Title: A key metabolic gene for recurrent freshwater colonization and radiation in fishes

Authors: Asano Ishikawa[1], Naoki Kabeya[2, 3], Koki Ikeya[4], Ryo Kakioka[1], Jennifer N. Cech[5], Naoki Osada[6], Miguel C. Leal[7, 8], Jun Inoue [9], Manabu Kume[1, 10], Atsushi Toyoda[11], Ayumi Tezuka[12], Atsushi J. Nagano[12], Yo Y. Yamasaki[1], Yuto Suzuki[13], Tomoyuki Kokita[13], Hiroshi Takahashi[14], Kay Lucek[7, 15, 16], David Marques[7, 15], Yusuke Takehana[17, 18], Kiyoshi Naruse[17], Seiichi Mori[19], Oscar Monroig[20], Nemiah Ladd[21], Carsten J. Schubert[21], Blake Matthews[22], Catherine L. Peichel[5, 15], Ole Seehausen[7, 15], Goro Yoshizaki[2], and Jun Kitano[1*]

Affiliations:

[1]Ecological Genetics Laboratory, National Institute of Genetics, Shizuoka, Japan.

[2]Department of Marine Biosciences, Tokyo University of Marine Science and Technology, Tokyo, Japan.

[3]Present address: Department of Aquatic Bioscience, The University of Tokyo, Japan.

[4]Gifu World Freshwater Aquarium, Gifu, Japan.

[5]Divisions of Human Biology and Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, USA.

[6]Graduate School of Bioengineering and Bioinformatics, Hokkaido University, Sapporo, Japan.

[7]Department of Fish Ecology and Evolution, Eawag Swiss Federal Institute of Aquatic Science and Technology, Centre for Ecology, Evolution and Biogeochemistry, Kastanienbaum, Switzerland.

[8]Present address: MARE – Marine and Environmental Sciences Centre, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal.

[9]Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan.

[10]Present address: Field Science Education and Research Center, Kyoto University, Kyoto, Japan.

[11]Comparative Genomics Laboratory, National Institute of Genetics, Shizuoka, Japan.

[12]Faculty of Agriculture, Ryukoku University, Otsu, Shiga, Japan.

[13]Department of Marine Bioscience, Fukui Prefectural University, Obama, Fukui, Japan.

[14]Department of Applied Aquabiology, National Fisheries University, Shimonoski, Yamaguchi, Japan.

[15]Institute of Ecology and Evolution, University of Bern, Bern, Switzerland.

[16]Present address: Department of Plant Ecology, University of Basel, Basel, Switzerland.

[17]Laboratory of Bioresources, National Institute for Basic Biology, Okazaki, Aichi, Japan.

[18]Present address: Department of Animal Bioscience, Nagahama Institute of Bioscience and Technology, Nagahama, Shiga, Japan.

[19]Biological Laboratory, Gifu-keizai University, Ogaki, Gifu, Japan.

[20]Instituto de Acuicultura Torre de la Sal (IATS-CSIC), Ribera de Cabanes, Castellón, Spain.

[21]Department of Surface Waters – Research and Management, Eawag Swiss Federal Institute of Aquatic Science and Technology, Centre for Ecology, Evolution and Biogeochemistry, Kastanienbaum, Switzerland.

[22]Department of Aquatic Ecology, Eawag Swiss Federal Institute of Aquatic Science and Technology, Centre for Ecology, Evolution and Biogeochemistry, Kastanienbaum, Switzerland.

*Correspondence to: Jun Kitano, jkitano@nig.ac.jp

Abstract: Colonization of new ecological niches has triggered large adaptive radiations. Although some lineages have made use of such opportunities, not all do so. The factors causing this variation among lineages are largely unknown. Here, we show that deficiency in docosahexaenoic acid (DHA), an essential ω-3 fatty acid, can constrain freshwater colonization by marine fishes. Our genomic analyses revealed multiple, independent duplications of the fatty acid desaturase gene *Fads2* in stickleback lineages that subsequently colonized and radiated in freshwater but not in close relatives that failed to colonize. Transgenic manipulation of *Fads2* in a marine stickleback increased their ability to synthesize DHA and survive on DHA-deficient diets. Multiple freshwater ray-finned fishes also show a convergent increase in *Fads2* copies, indicating its key role in freshwater colonization.

One Sentence Summary: Duplication of an ω-3 fatty acid metabolism gene plays a key role in colonization of freshwater habitats in fishes.

Main Text:

Empty niches can provide organisms with ecological opportunities to diversify (*1-3*). Many of the known large adaptive radiations followed invasion of underutilized habitats (*1, 2*). However, not all lineages appear to take advantage of such opportunities. For example, habitat shifts from marine to freshwater have repeatedly triggered radiations, but only in a limited number of fish lineages (*4-6*). The physiological and genetic factors causing this variation are unknown.

One of the under-appreciated constraints for freshwater colonization by marine animals is the poor nutritional quality of food in freshwater ecosystems. Generally, the food chain in marine environments is rich in ω-3 long-chain polyunsaturated fatty acids, especially docosahexaenoic acid (DHA) (Fig. S1A) (*7*), which is essential for animal development and health (*8, 9*). However, freshwater ecosystems contain very little DHA (Fig. S1A) (*7*). Here, we test and confirm that DHA deficiency can constrain freshwater colonization by marine fishes and identify genetic changes that appear to have enabled some lineages to overcome this constraint.

Threespine stickleback (*Gasterosteus aculeatus* species complex) are primarily marine or anadromous (hereafter we call both marine), but when new freshwater habitats emerged after glacial retreat, they successfully colonized freshwater and radiated into diverse ecotypes on multiple continents (*10*) (Fig. 1, Fig. S2, and Table S1-S2). In contrast, the closely related Japan Sea stickleback (*G. nipponicus*), which diverged from *G. aculeatus* around 0.68-1.5 million years ago (MYA) (*11*), failed to colonize freshwater environments and remains phenotypically homogenous (Fig. 1). Although *G. nipponicus* co-occur with Pacific Ocean populations of *G. aculeatus* (Pacific Ocean stickleback) in some localities in Japan (*12-14*), have geographical access to many freshwater habitats, and can use freshwater for spawning (*14*), our phylogenomic analysis showed that all known Japanese freshwater populations belonged to *G. aculeatus* (Fig. 1C, and Fig. S3) (*15-17*).

Since stickleback prey differs in DHA levels between marine and freshwater habitats (Fig. S1B), we tested the hypothesis that Japan Sea stickleback may have a lower physiological ability than Pacific Ocean stickleback to survive on DHA-free diets (*18*). Our rearing experiments showed that irrespective of salinity, Japan Sea had higher mortality than Pacific Ocean stickleback starting around 40 days after fertilization when fed DHA-free *Artemia* ($P < 0.01$) (Fig. 2A, Fig. S4, Table S3); this age is close to the timing of seaward migration in nature (*19*). Marine-derived diets or *Artemia* enriched with several fatty acids, including DHA, significantly improved survival of Japan Sea stickleback ($P < 0.01$) (Fig. 2A-2B). Further, Japan Sea stickleback had a lower DHA content than Pacific Ocean stickleback when fed DHA-free *Artemia* ($P < 0.01$ for both brain and eye) (Fig. 2C, and Fig. S5), suggesting that they have either lower DHA biosynthetic capabilities or higher rates of DHA degradation or secretion.

Our whole genome resequencing revealed that *Fatty acid desaturase 2* (*Fads2*), a gene encoding a key enzyme catalyzing desaturation in DHA biosynthesis (Fig. S6, Table S4) (*20-22*), has a higher copy number in Pacific Ocean stickleback than in Japan Sea stickleback (Fig. 2D and Fig. S7) ($F_{1,12} = 79.8$, $P < 0.01$). Higher *Fads2* copy numbers in females than males ($F_{1,12} = 11.7$, $P < 0.01$) (Fig. 2D and Fig. S7) are due to the X-linkage of *Fads2* (see below). RNA sequencing further revealed that Pacific Ocean stickleback express *Fads2* at higher levels than Japan Sea stickleback ($F_{1,12} = 5.3$, $P < 0.05$ for brain; $F_{1,12} = 7.0$, $P < 0.05$ for eyes) when fed only DHA-free *Artemia* (Fig. 2E and Fig. S8).

To directly demonstrate the effects of *Fads2* copy number on survival, we made transgenic Japan Sea stickleback overexpressing *Fads2*. When fed only DHA-free *Artemia*, the *Fads2*-trangenics showed higher survival rates (Fig. 2F) and higher DHA content at 40 days after fertilization than the control *GFP*-transgenics ($P < 0.01$) (Fig. 2G and Fig. S9). Analysis of an F_2 intercross further showed that hybrids with higher *Fads2* copy number had higher survival rates at 40-60 days after fertilization ($P < 0.01$: 10.0-12.1% of variance explained) and longer overall lifespan ($P < 0.05$) (Fig. S10). Higher survival rate of females than males in Japan Sea stickleback is consistent with the higher *Fads2* copy number in females ($P < 0.01$) (Fig. S11). These data suggest that the lower *Fads2* copy number may be a constraint to colonization of DHA-deficient freshwater niches by Japan Sea stickleback.

Fluorescence *in situ* hybridization (FISH) revealed that *Fads2* was located only on the X chromosome (Linkage Group [LG] 19) in Japan Sea, but on LG12 and LG19 in Pacific Ocean stickleback (Fig. 3A). This result was confirmed by linkage analysis of *Fads2* copy number using an F_2 intercross (Fig. S12). Genes flanking *Fads2* on LG19, but not on LG12, showed conserved synteny with other teleosts (Fig. S13). Furthermore, an outgroup, *G. wheatlandi*, another marine stickleback with no known freshwater populations (*10*), has *Fads2* on LG19 but not on LG12 (Table S5) and similar copy numbers to Japan Sea stickleback (Fig. S14). Thus, LG19 is the ancestral location of *Fads2*, and copy-and-paste transposition of *Fads2* from LG19 to LG12 increased the ability to synthesize DHA in *G. aculeatus*, but not in *G. nipponicus* or *G. wheatlandi*. At the LG12 locus where a 12kb-insertion containing *Fads2* and several types of transposons exist in *G. aculeatus* (Fig. 3B and Fig. S15), *G. wheatlandi* and *G. nipponicus* possess transposons without *Fads2* (Fig. S16-S18). This suggests that transposons might have mediated the *Fads2* transposition and/or that this locus is a hotspot of insertion-deletion mutations (*23*).

The estimated timing of *Fads2* duplication within *G. aculeatus* is 0.80 MYA (95% highest posterior density: 0.47-1.16 MYA) (Fig. S19), which is much earlier than the end of the last glacial period (0.012 MYA), when the majority of stickleback freshwater colonization occurred (*1, 10*). Marine sticklebacks from western North America and Europe also repeatedly colonized freshwater and radiated into diverse ecotypes (*10*). Our results show that they also have the extra-copy of *Fads2* on LG12 with copy numbers similar to those of the Pacific Ocean stickleback in Japan (Fig. 4A, Fig. S16 and S22C-S22F). These data confirm that transposition onto LG12 occurred before the split between the Pacific and the Atlantic Ocean lineages (0.3-0.5 MYA) (*24*). Thus, the pre-existing duplication of *Fads2* has likely given *G. aculeatus* an advantage over other *Gasterosteus* species in colonizing freshwater. However, our estimate suggests that *Fads2* on LG12 is younger than the oldest known freshwater *Gasterosteus* fossil (*10*). Ancient extinct freshwater species may therefore have possessed additional *Fads2* copies somewhere in the genome or adapted to DHA-deprived diets through other mutations.

To investigate whether there are any other loci involved in survivorship on DHA-deficient diets, we conducted quantitative trait locus (QTL) mapping of survival rates using an F_2 intercross between the Pacific Ocean and Japan Sea sticklebacks. In addition to a suggestive QTL overlapping the *Fads2* gene on LG12 (3.3% of variance in survival explained), one significant and two additional suggestive QTLs were found on different autosomes (Fig. S20). The QTL on LG12, but not other QTLs, explained *Fads2* copy number variation. Two other QTLs, including a significant one, showed overdominance rather than additive effects on the survival, which may reflect an epistatic interaction between inter-species alleles (*25*) (Fig. S21).

Although survival rate is a polygenic and complex trait, our unbiased QTL analysis confirmed that the additional *Fads2* copy of the Pacific Ocean stickleback on LG12 contributes to survivorship on DHA-deficient diets.

Because the Japanese Pacific Ocean stickleback also have increased survivorship with marine-derived diets (Fig. 2A), additional *Fads2* duplications beyond the LG12 copy may further increase DHA biosynthetic ability and be beneficial for permanent freshwater residency. Indeed, in Japan, freshwater populations had even higher *Fads2* copy numbers than Pacific Ocean populations ($\chi^2_2 = 17.1$, $P < 0.01$) (Fig. 4A, Fig. S22-S23). Even within freshwater populations, those that had a longer evolutionary history in freshwater had higher *Fads2* copy numbers ($\chi^2_3 = 35.3$, $P < 0.01$) (Fig. 1C and Fig. S22). Additional copy number increase also occurred in North American ($\chi^2_1 = 4.4$, $P = 0.035$) and European freshwater populations ($\chi^2_2 = 7.2$, $P = 0.028$) (Fig. 4A and Fig. S22). We confirmed that freshwater populations with additional copies of *Fads2* had more DHA than the Pacific Ocean or a freshwater population with fewer copies when fed only DHA-free diets ($F_{2,8} = 12.6$, $P < 0.01$) (Fig. S24). Both linkage analysis and long-read genome sequencing showed that tandem duplications on the X chromosome are responsible for additional copy number increase in both Japanese and Canadian freshwater populations (Fig. S25 and Fig. S26). Transposons near *Fads2* might have facilitated these tandem duplications (Fig. S26) (*26*).

To test the generality of the mechanism, we first investigated ninespine sticklebacks (genus *Pungitius*). The freshwater species, *P. tymensis* and *P. kaibarae*, had higher *Fads2* copy numbers than *P. pungitius* ($P < 0.05$), which inhabits only brackish environments in Japan. *P. sinensis*, which inhabits both freshwater and brackish environments (*27, 28*), had intermediate copy numbers between *P. pungitus* and freshwater ninespine sticklebacks (Fig. 4B, and Fig. S27). We finally investigated *Fads2* copy numbers in the ray-finned fishes whose whole genome sequences have been determined (Fig. S28). Fish species that form freshwater populations had significantly higher *Fads2* copy numbers than entirely marine species (Fig. 4C and Fig. S29; MCMCglmm accounting for phylogeny, pMCMC < 0.01), suggesting convergent increases of *Fads2* copies in diverse taxa that successfully colonized freshwater.

Gene duplications not only enhance overall gene expression levels but also allow duplicated copies to acquire new functions (*29*). Our yeast functional assay suggested that *Fads2* genes in the Pacific Ocean stickleback acquired an additional enzymatic function in the DHA synthetic pathway (Fig. S6 and Table S4). Some of the Pacific Ocean-specific amino acid changes were shared by other freshwater ray-finned fishes (Table S6), suggesting that they may be responsible for the acquisition of new enzymatic function. In addition to amino acid changes, both *cis*- and *trans*-regulatory changes cause expression differences between *Fads2* haplotypes (Fig. S30). Given that overexpression of *Fads2* rescued the lethality in Japan Sea sticklebacks (Fig. 2F-2G), differences in the copy number itself likely contribute to differences in survival on DHA-deficient diets, although the possible involvement of changes in *Fads2* protein sequence and regulation cannot be excluded.

Taken together, our data demonstrate that *Fads2* is a key metabolic gene important for overcoming the nutritional constraint associated with freshwater colonization in fishes. Intriguingly, *Fads* shows strong signatures of selection in human populations that colonized polar regions, suggesting the importance of *Fads* in even more diverse taxa, including humans (*30*).

References and Notes:

1. D. Schluter, *The Ecology of Adaptive Radiation*. (Oxford University Press, 2000).
2. J. B. Losos, Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* **175**, 623-639 (2010).
3. G. G. Simpson, *The Major Features of Evolution*. Columbia University biological series (Columbia University Press, 1953).
4. O. Seehausen, C. E. Wagner, Speciation in freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics* **45**, 621-651 (2014).
5. R. Betancur-R, G. Ortí, R. A. Pyron, Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol Lett* **18**, 441-450 (2015).
6. C. E. Lee, M. A. Bell, Causes and consequences of recent freshwater invasions by saltwater animals. *Trends Ecol Evol* **14**, 284-288 (1999).
7. M. Kainz, M. T. Brett, M. T. Arts, *Lipids in Aquatic Ecosystems*. (Springer Science & Business Media, 2009).
8. M. V. Bell, D. R. Tocher, Biosynthesis of polyunsaturated fatty acids in aquatic ecosystems: general pathways and new directions. 211-236 (2009).
9. D. Swanson, R. Block, S. A. Mousa, Omega-3 fatty acids EPA and DHA: health benefits throughout life. *Advances in Nutrition*, 1-7 (2012).
10. M. A. Bell, S. A. Foster, *The Evolutionary Biology of the Threespine Stickleback*. (Oxford University Press, 1994).
11. M. Ravinet *et al.*, The genomic landscape at a late stage of stickleback speciation: high genomic divergence interspersed by small localized regions of introgression. *PLoS Genet* **14**, e1007358 (2018).
12. J. Kitano, S. Mori, C. L. Peichel, Phenotypic divergence and reproductive isolation between sympatric forms of Japanese threespine sticklebacks. *Biol J Linn Soc* **91**, 671-685 (2007).
13. J. Kitano *et al.*, A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079-1083 (2009).
14. M. Higuchi, H. Sakai, A. Goto, A new threespine stickleback, *Gasterosteus nipponicus* sp. nov. (Teleostei: Gasterosteidae), from the Japan Sea region. *Ichthyol. Res.* **61**, 341-351 (2014).
15. T. Adachi *et al.*, Shifts in morphology and diet of non-native sticklebacks introduced into Japanese crater lakes. *Ecology and Evolution* **2**, 1083-1098 (2012).
16. M. Ravinet, N. Takeuchi, M. Kume, S. Mori, J. Kitano, Comparative analysis of Japanese three-spined stickleback clades reveals the Pacific Ocean lineage has adapted to freshwater environments while the Japan Sea has not. *PLoS ONE* **9**, e112404 (2014).
17. L. M. Cassidy, M. Ravinet, S. Mori, J. Kitano, Are Japanese freshwater populations of threespine stickleback derived from the Pacific Ocean lineage? *Evol Ecol Res* **15**, 295-311 (2013).
18. Further details are available in the supplementary materials.
19. M. Kume, S. Mori, Sea-run migratory behaviour in the Japan Sea form of three-spined stickleback *Gasterosteus aculeatus* in the tidal pool of eastern Hokkaido Island, Japan. *J. Fish Biol.* **75**, 2845-2850 (2009).
20. M. T. Nakamura, T. Y. Nara, Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. *Annu Rev Nutr* **24**, 345-376 (2004).

21. D. R. Tocher, Fatty acid requirements in ontogeny of marine and freshwater fish. *Aquac Res* **41**, 717-732 (2010).
22. L. F. C. Castro, D. R. Tocher, Ó. Monroig, Long-chain polyunsaturated fatty acid biosynthesis in chordates: Insights into the evolution of Fads and Elovl gene repertoire. *Prog. Lipid Res.* **62**, 25-40 (2016).
23. K. D. Makova, R. C. Hardison, The effects of chromatin organization on variation in mutation rates in the genome. *Nat Rev Genet* **16**, 213-223 (2015).
24. G. Ortí, M. A. Bell, T. E. Reimchen, A. Meyer, Global survey of mitochondrial DNA sequences in the threespine stickleback: Evidence for recent migrations. *Evolution* **48**, 608-622 (1994).
25. O. Seehausen, Hybridization and adaptive radiation. *Trends Ecol Evol* **19**, 198-207 (2004).
26. V. V. Kapitonov, J. Jurka, Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet* **23**, 521-529 (2007).
27. A. Ishikawa *et al.*, Speciation in ninespine stickleback: reproductive isolation and phenotypic divergence among cryptic species of Japanese ninespine stickleback. *J Evol Biol* **26**, 1417-1430 (2013).
28. H. Takahashi *et al.*, Species phylogeny and diversification process of Northeast Asian *Pungitius* revealed by AFLP and mtDNA markers. *Mol Phylogenet Evol* **99**, 44-52 (2016).
29. S. Ohno, *Evolution by gene duplication*. (Springer-Verlag, 1970).
30. R. Nielsen *et al.*, Tracing the peopling of the world through genomics. *Nature* **541**, 302-310 (2017).

Acknowledgments: We thank Tim Howes, Mark Ravinet, Kohta Yoshida, Koichi Kawakami, Kitano Lab members, many field assistants, the Aquatoto Gifu staff, Satoko Kondo, and Lina Kawaguchi for technical assistance, Bertil Borg, Hikaru Kanazawa, and Masumi Ozaki for providing fish, and Mike Bell for discussion. **Funding:** This project was supported by JSPS KAKENHI 15H02418 and 23113007 to J.K. and 16H06279 to A.T. and JSPS Postdoctoral Fellowships 11J04816 to A.I. and 16J06812 to N.K. **Author contributions:** A.I. and J.K. conceived and designed the study; all authors collected and analyzed data; A.I. and J.K. wrote the first draft with input from the other authors. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All generated sequence data are available from DDBJ (DRA006706; DRA007515-DRA007519; DRA007568; LC431555-413556), and fatty acid meta-analysis data are available from DRYAD (accession numbers here upon acceptance).

FIGURES

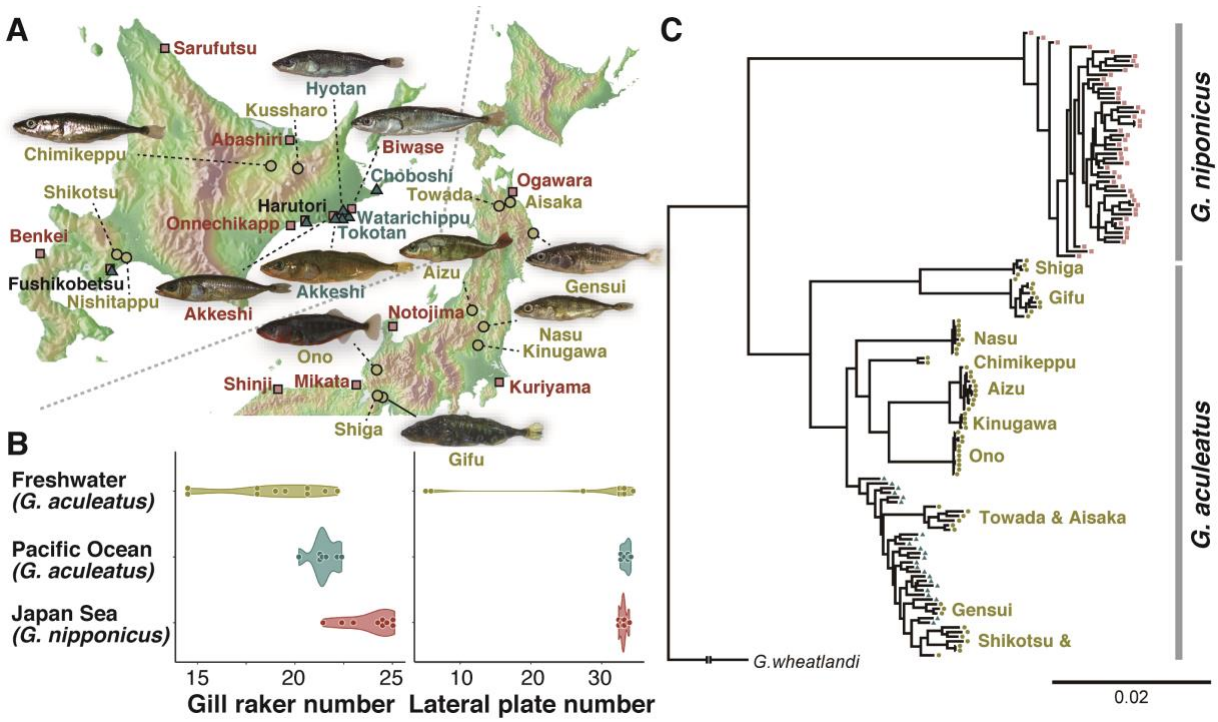

Fig. 1. Freshwater colonization and diversification in *G. aculeatus* but not in *G. nipponicus*. (**A**) Sampling sites in Japan: pink square, *G. nipponicus*; blue triangle, Pacific Ocean stickleback; green circle, freshwater. (**B**) Diversification of key foraging (gill raker number) and armor traits (lateral plate number) in freshwater populations. (**C**) ddRAD-seq phylogeny of Japanese *Gasterosteus* indicating that all freshwater populations (green circle) belonged to *G. aculeatus* (blue triangle) rather than *G. nipponicus* (pink square). Bar, the substitution rate. For bootstrap values, see Fig. S3.

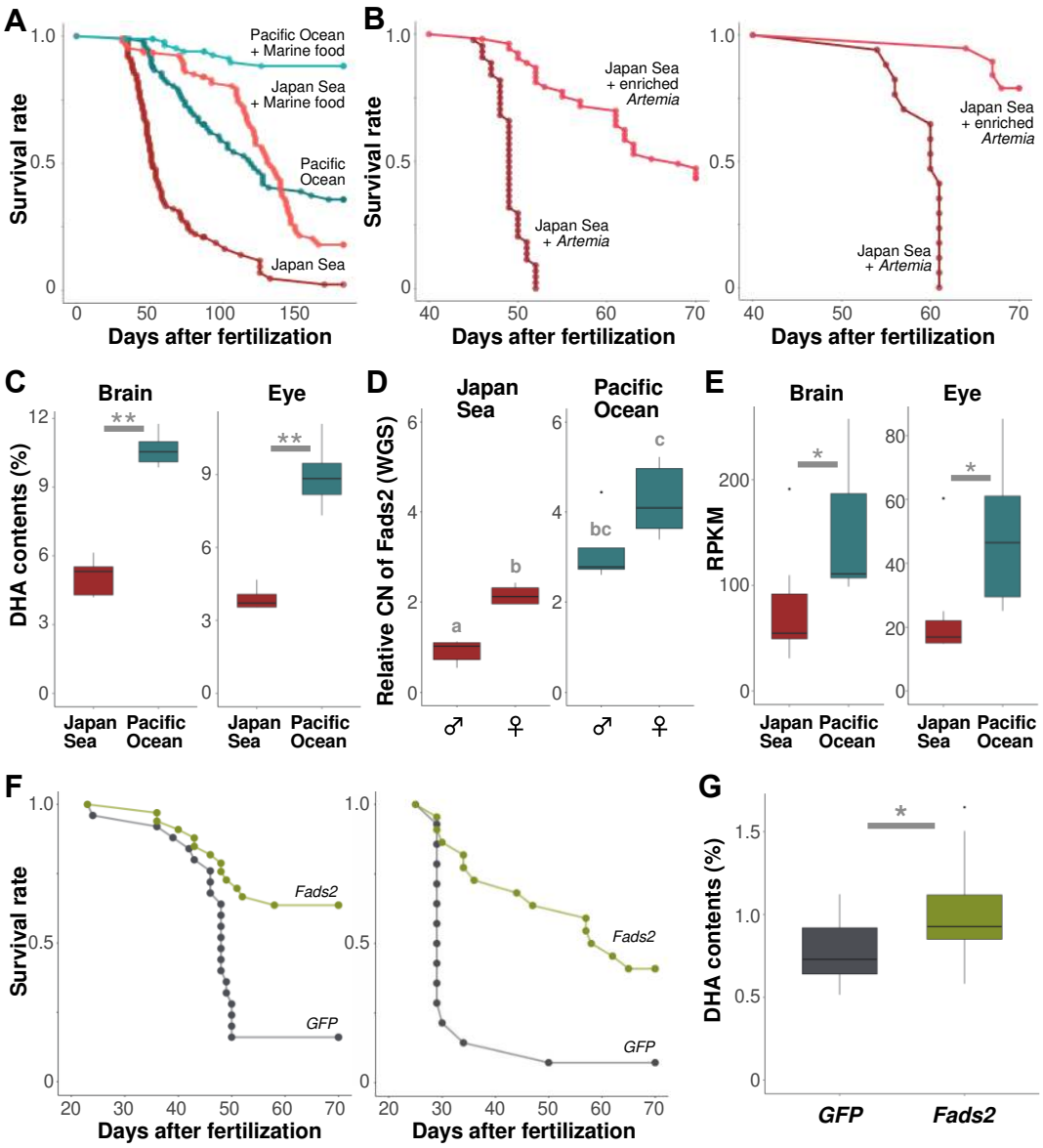

Fig. 2. Contribution of higher *Fads2* copy numbers to survival with DHA-free diets. (**A**) Survival curves of Japan Sea (red) and Pacific Ocean stickleback fed only DHA-free *Artemia* (blue) and Japan Sea (pink) and Pacific Ocean stickleback fed marine-derived diets (light blue). (**B**) Survival curves of Japan Sea fed with DHA-free *Artemia* (dark-red) and *Artemia* enriched with several fatty acids, including DHA (pink). Two panels indicate independent replicate crosses. (**C**) DHA contents in brain and eye of two species fed only DHA-free *Artemia*. ** $P < 0.01$. (**D**) Relative copy numbers of *Fads2* in males and females of Japan Sea and Pacific Ocean sticklebacks, estimated from whole genome resequencing data. Different letters above the boxes indicate significantly different pairs ($P < 0.05$). (**E**) Expression levels of *Fads2* at 40-60 days after fertilization. * $P < 0.05$. (**F**) Survival curves of *Fads2*- (yellow-green) and *GFP*-transgenic (grey) Japan Sea fish fed only DHA-free *Artemia*. Two panels indicate independent replicate crosses. (**G**) Whole body DHA content of *Fads2*- (yellow-green) and *GFP*-transgenics (grey) at 40 days after fertilization, when Japan Sea fish start to die with DHA-free diets. *$P < 0.05$.

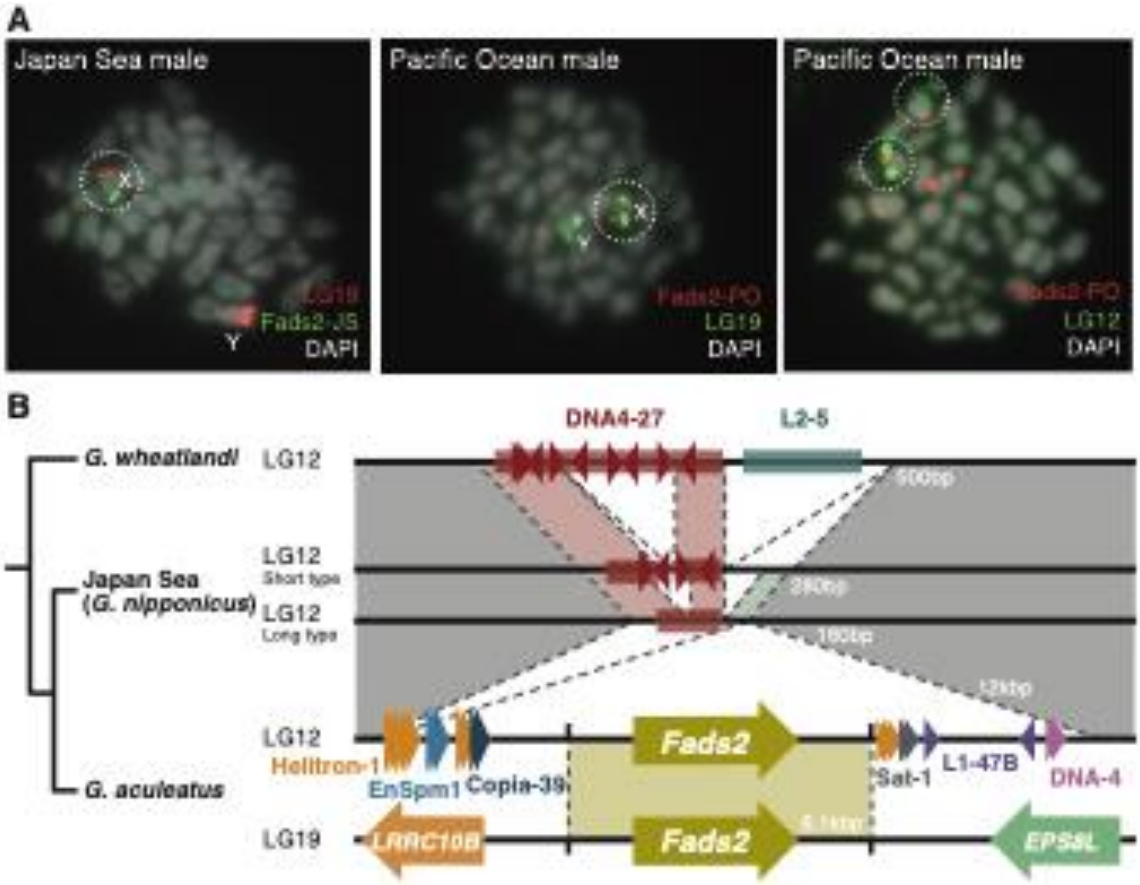

Fig. 3. Extra-copy of *Fads2* on LG12 in *G. aculeatus*. (**A**) FISH results with DAPI-nuclear staining. The left panel indicates a Japan Sea male with one *Fads2* copy (green) on the X chromosome (circled). In a Pacific Ocean male, *Fads2* (red) was located on the X chromosome (circle in the middle panel) and both copies of autosomal LG12 (circles in the right panel). The colors for *Fads2* and LG19 or LG12 are flipped between the Japan Sea and Pacific Ocean males. Note that LG19 is either the X or Y chromosome, and the LG19 probe detects the region retained on the Y chromosome. (**B**) Genome structure around *Fads2* on LG19 and LG12 of *G. aculeatus* and the corresponding region on LG12 of *G. nipponicus* and *G. wheatlandi*. Arrows and arrowheads indicate genes and repetitive sequences, respectively. White numbers indicate the insertion size.

