesimal model of evolution applied to mutations and may predict that large effect QTL may be made up of many tightly linked mutations of small effect acting on a trait. There have been recent efforts to understand the mutations underlying some of these large effect loci associated with morphological evolution. The large effect QTL controlling stickleback pelvic spine reduction has been associated with recurrent deletions of a pelvic fin enhancer of the transcription factor, *Pitx1* (8). However in another case, the large effect QTL controlling the loss of cuticle hair in *Drosophila* larvae is due to a combination of many linked mutations altering the regulation of a gene required for the formation of cuticle hair, *ovo/shaven-baby* (9). These two case studies demonstrate that many one to a small number of mutations can underlie large effect loci. However, it remains important to determine how generalizable these case studies are across evolution.

The relative contribution of *cis*-regulatory versus coding mutations underlying morphological evolution has been hotly debated in the field of evolution (10, 11). The *cis*-regulatory hypothesis proposes that changes in gene expression (e.g. through mutations in regulatory transcriptional enhancers) are a preferred substrate for morphological evolution because they provide a

mechanism for evolution to modify gene function with precise tissue and developmental stage specificity. This specificity is hypothesized to be able to bypass potential negative pleiotropy typically generated by coding mutations in a developmental regulatory gene that may modify all of gene's functions. Conversely, others have challenged this model suggesting that certain classes of coding mutations (e.g. coding mutations after gene duplication) may also bypass negative pleiotropy and could be used during evolution without fitness defects (12).

Recent genetic and comparative studies have helped identify and implicate many genes underlying evolutionary differences between species. It has become clear that morphological evolution can occur through mutations in *cis*-regulatory elements and also coding mutations. Furthermore, it has been discovered that evolutionarily distant linages evolving similar morphological traits often use the same pathways, the same genes, and sometimes even the same alleles to modify homologous structures (13, 14). These results suggest that there are constraints on the precise genes used for morphological evolution (13).

Despite the increasing number of examples of genes underlying evolution, the link between evolved morphological differences in vertebrates and the changes in their genomes remains elusive. Furthermore, how these genomic differences manifest during development to produce varying morphology is largely not understood. Are there constraints on when and how evolution can modify existing morphological structures during development? Are certain mutations preferred in different types of evolution (e.g. when a trait is lost or when a trait is gained in constructive evolution)? This thesis will help answer these broad questions by dissecting the developmental genetics of a constructive evolved trait.

1.2 Teeth as model systems for organ evolution and disease

Teeth are vertebrate innovations that are thought to have contributed to the evolution and success of the lineage (15). Teeth have been extensively studied as a model of organ evolution in vertebrates due to their excellent preservation in the fossil record. Tooth patterning has diversified greatly in vertebrates and in some cases help linages occupy novel trophic niches (16). During development, teeth arise from stereotypical interactions between cranial neural crest derived ectomesenchyme and an ectoderm or endoderm derived epithelium (17–19). Despite lineage specific elaborations, vertebrate teeth progress through four highly conserved core developmental stages: 1. epithelial cells thicken to form a placode, 2. epithelial layers of the placode bud into adjacent condensing mesenchyme, 3. epithelia of the bud fold around a mesenchymal core, and 4. ameloblast and odontoblasts differentiate from the epithelia and mesenchyme, respectively, to secrete extracellular matrix molecules that comprise the tooth (20, 21). A complex and not fully understood network of signaling interactions control progression through these stages involving members of the BMP, FGF, WNT, and HH pathways, which generate mature teeth of a particular shape, size, and number (22). The vast majority of

knowledge on tooth developmental genetics has come from studies using mice. However, since mice are monophydonts, having only one set of teeth without tooth replacement, the molecular genetic basis for tooth regeneration and replacement is far less understood. Reptile and fish model systems provide complementary systems to characterize the tooth replacement process (23).

An emerging model for tooth replacement is the continuously replacing fish pharyngeal teeth. Pharyngeal teeth lie in the pharynx of fish and are serial and phylogenetic homologs of mammalian oral teeth (24). These deeply homologous structures share similar morphological, developmental, and genetic bases (25). Over 300 genes are thought to be involved in mammalian tooth development and many homologous networks have been shown to be involved in fish tooth development (26, 27). Understanding the genetics of tooth development and replacement in fish offers a powerful opportunity to compare the development and genetic circuitry between species, providing new insight into tooth regeneration.

1.3 Roles of BMP signaling in tooth development

Bone morphogenetic proteins (BMPs), discovered due to their ability to induce ectopic bone when injected subcutaneously in rodent models, are master regulators of bone and skeletal development (28, 29). Since their discovery, BMP signaling has emerged as a major developmental orchestrator controlling a wide variety of developmental processes ranging from establishing the dorsal-ventral embryonic axis in vertebrates and invertebrates, to patterning vertebrate epithelial appendages, such as hair and teeth (30). In teeth, BMP signaling has been implicated in all stages of tooth development, including initiation and placement of a developing tooth, formation and patterning of a functional tooth, and tooth replacement.

Gene expression patterns early during mouse tooth development suggested that the specification of early tooth placement in mice is regulated by BMP signaling. *Bmp4* negatively regulates *Pitx2* and *Pax9*, early markers of tooth initiation, and restricts a positive-regulator of those makers, *Fgf8* (31, 32). *Pitx2* is required for early tooth specification in the maxillary molars; however *Pax9* is not required for tooth initiation (33, 34). These results suggest that BMP signaling can have an inhibitory role on tooth development at this early initiation stage.

During tooth development, *Bmp* ligands *2*-7*,* have been shown to be expressed in dynamic patterns in the developing tooth epithelium and/or mesenchyme (35). In addition to the possible inhibitory effects on early tooth placement described above, several lines of evidence suggest *Bmp* genes activate tooth development. For example, *Bmp4,* the most heavily studied BMP ligand in tooth development, is required for tooth formation to progress pass the bud stage in mice. In addition, *Bmp4* beads can increase the rate of molar tooth formation in mice (36). Mice mutant for the BMP receptor, *Bmpr1A*, in the developing tooth epithelia, and mice and humans with mutations in downstream effectors of BMP signaling, *Msx1*, fail to form functional teeth. The *Msx1* expression in mouse teeth is required for cell proliferation of the developing tooth mesenchyme (37). The loss of teeth in the *Msx1* mutant mice can be rescued by

the addition of recombinant *Bmp4* protein (37). These results imply that BMP signaling has an activating role in tooth formation.

The role of BMP signaling during tooth regeneration and replacement is starting to be explored. BMP ligands are expressed in replacement teeth in snakes and fish (38–40). In regenerating ball python teeth, *Bmp2* and *Bmp4* are expressed in the mesenchyme surrounding a regenerating tooth and consequently *pSMAD1/5/8* is detected strongly in the epithelium of replacement teeth (38). This BMP signaling in the mesenchyme modulates cell proliferation in the developing replacement tooth bud (38). Work in cichlid fish has also shown a role of BMP signaling in tooth replacement. Cichlid regenerating oral teeth fail to be replaced when soaked in the BMP antagonist, Dorsomorphin (40). These results point to a role of BMP signaling as inductive signals for tooth replacement. Understanding the positive and negative roles of BMP signaling in both primary and replacement teeth by dissecting its effects in both the epithelium and mesenchyme is a major goal for the field for the future.

BMP signaling also plays a role of stem cell regulation in teeth and other epithelial organs. *Bmp4* represses *Fgf3*, which helps maintain the labial cervical loop stem cell population (41). These results suggest that BMP signaling suppresses incisor stem cell proliferation. In developmentally homologous epithelial appendages such as hair, BMP signaling has been show to positively regulate stem cell quiescence and also to promote different differentiation paths of transit amplifying cell populations (42, 43). *Bmp6* ligand has been shown to be an important gene in hair follicle stem cell population and is thought to play a role in negatively regulating their proliferation (42). These results suggest a role BMP signaling in regulating epithelial appendage stem cell populations; however the precise roles of BMP signaling in teeth and in particular tooth replacement are still largely unknown.

1.4 Stickleback fish as genetic model for evolution

During adaptive radiations, populations of organisms diversify and adapt in response to novel ecological niches (44). A recent and dramatic example of an adaptive radiation in vertebrates is the threespine stickleback fish. Since the last ice age roughly 10,000 years ago, marine sticklebacks have colonized countless new freshwater lakes and streams and independently evolved in response to freshwater environments (45). Because of these relatively recent diversifications and the advent of new genomic and molecular resources, sticklebacks are emerging as a model for understanding the molecular basis of vertebrate evolution (46).

Despite the dramatic morphological variation that exists between marine and freshwater stickleback, derived freshwater stickleback populations can be crossed in the lab to ancestral marine stickleback to produce viable offspring. These F1 offspring can be incrossed to generate large F2 crosses, facilitating genetic mapping of evolved differences. This approach has proven successful in identifying many quantitative trait loci (QTL) underlying morphological evolution in stickleback (5, 6, 47, 48). These genetic mapping results along with genomic approaches identifying signals of selection have identified three stickleback

chromosomes (4, 20, and 21) as genomic hotspots of stickleback evolution where many evolved traits map and signals of selections are found (6, 49, 50). These clustered genomic hotspots may facilitate rapid adaptation of new freshwater environments by selection being able to easily increase the allele frequency of many linked freshwater traits at once.

An ultimate goal for QTL mapping in sticklebacks is to identify the genes and eventually the mutations responsible for evolutionary change. *Cis*-regulatory alleles of *Eda*, *Kitlg*, and *Pitx1* have been associated with evolved reductions in lateral plates, skin and gill pigmentation, and pelvic spines in freshwater stickleback, respectively (51–56). For *Eda* and *Pitx1*, candidate *cis*-regulatory mutations in developmental enhancers have been identified and are thought to contribute to the evolved skeletal differences. The *Pitx1* mutations are independently derived deletions of a pelvic enhancer within a fragile genomic region (8). In contrast, a shared allele with a SNP in a lateral plate enhancer of *Eda* underlies evolved lateral plate reduction. Both of these mutations reduce enhancer activity, which then presumably leads to pelvic spine and lateral plate loss, respectively (51, 56).

These first three cases where the causal gene has been identified that underlies a derived phenotype in freshwater sticklebacks are cases where skeletal elements are lost or reduced relative to the ancestral marine phenotype. Interestingly, however, a freshwater stickleback population has evolved a two-fold increase in the number of ventral pharyngeal teeth compared to their marine counterparts. This phenotype is intriguing because it is a constructive trait in the derived population. It is unclear whether the evolutionary mechanisms driving morphological gain are fundamentally different from the mechanisms driving morphological loss.

1.5 Stickleback evolved tooth gain

The first freshwater population that has been described to have evolved tooth gain is from Paxton Lake in Canada (6). This freshwater lake has a stickleback “species pair” containing benthic and limnetic morphs (meaning adapted to live on lake bottom and open water, respectively). Benthic and limnetic sticklebacks have evolved differences in trophic morphology likely in response to utilizing different ecological niches (57–59). Benthic sticklebacks are adapted to feeding on macroinvertebrates in the littoral zone of the lake. This diet is different from both the limnetic and ancestral marine stickleback populations that feed on smaller zooplankton. We hypothesize that the evolved tooth gain is an adaptive response to the benthic diet shift.

To date, evolved gain has been described only in the Paxton benthic freshwater population. However, many of the evolved traits in stickleback have evolved in multiple independently derived freshwater populations across the northern hemisphere (53, 60–62). It is possible that other freshwater populations that have experienced similar shifts in diet as the Paxton benthic fish may have evolved tooth gain. The presence of freshwater tooth gain in multiple populations would strongly suggest an adaptive role for the gain in tooth number. It remains an important goal to determine the frequency of freshwater tooth gain across

stickleback populations. Consistent with evolved tooth gain being adaptive in freshwater, we have discovered a second high-toothed freshwater population in nearby Cerrito Creek in El Cerrito.

Despite the vast diversity in tooth morphology across fish, the genetic basis for that variation is largely unknown. In fact, it has been shown that some variation in fish tooth morphology might be due to the plastic responses to diet (63–65). For example, studies comparing wild-caught cichlid fish feeding on a hard snail diet to lab-reared cichlids fed on a soft diet found changes in tooth shape and number between the different groups (64). Furthermore, the changes in tooth patterning can reverse within the lifetime of a cichlid upon shifting diet, showing a large phenotypic plasticity component. These results suggest diet may play an important role in tooth pattern. However, it remains an important goal to define to what extent environment and genetics determine tooth variation in vertebrates.

In previous work aimed at identifying the genetic basis for tooth gain in freshwater sticklebacks, a large (360 fish) F2 cross was made between a low-toothed marine population, Japanese Marine, and the high-toothed freshwater Paxton benthic population. These F2 fish were genotyped genome-wide and phenotyped to identify quantitative trait loci (QTL) responsible for the variation in tooth number. This study identified eight genomic regions that collectively explain a majority of the phenotypic variance seen in the cross (6). These results show a strong heritable component of freshwater tooth gain. The largest effect QTL is on chromosome 21 and explains ~30% of the variance in ventral tooth number. This genetic interval contains many genes, including two excellent candidate genes, *Bone morphogenetic protein 6 (Bmp6)* and *Transcription factor AP-2 alpha* (*Tfap2a)*, which have been implicated in various aspects of head skeletal development.

Major questions remain regarding the developmental and genetic basis of freshwater tooth gain in sticklebacks. (1) When and how during development does this tooth gain arise? (2) What gene or genes underlie the chromosome 21 tooth QTL? (3) How many and what type of mutations are responsible (4) Are there other freshwater populations that use a chromosome 21 tooth QTL?

Throughout my thesis, I have sought to answer these questions regarding tooth gain in sticklebacks. We have shown that freshwater tooth gain arises late in development accompanied by changes in tooth density and tooth field size, which are each under partially separable genetic control. We have shown that the chromosome 21 tooth QTL is associated with a late-acting *cis*-regulatory allele of *Bmp6* and that *Bmp6* is required for tooth development. Through comparative genomics, we have identified a candidate tooth enhancer of *Bmp6* containing tooth QTL-associated variants that may underlie aspects of the tooth QTL. Finally, we have identified an additional freshwater population that has a chromosome 21 tooth QTL, which shares the tooth QTL-associated variants in the tooth enhancer. Future work in the lab will determine if these variants have a functional consequence on enhancer activity in developing teeth and ultimately if they change tooth number. This thesis as a whole suggests constructive traits, like regressive traits, can be controlled by *cis*-regulatory mutations in genes

required for viability, providing additional support for the *cis*-regulatory hypothesis.

1.6 References

1. Fisher RA (1930) *The Genetical Theory of Natural Selection*.
2. Kimura M (1983) *The Neutral Theory of Molecular Evolution*.
3. Orr HA (1998) The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution (N Y)* 52(4):935–949.
4. Papa R, et al. (2013) Multi-allelic major effect genes interact with minor effect QTLs to control adaptive color pattern variation in Heliconius erato. *PLoS One* 8(3):e57033.
5. Albert AYK, et al. (2008) The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size. *Evolution (N Y)* 62:76–85.
6. Miller CT, et al. (2014) Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait Loci. *Genetics* 197(1):405–20.
7. Peichel CL, et al. (2001) The genetic architecture of divergence between threespine stickleback species. *Nature* 414(6866):901–905.
8. Chan YF, et al. (2010) Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. *Science (80-)* 327(5963):302–305.
9. Stern DL, Frankel N (2013) The structure and evolution of cis-regulatory regions: the shavenbaby story. *Philos Trans R Soc Lond B Biol Sci* 368(1632):20130028.
10. Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* 8(3):206–216.
11. Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61(5):995–1016.
12. Hoekstra HE, Coyne JA (2007) THE LOCUS OF EVOLUTION: EVO DEVO AND THE GENETICS OF ADAPTATION. *Evolution (N Y)* 61(5):995–1016.
13. Stern DL (2013) The genetic causes of convergent evolution. *Nat Rev Genet* 14(11):751–64.
14. Martin A, Orgogozo V (2013) The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67(5):1235–50.
15. Stock DW (2001) The genetic basis of modularity in the development and evolution of the vertebrate dentition. *Philos Trans R Soc B Biol Sci* 356(1414):1633–1653.
16. Muschick M, Indermaur a., Salzburger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol* 22:2362–8.
17. Soukup V, Epperlein H-H, Horácek I, Cerny R (2008) Dual epithelial origin of vertebrate oral teeth. *Nature* 455(7214):795–8.
18. Lumsden AG (1988) Spatial organization of the epithelium and the role of neural crest cells in the initiation of the mammalian tooth germ. *Development* 103 Suppl:155–69.

19. Mina M, Kollar EJ (1987) The induction of odontogenesis in non-dental mesenchyme combined with early murine mandibular arch epithelium. *Arch Oral Biol* 32(2):123–7.
20. Bei M (2009) Molecular genetics of tooth development. *Curr Opin Genet Dev* 19(5):504–510.
21. Huysseune A, Van der heyden C, Sire JY (1998) Early development of the zebrafish (Danio rerio) pharyngeal dentition (Teleostei, Cyprinidae). *Anat Embryol (Berl)* 198(4):289–305.
22. Catón J, Tucker AS (2009) Current knowledge of tooth development: patterning and mineralization of the murine dentition. *J Anat* 214(4):502–515.
23. Tucker AS, Fraser GJ (2014) Evolution and developmental diversity of tooth regeneration. *Semin Cell Dev Biol*.
24. Wise SB, Stock DW (2006) Conservation and divergence of Bmp2a, Bmp2b, and Bmp4 expression patterns within and between dentitions of teleost fishes. *Evol Dev* 8:511–523.
25. Stock DW (2007) Zebrafish dentition in comparative context. *J Exp Zool Part B Mol Dev Evol* 308B(5):523–549.
26. Fraser GJ, et al. (2009) An Ancient Gene Network Is Co-opted for Teeth on Old and New Jaws. *PLoS Biol* 7(2):e31.
27. Jackman WR, Draper BW, Stock DW (2004) Fgf signaling is required for zebrafish tooth development. *Dev Biol* 274(1):139–157.
28. Urist MR (1965) Bone: Formation by Autoinduction. *Science (80-)* 150(3698):893–899.
29. Sampath TK, Reddi AH (1981) Dissociative extraction and reconstitution of extracellular matrix components involved in local bone differentiation. *Proc Natl Acad Sci* 78(12):7599–7603.
30. Wang RN, et al. (2014) Bone Morphogenetic Protein (BMP) signaling in development and human diseases. *Genes Dis* 1(1):87–105.
31. Neubuser A (1997) Antagonistic Interactions between FGF and BMP Signaling Pathways: A Mechanism for Positioning the Sites of Tooth Formation. *Cell* 90(2):247–255.
32. St Amand TR, et al. (2000) Antagonistic Signals between BMP4 and FGF8 Define the Expression of Pitx1 and Pitx2 in Mouse Tooth-Forming Anlage. *Dev Biol* 217(2):323–332.
33. Lin CR, et al. (1999) Pitx2 regulates lung asymmetry, cardiac positioning and pituitary and tooth morphogenesis. *Nature* 401(6750):279–82.
34. Peters H, Neubuser A, Kratochwil K, Balling R (1998) Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev* 12(17):2735–2747.
35. Thesleff I, Mikkola M (2002) The role of growth factors in tooth development. *Int Rev Cytol* 217:93–135.
36. Kavanagh KD, Evans AR, Jernvall J (2007) Predicting evolutionary patterns of mammalian teeth from development. *Nature* 449(7161):427–32.

37. Bei M, Kratochwil K, Maas RL (2000) BMP4 rescues a non-cell-autonomous function of Msx1 in tooth development. *Development* 127(21):4711–4718.
38. Handrigan GR, Richman JM (2010) A network of Wnt, hedgehog and BMP signaling pathways regulates tooth replacement in snakes. *Dev Biol* 348(1):130–41.
39. Fraser GJ, Graham A, Smith MM (2004) Conserved deployment of genes during odontogenesis across osteichthyans. *Proc R Soc B Biol Sci* 271(1555):2311–2317.
40. Fraser GJ, Bloomquist RF, Streelman JT (2013) Common developmental pathways link tooth shape to regeneration. *Dev Biol* 377(2):399–414.
41. Wang X-P, et al. (2007) An integrated gene regulatory network controls stem cell proliferation in teeth. *PLoS Biol* 5(6):e159.
42. Genander M, et al. (2014) BMP Signaling and Its pSMAD1/5 Target Genes Differentially Regulate Hair Follicle Stem Cell Lineages. *Cell Stem Cell* 15(5):619–633.
43. Kandyba E, et al. (2013) Competitive balance of intrabulge BMP/Wnt signaling reveals a robust gene network ruling stem cell homeostasis and cyclic activation. *Proc Natl Acad Sci U S A* 110(4):1351–6.
44. Harder LD, Schluter D (2001) The Ecology of Adaptive Radiation. *Am J Bot* 88:1707.
45. Bell MA, Foster SA (1994) *The Evolutionary Biology of the Threespine Stickleback*.
46. Amemiya C, et al. (2004) New Genomic Tools for Molecular Studies of Evolutionary Change in Threespine Sticklebacks. *Behaviour* 141:1331–1344.
47. Colosimo PF, et al. (2004) The Genetic Architecture of Parallel Armor Plate Reduction in Threespine Sticklebacks. *PLoS Biol* 2(5):e109.
48. Wark AR, et al. (2012) Genetic Architecture of Variation in the Lateral Line Sensory System of Threespine Sticklebacks. *G3 Genes|Genomes|Genetics* 2(9):1047–1056.
49. Jones FC, et al. (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484(7392):55–61.
50. Hohenlohe PA, et al. (2010) Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLoS Genet* 6(2):e1000862.
51. Chan YF, et al. (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327:302–305.
52. Shapiro MD, et al. (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428:717–723.
53. Miller CT, et al. (2007) cis-Regulatory Changes in Kit Ligand Expression and Parallel Evolution of Pigmentation in Sticklebacks and Humans. *Cell* 131(6):1179–1189.

54. Harris MP, et al. (2008) Zebrafish eda and edar Mutants Reveal Conserved and Ancestral Roles of Ectodysplasin Signaling in Vertebrates. *PLoS Genet* 4(10):e1000206.
55. Colosimo PF (2005) Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science (80-)* 307(5717):1928–1933.
56. O'Brown NM, Summers BR, Jones FC, Brady SD, Kingsley DM (2014) A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene EDA. *Elife* 4:e05290.
57. Schluter D, McPhail JD (1992) Ecological character displacement and speciation in sticklebacks. *Am Nat* 140:85–108.
58. Caldecutt WJ, Bell MA, Buckland-Nicks JA (2001) Sexual Dimorphism and Geographic Variation in Dentition of Threespine Stickleback, Gasterosteus aculeatus. *Copeia* 2001(4):936–944.
59. McPhail JD (1992) Ecology and evolution of sympatric sticklebacks (Gasterosteus): evidence for a species-pair in Paxton Lake, Texada Island, British Columbia. *Can J Zool* 70(2):361–369.
60. Erickson PA, Glazer AM, Cleves PA, Smith AS, Miller CT (2014) Two developmentally temporal quantitative trait loci underlie convergent evolution of increased branchial bone length in sticklebacks. *Proc Biol Sci* 281(1788):20140822.
61. Glazer AM, Cleves PA, Erickson PA, Lam AY, Miller CT (2014) Parallel developmental genetic features underlie stickleback gill raker evolution. *Evodevo* 5:19.
62. Colosimo PF, et al. (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* 307(5717):1928–33.
63. Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evol Biol* 11:116.
64. Huysseune A (1995) Phenotypic plasticity in the lower pharyngeal jaw dentition of astatoreochromis alluaudi (teleostei: cichlidae). *Arch Oral Biol* 40(11):1005–1014.
65. Meyer A (1987) Phenotypic plasticity and heterochrony in Cichlasoma managuense (Pisces, Chichlidae) and their implications for speciation in Cichlid fishes. *Evolution (N Y)* 41(6):1357–1369.

Chapter 2

Evolved tooth gain in sticklebacks is associated with a *cis*-regulatory allele of *Bmp6*

The following chapter was originally published as an article:
PNAS, 2014.

Phillip A. Cleves[1], Nicholas A. Ellis[1], Monica T. Jimenez[1], Stephanie M. Nunez[2,3], Dolph Schluter[4], David M. Kingsley[2], Craig T. Miller[1]

[1]Department of Molecular and Cell Biology, University of California-Berkeley, Berkeley CA, 94720, USA
[2]Department of Developmental Biology and Howard Hughes Medical Institute, Stanford University, Stanford, CA, 94305, USA
[3] present address = University of Michigan School of Dentistry, Ann Arbor, MI, 48109, USA
[4]Department of Zoology, University of British Columbia, Vancouver, Canada V6T 1Z4

2.1 Abstract

Developmental genetic studies of evolved differences in morphology have led to the hypothesis that *cis*-regulatory changes often underlie morphological evolution. However, because most of these studies focus on evolved loss of traits, the genetic architecture and possible association with *cis*-regulatory changes of gain traits are less understood. Here we show that a derived benthic freshwater stickleback population has evolved an approximate two-fold gain in ventral pharyngeal tooth number compared to their ancestral marine counterparts. Comparing lab-reared developmental time courses of a low-toothed marine population and this high-toothed benthic population reveals that increases in tooth number and tooth plate area, and decreases in tooth spacing, arise at late juvenile stages. Genome-wide linkage mapping identifies largely separate sets of quantitative trait loci (QTL) affecting different aspects of dental patterning. One large effect QTL controlling tooth number fine-maps to a genomic region containing an excellent candidate gene, *Bone morphogenetic protein 6 (Bmp6)*. Stickleback *Bmp6* is expressed in developing teeth, and no coding changes are found between the high and low toothed populations. However, quantitative allele specific expression assays of *Bmp6* in developing teeth in F1 hybrids show that *cis*-regulatory changes have decreased the relative expression level of the freshwater benthic *Bmp6* allele at late, but not early, stages of stickleback development. Collectively our data support a model where a late-acting *cis*-regulatory down regulation of *Bmp6* expression underlies a significant increase in tooth number in derived benthic sticklebacks.

2.2 Introduction

Understanding the developmental genetic basis of morphological evolution is a long-standing goal in biology (1, 2). Evolved morphological differences can be "loss" (regressive) traits, where morphological features are lost or reduced, or "gain" (constructive) traits, where morphological features are gained or increased. Although many of the traits best understood at the molecular level involve loss traits (1, 2), recent studies have begun to genetically dissect some evolved gain traits (3-5). However, whether gain traits have similar genetic architectures as loss traits, and whether gain traits are also associated with *cis*-regulatory changes remains largely unknown.

