Supplemental Materials for:

Genetic dissection of a supergene implicates *Tfap2a* in craniofacial evolution of threespine sticklebacks

Priscilla A. Erickson, Jiyeon Baek, James C. Hart, Phillip A. Cleves, and Craig T. Miller[1]

Department of Molecular and Cell Biology; University of California, Berkeley; Berkeley, CA, 94720

Figures S1-S8
Tables S1-S3
Data available in File S1

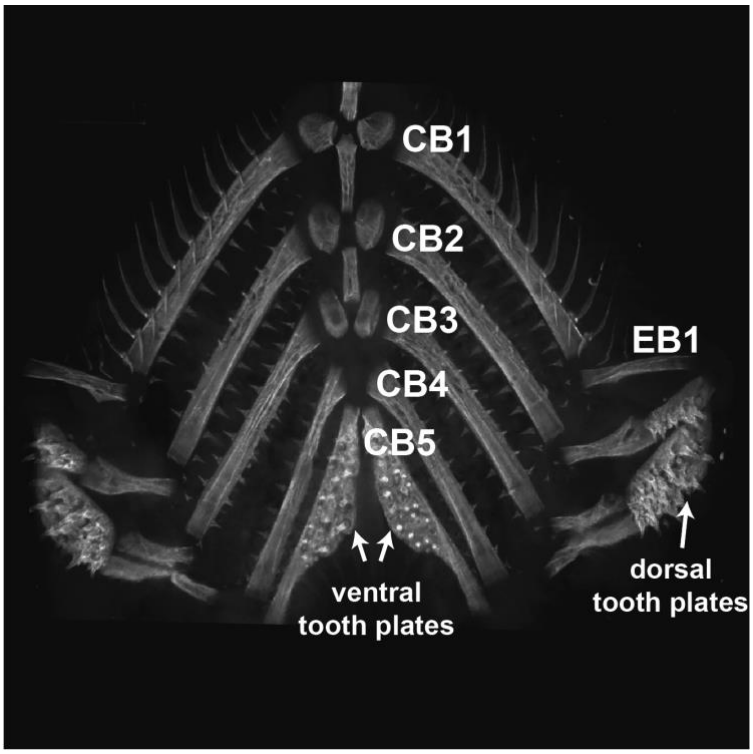

Figure S1: Stickleback branchial skeleton anatomy. Fluorescent image of flattened adult branchial skeleton stained with Alizarin red, with names of long endochondral bones and tooth plates. See Erickson *et al.*, 2014 and Cleves *et al.*, 2014 for pictures and quantification of evolved differences in bone length and tooth number, respectively. Anterior is oriented up. CB = ceratobranchial, EB = epibranchial.