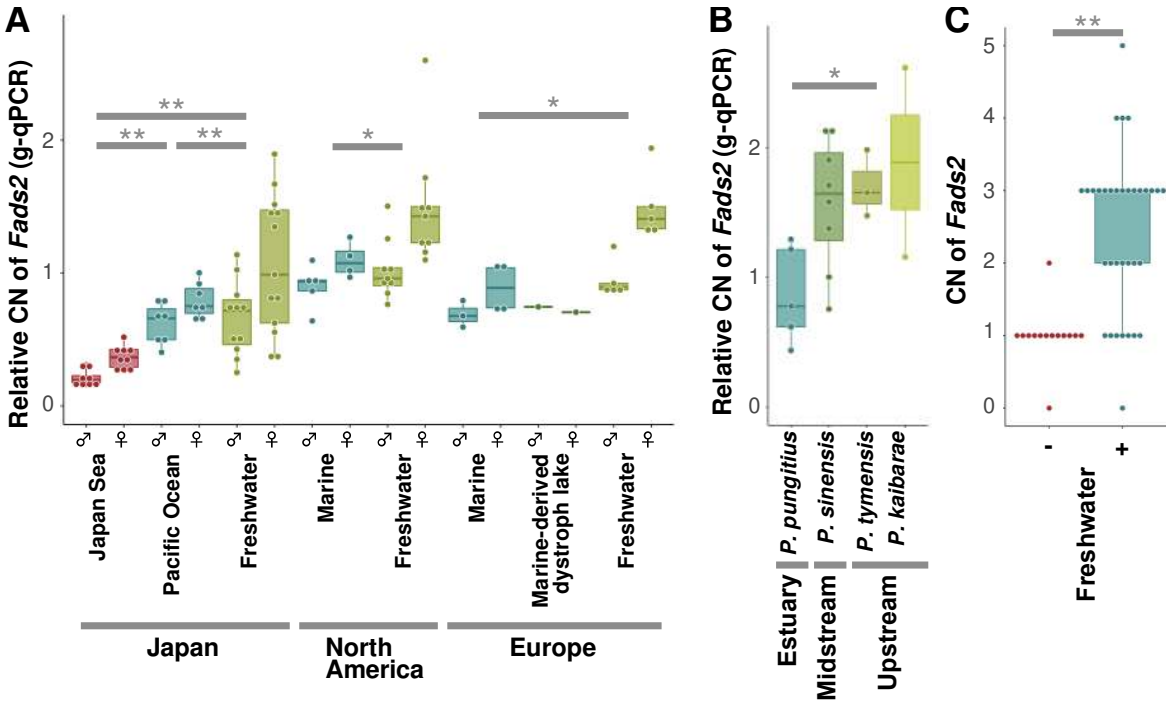

Fig. 4. Parallel increase in *Fads2* copies in freshwater fishes. (**A**) Relative copy numbers (CN) of *Fads2* in males and females of *Gasterosteus* populations: red, Japan Sea; blue, Japanese Pacific Ocean, North American and European marine ecotypes; green, freshwater ecotypes. **P* < 0.05, ***P* < 0.01. (**B**) Relative CN of *Fads2* in *Pungitius*. A single dot indicates the median of a single population in (A) and (B). (**C**) Comparison of *Fads2* copy numbers among ray-finned fishes utilizing marine (-) or freshwater (+) niches.

Supplementary Materials for

A key metabolic gene for recurrent freshwater colonization and radiation in fishes

Asano Ishikawa, Naoki Kabeya, Koki Ikeya, Ryo Kakioka, Jennifer N. Cech, Naoki Osada, Miguel C. Leal, Jun Inoue, Manabu Kume, Atsushi Toyoda, Ayumi Tezuka, Atsushi J. Nagano, Yo Y. Yamasaki, Yuto Suzuki, Tomoyuki Kokita, Hiroshi Takahashi, Kay Lucek, David Marques, Yusuke Takehana, Kiyoshi Naruse, Seiichi Mori, Oscar Monroig, Nemiah Ladd, Carsten Schubert, Blake Matthews, Catherine L. Peichel, Ole Seehausen, Goro Yoshizaki, and Jun Kitano

Correspondence to: jkitano@nig.ac.jp

This PDF file includes:

Materials and Methods
Figs. S1 to S30
Tables S1 to S7

Other Supplementary Materials for this manuscript include the following:

None

Materials and Methods

Fatty acid analysis of stickleback food items

For meta-analysis, we compiled data on the percentage of docosahexaenoic acid (DHA; 22:6 n-3) relative to total fatty acids for consumers (invertebrates that mostly are primary consumers, and omnivores) that are representative of the potential prey items available for stickleback in the pelagic and benthic habitats of freshwater and marine ecosystems. Overall, we compiled 130 measurements of DHA (63 freshwater, 67 marine) from 25 studies (*31-56*), which included a broad range of taxonomic diversity (72 species). We added this dataset to an existing meta-analysis (DRYAD depository:10.5061/dryad.67dg6) that compared the percentage of DHA in primary producers at the base of marine and freshwater ecosystems (*57*). We also compiled data of algae. These data will be deposited to DRYAD upon acceptance (accession no. here). The results are plotted in Figure S1A, and they show that (i) DHA content in freshwater consumers is lower than marine consumers, and (ii) high DHA content in freshwater ecosystems is primarily available in copepods, which are a common albeit evasive prey.

To directly compare DHA contents in stickleback prey items, stickleback prey items (about 1.5 g) were collected from natural habitats from May to August in 2013 and 2014, frozen in dry ice immediately, and kept at -80°C until analysis. Unsaturated fatty acid concentrations were measured using GC-MS analysis at Shokukanken (Maebashi, Japan); the minimum detectable quantity was 50 mg/100 g.

Ethics statement

All animal experiments were approved by the institutional animal care and use committee of the National Institute of Genetics (26-14, 27-4, 28-11, 29-11, 30-12). Before dissection, all fish were euthanized with ethyl 3-aminobenzoate methanesulfonate (MS-222).

Fish sampling

Threespine sticklebacks were collected with seine nets and minnow traps from each habitat in Japan as described previously (*11, 12*). Japan Sea and Pacific Ocean sticklebacks collected from Lake Akkeshi in 2012 were used for making pure crosses to compare survival rates between fish fed *Artemia* and those fed marine-derived diets, and for making an F_2 intercross to test the relationship between *Fads2* copy number and lethality. Japan Sea stickleback from Biwase collected in 2014 and 2015 were used for making crosses to conduct fish rearing experiments with DHA-enriched and DHA-free *Artemia* as well as to make transgenic fish. Collection of North American populations was described previously (*58, 59*). Collection of Japanese *Pungitius* samples was reported previously (*26, 27*). See Table S7 for further details of the sampling sites (11, 12, 14, 26, 27, *60, 61-68*).

Morphological analysis of Japanese sticklebacks

Morphological data for two Pacific Ocean populations (Akkeshi and Hyotan), one Japan Sea population (Akkeshi), and one freshwater population (Nishitappu) were reported previously (*63*). In the present study, we measured additional populations: eight from the Japan Sea, five from the Pacific Ocean, and nine freshwater threespine stickleback populations (Table S1). To exclude the effects of sex differences, we used only adult males for measurement (*63*). Sex was determined either by visual examination of the gonads or by PCR to detect sex specific markers in the isocitrate dehydrogenase gene (*IDH*) (*69*) or the leucine-rich repeat containing 61 gene (*LRRC61*); the latter was designed based on the X and Y sequences of the two species (*70, 71*),

because PCR at the *IDH* gene does not distinguish sexes in Japan Sea stickleback (*69*). The primers used were: *IDH* forward, GGGACGAGCAAGATTTATTG; *IDH* reverse, TTATCGTTAGCCAGGAGATGG; *LRRC61* forward, TGGGGAGTGTATGAGTTTAGAGAA; and *LRRC61* reverse, CTCAAAGGCTGAAGGCAGTG.

Standard length (SL), head length (HL), body depth (BD), upper-jaw length (JL), snout length (SnL), gape width (GW), first and second dorsal-spine length (DS1 and DS2), pelvic-spine length (PSL), and pelvic-girdle length (PGL) were measured with vernier calipers as described previously (*63*). Gill raker number (GRN) and lateral plate number (LPN) were counted under a dissecting microscope. For GRN and LPN, we used both males and females, because they are not sexually dimorphic in Japanese populations (*63, 72*). For LPN, fish with a standard length greater than 32 mm were used (*72*). We then calculated the population mean for each trait and conducted principal component analysis (PCA) using a correlation matrix with the statistical package R version 3.0.2 (*73*) .

Phylogenetic analysis of Japanese sticklebacks

We used four to twelve individuals per each population for the phylogenetic analysis. Genomic DNA was isolated with a DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, USA). Double digest RAD sequencing (ddRAD-seq) was performed as described previously (*74*). In brief, 10 ng of genomic DNA were digested with EcoRI and BglII, followed by adaptor ligation and amplification with uniquely barcoded primers. The libraries were run on HiSeq 2000 in a 50 bp single-end mode at Macrogen (Seoul, Korea). The sequence data are available from DNA Data Bank of Japan (DDBJ) (DRA007568). For Pacific Ocean sticklebacks collected from Lake Towada, Aisaka Stream, and Akkeshi, Japan Sea sticklebacks from Akkeshi, and *G. wheatlandi*, we used previously determined whole genome sequence (WGS) data (DRA001086, DRA001136, and DRA005065) (*71, 75*).

Short reads were trimmed with Trimmomatic (*76*) with the following parameters: LEADING:19 TRAILING:19 SLIDINGWINDOW:30:20 AVGQUAL:20 MINLEN:51. Trimmed reads were mapped to the BROADS S1 stickleback reference genome sequence using SMALT (http://www.sanger.ac.uk/science/tools/smalt-0) after building the hash index of the reference with a word length of 13 and a sampling step size of 2. In order to restrict the downstream analyses of WGS data to the genomic regions that were mapped with at least one RAD sequence, such regions were identified using bedtools (*77*). SNPs were called with FreeBayes (*78*). SNPs with a coverage of less than three were discarded. VCFtools (*79*) was used to include those SNPs that were observed in all of the populations with the maximum missing rate of 0.3 within a population, did not depart from Hardy–Weinberg equilibrium ($P \geq 0.001$), and located on autosomes (i.e. excluding chromosomes IX or XIX). A maximum likelihood phylogenetic tree was constructed using RAxML (*80*). We used K80 model of nucleotide substitution with ascertainment bias correction using the conditional likelihood method (*81*). We conducted a rapid bootstrap analysis with 100 replicates and a search for the best-scoring tree in a single run. The tree was visualized with FigTree (http://tree.bio.ed.ac.uk/software/figtree).

Fish rearing experiments under different dietary conditions

In order to compare the survival rates of fish subjected to two different feeding regimes, pure crosses of Japan Sea and Pacific Ocean sticklebacks were divided into three experimental groups after hatching: fish maintained in freshwater and fed *Artemia* ($N = 112$ for Japan Sea and $N = 92$

for Pacific Ocean sticklebacks), fish maintained in freshwater and fed marine diets (N = 107 for Japan Sea and N = 92 for Pacific Ocean sticklebacks), and fish maintained in seawater and fed *Artemia* (N = 66 for Japan Sea and N = 72 for Pacific Ocean sticklebacks). For fish fed marine-derived diets, we fed minced mackerel, minced horse mackerel, and frozen mysids in addition to *Artemia*. Mysids are natural food sources for marine sticklebacks, while minced mackerel and horse mackerel are not natural food sources, but are often used as diets for aquaculture of marine fishes. We monitored survival every day until 182 days after fertilization. For rescue experiments with DHA-enriched *Artemia*, two independent Japan Sea crosses were used. Each family was divided into two experimental groups after hatching: one group fed DHA-enriched *Artemia* (N = 20 and 54) and the other fed newly hatched *Artemia*, which is free of DHA (N = 18 and 44), with both groups maintained in freshwater. In order to prepare DHA-enriched and DHA-free *Artemia*, newly hatched *Artemia* nauplii (Marinetech Co., Ltd.) were transferred into two 5 L containers (70 individuals/mL) and enriched with Hyper Green (Marinetech Co., Ltd., Aichi, Japan) following manufacturer's instructions, or were left untreated as a DHA-free control. After 4 h culture in 80% sea water (final salinity of 25.6 g/L) at 28 °C, *Artemia* were harvested, washed with seawater and then freeze at -20 °C until further use. The lipid extraction and fatty acid methyl ester (FAME) preparation of *Artemia* were carried out following the previously described method (*82*). Using gas chromatography (GC), we confirmed that the enriched *Artemia*, but not the control *Artemia*, contained DHA (Table S3). The GC conditions for FAME analysis were described previously (*82*). It should be noted that the enriched *Artemia* also increased EPA, DPAn-6, and ARA (Table S3). EPA levels were higher in *Artemia*-fed Japan Sea stickleback compared to *Artemia*-fed Pacific Ocean stickleback (Fig. S5), suggesting that EPA was not deficient in Japan Sea stickleback fed with the control *Artemia*. As for DPAn-6 and ARA, our meta-analysis data showed that freshwater consumers generally contain significantly more DPAn-6 (ANOVA, $F_{1,76} = 9.26$, $P = 0.0032$) and ARA (ANOVA, $F_{1,109} = 10.93$, $P = 0.0013$) than marine consumers, suggesting that deficiency of DPAn-6 and ARA is unlikely to occur in freshwater environments.

To investigate whether survival rates differ between Japan Sea males and females, four independent crosses were used (N = 97, 53, 113 and 184). Sex was determined by PCR using the *LRRC61* gene (see above). The fish were fed only *Artemia*. We monitored survival rate every day until 205 days after fertilization. Survival curves were drawn by the Kaplan-Meier method and were compared using the log-rank test in R version 3.0.2 (*73*).

Fatty acid analysis of sticklebacks

The whole brain, eyeballs, and liver of the Japan Sea and Pacific Ocean juveniles were dissected at 50 days after fertilization following euthanasia with MS-222 and rapidly frozen in liquid nitrogen: two female and three male juveniles of Japan Sea stickleback and three female and three male juveniles of Pacific Ocean stickleback were analyzed. They were kept at -80°C until analysis. Since our preliminary analysis showed that the liver did not contain substantial amounts of DHA in the threespine stickleback, we analyzed only the brain and eyeballs. The collected tissues were lyophilized and then lipid extracted by the previously described method (*82*). The FAMEs were prepared and purified using fatty acid methylation kit and the methylated fatty acid purification kit (Nacalai Tesque, Inc., Kyoto, Japan), respectively. The GC conditions for FAME analysis were described previously (*82*). Student's t-test was used for statistical analysis. To compare DHA content of an old freshwater population from Gifu with those of Japan Sea and Pacific Ocean sticklebacks, eyeballs were dissected from adult Japan Sea (N = 3), Pacific Ocean

(N = 3), and Gifu freshwater sticklebacks (N = 5) raised in tanks and fed only DHA-free *Artemia*. The eyeballs were frozen in liquid nitrogen and kept at -80°C until analysis. GC analysis to measure unsaturated fatty acid concentrations was conducted at Toray Research Center, Inc. (Tokyo, Japan). In order to compare Marina Rheinhof and Le Bouveret from European populations, pure crosses were raised in tanks and fed only with *Artemia* and chironomids. Freeze-dried muscle tissue from 13 Marina Rheinhof individuals and 17 Le Bouveret individuals was ultrasonically extracted in 2:1 dichloromethane/methanol (4x) after the quantitation addition of n-$C_{19:0}$ fatty acid as a recovery standard. The resulting total lipid extracts were methylated with 1 mL of 14% BF_3 in methanol (2 hours at 100 °C) and quantified by gas chromatography - flame ionization detection (GC-FID), following previously described methods (*83*).

Measurement of relative copy number of *Fads2* from WGS data

Previously reported WGS data for five male and five female Japan Sea and one male and five female Pacific Ocean sticklebacks (*71*), and new WGS data for one female and three male Pacific Ocean sticklebacks (DRA001136) were mapped to the BROADS S1 reference genome sequence of an Alaskan lake female (http://www.ensembl.org) (*84*). New WGS data were generated using Illumina TruSeq DNA Library Construction Kit and HiSeq 2000 as described previously (*71*). Trimming and mapping of the new WGS data to the reference sequence were conducted as described previously (*71*). Since there are three *Fads2* genes in the reference genome (ENSGACG00000005827, ENSGACG00000005442, and ENSGACG00000018692), which are located on LG12, LG19, and Scaffold654 respectively, the total number of reads mapped to any of the three *Fads2* loci was counted using CLC Genomics Workbench Software 6.5 (CLC bio, Aarhus, Denmark). To compare relative copy numbers of *Fads2* between individuals, the number of reads mapped to *Fads2* was divided by genome-wide coverage. Two-way ANOVA, followed by Tukey's HSD test, was conducted using R version 3.0.2 (*73*). Because we found sequence differences between and within Japan Sea and Pacific Ocean sticklebacks in the *Fads2* cDNA (see below), we classified four different haplotypes based on the SNPs located at 37-45 bp after the start codon: EVK type, GAGGTCAAG; QVN type, CAGGTCAAC; LAN type, CTGGCCAAC; and LAK type, CTGGCCAGG. All sequence data are available from the DDBJ short read archive (DRA001136).

RNA sequencing (RNA-seq)

Whole brain, eyes, liver, and body trunk including muscle and skin were dissected from the pure crosses of Japan Sea and Pacific Ocean sticklebacks at 40-60 days after fertilization (three females and five males, respectively) following euthanasia with MS-222. They were fed only DHA-free *Artemia* after hatching. Total RNA was extracted from each tissue with a RNeasy Micro Kit (Qiagen). For each tissue, 100 ng of RNA was used for making RNA-seq libraries with a TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Libraries from different individuals were labelled with different barcode adaptors, pooled, and run on one lane of Illumina HiSeq 2000 with a paired end 2 x 100 bp mode at RIKEN GENESIS (Yokohama). From the sequence reads obtained (mean reads ± SE/fish = 4.06 ± 0.15 million for each individual), nucleotides with low quality scores (less than 20) and two or more ambiguous nucleotides at the ends were removed with CLC Genomics Software 6.5. The trimmed sequence reads were then mapped to the BROAD S1 reference genome sequence with gene annotations (http://www.ensembl.org) using RNAseq Analysis software included in the CLC Genomics Workbench to calculate reads per kilo base of exon per

million mapped reads (RPKM) of each transcript with the following parameters: maximum number of mismatches allowed = 2, minimum length fraction = 0.9, minimum similarity fraction = 0.8, and unspecific match limit = 10. Since there are three *Fads2* genes in the reference genome (ENSGACG00000005827, ENSGACG00000005442, and ENSGACG00000018692), RPKM of the transcripts mapped to any of these three genes were summed to calculate the total RPKM value. Sequence data are available from the DDBJ short read archive (DRA007516).

To examine which *Fads2* haplotype is expressed at higher levels within individuals of the Pacific Ocean stickleback, we sequenced the genome of *Fads2* for all Pacific Ocean sticklebacks used for RNA-seq. First, genomic PCR was conducted with two sets of primers: forward-1 (ACCCAGGTGTCAAAGTCTGC) and reverse-1 (TGTACACCTTCCGATCGATCAC) for distinguishing between EVK and QVN haplotypes; forward-2 (GCGACCATCATACAGGATTTCC) and reverse-2 (CTGAGAGGTTGCCAGCAGGA) for distinguishing between JS and PO haplotypes. PCR products were sequenced with capillary sequencers at Eurofins Genomics (Tokyo, Japan). Next, we conducted qPCR of genomic DNA as described below. Using these data, we determined which haplotypes each fish has. Next, using the SNPs in the RNA-seq reads, we counted the number of each haplotype expressed in each tissue of each fish. To test the association between the genotypes and the allele-specific expression levels, we used Pearson's correlation test.

Cloning of stickleback *Fads2* cDNA and sequence analysis

To investigate allelic sequence variation in *Fads2* between and within species, *Fads2* cDNA was cloned from Japan Sea and Pacific Ocean sticklebacks. Total RNA was extracted from the livers of eight Japan Sea and eight Pacific Ocean sticklebacks collected from Akkeshi Bay using TRIzol Reagent (Thermo Fisher Scientific Inc., Waltham, MA, USA), followed by further purification with an RNeasy Mini Kit (Qiagen). One microgram of total RNA was treated with DNase I (Thermo Fisher Scientific Inc.) to eliminate genomic DNA contamination and used for cDNA synthesis with PrimeScript Reverse Transcriptase (Takara Bio, Shiga, Japan). Forward (5′-ACCCAGGTGTCAAAGTCTGC-3′) and reverse primers (5′-TGTAGGTGCAGAGCAGAGGA-3′) were used to amplify *Fads2* cDNA fragments by polymerase chain reaction (PCR) with KOD -Plus- Neo following the manufacturer's instructions (Toyobo Biochemicals, Osaka, Japan). Nucleotide sequences were determined with a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA) using a DNA sequencer ABI PRISM 3130 (Applied Biosystems).

Heterologous expression of *Fads2* in *Saccharomyces cerevisiae*

The cDNAs of *Fads2* genes are PCR amplified using corresponding primers containing restriction enzyme sites to ligate into the yeast expression vector pYES2 (Invitrogen or Thermo Fisher Scientific). The primers used are as follows; forward primer for PO QVN type, PO EVK type and JS QVN type, 5′- CCCAAGCTTAGGATGGGAGGTGGAGGCCA-3′; reverse primer for PO QVN type and PO EVK type, 5′- CCGTCTAGATCATTTATGGAGATAAGCGTCAA-3′; reverse primer for JS QVN type, 5′- CCGTCTAGATCATTTATGGAGATATGCGTCAA-3′. The PCR were carried out using PrimeSTAR Max DNA polymerase (TaKaRa bio) following the manufacturer's instructions. The initial denaturation step and final extension step for all PCRs were 98ºC for 2 min and 72ºC for 2 min, respectively. The cycle reactions were repeated for 30 times, which comprised 98ºC for 10 sec, 55ºC for 5 sec and 72ºC for 1 min 30 sec for *Fads2*. The amplified fragment was digested by the corresponding restriction enzymes (TaKaRa bio)

and then purified from agarose gel (FastGene Gel/PCR Extraction Kit, NIPPON Genetics Co, Ltd). Subsequently, the digested fragment was ligated into similarly digested pYES2 using T4 DNA ligase (Promega) following the manufacturer's instructions. The yeast transformation and culture were described previously (*85*). The transgenic yeast were grown in the presence of ALA (18:3n–3), ETA (20:4n–3) or DPA (22:5n–3). In order to investigate *Fads2* desaturase activity towards TPA (24:5n–3), the yeast constitutive expression vector (p415TEF) containing zebrafish *elovl2* gene was co-transformed with each pYES2 containing *Fads2* gene following the previously described method (*86*). The resulting double-transgenic yeast were grown in the presence of ALA (18:3n–3) or TPA (22:5n–3). The yeast growing conditions were also described previously (*86*). The transgenic yeast were harvested, and the FAMEs were prepared and analyzed following the method described previously (*85, 86*).

In vivo functional assay of *Fads2* in Japan Sea stickleback

Transgenic sticklebacks were created by microinjection based on Hosemann et al. (2004) (*87*) and Chan et al. (2010) (*88*). In brief, stickleback eggs were stripped from gravid females, and testes were dissected from males and minced in Hank's solution to obtain sperm. The stripped eggs were then fertilized by adding the sperm solution. One-cell stage embryos were placed on a glass plate under a Leica M80 stereomicroscope, and then injection was conducted using an MN-153 Narishige micromanipulator and a FemtoJet express (Eppendorf, Hamburg, Germany). An injection needle was pulled with P-1000IVF (Sutter Instrument, Novato, CA).

The tol2 transposon system was used for stickleback transgenesis (*89*). A control plasmid (pT2AL200R150G) containing the EF1a promoter, *EGFP* cDNA, and SV40-poly-A, which are flanked by tol2 transposable elements, was provided by Dr. Koichi Kawakami (http://kawakami.lab.nig.ac.jp/trans.html). *EGFP* was replaced with a Japan Sea *Fads2* cDNA to construct the *Fads2*-expressing plasmid. The fertilized eggs were co-injected with either the control *EGFP*-expressing plasmid or the *Fads2*-expressing plasmid together with mRNA encoding the tol2 transposase. The transposase mRNA was transcribed *in vitro* from the NotI-linearized pCS-TP or pCS-zT2TP plasmid (donated by Dr. Koichi Kawakami) using an mMESSAGE mMACHINE kit (Life Technologies).

Transgenic fish were made from two independent crosses: Family1, *Fads2*-transgenic (N = 34) and *EGFP*-transgenic (N = 25); Family2, *Fads2*-transgenic (N = 23) and *EGFP*-transgenic (N = 15). In each family, one clutch was divided into two, with one half injected with the control *GFP* plasmid and the other with the *Fads2* plasmid. The transgenic fish were kept in 10% seawater and fed only DHA-free *Artemia*. We monitored their survival every day until 70 days after fertilization in 2014. Survival curves were drawn by the Kaplan-Meier method and were compared using the log-rank test in R version 3.0.2 (*73*). We did not screen for the fish that were successfully transformed, so the statistics also reflect noise from injections that failed.

For fatty acid analysis of transgenic sticklebacks, we created transgenic fish again in 2015 and kept them under the same conditions. We sampled them at two weeks (N = 11 and 12 for *Fads2* and *GFP*, respectively), four weeks (N = 12 and 12 for *Fads2* and *GFP*, respectively), six weeks (N = 19 and 20 for *Fads2* and *GFP*, respectively), and eight weeks (N = 13 and 12 for *Fads2* and *GFP*, respectively) after fertilization. Whole bodies were frozen in liquid nitrogen after removing the digestive tracts to avoid contamination by undigested stomach contents and kept at -80°C until analysis. GC-MS analysis was performed as described in the section entitled *Fatty acid analysis of sticklebacks*. Data were analyzed using Student's *t*-test.

Association between *Fads2* copy number and survivorship in F_2 hybrids

Three hundred and twenty-two F_2 individuals derived from a single Pacific Ocean grandmother and a single Japan Sea grandfather were kept in a tank of 10% seawater and fed only DHA-free *Artemia*. The survival time of each individual was recorded, and genomic DNA was extracted from the ethanol preserved body using a DNeasy Blood & Tissue Kit (Qiagen). Relative copy numbers of *Fads2* were determined by genomic qPCR. We designed forward and reverse primers and a TaqMan probe on a region conserved among *Fads2* haplotypes using Primer Express software v3.0 (Applied Biosystems): forward primer, 5′-TTAGTCTTGGACCCCCACTGA-3′; reverse primer, 5′-AAGTGTGGAAATCCTGTATGATGGT-3′; and TaqMan probe, [FAM]-TGGCGTCTGCTCGTCTCCGCT-[TAM]. We used *TSHβ2* as a reference control to quantify the relative copy number of *Fads2* because *TSHβ2* shows no signs of copy number variation. The sequences of the primers and the probe are as follows: forward primer, 5′-TTCATGATCCAGCGCAGCT-3′; reverse primer, 5′-CAGCCGGGCAAATGGAC-3′; and TaqMan probe, [FAM]-TGCCTCTCTCCCTGGTGTACCGAGC-[TAM]. TaqMan Real-Time PCR Master Mix (Applied Biosystems) was used for genomic qPCR reactions on a StepOnePlus (Applied Biosystems). Relative expression levels were calculated from standard curves drawn from the serially diluted DNA of one F_2 fish.

In order to estimate the exact number of *Fads2* copies, we sequenced *Fads2* of F_2 hybrids (125 males and 57 females) and distinguished among the haplotypes using the SNPs. First, we conducted microsatellite analysis of 125 F_2 males and confirmed that in this cross the Pacific Ocean QVN haplotype, Pacific Ocean EVK haplotype and Japan Sea QVN haplotype were located on LG12, LG19 and LG19, respectively. Microsatellite markers, *Stn140* (9,178,675-9,178,837 bp in the ensemble coordinate) and *Stn191* (10,374,707-10,374,923 bp in the ensemble coordinate) were amplified from the genomic DNA using a KAPA2G Fast Multiplex PCR Kit (KAPA Biosystems, Woburn, MA, USA) with the forward primers labeled with HEX or FAM and reverse primers tailed with GTTTCTT at the 5'-end to increase the accuracy of fragment analysis (*90*), as described previously (*91*).The amplified fragments were analyzed in the genotyping center of BEX Co. Ltd (Tokyo, Japan). Next, genomic PCR and capillary sequencing were conducted on F_2 hybrids (125 males and 57 females) as described above. Using the fluorescence intensity of Sanger sequencing, we estimated the copy number of each haplotype in each fish.

Then, we tested whether the total copy number of *Fads2* gene influenced the survival rates at 40, 50, and 60 days after fertilization, and the life duration. Logistic regression analyses and Pearson's product-moment correlation tests were conducted using R version 3.3.1 (*73*). The partial Nagelkerke R^2 was calculated by the rsq package (*92*). Sex was included as a covariate.

Fosmid library

For fluorescence in situ hybridization (FISH) and estimation of *Fads2* duplication timing (see below), genomic fosmid libraries of a Pacific Ocean male from Akkeshi and a Japan Sea male from a pond in Nemuro, Hokkaido, Japan, were constructed using the Copy Control Fosmid Library Production Kit (Epicentre, Madison, WI). Fosmid clones containing the *Fads2* gene were screened by qPCR.

The fosmid clones containing *Fads2* were sequenced with the shotgun method. In brief, fosmid DNA was extracted using the automated genomic DNA isolation system PI-200 (KURABO, Osaka, Japan) and was mechanically sheared into 1.5-4.0 kb fragments using the

HydroShear apparatus (GeneMachines, San Carlos, CA). The DNA fragments were inserted into the pCR-Blunt II-TOPO plasmid vector in the Zero Blunt TOPO PCR Cloning Kit (Invitrogen, Carlsbad, CA). Sequencing of the shotgun libraries was performed using the BigDye Terminator Cycle Sequencing Kit version 3 (Applied Biosystems, Waltham, MA) and ABI 3730xl capillary sequencers (Applied Biosystems, Waltham, MA). After assembly of all sequence data using KB basecaller/Phrap/Consed systems, gaps and low-quality regions in the initial assembly were closed and re-sequenced by primer walking and direct sequencing. These complete sequences are available from DDBJ under the accession numbers LC431555 (Pacific Ocean fish) and LC431556 (Japan Sea fish).

Fluorescence *in situ* hybridization analysis (FISH)

FISH was conducted using Pacific Ocean males and Japan Sea males as described previously (*93-96*). In brief, metaphase spreads were made from spleen and liver tissues from colchicine-injected fish. For FISH in the Japan Sea male, a BAC clone from LG19 (CHORI STB 213 188J19) was labelled with Alexa-568 (red), while a fosmid clone containing *Fads2* (P381) was labelled with Alexa-488 (green). For FISH in the Pacific Ocean male, BAC clones from LG19 (STB 213 101E08) and LG12 (STB 213 140B10) were labelled with Alexa-488 (green), while a fosmid clone containing *Fads2* (P381) was labelled with Alexa-568 (red). Images were obtained using a Nikon Eclipse 80i microscope (Nikon, Japan) with Chroma filters 31000v2 (DAPI), 41001 (FITC), and 41004 (Texas Red; Chroma, USA). Images were pseudo-colored with the NIS Elements imaging software (Nikon, Japan).

Quantitative trait loci (QTL) analyses of copy number variations in Pacific Ocean and Japan Sea sticklebacks

For QTL analysis of *Fads2* copy number variation and linkage mapping of *Fads2* haplotypes, we used a previously genotyped F_2 intercross derived from a Pacific Ocean female and a Japan Sea male, which is described in Kitano et al. (2009) (*12*). The relative copy numbers of *Fads2* in F_2 hybrids were determined by genomic qPCR as described above. To determine the *Fads2* haplotypes of the Pacific Ocean grandmother and the Japan Sea grandfather, we conducted PCR for *Fads2* using Takara Ex Taq following the manufacturer's instructions (Takara Bio, Shiga, Japan). The sequences of the primers are as follows: forward primer, 5′-ACCCAGGTGTCAAAGTCTGC-3′; reverse primer, 5′-TGTACACCTTCCGATCGATCAC-3′. The nucleotide sequences were determined with a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and a DNA sequencer ABI PRISM 3130 (Applied Biosystems). The Japan Sea grandfather possessed a SNP specific to Japan Sea fish at 57 bp after the start codon (C for Japan Sea and T for Pacific Ocean sticklebacks). In addition, the Pacific Ocean grandmother possessed at least one EVK haplotype of *Fads2*, which can be distinguished by the heterozygote of C/G at 37 bp after the start codon. We genotyped these SNPs in an F_2 intercross by PCR and subsequent capillary sequencing to investigate the linkage group with which these SNPs co-segregate. The R/qtl package was used for mapping QTL as described in Kitano et al. (2009) (*12*). The genome-wide thresholds of significant LOD scores ($P < 0.05$) were calculated by 1000 permutation tests.

Gene synteny analysis

To investigate the synteny around *Fads2* on LG12 and LG19, we used two teleost genomes, Medaka HdrR and Zebrafish GRCz10 reference genome sequences, in the Ensembl database

(http://www.ensembl.org). Because the *Fads2* gene is located on chromosome 6 of medaka and chromosome 25 of zebrafish, 0.2-0.4 Mb of sequences around *Fads2* were obtained from the database and compared with the stickleback BROADS S1 reference genome sequence.

