UNIVERSITY OF GENEVA

MONOGRAPH WORK OF THE BACHELOR IN BIOLOGY

CONSERVATION AND REGULATION OF *SHOX* GENES

Elise COCHAIN
No. 23-321-615
Complementary studies in sciences

Directors of the monograph
ROUCO GARCIA Raquel
ANDREY Guillaume
Faculty of medicine, CMU, Dep. of genetic medicine and development

Supervisor of the monograph
STEINER Florian
Faculty of sciences, Dep. of molecular and cellular biology

TABLE OF CONTENT

ABSTRACT

Homeobox genes are known to have a significant importance in research in the field of developmental biology by their involvement in regulating both spatially and temporally various target genes. Within this diverse group of genes, the *SHOX* genes family consists of two genes in human genome: *SHOX* and *SHOX2*. These genes play a crucial role during embryonic development by acting as transcriptional factors. In human, *SHOX* is located on both sex chromosomes, where it escapes from X chromosome inactivation, while *SHOX2* is located on the autosome chromosome 3. *SHOX* genes appeared in vertebrate and are found in the genome of most species where they are highly conserved, with a notable loss of *SHOX* gene in rodents during their evolution. Both *SHOX* and *SHOX2* show a similar expression pattern across species, *SHOX* being primarily expressed in limb buds and pharyngeal arches and *SHOX2* in the heart and nervous system. Being genes involved in embryonic development, they are highly regulated making them likely to cause disorders when they are disrupted.

1. INTRODUCTION

In the study of developmental biology, a crucial set of genes hold a significance importance for research: homeobox genes. These genes are characterized by the presence of a homeobox, a DNA sequence encoding a homeodomain (1). The homeodomain is a 60 amino acids conserved sequence arranged in a helix-loop-helix-turn-helix structure that serves as a DNA binding motif (2). Primarily involved in organogenesis and pattern formation during embryonic development, homeobox genes regulate both spatially and temporally various target genes, functioning either as transcriptional activators or repressors (3). That is why genetic mutations affecting these genes or their encoded protein can lead to structural abnormalities, physiological impairments and potentially embryonic death (4). In animals, homeobox genes are categorized into 11 classes (ANTP, PRD, LIM, POU, HNF, SINE, TALE, CUT, PROS, ZF, CERS), showing their expansion in the early evolution of Metazoans (5) (6). In the human genome, over 200 homeobox genes can be found, among which Hox genes are the most known due to of their implication along the antero-posterior axis (5). Another large and diverse group of homeobox genes is the short stature homeobox (SHOX) gene family, which in humans consists of two genes : *SHOX* and *SHOX2* (6). Initially, the *SHOX* gene was identified in a clinical context and became the focus of research due to its potential involvement in short stature (7). At the time, it was suggested that human sex chromosomes were involved in growth, so a gene involved in short stature was searched for in the Xp and Yp pseudoautosomal region 1 (PAR1) (7) (8). While *SHOX2* was discovered later because of its homology to *SHOX* and its higher homology to rodent *Shox2* gene. At first, it was initially named *SHOT* (SHOX HOmologous gene on chromosome Three) due to its localization on the third chromosome (9).

2. *SHOX* AND *SHOX2* CONSERVATION ACROSS SPECIES

It is suggested that *SHOX* genes first appeared in vertebrate species, as both *SHOX* and *SHOX2* seem to be absent from the genomes of invertebrate species (10). They exhibit a high conservation between various vertebrate species but they are not universally present in all species. In fact, *SHOX* is highly conserved in fish and chicken, but it is not detected in the genome of frog, rabbit and rodent. While *SHOX2* is detected in every vertebrate species except for dogs (Fig. 1) (10). The high conservation of *SHOX* genes between species suggests the evolutionary importance of maintaining the protein sequence among these genes (11).

