DNA methylation site loss for plasticity-led novel trait genetic fixation

Takafumi Katsumura[1,2,*], Suguru Sato[3], Kana Yamashita[3], Shoji Oda[4], Takashi Gakuhari[5], Shodai Tanaka[3], Kazuko Fujitani[6], Toshiyuki Nishimaki[1], Tadashi Imai[7], Yasutoshi Yoshiura[7], Hirohiko Takeshima[8], Yasuyuki Hashiguchi[9], Hiroshi Mitani[4], Motoyuki Ogawa[1], Hideaki Takeuchi[2] and Hiroki Oota[1,10]

[1]Department of Anatomy, Kitasato University School of Medicine, Japan

[2]The Graduate School of Natural Science and Technology, Okayama University, Japan

[3]Department of Biological Sciences, Kitasato University School of Science, Japan

[4]Department of Integrated Biosciences, The University of Tokyo, Japan

[5]Institute of Human and Social Sciences, Kanazawa University, Japan

[6]Gene Analysis Center, Kitasato University School of Medicine, Japan

[7]National Research Institute of Fisheries and Environment of Inland Sea, Japan Fisheries Research and Education Agency, Japan

[8]Department of Marine Biology, Tokai University School of Biological Sciences, Japan

[9]Department of Biology, Osaka Medical College, Japan

[10]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Japan

*Correspondence to:

Takafumi Katsumura, Department of Anatomy, Kitasato University School of Medicine, Japan, +81 42 778 9022, tk@med.kitasato-u.ac.jp

Abstract

Many organisms exhibit phenotypic plasticity that changes their traits in response to their environment. Although whether or not this plasticity contributes to adaptive evolution is a fundamental question in evolutionary biology, various studies report that natural populations adapt to rapid environmental changes via plasticity, which leads to novel adaptive traits as "triggers." Namely, phenotypic plasticity has considered allowing an accumulation of genetic mutations to fix the alternative phenotypes induced by nongenetic perturbations that include gene expression noise or epigenetic modification caused by environmental change. However, because the molecular mechanism of phenotypic plasticity is unknown, verification of the process from phenotypic plasticity to genetic fixation remains insufficient. Here we show that decrease in methylated CpG sites leads to loss of plasticity, which triggers genetic fixation of novel traits, in medaka fish (*Oryzias latipes*). We found that the gut length was correlated with the number of methylated CpG sites upstream of the *Plxnb3* gene, which is involved in the developmental process of nerve axons. The medaka, in which the methylated DNA region is deleted by CRISPR/Cas9, showed a loss of plasticity in gut length and a lower survival rate caused by nonoptimal feeding environments. Moreover, standing variation in the promoter region of another gene, *Ppp3r1*, which is also related to nerve axon development, raised the gene expression and made a longer gut stably in wild medaka groups that lost the gut-length plasticity. Furthermore, our phylogenetic analysis revealed the timing of these evolutionary events, indicating that the loss of phenotypic plasticity by nucleotide substitutions initiates the process of genetic fixation of the novel trait. That is, while phenotypic plasticity plays a role as a buffer of evolution and contributes to environmental adaptation, as previously thought, our molecular data suggest that mutation on CpG site causing the loss of phenotypic plasticity, is the trigger for a generation of novel traits.

Introduction

Genetic and environmental factors that alter a phenotype are driving-forces leading to novel traits. In current biology, a novel mutation that is considered to cause a novel trait spreads into a population through

generations if it is advantageous. Meanwhile, traits often show variances induced by environmental changes (phenotypic plasticity[1]). The altered trait, being exposed under selective pressure over generations, can be expressed regardless of environmental factors by subsequent genetic mutations through generations, even in nature[2]. This process is called, "plasticity-first" evolution (PFE), which is a candidate for an alternative pathway of adaptive evolution[3]. However, it has been discussed over the years whether the existing evolutionary framework can explain the PFE process[4] because the molecular mechanism of how novel traits that originated from plasticity lead to genetic fixation remains unknown.

To investigate the PFE process from plasticity to genetic fixation and the molecular mechanism involved in the process, it is rational to consider this process as taking place in two steps: (1) loss of phenotypic plasticity and (2) genetic fixation of an induced trait. The expression of phenotypic plasticity is considered to occur as a result of changes in gene expression patterns caused by epigenetic changes (such as DNA methylation[5, 6]). Moreover, the PFE hypothesis states that following random mutations fix plastic traits induced by environmental change and are transmitted to offspring. If this hypothesis is true, it is highly possible that the genetic mutation inducing the novel trait exists in the epigenetic modification region, which is related to the expression of phenotypic plasticity. Therefore, we examined the PFE process from the plastic trait to the genetic trait by finding the molecular basis of phenotypic plasticity and the genetic variation that fixes novel traits. To this end, we focus on gut length, which shows typical phenotypic plasticity and genetic variation in various animals[7, 8].

The medaka (*Oryzias latipes*), also known as Japanese rice fish, is an excellent model to investigate the molecular mechanism of PFE, because this small, omnivorous freshwater fish inhabits low to high latitudes in East Asia, adapts genetically to the various environments of each region, and shows geographic phenotypic plasticity and latitudinal variations for the traits associated with gut length[9, 10]. Moreover, the medakas' evolutionary history is well-studied[11, 12], and various molecular biological techniques can be adapted to medaka[13]. Thus, we aimed to show the PFE process underlying their gut length differences and reveal the molecular basis and evolutionary process from the loss of plasticity to genetic fixation.

This knowledge, coupled with two genome-wide approaches and subsequent genome-editing, allowed us to describe the molecular mechanism of PFE underlying gut-length variations of local medaka populations. First, we found that the evolution of the medaka gut fulfills the criteria of PFE by measuring the gut-length, which revealed seasonal and local geographical variations. Second, we examined each molecular mechanism involved in the two steps of the PFE process by genome-wide methylation and population genomic analyses. Finally, by analyzing two genes that we identified in the PFE process, we revealed that a longer gut became genetically expressed through fixation of mutations after the loss of gut-length plasticity due to a decrease in CpG sites. The new finding is that PFE is not a process of genetically fixing phenotypic plasticity by the following-mutation, but a process in which plasticity, a buffer that reduces environmental pressure, is lost by mutation and often replaced by an existing favorable mutation.

Results

We first verified whether the gut of medaka shows the phenotypic plasticity on ancestral lineages and genetic polymorphism on derived ones, both of which are requirements of PFE, respectively[3]. We investigated the gut length of wild medaka in three rivers, the Kabe, Ejiri, and Shihodo, in Kagawa Prefecture, Japan, over a 2-year period (Fig. 1A, fig. S1, and table S1), and found seasonal changes in the medaka gut length in the Kabe and Ejiri rivers. Medaka in the winter showed significantly shorter guts than did those in the summer, which was observed over that same 2-year period in the Kabe river (Fig. 1B, Bayesian generalized linear mixed model [GLMM], estimate = -0.44±0.05, 95% credible interval [CI] did not include zero, table S2). This seasonal variation in gut length must indicate phenotypic plasticity in response to the environment because no noticeable genetic changes or gene flow from different populations were detected between the summer and winter (fig. S2). To determine what environmental changes caused seasonal plasticity in gut length, we conducted stable isotope analysis and outdoor breeding experiments on captured wild medaka. The stable isotope analysis based on $\delta^{13}C$ and $\delta^{15}N$ showed no correlation between

gut length and food diversity or carnivorous habits (fig. S3). The outdoor breeding experiment using captured wild medaka from the Kabe river showed the same plasticity even under the condition of feeding with artificial food (fig. S4), indicating no relationship between the plasticity and the type of food.