Local genomic similarity analysis of *Fads2* regions

Local genomic similarity of the *Fads2* regions between LG12 and LG19 was analyzed using Yet Another Similarity Searcher (YASS) web server (http://bioinfo.lifl.fr/yass/yass.php) (*97*) with the following parameters: scoring matrix (match, transversion, transition, other) = +5, -4, -3, -4; gap costs (opening, extension) = -16, -4; E-value threshold = 10, X-drop threshold = 30. About 20,000 bp of non-repeat masked sequences around *Fads2* genes on LG12 and LG19 were obtained from the stickleback reference genome and analyzed. By comparing the sequence against itself on YASS, repetitive sequences were detected and the identity of the repetitive sequences was determined by screening them against known repeats in Repbase with CENSOR software (http://www.girinst.org/censor/index.php) (*98*).

Linkage analysis of the *Fads2* locus and measurement of *Fads2* relative copy number in the blackspotted stickleback *G. wheatlandi*.

For linkage analysis in *G. wheatlandi*, we used three pure crosses made in 2003 using fish from Wells, Maine, USA and three pure crosses made in 2007 using fish from Demarest Lloyd State Park, Massachusetts, USA. To find nucleotide differences in *Fads2* between parents, we obtained fragments of intron 6 in *Fads2* by PCR with KOD FX Neo following the manufacturer's instructions (Toyobo Biochemicals, Osaka, Japan). The primer sequences were as follows: forward primer, 5'-ACCACCAACACCAGTACTTCTTCCT-3'; and reverse primer, 5'-TGACCTGAGAAAAAGGGAAGA-3'. The nucleotide sequences were determined using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and a DNA sequencer ABI PRISM 3130 (Applied Biosystems). Three of the six crosses were used because their parents possessed a SNP (A/G) in intron 6, which could be used for linkage analysis. To investigate the linkage between the *Fads2* SNP and microsatellite markers located on threespine LG12 and LG19, 6 primer sets (LG12: *Stn287*, *Stn276*, *Pun2*; LG19: *Stn235*, *Stn284*, *Cyp19b1*) (*99-101*) labelled with HEX, NED, or FAM were used to genotype the crosses. Three differently colored primer sets were combined and the microsatellite loci were amplified from the genomic DNA of two crosses using fish from Massachusetts (41 individuals in total) by a KAPA2G Fast Multiplex PCR Kit (KAPA Biosystems, Woburn, MA, USA). The amplified fragments were analyzed at the genotyping center of BEX Co., Ltd. (Tokyo, Japan). The *Fads2* SNP was also genotyped by sequencing the intron 6 of *Fads2* as described above. We used Fisher's exact test in R (*102*) for a statistical test of correlation between *Fads2* SNP and microsatellite markers on LG12 and LG19.

The relative copy number of *Fads2* of *G. wheatlandi* was investigated by genomic qPCR as described above. Based on the previously determined WGS of *G. wheatlandi* (*71*), we designed forward and reverse primers and a TaqMan probe for *G. wheatlandi Fads2* and *TSHβ2* using Primer Express Software v3.0 (Applied Biosystems) as follows: *Fads2* forward primer, 5′-GAGACTGATCTCACGCTTTTGTTC-3′; *Fads2* reverse primer, 5′-TCTGAAGCCTCTGCTGATTGG-3′; *Fads2* TaqMan probe, [FAM]-TGGGCTCTGTCGCTGCCAGCTC-[TAM]; *TSHβ2* forward primer, 5′-CGGGAGGGAATTAATGAAACAG-3′; *TSHβ2* reverse primer, 5′-GGAAATGGACGGGTCGGT-3′; and *TSHβ2* TaqMan probe, [FAM]- GCCTGTCTCCCTGGT-[TAM].

Sequence comparison around the *Fads2* insertion site on LG12 between Pacific Ocean stickleback, Japan Sea stickleback, and blackspotted sticklebacks (*G. wheatlandi*)

To sequence the genomic regions of Japan Sea and blackspotted sticklebacks where *Fads2* has been inserted on LG12 of the Pacific Ocean lineage, PCR was conducted with KAPA2G Fast Multiplex Mix (Sigma-Aldrich, St. Louis, MO, USA) using the following primer set: F1 primer, 5′-CTTTGATGCTCAGACTTTGTACATATGATC-3′, and R1 primer, 5′-GGCAGCGGATATTCACAAGA-3′. The PCR products were electrophoresed on E-Gel EX agarose gels, 2% (Thermo Fisher Scientific Inc.) and detected with FAS-IV (Nippon Genetics, Tokyo, Japan). Since PCR failed in the Ogawara population, possibly because of the nucleotide changes at the primer sites, additional primers were designed at nearby sites as follows: F2 primer, 5′-AATGACACTGAACAGATGGGAATTTCTG-3′; R2 primer, 5′-TAAATATTACCGATATATGGTTAATAGTAC-3′; R3 primer, 5′-CAAATTACCAACCAGCAATTTACCA-3′; R4 primer, 5′-GTGTAGTACAGAAGACTGACCAC-3′; and R5 primer, 5′-TCCATGTTGAAACCACGGATGG-3′. PCR was conducted again using PrimeSTAR GXL DNA Polymerase (Takara Bio, Shiga, Japan). For *G. wheatlandi*, F2 and R1 primers were used for PCR, since the F1 primer sequence contained three SNPs. All PCR products were purified with a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA) or illustra ExoProStar (GE Healthcare Life Sciences, Chicago, IL, USA), followed by capillary sequencing at Eurofins Genomics (Tokyo, Japan). Neighbour-joining trees of nucleotide sequences were constructed using the ClustalX programme (*103*). They were bootstrapped 1000 times to estimate confidence in the nodes. Local genomic similarity was analyzed by YASS as described in the previous section.

Estimation of duplication timing of *Fads2* genes

For estimation of duplication timing, we compared *Fads2* genomic sequences of *G. nipponicus* and *G. aculeatus*. Fosmid sequences of Japan Sea *Fads2* (clone ID: J8-14-13; see above) and reference genome sequences of *G. aculeatus* (BROADS S1) were used in the estimation. Four sequences were aligned using the Muscle algorithm (*104*) implemented in MEGA 7.0 software (*105*). In total, 2270 nucleotide sites, including 1138 intron sites, were used for the divergence time estimation. The estimation was performed using the BEAST 2 software (*106*). The regions were partitioned into three codon positions and introns, and we used 7.1×10^{-9} for the mutation rate in introns, which was estimated using synonymous site divergence between ninespine and threespine sticklebacks (*107*). The HKY nucleotide substitution model was assumed, and the rate variation of substitution was modelled as the Gamma distribution with an estimated shape parameter. The posterior distribution of each parameter was generated by 45,000,000 iterations after a burn-in step of 5,000,000.

QTL analyses of survivorship with DHA-deficient diets

To investigate the genetic architecture underlying survivorship with DHA-deficient diets, we conducted QTL mapping of survival rates at 50 days after fertilization using the F_2 intercross resulting from a cross between a Japanese Pacific Ocean female and a Japan Sea male. For making a linkage map, the grandparents and three hundred and twenty-two F_2 fish were genotyped with ddRAD-seq. ddRAD-seq was conducted as described in the section of *Phylogenetic analysis of Japanese sticklebacks* except that we used a HiSeqX with a 150 bp-

paired end mode for sequencing. Short reads were trimmed with Trimmomatic: HEADCROP, 1; CROP, 140; LEADING, 3; TRAILING, 3: SLIDINGWINDOW, 4:15; MINLEN, 140. Trimmed reads were mapped to the BROADS1 reference sequence using Burrows-Wheeler Aligner (*108*). Mapped reads were analyzed with a Stacks pipeline (*109*) to produce a file for linkage mapping. Soft-clipped reads were discarded, and the minimum coverage for calling homozygous SNPs was set to 20x. Next, JoinMap 4 (*110*) was used to construct a linkage map using a regression model. Markers that showed segregation distortion (χ^2-test, $P < 0.001$, d.f. = 2), were co-segregating (similarity > 0.99), or had > 30% missing genotypes were removed. Linkage groups were identified with the LOD threshold of 10. Possible erroneous genotypes (*Genotype Probabilities* calculated by JoinMap 4 < 0.001) were removed. In total, we used 1689 loci.

We first conducted a single-QTL scan of survival rates at 50 days after fertilization using the R/qtl package as described above. Sex was included as an additive covariate. The thresholds of significant LOD scores ($P < 0.05$) were calculated by 1000 genome-wide permutation tests. QTLs with LOD scores greater than the genome-wide thresholds were considered as significant QTLs. QTLs with LOD > 3 were considered as suggestive QTLs (*111*). In addition to the survival rate, we conducted QTL mapping of the *Fads2* copy number. We calculated the 95% Bayesian credible intervals to test whether these QTLs overlap with *Fads2* gene. The same cross was used for testing the association between *Fads2* copy number and survivorship (see above). The relative copy numbers of *Fads2* were determined as described above. The PVE of four QTLs for the survival rates were calculated using the *fitqtl* and *refineqtl* function of R/qtl. Based on the estimated copy number of each *Fads2* haplotype and the linkage map of F_2 intercrosses, it turned out that the Pacific Ocean grandmother possessed three *Fads2* copies; one PO QVN on LG12, one PO QVN on LG19, and one PO EVK on LG19, while the Japan Sea grandfather possess only one *Fads2* copy (JS QVN) on LG19 (X chromosome).

A two-dimensional QTL mapping of the survival rate at 50 days after fertilization was also conducted using the R/qtl package. Sex was included as the additive covariate. The genome-wide thresholds of significant LOD scores ($P < 0.05$) were calculated by 1000 permutation tests. No significant interaction was found.

Comparison of relative copy number of *Fads2* using genomic qPCR

The relative copy numbers of *Fads2* of multiple Japan Sea, Pacific Ocean, and freshwater populations were determined by genomic qPCR as described above. For designing primers and probes for ninespine sticklebacks, we used a previously determined WGS of *Pungitius pungitius* (*112*) (DRA001085) and additional WGS of *P. sinensis* and *P. tymensis* collected from eastern Hokkaido, Japan, determined using HiSeq 2000 (DRA006706), as described previously (*112*). WGS data were mapped to the BROADS S1 stickleback genome sequence using CLC Genomics with the following parameters: deletion cost = 3, insertion cost = 3, length fraction = 0.5, mismatch cost = 2, and similarity fraction = 0.8. We designed forward and reverse primers and a TaqMan probe for the regions conserved among species with Primer Express Software v3.0 (Applied Biosystems): *Fads2* forward primer, 5′-TTGGAGAGCCACGGTTTTG-3′; *Fads2* reverse primer, 5′-CAGCCAGTCCTGGTGCTTCT-3′; *Fads2* probe, [FAM]-AATCATCTGCCCATGAACATCGACCA-[TAM]; *TSHβ2* forward primer, 5′-CACTAATTTGAAGGGACGTTTCG-3′; *TSHβ2* reverse primer, 5′-GGATAATACATCTGCGGGTTGAC-3′; and *TSHβ2* probe, [FAM]- CAGCTGTGTGCCTCT -[TAM]. We used linear mixed-effects regression (lmer) with the lme4 package in R (*102*) for

statistical analysis of copy number differences between the Japan Sea, Pacific Ocean, and freshwater threespine populations or between the ninespine stickleback species. For threespine sticklebacks, the statistical analysis was conducted for each geographic region (Japan, North America, and Europe). The fixed effects in the model were sex and either ecotype or species (among Japan Sea, Pacific Ocean, and freshwater populations, between marine and freshwater populations, or among the three ninespine stickleback species). Within the freshwater populations, we used Shiga, Gifu, Nasu, Chimikeppu, Aizu, Kinugawa and Ono as the old freshwater population, and Towada, Aisaka, Gensui, Shikotsu, and Nishikitappu as the young freshwater population, as based on the phylogenetic analysis (Fig. 1C and S3). The random effects were the populations. *P*-values were obtained by likelihood ratio tests of the full model against the model without the predictor ecotype or species. A Tukey's post-hoc test was conducted using the lsmeans package in R (*113*).

Estimation of *Fads2* copy number in freshwater populations from whole genome resequencing

To estimate the copy number of *Fads2* genes of freshwater populations, we conducted WGS of 5 females and 3 males from Nishikitappu, 6 females and 2 males from Chimikeppu, 3 females and 8 males from Gifu, 3 females and 6 males from Ono, and 2 females from Aizu. All genomic DNA was isolated with Qiagen DNeasy Blood & Tissue Kit. For WGS of three Gifu females and one Aizu female, genomic libraries were constructed with Illumina TruSeq DNA Prep Kit and run in a 100 bp-paired end mode of HiSeq 2000 (one fish per lane). For other fish, libraries were constructed with NEBNext Ultra DNA Prep Kit (NEB, Ipswitch, MA) and run in a 150 bp-paired end mode of HiSeqX (one Aizu fish per lane and other fishes per three lanes in total). Trimming and mapping to the reference sequence were conducted as described previously (*71*). The sequence reads were also mapped to the *Fads2* locus (ENSGACG00000005827) from the start codon to the last exon including introns between exons. To calculate the copy number, the coverage of the *Fads2* locus was divided by the genome-wide coverage for each individual. These data are available from DDBJ (DRA007515).

Long-read sequencing of *Fads2* loci in freshwater populations from Gifu and Little Campbell River streams

For PacBio sequencing, genomic DNA was extracted from the muscles of a male from the freshwater population in Gifu, Japan, and a female from the upstream of Little Campbell River, Canada, using the InnuPure C16 touch system (Analytik Jena AG, Jena. Germany) following the manufacturer's protocol. After size selection at 25 kb with BluePippin, a 30-kb library was constructed and analyzed on two SMRT cells in PacBio Sequel (Pacific Biosciences). This reaction generated 672,124 raw reads with mean read length of 10,205.28 bases for Gifu (available at DRA007518), and 88,964 raw reads with mean read length of 13,121.98 bases for Little Campbell River Stream (available at DRA007519).

For MinION Nanopore sequencing, genomic DNA was obtained from the trunk muscle of two freshwater males from Gifu with phenol-chloroform extraction. The DNA was treated by ribonuclease A (Qiagen, Hilden, Germany) and purified by Agencourt AMPure XP (Beckman Coulter, Bream, CA, USA). The quality of the genomic DNA was assessed by electrophoresis on a 0.6% agarose gel made with SeaPlaque GTG Agarose (Lonza, Basel, Switzerland) and NanoDrop 2000 (Thermo Fisher Scientific Inc.). Libraries were constructed with a Rapid sequencing kit, SQK-RA0002 (Oxford Nanopore Technologies, Oxford, UK), and run on a MinION MK 1B (Oxford Nanopore Technologies), according to the manufacturer's instructions.

We used one flow cell for one male and two flow cells for another male and obtained 423,105 raw reads with mean read length of 4,182.12 bases in total (available at DRA007517).

From the sequence reads generated by PacBio and MinION Nanopore sequencers, we conducted a BLAST search to find the reads containing *Fads2* using CLC Genomics Software 6.5. *Fads2* cDNA sequences in the stickleback reference genome (ENSGACG00000005442, ENSGACG00000005827, and ENSGACG00000018692) were used as query sequences. The following parameters were used: number of threads = 1, expect = 10, word size = 11, match = 2, mismatch = -2, and gap cost = Existence 5, Extension 2. Pairwise alignments were conducted between the reads containing *Fads2* and the non-repeat masked reference stickleback sequences around *Fads2* on LG12 and LG19 using YASS (*97*). The identity of repetitive sequences was determined by searching against a database of known repetitive sequences, the GIRI Repbase in CENSOR software (http://www.girinst.org/censor/index.php) (*98*).

Analysis of *Fads2* copy number in the ray-finned fishes

To obtain *Fads2* copy number data from multiple actinopterygian fishes, we used ORTHOSCOPE v.1.0.1 (*114*). ORTHOSCOPE is a recently developed web tool for identifying sets of orthologs from publicly available databases of protein-coding genes constructed from whole genome sequence data (*114*). ORTHOSCOPE only collects the longest transcript from a single locus when multiple transcripts, such as alternative splice variants, exist. *Fads2* sequences from four species (*Cyprinus carpio* XP018950032.1; *Salmo salar* XP014025867.1, *Gasterosteus aculeatus* ENSGACT00000007227.1, and *Oryzias latipes* ENSORLT00000007983.1) were used as a query (see arrows in Fig S28). In addition to 48 actinopterygian (ray-finned) fishes, gene models of non-actinopterygian vertebrates (*Callorhinchus milii*, *Rhincodon typus*, *Gallus gallus*, and *Homo sapiens*) were also searched in order to delineate an actinopterygian *Fads2* lineage. The following parameters were used: Mode = Search/Rearrangement; Focal group = Actinopterygii; E-value threshold = 1e-3; Number of hits to report per genome = 5; Dataset = DNA (Exclude 3rd). ORTHOSCOPE automatically generates a neighbor-joining tree of collected candidate orthologs and closely related genes. Using this tree, we selected *Fads2* orthologs based on the tree topology (see (*114*) for details) and counted the number of *Fads2* genes for each of 48 actinopterygian species. In total, we could obtain the data from 48 actinopterygian species (Fig. S29).

Finally, to further validate the phylogenetic relationships of the collected *Fads2* orthologs, we created a maximum likelihood tree using the *Fads2* orthologs collected above (Fig. S28). Protein sequences were aligned using MAFFT (*115*). The aligned protein sequences were trimmed by removing poorly aligned regions using TRIMAL 1.2 (*116*) with the option "gappyout." Corresponding coding sequences were aligned using PAL2NAL (*117*) with reference to the protein alignment and used for estimating the *Fads2* gene tree. The maximum likelihood tree was estimated using RAxML 8.2.6 (*80*) with the GTRGAMMA model, a general time-reversible model with gamma-distributed rate variations among sites (*118, 119*). We used the first and second codons (890 bp in total) and partitioned them.

Statistical analyses for association between *Fads2* copy number and habitat in the ray-finned fishes

To test the association between freshwater colonization and *Fads2* copy number, we used Bayesian inference for generalized linear mixed models (GLMM) that account for phylogeny as a covariate (*120*) using the MCMCglmm R package (*121*). We used two phylogenetic trees. One

is a previously published phylogenetic tree (*122*), which contained only 30 species whose *Fads2* copy numbers were estimated. To increase the statistical power, we also created a species tree with additional actinopterygians. To the tree of Betancur-R et al. (2017), we first added the topology and divergence time from published data on Salmoniformes (*123*), Carangiformes (*124*), Beloniformes (*125*), and Cyprinodontiformes (*126*). Data from these publications are indicated in Fig. S29. Divergence times which were not previously reported were estimated using whole mitochondrial genome sequences and can be found at four nodes marked with mtGenome in Fig. S29 (nodes between *S. grahami* and *C. carpio*, between *A. limnaeus* and *N. furzeri*, between *X. maculatus,* and between *P. reticulata*, and *P. formosa*). To estimate divergence time, we first obtained whole mitochondrial sequence data of 46 species from the NCBI database (https://www.ncbi.nlm.nih.gov/genome/browse/?report=5#!/overview/). For the four species whose sequences were not available, we used the sequence data of closely related species: *Paramormyrops gabonensis* instead of *Paramormyrops kingsleyae* (Mormyridae), *Astyanax paranae* instead of *Astyanax mexicanus* (Characidae), *Pseudolabrus eoethinus* instead of *Labrus bergylta* (Labridae), and *Austrolebias charrua* instead of *Austrofundulus limnaeus* (Cyprinodontiformes). The use of closely related species at the tip of the tree is not a problem for inferring divergence time at deeper nodes. Whole mitochondrial sequences were not available for *Stegastes partitus*, *Acanthochromis polyacanthus,* or their close relatives, so these two species were not used for time calibration. From the whole mitochondrial genome sequences, we obtained 13 protein-coding genes without 3rd codon positions (7558 bp in total) for the analysis. To estimate divergence time, we used the RelTime method (*127*) implemented in MEGA v.7.0 (*105*). Previously published divergence times (see above) were used as calibration constraints. We set five different evolutionary rates for each codon in a GTR model with a discrete Gamma distribution. All positions containing gaps and missing data were eliminated.

For the phylogenetic GLMM analysis, the *Fads2* copy numbers estimated above were used as a response variable. Habitat information was compiled from the Catalog of Fishes (*128*) and used as the predictor in the model. As for the habitat information, we used two models: a two-category model (species that have freshwater populations and those that lack freshwater populations, Fig. 4C) and a five-category model (1. marine only, 2. both marine and brackish, 3. marine, brackish, and freshwater, 4. both brackish and freshwater, and 5. freshwater only, Fig. S29). In addition, we tested models with and without whole genome duplication events (*129*) (Fig. S29) as covariates. This is because whole genome duplication can increase *Fads2* copy numbers independently of freshwater colonization. Thus, we tested four combinations of models.

Using the species tree that includes all 48 actinopterygians as a covariate, we first calculated DIC (deviance information criterion) for comparing the four models. An inverse Wishart prior (parameters V = 1 and nu = 0.002) was used. All models were run for 1050000 iterations with a burn-in of 50000 iterations and a thinning interval of 200. Comparison of DIC showed that the two-category habitat model as the predictor with whole genome duplication included as the covariate was the best: two-category with genome duplication, DIC = 120.24; two-category without genome duplication, DIC = 128.41; five-category with genome duplication, DIC = 124.15; five-category without genome duplication, DIC = 132.15. In the main text, we showed the results with the best model, but other models gave rise to qualitatively similar results.

We also conducted the same analyses using only the 30 species included in the tree of Betancur et al. 2017 (*122*) with the two-category model including whole genome duplication as a covariate. The effect of habitat was significant even with this smaller dataset: effect of habitat,

posterior mean = -1.0728, 95% credible interval = -1.8733–-0.2902, pMCMC < 0.01; effect of genome duplication, posterior mean = 2.1364, 95% credible interval = 0.7665–3.4838, pMCMC < 0.01.

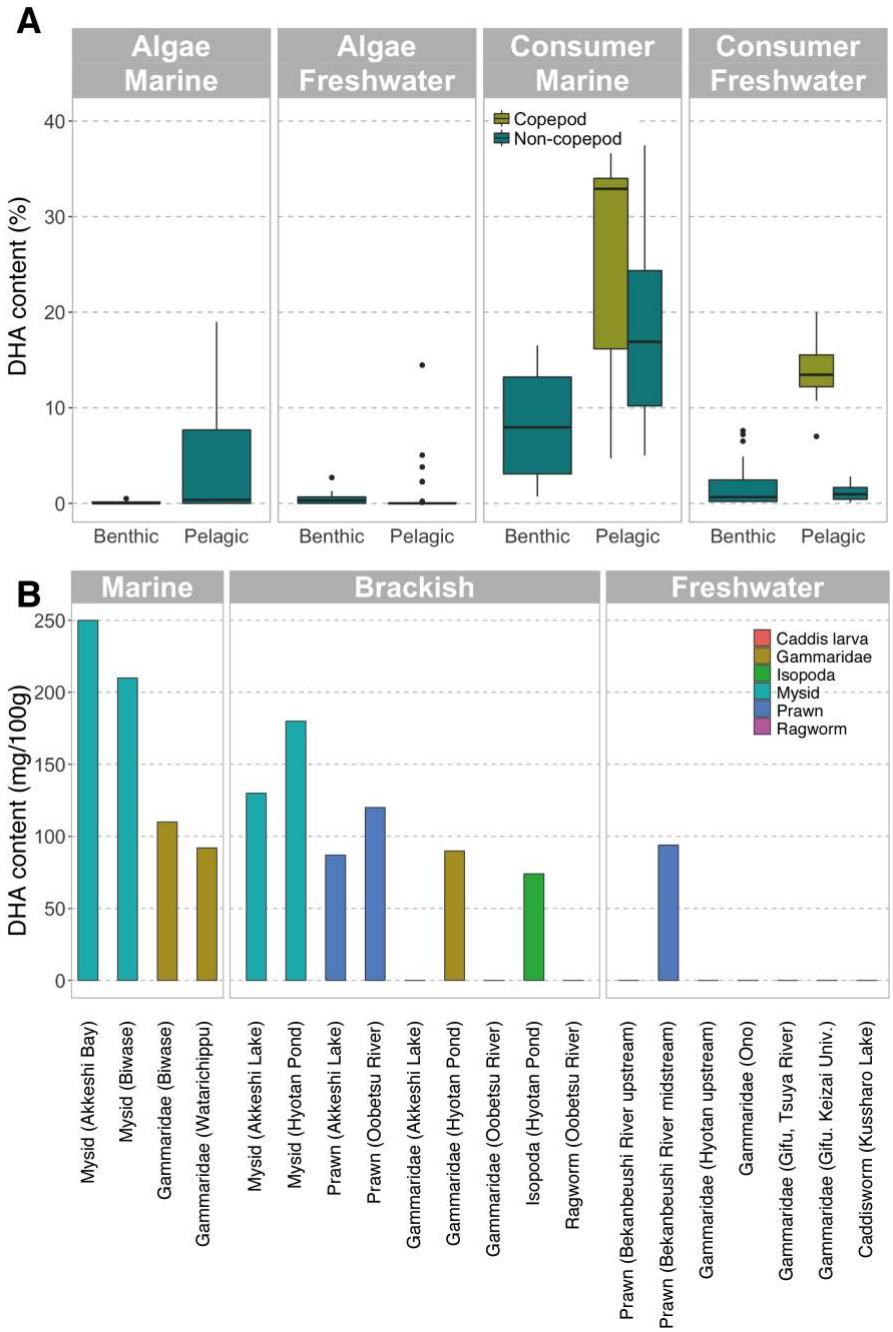

Fig. S1.

DHA contents in marine and freshwater ecosystems. (**A**) Meta-analysis of DHA contents in algae and consumers from marine and freshwater habitats. Percentage of DHA in total fatty acid were obtained from literatures. Marine ecosystems contain more DHA than freshwater ecosystems in both benthic and pelagic habitats. These differences are present at the base of the food chain (i.e. algae), but noticeable at the consumer level. Only major DHA source available in freshwater is copepods. (**B**) DHA contents in stickleback prey items collected from marine, brackish, and freshwater habitats, measured by GC-MS. Different color bars indicate different groups of prey items. None of the marine prey items analyzed were below the detection threshold (0/4), whereas 6 out of 7 freshwater prey items were undetectable (6/7) (Fisher's exact test, $P < 0.05$). In addition, although our data are based on $N = 1$ for each item, Wilcoxon rank sum test between marine and freshwater prey items showed a significance ($W = 1$, $P < 0.01$; $N = 4$ for marine and $N = 7$ for freshwater).

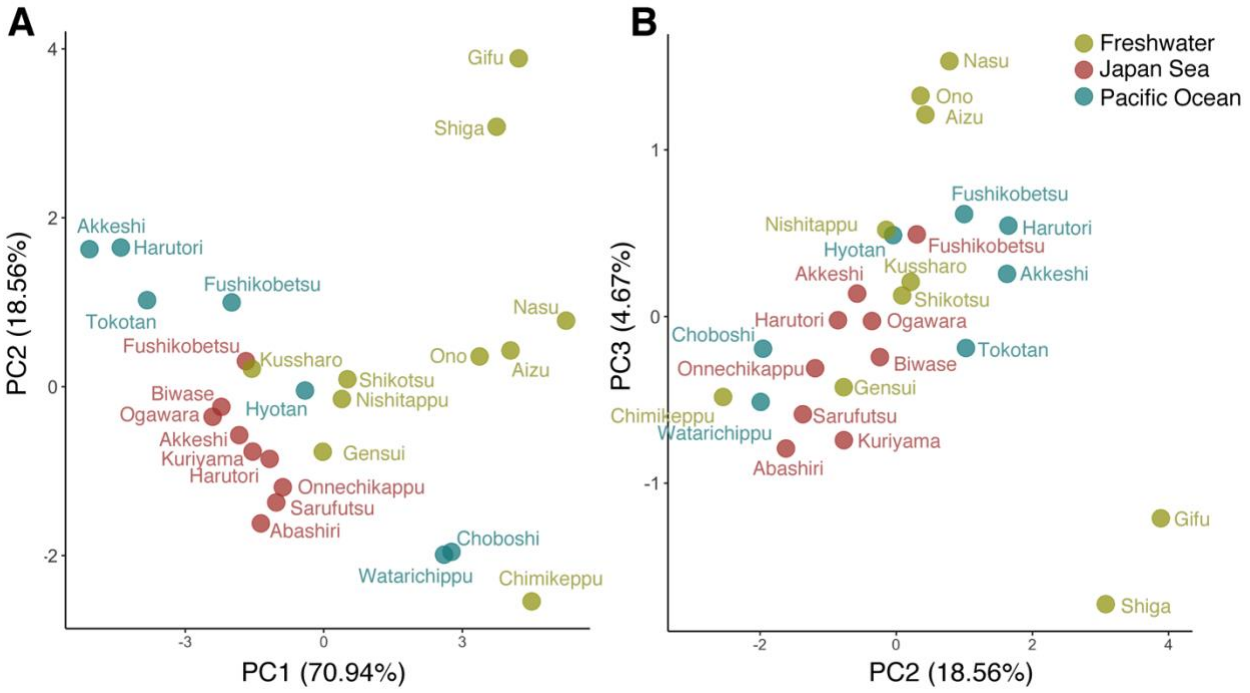

Fig. S2.

Principal component analysis (PCA) of morphological traits in Japanese *Gasterosteus* populations. (**A**) The X and Y axes indicate PC1 and PC2 of 13 morphological traits, respectively. PC1, representing overall body size, explains 70.94% of the variance. PC2, representing gill raker number, lateral plate number, and gape width, explains 18.56% of the variance. Red, blue, and yellow-green dots indicate the Japan Sea, Pacific Ocean, and freshwater populations, respectively. (**B**) The X and Y axes indicate PC2 and PC3, respectively. PC3, representing lateral plate number, gape width and dorsal and pelvic spine length, explains 4.67% of the variance. Red, blue, and yellow-green dots indicate the Japan Sea, Pacific Ocean, and freshwater populations, respectively. In both **A** and **B**, the freshwater populations show greater variation in traits than the Japan Sea and Pacific Ocean sticklebacks.

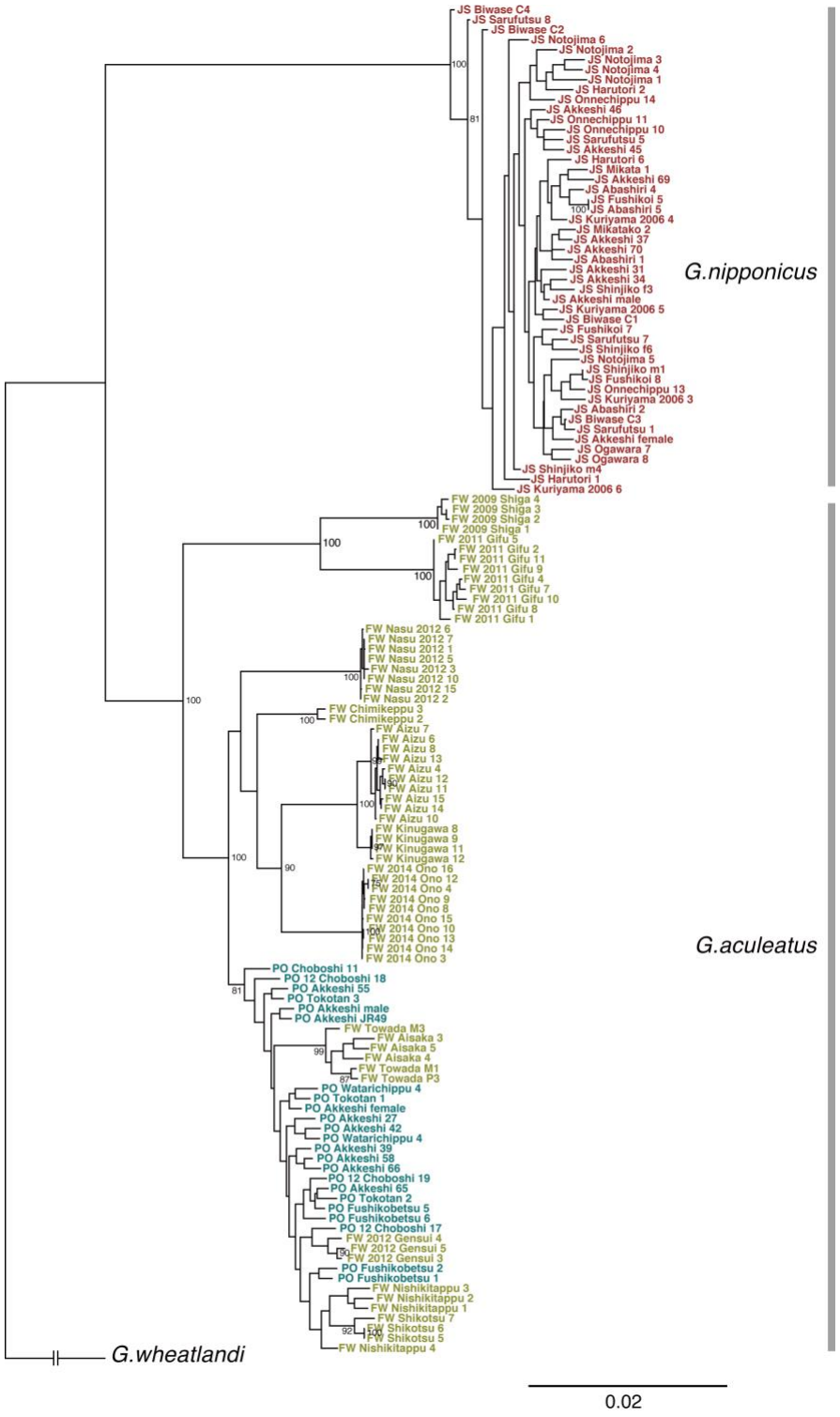

Fig. S3.