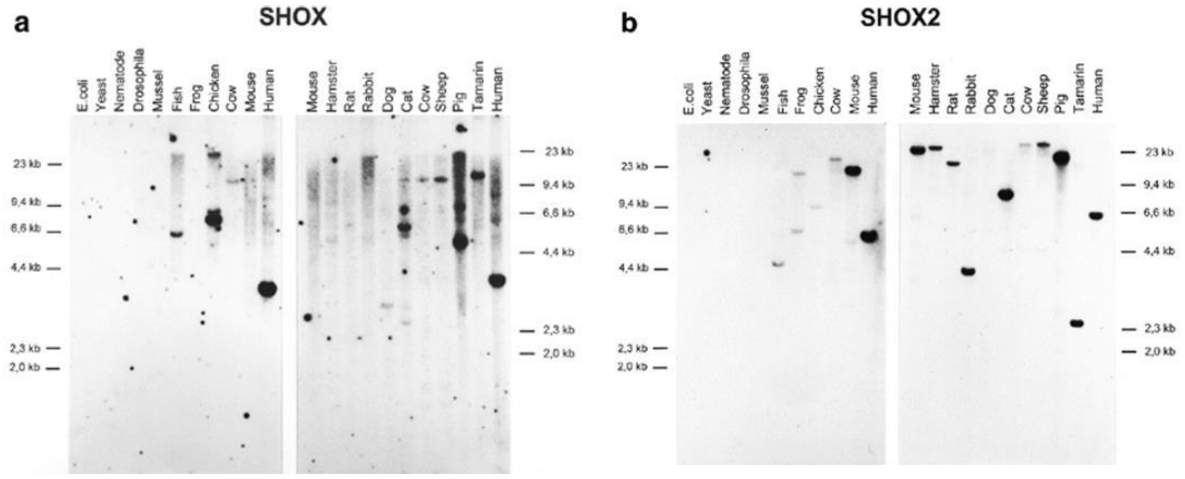

Figure 1: Zooblot hybridizations for SHOX (a) and SHOX2 genes (b). The DNA from several species was used for the zooblot: fish, frog, chicken, cow, mouse, hamster, rat, rabbit, dog, cat, sheep, pig and tamarin. Modified from Clement-Jones et al (10).

Human

In humans, both *SHOX* and *SHOX2* are present and expressed in the genome (10). The *SHOX* gene is located on the PAR1 on both X and Y chromosomes (3). Studies of *SHOX* expression in embryos have revealed its presence in various fetal tissues (muscle, skin, intestine, eye, brain and spinal cord) (Fig. 2) and adult tissues (placenta, skeletal muscle, bone marrow, adipose tissue and brain) (12) (3). But *SHOX* is first expressed during limb development, in the middle portion of the arm that later becomes the distal end of the humerus, the ulna, the radius and some bones of the wrist (10). The spatiotemporal expression pattern of *SHOX* in upper limbs development is similar as what is seen in lower limb development (10). Additionally, *SHOX* expression is observed in the first and second pharyngeal arches of the mesenchymal core, which later develop into the maxilla, mandible and middle ear (Fig. 2) (10). On the other hand, the *SHOX2* gene is located on an autosomal chromosome, the chromosome 3 (13). *SHOX2* expression is found in various tissues, including the upper limbs where it is first detected but more proximal than *SHOX* (Fig. 3) (3). But also the first, second and third pharyngeal arches, craniofacial region, nasal process, central nervous system, and the heart where it is involved in the development of cardiac system (10) (14).

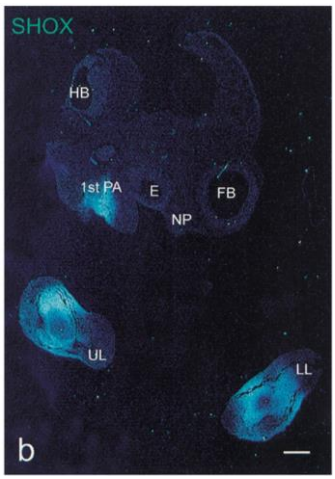

Figure 2: Lateral and sagittal sections of a human embryo at CS16 viewed by dark-field microscopy after an antisense prob hybridization to SHOX. An expression is detected in the upper limbs (UL), the lower limbs (LL), the first pharyngeal arch (1st PA), the hindbrain (HB), the eye (E), the nasal process (NP), and forebrain (FB). Modified from Clement-Jones et al., 2000 (10).