Teeth are a classic vertebrate model system for studying morphological evolution, due to their excellent preservation in the fossil record. Teeth are also a classic vertebrate model system for organogenesis, because teeth, like many other organs, develop through reciprocal signaling interactions between epithelia and mesenchyme. Continuing efforts have produced a rich understanding of the genetic networks that orchestrate tooth morphogenesis in model systems (6). However, despite the wealth of knowledge about tooth evolution and development, we still know little about the number and type of genetic changes that accompany diversification of dental patterning during evolution.

Pharyngeal jaws and teeth, used during mastication in fish, are located in the posterior branchial segments in the fish's throat (7, 8). In teleost fish, pharyngeal jaw patterning is an adaptive trait that covaries with diet and trophic niche (9). The rich phenotypic diversity of pharyngeal jaws and teeth in fish, coupled with the understanding of the genetic networks of tooth development from model organisms, offers an opportunity to understand the developmental genetic basis of evolved changes in tooth patterning.

The threespine stickleback (*Gasterosteus aculeatus)* fish has emerged as an excellent model system allowing for genetic dissection of evolutionary change in vertebrates (10). Sticklebacks have undergone an extensive adaptive radiation, independently colonizing thousands of freshwater lakes and creeks generated after widespread melting of glaciers at the end of the last ice age (11). The dietary shifts to larger prey accompanying freshwater adaptation have resulted in evolved changes in trophic morphology (12, 13). Despite striking morphological differences between marine and freshwater populations, hybrids are fertile, allowing forward genetic analysis of evolved differences. In several lakes, "species pairs" of benthic and limnetic stickleback morphs are found (13). In each of these lakes, a benthic species is adapted to feeding on macroinvertebrates in the littoral zone or deeper sediments. This derived diet differs from the diet of both the limnetic species and ancestral marine forms, both of which feed on smaller zooplankton. Benthic sticklebacks have evolved trophic adaptations matched for this specialized diet (13, 14). Here we describe evolved tooth gain, a heritable constructive increase in tooth number compared to ancestral marine fish, in a derived benthic stickleback population. We then apply quantitative genetics and developmental biology methods to begin to dissect the genetic and developmental basis of this evolved gain trait.

2.3 Results

2.3.1 Derived benthic fish have evolved increases in tooth number

Because benthic sticklebacks have undergone an adaptive shift in diet, and because aspects of pharyngeal jaw patterning correlate with trophic niche in other species (9), we hypothesized that wild benthic fish have evolved changes in tooth patterning compared to ancestral marine fish. To test this hypothesis, we first quantified adult ventral pharyngeal tooth number from wild benthic fish from Paxton Lake, Canada (PAXB, hereafter called “benthic”) and an ancestral marine population from Rabbit Slough, Alaska (RABS, hereafter called “marine”). In these samples, the wild benthic population has an approximate two-fold gain in tooth number compared to wild marine adults (Figure 2.1*A*, Table 2.1). To determine whether this striking difference in tooth number is heritable, we quantified tooth number in adult fish from each population in a common lab-r eared environment. The increased tooth number in benthic fish compared to marine fish is also seen in lab-reared stocks fed the same diet (Figure 2.1*B*, Table 2.1) showing that the tooth number differences have a large heritable component.

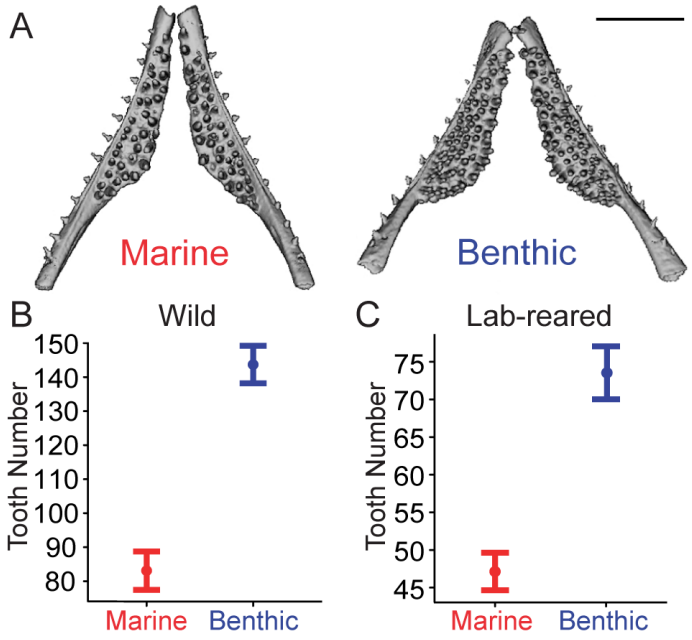


Figure 2.1. Heritable evolved tooth gain in derived benthic fish.
(A) MicroCT images of wild adult stickleback ventral pharyngeal tooth plates of marine fish from Rabbit Slough, Alaska (left) and benthic fish from Paxton Lake, Canada (right). Scale bar = 1 mm. (B-C) Total ventral pharyngeal tooth number in wild (B) and lab-reared (C) adults shows benthic fish have significantly higher tooth counts (P = 8.8x10-7 and P = 4.9x10-7 in two-tailed t-tests for wild and lab-reared, respectively). Error bars are standard error of the mean (SE).

Table 2.1. Wild and lab-reared tooth numbers of marine and freshwater fish

	Population	n	Mean standard length (mm)	Mean total tooth number	Comparison	*P* value
Wild	RABS	8	73.3 (1.4)	83 (16)	RABS-PAXB	< 0.00001
	JAMA	8	82.1 (2.5)	81 (19.7)	JAMA-RABS	0.98
	PAXB	20	58.0 (8.4)	144 (24.6)	PAXB-JAMA	< 0.00001
Lab-reared	RABS	21	38.8 (2)	47 (11.4)	RABS-PAXB	< 0.0001
	JAMA	23	38.5 (2.3)	58 (9.5)	JAMA-RABS	0.006
	PAXB	13	35.7 (1.4)	74 (12.7)	PAXB-JAMA	0.0004

Mean standard length in millimeters (mm) and total ventral pharyngeal tooth numbers with standard deviation (SD) for each are shown for Rabbit Slough marine (RABS), Japanese marine (JAMA), and Paxton benthic freshwater (PAXB) fish. *P* values from a Tukey's post-hoc test after an ANOVA for each population comparison are shown. The PAXB and RABS phenotypes are presented in Figure 2.1.

2.3.2 Evolved changes in tooth patterning occur late in development

To examine when during development this evolved increase in tooth number appears, we generated a dense developmental time course of lab-reared fish from both marine and benthic populations and quantified tooth number (Figure 2.2*A*). Tooth number between the two populations was not significantly different at early larval stages, but began diverging when fish reached a total length of about 20 millimeters (mm) (Figure 2.2*A*). Tooth number continued diverging after 20 mm, with

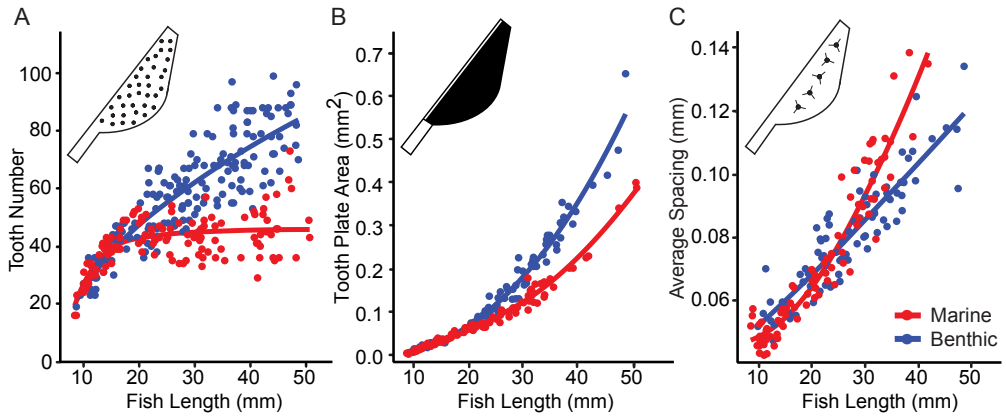


Figure 2.2. Evolved differences in tooth number, area, and spacing appear late during development.
Developmental time courses of lab-reared marine (red) and benthic (blue) fish of different total body lengths (x-axis) for tooth number (A), tooth plate area (B) and tooth spacing (C). All three traits have diverged after 20 millimeters (mm) fish length.

benthic fish continuing to add new teeth while marine fish tooth number plateaus, resulting in the approximate two-fold difference in tooth number seen in adults (Figure 2.2*A*, Table 2.2). The observed difference in tooth number between marine and benthic fish could arise either through an increase in the size of the tooth field and/or through an increased density of teeth in that field. We quantified tooth plate area (area of tooth-bearing portion of the fifth ceratobranchial bone) and average inter-tooth spacing throughout the developmental time courses. Relative to marine fish, in benthic fish tooth number and tooth plate area increased, while inter-tooth spacing decreased, with all three traits diverging late in development, after the 20 mm stage (Figure 2.2*B-C*, Table 2.2). Thus, the late increase in tooth number in derived benthic fish is accompanied by at least two other late developmental patterning changes: an expansion of the tooth field and an increase in tooth density within that field. These increases in tooth number and tooth plate area and decreased inter-tooth spaces were also observed in F1 marine by benthic hybrid fish, showing that the benthic phenotypes are at least partially dominant (Table 2.3).

Table 2.2. Divergence in tooth traits during development

TL range	Trait	Sample size	Marine	Benthic	*P* value
8-20mm					
	Tooth Number	94	0.60 (3.75)	-0.65 (4.04)	0.12
	Tooth Plate Area	59	8.24E-05 (0.004)	-0.0003 (0.005)	0.76
	Tooth Spacing	58	-0.001 (0.004)	0.002 (0.006)	0.02
20-30mm					
	Tooth Number	83	-7.24 (6.6)	3.89 (8.8)	6.91E-08
	Tooth Plate Area	62	-0.019 (0.013)	0.014 (0.019)	2.86E-10
	Tooth Spacing	59	2.58E-05 (0.008)	-1.77E-05 (0.008)	0.98
30-57mm					
	Tooth Number	126	-17.34 (8.7)	11.04 (10.7)	4.9E-31
	Tooth Plate Area	55	-0.051 (0.031)	0.045 (0.038)	3.97E-14
	Tooth Spacing	50	0.009 (0.012)	-0.006 (0.009)	9.27E-06

For each tooth trait, mean size-corrected phenotypic value with standard deviation (SD) for each population (marine = RABS, benthic = PAXB) at three total length (TL) ranges is shown after fish size was adjusted for using a linear model in R. *P* values from an ANOVA for population effect are shown.

Table 2.3. Tooth number, area, and spacing of marine by benthic F1 hybrids

Trait	Sample Size	Marine	F1 Hybrid	Benthic	*P* values
Tooth Number	52,31,89	-18.06 (8.53)	4.92 (7.52)	8.84 (10.57)	<0.00001, 0.12
Tooth Plate Area	28,27,37	-0.057 (0.033)	0.022 (0.037)	0.027 (0.042)	<0.00001, 0.86
Tooth Spacing	25,29,38	0.01 (0.012)	-0.004 (0.013)	-0.003 (0.009)	0.00005, 0.91

For each tooth trait, mean size-corrected phenotypic value with standard deviation (SD) is presented for each population (marine = RABS, F1 Hybrids = PAXBxRABS F1, benthic = PAXB). Sample sizes for marine, F1s, and benthic classes respectively are listed in “Sample Size” column. Animals from the benthic and marine time courses were selected to overlap the Hybrid F1 total length range (28-50mm). Fish size was adjusted for using a linear model in R. *P* values from Tukey’s posthoc test after an ANOVA are presented for the F1 hybrid comparisons to marine and benthic fish, respectively.

2.3.3 Genome-wide architecture of evolved changes in tooth number, tooth plate area, and spacing.

Previous work identified quantitative trait loci (QTL) controlling tooth number in a large genetic cross of 370 F2 fish derived from a Paxton benthic and a Japanese marine grandparent (15). When compared to Paxton benthic freshwater sticklebacks, Japanese marine sticklebacks, like Alaskan marine sticklebacks from the Rabbit Slough population, are low-toothed both in the wild and in the lab (Table 2.1). To test whether tooth number, tooth plate area, and spacing are genetically separable traits, we measured tooth plate area and tooth spacing in 272 F2 progeny of the Japanese Marine x Paxton benthic intercross. In the F2 progeny, tooth plate area and spacing are significantly correlated with tooth number, however tooth plate area and inter-tooth spacing are not correlated with each other (Figure 2.3). These results were robustly replicated in a Paxton benthic x Rabbit Slough marine F2 cross (Figure 2.3).

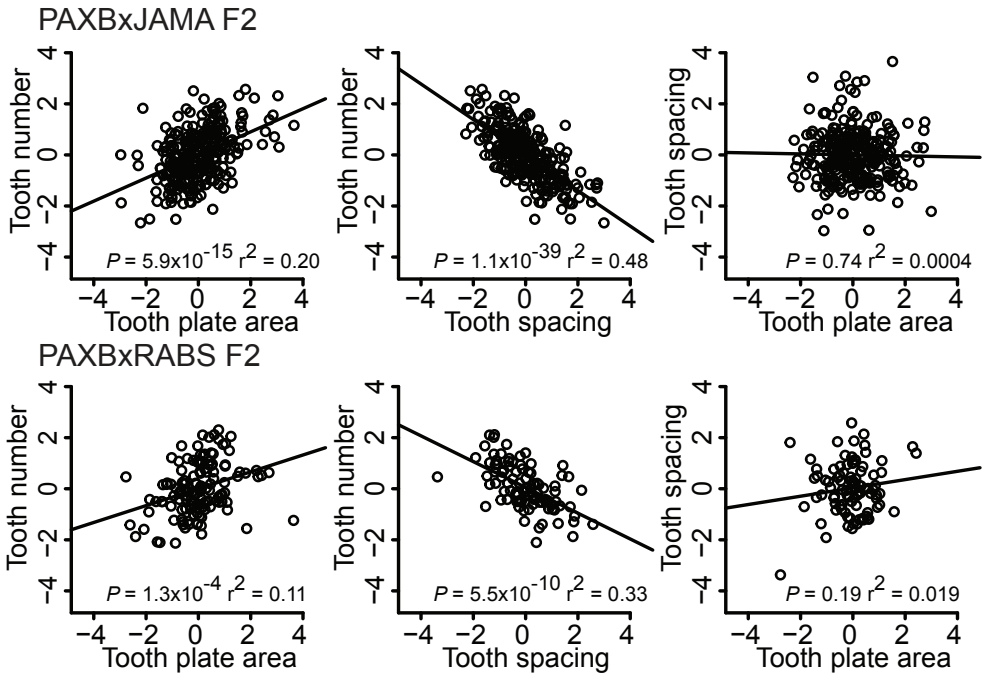


Figure 2.3. Correlations of tooth number, area, and spacing in two F2 genetic crosses

Pair-wise linear relationships for tooth number, tooth plate area, and tooth spacing for 272 fish from a Paxton benthic x Japanese Marine (top) and 142 fish from a Paxton benthic x Rabbit Slough Alaskan marine (bottom) F2 cross. Tooth number is significantly correlated with tooth plate area (left) and anti-correlated with tooth spacing (middle), suggesting that both area and spacing impact final tooth number. Conversely, tooth plate area and tooth spacing (right) are not correlated, suggesting that, despite each having an effect on total number, tooth plate area and spacing are genetically separable. The effect of fish length on each trait was removed using a linear regression and the residuals were z-scored in R. The *P* values (*P*) and r^2 values are shown for each comparison.

Furthermore, a principal component analysis of these three tooth traits revealed that tooth plate area and spacing load orthogonally onto the first principal component (Figure 2.4), suggesting that the genetic control of these tooth patterning phenotypes are at least partially separable.

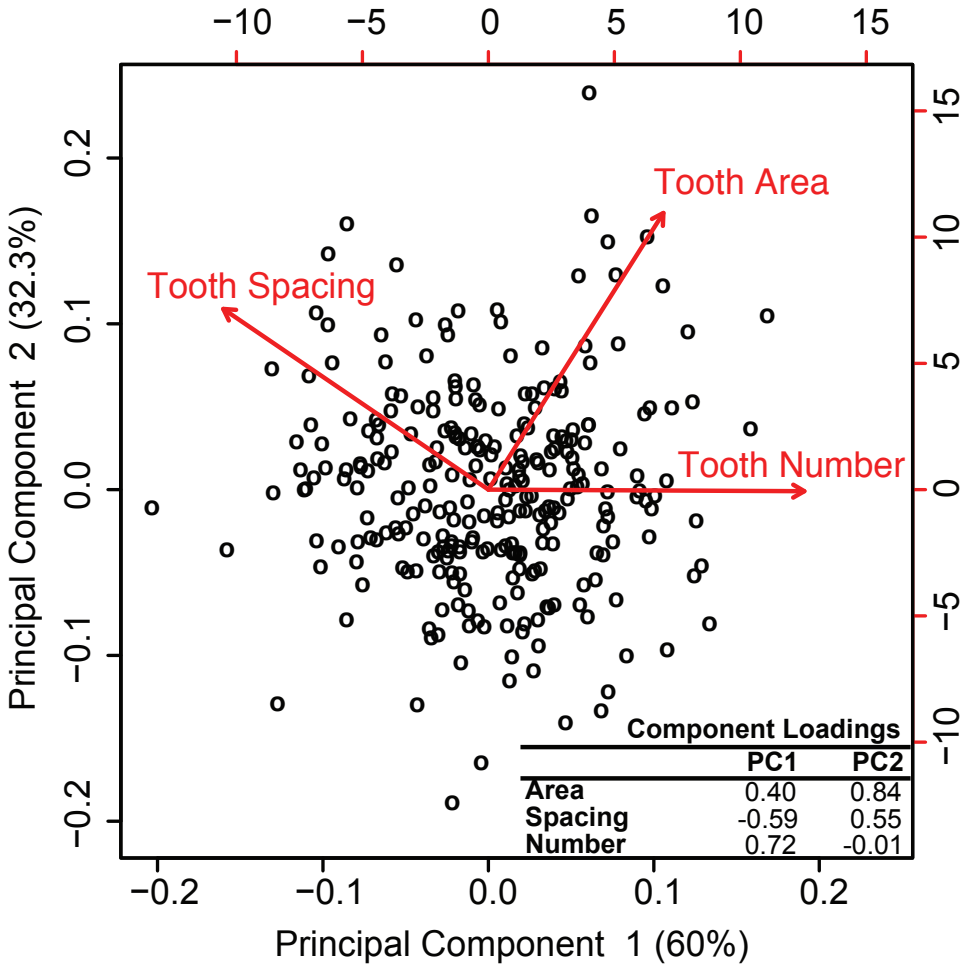


Figure 2.4. Biplot of principal component analysis of tooth patterning phenotypes
Scatter plot of the first two principal components from a principal component analysis of size-corrected tooth number, tooth plate area, and spacing phenotypes from 272 benthic by marine F2 fish. The first principal component (PC1) is explained primarily by variance in tooth number. Tooth area and spacing load orthogonally on PC1, suggesting genetic separability of area and spacing. The second principal component is explained largely by area and spacing with area loading stronger. Variable loadings are plotted for each trait in red and are presented in the inset. Percent variance explained is listed for each principal component on the axes.

To further test whether benthic sticklebacks have evolved more teeth by modifying the genetic programs controlling tooth plate area and/or tooth spacing, we used genome-wide linkage mapping to map loci controlling these three traits (Fig. 2.5). All three tooth phenotypes have a strong genetic component, with different chromosome regions having effects on one, two, or all three phenotypes. The QTL with the largest effect on tooth number maps to chromosome 21, to a region where benthic alleles confer not only more teeth, but also larger tooth plate area and smaller inter-tooth spacing (Fig. 2.5, Table 2.4). Other significant QTL had specific effects on one or two, but not all three, tooth phenotypes. For example, a chromosome 4 tooth number QTL overlaps a large spacing QTL but has no significant effect on tooth plate area. Conversely, QTL on chromosomes 1 and 7 control tooth plate area but had no significant effect on tooth number or spacing (Fig. 2.5, Table 2.4). Thus, the complex, polygenic architectures of tooth number, tooth plate area, and inter-tooth spacing are partially separable, and include a large effect QTL on chromosome 21controlling all three traits

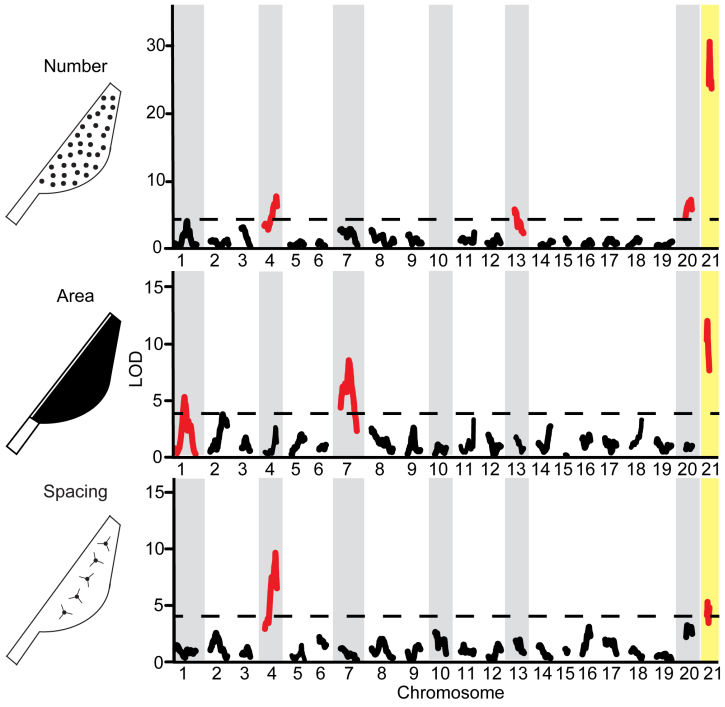


Figure 2.5. Genome-wide linkage mapping for tooth number, tooth plate area, and tooth spacing in a marine x benthic F2 genetic cross.
Genome-wide QTL mapping results for tooth number (A), tooth plate area (B), and inter-tooth spacing (C). All significant QTL are highlighted in red, and all chromosomes with significant effects on at least one tooth phenotype are shaded gray. The largest effect tooth QTL on chromosome 21 is highlighted in yellow. The dashed line is the significance threshold of $\alpha = 0.05$ determined by permutation tests (LOD scores of 4.1 for all three traits).

Table 2.4. Location and effect of Tooth Pattern QTL

Trait	LG	cM	Marker	LOD	MM	MB	BB	PVE
Tooth Number	4	60.4	Stn418	7.6	5.69	-0.52	-3.75	6.0
Tooth Number	10	11.7	Stn310	4.3	1.35	0.87	-3.80	3.3
Tooth Number	13	4.7	Stn153	6.0	-3.92	-0.05	3.39	4.6
Tooth Number	20	22.3	Gac1125	6.7	6.05	-1.39	-6.86	5.6
Tooth Number	21	4.0	Stn422	31.8	-9.93	-1.12	12.04	31.5
Tooth Spacing	4	57.0	Stn183	9.5	-0.013	-0.001	0.013	13.8
Tooth Spacing	21	3.2	Stn421	5.3	0.007	0.002	-0.010	7.3
Tooth Plate Area	1	48.9	Stn242	5.2	0.164	-0.010	-0.191	5.9
Tooth Plate Area	7	43.6	Stn79	8.7	-0.219	0.032	0.237	10.5
Tooth Plate Area	21	2.9	Stn222	11.9	-0.268	-0.005	0.293	14.7

For each QTL, Linkage Group (LG), genetic position in centimorgans (cM), non-interpolated peak marker name, peak LOD score, mean size-corrected residuals of phenotypic values for Marine homozygotes (MM), Marine-Benthic heterozygotes (MB), and Benthic homozygotes (BB), and percent variance explained (PVE) among the F2s are shown.

2.3.4 Bmp6 maps within the major effect QTL interval on chromosome 21.

The largest effect QTL controlling pharyngeal tooth patterning in this study maps to chromosome 21 and explains ~30% of the variance in pharyngeal tooth number (Figure 2.5) (15). To test whether this large effect tooth number QTL replicates in other wild-derived chromosomes from the Paxton lake benthic population, and to ask when during development this QTL acts, we analyzed an additional F2 genetic cross at three different developmental stages: before the tooth number divergence in the time course, around the time of divergence, and an adult stage after tooth number diverged in the time course. The results show robust replication of the chromosome 21 tooth QTL at the two later time points. Indeed, the effects of the chromosome 21 tooth QTL get more significant as fish develop (Figure 2.6), mirroring the late-onset developmental appearance of the tooth number differences (Figure 2.2).

To begin to understand the molecular genetic basis of evolved tooth gain, we fine-mapped this chromosome 21 QTL. We first genotyped 1004 additional F2 fish from the Paxton Benthic x Japanese marine cross (16), in order to identify individuals with recombination events within the original chromosome 21 tooth QTL (15). This genotyping identified 91 recombinant F2 fish, which were then selectively phenotyped for tooth number. These data localized the QTL to a 2.6 cM, 2.56 Mb 1.5 LOD candidate interval (Figure 2.7) containing 59 predicted genes, including one outstanding candidate gene (*Bmp6*) and one additional gene (*Tfap2a*) whose mammalian ortholog also has documented tooth expression (http://bite-it.helsinki.fi/). During tooth development, the *Bmp* pathway plays intimate roles in specifying tooth number, shape, and size (17-22), strongly motivating *Bmp6* as a candidate gene to underlie evolved tooth gain.

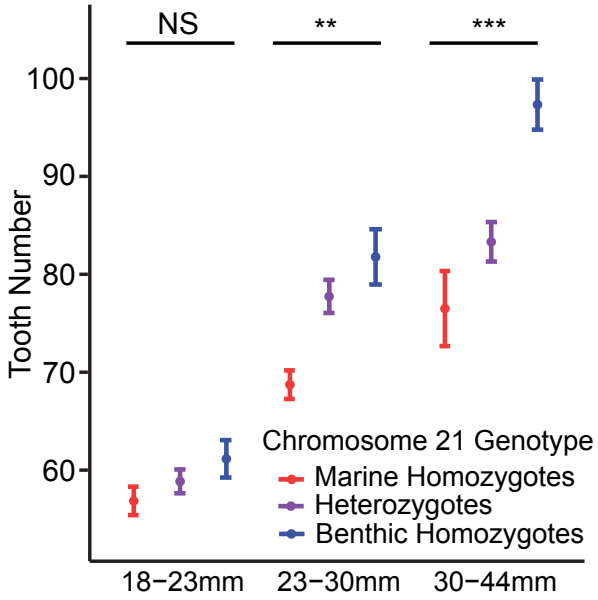


Figure 2.6. Developmental effects of the chromosome 21 tooth QTL

Effects of peak marker genotype on total tooth number (y-axis) in PAXB x RABS F2s at three different developmental time points (x-axis). The mean phenotypic value is shown for each genotypic class: marine homozygotes (red) heterozygotes (purple), and benthic homozygotes (blue). The effect of the QTL is not significant at the early larval time point, however at the later time points chromosome 21 genotype has significant effects on tooth number. The P values from a one-way ANOVA for each group are 0.32, 0.002, 0.0003, respectively. Error bars are standard error of the mean. Fish size is total length in millimeters (mm).