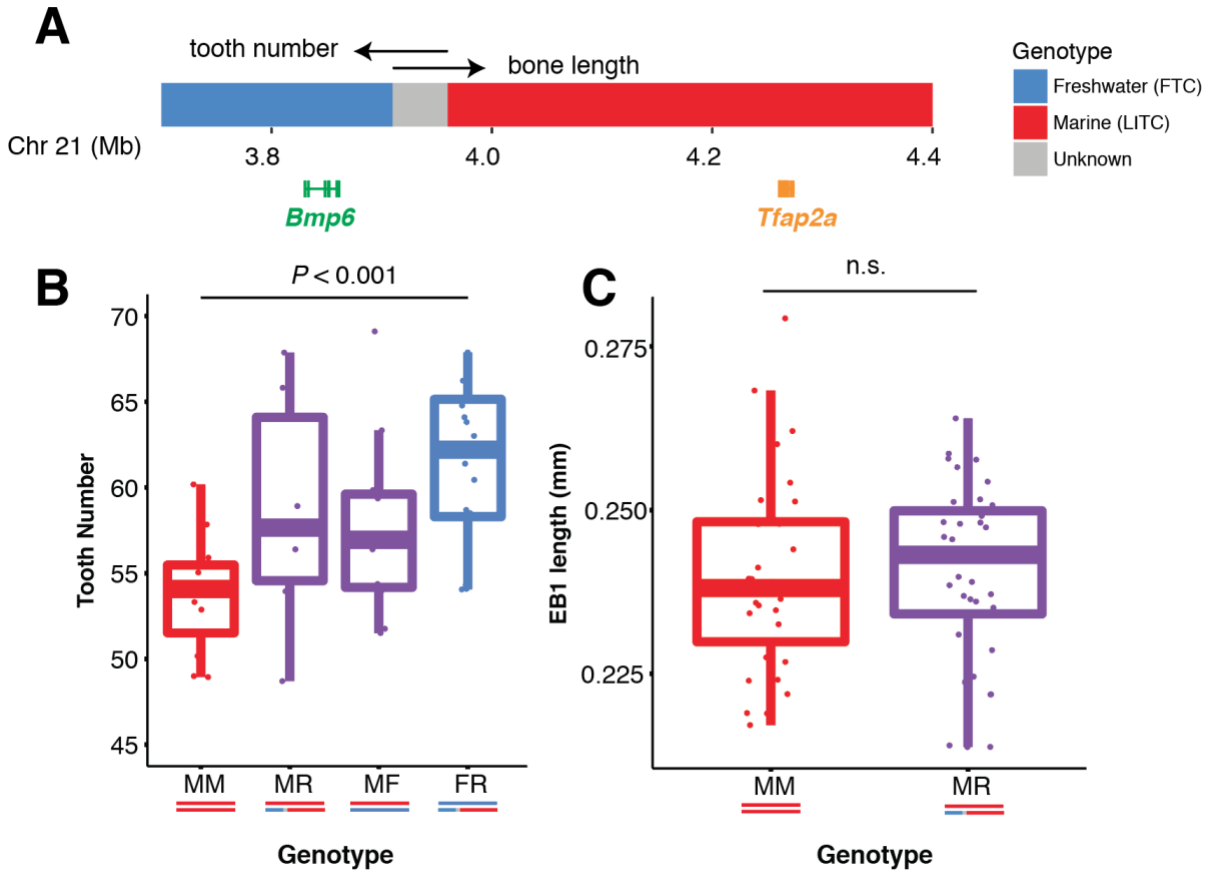

Figure S2: A FTC x LITC recombinant chromosome controls tooth number but not bone length. (A) Depiction of a region of chromosome 21 (3.7-4.4 Mb) indicating recombinant breakpoints, *Bmp6*, and *Tfap2a* (additional genes are not shown). This chromosome is RF_2 in Figure 1 & Table 1. (B) This chromosome affects pharyngeal tooth number in a small cross of ~120 dpf juveniles (ANOVA, df = 43, F = 5.785, P = 0.002), consistent with previous mapping of a pharyngeal tooth QTL to *Bmp6* (Cleves *et al*. 2014; 2018 in revision). (C) The recombinant chromosome does not control EB1 length at 28 dpf (T test, df = 64.289, t = -0.2287, P = 0.82). Bone lengths from Table 1 were back-transformed to a 10.5 mm individual. See Table 1 for additional data on this chromosome.