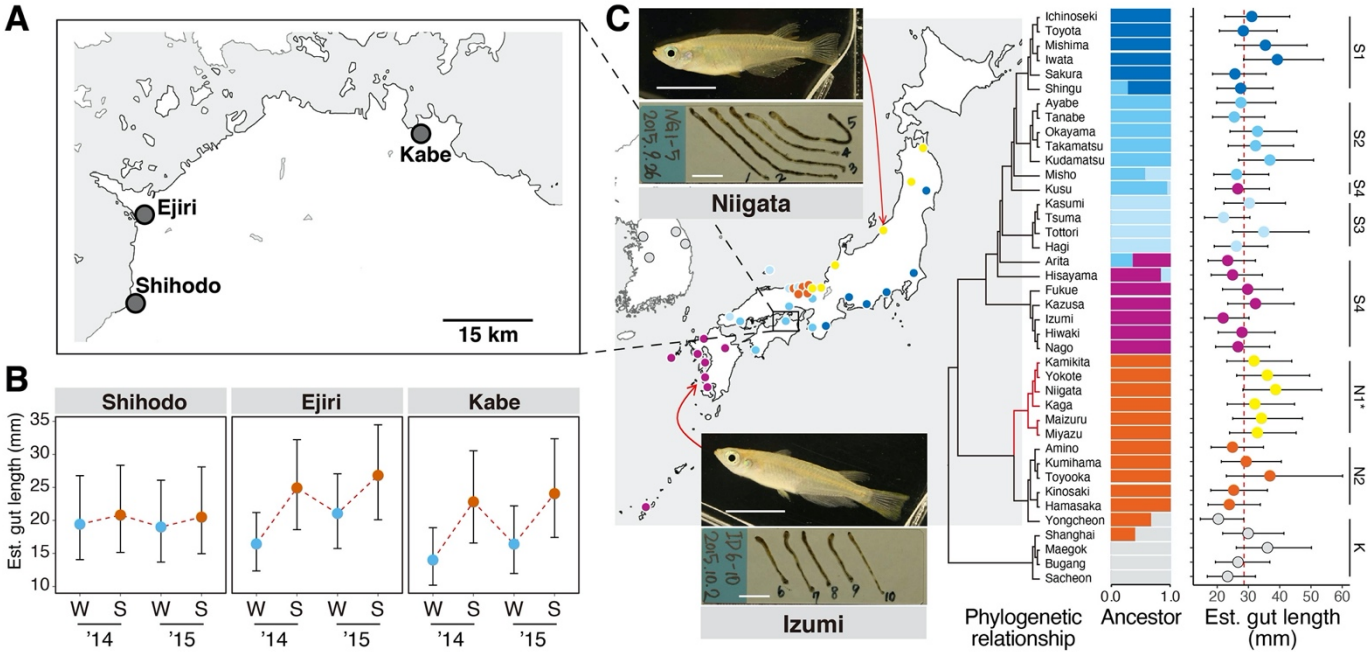


Figure 1 | Phenotypic plasticity and genetic variation of the gut length in medaka. (A) A map of sampling locations (light blue circles) in Kagawa Prefecture, Japan. (B) Seasonal gut length changes in the three rivers. (C) Geographic variations of gut length in 40 local populations with example photos (white bars: 1 cm). Phylogenetic tree and admixture graphs were constructed using genome-wide single nucleotide polymorphisms (SNPs) published in our prior study[11]. Points and error bars were fitted values with 95% credible intervals from the Bayesian GLMM. Abbreviations on the far right: S1–S4, N1, 2 and K are: SJPN1–4, NJPN1, 2 and KOR, respectively. The asterisk indicates that the NJPN1 medaka have a significantly longer gut than do the other subgroups.

To identify genetic polymorphisms concerning the variation of the gut in medaka, gut length was

measured in a wild-derived laboratory stocks consisting of 81 geographical populations[14]. Our previous genome-wide single nucleotide polymorphism (SNP) study showed that this stock genetically consists of seven subgroups (S1-4, N1, 2, and K) derived from three lineages from Korea and Japan[11]. Based on the population structure of wild medaka-derived laboratory stocks, we sampled 10 individuals from each of 40 populations to avoid bias from population structures (table S3). All local medaka populations were reared in the same environment. Nevertheless, the NJPN1 lineage showed a significantly longer gut compared to the average body length than did those of the other lineages (Fig. 1C, Bayesian GLMM, estimate = -0.11 – -0.26, 95% CIs did not include zero, table S4), indicating that a longer gut length of the NJPN1 lineage is determined by genetic factors. Meanwhile, the gut lengths within the SJPN and KOR groups vary, strongly suggesting that the gut length diversity in the SJPN and KOR types is defined by environmental factors rather than genetic factors. Thus, all medaka populations except the NJPN1 type show gut length variations due to phenotypic plasticity. Considering the seasonal plasticity of Kagawa medaka (including those in the Kabe river [Fig. 1B], which are in the SJPN group[9]), we found the evolution of the medaka gut fulfilled the PFE requirements. In that evolutionary framework, those medaka originally exhibit phenotypic plasticity in gut length, whereas the derived medaka subgroup has acquired a longer gut length for adaptation.

Under the PFE hypothesis, DNA methylation occurs as environmental responses and is stable over a long time[15], making it a potential candidate as a molecular basis of seasonal phenotypic plasticity. To test whether the DNA methylation is related to gut length plasticity, methyl-CpG binding domain sequencing (MBD-seq) was used to detect the methylation of DNA extracted from the gut of Kabe medakas which we observed as exhibiting the strongest seasonal plasticity (Fig. 1B). Exploring winter and summer methylated CpG sites revealed a total of 206 seasonally varying methylated regions, 14 regions (winter methylation), and 192 regions (summer methylation). A total 71 genes were present in the regions within 2 kb upstream and downstream (two genes methylated in winter, 69 genes methylated in summer) (Fig. 2A, table S5). We analyzed the gene lists using Reactome (https://reactome.org/) and Gene Ontology (GO) (http://geneontology.org/) analyses, and found that the semaphorin-plexin signaling related genes (*Plxna3*,

Plxna4, and *Plxnb3*) concerning the axon elongation was significantly enriched (Fig. 2A, table S6). Notably, the *Plxnb3* gene, which is involved in the inhibition of neurite outgrowth, had a CpG island (CGI) in the upstream region and showed seasonal methylation (Fig. 2B). A real-time quantitative polymerase chain reaction (PCR) showed that the gene expression of *Plxnb3* in summer was lower than that in winter ($p<0.01$), indicating those methylations were associated with controlling the seasonal plasticity (Fig. 2C).

