The phenotypic variation of *widefins* medaka is due to the insertion of a giant transposon containing a viral genome within *hoxca* cluster

Rina Koita[a,b], Shunsuke Otake[b], Natsuki Fukaya[b], Kenji Yamamoto[c], Akiteru Maeno[d], Haruna Kanno[a], Masaru Matsuda[b,*], Akinori Kawamura[a,*]

[a] Division of Life Science, Graduate School of Science and Engineering, Saitama University, Shimo-Okubo 255, Sakura-ku, Saitama 338-8570, Japan

[b] Center for Bioscience Research and Education, Utsunomiya University, Utsunomiya, Tochigi 321-8505, Japan.

[c] Kitatsuji Medaka, Kazo, Saitama 347-0023, Japan.

[d] Cell Architecture Laboratory, National Institute of Genetics, Yata 1111, Mishima, Shizuoka 411-8540, Japan.

[*] Author to whom correspondence should be addressed:

Akinori Kawamura, Ph.D.

Division of Life Science, Graduate School of Science and Engineering, Saitama University, Saitama 338-8570, Japan

Phone: +81-48-858-3808; E-mail: akawamur@mail.saitama-u.ac.jp

Masaru Matsuda, Ph.D.

Center for Bioscience Research and Education, Utsunomiya University, Utsunomiya 321-8505, Japan.

Phone: +81-28-649-8155; E-mail: matsuda@cc.utsunomiya-u.ac.jp

Keywords: phenotypic variation; dorsal and anal fins; *Hox* genes; medaka; transposon

Abstract

Phenotypic variation in species arises from genetic differences and environmental influences on gene expression. The differences in epigenetic modifications, such as histone modifications and DNA methylation, can also contribute to phenotypic differences, even in individuals with identical genetic information. However, the underlying molecular mechanisms remain not fully understood, particularly in vertebrates. The number of fin rays in teleosts, such as medaka, is a useful model for studying this variation. In our previous study, the teleost Hox code was shown to play a crucial role in determining the anterior-posterior identity required for the formation of dorsal and anal fins. In this study, we investigated *widefins* medaka, a spontaneous mutant displaying phenotypic variation in the number of dorsal and anal fin rays. Whole-genome sequencing revealed that an extremely large transposon, *Teratorn*, containing a herpesvirus genome, was inserted into the 3' UTR of the *hoxc12a* gene. This insertion decreased *hoxc12a* expression and, in some cases, also affected neighboring *hox* genes, resulting in variations in fin size and the presence or absence of dorsal fins. Additionally, *hoxc6a*, located 50 kb away from the insertion, was also downregulated in *widefins* medaka. These findings suggest that this large transposon insertion leads to a reduction in nearby *hox* gene expression, contributing to the phenotypic variation observed in *widefins* medaka. These results highlight the role of transposable elements and epigenetic regulation in generating phenotypic diversity in vertebrates.

Introduction

Phenotypic variation refers to observable differences in traits among individuals of a species, such as physical appearance, behavior, and learning ability. This variation arises from genetic differences and environmental factors influencing gene expression. Interestingly, phenotypic differences can also occur in individuals with identical genetic information due to epigenetic changes, such as histone modifications and DNA methylation (Peaston and Whitelaw, 2006; Triantaphyllopoulos et al., 2016). However, the underlying mechanisms of these phenomena, particularly in vertebrate development, remain not fully understood.

The number of fin rays in the dorsal and anal fins of teleosts provides a useful model for studying phenotypic variation, as quantitative analysis is possible. In our previous study with zebrafish and medaka, we demonstrated that the teleost *Hox* code is essential for establishing the anterior-posterior identity required for the formation of dorsal and anal fins (Adachi et al., 2024). During animal development, *Hox* genes are crucial for providing positional information along the anterior-posterior axis and are organized into clusters on specific chromosomes (Hubert and Wellik, 2023). Medaka possess seven *hox* clusters (Kurosawa et al., 2006). Our earlier findings revealed that *hox* genes define the regions where dorsal and anal fins can form (Adachi et al., 2024). In medaka, the loss of *hoxc11a* leads to the absence of the dorsal fin, while the loss of *hoxc12a* results in a posterior expansion of both the dorsal and anal fins. Additionally, *hoxc12a;hoxc13a* double mutants cause an even greater posterior expansion of these

fins, highlighting the important roles of these *hoxca* genes in dorsal and anal fin formation.

Medaka serves as a valuable model organism for genetic analysis in vertebrates. Since the 1950s, the late Dr. Hideo Tomita at Nagoya University in Japan collected over 80 spontaneous mutants from wild-type populations, showcasing a variety of intriguing phenotypes (Ishikawa and Araki, 2000; Tomita, 1992). Previous studies using these medaka from the Tomita collection have greatly enhanced our understanding of active transposons, such as *Tol1* and *Tol2,* in vertebrates (Koga et al., 1995; Koga et al., 1996), as well as of vertebrate developmental mechanisms (Fukamachi et al., 2001; Inohaya et al., 2010; Moriyama et al., 2012). Additionally, large-scale screenings of artificially induced mutants, similar to those conducted with zebrafish, have produced a diverse array of mutants (Furutani-Seiki et al., 2004), further contributing to the elucidation of new molecular mechanisms (Omran et al., 2008; Porazinski et al., 2015). Moreover, medaka is a popular ornamental fish, particularly in Japanese households. Since the 2000s, the number of new and improved medaka species resulting from natural crossings among breeders for ornamental purposes has rapidly increased, with reports indicating as many as 1,000 new variants. Consequently, annual medaka fairs have been held in Japan to showcase these new traits. Thus, medaka is recognized as a unique model fish for genetic analysis, benefiting from both laboratory-generated and a significant number of spontaneous mutants. These spontaneous mutants arise from mutations induced under natural conditions, revealing unexpected genetic changes that are difficult to create using genome editing technology, potentially leading to the discovery of novel biological phenomena. Furthermore, recent

advances in large-scale sequencing technology have simplified the identification of causative genomic regions in spontaneous mutants, and these mutants are expected to contribute significantly to our understanding of various biological phenomena in the future.

In this study, we analyzed *widefins* medaka, a spontaneous mutant identified by a Japanese enthusiast. Our results indicate phenotypic variation among the *widefins*: while most individuals exhibit expanded dorsal and anal fins, a small subset either lacks a dorsal fin or possesses even larger dorsal and anal fins. Whole genome analysis (WGA) revealed that a giant transposon containing the herpesvirus genome was inserted into *hoxc12a* locus, leading to decreased expression not only of the inserted gene but also of neighboring genes. Our findings suggest that the insertion of this extraordinarily large transposon with a viral genome is responsible for the phenotypic differences observed among the *widefins*.

Materials and methods

Medaka Husbandry

Medaka (*Oryzias latipes*, Hd-rR strain), provided by the National BioResource Project for Medaka (NBRP Medaka), were maintained at 25°C under a 14-hour light and 10-hour dark cycle. The *widefins* medaka, a spontaneous mutant isolated by a Japanese

breeding enthusiast in June 2016, have been continuously maintained outdoors through intercrosses over successive generations. For research purposes, the *widefins* were transferred to the laboratory at Utsunomiya University, where they were kept under the same conditions as the wild-type medaka. The *hoxc12a*^{MT1674} and *hoxc12a;hoxc13a* ^{MT1675} frameshift mutants used in this study were previously reported (Adachi et al., 2024). Embryos were collected from natural spawning, and developmental stages from embryo to adulthood were determined as previously described (Iwamatsu, 2004). All experiments involving medaka were conducted in accordance with the ARRIVE guidelines set forth by the Institutional Animal Care and Use Committee at Utsunomiya University and Saitama University.

Alizarin Red S Staining

Adult medaka were anesthetized with 2-phenoxyethanol, euthanized, and fixed in 10% neutral formalin in PBS for 3 days. They were then rinsed with D.W. for approximately 3 days, with several changes of D.W. Next, the specimens were incubated in a 2.5% trypsin solution with saturated sodium tetraborate for 1-3 days until the entire body became soft. After several washes with D.W., the specimens were soaked in 0.05% Alizarin Red S in a 0.5% KOH solution for at least 1 day. Subsequently, the stained specimens were cleared in a 1% KOH/0.5% H_2O_2 solution for several hours, followed by treatment with a 0.5% KOH/0.5% H_2O_2 solution for another 3 days, during which the 0.5% KOH was gradually replaced with 100% glycerol. For the stained fish, images were captured using a fluorescence stereomicroscope (M205 FA, Leica).

Ossified Bone Staining of Juvenile Medaka with Calcein

To visualize the calcified bones in larvae and juveniles, calcein staining was performed as previously described (Koita et al., 2024). After several washes, the images were captured using a fluorescence stereomicroscope (M205 FA, Leica).

