Supplementary Methods

Statistical analysis and phenotype corrections

Data were analyzed using R and Prism 5. For comparing tooth phenotypes between populations, one way analysis of variance (ANOVA) using a Tukey-Kramer *post hoc* test was performed for statistical analysis between greater than 2 groups unless otherwise noted. Total length (TL), tip of snout to end of tail, measurements were used in datasets where fish were <15mm, while standard length (SL), tip of snout to base of caudal peduncle, measurements were used when fish were >15mm. Phenotypes were size and/or sex corrected when appropriate.

Adult (~6 month old) lab-reared ventral pharyngeal tooth number, tooth plate area, and intertooth spacing phenotypes were all corrected for fish size (i.e. phenotypes were back transformed residuals for a regression to standard length for a mean standard length of 37 mm). Neither dorsal pharyngeal tooth plate tooth numbers correlated with standard length, so these were not size corrected. CERC x marine F2 cross phenotypes were regressed to standard length and/or corrected for sex and/or log-transformed when the transformation equalized variances by Levene's test for equality of variances and/or normalized the residuals by an Anderson-Darling test of normality when appropriate. Final corrected phenotypes for mapping QTL were ventral tooth number (raw), ventral tooth plate area and intertooth spacing (both log transformed and corrected for fish standard length), dorsal tooth plate 1 tooth number (raw), and dorsal tooth plate 2 tooth number (corrected for fish standard length).

Tooth germ number quantification

Germ number was quantified by counting un-erupted developing teeth on the ventral pharyngeal tooth plate in 6 µm serial sections of 4-6 individuals for each population and time point. Germs were sorted by stage (bud, cap, early to mid-bell, late bell) and germ area obtained by tracing the outer diameter of the outer dental epithelium (ODE) in ImageJ (Schneider et al., 2012). Torn or rippled germ sections were omitted from area measurements. Adult tooth width was measured by using the basement membrane of the epithelium (stratum compactum) as a landmark. Height was measured along a line perpendicular to the width measurement to the tip of the tooth in ImageJ.

Processing reads from grandparent resequencing and F2 GBS libraries

The Cerrito Creek (CERC) grandfather of the CERC x marine (Little Campbell River, LITC) F2 cross was sequenced using a Nextera DNA Sample Preparation kit (Illumina) followed by sequencing to ~6X coverage with 100 base, paired-end sequencing on an Illumina HiSeq 2000 sequencer (SRA accession # SRS951365). Reads were mapped to the reference genome with BWA (www.bio-bwa.sourceforge.net), SNPs were called with SAMtools (www.samtools.sourceforge.net), and these SNPs were filtered for positions homozygous for an alternate allele. Genotyping-by-Sequencing (GBS) was performed as previously described (Glazer et al., 2015) with the exception that only one grandparent was used for phasing F2 genotypes and phasing was performed separately in two half sibling families. First, SNP positions that were homozygous alternate in the CERC grandparent were phased in each family (n=7,606 and n=13,477 respectively), pooled into bins, and genotypes were calculated for each bin of SNPs. Bins did not span scaffold boundaries and scaffolds were equally divided

into bins with a maximum size of 500kb (see Glazer et al., 2015; see supplementary material Table S4). Second, these genotypes were used to further phase additional SNPs in the F2s. SNP positions that correlated above 80% with called genotypes were phased and included in the pooled genotypes. Three rounds of this phasing were repeated resulting in a total of 28,283 phased SNPs binned into 761 markers in family one and 50,000 phased SNPs binned into 999 markers in family two. Three fish that had missing data for over 50% of markers were removed from the analysis. Thirty markers with missing data for more than 40% of fish were removed from the analysis. This resulted in 974 markers and 171 fish in the merged data set with over 91% of all possible genotypes present. Genetic linkage maps were created with JoinMap 4.0 (Kyazma) with regression mapping and default settings. Markers with skewed genotypes were determined based on high confidence genotypes (determined from a minimum of 10x coverage). Genotype ratios that deviated significantly from the expected 1:2:1 ratio using a chi-squared test were dropped ($P < 0.01$) except in cases where multiple linked adjacent markers significantly deviated. Twenty-three markers were dropped as chi-squared outliers and 21 markers did not fit into the linkage map resulting in a map with 930 markers (see supplementary material Table S4).

