NPAS4L is involved in avian hemangioblast specification

by Wei Weng, Hiroki Nagai, Sofiane Hamidi, and Guojun Sheng

Haematologica 2020 [Epub ahead of print]

Citation: Wei Weng, Hiroki Nagai, Sofiane Hamidi, and Guojun Sheng. NPAS4L is involved in avian hemangioblast specification. Haematologica. 2020; 105:xxx
doi:10.3324/haematol.2019.239434

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.
Title: NPAS4L is involved in avian hemangioblast specification

Running title: NPAS4L in avian hemangioblast specification

Authors: Wei Weng¹, Hiroki Nagai¹, Sofiane Hamidi¹ and Guojun Sheng¹,#
¹: International Research Center for Medical Sciences (IRCMS), Kumamoto University, Kumamoto, Japan
#: Author for correspondence:
Guojun Sheng, Ph.D.
International Research Center for Medical Sciences (IRCMS), Kumamoto University
2-2-1 Honjo, Chuo-ku, Kumamoto 860-0811, Japan
Email: sheng@kumamoto-u.ac.jp
Phone: 81-96-373-6874

TO THE EDITOR,

Vertebrate primitive hematopoietic and vascular development is regulated by a conserved set of transcription factors. Their common precursors, the hemangioblasts, express stem cell leukemia/T-cell acute lymphoblastic leukemia 1 (SCL/TAL1) (1) and Lim only protein 2 (LMO2) (2) in all vertebrate groups examined. Hemangioblast specification from nascent mesoderm was reported to be less conserved, with Ets variant 2 (ETV2) and Neuronal PAS-domain containing protein 4-like (NPAS4L) identified as its master regulator in mammals (3) and zebrafish (4), respectively. We show here that the ortholog of NPAS4L, but not of ETV2, is present in the avian genome. Chicken NPAS4L is expressed in hemangioblasts prior to SCL/TAL1 and LMO2. CRISPR-Cas9-mediated ectopic expression of endogenous NPAS4L leads to ectopic SCL/TAL1 and LMO2, as with ectopic expression of zebrafish NPAS4L. We propose that the ancestral amniote genome had both NPAS4L and ETV2 genes. The ETV2 gene was lost in the avian lineage without affecting direct transcriptional regulation of SCL/TAL1 and LMO2 by NPAS4L (5, 6). The NPAS4L gene was lost in the mammalian lineage, with its roles partially replaced by ETV2.

Text: Vertebrate primitive hematopoietic and vascular systems are derived from the mesoderm germ layer (7, 8). Lineage specification events taking place between gastrulation and the onset of circulation are controlled by a set of evolutionarily-conserved transcription regulators (8, 9). In birds (10-12), as in fish, amphibians and mammals (13-17), common progenitors of blood and endothelial cells (the hemangioblasts) start to express transcription factors SCL/TAL1 and LMO2 at Hamburger and Hamilton stage 4+ (HH4+) (18), soon after their exit from posterior primitive streak where ventral mesoderm cells originate. This is followed by FGFR-mediated segregation of blood and endothelial lineages and functional differentiation of blood cells starting from HH7 (10), mediated by a conserved set of transcription factors including SCL/TAL1, LMO2, GATA-binding factor 2 (GATA2), LIM domain-binding protein 1 (LDB1) and transcription factor E2A (E2A) (19). After the onset of circulation from HH12/13, the hemangioblast markers SCL/TAL1 and
LMO2 become restricted to the blood and endothelial lineages, respectively.

Hemangioblast specification from their mesoderm precursors was reported to involve divergent transcriptional regulation, with ETV2 in mammals (3, 20) and NPAS4L in zebrafish (4) as the main driver. ETV2 ortholog is present in the zebrafish genome, but its function was reported to be under the control of NPAS4L (4, 6). No NPAS4L ortholog has been identified in any mammalian species, suggesting that this gene is not involved in hemangioblast specification in mammals. Mammalian NPAS4, a homolog of NPAS4L, was able to rescue fish cloche (npas4l) mutant phenotypes (4). Duplication of the NPAS4 and NPAS4L genes, however, took place before the divergence of Actinopterygians (bony fishes) and Sarcopterygians (including tetrapods) and NPAS4 has not been associated so far with any aspect of vertebrate hematopoietic development, suggesting that these two genes have different biological functions involving separate molecular regulatory networks.

Since the mammals and birds are closely related both phylogenetically (Fig.1A) and ontogenetically (Fig.1B), we investigated whether avian NPAS4L and ETV2 genes are involved in early hematopoietic and vascular development. Molecular phylogenetic analysis indicated that an NPAS4L ortholog was present in the chicken (G. gallus) genome (in both galGal5 and galGal6 assemblies) (Fig.1C). Although this gene is annotated as NPAS4 in the current assembly, syntenic analysis (Fig.1C) clearly indicated that it was the ortholog of NPAS4L in fish and other vertebrate groups (viewable through search term “npas4” in the NCBI genome data browser https://www.ncbi.nlm.nih.gov/genome/gdv?org=gallus-gallus or the chicken FANTOM dataset browser http://fantom.gsc.riken.jp/zenbu/gLyphs#config=b1ZI1gUFZ6mHX6:4Gvxx). Phylogenetic analyses also showed that the NPAS4L gene is present in all other bird species with their genomes fully or partially assembled and in non-avian reptiles with their genomes assembled (Anolis lizard shown as an example in Fig.1C). In contrast, the ETV2 ortholog is missing in the entire avian lineage, and also in crocodiles and turtles, suggesting a loss of this gene before avian evolution. The ETV2 ortholog, however, was found in some of the reptilian lineages (e.g., lizards and snakes) (Fig.1C,D). Taken together, our phylogenetic analyses suggest that birds have the NPAS4L, but not the ETV2, gene in their genomes.

