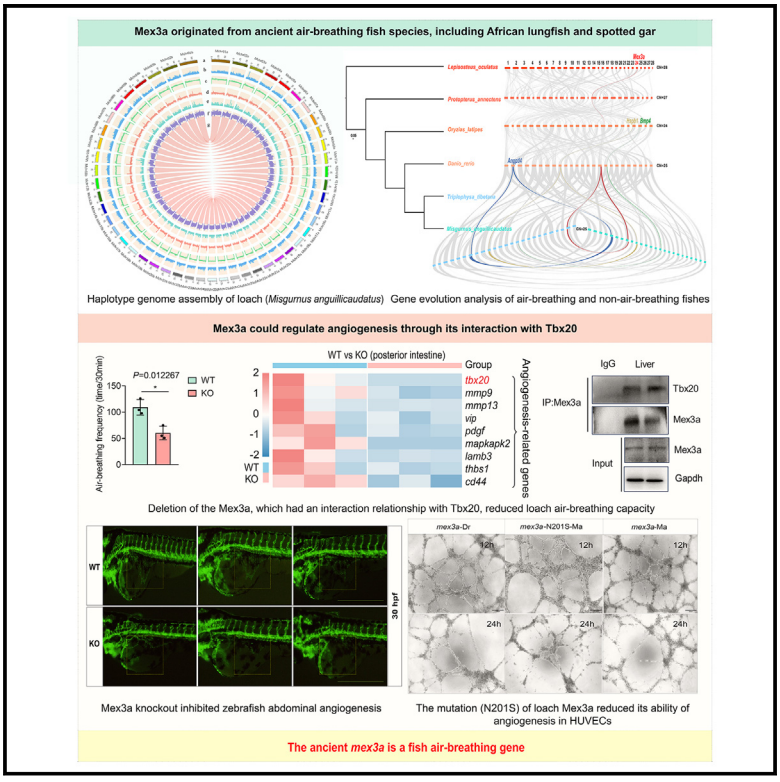
Cell Genomics

Article

The loach haplotype-resolved genome and the identification of Mex3a involved in fish air breathing

Graphical abstract



Highlights

- A high-quality haplotype genome of loach (an air-breathing fish) was assembled
- Mex3a originated from ancient air-breathing fish species
- Mex3a deletion impaired loach air-breathing capacity by inhibiting angiogenesis
- The ancient *mex3a* is a fish air-breathing gene

Authors

Bing Sun, Qingshan Li, Xinxin Xiao, ..., Yuwei Huang, Jian Gao, Xiaojuan Cao

Correspondence

gaojian@mail.hzau.edu.cn (J.G.), caoxiaojuan@mail.hzau.edu.cn (X.C.)

In brief

Sun et al. assembled a loach haplotype-resolved genome and identified a fish air-breathing gene, Mex3a, which could regulate angiogenesis through its interaction with Tbx20. Their findings hold significant implications for understanding the evolution of fish air breathing and provide a valuable resource for cultivating hypoxia-tolerant fish varieties.

Sun et al., 2024, Cell Genomics *4*, 100670
October 9, 2024
https://doi.org/10.1016/j.xgen.2024.100670

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

The loach haplotype-resolved genome and the identification of Mex3a involved in fish air breathing

Bing Sun,[1] Qingshan Li,[1] Xinxin Xiao,[1] Jianwei Zhang,[2] Ying Zhou,[1] Yuwei Huang,[1] Jian Gao,[1,*] and Xiaojuan Cao[1,3,*]
[1]College of Fisheries, Engineering Research Center of Green Development for Conventional Aquatic Biological Industry in the Yangtze River Economic Belt, Ministry of Education/Key Lab of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education/Key Lab of Freshwater Animal Breeding, Ministry of Agriculture and Rural Affairs, Huazhong Agricultural University, Wuhan 430070, China
[2]National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, China
[3]Lead contact
*Correspondence: gaojian@mail.hzau.edu.cn (J.G.), caoxiaojuan@mail.hzau.edu.cn (X.C.)
https://doi.org/10.1016/j.xgen.2024.100670

SUMMARY

Fish air breathing is crucial for the transition of vertebrates from water to land. So far, the genes involved in fish air breathing have not been well identified. Here, we performed gene enrichment analysis of positively selected genes (PSGs) in loach (*Misgurnus anguillicaudatus*, an air-breathing fish) in comparison to *Triplophysa tibetana* (a non-air-breathing fish), haplotype-resolved genome assembly of the loach, and gene evolutionary analysis of air-breathing and non-air-breathing fishes and found that the PSG *mex3a* originated from ancient air-breathing fish species. Deletion of Mex3a impaired loach air-breathing capacity by inhibiting angiogenesis through its interaction with T-box transcription factor 20. Mex3a overexpression significantly promoted angiogenesis. Structural analysis and point mutation revealed the critical role of the 201st amino acid in loach Mex3a for angiogenesis. Our findings innovatively indicate that the ancient *mex3a* is a fish air-breathing gene, which holds significance for understanding fish air breathing and provides a valuable resource for cultivating hypoxia-tolerant fish varieties.

INTRODUCTION

The evolution of life is a process filled with transformation and miracles. As the oldest vertebrates, fish have traversed a complex path of evolution, culminating in the emergence of humanity.[1,2] Among the myriad changes, the emergence of air breathing in fish stands as a pivotal adaptation, facilitating their transition from aquatic to terrestrial environments and liberating them from the confines of water. About 400 air-breathing fish species have been reported to date.[3–5] Histological observations of their air-breathing organs (ABOs) have revealed that high vascularization and a thin blood-gas barrier[6,7] make up the structural basis of fish air breathing, indicating a close linkage between angiogenesis and air-breathing ability.

A genomic study of African lungfish (a lobe-finned fish species with air-breathing ability) revealed that the ancient ancestors of bony fish already possessed a rudimentary ability for air breathing.[8] In-depth comparative genomic analysis of non-teleost ray-finned fishes, including bichir (*Polypterus senegalus*), paddlefish (*Polyodon spathula*), bowfin (*Amia calva*), and alligator gar (*Atractosteus spatula*), unveiled some ancient genes associated with lung development and angiogenesis that played pivotal roles in providing genetic and functional innovations for the evolution of air breathing in fish.[9] A genomic analysis of walking catfish (*Clarias batrachus*) with air-breathing ability revealed the expansion of certain genes related to angiogenesis and the binding and transportation of oxygen to meet the air-breathing demand of its terrestrial life.[10] In Indian catfish (*C. magur*), genes associated with vascular development and angiogenesis underwent positive selection to improve its air-breathing ability.[11] In addition, using transcriptome analyses of the gill and suprabranchial chamber (an ABO) of snakehead (*Channa argus*), researchers found that some genes (e.g., vascular endothelial growth factor [*vegf*]) related to angiogenesis were highly expressed in the ABO. Based on our prior transcriptome analyses of the posterior intestine (an ABO) in loach (*Misgurnus anguillicaudatus*), we also screened some angiogenesis-related genes.[12–14] Up to now, based on genomic analyses of single fish species with air-breathing ability, comparative genomic analyses of several non-teleost ray-finned fishes, and transcriptome studies of fish ABOs, some genes associated with angiogenesis involved in air breathing have been screened out. However, none of these genes have been confirmed as fish air-breathing genes.

CRISPR-Cas9, a gene editing technology, has rendered gene knockout (KO) more precise and efficient.[15–17] This technology allows researchers to explore the function of a gene within an organism by observing phenotypic changes. Compared to lungfish

and spotted gar, the loach is a better model for revealing the molecular mechanisms of fish air breathing due to the mature application of CRISPR-Cas9 editing technology on this species and its strong reproductive ability.[12,13,18] So, with this model, we can effectively uncover the intricacies of fish air-breathing formation, thus gaining deeper insights into the evolutionary journey of fish transitioning to land.

In our previous study, 120 positively selected genes (PSGs) were found in *M. anguillicaudatus* in comparison to *Triplophysa tibetana* (a non-air-breathing fish species).[19] Among these PSGs, there might be some associated with air breathing. In this study, combined with gene enrichment analysis for the PSGs, high-quality haplotype genomes of the loach, and gene evolutionary analyses of air-breathing and non-air-breathing fishes containing *Sarcopterygii* and *Actinopterygii* (teleost and non-teleost) species, the gene mex-3 RNA-binding family member a (*mex3a*) was screened out to be associated with intestinal air breathing in the loach. For functional analysis, deletion models of *mex3a* were constructed in both *M. anguillicaudatus* and vascular-visualization zebrafish (*Danio rerio*, a model fish without air breathing) by using CRISPR-Cas9 technology. Meanwhile, respiratory physiology experiments, transcriptomic analysis, protein interaction analysis, gene overexpression (OE), and knockdown (KD) in human umbilical vein endothelial cells (HUVECs) were performed. Furthermore, the functional site of Mex3a directly related to angiogenesis involvement in air breathing was determined through protein structural analysis, exogenous expression, and point mutations. Finally, the ancient *mex3a* was identified as a fish air-breathing gene. Our findings represent valuable assets in comprehending the evolution of fish air breathing, which aids in the adaptation of fish during the transition from water to land. Moreover, we offer a novel genetic resource for cultivating hypoxia-tolerant fish varieties.

RESULTS

Screening of genes associated with angiogenesis

Gene enrichment analysis for the PSGs

To further screen fish air-breathing-related genes, we conducted functional enrichment analysis for the 120 PSGs obtained from our previous study,[19] by using Gene Ontology (GO) and KEGG (Figure S1). Among the PSGs, we found that *mex3a* was annotated with blood vessel (BV) morphogenesis (GO: 0048514), heat shock protein beta-1 (*hspb1*) with the VEGF signaling pathway (ko04370), bone morphogenetic protein (*bmp4*) with the transforming growth factor beta (TGF-β) signaling pathway (ko04350), and angiopoietin-like 4 (*angptl4*) with a pathway involved in angiogenesis (GO: 0001525). Previous research indicated that MEX3A could promote angiogenesis in colorectal cancer via glycolysis.[20] Rezzola et al. reported that BMP4 could activate VEGFR2, leading to an angiogenic response.[21] Lee et al. found that soluble HSPB1 regulated VEGF-mediated angiogenesis through their direct interaction.[22] Chaube et al. reported that suppression of *angptl4* reprogramed endothelial cell metabolism and inhibited angiogenesis.[23] Kushwaha et al. believed that PSGs associated with vascular development and angiogenesis enhanced the air-breathing capacity of Indian catfish.[11] Hence, the positive selection of these four genes associated with angiogenesis may contribute to the loach intestinal air breathing (Table S1).

Loach haplotype genome assembly

A high-quality loach haplotype genome was assembled here for subsequent analysis. We assembled two haplotypes: the haplotype containing the longer-length homologous chromosomes was named "Mchra," while the other one was named "Mchrb." The final assembled Mchra and Mchrb genomes were approximately 1.04 and 1.06 Gb, respectively (Table S2). The scaffold N50 length of the loach surpassed those of *T. tibetana*,[24] *T. bleekeri*,[25] and *T. siluroides*[26] (Table S3). The Mchra and Mchrb genomes were evaluated with BUSCO, achieving completeness levels of 97% and 96.8%, respectively (Table S4). Additionally, as shown in Table S5, repetitive sequences accounted for approximately 59.87% in the Mchra genome and 59.54% in the Mchrb genome. Our data showed that the Mchra and Mchrb genomes were annotated with 24,400 and 24,544 protein-coding genes, respectively (Tables S6 and S7). To establish the chromosome orientation, we identified 126 and 128 bp centromere sequences in the loach and *T. tibetana* chromosomes, respectively, and then adjusted the chromosome to a consistent orientation (Figure S2A; Table S8). The similarity between the loach centromere sequences obtained here and the previously reported loach centromere sequences[27] was high at 80%, determined by using the NCBI alignment tool-blastn. It indicated the high accuracy of the loach centromere sequences in this study. According to the position of the centromere, 25 chromosomes of loach were divided into three types (metacentric chromosome [M; n = 5], submetacentric chromosome [SM; n = 2], and telocentric chromosome [T; n = 18]), which was consistent with the previous report by Li et al..[28] Furthermore, the Circos plot and genome collinearity revealed striking similarities in the characteristics of the two loach haplotype genomes (Figures 1 and S2B).

Genome synteny analysis and PSG location

To determine the positions of the four PSGs in the genomes of the loach and *T. tibetana*, the haplotype genome Mchra was used as the loach reference genome. Collinearity analysis indicated that there were some inversions between the genomes of the loach and *T. tibetana* (Figure S2C). Building upon this, through BLAST+ analysis, it was determined that the four PSGs (*mex3a*, *bmp4*, *hspb1*, and *angptl4*) were presented in both genomes of the loach and *T. tibetana*, albeit located on different chromosomes (Figure S2D). The analysis indicated that *mex3a* and *angptl4* were situated in the inversion region, while *hspb1* and *bmp4* were situated in the normal collinear region (Figure S2D).

