SI Appendix

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

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

This SI file includes:
Supplemental Materials and Methods and References
Supplementary Tables 1-5
Supplementary Figures 1-7

Supplemental Materials and Methods

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 (1) 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 tooth number, tooth plate area, and tooth spacing

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 $\alpha = 0.05$. 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 (3) 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 $\alpha = 0.05$. 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 (4).

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 S5 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.

Riboprobe design and synthesis

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 amplified 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.

References

(1) Walker MB & Kimmel CB (2007) A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotechnic & Histochemistry* 82(1):23-28.

(2) Broman KW & Sen S (2009) *A Guide to QTL Mapping with R/qtl* (Springer, New York).

(3) Miller CT*, et al.* (2014) Modular Skeletal Evolution in Sticklebacks Is Controlled by Additive and Clustered Quantitative Trait Loci. *Genetics* 197(1):405-20.

(4) Jones FC*, et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484(7392):55-61.

Supplementary Tables

Table S1. Wild and lab-reared tooth numbers of marine and freshwater fish

	Population	Sample size	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 1.

Table S2. 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 S3. 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 = PAXB x RABS 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.

Table S4. 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.

Table S5. 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.

Supplementary Figures

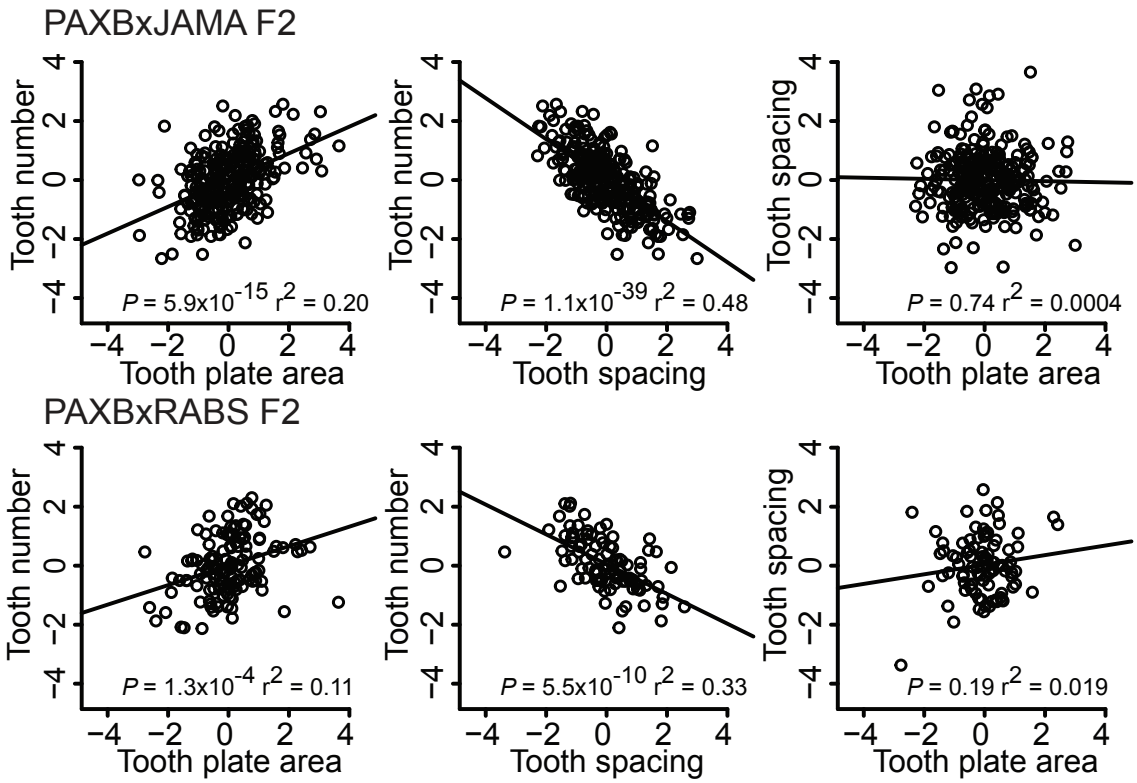

Figure S1. Correlations of tooth number, area, and spacing in two F2 genetic crosses

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.

