Development 136, 1665-1674 (2009) doi:10.1242/dev.032714

A cluster of three long-range enhancers directs regional *Shh* expression in the epithelial linings

Tomoko Sagai[1], Takanori Amano[1], Masaru Tamura[1], Yoichi Mizushina[1], Kenta Sumiyama[2] and Toshihiko Shiroishi[1,*]

The sonic hedgehog (*Shh*) pathway plays indispensable roles in the morphogenesis of mouse epithelial linings of the oral cavity and respiratory and digestive tubes. However, no enhancers that regulate regional *Shh* expression within the epithelial linings have been identified so far. In this study, comparison of genomic sequences across mammalian species and teleost fishes revealed three novel conserved non-coding sequences (CNCSs) that cluster in a region 600 to 900 kb upstream of the transcriptional start site of the mouse *Shh* gene. These CNCSs drive regional transgenic *lacZ* reporter expression in the epithelial lining of the oral cavity, pharynx, lung and gut. Together, these enhancers recapitulate the endogenous *Shh* expression domain within the major epithelial linings. Notably, genomic arrangement of the three CNCSs shows co-linearity that mirrors the order of the epithelial expression domains along the anteroposterior body axis. The results suggest that the three CNCSs are epithelial lining-specific long-range *Shh* enhancers, and that their actions partition the continuous epithelial linings into three domains: ectoderm-derived oral cavity, endoderm-derived pharynx, and respiratory and digestive tubes of the mouse. Targeted deletion of the pharyngeal epithelium specific CNCS results in loss of endogenous *Shh* expression in the pharynx and postnatal lethality owing to hypoplasia of the soft palate, epiglottis and arytenoid. Thus, this long-range enhancer is indispensable for morphogenesis of the pharyngeal apparatus.

KEY WORDS: Sonic hedgehog, Conserved non-coding sequence, Long-range enhancer, Epithelial lining, Mouse

INTRODUCTION

The *Shh* pathway plays a key role in multiple aspects of morphogenesis and development in vertebrate embryos (Chiang et al., 1996; Ingham and McMahon, 2001; Varjosalo and Taipale, 2008), including dorsoventral patterning of the central nervous system (CNS) (Echelard et al., 1993; Roelink et al., 1994) and anterorposterior patterning of the limb (Riddle et al., 1993). It also plays an important role in the morphogenesis in the epithelial flat sheet from the oral cavity to the hindgut, where a vast array of mature morphologies are exhibited (Chuong et al., 2000). Targeted deletion of *Shh* has shown a role for *Shh* signaling in the development of feathers (Yu et al., 2002), hair (St-Jacques et al., 1998), teeth (Dassule et al., 2000), lingual papillae (Hall et al., 2003), pharyngeal arches and pouches (Moore-Scott and Manley, 2005), foregut (Litingtung et al., 1998), and urogenital tracts (Haraguchi et al., 2001). However, the role of the *Shh* pathway in later stages of epithelial morphogenesis is poorly understood, because these tissues are severely malformed in *Shh* knockout (KO) mice.

A key step towards understanding the regulatory networks that control gene expression during morphogenesis and organogenesis is the identification of cis-regulatory elements of developmentally important genes. With the increasing availability of vertebrate genomic sequences, comparison of sequences across evolutionarily distant species has permitted the identification of numerous conserved non-coding sequences (CNCSs) (Abbasi et al., 2007; Boffelli et al., 2004; Frazer et al., 2004; Ghanem et al., 2003; Goode et al., 2005; Santagati et al., 2003; Woolfe and Elgar, 2007; Woolfe et al., 2007; Woolfe et al., 2005). Functional analysis of these elements has been carried out using BAC reporter transgenesis in mice (Gong et al., 2002; Jeong et al., 2006) and GFP reporter assays in zebrafish (Goode et al., 2005; McEwen et al., 2006; Woolfe et al., 2005).

We have previously reported that deletion of a CNCS located 840 kb upstream of the transcriptional start site of *Shh* results in a marked phenotype (Sagai et al., 2005). This CNCS is conserved among all tetrapods species examined, as well as in teleost fishes (Lettice et al., 2003; Sagai et al., 2004). Although the CNCS KO mice are viable, endogenous limb bud expression of *Shh* is completely lost, resulting in severe distal limb truncation indistinguishable from that observed in the KO mutant of the *Shh*-coding sequence (Chiang et al., 1996). Subsequent cis-trans tests verified that the CNCS contains a limb-specific *Shh* enhancer (Lettice et al., 2002; Sagai et al., 2004).

The 1 Mb genomic region spanning from the *Shh* coding region to the upstream limb-specific enhancer is unique in its low gene density and shows exceptionally long-range synteny between mammals and teleost fishes, with a number of CNCSs lining up in the same order and orientation in different species (Goode et al., 2005; Woolfe et al., 2005). Comparative sequence analysis and transgenic mouse reporter assays have uncovered three forebrain enhancers located 300 to 450 kb upstream of the *Shh*-coding sequence (Jeong and Epstein, 2003; Jeong et al., 2006). Displacement of these regulatory elements from the *Shh* promoter by chromosomal translocation is a likely cause of holoprosencephaly (HPE) in humans (Roessler et al., 1997), underscoring the importance of long-range enhancer elements in key development processes. Together with floor-plate enhancers near the transcriptional start site, these enhancers recapitulate *Shh* expression in the mouse embryonic central nervous system (Epstein et al., 1999; Jeong and Epstein, 2003; Jeong et al., 2006; Jeong et al., 2008). However, to date, epithelial linings-specific *Shh* enhancers have not been reported.

[1]Mammalian Genetics Laboratory, National Institute of Genetics, Yata-1111 Mishima Shizuoka-ken 411-8540, Japan. [2]Population Genetics Laboratory, National Institute of Genetics, Yata-1111 Mishima Shizuoka-ken 411-8540, Japan.

*Author for correspondence (e-mail: tshirois@lab.nig.ac.jp)

Accepted 10 March 2009

Here, we explore new CNCSs in the mouse 300 kb genomic region 600 to 900 kb upstream of the *Shh*-coding sequence by comparing the mouse genome with the genomes of other mammalian species and teleost fishes. We identify a cluster of three CNCSs that drive *lacZ* reporter expression in the epithelia of the oral apparatus, the pharyngeal apparatus, and the lung and gut. Interestingly, the co-linear genomic arrangement of the three CNCSs mirrors the anteroposterior order of their expression domains, partitioning the continuous epithelial lining into three *Shh* expression domains: the ectoderm-derived oral cavity, the anterior endoderm-derived pharynx, and posterior endoderm-derived respiratory and digestive tubes. We also generate KO mouse mutants that lack the CNCS that drives expression in the pharyngeal apparatus. In these animals, endogenous *Shh* expression is lost specifically in the pharyngeal epithelia, resulting in postnatal lethality owing to hypoplasia of the soft palate, epiglottis and arytenoid that are essential for respiratory and swallowing functions. These results demonstrate that the *Shh* pathway is essential for the morphogenesis and the development of the pharyngeal region in the mouse.

MATERIALS AND METHODS

Animals

The *Shh*-coding sequence KO mouse (*Shh*$^{-/-}$) was kindly provided by Dr P. Beachy, and is maintained at the National Institute of Genetics (NIG), Mishima, Japan. C57BL/6J and (C57BL/6×DBA/2)F1 mice were purchased from Japan Crea (Tokyo, Japan). The animal experiments in this study were approved by the Animal Care and Use Committee of the NIG.

Sequence analysis

For sequence alignment and homology comparisons, we used the ClustalW system (http://www.ddbj.nig.ac.jp/search/clustalw-j.html), VISTA program (http://genome.lbl.gov/vista/index.shtml), UCSC genome Browser Home (http://genome.ucsc.edu/) and Ensemble database (http://www.ensembl.org/index.html). The draft genome sequences of humans, mice and medaka fish have been previously described. For homology comparison, we used the following sequences from the UCSC database: mouse Chr5, 28,783,380-29,704,930; human Chr7, 155,294,520-156,378,6670; chicken Chr2, 8,031,550-8,431,420; Xenopus Chr28, 4,064,830-4,399,140; medaka Chr20, 17,738,740-17,852,820. The medaka genome sequence is also referenced from the NIG DNA sequence center (http://dolphin.lab.nig.ac.jp/medaka/index.php). The sequence data of the CNCSs MRCS1, MFCS3, MFCS4, MACS1 and MFCS2 have been submitted to DDBJ under Accession Numbers AB453051, AB453050, AB258402, AB453049 and AB453052, respectively. MFCS3 and MFCS4 are mouse orthologs of the previously described fugu or human sequences,SHH2 (Accession Number CR847489) and SHH1 (Accession Number CR847488) (Woolfe et al., 2005). The genomic positions of CNCSs in different species, which were used for the Vista analysis, are listed in Table S1 in the supplementary material.