Evolutionary analysis of mex3a, hspb1, bmp4, *and* angptl4

To identify genes involved in the intestinal air breathing of loach, we conducted evolutionary analyses on the genomes of ancient air-breathing fish, such as spotted gar (*Actinopterygii*) and African lungfish (*Sarcopterygii*), as well as non-air-breathing teleosts, like medaka, zebrafish, and *T. tibetana*, focusing on the genes *mex3a*, *bmp4*, *hspb1*, and *angptl4*. As shown in Figure 2, the gray lines represent the highly syntenic blocks, which were conserved genomic regions across these fish species. These blocks demonstrated a relatively stable gene arrangement and

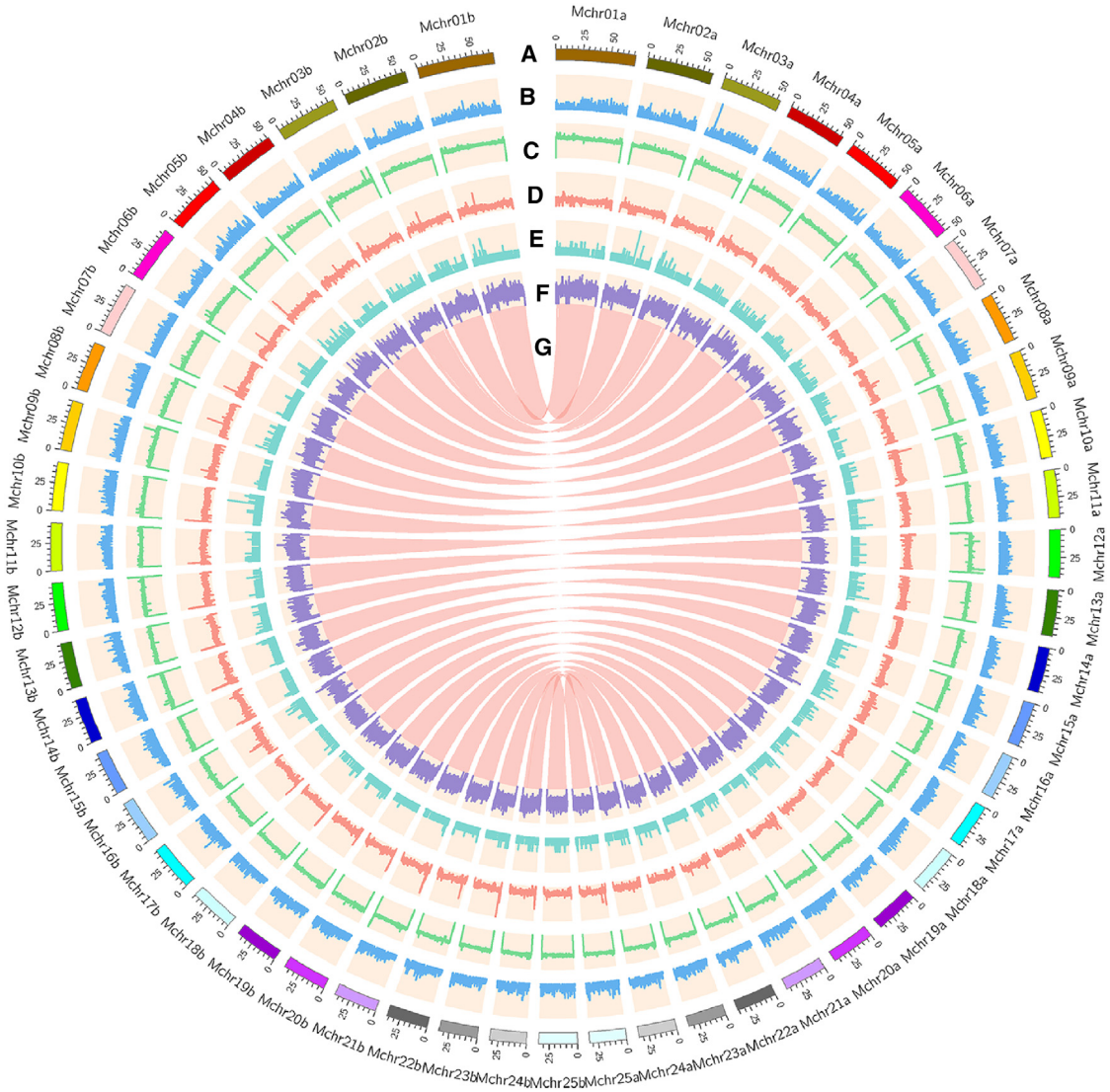


Figure 1. Genomic features of Mchra and Mchrb haplotype genomes

The Circos plot shows, from the outer to inner circle, (A) the chromosome number, (B) gene density, (C) GC content, (D) repetitive sequence density, (E) long terminal repeat (LTR) density, (F) allelic gene expression, and (G) haplotype genome collinearity.

order among these fish species, suggesting that these genes have been preserved throughout the evolutionary process. Evolutionary analysis of *hspb1* (yellow line), *bmp4* (green line), and *angptl4* (blue line) showed that these lines were not found in African lungfish and spotted gar, suggesting that these genes were not conserved and had not been observed from the ancient air-breathing fishes. For *mex3a* (red line), there were highly similar blocks among African lungfish, spotted gar, zebrafish, loach, and *T. tibetana*, indicating that *mex3a* was inherited from the ancient air-breathing fishes. The absence of the red line suggested that the *mex3a* gene orthologous region was not conserved between medaka and zebrafish (Figure 2). During the evolutionary process of fish, the ancestors of bony fish have possessed air-breathing ability.[8] Many ancient genes have provided crucial genetic and functional innovations for air breathing in fish, which have been inherited and elaborated upon by descendants throughout evolution.[8,9] Here, among the four PSGs of loach, only the *mex3a* gene was inherited from African lungfish and spotted gar. This suggested that the *mex3a* gene was ancient and likely played a role in fish air breathing.

Functional analysis of *mex3a*

Expression of Mex3a in fish ABOs

Mex3a is an RNA-binding protein possessing two functional domains: the RNA-binding KH domain and the RING-finger domain.[29] Recent studies have shown that MEX3A interacted with LAMA2 (laminin subunit alpha 2) and regulated the metastasis of lung adenocarcinoma cells through the PI3K/AKT signaling pathway.[30] In osteosarcoma, KD of MEX3A suppressed proliferation and migration.[31] In colorectal cancer,

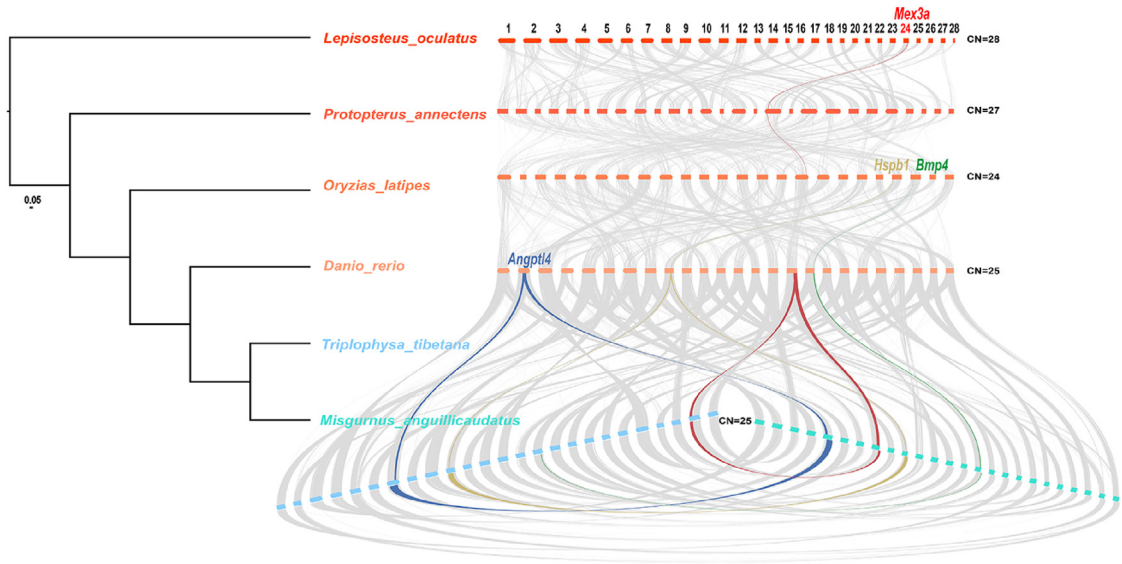


Figure 2. Gene evolution analysis of *mex3a*, *bmp4*, *hspb1*, and *angptl4*

Different colored lines represent collinearity of different genes, and CN denotes chromosome number. *mex3a*, mex-3 RNA binding family member a; *bmp4*, bone morphogenetic protein 4; *hspb1*, heat shock protein family b member 1; *angptl4*, angiopoietin-related 4.

MEX3A could promote angiogenesis via glycolysis.[20] So far, it has been demonstrated that Mex3a played a crucial role in regulating cell proliferation, invasion, and angiogenesis. However, whether Mex3a involved in fish air breathing remains unclear.

Three common air-breathing teleosts, namely *C. argus* (with an ABO: suprabranchial organ), *M. albus* (ABO: oropharynx), and loach (ABO: posterior intestine), were used to explore the expression patterns of *mex3a* in their ABOs (Figures S3A–S3D). When exposed to air, the expressions of Mex3a in the ABOs of these three fish species were significantly up-regulated, indicating that Mex3a was involved in the process of air-breathing fish adapting to low-oxygen environments. Tissue expression and localization analyses of loach *mex3a* were performed here. They showed that *mex3a* exhibited the highest expression level in the posterior intestine of the loach (Figure S3E). Early developmental signals were also predominantly concentrated in the intestinal region of the loach (Figure S3F). *In situ* hybridization and immunofluorescence further elucidated that Mex3a was strongly expressed in the capillaries of the mucosal layer and also expressed in the serosal layer in the posterior intestine of the loach, with minimal expression in its anterior intestine (Figures S3G and S3H). Similarly, previous genomic and transcriptomic studies on air-breathing fish, like African lungfish, walking catfish, *C. argus*, and Indian catfish, have indicated that certain genes involved in angiogenesis, which have undergone positive selection or are highly expressed in ABOs, may participate in the process of air breathing.[8–11,23] Therefore, our findings suggested that the ancient PSG *mex3a* was related to angiogenesis, involving intestinal air breathing in loach.

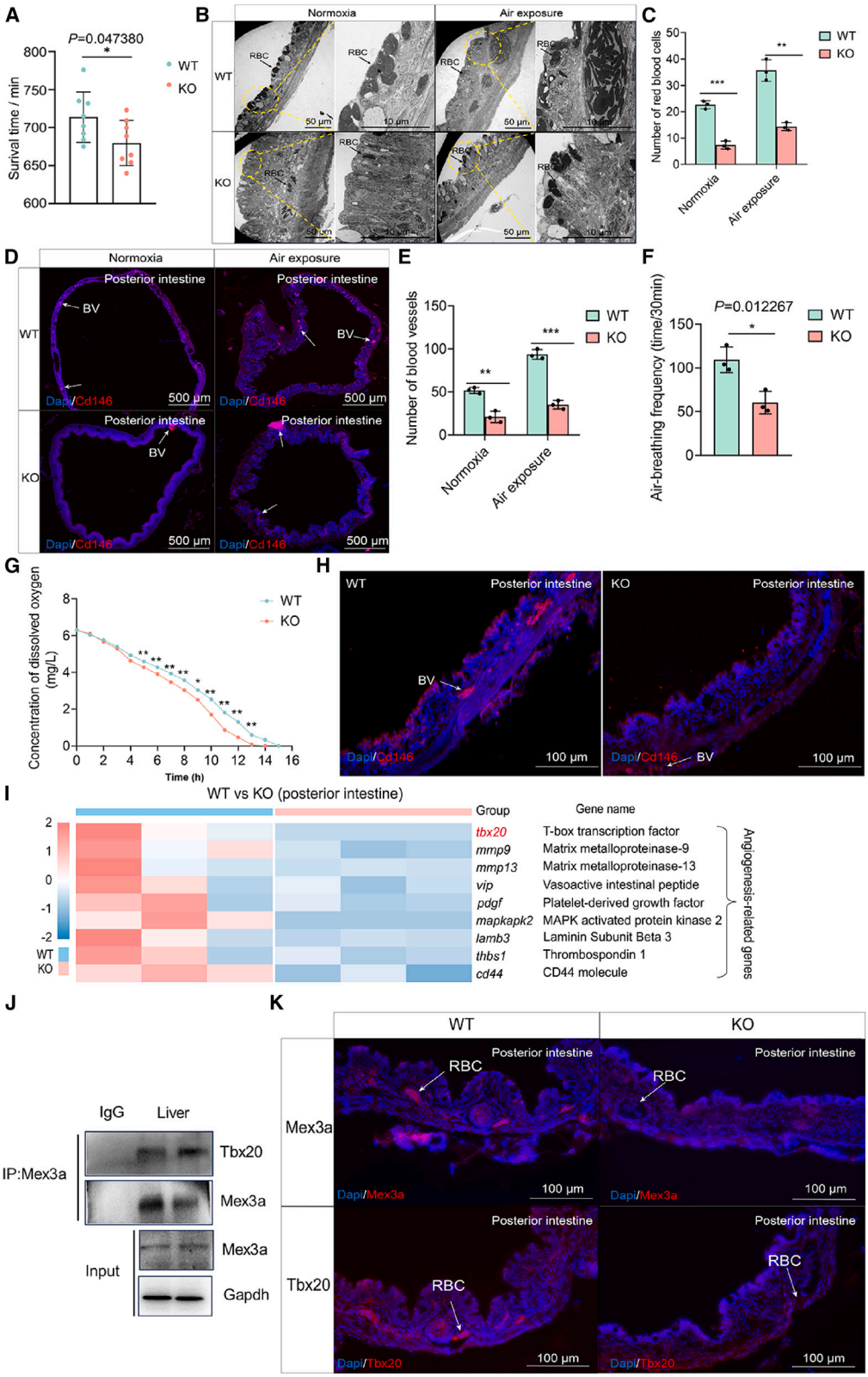
Construction of mex3a *KO model in loach*

By using the CRISPR-Cas9 system, we established a *mex3a* gene KO model in loach. As depicted in Figure S3I, the wild-type loach (WT loach) *mex3a* gene comprised two exons and one intron. We targeted the first exon for KO. The resulting mutant, *mex3a*$^{-/-}$ (KO loach), lacked 13 base pairs (GCGAGG ACAGCGG), causing premature termination of amino acid translation (with translation of 59 amino acids and mistranslation of 57 amino acids). Validation through real-time qPCR and western blot (WB) confirmed the successful generation of homozygous loach lacking Mex3a.

mex3a KO reduced the intestinal air-breathing capacity of loach

To investigate the impact of *mex3a* gene deletion on loach intestinal air breathing, we conducted an air exposure experiment and recorded the survival times of WT and KO loach. Results revealed a significant decrease in the survival time of KO loach compared with WT loach (Figure 3A). Histological analysis of the posterior intestines showed no significant structural changes between KO and WT loach (Figure S3J). However, the number of blood vessels in the posterior intestine of KO loach was slightly lower than that of WT loach (Figure S3K). Additionally, transmission electron microscopy (TEM) analysis of the posterior intestine indicated that the number of red blood cells (RBCs) in the mucosal layer of KO loach was significantly smaller than that of WT loach (Figures 3B and 3C).