Micro-CT scan

Micro-CT scan of an adult medaka was carried out as previously described (Adachi et al., 2024; Akama et al., 2020; Maeno et al., 2024). Briefly, adult medaka were fixed with 4% paraformaldehyde in PBS at 4 °C overnight and transferred to 70% ethanol. Using an X-ray micro-CT (ScanXmate-E090S105; Comscantechno), the fixed specimens were scanned at a tube peak voltage of 85 kV and a tube current of 90 μA. For higher-resolution whole-body scanning of medaka, the whole body was scanned in four parts for each sample. For each part, the specimens were rotated 360° in 0.3° steps, generating 1200 projection images of 992 × 992 pixels. The micro-CT data of each part were reconstructed using coneCTexpress software (Comscantechno) and stored as a dataset with an isotropic resolution of 14.2 to 18.3 μm. Finally, data from four locations were combined to generate a dataset of the entire specimen. Three-dimensional image analysis was performed using OsiriX MD software (Pixmeo).

Whole-Genome Sequencing of Widefins **Medaka**

After anesthesia with 2-phenoxyethanol, the adult male *widefins* medaka used for the complementation test was euthanized, and the entire body was fixed in 100% ethanol. Following fixation, the dissected testis was incubated overnight in a lysis buffer containing Proteinase K. Genomic DNA was then extracted using the Maxwell RSC Tissue DNA Kit (Promega) according to the manufacturer's protocol. After the quality check of the extracted genomic DNA, whole-genome sequencing of the *widefins* medaka was performed by Gene Nex (Chemical Dojin). The raw sequence data obtained from Gene Nex were initially trimmed for adapter sequences using Trimmomatic software (Java version “21.0.1,” 2023-10-17 LTS) to extract paired-end data. A filtering process was subsequently applied to retain only high-quality sequences. Reference genomes, specifically the *Oryzias latipes* Hd-rR strain (Oryzias_latipes.ASM223467v1.dna.toplevel.fa) and the HNI strain (Oryzias_latipes_hni.ASM223471v1.dna.toplevel.fa), were downloaded along with their respective annotation files (Oryzias_latipes.ASM223467v1.113.gff3 for the Hd-rR strain and Oryzias_latipes_hni.ASM223471v1.113.gff3 for the HNI strain). Next, an index of the reference genome was created using BWA software (0.7.18-r1243-dirty), and the sequence data were mapped against the reference genome using BWA-MEM, which is suitable for Illumina paired-end reads. The mapping results were stored in SAM files, which were then converted to BAM files using Samtools (1.21) and sorted. Finally, the reference file and the BAM file of the sequence data were opened in Integrative Genomics Viewer (2.18.4) to visualize the sequences mapped by BWA, while the annotation files were used to map protein-coding genes on the genome.

Whole-mount in situ hybridization

Whole-mount *in situ* hybridization was performed as previously described (Adachi et al., 2024). For the stained fish, the images were captured using a fluorescence stereomicroscope (M205 FA, Leica).

Results

The *widefins* medaka, a spontaneous mutant characterized by dorsal and anal fins that expand posteriorly, was isolated by a Japanese medaka breeder. This phenotype resembles that of the medaka *hoxc12a* and *hoxc12a;c13a* frameshift-induced mutants that we previously created using CRISPR-Cas9 (Adachi et al., 2024). To compare the phenotypes of *widefins* with those of the *hoxc12a* and *hoxc12a;c13a* mutants, we first analyzed the skeletal structures of adult *widefins* obtained from intercrosses using alizarin red staining and micro-CT scans (Fig. 1a-h). The analysis of fin ray counts revealed that some individuals of *widefins* exhibited varying phenotypes, unlike the consistent similarities observed in *hoxc12a* or *hoxc12a;c13a* mutants. Generally, the *widefins* medaka displayed a greater number of dorsal fin rays compared to the wild-type (n=9/11); many individuals had a fin ray count comparable to that of *hoxc12a* mutants (Fig. 1d, h, i), while a few had more fin rays, though still fewer than those of *hoxc12a;c13a* mutants (Fig. 1f, i). Notably, one individual examined among the *widefins* lacked a dorsal fin (Fig. 1e, i), sharply contrasting with the typical phenotype

of *widefins*. This dorsal fin-absent phenotype resembles that of the spontaneous *dorsalfinless* medaka, which encodes the hypomorphic allele of *hoxc11a* (Adachi et al., 2024). Additionally, the anal fin of the *widefins* exhibited a greater number of fin rays; typical individuals showed a comparable or slightly higher count than *hoxc12a* mutants (Fig. 1d, j). In some individuals, the number of fin rays was comparable to that of *hoxc12a;c13a* mutants (Fig. 1f, j). Furthermore, the number of vertebrae in *widefins* medaka also displayed variation similar to the phenotypes between *hoxc12a* mutants and *hoxc12a;c13a* mutants (Fig. 1k). These results indicate that *widefins* exhibit individual variations in the phenotypes of their dorsal and anal fins, as well as vertebrae. Overall, the phenotypes of *widefins* medaka are similar to those of *hoxc12a* mutants or *hoxc12a;c13a* mutants, showing a slightly stronger phenotype than *hoxc12a* mutants but weaker than *hoxc12a;c13a* mutants. Furthermore, in rare instances, *widefins* display a dorsalfinless phenotype resembling that of *hoxc11a* mutants.

To determine whether the abnormalities in *widefins* are associated with the loss of function of *hoxc12a* and *hoxc13a*, we conducted a genetic complementation test by crossing *widefins* with frameshift-induced *hoxc12a;c13a* heterozygous medaka. Analysis of 21 juveniles from this cross revealed that the *widefins* and *hoxc12a;c13a* mutations do not complement each other. We observed some individuals with a normal number of fin rays (*n*=9) and others with an increased number of fin rays (*n*=12) (Fig. 2a-g). Genotyping for frameshift mutations showed that all juvenile fish with a normal number of fin rays were identified as $hoxc12a^{+/wf-};c13a^{+/wf}$ while all juvenile fish with an increased number of fin rays were identified as $hoxc12a^{-/wf};c13a^{-/wf}$. These genetic

results indicate that *widefins* encodes a recessive mutation associated with the loss of function of the *hox* genes, specifically *hoxc12a* and *hoxc13a*.

Next, we performed whole-genome sequencing (WGS) to identify the genomic region responsible for the features observed in *widefins* medaka. Our results suggest that the mutation is likely located in the region encompassing *hoxc12a* and *hoxc13a* loci. The extensive sequence data acquired from WGS were mapped to the reference medaka genome, with the read sequences typically overlapping. However, we observed that all mapped read sequences were interrupted at a specific position in the 3' UTR of *hoxc12a*, indicating the insertion of a DNA element (Fig. 3a). Further examination of the sequences mapped to this region in *widefins* revealed that inserted DNA element corresponded to sequences at both ends of a DNA transposon known as *Teratorn*. Originally identified as an inserted DNA element in the medaka spontaneous mutant *Double analfin* (Moriyama et al., 2012), *Teratorn* is unique among DNA transposons as it is the largest, measuring 180 kb in length, resulting from the fusion of the herpes virus genome with *piggyBac*-type transposon elements (Inoue et al., 2017). Using PCR, we confirmed that the 5’ and 3’ ends of the *Teratorn* transposon sequences were indeed inserted into the 3’UTR of *hoxc12a* in *widefins* (Fig. 3b, c). Additionally, we confirmed that the coding sequences of *hoxc12a* and *hoxc13a* remain intact in *widefins*. To investigate whether the insertion of *Teratorn* affects the expression of *hoxc12a*, we next compared the expression patterns of *hox* genes in wild-type and *widefins* embryos using whole-mount *in situ* hybridization. We found that the expression of *hoxc12a* was significantly downregulated in *widefins* (Fig. 4c, d), which could lead to the expansion of dorsal and anal fins, as observed in *hoxc12a* mutants. Additionally, we examined the

expression of neighboring *hox* genes. Although not to the same extent as *hoxc12a*, the expression patterns of *hoxc11a* were also reduced in *widefins* (Fig. 4a, b). If the reduction in *hoxc11a* expression were more pronounced, the loss of *hoxc11a* function would result in the absence of a dorsal fin in *widefins* (Fig. 1e), because *hoxc11a* mutants lack the dorsal fin. Furthermore, we analyzed the expression pattern of *hoxc13a*, but could not detect significant differences (Fig. 4e, f). These results suggest that the insertion of the *Teratorn* transposon in the 3'UTR of *hoxc12a* significantly reduces the expression of *hoxc12a* and contributes to the decreased expression of neighboring *hox* genes in the *hoxca* cluster.