Phylogeny of Japanese *Gasterosteus* with *G. wheatlandi* as an outgroup. Each tip indicates one individual. The individual names starting with JS, PO, and FW with red, yellow-green, blue color, belong to Japan Sea, Pacific Ocean, and freshwater populations, respectively. Bootstrap values with > 70% are indicated at the nodes.

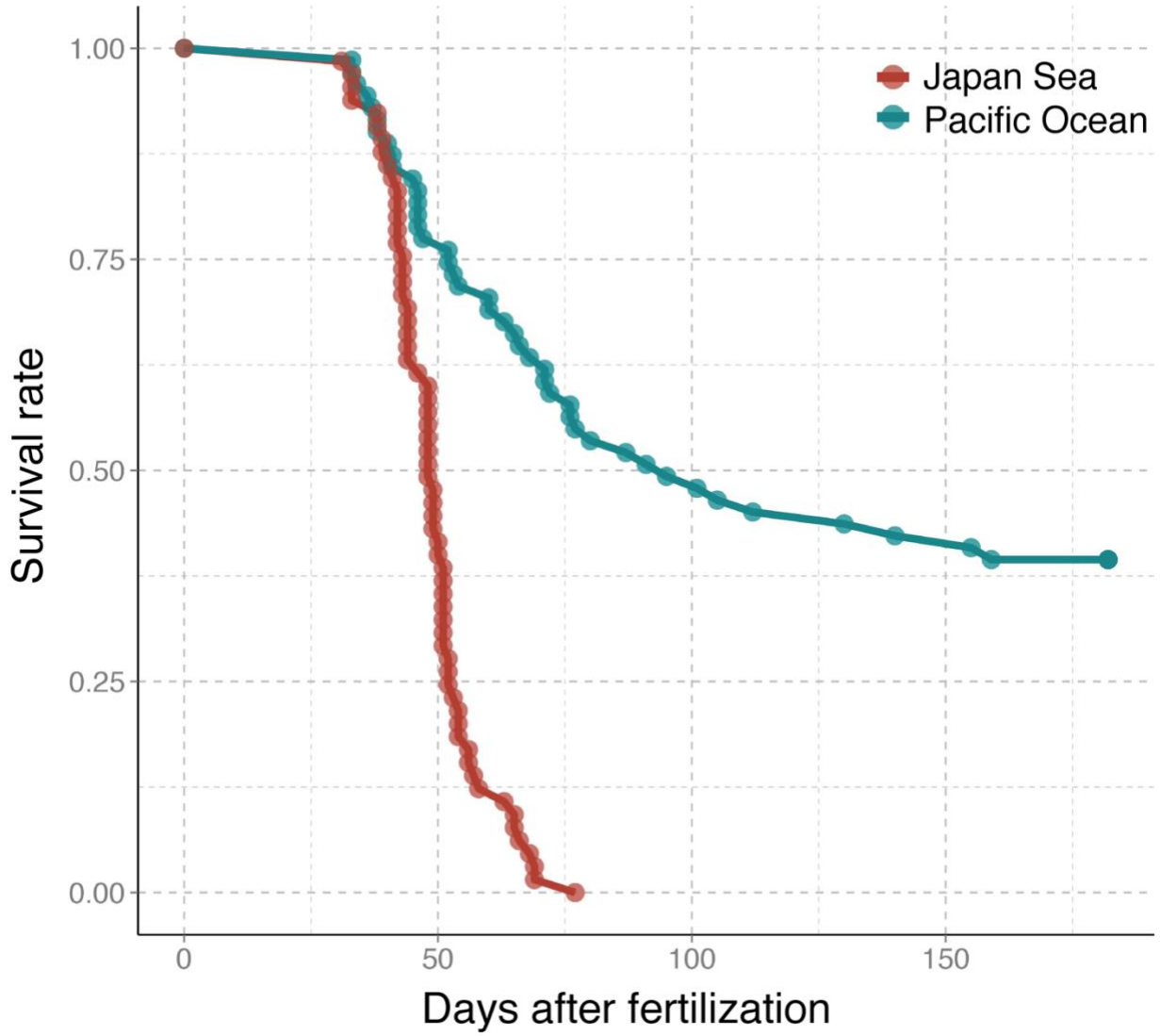

Fig. S4.

Survival curves of fish maintained in 100% seawater. Red and Blue lines indicate survival curves of Japan Sea and Pacific Ocean juveniles, respectively, fed only DHA-free *Artemia* and maintained in 100% seawater. The high level of mortality was observed in Japan Sea juveniles even when they were maintained in 100% seawater.

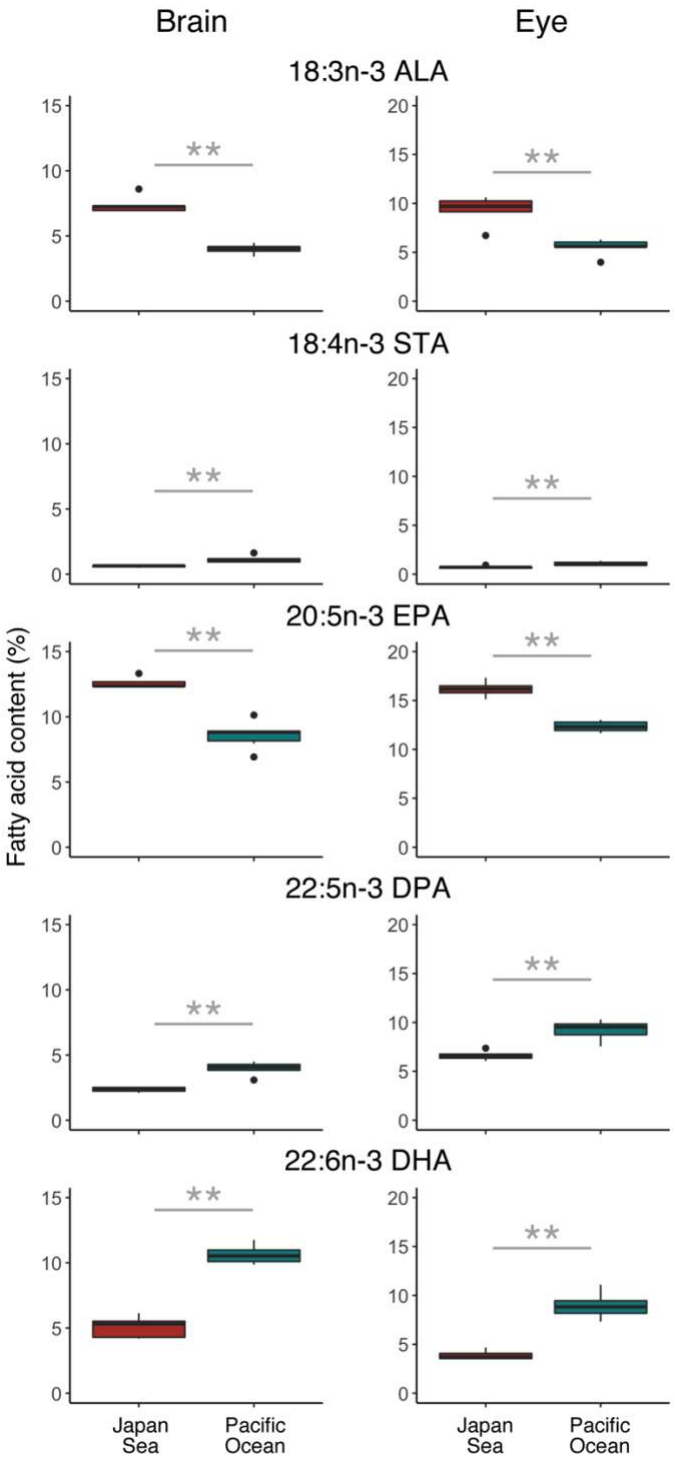

Fig. S5.

ω-3 Fatty acid content in the brains and eyeballs of Japan Sea and Pacific Ocean juveniles. α-Linoleic acid (ALA), stearidonic acid (STA), eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), and docosahexaenoic acid (DHA) were measured by GC-MS. Red and blue boxes indicate fatty acid content of Japan Sea and Pacific Ocean juveniles fed only DHA-free *Artemia* ($N = 5$ and 6, respectively). $^{**}P < 0.01$ (Student's *t*-test).

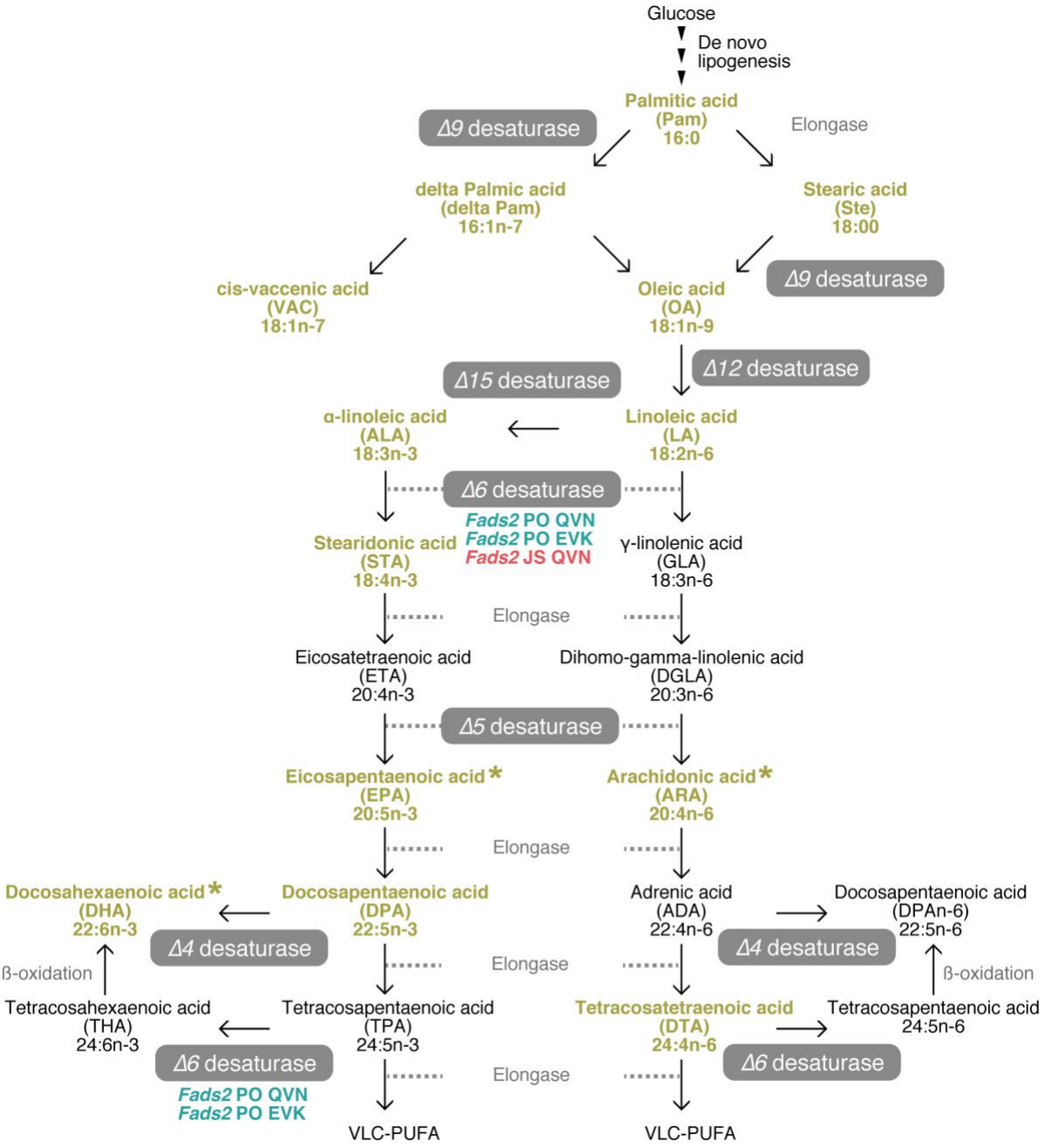

Fig. S6.

Proposed synthetic pathways of ω-3 and ω-6 polyunsaturated fatty acids (*130-132*). Fatty acids shown in yellow-green were quantified in the analysis of *Fads2*- and *GFP*-transgenic Japan Sea juveniles. Asterisks indicate fatty acids that showed significantly higher amounts in *Fads2*-transgenic Japan Sea juveniles than in *GFP*-transgenic juveniles at 6 weeks after fertilization (See Fig. S9). Our yeast experiments demonstrated that all three *Fads2* haplotypes found in Japan Sea (pink text) and Pacific Ocean stickleback (blue text) showed Δ6 desaturase activity from 18:3n-3 (ALA) to 18:4n-3 (STA), while two *Fads2* haplotypes only found in Pacific Ocean sticklebacks also showed Δ6 desaturase activity from 24:5n-3 (TPA) to 24:6n-3 (THA) (See also Table S4).

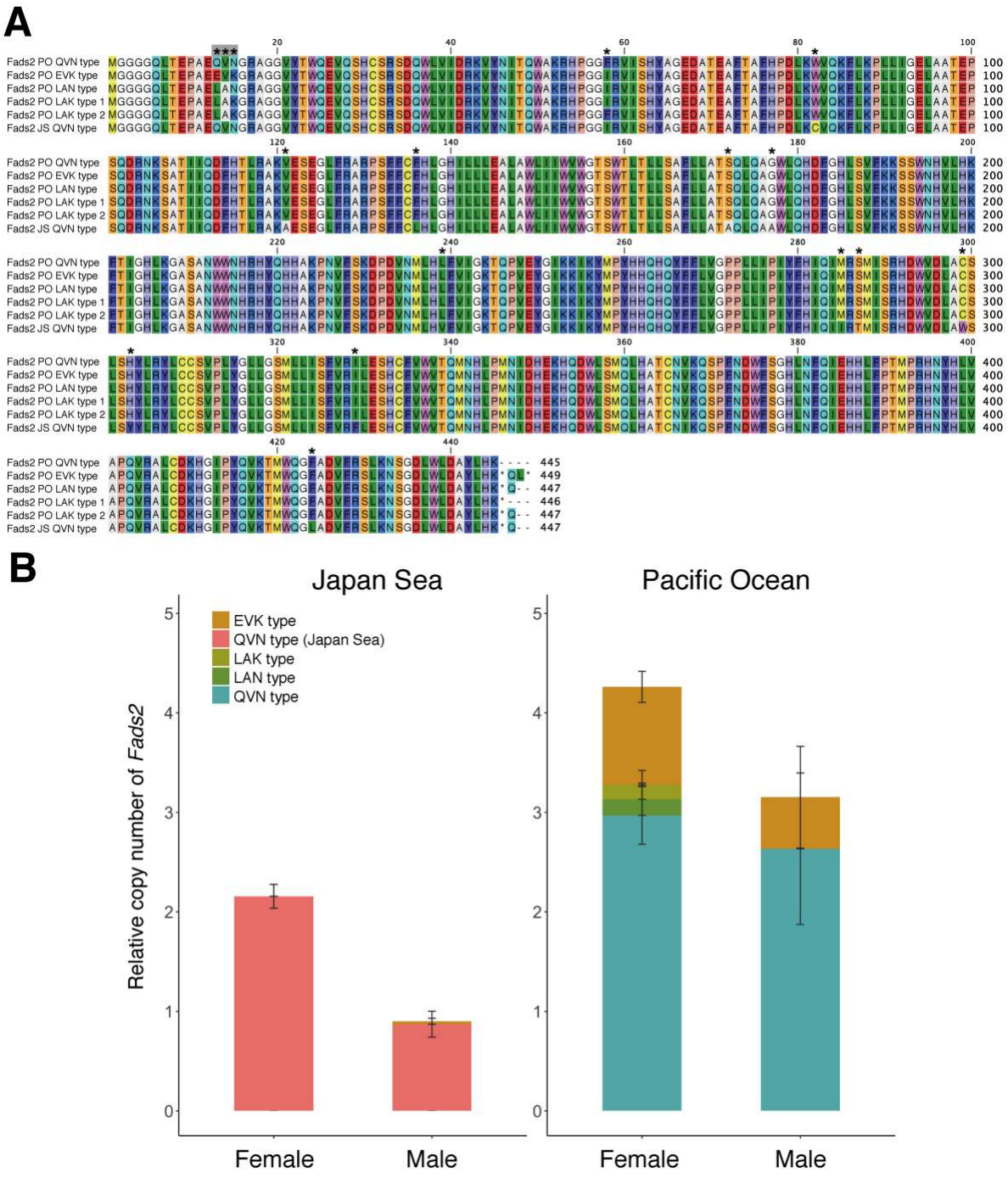

Fig. S7.

Alignment of amino acid sequences of different *Fads2* haplotypes and relative copy number of each haplotype in Japan Sea and Pacific Ocean sticklebacks. (**A**) Amino acid sequences of four *Fads2* haplotypes obtained from Pacific Ocean fish and one *Fads2* haplotype obtained from Japan Sea fish. Asterisks indicate variations in amino acids among the haplotypes. (**B**) Relative copy number of each *Fads2* haplotype in males and females of Japan Sea (5 males and 5 females) and Pacific Ocean sticklebacks (4 males and 6 females), estimated from the WGS data.

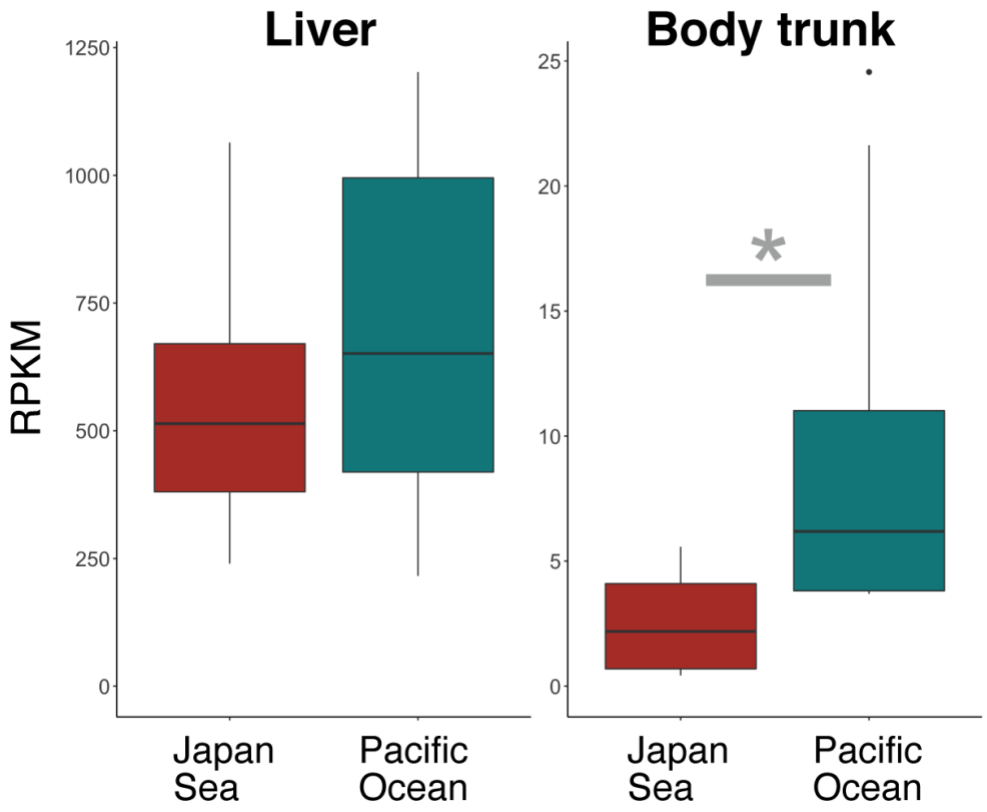

Fig. S8.

Expression levels of *Fads2* in the liver and body trunk of Japan Sea and Pacific Ocean fish at 40-60 days after fertilization. Body trunks included muscles and skins. * $P < 0.05$ (Student's *t*-test).

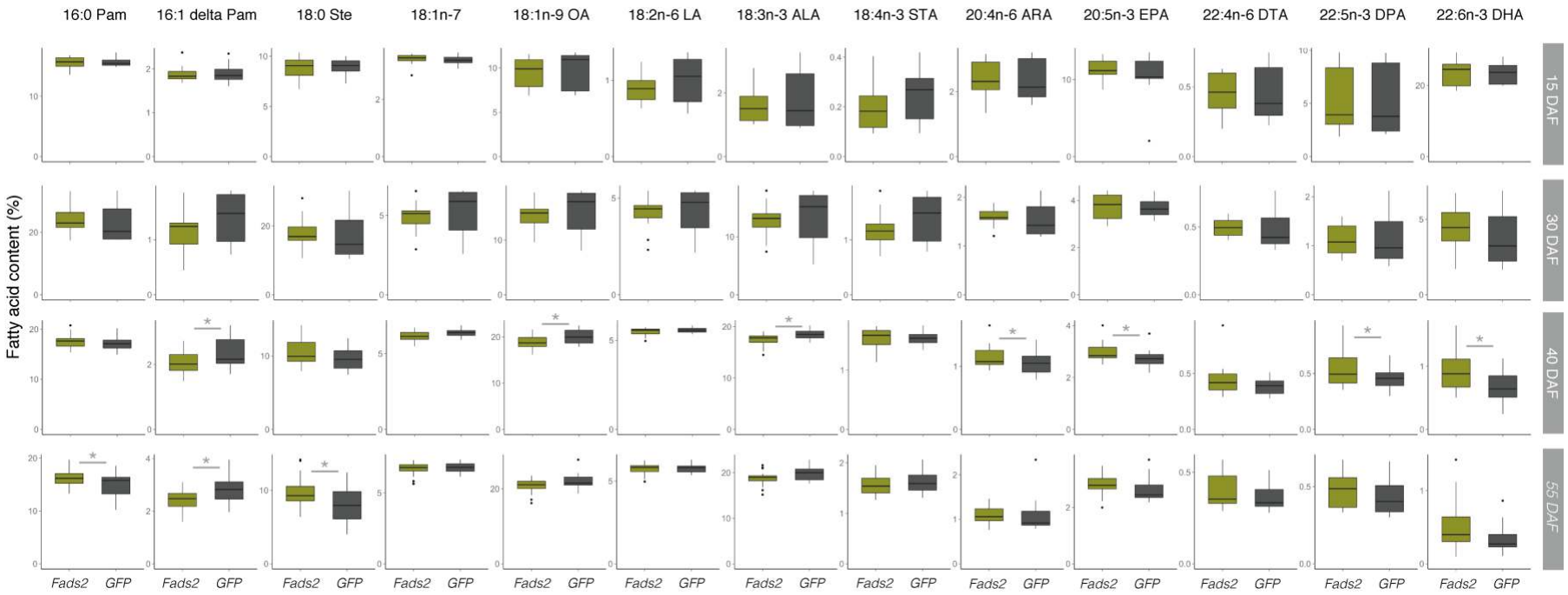

Fig. S9.

Fatty acid content of *Fads2*-transgenic and *GFP*-transgenic juveniles. Yellow-green (*Fads2*-transgenic) and grey (*GFP*-transgenic) boxes indicate fatty acid content of Japan Sea juveniles fed only DHA-free *Artemia*. *$P < 0.01$ (Student's *t*-test).

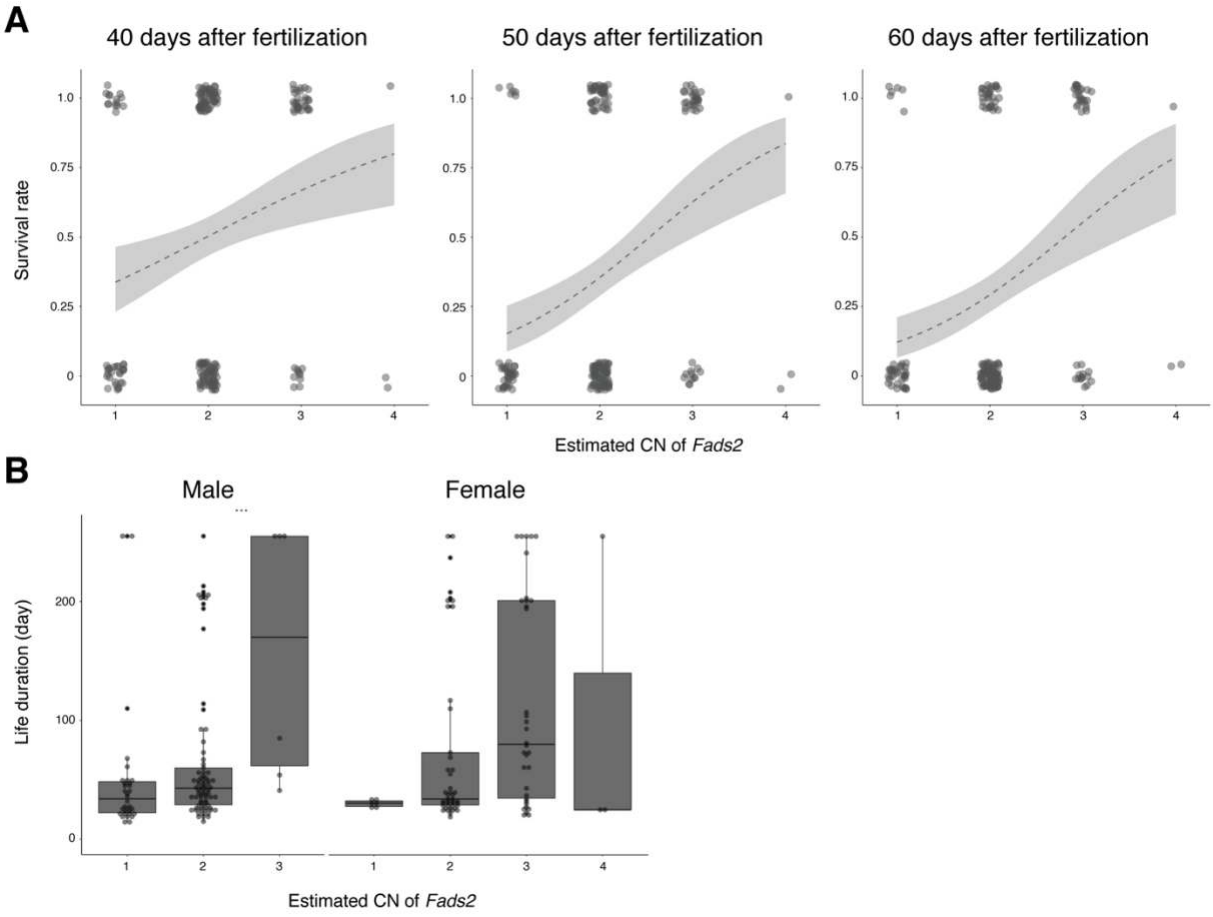

Fig. S10.

Association between *Fads2* copy number and juvenile lethality. **(A)** Association between *Fads2* copy number and survival rate at 40, 50, and 60 days after fertilization in F_2 hybrids of Japan Sea and Pacific Ocean sticklebacks. The X axis shows the estimated copy number of *Fads2*; the Y axis shows survival, with 1 indicating a survivor and 0 indicating a dead individual. The logistic regression curves indicate the probability of survival for a given copy number of *Fads2* (logistic regression; coefficients = 0.99, 1.15, and 1.05; Negelkerke $R^2 = 0.100$, 0.121, and 0.101 for 40, 50, and 60 days after fertilization, respectively; $P < 0.01$). Sex was used as a covariate in the regression model. This indicates that the percentage of variance in the survival rate at 40, 50, and 60 days after fertilization explained by the number of *Fads2* gene was 10.0, 12.1. and 10.1 %, respectively. **(B)** Correlation between the copy number of *Fads2* and the lifespan. The X axis shows the estimated copy number of *Fads2*; the Y axis shows the life duration (day). Each dot indicates a single F_2 individual. There is a positive correlation between copy number of *Fads2* and life duration (Pearson's product-moment correlation; $r = 0.1693$, $P < 0.05$).

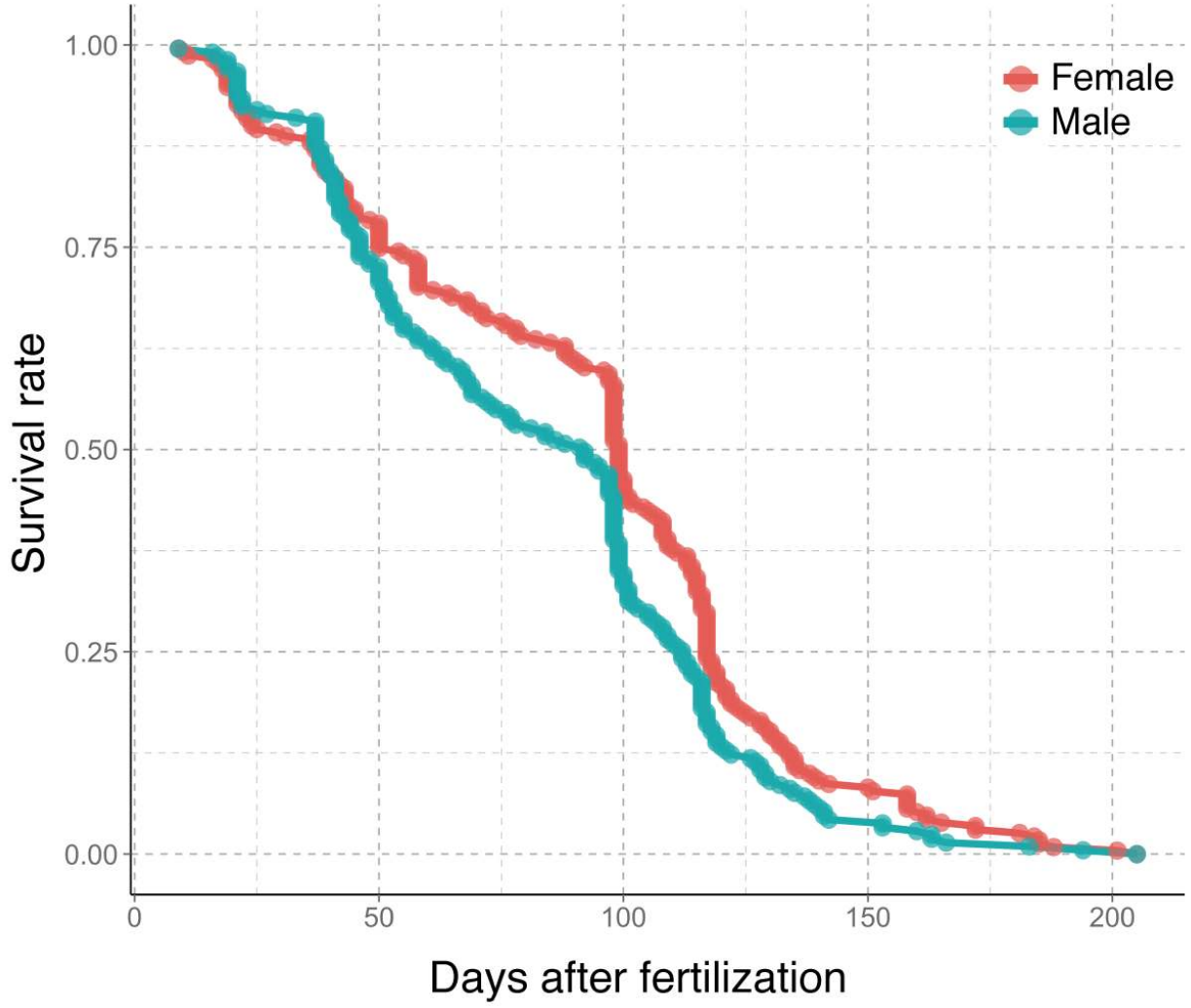

Fig. S11.

Survival curves of Japan Sea males and females. The pink and blue lines represent the survival curves of Japan Sea female and male juveniles, respectively, that were fed only DHA-free *Artemia*. Females showed higher rates of survival than males after about 50 days after fertilization (log-rank test, $P < 0.01$).