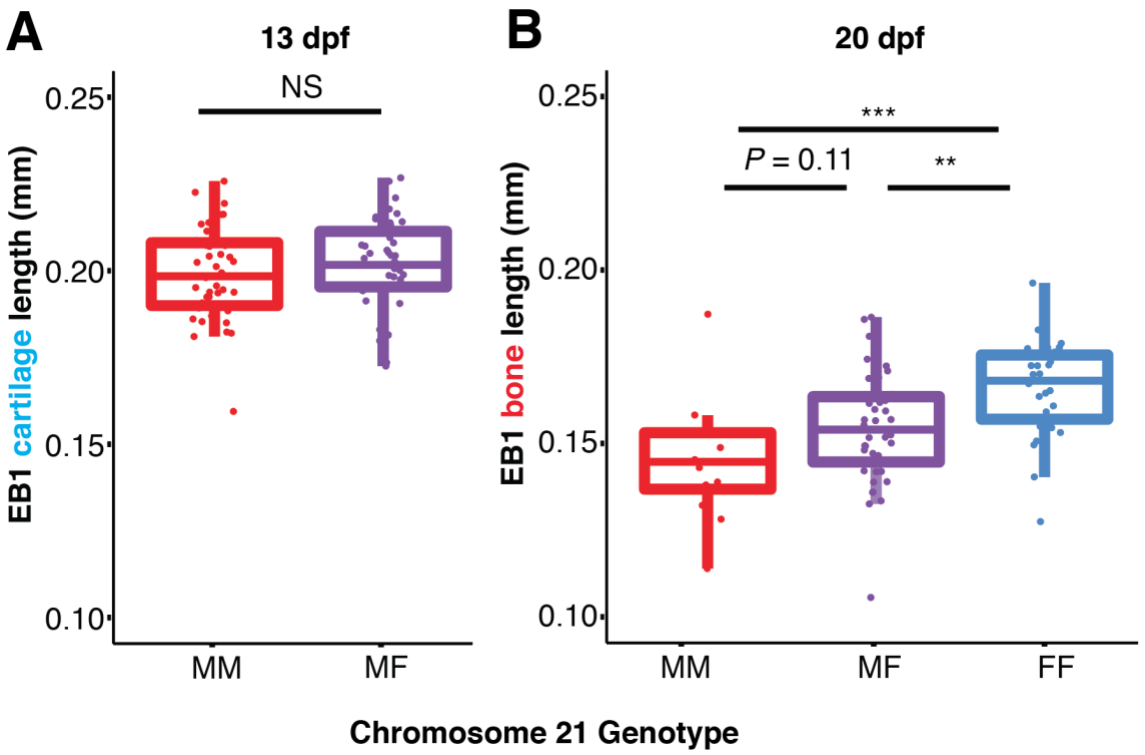

Figure S3: The chromosome 21 QTL has effects soon after initial EB1 bone formation. (A) At 13 dpf, chromosome 21 does not control the size of the EB1 cartilage template in F5 fish from the FTC x LITC cross (T test, df = 89.7, t = 1.09, P = 0.28). (B) By 20 dpf, a stage soon after the initial ossification of EB1 (Erickson *et al.* 2014), chromosome 21 controls bone length in related F5 fish (ANOVA, df = 86, F = 10.73 $P = 6.9 \times 10^{-5}$). Fish were genotyped for Stn423, and EB measurements were back transformed to fish of total length = 8.5 mm for (a) and standard length = 9 mm for (b).

```
Marine_stickleback          1 ----MKMLWKLTDNIKYEDCEDRHDGTSNGTARLPQLGGVGQSPYTSAPPLSHTPN
Freshwater_stickleback      1 ----MKMLWKLTDNIKYEDCEDRHDGTSNGTARLPQLGGVGQSPYTSAPPLSHTPN
Salmo_NP_001158795          1 ------------MLVHSFSAMDRHDGTSNGTARLPQLGGVGQSPYS-APPLSHTPN
Danio_NP_001306087          1 MYHIQKEETRMSLMGKMGDWQDRHDGTSNGTARLPQLGSVGQSPYTSAPPLSHTPN
Xenopus_NP_001089958        1 ------MLVH---AYSAMDRSEVLNG-AASGGRLSQLSSLNQGPYSSAPPLGHTPA
Mus_NP_035677               1 ------MLWKLTDNIKYEDCEDRHDGTSNGTARLPQLGTVGQSPYTSAPPLSHTPN
Homo_NP_001027451           1 ------------MLVHSFSAMDRHDGTSNGTARLPQLGTVGQSPYTSAPPLSHTPN

Marine_stickleback         57 PPYFPPPYQPI---YPQ-SQDPYSHVNDPYSL-NSLHAQPQPQHPGWPGQRQSQE-
Freshwater_stickleback     57 PPYFPPPYQPI---YPQ-SQDPYSHVNDPYSL-NSLHAQPQPQHPGWPGQRQSQE-
Salmo_NP_001158795         48 PPYFPPPYQPI---YSQ-SQDPYSHVNDPYSL-NSLHAQPQPQHPGWPGQRQSQE-
Danio_NP_001306087         61 PPYFPPPYQPI---YPQ-SQDPYSHVNDPYSI-NSLHAQSQPQHPGWPGQRQSQE-
Xenopus_NP_001089958       51 PPYFPPPYPQPPLSYSQSQESGYPHLGDPYSSINSIHH--QHQQPSWHTPRSRPEE
Mus_NP_035677              55 PPYFPPPYQPI---YPQ-SQDPYSHVNDPYSL-NPLHAQPQPQHPGWPGQRQSQE-
Homo_NP_001027451          49 PPYFPPPYQPI---YPQ-SQDPYSHVNDPYSL-NPLHAQPQPQHPGWPGQRQSQE-

Marine_stickleback        111 -HQHRGLPHQLCR-----EYRR-EVLLPSGHGIDTGLSD-SISLHGIPH-SLDDVQ
Freshwater_stickleback    111 -HQHRGLPHQLCR-----EYRR-EVLLPSGHGIDTGLSD-SISLHGIPH-SLDDVQ
Salmo_NP_001158795        102 -HQHRGLPHQLCR-----EYRR-EVLLPSGHGLDTGLSD-SIPIHGIHH-SLEDVQ
Danio_NP_001306087        115 -HQHRGLPHQLCR-----EYRR-EVLLPSGHGIDTGLTD-SIPIHGIPH-SLEDVQ
Xenopus_NP_001089958      109 SQTHRG-----LSLDPRRDYGGMSRLIPGLTDGGHSLADSSLSIHALSHHSLEDMQ
Mus_NP_035677             109 -HTHRGLPHQLSGLDPRRDYRRHEDLLHGPHGLGSGLG--DLPIHSLPH-AIEDVP
Homo_NP_001027451         103 -HTHRGLPHQLSGLDPRRDYRRHEDLLHGPHALSSGLG--DLSIHSLPH-AIEEVP

Marine_stickleback        159 ---VDDQGIHIPDQTVIKKGPVSLSKNN-N-ISSIPVNKDGLFGGVVNPNEVFCSV
Freshwater_stickleback    159 ---VDDQGIHIPDQTVIKKGPVSLSKNN-N-ISSIPVNKDGLFGGVVNPNEVFCSV
Salmo_NP_001158795        153 YLHIEDQGIHIPDQTVIKKGPVSLSKNN-N-VSTIGINKDGLFGGVVNPNEVFCSV
Danio_NP_001306087        163 ---VEDQGIHIPDQTVIKKGPVSISKNNSN-ISAIPINKDGLFGGVVNPNEVFCSV
Xenopus_NP_001089958      160 --LLDESGISILDQSVIKKVPISS-KNNSSMMSALSMNKESLIGGVSNPNEVFCSV