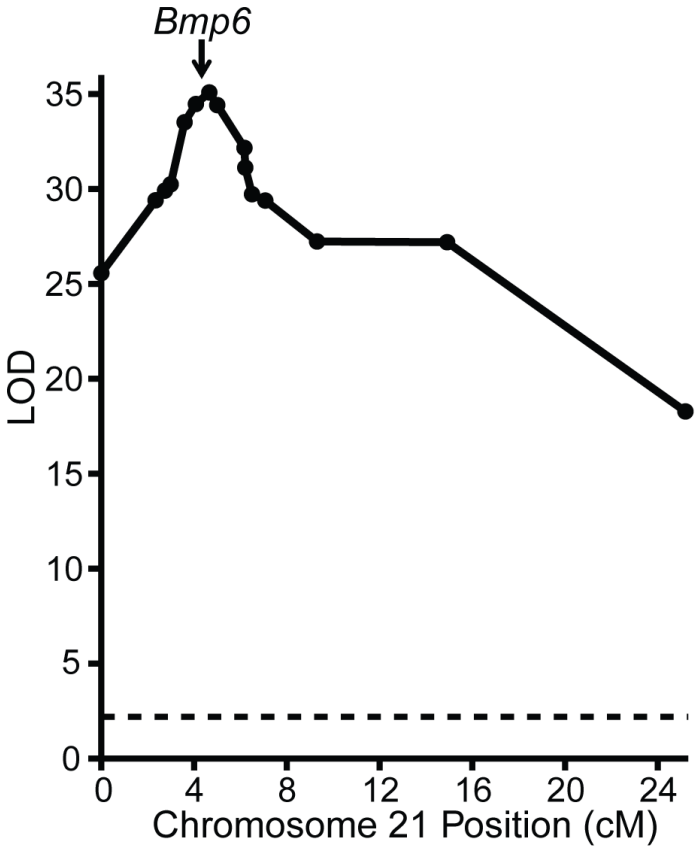


Figure 2.7. Fine mapping of the chromosome 21 tooth number QTL centers around *Bmp6*.
Genotype-phenotype association for genetic markers (circles) across chromosome 21 (x-axis). The position of *Bmp6* is marked with the arrow and the dashed line is the significance threshold of α = 0.05 determined by permutation tests (LOD score of 2.2).

2.3.5 *Bmp6* is expressed in developing stickleback teeth.

To test whether stickleback *Bmp6* is expressed in developing teeth, we performed *in situ* hybridization on marine and benthic embryos and larvae (Figure 2.8). We detected *Bmp6* mRNA in developing stickleback tooth germs, in addition to mRNA of two known markers for tooth development in fish and mammals: *Sonic hedgehog* (*Shh*) and *Pituitary homeobox 2*, *Pitx2* (Figure 2.8*A*-

C). In whole mount embryos, *Bmp6* expression at early tooth stages marked individual tooth germs similar to *Shh* and more restricted than *Pitx2,* which appeared to label the entire tooth-forming field (Figure 2.8*A -C*). Histological sections revealed that both *Shh* and *Pitx2* were expressed in epithelial cells (Figure 2.8*D-E*), similar to the epithelial expression of these two genes in developing teeth in mice and other fish species (19, 23-26).

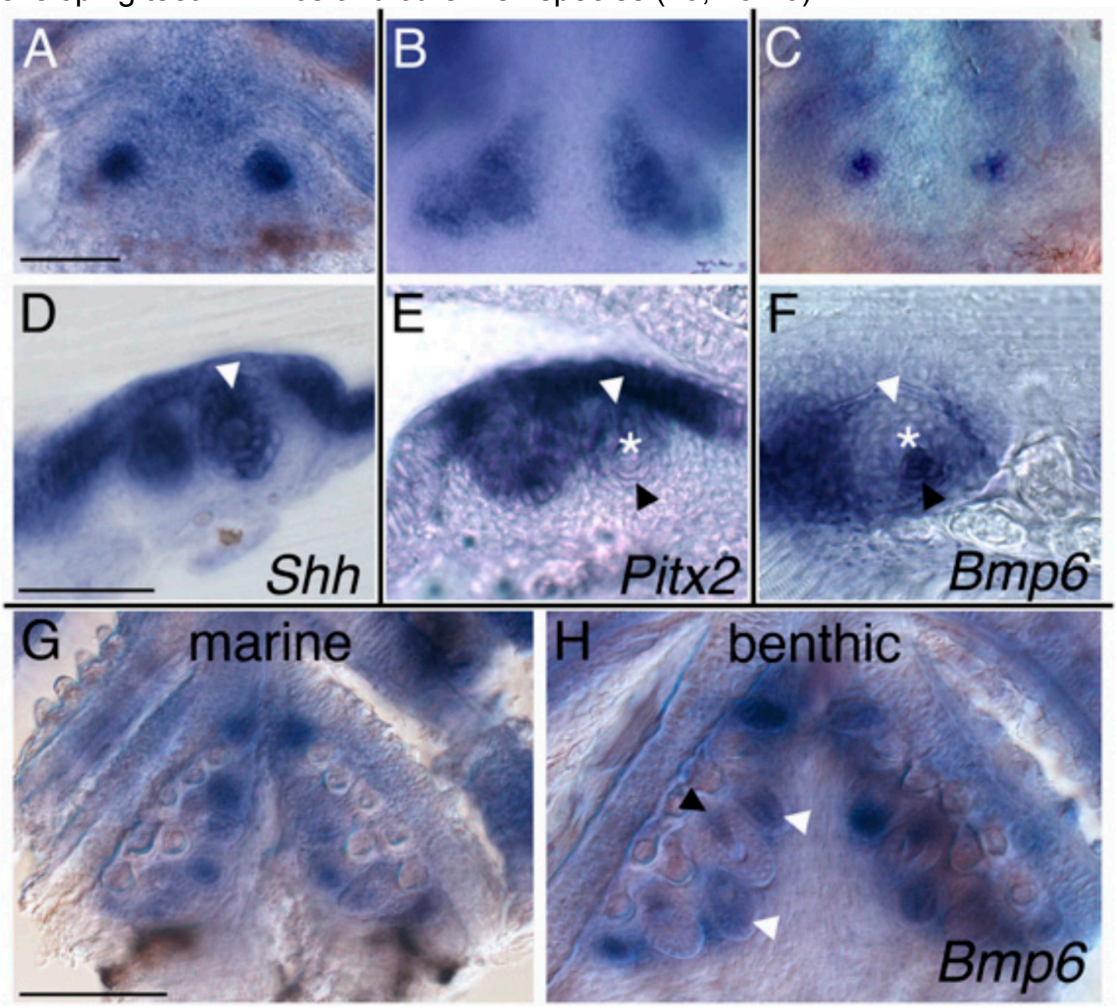


Figure 2.8. *Bmp6* is expressed in developing stickleback teeth.
Gene expression in developing benthic (A-F, H) and marine (G) stickleback teeth at 7.5 days post fertilization (dpf) (A-E) and 15 dpf (8 mm, F-H) revealed by in situ hybridization in whole mount (A-C, G-H) and 40 µm vibratome sections of comparably staged developing tooth germs (D-F, and see Figure S4). (A-F) Tooth markers *Shh* (A,B) and *Pitx2* (D,E) are detected in the odontogenic epithelium, while *Bmp6* is expressed dynamically in odontogenic epithelium early (C,H) and in odontogenic mesenchyme in newly ossifying teeth (F,H). (G-H) *Bmp6* continues to be expressed in teeth later in development in both marine and benthic larvae. White arrowheads: odontogenic epithelium, asterisks: newly mineralized developing teeth, black arrowheads: odontogenic mesenchyme. Scale bars = A-F = 50 uM, G,H = 100 µm.

In contrast, *Bmp6* expression was dynamically detected in odontogenic epithelial and mesenchymal cells (Figure 2.8, Figure 2.9) similar to most other *Bmp* genes in fish and mice (19, 27-30). During tooth development at larval stages, *Bmp6* showed complex but overall qualitatively similar expression patterns in marine and benthic fish (Figure 2.8*G-H*). Expression of *Bmp6* in newly developing teeth persists throughout later stages, including in putative replacement tooth germs (Figure 2.10). This *Bmp6* expression in developing teeth throughout embryonic and juvenile development supports the hypothesis that *Bmp6* underlies the chromosome 21 tooth QTL. In contrast, no expression of *Tfap2a* was detected in developing teeth (Figure 2.11).

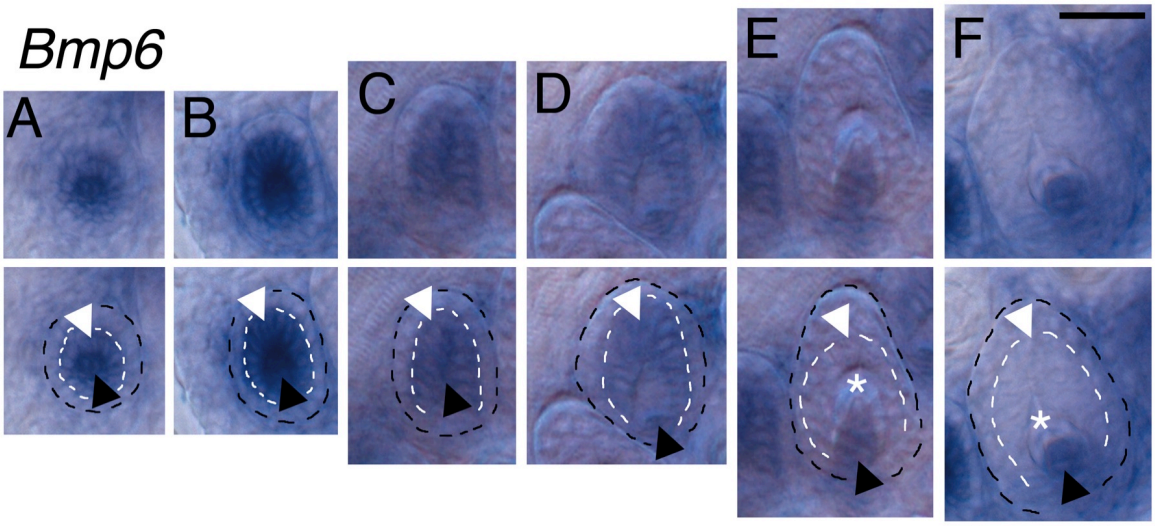


Figure 2.9. Dynamic *Bmp6* expression in developing teeth.
Successive stages of tooth development from left to right, showing early germ stage (A) through newly mineralized baby tooth (F). Unlabeled images at top, and labeled versions below. As the pharyngeal tooth field has multiple germs forming at different stages (see Figure 2.8G-H), example germs of different developmental stages from benthic fish 7.5-15 dpf are shown. In early germs (A), *Bmp6* expression is detected in inner dental epithelium (white arrowhead), and a rosette of odontogenic mesenchyme (black arrowhead). No expression is detected in the outer dental epithelium (area between white and black dotted lines). Slightly later in development (B-D), the tooth germ elongates and strong *Bmp6* expression is detected in the inner but not outer dental epithelium. As mineralization begins (E) and a baby tooth is formed (F), *Bmp6* expression is maintained in odontogenic mesenchyme, but is no longer detected in epithelia. For all germs, the black dashed line outlines the tooth germ, and the white dashed line outlines the boundary between the outer and inner dental epithelium. White arrowheads: inner odontogenic dental epithelium, asterisks: newly mineralized developing teeth, black arrowheads: odontogenic mesenchyme. Scale bar = 25 μm.

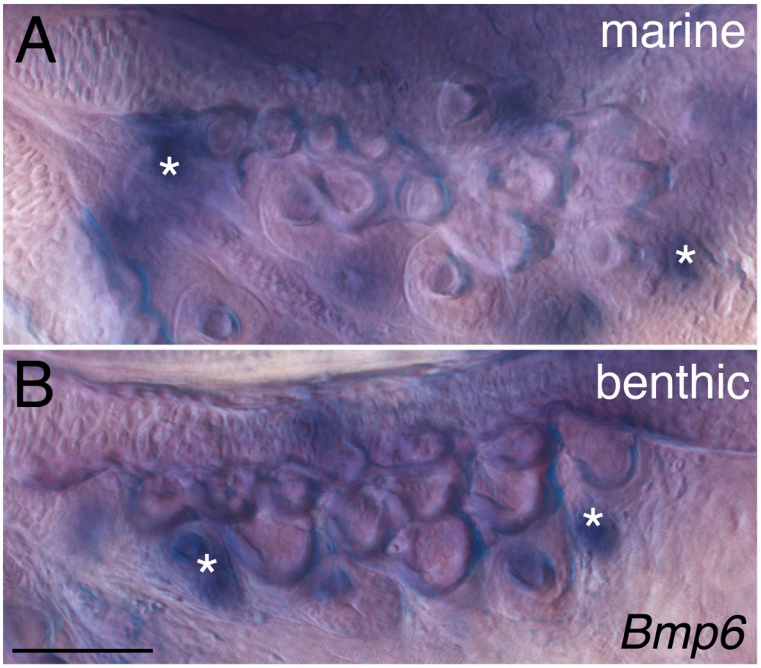


Figure 2.10. *Bmp6* expression during larval tooth development
In situ hybridization detecting *Bmp6* expression in developing pharyngeal teeth from (A) RABS marine and (B) PAXB benthic 30 day old (~10 mm total length) juvenile fish. *Bmp6* expression is detecting in developing tooth germs (white asterisks). Scale bar = 100 μm.

2.3.6 cis-Regulatory changes have lowered expression of the benthic *Bmp6* allele during tooth development.

We sequenced the exons of *Bmp6* in marine and benthic fish and found no nonsynonymous coding differences (Figure 2.12). To test for possible cis-acting regulatory differences in expression of marine and benthic alleles, we generated F1 hybrids between marine and benthic fish and used pyrosequencing assays to ask whether benthic and marine alleles made equal contributions to the overall level of *Bmp6* mRNA expression in F1 hybrid tooth plates. Allele-specific expression assays allow for the precise quantification of *cis*-regulatory differences between the two chromosomes in the same cells of the same fish in an identical *trans*-acting environment (31). We tested for a *cis*-regulatory change in *Bmp6* at three developmental time points, one prior to (larval), one during (juvenile), and one after (adult) the tooth number divergence. We detected no significant *cis*-regulatory

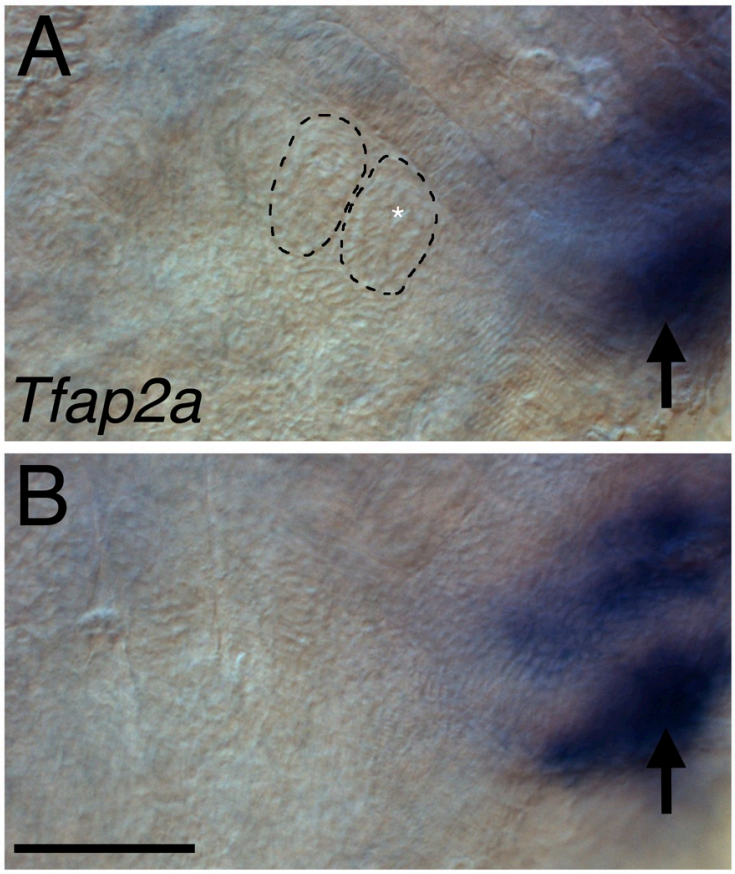


Figure 2.11. Expression of *Tfap2a* in 7.5 days post fertilization benthic fish. No *Tfap2a* expression was detected in developing stickleback tooth germs (A), although robust expression was detected in mesenchymal cells associated with the developing epibranchial cartilages. Shown are two focal planes, focused on (A) pharyngeal tooth germs and (B) dorsal pharyngeal mesenchyme (black arrow). In (A), two developing tooth germs are outlined with the black dashed lines, and a newly formed mineralized tooth is marked with the white asterisk. Scale bar = 50 μm.

difference in *Bmp6* at an early larval stage before the tooth number divergence in the time course (Figure 2.13). However, in both juveniles and adults, when tooth number differences are first being established and are further diverging between marine and benthic populations, we detected a highly significant allele-specific expression difference, with ~1.4-fold down-regulation of *Bmp6* expression from the benthic allele in F1 hybrid fish (Figure 2.13). This significant down-regulation of *Bmp6* at a later developmental stage mirrors both the late divergence in tooth number and the late-acting nature of the chromosome 21 QTL. These results support the hypothesis that a temporally regulated *cis*-regulatory difference in *Bmp6* expression drives the difference in tooth number between benthic and marine sticklebacks.

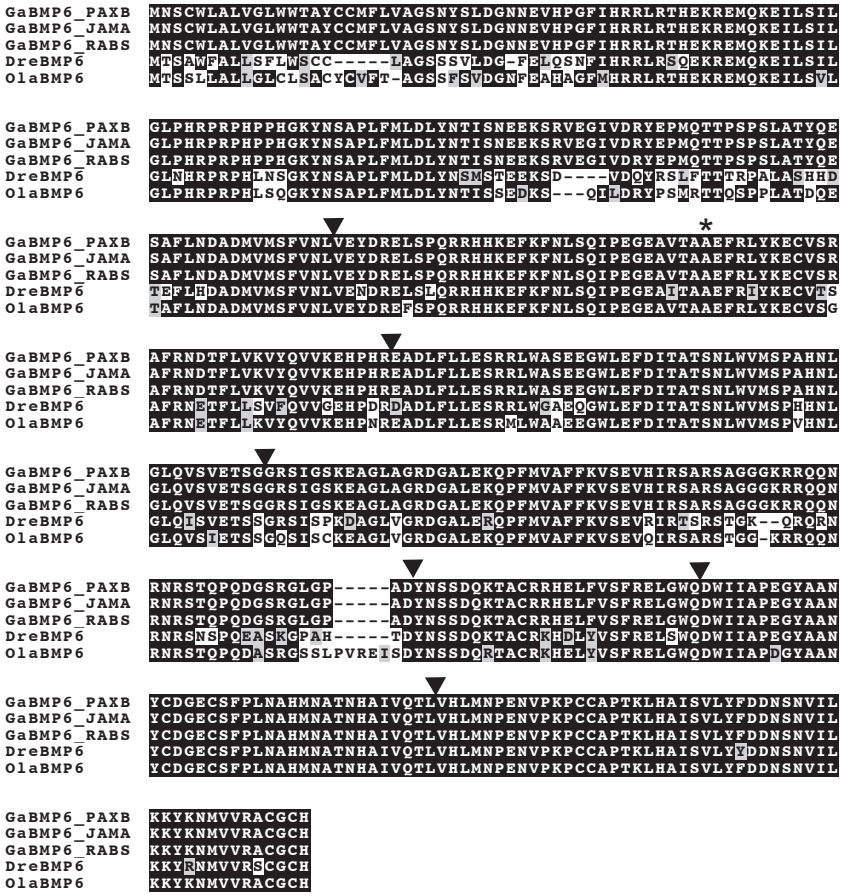


Figure 2.12. Predicted amino acid alignment of BMP6 in fish
BMP6 sequences of Paxton benthic freshwater (GaBMP6_PAXB), Japanese Pacific marine (GaBMP6_JAMA), and Rabbit Slough Alaskan marine (GaBMP6_RABS) are shown aligned to the BMP6 sequences of Zebrafish (DreBMP6) (Genbank accession number = NM_001013339.1) and Medaka (OlaBMP6) (Ensembl ENSORLT00000008205). Stickleback intron/exon boundaries are marked with arrowheads. The asterisk marks the position of the synonymous SNP used for the pyrosequencing assay. The predicted amino acid sequence is identical across the three stickleback populations.

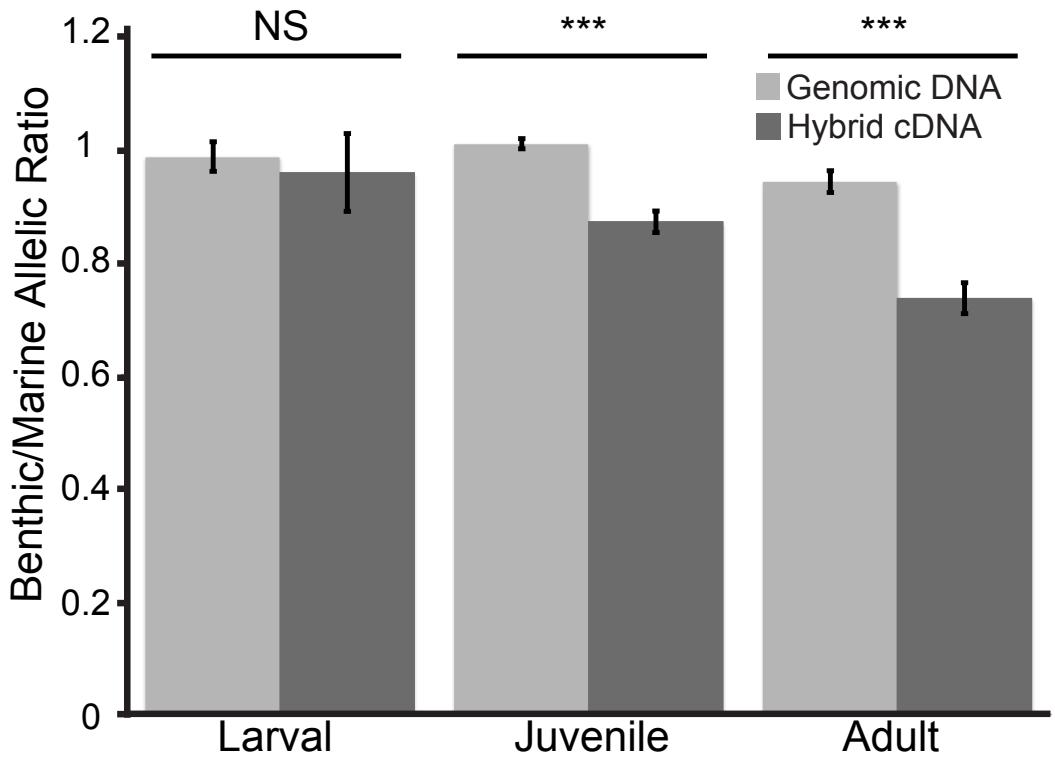


Figure 2.13. c*is*-regulatory down-regulation of the benthic allele of *Bmp6* in late not early stages of tooth development.
Shown are the ratios of benthic to marine alleles measured by pyrosequencing assays from either genomic DNA (light gray) or tooth plate cDNA (dark gray) from benthic x marine F1 hybrids at three different developmental stages. No significant difference in *Bmp6* expression was detected between marine and benthic alleles at the larval stage (left), but at the juvenile (middle) and adult stage (right) the benthic allele was significantly down-regulated (sample sizes and *P* values by Wilcoxon signed rank test for early, juvenile, and adult are n=12, *P* = 0.27, n=18, *P* = 0.0003 and n=13, *P* = 0.0005 respectively). Error bars are standard error of the mean.

2.4 Discussion

Our studies show that Paxton benthic freshwater sticklebacks have evolved major changes in tooth number, tooth plate area, and inter-tooth spacing that arise relatively late during development. Since sticklebacks, like most teleosts, retain the basal vertebrate condition of polyphyodonty (continuous tooth replacement) (32), the late divergence in tooth number could result from a change in the rate of the tooth regeneration program late in development, once the initial tooth pattern has been established. This late-forming increase in tooth number may match the time period when benthic fish begin to benefit from increased tooth number (i.e. perhaps wild benthic larvae do not normally begin exploiting a benthic diet until about 20-25 millimeters in length). Alternatively, developmental or genetic constraints may lead to late-forming divergence. For

example, altering the tooth developmental program at earlier stages may lead to deleterious pleiotropic consequences, or available standing genetic variation might primarily affect late not early development.

Although our lab-reared data show that major differences in tooth number are maintained between marine and freshwater fish when reared in a common lab environment, tooth numbers in both populations are reduced in lab-reared fish compared to wild fish. Differences in chronological age likely contribute to this difference, since wild fish are likely at least one year old, while our lab-reared adults were six months old. In addition, tooth number may be influenced by diet and rearing conditions, as has previously been reported in cichlids (33).

Previous quantitative genetic studies of stickleback pharyngeal tooth number revealed five quantitative trait loci (QTL) controlling ventral pharyngeal tooth number in a F2 genetic cross between an ancestral low-toothed Japanese marine fish and a derived high-toothed Paxton benthic freshwater fish (15). Our more detailed studies suggest that differences in total adult tooth number arise from a combination of several factors, including changes in the development programs controlling tooth number, the size of the tooth field, and the spacing of teeth within that field. This conclusion is supported by the statistical relationships between tooth number, area, and spacing in the F2 cross, and by the genome-wide linkage mapping results of all three phenotypes. We have identified at least seven quantitative trait loci (QTL) that have significant effects on tooth number, tooth plate size, or tooth spacing. Different QTLs affect one, two, or three different tooth phenotypes (tooth number, tooth spacing and tooth plate size), showing modular control of evolved changes in dental patterning.

