

Variations in the poly-histidine repeat motif of HOXA1 predispose individuals to bicuspid aortic valve

Gaëlle Odelin

Aix Marseille Univ <https://orcid.org/0000-0002-1395-7114>

Adèle Faucherre

Universtités Montpellier,

Damien Marchese

Université catholique de Louvain

Amelie Pinard

Aix Marseille Univ

Emilie Faure

Aix Marseille Univ

Marine Herbane

Aix Marseille Univ

Alexis Théron

Aix Marseille Université

Jean-François Avierinos

Aix Marseille Université

Chris Jopling

Institut de Génomique Fonctionnel

Gwenaëlle Collod-Bérout

Aix Marseille Université <https://orcid.org/0000-0003-4098-6161>

René Rezsöhazi

U C Louvain <https://orcid.org/0000-0003-4253-7431>

Stephane Zaffran (✉ stephane.zaffran@univ-amu.fr)

Aix Marseille Univ <https://orcid.org/0000-0002-0811-418X>

Article

Keywords: Valve defect, genetics, bicuspid aortic valve, HOXA1

Posted Date: March 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1415862/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on March 20th, 2023. See the published version at <https://doi.org/10.1038/s41467-023-37110-x>.

Abstract

Bicuspid aortic valve (BAV) is the most common cardiovascular malformation (0.5–1.2% of the population) and is often associated with premature aortic valve stenosis or insufficiency, aortic aneurysm and other congenital cardiac heart defects. Although highly heritable, few causal mutations have been identified in BAV patients. Here, we report the association of novel variants in the transcription factor HOXA1 with BAV in humans. Targeted sequencing of *HOXA1* in a cohort of 333 BAV patients identified rare indel variants in the homopolymeric histidine tract of HOXA1. *In vitro* analysis revealed that disruption of the histidine repeat motif causes a significant reduction in the half-life of the protein, and that these variants are associated with a defective transcriptional activity of the HOXA1 protein. Targeting the zebrafish *hoxa1a* ortholog *in vivo* resulted in aortic valve defects that could be rescued by expressing the wild-type human *HOXA1*. Furthermore, expression of *HOXA1* Histidine variants in zebrafish also impacted aortic valve development indicating a dominant negative effect of these mutated proteins. Lastly, homozygous knockout mice for *Hoxa1* develop a BAV phenotype, which is associated with a very small, rudimentary non-coronary leaflet. Genetic lineage analysis indicates that this defect is caused by a strong reduction of mesenchymal cells in the intercalated cushion due to a failure of neural crest cell migration toward the heart which is consistent with our transcriptomic analysis showing a down-regulation of premigratory and migratory neural crest markers in *Hoxa1*^{-/-} compared with control embryos. Together, these findings indicate that variants causing HOXA1 dysfunction play a significant role in the genetic cause of BAV in humans.

Introduction

Bicuspid aortic valve (BAV) is the most common congenital heart defect (CHD) with an overall frequency of 0.5-2%¹ and a third of those affected develop severe complications including aortic stenosis and regurgitation, endocarditis, ascending aortic aneurysm/dissection^{2,3}. BAV is defined as a spectrum of altered aortic valves presenting with two functional cusps and less than three zones of parallel apposition without fusion⁴. BAV may appear as either an isolated defect or it can be associated with other cardiac defects such as coarctation of the aorta, ventricular septal defect, and interruption of the aortic arch^{5,6}.

A normal tricuspid aortic valve is formed by three cusps or leaflets (right (R) and left (L) coronary and non-coronary (NC) leaflets), three sinuses at the aortic root, a commissure between each adjacent leaflets and an interleaflet triangle between each adjacent sinuses⁷. The BAV, which occurs when the aortic valve has two leaflets, has a variability in leaflet fusion phenotypes. Three major types of BAV are distinguished depending on the number of raphe (fusion), the spatial position of leaflets and functional status of the valve⁴. The categories of BAV are type 0 (valve with no raphe), type 1 (valve with one raphe) and type 2 (valve with two raphes). The most frequent BAV types are those with one raphe positioned between the right and left coronary (R-L) leaflets (~ 70%) and those with a raphe between the right and non-coronary (R-NC) leaflets (~ 35%). A purely BAV is when there is no raphe (type 0).

The aortic valve forms in the outflow tract region of the developing heart. As the outflow tract develops, the endocardial cushions become populated with mesenchymal cells, which are derived from different sources. A proportion of the cells originate from endothelial cells, which detach from one another to yield mesenchymal cells *via* a process named endothelial-to-mesenchymal transition (EndoMT)⁸. Neural crest cells are also responsible for populating the endocardial cushions and patterning the arterial valve leaflets⁹⁻¹¹. The last source of mesenchymal cells comes from the second heart field (SHF), which has been observed during the formation of the intercalated cushion^{12,13}. In the aortic valve, the right and left coronary leaflets derive from the main cushions (Neural crest/endothelial derived) while the non-coronary leaflet originates exclusively from the superior cushion (Neural crest/endothelial/SHF derived).

Genetic studies have established that BAV is a heritable trait, with about 9% of prevalence in first-degree relatives and up to 24% in larger families¹⁴. However, only a few mutations, particularly in the *NOTCH1* and *GATA5* genes, have been directly linked to BAV in humans^{3,15-17}. More recently, mutations in the *ROBO4* gene have also been shown to predispose individuals to BAV and thoracic aortic aneurysm¹⁸. BAV has also been reported in a number of mouse mutants, including *Nos3*, *Nkx2-5*, *Gata5*, *Krox20*, *Robo4*, *Alk2* and *Hoxa1*¹⁹. To determine whether deleterious *HOXA1* variants are associated with BAV in humans, we performed a targeted sequencing of this gene in a local cohort of 333 BAV patients. In this study we identify novel indel variants in the homopolymeric histidine tract of the HOXA1 protein in BAV patients. We show that modifying the length of the histidine repeat motif results in a decrease of protein half-life and produces a dominant negative HOXA1 variant. Furthermore, expressing dominant negative *HOXA1* indel variants in zebrafish results in defective aortic valve development. Lastly, we show that BAV present in *Hoxa1*-null mice is associated with a very small, rudimentary non-coronary leaflet caused by a reduced number of mesenchymal cells into the developing intercalated cushion.

Results

In total, 333 BAV unrelated patients were screened for genetic variability in *HOXA1*. The clinical characteristics of the study participants are summarized in Table 1. As previously reported²⁰, there are more men (225) presented in this cohort compared with women (80; sex ratio = 73.8%). Also, the average age, 54.1 ± 15.6 years, is slightly higher than previously reported for BAV cohorts²¹. Within our cohort the most common BAV morphology was type 1 with fusion between the left and right coronary cusps (R-L type) accounting for 69.2% of BAV. Additionally, 8.5% of the cases were considered as familial based on the diagnosis of bicuspid valves and/or aortic aneurysm within a patient pedigree.

We identified several potential pathogenic variants which alter the length of the poly-histidine tract present in HOXA1. A frequent variant we identified was the p.Arg73His (c.218G > A), in a poly-histidine tract (7.53×10^{-1} frequency in gnomAD) which alters the reference sequence His₆₅His₆₆His₆₇His₆₈His₆₉His₇₀His₇₁His₇₂**Arg**₇₃His₇₄ (WT^G; Table 2; Fig. 1) to result in an uninterrupted stretch of 10 histidines His₆₅His₆₆His₆₇His₆₈His₆₉His₇₀His₇₁His₇₂**His**₇₃His₇₄ (hereafter referred to as WT^A) (Table 2; Fig. 1). We also identified patients who displayed additional mutations which further

altered the stretch of poly-histidine (Table 2). The majority of length variations (n = 22) corresponds to the deletion of 1 histidine associated with the c.218G > A polymorphism (c.[213_215delCCA;218G > A]; named -1His). We also found 5 patients with an insertion of 1 histidine. Among these 5 patients 4 had an uninterrupted stretch (c.[213_215dup;218G > A]; named +1His) and 1 had a 9His1Arg1His stretch (c.[213_215dup;218G]; named +1His^{Arg}). Lastly, we found that 2 patients harboured a deletion of 3 histidines (1 homozygous and 1 heterozygous; c.[207_215delCCACCACCA;218G] named -3His^{Arg}). Reported frequencies in gnomAD for these insertions and deletions are in accordance with BAV frequency in the general population (Table 3). Clinical descriptions of all patients can be found in supplementary Table 1. Taken together our data indicates that rare modifications in the length of the poly-histidine stretch of HOXA1 are present in BAV patients.

To evaluate the functional consequences of these variants, we quantified protein levels in HEK293T cells transfected with either wild-type or mutated *HOXA1* cDNA constructs. Cycloheximide treatment revealed that the wild-type HOXA1 protein is very stable with a half-life of more than 24 h. Since no difference was detected between WT^A or WT^G HOXA1 protein, we chose to use WT^A HOXA1 as the wild-type reference for all experiments. In contrast, HOXA1 proteins harboring the different variants altering the poly-histidine stretch displayed a markedly reduced half-life ($t_{1/2}$ = 5h versus >24h, Fig. 2a,b). Next, we addressed whether variant HOXA1-dependent protein reduction was a result of proteasome-mediated degradation. Indeed, treating *HOXA1* transfected cells with the 20S proteasome inhibitor MG132 (7 μ M) for 20 h, dramatically increased the half-life of all the variants (Fig. 2c,d). Therefore, our results indicate that variations in the length of the homopolymeric histidine tract of HOXA1 drastically reduces protein half-life through degradation by a proteasome-dependent degradation pathway.

To investigate the impact of the histidine tract variations on the transcriptional activity of HOXA1, we used a physiological target enhancer derived from mouse gene *EphA2*. In mice, expression of *EphA2* is detected in the hindbrain and, in particular, its expression in the rhombomere "r4" is dependent on *Hoxa1*²². Based on this we assessed the activity of wild-type and mutated HOXA1 proteins using a luciferase reporter gene controlled by the "r4" enhancer of the *EphA2* gene (*EphA2-r4*). PBX1a and PREP proteins are known to be critical cofactors of HOXA1 and to play an important role in HOXA1-mediated transcriptional activation^{23,24}. HEK293T cells were co-transfected with different vectors for each of the HOXA1 proteins in conjunction with PBX1a and PREP and the *EphA2-r4* luciferase reporter plasmid (Fig. 2e,f). In comparison to wild-type HOXA1, all HOXA1 variants showed a significant decrease in transcription activity (Fig. 2e). These results indicate that variations in the length of the HOXA1 histidine motif impairs transactivation activity. Because the majority of BAV patients are heterozygous for their respective histidine variants, we next sought to determine whether these variants are associated with a dominant negative effect. To address this hypothesis, we co-transfected wild-type HOXA1 with either +1His, +1His^{Arg}, -1His, -1His^{Arg} or -3His^{Arg} HOXA1 variants. Consequently, we were able to determine that there was a significant decrease in wild-type HOXA1 transcriptional activity upon co-transfection with any of the histidine variants (Fig. 2f). Taken together these data indicate that the histidine variants act as dominant negatives.

The *in vivo* effects of the *HOXA1* variants were evaluated using zebrafish larvae. Zebrafish has one *HOXA1* ortholog, *hoxa1a* (Extended data Fig. 1)²⁵. *Hoxa1a* is expressed in the midbrain and anterior hindbrain during early zebrafish development, which corresponds to a more rostral extension of expression than for *HOXA1* in mammals (Extended data Fig. 1)²⁶. Moreover, functional data suggest that a paralog, *hoxb1b*, may actually perform the same hindbrain patterning role as the mouse *Hoxa1*²⁵. Therefore, we targeted both *hoxa1a* and *hoxb1b* using antisense morpholino oligonucleotides (MO) targeting the AUG start site to block translation. At 3 days post-fertilization (dpf), both *hoxa1a* and *hoxb1b* morphants displayed gross developmental defects including a heart edema (Extended data Fig. 2). We also analyzed the developing aortic valves to determine whether *hoxa1a* or *hoxb1b* is involved in this process. To achieve this, arterial valves were imaged at 7 dpf using two-photon microscopy as previously described²⁷. Wild-type larvae developed normally with two defined leaflets (Fig. 3a,d). In comparison, the valves in *hoxa1a* and *hoxb1b* morphants are highly dysmorphic and appear to be enlarged and mis-shaped (Fig. 3b,e). To validate that the morphant phenotype was a result of mRNA knockdown and not due to non-specific effects, we also targeted genomic *hoxa1a* or *hoxb1b* sequences using "crisprant" technology²⁸. Briefly, four non-overlapping gRNAs targeting either *hoxa1a* or *hoxb1b* were injected into single cell zebrafish embryos resulting in F0 knockouts for both genes (crisprants). Consistent with our observations in morphants, aortic valve development was also defective in *hoxa1a* and *hoxb1b* crisprants (Extended data Fig. 3). We next sought to rescue the *hoxa1a* and *hoxb1b* morphants using human WT *HOXA1* mRNA which is not recognized by either morpholino. We found that co-injection of human *HOXA1* mRNA at 25 pg with *hoxa1a*-MO could rescue the valvular defects observed in *hoxa1a* morphants (Fig. 3c,g). Interestingly, human *HOXA1* mRNA could not rescue the *hoxb1b*-MO phenotype (Fig. 3f,h; Extended data Fig. 4). These data indicate that loss of *hoxa1a* results in defective valve development which can be reversed by human wild-type *HOXA1* mRNA.