Transgenic assay

Mouse genomic DNA fragments, including the CNCSs were amplified from RP23-284A9 or RP23-428P20 BAC DNA. After sequencing, the amplified fragments were inserted into the *Hin*dIII or *Sal*I site of the hsp68/*LacZ* expression vector (Shashikant et al., 1995). Details of the primer pairs used for amplification of the inserts can be provided on request. To obtain the MFCS4 fragment lacking a 217 bp ultra-conserved sequence, inverted tail-to-tail primer pairs were used to amplify a basic whole vector, except for the 217 bp ultra-conserved sequence of MFCS4. The primer pair used was: F, 5′-AGATTGGGTTCACTGTGTGC-3′; R, 5′-CACAAGCCTCTTTA-GTCAGG-3′. Then, the deleted form of MFCS4 (ΔMFCS4) was subcloned into the *lacZ* reporter construct. The *Xho*I and *Not*I double-digested fragments were cut out from an 0.8% low melting agarose gel and digested with GELase enzyme (Epicentre Technologies, Madison, WI, USA) at 43°C overnight. After phenol and chloroform extraction, DNA was precipitated with ethanol and dissolved in a small volume of injection buffer [5 mM Tris-HCl (pH 7.5); 0.1 mM EDTA (pH 8.0)]. DNA (1-4 ng/μl) was purified using a filter unit and used in injection experiments. Transient transgenic embryos and stable transgenic mouse lines were generated by pronuclear injection into fertilized eggs derived from the (C57BL/6×DBA/2)F1 or C57BL/6 strain. Transgenic animals were selected using the following primer pairs for the *lacZ* gene: F, 5′-TCACCCTGCCATAAAGAAACT-3′; R, 5′-CTGTCGTCGTCCCCTCAAACT-3′. Whole-mount *lacZ* staining was carried out as previously described (Maas and Fallon, 2005). 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.

ES cell targeting

We used a previously described basic targeting vector to build the MFCS4 targeting construct (Sagai et al., 2005). The long arm (5478 bp) was amplified from BAC RP23-284A9 DNA with the primer pair 5′-ATG-GTACCAGGAGATATGCTGCATCCTC-3′ and 5′-TACTCGAGAGAA-CTGCGGTTTAACCTGC-3′, and the short arm (1824 bp) was amplified with the primer pair 5′-CCGGAATTCGCATTAGAAGCTGGGATGGA-3′ and 5′-CGCGAATTCGGACCTTACATACGTGAAGC-3′. The 999 bp genomic sequence, including mouse MFCS4, was replaced with the Neo cassette (see Fig. S1 in the supplementary material). The targeting vector was electroporated into TT2 ES cells, which originated from a (C57BL/6×CBA)F_1 mouse (Yagi et al., 1993). ES cells were screened with the following PCR primer pair: p1, 5′-AGTGCTGTCCCAGAGATAAG-3′ and p2, 5′-CATCGCATTGTCTGAGTAGG-3′. Positive clones were aggregated with eight-cell embryos from (DBA/2×C57BL/6)F_1 mice and transplanted into surrogate females. Male chimeras were mated with C57BL/6 females. Segregation of the targeted allele was determined using three PCR primers: p2, p3, 5′-TCTCAATCTGAACACTGGGC-3′, and p4, 5′-TCTCAATCTGAACACTGGGC-3′. Skeletal analysis of newborn mice was performed as previously described (Trokovic et al., 2003). For the histological analysis of the ΔMFCS4/ΔMFCS4 mutants, embryos at E18.5 were preserved in Bouin's fixative and embedded in paraffin. Serial sections of 5 μm were collected and stained with Hematoxylin-Eosin. We referred to an anatomical atlas (Kaufman, 1992) throughout the experiments.

In situ hybridization

For cryosectioning, embryos were fixed in 4% paraformaldehyde and immersed in 30% sucrose/phosphate-buffered saline overnight, embedded and frozen in OCT, and sectioned at 15 μm. For paraffin sections, embryos were fixed with 4% paraformaldehyde and dehydrated in a methanol series, then embedded in paraffin and sectioned at 8 μm. In situ hybridization was performed using digoxigenin-UTP-labeled riboprobes, as previously described (Makino et al., 2001). Whole-mount in situ hybridization was performed as previously described (Wilkinson, 1992).

RT-PCR

Total RNA was extracted from the anterior tongue, epiglottis-arytenoid swelling and lung tissues of C57BL/6 embryos at E13.5. One μg of each RNA sample was reverse transcribed into cDNA with SuperScript III Transcriptase (Invitrogen). One μl of the cDNA solution was used for PCR amplification. Details of primer pairs used can be provided on request.

RESULTS

Exploration of novel CNCSs 600-900 kb upstream of the *Shh* transcriptional start site

We compared the mouse genomic sequence 600-900 kb upstream of the *Shh* transcriptional start site with the syntenic regions of various tetrapod species and teleost fishes, using public genome databases. Previously, we have identified three CNCSs in this region that are conserved between mammals and teleost fishes and that we called Mammal Fish Conserved Sequence 1 to 3 (MFCS1 to 3) (Fig. 1) (Sagai et al., 2005). Here, we describe the identification of three additional CNCSs in this region (Fig. 1A) that are evolutionarily conserved at different level (Fig. 1B). MFCS4 is conserved between mammals and teleost fishes, albeit with slightly lower sequence

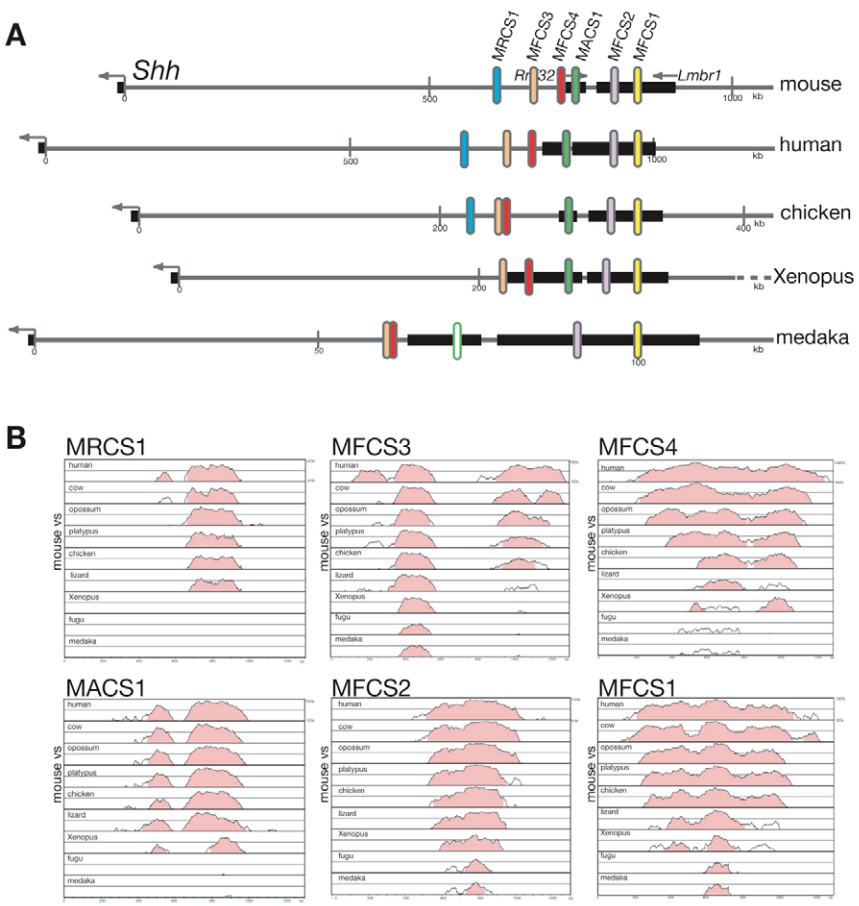

Fig. 1. Location of CNCSs in the interval between the *Shh*-coding sequence and *Lmbr1*. (**A**) The genomes of five evolutionarily distant species: mouse (chr5), human (chr7), chicken (chr2), *Xenopus* (chr28) and medaka (chr20). The scale is in kb and is expanded to 2.5-fold in the chicken and *Xenopus* genomes, and 10-fold in the medaka genome compared with the mammalian genomes. To compare the relative positions of the CNCSs, MFCS1 was right-aligned at the same position for all species. The colored ovals indicate the locations of MRCS1 (blue), MFCS3 (orange), MFCS4 (red), MACS1 (green), MFCS2 (purple) and MFCS1 (yellow). Putative medaka MACS1 is depicted with a white oval rimmed with green. (**B**) VISTA plots of the six CNCSs. Mouse CNCSs are compared with the homologs of the human, cow, opossum, platypus, chicken, lizard, *Xenopus*, fugu and medaka CNCSs. Constraints used are a 100 bp window length and 70% conservation level.

similarity than MFCS1-3 (Fig. 1A,B), and is an ortholog of a human-fugu conserved sequence called SHH1, which has been reported by another group (Goode et al., 2005; Woolfe et al., 2005). Mammal Reptile Conserved Sequence 1 (MRCS1) is conserved among mammals, chicken and lizard, but not in *Xenopus* and teleost fishes. Mammal Amphibian Conserved Sequence 1 (MACS1) is conserved among mammals, chicken, lizard and Xe*n*opus, and has very weak similarity to a short sequence in the medaka genome (Fig. 1B). MACS1 is a partial cDNA fragment previously registered as AK043126 in the FANTOM data set (Carninci and Hayashizaki, 1999) and is located in intron 8 of *Rnf32*, a gene with unknown function.

Thus, in total there are six CNCSs presently known to reside in the 600 to 900 kb region upstream of the *Shh* transcriptional start site (Fig. 1). Notably, the order and orientation of these CNCSs relative to direction of *Shh* transcription are conserved across evolutionarily distant species.