In other fish, pharyngeal jaw patterning is correlated with dietary niche, likely due to adaptive advantages of different morphologies in feeding success on different diets (9). Because benthic fish are well described as having trophic specializations for eating benthos (13), we hypothesize that the evolved tooth gain in benthic sticklebacks is also an adaptive trait that has been selected during an ecological shift to a benthic diet. We note that of the seven tooth patterning QTL, only three go in a direction that is concordant with the overall shift in tooth number in the parental populations (i.e. benthic alleles conferring more teeth) based on the developmental time courses. However, the QTL with largest phenotypic effect on chromosome 21 does act in a direction that is consistent with the overall trend in tooth number in the parental populations (benthic allele conferring more teeth). Perhaps the smaller effect QTL that have effects in the opposite direction result from chromosome 21's effect overshooting the adaptive peak for tooth patterning in this recently evolved population, with other loci evolving to bring tooth patterning closer to the adaptive peak (34). The mixed direction of effects of benthic alleles could alternatively result from pleiotropy (35), with QTL controlling other adaptive benthic phenotypes that might secondarily affect tooth patterning. For example, the large effect tooth spacing QTL on chromosome 4 overlaps the *Ectodysplasin (Eda)* gene which controls adaptive reductions in armor plate patterning (36), and is also well known to affect vertebrate tooth patterning (37, 38). Interestingly, *Eda* also plays a role in the spacing of hair placodes and tooth cusps in mice (39, 40), making

Eda an excellent candidate for underlying the tooth spacing QTL on chromosome 4. A third possibility is that some or all of these tooth traits could be changing due to genetic drift occurring after freshwater colonization. As several other species pairs and hundreds of other freshwater populations with trophic modifications have been described (12, 13, 41-43), one test of adaptive significance will be to ask whether other derived benthic lake or creek freshwater stickleback populations have also evolved increases in tooth number. Molecular genetic identification of the tooth patterning QTL that are segregating in the current cross, combined with population genetic tests of molecular variation surrounding causal loci, should also help distinguish these models.

To begin to study the molecular mechanisms behind evolved tooth gain, we fine-mapped the largest effect tooth number QTL on chromosome 21. A previous study identified a cluster of QTL on chromosome 21 controlling several derived freshwater skeletal traits (15). This QTL cluster mapped near a large genomic inversion previously shown to display strong worldwide patterns of divergence between marine and freshwater populations (44), suggesting that multiple phenotypes may be controlled by linked genetic changes within the chromosome inversion. Interestingly, we find that the 1.5 LOD candidate interval for the chromosome 21 tooth QTL maps over 1.5 Mb from the inversion, strongly suggesting that the molecular changes driving tooth gain map outside the inverted region.

The new fine-mapped interval for the tooth QTL contains an excellent candidate gene, *Bone morphogenetic protein 6* (*Bmp6*). We show that *Bmp6* is expressed in developing teeth in marine and benthic sticklebacks, has no predicted coding changes between populations, but has a late-onset *cis*-regulatory down-regulation in benthic fish. Since in other vertebrates, BMPs can act as activators and inhibitors of tooth development (45), we hypothesize that the lowered *Bmp6* expression observed in benthic sticklebacks contributes to their increased tooth number controlled by the chromosome 21 region. Bone Morphogenetic Proteins were originally identified based on their remarkable ability to induce ectopic bone when implanted at new sites in animals (46). Thus increases in tooth plate area could also result from lowered *Bmp6* expression. The divergence in tooth number and *Bmp6 cis*-regulation at late not early developmental stages might reflect a heterochronic shift in the benthic population, where the benthic tooth development and replacement program is “stuck” in the early rapid tooth-generating phase observed in early larval stages in both marine and benthic fish. While we parsimoniously favor the hypothesis that *Bmp6* underlies the evolved differences in tooth number, tooth plate area, and inter-tooth spacing, we note that the fine mapping was only done for tooth number, so it is possible that other genes underlie the evolved changes in tooth plate area and inter-tooth spacing.

The use of BMP ligands as major drivers of morphological evolution in vertebrates is striking. BMP family members have been implicated in several vertebrate evolved traits: size and shape of the beak in Darwin’s finches, size and shape of the jaw in cichlids, jaw and skull variation in brachycephalic dogs, and avian feather patterning (47-50). Although based on a limited number of

reported cases and possibly affected by ascertainment biases, this apparent reuse of the same signaling pathway across taxa may reflect a predisposition for *Bmp* genes to be used during morphological evolution, perhaps due to having complex, modular *cis*-regulatory architecture to generate evolutionary variation (51, 52).

Previous QTL mapping studies in sticklebacks have shown that major changes in pelvic hindfin development, armor plate formation, and body pigmentation are all due to alterations in key developmental signaling molecules and transcription factors (36, 53-55). In each of these previous cases, freshwater fish have evolved a major loss or reduction of skeletal structures that were originally present in marine ancestors. In all three cases, *cis*-regulatory changes are implicated, either directly (53, 54) or inferred (36). Here we show that a major gain in tooth number can also be genetically mapped to a relatively small number of chromosome regions. The QTLs with largest effects on tooth number control somewhat less of the overall variance than the previously identified QTL for armor plates, pelvis, and pigment (each of which controls 50% or more of the variance in the corresponding trait). Nevertheless, the overall effects of the tooth patterning QTLs are still quite large compared to classical predictions of nearly infinitesimal effects for genetic changes underlying evolved differences in natural populations. Finally, our results with *Bmp6* show that for both loss and gain traits, the chromosome regions with largest phenotypic effects show clear evidence of cis-acting regulatory changes in key developmental control genes. Although many more case studies will be needed to draw general conclusions, collectively these studies suggest that similar general principles may underlie the evolution of both loss and gain traits, and that regulatory changes in developmental control genes play an important role in both regressive and constructive evolution of the vertebrate skeleton.

2.5 Methods

Stickleback husbandry. Lab-reared fish were raised in 29 gallon tanks under common conditions (3.5g/l Instant Ocean salt, 0.4 ml/l NaHCO3) and fed live brine shrimp as larvae, then frozen daphnia, bloodworms, and Mysis shrimp as juveniles and adults. All experiments and field collections were done with the approval of the Institutional Animal Care and Use Committee from UC-Berkeley, Stanford University, or the University of British Columbia.

Skeletal morphology visualization and quantification
Adult lab-reared fish were fixed in 4% paraformaldehyde overnight, washed in H_2O, stained with 0.008% Alizarin Red in 1% KOH overnight, washed in H_2O, then cleared in 50% glycerol and 0.25% KOH. Branchial skeletons were dissected, cleared, and mounted, and Alizarin Red fluorescent teeth were quantified on a DM2500 Leica microscope using a TX2 filter. Significant differences between the populations in both lab and wild datasets were tested using two-tailed *t* tests. Wild PAXB and RABS fish, and F2 chromosome 21

recombinants from the PAXB x JAMA F2 cross were scanned unstained in ethanol using a Scanco uCT40 microcomputerized tomographer at 55 kVp at high resolution, averaging four frames, and teeth counted from digital volumes of ventral pharyngeal tooth plates.
For time courses, total fish length was measured after overnight fixation in 4% paraformaldehyde. Fish at different stages were stained with Alcian Blue and Alizarin Red stained using 100mM $MgCl_2$ as described or only with 0.008% Alizarin Red in 1% KOH and then dissected, mounted, and total (left plus right) ventral pharyngeal tooth number quantified as above. For both the area and spacing measurements, grayscale images of Alizarin red fluorescence of bilateral ventral tooth plates were acquired with a DFC340 FX camera on a Leica M165FC dissecting microscope using a rhodamine filter. The periphery of the tooth plates for both left and right sides was outlined excluding teeth not connected to the tooth plate and area was calculated using ImageJ. The average area for both plates is shown. The spacing measurements were calculated by placing a landmark on each tooth position on the left/right ventral pharyngeal tooth plate and measuring the distance to the closest three neighboring teeth using ImageJ and a custom Python script. The average of the bilateral spacing measurements is presented. Animals were only included in the spacing and area analysis if clear measurements could be made.

QTL Mapping. QTL mapping was done using R/qtl (56). To map QTL for adult tooth number, area, and spacing, we analyzed a subset (n=272 fish) of a previously described (16) Paxton Benthic and Japanese Marine F2 cross. 275 microsatellite markers were genotyped in each F2. Tooth number, area, and spacing were quantified in each F2. As all three traits were significantly correlated with fish total length, residuals from a linear regression were used for each of the three traits. We performed *scanone* in R/qtl (2) with Haley-Knott regressions to initially map QTL. For each phenotype, we performed one thousand permutations with *scantwo* to calculate the trait-specific LOD threshold at which α = 0.05. Conservatively, we used the highest of these LOD thresholds (4.1) for the significance threshold for all 3 traits. A forward-backward search was performed with *stepwiseqtl* to iteratively identify significant QTL with a main penalty of 4.1 and to identify the best fitting QTL model. Each QTL identified using *scanone* was also detected using *stepwiseqtl*. We calculated peak LOD and position for each QTL using *refineqtl* and percent variance explained with *fitqtl*. The LOD scores for chromosomes that did not have a significant effect in Figure 3 were determined with *addqtl*.

Fine mapping tooth number QTL
To fine map the chromosome 21 tooth QTL, 1004 F2s from four additional families from the initial mapping cross were genotyped for the two-LOD boundary markers (Stn484 and Stn491) of the initial chromosome 21 tooth QTL to screen for recombinants. All recombinants were then genotyped for a set of polymorphic microsatellite markers (Table S5) across the QTL interval. These combined genotypes were used to make a linkage map for chromosome 21 with JoinMap

4.0 (Kyazma). The effects of fish size and family on total ventral pharyngeal tooth number were corrected for using a linear model in R (www.r-project.org). In the two largest families, the effects of genotypes at two previously described unlinked tooth number QTL on chromosomes 4 and 20 were corrected for using *fitqtl* and genotypes at Gac4174 and Stn183 (chromosome 4) and Stn340 (chromosome 20). QTL mapping was done in the pooled dataset using *scanone* followed by *refineqtl* and *fitqtl*. *scanone* was used to perform 1000 simulations to calculate the LOD threshold (2.2) at which α = 0.05. Gene content of the 2.56 Mb fine mapping interval was determined by counting genes predicted by either the ENSEMBL gene prediction or tBLASTn human protein track on the UCSC genome browser (http://genome.ucsc.edu/) for the interval between chromosome 21: 2,564,997-5,120,542 base pairs in the stickleback genome assembly.

Mapping QTL in developmental time points of F2 cross

Lab-reared Paxton benthic (British Columbia, Canada) and Rabbit Slough (Alaska) fish were crossed by artificial fertilization and the F1s were intercrossed to generate F2s. F2s were sacrificed at 80 days post fertilization (dpf), 120 dpf, and adults (total n=142). These F2s were genotyped for the PAXB x JAMA F2 cross peak marker Stn489 (see Table 1.5 for primer sequences). The analyzed PAXB x RABS F2s total lengths ranged from 18-44mm. Animals were binned into three total length bins ranges so there were approximately equal numbers of animals in each bin (n=47-48) to compare across developmental time points. For the middle, but not early or late time points, tooth number significantly fit a linear regression to fish total length, so this effect was corrected for by taking the residuals of a linear regression to fish total length, followed by back-transforming values to fish with a length of 26.5 mm, the midpoint of the length bin. Effects of chromosome 21 genotype were tested using a one-way ANOVA in R.

***Bmp6* sequencing**

The seven predicted exons of *Bmp6* were amplified using RT-PCR from PAXB and RABS adult tooth plate cDNA using 5' UTR forward primer 5'-CTGCAGCTCCAAGAGAGACC -3' and 3'UTR reverse primer 5'-CTTTGCAAACCCCAACTTGT -3'. These primers amplified a ~1.3 kb PCR product, which was gel extracted (Qiagen kit) and sequenced, resulting in the predicted coding sequence shown in Fig. S7. To generate full exonic sequences of *Bmp6* from the different populations, the identified exons were PCR amplified from genomic DNA from PAXB, JAMA, and RABS fish. The reaction profile was 95^{o}C for 3 m, 34 cycles of 95^{o}C for 15 s, 56^{o}C for 15 s, and 72^{o}C for 30 s, followed by 3 m at 72^{o}C. The PCR fragments were purified using PCR purification kit (Qiagen) and sequenced. The primers used were: exon 1 5'-TAAGGGACTGCAGCTCCAAG-3' and 5'- GAAGTTCAACGATGACGATT -3'; exon 2 5'- GTGTGTGTTTCCATGCCACAG-3' and 5'-GAATCCACTCAAAGCTTCTT-3'; exon 3 5'- AAGTTGGGCTGCAGTTGTTT-3' and 5'- CGCGTGAGCTGGATCTCTTA -3'; exon 4 5'-CAACCTGTGGGTGATGAGC -3' and 5'-TCCTCTGTGCAACGAAACTG-3'; exon 5 5'- CTCCGAGCCTCTCTCTAGCA -3' and 5'-TCATATGCGTCAGAGGATGG-

3’; exon 6 5’-GCAGTTTGTCATCCAGCTGTT-3’ and 5’-AAGTCATGGCAAAGACGTG-3’; exon 7 5’-CTCGCTATACCAAACGTGAC-3’ and 5’-GATTTAAACCGGGAGTCTAGC-3’. Genbank accession numbers of *Bmp6* sequences from Rabbit Slough and Paxton benthic mRNA and genomic DNA are KM406380-KM406383.

In situ hybridization.
Marine and benthic embryos and larvae were sacrificed, fixed overnight in 4% paraformaldehyde in 1xPBS, then dehydrated and stored at -20^{o}C in methanol. For larvae older than 9 days post fertilization, ventral tooth plates were dissected after rehydration from methanol and prior to in situ hybridization. In situ hybridization was performed essentially as described. but with in situs done in tubes in a water bath not baskets, and using a two day hybridization for older larval stages. For sections, whole-mount in situs were fixed overnight in 4% paraformaldehyde in 1x PBS, embedded in gelatin-albumin cross-linked with 1.75% glutaraldehyde, and sectioned at 40 microns on a Pelco 101 Vibratome Series 1000**.** The plasmids used to synthesize the *Bmp6, Pitx2, and Shh* riboprobes were made by cloning RT-PCR amplicons into pBSII-SK+, amplified off random hexamer primed cDNA made from RNA of newly hatched (~8.5 dpf) Little Campbell marine fry. Amplicons were amplifying with the following primers (with added 5’ restriction enzyme cut sites underlined): *Bmp6* (~700bp) 5’-GCCGCTCGAGATGAACAGCTGCTGGCTTG-3’ and 5’-GCCGTCTAGACTCATCACCCACAGGTTGC-3’, *Pitx2* (~750bp) 5’ GCCGTCTAGACCTCAGTAACCCGTCTCTCAA-3’ and 5’-GCCGGGGCCCAAGCAGGCCTGGGTTCAT-3’, *Shh* (~720bp) 5’-GCCGCTCGAGCGGGAGCAAAATGAGACCTA-3’ and 5’-GCCGTCTAGAATGCAGACATGAGGCAGAAT-3’. Resulting amplicons for each gene were then digested with XhoI and XbaI (*Bmp6* and *Shh*) or XbaI and ApaI (*Pitx2*) to generate sticky ends to directionally clone into pBSII-SK(+). The resultant plasmids were linearized with XhoI (*Bmp6* and *Shh*) or XbaI (*Pitx2*) and antisense riboprobes transcribed using T3 polymerase for the *Bmp6 and Shh* probes or T7 polymerase for the *Pitx2* probe. The plasmid used to make the *Tfap2a* riboprobe was generated by first using the primers 5'-ATGGGAACTATTGCCAGCAC-3' and 5'-ACGAAGCGAAAAGAGGATGA-3' to amplify a ~760 bp amplicon by PCR off Little Campbell marine genomic DNA which was then TOPO TA (Invitrogen) cloned into pCR2.1. The resultant plasmid was linearized with HindIII and antisense riboprobe transcribed with T7 polymerase. ProbeDB sequences for Bmp6, Pitx2, Shh, and Tfap2 riboprobes are Pr032250589, Pr032250590, Pr032250591, and Pr032250592, respectively.

Pyrosequencing of the F1 Hybrids
For allele-specific expression experiments, Paxton Benthic freshwater fish were crossed with Rabbit Slough marine fish by in vitro fertilization to generate hybrid F1s. Hybrid ventral pharyngeal tooth plates were dissected on ice from larval, juvenile, and adult stages (~10-20mm, ~25-40mm, and >40mm in total length, respectively). RNA was isolated using TRI Reagent (Ambion) and cDNA reverse

transcribed with random hexamer primers using Superscript III (Invitrogen). For genomic DNA controls, total genomic DNA was isolated by phenol-chloroform extraction from caudal tail tissue. The second exon of *Bmp6* was amplified using an HPLC-purified 5' Biotin-labeled primer 5'-TTTGACTAGGAAGGTGTCGTT-3' and reverse primer of 5'-TCTCACAGATCCCAGAGGGC-3', which flank a synonymous SNP. PCR reactions of 20 uL used Phusion polymerase (NEB) and non-diluted cDNA or 10 total ng genomic DNA as template. The PCR conditions were 98^{o}C for 30 s, 34 cycles of 98^{o}C for 15 s, 56^{o}C for 15 s, and a 20 s extension at 72^{o}C followed by 5 m at 72^{o}C. The pyrosequencing was done by Epigendx (Hopkins, MA) with the sequencing primer 5'- CTTGTAGAGGCGGAAT -3'. Relative expression levels were calculated by the ratio of the expression level of the benthic allele to the marine allele. Significant differences in expression ratios between the genomic DNA controls and cDNA experimental groups were determined using a two-tailed t-test.

Table 2.5. Chromosome 21 microsatellites used for fine mapping

Stn Marker	Forward primer sequence (5' to 3')	Reverse primer sequence (5' to 3')	5' fluorophore	Accession number
Stn484	TGCAAAGCAAGTGTAACGAA	CTTCCCTTTGTCCGTTCCT	FAM	Pr032250564
Stn485	AGCAGGTGAAAGTGTGTATAACG	AGGGCACTATGGTTTCACAGG	FAM	Pr032250565
Stn486	CACAAGCCTTTGTGTTGGTG	GAAACGGGATTCTTTGACCA	NED	Pr032250566
Stn487	CACGGCAAACAGGTGAGAC	TCGATGGGCTGTAAATCCTC	NED	Pr032250567
Stn488	AATTACACTGCCTGCACTTGG	GTCAGATGGACGGACAGACG	NED	Pr032250568
Stn489	AGTGACGAATCCCTCTTCTGC	CACACCTTGTTGTGTTTGTAGC	FAM	Pr032250569
Stn490	ATGAGGTCACCCTGCCTAAC	CGCCTGTCATATACACATTGC	FAM	Pr032250570
Stn491	AACGTTAACCAGTTGCAGTCC	GATGTCGACACAGAATCTCTTAGC	HEX	Pr032250571
Stn492	CCGTATGCAGCCTGTTGG	AACCTGACCCTCCTCTGACC	FAM	Pr032250572
Stn493	ACGCCTTCCTCGATCAGACC	TTGTTACGCGTTCGTAGAGC	FAM	Pr032250573
Stn494	CTCTACTGCGCACGCTTAGG	GTCACATTTCCTCGGTTTGC	FAM	Pr032250574
Stn426	TTCCTCTTCTTCACAGGCTGA	CCTTTGATCCGCACAGTCA	FAM	Pr032250575

Primer sequences of new Stn microsatellite markers used for tooth QTL fine mapping. All forward primers were directly labeled with 5' fluorophores of FAM or HEX (IDT) or NED (ABI). Accession numbers in ProbeDB for each marker are listed.

2.6 References

1. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134(1):25-36.
2. Stern DL (2000) Evolutionary developmental biology and the problem of variation. *Evolution* 54(4):1079-1091.
3. Protas M*, et al.* (2008) Multi-trait evolution in a cave fish, *Astyanax mexicanus*. *Evol Dev* 10(2):196-209.
4. Wark AR*, et al.* (2012) Genetic architecture of variation in the lateral line sensory system of threespine sticklebacks. *G3* 2(9):1047-1056.
5. Werner T, Koshikawa S, Williams TM, & Carroll SB (2010) Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* 464(7292):1143-1148.
6. Jernvall J & Thesleff I (2012) Tooth shape formation and tooth renewal: evolving with the same signals. *Development* 139(19):3487-3497.
7. Lauder GV (1983) Functional Design and Evolution of the Pharyngeal Jaw Apparatus in Euteleostean Fishes. *Zool J Linn Soc-Lond* 77(1):1-38.
8. Wainwright PC (2005) Functional morphology of the pharyngeal jaw apparatus. In: Biomechanics of Fishes. (R. Shadwick and G. V. Lauder, eds) Academic Press.
9. Muschick M, Indermaur A, & Salzburger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol* 22(24):2362-2368 .
10. Kingsley DM & Peichel CL (2007) The molecular genetics of evolutionary change in sticklebacks. Biology of the three-spine stickleback. Ostlund-Nilsson, S., Mayer, I., and Huntingford, F.A. CRC Press: 41-81.
11. Bell MA & Foster SA (1994) *The evolutionary biology of the threespine stickleback* (Oxford University Press, Oxford ; New York) pp xii, 571 p.
12. Caldecutt WJ, Bell MA, & Buckland-Nicks JA (2001) Sexual dimorphism and geographic variation in dentition of threespine stickleback, *Gasterosteus aculeatus*. *Copeia* (4):936-944.
13. Schluter D & Mcphail JD (1992) Ecological character displacement and speciation in sticklebacks. *American Naturalist* 140(1):85-108.
14. McPhail JD (1992) Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): evidence for a species pair in Paxton Lake, Texada Island, British Columbia. *Can. J. Zool.* 70:361-369.
15. Miller CT*, et al.* (2014) Modular Skeletal Evolution in Sticklebacks Is Controlled by Additive and Clustered Quantitative Trait Loci. *Genetics* 197(1):405-20.
16. Colosimo PF*, et al.* (2004) The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol* 2(5):e109.
17. Andl T*, et al.* (2004) Epithelial Bmpr1a regulates differentiation and proliferation in postnatal hair follicles and is essential for tooth development. *Development* 131(10):2257-2268 .
18. Bei M, Kratochwil K, & Maas RL (2000) BMP4 rescues a non-cell-autonomous function of Msx1 in tooth development. *Development* 127(21):4711-4718.

19. Fraser GJ, Bloomquist RF, & Streelman JT (2013) Common developmental pathways link tooth shape to regeneration. *Dev Biol* 377(2):399-414.
20. Jia S*, et al.* (2013) Roles of Bmp4 during tooth morphogenesis and sequential tooth formation. *Development* 140(2):423-432.
21. Vainio S, Karavanova I, Jowett A, & Thesleff I (1993) Identification of BMP-4 as a signal mediating secondary induction between epithelial and mesenchymal tissues during early tooth development. *Cell* 75(1):45-58.
22. Wang Y*, et al.* (2012) BMP activity is required for tooth development from the lamina to bud stage. *J Dent Res* 91(7):690-695.
23. Bitgood MJ & McMahon AP (1995) Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. *Dev Biol* 172(1):126-138.
24. Fraser GJ, Berkovitz BK, Graham A, & Smith MM (2006) Gene deployment for tooth replacement in the rainbow trout (Oncorhynchus mykiss): a developmental model for evolution of the osteichthyan dentition. *Evol Dev* 8(5):446-457.
25. Fraser GJ, Graham A, & Smith MM (2004) Conserved deployment of genes during odontogenesis across osteichthyans. *Proc Biol Sci* 271(1555):2311-2317.
26. Mucchielli ML*, et al.* (1997) Mouse Otlx2/RIEG expression in the odontogenic epithelium precedes tooth initiation and requires mesenchyme-derived signals for its maintenance. *Dev Biol* 189(2):275-284.
27. Aberg T, Wozney J, & Thesleff I (1997) Expression patterns of bone morphogenetic proteins (Bmps) in the developing mouse tooth suggest roles in morphogenesis and cell differentiation. *Dev Dyn* 210(4):383-396.
28. Fraser GJ*, et al.* (2009) An ancient gene network is co-opted for teeth on old and new jaws. *PLoS Biol* 7(2):e31.
29. O'Connell DJ*, et al.* (2012) A Wnt-bmp feedback circuit controls intertissue signaling dynamics in tooth organogenesis. *Sci Signal* 5(206):ra4.
30. Wise SB & Stock DW (2006) Conservation and divergence of Bmp2a, Bmp2b, and Bmp4 expression patterns within and between dentitions of teleost fishes. *Evol Dev* 8(6):511-523.
31. Wittkopp PJ (2011) Using pyrosequencing to measure allele-specific mRNA abundance and infer the effects of cis- and trans-regulatory differences. *Methods Mol Biol* 772:297-317.
32. Huysseune A & Witten PE (2006) Developmental mechanisms underlying tooth patterning in continuously replacing osteichthyan dentitions. *J Exp Zoolog B Mol Dev Evol* 306(3):204-215.
33. Huysseune A (1995) Phenotypic Plasticity in the Lower Pharyngeal Jaw Dentition of Astatoreochromis-Alluaudi (Teleostei, Cichlidae). *Archives of Oral Biology* 40(11):1005-1014.
34. Orr HA (2005) The genetic theory of adaptation: a brief history. *Nat Rev Genet* 6(2):119-127.

35. Rogers SM*, et al.* (2012) Genetic signature of adaptive peak shift in threespine stickleback. *Evolution* 66(8):2439-2450.
36. Colosimo PF*, et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* 307(5717):1928-1933.
37. Aigler SR, Jandzik D, Hatta K, Uesugi K, & Stock DW (2014) Selection and constraint underlie irreversibility of tooth loss in cypriniform fishes. *Proc Natl Acad Sci U S A* 111(21):7707-7712 .
38. Haara O*, et al.* (2012) Ectodysplasin regulates activator-inhibitor balance in murine tooth development through Fgf20 signaling. *Development* 139(17):3189-3199.
39. Harjunmaa E*, et al.* (2012) On the difficulty of increasing dental complexity. *Nature* 483(7389):324-327.
40. Mustonen T*, et al.* (2004) Ectodysplasin A1 promotes placodal cell fate during early morphogenesis of ectodermal appendages. *Development* 131(20):4907-4919.
41. Gross HP & Anderson JM (1984) Geographic-Variation in the Gillrakers and Diet of European Threespine Sticklebacks, *Gasterosteus-Aculeatus*. *Copeia* (1):87-97.
42. Hagen DW & Gilbertson LG (1972) Geographic Variation and Environmental Selection in *Gasterosteus-Aculeatus L* in Pacific Northwest, America. *Evolution* 26(1):32-51.
43. McPhail JD (1993) Ecology and Evolution of Sympatric Sticklebacks (*Gasterosteus*) - Origin of the Species Pairs. *Can J Zool* 71(3):515-523 .
44. Jones FC*, et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484(7392):55-61.
45. Kavanagh KD, Evans AR, & Jernvall J (2007) Predicting evolutionary patterns of mammalian teeth from development. *Nature* 449(7161):427-432.
46. Reddi AH & Reddi A (2009) Bone morphogenetic proteins (BMPs): from morphogens to metabologens. *Cytokine & growth factor reviews* 20(5-6):341-342.
47. Abzhanov A, Protas M, Grant BR, Grant PR, & Tabin CJ (2004) Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305(5689):1462-1465.
48. Albertson RC, Streelman JT, Kocher TD, & Yelick PC (2005) Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. *Proc Natl Acad Sci U S A* 102(45):16287-16292.
49. Mou C*, et al.* (2011) Cryptic patterning of avian skin confers a developmental facility for loss of neck feathering. *PLoS Biol* 9(3):e1001028.
50. Schoenebeck JJ*, et al.* (2012) Variation of BMP3 contributes to dog breed skull diversity. *PLoS Genet* 8(8):e1002849.
51. Guenther C, Pantalena-Filho L, & Kingsley DM (2008) Shaping skeletal growth by modular regulatory elements in the Bmp5 gene. *PLoS Genet* 4(12):e1000308.