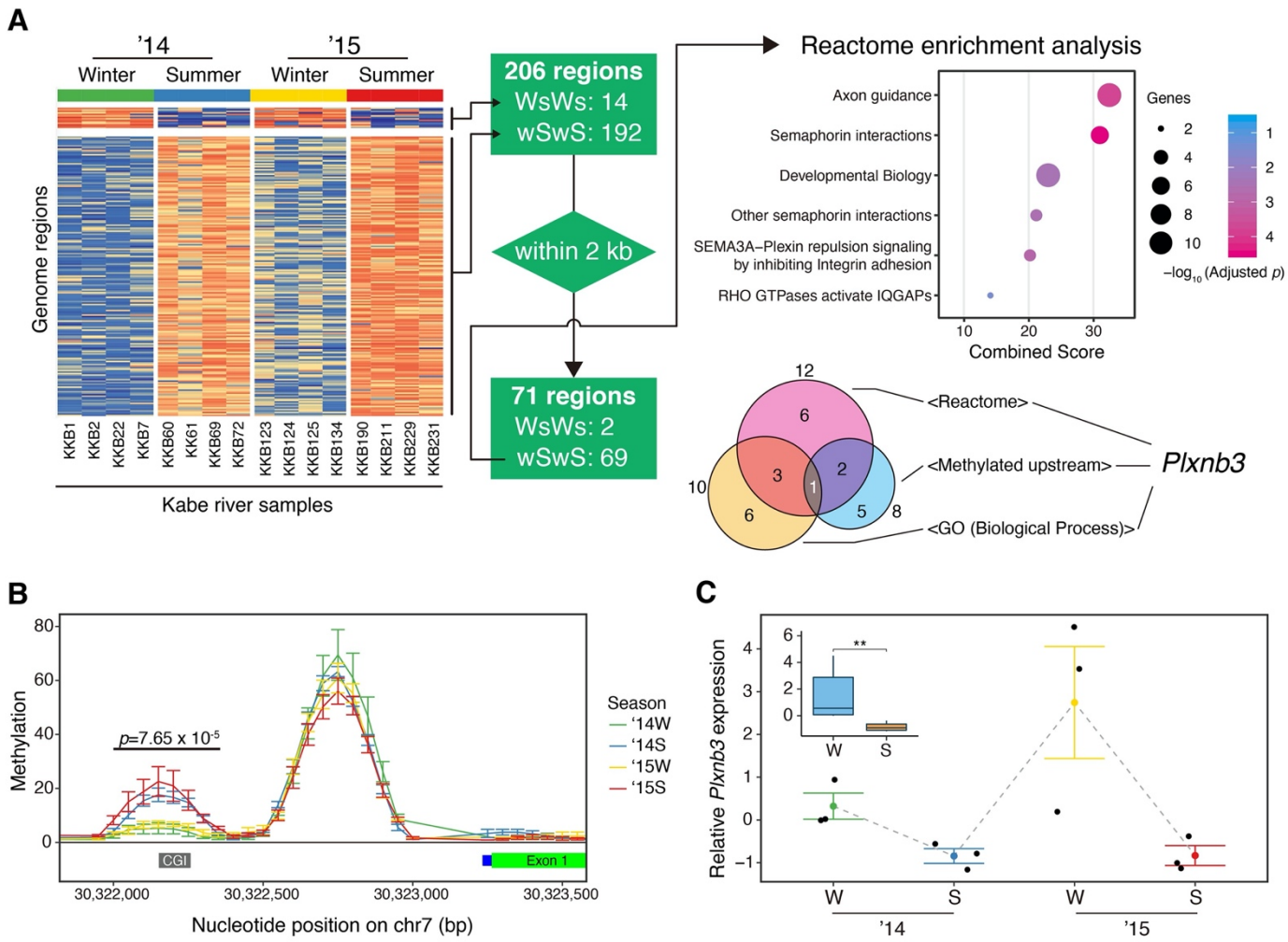
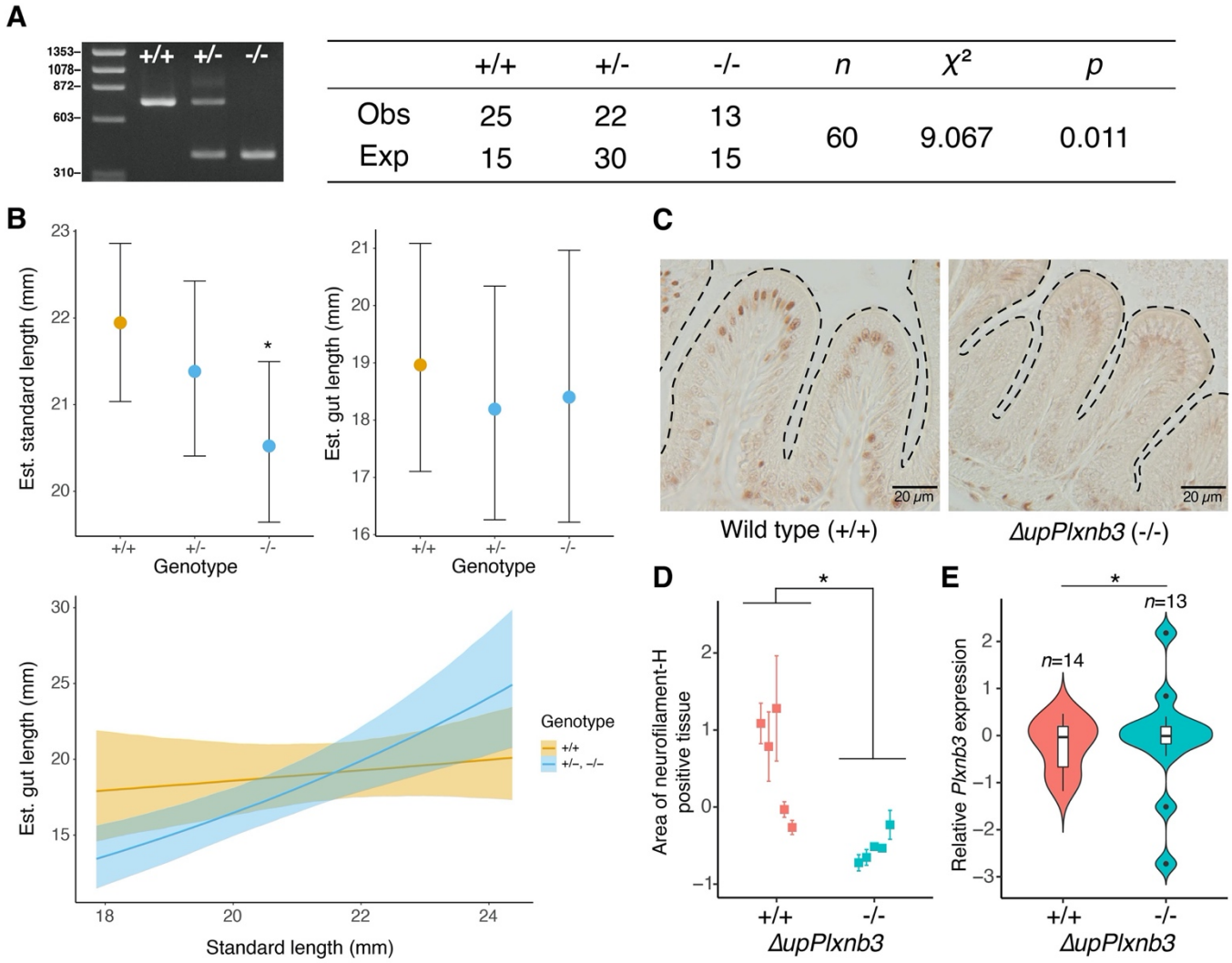


Figure 2 | DNA methylation map and *Plxnb3* gene expression pattern correlated with winter and summer. (A) Two hundred six differentially methylated regions (DMRs) were filtered using Reactome, GO analyses, and the position of the methylated region for the gene. The color difference of the heat map shows the difference in the degree of methylation, i.e., orange shows

the hypermethylated region. WsWs and wSwS represent hypermethylation in winter and summer, respectively. (B) The upstream region of *Plxnb3* showed seasonal methylation, which included the CpG island. (C) *Plxnb3* gene expression was suppressed in the summer. A significant difference in gene expression between seasons was detected using the Wilcoxon rank-sum test (**$p<0.01$).

To examine whether the upstream sequence of *Plxnb3* caused seasonal plasticity of the gut, we generated a genome-edited medaka, Δ*upPlxnb3*, in which the region showing seasonal DNA methylation was deleted using CRISPR/Cas9. At first, testing the segregation ratio of genotypes after a heterozygous mating showed the homozygous and heterozygous rates of medaka with the deletion were about half that compared with that of wild type medaka (Fig. 3A, $p=0.011$ using the chi-square test). No significant difference in gut length was detected, but a significant difference in body length was observed between wild type and mutant (Fig. 3B). By plotting body and gut lengths (Fig. 3B, fig. S5), we found dependency between those lengths in the mutants but not in the wild types, indicating that the mutant gut length is determined by body length not environment; i.e., gut plasticity was lost in the mutants (Bayesian generalized linear model [GLM], the effect size of the deletion and interaction between standard length and the deletion; estimates = -1.67±0.67 and 0.08±0.03; each 95% CI did not include zero, table S7). Next, using neurofilament H protein as a marker to measure the distribution of nerve axons involved in the *Plxnb3* gene, we found that the immunohistological signal disappeared at the tip of the intestinal villus in the mutants (Fig. 3C and D). Moreover, the *Plxnb3* expression level of the mutants showed significantly greater variance than that of the wild type, i.e., the expression regulation became unstable (Fig. 3E). These results indicate that disruption of the neural mechanism that senses the feeding environment (such as sensory-cell mediated food detection[16]) results in loss of gut length plasticity and ultimately inadequate nutrient absorption, leading to short body length and a lower survival rate (Fig. 3A and B). Thus, our data suggest that the role of methylation and demethylation in the upstream CpG island is a switch that regulates

the *Plxnb3* expression, which inhibits axon development, resulting in phenotypic plasticity of the gut.



	+/+	+/-	-/-	*n*	χ^2	*p*
Obs	25	22	13	60	9.067	0.011
Exp	15	30	15			

Figure 3 | Effects of the deletion of the *Plxnb3* upstream region on gut plasticity. (A) The deviation of the segregation ratio among genotypes was evaluated using the chi-squared test. (B) Comparisons of standard and gut lengths among genotypes were performed in the Bayesian GLM framework. Points and error bars were fitted values and 95% Bayesian credible intervals (full data in fig. S5). The asterisk indicates that -/- exhibited significantly shorter body length than did the other genotypes. (C) Immunochemical staining revealed the decrease of neurofilament-positive cells at the tip of the intestinal villus. (D) The difference in the degree of immunochemical staining (C) was statistically quantified and detected using the Wilcoxon rank-sum test (*p<0.05).

(E) *Plxnb3* expression showed the uneven variance between the wild types (+/+) and the mutants (-/-), which was evaluated using the F-test (*p<0.05), indicating the unstable *Plxnb3* expression in the mutant gut.

Under the PFE hypothesis, a genetic mutation should occur and fix a plasticity trait. Because the NJPN1 subgroup exhibited a long genetically fixed gut (Fig. 1C), the causative genetic mutations should be presumed to be highly frequent in the subgroup. In order to examine which gene is responsible for the longer gut of the NJPN1, 62 K SNPs obtained by RAD-seq[17] were used to search genomic regions that were differentiated in the NJPN1 subgroup and associated with the gut length (Fig. 4A). We found that a candidate signal strongly associated with length, resulting in a highly differentiated population, was detected in the *Ppp3r1* gene on chromosome 1, which is also involved in neuronal axon outgrowth[18]. Bayesian GLMM analysis confirmed that the T/T genotype had a significant effect compared to the C/C genotype (Fig. 4B, table S8). This mutation was a synonymous SNP, with no non-synonymous SNPs in the coding region within the NJPN1 subgroup, which predicted that the changes in the *Ppp3r1* expression contributed to the gut length (fig. S7). Given that a previous study of conditional *Ppp3r1* KO mice showed hypoplasia of the small intestine[19], we could hypothesize that the difference of this gene expression in medaka would lead to geographical differences in gut length. To test this hypothesis, we examined the relationship between *Ppp3r1* expression levels and gut length using outdoor-reared wild-derived laboratory stocks in the winter. We found that the gut length was significantly correlated with the *Ppp3r1* expression level (table S9), and the NJPN1 medakas with a homozygous T allele showed higher *Ppp3r1* expression than other subgroups with a homozygous C allele (Fig. 4C, p<0.01). These results indicate that the mutation associated with the up-regulation of the *Ppp3r1* expression causes the longest gut among the medaka subgroups.