Since the insertion of the *Teratorn* transposon decreased the expression of *hoxc12a* and nearby *hoxc11a*, we investigated whether this effect extends further. In medaka, *hoxc6a* locus is located approximately 50 kb away from the insertion site (Fig. 5a). Loss of function of medaka *hoxc6a* results in an anteriorization phenotype characterized by the absence of pleural ribs in the second vertebra, where rib formation should occur; instead, rib formation begins at the third vertebra (Maeno et al., 2024). Following a detailed examination of *widefins* medaka, we found that the most anteriorly formed rib, which differed from that of wild-type, was formed from the third vertebra (Fig. 5b-g; *n*=5/5). This phenotype resembles that of *hoxc6a* mutants, so we compared the expression patterns of *hoxc6a* between wild-type and *widefins* embryos. Using whole-mount in *situ* hybridization, we found that the expression of *hoxc6a* was decreased in *widefins* compared to wild-type (Fig. 5h, i; *n*=4/5). Taken together, these results suggest that the insertion of a giant *Teratorn* transposon containing viral genes reduces the expression of *hoxc6a*, which is located 50 kb away from the insertion site.

Discussion

In this study, we analyzed *widefins* isolated by a Japanese medaka enthusiast to create ornamental fish with novel traits. Our detailed analysis revealed that the phenotypes of *widefins* can be categorized into three groups based on individual differences in the number of dorsal and anal fin rays. This variation was particularly pronounced in the dorsal fin, where the number of fin rays differed significantly among individuals, with some *widefins* even lacking a dorsal fin entirely. Furthermore, WGS analysis identified a *Teratorn* transposon insertion, estimated at 180 kb, within the 3' UTR of *hoxc12a* in *widefins* medaka. Our results suggest that this transposon insertion decreases the expression of not only *hoxc12a* but also adjacent *hox* genes in some cases, leading to phenotypic variation.

To explain the phenotypic variation observed in *widefins* medaka, we propose a model in which the insertion of the *Teratorn* transposon stochastically downregulates the expression not only of the inserted *hoxc12a* but also of nearby *hox* genes. The *Teratorn* transposon is large, incorporating elements from a herpesvirus genome, which makes up approximately 80% of its total sequence (Inoue et al., 2017). Typically, the integration of foreign DNA, such as viral genomes and transposable elements, into mammalian cells and plants leads to the silencing of the foreign DNA through heterochromatin formation (Cabrera et al., 2022; Deniz et al., 2019; Slotkin and Martienssen, 2007). Epigenetic markers of heterochromatin, such as histone H3 trimethylation at lysine 9 (H3K9me3) and CpG methylation, spread to adjacent regions

due to positive feedback from the recruitment of histone methyltransferases at these methylation sites (Al-Sady et al., 2013; Grewal and Jia, 2007; Grewal, 2023). This process not only affects the insertion sites but also establishes repressive chromatin states in surrounding areas. However, the mere insertion of a transposon does not necessarily impact the surrounding region. In our previous study of spontaneous *dorsalfinless* medaka mutants, which lack a dorsal fin, we demonstrated that a transposon of about 12 kb was inserted into the 3' UTR of *hoxc11a* (Adachi et al., 2024), adjacent to *hoxc12a* where the *Teratorn* insertion was identified in *widefins* medaka in the current study. Interestingly, this 12 kb DNA transposon significantly reduces the expression of *hoxc11a*, but it does not appear to influence the expression of nearby *hox* genes, including *hoxc12a*. Given that *Teratorn* spans 180 kb (Inoue et al., 2017), it implies that around 150 kb of virus-derived long sequences have integrated into *hoxc12a*, suggesting a high capacity to form highly condensed heterochromatic regions. Since the 30-40 copies of *Teratoron* are distributed throughout the haploid medaka genome (Inoue et al., 2017) and cannot be individually distinguished, we are unable to provide direct evidence that heterochromatin is indeed formed at the *Teratorn* insertion into *hoxc12a* in *widefins*. However, based on the previous results, the insertion of *Teratorn* may generate heterochromatin at the insertion site, which, in some individuals, could have expanded into the surrounding region, significantly impacting the expression of neighboring *hox* genes.

Based on this proposed model, we presume that in most cases of *widefins* medaka, the severe effect of the *Teratorn* insertion is a substantial reduction in the expression of the inserted *hoxc12a*. It is difficult to distinguish whether the decrease in

hoxc12a mRNA is due to transcriptional repression and/or mRNA destabilization caused by the *Teratorn* insertion. While the expression of the other gene closest to the insertion site, *hoxc11a*, is also reduced, it is likely that the amount required to form the dorsal fin is maintained. Consequently, *widefins* medaka generally display a phenotype similar to that of *hoxc12a* mutants, characterized by wide dorsal and anal fins (Fig. 6a). We hypothesize that in some individuals, heterochromatin driven by *Teratorn* may extend over larger regions and impact neighboring genes. If *Teratorn* strongly affects not only the expression of the inserted *hoxc12a* but also that of closely located *hoxc11a*, both *hoxc11a* and *hoxc12a* could become functionally impaired. This may lead to the emergence of medaka exhibiting both phenotypes: loss of the dorsal fin due to the loss of *hoxc11a* function and a wide anal fin due to the loss of *hoxc12a* function (Fig. 6c). Similarly, if the *hoxc13a* locus is affected, both *hoxc12a* and *hoxc13a* could become dysfunctional, resulting in medaka with wider dorsal and anal fins, similar to *hoxc12a;hoxc13a* mutants (Fig. 6d). The variation in phenotypes among *widefins* individuals resembles position effect variation (PEV), which has been extensively studied in *Drosophila*. PEV is a variegation in gene expression caused by gene silencing in some cells due to abnormal chromosomal translocation or inversion of genes near heterochromatin (Elgin and Reuter, 2013). Our analysis suggests that the insertion of a giant transposon such as *Teratorn* can produce PEV-like phenomena in teleosts.

Additionally, the *hoxc6a* locus, located approximately 50 kb from the insertion site, was also found to be downregulated. It remains unclear whether this downregulation results from further expansion of heterochromatin regions, as vertebrate *Hox* genes are intricately regulated by enhancers both within and outside the *Hox*

clusters (Acemel et al., 2016; Afzal and Krumlauf, 2022; Andrey et al., 2013; Mallo and Alonso, 2013). It is possible that the enhancer of *hoxc6a* was inhibited by the *Teratorn* insertion. Furthermore, *widefins* have been selectively bred and maintained to create ornamental medaka with a novel trait, deliberately excluding the dorsal fin deletion and increased vertebra phenotypes. According to the breeder, the dorsal fin absent phenotype was more frequently observed when *widefins* were first identified. After several generations of continuous selection for the widefin trait, this phenotype became more pronounced. While this study describes the phenotypic differences observed among *widefins* individuals, variations in the proportion of phenotypes across generations were also noted, although these changes were not recorded. Epigenetic modifications have been shown to persist across generations in fish (Iwanami et al., 2020; Kelley et al., 2021), suggesting that a similar phenomenon may be occurring in *widefins* medaka.

Data Availability

All data necessary for confirming the conclusions of the article are present within the article, figures, and supplementary data. *widefins* medaka are commercially available in aquarium shops. The frozen sperm of *hoxc12a* and *hoxc12a;c13a* mutant medaka used in this study have been deposited in the NBRP medaka repository (https://shigen.nig.ac.jp/medaka/) and are available upon request. The induced mutations of *hoxc12a* and *hoxc12a;c13a* mutants, generated by CRISPR-Cas9, were previously described (Adachi et al., 2024). The raw sequence reads of *widefins* medaka

obtained through next-generation sequencing have been submitted to DDBJ and will be made available.

Acknowledgments

We thank the NBRP medaka for providing fish and preserving the mutant lines used in this study. This work was supported by KAKENHI Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology, Japan (23K05790 to A.K.), and by the National Institute of Genetics under the Joint Research and Research Meeting (NIG-JOINT) program (18A2022, 31A2023, 26A2024) to A.K.).

Author Contributions

A.K. and M.M. conceived and designed the experiments, R.K., S.O., N.F., and H.K. analyzed *widefins* medaka, K.Y. isolated *widefins* medaka, A.M. performed micro-CT scan, A.K. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

Acemel, R.D., Tena, J.J., Irastorza-Azcarate, I., Marletaz, F., Gomez-Marin, C., de la Calle-Mustienes, E., Bertrand, S., Diaz, S.G., Aldea, D., Aury, J.M.*, et al.* (2016). A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. Nat Genet *48*, 336-341.

Adachi, U., Koita, R., Seto, A., Maeno, A., Ishizu, A., Oikawa, S., Tani, T., Ishizaka, M., Yamada, K., Satoh, K.*, et al.* (2024). Teleost Hox code defines regional identities competent for the formation of dorsal and anal fins. Proc Natl Acad Sci U S A *121*, e2403809121.