To better visualize the changes in the blood vessels of the posterior intestines, we utilized the melanoma cell adhesion molecule (MCAM; Cd146/MCAM) to label the blood vessels.[32,33] We found that the number of blood vessels in the posterior intestine of KO loach was notably smaller compared with that of WT loach (Figures 3D and 3E). Furthermore, the expression levels of the genes involved in angiogenesis in KO loach, such as *vegf*, erythropoietin (*epo*), hemoglobin β chain (*hbb*), and enolase 3 (*eno3*), were significantly reduced in comparison to WT loach (Figures S3L and S3M). These results from the air exposure experiment suggested that the deletion of the *mex3a* gene



(legend on next page)

suppressed the formation of blood vessels in the posterior intestine of the loach, consequently diminishing the survival time of KO loach exposed to air.

Through further chronic hypoxia experiments, we once again investigated the effect of Mex3a deletion on intestinal air breathing of loach. Under chronic hypoxia (Figure S3N), the air-breathing frequency of KO loach was significantly lower than that of WT loach (Figure 3F), and the oxygen consumption in the water was markedly higher than that of WT loach (Figure 3G). These results indicated that KO loach oxygen uptake was dependent on dissolved oxygen in water, achieved through gill respiration. Results from Cd146 labeling of blood vessels in the posterior intestine also indicated a significant reduction in the number of blood vessels in KO loach compared with WT loach (Figure 3H). Additionally, genes related to angiogenesis, such as *vegf*, *epo*, *hbb*, and *eno3*, were found to be significantly down-regulated in KO loach compared with WT loach (Figure S3O). These results once again highlighted the crucial role of the *mex3a* gene in regulating angiogenesis in the posterior intestine of the loach, and its deletion led to a decreased air-breathing capacity in the loach.

Transcriptome analysis of posterior intestines from WT and KO loach

In organisms, oxygen served as a chemoattractant for vascular generation and was directly linked to vascular growth.[34] Studies have demonstrated that moderate hypoxia can stimulate angiogenesis, or the formation of new blood vessels.[35] To explore the regulatory mechanism of *mex3a* on vascular development, we conducted transcriptome sequencing of posterior intestines of the loach (WT vs. KO) under air exposure. The transcriptome analysis (KO vs. WT) revealed 920 differentially expressed genes (DEGs), comprising 390 up-regulated genes and 530 down-regulated genes (Figure S4A).

Further GO enrichment analysis indicated that some DEGs exhibited enrichment in the "morphogenesis of an epithelial fold" pathway, which played a crucial role in embryonic development and tissue formation, facilitating morphological changes and structural establishment essential for the development and function of specific organs or structures[36–38] (Figure S4B). KEGG enrichment analysis identified certain DEGs enriched in the "ECM-receptor interaction" pathway, regulating critical biological processes like cell adhesion, signal transduction, movement, and tissue establishment[39,40] (Figure S4C). Further analysis of DEGs revealed the significant down-regulation of known genes associated with angiogenesis, such as T-box transcription factor 20 (*tbx20*), matrix metalloproteinase-9 (*mmp9*), *mmp13*, vasoactive intestinal peptide (*vip*), platelet-derived growth factor (*pdgf*), MAPK activated protein kinase 2 (*mapkapk2*), laminin subunit beta 3 (*lamb3*), thrombospondin 1 (*thbs1*), and CD44 molecule (*cd44*), in KO loach (Figure 3I). Additionally, qPCR analysis of selected DEGs confirmed that the gene expression trends in the posterior intestine of WT and KO loach aligned with the transcriptome data, validating its reliability (Figure S4D). To delve deeper into the molecular mechanism by which *mex3a* regulated angiogenesis, we performed immunoprecipitation (IP) analysis. Based on the antibody quality and sequence homology, we analyzed interactions between Mex3a and Tbx20, Mmp9, Mmp13, and Mapkapk2 proteins, respectively (Figures S4E and S4F). Ultimately, IP analysis revealed an interaction between Mex3a and Tbx20 (Figure 3J).

Studies have reported that Tbx20 played a critical role in the regulation of various structures, including the heart, neural cells, and blood vessels, during embryonic development.[41–43] Depletion of Tbx20 in mouse embryos may lead to developmental abnormalities in the cardiovascular system, even resulting in mortality.[44] Additionally, Ji et al. performed single-cell sequencing and discovered that perivascular adipose stem cells had the potential to promote vascular remodeling, with Tbx20 OE significantly enhancing the differentiation of these cells.[45] Furthermore, we validated the expression levels of the Tbx20-Prok2-Prokr1 signaling pathway in the posterior intestine of loach. As depicted in Figure S4G, the deletion of Mex3a in loach significantly suppressed the expressions of Tbx20, Prok2, and Prokr1 genes. Immunofluorescence staining of the posterior intestine demonstrated that both Mex3a and Tbx20 were expressed in blood vessels of the loach. Mex3a deletion led to a significant decrease in Tbx20 expression of the loach, with only minimal expression signals detected in the blood vessels (Figure 3K). Meng et al. also found that inhibiting the expressions of the genes from the Tbx20-Prok2-Prokr1 pathway in zebrafish and mice caused vascular damage, while OE of the genes could rescue vascular injury.[46] Our findings suggested that Mex3a

Figure 3. Survival times, posterior intestine structural changes, and oxygen consumption of WT and KO loach

(A) Survival time under air exposure (each group with eight loaches). $^{*}p < 0.05$. Statistical significance is determined with Student's t test.

(B) Transmission electron microscopy (TEM) analysis of the posterior intestine under normoxia and air exposure (each group with three replicates). The golden circles indicate the magnified areas; RBC and black arrows represent red blood cells.

(C) Statistics of red blood cells in the posterior intestine under TEM analysis. $^{**}p < 0.01$ and $^{***}p < 0.001$. Statistical significance is determined with Student's t test.

(D) Cd146 marks the new blood vessels. White arrows indicate blood vessels. CD146, melanoma cell adhesion molecule.

(E) Statistics of blood vessels in the posterior intestine under immunofluorescence analysis (each group with three replicates). $^{*}p < 0.05$. Statistical significance is determined with Student's t test.

(F) Air-breathing frequency (each group with three replicates). Statistical significance is determined with Student's t test.

(G) Oxygen consumption detection. $^{*}p < 0.05$ and $^{**}p < 0.01$. Statistical significance is determined with Student's t test.

(H) Cd146 marks the blood vessels. White arrows indicate blood vessels (red blood cells).

(I) Transcriptome analysis of gene expression related to angiogenesis from WT and KO loach. *tbx20*, T-box transcription factor; *mmp9*, matrix metalloproteinase-9; *mmp13*, matrix metalloproteinase-13; *vip*, vasoactive intestinal peptide; *pdgf*, platelet-derived growth factor; *mapkapk2*, MAPK activated protein kinase 2; *lamb3*, laminin subunit beta 3; *thbs1*, thrombospondin 1; *cd44*, CD44 molecule.

(J) Immunoprecipitation (IP) analysis.

(K) Immunofluorescence analysis of Mex3a and Tbx20 proteins in the posterior intestine of loach. White arrows indicated blood vessels. RBC, red blood cells; BV, blood vessel.

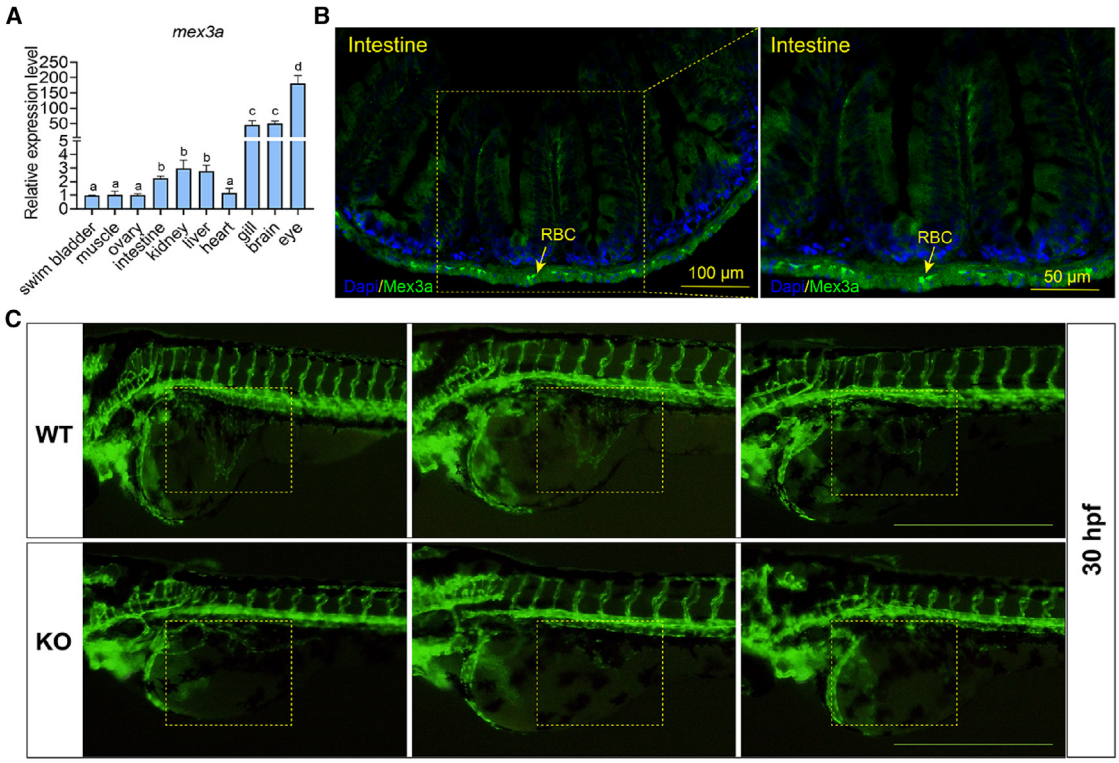


Figure 4. Tissue expression analysis of *mex3a* and observation of early developmental blood vessels in zebrafish

(A) *mex3a* tissue expression (three zebrafishes are used here). Different letters above the error bars indicate a significant difference among different groups ($p < 0.05$). "a" represents the lowest expression level, "b" represents a significantly higher expression level compared with a, "c" represents a significantly higher expression level compared with b, and "d" represents a significantly higher expression level compared with c. Statistical significance is determined with one-way ANOVA.

(B) Mex3a localization in the intestine. The golden arrows indicate expression signals in the intestine, RBC, red blood cells.

(C) Observation of early developmental blood vessels in WT and KO zebrafish (each group with three replicates). The golden frame represents the abdominal blood vessels of zebrafish. 30 hpf, 30 h post-fertilization.

interacted with Tbx20 to regulate angiogenesis in fish via the Tbx20-Prok2-Prokr1 pathway.

mex3a involved in the angiogenesis of fish

The transgenic zebrafish line (*fli1a*:EGFP) has become an indispensable experimental tool for studying angiogenesis.[47,48] Zebrafish *fli1* expression is highly ubiquitous, and its promoter can drive the expression of enhanced green fluorescent protein (EGFP) in all blood vessels throughout embryogenesis.[49,50] To validate our findings of the loach, we established a zebrafish model with a *mex3a* gene KO using the *fli1a*:EGFP line. As shown in Figure S5A, zebrafish (*fli1a*:EGFP) had a *mex3a* gene consisting of two exons and one intron. We targeted the first exon for KO, resulting in a mutant $mex3a^{-/-}$ (KO zebrafish) with a deletion of four base pairs (GGCG), leading to premature termination of amino acid translation (with translation of 54 amino acids and mistranslation of 449 amino acids). WB validation confirmed the successful generation of homozygous zebrafish lacking *mex3a* (Figure S5B).