We next asked whether NPAS4L plays a role in early hemangioblast specification in chick as was shown in zebrafish. For this purpose, we generated an RNA whole-mount in situ hybridization (WISH) probe for chicken NPAS4L and performed WISH using embryos from stage HH3 (early gastrulation) to stage HH12 (onset of circulation). Expression of chicken NPAS4L was detected in territories marking nascent hemangioblasts in ventral mesoderm (Fig.2A) from stage HH3+, the earliest among all hemangioblast-specific genes (e.g., SCL/TAL1 and LMO2 expression starts from stage HH4+). This observation was confirmed by WISH using left-right bisected embryos, with the left half stained for NPAS4L and the right half stained for SCL/TAL1 and Chordin (Fig.2C). Paraffin-sectioning of stained embryos (Fig.2B) showed that NPAS4L-positive cells are located in a subset of the mesoderm germ layer that will give rise to blood and endothelial cell lineages (red arrows; germ layers marked by arrowheads and brackets), as we had previously reported (10, 21). NPAS4L expression levels peaked at HH7 and declined soon afterwards (Fig.2A), suggesting that this gene is specifically and transiently involved in hemangioblast formation, but not in their differentiation.

We have previously generated the chicken promoterome database, spanning the entire 21-
day period of embryonic development (22). When we searched this database (http://fantom.gsc.riken.jp/zenbu/), NPAS4L was shown (Fig.3A) to be only expressed in a narrow time window with its peak expression at HH7, consistent with the WISH data. To evaluate its molecular function, we used CRISPRa (CRISPR-mediated gene activation; also known as CRISPR-on) (23) to ectopically express this gene. CRISPRa utilizes a modified Cas9 protein (with dead nuclease activity and fused with ten copies of VP16 transactivation domain) to recruit transcriptional machinery to targeted promoters mediated by single guide RNA (sgRNA). We had previously confirmed the effectiveness of CRISPRa system in the avian model by taking advantage of the single-nucleotide level resolution in transcription start site (TSS) mapping (22). Four sgRNA sequences located within the 500-base pair region preceding the NPAS4L TSS were selected (Fig.3B) (for interactive view of NPAS4L TSS, use the link http://fantom.gsc.riken.jp/zenbu/glyphs/#config=b1zI1gUFZ6mHX6-4Gvxr:loc=galGal5::chr3:4300587..4304021+) and cloned into expression construct pAC154-dual-dCas9VP160-sgExpression (Addgene #48240). Mesoderm precursors in the streak in HH2/3 embryos were targeted for electroporation [see (19) for electroporation protocol] with these four sgRNA expression constructs together with marker GFP expression construct (Fig.3C), and electroporated embryos were assessed for ectopic expression of endogenous NPAS4L and of two hemangioblast markers SCL/TAL1 and LMO2. NPAS4L CRISPRa constructs were able to ectopically activate endogenous NPAS4L (oval areas; Fig.3D) (11/12; 92%) in regions that are normally NPAS4L-negative (Fig.2A), as well as hemangioblast markers SCL/TAL1 (9/25; 36%) and LMO2 (6/21; 29%) (Fig.3E, oval areas in left two panels), albeit with reduced efficiency. Interestingly, similar inductive effect (5/13 for LMO2 and 4/9 for SCL/TAL1) was observed when we used zebrafish NPAS4L expression construct (4) (cloned into the pCAGGS expression vector) (Fig.3E, oval area in right panel), supporting partial molecular conservation between the zebrafish and chicken NPAS4L genes.

In conclusion, we present evidence that during chicken early development, NPAS4L, instead of ETV2, is involved in hemangioblast formation. Data from our molecular phylogenetic analyses support the hypothesis that both the NPAS4L and ETV2 genes were present in the common reptilian ancestor and likely also in the common amniote ancestor (Fig.3F). Conclusive answer to their epistatic relationship, however, requires additional evidence from gain-of-function of ETV2 (e.g., using a reptilian ETV2 ortholog) and loss-of-function of NPAS4L (e.g., through CRISPR-mediated transcription inhibition) studies. In birds and other reptilian lineages which lack the ETV2 ortholog in their genome, it is possible that other ETS family genes have been co-opted to play hemangioblast-specific roles of ETV2. Because ETV2 and NPAS4L are transcription factors with different DNA binding specificities and co-factor requirements, it remains to be shown how ETV2 took over molecular functions of NPAS4L during early mammalian evolution.