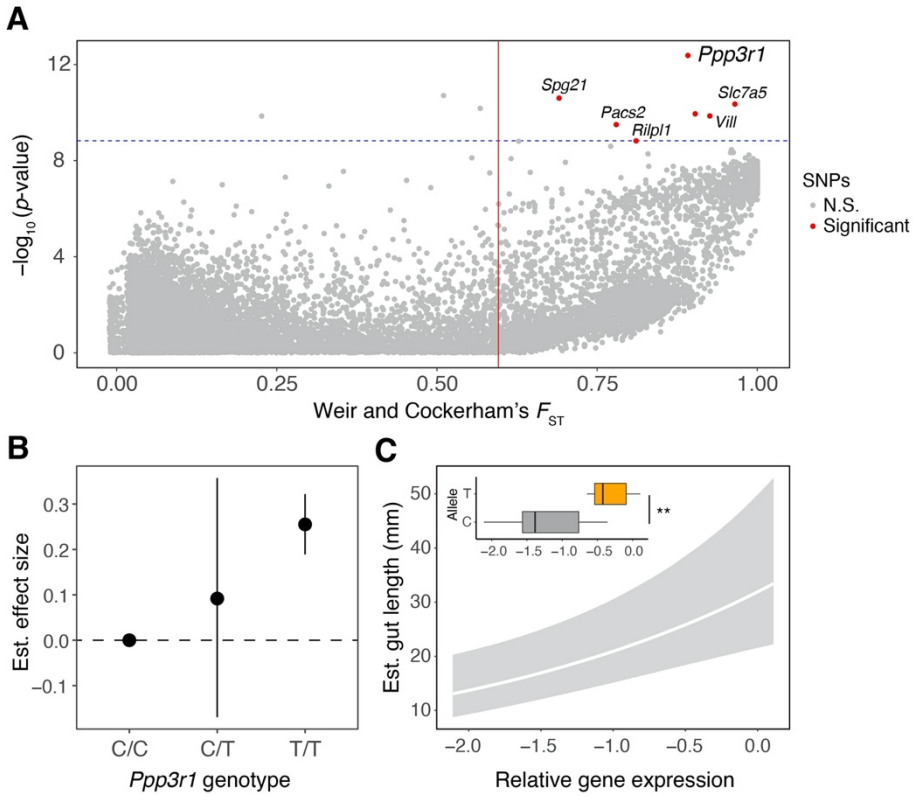


Figure 4 | Differentiation and association SNPs with gut length across the medaka populations. (A) F_{ST} was calculated between NJPN1 and the others. The red line indicates 0.59592 Weir and Cockerham's weighted F_{ST}. Association *p*-values with gut length were calculated using 22,235 SNPs filtered by a minor allele frequency of >0.05. The blue dashed line shows the *p*-value in which the top 10 SNPs distinguished from the others. The red points show that significant SNPs fulfilled the above criteria. (B) The effect size of the *Ppp3r1* mutation was estimated from the posterior mean of a Bayesian GLMM and is given relative to the model intercept at the C/C genotype (dashed line). Vertical bars indicate 95% CI. (C) *Ppp3r1* expression was significantly correlated with gut length estimated by Bayesian GLMM (effect size of the *Ppp3r1* expression: estimate =0.42±0.12; 95% CI did not include zero, full data in fig. S6). The inset graph shows the difference in the *Ppp3r1* expression between the C or T alleles, which were examined using winter-sampled individuals with homozygous C or T alleles. The statistical significance was evaluated using the Wilcoxon rank-sum test (**p<0.01).

Finally, we analyzed a molecular evolution of two identified regions based on the medaka population phylogeny. We determined the *Plxnb3* upstream sequences to estimate the evolution of methylated regions associated with plasticity in medaka populations (Fig. 5A, fig. S8). A comparison of the upstream sequences revealed that the number of CpG sites varied between the populations, indicating nucleotide substitutions resulted in the gain and/or loss of CpG sites (fig. S9). Ancestral sequence estimates indicated that the common ancestor of SJPN and NJPN had eight CpG sites in this upstream region, and the number of CpG sites diversified from seven to 10 upon branching to each SJPN lineage. On the other hand, the number of CpG sites was reduced to six in the common ancestor of NJPN1 and NJPN2, which did not meet the criteria of CpG islands. Since methylation at CpG sites is known to be a stochastic process[20], the number of CpG sites should influence the suppression of gene expression and the emergence of long guts. Supporting this prediction, we found that there was a positive correlation between the number of CpG sites and gut length among the populations without the *Ppp3r1* mutation (Bayesian GLMM, estimate =0.06, 95% CIs did not include zero, table S10) and that the long guts tended to develop depending on the number of CpG sites (Fig. 5B). In contrast, the *Ppp3r1* mutation had the effect of inducing a long gut in the populations with the lowest number of CpG sites (Fig. 5B). In addition, the seasonal plasticity of the gut length in the NJPN1 medaka was also lost (fig. S10), probably because of the small number of CpG sites that interfere with seasonal methylation (fig. S8). Because we observed the *Ppp3r1* mutation in both the NJPN1 and SJPN1 subgroups, the mutation existed in the common ancestors of the SJPN and NJPN, at least in the Japanese archipelago, as a standing variation. These data indicate that the mutation spread to the NJPN1 subgroup after loss of plasticity due to reduction of the number of CpG sites, after which a long gut became expressed in northern populations due to the standing variation of *Ppp3r1* (Fig. 5A).

