

ARTICLE

Received 7 Oct 2016 | Accepted 16 Dec 2016 | Published 3 Feb 2017

DOI: 10.1038/ncomms14300

OPEN

# Evolution of *Shh* endoderm enhancers during morphological transition from ventral lungs to dorsal gas bladder

Tomoko Sagai<sup>1</sup>, Takanori Amano<sup>1</sup>, Akiteru Maeno<sup>1</sup>, Tetsuaki Kimura<sup>2,†</sup>, Masatoshi Nakamoto<sup>2,†</sup>, Yusuke Takehana<sup>2,3</sup>, Kiyoshi Naruse<sup>2,3</sup>, Norihiro Okada<sup>4</sup>, Hiroshi Kiyonari<sup>5</sup> & Toshihiko Shiroishi<sup>1</sup>

*Shh* signalling plays a crucial role for endoderm development. A *Shh* endoderm enhancer, MACS1, is well conserved across terrestrial animals with lungs. Here, we first show that eliminating mouse MACS1 causes severe defects in laryngeal development, indicating that MACS1-directed *Shh* signalling is indispensable for respiratory organogenesis. Extensive phylogenetic analyses revealed that MACS1 emerged prior to the divergence of cartilaginous and bony fishes, and even euteleost fishes have a MACS1 orthologue. Meanwhile, ray-finned fishes evolved a novel conserved non-coding sequence in the neighbouring region. Transgenic assays showed that MACS1 drives reporter expression ventrally in laryngeal epithelium. This activity has been lost in the euteleost lineage, and instead, the conserved non-coding sequence of euteleosts acquired an enhancer activity to elicit dorsal epithelial expression in the posterior pharynx and oesophagus. These results implicate that evolution of these two enhancers is relevant to the morphological transition from ventral lungs to dorsal gas bladder.

<sup>1</sup>Mammalian Genetics Laboratory, Genetic Strains Research Center, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. <sup>2</sup>Interuniversity Bio-Backup Project Center, National Institute for Basic Biology, Okazaki, Aichi 444-8585, Japan. <sup>3</sup>Laboratory of Bioresources, National Institute for Basic Biology, Okazaki, Aichi 444-8585, Japan. <sup>4</sup>Department of Life Sciences, National Cheng Kung University, Tainan 701, Taiwan. <sup>5</sup>Laboratory for Animal Resources and Genetic Engineering, RIKEN Center for Developmental Biology (CDB), Kobe, Hyogo 650-0047, Japan. † Present addresses: Division of Human Genetics, Department of Integrated Genetics, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan (T.K.); Department of Marine Bioscience, Faculty of Marine Science, Tokyo University of Marine Science and Technology, Minato-ku, Tokyo 108-8477, Japan (M.N.). Correspondence and requests for materials should be addressed to T.S. (email: tshirois@nig.ac.jp).

Morphological evolution involves functional alteration of developmental genes that generally play pleiotropic roles during embryogenesis. It is now inferred that evolutionary changes in *cis*-regulatory elements (CREs) of pleiotropic developmental genes have contributed to morphological evolution more often than have changes in coding sequences<sup>1–4</sup>. This is because each of the pleiotropic functions in such developmental genes is defined by tissue-specific CREs<sup>5</sup>, and mutations in the CREs alter gene expression only in particular tissues without producing deleterious effects in other tissues<sup>6–8</sup>. Identification of the evolutionary changes in CREs that generate a novel expression pattern is thus particularly interesting, because they probably contribute to the appearance of phenotypic novelty. Recently, loss or substitution of CREs has been implicated in human-specific morphological traits<sup>9,10</sup>, and the contribution of sequence divergence of CREs to morphological evolution in *Drosophila*<sup>11</sup> and teleost fish<sup>8</sup> has been reported. However, establishing links between CRE evolution and particular morphological changes is still a major challenge in evolutionary developmental biology.

Sonic hedgehog (*Shh*) is a major pleiotropic developmental gene, and controls cell growth, cell survival and fate, and axial patterning in the vertebrate body plan<sup>12–16</sup>. Mouse *Shh* null mutation causes severe morphological defects in multiple organs including brain, foregut, axial skeleton and limb<sup>17,18</sup>. Conditional deletion of *Shh* expression in the respiratory endoderm epithelium causes respiratory failure with developmental defects in the lungs and tracheal-bronchial ring<sup>19</sup>. These defects may result from a disruption of mesenchymal growth in the foregut endoderm, which is regulated by *Shh* signalling triggered from the epithelium<sup>20</sup>. In the regulatory block spanning 1 Mb upstream of the *Shh* transcription start site, multiple tissue-specific enhancers are clustered<sup>21–26</sup>. These short- and long-range enhancers regulate different modes of *Shh* expression in different tissues<sup>27–29</sup>. We previously identified three epithelium-specific long-range enhancers, MRCS1, MFCS4 and MACS1, in the region 620–740 kb upstream of the *Shh* transcription start site, which direct *Shh* expression and thereby partition the continuous epithelial lining into three segments that give rise to the oral cavity, pharynx and lower respiratory-digestive organs<sup>24</sup>. Among them, the adjacent MFCS4 and MACS1 are separated by 24 kb, and drive reporter expression differentially in the endodermal organs. The phenotype of knockout (KO) mice indicated that MFCS4 is indispensable for morphogenesis of the pharyngeal structure<sup>24</sup>. On the other hand, MACS1 drives reporter expression widely in the epithelia of larynx, lung and intestinal and urogenital tracts<sup>24</sup>. Recently, we reported that a low-conserved enhancer, SLGE, also regulates *Shh* expression in these domains, excluding the larynx<sup>25</sup>. Thus, the laryngeal epithelium appears to be the only tissue where *Shh* expression is regulated solely by MACS1. Intriguingly, previous comparative genome analysis showed that terrestrial vertebrates with lungs have a MACS1 orthologue with high sequence similarity to the mouse MACS1, but such sequence was not explicitly found in the euteleost fishes that have a non-respiratory gas bladder (swim bladder)<sup>24</sup>. These observations suggested that evolutionary change in MACS1 correlated with morphological diversification of lungs and gas bladder, and prompted us to investigate the enhancer activity of MACS1 in the context of morphological evolution of the respiratory organ.

In this study, we first generated an MACS1 KO mouse strain. Phenotyping of the KO embryos clearly showed that MACS1 is an endoderm epithelial enhancer, and that the *Shh* signalling directed by MACS1 is indispensable for development of the larynx including the vocal folds (glottal valve), which is a valve-like laryngeal apparatus located between pharynx and lungs, and

is essential for efficient aerial respiration. Subsequently, we conducted comprehensive phylogenetic analyses of MACS1 and its surrounding genomic sequence for diverse vertebrate taxa. Unexpectedly, the results identified MACS1 orthologues in the cartilaginous fishes, as well as in the lobe-finned fish. Moreover, careful genome comparison revealed that even euteleost fishes have an MACS1-like sequence in the syntenic region. Transgenic assays showed that coelacanth, paddlefish and spotted gar have MACS1 orthologues with enhancer activity in mouse embryos to elicit reporter expression ventrally in the larynx. Meanwhile, teleost fish orthologues have lost the enhancer activity in laryngeal epithelia in both mouse and medaka larvae, implying that the enhancer activity of MACS1 has gradually diverged during evolution. In parallel, we newly identified a sequence that is conserved in the syntenic *rnf32* intron of the ray-finned fishes. Transgenic assays revealed that this sequence in teleost fishes acts as an enhancer to induce reporter expression of dorsal epithelium in mouse oesophagus, which is not driven by MACS1. A medaka transgenic assay confirmed that this CRE induced reporter expression dorsally in the posterior pharynx and oesophagus, from which the non-respiratory gas bladder develops. These findings collectively demonstrate that MACS1, conserved over a wide range of vertebrate taxa, has changed its enhancer activity in the ray-finned fish lineage; furthermore, the ray-finned fishes have evolved a new enhancer in the neighbouring region, which may be implicated in the morphological transition from the ventral lungs to the dorsal non-respiratory gas bladder.

## Results

### Elimination of MACS1 disrupts laryngeal development in mouse.

We eliminated the MACS1 sequence from the mouse genome by embryonic stem (ES) cell targeting (Supplementary Fig. 1), and analysed the KO mouse phenotype. In wild-type mouse embryos, the respiratory organ is composed of pharyngolaryngeal apparatuses surrounded by thyroid cartilage (Fig. 1a–c and j–l). We confirmed that homozygotes of the MFCS4 KO caused morphological defects in the pharyngeal organs including soft palate and epiglottis (Fig. 1d–f and m–o), as we previously reported<sup>24</sup>. All homozygotes of the MACS1 KO succumbed to respiratory problems within 2 days after birth (Supplementary Table 1). The mutation caused severe morphological defects in laryngeal structures, including the arytenoids, vocal folds and thyroid cartilages (Fig. 1g–i and p–r). The partial fistula between the larynx and oesophagus recapitulates the defects observed in a mouse KO mutant lacking the *Shh* coding sequence (yellow arrow in Fig. 1q)<sup>18</sup>. No visible defects were observed in the other *Shh* expression domains regulated by MACS1, including the lungs and the intestinal and urogenital tracts. A comparison of the defects in the MACS1 KO mutant with those in the MFCS4 KO mutant (Supplementary Table 2) clearly showed that MFCS4 and MACS1 are indispensable for the morphogenesis of the upper and lower respiratory structures in the pharynx (Fig. 1d–f and m–o) and larynx (Fig. 1g–i and p–r), respectively. To trace these defects to an earlier developmental stage, we next carried out histological analysis of MACS1 KO embryos at E13.5 and compared them with MFCS4 KO embryos (Supplementary Fig. 2). In the wild type and MFCS4 KO mutant, the arytenoid swelling was properly bifurcated and the laryngotracheal groove formed normally (Supplementary Fig. 2a–d and h–k). By contrast, the arytenoid swelling in the MACS1 KO mutant was hypoplastic, and did not bifurcate; consequently, the laryngotracheal groove failed to form (Supplementary Fig. 2o–r). These abnormalities led to subsequent malformation of the vocal folds, laryngeal cartilages and the septum between the oesophagus and larynx (Fig. 1q). Thus, the *Shh* signalling regulated by MACS1 is essential for

the proper morphogenesis of laryngeal structures including vocal folds.