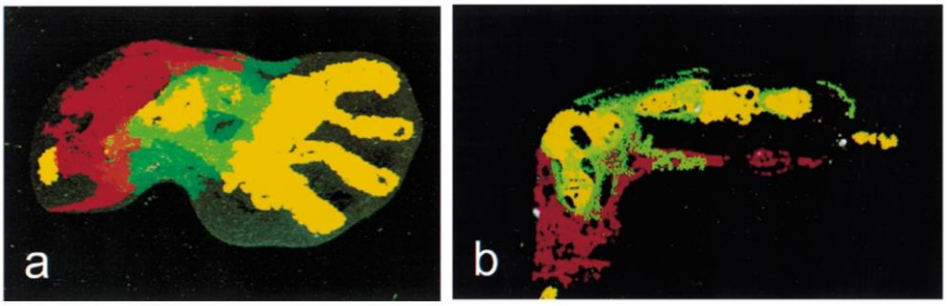

Figure 3 : Digital comparison of SHOX and SHOX2 expression in the upper limb at stage CS 18 (a) and CS 21 (b). Pattern expression of SHOX (green) and SHOX2 (red), as well as SOX9 (yellow) as positive control. SHOX2 expression is expressed more medially and proximally compared to SHOX. SHOX expression appears more concentrated around the developing bone. A certain overlap is noticeable around the connective tissue that surrounds the developing bone. Modified from Clement-Jones et al., 2000 (10).

Rodent

Rodents have lost their *SHOX* gene orthologue, along with other genes during evolution (6). So in rodents, only a *SHOX2* orthologue, initially called *Og12x*, is found in their genome (11). As in humans, it is located on chromosome 3, location that is syntenic between both species (11) (9). Mouse *Shox2* is strongly conserved, it shows 99% of common amino acid with the human *SHOX2* and 79% of common amino acid with the human *SHOX,* suggesting a common evolutionary origin between these genes (11). During embryo development between 9 and 12 days post-coitum (pc), *Shox2* mRNA is primarily observed in the developing heart, especially in the sinus venosus region. Later, it is also observed in the otic region, nasal process, limb buds where is restricted to the proximal and middle portion, dorsal root ganglia, diencephalon, mesencephalon, and first brachial arch (Fig. 4) (11). Overall, the tracking of *SHOX2* transcripts is restricted to these 4 structures : the central nervous system, the heart, the craniofacial tissues and limb buds (fore- ad hind- limbs) with temporal and spatial expression similar to human *SHOX2* (10). The loss of *SHOX* in rodents may explain their specific anatomy characterized by long feet and short bones suggesting the involvement of *SHOX* gene in limb extension and while *SHOX2* may be involved in limb development (8). But the lack of a *SHOX* orthologue in the rodent genome suggests that rodent *Shox2* might be involved in more functions during their embryogenesis than humans (4).

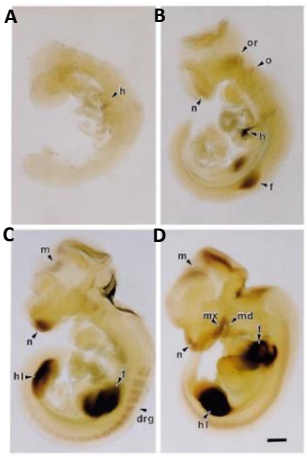

Figure 4: RNA probe in situ hybridization of Shox2 in mouse embryo, 9 days pc (bar ~222 µm) (A), 10 days pc (~357 µm) (B), 11 days pc (~400 µm) (C), 12 dayc pc (bar ~588 µm) (D). (h) heart, (n) nasal processes, (or) otic region, (o) otocyst region, (f) forelimbs, (m) midbrain, (hl) hindlimbs, (mx) maxilla, (md) mandibula, (drg) dorsal root ganglia. Modified from Semina EV et al, 1998 (11).

Chicken

In chicken, both *SHOX* and *SHOX2* orthologues are present in their genome. Compared to humans, the *Shox* gene is not located on a sex chromosome but on an autosomal chromosome, specifically chromosome 1 (15). Still, the conservation of *Shox* between both species is very high, sharing 94% of homology in their amino acid (15). During chick embryo development between stages 18 to 36, *Shox* is initially expressed in wing and leg buds, where its expression is found in the proximal region, being its expression reduced where cartilage is condensing as well as the vasculature. *Shox* is also expressed in wing and legs vasculature, visceral arches, neural tube, mandibular primordia, the brain, and around the eye (Fig. 5.1). However, its primary involvement is suggested to be in the development of wings and legs (15). For *SHOX2* orthologue, the gene is also located on an autosomal chromosome, chromosome 9 (15). During chick development, *Shox2* is found between stages 19 et 27. At the beginning, *Shox2* is first found in the sinus venosus where its expression is higher over time. A bit later, *Shox2* is also expressed in limb buds where its expression domain is bigger in the wing buds compared to the leg bud. Later, *Shox*2 is still expressed in the heart as well as the neural tube in the dorsal root ganglia (Fig. 5.2). In the limbs, *Shox2* is later expressed in the tissue surrounding muscles, and in tendons of the hand plate (15).