Supplementary references

Glazer, A. M., Killingbeck, E. E., Mitros, T., Rokhsar, D. S. and Miller, C. T. (2015) Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with Genotyping-by-Sequencing. *G3*, *in press*.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–75.

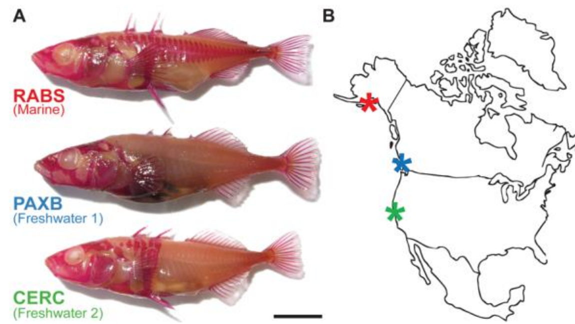

Figure S1: Description and location of independent stickleback populations.
(A) Representative example adult male from each population stained with Alizarin Red S marking bone. Scale bar is 10 mm. (B) Source of each population on a map of North America denoted by color coded asterisks. RABS (marine) is from Rabbit Slough, Alaska, PAXB (freshwater 1) is from Paxton Lake, British Colombia, CERC (freshwater 2) is from Cerrito Creek, California.

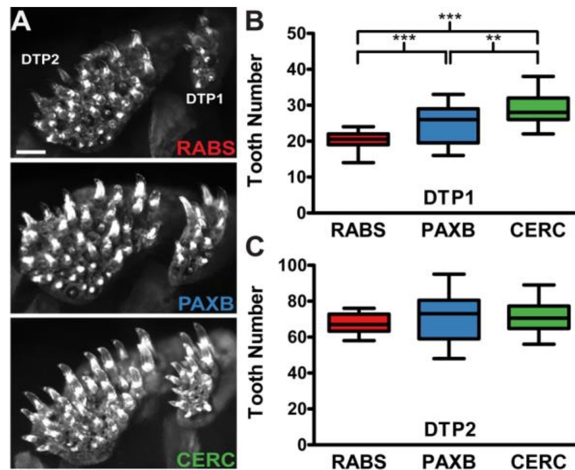

Figure S2: Tooth number for dorsal tooth plate 1, but not 2, differs between populations.
(A) Representative unilateral dorsal tooth plates. Scale bar is 200 μm. (B) Quantification of total DTP1 tooth number. (C) Quantification of total DTP2 tooth number. (B-C) Respective sample size for each trait: *n*=20 RABS, *n*=37 PAXB, *n*=25,26 CERC. ****P*<0.001, ***P*<0.01 (one-way ANOVA using a Tukey-Kramer *post hoc* test).