Next, we evaluated whether the expression of the dominant negative human *HOXA1* variants in the zebrafish could also affect aortic valve development. To achieve this, we injected either the WT human *HOXA1* or the variants at 5 pg or 25 pg of mRNA (Fig. 3i-n; Extended data Fig. 5). Expressing any of the dominant negative variants (+1His, -1His or -3His^{Arg}) also had a significant impact on aortic valve development when compared to WT *HOXA1* expression (Fig. 3n). Our results therefore show that expression of either +1His, -1His or -3His^{Arg} *HOXA1* disrupts aortic valve development in zebrafish.

Studies in humans found severe cardiovascular anomalies including tetralogy of Fallot, ventricular septal defect and BAV in patients carrying a homozygous truncating mutation in *HOXA1* [Bosley-Salih-Alorainy (BSAS)²⁹; Athabaskan Brainstem Dysgenesis Syndrome (ABDS)³⁰]. Interestingly, BAV has also been reported in *Hoxa1* knockout (*Hoxa1*^{-/-}) mice with a limited description³¹. Strikingly, however, the expression pattern of *Hoxa1* has been reported in cell types that contribute to cardiovascular development including neural crest and second heart field (SHF) cells, but not in the developing heart *per se*³¹⁻³⁵. Therefore, we further examined *Hoxa1*-null mice during development to better characterize BAV phenotype. Since the leaflets of the aortic valve form by a progressive excavation of the endocardial cushions after embryonic day (E) E12.5⁷, we examined *Hoxa1*^{-/-} embryos after this stage. At E13.5, we

observed the presence of three leaflets in both wild-type and *Hoxa1*^{-/-} mice, however, the non-coronary leaflet appeared smaller in *Hoxa1*^{-/-} (n = 14) compared to control (n = 3) embryos (Fig. 4b-e). At E15.5, one out of four *Hoxa1*^{-/-} hearts displayed abnormal aortic valve morphology with minor non-coronary leaflet (Extended data Fig. 6). At E18.5, we found that 27% of *Hoxa1*^{-/-} embryos (3/11) had BAV (Fig. 4a), consistent with a previous report³¹. Anatomical examination of *Hoxa1*^{-/-} aortic valves revealed a BAV with equal size between the two leaflets (Fig. 4g). Furthermore, three-dimensional (3D) reconstruction revealed a persistent small non-coronary leaflet in *Hoxa1*^{-/-} compared to wild-type hearts (Fig. 4h,i). No other structural abnormalities were evident at the level of the other valves or the septa. Based on this observation, we suggest that the persistence of a small non-coronary leaflet should also be considered during anatomical examinations of BAV patients.

The development of the arterial valves is characterized by the contribution of multiple cell types including derivatives of the endothelial, neural crest and SHF cells which are all crucial for proper valve development^{10-13,36}. We first assessed whether the number of cells in the aortic valve leaflets was affected in *Hoxa1* mutant mice (Fig. 5a-c). At E13.5, we found that the number of cells in the non-coronary and left coronary leaflets was significantly decreased in *Hoxa1*^{-/-} compared to control littermates (Fig. 5c; *p* < 0.001). Consistently, we observed a decrease in the relative surface of the aortic valve of *Hoxa1*^{-/-} mice compared to controls (Extended data Fig. 7; *p* < 0.001). However, cell density was unaffected in *Hoxa1*^{-/-} mice suggesting that the reduction in the number of cells is not compensated by the growth of these cells (Extended data Fig. 7).

To further analyze the contribution of neural crest, endothelial, and SHF-derived cells, we performed genetic lineage-tracing experiments using *Wnt1-Cre*, *Tie2-Cre* and *Tnnt2-Cre* transgenic mice, respectively (Fig. 5d-l). Quantification of *Wnt1-Cre*-labeled cells showed a significant decrease of neural crest derived cells in *Hoxa1*^{-/-} aortic valve leaflets (Fig. 5d-f; *p* = 0.004). Interestingly, the non-coronary leaflet was more affected than the right and left coronary leaflets (Fig. 5f; *p* = 0.008). Similarly, quantification of *Tie2-Cre;Rosa*^{tdTomato}-positive cells showed a significant decrease of endothelial derived cells in the non-coronary leaflet in *Hoxa1*^{-/-} mice (Fig. 5g-i; *p* = 0.002). Genetic lineage-tracing analysis of *Tnnt2-Cre;Rosa*^{tdTomato} mice revealed a decrease in the number of SHF-derived mesenchymal cells in the non-coronary leaflet in *Hoxa1*^{-/-} compared to wild-type littermates (Fig. 5j-l; *p* = 0.01). On the contrary, we were unable to detect any significant difference in the number of endothelial- and SHF-derived cells in the right and left coronary leaflets in *Hoxa1*^{-/-} compared to wild-type littermates (Fig. 5i,l). Taken together these findings indicate that the absence of *Hoxa1* specifically affects the ability of mesenchymal cells to populate the non-coronary leaflet of the aortic valve irrespective of their origin.

To explore the mechanisms leading to the BAV phenotype observed in absence of *Hoxa1*, we collected tissue from the pharyngeal (branchial) regions of *Hoxa1*^{-/-} and wild-type embryos for RNA-seq analysis. Since previous studies revealed that neural crest and SHF cells located in the 3rd to 6th pharyngeal arches contribute to cardiovascular development we chose to collect embryos at the early E9.5 stage

when these cells are actively migrating toward the developing heart (Fig. 6a). Principal Component Analysis (PCA) indicated that samples correlate depending on their genotype (Extended data Fig. 8). We identified 381 up-regulated genes and 205 down-regulated genes in the pharyngeal region of *Hoxa1* mutant embryos (Fig. 6b; Supplementary Table 2). We selected genes significantly down-regulated and up-regulated in *Hoxa1* mutants ($\log_2(\text{fold change})$ $p > 0.2$, $p < 0.05$) to perform GO analysis and found that down-regulated genes were enriched in biological processes involved in “cell fate commitment”, “epithelial cell proliferation” and “mesenchymal to epithelial transition” (Fig. 6c). In contrast, up-regulated genes were enriched in biological processes involved in “pattern specification process”, “cellular response to growth” and “negative regulation of cellular movement”. Interestingly, genes involved in neural crest migration were enriched in the list of down-regulated genes³⁷. Thus, we found that genes involved in premigratory and migratory neural crest cells such as *Crabp1*, *Tfap2a* (*AP-2a*) and *Sox10* were down-regulated in *Hoxa1*^{-/-} embryos (Fig. 6d; Fig. 7; Extended data Fig. 9). Consistent with abnormal specification of neural crest cells, we found that genes normally activated in the preplacodal region lateral to the neural crest domain, such as *Eya1* and its activators *Dlx5/6* which act as lateral neural border specifiers were down-regulated (Supplementary Table 2). Interestingly, *Dlx1* which is expressed in the entire mesenchyme of pharyngeal arches was also down-regulated in *Hoxa1*^{-/-}. To further characterize the expression of genes involved in neural crest cells migration we performed RNAscope fluorescent in situ hybridization (RNA-FISH) and immunofluorescent staining. At E9.5, RNA-FISH with *Sox10* probes showed that migration of neural crest cells is largely compromised (Fig. 7a,b). Interestingly, among the upregulated genes we identified several posterior *Hox* genes suggesting that *Hoxa1* function is necessary to establish or maintain the expression patterns of posterior *Hox* genes (Fig. 6d). In addition, we also found that *Robo1* and *Robo2*, which have been shown to affect neural crest cell migration^{38,39}, were upregulated in *Hoxa1*^{-/-} compared to controls (Fig. 6d, 7c,d). All together these data are consistent with previous studies showing that neural crest cell migration is affected when *Hoxa1* function is lost, a critical process for normal aortic valve formation¹⁰.

Hoxa1 is strongly expressed in the posterior hindbrain where cardiac neural crest cells arise^{32,35}. To determine the contribution of *Hoxa1*-positive cell to the formation of the outflow tract and the aortic valve, we performed genetic lineage-tracing experiment. We have already utilized *Hoxa1-enhIII-cre* transgenic mice to demonstrate that *Hoxa1* is expressed in the SHF progenitors that contribute to the atria and part of the outflow tract³². At E9.5, we found X-gal positive cells in the migrating neural crest cells that invade the pharyngeal region (Fig. 7g). Consistently, we detected cells from the *Hoxa1*-lineage in the outflow tract of E12.5 embryos (Fig. 7h). Moreover, *Cre*-labeled cells were detected in the arterial valves of E13.5 and E18.5 embryos (Extended data Fig. 10). These results confirm that *Hoxa1* is expressed in neural crest cells that contribute to development of the arterial valves.

Next, we investigated how loss of function of *Hoxa1* affects neural crest cells. Analysis of *Wnt1-Cre;Rosa^{lacZ}* in *Hoxa1*^{-/-} embryos revealed significant defects in cardiac neural crest cells migration (Fig. 7i,j). At E9.5, two discrete streams of *Wnt1-cre*-labelled cells were observed in control embryos (Fig. 7i), however, these streams were abnormal in *Hoxa1*^{-/-} embryos particularly into the 3rd to 6th

pharyngeal arches (Fig. 7j). Immunostaining for *Tfap2*, a marker of migrating neural crest cells⁴⁰, confirmed abnormal numbers of cells in the 3rd pharyngeal arch (Fig. 7e,f). Recently, we found that the -31/-23.5 *Krox20/lacZ* enhancer element labels a sub-population of cardiac neural crest cells migrating from rhombomere 5-6, which contributes to the development of the aortic valve¹¹. Based on these data, we generated *Hoxa1*^{-/-} mice that carry the -31/-23.5 *Krox20/lacZ* transgene to specifically examine this sub-population, and found that loss of *Hoxa1* results in aberrant migration of neural crest cells that normally migrate toward the outflow tract where they participate to aortic valve development (Fig. 7k,l). Taken together these results suggest that loss of *Hoxa1* function causes abnormal migration of neural crest cells that normally contribute to aortic valve leaflet development.

Discussion

In this study, we provide compelling evidence that heterozygous dominant negative mutations and homozygous loss of function of *HOXA1* are associated with BAV in human and mice, respectively. The spatiotemporal expression of *Hoxa1* in the mouse, the *Hoxa1* knockout mouse model and our RNA-seq data suggest a crucial role for *Hoxa1* in the migration of neural crest cells that contribute to the formation of the aortic valve leaflets and particularly the growth of the non-coronary leaflet.

Our study shows that length variation of the poly-histidine tract modifies the half-life and the activity of *HOXA1*. Single amino acid repeats (or homopolymeric tracts) are extremely important in eukaryotic proteomes⁴¹, and 15%-20% of human proteins contain at least one amino acid repeat motif. Among homopolymeric tracts, histidine repeats are relatively rare in human proteins⁴². However, poly-histidine tract is more frequently associated to developmental proteins⁴³. The variation of amino acid repeats in developmental transcription factors such as *PHOX2A/PHOX2B* and *LHX2/LHX9* has been shown to modify the transcriptional activity of these proteins similarly to *HOXA1*⁴⁴. Most significantly, extension or reduction of the histidine repeat in *HOXA1* protein act as dominant negative on the wild-type *HOXA1* protein function. This is consistent with the heterozygous condition found in patients with BAV. This dominant negative effect was also validated in the zebrafish model. Indeed, injection of *HOXA1* (-1His, +1His or -3His^{Arg}) mRNA in zebrafish embryos leads to abnormal arterial valve development, whereas injection of WT *HOXA1* mRNA has no effect (Fig. 3i-n). Interestingly, when we compare the proportion of aortic valve defects observed in zebrafish larvae, we observe that the strongest phenotype is associated with the largest reduction of histidine (Fig. 3n).