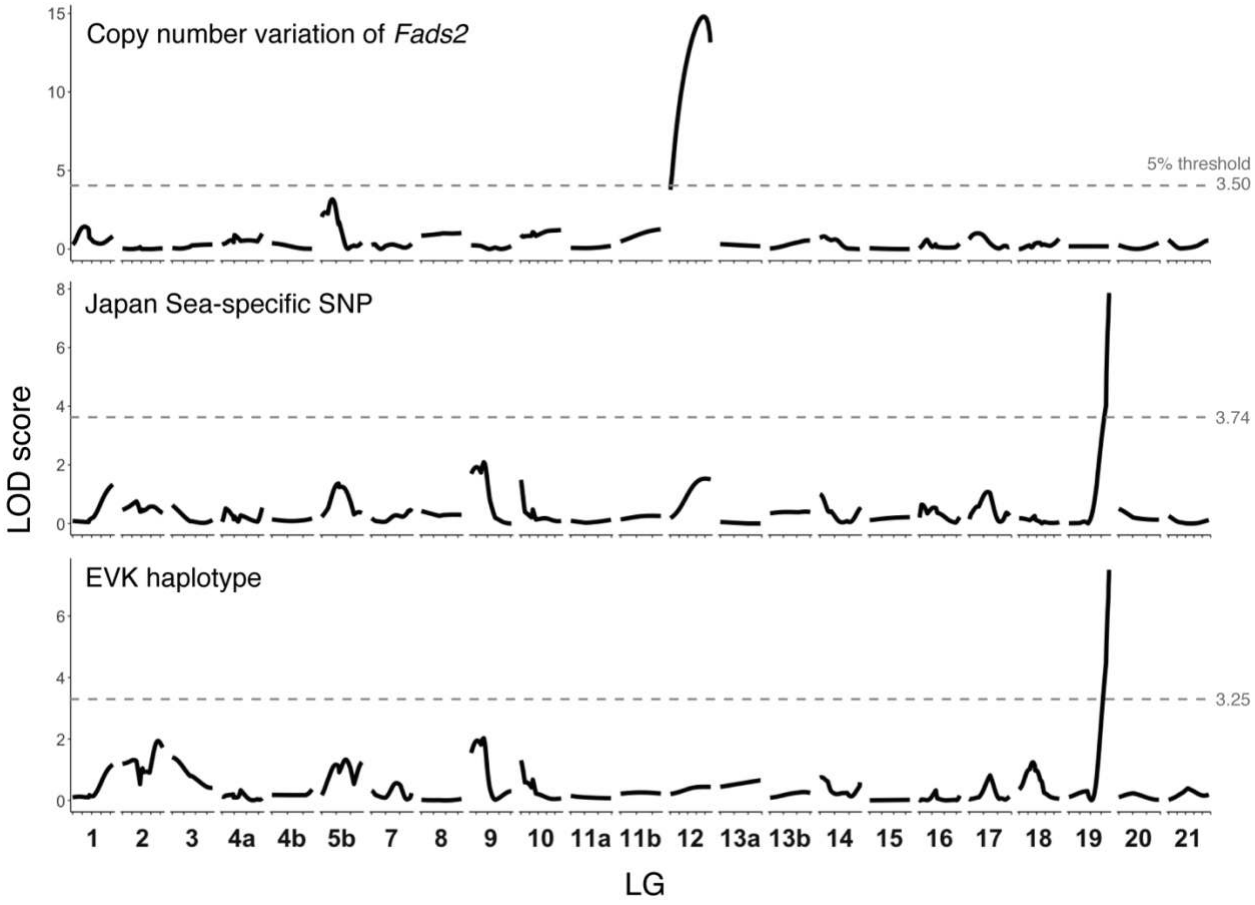

Fig. S12.

QTL mapping of copy number variation in *Fads2* and linkage mapping of SNPs within *Fads2* using F_2 hybrids of Japan Sea and Pacific Ocean sticklebacks. The X axis indicates the chromosomal position on each stickleback linkage group (LG), and the Y axis indicates the LOD score for copy number variation in *Fads2* (top panel), the presence/absence of a Japan Sea-specific SNP in *Fads2* derived from the Japan Sea grandfather (middle panel) and the presence/absence of the EVK haplotype of *Fads2* derived from the Pacific Ocean grandmother (bottom panel). Due to the small sample size of crosses, LG4, LG11, and LG13 were divided into two linkage groups, which are designated as 'a' and 'b'. The dashed lines indicate genome-wide significance thresholds determined by permutation tests (*$P < 0.05$). LG 12 (an autosome) significantly contributes to the copy number variation, and some SNPs within *Fads2* are linked also to the sex chromosome (LG19) in both Japan Sea and Pacific Ocean fishes.

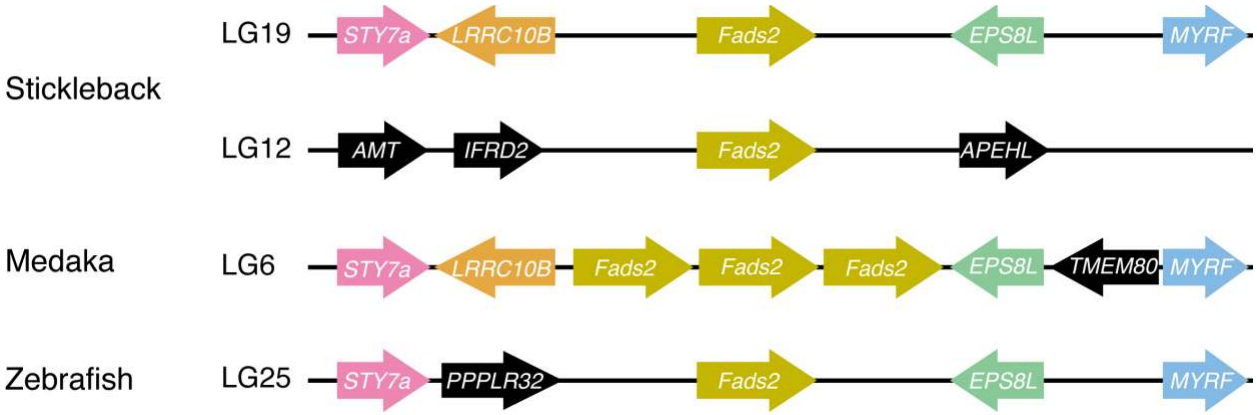

Fig. S13.

Synteny at *Fads2* loci in teleosts. Gene orders in the medaka and zebrafish were obtained from Medaka HdrR and Zebrafish GRCz10 reference genome sequences. The paralogues conserved among stickleback, medaka, and zebrafish are shown as color-matched arrows. Genes around *Fads2* on LG19 show conserved synteny with other teleosts, but those on LG12 do not.

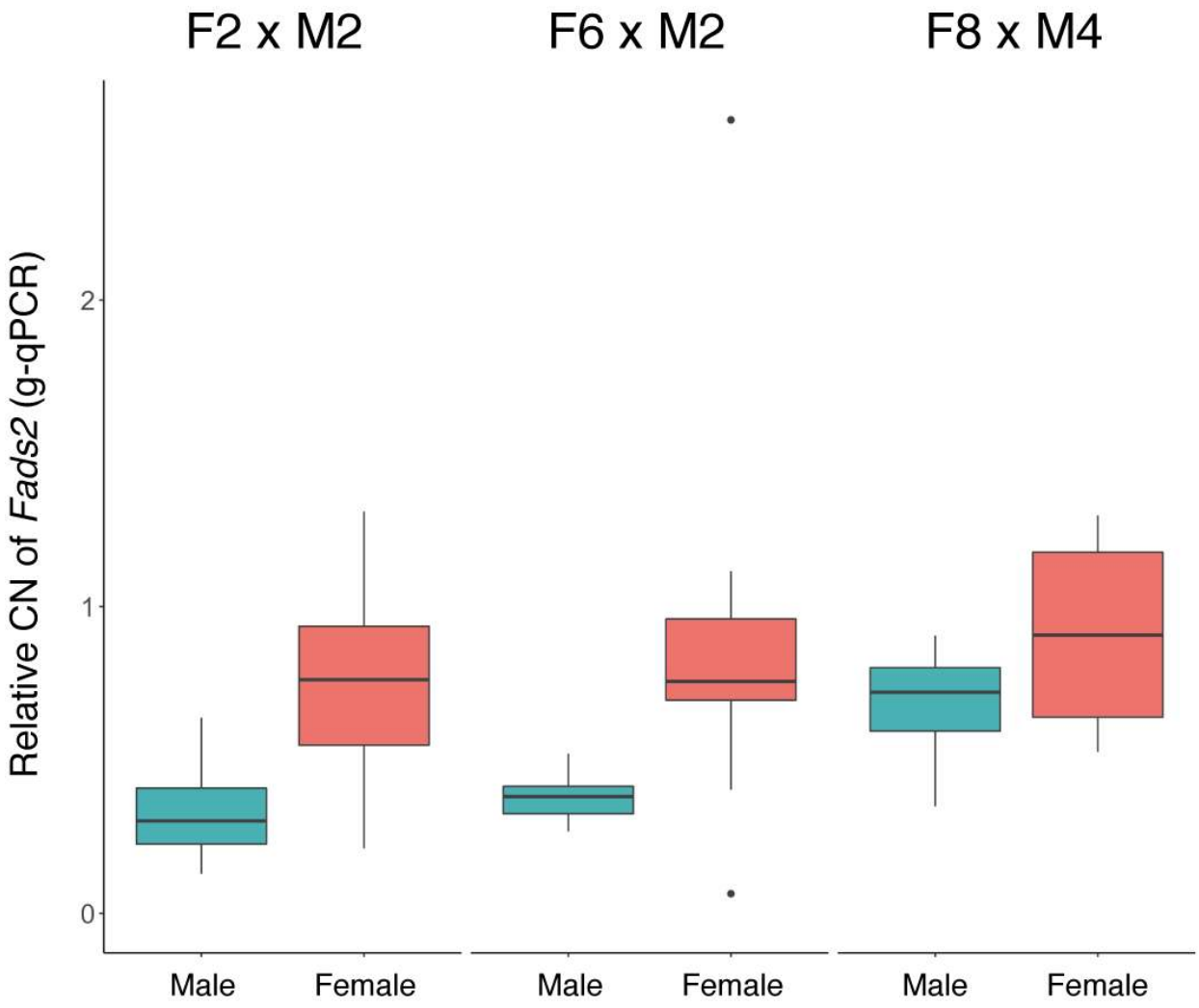

Fig. S14.

Relative copy numbers of *Fads2* in males and females of *G. wheatlandi*. Relative copy number was quantified by genomic qPCR. The three panels indicate three families used in the analysis. Reflecting the sex-linkage of LG19 in *G. wheatlandi* (*93*), *Fads2* copy number was higher in females than in males (one-way ANOVA followed by Tukey's HSD test, $P < 0.05$) (See also Fig. S22).

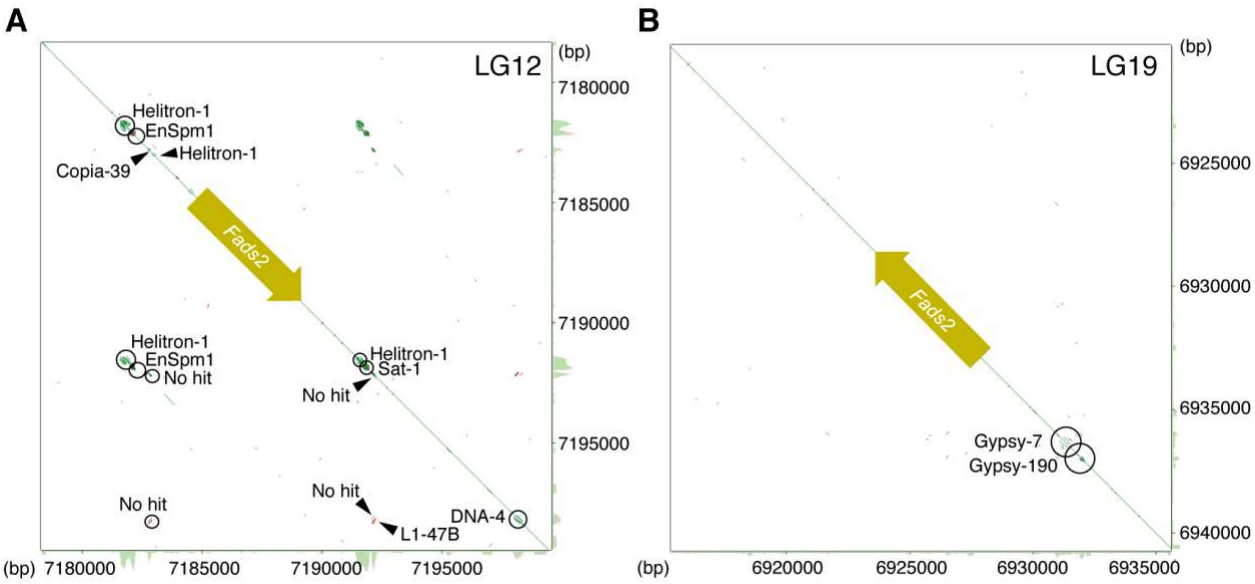

Fig. S15.

Analysis of repetitive sequences around *Fads2* on *G. aculeatus* LG12 and LG19. The dot-plot generated by YASS shows local pairwise alignments of the non-repeat-masked reference sequence around *Fads2* gene on LG12 (**A**) and LG19 (**B**) against itself. Chromosomal positions on LG12 and LG19 are given on both axes in bp. Yellow arrows indicate the *Fads2* gene. Green dots represent forward alignments and red dots represent reverse alignments. Circles and arrowheads indicate repetitive elements identified in Repbase by CENSOR.

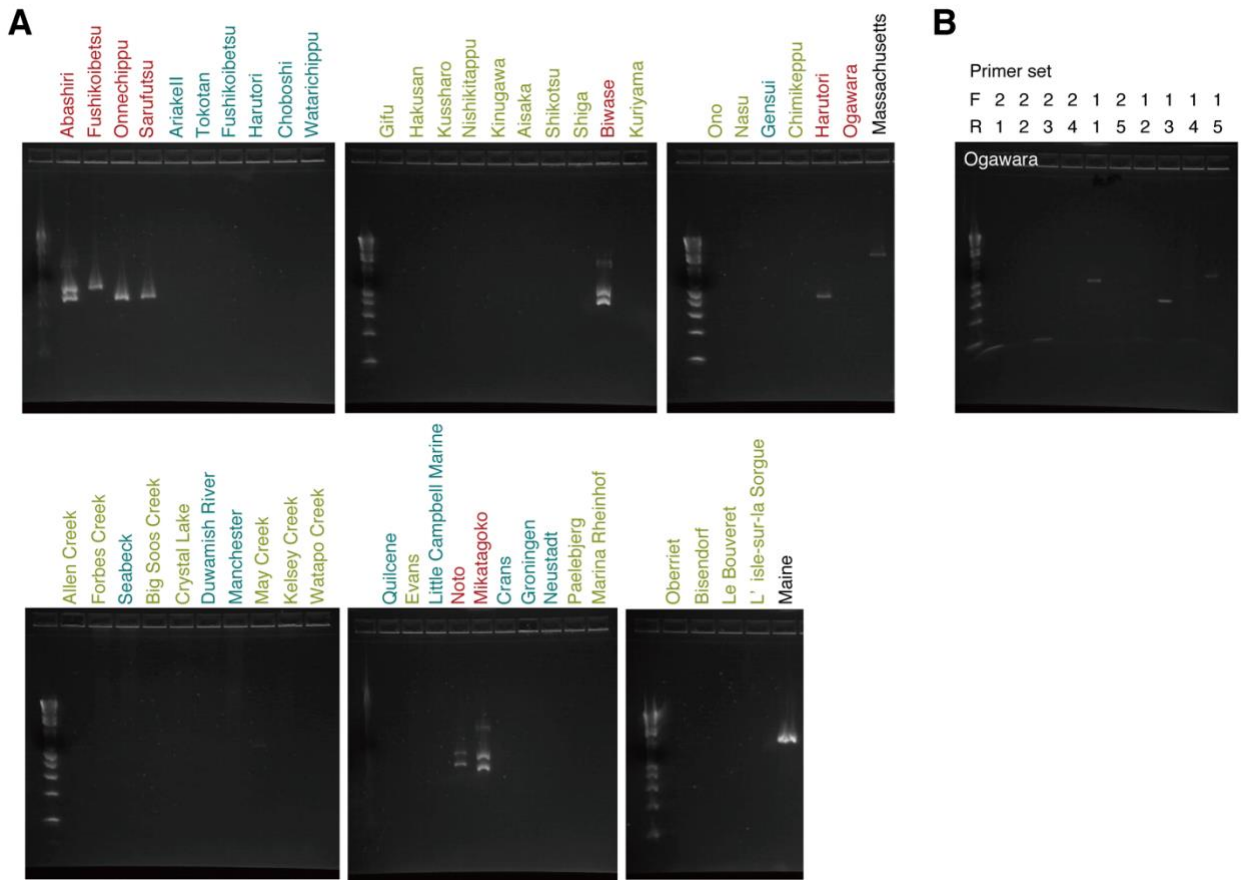

Fig. S16.

PCR detection of short sequences in the genomic region of Japan Sea stickleback and *G. wheatlandi* where *Fads2* is inserted in Pacific Ocean stickleback. (**A**) Agarose gel electrophoresis of PCR products obtained from Japan Sea, Pacific Ocean, and freshwater threespine stickleback populations, and *G. wheatlandi*. Red, blue, yellow-green, and black letters indicate Japan Sea, Pacific Ocean and freshwater threespine sticklebacks, and *G. wheatlandi*, respectively. Polymorphic sequences (160 or 290 bp) were obtained from all Japan Sea fish except for the Ogawara Lake population (see below), while no amplicon was obtained from Pacific Ocean and freshwater sticklebacks because of the 12 kb insertion including *Fads2*. *G. wheatlandi* had an amplicon about 1 kb, which includes additional 500 bp sequences at the 5' side because a primer set positioned at a different locus was used for PCR (see Methods). (**B**) Agarose gel electrophoresis of PCR products obtained from Japan Sea sticklebacks from Ogawara Lake using different primer sets. Since we could not obtain the PCR amplicon in the Ogawara Lake population possibly because of nucleotide changes at the primer sites in this population, additional PCR was conducted using different primer sets and another PCR enzyme (see Methods). The PCR amplicons were obtained when we used F1-R1, F1-R3, and F1-R5 primer pairs.

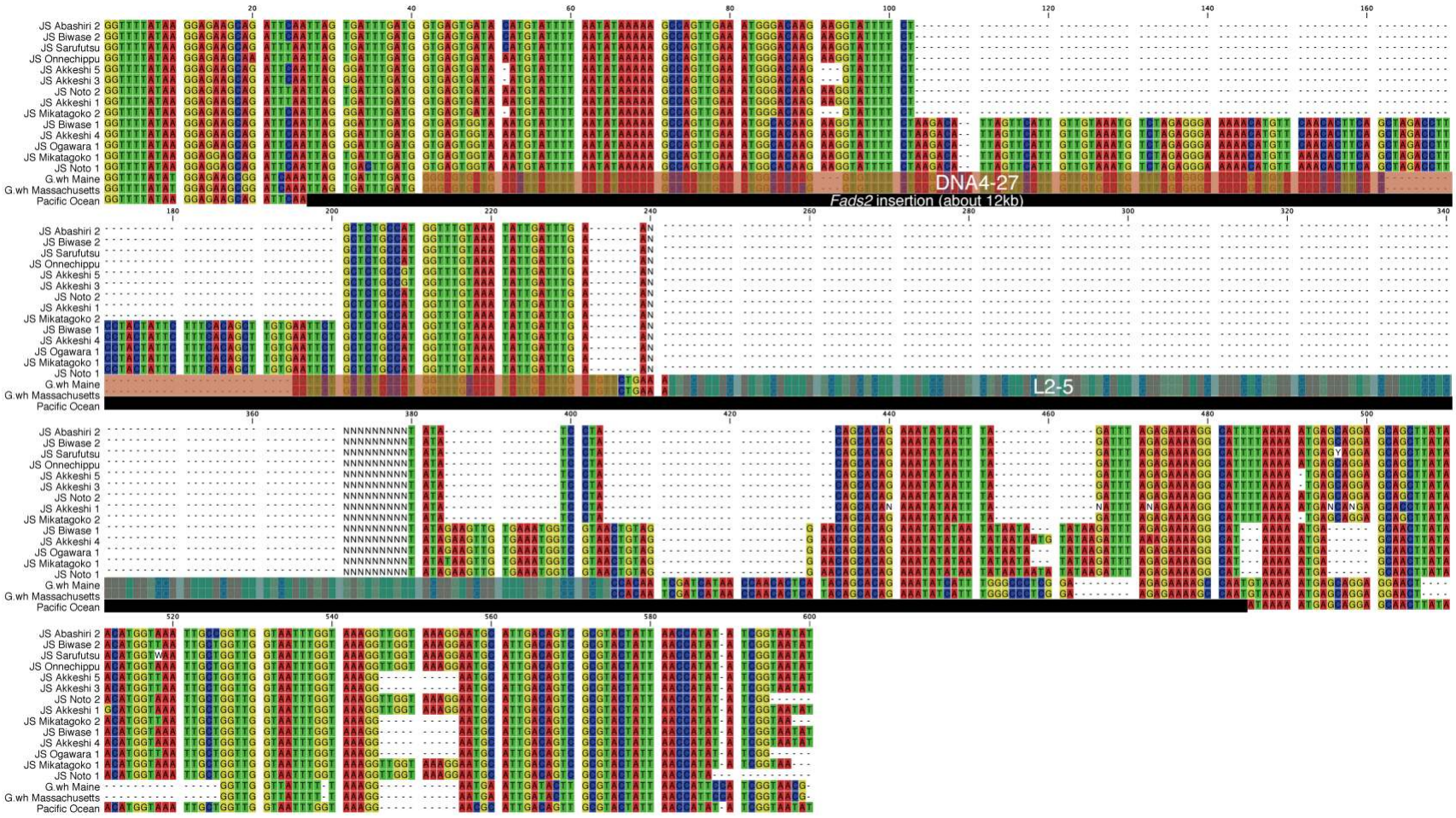

Fig. S17.

Multiple sequence alignments among multiple Japan Sea populations and *G. wheatlandi* for the region where *Fads2* is inserted in the Pacific Ocean lineage. The alignments were performed by CLC Sequence Viewer (CLC bio). The sequence name starting with JS or G.wh corresponds to that from Japan Sea or *G. wheatlandi*, respectively. Red-translucent and blue-translucent boxes on *G. wheatlandi* sequences are DNA4-27 transposon and L2-5 non-LTR retrotransposons. *G. wheatlandi* possessed the DNA transposon and non-LTR retrotransposon at this locus. Japan Sea sticklebacks had polymorphic short sequences, containing several inverted repeats (Fig. S18), which are similar to the DNA transposon found in *G. wheatlandi*.

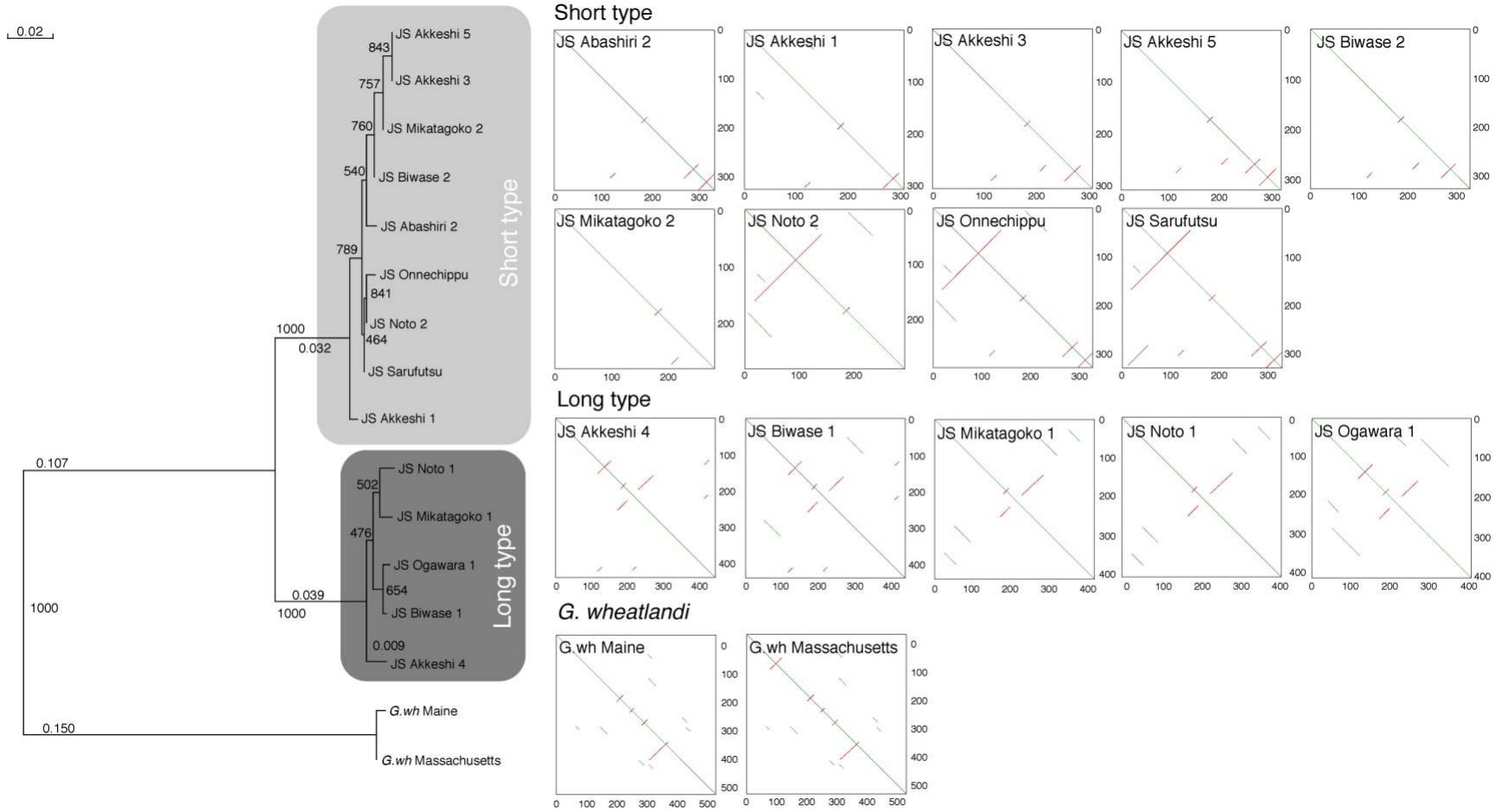

Fig. S18.

Polymorphic sequences among different Japan Sea and *G. wheatlandi* populations of the region where *Fads2* is inserted in the Pacific Ocean lineages. The left panel indicates the phylogenetic tree of polymorphic sequences found in Japan Sea and *G. wheatlandi* populations, which are located in the region where *Fads2* is inserted in the Pacific Ocean lineages. The sequences obtained from Japan Sea populations were divided into two groups (short and long types). The right panels indicate dot-plots generated by YASS showing local pairwise alignments of the haplotype sequences against themselves. Green dots represent forward alignments and red dots represent reverse alignments. Most of the sequences include several inverted repeats.

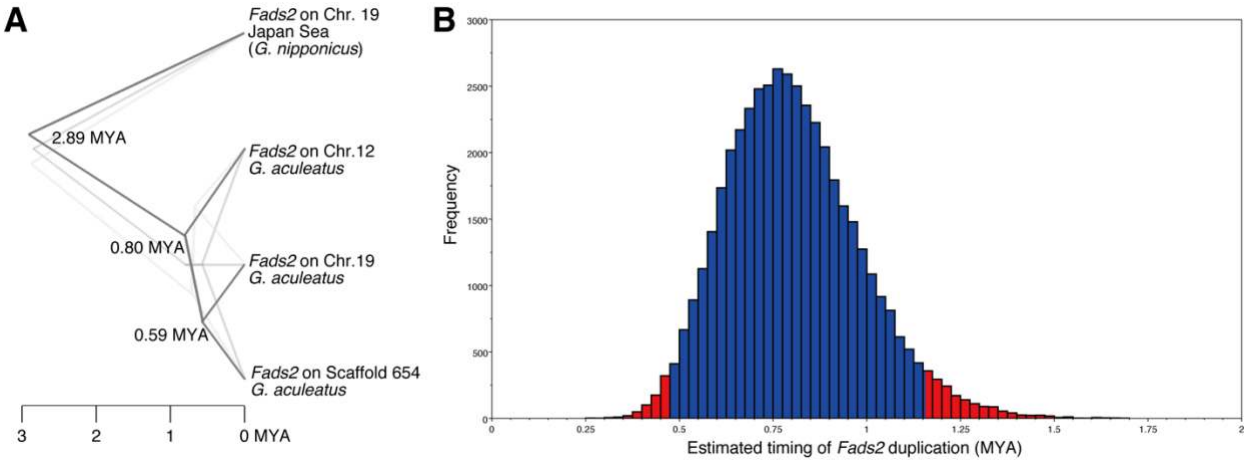

Fig. S19.

Phylogenetic tree of *Fads2* genes. (**A**) The inferred timing of duplication events is indicated on the horizontal axis. *Fads2* duplication was estimated to be 0.80 MYA, which is much earlier than the end of the last ice age (0.012 MYA), when the majority of stickleback colonization of freshwater occurred. (**B**) The posterior probability density of the estimated timing of *Fads2* duplication. Blue area indicates the 95% highest posterior density (0.47-1.16 MYA).

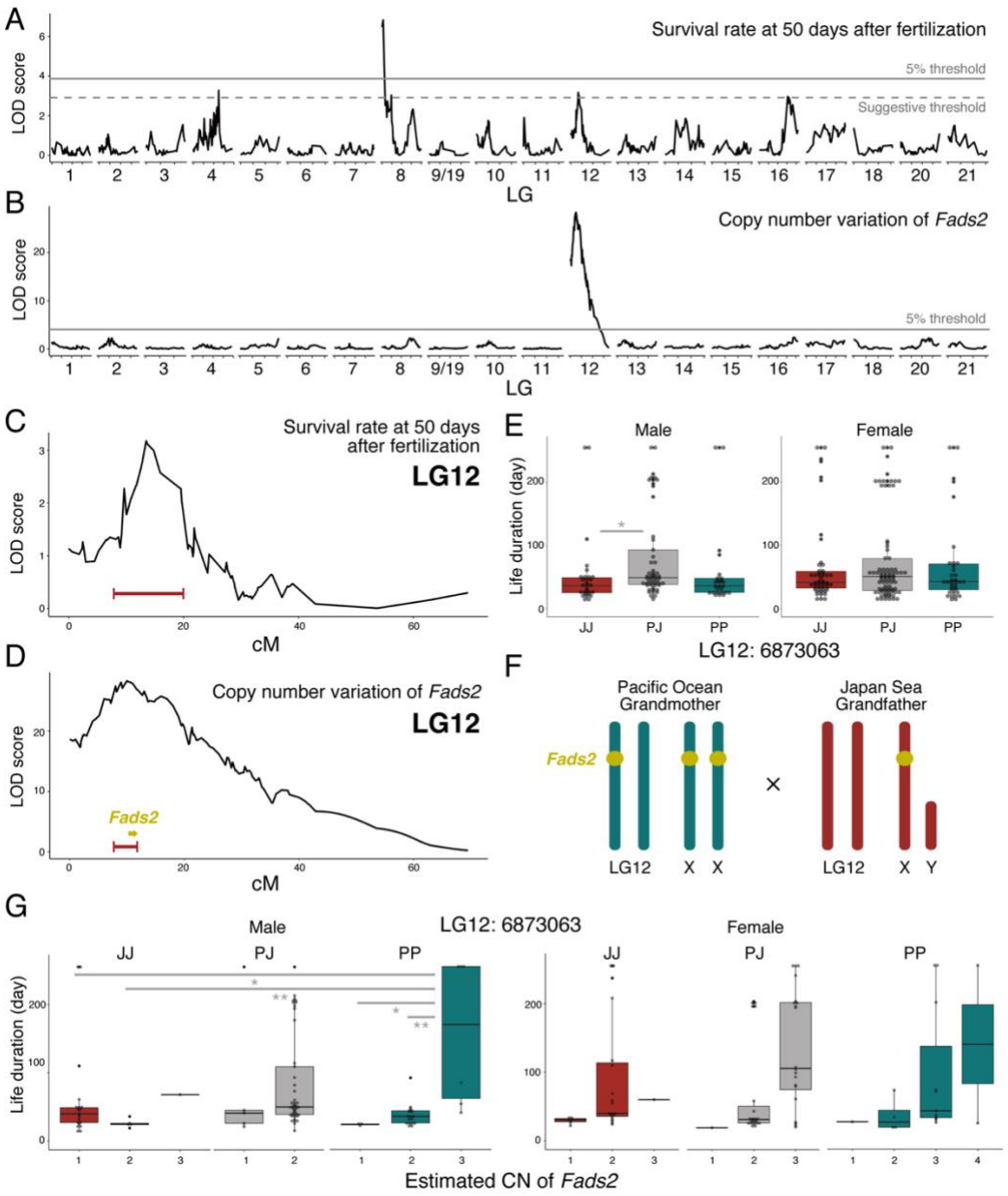

Fig. S20.

Genetic mapping of survival rates. (**A**-**D**) A single-QTL mapping of survivorship at 50 days after fertilization (A and C) and *Fads2* copy number (B and D) using the F_2 intercross resulting from a cross between a Japanese Pacific Ocean grandmother and a Japan Sea grandfather. The X axis indicates chromosomal position in each stickleback linkage group (LG), while the Y axis indicates the LOD score. The solid and dashed lines indicate genome-wide significance thresholds determined by genome-wide permutation tests (*P < 0.05) and thresholds of suggestive QTL (LOD > 3.0) (*111*), respectively. LG9 and LG19 are fused because F_2 males carry a Japan Sea neo-Y sex chromosome formed by the chromosomal fusion between LG9 and LG19 (*99*) and we used both males and females for making the linkage map. One significant and three suggestive QTLs were found for survivorship at 50 days after fertilization, whereas only one significant QTL was found on LG12 for the *Fads2* copy number. The significant QTL on LG8 and the suggestive QTLs on LG4, 12, and 16 explained 8.723, 3.309, 3.263, and 2.931 % of the phenotypic variance in survivorship, respectively. A two-dimensional QTL mapping did not detect any significant epistatic interactions. The 95% Bayesian credible intervals of one suggestive QTL on LG12 for the survival rates (red line in C) and the significant QTL for the copy number (red line in D) overlapped with the *Fads2* gene on LG12

(yellow arrow in D). It should be noted that no QTL was found on LG19, suggesting that changes in protein sequences or *cis*-regulatory regions of *Fads2* on LG19 are unlikely to contribute to survival rates. (**E**) Plot of association between SNP genotypes at a marker closest to *Fads2* on LG12 and life duration. The use of a marker nearest the QTL did not change the results qualitatively. Dark-red, grey and dark-blue boxes indicate homozygotes of Japan Sea alleles (JJ), heterozygotes of Pacific Ocean and Japan Sea alleles (PJ), and homozygotes of Pacific Ocean alleles (PP) at each locus, respectively. Asterisks indicate pairs with significant differences (one-way ANOVA followed by Tukey's HSD test, $*P < 0.05$, $**P < 0.01$). (**F**) Predicted genotypes of grandparents. In the Pacific Ocean grandmother, one LG12 allele possessed *Fads2*, while another LG12 allele lacked *Fads2*. Thus, in the F_2 intercross, both PJ and PP genotypes at the LG12 QTL can have different *Fads2* copy numbers. For example, F_2 males with PP genotypes at the LG12 QTL can have a single *Fads2* copy only on the X chromosome, two *Fads2* copies (one on the X chromosome and one on LG12), or three *Fads2* copies (one on the X chromosome and two on LG12). F_2 males with PJ genotypes at the LG12 QTL can have a single *Fads2* copy only on the X chromosome or two *Fads2* copies (one on the X chromosome and one on LG12). (**G**) Effects of *Fads2* copy numbers and genotypes at the *Fads2* locus on LG12 on the life duration. F_2 males with a higher copy number of *Fads2* had longer life duration when compared within each genotype (ANCOVA; effect of *Fads2* copy number: $F_{2,96} = 8.295$, $P < 0.01$, effect of SNP genotype: $F_{2,96} = 4.666$, $P < 0.05$, interaction: $F_{3,96} = 1.630$, $P > 0.05$). Because copy number variation among F_2 fish of the same sex is derived only from LG12 but not from LG19, these data suggest that higher copy numbers of *Fads2* on LG12 increase the survival rates. Asterisks indicate pairs with significant differences (Tukey's HSD test, $*P < 0.05$, $**P < 0.01$). One or a few JJ males with 2 or 3 *Fads2* copies are rare recombinants between *Fasd2* and the SNP marker. The F_2 females with higher copy numbers of *Fads2* also showed higher survival rates when compared within each genotype (ANCOVA; effect of estimated *Fads2* copy number: $F_{3,78} = 6.245$, $P < 0.01$, effect of genotype: $F_{2,78} = 0.527$, $P > 0.05$, interaction: $F_{4,78} = 0.362$, $P > 0.05$), although none of the pairs were significant with the posthoc Tukey's HSD test.