Afzal, Z., and Krumlauf, R. (2022). Transcriptional Regulation and Implications for Controlling Hox Gene Expression. J Dev Biol *10*.

Akama, K., Ebata, K., Maeno, A., Taminato, T., Otosaka, S., Gengyo-Ando, K., Nakai, J., Yamasu, K., and Kawamura, A. (2020). Role of somite patterning in the formation of Weberian apparatus and pleural rib in zebrafish. J Anat *236*, 622-629.

Al-Sady, B., Madhani, H.D., and Narlikar, G.J. (2013). Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. Mol Cell *51*, 80-91.

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science *340*, 1234167.

Cabrera, A., Edelstein, H.I., Glykofrydis, F., Love, K.S., Palacios, S., Tycko, J., Zhang, M., Lensch, S., Shields, C.E., Livingston, M.*, et al.* (2022). The sound of silence: Transgene silencing in mammalian cell engineering. Cell Syst *13*, 950-973.

Deniz, O., Frost, J.M., and Branco, M.R. (2019). Regulation of transposable elements by DNA modifications. Nat Rev Genet *20*, 417-431.

Elgin, S.C., and Reuter, G. (2013). Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. Cold Spring Harb Perspect Biol *5*, a017780.

Fukamachi, S., Shimada, A., and Shima, A. (2001). Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. Nat Genet *28*, 381-385.

Furutani-Seiki, M., Sasado, T., Morinaga, C., Suwa, H., Niwa, K., Yoda, H., Deguchi, T., Hirose, Y., Yasuoka, A., Henrich, T.*, et al.* (2004). A systematic genome-wide screen for

mutations affecting organogenesis in Medaka, Oryzias latipes. Mech Dev *121*, 647-658.
Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. Nat Rev Genet *8*, 35-46.
Grewal, S.I.S. (2023). The molecular basis of heterochromatin assembly and epigenetic inheritance. Mol Cell *83*, 1767-1785.
Hubert, K.A., and Wellik, D.M. (2023). Hox genes in development and beyond. Development *150*.
Inohaya, K., Takano, Y., and Kudo, A. (2010). Production of Wnt4b by floor plate cells is essential for the segmental patterning of the vertebral column in medaka. Development *137*, 1807-1813.
Inoue, Y., Saga, T., Aikawa, T., Kumagai, M., Shimada, A., Kawaguchi, Y., Naruse, K., Morishita, S., Koga, A., and Takeda, H. (2017). Complete fusion of a transposon and herpesvirus created the Teratorn mobile element in medaka fish. Nat Commun *8*, 551.
Ishikawa, Y., and Araki, K. (2000). [Medaka induced mutants]. Tanpakushitsu Kakusan Koso *45*, 2820-2828.
Iwamatsu, T. (2004). Stages of normal development in the medaka Oryzias latipes. Mech Dev *121*, 605-618.
Iwanami, N., Lawir, D.F., Sikora, K., C, O.M., Takeshita, K., Schorpp, M., and Boehm, T. (2020). Transgenerational inheritance of impaired larval T cell development in zebrafish. Nat Commun *11*, 4505.
Kelley, J.L., Tobler, M., Beck, D., Sadler-Riggleman, I., Quackenbush, C.R., Arias Rodriguez, L., and Skinner, M.K. (2021). Epigenetic inheritance of DNA methylation changes in fish living in hydrogen sulfide-rich springs. Proc Natl Acad Sci U S A *118*.
Koga, A., Inagaki, H., Bessho, Y., and Hori, H. (1995). Insertion of a novel transposable element in the tyrosinase gene is responsible for an albino mutation in the medaka fish, Oryzias latipes. Mol Gen Genet *249*, 400-405.
Koga, A., Suzuki, M., Inagaki, H., Bessho, Y., and Hori, H. (1996). Transposable element in fish. Nature *383*, 30.
Koita, R., Oikawa, S., Tani, T., Matsuda, M., and Kawamura, A. (2024). Live Visualization of Calcified Bones in Zebrafish and Medaka Larvae and Juveniles Using Calcein and Alizarin Red S. Bio Protoc *14*, e5142.
Kurosawa, G., Takamatsu, N., Takahashi, M., Sumitomo, M., Sanaka, E., Yamada, K., Nishii, K., Matsuda, M., Asakawa, S., Ishiguro, H.*, et al.* (2006). Organization and structure of hox gene loci in medaka genome and comparison with those of pufferfish and zebrafish genomes. Gene *370*, 75-82.
Maeno, A., Koita, R., Nakazawa, H., Fujii, R., Yamada, K., Oikawa, S., Tani, T., Ishizaka, M.,

Satoh, K., Ishizu, A., *et al.* (2024). The Hox code responsible for the patterning of the anterior vertebrae in zebrafish. Development *151*.

Mallo, M., and Alonso, C.R. (2013). The regulation of Hox gene expression during animal development. Development *140*, 3951-3963.

Moriyama, Y., Kawanishi, T., Nakamura, R., Tsukahara, T., Sumiyama, K., Suster, M.L., Kawakami, K., Toyoda, A., Fujiyama, A., Yasuoka, Y., *et al.* (2012). The medaka zic1/zic4 mutant provides molecular insights into teleost caudal fin evolution. Curr Biol *22*, 601-607.

Omran, H., Kobayashi, D., Olbrich, H., Tsukahara, T., Loges, N.T., Hagiwara, H., Zhang, Q., Leblond, G., O'Toole, E., Hara, C., *et al.* (2008). Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins. Nature *456*, 611-616.

Peaston, A.E., and Whitelaw, E. (2006). Epigenetics and phenotypic variation in mammals. Mamm Genome *17*, 365-374.

Porazinski, S., Wang, H., Asaoka, Y., Behrndt, M., Miyamoto, T., Morita, H., Hata, S., Sasaki, T., Krens, S.F.G., Osada, Y., *et al.* (2015). YAP is essential for tissue tension to ensure vertebrate 3D body shape. Nature *521*, 217-221.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet *8*, 272-285.

Tomita, H. (1992). The lists of the mutants and strains of the medaka, common, gambusia, silver crucian cacrp, goldfish, and golden venus fish maintained in the Laboratory of freshwater fish stocks, Nagoya university. Fish Biol J Medaka, 45-47.

Triantaphyllopoulos, K.A., Ikonomopoulos, I., and Bannister, A.J. (2016). Epigenetics and inheritance of phenotype variation in livestock. Epigenetics Chromatin *9*, 31.

Figure legends

Figure 1. *widefins* medaka exhibit a phenotypic variation, particularly in the dorsal fin. (a-d) Skeletal structures of the posterior body in wild-type (a; *n*=3), *hoxc12a*$^{-/-}$ (b; *n*=3), *hoxc12a*$^{-/-}$*;c13a*$^{-/-}$ (c; *n*=3), and *widefins* (d-f; *n*=11) fish. The ossified bones were stained with Alizarin Red S. The number of dorsal fin rays for each individual is

indicated in the upper right corner, while the number of anal fin rays is shown in the lower right corner. Three different phenotypes of *widefins* medaka are presented, illustrating individual variations in the number of dorsal and anal fin rays. Scale bar: 2 mm. (g, h) Micro-CT scan images of wild-type and *widefins* adult medaka. Scale bar: 2 mm. 3D movies are available in Supplementary Movies 1 and 2. (i-k) Comparison of the number of fin rays in the dorsal fin (i) and anal fin (j), as well as the number of vertebrae (k) among the mutants. (l, m) Comparison of the anterior and posterior ends of vertebrae attached to the fin rays of the dorsal fin (l) and anal fin (m) via the radials. The anterior end of the vertebra is indicated by a circle, while the posterior end is marked with a rhomboid. A single *widefins* fish with a missing dorsal fin (e) was excluded from the statistical analysis in (l). For all statistical analyses, the Tukey-Kramer test was performed on data obtained from Alizarin Red S-stained fish and micro-CT scans. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$.

Figure 2. *widefins* medaka do not complement *hoxc12a;c13a* mutations. (a-d) The calcified bones of juvenile fish, obtained by crossing male *widefins* with female *hoxc12a;c13a* heterozygous mutants, were visualized using calcein staining. Enlarged views (top) and the entire body (bottom) of the same representative fish are shown. After capturing images, the presence or absence of frameshift mutations in *hoxc12a* and *hoxc13a*, introduced by CRISPR-Cas9, was examined by PCR as previously described (Adachi et al., 2024). Scale bar: 1 mm. (e-g) Comparison of the number of fin rays in the dorsal fin (e) and anal fin (f), as well as the number of vertebrae (g). The data,

obtained from *hoxc12a*$^{+/wf-}$;*c13a*$^{+/wf}$ (n=9), *hoxc12a*$^{-/wf}$;*c13a*$^{-/wf}$ (n=12), *hoxc12a*$^{-/-}$ (n=6), and *hoxc12a*$^{-/-}$;*c13a*$^{-/-}$ (n=6) fish, were compared. Previously described data for *hoxc12a*$^{-/-}$ and *hoxc12a*$^{-/-}$;*c13a*$^{-/-}$ fish were also included for comparison (Adachi et al., 2024). The Tukey-Kramer test was performed. *$P < 0.05$, ***$P < 0.001$, and ****$P < 0.0001$.