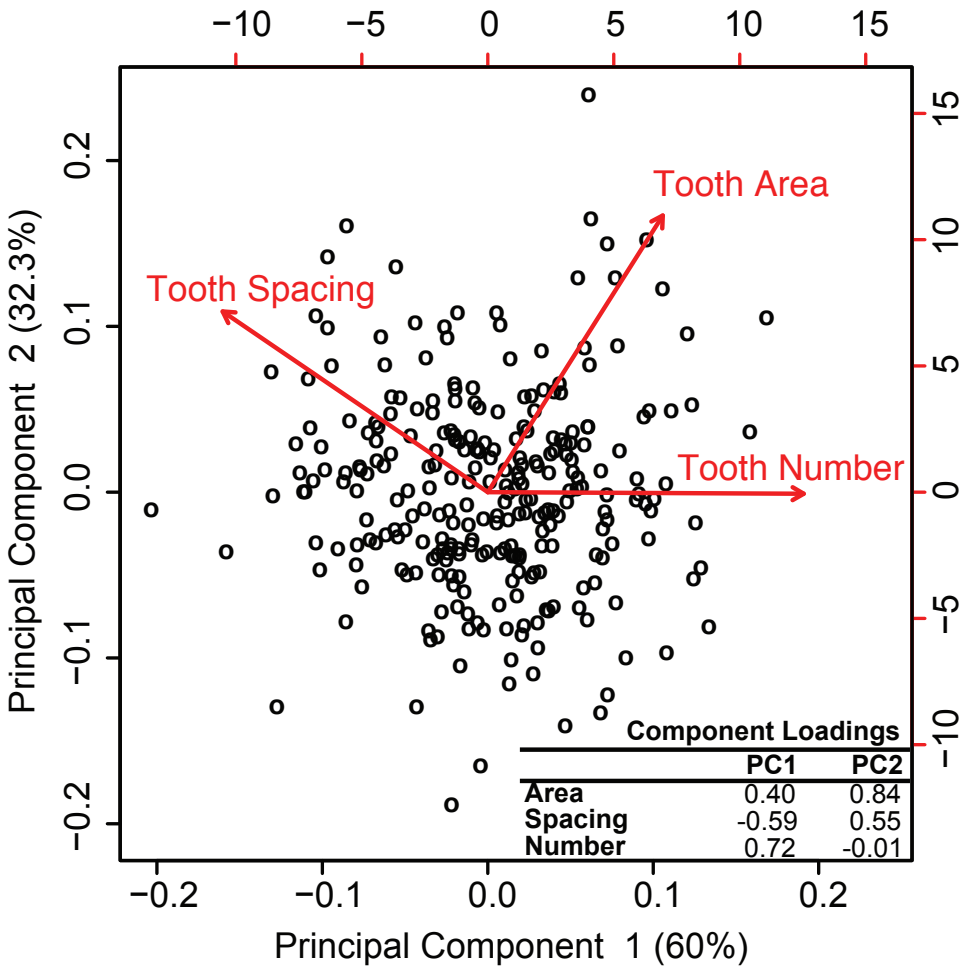

Figure S2. Biplot of principal component analysis of tooth patterning phenotypes

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.