Mus_NP_035677             161 --HVEDPGINIPDQTVIKKGPVSLSKSNSNAVSAIPINKDNLFGGVVNPNEVFCSV
Homo_NP_001027451         155 --HVEDPGINIPDQTVIKKGPVSLSKSNSNAVSAIPINKDNLFGGVVNPNEVFCSV

                                                        *
Marine_stickleback        214 SLLSSTSKYKVTVAEVQRRLSPPECLNASLLGGVLRRAKSKNGGRSLREKLDKIGL
Freshwater_stickleback    214 SLLSSTSKYKVTVAEVQRRLSPPECLNASLLGGVLRRAKSKNGGRSLREKLDKIGL
Salmo_NP_001158795        211 SLLSSTSKYKVTVAEVQRRLSPPECLNASLLGGVLRRAKSKNGGRSLREKLDKIGL
Danio_NP_001306087        219 SLLSSTSKYKVTVAEVQRRLSPPECLNASLLGGVLRRAKSKNGGRSLREKLDKIGL
Xenopus_NP_001089958      217 SLLSSTSKYKVTVGEVQRRLSPPECLNASLLGGVLRRAKSKNGGRCLRERLEKIGL
Mus_NP_035677             219 SLLSSTSKYKVTVAEVQRRLSPPECLNASLLGGVLRRAKSKNGGRSLREKLDKIGL
Homo_NP_001027451         213 SLLSSTSKYKVTVAEVQRRLSPPECLNASLLGGVLRRAKSKNGGRSLREKLDKIGL

Marine_stickleback        274 GRRKAANVTLLTSLVEAGEAVHLARDFGYVCETEFPAKAVAEYVNRQHSDPNEQVQ
Freshwater_stickleback    274 GRRKAANVTLLTSLVEAGEAVHLARDFGYVCETEFPAKAVAEYVNRQHSDPNEQVQ
Salmo_NP_001158795        271 GRRKAANVTLLTSLVE-GEAVHLARDFGYVCETEFPAKAVAEYTNRQHSDPNEHVQ
Danio_NP_001306087        279 GRRKAANVTLLTSLVE-GEAVHLARDFGYVCETEFPAKAIAEYMNRQHSDPNEQVQ
Xenopus_NP_001089958      277 GRRKAANVTLLTSLVE-GEAVHLARDFGYVCETEFPAKAAAEYLCRQHSDPTELHA
Mus_NP_035677             279 GRRKAANVTLLTSLVE-GEAVHLARDFGYVCETEFPAKAVAEFLNRQHSDPNEQVA
Homo_NP_001027451         273 GRRKAANVTLLTSLVE-GEAVHLARDFGYVCETEFPAKAVAEFLNRQHSDPNEQVT

Marine_stickleback        334 LLATKQVCKEFTDLLSQDRSPLGNSRPQPILEPGIQSCLTHFSLISHGFGTPALCA
Freshwater_stickleback    334 LLATKQVCKEFTDLLSQDRSPLGNSRPQPILEPGIQSCLTHFSLISHGFGTPALCA
Salmo_NP_001158795        330 LLATKQICKEFTDLLSQDRTPLGNSRPQPILEPGIQSCLTHFSLISHGFGTPAMCA
Danio_NP_001306087        338 LLATKQICKEFTDLLSQDRSPLGNSRPQPILEPGIQSCLTHFSLISHGFGTPAVCA
Xenopus_NP_001089958      336 LLATKQICKEFADLLAQDRSPLGNSRPSLILEPGVQSCLTHFSLITHGFGGPAICA
Mus_NP_035677             338 LLATKQICKEFTDLLAQDRSPLGNSRPNPILEPGIQSCLTHFNLISHGFGSPAVCA
Homo_NP_001027451         332 LLATKQICKEFTDLLAQDRSPLGNSRPNPILEPGIQSCLTHFNLISHGFGSPAVCA

Marine_stickleback        394 LQNYLTEAIKAMDKMYLNNNPNSHSDNGTKGGDCDQKKKK
Freshwater_stickleback    394 LQNYLTEAIKAMDKMYLNNNPNSHSDNGTKGGDCDQKKKK
Salmo_NP_001158795        390 LQNYLTEAIKAMDKMYLNNN--SHSDSGTKGGDKDEKHRK
Danio_NP_001306087        398 LQNYLTEAIKAMDKMYLNNNPNSHSETGSKAGDKDEKHRK
Xenopus_NP_001089958      396 FQNYLLESLKGMDKIFMSSTGNGHSAAESKS-EKDIKHRK
Mus_NP_035677             398 LQNYLTEALKAMDKMYLSNNPNSHTDNSAKSSDKEEKHRK
Homo_NP_001027451         392 LQNYLTEALKAMDKMYLSNNPNSHTDNNAKSSDKEEKHRK
```

Figure S4: Amino acid alignment of marine and freshwater stickleback TFAP2A with other vertebrate TFAP2A sequences. Stickleback predicted protein sequence is based on the predicted Ensembl transcript ENSGACT00000003048.1. The RefSeq accession number is given next to each non-stickleback genus name. The red asterisk indicates the location of the synonymous polymorphism used for ASE, and red bars indicate TALEN binding sites.