Variation of simple amino acid repeat tracts is often associated with human disorders^{45,46}. The list of developmental and degenerative diseases that are caused by expansion of unstable repeats keeps expanding⁴⁷. Repeat expansion diseases share number of similarities such as formation of protein aggregation and cell death. Disease-associated repeat expansions in transcription factors including *HOXA13*, *HOXD13*, and *RUNX2*, have been found to alter their capacity to interact with the transcription machinery to control gene expression⁴⁸. Therefore, in addition to affecting the stability of the protein,

length variation of the histidine repeat in HOXA1 could alter its interaction with other proteins involved in its transcriptional function.

Analysis of aortic valve defect observed in *Hoxa1* knockout mice reveals a novel morphological feature of BAV. Rather than only purely BAV with two leaflets and no raphe, our data demonstrate that, in *Hoxa1*-null mice, a small third cushion in non-coronary position is still present, although the other two leaflets appose one another across the lumen. The presence of such structure may compromise valvular function after birth. Interestingly, the anterior leaflet of the pulmonary valve, which arises from the intercalated cushion, is not affected. Although the contribution of SHF mesenchymal cells is equivalent in both intercalated cushions, the pulmonary anterior leaflet receives fewer neural crest derived cells than the aortic non-coronary leaflet^{10,12}. This difference is likely to have a morphogenetic impact on the developing aortic and pulmonary intercalated cushions which influence their differential predisposition to pathology.

We provide evidence that neural crest cells populating the intercalated cushion are essential to promote non-coronary leaflet development. Here we show that loss of *Hoxa1* function disrupts neural crest cell migration. Previous studies have shown that *Hoxa1*^{-/-} mice have cardiovascular defects including interrupted aortic arch type B, aberrant origin of the right subclavian artery and cervical aortic arch^{31,33,34}. Several studies have shown that BAV is often associated with aortopathy, aneurysmal dilation of the ascending aorta or coarctation of the aorta⁴⁹. Since BAV and aortopathy are observed in *Hoxa1*^{-/-} mice, this suggests that a deficiency in neural crest cells migration can be the common origin of these anomalies. Indeed, neural crest cells contribute to mesenchymal cells of the aortic valve leaflets but also to smooth muscle cells of the proximal thoracic aorta⁵⁰. Significantly, abnormal condensation of neural crest cells in outflow tract cushions during septation results in irregular sized leaflet and subsequently to BAV¹⁰. In addition, we have previously shown that excess of neural crest-derived cells within the aortic valve leaflets leads to BAV¹¹. Therefore, we believe that reduction of neural crest cells in *Hoxa1*-null mice is the main cause of abnormal growth of the intercalated cushion and subsequent non-coronary leaflet. Unlike the main cushions, intercalated cushion formation involves three cell types including the neural crest and SHF cells^{12,13}. Here, we demonstrate that the absence of *Hoxa1* specifically affects the ability of mesenchymal cells to populate the non-coronary leaflet of the aortic valve irrespective of their origin. Several studies have shown that interaction between neural crest and SHF cells is required for normal outflow tract development⁵¹⁻⁵³. Interaction between neural crest and SHF-derived mesenchymal cells could also be important for the growth of the intercalated cushion. However, the regulation of the interaction between these derivatives remains to be characterized. The TGF-beta/BMP signaling pathway might be involved in the interaction between the mesenchymal cells of neural crest and SHF origins. Indeed, it is important to mention that the presence of only a very small, rudimentary non-coronary leaflet has also been reported when *Alk2*, a BMP-type I receptor, is specifically deleted in post-EndoMT mesenchymal cells⁵⁴.

Our transcriptomic analysis uncovered part of the cellular and molecular mechanisms by which *Hoxa1* controls neural crest cells behavior. *Hoxa1* deletion triggers downregulation of markers reminiscent of the

regulatory network controlling premigration of neural crest³⁷. For instance, *Tfap2a* and *Sox10*, which are part of the gene regulatory network of the premigratory and migratory states, are downregulated in *Hoxa1* mutant mice. Furthermore, out of the transcriptional changes induced by the loss of *Hoxa1*, several lines of evidence support the idea that the mis-regulation of transmembrane receptor genes *Robo1*, *Robo2* and *Nrp2* is likely to play a predominant role in the decrease of neural crest cells migrating toward the pharyngeal arches 4–6. Indeed, the neural crests respond to a variety of signals that target them to a particular arch (see review⁵⁵). Neuropilins are important not only for normal migration but also for condensation of the cardiac neural crest cells in the distal outflow tract. This later may explain the cardiovascular defects observed in *Hoxa1* mutant mice. Interestingly, the Slit/Robo signaling seems to be an attractant for arch 4 colonizing crest cells and repellent in trunk crest. We found that *Robo1* and *Robo2* are up-regulated following the loss of *Hoxa1*, which may contribute to the disruption of normal neural crest cells migration in this environment.

In conclusion, we identify novel indel variants in the poly-histidine tract of the HOXA1 protein in BAV patients and show that length variation of this poly-histidine motif modifies the half-life and the activity of HOXA1. In addition, using the zebrafish model we find that these indel variants act as a dominant negative resulting in defective aortic valve development. Finally, we find that deletion of *Hoxa1* in the mouse results in BAV phenotype associated with abnormal growth of the intercalated cushion and subsequent non-coronary leaflet. Based on this last observation, we believe that it would be important to reassess the classification of BAV types in humans.

Methods

Patient recruitment and DNA extraction

A group of 333 BAV index cases (249 males, 84 females) were prospectively recruited from La Timone Hospital, Marseille in strict compliance with all relevant ethical regulations (approved by the Marseille ethic committee n°13.061). All DNA and research protocols were collected in compliance with the Institutional Review Board after informed consent was obtained. Patients underwent clinical evaluation that included family history, physical examination and 2D-echocardiography.

Genetic analysis

Genomic DNA was extracted from peripheral blood leukocytes using the FlexiGene DNA kit (Qiagen) following the manufacturer's protocol. DNA purity and quantity were assessed with Nanodrop spectrophotometer (Thermo Scientific).

The two exons and intronic *HOXA1* regions were sequenced bidirectionally to search for sequence variations. Primers were designed according to *HOXA1* genomic reference sequence NM_005522. PCR products were sequenced using BigDye Terminator Cycle Sequencing kit (Applied Biosystems) and run-on automatic sequencer ABI 3130XL (Applied Biosystems). Patient sequences were aligned to the *HOXA1* reference sequence using Seqscape software (Applied Biosystems).

Variant pathogenicity was estimated using UMD-Predictor (<http://umd-predictor.eu>)⁵⁶ and Human Splicing Finder⁵⁷. Allele frequencies for each variant were obtained from the Genome Aggregation Database (gnomAD)⁵⁸, which provides high-quality genetic data on 141,456 individuals (125,748 exomes and 15,708 whole genomes).

Mice

All mouse experiments were carried out using protocols approved by the “comité d'éthique pour l'expérimentation animale” (Marseille ethical committee, Protocol N°32-08102012). The null allele of *Hoxa1*^{neo} (hereafter referred to as *Hoxa1*) has been described previously⁵⁹. All genotypes were observed at the expected Mendelian ratios (supplementary material, Table S2). *Hoxa1-enhIII-Cre*, *Wnt1-Cre*, *Rosa^{LacZ}*, and *Gt(ROSA)26Sor^{tm9(CAG-tdTomato)Hze} (Rosa^{tdTomato})* transgenic lines have been previously described^{32,60-62}. Mice were genotyped by PCR and embryos staged taking the morning of the vaginal plug as embryonic day (E) 0.5.

Plasmid construction

Plasmids pCS2-Prep1 [1], pCMV-PBX1a [2], pML-EPHA2-r4-Luc [3] pGL4.74 [from Promega] have been described elsewhere. Plasmids coding for wild type HOXA1 (human HOXA1) and hHOXA1 mutants 9A, 9G and 7G were generated with the Gateway® cloning system. Gateway® entry vectors were obtained using the following primers for PCR delineating the *hHOXA1* sequence: 5'-GGGGACAACCTTTGTACAAAAAAGTTGGCATGGACAATGCAAGAATGAACTCC-3' and 5'-GGGGACAACCTTTGTACAAGAAAGTTGGGTAGTGGGAGGTAGTCAGAGTGTC-3'.

For plasmids coding for the hHOXA1 mutants +1His and +1His^{Arg}, mutagenesis was carried out by NEBaseChanger using the following mutagenic primers forward 5'-CACCATCACCACCCAGCCG-3', reverse 5'-GTGGTGGTGGTGGTGGTG-3'

For plasmids coding for the hHOXA1 variants -1His, -1His^{Arg} and -3His^{Arg}, mutagenesis was carried out by overlapping triple PCRs using the above primers as well as the following mutagenic primers: +1His forward 5'-CAGATCGGTTGCGCCCAACCACCACCACCACCACCATCACCACCCAGCCGGCTACC-3', reverse 5'-GGTAGCCGGCTGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGCGAACCGATCTG-3'; -1His, forward 5'-CAGATCGGTTGCGCCCAACCACCACCACCACCACCATCACCACCCAGCCGGCTACC-3', reverse 5'-GGTAGCCGGCTGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGCGAACCGATCTG-3'; -3His forward 5'-CAGATCGGTTGCGCCCAACCACCACCACCACCATCACCACCCAGCCGGCTACC-3', reverse 5'-GGTAGCCGGCTGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGCGAACCGATCTG-3'. BP-clonase® reactions were achieved using pDONR223 vector and amplified cDNA to generate Gateway® entry pEnt plasmids. These pEnt plasmids were involved in LR-clonase® reactions with pDest-FLAG N-ter destination vector [4] to generate pExp-FLAG-hHOXA1 expression vectors for the WT and mutant hHOXA1 sequences.

Cell culture

Human embryonic kidney cells, HEK293T, were maintained in culture flasks at 37 °C under 5% of CO₂. D-MEM medium (Gibco, Hek293T: #61965-059; COS7: #31885-023) was supplemented with 10% fetal albumin serum (Invitrogen, #10270-106) and 100 U/ml penicillin-streptomycin (Gibco, #15140-122), 1 mM sodium pyruvate (Gibco, #11360-070). Embryonic carcinoma cells, P19, were maintained in culture flasks at 37°C under 5% of CO₂. D-MEM medium (Gibco, Hek293T: #61965-059; COS7: #31885-023) was supplemented with 10% fetal albumin serum (Invitrogen, #10270-106) and 100 U/ml penicillin-streptomycin (Gibco, #15140-122).

Luciferase reporter assays

For Luciferase reporter assays, HEK293T cells were seeded in 96-well plates (1.4×10^4 cells per well) and transfected after 24 hours using JetPrime transfection reagent (Polyplustransfection, #114-07), with 65 ng of target luciferase reporter plasmid (pML-EPHA2-r4-Luc), 25 ng of each FLAG-hHOXA1 expression plasmids (WT, +1His, +1His^{Arg}, -1His, or -3His^{Arg}), 12.5 ng of PREP and PBX expression plasmids each, and 5 ng of constitutive reporter plasmid pGL4.74. Empty vector was added when needed so that each transfection included a total amount of 120 ng of DNA. Forty-Eight hours after transfection, cell lysis and enzymatic activity analysis were performed with Dual-Glo® Luciferase Assay System (Promega E2920) following manufacturer's instructions. For each transfection, the constitutively active pGL4.74 reporter was used as an internal standard for target reporter activity normalization. The relative luciferase activity was established as the ratio between target and constitutive luciferase activities.

Protein half-life analysis

HEK293T cells were plated on 6-well plates, 700 000 cells per well. The cells were transfected with 200 ng of specific pExp plasmids coding for FLAG-tagged WT or mutant HOXA1 proteins using JetPrime transfection reagent (Polyplustransfection, #114-07). For translation inhibition, 24 h after transfection, cells were treated with 200 µg/ml of cycloheximide (Sigma #01810). For both translation and proteasome inhibition, cells were pre-treated 20h after transfection with 7 µM of MG132 (Merck Milipore #474790) and 24 h after transfection were treated with 200 µg/ml of cycloheximide and with 7 µM of MG132. HEK293T cells were lysed with IPLS buffer (20 mM TrisHCl pH 7.5, 120 mM NaCl, 0.5 mM EDTA, 0.5% NP40, 10% glycerol) supplemented with Complete™ protease inhibitor (Roche, #11873580001) for 20 min under agitation on ice. Cell lysates were centrifuged 5 min at 16,000 g at 4°C. 50 µl of each lysate was mixed to Laemmli 6×, boiled 5 min at 95°C and run on a 10% SDS-PAGE. Western blotting was used to visualize FLAG-tagged proteins (α-Flag: Sigma-Aldrich #F1804; α-mouse: Santacruz #sc-516102), as well as ACTIN (α-β-actin-HRP, Sigma-Aldrich #A3854). ImageJ software was used to estimate protein abundance.