In a previous transgenic assay, we showed that MACS1 drives reporter expression in the continuous epithelial lining of the endoderm, including the laryngeal epithelium<sup>24</sup>. We now examined whether the elimination of MACS1 from the genome alters the *Shh* expression pattern in mouse embryos at E12.5. At this stage, wild type embryos showed intense *Shh* expression on the ventral side of the pharyngeal and laryngeal epithelia through the laryngotracheal groove in the arytenoid swelling (Fig. 2a,b), whereas in MACS1 KO homozygotes, *Shh* expression in the laryngeal epithelium was abrogated (Fig. 2c,d). Yellow arrow in Fig. 2d depicts the faded *Shh* expression. On the other hand, other MACS1-regulated *Shh* expression domains, namely the lungs and intestinal and urogenital tracts, were unaffected in the MACS1 KO embryos (Supplementary Fig. 3). These results indicate that MACS1 acts as an enhancer to regulate *Shh* expression in the laryngeal epithelium.

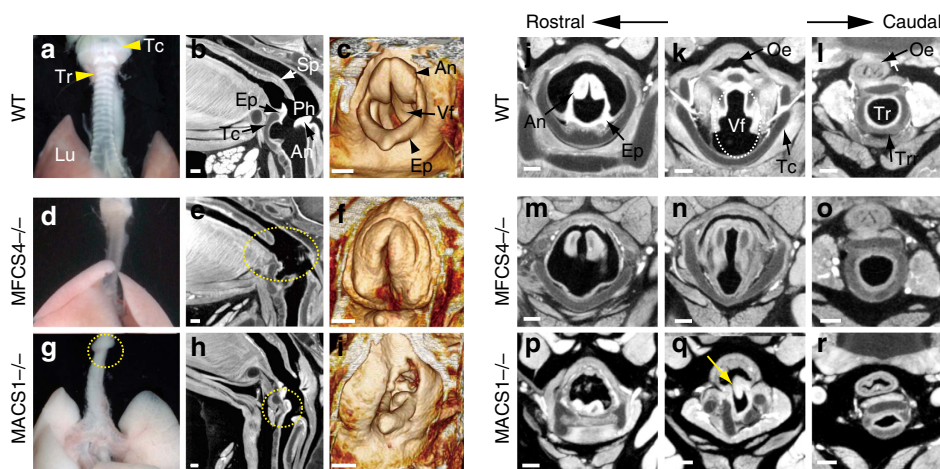
To elucidate the role of Shh signalling in laryngeal development, we examined cell death and cell proliferation in the pharyngeal arches of the MACS1 KO homozygotes. The vertebrate pharyngeal and laryngeal apparatuses, which serve the dual functions of respiration and swallowing, have their embryonic origins in the pharyngeal arches. The laryngeal muscle and cartilage are derivatives of pharyngeal arches 4 and 6 (ref. 30). A TUNEL assay of E11.0 embryos revealed more apoptotic cells in the endodermal epithelia surrounding arch 4 and the surrounding area of the KO homozygotes than in the corresponding regions of the wild type embryos (Fig. 2e–h). A statistically significant difference in the apoptotic cell number was observed between the two groups (Fig. 2i). On the other hand, a BrdU incorporation assay of E11.0 embryos showed no statistically significant difference in cell proliferation around the corresponding areas between the MACS1 KO homozygotes and the wild type embryos (Fig. 2j–l). These results indicated that the loss of MACS1-mediated Shh signalling increases cell death in the epithelia of pharyngeal arch 4-derived organs.

### Fox-binding motif is essential for enhancer activity of MACS1.

We attempted to define a core region and sequence motif in mouse MACS1 that are responsible for the enhancer activity. Comparison of the *rnf32* intronic sequence between human and mouse revealed that an 807-bp fragment of mouse MACS1 contains two stretches of sequence that are highly conserved in the two species (Fig. 3a). To identify the core region, we made transgenic constructs with serial deletions in this 807-bp or a larger 1,264-bp fragment (Fig. 3a), and assayed their ability to drive *LacZ* reporter expression in mouse embryos (Supplementary Fig. 4). Results from the assays with constructs c1 and c2 indicated that the region designated Block-p contains a regulatory sequence necessary for expression in the gut, while those with constructs c2 and c3 indicated that Block-m contains a regulatory sequence necessary for expression in the lungs. Assays with constructs c3, c4 and c5 revealed that a sequence necessary for expression in the laryngeal epithelium was confined to a 29-bp segment (coloured in green in Fig. 3a; see also Supplementary Fig. 4). We found the binding core motif of the Forkhead box (Fox) proteins in this 29-bp segment (Fig. 3a and Supplementary Fig. 5). Finally, to pinpoint the core regulatory sequence required for expression in the laryngeal epithelium, we made transgenic constructs with serial deletions within the 29-bp segment of the 807-bp fragment (Fig. 3a). The transgenic reporter assay with c6–c10 showed that deleting the Fox-binding core motif and its flanking nucleotides caused abrogation of the *LacZ* reporter expression in the laryngeal epithelia (Fig. 3a,b Supplementary Fig. 4). The indispensability of the Fox-binding motif was confirmed in a transgenic assay using c11 that harboured three base substitutions within the core sequence of 7 bp (Fig. 3a,b Supplementary Fig. 4).

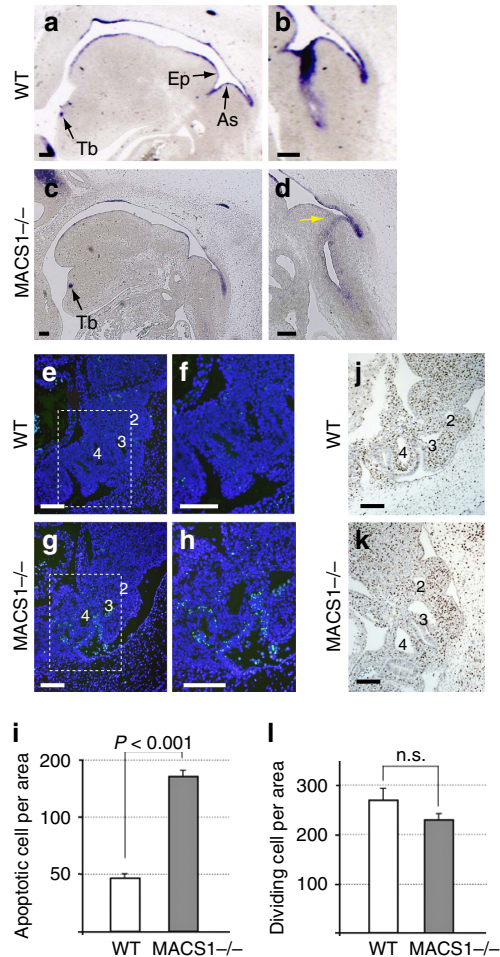
### Non-coding sequences are conserved in the *rnf32* intron.

Our earlier study<sup>24</sup> showed that MACS1 is localized in intron 8 of the mouse *Rnf32* gene, and orthologues have been identified in the syntenic region of all examined tetrapods. In this study,



**Figure 1 | MACS1 is indispensable for morphogenesis of the respiratory apparatus in mouse larynx.** Phenotypes of the respiratory apparatus in wild type (WT) (a–c and j–l), MFCS4<sup>-/-</sup> (d–f and m–o) and MACS1<sup>-/-</sup> (g–i and p–r) at E18.5. Deformed structures are encircled in yellow dashed lines. External appearance of the respiratory organs in WT (a), MFCS4<sup>-/-</sup> (d) and MACS1<sup>-/-</sup> (g). The thyroid cartilage is deformed in MACS1<sup>-/-</sup> (g). X-ray micro-CT sagittal image of WT (b), MFCS4<sup>-/-</sup> (e) and MACS1<sup>-/-</sup> (h). Morphological defects are apparent in the pharyngeal apparatus (soft palate and epiglottis) in MFCS4<sup>-/-</sup> (circled with yellow dashed line in e) and the laryngeal apparatus (arytenoids) in MACS1<sup>-/-</sup> (circled with yellow dashed line in h). 3D X-ray micro-CT images (c,f,i). WT develops arytenoids and epiglottis over the vocal folds (c). The epiglottis in MFCS4<sup>-/-</sup> is truncated (f), and the arytenoids and vocal folds in MACS1<sup>-/-</sup> (i) are deformed. X-ray micro-CT transverse images of the laryngeal apparatus (j–r). Arytenoids, epiglottis, vocal folds (circled with white dashed line in k) and thyroid cartilage are well developed in WT (j–l). The epiglottis in MFCS4<sup>-/-</sup> is truncated (m), and the thyroid cartilage and vocal fold are disrupted in MACS1<sup>-/-</sup> (q). Scale bars, 200  $\mu$ m. An, arytenoids; As, arytenoid swelling; Ep, epiglottis; Lu, lung; Oe, oesophagus; Ph, pharynx; Sp, soft palate; Tc, thyroid cartilage; Tr, trachea; Trr, tracheal ring; Vf, vocal folds.

we expanded phylogenetic analysis to representatives of diverse vertebrate taxa including cartilaginous fishes, coelacanth, paddlefish and spotted gar, which diverged more deeply in the ray-finned fish phylogeny and are referred to as non-teleost ray-finned fishes, and teleost fishes including euteleosts. The analysed species are listed in Methods, and the VISTA plots for the intronic sequences of the *rnf32* gene are summarized in Fig. 4. These results revealed that coelacanth, paddlefish, spotted gar, golden arowana and Japanese eel all have an MACS1 orthologue in the syntenic region of their *rnf32* intron (Fig. 4a). Moreover,



**Figure 2 | MACS1 is essential for *Shh* expression and cell survival in the laryngeal epithelia.** Endogenous *Shh* expression in sagittal sections of the pharyngeal and laryngeal epithelia at E12.5 (a–d). (b,d) are magnified images of the laryngotracheal grooves of the WT and *MACS1*<sup>-/-</sup>, respectively. *Shh* expression is detected in the pharyngeal and laryngeal epithelia in WT (a,b). On the other hand, *Shh* expression was abrogated from epithelium of the arytenoid swelling in *MACS1*<sup>-/-</sup> (yellow arrow in d). In the TUNEL assay at E11.0, apoptotic cells were labeled with green FITC (e–h). Apoptotic cells are scarcely detected in WT (e,f). In *MACS1*<sup>-/-</sup>, many apoptotic cells are detected in the epithelium surrounding pharyngeal arch 4 (g,h). (f,h) represent magnified images of the area surrounded with dashed lines in e,g, respectively. Numerals in e,g depict the pharyngeal arch arteries. Significant difference ( $P < 0.001$ ) is observed between WT and *MACS1*<sup>-/-</sup> (i). Data are presented as mean  $\pm$  s.d.,  $n = 3$ . BrdU incorporation assay showed that cell proliferation in the pharyngeal arch of the *MACS1*<sup>-/-</sup> is similar to that of WT (j,k). Numerals in (j,k) depict the pharyngeal arch arteries. No difference is observed between WT and *MACS1*<sup>-/-</sup> (l). Data are presented as mean  $\pm$  s.d.,  $n = 3$ . Scale bars, 200  $\mu$ m. Tb, tooth bud.

even elephant shark and skate also have an MACS1 orthologue<sup>31</sup>. Thus, the origin of MACS1 appears to predate the divergence of cartilaginous fishes and bony fishes. To close the gap between tetrapods and euteleost fishes, we carried out Southern blot analysis with genomic DNAs of an expanded panel of fish using the spotted gar MACS1 sequence as the probe (Supplementary Fig. 6). The blot showed that MACS1 is also conserved in Siberian sturgeon and white sturgeon, bowfin, silver arowana and Japanese eel. Genomic DNAs of catfish, goldfish and carp yielded very faint bands (arrows in Supplementary Fig. 6), whereas zebrafish, salmon, trout, stickleback and medaka did not yield any band.