Furthermore, tissue expression analysis revealed that the zebrafish (AB strain) *mex3a* gene was primarily expressed in tissues rich in blood vessels, including the eyes, gills, and brain, with relatively low expression levels in the intestine (Figure 4A). Meanwhile, the expression pattern of the *mex3a* gene differed between zebrafish and the loach. Immunofluorescence analysis of zebrafish (AB strain) intestine tissues revealed that Mex3a was highly expressed in the serosal layer and also expressed in the mucosal layer (Figure 4B), different from the results in the loach (Figure S3H). Through vascular visualization analysis, we observed significant inhibition of abdominal blood vessel growth in *mex3a* deletion zebrafish (*fli1a*:EGFP) at 30 h post-fertilization (hpf) compared with WT zebrafish (*fli1a*:EGFP) (Figure 4C). Furthermore, the deletion of Mex3a significantly inhibited the expressions of Tbx20, Prok2, and Prokr1 genes (Figure S5C). These findings served to further solidify the pivotal role of the ancient gene *mex3a* in the process of angiogenesis.

Validation of Mex3a *function on angiogenesis in HUVECs*

Angiogenesis is a complex process involving the proliferation and migration of endothelial cells, regulated by angiogenic factors and the extracellular matrix.[51–53] Endothelial cell migration is one of the early steps in angiogenesis and serves as a hallmark of vascular development.[54] To mimic the process of vascular development *in vitro*, endothelial cell migration and a three-dimensional model of the vascular morphogenesis assay are commonly employed

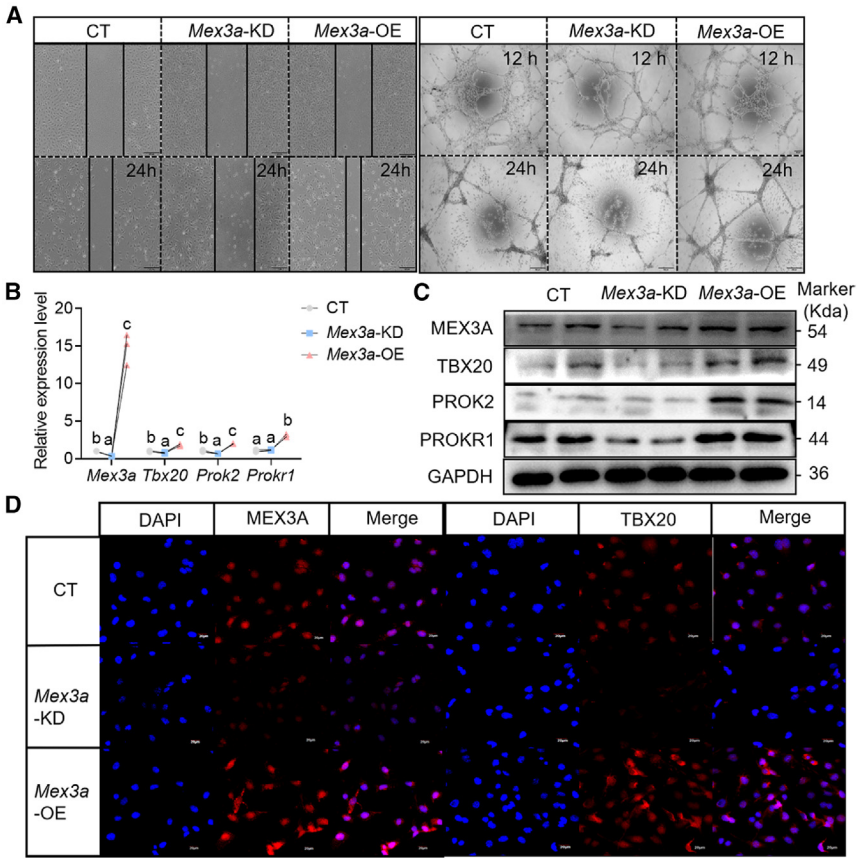


Figure 5. The effects of *Mex3a* overexpression and knockdown on the migration and tube-forming ability of HUVECs

(A) HUVEC migration (left) and HUVEC tube formation (right) (each group with three replicates).

(B) qPCR (each group with three replicates). Different letters above the error bars indicate a significant difference among different groups ($p < 0.05$). "a" represents the lowest expression level, "b" represents a significantly higher expression level compared with a, and "c" represents a significantly higher expression level compared with b. Statistical significance is determined with one-way ANOVA.

(C) Western blot.

(D) Immunofluorescence.

to assess angiogenesis. Endothelial cells such as HUVECs are frequently used in studies related to hypoxia, cell migration, and tumor-associated angiogenesis applications.[55] To verify the regulatory role of *Mex3a* in angiogenesis, we utilized HUVECs to establish models of *Mex3a* KD and OE. Figure S6A shows the successful transfection of *Mex3a* into HUVECs.

Through cell migration experiments under normoxia, we observed that knockdown of *Mex3a* (*Mex3a*-KD group) inhibited HUVEC migration compared with the control group (CT group), while OE of *Mex3a* (*Mex3a*-OE group) promoted HUVEC migration (Figure 5A). Additionally, a tube formation assay on HUVECs showed no significant difference in tube-forming ability among the *Mex3a*-KD, *Mex3a*-OE, and CT groups at 12 h incubation. However, at 24 h, the tube-forming ability of the *Mex3a*-OE group was significantly stronger than that of the CT and *Mex3a*-KD groups. In the *Mex3a*-OE group, most blood vessels remained intact, with a few dead cells, while in the CT and *Mex3a*-KD groups, the majority of blood vessels could not maintain integrity, with a large number of dead cells (Figure 5A). Through qPCR, WB, and immunofluorescence, we validated the TBX20-PROK2-PROKR1 pathway in HUVECs. It showed that *Mex3a* KD significantly inhibited the expression of the TBX20-PROK2-PROKR1 pathway, while *Mex3a* OE significantly promoted the expression of the TBX20-PROK2-PROKR1 pathway (Figures 5B and 5C). Immunofluorescence results demonstrated that the expression of TBX20 was obviously suppressed in the *Mex3a*-KD group and notably enhanced in the *Mex3a*-OE group (Figure 5D).

As shown in Figures S6B–S6E, low oxygen treatment inhibited the proliferation of HUVECs. However, results from the three HUVEC groups (*Mex3a*-KD, *Mex3a*-OE, and CT groups) under hypoxia were consistent with those under normoxia, indicating that *Mex3a* KD inhibited HUVEC migration, while *Mex3a* OE promoted HUVEC migration. Tube formation results of HUVECs showed that at 12 h incubation, there was no obvious difference in tube-forming ability among the three groups. However, at 24 h, most blood vessels in the *Mex3a*-OE group remained intact, while in the CT and *Mex3a*-KD groups, most blood vessels could not maintain integrity, and cells gradually underwent death (Figure S6C). These results suggested that the tube-forming ability of the *Mex3a*-OE group was stronger than that of the CT and *Mex3a*-KD groups. Additionally, *Mex3a* knockdown significantly inhibited the expression of the TBX20-PROK2-PROKR1 pathway, while *Mex3a* OE significantly promoted the expression of the TBX20-PROK2-PROKR1 pathway (Figures S6D and S6E). Our findings indicated that, similar to fish, mammalian Mex3a also possessed the ability to regulate angiogenesis through the TBX20-PROK2-PROKR1 signaling pathway.

The effect of loach Mex3a mutation N201S on the migration and tube-forming ability of HUVECs

Conservation analysis of Mex3a

The protein structure determines its function.[56,57] By comparing the amino acid sequences of Mex3a in the loach, African lungfish, *Erpetoichthys calabaricus*, *Mastacembelus armatus*, *C. argus*, *Kryptolebias marmoratus*, *Ctenopharyngodon idellus*, *Labeo rohita*, zebrafish, *T. tibetana*, *Latimeria chalumnae*, and *L. oculatus*, it was found that Mex3a exhibited a high degree of conservation (Figure S6F), suggesting that Mex3a may play a role in regulating angiogenesis in fish species. We wondered why zebrafish and *T. tibetana* could not survive in air. With

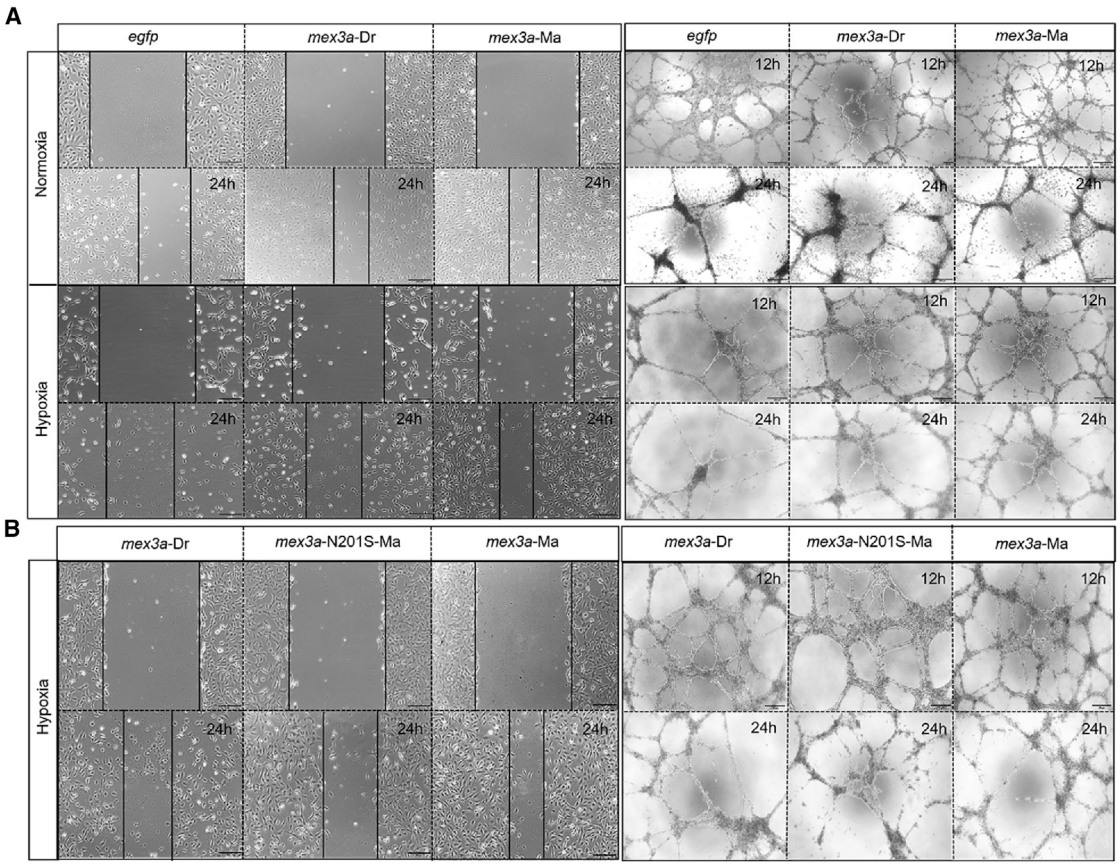


Figure 6. HUVEC migration and tube formation ability

(A) HUVEC migration and tube formation ability under normoxia (top) and hypoxia (bottom) (each group with three replicates). *egfp* represents the control group, *mex3a*-Dr represents the zebrafish *mex3a* transfection group, and *mex3a*-Ma represents the loach *mex3a* transfection group.

(B) HUVEC migration and tube formation ability under hypoxia (each group with three replicates). *mex3a*-Dr represents the zebrafish *mex3a* transfection group, *mex3a*-N201S-Ma represents the loach *mex3a* mutation transfection group, and *mex3a*-Ma represents the loach *mex3a* transfection group.

positive selection analysis, it was found that the 201st amino acid of loach Mex3a (the position after the sequence alignment was 188) underwent positive selection and mutation, evolving into asparagine (N), which was different from serine (S) in zebrafish and *T. tibetana* (Figure S6G; Table S9). The primary structure of a protein determines its spatial structure. By predicting the tertiary structure of the Mex3a protein, it was found that its spatial conformations varied among species, including African lungfish, spotted gar, *M. albus*, *C. argus*, the loach, *T. tibetana*, zebrafish, mouse, and human, especially in the RING-finger domain (Figures S6H–S6P). Previous research indicated that the RING-finger domain was crucial in regulating processes such as the cell cycle, migration, and gene transcription, and its defects were often associated with the development of cancer.[58,59] In Cypriniformes, the protein spatial conformations between *T. tibetana* (*Tt*) and zebrafish (*Dr*) are similar, with differences observed only in the RING-finger domain compared to loach (*Ma*), as shown in Figure S6Q. Furthermore, after mutating the 201st amino acid of loach Mex3a from asparagine (201N) to serine (*Ma_N201S*), the conformation of the Mex3a protein in the RING-finger domain became similar to those of *T. tibetana* and zebrafish (Figure S6R). Further analysis of the surrounding amino acid conformations of loach Mex3a 201N, *T. tibetana* 201S, zebrafish 197S, and loach Mex3a N201S showed that the conformation of zebrafish 197S differed from that of loach 201N, *T. tibetana* 201S, and loach Mex3a N201S (Figures S6S–S6U); The difference in conformation between loach 201N and *T. tibetana* 201S was associated with the different number of hydrogen bonds formed by surrounding amino acids (Figures S6S and S6T); the conformation of the N201S was closer to that of *T. tibetana* 201S (Figures S6T and S6V). Our findings suggested that the change of the 201st amino acid (N or S) in the loach Mex3a protein could affect its spatial structure, potentially resulting in functional alterations.