Zebrafish studies

AB wild-type (ABwt) zebrafish were maintained under standardized conditions and experiments were conducted in accordance with the European Communities council directive 2010/63. Morpholino

oligonucleotides (MOs) were obtained from Gene Tools (Philomath, OR, USA) and injected into one-cell stage embryos. The sequence of the injected MOs are the following:

hoxa1a MO: 5'-CTAAGAATGTGCTCATTGTGTCATC-3', 7.5 ng

hoxb1b MO: 5'-ATTGCTGTGTCCTGTTTTACCCATG-3', 1 to 2 ng

Rescue experiments were performed by co-injecting *hoxa1a* MO (7.5 ng) or *hoxb1b* MO (1.5 ng) with 25 pg of human wild-type *HOXA1* (WT) mRNA. Injections of human *HOXA1* mRNA alone were performed with either 5pg (WT, -1His, -3His^{Arg}) or 25 pg (WT, +1His) of *in vitro* transcribed 5' capped sense RNAs synthesis using the mMessage mMachine kit (Ambion).

For aortic valve imaging, larvae were incubated one day prior imaging in 0.2 μ M BODIPY-FL Ceramide (Invitrogen D3521) in Embryo medium + PTU (0.003% 1-phenyl-2-thiourea). 7 dpf larvae were then anesthetized with Tricaine (0.16 g/L) and mounted in low melting agarose. Imaging was performed with a Zeiss LSM710 two-photon microscope coupled to a Ti:sapphire laser (Spectra-Physics, Santa Clara, CA, USA) and a water immersion X objective.

Histological, immunohistochemistry and X-gal staining

Standard histological procedures were used⁶³. Heart tissues from *Hoxa1*^{-/-} and littermate controls were paraffin-embedded and cut at 8 μ m per tissue sections. Sections were stained with H&E (Sigma) according to the manufacturer's instruction. CD31 (PECAM) (1:100, Pharmingen) immunohistochemistry were performed using 4% paraformaldehyde fixed tissue and Alexa fluorescent-conjugated antibodies (Life Technologies) were used at 1:500. X-gal staining and *in situ* hybridization were performed on 12 μ m thick frozen sections of hearts from E13.5 and E18.5 as described previously⁶³. For each experiment a minimum of 3 embryos of each genotype was scored. Embryos were examined using an AxioZoom.V16 (Zeiss) and photographed with an AxioCam digital camera (Zen 2011, Zeiss).

RNAFISH

RNA-FISH was performed according to the protocol of the RNAscope Multiplex Fluorescent v2 Assay (Acdbio; cat. no.323110), which detects single mRNA molecules. In briefly, E9.5 embryos were fixed for 20-30 h in 4% paraformaldehyde and then dehydrated in methanol. Whole-mount RNA-FISH was performed as previously described⁶⁴. Embryos were imaged using an AxioZoom.V16 microscope (Zeiss). The following probes were used: mm-*Sox10*-C2 (Acdbio; cat n° 43591-C2), mm-*Robo2*-C1 (Acdbio; cat n° 475961).

RNA-seq

Total RNA was isolated from the pharyngeal region and sorted cells with NucleoSpin RNA XS (Macherey-Nagel) following the protocol of the manufacturer. RNA-seq libraries were created on RNA samples (500 ng) with RNA integrity number (RIN) > 7 using 'kapa mRNA HyperPrep Kit' (Roche). Final cDNA libraries

were checked for quality and quantified using 2100 Bioanalyzer (Agilent Technologies). The libraries were loaded in the flow cell at 1.8 pM concentration and clusters were generated in the Cbot and sequenced in the Illumina NextSeq 500 as 75 bp pair-end reads following Illumina's instructions. Reads were mapped onto the GRCh38 assembly of the *Mus musculus* genome using STAR v2.5.3a. Aligned reads in BAM format were annotated against the protein-coding mRNA regions. Stringtie v.1.3.1c platform was utilized to visualize the annotated/mapped read, quantify data (using corrected log₂ transformed Reads Per Million Values of nonstrand specific, unmerged isoform), and perform percentile normalization. An intensity difference analysis method was used to identify differentially expressed genes (DEGs) on the normalized quantification data.

Three-dimensional (3D) reconstructions

Fiji software was used to make the 3D reconstructions presented in our manuscript. At E13.5, E15.5 and E18.5, images of 20-30 8 µm paraffin sections were manually aligned and taken to generate 3D reconstruction.

Statistical analysis

All parametric data are expressed as mean ±SEM. Statistic was determined using the Student's *t* test to compare variances. For non-parametric data, the Mann-Whitney test was used to calculate significant between the medians. A *p* value of less than 0.05 was considered significant.

Declarations

FUNDINGS

This work was supported by the "Association Française contre les Myopathies" [NMH-Decrypt Project], the "Fondation pour la Recherche Médicale" [DPC20111123002], the "Institut National de la Santé et de la Recherche Médicale" and "la Fondation Leducq" to S.Z. This work was supported by the "Fonds de la recherche Scientifique FNRS" [Crédit de recherche (CDR) J.0157.21] and the "Fonds spéciaux de recherche" (FSR, UCLouvain) to R.R. A.P. received PhD fellowships from the "Association Française du syndrome de Marfan et apparentés". A.F. and C.J. are members of the Laboratory of Excellence Ion Channel Science and Therapeutics supported by a grant from the ANR. Work in the C.J lab is supported by a grant from the "la Fondation Leducq".

ACKNOWLEDGEMENTS

We thank Pr. Pierre Chambon for his authorization to use the *Hoxa1^{neo}* mouse strain. We are grateful to the GBiM MMG Platform and particularly to Nicolas Lenfant to his help in the analysis of the RNA-seq data. We thank the MMG animal core facility for their help in the breeding. We thank IPAM and FranceBioImaging (ANR-10-INSB-04 to P.M.).

References

1. Michelena, H. I. *et al.* Incidence of aortic complications in patients with bicuspid aortic valves. *JAMA* **306**, 1104–1112 (2011).
2. Hinton, R. B. Bicuspid aortic valve and thoracic aortic aneurysm: three patient populations, two disease phenotypes, and one shared genotype. *Cardiol. Res. Pract.* 2012, 926975 (2012).
3. Padang, R., Bagnall, R. D. & Semsarian, C. Genetic basis of familial valvular heart disease. *Circ. Cardiovasc. Genet.* **5**, 569–580 (2012).
4. Sievers, H.-H. & Schmidtke, C. A classification system for the bicuspid aortic valve from 304 surgical specimens. *J. Thorac. Cardiovasc. Surg.* **133**, 1226–1233 (2007).
5. Duran, A. C. *et al.* Bicuspid aortic valves in hearts with other congenital heart disease. *J. Heart Valve Dis.* **4**, 581–590 (1995).
6. Hor, K. N., Border, W. L., Cripe, L. H., Benson, D. W. & Hinton, R. B. The presence of bicuspid aortic valve does not predict ventricular septal defect type. *Am. J. Med. Genet. A.* **146A**, 3202–3205 (2008).
7. Anderson, R. *et al.* Myths and Realities Relating to Development of the Arterial Valves. *J. Cardiovasc. Dev. Dis.* **1**, 177–200 (2014).
8. Markwald, R. R., Fitzharris, T. P. & Manasek, F. J. Structural development of endocardial cushions. *Am. J. Anat.* **148**, 85–119 (1977).
9. Kirby, M. L., Gale, T. F. & Stewart, D. E. Neural crest cells contribute to normal aorticopulmonary septation. *Science* **220**, 1059–1061 (1983).
10. Phillips, H. M. *et al.* Neural crest cells are required for correct positioning of the developing outflow cushions and pattern the arterial valve leaflets. *Cardiovasc. Res.* **99**, 452–460 (2013).
11. Odelin, G. *et al.* Krox20 defines a subpopulation of cardiac neural crest cells contributing to arterial valves and bicuspid aortic valve. *Dev. Camb. Engl.* **145**, dev151944 (2018).
12. Eley, L. *et al.* A novel source of arterial valve cells linked to bicuspid aortic valve without raphe in mice. *eLife* **7**, e34110 (2018).
13. Mifflin, J. J., Dupuis, L. E., Alcalá, N. E., Russell, L. G. & Kern, C. B. Intercalated cushion cells within the cardiac outflow tract are derived from the myocardial troponin T type 2 (Tnnt2) Cre lineage: Tnnt2-Cre Lineage form Intercalated Cushions. *Dev. Dyn.* **247**, 1005–1017 (2018).
14. Cripe, L., Andelfinger, G., Martin, L. J., Shoener, K. & Benson, D. W. Bicuspid aortic valve is heritable. *J. Am. Coll. Cardiol.* **44**, 138–143 (2004).
15. Garg, V. *et al.* Mutations in NOTCH1 cause aortic valve disease. *Nature* **437**, 270–274 (2005).
16. Bonachea, E. M. *et al.* Rare GATA5 sequence variants identified in individuals with bicuspid aortic valve. *Pediatr. Res.* **76**, 211–216 (2014).
17. Mohamed, S. A. *et al.* Novel missense mutations (p.T596M and p.P1797H) in NOTCH1 in patients with bicuspid aortic valve. *Biochem. Biophys. Res. Commun.* **345**, 1460–1465 (2006).

18. Baylor-Hopkins Center for Mendelian Genomics *et al.* ROBO4 variants predispose individuals to bicuspid aortic valve and thoracic aortic aneurysm. *Nat. Genet.* **51**, 42–50 (2019).
19. Laforest, B. & Nemer, M. Genetic Insights into Bicuspid Aortic Valve Formation. *Cardiol. Res. Pract.* 2012, 1–8 (2012).
20. Théron, A. *et al.* Clinical insights into a tertiary care center cohort of patients with bicuspid aortic valve. *Int. J. Cardiovasc. Imaging* (2021) doi:10.1007/s10554-021-02366-1.
21. Kong, W. K. F. *et al.* Sex Differences in Phenotypes of Bicuspid Aortic Valve and Aortopathy: Insights From a Large Multicenter, International Registry. *Circ. Cardiovasc. Imaging* **10**, e005155 (2017).
22. Chen, J. & Ruley, H. E. An enhancer element in the EphA2 (Eck) gene sufficient for rhombomere-specific expression is activated by HOXA1 and HOXB1 homeobox proteins. *J. Biol. Chem.* **273**, 24670–24675 (1998).
23. Lambert, B. *et al.* Pentapeptide insertion mutagenesis of the Hoxa1 protein: Mapping of transcription activation and DNA-binding regulatory domains. *J. Cell. Biochem.* n/a-n/a (2010) doi:10.1002/jcb.22563.
24. Di Rocco, G. Functional dissection of a transcriptionally active, target-specific Hox-Pbx complex. *EMBO J.* **16**, 3644–3654 (1997).
25. McClintock, J. M., Carlson, R., Mann, D. M. & Prince, V. E. Consequences of Hox gene duplication in the vertebrates: an investigation of the zebrafish Hox paralogue group 1 genes. *Dev. Camb. Engl.* **128**, 2471–2484 (2001).
26. Shih, L. J., Tsay, H. J., Lin, S. C. & Hwang, S. P. Expression of zebrafish Hoxa1a in neuronal cells of the midbrain and anterior hindbrain. *Mech. Dev.* **101**, 279–281 (2001).
27. Faucherre, A. *et al.* Piezo1 is required for outflow tract and aortic valve development. *J. Mol. Cell. Cardiol.* **143**, 51–62 (2020).
28. Kroll, F. *et al.* A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife* **10**, e59683 (2021).
29. Tischfield, M. A. *et al.* Homozygous HOXA1 mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. *Nat. Genet.* **37**, 1035–1037 (2005).
30. Holve, S. *et al.* Athabaskan brainstem dysgenesis syndrome. *Am. J. Med. Genet. A.* **120A**, 169–173 (2003).
31. Makki, N. & Capecchi, M. R. Cardiovascular defects in a mouse model of HOXA1 syndrome. *Hum. Mol. Genet.* **21**, 26–31 (2012).
32. Bertrand, N. *et al.* Hox genes define distinct progenitor sub-domains within the second heart field. *Dev. Biol.* **353**, 266–274 (2011).
33. Roux, M. *et al.* Hoxa1 and Hoxb1 are required for pharyngeal arch artery development. *Mech. Dev.* **143**, 1–8 (2017).
34. Roux, M., Laforest, B., Capecchi, M., Bertrand, N. & Zaffran, S. Hoxb1 regulates proliferation and differentiation of second heart field progenitors in pharyngeal mesoderm and genetically interacts