Author contributions
GS, WW, HN and SH designed the experiments. WW, HN, SH and GS performed the experiments. GS, WW, HN, SH analyzed the data. GS wrote the paper, all authors agreed on the content.

Conflict of interest disclosure
The authors declare that there is no competing financial interest.
References

Figure legends

Figure 1: The chicken genome has NPAS4L, but not ETV2 ortholog. A) A simplified vertebrate phylogenetic tree. B) Schematic view of blood and endothelial cell differentiation from mesoderm precursors in the streak. NPAS4L and ETV2 are proposed to function during hemangioblast specification in the ventral mesoderm. C) The chicken genome has the NPAS4L orthologous gene flanked by KLRHDC3 gene on one side and RRP36 and TMEM121L genes on the other. Similar syntenic organization is seen in lizard A. carolinensis and zebrafish D. rerio. These genes are missing in mammalian genomes. The chicken genome does not have ETV2 ortholog. The lizard genome has ETV2 and FLI1B as in the zebrafish genome. Mammals have ETV2, but not FLI1B. It is to be noted, not shown here, that vertebrate genomes have three copies of such tandemly duplicated ETS family genes. In addition to the ETV2-FLI1B couplet which is the least conserved, the other two couplets (ETS1-FLI1 and ETS2-ERG) are well-conserved. D) Summary of presence and absence of NPAS4L, ETV2, SCL/TAL1, LMO2 and NPAS4 genes in various vertebrate groups. The following protein sequences were used for comparison. For SCL/TAL1: NP_001274276.1 (human), NP_001274317.1 (mouse), XP_001374963.1 (opossum), DNA clone XX-200B24 (platypus), NP_990683.1 (chicken), XP_030427307.1 (desert turtle, sequence in Chinese soft-shell turtle is incomplete), XP_008114556.1 (lizard), NP_001081746.1 (Xenopus) and NP_998402.1 (zebrafish); for LMO2: AAH42426.1 (human), AAH57880.1 (mouse), XP_027693653.1 (opossum), XP_028917173.1 (platypus), AAL78036.1 (chicken), XP_030415938.1 (turtle), XP_003225211.1 (lizard), NP_001081112.1 (Xenopus) and AAH93136.1 (zebrafish); for ETV2: NP_055024.2 (human), NP_031985.2 (mouse), XP_007491908.1 (opossum), XP_028921116.1 (platypus), XP_008119144.1 (lizard), NP_001089600.1 (Xenopus) and NP_001032452.1 (zebrafish); for NPAS4L: EntrezID 101750093 (chicken), XP_008103134.1 (lizard), XP_008165306.1 (turtle) and NP_001316841.1 (zebrafish).

Figure 2: Chicken NPAS4L gene expression during hemangioblast specification. A) WISH of NPAS4L from HH3 to HH10. Top: white background for expression visualization; bottom: dark background for stage visualization. Black lines indicate section levels shown in panel B. B) Section of embryos shown in A. NPAS4L expressing cells indicated by red arrows. Germ layers marked by black arrowheads (ectoderm and endoderm) and brackets (mesoderm). C) Chicken NPAS4L is expressed starting from HH3+, earlier than SCL/TAL1. Embryos were fixed and processed to the pre-hybridization step (left panels) and were cut into left (stained for NPAS4L) and right (stained for SCL/TAL1 and Chordin together) halves. Stained half embryos were then photographed together (middle panels: white background showing both halves; right panels: dark background showing both halves). Chordin expression was used to mark precise embryo stages. At HH3+ (top row) and HH4 (middle row), NPAS4L is expressed and SCL/TAL1 is not expressed. At HH5 (bottom row), both NPAS4L and SCL/TAL1 are expressed.

Figure 3: NPAS4L is involved in chicken hemangioblast specification. A) A screenshot of NPAS4L locus in chicken promoterome database (see text for web link). NPAS4L transcription start site (TSS) is indicated by red label and black arrow. TSS activity levels at different developmental stages are shown at the bottom. The highest expression is seen at HH7. B) Design of sgRNA for chicken NPAS4L CRISPRa. Mapped TSS is TCAGCAGG (underlined). Preceding 500 bp of promoter region is shown, with sgRNA sequences highlighted in red and PAM sequences in blue. C) Schematic diagram of how embryos are electroporated and cultured. D) NPAS4L CRISPRa constructs activate endogenous NPAS4L
expression ectopically (oval). Top: NPAS4L expression only; bottom: NPAS4L expression together with anti-GFP staining marking the electroporated territories (brown). E) NPAS4L CRISPRa constructs activate endogenous LMO2 (left) and SCL/TAL1 (middle) ectopically (oval). Zebrafish NPAS4L is also capable of activating hemangioblast markers (SCL/TAL1 shown in right panel, oval). F) Hypothetic scenario of hemangioblast specification in the ancestral amniote and ancestral reptile. It is proposed that the zebrafish scenario (both NPAS4L and ETV2 genes are present, with NPAS4L functioning upstream of ETV2) is the default one. Mammals have lost NPAS4L; and birds have lost ETV2.