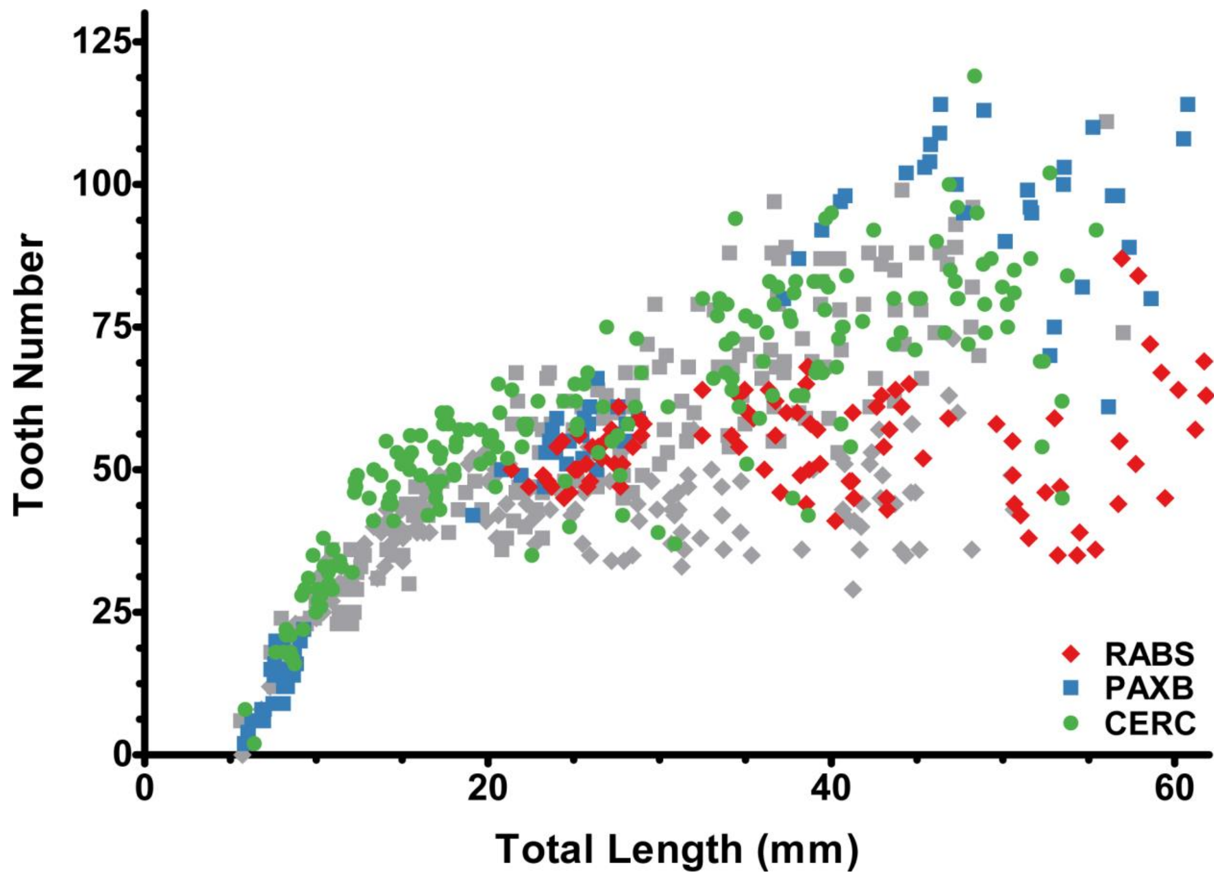

Figure S3: Previously published PAXB and RABS data points.
Points denoted in gray were previously published in Cleves *et al.,* 2014 and included for comparison to CERC and later time points.

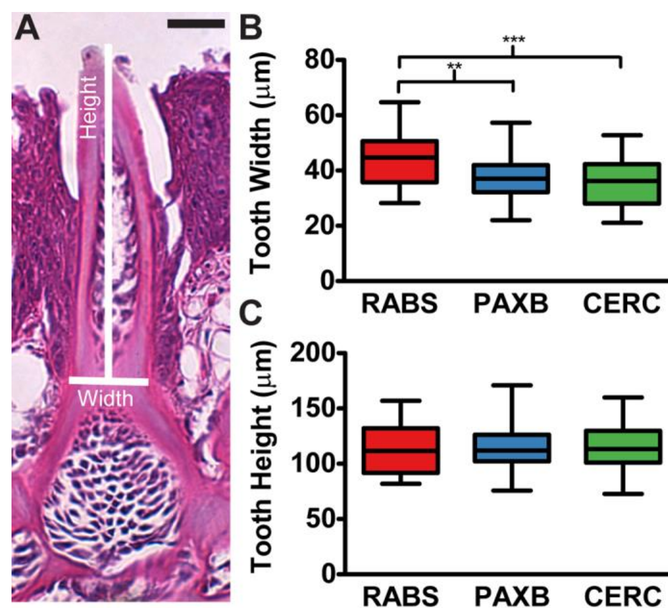

Figure S4: Adult marine and freshwater teeth vary in width, but not height.
(A) Phenotyping example of adult teeth. Scale bar is 25 μm. (B) Tooth width. (C) Tooth height. (B-C) Sample Size: n=29 RABS, n=44 PAXB, n=37 CERC. ***P<0.001, **P<0.01 (one-way ANOVA using a Tukey-Kramer *post hoc* test).