Transgenic reporter assay of the CNCSs

To test the function of the newly identified CNCSs, we carried out transgenic assays by examining β-galactosidase activity in mouse embryos carrying a ~1 kb mouse genomic sequence containing the five different CNCSs linked to the *lacZ* reporter gene. We first examined reporter gene expression in whole-mount transgenic embryos at embryonic day (E) 9.0-15.5. As summarized in Table 1, MFCS3, MFCS4, MRCS1 and MACS1 directed reproducible tissue-specific reporter gene expression, whereas MFCS2 did not show tissue-specific expression. MFCS3 drives *lacZ* expression in the brain and motor neurons at E11.5 (see Fig. S2 in the supplementary material), whereas the other three CNCSs (MFCS4, MRCS1 and MACS1) drive expression in different epithelial linings.

MRCS1 drives strong *lacZ* expression in the epithelia of the incisor and molar teeth at E12.5 (Fig. 2A,B) and weak expression in the lingual papillae (Fig. 2B). At E13.5, *lacZ* signal appears in the whisker buds, tooth primordia, the rugae of the hard palate (Fig. 2C). Expression in the fungiform papilla of the anterior tongue also becomes more prominent at this stage (Fig. 2D). Expression in these domains continues at least until the palatal structure is established at E15.5 (Fig. 2E-G). At E14.5-15.5, the reporter signals were also detected in the hair and nail buds (Fig. 2E).

In the MFCS4 transgenic mice, very weak signal was detected in the oral epithelium at E11.0 (data not shown). At E11.5, weak and transient signal was detected in the dental germ (Fig. 2H). Strong expression was observed in the pituitary fossa region and in the tympanic tube and recess (Fig. 2H). In the lower facial area,

Table 1. *lacZ* reporter expression directed by CNCSs in transgenic mice

	Number of embryos with specific reporter expression/ total number of embryos examined				
	MFCS2	MFCS3	MFCS4	MRCS1	MACS1
Motor neuron	–	10/13	–	–	–
Tooth, lingual papilla	–	–	–	9/10	–
Epiglottis, soft palate	–	–	18/20	–	–
Lung, urogenital organ	–	–	–	–	7/10

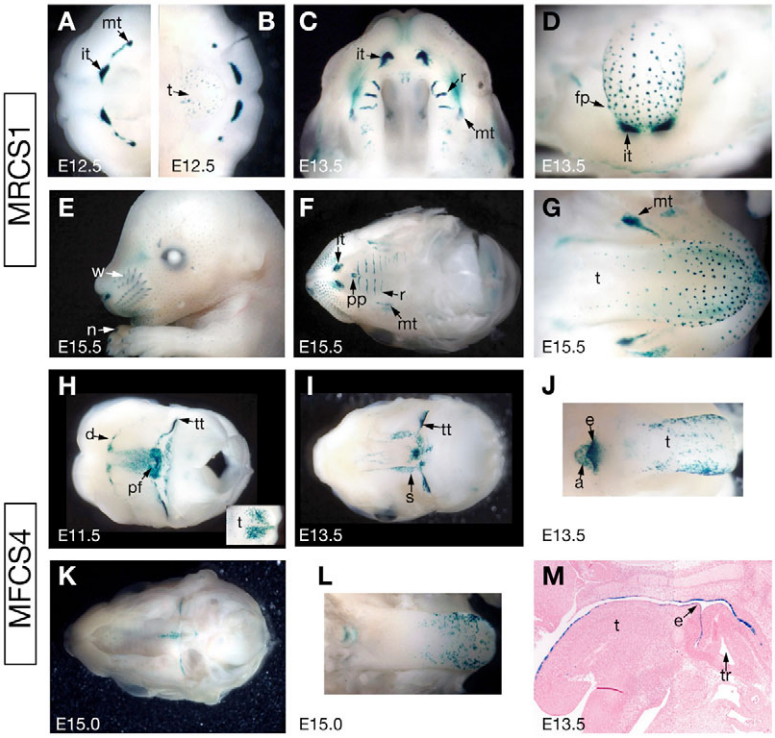

Fig. 2. *lacZ* reporter expression driven by MRCS1 and MFCS4 in transgenic mouse embryos. (**A-G**) Reporter signal in MRCS1 transgenic embryos. X-gal staining is detected in the dental placode of the maxilla (A) and mandible (B) at E12.5. Signal is also detected in the fungiform papillae (B), whisker buds, teeth, palatal rugae and fungiform papillae (C,D) at E13.5. At E15.5, expression is detected in the whisker, hair and nail buds (E); teeth, primary palate and palatal rugae (F); molar teeth and fungiform papillae (G). (H-M) Reporter signal in MFCS4 transgenic embryos. (**H**) At E11.5, transient signal is detected in the dental placode, and intense signal around the pituitary fossa, the tympanic tube and recess. Punctate signal is detected on the tongue surface (inset in H). (**I**) At E13.5, expression is detected in the soft palatal edge and tympanic tube in the skull base. (**J**) Strong signal is detected in the epiglottis and arytenoid swelling. (**K**,**L**) At E15.0, expression is observed in the midline of the soft palate (K), but the intensity of the signal is depressed in the epiglottis and arytenoid swelling (L). (**M**) A mid-sagittal section of an E13.5 embryo shows reporter signal in the epithelia of the tongue, epiglottis, arytenoid swelling and dorsal pharyngeal wall. a, arytenoid swelling; d, dental placode; e, epiglottis; it, incisor tooth; fp, fungiform papillae; mt, molar tooth; n, nail bud; pf, pituitary fossa; pp, primary palate; r, palatal rugae; s, soft palate; t, tongue; tr, trachea; tt, tympanic tube; w, whisker bud.

punctuate expression was detected on the tongue surface, but not in the fungiform papillae (inset in Fig. 2H). *lacZ* expression persists in these regions until E13.5 (Fig. 2I,J). Notably, signal was detected in the palatal shelves of the soft palate, but not in the anterior half of the palate or in the hard palate (Fig. 2I), where MRCS1 drives reporter expression (Fig. 2C,F). Strong expression was detected in the epiglottis and arytenoid swelling at E13.5 (Fig. 2J), which are essential for respiratory and swallowing functions. At E15.0, palatal formation is completed and signal was observed in the midline of the soft palate (Fig. 2K), but was decreased in the epiglottis and arytenoid swelling (Fig. 2L). Histological analysis showed that MFCS4-driven expression is restricted to the epithelial lining (Fig. 2M).

MACS1 drives *lacZ* reporter expression in the gut endoderm and caudal region at E9.0 (Fig. 3A). At E9.5, intense signal marks the region where the future respiratory tube will evaginate from the primitive gut (Fig. 3B). As shown in the sagittal section, signal was detected throughout the continuous lining of the gut endoderm at E9.5 (Fig. 3C). At E10.5, signal was observed in the gut and cloaca (Fig. 3D), as well as the primary lung buds (Fig. 3E). At E10.5,

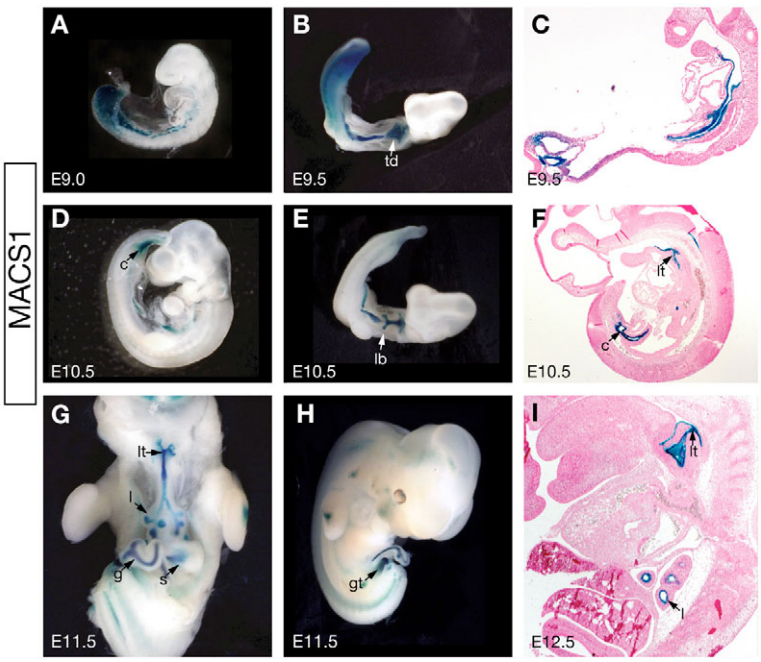

Fig. 3. MACS1-driven *lacZ* reporter expression in transgenic embryos. (**A**) *lacZ* reporter expression is observed throughout the entire gut tube at E9.0. (**B**) The condensed signal is observed in the tracheal diverticulum at E9.5. (**C**) A mid-sagittal section at E9.5 shows that signal is confined to the epithelial lining of the endoderm. (**D**) At E10.5, strong caudal expression disappears and signal is detected in the gut endoderm and cloaca. (**E**) Clear signal is detected in the primary lung buds. (**F**) A mid-sagittal section of an E10.5 embryo shows reporter signal in the epithelia of the ventral pharyngeal wall, laryngotracheal tube and cloaca. (**G**,**H**) At E11.5, reporter signal is widely observed in the laryngotracheal tube, lung, gut (G) and genital tubercle (H). (**I**) A mid-sagittal section of an E12.5 embryo shows strong signal in the epithelia of the laryngotracheal tube and lung. c, cloaca; g, gut; gt, genital tubercle; l, lung; lb, lung bud; lt, larybgotracheal tube; s, stomach; td, tracheal diverticulum.