52. Kingsley DM (1994) What do BMPs do in mammals? Clues from the mouse short-ear mutation. *Trends in Genetics* 10(1):16-21.
53. Chan YF*, et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327(5963):302-305.
54. Miller CT*, et al.* (2007) cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* 131(6):1179-1189.
55. Shapiro MD*, et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428(6984):717-723.
56. Broman KW & Sen S (2009) *A Guide to QTL Mapping with R/qtl* (Springer, New York).
57. Thisse C & Thisse B (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature protocols* 3(1):59-6

Chapter 3

Comparative genomics and reverse genetics of evolved tooth gain and *Bmp6* in sticklebacks

3.1 Abstract

Understanding how changes to developmental genetic networks leads to morphological evolution in nature remains an outstanding and largely unsolved biological problem. The threespine stickleback fish offers several genomic and forward and reverse genetic advantages to dissect how morphology evolves in nature. Marine sticklebacks have repeatedly invaded and adapted to numerous freshwater environments throughout the Northern hemisphere. In response to new diets in freshwater habitats, changes in craniofacial pattern have evolved in derived freshwater populations. A freshwater population adapted to live on the bottom of Paxton Lake, Canada has evolved a near two-fold increase in tooth number that arises late in development. This evolved tooth gain is heritable, and largely controlled by a quantitative trait locus (QTL) on chromosome 21 that contains a *cis*-regulatory allele of *Bone Morphogenetic Protein 6* (*Bmp6*). Here we fine-map the chromosome 21 tooth QTL, further supporting *Bmp6* as underlying the QTL. Comparative genomics of marine and freshwater chromosomes with and without the tooth QTL identified a cluster of sequence variants in intron 4 of *Bmp6* which surround a robust tooth and fin enhancer. We induced mutations in *Bmp6* with TALENs, which revealed a required role for stickleback *Bmp6* in tooth patterning. Lastly, transcriptional profiling of *Bmp6* mutants reveals that *Bmp6* regulates the TGF-β pathway in developing tooth plate tissue. Furthermore, there is also significant downregulation of genes involved in mouse hair stem cell quiescence. Collectively these data support a model where mutations around *a Bmp6* intronic tooth enhancer contribute to evolved tooth gain, and suggest ancient shared genetic circuitry might regulate the regeneration of diverse vertebrate epithelial appendages including mammalian hair and fish teeth.

3.2 Introduction

Finding the genes and ultimately the mutations that drive the evolution of animal form remains an important goal in biology (1). The *cis*-regulatory hypothesis proposes that *cis*-regulatory changes are the preferred substrate for morphological evolution because these mutations are more likely to bypass potential negative pleiotropy typically generated by coding mutations (2). Although morphological evolution has been shown to act through both *cis*-regulatory and coding mutations, the predisposition towards either type of mutation as the major evolutionary driver remains unclear due to relatively few examples (3, 4). Furthermore, direct tests of negative pleiotropy, which would prevent certain types of coding mutations, have generally not been explored in non-model organisms.

Morphological evolution can occur through constructive and regressive changes during evolution. The majority of work aiming to understand the molecular basis of evolution has focused on regressive traits (loss of tissue) (5–7). Recent work has begun to identify the genetic basis of constructive traits (gain of tissue) in evolution (8, 9). However, the extent that constructive evolution acts through similar developmental and genetic mechanisms as regressive evolution is still largely unknown. Given that during development, repression is just as common as activation (10, 11)**,** there may be no fundamental distinction between the typical genetic basis of loss and gain traits.

Teeth are a classic model system for studying organ development and evolution in vertebrates (12). During tooth development, epithelial and mesenchymal cells reciprocally signal to each other, integrating the BMP, TGF-β, FGF, SHH, NOTCH, EDA, and WNT signaling pathways to orchestrate the formation of an adult tooth (13). Bone Morphogenetic Protein (BMP) signaling plays an especially intimate role in tooth development. At the early initiation stage, epithelial *Bmp4* inhibits developmental markers of the early forming tooth placode, *Pax9* and *Pitx2* (14, 15). These results suggest an inhibitory role of BMP signaling on tooth development. However, several lines of evidence support an activating role of BMPs on tooth development. For example, exogenous *Bmp4* can rescue tooth development in *Msx1* mutant mice and accelerate tooth development in cultured tooth mandibles, suggesting an activating role of BMP signaling (16, 17). Furthermore, mice homozygous for mutations in the BMP receptor, *Bmpr1a*, or transgenic for a construct overexpressing a BMP antagonist, *Noggin*, in tooth epithelium have tooth arrest at the bud and placode stage, respectively (18, 19). Together, these results suggest that there are both activating and inhibitory roles of BMP signaling during tooth development. However, the roles of specific BMP signaling components are not fully understood. Furthermore, the genetic pathways of early tooth pattern and initiation have been extensively studied and well characterized in mice. Because mice are monophyodont rodents that do not replace their teeth, significantly less is known about the developmental genetic basis of tooth replacement. Polyphyodont fish and reptiles that continuously replace their teeth offer an

opportunity to study the genetic and developmental basis of tooth regeneration (20).

Threespine stickleback fish (*Gasterosteus aculeatus*) are an excellent model for understanding the developmental and genetic basis of natural variation (21). Sticklebacks have undergone a dramatic adaptive radiation in which ancestral marine stickleback have colonized newly-formed freshwater lakes and streams throughout the Northern hemisphere (22). Recent genetic studies have implicated *cis*-regulatory changes of major developmental signaling molecules with stickleback morphological evolution (5–8, 23). Genome-wide searches for regions under selection during freshwater adaptation have found an enrichment in non-coding elements of the genome, further implicating *cis*-regulatory changes in underlying stickleback evolution (24, 25)

Freshwater sticklebacks have evolved several morphological adaptations to their head skeleton, some likely due to the shift to feeding on larger prey in freshwater niches (26). Recent studies have reported a major constructive trait of evolved tooth gain, in which a nearly two-fold gain in ventral pharyngeal tooth number is seen in the freshwater benthic (adapted to lake bottom) morph from Paxton Lake, Canada (8). Pharyngeal teeth lie in the pharynx of fish and are serial and phylogenetic homologs of mammalian oral teeth (27). Pharyngeal jaw patterning is an adaptive trait in fish that covaries with diet and ecological niche (28). Over 300 genes are thought to be involved in mammalian tooth development (29). Homologous genetics networks are involved in fish tooth development (30, 31), making stickleback evolved tooth gain a powerful opportunity to understand the evolutionary genetics of tooth development and replacement.

Evolved tooth gain in Paxton benthic freshwater fish is accompanied by an increase in the size of the tooth field, a decrease in tooth spacing, and an increase in tooth replacement rate late in development (8) (Ellis et al., submitted). This derived tooth pattern is partially explained by a large effect quantitative trait locus (QTL) on chromosome 21 that is associated with a late-acting *cis*-regulatory downregulation of *Bmp6* mRNA in the tooth plates of benthic fish (8). These results make *Bmp6* an excellent candidate gene for underlying evolved tooth gain by regulating tooth patterning and replacement. Here, we use a combination of recombinant mapping, comparative genomics, genome editing, and transcription profiling to further dissect the molecular genetic basis of evolved tooth gain and the role of *Bmp6* during tooth development in stickleback fish.

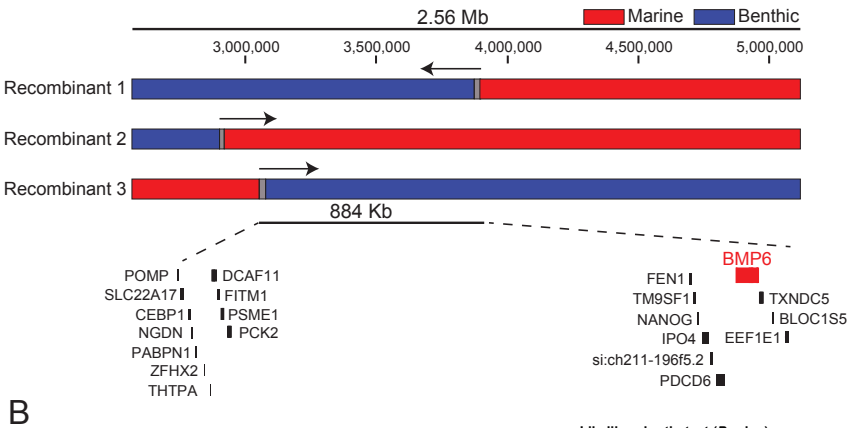
3.3 Results

3.3.1 Recombinant mapping of chromosome 21 tooth number QTL supports *Bmp6*

Recent work fine-mapped a large effect tooth number QTL to a 2.56 Mb 1.5 LOD interval on stickleback chromosome 21 containing an excellent candidate gene, *Bone Morphogenetic Protien 6* (*Bmp6*), along with 58 other predicted genes (8). To further fine-map this QTL, we identified three recombinant chromosomes with

marine-benthic recombination events within the 2.56 Mb fine-mapped interval (Figure 3.1A).

A



B

Recombinant Cross	Number	Genotypic Classes				Likelihood-ratio test (*P* value) Acts as marine	Acts as benthic	Direction
Recombinant 1 (MR x MB)		**MM**	**MB**	**MR**	**BR**			
Tooth Number	109	-5.12 (0.84)	1.12 (1.33)	-0.83 (1.33)	5.67 (1.35)	0.24	**0.0003**	**Left**
Recombinant 2 (BR x MB)		**MR**	**MB**	**BR**	**BB**			
Tooth Number	147	-2.77 (0.79)	0.93 (0.75)	0.27 (0.94)	2.07 (1.03)	**0.001**	0.55	**Right**
Recombinant 3 (MR x MB)		**MM**	**MB**	**MR**	**BR**			
Tooth Number	102	-4.52 (1.32)	-0.3 (1.25)	1.81 (1.32)	2.67 (1.05)	0.25	**0.0003**	**Right**

Figure 3.1. Recombinant mapping of chromosome 21 tooth QTL supports a narrowed region containing *Bmp6*

The chromosome 21 tooth QTL was previously fine mapped to a 2.56 Mb region containing *Bmp6* along with 58 other ENSEMBL predicted genes (8). The three recombinant chromosome 21s that were tested are shown (A). Genotypes are colored red for marine, blue for benthic, and grey for unresolved. Arrows denote position of tooth QTL supported by each recombinant chromosome. The final recombinant mapped interval is 884 Kb containing 21 genes, including *Bmp6.* Gene content was determined by hand annotating the Ensembl predicted gene list (A). Size-corrected total ventral pharyngeal tooth number and standard error are shown for the genotypic classes within each of the recombinant crosses (B). For each cross, parental genotypes of the tooth QTL are listed and coded: marine (M), benthic (B), or recombinant (R). *P* values from likelihood-ratio tests determining whether recombinant chromosomes behave like marine or benthic chromosomes are shown.

Fish with each of these recombinant chromosomes were crossed to marine/benthic heterozygotes to generate large (>100 fish each) crosses to test these recombinant chromosomes for effects on tooth number (Table 3.1). Recombinant chromosomes that increase tooth number compared to the marine chromosome would suggest that the tooth controlling region of chromosome 21 lies within the benthic portion of the recombinant chromosome. We used a

Table 3.1. Summary of recombinant crosses with genotyping markers

Recombinant Cross	Sample size	Correction	Recombinant left marker	Recombinant right marker	Marker type
1	109	Sex	CACTGAAGCCGGAGGAGAGG, ATCAGAGAGGGTCCAGAACG	AGTCCGCCACTTGTCTTTCC, GTCATGCAGACCATGATTCC	Size polymorphism
2	147	Fish length and Sex	TGAACCAATTGTTTGGAACATC, AATCGCCATGTCAAATTCCT	CCCGCAAGAAAGCAATTTAT, TTTGTTTCCTGCCTTCGAGT	Size polymorphism
3	102	Fish length And Sex	ATCCAGCCCAGAGTGAAATG, GGCCTACCAACTTGACCGTA	TGTGTGCAAACACACAGCAT, TCTGCTCTGCTTTGCTTCTTC	RFLP (AvuII left and SalI right)

Sample sizes of the Paxton benthic x Little Campbell marine recombinant crosses are shown along with the sequences for left and right genotyping markers used as the boundaries for the recombination breakpoint. The markers for recombinant 1 and 2 are size polymorphisms and the markers for recombinant 3 are restriction fragment length polymorphisms (RFLPs) using the restriction nuclease shown. Standard fish length and sex were corrected for when appropriate and corrections performed for each cross are listed.

likelihood ratio test to determine whether each recombinant chromosome behaves more like a marine or benthic allele. Recombinant chromosomes one and three increased tooth number, behaving like a benthic chromosome (*P* value from likelihood ratio test = 0.0003 and 0.0003, respectively) (Figure 3.1B). Recombinant chromosome two did not increase tooth number behaving like a marine chromosome (*P* = 0.001 from likelihood ratio test) (Figure 3.1B). Together, these recombinant crosses support a new smaller genetic interval, 884 kb in the stickleback reference genome assembly, that contains 21 predicted genes including *Bmp6* (Figure 3.1A), reducing the physical size and number of genes by 65% and 64%, respectively.

3.3.2 Eight out of nine derived benthic chromosomes have a large effect tooth QTL

To estimate the frequency of the chromosome 21 tooth QTL within the wild Paxton benthic population, we generated six marine by benthic F2 crosses testing nine wild-derived benthic chromosomes with different genotypes at three microsatellite loci across the tooth QTL (Figure 3.2, Table 3.2, see Material and Methods). We found that eight of these nine benthic chromosome 21s had significant effects on tooth number with the same direction and similar magnitude of effect (Figure 3.2, Table 3.2). The benthic chromosome tested in cross 6 had no effects on tooth number (Figure 3.2, Table 3.2). These results together suggest that the tooth number QTL on chromosome 21 is at high frequency in the Paxton benthic population and likely plays a major role in their evolved tooth gain phenotype (8), but there are also lower-frequency benthic chromosomes that lack a tooth QTL.

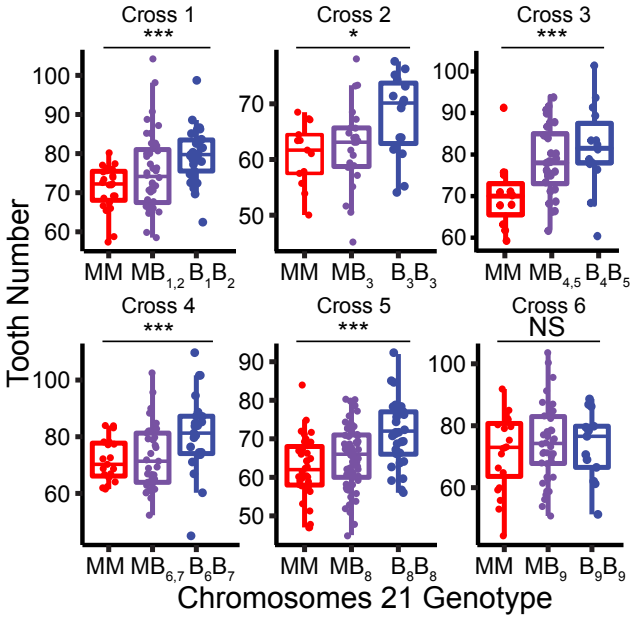


Figure 3.2. Eight out of nine benthic chromosome 21's have a tooth QTL

Results from six benthic by marine F2 crosses testing nine benthic chromosome 21s (B1-9) are shown. Benthic chromosomes 1-8 have strong effects on tooth number; however the benthic chromosome 9 has no detectable effects on tooth number. Total tooth numbers from marine homozygous (red), heterozygous (purple), and benthic homozygous (blue) animals for chromosome 21 are presented. *P* values from an ANOVA for cross 1-6 are 0.002, 0.024, 0.003, 0.004, 2.11×10^{-5}, 0.69, respectively. F2 crosses 1, 3, and 4 are testing two benthic chromosomes each and crosses 2, 5, and 6 each are testing one.

3.3.3 Whole genome resequencing reveals a cluster of QTL-associated SNPs in intron 4 of *Bmp6*

We parsimoniously hypothesized that the Paxton benthic chromosomes with the chromosome 21 tooth QTL may have shared sequence polymorphisms, which underlie evolved tooth gain, that are not present on marine chromosomes or the benthic chromosome without the QTL. To identify QTL-associated SNPs, we resequenced the genomes of the nine benthic chromosomes and the three marine grandparents from crosses 2, 5, and 6 tested in Figure 3.2. We identified 373 homozygous sequence variants within the 884 kb fine mapped genetic interval between the benthic with the QTL and marine fish (Figure 3.3A). We gave variants a QTL concordance score, the number of times a variant is found in the benthic animals with a chromosome 21 tooth QTL minus the number of times the same variant was found in animals without a tooth QTL, normalized to 1. Ten of these variants were perfectly associated with the presence of the tooth QTL (Figure 3.3). Strikingly, all of these variants lie within a ~4.4 kb region of *Bmp6* intron 4 that is partially conserved in other fish and mice (Figure 3.3B).

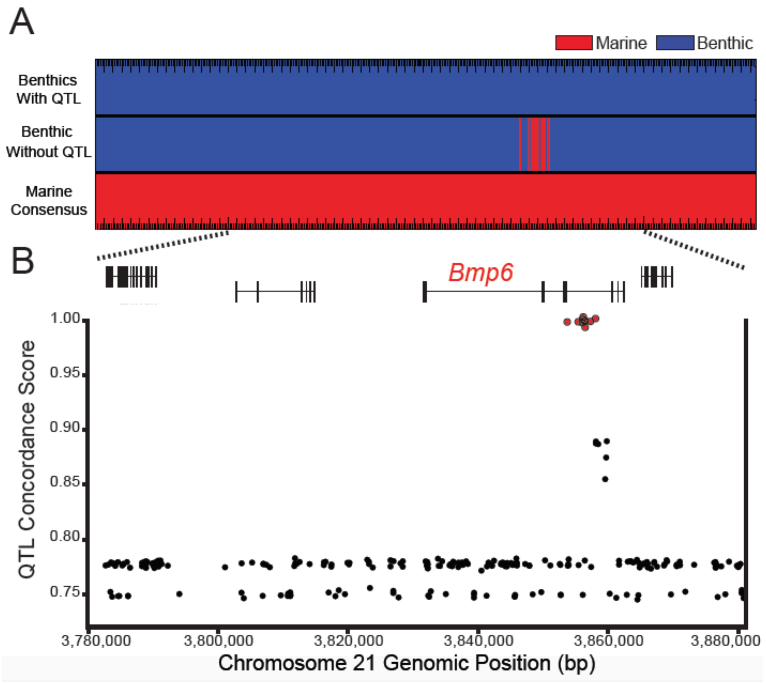


Figure 3.3. Comparative genomics reveal QTL-associated SNPs in intron 4 of *Bmp6*

Comparing genomic sequences of the fine-mapped tooth QTL between marine and benthic chromosomes with the tooth QTL identified a set of SNPs with opposite homozygous genotypes, colored red for marine (bottom) and blue for benthic (top). The benthic chromosome without the QTL had a cluster of SNPs sharing the consensus marine genotype (middle) (A). The SNPs (red points) with perfect QTL association lie within intron 4 of *Bmp6* (B). The y-axis shows a QTL concordance score (see Methods), a metric of concordance between genotype and presence or absence of tooth QTL.

The association of this intron 4 haplotype with the presence of a chromosome 21 tooth QTL within Paxton lake suggests that these SNPs might predict the QTL in independently evolved freshwater populations. To test this hypothesis, we tested two marine by freshwater F2 crosses using the Fishtrap Creek freshwater population in Washington, USA, where the freshwater grandparents were mostly homozygous for the two different alleles of intron 4 (Figure 3.4A). Supporting this hypothesis, we identified a large effect tooth QTL in the F2 cross with the tooth QTL-associated SNP pattern and no significant QTL in the F2 cross with the alternative allele (Figure 3.4B). These results show a chromosome 21 tooth QTL in an independently derived freshwater population and support the model that the SNPs in intron 4 of *Bmp6* contribute to the tooth QTL in this additional population.

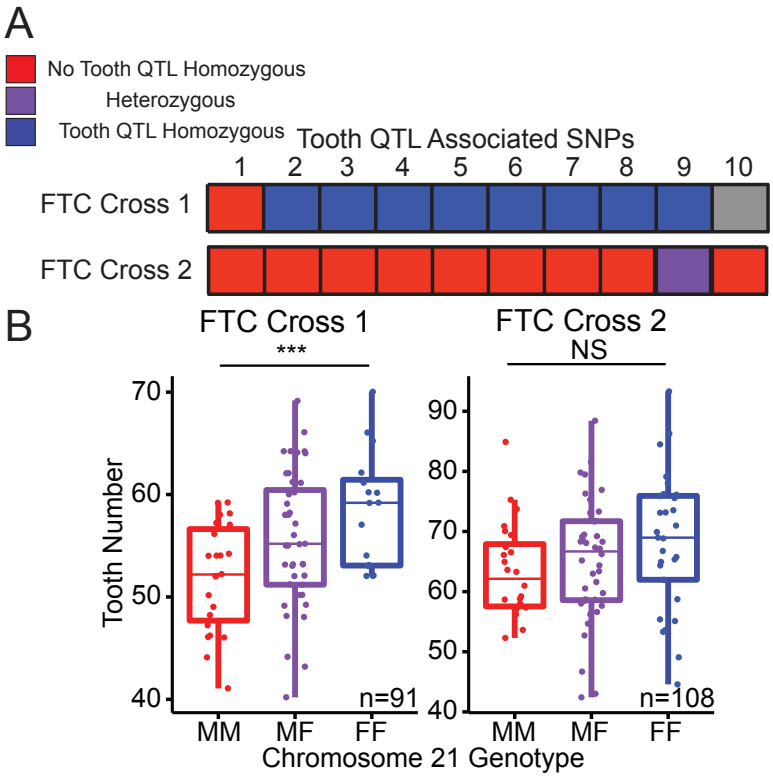


Figure 3.4. Intron 4 SNP pattern predicts a chromosome 21 tooth QTL in an independently derived freshwater population

Table 3.2. Wild Benthic F2 cross summary

	Populations	Number	Marker	MM	MB	BB	Effect	Correction	Unique Benthic Chromosomes
Cross 1	Paxton Benthic x Japanese Marine	92	CM1440	71 (1.2)	75 (1.6)	80 (1.3)	9	Fish Length	2
Cross 2	Paxton Benthic x Japanese Marine	51	CM1440	61 (1.6)	63 (1.6)	68 (1.9)	7	Fish Length	1
Cross 3	Paxton Benthic x Rabbit Slough Marine	54	Stn489	70 (2)	78 (1.6)	80 (2.3)	10	Fish Length	2
Cross 4	Paxton Benthic x Japanese Marine	77	Stn489	72 (1.7)	73 (2)	81 (2.9)	9	Fish Length	2
Cross 5	Paxton Benthic x Little Campbell Marine	138	Stn487	62 (1.4)	65 (0.9)	72 (1.5)	10	None	1
Cross 6	Paxton Benthic x Little Campbell Marine	75	Stn487	72 (3.1)	74 (1.9)	74 (2.7)	NA	Fish Length	1

Results from Marine by Benthic F2 crosses testing nine wild-derived benthic chromosomes. Populations, number of F2 animals, and chromosome 21 marker used for each cross are presented. Standard fish length effects on total tooth number was corrected for when appropriate and residuals were back-transformed to the mean standard fish length within each cross. Mean and standard error of corrected tooth number are shown for marine homozygotes (MM), heterozygotes (MB), and benthic homozygotes (BB). The effect of chromosome 21 in each cross is shown as the differences in means between benthic and marine homozygotes. Crosses 1, 3, and 4 tested two distinct benthic chromosomes and crosses 2, 5, and 6 tested a single benthic chromosome.

Shown are genotypes for the ten tooth QTL associated SNPs in the freshwater grandparents from two Fishtrap Creek (FTC) by Little Campbell marine F2 crosses (A). The FTC cross 1 grandparent is homozygous for the second through ninth tooth QTL associated variants (Blue) and the FTC cross 2 grandparent is mostly homozygous for the variant pattern seen in fish without the QTL (Red). Heterozygous genotypes are shown in purple and positions with missing data as grey. Chromosome 21 genotype has strong effects on tooth number in FTC Cross 1, where the freshwater homozygotes (blue) have more teeth than the marine homozygotes (Red) (*P* value from ANOVA is 0.004) (B). However, FTC cross 2 had no significant effects on tooth number (*P* value from ANOVA is 0.13).

3.3.4 A conserved tooth and fin enhancer in intron 4 of *Bmp6*

A *cis*-regulatory allele of *Bmp6* is associated with the chromosome 21 tooth QTL in Paxton benthic fish (8). We hypothesized that the intron 4 region containing tooth QTL specific SNPs is a tooth enhancer of *Bmp6* (Figure 3.4B). To test for enhancer function, we cloned a 2022 bp intron 4 genomic fragment into a reporter construct. Transgenic fish for this construct expressed GFP in the distal tips of developing pectoral and caudal fins at 8 days post fertilization (dpf) (Figure 3.5A) and pharyngeal (Figure 3.5B,D) and oral teeth (Figure 3.5C) at 10 dpf. These domains have been previously shown to be endogenous sites of *Bmp6* expression in developing sticklebacks (8, 32). These results demonstrate that intron 4 contains an enhancer active in developing teeth and fins.