Effect of exogenous loach and zebrafish mex3a *gene on the migration and tube formation of HUVECs*

Through cell migration experiments, it was found that compared to the CT group (*egfp*), both loach (*mex3a*-Ma group) and zebrafish (*mex3a*-Dr group) *mex3a* could promote the migration and tube-forming ability of HUVECs under normoxia. However, there was no significant difference between the loach and zebrafish (Figure 6A, top). Hypoxia is one of the main factors inducing

angiogenesis.[36] Furthermore, through hypoxia induction, in terms of cell migration in HUVECs, we found that compared to the *mex3a*-Dr and *egfp* groups, the cell density in the *mex3a*-Ma group was higher, and it had a longer cell migration distance. The integrity and density of tube formation in the *mex3a*-Ma group were also higher than those of the *mex3a*-Dr and *egfp* groups (Figure 6A, bottom). These findings suggest that both loach and zebrafish *mex3a* genes can enhance HUVEC migration and tube formation under normoxia and hypoxia. Furthermore, the loach *mex3a* appeared to have a stronger pronounced promoting effect under hypoxia compared to zebrafish. The positive selection of ancient *mex3a* represented a strategic adaptation to terrestrial environments by enhancing loach air-breathing ability.

Under hypoxia, loach Mex3a demonstrated superior promotion of HUVEC migration and tube formation. This suggested that the 201N position of Mex3a was the primary distinguishing factor in terms of HUVEC migration and tube formation between the loach and zebrafish. Furthermore, we established a mutant loach Mex3a (N201S) model in HUVECs, exhibiting a similar trend to zebrafish Mex3a in promoting HUVEC migration and tube formation (Figure 6B). These findings suggested that the amino acid residue at position 201 (201N) in the Mex3a protein was a key distinguishing factor in promoting angiogenesis between the loach and zebrafish.

Conclusions

In this study, we assembled two high-quality haplotypes of the loach *M. anguillicaudatus* and then conducted gene evolutionary analysis of air-breathing and non-air-breathing fishes containing *Actinopterygii* and *Sarcopterygii* (teleost and non-teleost) species, through which *mex3a* associated with air breathing was screened out. Subsequently, through gene KO, KD, and OE, respiratory physiology experiments, transcriptomic analysis, protein interaction/structure analysis, etc., we found that Mex3a deletion inhibited angiogenesis, impairing loach air-breathing capacity. Furthermore, the 201st amino acid in loach Mex3a directly related to its role in angiogenesis. Finally, the ancient *mex3a* was innovatively identified as a fish air-breathing gene. Our findings provide new insights into the evolution of fish air breathing and offer an important theoretical basis for developing new fish varieties of hypoxia tolerance. Additionally, it has been reported that humans evolved from fish, and therefore the ancient *mex3a* gene, an air-breathing gene, may hold potential for the treatment of human hypoxia-related diseases.

Limitations of the study

This study aims to elucidate the function and regulatory mechanisms of the ancient air-breathing gene *mex3a*, with the goal of using it as a candidate gene for breeding hypoxia-tolerant fish species or assisting in solving human hypoxia-related diseases. We have created a *mex3a* KO model in loach and found that *mex3a* indeed regulated angiogenesis. Ideally, constructing a *mex3a* OE model in loach would provide additional insights. However, this process was challenging and time consuming and had a very low success rate. Despite these limitations, our *mex3a* OE model in HUVECs has also confirmed the gene's role in promoting angiogenesis. Our findings will aid future studies in understanding fish air breathing and provide a valuable resource for cultivating hypoxia-tolerant fish varieties.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xiaojuan Cao (caoxiaojuan@mail.hzau.edu.cn).

Materials availability

The materials that support the findings of this study are available from the corresponding authors upon reasonable request. Please contact the lead contact, Xiaojuan Cao (caoxiaojuan@mail.hzau.edu.cn), for additional information.

Data and code availability

The assembled haplotype genomes reported in this paper have been deposited in the Genome Warehouse in National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation, under accession numbers GWHERQE00000000 (Mchra) and GWHERQD00000000 (Mchrb), publicly accessible at https://ngdc.cncb.ac.cn/gwh. WT and KO loach intestine tissue RNA sequencing (RNA-seq) data were deposited in the NCBI Sequence Read Archive (SRA) under accession number PRJNA1091403.

ACKNOWLEDGMENTS

The authors would like to thank Wuhan Onemore-tech Co., Ltd., for their assistance with genome sequencing and analysis. The authors would like to thank the National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, for providing the computing platform. Most bioinformatic analyses and computations in this paper were run on the platform. We gratefully acknowledge Limin He (Huazhong Agricultural University) for her help with the TEM. This study was supported by the National Natural Science Foundation of China (U23A20245 and 32172962) and the Key R&D Program Project in Hubei Province of China (2023BBB054).

AUTHOR CONTRIBUTIONS

X.C. and J.G. designed the research; B.S., X.X., and Y.Z. performed the research; B.S., Q.L., and J.Z. analyzed the data; B.S. wrote the paper; X.C. and J.G. obtained the research funding; X.C. and J.G. were involved in the discussion; and X.C. and J.G. revised the paper. All authors have reviewed and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animals
  - Animal models
  - HUVEC cells
- METHOD DETAILS
  - Screening of genes associated with angiogenesis
  - Functional analysis of Mex3a
  - The effect of loach Mex3a mutation N201S on the migration and tube-forming ability of HUVECs
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2024.100670.

Received: May 29, 2024
Revised: July 30, 2024
Accepted: September 13, 2024
Published: October 9, 2024

REFERENCES

1. Andreev, P.S., Sansom, I.J., Li, Q., Zhao, W., Wang, J., Wang, C.C., Peng, L., Jia, L., Qiao, T., and Zhu, M. (2022). The oldest gnathostome teeth. Nature *609*, 964–968. https://doi.org/10.1038/s41586-022-05166-2.
2. Andreev, P.S., Sansom, I.J., Li, Q., Zhao, W., Wang, J., Wang, C.C., Peng, L., Jia, L., Qiao, T., and Zhu, M. (2022). Spiny chondrichthyan from the lower Silurian of South China. Nature *609*, 969–974. https://doi.org/10.1038/s41586-022-05233-8.
3. Wright, P.A., and Turko, A.J. (2016). Amphibious fishes: evolution and phenotypic plasticity. J. Exp. Biol. *219*, 2245–2259. https://doi.org/10.1242/jeb.126649.
4. Martin, K.L. (2014). Theme and variations: amphibious air-breathing intertidal fishes. J. Fish. Biol. *84*, 577–602. https://doi.org/10.1111/jfb.12270.
5. Bressman, N.R. (2022). Terrestrial capabilities of invasive fishes and their management implications. Integr. Comp. Biol. *62*, icac023-1394. https://doi.org/10.1093/icb/icac023.
6. Zaccone, G., Lauriano, E.R., Capillo, G., and Kuciel, M. (2018). Air- breathing in fish: Air- breathing organs and control of respiration: Nerves and neurotransmitters in the air-breathing organs and the skin. Acta Histochem. *120*, 630–641. https://doi.org/10.1016/j.acthis.2018.08.009.
7. Kuciel, M., Rita Lauriano, E., Silvestri, G., Zuwała, K., Pergolizzi, S., and Zaccone, D. (2014). The structural organization and immunohistochemistry of G-protein alpha subunits in the olfactory system of the air-breathing mudskipper, *Periophthalmus barbarus* (Linnaeus, 1766) (Gobiidae, Oxudercinae). Acta Histochem. *116*, 70–78. https://doi.org/10.1016/j.acthis.2013.05.005.
8. Wang, K., Wang, J., Zhu, C., Yang, L., Ren, Y., Ruan, J., Fan, G., Hu, J., Xu, W., Bi, X., et al. (2021). African lungfish genome sheds light on the vertebrate water-to-land transition. Cell *184*, 1362–1376.e18. https://doi.org/10.1016/j.cell.2021.01.047.
9. Bi, X., Wang, K., Yang, L., Pan, H., Jiang, H., Wei, Q., Fang, M., Yu, H., Zhu, C., Cai, Y., et al. (2021). Tracing the genetic footprints of vertebrate landing in non-teleost ray-finned fishes. Cell *184*, 1377–1391.e14. https://doi.org/10.1016/j.cell.2021.01.046.
10. Li, N., Bao, L., Zhou, T., Yuan, Z., Liu, S., Dunham, R., Li, Y., Wang, K., Xu, X., Jin, Y., et al. (2018). Genome sequence of walking catfish (*Clarias batrachus*) provides insights into terrestrial adaptation. BMC Genom. *19*, 952. https://doi.org/10.1186/s12864-018-5355-9.
11. Kushwaha, B., Pandey, M., Das, P., Joshi, C.G., Nagpure, N.S., Kumar, R., Kumar, D., Agarwal, S., Srivastava, S., Singh, M., et al. (2021). The genome of walking catfish *Clarias magur* (Hamilton, 1822) unveils the genetic basis that may have facilitated the development of environmental and terrestrial adaptation systems in air-breathing catfishes. DNA Res. *28*, dsaa031. https://doi.org/10.1093/dnares/dsaa031.
12. Sun, B., Huang, S., Huang, L., Yang, L., Gao, J., and Cao, X. (2021). Fibronectin 1B gene plays an important role in loach barbel air-breathing. Int. J. Mol. Sci. *22*, 11928. https://doi.org/10.3390/ijms222111928.
13. Sun, B., Gao, J., Yang, L., Huang, S., and Cao, X. (2022). Depletion of LOXL2 improves respiratory capacity: From air-breathing fish to mammal under hypoxia. Int. J. Biol. Macromol. *209*, 563–575. https://doi.org/10.1016/j.ijbiomac.2022.04.040.
14. Huang, S., Cao, X., and Tian, X. (2016). Transcriptomic analysis of compromise between air-breathing and nutrient uptake of posterior intestine in loach (*Misgurnus anguillicaudatus*), an air-breathing fish. Mar. Biotechnol. *18*, 521–533. https://doi.org/10.1007/s10126-016-9713-9.
15. Ablain, J., Durand, E.M., Yang, S., Zhou, Y., and Zon, L.I. (2015). A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. Dev. Cell *32*, 756–764. https://doi.org/10.1016/j.devcel.2015.01.032.
16. Naert, T., and Vleminckx, K. (2018). CRISPR/Cas9 disease models in zebrafish and Xenopus: The genetic renaissance of fish and frogs. Drug Discov. Today Technol. *28*, 41–52. https://doi.org/10.1016/j.ddtec.2018.07.001.
17. Wang, S.W., Gao, C., Zheng, Y.M., Yi, L., Lu, J.C., Huang, X.Y., Cai, J.B., Zhang, P.F., Cui, Y.H., and Ke, A.W. (2022). Current applications and future perspective of CRISPR/Cas9 gene editing in cancer. Mol. Cancer *21*, 57. https://doi.org/10.1186/s12943-022-01518-8.
18. Zhang, Y., Lv, M., Jiang, H., Li, H., Li, R., Yang, C., Huang, Y., Zhou, H., Mei, Y., Gao, J., and Cao, X. (2023). Mitotic defects lead to unreduced sperm formation in cdk1$^{(-/-)}$ mutants. Int. J. Biol. Macromol. *242*, 125171. https://doi.org/10.1016/j.ijbiomac.2023.125171.
19. Sun, B., Huang, Y., Castro, L.F.C., Yang, S., Huang, S., Jin, W., Zhou, H., Ijiri, S., Luo, Y., Gao, J., and Cao, X. (2023). The chromosome-level genome and key genes associated with mud-dwelling behavior and adaptations of hypoxia and noxious environments in loach (*Misgurnus anguillicaudatus*). BMC Biol. *21*, 18. https://doi.org/10.1186/s12915-023-01517-1.
20. Lu, Y., Bi, T., Zhou, S., and Guo, M. (2023). MEX3A promotes angiogenesis in colorectal cancer via glycolysis. Libyan J. Med. *18*, 2202446. https://doi.org/10.1080/19932820.2023.2202446.
21. Rezzola, S., Di Somma, M., Corsini, M., Leali, D., Ravelli, C., Polli, V.A.B., Grillo, E., Presta, M., and Mitola, S. (2019). VEGFR2 activation mediates the pro-angiogenic activity of BMP4. Angiogenesis *22*, 521–533. https://doi.org/10.1007/s10456-019-09676-y.
22. Lee, Y.J., Lee, H.J., Choi, S.H., Jin, Y.B., An, H.J., Kang, J.H., Yoon, S.S., and Lee, Y.S. (2012). Soluble HSPB1 regulates VEGF-mediated angiogenesis through their direct interaction. Angiogenesis *15*, 229–242. https://doi.org/10.1007/s10456-012-9255-3.
23. Chaube, B., Citrin, K.M., Sahraei, M., Singh, A.K., de Urturi, D.S., Ding, W., Pierce, R.W., Raaisa, R., Cardone, R., Kibbey, R., et al. (2023). Suppression of angiopoietin-like 4 reprograms endothelial cell metabolism and inhibits angiogenesis. Nat. Commun. *14*, 8251. https://doi.org/10.1038/s41467-023-43900-0.
24. Yang, X., Liu, H., Ma, Z., Zou, Y., Zou, M., Mao, Y., Li, X., Wang, H., Chen, T., Wang, W., and Yang, R. (2019). Chromosome-level genome assembly of *Triplophysa tibetana*, a fish adapted to the harsh high-altitude environment of bthe Tibetan Plateau. Mol. Ecol. Resour. *19*, 1027–1036. https://doi.org/10.1111/1755-0998.13021.
25. Yuan, D., Chen, X., Gu, H., Zou, M., Zou, Y., Fang, J., Tao, W., Dai, X., Xiao, S., and Wang, Z. (2020). Chromosomal genome of *Triplophysa bleekeri* provides insights into its evolution and environmental adaptation. GigaScience *9*, giaa132. https://doi.org/10.1093/gigascience/giaa132.
26. Yang, L., Wang, Y., Wang, T., Duan, S., Dong, Y., Zhang, Y., and He, S. (2019). A chromosome-scale reference assembly of a tibetan loach, *Triplophysa siluroides*. Front. Genet. *10*, 991. https://doi.org/10.3389/fgene.2019.00991.
27. Fujimoto, T., Yamada, A., Kodo, Y., Nakaya, K., Okubo-Murata, M., Saito, T., Ninomiya, K., Inaba, M., Kuroda, M., Arai, K., and Murakami, M. (2017). Development of nuclear DNA markers to characterize genetically diverse groups of *Misgurnus anguillicaudatus* and its closely related species. Fish. Sci. *83*, 743–756. https://doi.org/10.1007/s12562-017-1108-y.
28. Li, Y.J., Tian, Y., Zhang, M.Z., Tian, P.P., Yu, Z., Abe, S., and Arai, K. (2010). Chromosome banding and FISH with rDNA probe in the diploid and tetraploid loach *Misgurnus anguillicaudatus*. Ichthyol. Res. *57*, 358–366. https://doi.org/10.1007/s10228-010-0168-0.
29. Buchet-Poyau, K., Courchet, J., Le Hir, H., Séraphin, B., Scoazec, J.Y., Duret, L., Domon-Dell, C., Freund, J.N., and Billaud, M. (2007).

Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. Nucleic Acids Res. *35*, 1289–1300. https://doi.org/10.1093/nar/gkm016.

30. Wang, B., Hong, Z., Zhao, C., Bi, Q., Yuan, J., Chen, J., and Shen, Y. (2021). The effects of MEX3A knockdown on proliferation, apoptosis and migration of osteosarcoma cells. Cancer Cell Int. *21*, 197. https://doi.org/10.1186/s12935-021-01882-3.

31. Liang, J., Li, H., Han, J., Jiang, J., Wang, J., Li, Y., Feng, Z., Zhao, R., Sun, Z., Lv, B., and Tian, H. (2020). Mex3a interacts with LAMA2 to promote lung adenocarcinoma metastasis via PI3K/AKT pathway. Cell Death Dis. *11*, 614. https://doi.org/10.1038/s41419-020-02858-3.

32. Jiang, Y., Feng, S., Xu, J., Zhang, S., Li, S., Sun, X., and Xu, P. (2016). Comparative transcriptome analysis between aquatic and aerial breathing organs of *Channa argus* to reveal the genetic basis underlying bimodal respiration. Mar. Genomics *29*, 89–96. https://doi.org/10.1016/j.margen.2016.06.002.

33. Leroyer, A.S., Blin, M.G., Bachelier, R., Bardin, N., Blot-Chabaud, M., and Dignat-George, F. (2019). CD146 (Cluster of Differentiation 146). Arterioscler. Thromb. Vasc. Biol. *39*, 1026–1033. https://doi.org/10.1161/atvbaha.119.312653.

34. Xue, B., Wang, P., Yu, W., Feng, J., Li, J., Zhao, R., Yang, Z., Yan, X., and Duan, H. (2022). CD146 as a promising therapeutic target for retinal and choroidal neovascularization diseases. Sci. China Life Sci. *65*, 1157–1170. https://doi.org/10.1007/s11427-021-2020-0.

35. Fraisl, P., Mazzone, M., Schmidt, T., and Carmeliet, P. (2009). Regulation of angiogenesis by oxygen and metabolism. Dev. Cell *16*, 167–179. https://doi.org/10.1016/j.devcel.2009.01.003.

36. Brahimi-Horn, M.C., and Pouysségur, J. (2005). The hypoxia-inducible factor and tumor progression along the angiogenic pathway. Int. Rev. Cytol. *242*, 157–213. https://doi.org/10.1016/s0074-7696(04)42004-x.

37. St Johnston, D., and Sanson, B. (2011). Epithelial polarity and morphogenesis. Curr. Opin. Cell Biol. *23*, 540–546. https://doi.org/10.1016/j.ceb.2011.07.005.

38. Urbano, J.-M., Naylor, H.W., Scarpa, E., Mureşan, L., and Sanson, B. (2018). Suppression of epithelial folding at actomyosin-enriched compartment boundaries downstream of Wingless signalling in *Drosophila*. Development (Cambridge, England) *145*, dev155325. https://doi.org/10.1242/dev.155325.

39. Sainio, A., and Järveläinen, H. (2020). Extracellular matrix-cell interactions: Focus on therapeutic applications. Cell. Signal. *66*, 109487. https://doi.org/10.1016/j.cellsig.2019.109487.

40. Nersisyan, S., Novosad, V., Engibaryan, N., Ushkaryov, Y., Nikulin, S., and Tonevitsky, A. (2021). ECM-Receptor Regulatory Network and Its Prognostic Role in Colorectal Cancer. Front. Genet. *12*, 782699. https://doi.org/10.3389/fgene.2021.782699.

41. Chen, Y., Xiao, D., Zhang, L., Cai, C.L., Li, B.Y., and Liu, Y. (2021). The role of Tbx20 in cardiovascular development and function. Front. Cell Dev. Biol. *9*, 638542. https://doi.org/10.3389/fcell.2021.638542.

42. Boogerd, C.J., Zhu, X., Aneas, I., Sakabe, N., Zhang, L., Sobreira, D.R., Montefiori, L., Bogomolovas, J., Joslin, A.C., Zhou, B., et al. (2018). Tbx20 is required in Mid-Gestation Cardiomyocytes and plays a central role in atrial development. Circ. Res. *123*, 428–442. https://doi.org/10.1161/circresaha.118.311339.

43. Tang, Y., Aryal, S., Geng, X., Zhou, X., Fast, V.G., Zhang, J., Lu, R., and Zhou, Y. (2022). TBX20 improves contractility and mitochondrial function during direct human cardiac reprogramming. Circulation *146*, 1518–1536. https://doi.org/10.1161/circulationaha.122.059713.

44. Shen, T., Aneas, I., Sakabe, N., Dirschinger, R.J., Wang, G., Smemo, S., Westlund, J.M., Cheng, H., Dalton, N., Gu, Y., et al. (2011). Tbx20 regulates a genetic program essential to adult mouse cardiomyocyte function. J. Clin. Invest. *121*, 4640–4654. https://doi.org/10.1172/jci59472.

45. Ji, Y., Ma, Y., Shen, J., Ni, H., Lu, Y., Zhang, Y., Ma, H., Liu, C., Zhao, Y., Ding, S., et al. (2021). TBX20 contributes to balancing the differentiation of perivascular adipose-derived stem cells to vascular lineages and neointimal hyperplasia. Front. Cell Dev. Biol. *9*, 662704. https://doi.org/10.3389/fcell.2021.662704.

46. Meng, S., Gu, Q., Yang, X., Lv, J., Owusu, I., Matrone, G., Chen, K., Cooke, J.P., and Fang, L. (2018). TBX20 regulates angiogenesis through the Prokineticin 2-Prokineticin Receptor 1 pathway. Circulation *138*, 913–928. https://doi.org/10.1161/circulationaha.118.033939.

47. He, Y., Kam, H., Wu, X., Chen, Q., and Lee, S.M.Y. (2023). Dual effect of aucubin on promoting VEGFR2 mediated angiogenesis and reducing RANKL-induced bone resorption. Chin. Med. *18*, 108. https://doi.org/10.1186/s13020-023-00786-w.

48. Wong, W., Huang, Y., Wu, Z., Kong, Y., Luan, J., Zhang, Q., Pan, J., Yan, K., and Zhang, Z. (2021). Mvda is required for zebrafish early development. Biol. Res. *54*, 17. https://doi.org/10.1186/s40659-021-00341-7.

49. Harrison, M.R.M., Bussmann, J., Huang, Y., Zhao, L., Osorio, A., Burns, C.G., Burns, C.E., Sucov, H.M., Siekmann, A.F., and Lien, C.L. (2015). Chemokine-guided angiogenesis directs coronary vasculature formation in zebrafish. Dev. Cell *33*, 442–454. https://doi.org/10.1016/j.devcel.2015.04.001.

50. Lawson, N.D., and Weinstein, B.M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. Dev. Biol. *248*, 307–318. https://doi.org/10.1006/dbio.2002.0711.

51. Griffioen, A.W., and Dudley, A.C. (2021). Angiogenesis: a year in review. Angiogenesis *24*, 195–196. https://doi.org/10.1007/s10456-021-09798-2.

52. Dudley, A.C., and Griffioen, A.W. (2023). The modes of angiogenesis: an updated perspective. Angiogenesis *26*, 477–480. https://doi.org/10.1007/s10456-023-09895-4.

53. Folkman, J. (2006). Angiogenesis. Annu. Rev. Med. *57*, 1–18. https://doi.org/10.1146/annurev.med.57.121304.131306.

54. Lamalice, L., Le Boeuf, F., and Huot, J. (2007). Endothelial cell migration during angiogenesis. Circ. Res. *100*, 782–794. https://doi.org/10.1161/01.RES.0000259593.07661.1e.

55. Gao, L., Yang, J., Li, Y., Liu, K., Sun, H., Tang, J., Xia, Z., Zhang, L., and Hu, Z. (2022). Long noncoding RNA SCIRT promotes HUVEC angiogenesis via stabilizing VEGFA mRNA induced by hypoxia. Oxid. Med. Cell. Longev. *2022*, 9102978. https://doi.org/10.1155/2022/9102978.

56. Taha, K., and Yoo, P.D. (2016). Predicting the functions of a protein from its ability to associate with other molecules. BMC Bioinf. *17*, 34. https://doi.org/10.1186/s12859-016-0882-3.

57. Senior, A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A.W.R., Bridgland, A., et al. (2020). Improved protein structure prediction using potentials from deep learning. Nature *577*, 706–710. https://doi.org/10.1038/s41586-019-1923-7.

58. Chasapis, C.T., and Spyroulias, G.A. (2009). RING finger E(3) ubiquitin ligases: structure and drug discovery. Curr. Pharm. Des. *15*, 3716–3731. https://doi.org/10.2174/138161209789271825.

59. Liu, L., Wong, C.C., Gong, B., and Yu, J. (2018). Functional significance and therapeutic implication of ring-type E3 ligases in colorectal cancer. Oncogene *37*, 148–159. https://doi.org/10.1038/onc.2017.313.

60. Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. Nat. Methods *18*, 170–175. https://doi.org/10.1038/s41592-020-01056-5.

61. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden, E.L. (2017). *De novo* assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95. https://doi.org/10.1126/science.aal3327.

62. Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate *de novo* genome assembly from long uncorrected reads. Genome Res. *27*, 737–746. https://doi.org/10.1101/gr.214270.116.

63. Seppey, M., Manni, M., and Zdobnov, E.M. (2019). BUSCO: Assessing genome assembly and annotation completeness. Methods Mol. Biol. *1962*, 227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.

64. Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. *27*, 573–580. https://doi.org/10.1093/nar/27.2.573.

65. Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. USA *117*, 9451–9457. https://doi.org/10.1073/pnas.1921046117.

66. Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Curr. Protoc. Bioinformatics *Chapter 4*, 4.10.1–4.10.14. https://doi.org/10.1002/0471250953.bi0410s25.

67. Mario, S., Mark, D., Robert, B., and David, H. (2008). Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics *24*, 637–644. https://doi.org/10.1093/bioinformatics/btn013.

68. Borodovsky, M., and Lomsadze, A. (2011). Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. Curr. Protoc. Bioinformatics *Chapter 4*, 4.6.1–4.6.10. https://doi.org/10.1002/0471250953.bi0406s35.

69. Keilwagen, J., Hartung, F., and Grau, J. (2019). GeMoMa: Homology-based gene prediction utilizing intron position conservation and RNA-seq data. Methods Mol. Biol. *1962*, 161–177. https://doi.org/10.1007/978-1-4939-9173-0_9.

70. Gremme, G., Brendel, V., Sparks, M.E., and Kurtz, S. (2005). Engineering a software tool for gene structure prediction in higher organisms. Inf. Software Technol. *47*, 965–978.

71. Shumate, A., Wong, B., Pertea, G., and Pertea, M. (2022). Improved transcriptome assembly using a hybrid of long and short reads with StringTie. PLoS Comput. Biol. *18*, e1009730. https://doi.org/10.1371/journal.pcbi.1009730.

72. Jo, Y., Lian, S., Cho, J.K., Choi, H., Chu, H., and Cho, W.K. (2015). *De novo* transcriptome assembly of two different Prunus mume cultivars. Genom. Data *6*, 273–274. https://doi.org/10.1016/j.gdata.2015.10.011.

73. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. *9*, R7. https://doi.org/10.1186/gb-2008-9-1-r7.

74. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

75. Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., and Chen, J. (2020). RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. PeerJ. Comput. Sci. *6*, e251. https://doi.org/10.7717/peerj-cs.251.

76. Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. *40*, e49. https://doi.org/10.1093/nar/gkr1293.

77. Delcher, A.L., Salzberg, S.L., and Phillippy, A.M. (2003). Using MUMmer to identify similar regions in large sequence sets. Curr Protoc Bioinformatics *Chapter* 10, Unit 10.13. https://doi.org/10.1002/0471250953.bi1003s00.

78. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

79. Wlodzimierz, P., Rabanal, F.A., Burns, R., Naish, M., Primetis, E., Scott, A., Mandáková, T., Gorringe, N., Tock, A.J., Holland, D., et al. (2023). Cycles of satellite and transposon evolution in Arabidopsis centromeres. Nature *618*, 557–565. https://doi.org/10.1038/s41586-023-06062-z.

80. Carty, B.L., and Dunleavy, E.M. (2020). Centromere assembly and non-random sister chromatid segregation in stem cells. Essays Biochem. *64*, 223–232. https://doi.org/10.1042/ebc20190066.

81. Li, W., Li, S., Li, Y., Lin, X., Hu, Y., Meng, T., Wu, B., He, R., and Feng, D. (2019). Immunofluorescence staining protocols for major autophagy proteins including LC3, P62, and ULK1 in mammalian cells in response to normoxia and hypoxia. Methods Mol. Biol. *1854*, 175–185. https://doi.org/10.1007/7651_2018_124.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-Mex3a	Abclonal Technology	Cat# A22102; RRID: AB_3662045
Rabbit polyclonal anti-Tbx20	Abclonal Technology	Cat# A15994; RRID:AB_2763434
Rabbit polyclonal anti-Mmp9	Abclonal Technology	Cat# A0289; RRID:AB_2757101
Rabbit polyclonal anti-Mmp13	Abclonal Technology	Cat# A1606; RRID:AB_2763499
Rabbit polyclonal anti-Mapkapk2	Abclonal Technology	Cat# A22183; RRID: AB_3662047
HRP-conjugated IgG Fraction Monoclonal Mouse Anti-Rabbit IgG, Light Chain Specific antibody	Proteintech Company	Cat# SA00001-7L; RRID:AB_2890988
Rabbit polyclonal anti-Prok2	Abclonal Technology	Cat# A6705; RRID:AB_2767288
Rabbit polyclonal anti-Prokr1	Cusabio Technology	Cat# CSB-PA003812; RRID: AB_3662046
Anti-Gapdh	ABclonal Technology	Cat# AC001; RRID:AB_2619673
Rabbit polyclonal anti-CD146	ABclonal Technology	Cat# A13927; RRID:AB_2760779
Bacterial and virus strains		
pcDNA3.1(+)eGFP	Wuhan Tianyi Huayu Gene Technology	N/A
Critical commercial assays		
MAXIscript® SP6/T7 Kit	Thermo Fisher Scientific	Cat# AM1320
Protein A + G Agarose (Fast Flow, for IP) kit	Beyotime Biotechnology Company	Cat# P2055-2mL
RFectPMsiRNA/miRNA transfection reagent	Changzhou Biogenerating Biotechnology	N/A
Hieff Mut^{TM} Site-Directed Mutagenesis Kit	Yeasen Biotechnology	Cat# 11003ES10
Deposited data		
Loach haplotype-Mchra	This paper	GWHERQE00000000
Loach haplotype-Mchrb	This paper	GWHERQD00000000
RNA_seq: WT and KO loach intestine tissues	This paper	PRJNA1091403
Experimental models: Cell lines		
HUVEC	Wuhan Pricella Biotechnology	CL-0675
Experimental models: Organisms/strains		
Loach	This paper	N/A
AB strain zebrafish	China Zebrafish Resource Center	CZ1
Tg (*fli1a*:EGFP) transgenic zebrafish	China Zebrafish Resource Center	CZ-ID:CZ55
Oligonucleotides		
Primers	This paper	N/A
Software and algorithms		
GraphPad Prism 8.0 software	GraphPad Software Inc., San Diego, CA, USA	N/A
IBM SPSS Statistics v 26.0	SPSS Inc., USA	N/A
Hifiasm	Li et al.[60]	https://github.com/chhylp123/hifiasm
3D-DNA	Dudchenko et al.[61]	https://github.com/aidenlab/3d-dna
Racon	Vaser et al.[62]	https://github.com/isovic/racon
BUSCO	Seppey et al.[63]	https://busco.ezlab.org/
TRF	Benson[64]	https://tandem.bu.edu/trf/trf.html
RepeatModeler	Flynn et al.[65]	https://github.com/Dfam-consortium/RepeatModeler

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RepeatMasker	Tarailo-Graovac and Chen[66]	https://github.com/rmhubley/RepeatMasker
AUGUSTUS	Mario et al.[67]	https://github.com/Gaius-Augustus/Augustus
GeneMark-ES	Borodovsky and Lomsadze[68]	https://github.com/gatech-genemark/GeneMark-EP-plus
GeMoMa	Keilwagen et al.[69]	https://www.jstacs.de/index.php/GeMoMa-Docs
GenomeThreader	Gremme et al.[70]	https://github.com/genometools/genomethreader
StringTie	Shumate et al.[71]	https://github.com/gpertea/stringtie/releases
TransDecoder	Jo et al.[72]	https://github.com/TransDecoder/TransDecoder
EVidenceModeler	Haas et al.[73]	https://github.com/EVidenceModeler/EVidenceModeler
SAMtools	Li et al.[74]	https://github.com/samtools/samtools
R	Hao et al.[75]	https://www.r-project.org/
MCscan	Wang et al.[76]	https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version)
MUMmer	Delcher et al.[77]	https://github.com/mummer4/mummer
RSEM	Li and Dewey[78]	https://github.com/deweylab/RSEM

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

The adult loach (one year old; the average body length was 10.5 cm) used in this study were from a laboratory-reared population. The adult AB strain zebrafish and Tg (*fli1a*:EGFP) transgenic zebrafish (CZ-ID:CZ55) were purchased from China Zebrafish Resource Center. All the loach and zebrafish were cultured in a recirculating water system with a temperature of 24°C–26°C, pH of 7.0–7.5, and dissolved oxygen (DO) of 6.0–6.5 mg/L. All experimental protocols in this study were approved by the Animal Experimental Ethical Inspection of Laboratory Animal Center, Huazhong Agricultural University, Wuhan, China (HZAUFI-2021-0032). All efforts were made to minimize the suffering of the fish.

Animal models

We constructed the *mex3a* knockout loach using the CRISPR/Cas9 system. The target sequence for loach *mex3a* was "GTTGGGTCTTGGAGAGGGCGA". To validate the construction efficiency of the *mex3a* knockout loach model, we collected posterior intestine tissues from WT and KO loaches for qPCR and WB analysis. Furthermore, a *mex3a* knockout model in the transgenic zebrafish (*fli1a*:EGFP) was generated. The target sequence for zebrafish *mex3a* was "GCTGGGTCTTGGAGAGGGCG". To validate the construction efficiency, muscle tissues were collected for WB analysis. The primers are listed in Table S10.

HUVEC cells

HUVECs (obtained from the Wuhan Pricella Biotechnology Co., Ltd.) were maintained in HUVEC-specific culture medium (Wuhan Pricella Biotechnology Co., Ltd.), and kept at 37°C in a 0.5% CO2 incubator.

METHOD DETAILS

Screening of genes associated with angiogenesis

Gene enrichment analysis

The GO and KEGG functional enrichment analysis of 120 PSGs obtained from our previous genomic analysis of loach (GCF_027580225.1),[19] were conducted for further screening candidate genes associated with angiogenesis involvement in air-breathing.

Loach haplotype genome assembly and annotation

Due to the inability to identify centromere sequences in most chromosomes of the loach genome (GCF_027580225.1) assembled using PacBio CLR data,[19] we supplemented PacBio HiFi data by using muscle tissue samples from the loach to further enhance the completeness of the genome. Firstly, Hifiasm (v0.16.1-r375)[60] was employed as the primary assembly software to construct two sets of chromosome-level haploid genomes of loach, by utilizing PacBio HiFi sequencing data and Hi-C sequencing data. The assembly process utilized the HiFi reads + Hi-C reads mode with parameters "-o output -t 10 HiFi reads.fq –h1 HiC-R1.fq –h2 HiC-R1.fq". Secondly, pseudochromosomes of the loach were obtained using Hi-C technology. 3D-DNA was used to organize

and orient the generated contigs into pseudochromosomes.[61] The Hi-C contact matrix was visualized using Juicebox, with manual adjustments based on neighboring interactions. Finally, Racon (v1.4.30) was used for genome polish,[62] and BUSCO was used to evaluate genome completeness by using actinopterygii_odb10 database.[63]

For the annotation of tandem repeats and transposable elements (TEs), Tandem Repeats Finder (v 4.09; TRF; default parameters)[64] was employed. Simultaneously, both *de novo* and homologous search strategies were used to identify TEs from the loach genome. The *de novo* repeat library of loach was identified and constructed using RepeatModeler (v 2.0.1; default parameters).[65] Subsequently, RepeatMasker (v 4.1.0)[66] was utilized to align homologous TEs against the Repbase database and integrate them with the *de novo* repeat library.

The genome structural annotation process involved three main methods: *de novo* prediction, homology-based prediction, and RNA-seq assisted prediction. For *de novo* prediction, AUGUSTUS (v 3.3.3),[67] and GeneMark-ES (v 4.68)[68] software were used with a repeat-masked genome as input. For homology-based prediction, GeMoMa (v 1.9)[69] and GenomeThreader (GTH, v 1.7.1)[70] were employed. For RNA-seq assisted prediction, HISAT2 (v 2.2.1), StringTie (v 2.1.4),[71] and TransDecoder[72] were used for transcript alignment, assembly and prediction. Finally, the EVidenceModeler (EVM, v 2.1.0)[73] software was used to integrate the predictions from the three methods, considering the reliability and accuracy of each prediction source, to generate a comprehensive genome annotation.

Gene function annotation primarily relied on external protein databases (NR, SwissProt, TrEMBL, KEGG, GO, InterPro, and Pfam) to annotate the functions of proteins in the gene set. During the annotation process of non-coding RNAs, tRNA sequences in the genome were identified using tRNAscan-SE based on the structural characteristics of tRNAs. rRNA sequence was identified by using blastn with an E-value of 1e−5. Additionally, miRNA and snRNA sequences on the loach genome were predicted using Rfam 9.1 and INFERNAL.

Identification of centromere sequences and chromosome orientation adjustment

Centromeres are specialized DNA sequences that connect sister chromatids.[79] Centromere structures in most animals and plants are complex, primarily composed of highly repetitive satellite DNA and interspersed retrotransposon sequences.[79,80] In this study, centromere sequences in the loach genome were identified using TRF (v 4.0.9).[64] Initially, each chromosome sequence of the loach genome was individually extracted using SAMtools (v 1.9)[74] and then utilized as input for TRF with parameters "2 7 7 80 10 50 500 -f -d -m". As the centromere regions in the loach are mainly comprised of tandem repeat sequences, and the centromere sequences on each chromosome within a species are highly homologous, tblastn was further employed to align the tandem repeat sequences identified on the 25 chromosomes. Subsequently, alignment results were filtered based on 85% identity. Finally, following the general principle "short arm on the left, long arm on the right", chromosome orientation could be further adjusted based on the positions of the centromere candidate sequences. Additionally, chromosomes were numbered based on their lengths, arranged from longest to shortest. The R language and the RIdeogram package were utilized for plotting chromosome centromere maps.[75]

PSGs in the loach and Triplophysa tibetana genomes

Firstly, a local BLAST database of *T. tibetana* genome sequence was established by using the makeblastdb command from BLAST+. Second, gene sequences were extracted from the loach genome as reference sequences; subsequently, the blastn command was used to compare the gene sequences with the *T. tibetana* genome. Finally, based on the alignment result file, the information was parsed to determine their positions on the chromosome.

MCscan (v 1.0; Python version)[76] was used with parameters "–cscore = 0.99; –minspan = 30" for genome synteny analysis between the loach (GCF_027580225.1) and *T. tibetana*. Based on the results of genome synteny analysis of the two species, we conducted a step-by-step visual localization analysis of intestinal air-breathing-related genes using MUMmer (v 4.0.0beta2)[77] with delta-filter parameters "-i 90 -L 100 -m"[10] and LINKVIEW2 from the loach genome and *T. tibetana* genome.

Evolutionary analysis of intestinal air-breathing-related genes

We selected six fish species including African lungfish, spotted gar, *Oryzias latipes*, zebrafish, *T. tibetana*, and the loach for whole-genome synteny analysis. The analysis was conducted using the MCscan software, and jcvi scripts were employed to identify syntenic blocks among the six fishes with parameters "–cscore = 0.50; –minspan = 30". Subsequently, the synteny results were further visualized for intestinal air-breathing-related genes.

Functional analysis of Mex3a

Air exposure

A towel was laid at the base of a glass tank and stayed moist. The fish (three kinds of air-breathing fishes, namely the loach, *C. argus*, and *M. albus*) were put on the moist towel. After 0 and 6 h of exposure, the air-breathing organs were sampled from the fish for *mex3a* expression analysis (Figure S3A). For expression analysis, qPCR (2x FonsSuper SYBR Green qPCR Mix (With ROX) were purchased from Wuhan Fonsber Biotechnology Co., Ltd. Cat no: FS3310-01) and WB were performed. The Mex3a antibody (ABclonal Technology Co., Ltd. Cat no: A22102; Rabbit) and glyceraldehyde-3-phosphate dehydrogenase (Gapdh) antibody (ABclonal Technology Co., Ltd. Cat no: AC001) were diluted 1000 and 10000 times with PBS buffer, respectively. The primers are listed in Table S10. The qPCR and WB experiments follows procedures described by Sun et al.[21]

Expression and location analysis of loach Mex3a

Eleven tissues including intestine, eye, skin, spleen, gill, kidney, brain, fin, barbel, heart, and liver from the loach were collected and used for qPCR analysis (three loaches were used here). The relative expression levels were normalized to the endogenous control

genes, namely *β-actin* and *gapdh*. To investigate the expression pattern of *mex3a* during the early development of loach, we conducted whole-mount *in situ* hybridization. Additionally, we analyzed the expression signals of Mex3a in the intestine of loach using tissue *in situ* hybridization and immunofluorescence. The primers are listed in Table S10. The procedures of whole-mount *in situ* hybridization and immunofluorescence followed to Sun et al.[19]

Construction of mex3a knockout loach

We constructed the *mex3a* knockout loach using the CRISPR/Cas9 system.