- with *Hoxa1* during cardiac outflow tract development. *Dev. Biol.* **406**, 247–258 (2015).
35. Makki, N. & Capecchi, M. R. *Hoxa1* lineage tracing indicates a direct role for *Hoxa1* in the development of the inner ear, the heart, and the third rhombomere. *Dev. Biol.* **341**, 499–509 (2010).
 36. Jain, R. *et al.* Cardiac neural crest orchestrates remodeling and functional maturation of mouse semilunar valves. *J. Clin. Invest.* **121**, 422–430 (2011).
 37. Simões-Costa, M. & Bronner, M. E. Establishing neural crest identity: a gene regulatory recipe. *Development* **142**, 242–257 (2015).
 38. Jia, L., Cheng, L. & Raper, J. Slit/Robo signaling is necessary to confine early neural crest cells to the ventral migratory pathway in the trunk. *Dev. Biol.* **282**, 411–421 (2005).
 39. Shiao, C. E. & Bronner-Fraser, M. N-cadherin acts in concert with Slit1-Robo2 signaling in regulating aggregation of placode-derived cranial sensory neurons. *Dev. Camb. Engl.* **136**, 4155–4164 (2009).
 40. Mitchell, P. J., Timmons, P. M., Hébert, J. M., Rigby, P. W. & Tjian, R. Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes Dev.* **5**, 105–119 (1991).
 41. Karlin, S., Brocchieri, L., Bergman, A., Mrazek, J. & Gentles, A. J. Amino acid runs in eukaryotic proteomes and disease associations. *Proc. Natl. Acad. Sci.* **99**, 333–338 (2002).
 42. Salichs, E., Ledda, A., Mularoni, L., Albà, M. M. & de la Luna, S. Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment. *PLoS Genet.* **5**, e1000397 (2009).
 43. Pelassa, I. *et al.* Compound Dynamics and Combinatorial Patterns of Amino Acid Repeats Encode a System of Evolutionary and Developmental Markers. *Genome Biol. Evol.* **11**, 3159–3178 (2019).
 44. Radó-Trilla, N. *et al.* Key Role of Amino Acid Repeat Expansions in the Functional Diversification of Duplicated Transcription Factors. *Mol. Biol. Evol.* **32**, 2263–2272 (2015).
 45. Brown, L. Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends Genet.* **20**, 51–58 (2004).
 46. Albrecht, A. & Mundlos, S. The other trinucleotide repeat: polyalanine expansion disorders. *Curr. Opin. Genet. Dev.* **15**, 285–293 (2005).
 47. Gatchel, J. R. & Zoghbi, H. Y. Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nat. Rev. Genet.* **6**, 743–755 (2005).
 48. Basu, S. *et al.* Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. *Cell* **181**, 1062–1079.e30 (2020).
 49. Tzemos, N. *et al.* Outcomes in adults with bicuspid aortic valves. *JAMA* **300**, 1317–1325 (2008).
 50. Waldo, K. L. *et al.* Secondary heart field contributes myocardium and smooth muscle to the arterial pole of the developing heart. *Dev. Biol.* **281**, 78–90 (2005).
 51. El Robrini, N. *et al.* Cardiac outflow morphogenesis depends on effects of retinoic acid signaling on multiple cell lineages. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **245**, 388–401 (2016).
 52. Kodo, K. *et al.* Regulation of *Sema3c* and the Interaction between Cardiac Neural Crest and Second Heart Field during Outflow Tract Development. *Sci. Rep.* **7**, 6771 (2017).

53. Darrigrand, J.-F. *et al.* Dullard-mediated Smad1/5/8 inhibition controls mouse cardiac neural crest cells condensation and outflow tract septation. *eLife* **9**, e50325 (2020).
54. Thomas, P. S., Sridurongrit, S., Ruiz-Lozano, P. & Kaartinen, V. Deficient Signaling via Alk2 (Acvr1) Leads to Bicuspid Aortic Valve Development. *PLoS ONE* **7**, e35539 (2012).
55. Kirby, M. L. & Hutson, M. R. Factors controlling cardiac neural crest cell migration. *Cell Adhes. Migr.* **4**, 609–621 (2010).
56. Salgado, D. *et al.* UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution. *Hum. Mutat.* **37**, 439–446 (2016).
57. Desmet, F.-O. *et al.* Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37**, e67 (2009).
58. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
59. Lufkin, T., Dierich, A., LeMeur, M., Mark, M. & Chambon, P. Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell* **66**, 1105–1119 (1991).
60. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70–71 (1999).
61. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
62. Jiang, X., Rowitch, D. H., Soriano, P., McMahon, A. P. & Sucov, H. M. Fate of the mammalian cardiac neural crest. *Dev. Camb. Engl.* **127**, 1607–1616 (2000).
63. Odelin, G. *et al.* Loss of Krox20 results in aortic valve regurgitation and impaired transcriptional activation of fibrillar collagen genes. *Cardiovasc. Res.* **104**, 443–455 (2014).
64. Stefanovic, S. *et al.* Hox-dependent coordination of mouse cardiac progenitor cell patterning and differentiation. *eLife* **9**, (2020).

Tables

Table 1: Baseline patient characteristics in 333 patients with bicuspid aortic valve

Patients (n)	333
Mean age (years)	52.39±14.94
Male (%)	249 (74.77%)
Family history of BAV (%)	30 (9.01%)
Type of BAV according to Sievers classification	
Type 0	44 (13.22%)
Type 1, R-L	227 (68.16%)
Type 1, R-NC	38 (11.42%)
Type 1, L-NC	9 (2.70%)
Type 2	5 (1.50%)
Type 1, Not determined	1 (0.30%)
Not determined	9 (2.70%)

BAV= Bicuspid Aortic Valve

Type 0= no fusion

Type 1, R-L= fusion of the right and left coronary (R-L) leaflets.

Type 1, R-NC= fusion of the right and non-coronary (R-NC) leaflets.

Type 1, L-NC= fusion of the left and non-coronary (R-NC) leaflets.

Type 2= fusion of the right and left coronary (R-L) leaflets and of the left and non-coronary (R-NC) leaflets.

Table 2: Distribution of *HOXA1* Histidine stretch variations in our BAV population (n=333 Index Cases).

Variations	Impact on Histidine stretch	Short Names	Number of studied alleles	Frequencies in our BAV cohort (%)
c.[213_215dup;218G>A]	11His	+1His	4	0.60
c.[213_215dup;218G]	9His1Arg1His	+1His ^{Arg}	1	0.15
c.218G>A	10His	WT ^A	537	80.63
c.218G	8His1Arg1His	WT ^G	99	14.87
c.[213_215delCCA;218G>A]	9His	-1His	22	3.30
c.[213_215delCCA;218G]	7His1Arg1His	-1His ^{Arg}	<i>Not found in this cohort</i>	
c.[207_215delCCACCACCA;218G>A]	7His	-3His	<i>Not found in this cohort</i>	
c.[207_215delCCACCACCA;218G]	5His1Arg1His	-3His ^{Arg}	3	0.45

Table 3: Reported *HOXA1* variation frequencies in gnomAD

Mutation names	Mutation name at protein level	Short Names	gnomAD allele frequencies
c.213_215dupCCA	p.His72dup	+1His or +1His ^{Arg}	4.25 10 ⁻³
c.213_215delCCA	p.His72del	-1His or -1His ^{Arg}	7.78 10 ⁻⁴
c.207_215delCCACCACCA	p.His70_His72del	-3His or -3His ^{Arg}	3.29 10 ⁻³
c.218G>A	p.Arg73His		7.53 10 ⁻¹

Extended Figure Legends

Extended data Figure 1: Conservation of Hoxa1 ortholog and paralogs. **a**, Amino acid alignment of sequence from zebrafish (Q98SI1), human (P49639) and mouse (P09022) Hoxa1 orthologs. Identical residues in red, conserved changes in blue. Green line shows the homeodomain and blue line underlines the hexapeptide motif. **b**, Neighbor-joining tree showing phylogenetic relationships between Hoxa1 and Hoxb1 paralog amino acid sequences. As expected Hoxb1 provides an outgroup to all Hoxa1 sequences. **c-e**, Whole-mount in situ hybridization analysis of zebrafish *hoxa1a* (**c**), *hoxb1a* (**d**) and *hoxb1b* (**e**) genes. **c**, At 24 h post-fertilization (24 hpf) *hoxa1a* expression is localized to discrete bilateral clusters of cells in

the anterior hindbrain and ventral midbrain. **d**, At the equivalent stage, *hoxb1a* is at the equivalent stage is expressed in rhombomere 4 (r4). **e**, *hoxb1b* expression is already retreated posteriorly in an anterior subpopulation of neurons. mhb, midbrain/hindbrain boundary.

Extended data Figure 2: Targeting *hoxa1a* and *hoxb1a* alleles. **a**, One-cell stage zebrafish embryos were injected with 7.5 ng of *hoxa1a* morpholino (MO) and examined at 72 h post-fertilization (hpf) in comparison to non-injected controls (NI). **b**, One-cell stage zebrafish embryos were injected with 1 ng and 2 ng of *hoxb1a* MO morpholino (MO) and examined at 72 hpf in comparison to non-injected controls (NI).

Extended data Figure 3: Injection of gRNAs targeting either *hoxa1a* and *hoxb1a* disrupts aortic valve development. Two-photon images of aortic valves in 7 days post-fertilization (dpf) zebrafish larvae labelled with BODIPY. **a**, Representative images of aortic valve leaflets in a wild-type, non-injected larvae (ABwt, NI), valves are outlined with a dashed white line (n=17). **b**, gRNAs targeting *hoxa1a* (crispant) were injected into single cell zebrafish embryos (n=18). *hoxa1a* crispants show hypertrophic aortic valve. **c**, gRNAs targeting *hoxb1b* (crispant) were injected into single cell zebrafish embryos (n=19). *hoxb1b* crispants show hypertrophic aortic valve. **d**, Graph showing quantification of aortic valve defects observed in a 7 dpf larvae after injection of gRNA to knockout *hoxa1a* or *hoxb1b* genes. Approximately 80% of *hoxa1a* and 70% of *hoxb1a* crispants exhibited aortic valve defects.

Extended data Figure 4: Rescue experiment of *hoxb1b* morpholino. **a-b**, One-cell stage zebrafish embryos were injected with either 1.5 ng (**a**) or 2 ng (**b**) of *hoxb1b* MO alone or in combination with human wild-type *HOXA1* mRNA at 25 pg, and examined at 72 h post-fertilization (hpf) in comparison to non-injected controls (NI) Injection of human wild-type (wt) *HOXA1* mRNA cannot rescue the mild (**a**) or strong (**b**) phenotype induced by *hoxb1b* MO.

Extended data Figure 5: Titration experiment of *HOXA1* variants. **a**, One-cell stage zebrafish embryos were injected with 25 pg of *HOXA1* wild-type (WT) or mutated (+1His, -1His, -1His^{Arg}, -3His^{Arg}), and examined at 72 hpf in comparison to non-injected controls (NI). **b**, One-cell stage zebrafish embryos were injected with 1 pg, 5 pg or 10 pg of *HOXA1* -3His^{Arg}, in order to determine the amount of mRNA to use for valve leaflet observation. Injecting 5 pg gave a phenotype mild enough to carry the experiment at 7 dpf as presented in Figure 3.

Extended data Figure 6: Morphology of the aortic valve of *Hoxa1*^{-/-} embryos at E15.5. **a-b**, Transversal section of E15.5 hearts immunostained with an endothelial marker, anti-Pecam antibody (red), and stained for nuclei marker (DAPI, blue). Note the bicuspid aortic valve in *Hoxa1*^{-/-} (**b**). **c-d**, 3D reconstruction of histological images at E15.5 showing three aortic valve leaflets in the wild-type (WT) embryo (**c**), whereas a persistent small non-coronary leaflet is seen in *Hoxa1*^{-/-} embryo (**d**). Left coronary (L; pink), right coronary (R; yellow), non-coronary (NC; blue) leaflets. Scale bars: 100 μm (**a,b**).

Extended data Figure 7: Relative surface and cell density of the aortic valve leaflets. **a**, The relative surface of the three aortic valve leaflets was measured. There is a significant decrease in the *Hoxa1*^{-/-} (n=16) compared to the wild-type (WT; n=14) littermates. **b**, The cell density was measured for each leaflet and no significant difference was observed between in the *Hoxa1*^{-/-} (n=16) compared to WT (n=14) littermates. Data are shown as mean +/- SEM. Statistical values were obtained using the Mann-Whitney test. LC: left coronary leaflet; NC: non-coronary leaflet; RC: right coronary leaflet.

Extended data Figure 8: Quality assessment of RNA-seq data performed with wild-type and *Hoxa1*^{-/-} embryos. Principal component analysis (PCA) of RNA-sequencing datasets from wild-type (WT; blue) and *Hoxa1*^{-/-} (red) embryos.