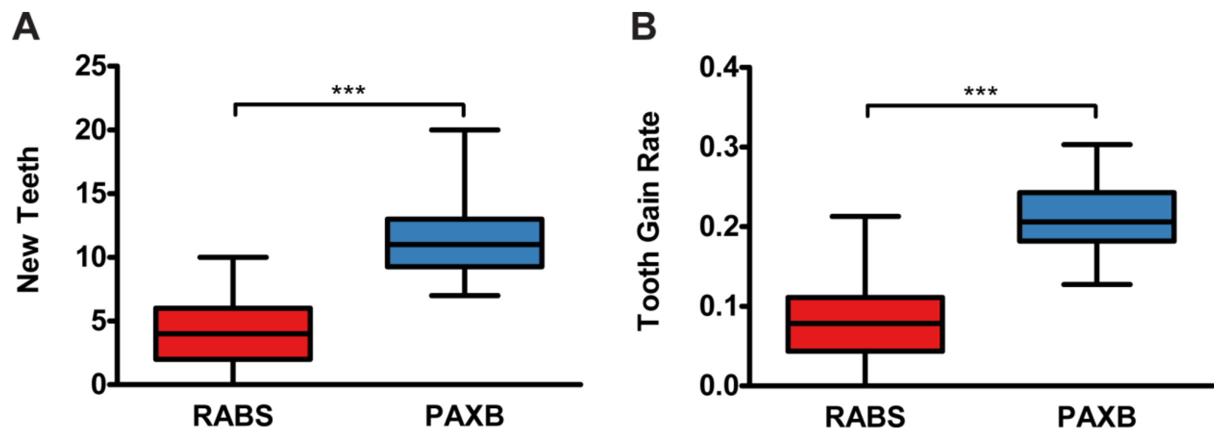

Figure S5: Early pulse-chase reveals tooth gain rates are relatively fixed early in marine and freshwater sticklebacks.
(A-B) Pulse-chase performed on two month, ~20 mm, PAXB and RABS sticklebacks show new tooth number (A) and tooth gain rate (B) differences are established early and are comparable to adult gain rates. ****P*<0.001 (two-tailed t-test).