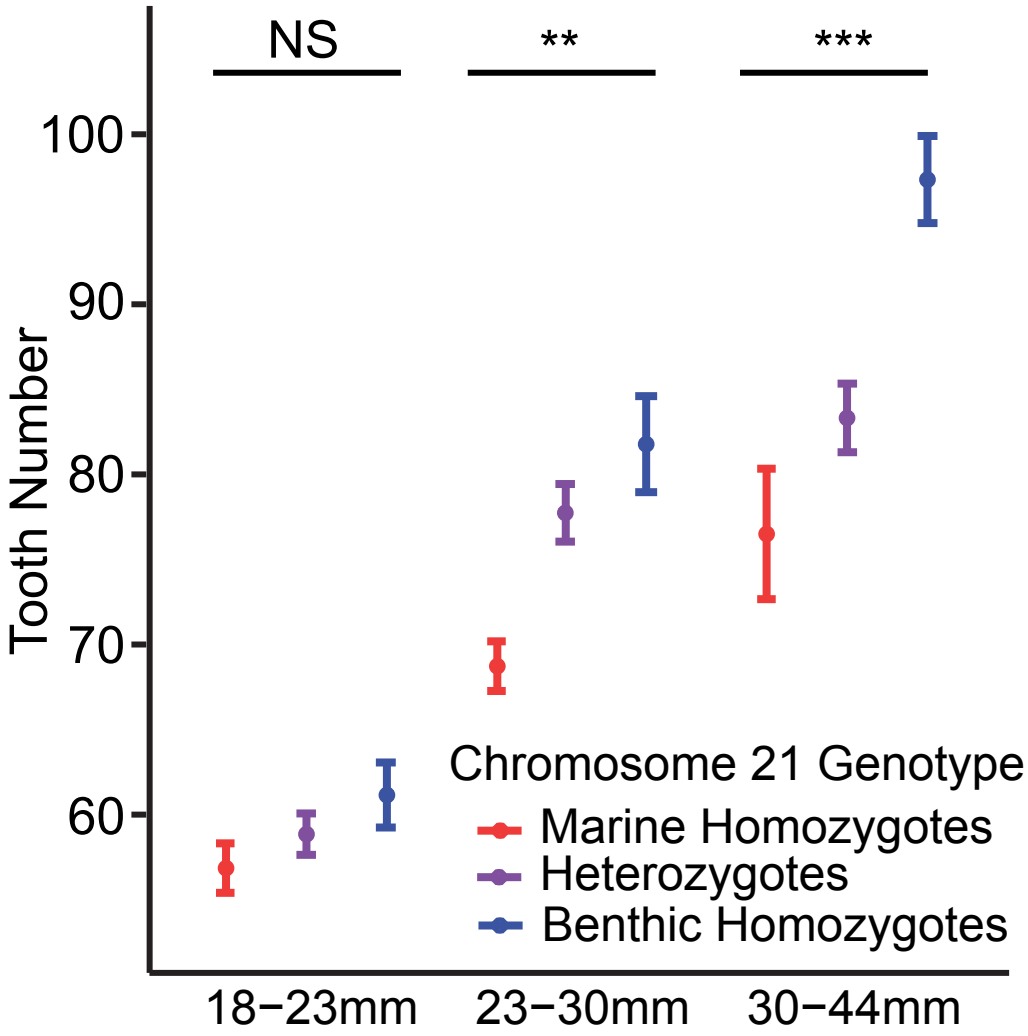

Figure S3. Developmental effects of the chromosome 21 tooth QTL

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).