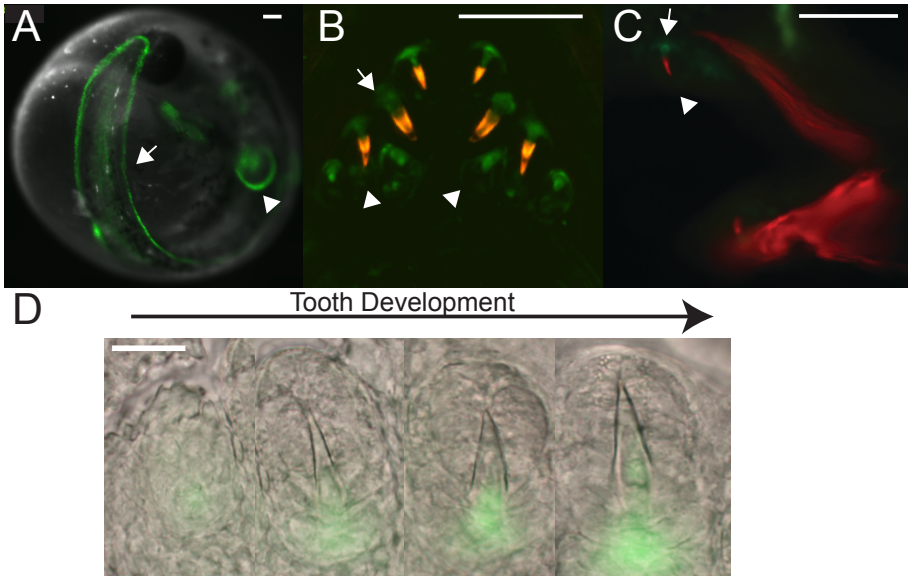


Figure 3.5. 2022 bp intron 4 region is an enhancer active in developing fins and teeth

The marine 2022 bp intronic element drives expression at 8 dpf in the distal tips of the developing median fin (arrow) and pectoral fin (arrowhead) (A). By 10 dpf, the enhancer drives GFP expression in tooth mesenchyme (arrow) and diffusely in the tooth epithelium (arrowhead) in both pharyngeal (B) and oral (C) jaws. This

2022 bp enhancer controls dynamic expression throughout development, becoming more restricted to the mesenchyme as the tooth matures (D). Scale bars are 100 μm (A-C) and 50 μm (D).

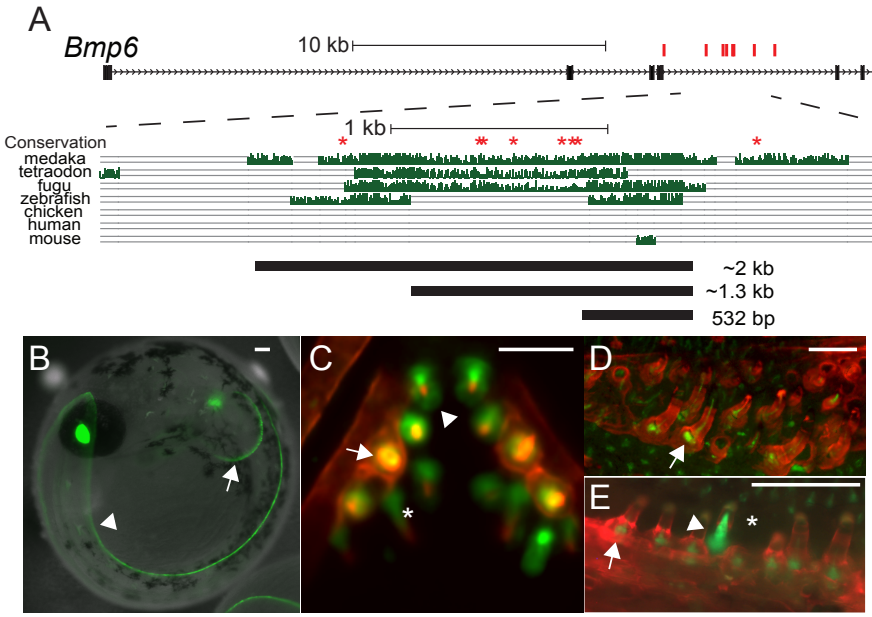


Figure 3.6. Intron 4 region with QTL-associated SNPs, contains a tooth and fin enhancer

The coding region of *Bmp6*'s seven exons (thick bars) and six intronic regions (arrowed line) are shown. All ten QTL-associated SNPs (red ticks) are located within intron 4 of *Bmp6*. Eight of these SNPs (red asterisks) are in conserved portions of the fourth intron of *Bmp6.* The conservation in vertebrates is shown from the UCSC genome browser. Shown with black bars are the ~2 kb, 1.3 kb, and the 528 bp enhancer subclones used in this study (A). GFP reporter expression from the marine 528 bp enhancer in a stable transgenic 8 dpf fish (B). Expression was detected in the developing distal edge of the pectoral (arrow) and median fins (arrowhead). By 13 dpf, there is reporter expression in developing tooth epithelia (arrowhead) and strong expression in mesenchyme of early stage (asterisk) and fully-formed teeth (arrow) (B). By late juvenile stages, the tooth expression persists in developing pharyngeal (D) and oral (E) teeth. GFP was detected in the core of fully-formed teeth (arrows) (D-E). The asterisk marks a tooth position being replaced and the arrowhead marks a newly forming tooth with GFP-positive mesenchyme extending to the tooth plate (E). Panels C and D are confocal micrographs. Scale bars are 100 μm.

We hypothesized that the tooth and fin expression driven by this ~2 kb enhancer might be due to separable tissue specific enhancers. We subcloned the ~2 kb fragment into two smaller fragments of ~1.3 kb and the 532 bp, based on patterns of sequence conservation in vertebrates (Figure 3.6A). The 532 bp construct, which is highly conserved and contains no QTL-specific SNPs, drove expression in the distal tips of the caudal and pectoral fins at 8 dpf (Figure 3.6B).

By 13 dpf, this minimal enhancer drove expression in mesenchymal cell in developing pharyngeal teeth (Figure 3.6B), as well as distal expression in tooth epithelium that is generally weaker than the mesenchymal expression (Figure 3.6C). In developing teeth, the GFP-positive mesenchyme population extends from each tooth germ deep into the tooth plate (Figure 3.6C, E). This tooth expression continued into late juvenile stages when the pharyngeal tooth number differences arise between marine and freshwater populations (Figure 3.6D) (8). GFP expression was also detected in late juvenile oral teeth (Figure 3.6E). These results demonstrate that the intron 4 tooth QTL-associated variants surround an enhancer sufficient to drive expression in developing fins and teeth, and that these two expression domains were not separable in the subclones tested.

3.3.5 Intron 4 enhancer *Bmp6* controls modular and distinct expression from 5' enhancer

We previously identified a TGFβ-responsive 5' enhancer of *Bmp6* that also drives expression in developing teeth and fins in sticklebacks (29). This enhancer drives epithelial expression in developing teeth and mesenchymal cells in cores of maturing teeth (32). Because stickleback *Bmp6* tooth expression is spatially and temporally complex (8), we hypothesized that the two regulatory elements may control distinct aspects of *Bmp6* expression in teeth. To test this hypothesis, we compared reporter gene expression in stable transgenic lines of the190 bp 5' tooth enhancer to the 532 bp intron 4 tooth enhancer (Figure 3.7A-C). As previously described, we found that the 5' enhancer drives robust expression in developing tooth epithelium and adjacent tooth mesenchyme (Figure 3.7A, D). We found that the intronic enhancer lines had expression that at some stages of tooth development was distinct from the 5' enhancer (Figure 3.7C, E). The intronic enhancer drives expression in the mesenchymal cores of mature teeth similar to what is seen for the 5' element. However, the intronic enhancer's mesenchymal expression was expanded broadly around the base of the developing tooth compared to the 5' enhancer (Figure 3.7A-E).

To directly compare the fin and tooth expression domains driven by the two enhancers, we performed transient injections of ~1.3 kb intron 4 enhancer mCherry reporter construct into stable GFP reporter lines of the 190 bp 5' tooth enhancer (Figure 3.7F). The tooth expression domains were partially complementary between the two enhancers in developing teeth (Figure 3.7G-H). As was seen comparing the stable lines, both enhancers drive similar mesenchymal expression at early stages of tooth development, but at later stages of tooth development the 5' enhancer appears to drive more distal mesenchymal expression than the intronic enhancer (Figure 3.7G-H). As tooth development progresses, mesenchymal expression of the intronic enhancer expands to an apparently larger mesenchymal cell population past the edge of the mineralized tooth extending to the tooth plate (Figure 3.7E-F). These results demonstrate that *Bmp6* epithelial and mesenchymal tooth expression is driven by modular enhancers, which drive overlapping and distinct expression pattern.

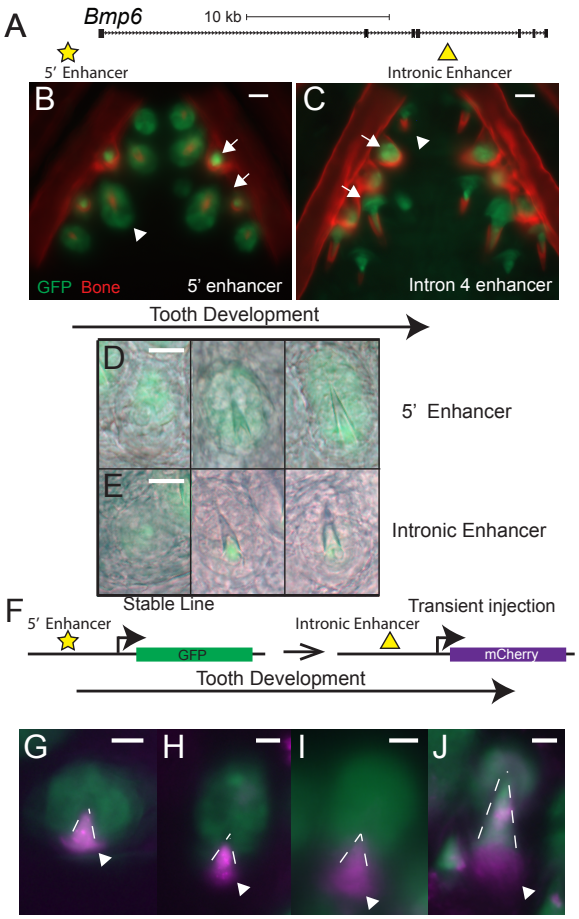


Figure 3.7: Bmp6 intron 4 enhancer controls modular expression from 5' enhancer

The 5' (star) and intron 4 (triangle) enhancers of *Bmp6* that drive expression in developing fins and teeth are shown (A). Comparing transgenic lines for the 190 bp 5' (B, D) and 528 bp intron 4 (C, E) enhancer in developing teeth shows overlapping, but distinct expression domains. Arrowheads mark developing tooth epithelia and arrows mark developing tooth mesenchyme (B-C). We injected the ~1.3 kb intron 4 enhancer driving mCherry (purple) into a stable line for the 190 bp 5' enhancer driving GFP (green) (C). Similar to the transgenic lines, at early stages of tooth development the intronic enhancer is expressed in the mesenchyme core of the developing tooth (G). As development progresses, the intron 4 positive mesenchymal cell population (arrowheads) expands out to form a lawn at the base of the developing tooth (G-I). The 5' enhancer drives expression in the core mesenchyme of a developing tooth (G-I). The epithelial expression of intron 4 arises at the distal tip of the developing tooth late in development (I). Dashed lines mark the mineralizing tooth. Scale bars are 20 μm (B-E) and 10 μm (G-J).

3.3.6 Generation of a loss-of function allele of *Bmp6* in stickleback fish

To test whether *Bmp6* is required for tooth patterning in sticklebacks, we used transcription activator-like effector nucleases (TALENs) to generate two predicted loss-of-function mutations in stickleback *Bmp6* (Figure 3.8A, Table 3.4). We designed a TALEN pair to target the highly conserved second exon of *Bmp6*, which is 5' to the exons encoding the predicted secreted ligand. Thus early stop codons would be predicted to generate strong loss-of-function alleles. Injection of these TALEN RNAs into stickleback embryos efficiently induced mutations in the *Bmp6* target sequence. Of injected F_o stickleback embryos, 24-57% had detectable deletions using a restriction enzyme assay, with up to 12% of these animals appearing to have biallelic mutations (Figure 3.9A). To identify TALEN mutations from F_o injected animals and mutations transmitted through the germline in F_1 animals, we PCR amplified the surrounding sequence around the target site, digested this amplicon with EcoRI, then gel extracted and sequenced the uncut band. Consistent with previous studies using TALENs in fish, we identified a spectrum of insertions and deletions at the target site (Figure 3.8B) (33). We generated two stable lines: one harboring a 13 bp deletion, and a second with a 3 bp deletion and 4 bp insertion (Figure 3.8B). Both of these mutations are predicted to produce frameshifts 5' to the secreted BMP ligand and thus are both likely strong loss-of-function alleles (Figure 3.10). Furthermore, we found that the 13 bp deletion allele had reduced expression by RNA-seq as would be predicted by non-sense mediated mRNA degradation (Figure 3.11) (34).

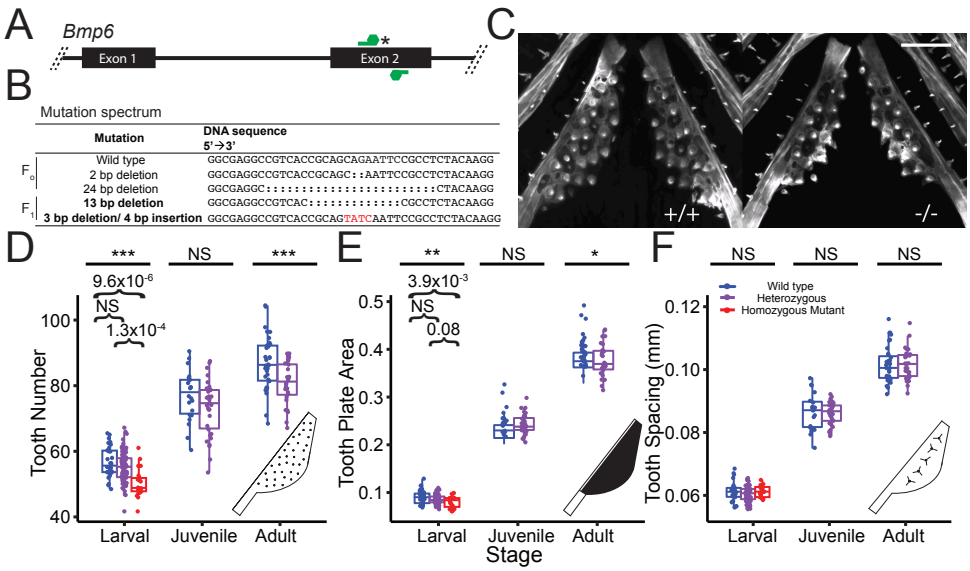


Figure 3.8: *Bmp6* is required for proper ventral pharyngeal tooth pattern in stickleback fish

The TALEN pair targeting the second exon of *Bmp6* designed around an EcoRI site (asterisk) that would be destroyed by an induced mutation (A). Sanger sequencing of F_o or F_1 fish revealed a spectrum of genomic insertions and deletions in *Bmp6.* The two mutation lines used in this study are in bold (B). Confocal images of comparable wild-type and homozygous mutant tooth plates show altered tooth pattern in mutant fish (C). Developmental time course of wild-type (blue), heterozygous (purple), and homozygous mutant (red) fish of different

total body lengths (x-axis) and tooth number (D), tooth plate area (E), and tooth spacing (F) (y-axis). Homozygous fish have recessive reduction of tooth number and tooth plate area at the early larval stage (Tukey post-hoc P values comparing wild-type to homozygous mutant are 8.8 $x10^{-6}$ and 0.005, respectively) (D-E). Tooth number and area diverges late in development between wild-type and heterozygous fish (D-E). Tooth spacing is not changed in the mutant at any stage (F). The late stage juvenile and adult crosses were heterozygous mutant backcrossed to wild-type fish. Scale bar is 200 μm (C).

A

Mutation Efficency	F_o *Bmp6* genotype		
Sample Size	+/+	+/-	-/-
N=17	59% (N = 10)	29% (N = 5)	12% (N = 2)
N=47	43% (N = 20)	57% (N = 27)	0% (N = 0)
N=72	76% (N = 55)	18% (N = 13)	6% (N = 4)

B

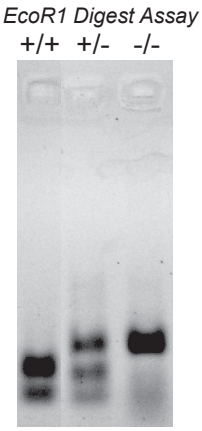


Figure 3.9: Efficacy of *Bmp6* TALENs in stickleback embryos
Frequencies of wild-type (+/+), heterozygous (+/-), and homozygous (-/-) mutant F_0-injected 3 days post fertilization (dpf) embryos are shown for three independent injection rounds (A). An EcoRI site is destroyed by induced mutations. Representative EcoRI digest assays on genomic DNA from a wild-type (left, +/+), heterozygous (middle, +/-), and homozygous (right. -/-) injected embryo are shown (B).

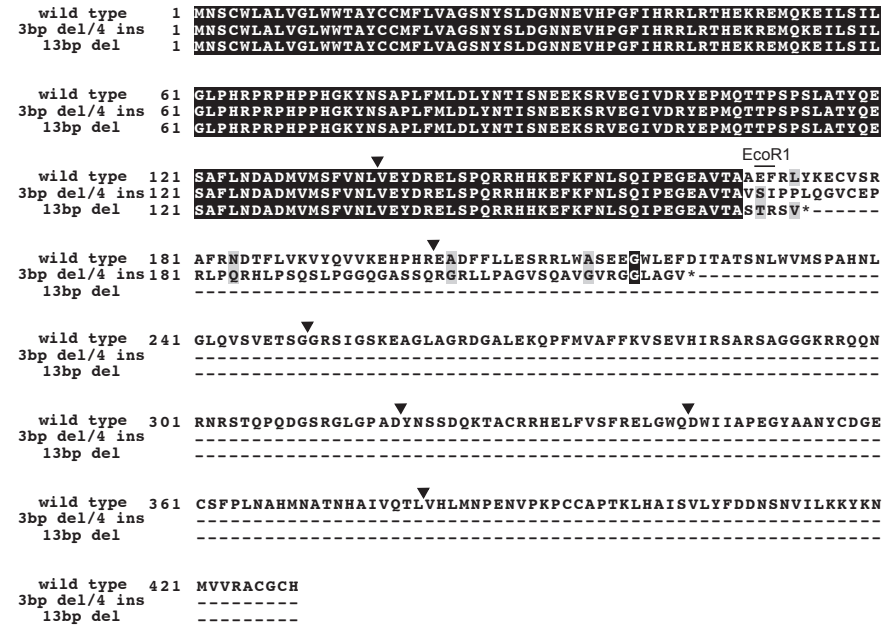


Figure 3.10: Predicted amino acid alignments of the wild-type, 13bp deletion, and the 3 bp deletion/4 bp insertion alleles of BMP6
Predicted mutant BMP6 sequences, 3bp deletion/4bp insertion (middle) and 13bp deletion (bottom), aligned to wild-type (top) BMP6 sequence. The 13bp deletion and the 3bp deletion/4bp insertion generate frameshifts that result in premature stop codons (marked by asterisk) in the 2nd and 3rd exons, respectively, predicted to truncate the protein. Wild-type BMP6 sequences and intron/exon boundaries (marked with arrowheads) as previously described (Cleves et al., 2014). The position of EcoRI site used as the genotyping assay is noted.

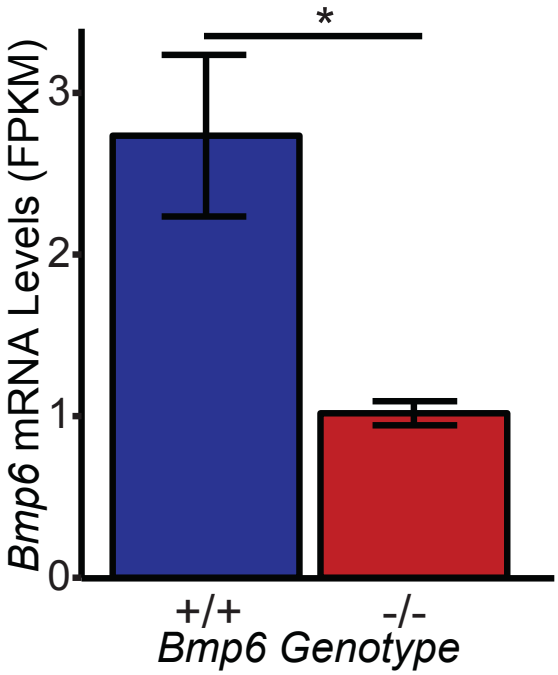


Figure 3.11: *Bmp6* mRNA level is reduced in the mutant
Normalized *Bmp6* mRNA expression levels in tooth plates from wild-type (+/+) and homozygous (-/-) stickleback fish measured using RNAseq. Levels are presented as fragments per kilobase of exon per million fragment (FPKM). There is a significant reduction of *Bmp6* expression of the homozygous mutant compared to wild-type animals (ANOVA *P* value = 0.027).

Table 3.3 Reporter construct cloning primers

Construct	Forward primer	Reverse primer	Enzyme	Orientation
~2.2 kb enhancer GFP	GCCGGCTAGCACCGACACAG CTGTACTTGG	GCCGGCTAGCAGAGTCCTG ATGGCCTCTCC	NheI	Minus
~1.3 kb enhancer GFP	GCCGGCTAGCGAGAGCATCC GTCTTGTGGG	GCCGGCTAGCAGAGTCCTG ATGGCCTCTCC	NheI	Minus
~532 bp enhancer GFP	GCCGGCTAGCGTGTGTGCGC GGTGGAAAATG	GCCGGCTAGCAGAGTCCTG ATGGCCTCTCC	NheI	Minus
~1.3 kb enhancer mCherry	GCCGGCTAGCGAGAGCATCC GTCTTGTGGG	GCCGGGATCCAGAGTCCTG ATGGCCTCTCC	NheI and BamHI	Plus

Sequences of forward and reverse primers used to clone each construct are listed 5' to 3' along with the restriction enzyme used to digest the PCR amplicon. The orientation of the inserts in the GFP constructs were in the minus direction relative to the promoter to mirror endogenous position relative to the *Bmp6* promoter. The mCherry construct was cloned in the plus orientation to mirror the orientation for the 5' tooth enhancer transgenic line used in the co-labeling experiment (see Fig 5)

Table 3.4. Custom TALEN design and targets

Target	TAL1 (left), TAL2 (right) RVDs	TAL1,TAL2 Lengths	Spacer	Target Stickleback Sequence
Bmp6 exon 2	TAL1: HD HD HD NI NN NI NN NN NN HD NN NI NN NN HD HD NN NG HD TAL2: HD NI HD NI HD NI HD NG HD HD HG HG NN NG NI NN NI NN NN	19bp	17 bp	5'-CCCAGAGGGCGAGGCCGTC accgcagca**gaattc**cg CCTCTACAAGGAGTGTGTG-3'

RVDs used to generate left (TAL1) and right (TAL2) nuclease pairs targeting the second exon of *Bmp6*, and stickleback target sequence is listed. Underlined nucleotides correspond to the 19bp TAL1 and TAL2 targets flanked by the 17bp spacer containing an EcoRI restriction site (bold).

!

3.3.7 *Bmp6* specifies early and late stage ventral pharyngeal tooth number and tooth plate area, but not inter-tooth spacing

To test for tooth and skeletal patterning phenotypes in *Bmp6* mutants, we intercrossed fish that were heterozygous for the 13 bp deletion, and raised dense developmental time courses of fish with all three possible genotypic classes. Homozygous mutants were underrepresented from expected ratios at later developmental stages, suggesting late larval lethality (Table 3.6). The surviving homozygous mutants tended to be slightly smaller (Table 3.6). Because of the late stage lethality, we continued the *Bmp6* mutant time course with heterozygous backcrosses for juvenile and adult stages. To test for required roles of *Bmp6* in tooth patterning, we quantified ventral pharyngeal tooth number, tooth plate area, and inter-tooth spacing, three phenotypes all controlled by the chromosome 21 tooth QTL (8) in the *Bmp6* mutant time course (Figure 3.8). At the early larval stage, there was a recessive reduction of tooth number in the homozygous mutant when compared to wild-type and heterozygous fish for both tooth number and tooth plate area, but not inter-tooth spacing (Figure 3.8D-F). At this stage, tooth number and area were not significantly different between wild-type and heterozygous fish. However, beginning in juveniles, when tooth replacement likely begins, heterozygous fish had fewer ventral teeth and smaller tooth plate area, which was significantly reduced at later time points including adults (Figure 3.8D-E). There were no significant differences in inter-tooth spacing at any stage (Figure 3.8F). These results show that *Bmp6* is required for specifying tooth number and the size of the tooth field.

Stickleback pharyngeal teeth are composed of two bilateral dorsal tooth plates, dorsal tooth plate 1 (DTP1) and dorsal tooth plate 2 (DTP2), and one bilateral ventral tooth plate. We asked if *Bmp6* was also required for DTP1-2 tooth number as it was for ventral tooth number (Figure 3.12). We found that that there were no significant differences in tooth number of either dorsal tooth plate at early developmental stages (Figure 3.12C). In adults, DTP2 tooth number was lowered in the heterozygous mutants, but to a lesser degree than the ventral tooth number differences at the same stage (Figure 3.12C and 3.8D). For both dorsal tooth plates, tooth numbers trended in the same direction as seen for the ventral tooth plates, with fewer teeth in mutants than wild types. These results demonstrate that like the chromosome 21 tooth QTL (35), *Bmp6* has stronger effects on ventral pharyngeal tooth number than dorsal pharyngeal tooth number.