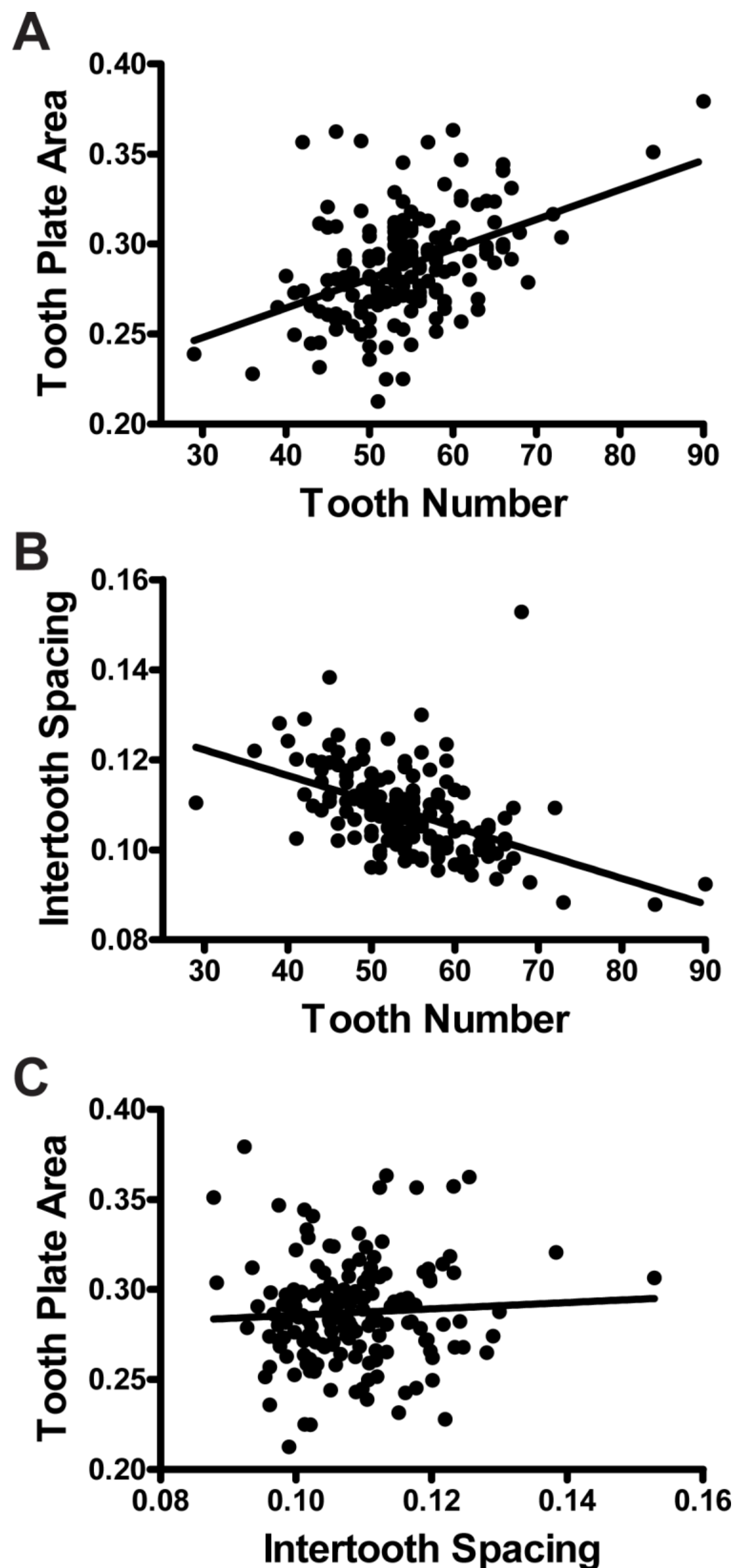

Figure S6: Correlation of tooth number, tooth plate area, and intertooth spacing phenotypes in F2 cross.
(A-C) Pairwise correlations of three ventral pharyngeal tooth patterning traits from CERC x marine F2 cross. (A) Tooth plate area and tooth number are positively correlated (P<0.001, r^2=0.20). (B) Intertooth spacing and tooth number are negatively correlated (P<0.001, r^2=0.23). (C) Tooth plate area and intertooth spacing are not correlated (P=0.47, r^2=0.003). (linear regression).