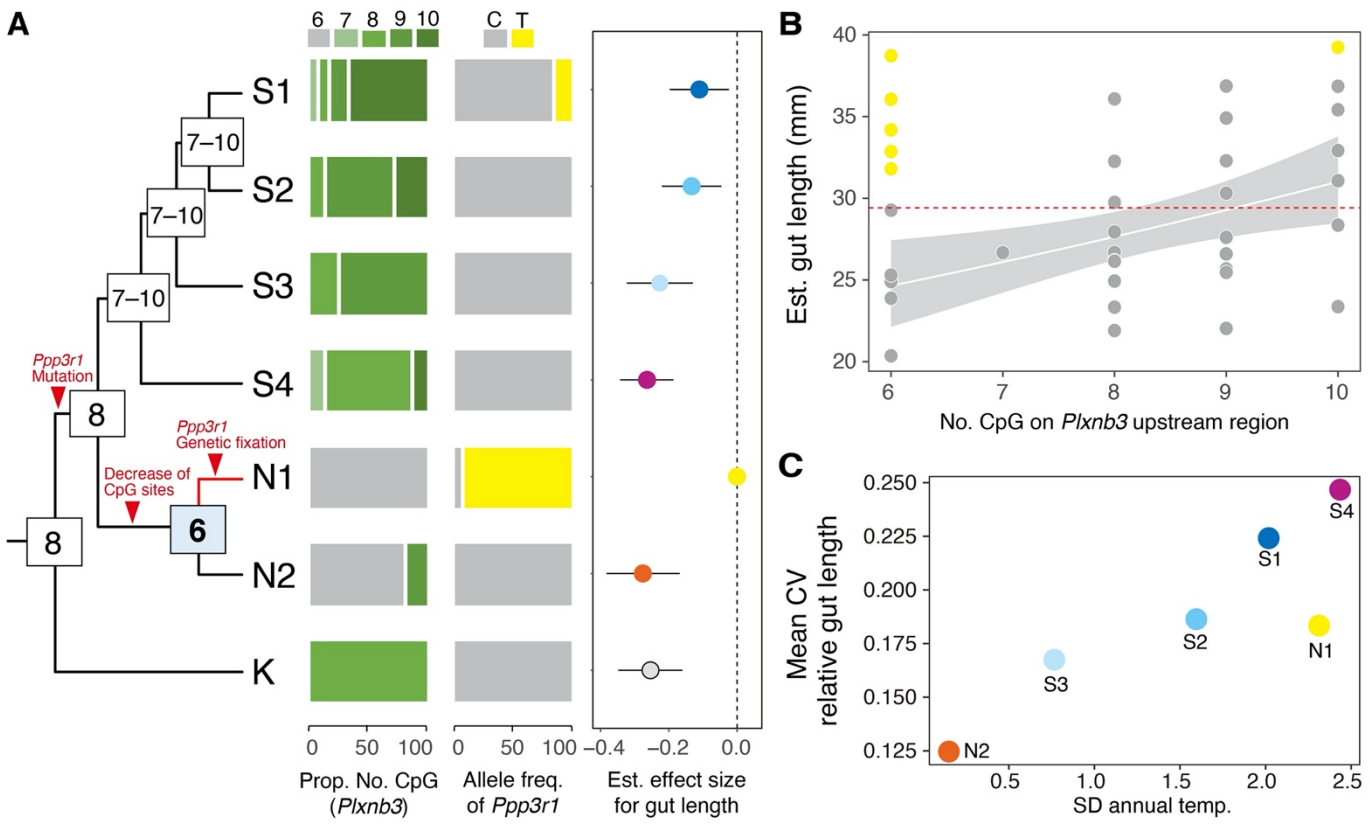


Figure 5 | Evolutionary relationship between the *Plxnb3* upstream region and the *Ppp3r1* mutation. (A) The number of CpG sites increased or decreased at each node on the population tree. The red arrowheads indicate evolutionary events, which were inferred from the current proportion of the number of CpG sites on the *Plxnb3* upstream region and the allele frequency of the *Ppp3r1* mutation. The right plot shows the effect sizes of each genetic background estimated as a fixed effect in the Bayesian GLMM framework. The effect size is given relative to the model intercept at NJPN1 (dashed line), and each 95% CI of effect size did not include zero, indicating the long gut are fixed genetically in the NJPN1 subgroup. (B) A correlation between the number of CpG sites in the *Plxnb3* upstream and gut length in each local population. Gray and yellow indicate *Ppp3r1* alleles (C or T) fixed in each population. (C) A relationship between gut length and annual temperature among subgroups (r=0.869, p=0.025). X-axis and Y-axis indicate the standard deviation (of annual temp.) of their original habitats and the average of coefficient of variation (of relative gut length) within genetic group, respectively. The

abbreviations in the figures are the same as those in Fig. 1C.

Discussion

The results of this study show that two genes involved in neural axons development contribute to phenotypic plasticity and genetic polymorphism of gut length. It was clarified at the molecular level that the plasticity was caused by methylation and demethylation of CpG in the *Plxnb3* gene upstream and that the genetic mutation in *Ppp3r1* spread to populations that the methylation site had reduced by the mutation (i.e., nucleotide substitution). In short, these molecular mechanisms and evolutionary processes indicate that the mutation-driven PFE occurred in gut length. This differs from "plasticity-driven evolution," which is based on phenotypic plasticity, in which the adaptive trait is led by natural selection and the following mutation fixes it (summarized as "selection-driven evolution" in Nei [2013][21]). The phenotypic plasticity controlled by CpG methylation plays the role of a buffer for selective pressure, which contributes to the accumulation and maintenance of *de novo* mutations as standing variations. When plasticity is lost, however, the adaptive mutation is selected from standing variations, and then the adaptive phenotype is genetically fixed. Our data indicate that this process is likely to be a factor in PFE, and the hereditary phenomenon of the environmentally-induced trait could be explained by the genetic framework of loss of CpG sites and fixation by standing variations.

Why has a genetic fixation of the long gut occurred in the NJPN1 medaka? SJPN and NJPN medaka originate in northern Kyushu and the Tajima-Tango region, respectively[11]. In SJPN medaka, even populations with similar genetic backgrounds differ in their climatic environment habitat. Because the foraging environment is susceptible to annual temperature changes, and the gut length has become varied in response to the average temperature variation (Fig. 5C), the gut length would have been plastically adjusted for efficient digestion and absorption in various foraging environments. On the other hand, an NJPN medaka is a group that has spread rapidly to the north from small habitats where the environment is relatively uniform[11]. In the NJPN2 habitats, both temperature and gut length show minimal variability (Fig.

5C), while the NJPN1 habitats show temperature fluctuations. And the growing period of those NJPN2 habitats is short because of high latitudes[9]. Therefore, it may be not necessary to change the gut length in the NJPN2 medaka, but the NJPN1 medaka must feed during the short summer and prepare for the winter. Under these circumstances for the NJPN1 medaka, maintaining a long gut throughout the season may have served as food storage, rather than adjusting the gut length to increase absorption efficiency as in the SJPN medaka. Moreover, the foraging behavior of NJPN1 medaka that has expanded into the higher latitude regions is more frequent than that of medaka in the lower latitude regions[9]. This indicates that genetic mutations affecting gut length may be driving geographic differences in behavior. Indeed, the neurofilament protein involved in sensory neural regulation in the gut[22] was not stably expressed in medaka that had lost plasticity (Fig. 3C). This result suggests that the gut could be not detecting the appropriate amount of feeding, which may lead to excessive feeding behavior. I.e., the gut-brain interaction may have enhanced the genetic fixation of advantageous mutations.

The identified molecular mechanisms and the above evolutionary inference leads to the conclusion that plasticity is easily lost under a stable environment, and that after loss of plasticity, a favorable mutation can easily be fixed on foraying into harsh habitats. In a macroscopic-type view, this phenomenon may appear as though the evolution of acquired traits has occurred because it occurs continuously.

Materials and Methods

Ethics statement

The Institutional Animal Care and Use Committee of Kitasato University and The University of Tokyo approved all the experimental procedures (No. 4111 and C-09-01, respectively).

Sampling and measurement of gut length

To detect the seasonal plasticity of gut length, we collected 228 medaka individuals from three rivers in Kagawa prefecture from January 2014 to August 2015 (Fig. 1A). The details are in the supplementary table

S1. To explore the genetically-fixed gut length in medaka populations, we sampled 400 individuals from 40 wild-derived laboratory stocks that originated from wild populations, which had been maintained in a constant environment at an outdoor breeding facility, The University of Tokyo[11, 23]. The sampled stocks were: SJPN1 (PO.NEH): Ichinoseki, Toyota, Mishima, Iwata, Sakura, Shingu; SJPN2 (Sanyo/Shikoku/Kinki): Ayabe, Tanabe, Okayama, Takamatsu, Kudamatsu, Misho; SJPN3 (San-in): Kasumi, Tsuma, Tottori, Hagi; SJPN4 (Northern and Southern Kyushu): Kusu, Arita, Hisayama, Fukue, Kazusa, Izumi, Hiwaki, Nago; NJPN1 (derived N.JPN): Kamikita, Yokote, Niigata, Kaga, Maizuru, Miyazu; NJPN2 (ancestral NJPN): Amino, Kumihama, Toyooka, Kinosaki, Hamasaka; KOR (KOR/CHN): Yongcheon, Maegok, Bugang, Sacheon, Shanghai. The subgroup names in parentheses are defined in our prior study[11]. We isolated the gut (from the esophagus to the anus), fixed each one in 4% paraformaldehyde for 1 hour on ice, and took photos of them laid out on glass slides to measure their lengths. Unfortunately, 32 medaka individuals died before this analysis, and 27 individuals were removed because of presenting with an abnormal form due to aging (e.g., spondylosis). Finally, we obtained the standard and gut lengths from photos of 341 individuals (table S3) using Image J software[24]. Statistical analyses were performed using R software[25].

Detecting gut length changes using a Bayesian GLMM framework

The relationships between gut length and seasons, genetic background, *PlxnB3* upstream, and *Ppp3r1* genotypes were modeled using a GLMM framework with the MCMC method in the R package brms[26] in R v3.5.2. The gut length was modeled with a gamma error structure and log link function. Fixed effects included the standard length and sex in each analysis, and the random effect was subject identity. The MCMC conditions are described in each supplementary table.