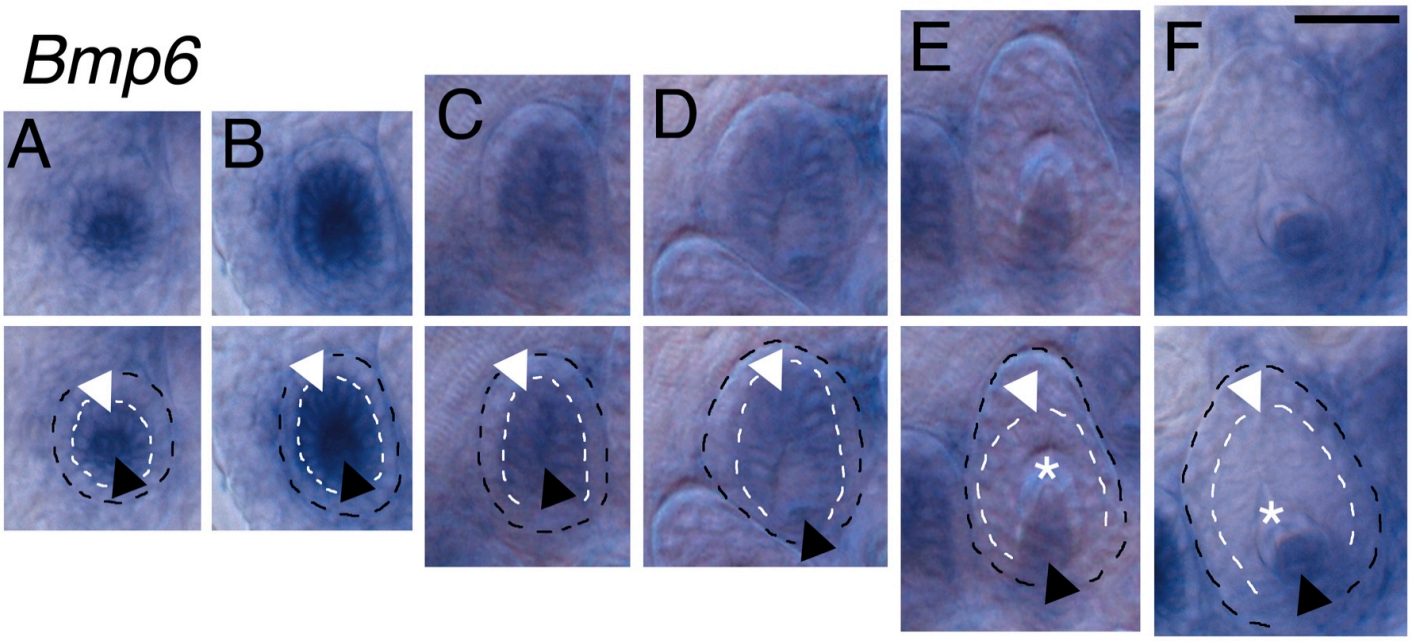

Figure S4. Dynamic *Bmp6* expression in developing teeth. Successive stages of tooth development from left to right, showing early germ stage (A) through newly mineralized baby tooth (F). Unlabeled images at top, and labeled versions below. As the pharyngeal tooth field has multiple germs forming at different stages (see Fig. 5G-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.

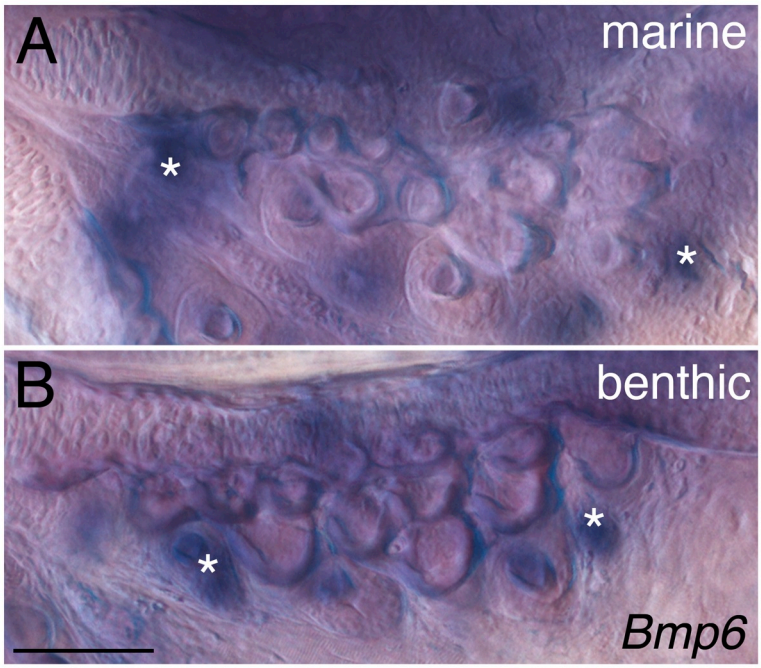

Figure S5. *Bmp6* expression during larval tooth development

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.