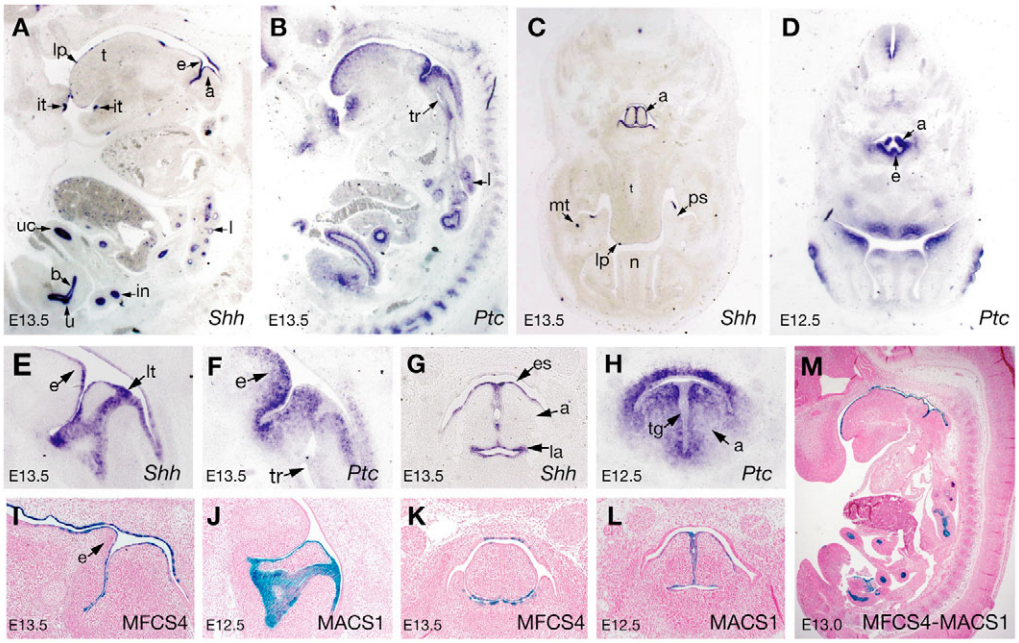

Fig. 4. *Shh* is expressed in the epithelial linings along the anteroposterior axis. (**A-H**) Endogenous *Shh* and *Ptch1* expression in sagittal (A,B,E,F) and cross (C,D,G,H) sections of wild-type embryos. At E13.5, *Shh* expression is widely detected in the epithelia of the teeth, palatal shelf, lingual papillae, epiglottis, arytenoid swelling, pharyngeal wall, lung, gut and urogenital tract (A,C). *Ptch1* expression is mainly detected in the mesenchyme adjacent to the *Shh*-expressing epithelia (B,D). (E-L) Magnified images in the pharyngeal apparatus. *Shh* expression is detected in the epithelia of the epiglottis, arytenoid swelling and laryngotracheal tube (E,G). *Ptch1* expression is mainly observed in the mesenchyme adjacent to the *Shh*-expressing epithelia (F,H). (**I-L**) Reporter expression in MFCS4 and MACS1 transgenic embryos. MFCS4 reporter expression (I,K) partially overlapped with MACS1-driven expression (J,L) in the epithelia of the arytenoid swelling. (**M**) Reporter expression in the transgenic mice with a tandem array of MFCS4 and MACS1. The linked sequence drives the signal in both expression domains by MFCS4 and MACS1. a, arytenoid swelling; b, bladder; e, epiglottis; it, incisor tooth; in, intestine; l, lung; la, laryngeal aditus; lp, lingual papillae; lt, laryngotracheal tube; mt, molar tooth; n, nasal septum; p, pharynx; ps, palatal shelf; t, tongue; tr, trachea; u, urethra; uc, umbilical cord.

intense signal was apparent in the epithelia of the laryngo-tracheal tube and cloaca. At E11.5, strong expression was observed in the laryngotracheal tube, lung, gut (Fig. 3G) and genital tubercle (Fig. 3H). A mid-sagittal section confirmed the intense signal in the epithelia of the laryngotracheal tube, lung, digestive tube and urogenital tract (Fig. 3I; data not shown) at E12.5. Thus, MACS1 strictly specifies the expression in the epithelia of the respiratory and digestive tubes, which are derived from the primitive gut endoderm.

Comparison of *lacZ* reporter expression with endogenous *Shh* expression in the epithelial linings

We observed endogenous expression of *Shh* and its downstream gene *Ptch1*, and compared it with *lacZ* reporter expression (Fig. 4). Endogenous *Shh* expression is observed in the epithelia of the tooth primordia, lingual papillae, palate, epiglottis, arytenoid swelling, lung, gut and urogenital tracts at E12.5-13.5 (Fig. 4A,C,E,G), whereas *Ptch1* endogenous expression is mainly observed in the mesenchyme of the tooth primordia, tongue, epiglottis, arytenoid swelling, lung and gut at these stages (Fig. 4B,D,F,H). As described above, the reporter signals driven by MRCS1, MFCS4 and MACS1 are confined to the epithelial linings (Fig. 2M; Fig. 3C,F,I), and mirror the endogenous *Shh* expression domains. Even though the MFCS4 and MACS1 expression domains partially overlap at their borders in the arytenoid swelling (Fig. 4I-L), they are distinct, with MFCS4 driving expression in the epiglottis (Fig. 2M; Fig. 4I) and MACS1 driving expression in the laryngotracheal tube (Fig. 4J,L). The MFCS4 and MACS1 tandem construct drives reporter expression throughout the pharyngeal apparatus and the respiratory and digestive tubes (Fig. 2M; Fig. 3I; Fig. 4M). Thus, these three CNCSs appear to contain enhancers that together recapitulate the endogenous pattern of *Shh* expression in the epithelial linings.

Generation of MFCS4 KO mouse mutant

To examine whether the CNCSs act as functional elements in vivo, we performed gene targeting, focusing on MFCS4. Following the same strategy used to generate MFCS1 knockout mice (Sagai et al., 2005), we replaced a ~1 kb genomic region, including MFCS4 with a *neo* cassette. We successfully generated two lines of germline chimeras, then, heterozygotes (ΔMFCS4/+) were intercrossed to generate homozygous mutants (ΔMFCS4/ΔMFCS4) (Table 2). ΔMFCS4/ΔMFCS4 mice were present at E18.5 and early postnatal stages, but all of them died within a few days of birth (Table 2). Although ΔMFCS4/+ mice are viable and grossly indistinguishable from wild-type mice, many ΔMFCS4/ΔMFCS4 neonates have

Table 2. Viability of MFCS4 KO mouse mutants

	Number of mice with MFCS4 genotype		
Age	+/+	+/–	–/–
E18.5	22	40	11
P0	6 (1)*	10	8 (4)*
P1	15	28	12 (10)*
P2	4	10	5 (5)*
1-2 months	37	71	0

*In parentheses is the number of mice that died on the day.