Detecting seasonal methylated regions and filtering using Reactome and GO databases

We extracted genomic DNA from two male and two female guts per season (total 16 guts) from the Kabe

river medaka using the phenol/chloroform method described previously[11]. The quality and quantity of genomic DNA were assessed using 0.5% agarose electrophoresis, a nanophotometer (IMPLEN), and the Qubit BR dsDNA Assay Kit (Thermo Fisher Scientific). After adjusting the DNA concentration to 40 ng/µl, 1.2 µg of DNA was sheared to target 150–200 bp by using two protocols of 400 bp and 200 bp implemented in the S220 Focused-ultrasonicator (Covaris). Sheared DNAs were purified using the MinElute PCR Purification Kit according to the manufacture's protocol, and their quality and quantity were measured by TapeStation and Qubit dsDNA BR. A total of 500 ng of sheared DNAs was input for methyl-CpG binding domain (MBD) enrichment using the EpiXplore Methylated DNA Enrichment Kit (Clontech) according to the manufacturer's instruction. Subsequently, 30 ng of methylated DNA was used to generate the NGS (Next-Generation Sequencing) library by using the NEBNext ULTRA DNA Library Prep and the NEBNext Multiplex Oligos Kits for Illumina Sequencing according to the manufacturers' instructions. The quality and quantity of genomic DNA were assessed using the TapeStation D1000HS screen tape (Agilent Technologies) and the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). The sequencing process was outsourced to Macrogen Japan in Kyoto and was performed by Illumina Hiseq 2500 with 51 bp single-end reads to obtain an average of 40 million reads per sample. Our single-end reads were filtered using the FASTQ Quality Trimmer and Filter in FASTX-Toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/download.html) using the following options: “-t 20 -l 30 -Q 33” and “-Q 33 -v -z -q 20 -p 80”, respectively. The draft genome of the medaka sequenced by the PacBio sequencer (Medaka-Hd-rR-pacbio_version2.2.4.fasta; http://utgenome.org/medaka_v2/#!Assembly.md) was used to align the reads using BWA (Burrows-Wheeler Aligner) backtrack 0.7.15-r1140[27] using the “-n 0.06 -k 3” option. After the mapping process, the multi-mapped reads were removed using SAMtools v1.9[28] and the “-q 1 -F 4 -F 256 -F 2048” option. Peak call was performed using MethylAction[29] with "fragsize=200, winsize=50" on R version 3.4.4 in RStudio version 1.0.136. After filtering by states that were TRUE in "frequent column" and showed a seasonal correlation, we listed genes within 2,000 bp upstream and downstream of methylated regions using BEDTools[30].

To identify the biological processes of genes uniquely associated with each seasonally-correlated region, we performed Reactome pathway enrichment analyses using Enrichr[31, 32], GO terms, and human genome annotations. Before this Enrichr analysis, we estimated medaka orthologous genes to human genes using biomart in Ensembl database and ORTHOSCOPE[33]. Significant Reactome pathways (Adjusted *p*-value <0.05) and the top 10 GO terms for "biological processes" were used as the filter to list the candidate genes. Adding the gene list with the seasonally methylated region on their upstream, we visualized as a Venn diagram and identified the functional gene, *Plxnb3*, associated with seasonal plasticity (Fig. 2A). Finally, we identified a CpG island on upstream of this gene using Meth primer[34] with the following setting: GC >50%, O/E >0.6, and length >100 bp (Fig. 2B).

Finding the high-differentiated and gut-length-associated SNPs

For 341 gut-length-measured medaka, genomic DNAs were extracted from one-third of the medaka body using NucleoSpin Tissue (Macherey-Nagel) according to the manufacturer's protocol. After the quality check using a Nanophotometer (IMPLEN) and 0.5% agarose gel electrophoresis, six individuals were removed because of low DNA concentrations. For 335 individuals, we arranged the original RAD-seq protocol adding an indexing PCR step to adjust the sample size and generated 14 RAD-seq libraries (24 individuals per library, 23 individuals in the last library) by the following method. The designed P1-adaptor included 24 in-line barcodes which had two nucleotide differences in each, and the adaptors of Sbf I-HF (New England Biolabs)-digested DNAs for 90 min at 37˚C were ligated by T4 ligase (NEB). After DNA pooling, DNAs were sonicated using an S220 Focused-ultrasonicator (Covaris) to target 300 bp and purified using the GeneRead Size Selection Kit (Qiagen) according to the manufacturers' protocols. The End-repair, A-tailing, and P2 adaptor ligation steps were performed using the NEBNext Ultra DNA Library Prep Kit for Illumina (NEB). After size selection using AMPure XP beads (Beckman Coulter), indexing PCR was performed using Q5 (NEB) under the following conditions: initial denaturing step at 90˚C for 30 sec, 14 cycles of denaturation at 98˚C for 10 sec, annealing at 68˚C for 30 sec, extension at 72˚C for 20 sec,

and a final extension step at 72˚C for 5 min. The PCR products were purified with AMPure XP beads and then were validated using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) for DNA concentration, TapeStation (Agilent Technologies) for fragment length, and Miseq (Illumina) for library quality. Finally, RAD-seq data were generated using three lanes of Hiseq 2500 (Illumina) with 51 bp single-end reads settings conducted by Macrogen Japan.

Our single-end reads were filtered by Cutadapt ver1.12[35] using the following options: "-m 50 -e 0.2," and were demultiplexed by process_radtags (v1.44) implemented in Stacks[36] using the following options: "-c -r -t 44 -q -s 0 --barcode_dist_1 2." After quality filtering and check, one individual (from the Kazusa population) was removed because of low reads. The draft genome of the medaka sequenced by the PacBio sequencer (Medaka-Hd-rR-pacbio_version2.2.4.fasta; http://utgenome.org/medaka_v2/#!Assembly.md) was used to align the reads using BWA backtrack 0.7.15-r1140[27] using the "-n 0.06 -k 3" option. After the mapping process, the multi-mapped reads were removed using SAMtools v1.9[28] and the "-q 1 -F 4 -F 256 -F 2048" option. SNP call was performed by Stacks pipeline: pstacks -m 2 --model_type snp --alpha 0.05; cstacks -g -n 1; sstacks -g; rxstacks --lnl_filter --lnl_lim -10 --conf_filter --conf_lim 0.75 --prune_haplo --model_type bounded --bound_low 0 --bound_high 0.1; cstacks -g -n 1; sstacks -g; populations -r 0.5 -p 7 -m 6 -f p_value -a 0.0 --p_value_cutoff 0.1 --lnl_lim -10. Assigned medaka populations into seven genetic groups based on our prior study[11], genotype imputation was performed by Beagle (v4.1) using GT format[37]. Finally, data set was generated using VCFtools[38] to merge the seven genotype-imputed data, and included 63,265 SNPs.

To detect high-differentiated and gut-length-associated genes, we calculated Weir and Cockerham's F_{ST} and the degree of association with the gut length using VCFtools[38] and Plink 1.9 (https://www.cog-genomics.org/plink/1.9/) using the SNP data. Although the gut length was highly correlated with the population structure of medaka (Fig. 1C), we did not control the structure information because this reduced the power of the analysis and could lead to false negatives[39]. To avoid the false positives without controlling the population structure as much as possible, we designed the sampling strategy to choose

multiple subgroups from each group and to use balancing samples across subgroup populations to homogenize allele frequencies[39]. Focusing on SNPs with a minor allele frequency of more than 0.05, we extracted the top 10 SNPs above the global F_{ST} value (weighted F_{ST}) and with the highest -$\log_{10}$ p-value from 22,235 SNPs. Association of gut length and SNPs were assessed by permutation test with 1,000,000 times resampling.

Generating a genome-editing medaka by the CRISPR/Cas9 system

The d-rR medaka line was used to generate mutants deleted from the seasonal methylation region. Fish were maintained at 28°C under a 14-h light/10-h dark cycle. We collected the d-rR eggs every morning. Microinjection was carried out in medaka embryos at the one-cell stage using FemtoJet and InjectMan NI 2 (Eppendorf). Cas9 protein (500 ng/μL), tracrRNA (200 ng/μL), two crRNAs (each 200 ng/μL), and saturated phenol red solution were mixed in a 4:2:2:1 ratio, and injected into the embryos. The crRNAs were designed to delete the seasonal methylation region of 334 bp using CRISPRdirect[40], which identified the specific regions on the medaka genome, and their crRNA sequences were: oPlxnB3upCGI_F1, GAUGCUGUUGCCUGGCCAUAguuuuagagcuaugcuguuuug; oPlxnB3upCGI_R1, GAUGACACAGUGGAGCAUCGguuuuagagcuaugcuguuuug. The crRNAs, tracrRNA, and recombinant Cas9 protein were obtained from FASMAC. After injection, the embryos (F0) were maintained at 28°C using an air incubator. After hatching and maturing, the injected medaka were intercrossed for the identification of the germline founders with mutations in each target locus by genotyping their fertile eggs. The genotyping was performed using PCR and 1.5% agarose gel electrophoresis using EX Taq Hot Start Version (TaKaRa Bio) according to the manufacturer's protocol and two designed primers: oPlxnb3CpG1F2, 5'-TTGCATCTGGCTTTGCATGAATC-3'; oPlxnb3CpG1R2, 5'-ATAACATCACCATGGCAACAACG-3'. The germline founders (F1) were outcrossed and generated their off-springs (F2). Heterozygous mutation carriers in the F2 generation were identified by above PCR-based genotyping using their fin-extracted DNAs. Finally, sixty F3 individuals

were used for subsequent analyses.