Figure 3. Insertion of the *Teratorn* transposon in the 3' UTR of *hoxc12a*. (a) Schematic representation of the medaka *hoxc12a* genomic structure, consisting of two exons. Below the diagram, short read sequences mapped to the reference genome in exon 2 of *hoxc12a* are displayed using Integrative Genomics Viewer software. The genomic region where the mapped read sequences terminate abnormally is outlined in a red frame, indicating the insertion of the *Teratorn* transposon into the 3' UTR of *hoxc12a*. (b) A schematic of the *hoxc12a* 3' UTR comparing wild-type and *widefins* medaka, showing the positions of the primers used for PCR-based genotyping. (c) Genotyping of wild-type and *widefins* medaka was performed using three primers (indicated on the right side of the panel) for each end of the *Teratorn* transposon insertion site. Genomic DNA was extracted from three individuals of both wild-type and *widefins* medaka and used as templates for PCR. Following the reaction, PCR products were separated on a 2% agarose gel in 0.5x TBE buffer. In *widefins*, the PCR products could not be amplified using the c12a-3'UTR-s and c12a-3'UTR-as primers due to the *Teratorn* insertion.

Figure 4. Significantly decreased expression of *hoxc12a* and reduced expression of *hoxc11a* in *widefins* medaka. (a-f) Expression patterns of *hoxc11a*, *hoxc12a*, and *hoxc13a* were analyzed between wild-type and *widefins*. Whole-mount *in situ* hybridization was performed on embryos at stages 29-30. For each staining, at least three different specimens were used to confirm reproducibility. Lateral view of the posterior body. Scale bar: 100 μm.

Figure 5. The expression of *hoxc6a* is decreased, resulting in anteriorizing vertebral patterning in *widefins* medaka. (a) Schematic representation of medaka *hoxca* cluster. Boxes represent the *hox* genes, and arrows indicate their transcriptional orientation. The arrowhead marks the insertion site of the *Teratorn* transposon in the 3'UTR of *hoxc12a*. (b, c) Skeletal structures of the anterior vertebrae in wild-type (n=6) and *widefins* (n=4). The ossified bones were stained with Alizarin Red S. Numbers indicate the number from the most anterior vertebra. Yellow arrowheads point to the anteriormost pleural rib in both wild-type and *widefins.* Scale bar: 2 mm. (d-g) Micro-CT scanning of the anterior vertebrae in wild-type and *widefins.* Scale bar: 2 mm. (h, i) Expression patterns of *hoxc6a* in wild-type and *widefins* embryos. Whole-mount *in situ* hybridization was performed on embryos at stages 29-30. At least three different specimens were used for each staining to confirm reproducibility. Lateral view of the posterior body. Scale bar: 100 μm.

Figure 6. Proposed model illustrating phenotypic variation among *widefins* individuals. (a) In medaka, the genomic loci of *hoxc11a*, *hoxc12a*, and *hoxc13a* are aligned in the same transcriptional direction (arrow in right schematic). In the posterior body, these *hox* genes define regional identities competent for the formation of dorsal and anal fins (left schematic). (b) The insertion of a giant *Teratorn* transposon containing a viral genome within *hoxc12a* (arrowhead) leads to the formation of heterochromatin in the surrounding region, significantly reducing *hoxc12a* expression (dashed arrow). This reduction results in wide dorsal and anal fins, resembling the phenotype observed in *hoxc12a* mutants (Adachi et al., 2024). (c, d) In some instances, the insertion of the *Teratorn* transposon may expand the heterochromatin to a broader genomic region, affecting nearby *hox* genes. When the heterochromatin expands to the *hoxc11a* locus, *hoxc11a* expression is considerably reduced, resulting in a wide anal fin due to *hoxc12a* deficiency, as well as the absence of a dorsal fin seen in *hoxc11a* mutants (Adachi et al., 2024). If the heterochromatin extends to the *hoxc13a* locus, the expression of *hoxc13a* is also reduced, leading to fish with even wider dorsal and anal fins, as observed in *hoxc12a;hoxc13a* mutants (Adachi et al., 2024).

Supplementary Movie 1. Micro-CT scanning of the whole skeletal structure of wild-type medaka.

Micro-CT scanning was performed on adult wild-type medaka to visualize the entire skeletal structure.

Supplementary Movie 2. Micro-CT scanning of the whole skeletal structure of *widfeins* medaka.

Micro-CT scanning was conducted on adult *widefins* medaka to show the entire skeletal structure. The most common phenotype of *widefins* medaka, which resembles that of *hoxc12a* mutants, is presented.

Supplementary Movie 3. Micro-CT scanning of the anterior vertebrae of wild-type and *widefins* medaka.

Micro-CT scanning of the anterior vertebrae of wild-type and *widefins* medaka was conducted using the same specimens featured in Supplementary Movies 1 and 2.