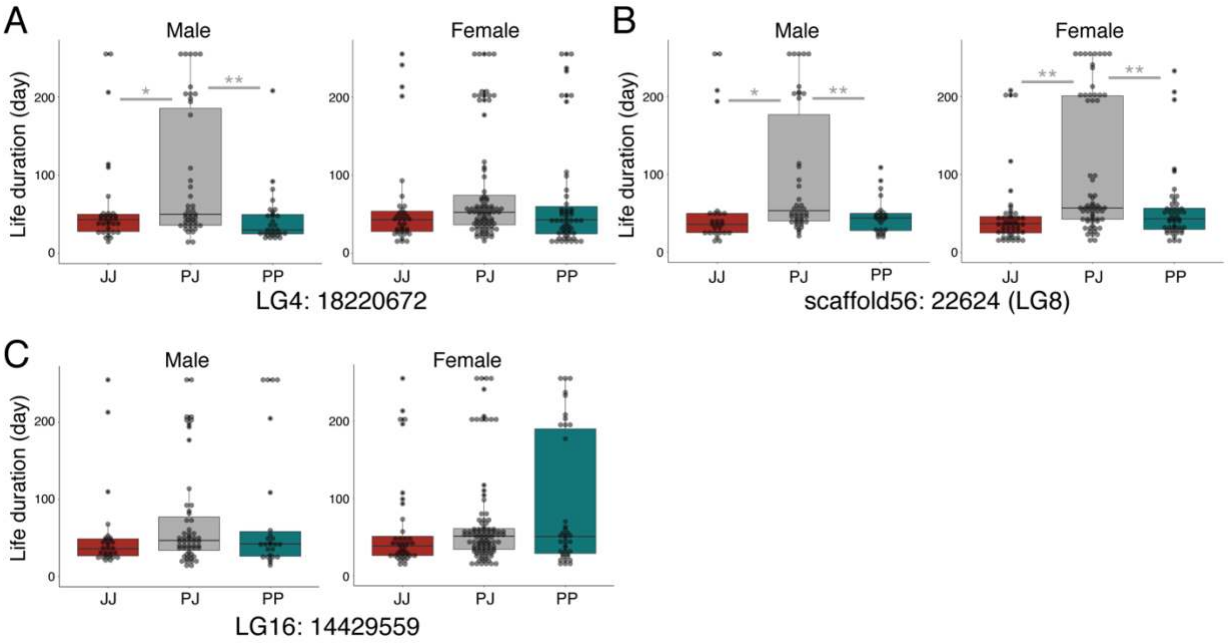

Fig. S21

Allelic effects on the life duration at markers closest to the significant and suggestive QTLs on LG4 (A), 8 (B), and 16 (C). Dark-red, grey and dark-blue boxes indicate homozygotes of Japan Sea alleles (JJ), heterozygotes of Pacific Ocean and Japan Sea alleles (PJ), and homozygotes of Pacific Ocean alleles (PP) at each locus, respectively. Two QTLs on LG4 and LG8 showed overdominance. Asterisks indicate pairs with significant differences (one-way ANOVA followed by Tukey's HSD test, $*P < 0.05$, $**P < 0.01$)

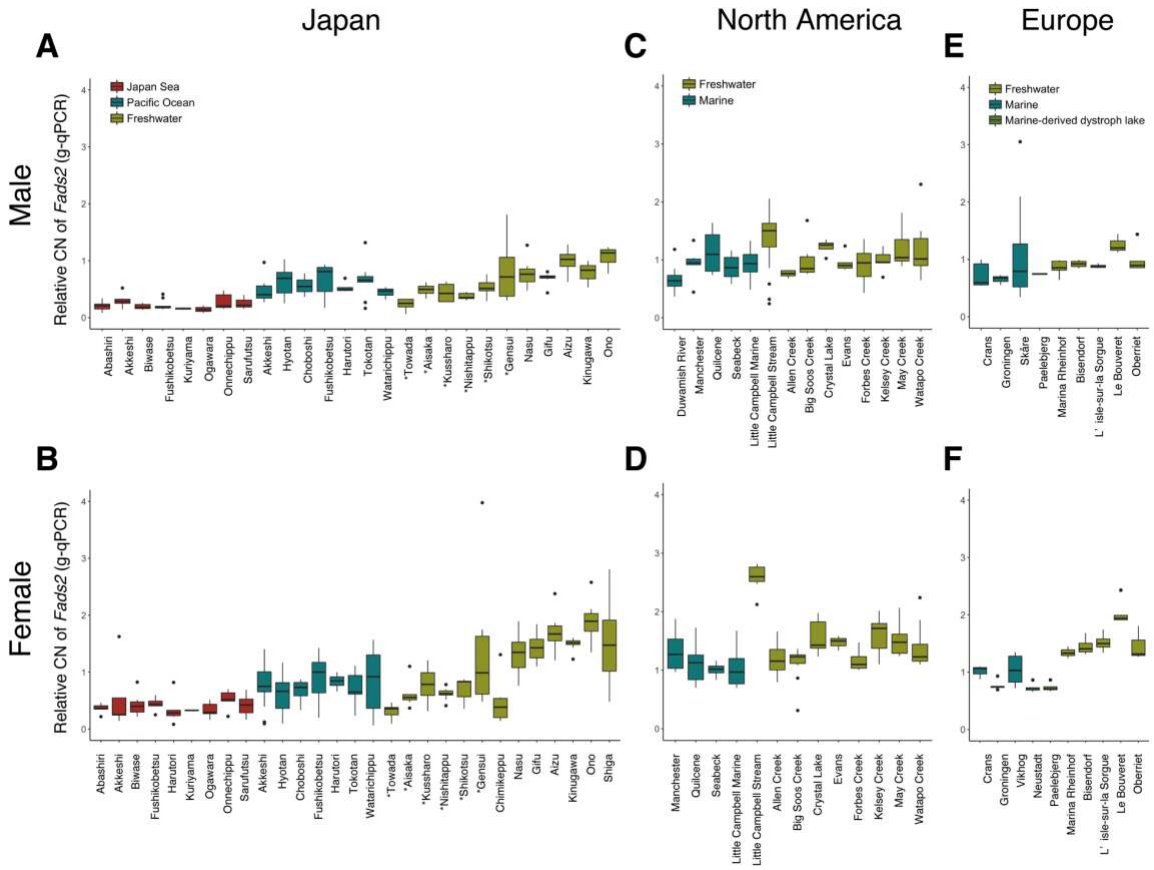

Fig. S22.

Copy number (CN) variation in *Fads2* shown for each population. (**A**) Copy number variation in Japanese males. (**B**) Copy number variation in Japanese females. (**C**) Copy number variation in North American males. (**D**) Copy number variation in North American females. (**E**) Copy number variation in European males. (**F**) Copy number variation in European females. Dark-red, dark-blue, and yellow-green boxes indicate Japan Sea, Pacific Ocean, and freshwater sticklebacks, respectively in (**A** and **B**). Dark-blue and yellow-green boxes indicate marine and freshwater sticklebacks in (**C**-**F**). The freshwater populations with asterisk (*) in (**A** and **B**) colonized freshwater more recently than others (see Fig. 1C). The dark green box in (F) indicates the population in a marine-derived dystrophic lake (Paelebjerg), which was landlocked within the last 400 years.

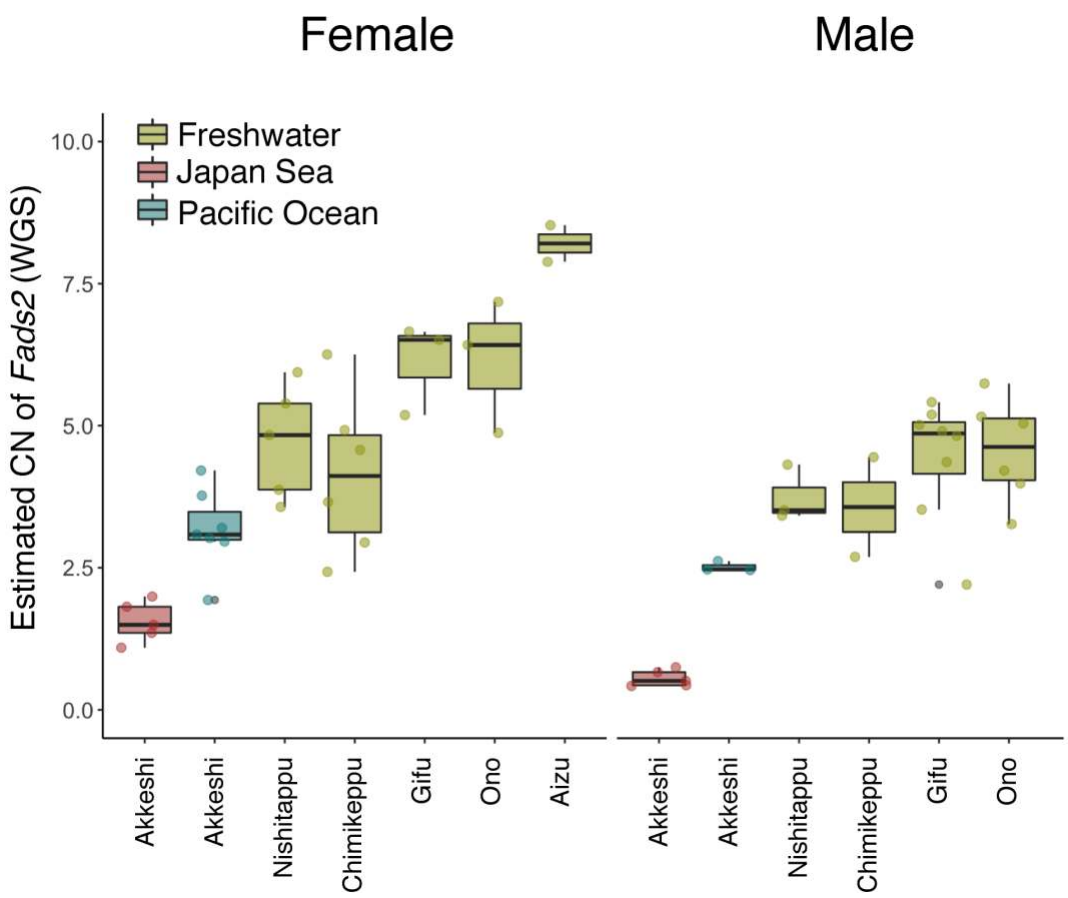

Fig. S23.

Estimated copy number of *Fads2* from WGS data. Left, copy number variation in female. Right, copy number variation in males. Red, blue, and yellow-green boxes indicate Japan Sea, Pacific Ocean, and freshwater sticklebacks, respectively. Each dot indicates a single individual.

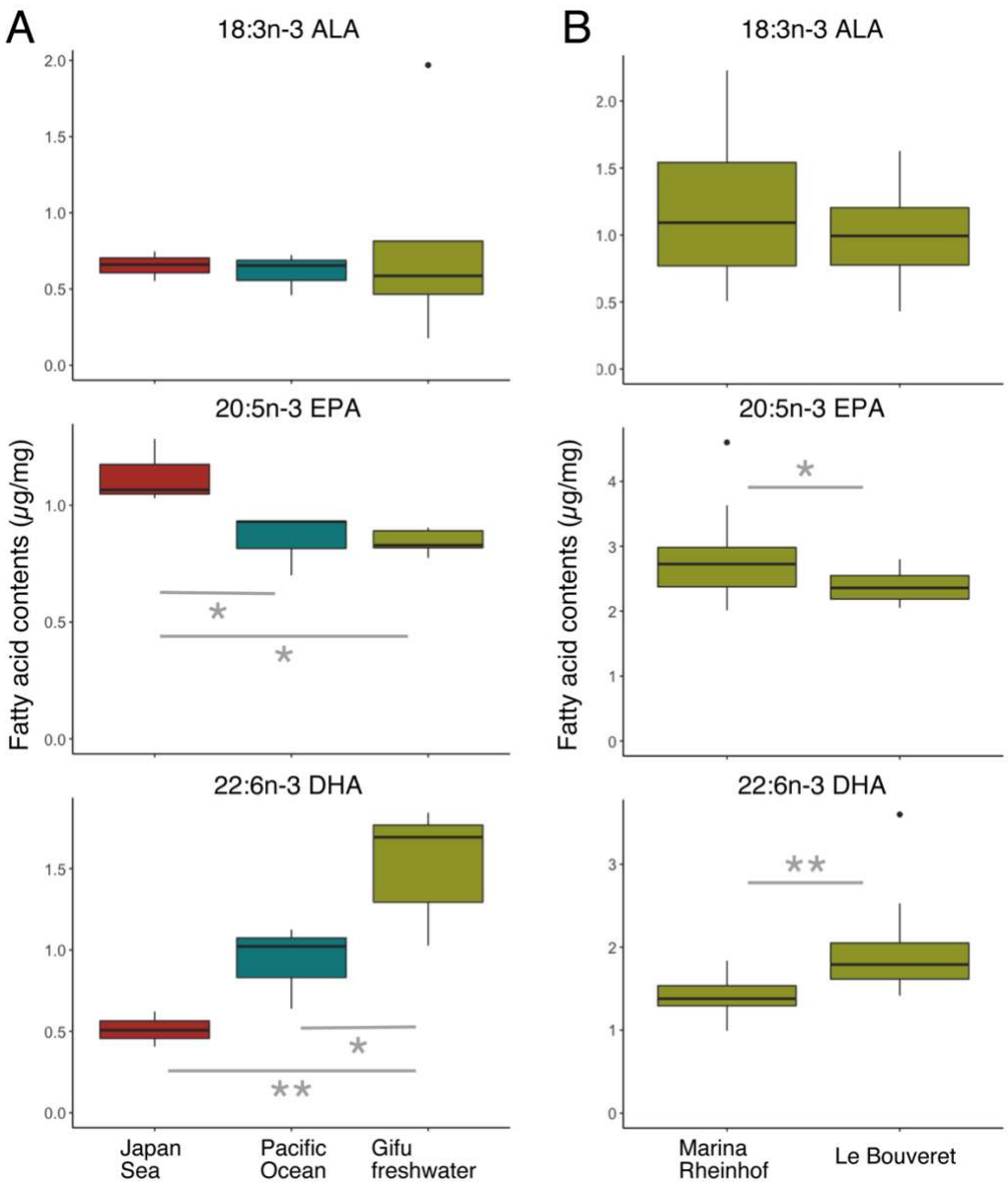

Fig. S24.

ω-3 Fatty acid contents of freshwater populations. **(A)** ω-3 Fatty acid content in the eyeballs of a freshwater population from Gifu compared with those of Japan Sea and Pacific Ocean sticklebacks (per wet weight in eyeball). ALA, EPA and DHA were measured with GC-MS. Dark-red, dark-blue, and yellow-green boxes indicate the fatty acid contents of Japan Sea, Pacific Ocean, and freshwater fish from Gifu fed only DHA-free *Artemia* (N = 3, 3, and 5, respectively). Asterisks indicate significant differences between populations (one-way ANOVA followed by Tukey's HSD test, *P < 0.05, **P < 0.01). **(B)** ω-3 Fatty acid content of the muscle tissue of Marina Rheinhof and Le Bouvert. ALA, EPA and DHA were measured with GC-FID (N = 13 for Marina Rheinhof and N = 17 for Le Bouvert) (per dry weight in muscle). Asterisks indicate significant differences between the populations (Student's t-test, *P < 0.05, **P < 0.01). It should be noted that the Le Bouvert population has higher *Fads2* copy numbers than the Marina Rheinhof population (see Fig. S22).

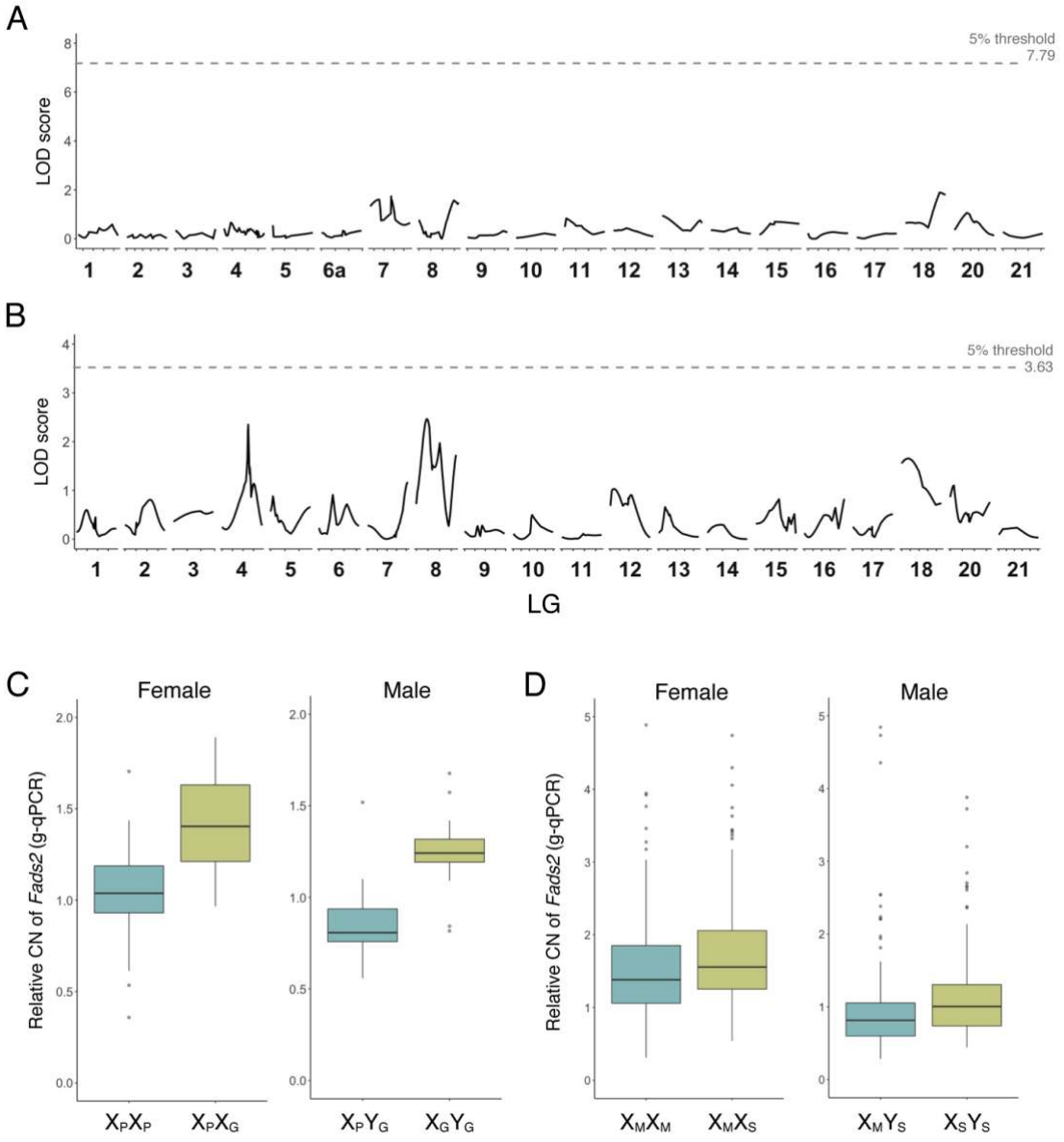

Fig. S25.

Linkage analyses of copy number variation in *Fads2* using F_2 hybrids of Pacific Ocean and Gifu sticklebacks (**A** and **C**) and those of Little Campbell River marine and stream sticklebacks (**B** and **D**). (**A**-**B**) The X axis indicates the chromosomal position on each stickleback linkage group (LG), and the Y axis indicates the LOD score for copy number variation in *Fads2*. The dashed lines indicate genome-wide significance thresholds determined by permutation tests (*$P < 0.05$). No significant QTLs were found on autosomes. (**C**) The copy number differences between different genotypes at a SNP marker on LG19 (groupXIX:5500988) in F_2 hybrids of Pacific Ocean and Gifu fishes. The F_2 individuals with the X chromosome from Gifu grandfather (X_G) possessed higher copy number than those with the X chromosome from Pacific Ocean grandmother (X_P) when compared within each sex (two-way ANOVA; Genotype: $F_{1,67} = 18.0$, $P < 0.01$; Sex: $F_{1,67} = 3.70$, $P > 0.05$). (**D**) The copy number differences between different genotypes at a SNP marker on LG19 (groupXIX: 10118548) in F_2 hybrids of Little Campbell River marine and stream ecotypes. The F_2 individuals with the X chromosome from the stream grandfather (X_S) possessed a higher copy number than those with the X chromosome from a marine grandmother (X_M) when compared within each sex (two-way ANOVA; Genotype: $F_{1,407} = 15.6$, $P < 0.01$; Sex: $F_{1,407} = 32.7$, $P < 0.01$).

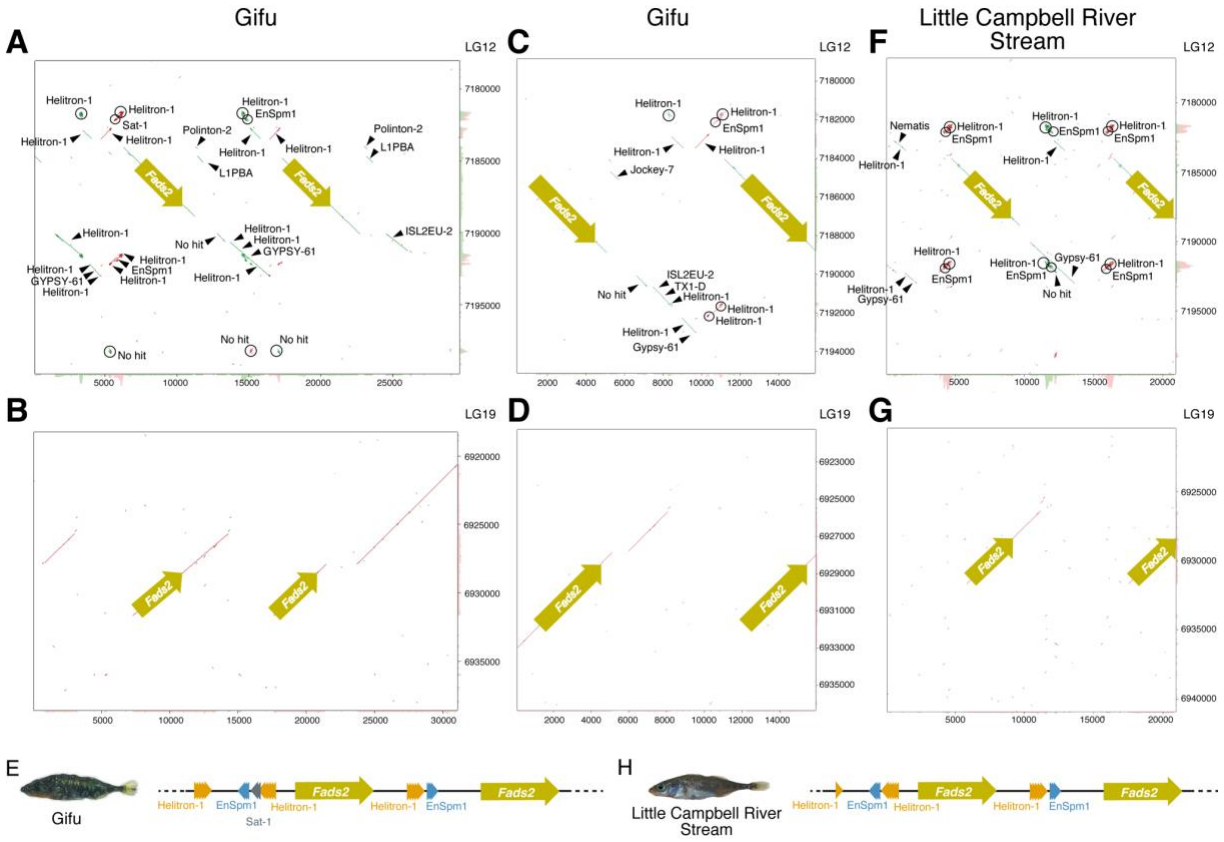

Fig. S26.

Tandem duplication of *Fads2* in freshwater sticklebacks from Gifu and Little Campbell River streams, identified by long-read sequencing. Dot-plots of long-read sequences of the freshwater population from Gifu obtained by PacBio sequencing (**A**-**B**) and MinION Nanopore sequencing (**C**-**D**) are shown. Dot-plots of long-read sequences of the freshwater population from the upstream of Little Campbell River obtained by PacBio sequencing (**F**-**G**). Long reads containing *Fads2* were analyzed against *Fads2* loci on LG12 (**A**, **C** and **F**) and LG19 (**B**, **D** and **G**) from the stickleback reference genome. Positions on the long read sequences and the chromosomal positions on LG12 and LG19 of the stickleback reference genome are indicated on the X and Y axes, respectively. Yellow arrows indicate *Fads2*. Green dots represent forward alignments and red dots represent reverse alignments. Circles and arrowheads indicate repetitive elements identified in Repbase by CENSOR. (**E** and **H**) Genomic structure of *Fads2* loci of freshwater stickleback from Gifu (**E**) and Little Campbell River Stream (**H**). Yellow-green arrows indicate tandemly duplicated *Fads2* genes. Arrowheads indicate repetitive elements identified in Repbase by CENSOR, which are found in both PacBio and Nanopore sequencing, suggesting that the detected tandem duplications are unlikely to be sequencing errors.

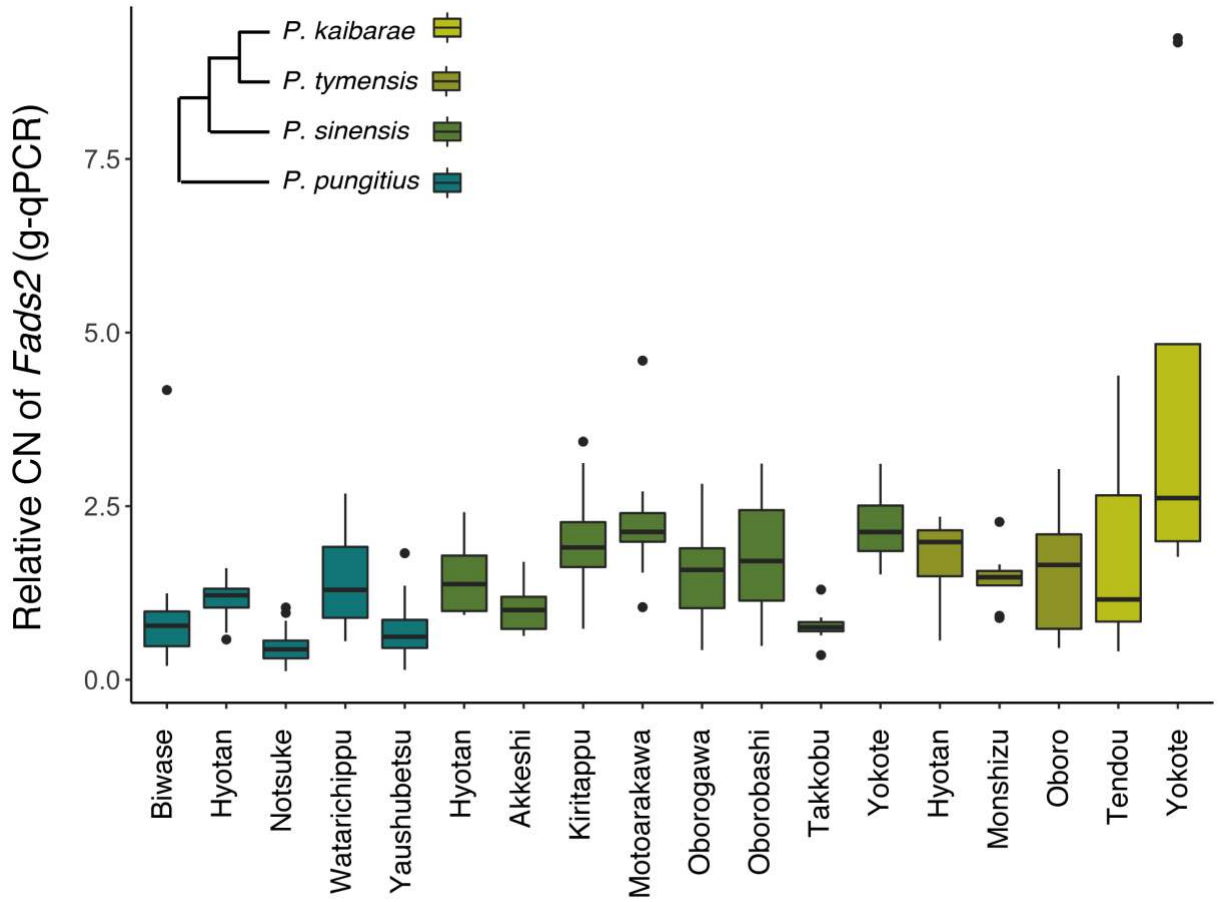

Fig. S27.

Copy number (CN) variation of *Fads2* in ninespine sticklebacks. Dark-blue, green, yellow-green and yellow boxes indicate copy number variation in *Pungitius pungitius*, *P. sinensis*, *P. tymensis*, and *P. kaibarae*. *P. pungitius* inhabits only brackish environments in Japan, while *P. sinensis* inhabits both freshwater and brackish environments. Both *P. tymensis* and *P. kaibarae* are entirely freshwater species. The phylogenetic relationship of the four species is shown in the upper-left corner (*27*).

Fig. S28.
Maximum likelihood tree of *Fads* genes. Arrows indicate species used as query sequences in the ORTHOSCOPE analysis. Numbers beside nodes indicate bootstrap probabilities with only those > 50% shown. Blue and red letters indicate genes from fishes that have freshwater populations and those that lack freshwater populations, respectively.