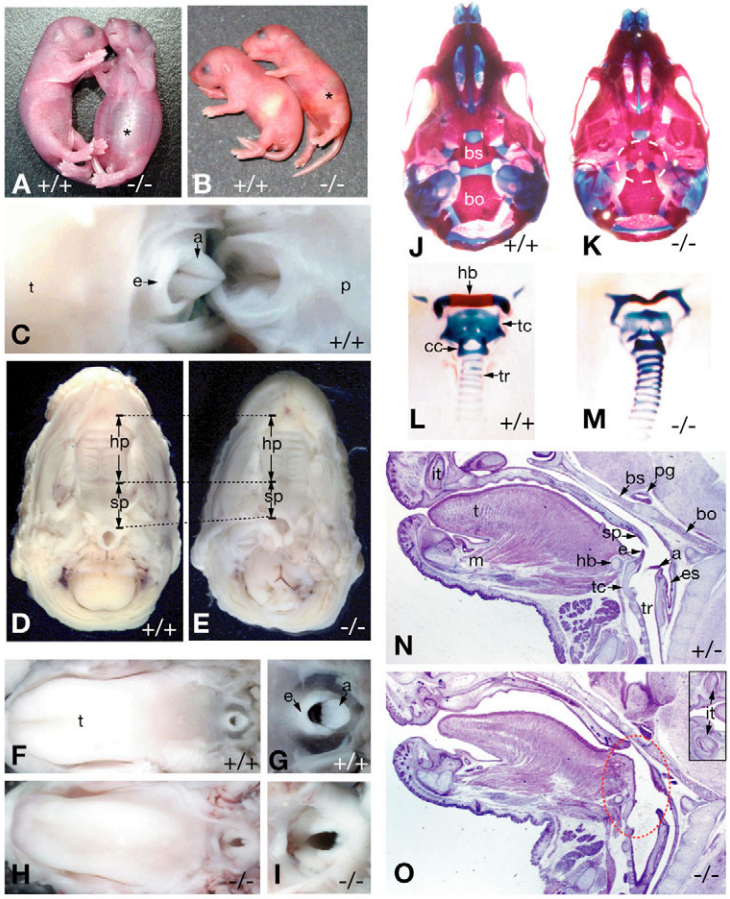

Fig. 5. Phenotypes of ΔMFCS4/ΔMFCS4 mice. (**A**) Many ΔMFCS4/ΔMFCS4 mice had a dilated belly (asterisk). (**B**) A ΔMFCS4/ΔMFCS4 neonate whose stomach is not filled with milk (asterisk) (B). (C-M) Phenotypes of embryos at E18.5. (**C**) Wild-type oro- and naso-pharyngeal junction, upper right. The epiglottis and arytenoid are clearly recognizable. (**D**,**E**) Lower aspect of the hard and soft palates, and the naso-pharyngeal opening. In wild-type mice, the hard palate (hp, double arrow) and soft palate (sp, double arrow) are well formed (D). In ΔMFCS4/ΔMFCS4 mice, the palates are fused at the midline. The hard palate is formed normally, but the soft palate terminates prematurely. As a consequence, the naso-pharyngeal opening in ΔMFCS4/ΔMFCS4 mice is misaligned to the anterior side (E). (**F-I**) Views of the dorsal aspect of the tongue and the oro-pharyngeal opening. (G,I) Magnified images of the oro-pharyngeal opening. In ΔMFCS4/ΔMFCS4 mice, the tongue is misshapen (H) compared with wild-type (F), the epiglottis and arytenoid are hypoplastic (I) compared with wild-type (G). (**J**,**K**) Ventral view of the skull. Mutant animals have minor morphological anomalies in the basisphenoid and basioccipital bones, and have a hole in the spheno-occipital synchondrosis just inferior to the pituitary (K). (**L**,**M**) Dorsal views of the laryngeal cartilages. In mutant animals, the hyoid bone is malformed and the thyroid cartilage appeared to be hypoplastic (M). (**N**,**O**) Mid-sagittal sections of the oral and pharyngeal structures of heterozygous and homozygous mutant animals at E18.5. In ΔMFCS4/+ mice, the posterior edge of the soft palate extends alongside the posterior pharyngeal wall and is situated close to the epiglottis. The epiglottis and arytenoid are well formed (N). By contrast, in ΔMFCS4/ΔMFCS4 mice, the soft palate terminates prematurely, and the posterior tongue, epiglottis and arytenoid are severely deformed. The cross-sectional diameter of the body of the hyoid bone is decreased (thinner), and the bone and cartilage beneath the pituitary gland show morphological abnormalities (O). Red outline indicates the affected area in the ΔMFCS4/ΔMFCS4 mutant. a, arytenoid; bo, basioccipital bone; bs, basisphenoid; cc, cricoid cartilage; e, epiglottis; es, esophagus; hb, hyoid bone; hp, hard palate; it, incisor tooth; m, mandible; p, palate; pg, pituitary gland; sp, soft palate; t, tongue; tc, thyroid cartilage; tr, tracheae.

bloated bellies, which might be caused by air accumulation in the stomach and bowels (Fig. 5A), and their stomachs are always devoid of milk (Fig. 5B). The tongue, soft palate, epiglottis and arytenoid comprise the pharyngeal apparatus, which is essential for respiration and swallowing (Fig. 5C,D,F,G). In the ΔMFCS4/ΔMFCS4 mice, truncation of the soft palate (Fig. 5E), loss or reduced size of the epiglottis and hypotrophy of the arytenoid (Fig. 5H,I), and tongue deformation (Fig. 5H) are observed. Moreover, deletion of MFCS4 appears to result in minor morphological abnormalities of the basisphenoid and basioccipital bones, as is visible from a ventral view of the skull (Fig. 5J,K). The cartilage between these bones, which is probably the pituitary fossa region that closes after formation of the pituitary, has a hole (marked with the broken white line in Fig. 5K). The ΔMFCS4/ΔMFCS4 mice display defects in the hyoid bone and thyroid cartilage (Fig. 5L,M). Although there is variability in the severity of the defects, the major phenotype was completely penetrant, with abnormalities observed in the aforementioned structures in all the ΔMFCS4/ΔMFCS4 mice examined (n=17).

Histological analysis of the sagittal sections showed that the ΔMFCS4/ΔMFCS4 embryos have deformed posterior tongues, hypoplasia of the hyoid cartilage, loss of the epiglottis, hypoplasia of the arytenoid and shortening of the soft palate (Fig. 5O), which eventually resulted in atypical junctions between the nasopharynx and oropharynx. The posterior tongue edge, including hyoid bone and thyroid cartilage was deformed (Fig. 5M,O), and the base of

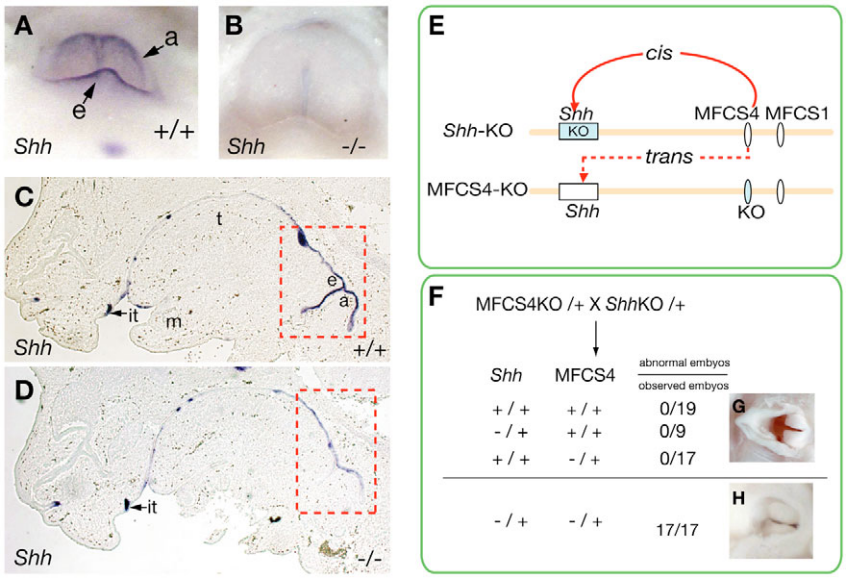

Fig. 6. MFCS4 is a cis-acting regulator of *Shh*. (A-D) *Shh* expression in wild-type and ΔMFCS4/ΔMFCS4 embryos. (**A**) In wild-type embryos, *Shh* expression is detected in the epiglottis and arytenoid swelling at E13.5. (**B**) In ΔMFCS4/ΔMFCS4 embryos, *Shh* expression is lost in both structures. (C,D) Section in situ hybridization for *Shh* in wild-type and ΔMFCS4/ΔMFCS4 embryos at E13.0. (**C**) In wild-type embryos, *Shh* signal is broadly detected in the pharyngeal epithelium. (**D**) In ΔMFCS4/ΔMFCS4 embryos, *Shh* expression is downregulated in the epithelium of the pharyngeal wall, which is the future oro-pharynx area, and expression almost completely disappears in the epithelium of the epiglottis, the thyroid duct and the arytenoid swelling. *Shh* signal is detected at wild-type levels in the incisor tooth primordium of ΔMFCS4/ΔMFCS4 embryos (D). Boxed areas indicate *Shh* expression in the pharyngeal region of the wild type and the ΔMFCS4/ΔMFCS4 mutant. (E-H) Genetic test for cis-activity of MFCS4. (**E**) Diagram shows how cis or trans regulation of *Shh* expression by MFCS4 would function in the pharyngeal epithelium. If MFCS4 were a trans-acting regulator of *Shh* expression, an intact MFCS4 sequence located in a cis position relative to the *Shh* KO allele should be able to activate transcription of the intact *Shh* gene on the trans allele. Therefore, the ΔMFCS4/ΔMFCS4 phenotype should be rescued in *Shh* KO: MFCS4 KO compound heterozygotes. Conversely, if MFCS4 were a cis-acting regulator, the intact MFCS4 sequence would not be able to activate transcription of the *Shh* gene on the trans allele, and the compound heterozygotes would have the same phenotype as individuals with the single MFCS4 mutation. (**F-H**) Phenotypes of E18.5 embryos from the cross mating of $Shh^{+/-}$ mice to ΔMFCS4/+ mice. The four expected genotypes of the progeny segregated in a Mendelian fashion. Among them, the wild-type and the $Shh^{+/-}$ or ΔMFCS4/+ mice had normal phenotypes with respect to the oro-pharyngeal openings (G). However, all compound heterozygotes had abnormal openings, as did single ΔMFCS4/ΔMFCS4 mutants (H). a, arytenoid swelling; e, epiglottis; it, incisor tooth; m, mandible; t, tongue.

tongue seemed to be affected as well, which eventually result in the atypical tongue (Fig. 5H). Though the transgenic MFCS4 reporter signal was detected in the tympanic tube and recess, visible defects were not obvious in the auditory organs of the ΔMFCS4/ΔMFCS4 neonates. Most of the ΔMFCS4/ΔMFCS4 had normal hard palate, and a few neonates (5/37) exhibited cleft palate (data not shown). Notably, in the ΔMFCS4/ΔMFCS4 neonates, the tooth and whisker, in which the MFCS4-mediated *lacZ* expression was not observed, but the MRCS1-mediated *lacZ* expression was observed, were not affected (inset in Fig. 5O; data not shown). As described before, the reporter expression domains driven by MFCS4 and MACS1 are partially overlapped in the arytenoid swelling. It is consistent with the fact that the posterior border of the defects observed in the ΔMFCS4/ΔMFCS4 neonates corresponds to the posterior end of the pharynx. The ΔMFCS4/ΔMFCS4 neonates showed no visible defects in the cricoid cartilage, trachea, lung, esophagus, digestive tube and urogenital organs (see Table S2 in the supplementary material). Thus, MFCS4 is crucial for development specifically of the pharyngeal apparatus.