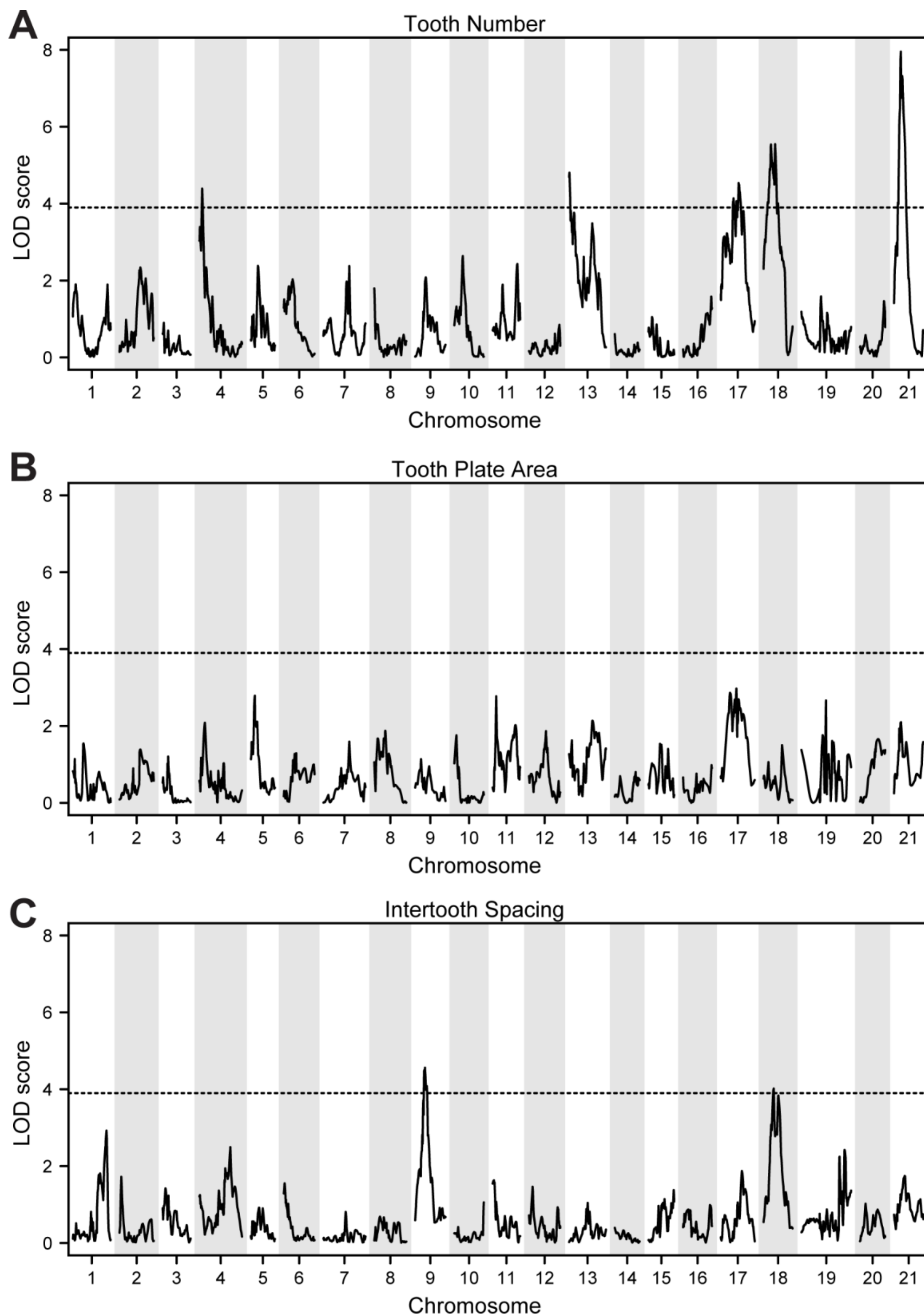

Figure S7: Genome wide QTL scans for tooth number, area, and intertooth spacing.
(A-C) Manhattan plots for tooth number (A), tooth plate area (B), and intertooth spacing (C). The y-axis is the logarithm of the odds (LOD) score of the association between genotype and phenotype. The dotted line is the genome wide significance threshold of alpha=0.05 determined by permutation tests.