As shown in the VISTA plot using the mouse MACS1 sequence as a reference, no apparent MACS1 orthologue was found in euteleost fishes (Fig. 4a). However, euteleost fishes are known to express *shh* in epithelia of the posterior pharynx, gut and gas bladder<sup>32</sup>, suggesting that they have alternative enhancer(s) to regulate *shh* expression in the epithelia of these organs. In the VISTA analysis using the medaka *rnf32* intronic sequence as a reference, we found that three euteleost fishes, medaka, stickleback and Nile tilapia, have two conserved sequence blocks in the syntenic region of the *rnf32* intron, and we named them Block-1 and Block-2 (Fig. 4b). We investigated how the Block-1 sequence diverged during the evolution of ray-finned fishes. To do this, we conducted a VISTA analysis, using the spotted gar *rnf32* partial sequence as a reference, from which it appeared that an MACS1-like sequence was present in salmon at the syntenic region in the *rnf32* intron (Fig. 4c). We further compared a 77-bp sequence that is conserved in salmon, spotted gar and medaka, and found that the overall nucleotide identity among the three sequences is 66% (51/77) (Supplementary Fig. 7). Moreover, the salmon MACS1-like sequence displays 75% (58/77) identity to the spotted gar MACS1, and 78% (60/77) identity to the medaka Block-1. Thus, salmon bridged the Block-1 sequence of the euteleost fishes and the MACS1 orthologues of the remaining taxa. This result indicated that Block-1 is a diverged form of MACS1. This is consistent with the results of Southern blot analysis.

Next, we explored the origin of Block-2 in ray-finned fish evolution. VISTA analysis using the salmon *rnf32* intronic sequence as a reference revealed that the core sequence of Block-2 resides in the syntenic region of the *rnf32* intron of spotted gar, golden arowana and Japanese eel, as well as three euteleost fishes (Fig. 3d). The overall sequence identity of Block-2 among salmon, spotted gar and medaka is 55% (59/108) (Supplementary Fig. 8), and the salmon sequence shows high identity to those of medaka (73%) and spotted gar (73%). Thus, salmon again bridges the Block-2 sequence of the non-teleost ray-finned fishes and that of euteleosts. Since Block-2 was not found in the cartilaginous fishes, paddlefish, coelacanth and tetrapods, it most likely emerged sometime in the non-teleost ray-finned fish lineage.

**Enhancer activity of MACS1 and Block-2 orthologues in mouse.** To examine whether the MACS1 orthologues have enhancer activity, we performed a transgenic assay in mouse embryos, focusing on reporter expression in pharyngeal and laryngeal epithelia (Fig. 5). The MACS1 orthologues of two cartilaginous fishes, skate and elephant shark, which have neither lungs nor a gas bladder, did not induce any reproducible *LacZ* reporter expression in the pharyngeal or laryngeal epithelia, or in the digestive organs (skate, 0 of 39 transgene-positive embryos; elephant shark, 0 of 18 transgene-positive embryos). By contrast, coelacanth and paddlefish MACS1 orthologues induced reporter expression in the laryngeal epithelium and laryngotracheal groove (Fig. 5a). Expression of these reporters was indistinguishable from

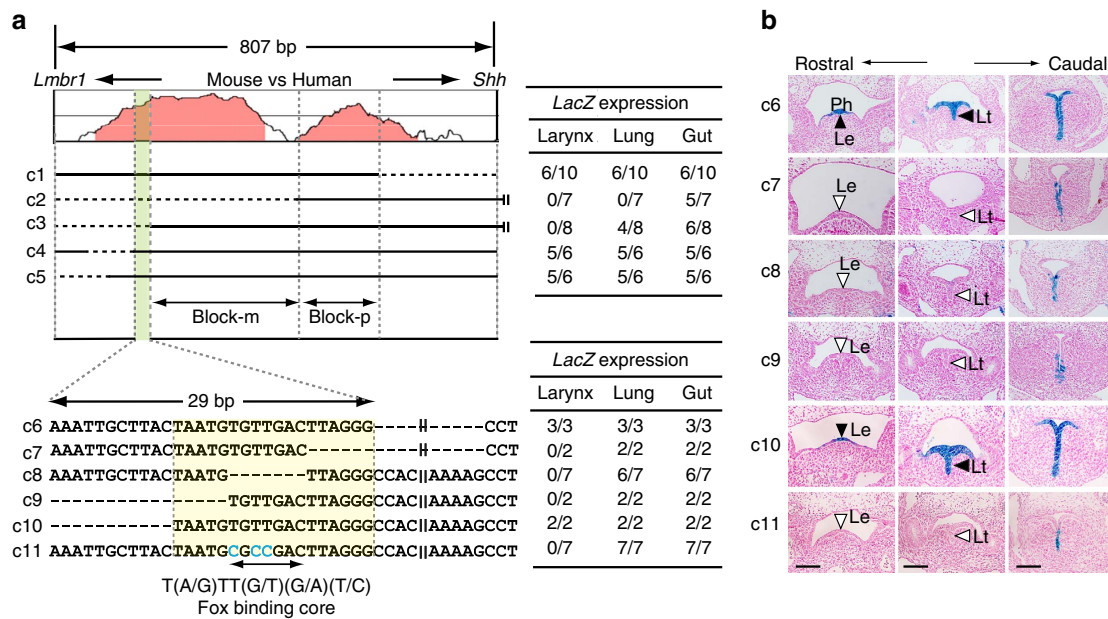
that induced by mouse *MACS1*. The *MACS1* orthologue of spotted gar also induced reporter expression in the laryngeal epithelium and laryngotracheal groove, whereas the *MACS1* orthologue of Japanese eel and medaka did not induce expression in the laryngeal epithelium or laryngotracheal groove, as shown by the white arrowheads in Fig. 5a. These results indicated that although *MACS1* is conserved throughout all vertebrate taxa examined, enhancer activity differs among the orthologues, suggesting that their regulatory functions have diverged. The enhancer activity of the *MACS1* orthologues, which induced laryngeal reporter expression in mouse embryos, is restricted to tetrapods, coelacanth and the non-teleost ray-finned fishes.

We next conducted transgenic assays with *Block-2* orthologues to test their enhancer activity in mouse embryos (Fig. 5b). First, we examined reporter expression induced by the *Block-2* orthologue of spotted gar, and found that it induced reporter expression neither in the pharynx nor in oesophagus. By contrast, the *Block-2* orthologue of Japanese eel and medaka *Block-2* induced intense reporter expression specifically in the dorsal epithelia of the posterior pharynx and the oesophagus (Fig. 5b). It is of interest to note that expression in ventral epithelia of the larynx, which was observed in the assay with *MACS1* and its orthologues (Fig. 5a), was not induced by *Block-2* (Fig. 5b). *Block-2* has two conserved blocks, designated *core1* and *core2* (Supplementary Fig. 9a,b,d,e). Elimination of *core1*, which harbours a Fox-binding core motif (TGTTGAC), from medaka *Block-2* abrogated reporter expression in mouse embryos (Supplementary Fig. 9e). By contrast, elimination of *core2* did

not affect the enhancer activity of *Block-2* (Supplementary Fig. 9e). These results showed that *core1* contains an element that is indispensable for the enhancer activity of *Block-2* in mouse embryos.

**Medaka *Block-2* has enhancer activity in medaka dorsal epithelia.**

To verify endogenous enhancer activity of the *MACS1* orthologue and *Block-2* in euteleost fishes, we conducted transgenic assays using a green fluorescent protein (GFP) reporter in medaka larvae (Fig. 6). To do this, we generated three different reporter constructs with a fragment of the whole medaka *rnf32* intron, the medaka *MACS1* orthologue and medaka *Block-2* (Fig. 6a). We found that the whole intron fragment consisting of the *MACS1* orthologue and *Block-2* induced GFP reporter expression specifically in the posterior pharynx and oesophagus at stages 2–10 dpf (Fig. 6b–f). Expression was also detected in the pneumatic duct, which develops dorsally from the posterior pharynx and connects the oesophagus and gas bladder in the larvae stage (Fig. 6d). A transgenic assay with the *MACS1* orthologue alone induced no detectable expression (0 of 312 injected embryos) (Fig. 6g). By contrast, *Block-2* alone yielded GFP expression (7 of 37 injected embryos) (Fig. 6h), overlapping the expression induced by the whole intron fragment (Fig. 6f). Immunostaining with anti-GFP antibody confirmed that the GFP expression is detected in the epithelium of the oesophagus (Fig. 6i–k). Finally, we examined whether the Fox-binding core motif in *Block-2* is indispensable for enhancer activity in medaka larvae. Elimination



**Figure 3 | A core regulatory sequence is necessary for *LacZ* reporter expression in the laryngeal epithelium.** Two series of stepwise deletion constructs generated for functional dissection of *MACS1*, and results of the *LacZ* reporter assay (**a**). Upper left: a series of the first deletion constructs in the 807-bp DNA fragment that is sufficient for enhancer activity of *MACS1* (c1, c4, c5), or in a longer 1,264-bp fragment (c2, c3). Dashed lines mark the deleted segment in each construct. Upper right: result of the transgenic assays for the deletion constructs using mouse embryos. For each construct, the fraction of embryos exhibiting reproducible reporter expression in the larynx, lung and gut relative to the total number of transgene-positive embryos at E11.5 is shown. The result indicated that the green coloured area of 29-bp is a critical region necessary for laryngeal expression. *Block-m* and *Block-p* contain the regulatory sequences necessary for expression in the lung and gut, respectively. Lower left: a second series of deletion and mutant constructs (c6 – c10) in the 807-bp fragment to delineate a core sequence (yellow coloured area) within the highly conserved 29-bp sequence (see also Supplementary Fig. 5). In the construct c11, three bases in the Fox-binding core motif are substituted from T to C (blue letters). Lower right: result of the transgenic assays for the second deletion constructs using mouse embryos. Representative reporter expression patterns in transverse sections of the transgenic embryos (**b**). Labels to the left of each image identify the transgenic constructs in (**a**). Black and white arrowheads depict presence and absence of the reporter expression specifically in the ventral laryngeal epithelium, respectively. Deletions that removed the Fox-binding core motif and its flanking sequences (c7–c9), and base substitutions in the Fox-binding core motif (c11), specifically abolished reporter expression in the laryngeal epithelia. Scale bars, 200 μm. Le, laryngeal epithelia; Lt, laryngotracheal groove.