Analysis of *Plxnb3* upstream-deleted medaka

To examine a phenotype of *Plxnb3* upstream-deleted (Δ*upPlxnb3*) medaka, we conducted three tests: segregation of genotypes, a comparison of gut and body length, and an immunohistological comparison. For these tests, we performed heterozygous mating using the F1 generation and obtained their fertile eggs. We bred the medaka larvae in a tank (W52×D21×H23cm) and raised them until 2 weeks after hatching, divided the juveniles into groups of four to six individuals, and transferred them to 2 L tanks to raise them for an average of 103 days after hatching. Each genotype was determined using the PCR-based method described above, and the segregation ratio was evaluated using the chi-squared test. The significant differences in gut lengths between genotypes were detected using a Bayesian GLM, performed with the brms R package with the default MCMC setting on RStudio 1.3. The gut length was modeled with a gamma error structure and log link function. Fixed effects included the standard length and sex in each analysis. For immunohistological analysis, 5 μm serial tissue sections from wild types and mutants (n=5 each) were prepared using our previously described method[41]. The sections were incubated with a mouse monoclonal anti-phosphorylated neurofilament H (BioLegend, #smi-31, 1:1000) of which cross-reactivity for medaka had been confirmed in a previous study[42]. Quantitative analysis of an immunohistological-staining signal was performed using ImageJ software[18]. We obtained an average of six images per individual at a 100x objective using light microscopy (BX63, Olympus) equipped with a digital camera (DP74, Olympus). We then split each image into red, green, and blue using the color channels option and subtracted green from the red image with the image calculator. Setting the same threshold, we measured the proportion of the stained area for one intestinal villus of each individual. After standardizing the percentage of the stained area in each experiment, we evaluated the difference of the immunohistological-staining signals between the wild types and the mutants using the Wilcoxon rank-sum test (p<0.05).

RT-qPCR

We performed reverse transcription-quantitative PCR (RT-qPCR) analyses to evaluate the gene expressions of the *Plxnb3* and *Ppp3r1* genes. We sampled three individuals from each season in the Kabe river (12 individuals totally) for the *Plxnb3* gene expression, and four individuals from four geographical populations (Kamikita and Yokote from NJPN1, Amino from NJPN2, and Izumi from SJPN4) for the *Ppp3r1* gene expression. The guts were isolated from ethanol-fixed Kabe river samples stored at -20˚C and wild-derived laboratory stocks in winter (29 Jan. 2018). The gut was homogenized using a BioMasher (Nippi), and total RNA was isolated from each sample using a combination of the TRIzol (Invitrogen) and NucleoSpin RNA (Macherey-Nagel).

cDNA, which was used as a template for qPCR, was synthesized using a PrimeScript RT reagent kit with gDNA Eraser (TaKaRa Bio) according to the manufacturer's protocol. Ribosomal protein L7 (*Rpl7*) was used as an internal reference gene because medaka *Rpl7* was expressed with the less variance among the same tissue samples, in different tissues and stages of development[43], and between winter and summer[44]. The negative control was treated with RNase free water and RT minus products. The reaction mixture consisted of 2 µl of cDNA template, 10 µl of TB Green® Premix Ex Taq™ (Tli RNaseH Plus) (TaKaRa Bio), 0.4 µl of forward and reverse primers (10 µM), and 7.2 µl of RNase-free water. The reaction procedure in a LightCycler® 480 Real-Time PCR System (Roche Diagnostic) was: 95˚C, 30 sec; 95˚C, 5 sec + 60˚C, 30 sec (for amplification, 40 cycles); 95˚C, 5 sec + 60˚C, 60 sec + 97˚C, 0.11˚C/sec (for the melting curve); 40˚C, 30 sec (for cooling). The primer pairs were:
oPlxnb3ex32F1:TGGTGAAGAGCAGTGAAGATCC;
oPlxnb3ex33R1:TAGTGTTCCCTTCATGGACAGC for *Plxnb3* (PCR efficiency: 105%),
oCANB1ex4F1:TAAACATGAAGGGAAGGCTGGAC;
oCANB1ex3R1:ATTCGCCTTCAGGATCTACGAC, for *Ppp3r1* (PCR efficiency: 95%), and oRPL7ex3-4F1:ATCCGAGGTATCAACGGAGTC; oRPL7ex5R1:TGCCGTAGCCACGTTTGTAG for *Rpl7* (PCR efficiency: 95%). The specificity of all PCR primer pairs was confirmed as a single peak in the melting

curve analysis. Two technical replicates were run for the RT-qPCR analyses, and the relative expressions of each sample were calculated using the ΔΔCt method with the LightCycler® (Roche Diagnostic).

Outdoor breeding experiment

To test whether the food quality affects the gut length, we transported the wild medaka from the Kabe river into the outdoor breeding facility on the Kashiwa campus, The University of Tokyo, in August 2015. Then in September 2016 and February 2017, we sampled and compared the gut lengths of bred and wild medakas in the outdoor breeding facility and the Kabe river, respectively. Their gut lengths were measured according to the method described above.

Common garden experiment

To examine whether NJPN1 medaka show a shorter gut in winter than in summer, we sampled and measured the medaka gut lengths every 2 months from October to February using wild-derived laboratory stocks. The sampled stocks were: NJPN1: Kamikita (n=9, 4, 3), Maizuru (n=10, 4, 4), Kaga (n=7, 4, 4); SJPN2 as control: Tanabe (n=9, 4, 4), Okayama (n=10, 4, 4), Misho (n=7, 2, 4). We used the Bayesian GLMM to detect the gut length variations using the brms package, as described above. Fixed effects included the standard length, sex, and sampling date (October, December, and February). Random effects were subject identity and the local population. Four MCMC chains were run for 105,000 samples, with a 5,000 sample burn-in, by using every fourth sample. This gave a total of 100,000 used samples. R-hat is the Germany-Rubin convergence statistic for estimating the degree of convergence of a random Markov chain and indicates insufficient convergence at values greater than 1.1.

Principal component analysis

To test whether the population structure had changed in the Kabe river between winter and summer, we performed a PCA (principal component analysis) using the SNPRelate program[45] in R version 3.2.2. A total

of 2,167,650 bi-allelic SNPs were generated from the bam files used to detect the methylation regions. The procedure was: SNPs were extracted using SAMtools mpileup and filtered vcfutils.pl with varFilter -d 6 option, and then only bi-allelic sites were extracted using VCFtools[38]. The PCA graph was drawn using ggplot2[46].

Relationship between the number of CpG sites on the upstream of *Plxnb3* and the mutation in the *Ppp3r1* region

To examine the genetic variation on the upstream of *Plxnb3* among geographic populations, we performed PCR direct sequencing as described previously[23]. Two primers were: oPlxnb3CpG1F1:GTTCATTTAAGGGAGGAACCAAAGG; oPlxnb3CpG1R1:CTCCACTGTGTCATCAAAAGAAGC. Sequences were determined and aligned across 40 populations, and then CpG sites on these sequences were counted using MEGA X software[47]. Note that sequences obtained from Kaga, Shingu, and Shanghai populations were excluded from the analysis because their CpG sites of upstream sequences or *Ppp3r1* mutations existed as polymorphisms. Thus, the regression line between the number of CpG sites and gut lengths was calculated using the populations with homozygotes of the C allele on *Ppp3r1*. A phylogenetic tree was reconstructed to estimate the number of CpG sites on ancestral upstream sequences of *Plxnb3* using a maximum likelihood method under the GTR model implemented in MEGA X[47]. Those numbers were plotted on the medaka genetic group tree obtained from the previous study, and the allele frequencies of the *Plxnb3* upstream sequence and *Ppp3r1* mutations in each group were calculated as the number of alleles divided by the number of chromosomes. Note that data obtained from Yongcheon and Shanghai populations were excluded from this analysis because those populations might have experienced artificial gene-flow via the strain maintenance process mentioned in our prior study[9].

To test whether the dispersion of the gut length was large, according to the dispersion of the air temperature of each region, we calculated the coefficient of variation (CV) of the gut length and standard

deviation of the average annual temperature for each group using the Global Solar Atlas database (https://globalsolaratlas.info). We plotted the standard deviation of the average annual temperature on the X-axis and the mean CV of the gut length on the Y-axis and performed a Pearson's product-moment correlation test in R and RStudio.

Acknowledgments

We thank Mrs. Shizuko Chiba, Mrs. Sumiko Tomizuka, Dr. Atsuko Shimada and Prof. Emeritus Akihiro Shima (The University of Tokyo) for maintaining the medaka stocks from wild populations. We also thank Robert E. Brandt, Founder, CEO, and CME, of MedEd Japan, for editing and formatting the manuscript.

This study was supported in part by grants-in-aid for Young Scientists (B) no. JP16K21352, Early-Career Scientists no. JP19K16201, Scientific Research on Innovative Areas no. JP19H05737, Scientific Research (A) no. JP17H01453, and Scientific Research (B) no 17H03738. T.K. was also supported by a grant-in-aid from JSPS Research Fellowships no. JP16J07227.