Phenotypic analysis of mex3a knockout loach

Firstly, we conducted air exposure experiments with WT and KO loaches and recorded the survival time of each loach (eight loaches were sampled from each group). To determine the death of loach, we placed the loaches back into the water, and if they did not swim within half an hour, we considered they were dead. Posterior intestines of WT and KO loaches exposed to air for 0 and 6 h were collected for H&E staining, TEM, immunofluorescence, and gene expression analysis. For immunofluorescence, CD146 antibody (ABclonal Technology Co., Ltd. Cat no: A13927; diluted 50 times with PBS buffer; Rabbit) was used to label new-formed blood vessels.

Furthermore, we conducted a chronic hypoxia treatment. Two 2 L-conical flasks containing 1.5 L of water (with an oxygen content of 6.0 mg/L) were utilized, with a layer of peanut oil added to create a barrier between the water and the air (Figure S3N). Four same-sized WT and KO loaches were respectively placed into the conical flasks. During the experiment, a dissolved oxygen analyzer was employed to continuously monitor fluctuations in water oxygen levels. Simultaneously, the occurrences of loach air-breathing behavior were recorded (loach air-breathing behavior defined as when the head of the loach emerged above the water surface). When the oxygen content reached 0 mg/L, the experiment was stopped. Posterior intestine tissues of WT and KO loaches were collected for immunofluorescence and gene expression analysis.

RNA_seq analysis

After exposed to air for 6 h, posterior intestine tissues were sampled from WT and KO loaches (3 loaches per replicate, 3 replicates per group) for RNA-seq analysis. Quantitative analysis of gene and transcript expression levels were conducted using the RSEM (v 1.3.3) software,[78] with expression levels measured in Fragments Per Kilobase per Million reads (FPKM). Gene with expression levels presenting $\log_2$(FoldChange) ≥ 2 and $p \leq 0.05$ were considered as DEGs. Enrichment analyses for GO and KEGG pathways were performed for the identified DEGs.

IP analysis

The main steps of the IP experiment referred to the Protein A + G Agarose (Fast Flow, for IP) kit (Cat no: P2055-2mL) from Beyotime Biotechnology Company. Firstly, protein extraction was conducted using liver tissues obtained from wild-type loaches. The extracted protein was subsequently incubated with 4 μL of Mex3a antibody and gently shaken overnight at 4°C. Then, 40 μL Protein A + G Agarose was added and incubated for 3 h with slow shaking at 4°C. For WB analysis, Tbx20 (ABclonal Technology Co., Ltd. Cat no: A15994; Rabbit; diluted 1000 times), Mmp9 (ABclonal Technology Co., Ltd. Cat no: A0289; Rabbit; diluted 700 times), Mmp13 (ABclonal Technology Co., Ltd. Cat no: A1606; Rabbit; diluted 700 times), and Mapkapk2 (ABclonal Technology Co., Ltd. Cat no: A22183; Rabbit; diluted 2000 times) were used as the primary antibodies, respectively. To avoid interference from heavy chains, the HRP-conjugated IgG Fraction Monoclonal Mouse Anti-Rabbit IgG, Light Chain Specific antibody (diluted 10,000 times) from Proteintech China Biotechnology Company (Cat no: SA00001-7L) was utilized as the secondary antibody.

Tbx20-Prok2-Prokr1 pathway expression analysis

We collected the posterior intestine tissues from WT and KO loaches for qPCR, WB, and immunofluorescence analysis. Prok2 antibody (ABclonal Technology Co., Ltd. Cat no: A6705) and Prokr1 antibody (Cusabio Technology LLC. Cat no: CSB-PA003812) were diluted 1000 times with PBS buffer. For immunofluorescence analysis, Mex3a and Tbx20 antibodies were diluted 50 times with PBS buffer. The primers are listed in Table S10.

Expression and location analysis of zebrafish Mex3a

Ten tissues including swim bladder, muscle, ovary, intestine, kidney, liver, heart, gill, brain, and eye from the zebrafish (AB strain) were collected and used for qPCR analysis (three zebrafishes were used here). The relative expression levels were normalized to the endogenous control genes, namely *β-actin* and *gapdh*. The primers are listed in Table S10. Additionally, we analyzed the expression signals of Mex3a in the intestine of zebrafish (AB strain) using immunofluorescence. The procedures of immunofluorescence followed to Sun et al.[19]

Construction of mex3a knockout model in transgenic zebrafish (fli1a:EGFP)

We constructed a *mex3a* knockout model in transgenic zebrafish (*fli1a*:EGFP) using the CRISPR/Cas9 system.

Detection of early vascular development of zebrafish

Following the successful establishment of *mex3a* knockout zebrafish, we observed and captured images of WT and KO zebrafishes at 30 h postfertilization (hpf) under a green fluorescence microscope following the method described by Meng et al.[46] Meanwhile, Tbx20-Prok2-Prokr1 pathway expression analyses in WT and KO zebrafishes were performed. The intestine tissues from WT and KO zebrafishes were also sampled for qPCR and WB analysis. Prok2 antibody and Prokr1 antibody were diluted 1000 times with PBS buffer. The primers are listed in Table S10.

Construction of HUVEC Mex3a knockdown and overexpression models

When the cell density of HUVECs reached 60–80%, passaging was conducted. siRNA transfection was performed using RFectPMsiRNA/miRNA transfection reagent (Changzhou Biogenerating Biotechnology Co., Ltd.). The volume ratio of si:Mex3a to RFectPMsiRNA/miRNA transfection reagent was 6:1. HUVEC Mex3a-KD model was established.

For HUVEC *Mex3a*-OE model construction, the overexpression vector for human Mex3a (Mex3a-pcDNA3.1(+)eGFP; *Mex3a*-OE group) was constructed by Wuhan Tianyi Huayu Gene Technology Co., Ltd. Plasmid extraction was conducted using the MolPure Endo-free Plasmid Mini Kit (Yeasen Biotechnology Co., Ltd.), following the manufacturer's instructions. Hieff Trans Liposomal Transfection Reagent (Yeasen Biotechnology Co., Ltd.) was used for plasmid cell transfection. The volume ratio of Mex3a-pcDNA3.1(+) eGFP and pcDNA3.1(+) eGFP (the Control group) to Hieff Trans Liposomal Transfection Reagent was 2.5 μg:6 μL. Each group had 3 replicates. During transfection, HUVECs were cultured in Opti-MEM medium (Thermo Fisher Scientific Inc.). After 6 h, the medium was replaced with HUVEC-specific culture medium (Wuhan Pricella Biotechnology Co., Ltd.), and the transfected cells were cultured at 37°C for 48 h, and then used for WB analysis. To clearly verify the success of transfection and expression of exogenous genes, fluorescence microscopy, qPCR, and WB analysis were performed. Three replicates were performed. The primers are listed in Table S10.

HUVEC migration analysis under normoxia

After successful constructions of the HUVEC *Mex3a*-KD and *Mex3a*-OE model, cell migration experiment was conducted after 24 h incubation. Firstly, a consistent-width straight line was drawn by using the 1 mL pipette tip. Then, the medium was replaced with new HUVEC culture medium. Finally, the initial (0 h) line distance was recorded under an optical microscope. After 24 h incubation, the cell migration distance was recorded again and analyzed. Three replicates were performed.

HUVEC tube-formation assay under normoxia

The HUVEC *Mex3a*-KD and *Mex3a*-OE group, along with the control group (CT group), were used to carry out the tube-formation assay. The ABW Matrigengel (Shanghai Nova Pharmaceutical Technology Co., Ltd) was used. 96-well plate was utilized, with 40 μL of Matrigel per well. Following a 30 min incubation at 37°C cell culture incubator, 50 μL of cell suspension containing approximately 30,000 cells was added to each well. Subsequently, the plate was put into the 37°C cell culture incubator. Photographs were captured at 6 h and 24 h post-incubation. Three replicates were performed.

Expression analysis of TBX20-PROK2-PROKR1 pathway under normoxia

Cells from the CT, *Mex3a*-KD, and *Mex3a*-OE group were harvested for qPCR analysis after 24 h culture and for WB analysis after 48 h culture. Three replicates were performed. The primers are listed in Table S10.

Immunofluorescence

The procedures of cell immunofluorescence were performed as described by Li et al. with slight changes.[81] The polyclonal antibody of Mex3a and Tbx20 was used as the primary antibody and Cy3-labeled goat anti-rabbit antibody (ABclonal, China) used as second antibody. The nuclei were stained with DAPI. Tissues were observed under a laser scanning confocal microscope (Leica DMi8, Germany). The immunofluorescences of cells from the CT, *Mex3a*-KD, and *Mex3a*-OE group were performed after 48 h incubation.

HUVEC migration analysis under hypoxia

After delineating line on the three groups (CT, *Mex3a*-KD, and *Mex3a*-OE), cells were initially incubated in a hypoxic chamber (1% O_2, 37°C) for 4 h, followed by transfer to a 37°C cell culture incubator for continued incubation. Three replicates were performed.

HUVEC tube-formation assay under hypoxia

After 24 h incubation, cells from the CT, *Mex3a*-KD, and *Mex3a*-OE group were exposed to a chamber (1% O_2, 37°C) for 4 h before proceeding to the tube-formation assay. Three replicates were performed.

Expression analysis of TBX20-PROK2-PROKR1 pathway under hypoxia

After 24 h culture, cells from the CT, *Mex3a*-KD, and *Mex3a*-OE group were placed in a hypoxic chamber (1% O_2, 37°C) for 4 h before being collected for qPCR analysis. After 48 h culture, the cells were placed in a hypoxic chamber (1% O_2, 37°C) for 4 h before being collected for WB analysis. Three replicates were performed. The primers are listed in Table S10.

The effect of loach Mex3a mutation N201S on the migration and tube-forming ability of HUVECs

Prediction and comparative analysis of Mex3a protein structure

The conservation and divergence of Mex3a amino acid sequences from five fish species including zebrafish, *T. tibetana*, the loach, African lungfish, and spotted gar were analyzed by using the online software NCBI-BLAST-Multiple Alignment (COBALT: Multiple Alignment Tool (nih.gov). Additionally, we selected Mex3a amino acid sequences from 11 fish species (zebrafish, spotted gar, *T. tibetana*, *C. argus*, *L. chalumnae*, *E. calabaricus*, *M. armatus*, *K. marmoratus*, *C. idellus*, *L. rohita*, and the loach) for analysis. The sequences were aligned using the Jalview software subroutine Mafft (default parameters). Positive selection sites were identified based on the alignment results combined with positive selection analysis results (188N, the Asparagine at position 188 was under positive selection) (Table S9).

Prediction of the tertiary structure of Mex3a protein

We used the sub-tool Protein Structure Prediction (AlphaFold2.3.0) within the online software WeMol for predicting tertiary structures (https://wemol.wecomput.com/ui/#/frontend/workflow/workflow-modules). Following the prediction, we selected the result with the highest score for subsequent protein structure comparative analysis.

Prediction of tertiary structure after Mex3a mutation (N201S) of loach

We conducted the analysis of the amino acid at position 201st of Mex3a of loach (201N), *T. tibetana* (201S), and zebrafish (201S) to compare their differences. After mutating Mex3a amino acid (N201S) of loach, we performed Mex3a protein tertiary structure prediction among loach (201S), *T. tibetana* (201S), and zebrafish (201S).

Effect of exogenous loach and zebrafish mex3a gene on the migration and tube formation of HUVECs

Overexpression with zebrafish *mex3a* (*mex3a*-Dr group) or loach *mex3a* (*mex3a*-Ma group) in HUVEC were conducted using the pcDNA3.1(+)eGFP plasmid. In comparison to the control group (pcDNA3.1(+)eGFP; *egfp* group), we assessed the effect of the *mex3a*-Dr and *mex3a*-Ma on HUVEC migration and tube formation under normoxic and hypoxic conditions. Three replicates were performed. The primers are listed in Table S10.

Effect of Mex3a mutation (N201S) of loach on the migration and tube formation of HUVECs

Loach-Mex3a-mutation (N201S, *mex3a*-N201S-Ma) HUVEC model was established using the Hieff MutTM Site-Directed Mutagenesis Kit (Yeasen Biotechnology Co., Ltd.). We assessed the effects of the *mex3a*-N201S-Ma, *mex3a*-Dr, and *mex3a*-Ma on HUVEC migration and tube formation under hypoxic condition. Three replicates were performed. The primers are listed in Table S10.

QUANTIFICATION AND STATISTICAL ANALYSIS

All data was presented as mean ± standard deviation (SD). The data was statistically analyzed and graphically represented using GraphPad Prism 8.0 software. IBM SPSS Statistics v 26.0 was used for statistical evaluation using Student's t test or One-way ANOVA. $p < 0.05$ was considered statistically significant, while $p < 0.01$ and <0.001 indicated very significant and extremely significant differences, respectively ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). Different letters above error bars indicated significant a difference among different groups ($p < 0.05$).