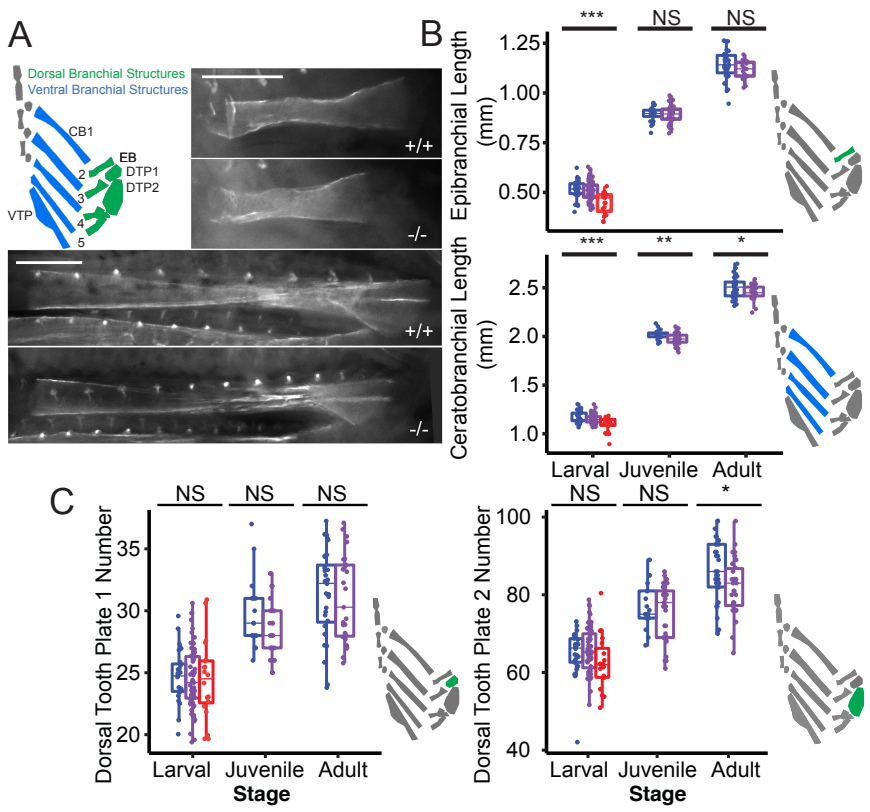


Figure 3.12. Ventral branchial bones and teeth affected more by *Bmp6* mutation

The dorsal ebibranical (EB) and the five serially homologous ventral ceratobranchial (CB) bones make up the major supports of the fish pharyngeal skeleton. Representative fluorescent images of wild-type and homozygous mutant EBs (top right) and CBs (bottom) are shown (A). Both the EB length (top) and the average length of the five CB bones (CB1-5) (bottom) are significantly shorter in the homozygous mutant compared to wild-type fish at the early larval stage (Tukey post-hoc *P* values comparing wild-type to homozygous mutant are 5 x10^{-5} and 2 x10^{-5}, respectively) (A-B). However, in juveniles and adults, the CBs, but not the EBs, are significantly shorter in the heterozygotes compared to the wild-type fish (larval, juvenile, and adult stage ANOVA *P* values are 0.4, 0.007, 0.04 and 0.91, 0.72, 0.19 for average CB1-5 length and EB length, respectively) (B). Similarly, pharyngeal tooth number on dorsal tooth plate 1 (DTP1) (left) are not significantly different between homozygous mutant, heterozygote, and wild-type fish at any stage and the dorsal tooth plate 2 (DTP2) tooth numbers are only significant at the adult stage (ANOVA *P* = 0.028) (right) in contrast to the ventral pharyngeal teeth (VTP) results (A, C). Scale bars are 0.2 mm.

3.3.8 *Bmp6* is required for viability and for patterning the axial, fin ray, and craniofacial skeleton

In mice, *Bmp6* is required for axial skeletal patterning (36), and in zebrafish *Bmp6* is expressed during development of the dermal median fin rays (37). To test whether stickleback *Bmp6* is also required for patterning other regions of the skeleton including the axial and dermal skeleton, we next examined the external skeletons of wild types and homozygous mutants (Figure 3.13). Compared to wild types, homozygous mutants displayed a recessive complete loss of epipleural ribs at early stages (Fig 3.13B-C). In addition, the first dorsal spine, and its underlying skeletal support, the second pterygiophore, were also missing in homozygous mutants compared to wild types (Fig 3.13D-E). This phenotype was less penetrant than the completely penetrant loss of epipleural ribs. In the developmental time courses, wild-type fish with fully-formed ribs and spines were of comparable fish length to homozygous mutants without these bones, supporting the model that these skeletal phenotypes are not due to developmental delay, but to specific roles of *Bmp6* during development.

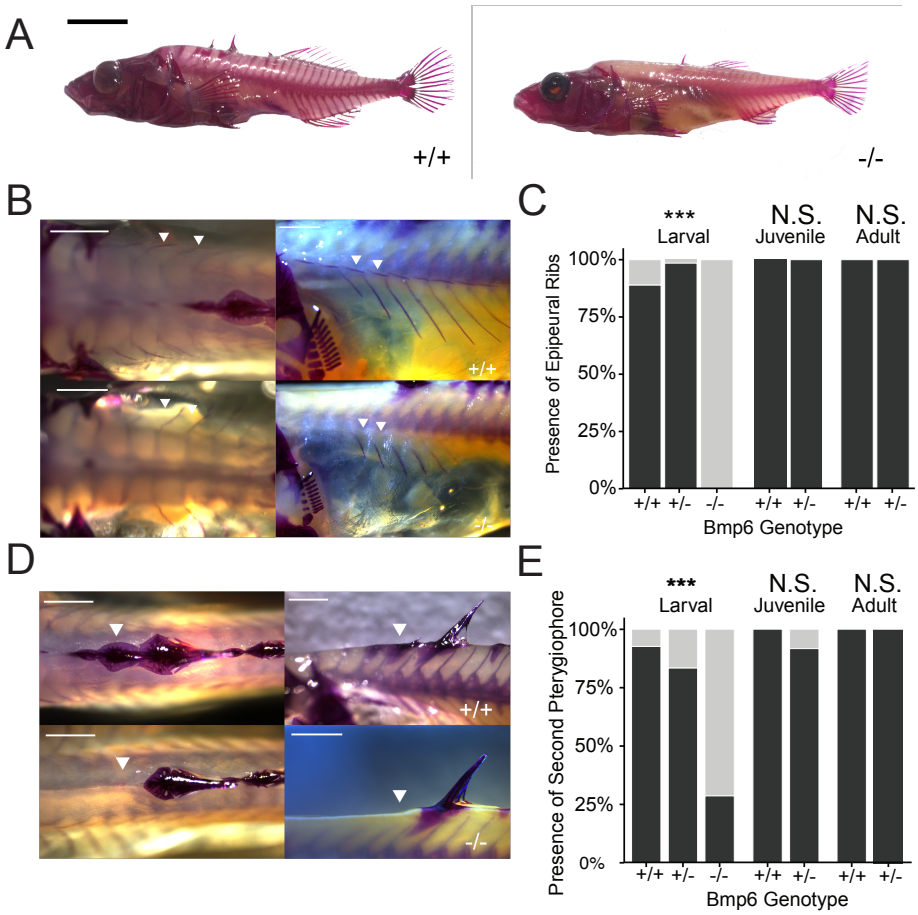


Figure 3.13. *Bmp6* is required for proper development of ribs and dorsal spines

Representative wild-type (left) and homozygous mutant stickleback (A). The *Bmp6* mutation causes a recessive loss of epipleural ribs by early larval stages of development with complete penetrance (B-C). *Bmp6* is required for the formation of the base of the first dorsal spine, the second pterygiophore (D-E). The first dorsal spine is absent in all homozygous animals without the second

pterygiophore. Dorsal (right) and lateral views (left) of both the wild-type (top) and homozygous mutants (bottom) are shown (B, D). Significant *P* values from a Fisher's exact test for the larval comparisons are 2.9 $x10^{-12}$ (C) and 0.01 (D). Scale bars are 10 mm (A) and 1 mm (B, D).

The genomic region on chromosome 21 containing stickleback *Bmp6* has been associated with several traits (e.g. number of dorsal and ventral pharyngeal teeth and the length of the branchial bones) in sticklebacks (8, 35, 38). For many of these skeletal differences, chromosome 21 has modular effects on dorsal and ventral structures (32, 33). Furthermore, BMP signaling has been shown to play a role in early specification of dorsal-ventral fate of the pharyngeal skeleton (39). To ask if *Bmp6* has different effects on dorsal and ventral skeletal structures, we measured the length of dorsal branchial bones, epibranchials (EB), and ventral branchial bones, ceratobranchials (CB), in the *Bmp6* mutant time course (Figure 3.12A-B). Strikingly, there was a significant recessive reduction of the size-corrected lengths of both the EBs and CBs bones in the homozygous mutants at early larval stages (Figure 3.12A-B). There were no differences in either of the bone lengths between wild-type and heterozygous fish at this stage. However at juvenile and adult stages of development, the ventral CBs were significantly shorter in heterozygous mutants, while there were no differences in length of the dorsal EB bones at these later stages (Figure 3.12A-B). These results support the model that *Bmp6* has stronger effects on ventral bones than dorsal bones late in development similar to what was seen with the pharyngeal teeth.

To test if off-target mutations were driving the phenotypic differences in the *Bmp6* mutants compared to the wild-type sticklebacks, we generated a trans-heterozygote cross using the two different *Bmp6* mutant alleles and found that fish trans-heterozygous for the two different mutations had similar skeletal and tooth phenotypes as fish homozygous for the 13 bp deletion (Table 3.5). Thus, off-target mutations are likely not responsible for the phenotypes.

3.3.9 *Bmp6* regulates TGFβ signaling and stem cell genes in tooth plates

To begin to identify the genetic networks downstream of *Bmp6*, we performed RNA-seq of three juvenile wild-type and homozygous mutant bilateral pharyngeal tooth plates. Following read mapping and gene expression quantification, we performed principle component analysis of normalized read count of the entire dataset. The first two principle components (PC1-2) explained a large fraction of the variance in the data explaining 31.15% and 21.08%, respectively (Figure 3.14A). As expected, PC1 discriminates between the *Bmp6* homozygous wild-type and mutant samples (Figure 3.14A).

The TOOTHcode project has identified effector genes downstream of BMP signaling in developing tooth epithelium or mesenchyme in mice (29). We tested whether stickleback orthologs of these two gene sets were differentially affected in *Bmp6* mutant toothplate tissue. Mesenchymal BMP effectors displayed significantly reduced expression in *Bmp6* mutants ($P = 8.4 \times 10^{-3}$), while epithelial BMP effectors were not significantly affected (Figure 3.14B). We also looked at a

Table 3.5. Trans-heterozygous effects on skeletal traits

Continuous Traits	Homozygous mutant (n=8)	Heterozygous (n=27)	Wildtype (n=9)	WW-WD	WD-DD	WW-DD
Ventral Tooth Number	49.73 (4.67)	54.46 (2.72)	54.83 (3.34)	0.9537	**0.0024**	**0.0070**
Dorsal Tooth Plate 1 Number	24.14 (2.6)	25.4 (1.99)	24.35 (2.64)	0.4509	0.3548	0.9804
Dorsal Tooth Plate 2 Number	64.47 (6.07)	67.5 (5.45)	66.57 (5.2)	0.9023	0.3777	0.7143
Tooth Plate Area	0.07 (0.01)	0.08 (0.01)	0.08 (0.01)	0.6802	**0.0260**	**0.0157**
Average Tooth Spacing	0.06 (0)	0.06 (0)	0.06 (0)	0.8967	0.6829	0.9384
Ceratobranchial 1 Length	1.16 (0.04)	1.2 (0.03)	1.21 (0.04)	0.7317	**0.0233**	**0.0170**
Ceratobranchial 2 Length	1.11 (0.04)	1.15 (0.04)	1.16 (0.03)	0.6182	**0.0345**	**0.0165**
Ceratobranchial 3 Length	1.05 (0.04)	1.11 (0.04)	1.12 (0.03)	0.8747	**0.0007**	**0.0016**
Ceratobranchial 4 Length	1.1 (0.04)	1.14 (0.03)	1.16 (0.03)	0.4958	**0.0020**	**0.0009**
Ceratobranchial 5 Length	1.08 (0.03)	1.11 (0.03)	1.14 (0.03)	0.0606	**0.0570**	**0.0013**
EB Length	0.43 (0.04)	0.48 (0.03)	0.48 (0.05)	0.9987	**0.0129**	**0.0407**

Discrete Traits	Fisher's Exact Test *P* value
Epipleural Ribs	**5.64e-09**
Second Pterygiophores	**0.0007**

Analysis of a trans-heterozygous cross, 13bp deletion by the 3bp deletion/4bp insertion, for all skeletal traits used in this study. There are significant recessive differences in the homozygous mutant class for all traits besides dorsal tooth plate 1 and 2 numbers and average tooth spacing, consistent with the results of the mutant time course. Continuous trait means and standard deviations (shown in brackets) for each genotypic class along with *P* values from a Tukey post-hoc test are shown. For the two discrete traits, *P* values from a Fisher's Exact Test are shown.

Table 3.6. *Bmp6* mutant class survival and fish length

Intercross	Genotype	Sample size	Mortality *P* value	Fish Length	Standard Deviation	Length *P* Value
A	+/+	6	0.75	14.53	0.80	**0.04**
	+/-	16	-	14.59	0.55	-
	-/-	6	-	13.95	0.72	-
B	+/+	9	0.31	20.56	1.31	**0.04**
	+/-	27	-	20.26	1.44	-
	-/-	8	-	19.21	1.20	-
C	+/+	25	**0.005**	25.47	4.63	**0.03**
	+/-	55	-	25.85	4.71	-
	-/-	9	-	22.15	2.60	-
Backcross						
D	+/+	32	0.35	37.61	3.07	0.52
	+/-	25	-	38.21	3.77	-
E	+/+	18	0.87	41.34	2.25	**0.05**
	+/-	17	-	39.38	3.15	-

Sample sizes, mean total fish lengths, and standard deviations for cross generating wild-type, heterozygous, and homozygous mutant fish (intercross, top) or wild-type and heterozygotes (backcross, bottom) are shown. Mortality *P* values from a Chi-square test expecting a 1:2:1 ratio for the homozygous clutches and a 1:1 ratio for the heterozygous clutches are shown. There is significant mortality in the homozygous mutant animals in clutch C, where the fish are the largest. Length *P* values from an ANOVA are shown for a recessive model (Wild-type and heterozygous classes are merged and compared to the homozygous mutants). The homozygous mutant fish are smaller than the heterozygous and the wild-type animals. One of the backcross clutches has a significant recessive size defect, which is not seen in the other clutch.

core set of effectors of BMP signaling revealed by a recent meta-analysis (78). We hypothesized that stickleback orthologs of these BMP effectors may be differentially regulated in *Bmp6* mutant sticklebacks. We found these effectors were downregulated in the *Bmp6* mutant as a whole ($P = 1.0 \times 10^{-4}$) (Figure 3.14B). These results show that stickleback *Bmp6* regulates a conserved battery of BMP-responsive genes, and that *Bmp6* is required for proper mesenchymal BMP signaling.

We hypothesized that the *Bmp6* tooth number phenotype may be explained by changes in signaling pathways known to be involved in tooth development (13,20). The TOOTHcode project manually curated a list of genes involved in mouse tooth development that function in one of eight major signaling pathways (BMP, FGF, SHH, WNT, Activin, TGFβ, NOTCH, and EDA) important for tooth development in mice (29). We asked if stickleback genes annotated as being in each of these pathways are concertedly differentially expressed in *Bmp6* mutants compared to wild types. We found the TGF-β signaling pathway to be significant down-regulated ($P = 4.7 \times 10^{-3}$) in *Bmp6* mutant tooth plates (Figure 3.14C). Strikingly, all ten TGF-β components tested had reduced expression in *Bmp6* mutant tooth plates (Figure 3.14D). We detected no significant differences in the other pathways (Figure 3.14C). Together these data suggest that *Bmp6* positively and specifically regulates TGF-β signaling in stickleback tooth plates.

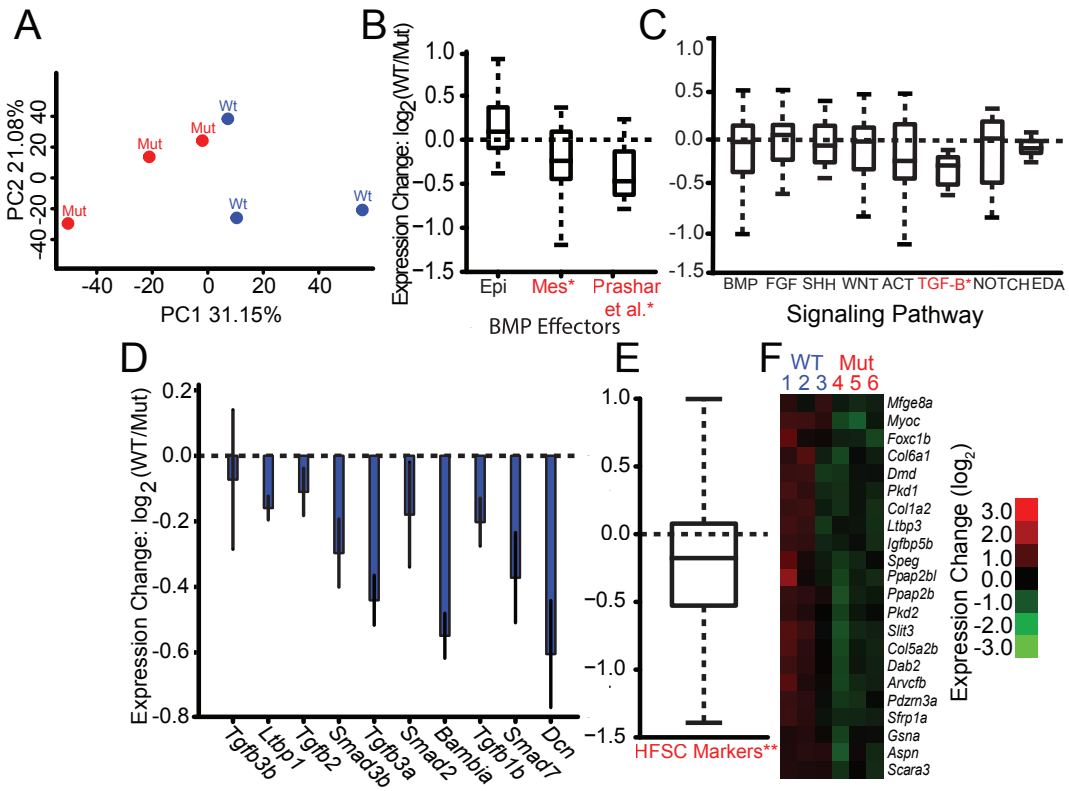


Figure 3.14. Transcriptional profiling reveals a down-regulation of TGF-β signaling and genes differentially expressed in mouse hair follicle stem cells in *Bmp6* mutant tooth plates

Principle components analysis shows that *Bmp6* genotype loads strongly on PC1 (A) Known BMP effectors from the TOOTHcode study (left two bars) are significantly downregulated in the mutant mesenchyme ($P = 8.4 \times 10^{-3}$), but not in the epithelium (error bars of SE of the mean). A set of BMP target genes (78) is significantly downregulated in mutants. Homozygous mutant fish (Mut) have significantly lower TGF-β pathway expression compared to wild-type fish (WT) ($P = 4.7 \times 10^{-3}$) (C). None of the other pathways show significant differences. Each of the TOOTHcode TGF-β genes is downregulated in the mutant (D). Genes differentially expressed in the hair follicle stem cell niche are downregulated in mutants ($P = 8.5 \times 10^{-12}$) (E). Genes differentially expressed in hair follicle stem cells were also significantly enriched for genes with differential expression at the genome-wide level ($P = 2.6 \times 10^{-9}$) (F).

In mice, *Bmp6* has been shown to inhibit the proliferation of hair follicle stem cells (40, 41). Teeth and hair are epithelial appendages with deep developmental and genetic homology (42). Thus, we hypothesized that *Bmp6* may play a conserved role of mediating stem cell quiescence during tooth replacement. A recent study characterized a set of genes that are upregulated expressed between the stem cell niche compared to surrounding tissue in the hair follicle (43). *Bmp6* mutants showed a highly significant ($P = 8.5 \times 10^{-12}$) decrease in the expression of these hair follicle stem cell markers (Figure 3.14E). Hair follicle stem cell markers were also significantly enriched for genes displaying differential expression at the genome-wide level ($P = 2.6 \times 10^{-9}$, hypergeometric test) (Figure 3.14E-F). The reduced expression of these hair follicle stem cell markers supports the hypothesis that *Bmp6* regulates stem cell quiescence during tooth replacement.

3.4 Discussion

In this study, we present genetic mapping, genome editing, and transcriptional profiling data that together further support a model where cis-regulatory changes in stickleback *Bmp6* underlie evolved increases in tooth number in derived freshwater fish.

A *cis*-regulatory down-regulation of *Bmp6* has been associated with the chromosome 21 tooth QTL (8). This QTL controls a late-acting (>25mm) tooth number increase with coinciding increases in tooth plate area with smaller changes tooth spacing. This QTL has stronger effects on ventral tooth number than dorsal tooth number (8, 35). Strikingly, the stickleback *Bmp6* mutation also has a late-acting effect on tooth number and tooth plate area with stronger effects on ventral tooth number mirroring the hallmarks of the freshwater tooth QTL supporting the model that *Bmp6* underlies aspects of evolved freshwater tooth gain. However, the direction of the *cis*-regulatory allele would predict that a mutation that lowers *Bmp6* mRNA levels would increase tooth number. One explanation is that the *Bmp6* mutation was made in a freshwater genetic background with already low levels of *Bmp6* and the further lowering of *Bmp6* inhibits tooth development. A test of this would be to analyze *Bmp6* role on tooth

development in marine sticklebacks. Alternatively, the TALEN-induced mutant alleles of *Bmp6* might not recapitulate the evolved *cis*-regulatory differences between marine and freshwater fish. *Bmp6* is expressed dynamically in the epithelium and mesenchyme at different stages of tooth development controlled by at least partially distinct two *cis*-regulatory elements (32, This study). The evolved *cis*-regulatory allele of *Bmp6* may change the spatiotemporal pattern and/or levels of *Bmp6* mRNA in different tissues, leading to different phenotypes than the exon mutation. A test of this model would be to induce mutations in the two known stickleback *Bmp6* enhancers (32, This study) and assess potential changes in tooth patterning.

The *cis*-regulatory hypothesis predicts that morphological evolution may typically proceed through *cis*-regulatory mutations that avoid the negative pleiotropy typical of coding mutations (2, 46–48). Recent studies have shown that cis-regulatory and coding mutations can drive morphological evolution and the type of mutation may depend of the degree of pleiotropy of the gene of interest (7, 8, 49–51). The partial lethality of *Bmp6* coding mutations in sticklebacks could explain why *cis*-regulatory changes of *Bmp6* have been used to evolve increases in tooth number.

There were no significant differences in tooth pattern at early developmental stages in heterozygous *Bmp6* mutant fish. However, as these heterozygous fish continue to develop to adult stages, when newly forming teeth are likely replacement teeth, the reduction of tooth number and tooth plate area became more dramatic, suggesting that tooth development at late stages are more sensitive to the dosage of *Bmp6.* These differences could be due to different developmental or genetic constraints at the early larval and late adult stages of tooth patterning. For example, there could be more functional redundancy of *Bmp6* with other BMP ligands in teeth at early developmental stages that compensate in *Bmp6* heterozygous mutants. Alternatively, these differences may signify differing roles of *Bmp6* in primary compared to replacement tooth formation. However, homozygous mutants have significantly fewer teeth at early larval stages, suggesting *Bmp6* is also required for formation of primary teeth.

We identified novel roles of *Bmp6* in patterning endochondral bones in the craniofacial branchial skeleton, as well as dermal bones in the axial skeleton (epipleural ribs), and median fin skeleton (the first dorsal spine and its underlying supporting pterygiophore). Dorsal spine lengths evolve in response to predator abundance or absence, and reduced spine lengths are a classic recurrent freshwater adaptation in sticklebacks (50). A dorsal spine length QTL was previously mapped to chromosome 21, to a broad region overlapping *Bmp6,* in a large marine x Paxton benthic F2 genetic cross (35). Based upon the dorsal spine phenotype reported here, we hypothesize that *Bmp6* underlies adaptive changes in dorsal spine lengths as well, and might represent a supergene, where multiple trophic and defense adaptive phenotypes (tooth number and spine length) are controlled by the *Bmp6* genomic region. *Bmp6* in mice has been implicated in sternum and long bone development (36, 51) suggesting a shared mechanism of bone formation of *Bmp6* in fish and mammals.

We demonstrate that eight out of nine molecularly distinct Paxton benthic chromosomes that we tested in genetic crosses have a strong effect on ventral pharyngeal tooth number. These results, along with the previously described large marine by benthic F2 cross (35), show that eight out of nine (89%) benthic chromosomes have strong effects on tooth number. The relatively high frequency of a large effect tooth QTL on chromosome 21 suggest that this QTL plays a major contribution to evolved tooth gain seen in wild Paxton benthic fish (8). We hypothesize that natural selection for more teeth drove the high frequency of the tooth QTL in Paxton lake. Alternatively, selection on other traits controlled by chromosome 21 or genetic drift may be responsible. Future genotyping across chromosome 21 in additional wild benthic fish could test whether this genomic region is under selection.

Genome resequencing of marine and freshwater sticklebacks identified a cluster of tooth QTL-associated SNPs in the fourth intron of *Bmp6*. The 10 SNP lie within a 4.4 kb interval. These SNPs also predicted the presence of a tooth QTL in one marine x freshwater cross using an independently derived freshwater population (FTC), as well as the absence of a tooth QTL in a second marine x freshwater cross from the same population where the grandparent was homozygous for the consensus marine genotype at the intron 4 tooth QTL-associated SNPs. Thus, this intron 4 haplotype is shared in at least one other freshwater population, and is thus likely present at low frequency in marine populations like other freshwater adaptive alleles (5, 24).

This association of multiple SNPs in intron 4 of *Bmp6* with the chromosome 21 tooth QTL, together with our data showing intron 4 contains a robust tooth enhancer, suggest a model where these QTL-associated SNPs at least partially underlie the tooth QTL. Although all of the tooth QTL-associated SNPs are outside of the minimally sufficient 532 bp tooth enhancer, we propose that some or all of these SNPs underlie the cis-regulatory changes in *Bmp6*. Future experiments will dissect what signals this intronic tooth enhancer is responding to, as well as what phenotypic consequences, if any, result from mutations in this enhancer. Future experiments will also test whether the marine and freshwater versions of the enhancer have different activity.

These tooth QTL-associated SNPs lie within a conserved island that contains a tooth and fin enhancer. We recently discovered and functionally dissected another tooth and fin enhancer of stickleback *Bmp6* (32). Two additional tooth and limb enhancers have been described in other species: a fin and tooth enhancer of *Dlx2b* in zebrafish and a tooth and limb bud mesenchyme enhancer of *Bmp4* in mouse (52, 53). These results add to the growing evidence that some of the genetic pathways underlying limb and tooth development are shared and regulated by shared tooth/fin cis-regulatory elements.