Authors' contributions

T.K., H.Takeshima, Y.H. H.Takeuchi, and H.O. designed the experiments; T.K., S.S., K.Y., S.O., T.G., S.T., K.F., T.N., T.I., Y.H., and H.Takeshima performed the sampling and experiments; T.K., S.S., and K.Y. analyzed the data; T.K., T.G., T.N., H.M., M.O., H.Takeuchi, and H.O. interpreted the results of the experiments; T.K. and H.O. wrote the manuscript. All authors have read and approved the final manuscript.

References

1. Gilbert, S. F., Bosch, T. C. G. & Ledón-Rettig, C. Eco-Evo-Devo: Developmental Symbiosis and Developmental Plasticity as Evolutionary Agents. *Nat. Rev. Genet.* **16**, 611–622 (2015).
2. Levis, N. A., Isdaner, A. J. & Pfennig, D. W. Morphological Novelty Emerges from Pre-Existing Phenotypic Plasticity. *Nat. Ecol. Evol.* **2**, 1289–1297 (2018).

3. Levis, N. A. & Pfennig, D. W. Evaluating 'Plasticity-First' Evolution in Nature: Key Criteria and Empirical Approaches. *Trends Ecol. Evol.* **31**, 563–574 (2016).
4. Laland, K. *et al.* Does Evolutionary Theory Need a Rethink? *Nature* **514**, 161–164 (2014).
5. Baerwald, M. R. *et al.* Migration-Related Phenotypic Divergence is Associated with Epigenetic Modifications in Rainbow Trout. *Mol. Ecol.* **25**, 1785–1800 (2016).
6. Kronholm, I., Johannesson, H. & Ketola, T. Epigenetic Control of Phenotypic Plasticity in the Filamentous Fungus Neurospora crassa. *G3 (Bethesda)* **6**, 4009–4022 (2016).
7. Matthias Starck, J. Shaping Up: How Vertebrates Adjust Their Digestive System to Changing Environmental Conditions. *Anim. Biol. Leiden Neth.* **53**, 245–257 (2003).
8. Zandonà, E., Auer, S. K., Kilham, S. S. & Reznick, D. N. Contrasting Population and Diet Influences on Gut Length of an Omnivorous Tropical Fish, the Trinidadian Guppy (Poecilia reticulata). *PLoS One* **10**, e0136079 (2015).
9. Suzuki, Y., Miyake, T. & Yamahira, K. An Acquisition Trade-Off with Fast Growth in a Fish, the Medaka Oryzias latipes: Why Do Low-Latitude Ectotherms Grow More Slowly? *Evol. Ecol.* **24**, 749–759 (2010).
10. Yamahira, K. & Nishida, T. Latitudinal Variation in Axial Patterning of the Medaka (Actinopterygii: Adrianichthyidae): Jordan's Rule is Substantiated by Genetic Variation in Abdominal Vertebral Number. *Biol. J. Linn. Soc. Lond.* **96**, 856–866 (2009).
11. Katsumura, T., Oda, S., Mitani, H. & Oota, H. Medaka Population Genome Structure and Demographic History Described via Genotyping-by-Sequencing. *G3 (Bethesda)* **9**, 217–228 (2019).
12. Hilgers, L. & Schwarzer, J. The Untapped Potential of Medaka and Its Wild Relatives. *Elife* **8**, (2019).
13. Kirchmaier, S., Naruse, K., Wittbrodt, J. & Loosli, F. The Genomic and Genetic Toolbox of the Teleost Medaka (Oryzias latipes). *Genetics* **199**, 905–918 (2015).
14. Shima, A., Shimada, A., Sakaizumi, M. & Egami, N. First Listing of Wild Stocks of the Medaka Oryzias latipes Currently Kept by Zoological Institute, Faculty of Science, University of Tokyo. *J.*

Fac. Sci. Univ. Tokyo Sec. IV **16**, 27–35 (1985).

15. Artemov, A. V. *et al.* Genome-Wide DNA Methylation Profiling Reveals Epigenetic Adaptation of Stickleback to Marine and Freshwater Conditions. *Mol. Biol. Evol.* **34**, 2203–2213 (2017).
16. Schneider, C., O'Leary, C. E. & Locksley, R. M. Regulation of Immune Responses by Tuft Cells. *Nat. Rev. Immunol.* **19**, 584–593 (2019).
17. Baird, N. A. *et al.* Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS One* **3**, e3376 (2008).
18. Graef, I. A. *et al.* Neurotrophins and Netrins Require Calcineurin/NFAT Signaling to Stimulate Outgrowth of Embryonic Axons. *Cell* **113**, 657–670 (2003)
19. Fujita, M. *et al.* Calcineurin B1 Deficiency in Glial Cells Induces Mucosal Degeneration and Inflammation in Mouse Small Intestine. *Biol. Pharm. Bull.* **41**, 786–796 (2018).
20. Feinberg, A. P. & Irizarry, R. A. Evolution in Health and Medicine Sackler Colloquium: Stochastic Epigenetic Variation as a Driving Force of Development, Evolutionary Adaptation, and Disease. *Proc. Natl. Acad. Sci. U. S. A.* **107 Suppl 1**, 1757–1764 (2010).
21. Nei, M. *Mutation-Driven Evolution*. (OUP, Oxford, 2013).
22. Latorre, R., Sternini, C., De Giorgio, R. & Greenwood-Van Meerveld, B. Enteroendocrine Cells: A Review of Their Role in Brain-Gut Communication. *Neurogastroenterol. Motil.* **28**, 620–630 (2016).
23. Katsumura, T. *et al.* Natural Allelic Variations of Xenobiotic-Metabolizing Enzymes Affect Sexual Dimorphism in Oryzias latipes. *Proc. Biol. Sci.* **281**, (2014).
24. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **9**, 671–675 (2012).
25. R Core Team. R: A Language and Environment for Statistical Computing. Foundation for Statistical Computing, Vienna. (2018).
26. Bürkner, P.-C. Advanced Bayesian Multilevel Modeling with the R Package brms. *R J.* **10**, 395–411 (2018).

27. Li, H. & Durbin, R. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **25**, 1754–1760 (2009).

28. Li, H. *et al.* The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

29. Bhasin, J. M., Hu, B. & Ting, A. H. MethylAction: Detecting Differentially Methylated Regions That Distinguish Biological Subtypes. *Nucleic Acids Res.* **44**, 106–116 (2016).

30. Quinlan, A. R. & Hall, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features *Bioinformatics* **26**, 841–842 (2010).

31. Kuleshov, M. V. *et al.* Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update. *Nucleic Acids Res.* **44**, W90–7 (2016).

32. Chen, E. Y. *et al.* Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool. *BMC Bioinformatics* **14**, 128 (2013).

33. Inoue, J. & Satoh, N. ORTHOSCOPE: An Automatic Web Tool for Phylogenetically Inferring Bilaterian Orthogroups with User-Selected Taxa. *Mol. Biol. Evol.* **36**, 621–631 (2019).

34. Li, L.-C. & Dahiya, R. MethPrimer: Designing Primers for Methylation PCRs. *Bioinformatics* **18**, 1427–1431 (2002).

35. Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal* **17**, 10–12 (2011).

36. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. Stacks: An Analysis Tool Set for Population Genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).

37. Browning, S. R. & Browning, B. L. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies by Use of Localized Haplotype Clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).

38. Danecek, P. *et al.* The Variant Call Format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

39. Brachi, B., Morris, G. P. & Borevitz, J. O. Genome-Wide Association Studies in Plants: The Missing

Heritability is in the Field. *Genome Biol.* **12**, 232 (2011).

40. Naito, Y., Hino, K., Bono, H. & Ui-Tei, K. CRISPRdirect: Software for Designing CRISPR/Cas Guide RNA with Reduced Off-Target Sites. *Bioinformatics* **31**, 1120–1123 (2015).
41. Nagata, K. *et al.* In vivo 3D Analysis of Systemic Effects after Local Heavy-Ion Beam Irradiation in an Animal Model. *Sci. Rep.* **6**, 28691 (2016).
42. Uemura, N. *et al.* Viable Neuronopathic Gaucher Disease Model in Medaka (Oryzias latipes) Displays Axonal Accumulation of Alpha-Synuclein. *PLoS Genet.* **11**, e1005065 (2015).
43. Zhang, Z. & Hu, J. Development and Validation of Endogenous Reference Genes for Expression Profiling of Medaka (Oryzias latipes) Exposed to Endocrine Disrupting Chemicals by Quantitative Real-Time RT-PCR. *Toxicol. Sci.* **95**, 356–368 (2007).
44. Shimmura, T. *et al.* Dynamic Plasticity in Phototransduction Regulates Seasonal Changes in Color Perception. *Nat. Commun.* **8**, 412 (2017).
45. Zheng, X. *et al.* A High-Performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. *Bioinformatics* **28**, 3326–3328 (2012).
46. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York (2016).
47. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).