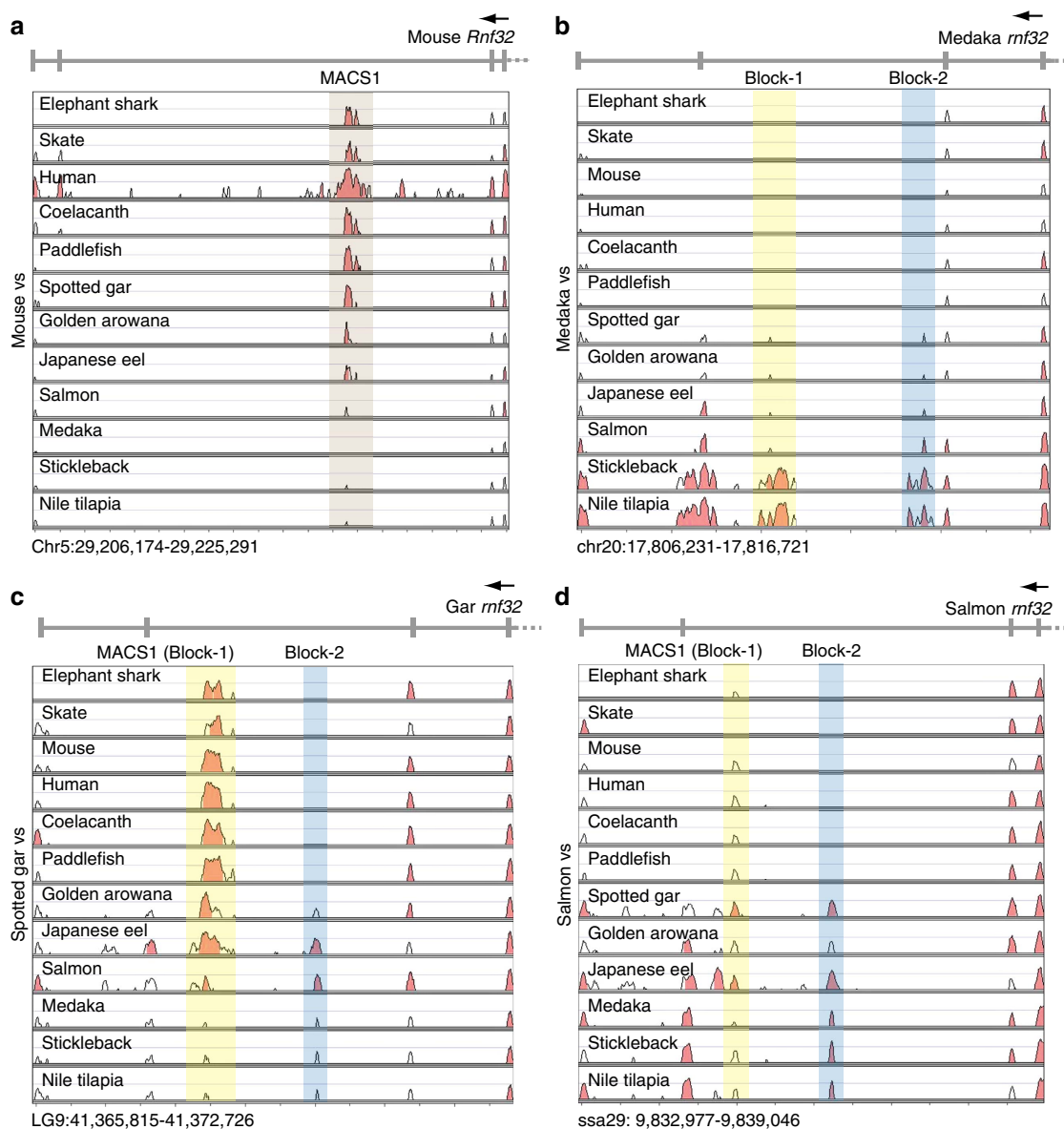
of the Fox-binding core motif (blue arrow in Supplementary Fig. 9e) from Block-2 did not affect GFP reporter expression in medaka larvae (Supplementary Fig. 10), indicating that the Fox-binding core motif is not crucial for the endogenous enhancer activity of Block-2.

## Discussion

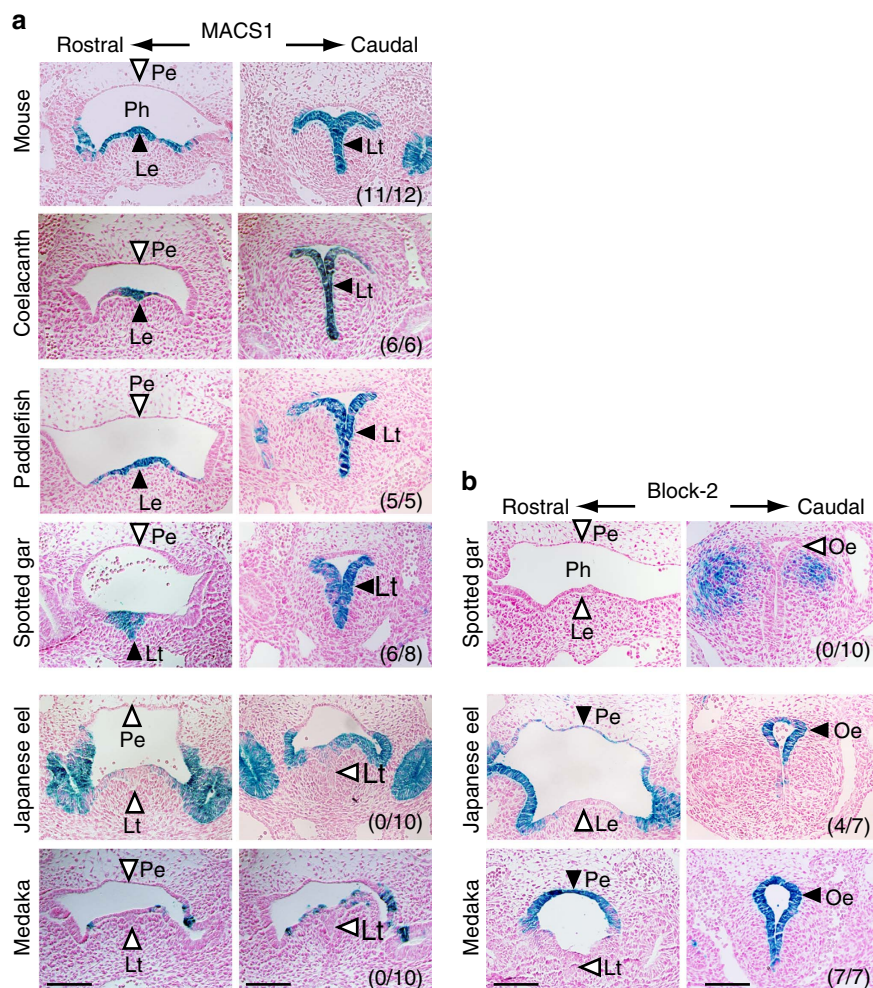
Coordinated movements of the multiple respiratory organs in the pharynx and larynx are essential for respiration and swallowing. Defects in the structures and functions of these organs cause fatal respiratory problems. A mouse KO mutant lacking the *Shh* coding sequence exhibits oesophageal atresia/stenosis, tracheo-oesophageal fistula and tracheal and lung anomalies<sup>17,18,33</sup>. These features are similar to those observed in human foregut defects<sup>18,34</sup>. Since such human malformations are relatively common, occurring in around 1 in 3,500 births<sup>18</sup>, it is important to elucidate the molecular mechanisms underlying these defects. In this study, we have presented evidence that

MACS1 is an enhancer that regulates *Shh* expression in the laryngeal epithelia. A mouse MACS1 KO mutant displayed abrogated *Shh* expression in the laryngeal epithelium, which resulted in developmental defects in the larynx, especially the laryngotracheal groove. This phenotype, associated with laryngeal-oesophageal fistula, recapitulates the defects in the KO mutant lacking the *Shh* coding sequence<sup>18</sup>. The defect in laryngeal development is likely to be caused by elevated apoptosis due to abrogation of Shh signalling in the endodermal epithelia of pharyngeal arch 4 and the surrounding area, from which the laryngeal muscle and cartilage develop. In this regard, it is notable that Shh signalling is known to play an anti-apoptotic role in the proliferation of ameloblastoma and colorectal tumour cells<sup>35,36</sup>.

We previously identified another epithelial lining-specific enhancer, MFCS4, which is conserved from euteleost fishes to mammals and regulates *Shh* expression in the pharyngeal epithelium in mouse embryos<sup>24</sup>. Elimination of this enhancer led to severe truncation of the upper respiratory organs including



**Figure 4 | Phylogenetic analysis of two enhancers in syntenic regions of the *rnf32* intron.** VISTA plots of 13 partial *rnf32* sequences in diverse vertebrate taxa (a–d). The partial *rnf32* sequences of mouse, medaka, spotted gar and salmon were used as the reference genomes. All sequences used for this analysis are summarized in Supplementary Table 4. Brown and blue shades depict MACS1 and Block-2 orthologues, respectively. Yellow shade depicts MACS1 orthologues (Block-1). MACS1 orthologues were identified not only in the cartilaginous fishes, but also coelacanth and the ray-finned fishes (c,d).