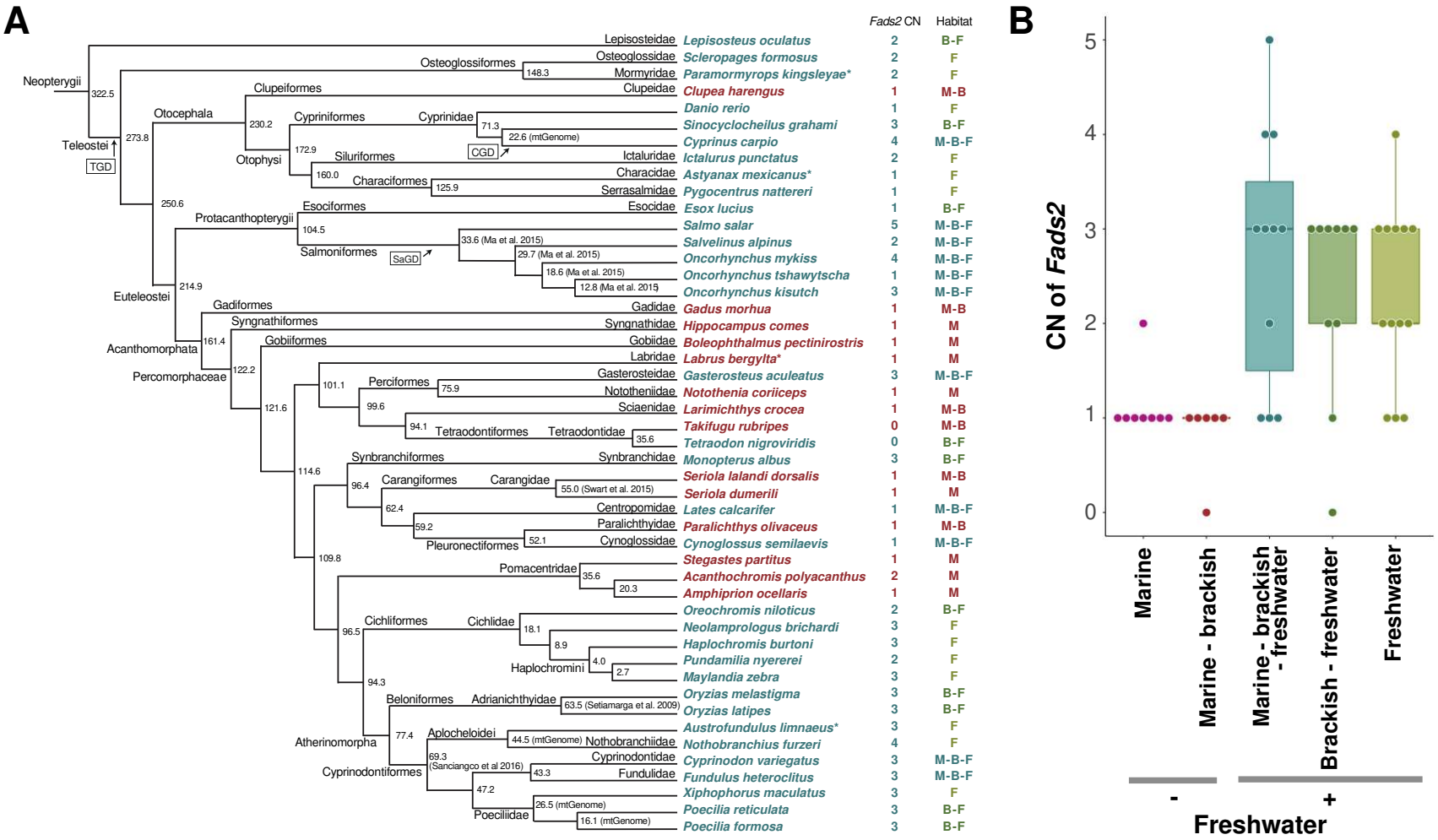
A
Fads2 CN
Habitat
Neopterygii
322.5
Teleostei
TGD
273.8
Otocephala
230.2
Cypriniformes
Cyprinidae
71.3
22.6 (mtGenome)
CGD
Otophysi
172.9
Siluriformes
Ictaluridae
160.0
Characiformes
Characidae
125.9
Serrasalmidae
250.6
Protacanthopterygii
Esociformes
Esocidae
104.5
Salmoniformes
SaGD
33.6 (Ma et al. 2015)
29.7 (Ma et al. 2015)
18.6 (Ma et al. 2015)
12.8 (Ma et al. 2015)
214.9
Euteleostei
Clupeiformes
Clupeidae
Osteoglossiformes
Osteoglossidae
148.3
Mormyridae
Gadiformes
Gadidae
Syngnathiformes
Syngnathidae
161.4
Acanthomorphata
Gobiiformes
Gobiidae
Labridae
122.2
Percomorphaceae
101.1
Perciformes
75.9
Gasterosteidae
Nototheniidae
99.6
Sciaenidae
121.6
94.1
Tetraodontiformes
Tetraodontidae
35.6
Synbranchiformes
Synbranchidae
114.6
96.4
Carangiformes
Carangidae
55.0 (Swart et al. 2015)
62.4
Centropomidae
59.2
Paralichthyidae
Pleuronectiformes
52.1
Cynoglossidae
109.8
Pomacentridae
35.6
20.3
96.5
Cichliformes
Cichlidae
18.1
8.9
Haplochromini
4.0
2.7
94.3
Beloniformes
Adrianichthyidae
63.5 (Setiamarga et al. 2009)
77.4
Atherinomorpha
Aplocheloidei
44.5 (mtGenome)
Nothobranchiidae
69.3
(Sanciangco et al 2016)
Cyprinodontidae
Cyprinodontiformes
43.3
Fundulidae
47.2
26.5 (mtGenome)
Poeciliidae
16.1 (mtGenome)
Lepisosteidae Lepisosteus oculatus 2 B-F
Osteoglossidae Scleropages formosus 2 F
Mormyridae Paramormyrops kingsleyae* 2 F
Clupeidae Clupea harengus 1 M-B
Danio rerio 1 F
Sinocyclocheilus grahami 3 B-F
Cyprinus carpio 4 M-B-F
Ictaluridae Ictalurus punctatus 2 F
Characidae Astyanax mexicanus* 1 F
Serrasalmidae Pygocentrus nattereri 1 F
Esocidae Esox lucius 1 B-F
Salmo salar 5 M-B-F
Salvelinus alpinus 2 M-B-F
Oncorhynchus mykiss 4 M-B-F
Oncorhynchus tshawytscha 1 M-B-F
Oncorhynchus kisutch 3 M-B-F
Gadidae Gadus morhua 1 M-B
Syngnathidae Hippocampus comes 1 M
Gobiidae Boleophthalmus pectinirostris 1 M
Labridae Labrus bergylta* 1 M
Gasterosteidae Gasterosteus aculeatus 3 M-B-F
Nototheniidae Notothenia coriiceps 1 M
Sciaenidae Larimichthys crocea 1 M-B
Takifugu rubripes 0 M-B
Tetraodontidae Tetraodon nigroviridis 0 B-F
Synbranchidae Monopterus albus 3 B-F
Seriola lalandi dorsalis 1 M-B
Carangidae Seriola dumerili 1 M
Centropomidae Lates calcarifer 1 M-B-F
Paralichthyidae Paralichthys olivaceus 1 M-B
Cynoglossidae Cynoglossus semilaevis 1 M-B-F
Stegastes partitus 1 M
Pomacentridae Acanthochromis polyacanthus 2 M
Amphiprion ocellaris 1 M
Oreochromis niloticus 2 B-F
Cichlidae Neolamprologus brichardi 3 F
Haplochromis burtoni 3 F
Haplochromini Pundamilia nyererei 2 F
Maylandia zebra 3 F
Oryzias melastigma 3 B-F
Adrianichthyidae Oryzias latipes 3 B-F
Austrofundulus limnaeus* 3 F
Nothobranchiidae Nothobranchius furzeri 4 F
Cyprinodontidae Cyprinodon variegatus 3 M-B-F
Fundulidae Fundulus heteroclitus 3 M-B-F
Xiphophorus maculatus 3 F
Poeciliidae Poecilia reticulata 3 B-F
Poecilia formosa 3 B-F
B
CN of Fads2
0
1
2
3
4
5
Marine
Marine - brackish
Marine - brackish - freshwater
Brackish - freshwater
Freshwater
-
+
Freshwater

Fig. S29.
Fads2 copy number in actinopterygians (ray-finned fishes). **(A)** *Fads2* copy numbers and habitat information were mapped on the actinopterygian species tree. Estimated divergence time (in million years) is shown at each node. Literature from which the divergence time was taken is shown at each node, while divergence time estimated from whole mitochondrial genome sequences is marked with mtGenome. Other data were taken from Betancur et al. 2017 (*122*). Teleost-specific whole genome duplication (TGD), salmonid-specific whole genome duplication (SaGD), and carp-specific whole genome duplication (CGD) are indicated in the tree (*129*). Blue and red color indicate species that have freshwater populations and those that lack freshwater populations, respectively. Habitat information was classified according to the Catalog of Fishes (*128*); M, marine only (pink); M-B, marine -brackish (red); M-B-F: marine – brackish - freshwater (blue); B-F, brackish-freshwater (green); F, freshwater only (yellow). For species with asterisks (*), we used the sequence data of closely related species for making this species tree (see the Methods). **(B)** Boxplot of *Fasd2* copy number (CN) for fish species inhabiting water ranging from marine to freshwater environments. Fish species that have freshwater populations (Freshwater +) possess higher *Fads2* copy number than species that lack freshwater populations (Freshwater -) (habitat: post.mean = -1.1172, 95% credible interval = -1.8495 - -0.4907, pMCMC < 0.01; genome duplication: post.mean=1.6312, 95% credible interval = 0.4875 – 2.8238, pMCMC < 0.01).

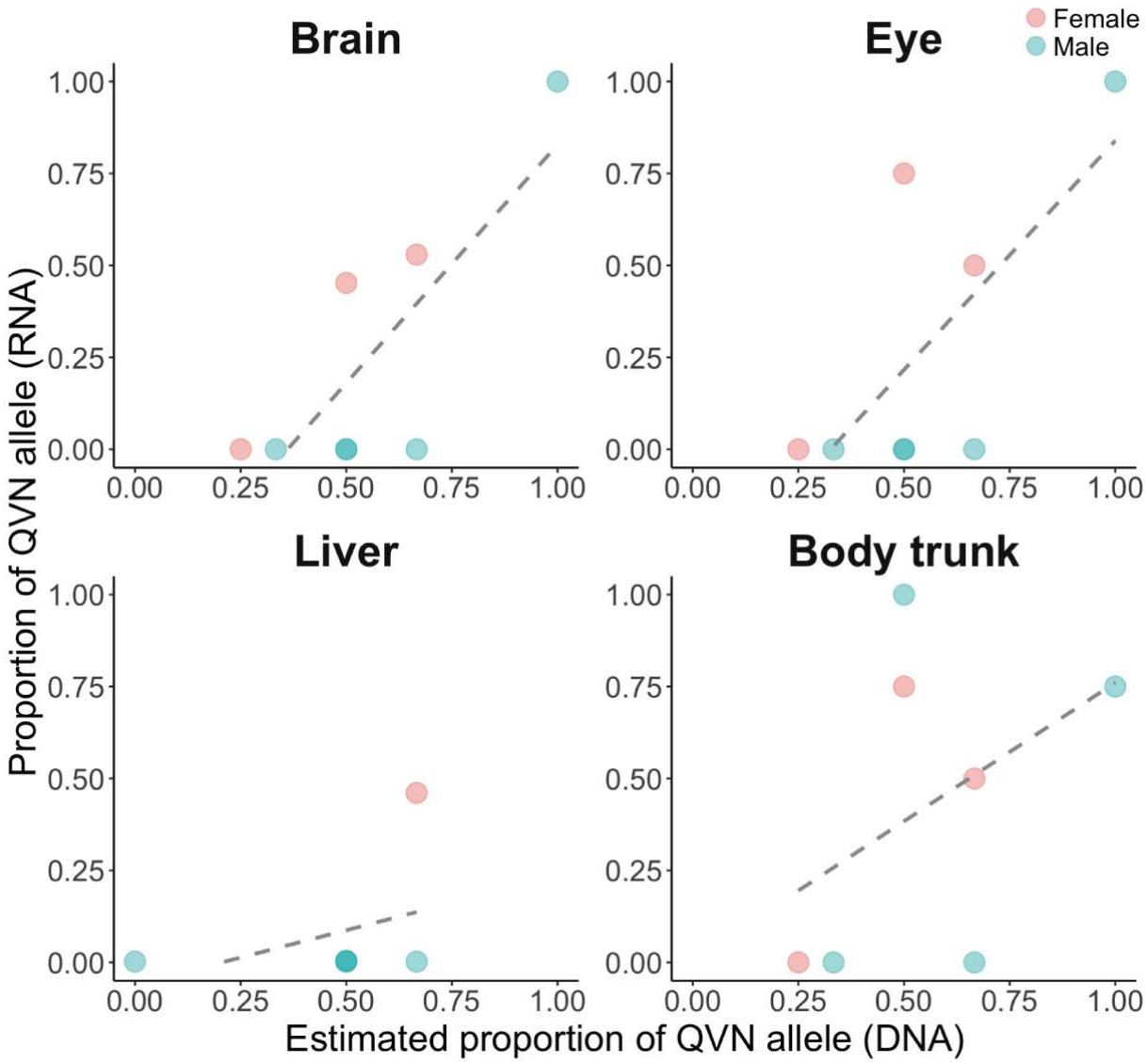

Fig. S30. Correlation between copy number and expression of *Fads2* haplotypes.

Allele-specific expression of *Fads2* haplotypes was analyzed in Pacific Ocean stickleback at 40-60 days after fertilization. The total expression levels are shown in Fig. 2E and Fig. S8. Pink and blue circles indicate female and male fish, respectively. There is a positive correlation between the proportion of *Fads2* QVN allele among all haplotypes and their expression levels in the brain (Pearson's product-moment correlation; $N = 8$, $r = 0.7953$, $P < 0.05$). A similar trend was observed in the eyes ($N = 8$, $P = 0.0524$), but not in the liver ($N = 6$, $P > 0.05$) or in the body trunk ($N = 8$, $P > 0.05$). There is also some variation in expression levels among *Fads2* QVN haplotypes within individuals and among individuals within haplotypes, suggesting that *cis*-regulatory mutations also exist in each haplotype.

Table S1.

Morphological data (mean) of Japanese threespine sticklebacks.

Population ID	Species/ ecotype/ spawning site	Collection year and month	SL	HL	BD	JL	SnL	GW	DS1	DS2	PSL	PGL	*N*	GRN	*N*	LPN	*N*	PC1 (70.94%)	PC2 (18.56%)	PC3 (4.67%)	PC4 (2.65%)	PC5 (1.75%)
FW Aizu	*G.a*/stream/FW	1969 June	45.85	16.63	11.51	4.30	4.43	3.46	3.45	4.05	5.75	8.93	8	18.0	8	32.9	8	4.04	0.43	1.21	0.42	-0.35
FW Chimikeppu	*G.a*/lake/FW	2011 May	45.73	13.66	10.09	3.17	4.11	2.58	4.89	5.22	7.09	9.61	2	20.7	22	33.0	6	4.50	-2.54	-0.48	-0.48	-0.16
FW Gensui	*G.a*/stream/FW	2012 Aug	57.12	17.57	14.47	4.44	4.83	3.54	6.10	6.70	9.24	13.81	5	20.6	5	33.2	5	-0.02	-0.77	-0.42	0.01	-0.93
FW Gifu	*G.a*/stream/FW	2009 Sep	46.71	16.84	11.90	4.75	5.25	4.18	3.25	4.14	5.41	9.23	8	14.5	8	5.1	8	4.22	3.89	-1.21	-0.08	0.45
FW Kussharo	*G.a*/crater lake/FW	2003 June	64.41	20.16	14.63	5.15	6.33	4.03	6.14	6.90	9.01	14.08	5	22.2	11	32.4	9	-1.56	0.21	0.21	0.19	0.02
FW Nasu	*G.a*/stream/FW	1967 unknown	41.88	15.06	10.89	4.19	4.53	3.53	2.93	3.53	4.69	8.79	8	18.1	8	27.4	8	5.25	0.78	1.53	0.20	0.23
FW Nishitappu*	*G.a*/stream/FW	2003 June	57.37	18.29	13.04	4.86	5.73	3.54	4.84	5.56	8.28	12.4	20	21.6	20	34.5	20	0.39	-0.15	0.52	0.48	0.06
FW Ono	*G.a*/stream/FW	2014 Nov	47.64	16.40	11.51	4.33	5.23	3.55	3.38	4.08	6.18	9.63	8	19.5	8	32.6	8	3.37	0.36	1.33	0.27	0.14
FW Shiga	*G.a*/stream/FW	1985 Apr	50.64	16.83	11.60	4.44	5.21	3.79	3.70	4.31	6.34	10.16	8	14.5	8	5.9	8	3.74	3.08	-1.73	-0.05	0.01
FW Shikotsu	*G.a*/crater lake/FW	2010 Sep	54.53	17.02	12.39	4.52	5.46	4.33	5.69	6.71	8.70	12.45	16	19.0	5	33.2	5	0.51	0.09	0.13	-1.02	-0.67
JS Abashiri	*G.n*/marine/FW-BW	2008 May	63.00	18.89	14.07	4.86	6.08	2.84	7.40	7.97	9.88	13.75	23	25.0	20	32.4	5	-1.36	-1.62	-0.79	0.86	0.60
JS Akkeshi*	*G.n*/marine/BW	2004 May	61.6	19.51	14.11	5.03	6.6	4	7.24	7.67	9.45	13.78	35-39	25.1	21	33.7	21	-1.83	-0.57	0.14	-0.27	0.69
JS Biwase	*G.n*/marine/BW	2016 May	64.89	19.66	14.44	4.65	6.67	4.54	8.23	8.75	9.85	14.17	5	22.4	5	33.2	5	-2.22	-0.24	-0.24	-1.08	-0.08
JS Fushikobetu	*G.n*/marine/ND	2004 June	62.12	19.68	14.43	5.00	6.87	4.61	6.82	7.14	8.77	13.45	8	23.0	5	33.0	5	-1.69	0.30	0.49	-0.63	0.36
JS Harutori	*G.n*/marine/BW	2006 June	61.88	18.37	13.65	4.65	5.86	4.07	7.26	7.84	8.93	13.22	7	24.7	7	33.0	7	-1.17	-0.86	-0.02	-0.72	0.49
JS Kuriyama	*G.n*/marine/FW	2006 Mar	66.82	19.83	15.05	5.13	5.64	3.03	7.16	7.36	9.92	14.31	7	21.4	16	33.9	16	-1.54	-0.77	-0.74	1.16	-0.43
JS Ogawara	*G.n*/marine/BW	2002 Mar	66.75	19.99	14.79	5.12	6.60	4.13	7.69	8.06	9.47	14.67	3	24.5	13	33.2	5	-2.41	-0.36	-0.03	-0.25	0.44
JS Onnechikappu	*G.n*/marine/BW	2001 June	60.01	18.46	13.76	4.92	5.97	3.16	6.82	7.32	9.20	13.13	4	24.5	19	32.8	5	-0.89	-1.19	-0.31	0.53	0.63
JS Sarufutu	*G.n*/marine/ND	2005 June	60.40	18.35	13.72	4.67	5.81	3.33	7.52	8.34	9.53	12.82	5	24.3	20	32.4	5	-1.03	-1.37	-0.59	-0.02	0.45

PO Akkeshi*	*G.a*/marine/FW-BW	2005 May	75.88	24.55	18.05	6.37	8.45	5.25	8.05	8.54	1.51	17.27	12-14	21.4	11	34.1	11	-5.07	1.63	0.26	0.13	-0.40
PO Choboshi	*G.a*/marine/BW	2012 May	49.09	15.27	11.77	3.60	4.41	2.95	4.69	5.69	7.90	10.75	4	22.4	5	32.6	5	2.77	-1.96	-0.19	-0.23	-0.11
PO Fushikobetu	Pacific Ocean/ND	2004 June	64.09	20.94	15.88	5.51	7.23	4.23	5.70	6.89	8.25	14.52	7	21.4	15	32.8	5	-1.99	1.00	0.62	0.49	0.09
PO Harutori	*G.a*/marine/FW	1999 May	76.46	23.11	17.15	5.82	8.35	5.73	7.51	8.01	10.45	16.46	5	21.6	5	33.8	5	-4.39	1.65	0.55	-0.54	-0.28
PO Hyotan*	*G.a*/marine/FW-BW	2003 June	57.06	19.28	13.63	4.75	6.24	3.88	5.48	6.14	8.6	13.04	24-26	22.2	20	33.9	20	-0.41	-0.05	0.49	0.03	0.12
PO Tokotan	*G.a*/marine/BW	2006 June	73.70	23.53	18.31	5.55	7.85	4.24	7.63	8.45	10.59	14.84	7	20.2	5	33.6	5	-3.83	1.02	-0.19	0.61	-0.66
PO Watarichippu	*G.a*/marine/BW	2012 May	49.37	15.48	12.13	3.72	3.95	2.89	4.94	5.77	8.65	11.11	2	21.2	5	33.2	5	2.60	-1.99	-0.51	-0.02	-0.72

FW, freshwater; JS, Japan Sea, PO, Pacific Ocean; *G.a.*, *G. aculeatus*; *G.ni.*; *G. nipponicus* (Japan Sea); BW, brackish water; SW, seawater; ND, not determined. Marine ecotypes refer to any fish that spend winter in the sea regardless of whether their spawning sites are FW, BW, or SW. SL, Standard length; HL, head length; BD, body depth; JL, upper jaw length; SnL, snout length; GW, gape width; DS1, first dorsal spine length; DS2, second dorsal spine length; PSL, pelvic spine length; PGL, pelvic girdle length; GRN, gill raker number; LPN, lateral plate number. PC, principal component with the parenthesis showing the variance explained by each PC. *, data taken from (*63*).

Table S2.

Component loadings and percentage of variance explained by each principal component (PC) for the morphology of Japanese threespine sticklebacks.

Morphology	PC1	PC2	PC3	PC4	PC5
SL	-0.338	0.038	-0.117	0.092	-0.053
HL	-0.317	0.221	0.089	0.264	-0.031
BD	-0.326	0.121	0.002	0.284	-0.282
JL	-0.278	0.331	0.152	0.409	0.232
SnL	-0.306	0.245	0.141	-0.021	0.366
GW	-0.196	0.444	0.344	-0.745	-0.086
DS1	-0.311	-0.214	-0.277	-0.227	0.067
DS2	-0.314	-0.197	-0.281	-0.237	0.041
PSL	-0.308	-0.213	-0.287	-0.066	-0.326
PGL	-0.336	-0.038	-0.095	-0.006	-0.163
GRN	-0.210	-0.473	0.205	-0.049	0.662
LPN	-0.151	-0.462	0.726	0.062	-0.379
% variance explained	70.94	18.56	4.67	2.47	1.72

For abbreviations, see Table S1.

Table S3.

Fatty acid analysis of *Artemia* nauplii.

	DHA-free	DHA-enriched
Total lipid	1.27 ± 0.01	2.29 ± 0.06
18:3n-3, ALA	36.68 ± 0.10	31.98 ± 0.14
18:4n-3, STA	6.96 ± 0.07	6.06 ± 0.03
20:4n-3, ETA	1.34 ± 0.02	1.23 ± 0.01
20:5n-3, EPA	1.23 ± 0.03	2.54 ± 0.05
22:5n-3, DPA	n.d.	n.d.
22:6n-3, DHA	n.d.	3.08 ± 0.11
Total n-3 PUFA	**46.21 ± 0.05**	**44.89 ± 0.07**
18:2n-6, LA	4.95 ± 0.05	4.46 ± 0.05
18:3n-6, GLA	0.70 ± 0.06	0.67 ± 0.01
20:3n-6, DGLA	n.d.	n.d.
20:4n-6, ARA	0.49 ± 0.02	0.97 ± 0.06
22:4n-6, DTA	0.44 ± 0.00	n.d.
22:5n-6, DPAn-6	n.d.	1.16 ± 0.03
Total n-6 PUFA	**6.58 ± 0.12**	**7.26 ± 0.14**
Total PUFA	**52.78 ± 0.09**	**52.15 ± 0.31**

Total lipid content (% of wet weight ± S.E.) and fatty acid composition (% of total fatty acids) of *Artemia* used for feeding trial (N = 3 for each group). n.d., not detected.

Table S4.

Functional characterization of stickleback *Fads2* genes in yeast.

Fads2 genotype	Δ6 desaturase (From ALA to STA)	Δ5 desaturase (From ETA to EPA)	Δ4 desaturase (From DPA to DHA)	Δ6 desaturase (From TPA to THA)
PO QVN type	1.68	n.d.	n.d.	1.50 ± 0 .31*
PO EVK type	0.91	n.d.	n.d.	0.88 ± 0.98*
JS QVN type	4.53	n.d.	n.d.	0*

Two haplotypes found only in the Pacific Ocean stickleback (PO QVN type and PO EVK type) showed Δ6 desaturase activity from 24:5n-3 to 24:6n-3 in addition to Δ6 activity from 18:3n-3 to 18:4n-3, while only the latter was observed in the Japan Sea haplotype (JS QVN type). This suggests that amino acid changes might also contribute to the higher ability to synthesize DHA in PO fish. Conversion rates are shown. n.d., not detected. *average ± S.E. from triplicate experiments.

Table S5.

Linkage analysis of *Fads2* haplotypes in *Gasterosteus wheatlandi.*

	Cross	F1xM1																						F1xM2																									*P*- value (Fisher's exact test)
	Individual ID	3	7	11	12	13	14	16	18	20	23	24	25	26	27	28	31	1	4	5	6	7	11	16	17	18	20	21	22	23	28	32	36	37	38	39	40	43	44	46	49	50								
	***Fads2* genotype**	**A/G**	**A/G**	**G**	**G**	**G**	**G**	**G**	**A**	**G**	**A/G**	**G**	**A/G**	**G**	**A/G**	**G**	**A**	**G**	**G**	**A/G**	**G**	**A**	**G**	**A/G**	**G**	**G**	**A/G**	**A/G**	**A/G**	**A/G**	**G**	**A**	**A**	**G**	**A**	**A**	**A**	**A/G**	**A/G**	**G**	**A/G**	**G**								
Microsatellite markers (LG)	Cyp19 (LG19)	b	b	a	a	a	a	a	b	a	b	a	b	a	b	a	b	a	a	a	a	b	a	a	a	a	b	a	b	b	a	b	b	a	b	b	b	a	a	a	a	a								**4.96E-08**
	Stn284 (LG19)	b	b	a	a	a	b	a	a	a	b	a	b	a	b	a	b	a	a	a	a	b	a	a	a	a	a	b	a	b	b	b	b	a	b	b	b	a	a	a	a	b								**0.001249**
	Stn235 (LG19)	b	a	a	b	a	b	b	a	a	b	b	b	a	b	a	b	a	b	b	a	a	a	b	a	a	a	b	a	b	b	b	b	a	b	b	b	b	b	a	a	b								0.07044
	Stn287 (LG12)	a	a	a	a	b	b	a	a	a	b	a	a	b	a	b	a	b	a	b	a	a	b	b	b	b	b	b	b	b	b	a	a	a	b	b	b	b	a	a	a	b								0.7204
	Stn276 (LG12)	a	a	a	a	b	b	a	a	a	b	a	a	b	a	b	a	b	a	b	a	a	b	b	b	b	b	b	b	b	b	a	a	a	b	b	b	b	a	a	a	b								0.7204
	Pun2 (LG12)	a	a	a	a	a	b	a	a	a	a	b	a	b	a	b	a	b	a	b	a	b	b	b	b	b	a	b	b	b	b	a	b	a	b	b	b	b	b	a	a	b								0.8464

a/b indicates the different genotypes.

Table S6.

Pacific Ocean-specific amino acid changes shared in other ray-finned fishes.

Species	Habitat	Freshwater +/-	C82W	A121V	L136F	A172S	A177G	V239L	I285M	T287S	W299C	Y303H	F329I	L424F
Lepisosteus oculatus	Freshwater, brackish	+		V	F	S	G	L						F
Scleropages formosus	Freshwater	+		V		S	G		M					F
Paramormyrops kingsleyae	Freshwater	+		V		S	G		M					F
Clupea harengus	Brackish, marine	-				S	G		M					
Ictalurus punctatus	Freshwater	+				S	G							
Astyanax mexicanus	Freshwater	+					G							
Pygocentrus nattereri	Freshwater	+					G							
Danio rerio	Freshwater	+					G							F
Sinocyclocheilus grahami	Freshwater, brackish	+					G							F
Cyprinus carpio	Freshwater, brackish, marine	+					G							F
Esox lucius	Freshwater, brackish	+				S	G		M					F
Salmo salar	Freshwater, brackish, marine	+		V		S	G							
Salvelinus alpinus	Freshwater, brackish, marine	+				S	G							
Oncorhynchus mykiss	Freshwater, brackish, marine	+		V		S	G							
Oncorhynchus tshawytscha	Freshwater, brackish, marine	+				S	G							
Oncorhynchus kisutch	Freshwater, brackish, marine	+		V		S	G							
Gadus morhua	Brackish, marine	-					G							
Notothenia coriiceps	Marine	-					G							
Larimichthys crocea	Brackish, marine	-					G	L						
Labrus bergylta	Marine	-					G							
Tetraodon nigroviridis	Freshwater, brackish	+												
Takifugu rubripes	Brackish, marine	-					G							
Cynoglossus semilaevis	Freshwater, brackish, marine	+							M					
Paralichthys olivaceus	Brackish, marine	-					G							
Lates calcarifer	Freshwater, brackish, marine	+					G		M					
Seriola lalandi dorsalis	Brackish, marine	-					G		M					
Seriola dumerili	Marine	-					G		M					
Stegastes partitus	Marine	-					G		M					
Amphiprion ocellaris	Marine	-					G		M					
Acanthochromis polyacanthus	Marine	-					G		M					F
Oreochromis niloticus	Freshwater, brackish	+		V			G	L	M					F
Neolamprologus brichardi	Freshwater	+		V			G	L	M					F
Pundamilia nyererei	Freshwater	+		V			G	L	M					F
Maylandia zebra	Freshwater	+		V			G	L	M					F
Haplochromis burtoni	Freshwater	+		V			G	L	M					F
Hippocampus comes	Marine	-					G		M					
Boleophthalmus pectinirostris	Marine	-		V										
Monopterus albus	Freshwater, brackish	+					G		M					
Nothobranchius furzeri	Freshwater	+		V		S	G		M		C			
Austrofundulus limnaeus	Freshwater	+		V		S	G		M					

Cyprinodon variegatus	Freshwater, brackish, marine	+	V	S	G		M
Fundulus heteroclitus	Freshwater, brackish, marine	+	V	S	G		M
Xiphophorus maculatus	Freshwater	+	V	S	G		M
Poecilia formosa	Freshwater, brackish	+	V	S	G		M
Poecilia reticulata	Freshwater, brackish	+	V	S	G		M
Oryzias melastigma	Freshwater, brackish	+		S	G	L	M
Oryzias latipes	Freshwater, brackish	+		S	G	L	M

The positions of amino acid substitutions correspond to Fig. S7. Habitat information was classified according to the Catalog of Fishes (*128*). Amino acid sequences were obtained by the Orthoscope analysis.

Table S7.
Habitat information and sample sizes used for phylogenetic analysis and *Fads2* CNV analysis.