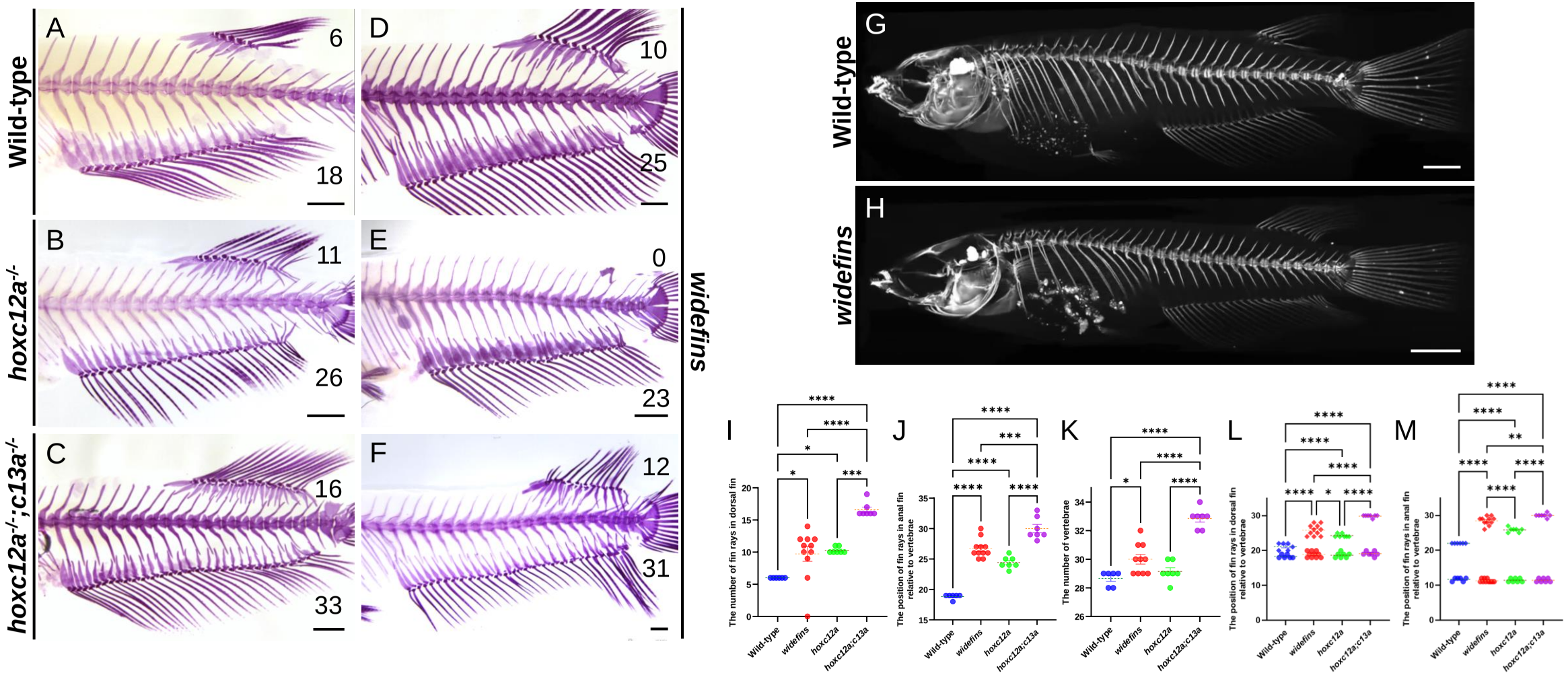

Figure 1. *widefins* medaka exhibit a phenotypic variation, particularly in dorsal fin. (A-D) Skeletal structures of the posterior body in wild-type (A; *n*=3), *hoxc12a*$^{-/-}$ (B; *n*=3), *hoxc12a*$^{-/-}$*;c13a*$^{-/-}$ (C; *n*=3), and *widefins* (D-F; *n*=11) fish. The ossified bones were stained with Alizarin Red S. The number of dorsal fin rays for the individuals is indicated in the upper right corner, while the number of anal fin rays is shown in the lower right corner. Three different phenotypes of *widefins* medaka were presented because *widefins* exhibited individual variations in the number of dorsal and anal fin rays. Scale bar: 2 mm. (G, H) Micro-CT scan images of wild-type and *widefins* adult medaka. Scale bar: 2 mm. 3D movies are available in Supplementary Movies 1 and 2. (I-K) Comparison of the number of fin rays in the dorsal fin (I) and anal fin (J), as well as the number of vertebrae (K) among the mutants. (L, M) Comparison of the anterior and posterior ends of vertebrae attached to the fin rays of the dorsal fin (L) and anal fin (M) via the radials. The anterior end of the vertebra is indicated by a circle, while the posterior end is marked with a rhomboid. A single *widefins* fish with a missing dorsal fin (E) was excluded from this statistical analysis in L. For all statistical analysis, the Tukey-Kramer test was performed on data obtained from Alizarin Red S-stained fish and micro-CT scans. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$.

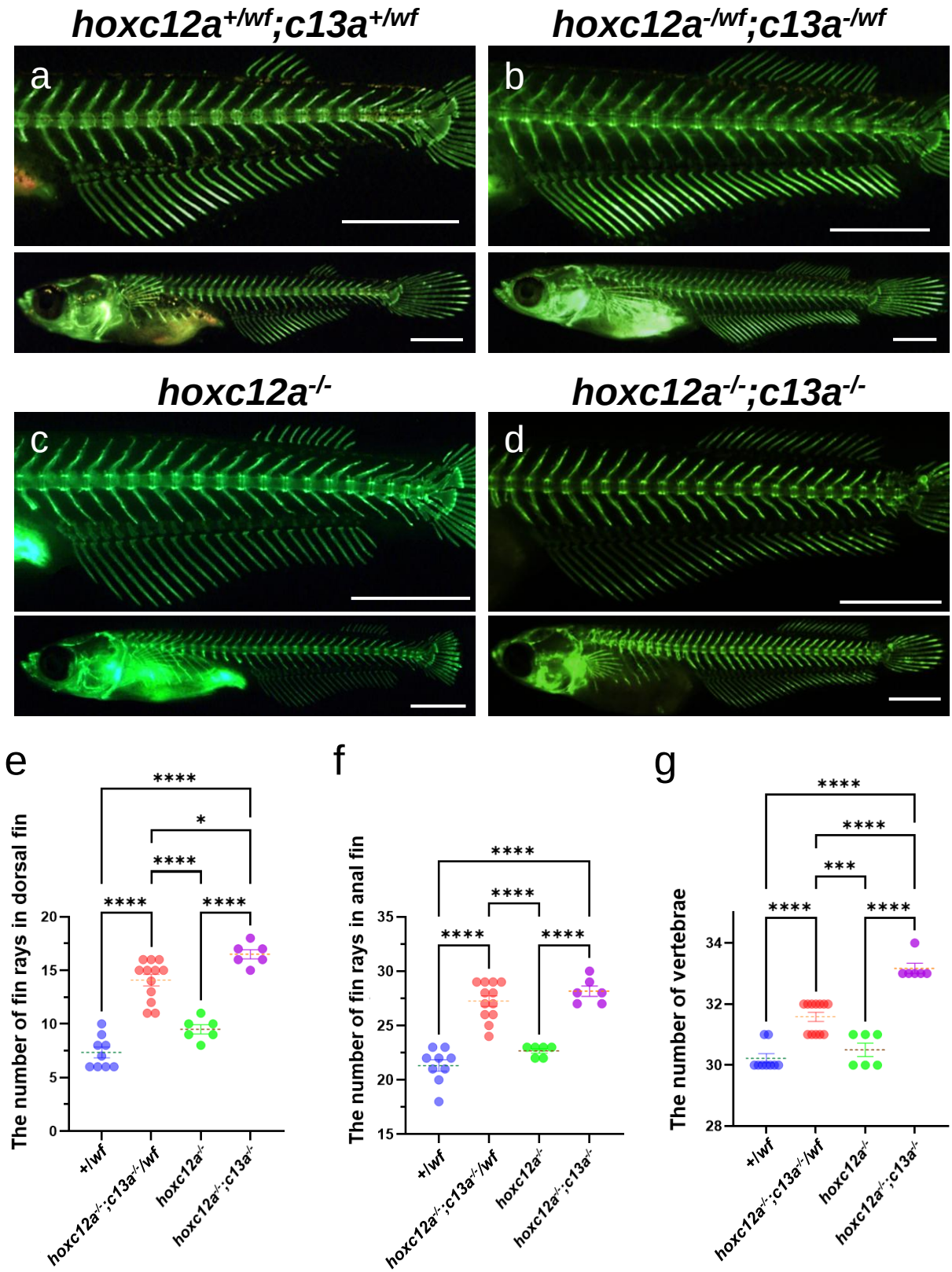

Figure 2. Medaka *widefins* does not complement *hoxc12a;c13a* mutations. (a-d) The calcified bones of juvenile fish, obtained by crossing *widefins* with *hoxc12a;c13a* heterozygous mutants, were visualized using calcein staining. Enlarged views (top) and the entire body (bottom) of the same representative fish are shown. After capturing images, the presence or absence of frameshift mutations in *hoxc12a* and *hoxc13a*, introduced by CRISPR-Cas9, were examined by PCR as previously described. Scale bar: 1 mm. (e-g) Comparison of the number of fin rays in the dorsal fin (e) and anal fin (f), as well as the number of vertebrae (g). The data, obtained from *hoxc12a*$^{+/wf-}$*;c13a*$^{+/wf}$ (n=9), *hoxc12a*$^{-/wf}$*;c13a*$^{-/wf}$ (n=12), *hoxc12a*$^{-/-}$ (n=6), and *hoxc12a*$^{-/-}$*;c13a*$^{-/-}$ (n=6) fish, were compared. For *hoxc12a*$^{-/-}$ and *hoxc12a*$^{-/-}$*;c13a*$^{-/-}$ fish, the previously described data were used for comparison. The Tukey-Kramer test was performed. *$P < 0.05$, ***$P < 0.001$, and ****$P < 0.0001$.