Functionally redundant shadow enhancers have been shown to be compensating for each other and provide robust expression during normal development, as well as during development in more variable environmental conditions (54, 55). Here we show that the intronic enhancer of *Bmp6* has partially overlapping domains with the 5' enhancer in the mesenchymal cores of

developing teeth. This redundancy of expression driven by the two enhancers may provide robust buffering of *Bmp6* expression to genetic and/or environmental changes. In addition to overlapping domains, these two enhancers also have some non-overlapping expression domains. Relative to the 5' enhancer, the intronic enhancer drives broader mesenchymal expression and more restricted epithelial expression. This modular expression from the two enhancers suggests different signaling inputs controlling the mesenchymal and epithelial expression of *Bmp6* in developing teeth. This decoupling of *Bmp6* expression in these two tooth tissues may allow for evolution to modify epithelial and mesenchymal *Bmp6* expression independently in developing teeth.

To test what genes and pathways are downstream of *Bmp6* signaling, we used RNA-seq to compare genome-wide transcriptional profiles of wild-type and homozygous mutant *Bmp6* tooth plates. Seven signaling pathways were not significantly different in this contrast, perhaps surprising given the significant difference in total tooth number in these samples. However, we found that there is a concerted and specific down-regulation of the TGF-β signaling pathway components in homozygous mutants. TGF-β signaling is known to be required for tooth development (56). Furthermore, TGF-β signaling has been shown to regulate *Bmp6* signaling in stickleback teeth through the 5' tooth enhancer (32). These results suggest that TGF-β signaling might be involved in tooth replacement and that *Bmp6* may auto-regulate itself through TGF-β signaling in developing teeth. Consistent with this model, we previously observed severe loss of *Bmp6* expression in sticklebacks with induced mutations in the 5' enhancer (32).

During mouse tooth development, reciprocal signaling events involving *Bmp4* and *Msx1* were shown to occur between developing tooth epithelium and mesenchyme. *Bmp4* expression initially comes in dental epithelium, and is required to induce *Msx1* expression in underlying mesenchyme, which in turn is required to induce *Bmp4* expression in dental mesenchyme (17, 57–59), Thus, *Bmp4* is thought to play critical roles during tooth development in both dental epithelium and mesenchyme. A large mouse gene expression study revealed sets of genes regulated by Bmp2/4/7 in dental epithelium and mesenchyme (29). We hypothesized that stickleback Bmp6 regulates genetic circuitry conserved to mice in both dental epithelium and mesenchyme, which we tested by asking whether Bmp6 mutant toothplate tissue had differential regulation of homologous genetic circuitry. Surprisingly, we found significantly reduced expression of the set of genes responsive to BMP signaling in mouse dental mesenchymal cells, while the set of genes responsive to BMP signaling in mouse dental epithelial cells was not significantly altered.

In other polyphyodonts that undergo tooth replacement, dental stem cells have been proposed to mediate tooth replacement (60–62). Teeth develop from placodes, transient embryonic thickenings that grow outwards or inwards to form epithelial appendages (42, 63). Teeth are developmentally homologous to other placode-derived organs, such as mammalian hair (42). Mammalian hairs, like fish teeth, are constantly replaced throughout adult life. During mammalian hair regeneration, *Bmp6* has been shown to regulate stem cell quiescence in the hair

follicle stem cell niche (41, 43). Thus, we hypothesized that stickleback *Bmp6* might regulate similar genetic pathways during tooth replacement as during hair regeneration. Supporting this hypothesis, we found a significant down-regulation in *Bmp6* mutant toothplate tissue of a set of genes previously described to be differentially regulated in hair follicle stem cells as they differentiate (40). This result suggests that the genetic circuitry regulating stem cell quiescence in continuously regenerating mammalian hair may be shared during constant tooth replacement in fish. This shared gene set might reflect an ancient highly conserved pathway regulating epithelial appendage regeneration. Furthermore, it is possible that the tooth number differences seen in both the *Bmp6* mutants and between stickleback populations could be due to changes in stem cell quiescence during tooth replacement, which future experiments can test.

3.5 Methods

Stickleback Husbandry

Stickleback fish were raised in 29-gallon tanks in ~1/10th ocean water (3.5 g/l Instant Ocean salt, 0.4 mL/l NaHCO3) and fed live brine shrimp as larvae, then frozen daphnia, bloodworms, and Mysis shrimp as juveniles and adults. All fish crosses were conducted using artificial fertilization. All experiments were done with the approval of the Institutional Animal Care and Use Committee from University of California, Berkeley (protocol #R330).

Recombinant mapping

Recombinant fish in derived from a Paxton Benthic freshwater by Little Campbell marine F2 were identified using microsatellite markers, Stn487 and Stn489, flanking the genetic interval surrounding *Bmp6*. Caudal fin tissue was genotyped by first isolating DNA by incubating for 20' at 94°C, then digesting with 2.5 uL of 20mg/ml proteinase K in lysis buffer (10mM Tris, pH 8.3 ; 50 mM KCL ; 1.5 mM $MgCL_2$; 0.3% Tween-20 0.3% NP-40) for 60' at 55°C followed by 20' at 94°C. One ul of undiluted DNA was used directly in the genotyping PCR. Once recombinant animals were identified, recombinant breakpoints were further mapped using a combination of microsatellite markers and RFLPs. The left and right markers used to refine each recombinant chromosome used in this study are shown in Table 3.1.

Recombinant fish were then crossed to fish heterozygous for marine and benthic chromosome 21 that were also derived from the same F2 grandparents. Recombinant crosses were raised to ~30 mm standard fish length. Fish were stained for bone with Alizarin Red, cleared, and ventral pharyngeal teeth were quantified as previously described (8). If tooth number was significantly correlated with standard fish length, sex, or family, we corrected for each using a linear model and used residuals from that regression for statistical analysis. We performed a likelihood-ratio test on each cross comparing the two models of the recombinant behaving like a marine or benthic chromosome to determine the direction of effect. As a second statistical validation for each recombinant cross,

we used the Akaike information criterion (AIC) to compare the two models. These test produced AIC differences of 11.4, 10, and 7.6 between the models for recombinant crosses 1-3, respectively. These results strongly support the direction of effect determined by the likelihood-ratio test.

Benthic F2 crosses

The benthic animals used for the F2 crosses were generated by incrossing a number of wild fish from Paxton Benthic lake, CA. Five benthic fish were crossed to marine fish and subsequently incrossed to generate six F2 crosses. The specifics of marine populations used in each cross are presented in Table 3.2. Three microsatellite markers spanning the chromosome 21 tooth QTL were genotyped: cm1440, Stn489, and Stn1488. Genotypes at these loci were used to determine the relatedness of the grandparents of these F2 crosses. F2 crosses 5 and 6 shared a benthic grandparent. This marker analysis suggests that there are nine molecularly distinct chromosome 21s in the five benthic grandparents.

To determine chromosome 21s effect on tooth number, the F2 crosses were genotyped using microsatellites markers on chromosome 21 near the tooth QTL (see Table 3.2 for details). The effects of fish size on tooth number were removed by linear regression and the residuals were back-transformed to the mean standard fish length in each cross. Statistical association between chromosome 21 genotype and back-transformed phenotypes was tested using an ANOVA in R.

Genome sequences of marine and benthic stickleback fish

We genome resequenced the four benthic grandparents from cross 1-4 and F2 animals homozygous for chromosome B_9 and B_{10} and phased sequence for the two chromosomes from the benthic grandparent of crosses 5-6. We also sequenced the marine Little Campbell grandparents from crosses 5-6, the Japanese marine grandparent from cross 3, and the grandparents the two Fishtrap creek crosses (Figure 3.2). Caudal fins were digested overnight at 55°C in Tail Digestion Buffer (10 mM Tris, pH 8.0, 100 mM NaCl, 10 mM EDTA, pH 8.0, 0.5% SDS, 10ul of 20mg/ml proteinase K). Genomic DNA was purified with a phenol:choloroform extraction followed by ethanol precipitation. Genomic libraries were generated using the Nextera DNA Sample Prep Kit (Epicentre Biotechologies), the Nextra DNA Sample Preparation Kit, or the Nextera XT DNA Library Preparation Kit (Illumina). 100 bp paired-end reads were sequenced using an Illumina HiSeq2000

Variant calling and tooth specific variant identification

Resulting reads were aligned to the reference stickleback genome (Feb. 2006 Broad Assembly) using the bwa aln and bwa sampe modules of the burrows-wheeler aligner (64). Samtools (version 0.1.17) (65) was used to create a sorted and indexed BAM file, and picard tools (version 1.51) (http://broadinstitute.github.io/picard/) was used to fix mate information, add read groups, and remove PCR duplicates. GATK's Unified Genotyper (parameters: '--

genotype_likelihoods_model INDEL', '-stand_call_conf 25', and '-stand_emit_conf 25') RealignerTargetCreator, IndelRealigner (parameter: '-LOD 0.4') was used to call potential target indels and perform realignment around indels. Base quality recalibration was accomplished using BaseRecalibrator. HaplotypeCaller (parameters: '-emitRefConfidence GVCF', '--variant_index_type LINEAR', and '--variant_index_parameter 128000') was used to generating a genomic VCF (gVCF) file for each library. The resulting gVCFs were merged and variants were called using the GenotypeGVCFs module (66-68).
High quality variants were selected using the following criteria: 1) Variants must have a variant quality score greater than 400. 2) Variants must not be called 'missing' or have a quality score of less than 10 in either high-coverage benthic genome. 3) Variants must not be called 'missing' or have a quality score of less than ten in no more than two genomes. In ensure high quality genotypes, variants were called on a repeat masked genome. To future remove stickle specific repeats, we removed variants with >99% of the 100bp flanking sequence matching more than 6 places in the genome using blastn with an e-value of less than $1x10^{-30}$ (69). QTL concordance score is the absolute value of the number of times a variant was present in benthic animals with a chromosome 21 tooth QTL minus the number of times the same variant was found in animals without a tooth QTL. Concordance scores were normalized such that the max possible score is 1. QTL Concordance scores were calculated using a custom python script.

Generation of transgenic enhancer stickleback lines
To generate GFP reporter constructs, each of the intron 4 fragments from the benthic and marine grandparents from cross 5 were cloned upstream of the Hsp70 promoter in a Tol2 expression construct using NheI (32). For the mCherry constructs, we cloned mCherry into the Hsp70 reporter construct using SalI and ClaI and the inserts were cloned upstream using NheI and BamHI. Primers for construct generation and sequencing are shown in Table 3.3.

To generate transgenic stickleback, transposase messenger RNA was synthesized from pCS2-TP plasmid linearized with NotI and transcribed using the mMessage SP6 *in vitro* transcription kit (Ambion) and purified using the Qiagen RNeasy column. 1-cell marine stickleback embryos were injected with a mixture of 37.6 ng/uL plasmid DNA and 75 ng/uL RNA with 0.05% phenol red as previously described (32).

Generation of TALEN construct targeting stickleback *Bmp6*
To generate a TALEN pair to target the stickleback *Bmp6* gene, we used the TAL effector Nucleotide Targeter 2.0 (https://tale-nt.cac.cornell.edu/node/add/talen) to scan the second exon sequence of *Bmp6* for potential target sites (32, 33). We chose TALEN parameters as described (33). We chose a target site that is unique to *Bmp6* in the stickleback genome and contains a common restriction site, EcoRI, which can be used to detect molecular deletions. We assembled the two TALEN constructs using Gateway cloning into the destination vectors pCS2TALDD and pCS2TALRR and verified correct assembly using Sanger

sequencing as described (33). See table 3.4 for the specifics of the *Bmp6* TALEN design.

Synthesis and injection of TALEN RNA into stickleback embryos
5'-capped mRNA for each TALEN pair was transcribed using the SP6 mMessage Machine (Ambion) after the TALEN plasmid templates had been linearized with NotI. Pooled TALEN mRNA was injected into 1-cell stickleback embryos at a concentration of 40 ng/uL for each mRNA with 0.05% phenol red.

Talen mutation identification
To genotype fish for TALEN induced mutations, we clipped a small fragment of caudal tail fin from adult fish or homogenized whole 3 dpf embryos. DNA was extracted as described in the recombinant section and the genotyping PCR was performed using forward primer 5'- ACAAGCCGCTAAAAAGGACA-3' and reverse primer 5'- GCACGTGTGCATGCTTTAGA -3'. The reaction profile for the NEB Phusion reaction was 98°C for 30'', 39 cycles of 98°C for 10'', 58°C for 15'', 72°C for 30'', and 72°C for 10'. The PCR products were cut directly with EcoRI. The product is cut from the wild-type allele and uncut from the mutant allele.

Axial, fin, and pharyngeal skeleton morphology and quantification
Axial and fin skeletal phenotypes, were scored as presence (having an Alizarin positive skeletal element) or absence (lacking an Alizarin stained element) using a Lecia S6E dissecting microscope. Dorsal and ventral pharyngeal tooth number was quantified on a DM2500 Leica microscope using a TX2 filter as previously described (8). Tooth plate area and spacing of the ventral pharyngeal tooth plate was quantified from a gray scale image taken with a DFC340 Fx camera on a Leica M165FC as previously described (8). Total tooth number is the sum of the left and right sides for ventral and dorsal pharyngeal tooth plates. Area and spacing of the ventral pharyngeal tooth plates are the averages of the left and right tooth plate. Brachial bones measurements were measured from an image taken with a DFC340 Fx camera on a Leica M165FC as previously described (38). Bone measurements were presented as the average of the left and right side. Skeletal traits were binned by total fish length for three stages: larval <27 mm, juvenile 27-37 mm, and adults >37 mm. For the qualitative traits, ribs and pterygiophore, we conducted a Fisher's exact test within the bins. For the continuous traits, tooth number and bone lengths, if the trait was significantly correlated with fish length within these bins, we removed the effects of fish size with a linear regression and then back-transformed to average fish size within the bin.

RNA purification, sequencing, and alignment
Ventral tooth plates from three wild-type and homozygous mutant *Bmp6* female sticklebacks (standard length ~25 mm) were dissected, placed into TRI reagent (Sigma-Aldrich) on ice, and ground with a disposable pestle. RNA was extracted, isopropanol precipitated, and resuspended in DEPC-treated water. 200 ng of purified RNA was used with Illumina's Truseq Stranded mRNA Library Prep Kit to create sequencing libraries. The resulting bar-coded libraries were

pooled and 100 bp paired end reads were generated using the Illumina HiSeq2000. Reads were mapped to the stickleback reference genome (Feb. 2006 Broad Assembly) using STAR (parameters: '--alignIntronMax 200000' '--alignMatesGapMax 200000' '--outFilterMultimapNmax 8') (72). BAM files were created, sorted, and indexed using Samtools (version 0.1.17) (65). Picard tools (version 1.51) was used to fix mate information, add read groups, and remove PCR duplicates (http://broadinstitute.github.io/picard/) . Using the ENSEMBLE reference transcriptome, transcripts were quantified using cuffquant version 2.2.1 (parameters: '-u' '--library-type fr-firststrand'), normalized using cuffnorm, and differential expression was detected using an FDR of 0.10 (73, 74). Principal component analysis of the resulting transcript abundances was done using the PCA package of FactoMineR (http://factominer.free.fr/index.html) in R, and was plotted in R. Additional figures and analyses were done using custom python scripts and figures created using matplotlib. Hierarchical clustering was done using Cluster3.0 (parameters: '-l' '-cg a' '-g 2' '-e 0' '-m c') (75), and the results were visualized using JavaTreeView (version 1.1.6r4) (76).

Gene Set Enrichment Analysis:
Gene sets were downloaded from the ToothCode database (http://compbio.med.harvard.edu/ToothCODE/). ToothCode identified downstream targets of Bmp signaling by literature mining manipulations of *Bmp2, Bmp4, and Bmp7*. Targets that were upregulated when BMP signaling increased or downregulated when BMP signaling was decreased were termed BMP effectors. Hair follicle stem cell expression data was obtained from Kandyba 2013 (43). Genes upregulated in the bulge relative to the hair germ were termed hair follicle stem cell markers. Orthologous stickleback genes were identified using ENSEMBLE predictions. Statistical enrichment was done similar to the methods as previously described (77). Each gene in a set was subject to a t-test, obtaining a list of z-scores. The null hypothesis, that the gene set displays no differential expression enrichment, (i.e. t-test z-scores are drawn from a standard normal distribution) was tested using a 1-sample t-test, with resulting *P* values subject to a Bonferronni correction. These results were confirmed by a simulation using 10,000 permutations of an equal number of genes as in the gene set, randomly chosen without replacement, subject to a Bonferronni corrected *P* value less than .05. Analysis was done using a set of custom python scripts.

3.6 References

1. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134(1):25–36.
2. Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* 8(3):206–216.
3. Martin A, Orgogozo V (2013) The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67(5):1235–50.
4. Stern DL (2013) The genetic causes of convergent evolution. *Nat Rev Genet* 14(11):751–64.
5. Colosimo PF, et al. (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* 307(5717):1928–33.
6. Chan YF, et al. (2010) Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. *Science (80-)* 327(5963):302–305.
7. Miller CT, et al. (2007) cis-Regulatory Changes in Kit Ligand Expression and Parallel Evolution of Pigmentation in Sticklebacks and Humans. *Cell* 131(6):1179–1189.
8. Cleves PA, et al. (2014) Evolved tooth gain in sticklebacks is associated with a cis-regulatory allele of Bmp6. *Proc Natl Acad Sci U S A* 111(38):13912–7.
9. Wark AR, et al. (2012) Genetic Architecture of Variation in the Lateral Line Sensory System of Threespine Sticklebacks. *G3 Genes|Genomes|Genetics* 2(9):1047–1056.
10. Gray S, Levine M (1996) Transcriptional repression in development. *Curr Opin Cell Biol* 8(3):358–64.
11. Payankaulam S, Li LM, Arnosti DN (2010) Transcriptional repression: conserved and evolved features. *Curr Biol* 20(17):R764–71.
12. Jernvall J, Thesleff I (2012) Tooth shape formation and tooth renewal: evolving with the same signals. *Development* 139:3487–3497.
13. Catón J, Tucker AS (2009) Current knowledge of tooth development: patterning and mineralization of the murine dentition. *J Anat* 214(4):502–515.
14. Neubuser A (1997) Antagonistic Interactions between FGF and BMP Signaling Pathways: A Mechanism for Positioning the Sites of Tooth Formation. *Cell* 90(2):247–255.
15. St Amand TR, et al. (2000) Antagonistic Signals between BMP4 and FGF8 Define the Expression of Pitx1 and Pitx2 in Mouse Tooth-Forming Anlage. *Dev Biol* 217(2):323–332.
16. Kavanagh KD, Evans AR, Jernvall J (2007) Predicting evolutionary patterns of mammalian teeth from development. *Nature* 449(7161):427–32.
17. Bei M, Kratochwil K, Maas RL (2000) BMP4 rescues a non-cell-autonomous function of Msx1 in tooth development. *Development* 127(21):4711–4718.

18. Andl T (2004) Epithelial Bmpr1a regulates differentiation and proliferation in postnatal hair follicles and is essential for tooth development. *Development* 131(10):2257–2268.
19. Wang Y, et al. (2012) BMP activity is required for tooth development from the lamina to bud stage. *J Dent Res* 91(7):690–5.
20. Tucker AS, Fraser GJ (2014) Evolution and developmental diversity of tooth regeneration. *Semin Cell Dev Biol*.
21. Kingsley DM, Peichel CL (2007) The molecular genetics of evolutionary change in sticklebacks. The Biology of the Threespine Stickleback. Sara Ostlund-Nilsson, I. Mayer, and F.A. Huntingford. *CRC Press*:41–81.
22. Bell MA, Foster SA (1994) *The Evolutionary Biology of the Threespine Stickleback*.
23. O'Brown NM, Summers BR, Jones FC, Brady SD, Kingsley DM (2014) A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene EDA. *Elife* 4:e05290.
24. Jones FC, et al. (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484(7392):55–61.
25. Hohenlohe PA, et al. (2010) Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLoS Genet* 6(2):e1000862.
26. Schluter D, McPhail JD (1992) Ecological character displacement and speciation in sticklebacks. *Am Nat* 140:85–108.
27. Wise SB, Stock DW (2006) Conservation and divergence of Bmp2a, Bmp2b, and Bmp4 expression patterns within and between dentitions of teleost fishes. *Evol Dev* 8:511–523.
28. Muschick M, Indermaur a., Salzburger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol* 22:2362–8.
29. O'Connell DJ, et al. (2012) A Wnt-bmp feedback circuit controls intertissue signaling dynamics in tooth organogenesis. *Sci Signal* 5(206):ra4.
30. Fraser GJ, et al. (2009) An Ancient Gene Network Is Co-opted for Teeth on Old and New Jaws. *PLoS Biol* 7(2):e31.
31. Jackman WR, Draper BW, Stock DW (2004) Fgf signaling is required for zebrafish tooth development. *Dev Biol* 274(1):139–157.
32. Erickson PA, et al. (2015) A 190 base pair, TGF-β responsive tooth and fin enhancer is required for stickleback Bmp6 expression. *Dev Biol*.
33. Dahlem TJ, et al. (2012) Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet* 8(8):e1002861.
34. Chang Y-F, Imam JS, Wilkinson MF (2007) The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* 76:51–74.
35. Miller CT, et al. (2014) Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait Loci. *Genetics* 197(1):405–20.
36. Solloway MJ, et al. (1998) Mice lacking Bmp6 function. *Dev Genet* 22(4):321–39.

37. Smith A, Avaron F, Guay D, Padhi BK, Akimenko MA (2006) Inhibition of BMP signaling during zebrafish fin regeneration disrupts fin growth and scleroblast differentiation and function. *Dev Biol* 299(2):438–454.
38. Erickson PA, Glazer AM, Cleves PA, Smith AS, Miller CT (2014) Two developmentally temporal quantitative trait loci underlie convergent evolution of increased branchial bone length in sticklebacks. *Proc Biol Sci* 281(1788):20140822.
39. Alexander C, et al. (2011) Combinatorial roles for BMPs and Endothelin 1 in patterning the dorsal-ventral axis of the craniofacial skeleton. *Development* 138(23):5135–46.
40. Genander M, et al. (2014) BMP Signaling and Its pSMAD1/5 Target Genes Differentially Regulate Hair Follicle Stem Cell Lineages. *Cell Stem Cell* 15(5):619–633.
41. Blanpain C, Lowry WE, Geoghegan A, Polak L, Fuchs E (2004) Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* 118(5):635–648.
42. Biggs LC, Mikkola ML (2014) Early inductive events in ectodermal appendage morphogenesis. *Semin Cell Dev Biol* 25-26:11–21.
43. Kandyba E, et al. (2013) Competitive balance of intrabulge BMP/Wnt signaling reveals a robust gene network ruling stem cell homeostasis and cyclic activation. *Proc Natl Acad Sci U S A* 110(4):1351–6.
44. Carroll SB (2000) Endless forms: the evolution of gene regulation and morphological diversity. *Cell* 101:577–580.
45. Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution (N Y)* 61(5):995–1016.
46. King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.
47. Chan YF, et al. (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327:302–305.
48. Stern DL, Frankel N (2013) The structure and evolution of cis-regulatory regions: the shavenbaby story. *Philos Trans R Soc Lond B Biol Sci* 368(1632):20130028.
49. Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313(5783):101–4.
50. Reimchen TE (1980) Spine deficiency and polymorphism in a population of Gasterosteus aculeatus : an adaptation to predators ? *Can J Zool* 58:1232–1244.
51. Perry MJ, McDougall KE, Hou S-C, Tobias JH (2008) Impaired growth plate function in bmp-6 null mice. *Bone* 42(1):216–25.
52. Jumlongras D, et al. (2012) An evolutionarily conserved enhancer regulates Bmp4 expression in developing incisor and limb bud. *PLoS One* 7(6):e38568.

53. Jackman WR, Stock DW (2006) Transgenic analysis of Dlx regulation in fish tooth development reveals evolutionary retention of enhancer function despite organ loss. *Proc Natl Acad Sci U S A* 103(51):19390–5.
54. Perry MW, Boettiger AN, Bothma JP, Levine M (2010) Shadow enhancers foster robustness of drosophila gastrulation. *Curr Biol* 20(17):1562–1567.
55. Hong J-W, Hendrix DA, Levine MS (2008) Shadow enhancers as a source of evolutionary novelty. *Science* 321(5894):1314.
56. Klopcic B, et al. (2007) TGF-beta superfamily signaling is essential for tooth and hair morphogenesis and differentiation. *Eur J Cell Biol* 86(11-12):781–99.
57. Chen Y, Bei M, Woo I, Satokata I, Maas R (1996) Msx1 controls inductive signaling in mammalian tooth morphogenesis. *Development* 122:3035–3044.
58. Bei M (2009) Molecular genetics of tooth development. *Curr Opin Genet Dev* 19(5):504–510.
59. Vainio S, Karavanova I, Jowett A, Thesleff I (1993) Identification of BMP-4 as a signal mediating secondary induction between epithelial and mesenchymal tissues during early tooth development. *Cell* 75(1):45–58.
60. Handrigan GR, Richman JM (2010) A network of Wnt, hedgehog and BMP signaling pathways regulates tooth replacement in snakes. *Dev Biol* 348(1):130–41.
61. Handrigan GR, Leung KJ, Richman JM (2010) Identification of putative dental epithelial stem cells in a lizard with life-long tooth replacement. *Development* 137(21):3545–3549.
62. Wu P, et al. (2013) Specialized stem cell niche enables repetitive renewal of alligator teeth. *Proc Natl Acad Sci U S A* 110(22):E2009–18.
63. Ahn Y (2015) Signaling in tooth, hair, and mammary placodes. *Curr Top Dev Biol* 111:421–59.
64. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.
65. Li H, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–9.
66. McKenna A, et al. (2010) The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303.
67. Van der Auwera GA, et al. (2013) From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinforma*.
68. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498.
69. Camacho C, et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10(1):421.
70. Cermak T, et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 39(12):e82.

71. Doyle EL, et al. (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res* 40(Web Server issue):W117–22.
72. Dobin A, et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15–21.
73. Trapnell C, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28(5):511–5.
74. Trapnell C, et al. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31(1):46–53.
75. De Hoon MJL, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. *Bioinformatics* 20(9):1453–4.
76. Saldanha AJ (2004) Java Treeview--extensible visualization of microarray data. *Bioinformatics* 20(17):3246–8.
77. Irizarry RA, Wang C, Zhou Y, Speed TP (2009) Gene set enrichment analysis made simple. *Stat Methods Med Res* 18(6):565–575.
78. Prashar P, Yadav PS, Samarjeet F, Bandyopadhyay A (2014) Microarray meta-analysis identifies evolutionarily conserved BMP signaling targets in developing long bones. *Dev Biol* 389(2):192–207.