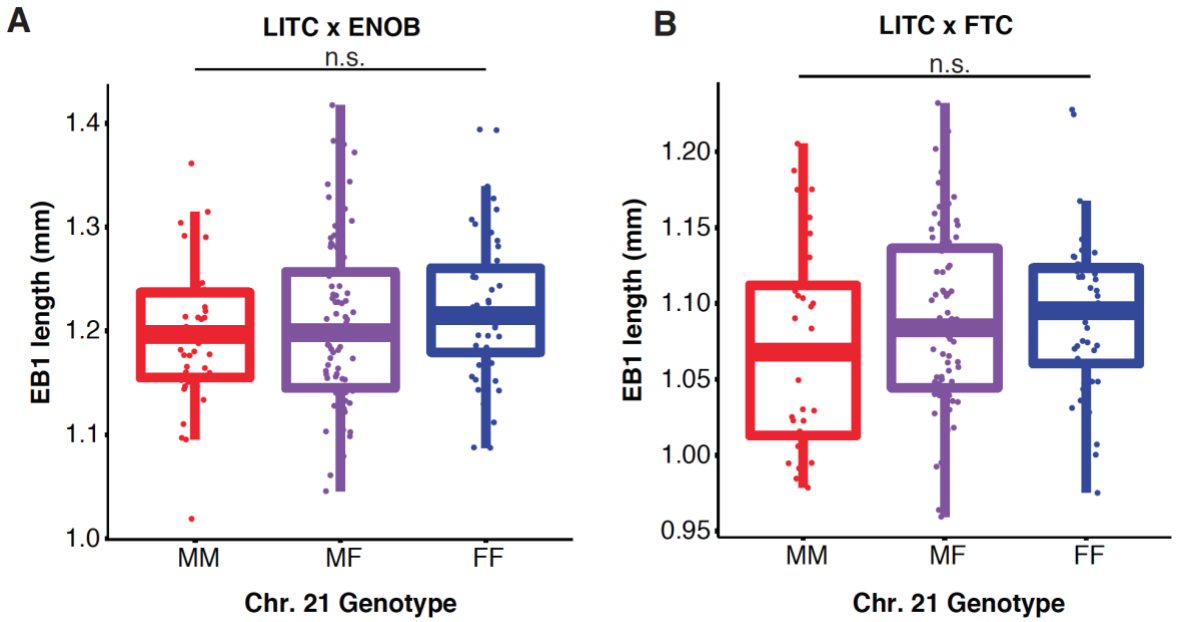

Figure S5: Two crosses lack the chromosome 21 bone length QTL near *Tfap2a*. (A) A LITC x Enos Benthic (ENOB) cross was reported to have a suggestive bone length QTL near the end of chromosome 21 (Erickson *et al.* 2016). However, the marker containing *Tfap2a* does not have a significant effect on EB1 bone length in this cross (ANOVA; n=176, F = 1.597, *P* = 0.2). (B) Epibranchial 1 length was measured in 210 fish from a LITC x FTC cross different from the one used in this study (Glazer *et al.* 2015). No significant relationship between genotype and phenotype was observed (ANOVA, n=158, F = 1.135, *P* = 0.3). The GBS binned marker containing *Tfap2a* (marker 16_9, chrXXI:4170348-4654382) was used to test for the presence of a QTL at *Tfap2a* in both crosses.