We next examined whether the *Shh* pathway is altered in the ΔMFCS4/ΔMFCS4 embryos by whole-mount and section in situ hybridization using a *Shh* riboprobe. Strong *Shh* expression was detected in the epithelium of the epiglottis and arytenoid swelling of wild-type embryos at E13.0 and E13.5 (Fig. 4A,C; Fig. 6A,C), whereas the ΔMFCS4/ΔMFCS4 embryos lost almost all *Shh* expression in the epithelia of the epiglottis and arytenoid swelling (Fig. 6B,D). However, *Shh* expression in the tooth primordia was not altered in the mutant embryos (Fig. 6D), in accordance with the normal tooth development in the ΔMFCS4/ΔMFCS4 embryos (Fig. 5O).

Finally, we carried out a cis-trans test by mating animals heterozygous for the *Shh*-coding sequence KO allele ($Shh^{+/-}$) to the ΔMFCS4/+ mouse. All compound heterozygotes exhibited hypoplasia of the epiglottis and arytenoids, indicating that MFCS4 is *cis*-acting to the *Shh*-coding sequence (Fig. 6E,F). Together, our results indicate that MFCS4 is a pharyngeal epithelium-specific enhancer of *Shh*, and is indispensable for morphogenesis of the pharyngeal structures in the mouse.

Evolutionary rigidity assay (ERA) of MFCS4

Alignment of the MFCS4 nucleotide sequences across distant species revealed a 217 bp ultra-conserved sequence, which is conserved from mammals to teleost fish medaka (Fig. 1A; see Fig. S3 in the supplementary material). To assess importance of the sequence for the MFCS4 enhancer activity, we carried out a transgenic assay for the *lacZ* reporter construct, in which the MFCS4 fragment lacks the 217 bp sequence (see Fig. S3 in the supplementary material). The result showed that this deletion form of MFCS4 drives no reporter expression in the relevant tissues of the transgenic embryos (0/5). It supports the fact that the 217 bp ultra-

conserved sequence is indispensable for the MFCS4 enhancer activity. To explore transcription factors that bind to this sequence, we carried out the evolutionary rigidity assay (ERA) for the 217 bp sequence. We identified four nearly perfect match motifs (NPMMs) (see Fig. S3 in the supplementary material), which are potentially binding sites for Pbx1 (MFCS4-MF-A), Sox5 (MFCS4-MF-B), TCFs (MFCS4-MF-C), inner-cell mediators of Wnt signaling and Hes1 (MFCS4-MF-D), a well known target of Notch signaling.

A number of reports suggest that Pbx, Sox, Wnt and Notch signaling pathways are indispensable for normal development of epithelial tissues such as tooth, hair follicle, taste papillae, lung and gastrointestinal tract (Ito et al., 2000; Iwatsuki et al., 2007; Li et al., 2005; Okubo et al., 2006; Schnabel et al., 2001). We examined expression of many genes involving these signaling pathways in the tongue, epiglottis and lung. As shown in Fig. S4 in the supplementary material, most of the genes are activated in the epiglottis and arytenoid swelling at E13.5. The results suggested that they are good candidates for the MFCS4 activity.

DISCUSSION

Three clustered enhancers direct co-linear *Shh* expression along the anteroposterior axis in the epithelial linings

In this study, we have identified three novel CNCSs in the region 620 to 740 kb upstream of the *Shh* transcriptional start site. Their sequences, order and orientation are highly conserved between evolutionarily distant species (Fig. 1A; Fig. 7A). These CNCSs direct regional expression of *Shh* in the epithelial linings from the oral cavity to the hindgut along the anteroposterior body axis (Fig. 7B). At the most anterior extent, MRCS1 directs *Shh* expression in the epithelia of the hair and whisker buds, dental placode, rugae of the hard palate and fungiform papillae of the anterior tongue, all of which are presumptive derivatives of the oral ectoderm. MFCS4 drives expression in the soft palate, epiglottis, arytenoid swelling and other pharyngeal tissues, forming a border with the more anterior MRCS1-directed expression domain. Finally, MACS1 drives expression in the respiratory and digestive tubes, with the anterior border of expression partially overlapping the posterior border of MFCS4-driven expression in the pharyngeal structures. Thus, this study clearly shows that three distinct enhancers regulate *Shh* expression along the long continuous epithelial linings from the oral cavity to the hindgut. Interestingly, the genomic order of these three enhancers is co-linear with the regional control of *Shh* expression along the anteroposterior body axis (Fig. 7). This suggests that, at least with regard to the *Shh* pathway, morphogenesis of the epithelial linings is roughly partitioned into three components: the oral ectoderm, pharyngeal endoderm and gut endoderm. At present, the biological implications of the co-linearity of these enhancers are uncertain, but it is possible that clustering of the three enhancers is required for the proper spatiotemporal regulation of *Shh* expression.

Endogenous *Shh* expression is initiated progressively along the continuous epithelial linings spanning from the oral cavity to the hindgut (Bitgood and McMahon, 1995; Iseki et al., 1996; Varjosalo and Taipale, 2008). In the anterior ectoderm-derived epithelial lining, *Shh* expression starts around E11.5-12.5 and is implicated in development of the oral cavity, the tooth and the tongue (Cobourne et al., 2004; Hall et al., 1999). At E11.5-12.5, the pharyngeal arches become broadened and flattened externally, and form the neck of the embryo. Around this stage, *Shh* expression is observed in the pharyngeal epithelial lining (Fig. 4E,G) (Rice et al., 2006). In the primary gut, the earliest *Shh* expression is detected around E8.0 and

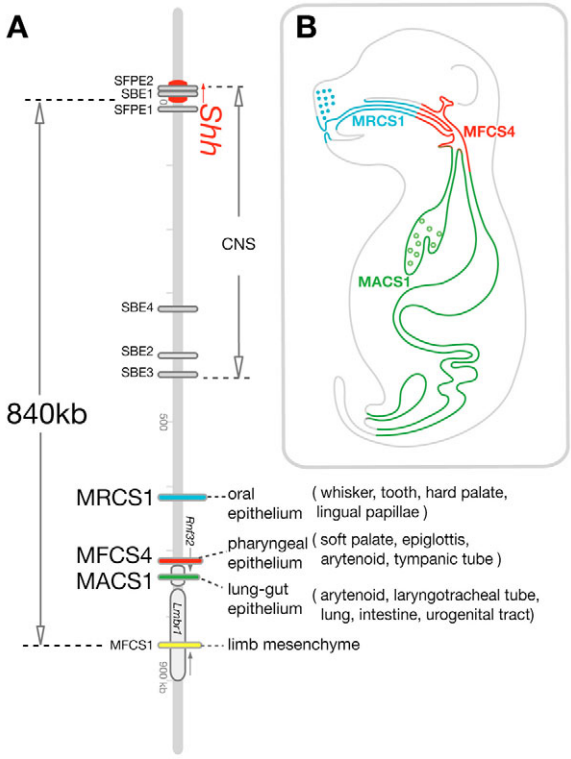

Fig. 7. Genomic map of *Shh* enhancers. (**A**) The 840 kb genome region upstream of the *Shh* transcriptional start site contains numerous regulatory elements. Previously described enhancers directing *Shh* expression in the CNS are also shown (Jeong et al., 2006). (**B**) Schematic diagram showing the expression domains of *Shh* in epithelial lining regulated by MRCS1 (blue), MFCS4 (red) and MACS1 (green) along the anteroposterior axis.

is necessary for the formation of the lung, gut and urogenital organs (Haraguchi et al., 2007; Litingtung et al., 1998). Our transgenic experiments showed that timing of reporter expression onset driven by MRCS1, MFCS4 and MACS1 in each lining is consistent with the endogenous temporal expression pattern of *Shh*, suggesting that the three enhancers identified in this study regulate endogenous *Shh* expression. However, it is notable that the co-linearity in the temporal expression patterns driven by the three enhancers is not as obvious as that observed for the regional expression patterns.

At present we have no compelling interpretation for the biological implication of the long-range enhancers that regulate the *Shh* expression. Our very recent study (Amano et al., 2009) demonstrated that chromosome dynamics is involved in *Shh* expression in the mouse developing limb bud. A long-range limb bud-specific enhancer, MFCS1, specifically interacts with the *Shh* promoter via change of the chromosome conformation. The *Shh* expression regulated by the epithelial lining-specific enhancers identified in this study is probably exerted by the similar mechanism. Therefore, it is of interest to investigate how the co-linear *Shh* expression in the epithelial linings along the anteroposteior body axis is regulated by orchestrated chromosome dynamics.