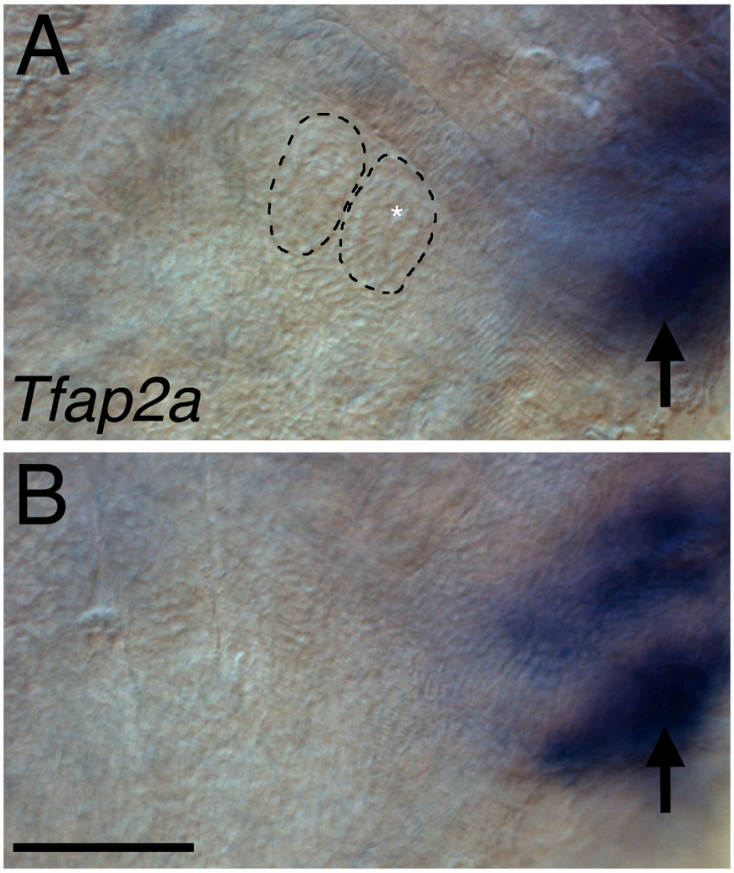

Figure S6. Expression of *Tfap2a* in 7.5 days post fertilization benthic fish. No *Tfap2a* expression was detected in developing stickleback tooth germs (A), although robust expression was detected in mesenchymal cells associated with the developing epibranchial cartilages. Shown are two focal planes, 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.

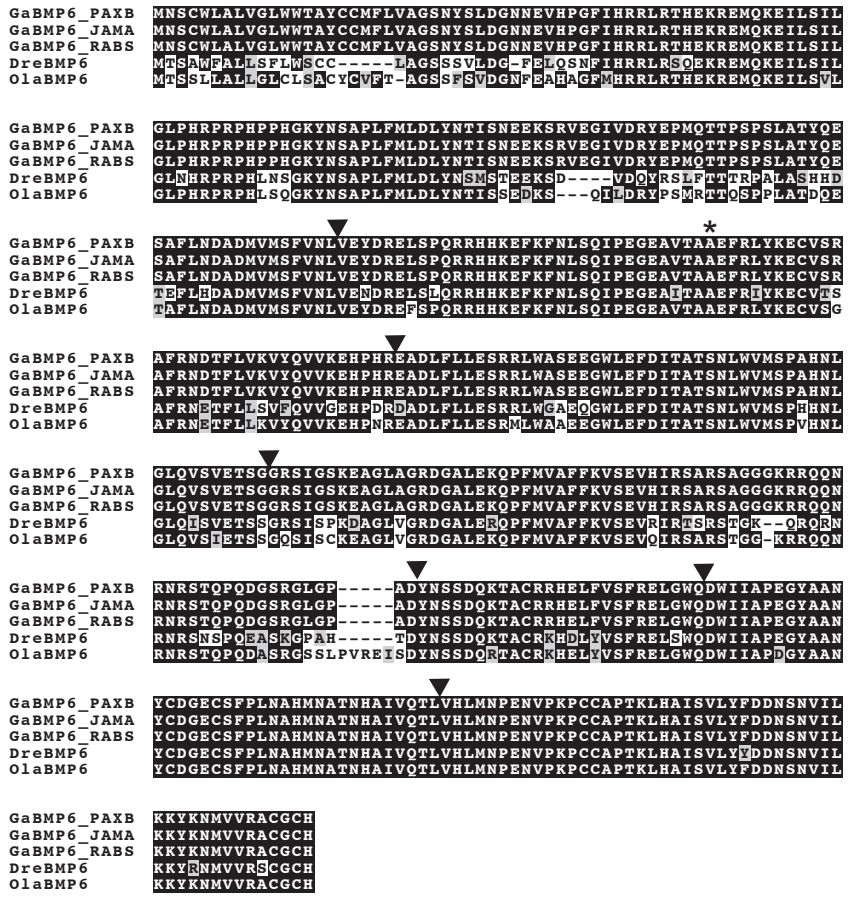

Figure S7. Predicted amino acid alignment of BMP6 in fish

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.