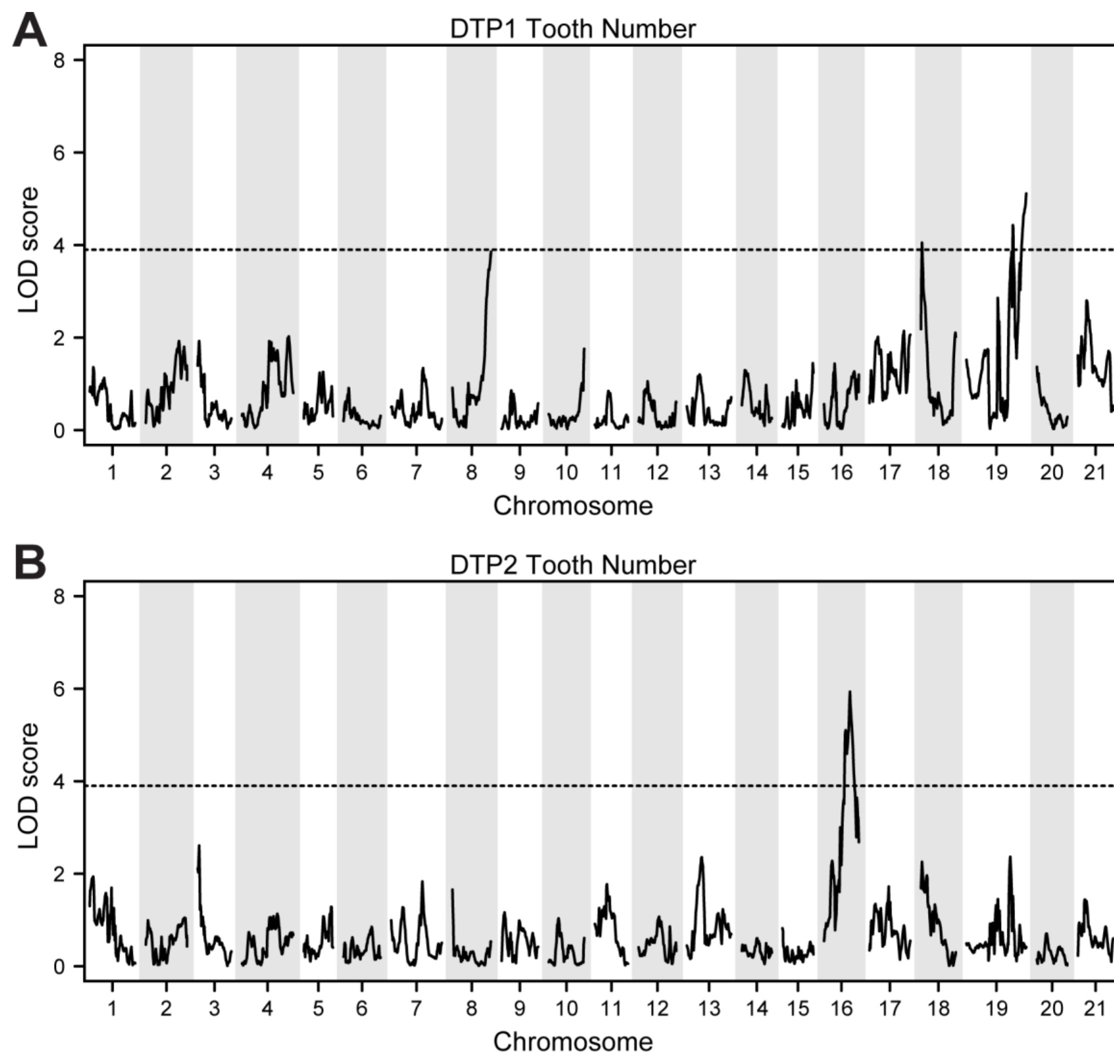

Figure S8: Genome wide QTL scans for dorsal pharyngeal tooth number.
(A-B) Manhattan plots for dorsal pharyngeal tooth number on DTP1 (A) and DTP2 (B). The y-axis is the logarithm of the odds (LOD) score of the association between genotype and phenotype. The dotted line is the genome wide significance threshold of alpha=0.05 determined by permutation tests.

Supplementary Table 1. Spatial location of teeth by population

Population	Total Teeth	Teeth off VTP	Teeth on VTP	New teeth off VTP	New teeth on VTP
RABS	59.4 (± 1.5)	5.6 (± 0.6)	53.7 (± 1.6)	1.2 (± 0.3)	2.4 (± 0.5)
PAXB	98.9 (± 2.6)	12.6(± 0.9)	86.3 (± 2.2)	7.5 (± 0.7)	10.6 (± 1.2)
CERC	79.3 (± 2.2)	0.9 (± 0.3)	78.4 (± 2.1)	0.5 (± 0.2)	10.1 (± 1.0)

Mean values are given for each trait ± standard error (VTP= Ventral Tooth Plate). All pairwise comparisons between populations for total teeth, teeth off VTP, and teeth on VTP are significantly different ($P < 0.001$, with the exception of PAXB vs. CERC teeth on VTP being $P < 0.05$). New teeth off and on VTP are plotted in Figure 5B,C (Sample size: n=14 for each population).