Indispensable role of MFCS4 in morphogenesis of pharyngeal apparatus

The anterior endoderm-derived pharynx is the region of the digestive tube anterior to the point where the respiratory tube branches off. Loss-of-function and misexpression experiments have

shown that the *Shh* pathway functions in the pharyngeal endoderm to generate multiple organs, including the pancreas, pituitary gland, parathyroid gland and jaw (Brito et al., 2006; Hebrok, 2003; Litingtung et al., 1998; Sbrogna et al., 2003; Treier et al., 2001). However, owing to severe abnormalities in the *Shh* KO mutant mouse, the function of *Shh* at later stages of the pharyngeal morphogenesis has been poorly understood. In this study, we examined a functional requirement for the pharyngeal epithelium-specific *Shh* enhancer, MFCS4, by targeted deletion of MFCS4 in mice. The result clearly showed that MFCS4 is essential for morphogenesis of the pharyngeal structures necessary for respiration and swallowing, including the soft palate, epiglottis and arytenoid. As shown in Table S2 in the supplementary material, relevance between domains of the reporter expression driven by the three epithelial *Shh* enhancers and the phenotype in the ΔMFCS4/ΔMFCS4 neonates shows a crucial role of the *Shh* pathway in the morphogenesis of the pharyngeal epithelium. Moreover, detailed characterization of the phenotype of the ΔMFCS4/ΔMFCS4 embryos suggested that the *Shh* pathway is involved in regional segmentation of the epithelial linings. This will be confirmed by future studies with knockout mutants of MRCS1 and MACS1.

We found that the 217 bp ultra-conserved sequence of MFCS4 has the four motifs for the transcription factors involved in the Pbx, Sox, Wnt and Notch signaling pathways. As we revealed that many genes encoding these transcription factors are expressed in the epiglottis and arytenoid swelling of the E13.5 embryos, in which we observed the highest level of the MFCS4-mediated *lacZ* reporter expression, these transcription factors most probably act as direct upstream regulators of the *Shh* expression in the pharyngeal epithelium.

Here, we need to pay attention to influence of the adjacent genes *Lmbr1* and *Rnf32* on the phenotype of the ΔMFCS4/ΔMFCS4 embryos, because the deletion of MFCS4 may disrupt the functions of these genes. The *Lmbr1* null mutation has been reported previously (Clark et al., 2000). The *Lmbr1* KO homozygotes show limb defects, but they are viable and fertile. It suggests that disruption of the *Lmbr1* gene cannot cause crucial pharyngeal defects, as shown in the ΔMFCS4/ΔMFCS4 embryos. The *Rnf32* mutant mouse has not been reported thus far, and the function of this gene remains unclear. Thus, influence of the MFCS4 disruption on the pharyngeal phenotype is currently undeniable.

Evolutionary diversification of the enhancers

Morphological variation in the epithelial architecture is well exemplified in the evolutionary distant species (Botella et al., 2007; Brainerd and Owerkowicz, 2006; Delgado et al., 2005; Ichim et al., 2007; Iwasaki, 2002; Mitsiadis et al., 2003). We showed that the oral epithelium-specific enhancer MRCS1 is conserved in birds and reptiles at the almost same level as in mammals, whereas its homolog has not been identified in amphibians and teleost fishes. The lung-gut epithelium-specific enhancer MACS1 is conserved in amphibians, but not in teleost fishes. It is likely that the various epithelial derivatives evolved from a flat epithelial structure (Chuong and Edelman, 1985; Chuong et al., 2000), and that their shapes and sizes can be altered, depending on the timing and location of Shh signaling. Thus, regulation of *Shh* expression probably influences the architecture of the epithelial derivatives. At present, it is not clear how the oral and lung-gut epithelial architecture develop in the species without MRCS1 and MACS1. However, one possible explanation is that, in the past, MFCS4 specified *Shh* expression in both the ectoderm-derived oral cavity and the anterior endoderm-derived pharynx. Indeed, we found that mouse MFCS4 drives transient reporter expression in the dental placode (Fig. 2H). Alternatively, there may be unidentified, species-specific enhancers that regulate *Shh* expression in a tissue-specific manner in amphibians and teleost fishes. It would be of interest to explore whether such CNCSs exist within different teleost fishes or between amphibian and teleost fishes.

We are grateful to Drs Y. Katori, T. Kobayashi, S. Iseki and N. Wada for helpful comments on anatomy, to Drs T. Takada, S. Tanaka, N. Sakai, M. Shinya, H. Kokubo and M. Okabe for useful discussion throughout this study, and to Dr S. Kitajima for technical advice on ES cell manipulation. We thank Ms N. Yamatani, H. Nakazwa, A. Okagaki and Y. Kato for their kind technical support. We are also grateful to Dr P. Beachy for providing us with the *Shh* knockout mice, to Dr A. McMahon for providing the *Shh* probe and to Dr M. P. Scott for the *Ptch1* probe. This study was supported in part by grants-in-aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan. This paper is contribution number 2513 from the National Institute of Genetics, Japan.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/cgi/content/full/136/10/1665/DC1

References

Abbasi, A. A., Paparidis, Z., Malik, S., Goode, D. K., Callaway, H., Elgar, G. and Grzeschik, K. H. (2007). Human GLI3 intragenic conserved non-coding sequences are tissue-specific enhancers. *PLoS ONE* **2**, e366.

Amano, T., Sagai, T., Tanabe, H., Mizushina, Y., Nakazawa, H. and Shiroishi, T. (2009). Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47-57.

Bitgood, M. J. and McMahon, A. P. (1995). Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. *Dev. Biol.* **172**, 126-138.

Boffelli, D., Nobrega, M. A. and Rubin, E. M. (2004). Comparative genomics at the vertebrate extremes. *Nat. Rev. Genet.* **5**, 456-465.

Botella, H., Blom, H., Dorka, M., Ahlberg, P. E. and Janvier, P. (2007). Jaws and teeth of the earliest bony fishes. *Nature* **448**, 583-586.

Brainerd, E. L. and Owerkowicz, T. (2006). Functional morphology and evolution of aspiration breathing in tetrapods. *Respir. Physiol. Neurobiol.* **154**, 73-88.

Brito, J. M., Teillet, M. A. and Le Douarin, N. M. (2006). An early role for sonic hedgehog from foregut endoderm in jaw development: ensuring neural crest cell survival. *Proc. Natl. Acad. Sci. USA* **103**, 11607-11612.

Carninci, P. and Hayashizaki, Y. (1999). High-efficiency full-length cDNA cloning. *Methods Enzymol.* **303**, 19-44.

Chiang, C., Litingtung, Y., Lee, E., Young, K. E., Corden, J. L., Westphal, H. and Beachy, P. A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383**, 407-413.

Chuong, C. M. and Edelman, G. M. (1985). Expression of cell-adhesion molecules in embryonic induction. II. Morphogenesis of adult feathers. *J. Cell Biol.* **101**, 1027-1043.

Chuong, C. M., Patel, N., Lin, J., Jung, H. S. and Widelitz, R. B. (2000). Sonic hedgehog signaling pathway in vertebrate epithelial appendage morphogenesis: perspectives in development and evolution. *Cell Mol. Life Sci.* **57**, 1672-1681.

Clark, R. M., Marker, P. C. and Kingsley, D. M. (2000). A novel candidate gene for mouse and human preaxial polydactyly with altered expression in limbs of Hemimelic extra-toes mutant mice. *Genomics* **67**, 19-27.

Cobourne, M. T., Miletich, I. and Sharpe, P. T. (2004). Restriction of sonic hedgehog signalling during early tooth development. *Development* **131**, 2875-2885.

Dassule, H. R., Lewis, P., Bei, M., Maas, R. and McMahon, A. P. (2000). Sonic hedgehog regulates growth and morphogenesis of the tooth. *Development* **127**, 4775-4785.

Delgado, S., Davit-Beal, T., Allizard, F. and Sire, J. Y. (2005). Tooth development in a scincid lizard, Chalcides viridanus (Squamata), with particular attention to enamel formation. *Cell Tissue Res.* **319**, 71-89.

Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A. and McMahon, A. P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417-1430.

Epstein, D. J., McMahon, A. P. and Joyner, A. L. (1999). Regionalization of Sonic hedgehog transcription along the anteroposterior axis of the mouse central nervous system is regulated by Hnf3-dependent and -independent mechanisms. *Development* **126**, 281-292.

Frazer, K. A., Tao, H., Osoegawa, K., de Jong, P. J., Chen, X., Doherty, M. F. and Cox, D. R. (2004). Noncoding sequences conserved in a limited number of