Extended data Figure 9: Abnormal cardiac neural crest migration in *Hoxa1*^{-/-}. **a,b** Whole-mount *in situ* hybridization for *Crabp1* transcripts (marking neural crest cells) in wild-type (WT; **a**), and *Hoxa1*^{-/-} (**b**) embryos at E9.5. **a**, Arrowheads indicate normal post-otic streams of migrating neural crest cells. **b**, The asterisk indicates reduction of *Crabp1* expression in the 3rd pharyngeal arch of *Hoxa1*^{-/-} embryos compared to WT. ba: branchial arch; rv: right ventricle.

Extended data Figure 10: Genetic lineage analysis of the *Hoxa1*-positive cells. **a-b**, *Hoxa1*-lineage visualized by X-gal staining of *Hoxa1-enhIII-cre;R26R-lacZ* embryos. **a**, Transversal section of the outflow tract at E13.5. β -galactosidase activity is observed in the outflow tract cushions. **b**, At E18.5, β -galactosidase activity is observed in cells forming the aortic valve leaflets. AoV, aortic valve; PV, pulmonary valve. Scale bars: 100 μ m (**a,b**).

Figures

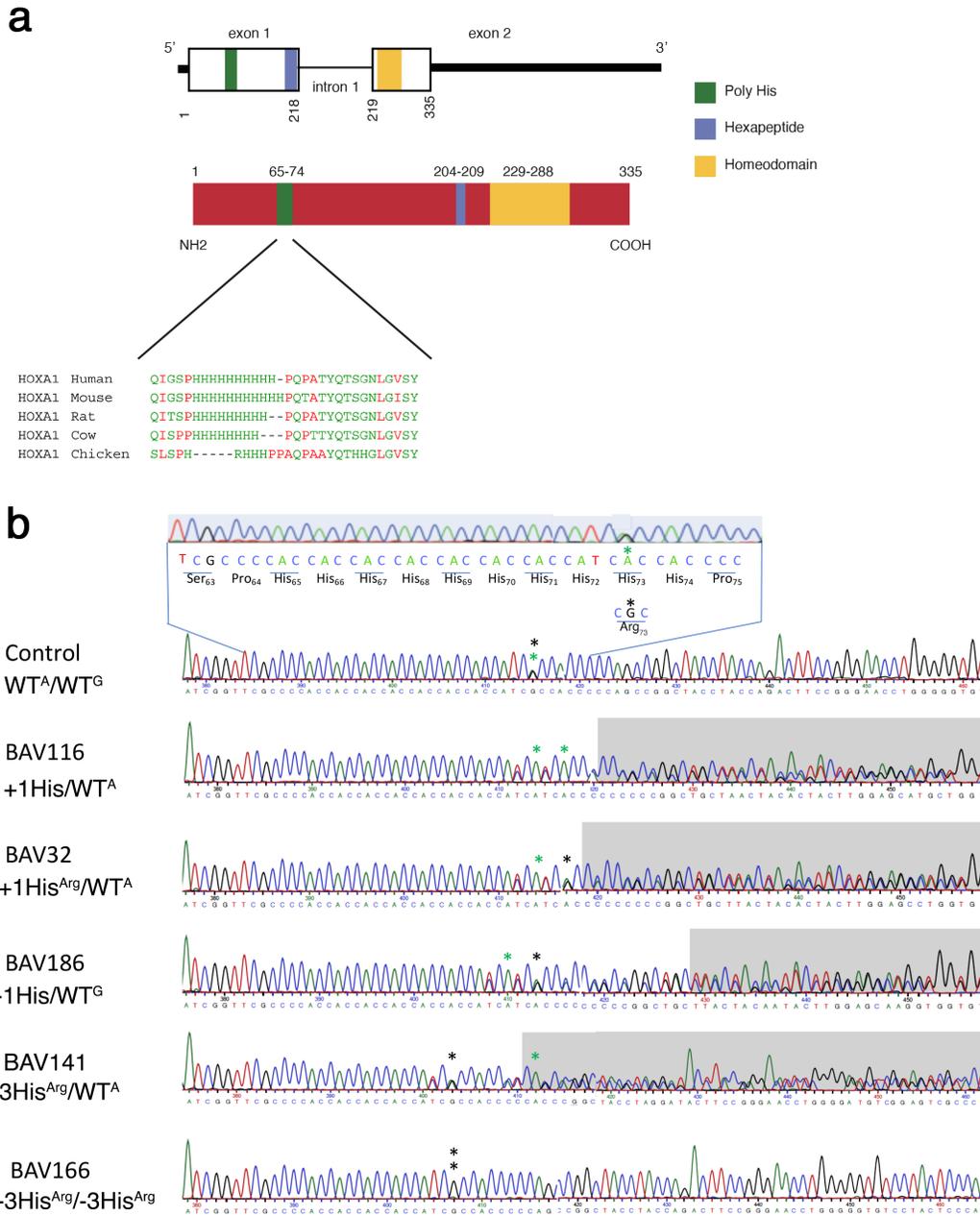


Figure 1

Figure 1

Identification of *HOXA1* variants in index cases with BAV. **a**, Schematic representation of the human *HOXA1* gene and corresponding predicted protein. The *HOXA1* protein contains the histidine repeat motif (green; amino acids 65-74), the hexapeptide motif involved in the interaction with PBX proteins (blue; amino acids 204-209), encoded in exon 1, and the DNA-binding homeodomain (yellow; amino acids 229-288), encoded in exon 2. Alignment of the *HOXA1* histidine repeat sequences from the human, mouse, rat,

cow and chicken species is represented. The NCBI GenBank accession numbers that were utilized for the alignment are as follows: Human-P49639, Mouse-P09022, Rat-O08656, Cow-E1B918, and Chicken-R4GHI8. **b**, Sanger sequences showing the insertion or deletion variants observed in exon 1 of *HOXA1*. The polymorphism c.218G>A (p.Arg73His) is shown with asterisks (black, c.218G; green, c.218G>A). We identified different variants in this region (c.[213_215dup;218G>A] named +1His; c.[213_215dup;218G] named +1His^{Arg}; c.[213_215delCCA;218G>A] named -1His and c.[207_215delCCACCACCA;218G] named -3His^{Arg}) with low frequency in gnomAD (see Table 3).

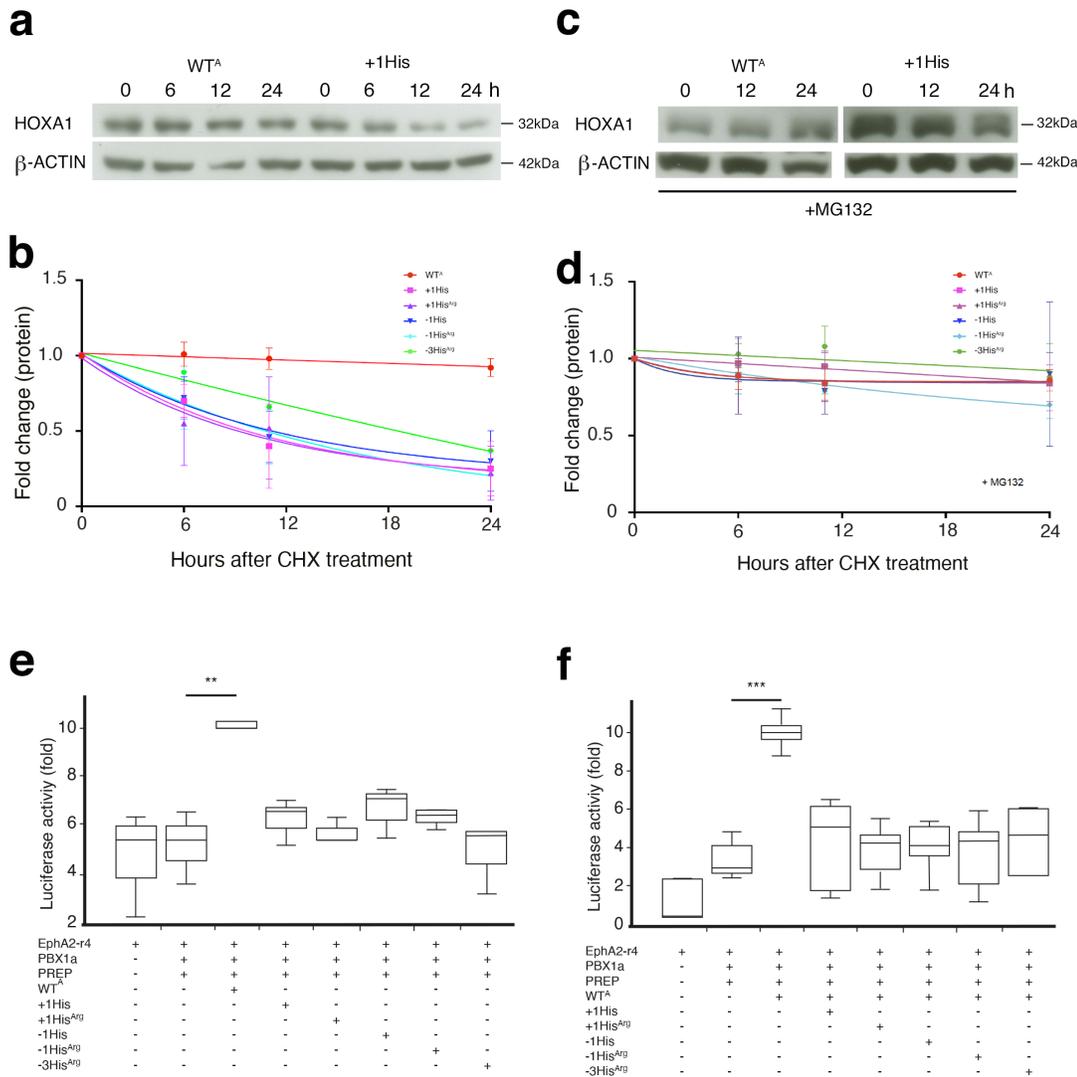


Figure 2

Figure 2

HOXA1 variants induce proteasome-dependent protein degradation and reduce transcriptional activity of the protein. **a-d**, HEK293T cells were transfected with expression vectors coding for FLAG-HOXA1 (WT^A, +1His, +1His^{Arg}, -1His, -1His^{Arg} and -3His^{Arg}). To determine the HOXA1 half-life, cells were treated with cycloheximide (CHX; 200µg/ml) for the indicated time. **a**, Cell lysates were collected every 6 hours and then subjected to immunoblot analysis with antibodies against FLAG and β-ACTIN. **b**, Intensity of the

HOXA1 bands observed upon cycloheximide treatment was quantified and the relative HOXA1/ β -ACTIN ratios were plotted (n>6). **c**, Involvement of the proteasome in the HOXA1 degradation was assayed by treating cells with proteasome inhibitor (MG132; 7 μ M) for 20 hours prior inhibition of protein translation with CHX (n=3). Cell lysates were collected every 12 hours and then subjected to immunoblot analysis with antibodies against FLAG and β -ACTIN. **d**, The intensity of the HOXA1 bands observed upon cycloheximide treatment was quantified and the relative HOXA1/ β -ACTIN ratios were plotted. **e**, Luciferase assay showing transcriptional activity of wild-type (WT^A) and mutated hHOXA1 (+1His, +1His^{Arg}, -1His, -1His^{Arg}, and -3His^{Arg}). HEK293T cells were transfected with the *EphA2-r4*-luciferase reporter construct, alone (*EphA2-r4*) or together with expression plasmids for PREP1, PBX1a, and FLAG-HOXA1 (WT^A, +1His, +1His^{Arg}, -1His, -1His^{Arg}, and -3His^{Arg}). Same quantity of DNA was transfected for each plasmid. **f**, In order to characterize the dominant negative effect of mutated human HOXA1 proteins, cells were transfected with the *EphA2-r4*-luciferase reporter together with both wild-type and variant HOXA1 coding vectors. ** p <0.01; *** p <0.001. Statistical values were obtained using Pairwise t test with Holm P value adjustment method test. (n=3, each experiment was performed in duplicate).