**Figure 5 | Enhancer activity of the MACS1 and Block-2 orthologues in mouse embryos.** The *LacZ* reporter expression driven by orthologues of MACS1 (**a**) and Block-2 (**b**) in mouse embryos. Black and white arrowheads depict presence and absence of the reporter expression specifically in the ventral laryngeal epithelium or dorsal esophageal epithelium, respectively. The MACS1 orthologues of coelacanth, paddlefish and spotted gar drove reporter expression in the laryngeal epithelia (Le) and laryngotracheal groove (Lt), which was indistinguishable from that driven by the mouse MACS1 (**a**). They did not induce reporter expression in the dorsal epithelia of pharynx. The MACS1 orthologues of Japanese eel and medaka did not induce reporter expression in the laryngeal epithelia or laryngotracheal groove (**a**). They did not induce reporter expression in the dorsal epithelia of pharynx, although expression was sometimes observed in the ventral epithelia of pharynx (**a**). The Block-2 orthologue of spotted gar did not drive any reporter expression, despite its sequence similarity to those of Japanese eel and euteleosts (**b**). Block-2 orthologue of Japanese eel and medaka Block-2 drove reporter expression in the dorsal epithelia of pharynx and oesophagus (**b**). The number of *LacZ* positive embryos over total number of transgenic embryos is shown in right bottom of each caudal section. Scale bars, 200  $\mu\text{m}$ . Pe, pharyngeal epithelia.

the soft palate and epiglottis, and the animals died soon after birth, probably due to respiration problems. Thus, *Shh* signalling regulated by these two different enhancers is indispensable for morphogenesis of mammalian respiratory organs. Considering their characteristic phenotypes, the mouse MFCS4 and MACS1 KO mutants should be good animal models for human congenital foregut anomalies. MFCS4 and MACS1 partition the *Shh* expression domain in the continuous epithelial lining into two parts, the pharynx and larynx. Their distinct phenotypes—severe truncation of the upper respiratory organs in the pharynx of the MFCS4 KO mutant; severe disruption of the lower respiratory organs in the larynx of the MACS1 KO mutant—reflect well the compartmentalization of a continuous regulatory domain by the independent actions of the two enhancers.

Homozygous MACS1 KO embryos show no visible defects in tissues or organs other than the larynx. This may be due to compensatory functions of additional enhancer(s). Another enhancer, SLGE, which shows no marked sequence conservation

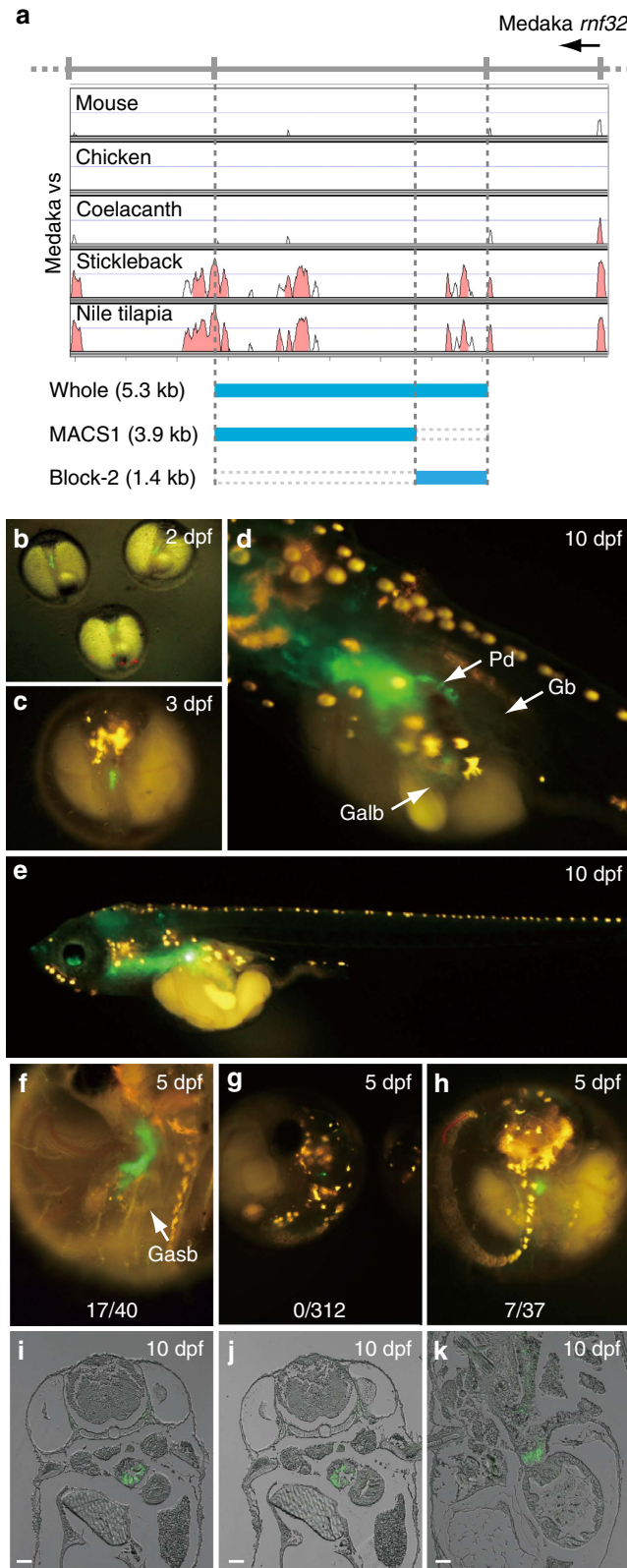
even among mammalian species, regulates *Shh* expression in the epithelia of lungs, guts and urogenital tracts<sup>25</sup>. This redundancy of expression domains for the two enhancers may explain the lack of a visible phenotype in organs other than the larynx in the MACS1 KO embryos<sup>37,38</sup>. It also implies that *Shh* expression in the laryngeal epithelium is solely regulated by MACS1. Our transgenic reporter assay clearly showed that the Fox-binding motif in MACS1 is required for *Shh* regulation in the mouse embryonic larynx. It was reported that depletion of *Foxp4* causes elevated cell death around the laryngotracheal groove, leading to formation of a large cavity in this region<sup>39</sup>. Notably, downregulation of *Shh* was observed specifically in the epithelium around the cavity, but not in the other endodermal epithelia. Considering the laryngeal defects and downregulation of *Shh*, which were commonly observed in the *Foxp4* KO and the MACS1 KO embryos, *Foxp4* is a good candidate of upstream transcription factor for *Shh* regulation through binding to MACS1 in the mouse larynx.

The phylogeny of MACS1 and Block-2 together with the results of transgenic reporter assays and morphological diversity of the posterior pharynx over a wide range of vertebrate taxa are summarized in Fig. 7. At the outset of this study, we inferred that MACS1 emerged in terrestrial animals along with the innovation of a laryngeal apparatus, perhaps including the glottal valve<sup>40</sup>,

which was required for pulmonary respiration and for adaptation to terrestrial life. However, our comprehensive phylogenetic analyses clearly showed that this was not the case. Instead, the evolutionary origin of MACS1 is very old, predating the divergence of the cartilaginous fishes and the bony fishes. Importantly, extant cartilaginous fishes, such as skate and elephant shark, have neither lungs nor a gas bladder. Moreover, conservation of MACS1 is observed throughout diverse vertebrate taxa in the ray-finned fishes and weak conservation was found even in the euteleost fishes, which have only a non-respiratory gas bladder. Thus, MACS1 did not emerge along with the innovation of the laryngeal apparatus, which was prerequisite for terrestrial life.

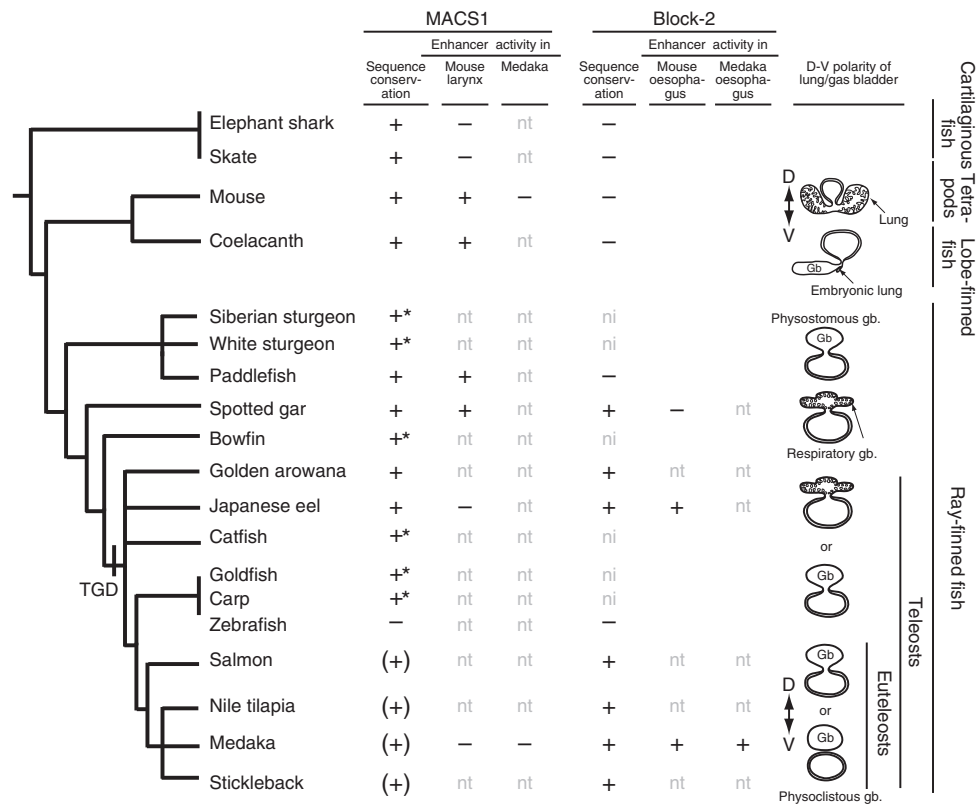
The mouse transgenic assay showed that MACS1 of coelacanth, paddlefish and spotted gar induced reporter expression ventrally in the laryngeal epithelium and laryngo-tracheal groove, where the larynx develops and endogenous *Shh* expression is observed in mouse embryos, and that the expression pattern was indistinguishable from that induced by mouse MACS1. In contrast, the MACS1 orthologues of Japanese eel and medaka did not induce reporter expression in the laryngeal epithelia, although non-reproducible reporter expression was often observed in lateral epithelia of the posterior pharynx. Likewise, the MACS1 orthologues of the cartilaginous fishes did not induce any reporter expression in laryngeal epithelia or other tissues, despite sharing a 29-bp core sequence. A molecular phylogenetic tree shows that the MACS1 sequences of skate and elephant shark and the MACS1 orthologues of Japanese eel and medaka are far diverged from those of tetrapods and coelacanth, although the 29-bp core sequence is partially conserved (Supplementary Figs 7 and 11). This may explain why the MACS1 orthologues of the cartilaginous and teleost fishes could not drive reporter expression in mouse embryos. It is possible that other motifs for as-yet-unknown transcription factors whose binding sites are located outside the 29-bp core sequence are involved in *Shh* regulation in those species. These results indicate that despite the ancient origin and extensive conservation of the MACS1 sequence, the responsiveness of these various MACS1 orthologues to mouse transcriptional factors differs from one to another. Given that the medaka MACS1 has no enhancer activity in the medaka posterior pharynx, we infer that regulatory activity of the MACS1 orthologues has diminished in the euteleost fishes, which might be correlated with the sequence divergence.