Population ID	Species	Ecotype (spawning site)	Continent	Latitude	Longitude	Reference for the population	Collection moth and year	Phylogenic (*N*)	*Fads2* copy number analysis	
									Male (*N*)	Female (*N*)
FW Aisaka	*G. aculeatus*	Stream (FW)	Asia	40°35'31.2"N	141°13'15.6"E	(*16*, *17*)	Oct 2010		7	9
FW Aizu	*G. aculeatus*	Stream (FW)	Asia	37°30'39.6"N	139°51'57.6"E	(*16*)	May 2009	10	4	11
FW Chimikeppu	*G. aculeatus*	Lake (FW)	Asia	43°37'44.4"N	143°53'06.0"E	(*16*)	May-June 2011	2		6
FW Gensui	*G. aculeatus*	Stream (FW)	Asia	39°21'54.0"N	141°53'49.2"E	(*16*)	July 2012	3	7	9
FW Gifu	*G. aculeatus*	Stream (FW)	Asia	35°15'13.2"N	136°34'10.9"E	This study	Aug 2011	9	7	12
FW Kinugawa	*G. aculeatus*	River (FW)	Asia	36°40'35.9"N	139°56'47.8"E	This study	July 2010	4	3	10
FW Kussharo	*G. aculeatus*	Crater lake (FW)	Asia	43°36'00.0"N	144°20'52.8"E	(*16*, *17*)	May 2003		4	12
FW Nasu	*G. aculeatus*	Stream (FW)	Asia	36°49'15.4"N	140°00'30.1"E	This study	Apr 2012	8	7	13
FW Nishitappu	*G. aculeatus*	Stream (FW)	Asia	42°38'45.6"N	141°28'33.6"E	(*16*)	Sep 2010	4	5	11
FW Ono	*G. aculeatus*	Stream (FW)	Asia	35°58'31.1"N	136°29'29.6"E	This study	Dec 2014	10	9	10
FW Shiga	*G. aculeatus*	Stream (FW)	Asia	A stream in Samegai*		This study	Oct 2008/Aug 2009	4		15
FW Shikotsu	*G. aculeatus*	Crater lake (FW)	Asia	42°46'30.0"N	141°24'00.0"E	(*16*, *17*)	Oct 2010		9	7
FW Towada	*G. aculeatus*	Crater lake (FW)	Asia	40°26'42.0"N	140°50'31.2"E	(*16*, *17*)	Oct 2010		5	11
JS Abashiri	*G. nipponicus*	Marine (FW-BW)	Asia	43°57'36.0"N	44°12'00.0"E	(*16*)	May 2008	4	14	8
JS Akkeshi	*G. nipponicus*	Marine (BW)	Asia	43°04'04.8"N	144°53'02.4"E	(13, 16, *60*)	June-July 2006	9	11	5
JS Biwase	*G. nipponicus*	Marine (BW)	Asia	43°03'27.0"N	145°04'57.6"E	This study	June 2010	3	8	9
JS Fushikobetu	*G. nipponicus*	Marine (ND)	Asia	42°32'24.0"N	141°19'58.8"E	(*11*)	June 2004	3	9	6
JS Harutori	*G. nipponicus*	Marine (BW)	Asia	42°58'08.4"N	144°23'45.6"E	(*16*)	June 2010	3		9
JS Kuriyama	*G. nipponicus*	Marine (FW)	Asia	35°36'54.0"N	140°31'51.6"E	(*61*)	Feb-Apr 2005/Mar 2006	4	3	7
JS Mikata	*G. nipponicus*	Marine (BW)	Asia	35°36'54.0"N	140°31'51.6"E	This study	March 2011	2		
JS Notojima	*G. nipponicus*	Marine (ND)	Asia	37°07'37.2"N	137°02'38.4"E	(*16*)	May 2007	6		
JS Ogawara	*G. nipponicus*	Marine (BW)	Asia	40°50'24.0"N	141°22'19.2"E	(*16*)	Mar 2002	2	3	8
JS Onnechikappu	*G. nipponicus*	Marine (BW)	Asia	42°57'57.6"N	144°06'57.6"E	(*16*)	June 2001	4	10	6
JS Shinji	*G. nipponicus*	Marine (BW)	Asia	35°27'00.0"N	132°57'36.0"E	This study	Mar 2014	4		

JS Sarufutu	*G. nipponicus*	Marine (ND)	Asia	45°15'18.0"N	142°14'20.4"E	(*16*)	June 2005	4	3	13
PO Akkeshi	*G. aculeatus*	Marine (FW)	Asia	43°05'56.4"N	144°52'51.6"E	(13, *60*)	June-July 2006	10	6	15
PO Choboshi	*G. aculeatus*	Marine (BW)	Asia	43°15'28.4"N	145°33'21.6"E	(*16*)	May 2012	4	3	13
PO Fushikobetu	*G. aculeatus*	Marine (ND)	Asia	42°32'24.0"N	141°19'58.8"E	(*11*)	June 2004	4	7	9
PO Harutori	*G. aculeatus*	Marine (FW)	Asia	42°58'08.4"N	144°23'45.6"E	(*16*)	May 1999		5	4
PO Hyotan	*G. aculeatus*	Marine (FW-BW)	Asia	43°01'55.2"N	144°50'38.4"E	(*16*)	May 2011		8	6
PO Tokotan	*G. aculeatus*	Marine (BW)	Asia	42°59'56.4"N	144°52'04.8"E	This study	June 2006	3	9	7
PO Watarichippu	*G. aculeatus*	Marine (BW)	Asia	43°02'11.0"N	145°03'11.5"E	(*16*)	June 2011/May 2012	2	7	3
Allen Creek	*G. aculeatus*	Stream (FW)	North America	48°03'06.3"N	122°09'51.4"W	This study	May-June 2005		5	11
Big Soos Creek	*G. aculeatus*	Stream (FW)	North America	47°22'29.1"N	122°08'39.3"W	(*59*)	May-June 2005		6	10
Crystal lake	*G. aculeatus*	Lake (FW)	North America	49°02'40.9"N	123°57'35.2"W	(*63*)	June 2004		4	12
Duwamish River	*G. aculeatus*	Marine (BW)	North America	47°31'51.9"N	122°19'10.7"W	(*62*)	May-June 2006		14	
Evans Creek	*G. aculeatus*	Stream (FW)	North America	47°40'11.5"N	122°04'44.6"W	(*59*)	May-June 2007		5	3
Forbes Creek	*G. aculeatus*	Stream (FW)	North America	47°41'47.6"N	122°12'34.1"W	(*62*)	May-June 2005		12	4
Kelsey Creek	*G. aculeatus*	Stream (FW)	North America	47°36'05.6"N	122°10'18.8"W	(*62*)	May-June 2007		7	9
Little Campbell River Marine	*G. aculeatus*	Marine (BW)	North America	49°00'58.1"N	122°46'46.1"W	(*59*)	June 2004		10	6
Little Campbell River Stream	*G. aculeatus*	Stream (FW)	North America	49°00'43.8"N	122°37'30.8"W	(*59*)	May 2004		16	8
Manchester	*G. aculeatus*	Marine (SW)	North America	47°34'26.8"N	122°32'49.2"W	(*62*)	May-June 2006		6	10
May Creek	*G. aculeatus*	Stream (FW)	North America	47°30'54.7"N	122°08'35.3"W	(*62*)	May-June 2007		8	8
Quilcene	*G. aculeatus*	Marine (BW)	North America	47°49'40.6"N	122°51'28.2"W	(*62*)	May-June 2006		4	8
Seabeck	*G. aculeatus*	Marine (BW)	North America	47°38'25.0"N	122°49'45.1"W	(*62*)	May-June 2006		12	4
Wapato Creek	*G. aculeatus*	Stream (FW)	North America	47°13'39.0"N	122°20'33.7"W	This study	May-June 2005		7	9
Bissendorf/Weser	*G. aculeatus*	Stream (FW)	Europe	52°14'11"N	8°10'25"E	(*64*)	July 2009		4	4
Crans	*G. aculeatus*	Marine (FW)	Europe	53°32'46"N	9°45'26"E	This study	July 2009		5	3
Groenningen	*G. aculeatus*	Marine (BW)	Europe	55°27'22"N	8°21'11"E	(*67*)	July 2007/July 2008		3	5
L'isle-sur-la Sorgue	*G. aculeatus*	Stream (FW)	Europe	43°54'34"N	5° 5'10"E	(*65*)	July 2012		4	4
Le Bouveret	*G. aculeatus*	Stream (FW)	Europe	46°23'10"N	6°51'6"E	(*64*)	May-June 2007		3	5
Marina Rheinhof	*G. aculeatus*	Potamodromous (FW)	Europe	47°29'54"N	9°33'28"E	(*67*)	Apr 2013		5	3
Neustadt	*G. aculeatus*	Marine (SW, 20‰)	Europe	54° 5'44"N	10°48'35"E	(*68*)	July 2010			6

Oberriet	*G. aculeatus*	Stream (FW)	Europe	47°19'38"N	9°34'14"E	(*64*)	2009	5		3
Paelebjerg	*G. aculeatus*	Lake (FW)	Europe	55°23'19"N	8°24'33"E	This study	July 2009	4		4
Skåre	*G. aculeatus*	Marine (SW, 10-12‰)	Europe	55°22'35"N	13°3'8"E	(*66*)	Dec 2014	8		
Vikhog	*G. aculeatus*	Marine (SW, 10-12‰)	Europe	55°43'46"N	12°57'22"E	(*66*)	Feb 2015			6
P.pun Biwase	*P. pungitius*	Estuarine resident (BW)	Asia	43°03'27.0"N	145°04'57.6"E	(*26*)	May 2011/2016	5		12
P.pun Hyotan	*P. pungitius*	Estuarine resident (BW)	Asia	43°01'55.7"N	144°50'38.4"E	(*26*)	June 2011	3		9
P.pun Notsuke	*P. pungitius*	Estuarine resident (BW)	Asia	43°37'42.4"N	144°50'36.7"E	This study	Apr 2000		24	
P.pun Watarichippu	*P. pungitius*	Estuarine resident (BW)	Asia	43°02'11.0"N	145°03'11.5"E	This study	June 2011	5		5
P.pun Yaushubetsu	*P. pungitius*	Estuarine resident (BW)	Asia	43°20'34.0"N	145°13'13.5"E	This study	Apr 2000		24	
P.sin Hyotan	*P. sinensis*	Midstream (FW-BW)	Asia	43°01'22.8"N	144°51'05.3"E	(*26*)	June 2011	4		8
P.sin Akkeshi	*P. sinensis*	Midstream (FW-BW)	Asia	43°05'56.4"N	144°52'51.6"E	This study	June 2014	5		5
P.sin Kiritappu	*P. sinensis*	Midstream (FW-BW)	Asia	43°05'07.4"N	145°03'58.1"E	(*26*)	June 2011	8		4
P.sin Motoarakawa	*P. sinensis*	Stream (FW)**	Asia	36°48'00.0"N	139°14'24.0"E	(*27*)	June 2012	5		4
P.sin Oborogawa	*P. sinensis*	Stream (FW)	Asia	43°01'45.5"N	144°43'58.7"E	(*26*)	May 2011	8		8
P.sin Oborobashi	*P. sinensis*	Midstream (FW-BW)	Asia	43°04'32.1"N	144°50'16.3"E	(*26*)	June 2011	9		9
P.sin Takkobu	*P. sinensis*	Lake (FW)	Asia	43°05'59.9"N	144°29'02.0"E	This study	May 2017		16	
P.sin Yokote	*P. sinensis*	Stream (FW)	Asia	39°16'31.4"N	140°30'12.8"E	This study	Sep 1998		8	
P.tym Hyotan	*P. tymensis*	Stream (FW)	Asia	43°01'09.0"N	144°50'48.3"E	(*26*)	June 2011	5		7
P.tym Monshizu	*P. tymensis*	Pond (FW)	Asia	43°02'00.9"N	144°47'51.1"E	This study	June 2013	5		5
P.tym Oborogawa	*P. tymensis*	Stream (FW)	Asia	43°01'45.5"N	144°43'58.7"E	(*26*)	May 2011	4		10
P.kai Tendou	*P. kaibarae*	Stream (FW)	Asia	38°23'08.0"N	140°21'02.9"E	This study	Dec 2012	9		8
P.kai Yokote	*P. kaibarae*	Stream (FW)	Asia	39°16'31.4"N	140°30'12.8"E	This study	Sep 1998		8	

FW, freshwater; BW, brackish water; SW, seawater; ND, not determined. Marine ecotypes refer to any fish that spend winter in the sea, and their spawning sites are either FW, BW, SW, or in a brackish sea (e.g., Baltic Sea). *Detailed sampling point is not shown for conservation purposes. Please contact the corresponding author. **Outbred stocks maintained in the Center for Environmental Science in Saitama for conservation were used.

Reference and Notes

1. D. Schluter, *The Ecology of Adaptive Radiation*. (Oxford University Press, 2000).
2. J. B. Losos, Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* **175**, 623-639 (2010).
3. G. G. Simpson, *The Major Features of Evolution*. Columbia University biological series (Columbia University Press, 1953).
4. O. Seehausen, C. E. Wagner, Speciation in freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics* **45**, 621-651 (2014).
5. R. Betancur-R, G. Ortí, R. A. Pyron, Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecol Lett* **18**, 441-450 (2015).
6. C. E. Lee, M. A. Bell, Causes and consequences of recent freshwater invasions by saltwater animals. *Trends Ecol Evol* **14**, 284-288 (1999).
7. M. Kainz, M. T. Brett, M. T. Arts, *Lipids in Aquatic Ecosystems*. (Springer Science & Business Media, 2009).
8. M. V. Bell, D. R. Tocher, Biosynthesis of polyunsaturated fatty acids in aquatic ecosystems: general pathways and new directions. 211-236 (2009).
9. D. Swanson, R. Block, S. A. Mousa, Omega-3 fatty acids EPA and DHA: health benefits throughout life. *Advances in Nutrition*, 1-7 (2012).
10. M. A. Bell, S. A. Foster, *The Evolutionary Biology of the Threespine Stickleback*. (Oxford University Press, 1994).
11. M. Ravinet *et al.*, The genomic landscape at a late stage of stickleback speciation: high genomic divergence interspersed by small localized regions of introgression. *PLoS Genet* **14**, e1007358 (2018).
12. J. Kitano, S. Mori, C. L. Peichel, Phenotypic divergence and reproductive isolation between sympatric forms of Japanese threespine sticklebacks. *Biol J Linn Soc* **91**, 671-685 (2007).
13. J. Kitano *et al.*, A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079-1083 (2009).
14. M. Higuchi, H. Sakai, A. Goto, A new threespine stickleback, *Gasterosteus nipponicus* sp. nov. (Teleostei: Gasterosteidae), from the Japan Sea region. *Ichthyol. Res.* **61**, 341-351 (2014).
15. T. Adachi *et al.*, Shifts in morphology and diet of non-native sticklebacks introduced into Japanese crater lakes. *Ecology and Evolution* **2**, 1083-1098 (2012).
16. M. Ravinet, N. Takeuchi, M. Kume, S. Mori, J. Kitano, Comparative analysis of Japanese three-spined stickleback clades reveals the Pacific Ocean lineage has adapted to freshwater environments while the Japan Sea has not. *PLoS ONE* **9**, e112404 (2014).
17. L. M. Cassidy, M. Ravinet, S. Mori, J. Kitano, Are Japanese freshwater populations of threespine stickleback derived from the Pacific Ocean lineage? *Evol Ecol Res* **15**, 295-311 (2013).
18. Further details are available in the supplementary materials.
19. M. Kume, S. Mori, Sea-run migratory behaviour in the Japan Sea form of three-spined stickleback *Gasterosteus aculeatus* in the tidal pool of eastern Hokkaido Island, Japan. *J. Fish Biol.* **75**, 2845-2850 (2009).
20. M. T. Nakamura, T. Y. Nara, Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. *Annu Rev Nutr* **24**, 345-376 (2004).

21. D. R. Tocher, Fatty acid requirements in ontogeny of marine and freshwater fish. *Aquac Res* **41**, 717-732 (2010).
22. L. F. C. Castro, D. R. Tocher, Ó. Monroig, Long-chain polyunsaturated fatty acid biosynthesis in chordates: Insights into the evolution of Fads and Elovl gene repertoire. *Prog. Lipid Res.* **62**, 25-40 (2016).
23. K. D. Makova, R. C. Hardison, The effects of chromatin organization on variation in mutation rates in the genome. *Nat Rev Genet* **16**, 213-223 (2015).
24. G. Ortí, M. A. Bell, T. E. Reimchen, A. Meyer, Global survey of mitochondrial DNA sequences in the threespine stickleback: Evidence for recent migrations. *Evolution* **48**, 608-622 (1994).
25. O. Seehausen, Hybridization and adaptive radiation. *Trends Ecol Evol* **19**, 198-207 (2004).
26. V. V. Kapitonov, J. Jurka, Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet* **23**, 521-529 (2007).
27. A. Ishikawa *et al.*, Speciation in ninespine stickleback: reproductive isolation and phenotypic divergence among cryptic species of Japanese ninespine stickleback. *J Evol Biol* **26**, 1417-1430 (2013).
28. H. Takahashi *et al.*, Species phylogeny and diversification process of Northeast Asian *Pungitius* revealed by AFLP and mtDNA markers. *Mol Phylogenet Evol* **99**, 44-52 (2016).
29. S. Ohno, *Evolution by gene duplication*. (Springer-Verlag, 1970).
30. R. Nielsen *et al.*, Tracing the peopling of the world through genomics. *Nature* **541**, 302-310 (2017).
31. J. O. Evjemo, N. Tokle, O. Vadstein, Y. Olsen, Effect of essential dietary fatty acids on egg production and hatching success of the marine copepod Temora longicornis. *J Exp Mar Biol Ecol* **365**, 31-37 (2008).
32. B. McMeans, M. Arts, S. Rush, A. Fisk, Seasonal patterns in fatty acids of Calanus hyperboreus (Copepoda, Calanoida) from Cumberland Sound, Baffin Island, Nunavut. *Marine Biology* **159**, 1095-1105 (2012).
33. T. van der Meeren, R. E. Olsen, K. Hamre, H. J. Fyhn, Biochemical composition of copepods for evaluation of feed quality in production of juvenile marine fish. *Aquaculture* **274**, 375-397 (2008).
34. E. Prato, A. Danieli, M. Maffia, F. Biandolino, Lipid contents and fatty acid compositions of *Idotea baltica* and *Sphaeroma serratum* (Crustacea: Isopoda) as indicators of food sources. *Zool Stud* **51**, 38-50 (2012).
35. U. Braeckman, P. Provoost, K. Sabbe, K. Soetaert, J. Middelburg, M. Vincx, J. Vanaverbeke, Temporal dynamics in the diet of two marine polychaetes as inferred from fatty acid biomarkers. *J Sea Res* **68**, 6-19 (2012).
36. V. I. Kharlamenko, S. I. Kiyashko, S. A. Rodkina, A. B. Imbs, Determination of food sources of marine invertebrates from a subtidal sand community using analyses of fatty acids and stable isotopes. *Russ J Mar Biol* **34**, 101-109 (2008).
37. C. M. C. Woods, Caprellid amphipods: An overlooked marine finfish aquaculture resource? *Aquaculture* **289**, 199-211 (2009).
38. C. F. Phleger, M. M. Nelson, B. D. Mooney, P. D. Nichols, Interannual and between species comparison of the lipids, fatty acids and sterols of Antarctic krill from the US

AMLR Elephant Island survey area. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **131**, 733-747 (2002).
39. P. Virtue, P. Mayzaud, E. Albessard, P. Nichols, Use of fatty acids as dietary indicators in northern krill, Meganyctiphanes norvegica, from northeastern Atlantic, Kattegat, and Mediterranean waters. *Can J Fish Aquat Sci* **57**, 104-114 (2000).
40. M. Patrick, V. Patti, A. Eric, Seasonal variations in the lipid and fatty acid composition of the euphausiid *Meganyctiphanes norvegica* from the Ligurian Sea. *Mar Ecol Prog Ser* **186**, 199-210 (1999).
41. E. Baeza-Rojano, I. Hachero-Cruzado, J. M. Guerra-García, Nutritional analysis of freshwater and marine amphipods from the strait of gibraltar and potential aquaculture applications. *J Sea Res* **85**, 29-36 (2014).
42. Y. Ando, M. Nozaki, Lipid profile of mysid *Neomysis intermedia* and its changes by feeding DHA-rich marine lipids. *Journal of Oleo Science* **56**, 471-478 (2007).
43. D. Mohan Stephanie, L. Connelly Tara, M. Harris Carolynn, H. Dunton Kenneth, W. McClelland James, Seasonal trophic linkages in Arctic marine invertebrates assessed via fatty acids and compound-specific stable isotopes. *Ecosphere* **7**, e01429 (2016).
44. R. J. Shields, J. G. Bell, F. S. Luizi, B. Gara, N. R. Bromage, J. R. Sargent, Natural copepods are superior to enriched artemia nauplii as feed for halibut larvae (*Hippoglossus hippoglossus*) in terms of survival, pigmentation and retinal morphology: relation to dietary essential fatty acids. *The Journal of Nutrition* **129**, 1186-1194 (1999).
45. H. Saito, Y. Kotani, J. M. Keriko, C. Xue, K. Taki, K. Ishihara, T. Ueda, S. Miyata, High levels of n-3 polyunsaturated fatty acids in Euphausia pacifica and its role as a source of docosahexaenoic and icosapentaenoic acids for higher trophic levels. *Mar Chem* **78**, 9-28 (2002).
46. O. Makhutova, N. Sushchik, M. Gladyshev, A. V Ageev, E. Pryanichnikova, G. S Kalachova, Is the fatty acid composition of freshwater zoobenthic invertebrates controlled by phylogenetic or trophic factors? *Lipids* **46**, 709-721 (2011).
47. H. L. Mariash, M. Cusson, M. Rautio, Fall composition of storage lipids is associated with the overwintering strategy of *Daphnia*. *Lipids* **52**, 83-91 (2017).
48. L. Ravet Joseph, T. Brett Michael, B. Arhonditsis George, The effects of seston lipids on zooplankton fatty acid composition in Lake Washington, Washington, USA. *Ecology* **91**, 180-190 (2010).
49. O. N. Makhutova, M. I. Gladyshev, N. N. Sushchik, O. P. Dubovskaya, Z. F. Buseva, E. B. Fefilova, V. P. Semenchenko, G. S. Kalachova, O. N. Kononova, M. A. Baturina, Comparison of fatty acid composition of cladocerans and copepods from lakes of different climatic zones. *Contemporary Problems of Ecology* **7**, 474-483 (2014).
50. N. N. Sushchik, M. I. Gladyshev, A. V. Moskvichova, O. N. Makhutova, G. S. Kalachova, Comparison of fatty acid composition in major lipid classes of the dominant benthic invertebrates of the Yenisei river. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **134**, 111-122 (2003).
51. O. N. Makhutova, S. P. Shulepina, T. A. Sharapova, O. P. Dubovskaya, N. N. Sushchik, M. A. Baturina, E. G. Pryanichnikova, G. S. Kalachova, M. I. Gladyshev, Content of polyunsaturated fatty acids essential for fish nutrition in zoobenthos species. *Freshwater Science* **35**, 1222-1234 (2016).
52. M. I. Gladyshev, N. N. Sushchik, S. P. Shulepina, A. V. Ageev, O. P. Dubovskaya, A. A. Kolmakova, G. S. Kalachova, Secondary production of highly unsaturated fatty acids by

zoobenthos across rivers contrasting in temperature. *River Res Appl* **32**, 1252-1263 (2016).
53. O. Hessen Dag, E. V. A. Leu, Trophic transfer and trophic modification of fatty acids in high Arctic lakes. *Freshwater Biol* **51**, 1987-1998 (2006).
54. J. Persson, T. Vrede, Polyunsaturated fatty acids in zooplankton: variation due to taxonomy and trophic position. *Freshwater Biol* **51**, 887-900 (2006).
55. W. Kolanowski, A. Stolyhwo, M. Grabowski, Fatty acid composition of selected fresh water Gammarids (*Amphipoda, Crustacea*): A potentially innovative source of omega-3 LC PUFA. *Journal of the American Oil Chemists' Society* **84**, 827-833 (2007).
56. P. Mayzaud, P. Virtue, E. Albessard, Seasonal variations in the lipid and fatty acid composition of the euphausiid *Meganyctiphanes norvegica* from the Ligurian Sea. *Mar Ecol Prog Ser* **186**, 199-210 (1999).
57. W. Twining Cornelia, J. T. Brenna, G. Hairston Nelson, S. Flecker Alexander, Highly unsaturated fatty acids in nature: what we know and what we need to learn. *Oikos* **125**, 749-760 (2015).
58. J. Kitano, D. I. Bolnick, D. A. Beauchamp, M. M. Mazur, S. Mori, T. Nakano, C. L. Peichel, Reverse evolution of armor plates in the threespine stickleback. *Curr. Biol.* **18**, 769-774 (2008).
59. J. Kitano, S. C. Lema, J. A. Luckenbach, S. Mori, Y. Kawagishi, M. Kusakabe, P. Swanson, C. L. Peichel, Adaptive divergence in the thyroid hormone signaling pathway in the stickleback radiation. *Curr Biol* **20**, 2124-2130 (2010).
60. M. Kume, J. Kitano, S. Mori, T. Shibuya, Ecological divergence and habitat isolation between two migratory forms of Japanese threespine stickleback (*Gasterosteus aculeatus*). *J Evol Biol* **23**, 1436-1446 (2010).
61. M. Kume, Japan Sea form of threespine stickleback, *Gasterosteus aculeatus*, from the Isumi and Kuriyama Rivers in Chiba Prefecture: possibility of the southernmost population. *Biol Inland Waters* **23**, 21-26 (2008).
62. J. Kitano, S. Mori, C. L. Peichel, Divergence of male courtship displays between sympatric forms of anadromous threespine stickleback. *Behaviour* **145**, 443-461 (2008).
63. J. Kitano, S. Mori, C. L. Peichel, Sexual dimorphism in the external morphology of the threespine stickleback (*Gasterosteus aculeatus*). *Copeia*, 336-349 (2007).
64. K. A. Y. Lucek, D. Roy, E. Bezault, A. Sivasundar, O. L. E. Seehausen, Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Mol. Ecol.* **19**, 3995-4011 (2010).
65. K. Lucek, O. Seehausen, Distinctive insular forms of threespine stickleback (*Gasterosteus aculeatus*) from western Mediterranean islands. *Conserv Genet* **16**, 1319-1333 (2015).
66. S. S. Madsen, C. Weber, A. M. Nielsen, M. Mohiseni, M. C. Bosssus, C. K. Tipsmark, B. Borg, Sexual maturation and changes in water and salt transport components in the kidney and intestine of three-spined stickleback (*Gasterosteus aculeatus* L.). *Comp. Biochem. Physiol., Part A Mol. Integr. Physiol.* **188**, 107-119 (2015).
67. D. A. Marques, K. Lucek, J. I. Meier, S. Mwaiko, C. E. Wagner, L. Excoffier, O. Seehausen, Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genet* **12**, e1005887 (2016).
68. D. A. Marques, thesis, University of Bern, (2016).

69. C. L. Peichel, J. A. Ross, C. K. Matson, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, S. Mori, D. Schluter, D. M. Kingsley, The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr Biol* **14**, 1416-1424 (2004).
70. K. Yoshida, T. Makino, J. Kitano, Accumulation of deleterious mutations on the neo-Y chromosome of Japan Sea stickleback (*Gasterosteus nipponicus*). *J Hered* **108**, 63-68 (2017).
71. K. Yoshida, T. Makino, K. Yamaguchi, S. Shigenobu, M. Hasebe, M. Kawata, M. Kume, S. Mori, C. L. Peichel, A. Toyoda, A. Fujiyama, J. Kitano, Sex chromosome turnover contributes to genomic divergence between incipient stickleback species. *PLoS Genet* **10**, e1004223 (2014).
72. M. A. Bell, W. E. Aguirre, N. J. Buck, Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**, 814-824 (2004).
73. R Development Core Team. (2013).
74. S. Sakaguchi, T. Sugino, Y. Tsumura, M. Ito, M. D. Crisp, D. M. J. S. Bowman, A. J. Nagano, M. N. Honjo, M. Yasugi, H. Kudoh, Y. Matsuki, Y. Suyama, Y. Isagi, High-throughput linkage mapping of Australian white cypress pine (*Callitris glaucophylla*) and map transferability to related species. *Tree Genetics & Genomes* **11**, 121 (2015).
75. K. Yoshida, R. Miyagi, S. Mori, A. Takahashi, T. Makino, A. Toyoda, A. Fujiyama, J. Kitano, Whole-genome sequencing reveals small genomic regions of introgression in an introduced crater lake population of threespine stickleback. *Ecol Evol* **6**, 2190-2204 (2016).
76. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, (2014).
77. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
78. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. *arXiv* **1207**, (2014).
79. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
80. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).
81. A. D. Leaché, B. L. Banbury, J. Felsenstein, A. n.-M. de Oca, A. Stamatakis, Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Syst Biol* **64**, 1032-1047 (2015).
82. N. Kabeya, Y. Takeuchi, R. Yazawa, Y. Haga, S. Satoh, G. Yoshizaki, Transgenic modification of the n-3 HUFA biosynthetic pathway in nibe croaker larvae: improved DPA (docosapentaenoic acid; 22:5n-3) production. *Aquac Nutr* **22**, 472-478 (2016).
83. S. N. Ladd, N. Dubois, C. J. Schubert, Interplay of community dynamics, temperature, and productivity on the hydrogen isotope signatures of lipid biomarkers. *Biogeosciences* **14**, 3979-3994 (2017).
84. F. C. Jones, M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, E. S. Lander, F. Di Palma, K. Lindblad-

Toh, D. M. Kingsley, The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55-61 (2012).
85. N. Kabeya, Y. Yamamoto, S. Cummins, A. Elizur, R. Yazawa, Y. Takeuchi, Y. Haga, S. Satoh, G. Yoshizaki, Polyunsaturated fatty acid metabolism in a marine teleost, Nibe croaker Nibea mitsukurii: Functional characterization of Fads2 desaturase and Elovl5 and Elovl4 elongases. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **188**, 37-45 (2015).
86. A. Oboh, N. Kabeya, G. Carmona-Antoñanzas, L. F. C. Castro, J. R. Dick, D. R. Tocher, O. Monroig, Two alternative pathways for docosahexaenoic acid (DHA, 22:6n-3) biosynthesis are widespread among teleost fish. *Scientific Reports* **7**, 3889 (2017).
87. K. E. Hosemann, P. F. Colosimo, R. S. Brian, D. M. Kingsley, A simple and efficient microinjection protocol for making transgenic sticklebacks. *Behaviour* **141**, 1345-1355 (2004).
88. Y. F. Chan, M. E. Marks, F. C. Jones, G. Villarreal, M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, D. Petrov, B. Jónsson, D. Schluter, M. A. Bell, D. M. Kingsley, Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302-305 (2010).
89. K. Kawakami, Tol2: a versatile gene transfer vector in vertebrates. *Genome Biol* **8**, S7 (2007).
90. L. Ballard, P. Adams, Y. Bao, D. Bartley, D. Bintzler, L. Kasch, L. Petukhova, C. Rosato, Strategies for genotyping: Effectiveness of tailing primers to increase accuracy in short tandem repeat determinations. *Journal of Biomolecular Techniques* **13**, 20–29 (2002).
91. T. Adachi, A. Ishikawa, S. Mori, W. Makino, M. Kume, M. Kawata, J. Kitano, Shifts in morphology and diet of non-native sticklebacks introduced into Japanese crater lakes. *Ecology and Evolution* **2**, 1083-1098 (2012).
92. D. Zhang. (2018).
93. J. A. Ross, J. R. Urton, J. Boland, M. D. Shapiro, C. L. Peichel, Turnover of sex chromosomes in the sickleback sishes (Gasterosteidae). *PLoS Genet* **5**, e1000391 (2009).
94. J. N. Cech, C. L. Peichel, Identification of the centromeric repeat in the threespine stickleback fish (*Gasterosteus aculeatus*). *Chromosome Res.* **23**, 767-779 (2015).
95. J. A. Ross, C. L. Peichel, Molecular cytogenetice evidence of rearrangements on the Y chromosome of the threespine stickleback fish. *Genetics* **179**, 2173 (2008).
96. J. N. Cech, C. L. Peichel, Centromere inactivation on a neo-Y fusion chromosome in threespine stickleback fish. *Chromosome Res.* **24**, 437-450 (2016).
97. L. Noé, G. Kucherov, YASS: enhancing the sensitivity of DNA similarity search. *Nucleic Acids Res.* **33**, W540-W543 (2005).
98. O. Kohany, A. J. Gentles, L. Hankus, J. Jurka, Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* **7**, 474 (2006).
99. J. Kitano, J. A. Ross, S. Mori, M. Kume, F. C. Jones, Y. F. Chan, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, D. M. Kingsley, C. L. Peichel, A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079-1083 (2009).
100. C. L. Peichel, K. S. Nereng, K. A. Ohgi, B. L. E. Cole, P. F. Colosimo, C. A. Buerkle, D. Schluter, D. M. Kingsley, The genetic architecture of divergence between threespine stickleback species. *Nature* **414**, 901-905 (2001).

101. M. D. Shapiro, B. R. Summers, S. Balabhadra, J. T. Aldenhoven, A. L. Miller, C. B. Cunningham, M. A. Bell, D. M. Kingsley, The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. *Curr. Biol.* **19**, 1140-1145 (2009).
102. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **1**, (2015).
103. M. A. Larkin, G. Blackshields, N. P. Brown, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
104. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797 (2004).
105. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870-1874 (2016).
106. R. Bouckaert, J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, A. J. Drummond, BEAST 2: A software platform for bayesian evolutionary analysis. *PLoS Computational Biology* **10**, e1003537 (2014).
107. B. Guo, F. J. J. Chain, E. Bornberg-Bauer, E. H. Leder, J. Merilä, Genomic divergence between nine- and three-spined sticklebacks. *BMC Genomics* **14**, 756 (2013).
108. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
109. J. Catchen, P. A. Hohenlohe, S. Bassham, A. Amores, W. A. Cresko, Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124-3140 (2013).
110. J. W. Van Ooijen, *JoinMap® 4.0: software for the calculation of genetic linkage maps in experimental population. Kyazma BV.* (2006).
111. N. E. Morton, Significance Levels in Complex Inheritance. *The American Journal of Human Genetics* **62**, 690-697 (1998).
112. M. A. White, J. Kitano, C. L. Peichel, Purifying selection maintains dosage-sensitive genes during degeneration of the threespine stickleback Y chromosome. *Mol. Biol. Evol.* **32**, 1981-1995 (2015).
113. R. V. Lenth, Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software; Vol 1, Issue 1 (2016)*, (2016).
114. J. Inoue, N. Satoh, ORTHOSCOPE: An automatic web tool for phylogenetically inferring bilaterian orthogroups with user-selected taxa. *Mol. Biol. Evol.* **36**, 621-631 (2018).
115. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780 (2013).
116. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford, England)* **25**, 1972-1973 (2009).
117. M. Suyama, D. Torrents, P. Bork, PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609-W612 (2006).
118. Z. Yang, Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* **39**, 105-111 (1994).
119. Z. Yang, Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* **39**, 306-314 (1994).

120. J. D. Hadfield, S. Nakagawa, General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J Evol Biol* **23**, 494-508 (2010).
121. J. D. Hadfield, MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *2010* **33**, 22 (2010).
122. R. Betancur-R, E. O. Wiley, G. Arratia, A. Acero, N. Bailly, M. Miya, G. Lecointre, G. Ortí, Phylogenetic classification of bony fishes. *BMC Evol Biol* **17**, 162 (2017).
123. B. Ma, H. Jiang, P. Sun, J. Chen, L. Li, X. Zhang, L. Yuan, Phylogeny and dating of divergences within the genus Thymallus (Salmonidae: Thymallinae) using complete mitochondrial genomes. *Mitochondrial DNA Part A* **27**, 3602-3611 (2016).
124. B. L. Swart, S. von der Heyden, A. Bester-van der Merwe, R. Roodt-Wilding, Molecular systematics and biogeography of the circumglobally distributed genus Seriola (Pisces: Carangidae). *Mol. Phylogenet. Evol.* **93**, 274-280 (2015).
125. H. E. Setiamarga Davin, M. Miya, Y. Yamanoue, Y. Azuma, G. Inoue Jun, B. Ishiguro Naoya, K. Mabuchi, M. Nishida, Divergence time of the two regional medaka populations in Japan as a new time scale for comparative genomics of vertebrates. *Biol Lett* **5**, 812-816 (2009).
126. M. D. Sanciangco, K. E. Carpenter, R. Betancur-R, Phylogenetic placement of enigmatic percomorph families (Teleostei: Percomorphaceae). *Mol. Phylogenet. Evol.* **94**, 565-576 (2016).
127. K. Tamura, F. U. Battistuzzi, P. Billing-Ross, O. Murillo, A. Filipski, S. Kumar, Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19333 (2012).
128. R. Fricke, W. N. Eschmeyer, R. v. d. Laan, *Eschmeyer's catalog of fishes: genera, species, references*. (Available at: https://www.calacademy.org/scientists/projects/catalog-of-fishes Last accessed 15 Jan 2019., 2019).
129. I. Braasch, J. H. Postlethwait, in *Polyploidy and Genome Evolution,* P. S. Soltis, D. E. Soltis, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), pp. 341-383.
130. D. R. Tocher, Fatty acid requirements in ontogeny of marine and freshwater fish. *Aquac Res* **41**, 717-732 (2010).
131. L. F. Castro, O. Monroig, M. J. Leaver, J. Wilson, I. Cunha, D. R. Tocher, Functional desaturase Fads1 (Delta5) and Fads2 (Delta6) orthologues evolved before the origin of jawed vertebrates. *PLoS ONE* **7**, e31950 (2012).
132. M. T. Nakamura, T. Y. Nara, Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. *Annu Rev Nutr* **24**, 345-376 (2004).