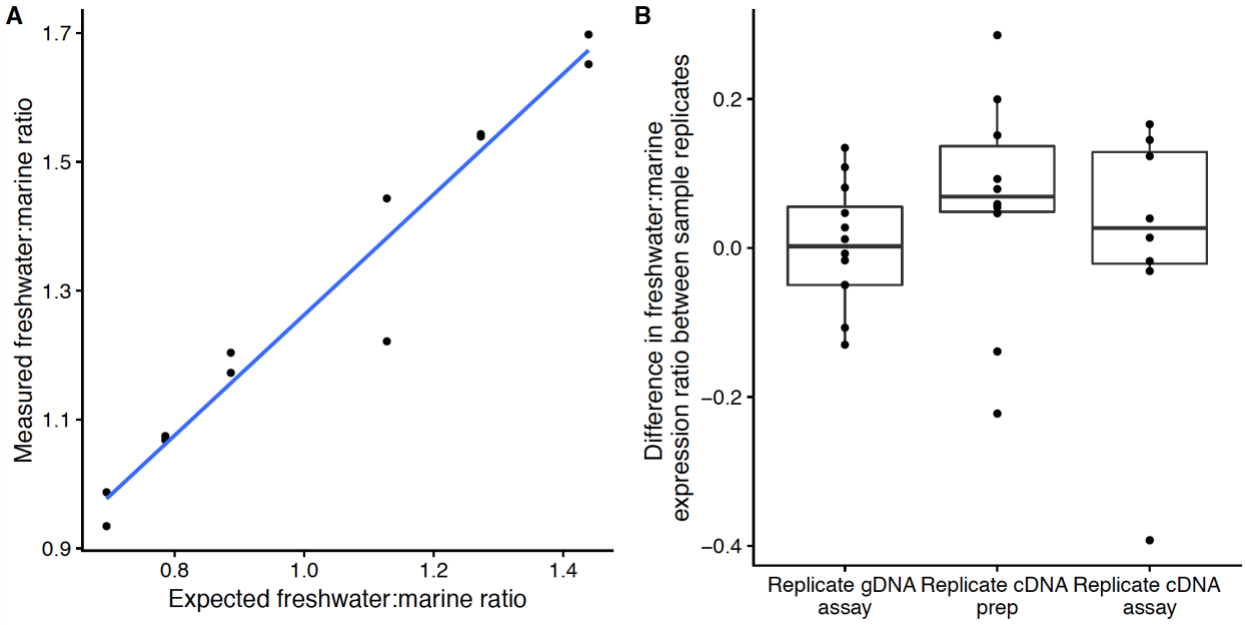

Figure S6: Allele-specific expression assay controls. A) Genomic DNA samples from a homozygous freshwater individual and a homozygous marine individual were each diluted to 50 ng/uL. The two DNA samples were mixed at fixed ratios of 4.1:5.9, 4.4:5.6, 4.7:5:3, 5.3:4.7, 5.6:4.4 and 5.9:4.1, and the *Tfap2a* allelic ratio was measured. The measured freshwater:marine allelic ratio is plotted with respect to the expected ratio. A range of allelic ratios is robustly detected by the assay. The measured ratios are always higher than expected in this titration, likely due to an initial difference in the starting DNA concentrations of marine and freshwater samples, since genomic DNA from known heterozygous individuals gives a robust 1:1 ratio (see Figure 4). B) Replicate controls: for each control, the difference in allelic ratio between two replicates from the same biological sample is plotted. Left: The same genomic DNA sample was measured in replicate assays with different barcodes. Middle: cDNA was prepared twice from the same RNA sample and assayed separately. Right: the same cDNA sample was assayed with two different barcodes. Although there is variability in the assay, the difference in allelic ratio of replicate samples is not significantly different from zero for any of the three controls (Wilcoxon signed rank test, $P > 0.23$ for all)

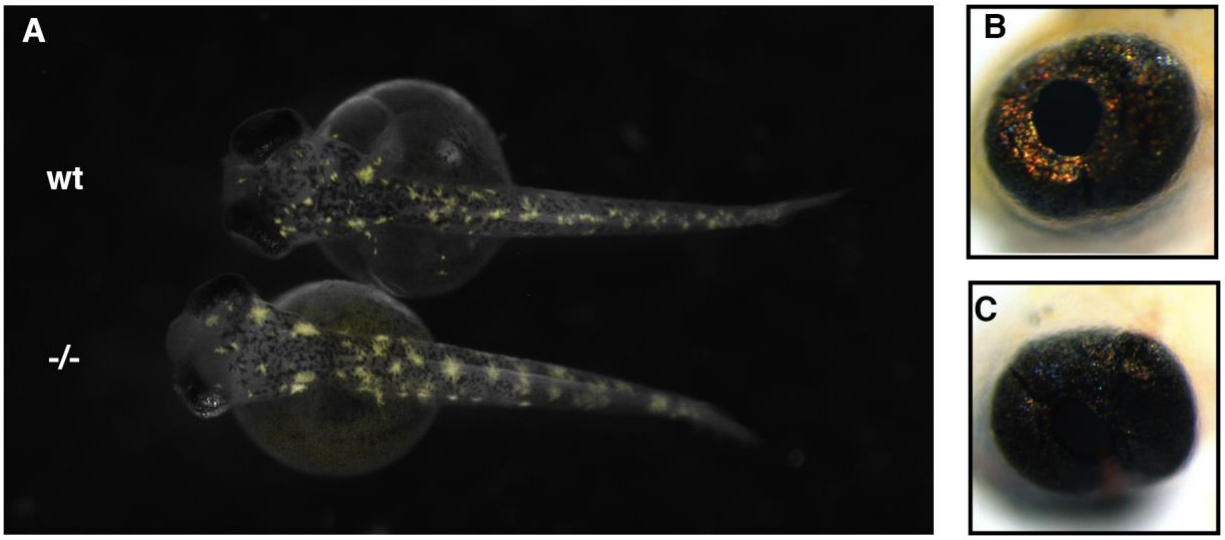

Figure S7. Iridophore but not xanthophore defects in stickleback *Tfap2a* mutants. (A) Xanthophores appear unaffected in FTC *Tfap2a* homozygous mutants relative to wild-type. (B, C) Iridophores were reduced but present in the eyes of mutants (C) relative to wild-type (B). Xanthophores were photographed using a GFP filter and iridophores were imaged under brightfield light. All fish were imaged at 9 dpf.