Supplementary Table 2. Summary of CERC tooth patterning QTL

Trait	Chr	Peak position (cM)	LOD score	PVE	1.5 LOD interval (cM)	1.5 LOD interval (Mb)	Mean Phenotype ± Standard Error		
							MM	MF	FF
VTP Tooth Number	4	7.7	4.2	6.8	0-15.4	0-3.1	50.7 (± 1.4)	55.1 (± 0.8)	54.8 (± 1.4)
VTP Tooth Number	13	0	4.8	7.8	0-8	1.0-1.9	52.5 (± 1.4)	53.5 (± 0.8)	58.4 (± 1.4)
VTP Tooth Number	17	51	4.5	7.3	32.8-71.7	3.7-11.7	52.2 (± 1.2)	53.6 (± 0.8)	58.9 (± 1.4)
VTP Tooth Number	18	32.7	5.5	9.0	13.8-38.2	2.9-11.2	57.6 (± 1.2)	54.7 (± 0.8)	50.4 (± 1.1)
VTP Tooth Number	21	19	8.0	13.5	14.6-31.3	3.7-9.0	49.3 (± 1.2)	55.4 (± 0.8)	56.8 (± 1.2)
VTP Intertooth Spacing	9	27.7	4.7	11.3	21.2-35.9	6.7-16.4	0.105 (± 0.001)	0.108 (± 0.001)	0.114 (± 0.002)
VTP Intertooth Spacing	18	28.7	4.1	9.9	17.8-49.4	4.8-13.2	0.104 (± 0.002)	0.108 (± 0.001)	0.113 (± 0.001)
DTP1 Tooth Number	18	2.5	4.1	9.3	1.1-13.8	0-3.4	23.2 (± 0.5)	20.8 (± 0.4)	20.7 (± 0.5)
DTP1 Tooth Number	19	140.6	5.7	13.1	102.4-140.6	2.4-17.8	19.7 (± 0.6)	21.9 (± 0.4)	25.3 (± 1.8)
DTP2 Tooth Number	16	61	5.7	14.2	47.1-72.0	13.2-17.6	59.5 (± 1.3)	64.3 (± 0.9)	69.2 (±1.3)

Genotypic classes of F2 fish are abbreviated: MM = homozygous marine, MF = heterozygous, FF = homozygous freshwater. LOD is the logarithm of the odds and PVE is the percentage of phenotypic variance explained (cM=centiMorgans, Mb=megabases, VTP= Ventral Tooth Plate, DTP1= Dorsal Tooth Plate 1, DTP2=Dorsal Tooth Plate 2). Intertooth spacing is measured in mm.

Supplementary Table 3. CERC QTL interval details

Trait	Chr	Peak position (cM)	Peak Marker	Peak LOD score	Left 1.5 LOD			Right 1.5 LOD			FDR *P*-value
					Marker	Position (cM)	Score	Marker	Position (cM)	Score	
VTP Tooth Number	4	7.7	20_3	4.2	65_1	0	3.0	20_4	15.4	1.6	0.026
VTP Tooth Number	13	0	51_3	4.8	51_3	0	4.7	52_1	8.0	3.3	0.0076
VTP Tooth Number	17	51	18_7	4.5	25_7	32.8	3.0	18_11	71.7	1.9	0.014
VTP Tooth Number	18	32.7	21_12	5.5	29_7	13.8	4.0	21_13	38.2	3.7	0.0024
VTP Tooth Number	21	19	16_11	8.0	16_8	14.6	5.9	16_18	31.3	5.4	0.0001
VTP Intertooth Spacing	9	27.7	8_9	4.7	8_25	21.2	2.6	8_6	35.9	2.8	0.010
VTP Intertooth Spacing	18	28.7	21_10	4.1	21_1	17.8	2.1	32_2	49.4	2.3	0.035
DTP1 Tooth Number	18	2.5	29_3	4.1	29_1	1.1	2.6	29_7	13.8	1.9	0.036
DTP1 Tooth Number	19	140.6	3_28	5.7	34_5	102.4	3.2	3_28	140.6	5.1	0.0021
DTP2 Tooth Number	16	61	44_4	5.7	14_17	47.1	3.5	44_3	72.0	3.8	0.0017

LOD is the logarithm of the odds and FDR is the false discovery rate (cM=centiMorgans, VTP= Ventral Tooth Plate, DTP1= Dorsal Tooth Plate 1, DTP2= Dorsal Tooth Plate 2).

Supplementary Table 4.

Click here to Download Table S4