Compared with MACS1, Block-2 is less conserved and not found in the cartilaginous fishes and the lobe-finned fishes including tetrapods. Therefore, we infer that it emerged sometime in the non-teleost ray-finned fish lineage. The mouse transgenic



**Figure 6 | Medaka Block-2 has endogenous enhancer activity.** Enhancer activity of three fragments of the Medaka *rnf32* intron was tested by transgenic reporter assay. The tested fragments were 3.9 kb of MACS1, 1.4 kb of Block-2, and the whole intronic fragment encompassing MACS1 and Block-2 (a). GFP reporter expression (green) was monitored in medaka larvae (b–k). The whole intronic fragment drove reporter expression in the posterior pharynx and oesophagus at 2 dpf (b), 3 dpf (c) and 10 dpf (d,e). At 10 dpf, reporter expression was also detected in the dorsal pneumatic duct connecting to the gas bladder (d). Reporter expression driven by the whole intronic fragment (f), MACS1 (g) and Block-2 (h) at 5 dpf. The whole intronic fragment and Block-2 yielded similar signals in the digestive tube (f,h). MACS1 did not drive reproducible expression in medaka larvae (g). Immunohistochemistry for GFP reporter signals driven by the whole intronic fragment at 10 dpf (i–k). Signals were detected in the epithelium of the digestive tube (i–k). The number of larvae exhibiting reproducible reporter expression among the total number of injected eggs is indicated at the bottom in each image (f–h). Scale bars, 500  $\mu$ m. Galb, gall bladder; Gb, gas bladder; Pd, pneumatic duct.





**Figure 7 | Phylogeny of endoderm enhancers along with transition from ventral lungs to dorsal gas bladder.** The figure summarizes phylogeny of the two enhancers, MACS1 orthologues and Block-2, integrating their enhancer activity in mouse and medaka. MACS1 orthologues are present in the cartilaginous fish and bony fish lineages. Block-2 emerged in the non-teleost ray-finned fish lineage, and is conserved in the teleost fishes including euteleost fishes. Spotted gar and teleost fishes have both MACS1 and Block-2 in the same *mf32* intron. Exceptionally, both enhancers are absent in zebrafish. Notably, the transgenic reporter assay indicated that MACS1 induced reporter expression ventrally in the laryngeal epithelium and/or the laryngotracheal groove in mouse embryos, but not in medaka larvae, whereas Block-2 induced expression in dorsal epithelia of the pharynx and oesophagus in both mouse embryos and medaka larvae. This dorsoventral axis polarity is relevant to the position where the lungs and gas bladder evaginate from the posterior pharynx. Asterisks mark species analysed only by Southern blotting. Highly diverged MACS1 orthologues are noted in parentheses. Physostomous fishes retain pneumatic duct until adulthood, and many of them have lost respiratory function in the gas bladder<sup>41</sup>. Euteleost fishes including medaka, stickleback, Nile tilapia have physoclistous gas bladder, which is completely separated from the foregut in adulthood<sup>70</sup>. ni, no information; nt, not tested.

assay revealed that Block-2 of Japanese eel and medaka induced reporter expression in dorsal epithelia of the posterior pharynx and esophagus, but not ventrally in developing laryngeal epithelium, in mouse embryos. Furthermore, the medaka transgenic assay showed that the medaka Block-2 induced reporter expression in the pneumatic duct, which is formed dorsally from the posterior pharynx. The dorsoventral axis polarity of the reporter expression driven by MACS1 and Block-2 corresponds to the position where lungs and the non-respiratory gas bladder develop in the posterior pharynx. Indeed, positional difference has been highlighted by the fact that lungs are ventrally evaginated from the posterior pharynx, whereas the gas bladder is evaginated dorsally<sup>41</sup>.

The evolutionary relationship between ventral paired lungs and the dorsal gas bladder (swim bladder) of most ray-finned fishes has long been debated<sup>41-44</sup>. For instance, signalling pathways of many developmental genes (*Shh*, *Fgf*, *Wnt*, *Nkx*, *Fox*) are commonly involved in the development of lungs and gas bladders<sup>32,45-47</sup>. Both organs develop from an out-pocketing of the posterior pharynx<sup>48</sup>, having similar histology<sup>49</sup>, and produce similar surfactant proteins<sup>50</sup>. An anatomical study showed that there is no clear criterion for structural distinction between lungs and respiratory gas bladder of spotted gar<sup>51</sup>. More recently, ancestral reconstructions supported the homology of lungs and

gas bladders due to their shared vasculature supply<sup>52</sup>. The presence of a well-developed and potentially functional lung was reported in the early embryonic stage of coelacanth<sup>53</sup>. Since the parallel development of a fatty organ for buoyancy control was found in the embryos, coelacanth may have this lung transiently in addition to the gas bladder. Another study of a non-teleost ray-finned fish, the bichir, which has ventral lungs, showed that its lung development resembles that of tetrapods from the aspects of histology and the expression patterns of genes that play key roles in the lung development<sup>54</sup>. This supports the notion that the lungs emerged in the common ancestor of the lobe-finned fish and the ray-finned fishes. The lungs and respiratory gas bladder with a glottal valve are observed in certain extant ray-finned fishes, including bichir, bowfin and spotted gar<sup>55,56</sup>. Considering the results and findings of these studies, the lung and the gas bladder are most likely homologous organs, and the gas bladder can be regarded as an evolutionary modification of lung. In this context, we infer that Block-2 emerged in the non-teleost ray-finned fish lineage along with innovation of the non-respiratory gas bladder.

Functional differentiation of MACS1, acquisition of the extant enhancer activity of Block-2, and the corresponding changes of transcriptional factor(s) probably postdated teleost genome duplication (TGD), which occurred after the teleost fishes

branched off from the non-teleost ray-finned fishes (Fig. 7). MACS1 of spotted gar and paddlefish induced reporter expression in the ventral epithelia of the mouse larynx, although the gas bladder in these species is positioned dorsally in the posterior pharynx. In ventral lung development, *Shh* signalling is necessary for lung growth rather than for budding of lungs from foregut<sup>19</sup>. Therefore, a regulatory signalling other than *Shh* is required for the ventral development, and change of the relevant signalling might predate the loss of MACS1 enhancer activity in the ray-finned fish lineage. This may interpret for the ventral enhancer activity of the spotted gar MACS1 in the transgenic mouse embryos. Alternatively, it is also possible that the spotted gar MACS1 induced the ectopic ventral expression of the transgenic reporter in mouse embryos, due to incompatibility between the MACS1 *cis*-element(s) of the spotted gar and xenotropic transcriptional environment of mouse. It is also noted that Block-2 of spotted gar has no regulatory activity dorsally in the oesophagus. All together, transition from the MACS1-directed ventral *shh* regulation to Block-2-directed dorsal *shh* regulation might be established in the teleost lineage after TGD, but not in the non-teleost ray-finned fishes. This indicates that there is a gap between emergence of the Block-2 sequence and acquisition of the enhancer activity, implying that the non-teleost ray-finned fishes have transitional status with regard to *shh* regulation in the posterior pharynx. Indeed, a recent report clearly showed that spotted gar has a unique position in evolution, bridging teleosts to tetrapods, and genome of spotted gar provides connectivity of vertebrate regulatory elements between teleosts and tetrapods<sup>57</sup>.

Many lines of evidence indicate that genome differentiation was accelerated in the teleost fish lineage by the TGD<sup>58–60</sup>. The ray-finned fishes comprise over 28,000 species, nearly half of the total number of extant vertebrate species<sup>61</sup>. Euteleost is the crown group of the ray-finned fishes, comprising 17,000 species<sup>62</sup>. In this group, the gas bladder primarily functions as a hydrostatic organ to maintain neutral buoyancy. Dorsal position of gas bladder stabilizes posture, because the centre of fish mass is below the gas bladder, and it therefore contributes to the ability of fish to remain at their current water depth without expending energy on swimming. Although it is uncertain how a change in *Shh* regulation in the posterior pharynx influenced the innovation of a non-respiratory gas bladder in the ray-finned fish lineage, the relevance of the dorsoventral axis polarity of *Shh* regulation to this innovation will be an important issue for consideration in future studies. Intriguingly, the Fox-binding core motif is dispensable for the enhancer activity of Block-2 in medaka larvae. The change in *Shh* regulation, which allowed switching of the dorsoventral axis polarity of gene expression, may have been accompanied by changes in transcription factor(s).

## Methods

**Mice.** C57BL/6, ICR and (C57BL/6 × DBA/2)F<sub>1</sub> mice for transgenic reporter assays were purchased from Japan Clea (Tokyo, Japan). The KO mutant strains are maintained at the National Institute of Genetics (NIG), Mishima, Japan. Animal experiments in this study were approved by the Animal Care and Use Committee of NIG and the Animal Care and Use Committee of RIKEN Kobe Branch.