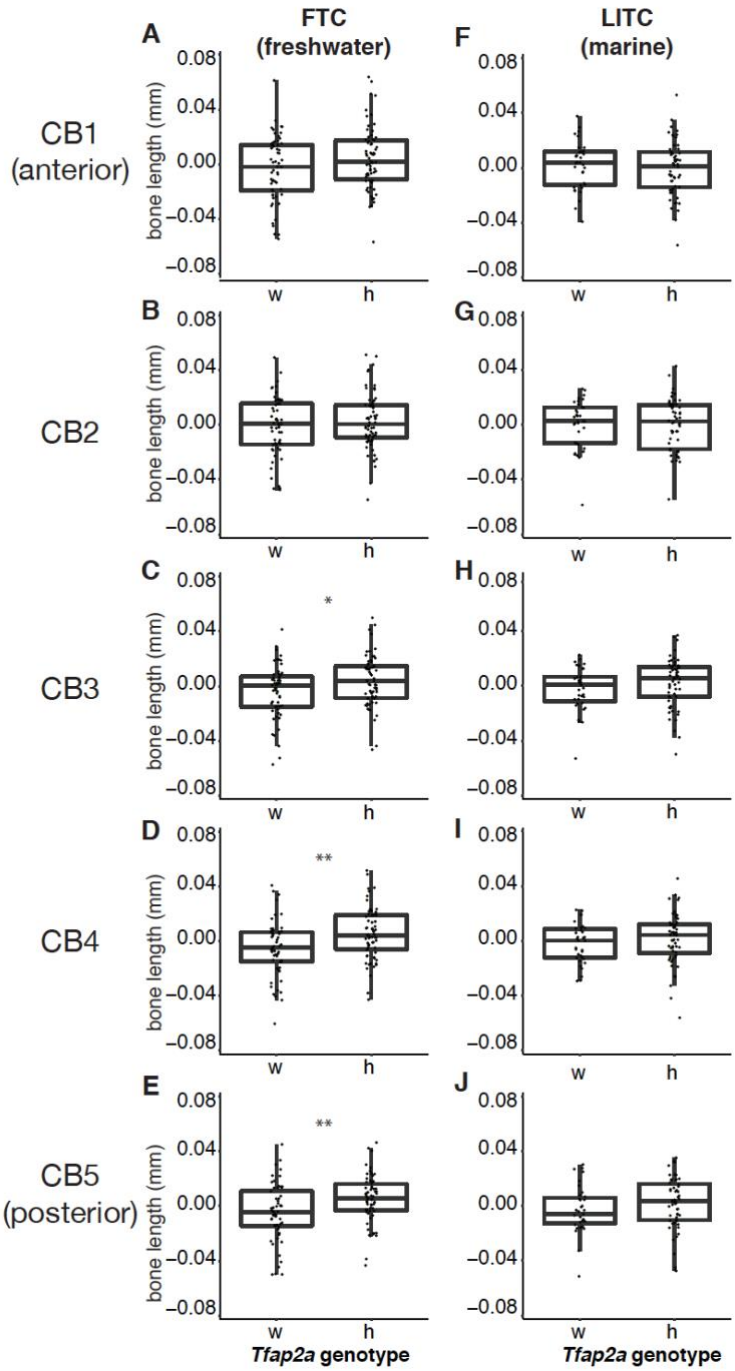

Figure S8. Heterozygous *Tfap2a* mutation affects posterior ventral bone length in freshwater but not marine fish. (A-E) FTC freshwater fish heterozygous for *Tfap2a* mutations have significantly longer ventral posterior bones [CB3-CB5, t-test $P < 0.05$ (*) and $P < 0.01$ (**)]. (F-J) No significant effect of *Tfap2a* genotype on bone length was seen in any ventral bone in the LITC marine background. Bone length was size-corrected for standard length using a linear model; residuals are shown. Standard length did not differ between genotypes in freshwater or marine fish. w = wild-type; h = heterozygous. See Figure S1 for branchial skeleton anatomy.