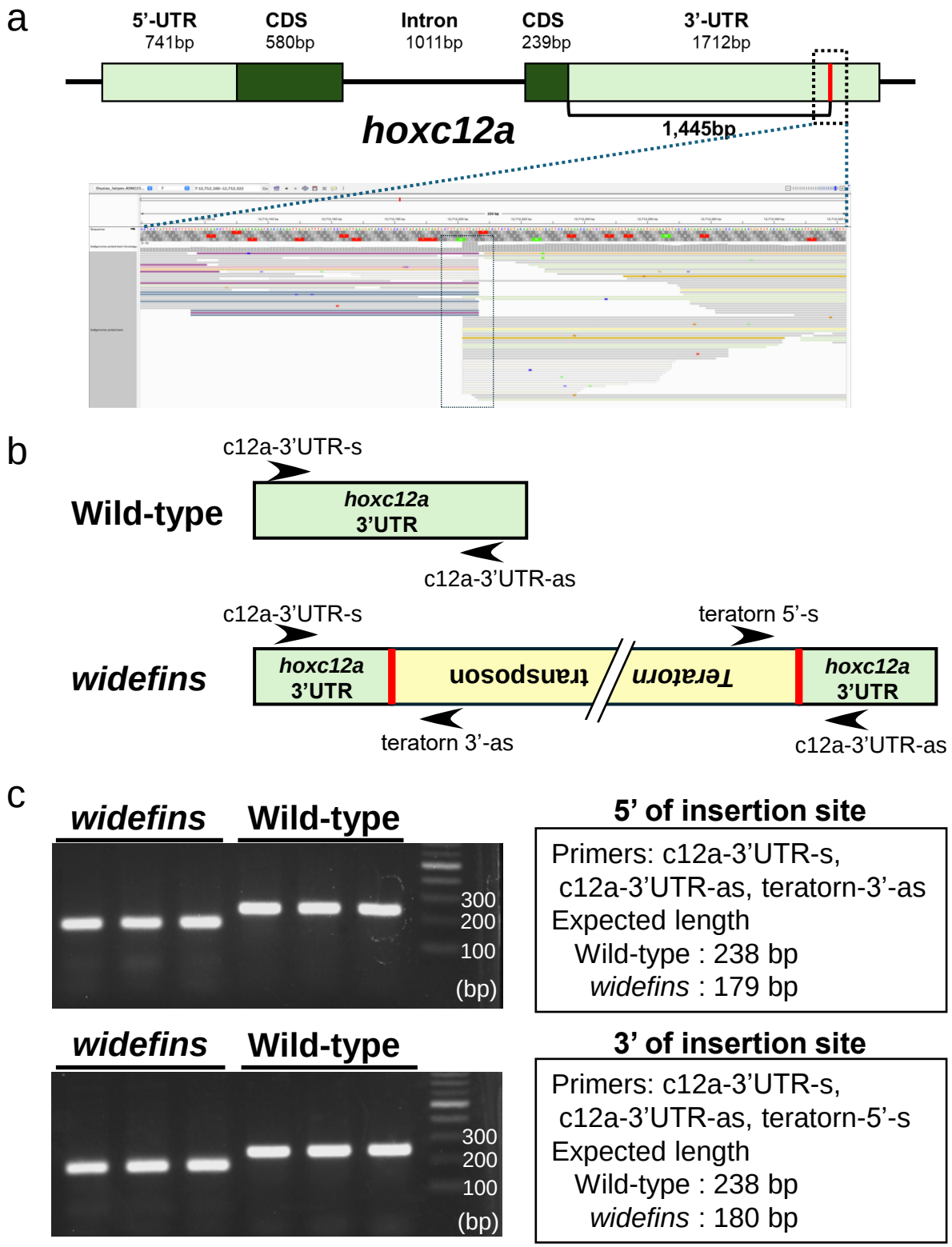

Figure 3. Insertion of the *Teratorn* transposon in the 3' UTR of *hoxc12a*. (a) Schematic representation of the medaka *hoxc12a* genomic structure, which consists of two exons. Below the diagram, short read sequences mapped to the reference genome in exon 2 of *hoxc12a* are displayed using Integrative Genomics Viewer software. The genomic region where the mapped read sequences terminate abnormally is outlined in a red frame, indicating the insertion of the *Teratorn* transposon into the 3' UTR of *hoxc12a*. (b) A schematic of the *hoxc12a* 3' UTR comparing wild-type and *widefins* medaka, showing the positions of the primers used for PCR-based genotyping. (c) Genotyping of wild-type and *widefins* medaka was performed using three primers (indicated on the right side of the panel) for each end of the *Teratorn* transposon insertion site. Genomic DNA was extracted from three individuals of both the wild-type and *widefins* medaka and used as templates for PCR. Following the reaction, PCR products were separated on a 2% agarose gel in 0.5x TBE buffer. In *widefins*, the PCR products could not be amplified using the c12a-3'UTR-s and c12a-3'UTR-as primers due to the *Teratorn* insertion.

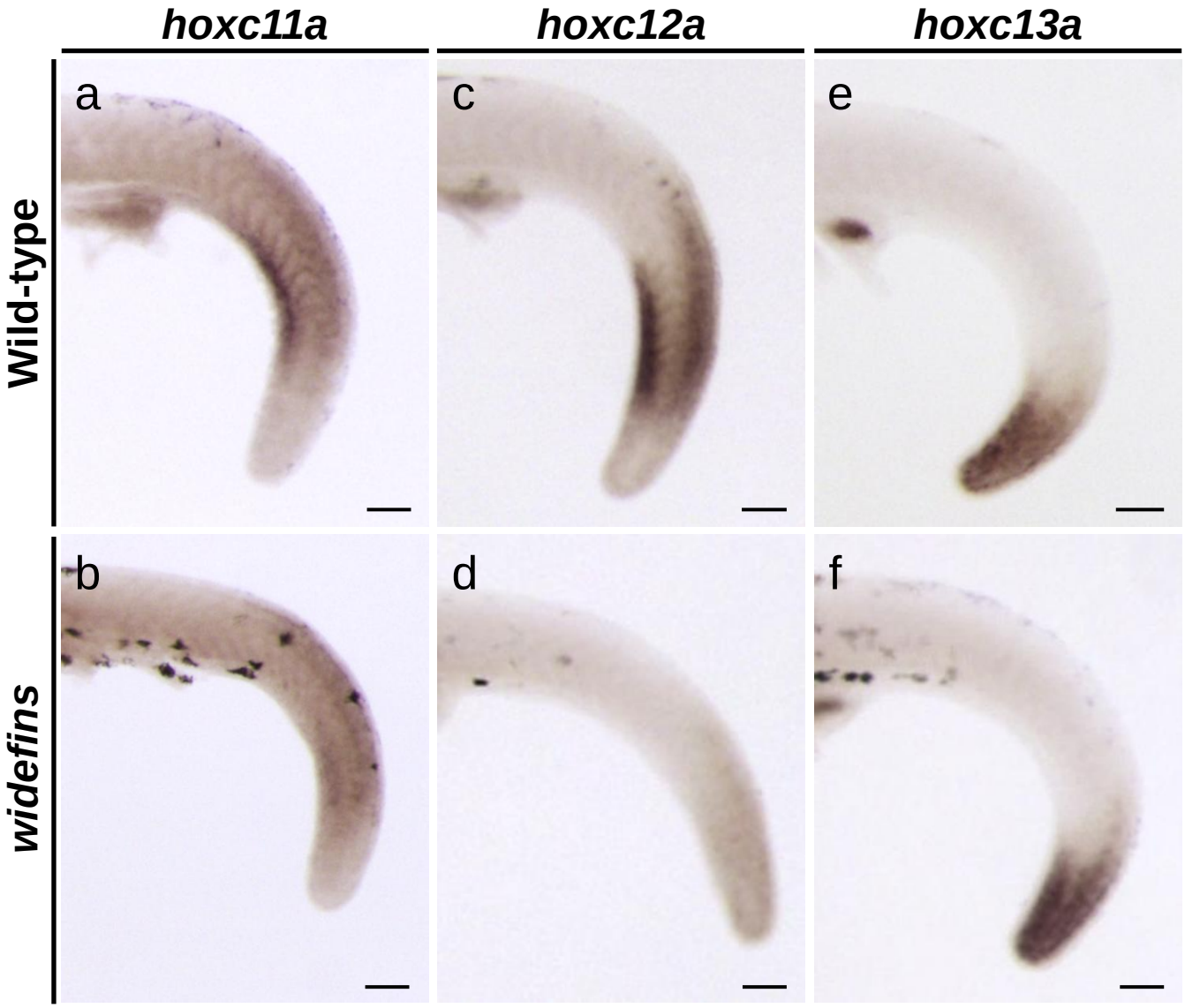

Figure 4. Significantly decreased expression of *hoxc12a* and reduced expression of *hoxc11a* in *widefins* medaka. (a) Schematic representation of medaka *hoxca* cluster. Black boxes represent the *hox* genes. Arrows indicate the transcriptional orientation. Yellow box indicates the insertion site of the *Teratorn* transposon. (b-g) Expression patterns of *hoxc11a*, *hoxc12a*, and *hoxc13a* were analyzed between wild-type and *widefins*. Whole-mount *in situ* hybridization was performed on embryos at stages 29-30. For each staining, at least three different specimens were used to confirm reproducibility. Lateral view of the posterior body. Scale bar: 100 μm.