**DNAs.** A piece of frozen muscle of coelacanth (*Latimeria chalumnae*) was obtained from the Foundation for Advancement of International Science (Tsukuba, Japan). Zebrafish (*Danio rerio*), stickleback (*Gasterosteus aculeatus*) and medaka (*Orizias latipes*) were obtained from NIG (National Institute of Genetics) and NIBB (National Institute for Basic Biology) of Japan, respectively. Skate (*Rajidae sp.*), paddlefish (*Polyodon spathula*), Siberian sturgeon (*Acipenser baerii*), white sturgeon (*Acipenser transmontanus*), spotted gar (*Lepisosteus oculatus*), bowfin (*Amia calva*), silver arowana (*Osteoglossum bicirrhosum*), Japanese eel (*Anguilla japonica*), catfish (*Silurus asotus*), goldfish (*Carassius auratus*), carp (*Cyprinus carpio*), silver salmon (*Oncorhynchus kisutch*) and rainbow trout (*Oncorhynchus mykiss*) were purchased from fish shops. Genomic DNA was extracted from fins and muscles with the standard phenol-chloroform method. An 8.5-kb partial sequence of skate *rnf32* and a 4.3-kb partial sequence of

paddlefish *rnf32* were determined on the basis of the conserved sequence in the *rnf32* gene. The sequences were ascribed as LC009631 and LC128331, respectively, by DNA Data Bank of Japan (DDBJ). The 1.3-kb MACS1 sequence of elephant shark (*Callorhynchus milii*) was synthesized by GenScript Japan.

**Comparative analysis.** Multiple sequences of the analysed vertebrates were retrieved from the UCSC Genome Browser (<http://genome.ucsc.edu/>), Ensembl (<http://asia.ensembl.org/index.html>), little skate database (<http://skatebase.org/skate>), salmon genome blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), NAGRP Blast Center (<http://www.animalgenome.org/blast/>), A. iaponika genome assembly (<http://www.zfgenomics.com/sub/eel>)<sup>63</sup> and Genbank (<http://www.ncbi.nlm.nih.gov/genbank/>). For comparative analysis, the ClustalW system (<http://clustalw.ddbj.nig.ac.jp/>), VISTA program (<http://genome.lbl.gov/vista/mvista/submit.shtml>) and MEGA7.0.18 were applied.

**Generation of the MACS1 knockout mouse.** The basic targeting vector was constructed by inserting pKO Neo and pKO DT sequences into the pKO Scrambler V901 vector<sup>24</sup> (Supplementary Fig. 1b). The neomycin gene in the pKO Neo cassette was replaced with the loxP-neo-loxP sequence. The long and short arm fragments were amplified from RP23-284A9 BAC DNA. All primer pairs are listed in Supplementary Table 3. A 1,554-bp fragment (chr5: 29,211,522–29,213,075 mm10) including mouse MACS1 was replaced with the loxP-neo-loxP cassette by homologous recombination in TT2 ES cells<sup>64</sup>. Correct homologous recombination was confirmed by Southern blot analysis (Supplementary Fig. 1c). Germline transmission of chimera mice was affirmed by cross mating with C57BL/6 mice. The neomycin cassette in the mutant allele was removed by crossing with strain B6-Tg (CAG-Cre)<sup>65</sup>, and genotyping was carried out by PCR (Supplementary Fig. 1d). The established MACS1 KO mouse strain is maintained as the heterozygote with C57BL/6 strain (No. CDB0694K in <http://www.cdb.riken.jp/arg/mutant%20mice%20list.html>).

**X-ray micro-CT analysis.** X-ray micro-CT analysis was carried out as previously described<sup>66</sup>. Briefly, mouse embryos were scanned using X-ray micro-CT (ScanXmate-E090S) (Comscan Techno, Tokyo, Japan) at a tube voltage peak of 60 kVp and a tube current of 130 μA. Samples were rotated 360° in steps of 0.18°, generating 2,000 projection images of 992 × 992 pixels. The micro-CT data were reconstructed at an isotropic resolution of 5.3 × 5.3 × 5.3 μm. Before scanning, embryos were soaked in contrast agent, a 1:3 mixture of Lugol's solution and deionized distilled water. Three-dimensional tomographic images were obtained using the OsiriX program ([www.osirix-viewer.com](http://www.osirix-viewer.com)).

**Section in situ hybridization.** Embryos were fixed with 4% paraformaldehyde, and paraffin sections (8 μm thickness) were processed for RNA *in situ* hybridization using a digoxigenin-labelled *Shh* riboprobe<sup>12</sup>; the *Shh* riboprobe was gifted from Dr A. McMahon. Section *in situ* hybridization was carried out using standard methods.

**Transgenic reporter assay.** Mouse and medaka genomic DNA fragments for the transgenic assay were amplified from RP23-284A9 (mouse) and Ola1-012F13 (medaka) BAC DNAs. The other fragments were amplified from genomic DNA. Primer pairs for amplification of MACS1 orthologues are listed in Supplementary Table 3. Genomic DNA fragments were subcloned into the reporter cassettes hsp promoter-*LacZ*<sup>67</sup>, zebrafish *shh* promoter-*LacZ* and zebrafish *shh* promoter-*EGFP*<sup>68</sup>. Transgenic mice were generated by the pronuclear microinjection procedure. For *LacZ* staining, mouse embryos were fixed in 2% formaldehyde and 0.2% glutaraldehyde containing 0.2% Nonidet P-40 for 1 h at 4 °C, and then washed with PBS several times. Staining was performed in a solution containing 0.5 mg ml<sup>-1</sup> X-gal, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2 mM MgCl<sub>2</sub> and 0.2% Nonidet P-40 in PBS at 37 °C overnight with shaking. For histological analysis of transgenic embryos, embryos were fixed overnight in 4% paraformaldehyde, dehydrated in an ethanol series and embedded in paraffin. Sections were cut at 5 μm and counterstained with acidic Fast Red<sup>24</sup>. A transgenic assay in medaka followed a previously described method<sup>69</sup>. Briefly, medaka strain d-rR was maintained under controlled lighting conditions (14 h light and 10 h dark) at 26 °C. Plasmids (25 ng μl<sup>-1</sup>) were injected into the cytoplasm of fertilized eggs before the first cleavage. The injected eggs were incubated at 28 °C, and signal-positive eggs were raised to adulthood to obtain germline-transformed individuals.

**Analysis of cell death and cell proliferation.** For detection of apoptotic cells, a TUNEL assay was carried out using an In Situ Cell Death Detection Kit, POD (Roche). FITC signals were detected in the apoptotic cells at E11.0. For analysis of cell proliferation, 5-bromodeoxyuridine (BrdU, Sigma B5002) was intraperitoneally injected into pregnant females (40 mg kg<sup>-1</sup>) at E11.0. Two hours later, embryos were collected and fixed with 4% paraformaldehyde in PBS and embedded in paraffin. Immunostaining was carried out as described elsewhere. Briefly, the deparaffinized sections (5 μm thickness) were treated with 0.3% H<sub>2</sub>O<sub>2</sub>, 2 N HCl,

trypsin and reacted with monoclonal mouse BrdU antibody (Sigma clone B8434) and biotinylated mouse anti-IgG (Vector M.O.M. peroxidase Kit PK-2200). For colouring, a DAB tablet (Wako 049-22831) was used.

**Southern blotting.** Fish fin or muscle was digested with proteinase K and genomic DNA was extracted with phenol-chloroform. Genomic DNA (10 µg) was double-digested with EcoRI and PstI. After electrophoresis and blotting onto Hybond N+, the genome fragments were hybridized with the Dig-labelled MACS1 sequence of spotted gar and the blot was washed successively in low-stringency buffer (0.1% SDS, 2XSSC) at room temperature and high-stringency buffer (0.1% SDS, 0.5XSSC) at 55 °C. For signal detection, a non-radioisotope system was used (Roche).

**Immunohistochemistry of medaka larvae.** Medaka larvae at 10 dpf were fixed in 4% paraformaldehyde in 0.85X PBS at 4 °C overnight. After fixation, the specimens were embedded in paraffin and sectioned serially at 5 µm. The deparaffinized sections were treated with anti-GFP antibody (1:500, Invitrogen, A11122) at room temperature for 3 h. After several washes with PBS, sections were incubated with Alexa 488-conjugated anti-rabbit IgG secondary antibody (1:500, Invitrogen, A21206) for 1 h. The sections were counterstained with DAPI.

**Quantification of apoptotic cells and proliferating cells.** The number of apoptotic and proliferating cells was counted and normalized to an area around the pharyngeal arch 4, a square with one side of 500 µ. Statistical analyses were performed using the two-tailed Student's *t*-test with Excel (Microsoft). A *P* value of <0.05 was considered statistically significant.

**Data availability.** Sequences of skate and paddlefish are ascribed as LC009631 and LC128331 respectively, by DDBJ. All relevant data are available from the authors.