Table S1: Primers used in this study

PRIMER NAME	CHR 21 POSITION (MB)	FORWARD PRIMER (5'→3')	REVERSE PRIMER (5'→3')	PURPOSE
PAE450/451	2.75	GGGGTAGTTCGGATGTTTCA	CCCTGTTCTGCAAGGTGAAT	indel marker, F
PAE444/445	3.02	AAAGGCAGTGTCGATGTTGG	TTGGGAAGATGTGTGGACAG	indel marker, P+F
PAE309/310	3.9	CAGGGTGATTGTCAGCCATT	CCCGATTTATTAAAATGCCAAA	indel marker, P+F
PAE496/497	3.963	ACAATATTTGACAGGTTATCTAGTGTCC	GTGAGAGTTGTGAGGACACAAAACAA	indel marker, F
PAE311/312	4.15	AACGTCCAAAGCAATGTTGG	CGGACTAGCAGAGCAATGTG	indel marker, P+F
PAE428/429	4.2	AAAAACAGATGTGCCCCTTTT	TACAGGCCTCTCAGGCAAAT	indel marker, P
PAE540/541	4.21	AAAACTGGTGGCGCGTGTTT	CAACTGCCTTTCGTCAGTCTGC	indel marker, F
PAE542/543	4.231	ACCGACGTGCACAACACAA	ACCTGCCACCAGCAGAGTCT	indel marker, F
PAE436/437	4.32	TGATGCAAGAGAGCTGGAAA	GCTGGGTTGGAAACGATAAA	indel marker, P
PAE438/439	4.35	ACGTGTTTTGGTCCGATACC	CCTCCAATTGGCCCTTAATA	indel marker, P
PAE317/318	4.37	GGCCGTACTTCCACATTATGA	ACTCCATGGCAACCAGATCC	indel marker P+F
PAE323/324	4.42	AGAGGGAGCCTTGTTGCAT	GCTCTGGTCTGTCAGCGTAA	indel marker, P
PAE325/326	4.44	ATGCTTCCCGATAAATCGATAC	TGCCCAAAATAGGAAATGACA	indel marker, P
PAE511/512	4.59	GTGCCAGAAATAGCGTCGGC	GGGTATTACGTACCGTGATGATGATT	indel marker, P
PAE513/514	4.72	CATAGCTGCCGATTAATCCCCCT	TGAGCGTGCACTACTGGGGA	indel marker, P
PAE462/463	4.98	TGCAGGTACCAAAATAGCAAGA	GCTCCGGGACTAGTTGCAC	indel marker, F
PAE464/465	5.03	GATTTGGCTCTTGGGGTGTA	AGGCAAATGAAGCACAGAGG	indel marker, F
PAE363/364	5.6	TGATCACATGGTTCTTGTCCA	AAACAGGGTTTGGGACTGG	indel marker, P+F
PAE349/350	5.96	CCTGCAAGATGTCCCCTAAA	CCTCCCCTAGTTGCATTGTG	indel marker, P+F
PAE414/416	4.26	gccggcagc AACAACAACATCTCCTCCATCCCC	gccggcagc CCCAGCAGGGAGGCGTTGAG	*Tfap2a* ASE
PAE379/381	4.26	GGGTCGTTGACGTGCGAGTAA	AGCGGGACAACGTCATCACTTA	TALEN genotyping

For genotyping, primers were designed to amplify insertion/deletions (indels) that differed between the grandparents of the PAXB x LITC and FTC x LITC crosses (indicated as P and F, respectively, in the purpose column). Lowercase letter indicate the *ApeK*I restriction enzyme tag for the ASE assay.

Table S2: TALEN RVDs

	TALEN 1			TALEN 2		
POSITION	***pFUS_A 5'***	***pFUS_B 5'***	***pLR 5'***	***pFUS_A 3'***	***pFUS_B 3'***	***pLR 3'***
1	HD	HD	HD	NN	NN	HD
2	NN	NG		NN	NN	
3	NG	HD		NN	HD	
4	NN	NG		NI	NG	
5	NG	NN		HD	NI	
6	NI	NN		NI	HD	
7	NN			NN	HD	
8	NN			HD	NG	
9	NN			HD		
10	NI			NI		

See Cermak *et al*. 2011 for details.

Table S3: *Tfap2a* mutations studied

DESCRIPTION	SEQUENCE	PRODUCT
wild-type	ACAGCCAGGCTACCTCAGCTGGGCGGCGTGGGCC	427 AA
6 bp deletion/22 bp insertion (*)	ACAGCCAGGCTACCTCAGC******GCGTGGGCC	early stop after 158 AA
13 bp deletion	ACAGCC---C-A-CTCA--------GCGTGGGCC	early stop after 142 AA
22 bp deletion	ACAGC----------GTGGGCC	early stop after 139 AA
8 bp deletion	ACAGCCAGGCTACCTCAGC--------GTGGGCC	early stop after 150 AA
10 bp deletion	ACAGCCAGGCTACCTCAG----------TGGGCC	early stop after 143 AA

Sequenced *Tfap2a* mutations induced by TALENs. ** indicates 22 bp inserted sequence: ACCTCACCTCAGCTCCCTACAC. Product indicates predicted protein product length, which is frame-shifted and results in an early stop for all mutations (AA = amino acids).