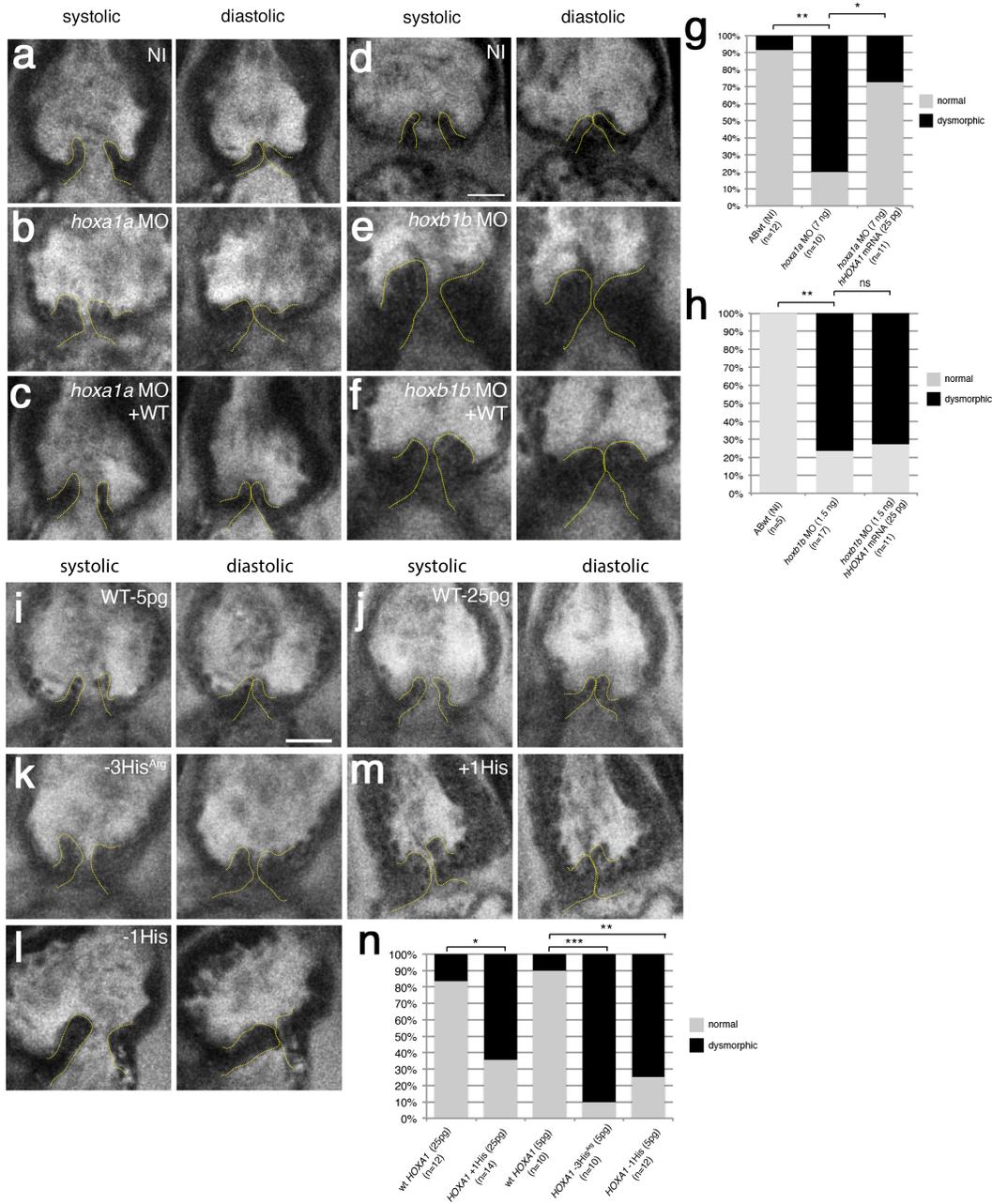


Figure 3

Figure 3

Expression of HOXA1 variants *in vivo* disrupts aortic valve development. Two-photon images of aortic valves in 7 dpf zebrafish larvae labelled with BODIPY. **a,d** Representative image of aortic valve leaflets in a wild-type, non-injected larvae (ABwt, NI), valves are outlined with a dashed yellow line (a: n=12; d: n=5). **b**, *hoxa1a* morphant (n=10). **c**, *hoxa1a* morphant co-injected with wild-type human *HOXA1* mRNA (n=11). **e**, *hoxb1b* morphant (n=17). **f**, *hoxb1b* morphant co-injected with wild-type human *HOXA1* mRNA (n=11).

g, Graph showing quantification of aortic valve defects observed in 7 dpf larvae after knockdown of *hoxa1a*. Approximately 80% of *hoxa1a* morphants exhibited aortic valve defects (** $p=0.0015$; * $p=0.03$). *HOXA1* human mRNA rescued the *hoxa1a* morpholino aortic valve phenotype. **h**, Graph showing quantification of aortic valve defects observed in 7 dpf larvae after knockdown of *hoxb1b* (** $p=0.0048$). human *HOXA1* mRNA does not rescue the *hoxb1b* morpholino aortic valve phenotype. **i-j**, Representative images of aortic valve leaflets after injection of WT human *HOXA1* mRNA (WT) at 5 pg (n=10) or 25 pg (n=12), respectively. **k**, Injection of the -3His^{Arg} variant of human *HOXA1* at 5 pg (n=10). **l**, Injection of the -1His variant of human *HOXA1* at 5 pg (n=10). **m**, Injection of the +1His variant of human *HOXA1* at 25 pg (n=12). **n**, Graph showing quantification of aortic valve defects observed in 7 dpf larvae after injection of the variants of human *HOXA1* (** $p=0.001$; ** $p=0.0037$; * $p=0.021$). Statistical values were obtained using Fisher exact test. Scale bar=20 μ m.

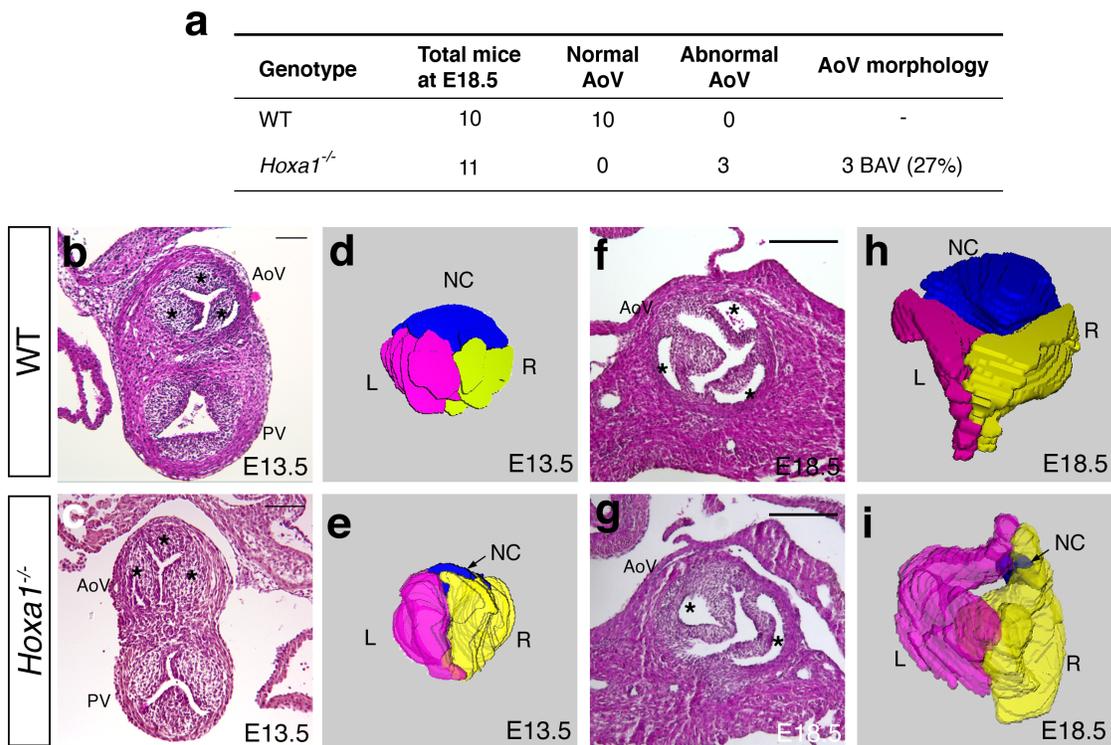


Figure 4

Figure 4

***Hoxa1* null embryos display bicuspid aortic valves.** **a**, Table depicting penetrance of bicuspid aortic valve (BAV) in *Hoxa1*^{-/-} mice at E18.5. **b-c**, Hematoxylin and eosin (H&E) staining images showing representative transversal section of wild-type (WT) and *Hoxa1*^{-/-} outflow tract at E13.5. **d-e**, Three-dimensional (3D) reconstructions of histological images at E13.5. A slight difference between WT (**d**) and *Hoxa1*^{-/-} (**e**) aortic valve (AoV) structure is observed. Indeed, the non-coronary leaflet (NC) seems smaller in *Hoxa1*^{-/-} when compared to wild-type embryos. **f-g**, Cross-sectional H&E images through the aortic valve of WT (**f**) and *Hoxa1*^{-/-} (**g**) embryos at E18.5. Normal valve with three leaflets (asterisks) is observed in WT embryos (**f**), whereas bicuspid aortic valve is detected in the mutant (**g**). **h-i**, 3D reconstruction of histological images at E18.5 showing three aortic valve leaflets in the WT embryo (**h**), whereas a persistent small non-coronary leaflet is seen in *Hoxa1*^{-/-} embryos (**i**). Left coronary (L; pink), right coronary (R; yellow), non-coronary (NC; blue) leaflets; PV, pulmonary valve. Scale bars: 100 μm (**b,c**); 200 μm (**f,g**).

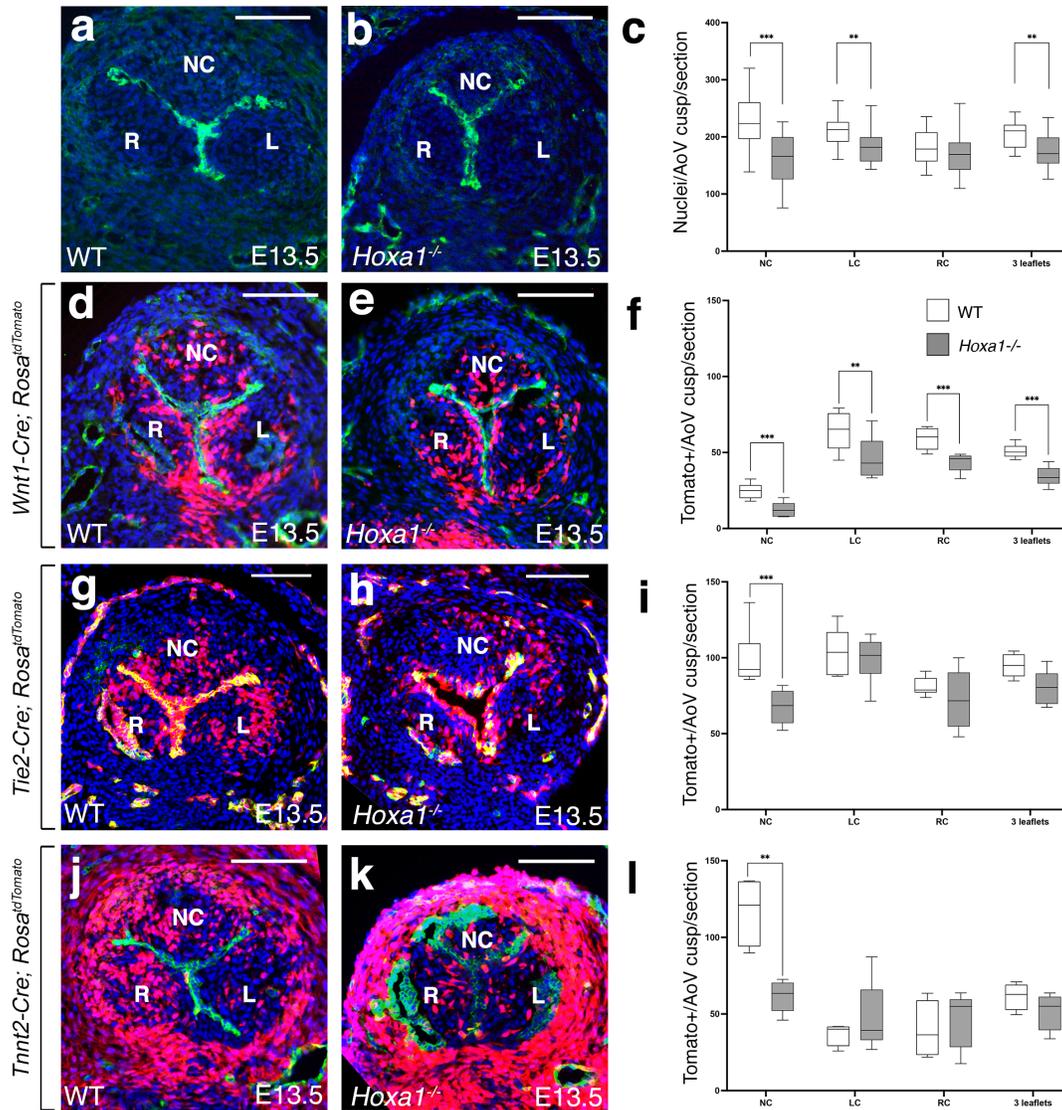


Figure 5

Figure 5

Analysis of cell lineage contribution in *Hoxa1*-null mice. **a-b**, Transversal sections of E13.5 hearts immunostained with an endothelial marker (anti-Pecam antibody, green) and DAPI staining for nuclei (blue). Note the smaller non-coronary (NC) leaflet in *Hoxa1*^{-/-} (**b**) compared with WT (**a**) embryos. **c**, Total nuclei were counted in the three forming aortic valve leaflets from *Hoxa1*^{-/-} (n=20) and WT (n=18) embryos at E13.5 spanning a 180 μm depth. **d-l**, Fate-mapping of the neural crest-, endothelial- and