## References

- Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69 (2012).
- Wray, G. A. The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* **8**, 206–216 (2007).
- Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
- Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A. & Bejerano, G. Enhancers: five essential questions. *Nat. Rev. Genet.* **14**, 288–295 (2013).
- Montavon, T. & Duboule, D. Landscapes and archipelagos: spatial organization of gene regulation in vertebrates. *Trends Cell Biol.* **22**, 347–354 (2012).
- Sagai, T., Hosoya, M., Mizushima, Y., Tamura, M. & Shiroishi, T. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797–803 (2005).
- Cretekos, C. J. *et al.* Regulatory divergence modifies limb length between mammals. *Genes Dev.* **22**, 141–151 (2008).
- Chan, Y. F. *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science (New York, N.Y.)* **327**, 302–305 (2010).
- Prabhakar, S. *et al.* Human-specific gain of function in a developmental enhancer. *Science (New York, N.Y.)* **321**, 1346–1350 (2008).
- McLean, C. Y. *et al.* Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* **471**, 216–219 (2011).
- Frankel, N. *et al.* Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* **474**, 598–603 (2011).
- Echelard, Y. *et al.* Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417–1430 (1993).
- Hammerschmidt, M., Brook, A. & McMahon, A. P. The world according to hedgehog. *TIG* **13**, 14–21 (1997).
- Ingham, P. W. & McMahon, A. P. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* **15**, 3059–3087 (2001).
- McMahon, A. P., Ingham, P. W. & Tabin, C. J. Developmental roles and clinical significance of hedgehog signaling. *Curr. Top. Dev. Biol.* **53**, 1–114 (2003).
- Varjosalo, M. & Taipale, J. Hedgehog: functions and mechanisms. *Genes Dev.* **22**, 2454–2472 (2008).
- Chiang, C. *et al.* Cyclopia and defective axial patterning in mice lacking sonic hedgehog gene function. *Nature* **383**, 407–413 (1996).
- Litingtung, Y., Lei, L., Westphal, H. & Chiang, C. Sonic hedgehog is essential to foregut development. *Nat. Genet.* **20**, 58–61 (1998).
- Miller, L. A. *et al.* Role of sonic hedgehog in patterning of tracheal-bronchial cartilage and the peripheral lung. *Dev. Dyn.* **231**, 57–71 (2004).
- Mao, J., Kim, B. M., Rajurkar, M., Shivdasani, R. A. & McMahon, A. P. Hedgehog signaling controls mesenchymal growth in the developing mammalian digestive tract. *Development* **137**, 1721–1729 (2010).
- Lettice, L. A. *et al.* A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735 (2003).
- Sagai, T. *et al.* Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mamm. Genome* **15**, 23–34 (2004).
- Jeong, Y., El-Jack, K., Roessler, E., Muenke, M. & Epstein, D. J. A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* **133**, 761–762 (2006).
- Sagai, T. *et al.* A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. *Development* **136**, 1665–1674 (2009).
- Tsukiji, N., Amano, T. & Shiroishi, T. A novel regulatory element for Shh expression in the lung and gut of mouse embryos. *Mech. Dev.* **131**, 127–136 (2014).
- Yao, Y. *et al.* Cis-regulatory architecture of a brain signaling center predates the origin of chordates. *Nat. Genet.* **48**, 575–580 (2016).
- Amano, T. *et al.* Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47–57 (2009).
- Lettice, L. A. *et al.* Opposing functions of the ETS factor family define Shh spatial expression in limb buds and underlie polydactyly. *Dev. Cell* **22**, 459–467 (2012).
- Anderson, E., Devenney, P. S., Hill, R. E. & Lettice, L. A. Mapping the Shh long-range regulatory domain. *Development* **141**, 3934–3943 (2014).
- Frisdal, A. & Trainor, P. A. Development and evolution of the pharyngeal apparatus. *Wiley Interdiscip. Rev. Dev. Biol.* **3**, 403–418 (2014).
- Venkatesh, B. *et al.* Ancient noncoding elements conserved in the human genome. *Science (New York, N.Y.)* **314**, 1892 (2006).
- Winata, C. L. *et al.* Development of zebrafish swimbladder: the requirement of hedgehog signaling in specification and organization of the three tissue layers. *Dev. Biol.* **331**, 222–236 (2009).
- Pepicelli, C. V., Lewis, P. M. & McMahon, A. P. Sonic hedgehog regulates branching morphogenesis in the mammalian lung. *Curr. Biol.* **8**, 1083–1086 (1998).
- Spitz, L. Oesophageal atresia. *Orphanet J. Rare Dis.* **2**, 24 (2007).
- Qualtrough, D., Buda, A., Gaffield, W., Williams, A. C. & Paraskeva, C. Hedgehog signalling in colorectal tumour cells: induction of apoptosis with cyclopamine treatment. *Int. J. Cancer* **110**, 831–837 (2004).
- Kanda, S. *et al.* Anti-apoptotic role of the sonic hedgehog signaling pathway in the proliferation of ameloblastoma. *Int. J. Oncol.* **43**, 695–702 (2013).
- Hong, J. W., Hendrix, D. A. & Levine, M. S. Shadow enhancers as a source of evolutionary novelty. *Science (New York, N.Y.)* **321**, 1314 (2008).
- Frankel, N. *et al.* Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493 (2010).
- Li, S., Zhou, D., Lu, M. M. & Morrisey, E. E. Advanced cardiac morphogenesis does not require heart tube fusion. *Science (New York, N.Y.)* **305**, 1619–1622 (2004).
- Graham, J. B. *et al.* Spiracular air breathing in polypterid fishes and its implications for aerial respiration in stem tetrapods. *Nat. Commun.* **5**, 3022 (2014).
- Perry, S. F., Wilson, R. J., Straus, C., Harris, M. B. & Remmers, J. E. Which came first, the lung or the breath? *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 37–47 (2001).
- Darwin, C. The origin of species by means of natural selection. (London: Murray, 1859).
- Graham, J. B. An evolutionary perspective for bimodal respiration: A biological synthesis of fish air-breathing. *Am. Zool.* **34**, 229–237 (1994).
- Graham, J. B. & Lee, H. J. Breathing air in air: in what ways might extant amphibious fish biology relate to prevailing concepts about early tetrapods, the evolution of vertebrate air breathing, and the vertebrate land transition? *Physiol. Biochem. Zool.* **77**, 720–731 (2004).
- Korzh, S. *et al.* The interaction of epithelial Ihha and mesenchymal Fgf10 in zebrafish esophageal and swimbladder development. *Dev. Biol.* **359**, 262–276 (2011).
- Yin, A., Korzh, S., Winata, C. L., Korzh, V. & Gong, Z. Wnt signaling is required for early development of zebrafish swimbladder. *PLoS ONE* **6**, e18431 (2011).
- Cass, A. N., Servetnick, M. D. & McCune, A. R. Expression of a lung developmental cassette in the adult and developing zebrafish swimbladder. *Evol. Dev.* **15**, 119–132 (2013).
- Edwin, S. G. *Studies on the Structure and Development of Vertebrates* Vol. II, 487–837 (Macmillan, 1930).
- Graham, J. B. Air-breathing Fishes: Evolution, Diversity, and Adaptations. (Academic Press, 1997).
- Daniels, C. B. *et al.* The origin and evolution of the surfactant system in fish: insights into the evolution of lungs and swim bladders. *Physiol. Biochem. Zool.* **77**, 732–749 (2004).
- Icardo, J. M. *et al.* The structure of the gas bladder of the spotted gar, *Lepisosteus oculatus*. *J. Morphol.* **276**, 90–101 (2014).

52. Longo, S., Riccio, M. & McCune, A. R. Homology of lungs and gas bladders: insights from arterial vasculature. *J. Morphol.* **274**, 687–703 (2013).
53. Cupello, C. *et al.* Allometric growth in the extant coelacanth lung during ontogenetic development. *Nat. Commun.* **6**, 8222 (2015).
54. Tatsumi, N. *et al.* Molecular developmental mechanism in polypterid fish provides insight into the origin of vertebrate lungs. *Sci. Rep.* **6**, 30580 (2016).
55. Davis, P. J., Hedrick, M. S. & Jones, D. R. Neuromuscular control of the glottis in a primitive air-breathing fish *Amia calva*. *Am. J. Physiol.* **264**, R204–R210 (1993).
56. Zaccone, D. *et al.* The simultaneous presence of neuroepithelial cells and neuroepithelial bodies in the respiratory gas bladder of the longnose gar, *Lepisosteus osseus*, and the spotted gar, *L. oculatus*. *Acta Histochem.* **114**, 370–378 (2012).
57. Braasch, I. *et al.* The spotted gar genome illuminates vertebrate evolution and facilitates human–teleost comparisons. *Nat. Genet.* **48**, 427–437 (2016).
58. Ohno, S. *Evolution by Gene Duplication* (Springer-Verlag, 1970).
59. Lee, A. P., Kerk, S. Y., Tan, Y. Y., Brenner, S. & Venkatesh, B. Ancient vertebrate conserved noncoding elements have been evolving rapidly in teleost fishes. *Mol. Biol. Evol.* **28**, 1205–1215 (2011).
60. Glasauer, S. M. & Neuhauss, S. C. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol. Genet. Genomics* **289**, 1045–1060 (2014).
61. Zou, M., Guo, B., Tao, W., Arratia, G. & He, S. Integrating multi-origin expression data improves the resolution of deep phylogeny of ray-finned fish (Actinopterygii). *Sci. Rep.* **2**, 665 (2012).
62. Nelson, J. S. *Fishes of the World* (John Wiley and Sons, Inc., 2006).
63. Henkel, C. V. *et al.* First draft genome sequence of the Japanese eel, *Anguilla japonica*. *Gene* **511**, 195–201 (2012).
64. Yagi, T. *et al.* A novel ES cell line, TT2, with high germline-differentiating potency. *Anal. Biochem.* **214**, 70–76 (1993).
65. Sasaki, K. & Miyazaki, J. A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *BBRC* **237**, 318–324 (1997).
66. Tamura, M. *et al.* Overdosage of Hand2 causes limb and heart defects in the human chromosomal disorder partial trisomy distal 4q. *Hum. Mol. Genet.* **22**, 2471–2481 (2013).
67. Shashikant, C. S. *et al.* Regulation of Hoxc-8 during mouse embryonic development: identification and characterization of critical elements involved in early neural tube expression. *Development* **121**, 4339–4347 (1995).
68. Chang, B. E., Blader, P., Fischer, N., Ingham, P. W. & Strähle, U. Axial (HNF3 $\beta$ ) and retinoic acid receptors are regulators of the zebrafish sonic hedgehog promoter. *EMBO J.* **16**, 3955–3964 (1997).
69. Kinoshita, M., Kani, S., Ozato, K. & Wakamatsu, Y. Activity of the medaka translation elongation factor 1 $\alpha$ -A promoter examined using the GFP gene as a reporter. *Dev. Growth Differ.* **42**, 469–478 (2000).
70. Smith, F. M. & Croll, R. P. Autonomic control of the swimbladder. *Auton. Neurosci: basic & clinical* **165**, 140–148 (2011).

## Acknowledgements

We thank Dr T. Abe for ESC targeting of MACS1. We thank Y. Mizushima, K. Fukunaga, A. Kondo, H. Nakazawa and A. Okagaki for generating many transgenic mice. We thank the Foundation for Advancement of International Science (Tsukuba, Japan) for coelacanth tissue. We thank Dr A. McMahon for the *Shh* riboprobe, Dr J. Kitano for stickleback and Dr N. Sakai for zebrafish. We thank Dr T. Takada and Dr K. Sumiyama for helpful genomic analysis and discussion, and we also thank all members of our laboratory for helpful discussion. We are grateful to NBRP Medaka for providing the d-rR strain (ID: MT837). This study was supported in part by a grant from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## Author contributions

T. Sa. and T. Sh. were involved in all aspects of the project; T. Sa. performed the majority of the mouse experiments; T.A. was involved in the cis-motif exploration; A.M. performed the micro-CT analysis; T.K., M.N., Y.T. and K.N. designed and performed the fish experiments; N.O. provided information about the coelacanth genome and the coelacanth tissue sample; H.K. performed the ES cell targeting of MACS1; all authors contributed to the preparation of the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Sagai, T. *et al.* Evolution of *Shh* endoderm enhancers during morphological transition from ventral lungs to dorsal gas bladder. *Nat. Commun.* **8**, 14300 doi: 10.1038/ncomms14300 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017