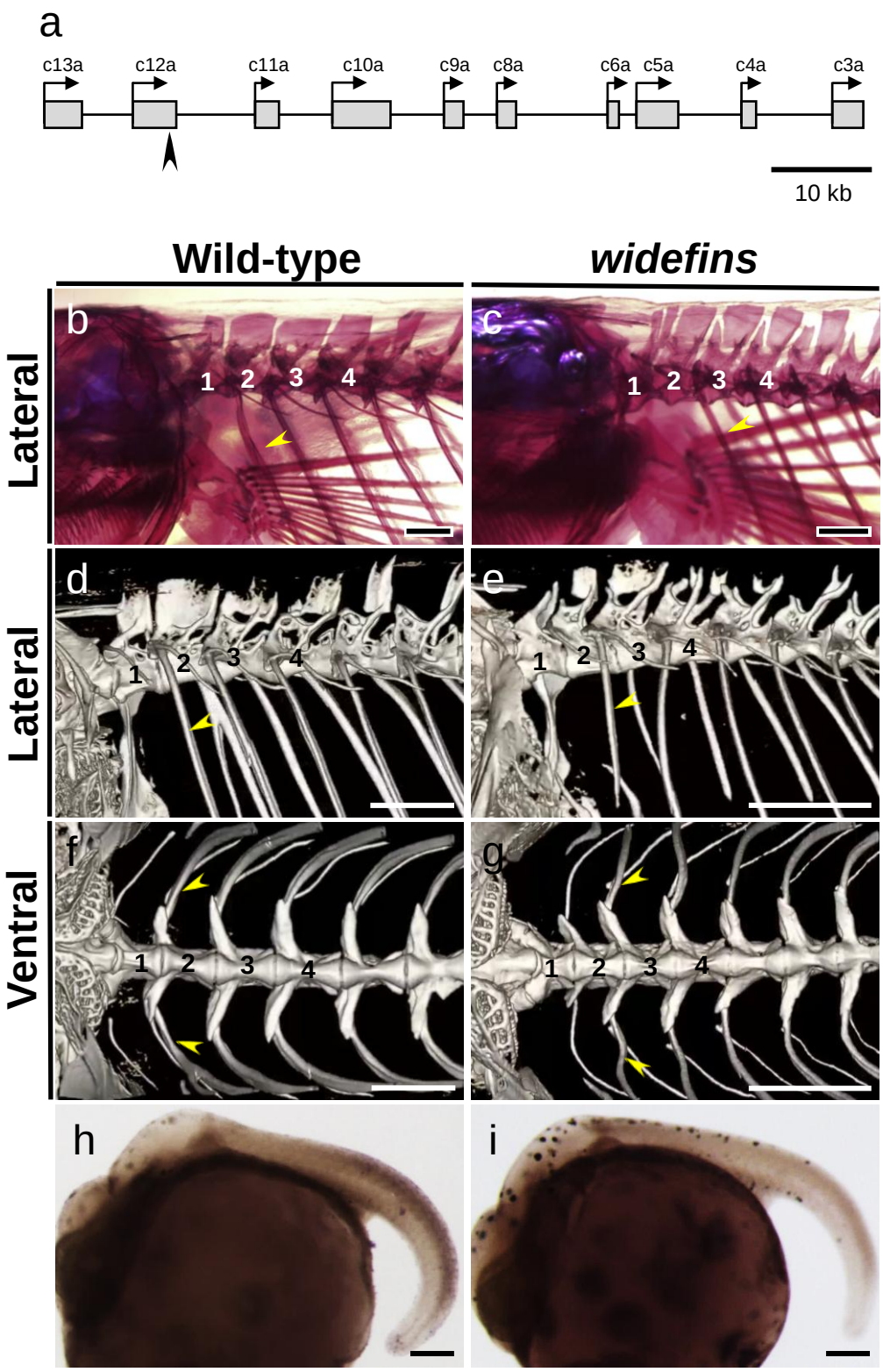

Figure 5. The expression of *hoxc6a* is decreased, resulting in anteriorizing vertebral patterning in *widefins* medaka. (a) Schematic representation of medaka *hoxca* cluster. Boxes represent the *hox* genes, and arrows indicate their transcriptional orientation. The arrowhead marks the insertion site of the *Teratorn* transposon in the 3'UTR of *hoxc12a*. (b, c) Skeletal structures of the anterior vertebrae in wild-type (*n*=6) and *widefins* (*n*=4). The ossified bones were stained with Alizarin Red S. Numbers indicate the number from the most anterior vertebra. Yellow arrowheads point to the anteriormost pleural rib in both wild-type and *widefins*. Scale bar: 2 mm. (d-g) Micro-CT scanning of the anterior vertebrae in wild-type and *widefins*. Scale bar: 2 mm. (h, i) Expression patterns of *hoxc6a* in wild-type and *widefins* embryos. Whole-mount *in situ* hybridization was performed on embryos at stages 29-30. At least three different specimens were used for each staining to confirm reproducibility. Lateral view of the posterior body. Scale bar: 100 μm.

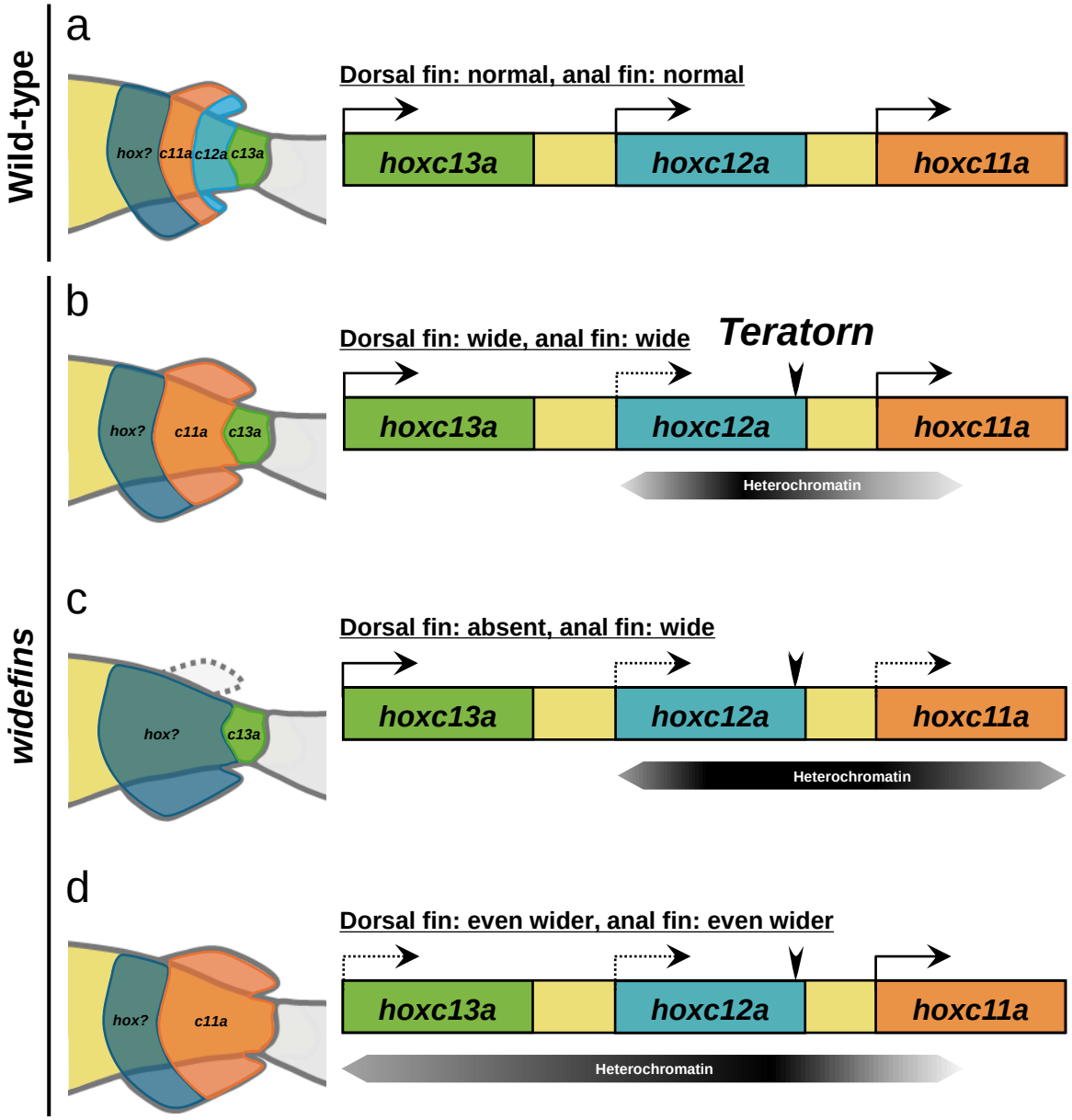

Figure 6. Proposed model illustrating phenotypic variation among *widefins* individuals. (a) In medaka, the genomic loci of *hoxc11a, hoxc12a, and hoxc13a* are aligned in the same transcriptional direction (arrow in right schematic). In the posterior body, these *hox* genes define regional identities competent for the formation of dorsal and anal fins (left schematic). (b) The insertion of a giant *Teratorn* transposon containing a viral genome within *hoxc12a* (arrowhead) leads to the formation of heterochromatin in the surrounding region, significantly reducing *hoxc12a* expression (dashed arrow). This reduction results in wide dorsal and anal fins, resembling the phenotype observed in *hoxc12a* mutants (Adachi et al., 2024). (c,d) In some instances, the insertion of the *Teratorn* transposon may expand the heterochromatin to a broader genomic region, affecting nearby *hox* genes. When the heterochromatin expands to *hoxc11a* locus, *hoxc11a* expression is considerably reduced, resulting in a wide anal fin due to *hoxc12a* deficiency, as well as the absence of a dorsal fin seen in *hoxc11a* mutants (Adachi et al., 2024). If the heterochromatin extends to *hoxc13a* locus, the expression of *hoxc13a* is also reduced, leading to fish with even wider dorsal and anal fins, as observed in *hoxc12a;hoxc13a* mutants (Adachi et al., 2024).