mammals in the SIM2 interval are frequently functional. *Genome Res.* **14**, 367-372.
Ghanem, N., Jarinova, O., Amores, A., Long, Q., Hatch, G., Park, B. K., Rubenstein, J. L. and Ekker, M. (2003). Regulatory roles of conserved intergenic domains in vertebrate Dlx bigene clusters. *Genome Res.* **13**, 533-543.
Gong, S., Yang, X. W., Li, C. and Heintz, N. (2002). Highly efficient modification of bacterial artificial chromosomes (BACs) using novel shuttle vectors containing the R6Kgamma origin of replication. *Genome Res.* **12**, 1992-1998.
Goode, D. K., Snell, P., Smith, S. F., Cooke, J. E. and Elgar, G. (2005). Highly conserved regulatory elements around the SHH gene may contribute to the maintenance of conserved synteny across human chromosome 7q36.3. *Genomics* **86**, 172-181.
Hall, J. M., Hooper, J. E. and Finger, T. E. (1999). Expression of sonic hedgehog, patched, and Gli1 in developing taste papillae of the mouse. *J. Comp. Neurol.* **406**, 143-155.
Hall, J. M., Bell, M. L. and Finger, T. E. (2003). Disruption of sonic hedgehog signaling alters growth and patterning of lingual taste papillae. *Dev. Biol.* **255**, 263-277.
Haraguchi, R., Mo, R., Hui, C., Motoyama, J., Makino, S., Shiroishi, T., Gaffield, W. and Yamada, G. (2001). Unique functions of Sonic hedgehog signaling during external genitalia development. *Development* **128**, 4241-4250.
Haraguchi, R., Motoyama, J., Sasaki, H., Satoh, Y., Miyagawa, S., Nakagata, N., Moon, A. and Yamada, G. (2007). Molecular analysis of coordinated bladder and urogenital organ formation by Hedgehog signaling. *Development* **134**, 525-533.
Hebrok, M. (2003). Hedgehog signaling in pancreas development. *Mech. Dev.* **120**, 45-57.
Ichim, I., Kieser, J. and Swain, M. (2007). Tongue contractions during speech may have led to the development of the bony geometry of the chin following the evolution of human language: a mechanobiological hypothesis for the development of the human chin. *Med. Hypotheses* **69**, 20-24.
Ingham, P. W. and McMahon, A. P. (2001). Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* **15**, 3059-3087.
Iseki, S., Araga, A., Ohuchi, H., Nohno, T., Yoshioka, H., Hayashi, F. and Noji, S. (1996). Sonic hedgehog is expressed in epithelial cells during development of whisker, hair, and tooth. *Biochem. Biophys. Res. Commun.* **218**, 688-693.
Ito, T., Udaka, N., Yazawa, T., Okudela, K., Hayashi, H., Sudo, T., Guillemot, F., Kageyama, R. and Kitamura, H. (2000). Basic helix-loop-helix transcription factors regulate the neuroendocrine differentiation of fetal mouse pulmonary epithelium. *Development* **127**, 3913-3921.
Iwasaki, S. (2002). Evolution of the structure and function of the vertebrate tongue. *J. Anat.* **201**, 1-13.
Iwatsuki, K., Liu, H. X., Gronder, A., Singer, M. A., Lane, T. F., Grosschedl, R., Mistretta, C. M. and Margolskee, R. F. (2007). Wnt signaling interacts with Shh to regulate taste papilla development. *Proc. Natl. Acad. Sci. USA* **104**, 2253-2258.
Jeong, Y. and Epstein, D. J. (2003). Distinct regulators of Shh transcription in the floor plate and notochord indicate separate origins for these tissues in the mouse node. *Development* **130**, 3891-3902.
Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M. and Epstein, D. J. (2006). A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* **133**, 761-772.
Jeong, Y., Leskow, F. C., El-Jaick, K., Roessler, E., Muenke, M., Yocum, A., Dubourg, C., Li, X., Geng, X., Oliver, G. et al. (2008). Regulation of a remote Shh forebrain enhancer by the Six3 homeoprotein. *Nat. Genet.* **40**, 1348-1353.
Kaufman, M. H. (1992). *The Atlas of Mouse Development*. San Diego, CA: Academic Press.
Lettice, L. A., Horikoshi, T., Heaney, S. J., van Baren, M. J., van der Linde, H. C., Breedveld, G. J., Joosse, M., Akarsu, N., Oostra, B. A., Endo, N. et al. (2002). Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci. USA* **99**, 7548-7553.
Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E. and de Graaff, E. (2003). 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.
Li, C., Hu, L., Xiao, J., Chen, H., Li, J. T., Bellusci, S., Delanghe, S. and Minoo, P. (2005). Wnt5a regulates Shh and Fgf10 signaling during lung development. *Dev. Biol.* **287**, 86-97.
Litingtung, Y., Lei, L., Westphal, H. and Chiang, C. (1998). Sonic hedgehog is essential to foregut development. *Nat. Genet.* **20**, 58-61.
Maas, S. A. and Fallon, J. F. (2005). Single base pair change in the long-range Sonic hedgehog limb-specific enhancer is a genetic basis for preaxial polydactyly. *Dev. Dyn.* **232**, 345-348.
Makino, S., Masuya, H., Ishijima, J., Yada, Y. and Shiroishi, T. (2001). A spontaneous mouse mutation, mesenchymal dysplasia (mes), is caused by a deletion of the most C-terminal cytoplasmic domain of patched (ptc). *Dev. Biol.* **239**, 95-106.
McEwen, G. K., Woolfe, A., Goode, D., Vavouri, T., Callaway, H. and Elgar, G. (2006). Ancient duplicated conserved noncoding elements in vertebrates: a genomic and functional analysis. *Genome Res.* **16**, 451-465.
Mitsiadis, T. A., Cheraud, Y., Sharpe, P. and Fontaine-Perus, J. (2003). Development of teeth in chick embryos after mouse neural crest transplantations. *Proc. Natl. Acad. Sci. USA* **100**, 6541-6545.
Moore-Scott, B. A. and Manley, N. R. (2005). Differential expression of Sonic hedgehog along the anterior-posterior axis regulates patterning of pharyngeal pouch endoderm and pharyngeal endoderm-derived organs. *Dev. Biol.* **278**, 323-335.
Okubo, T., Pevny, L. H. and Hogan, B. L. (2006). Sox2 is required for development of taste bud sensory cells. *Genes Dev.* **20**, 2654-2659.
Rice, R., Connor, E. and Rice, D. P. (2006). Expression patterns of Hedgehog signalling pathway members during mouse palate development. *Gene Expr. Patterns* **6**, 206-212.
Riddle, R. D., Johnson, R. L., Laufer, E. and Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401-1416.
Roelink, H., Augsburger, A., Heemskerk, J., Korzh, V., Norlin, S., Ruiz i Altaba, A., Tanabe, Y., Placzek, M., Edlund, T., Jessell, T. M. et al. (1994). Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog expressed by the notochord. *Cell* **76**, 761-775.
Roessler, E., Ward, D. E., Gaudenz, K., Belloni, E., Scherer, S. W., Donnai, D., Siegel-Bartelt, J., Tsui, L. C. and Muenke, M. (1997). Cytogenetic rearrangements involving the loss of the Sonic Hedgehog gene at 7q36 cause holoprosencephaly. *Hum. Genet.* **100**, 172-181.
Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T. and Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mamm. Genome* **15**, 23-34.
Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. and Shiroishi, T. (2005). 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.
Santagati, F., Abe, K., Schmidt, V., Schmitt-John, T., Suzuki, M., Yamamura, K. and Imai, K. (2003). Identification of Cis-regulatory elements in the mouse Pax9/Nkx2-9 genomic region: implication for evolutionary conserved synteny. *Genetics* **165**, 235-242.
Sbrogna, J. L., Barresi, M. J. and Karlstrom, R. O. (2003). Multiple roles for Hedgehog signaling in zebrafish pituitary development. *Dev. Biol.* **254**, 19-35.
Schnabel, C. A., Selleri, L., Jacobs, Y., Warnke, R. and Cleary, M. L. (2001). Expression of Pbx1b during mammalian organogenesis. *Mech. Dev.* **100**, 131-135.
Shashikant, C. S., Bieberich, C. J., Belting, H. G., Wang, J. C., Borbely, M. A. and Ruddle, F. H. (1995). Regulation of Hoxc-8 during mouse embryonic development: identification and characterization of critical elements involved in early neural tube expression. *Development* **121**, 4339-4347.
St-Jacques, B., Dassule, H. R., Karavanova, I., Botchkarev, V. A., Li, J., Danielian, P. S., McMahon, J. A., Lewis, P. M., Paus, R. and McMahon, A. P. (1998). Sonic hedgehog signaling is essential for hair development. *Curr. Biol.* **8**, 1058-1068.
Sumiyama, K., Kim, C. B. and Ruddle, F. H. (2001). An efficient cis-element discovery method using multiple sequence comparisons based on evolutionary relationships. *Genomics* **71**, 260-262.
Treier, M., O'Connell, S., Gleiberman, A., Price, J., Szeto, D. P., Burgess, R., Chuang, P. T., McMahon, A. P. and Rosenfeld, M. G. (2001). Hedgehog signaling is required for pituitary gland development. *Development* **128**, 377-386.
Trokovic, N., Trokovic, R., Mai, P. and Partanen, J. (2003). Fgfr1 regulates patterning of the pharyngeal region. *Genes Dev.* **17**, 141-153.
Varjosalo, M. and Taipale, J. (2008). Hedgehog: functions and mechanisms. *Genes Dev.* **22**, 2454-2472.
Wilkinson, D. G. (1992). *In Situ Hybridization: A Practical Approach*. Oxford: Oxford University Press.
Woolfe, A. and Elgar, G. (2007). Comparative genomics using Fugu reveals insights into regulatory subfunctionalization. *Genome Biol.* **8**, R53.
Woolfe, A., Goodson, M., Goode, D. K., Snell, P., McEwen, G. K., Vavouri, T., Smith, S. F., North, P., Callaway, H., Kelly, K. et al. (2005). Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol.* **3**, e7.
Woolfe, A., Goode, D. K., Cooke, J., Callaway, H., Smith, S., Snell, P., McEwen, G. K. and Elgar, G. (2007). CONDOR: a database resource of developmentally associated conserved non-coding elements. *BMC Dev. Biol.* **7**, 100.
Yagi, T., Tokunaga, T., Furuta, Y., Nada, S., Yoshida, M., Tsukada, T., Saga, Y., Takeda, N., Ikawa, Y. and Aizawa, S. (1993). A novel ES cell line, TT2, with high germline-differentiating potency. *Anal. Biochem.* **214**, 70-76.
Yu, M., Wu, P., Widelitz, R. B. and Chuong, C. M. (2002). The morphogenesis of feathers. *Nature* **420**, 308-312.