second heart field-lineages in the aortic valves of WT (**d,g,j**), and *Hoxa1*-null (**e,h,k**) embryos at E13.5. Tomato-reporter is visualized in red and nuclei (DAPI) appear in blue. **d-e**, Immunofluorescence staining using *Wnt1-Cre,Rosa^{tdTomato/+}* reporter mice shows a decrease in the number of neural crest derivatives in absence of *Hoxa1*. **f**, Quantification of neural crest-derived cells demonstrates significant decrease in *Hoxa1*^{-/-} (n=5) compared to WT (n=6) littermate embryos (******* $p < 0.004$; ****** $p = 0.03$). **g-h**, Immunofluorescence staining using *Tie2-Cre,Rosa^{tdTomato/+}* reporter mice shows a decreased contribution of endothelial lineage to the non-coronary leaflet of *Hoxa1*^{-/-} (**h**) compared to WT (**g**) littermates. **i**, Quantification confirms the decreased number of endothelial-derived cells in the non-coronary leaflet in *Hoxa1*^{-/-} (n=6) compared to WT (n=6) littermates ($p = 0.002$). **j-k**, Immunofluorescence staining using *Tnnt2-Cre,Rosa^{tdTomato/+}* reporter mice demonstrates a decreased contribution of the second heart field lineage to the non-coronary leaflet in *Hoxa1*^{-/-} (**k**) compared to WT (**j**) littermates. **l**, Quantification confirms the reduced number of second heart cells-derived cells in the non-coronary leaflet in *Hoxa1*^{-/-} (n=5) compared to WT (n=4) littermates ($p = 0.01$). Data are shown as mean +/- SEM. Statistical values were obtained using the Mann-Whitney test. LC: left coronary leaflet; NC: non-coronary leaflet; RC: right coronary leaflet.

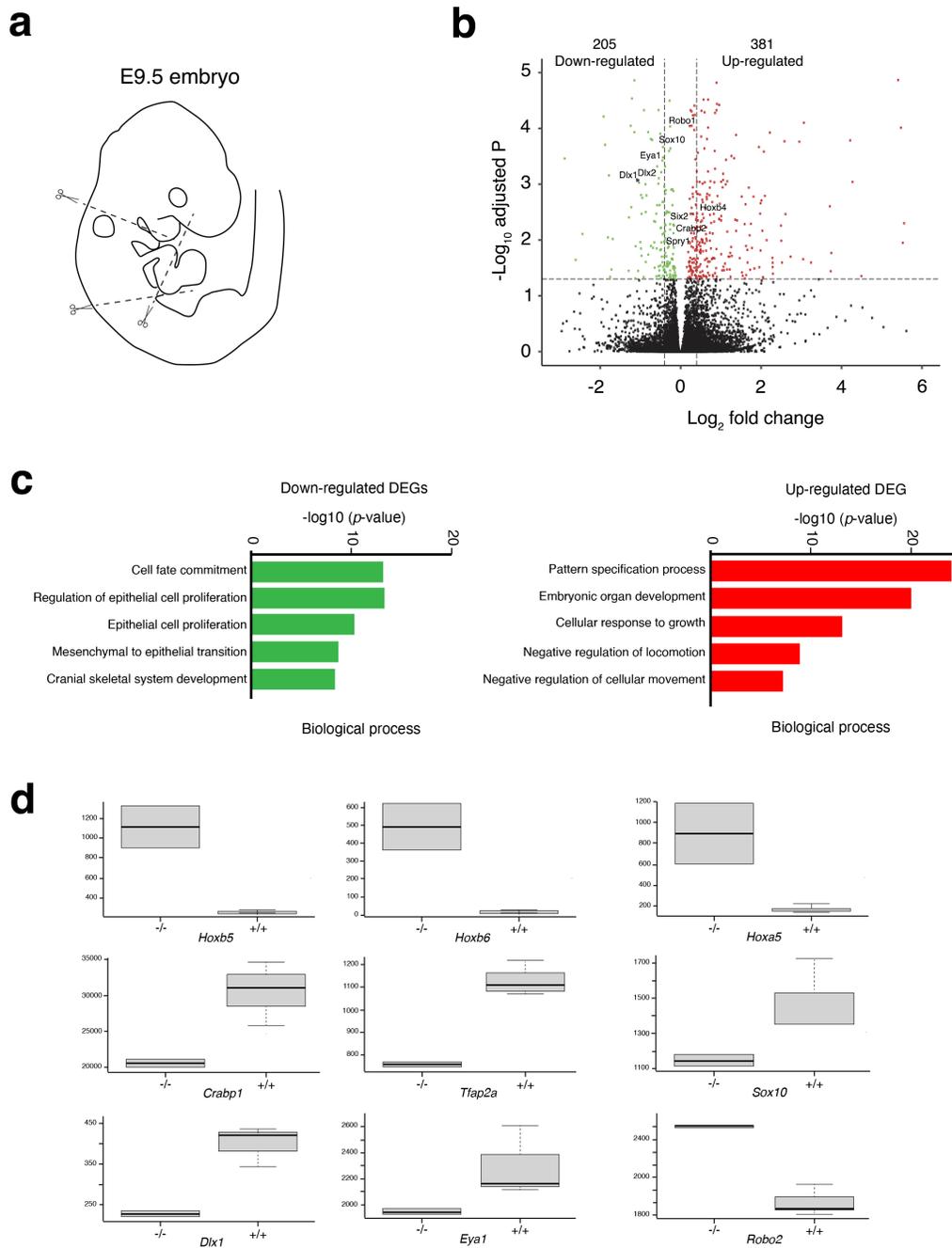


Figure 6

Figure 6

Transcriptomic analysis identifies *Hoxa1* targets involved in migration of neural crest cells. **a**, Scheme showing which region was used to perform RNA-seq analysis on *Hoxa1*^{-/-} vs WT embryos. **b**, Volcano plot of fold changes and corresponding FDR (False Discovery Rate) value for each RNA after comparison of *Hoxa1*^{-/-} and WT embryos. The red and green dots represent RNAs with fold change >2.0 and FDR<0.05. The black dots represent RNAs with fold change <2.0 and FDR> 0.05. **c**, Gene ontology (GO) analysis

performed with ClusterProfiler system showing enrichment of up-regulated (red) and down-regulated (green) genes in the mutant with ranked by $-\log_{10}(p\text{-value})$. **d**, Boxplot of FPKM (Fragments Per Kilo base of transcript per Million mapped fragments) expression value for nine selected DE genes in RNA-seq. The y-axis represents the FPKM expression level, and the x-axis the genotype. Data are shown as mean \pm SEM.

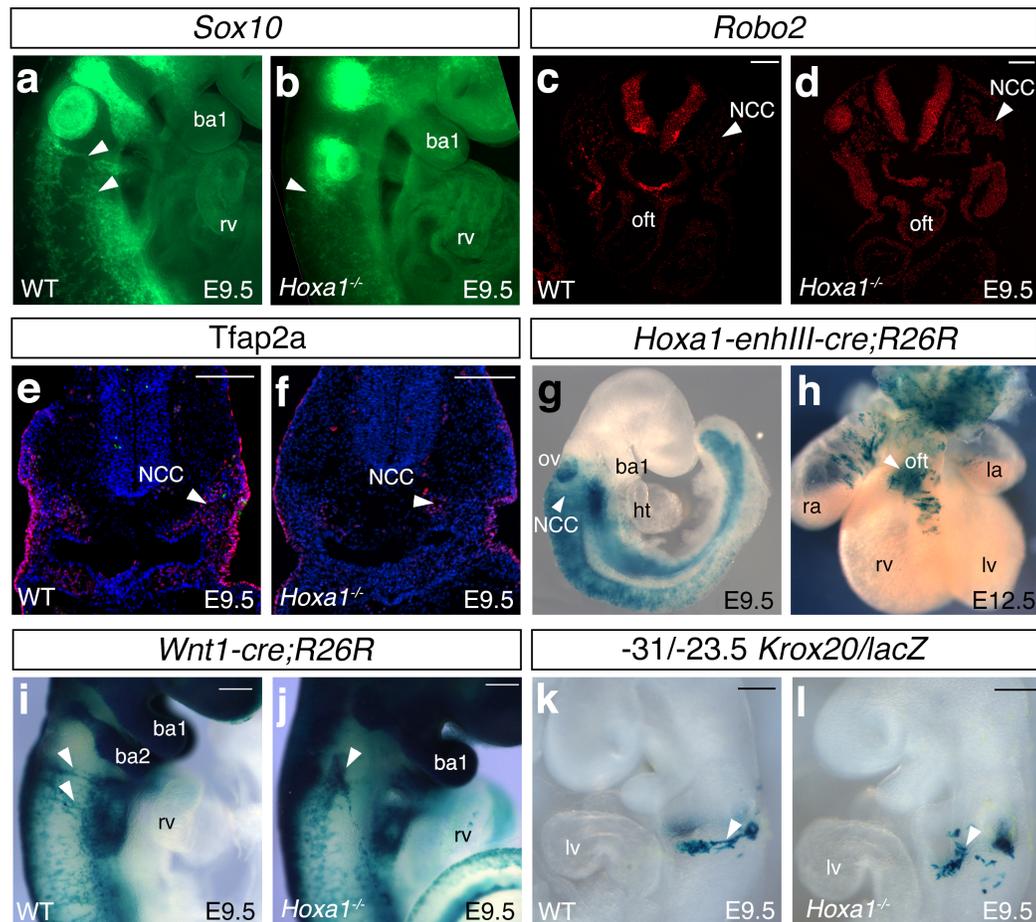


Figure 7

Figure 7

Neural crest cell defects in *Hoxa1* null mice. **a-b**, RNA-FISH showing the expression of *Sox10* (green) in wild-type (WT) (**a**) and *Hoxa1*^{-/-} (**b**) E9.5 embryos. Lateral view of *Sox10* staining revealed a decrease of expression in the migrating neural crest cells (arrowheads). **c-d**, RNA-FISH showing *Robo2* expression (green) in wild-type (**c**) and *Hoxa1*^{-/-} (**d**) E9.5 embryos. Expression of *Robo2* is higher in *Hoxa1*^{-/-} than wild-type littermates. **e-f** Cross-sections of wild-type (**e**) and *Hoxa1*^{-/-} embryos (**f**) immunostained with anti-Tfap2-alpha antibody (red). Note the reduction of Tfap2-alpha positive cells in the mutant (arrowhead). **g-h**, *Hoxa1*-lineage visualized by X-gal staining of *Hoxa1*-*enhIII*-*cre;R26R-lacZ* embryos. Lateral view showing β -galactosidase activity in migrating neural crest cells (NCC; arrowhead) in transgenic embryo at

E9.5. At E12.5, β -galactosidase activity is observed in the outflow tract (OFT). **i-j**, Whole-mount X-gal staining of *Wnt1-cre;R26R-lacZ* WT (**i**) and *Hoxa1^{-/-}* (**j**) E9.5 embryos. Migration of neural crest cells (arrowhead) is disrupted in *Hoxa1^{-/-}* embryos (**j**) compared to WT (**i**) littermate embryos. **k-l**, Whole-mount X-gal staining of -31/-23.5 *Krox20/lacZ* WT (**k**) and *Hoxa1^{-/-}* (**l**) E9.5 embryos. Migration of β -galactosidase-positive cells (arrowhead) is disrupted *Hoxa1^{-/-}* (**l**) compared to WT (**k**) littermate embryos. ba: branchial arch; ht, heart tube; la, left atrium; lv, left ventricle; NCC, neural crest cells; ov, otic vesicle; oft, outflow tract; ra, right atrium; rv, right ventricle. Scale bars: 100 μ m (**c-f**), 200 μ m (**i-l**).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFiguresOdelin.pdf](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.xlsx](#)