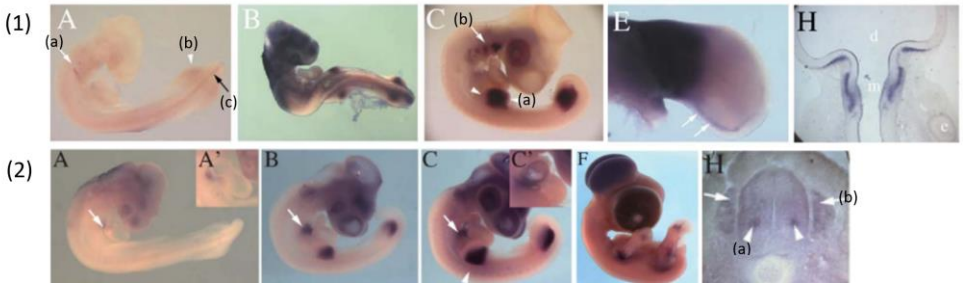

Figure 5 : Expression of Shox and Shox2 in chick embryos by in situ hybridization. (1) Shox expression in chick embryo. (A) Branchial arches (a), hindlimbs buds (b), and neural tube (c) at stage 19. (B) Head and limbs buds at stade 20. (C) Central region of the limb bud (a), central region of brachial arches (b) at stage 22. (E) Vasculature (arrows) in the posterior bud at stage 25. (H) Brain, in the diencephalon (d) and metacoele (m), and the eye (e) on a sagittal section at stage 25. (B) Shox2 expression in chick embryos. (A) Sinus venosus (arrow) and (A') magnification at stade 19. (B) Sinus venosus (arrow) at stage 21. (C) Central region of the limb bud, root ranglia and sinus venosus, and (C') atria at stage 23. (F) Neural tube (a) and dorsal root ganglia at stage 23. (H) Neural tube (a) and dorsal root ganglia by section at stage 23). Modified from Tiecke E et al, (2006) (15).

Zebrafish

In zebrafish, both *SHOX* and *SHOX2* orthologues are present in their genome. The *shox* gene is located in chromosome 9 (8). The zebrafish SHOX protein is also very well conserved, sharing 84% homology in amino acid with human SHOX protein (16). The first signs of *shox* gene expression are quite early during zebrafish development, at 24 hours post fertilization (hpf), where it is first found in blood, putative heart, hatching gland and brain. A bit later, *shox* is also expressed in the pharyngeal arch, the olfactory epithelium, and the fin bud (Fig. 6) (8). The *shox2* gene is located on chromosome 15. Its expression is detected in otic placode, diencephalon, cranial ganglion neurons, hindbrain, optic tectum, and inner ear, as well as the developing heart (17) (18) (19).

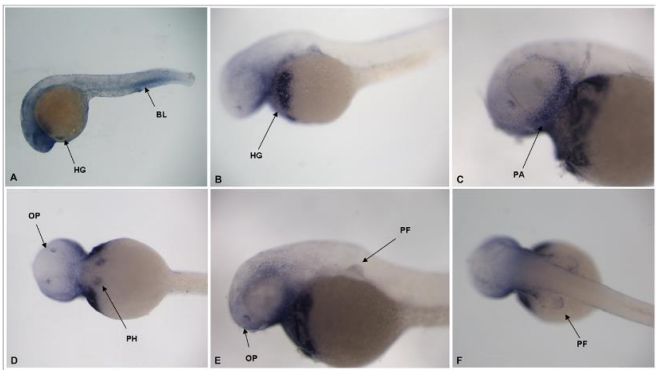

Figure 6. In situ hybridization of Shox in whole zebrafish embryos (8). (A) Hatching gland (HG) and blood (BL) at 24 hpf. (B) HG at 72 hpf. (C) Pharyngeal arch (PA) at 72 hpf. (D) Olfactory pits (OP) and putative heart (PH) at 72 hpf in a ventral view. (E) Pectoral fin (PF) at72 hpf. (F) PF at 72hpf in a dorsal view. Modified from E. Kenyon et al, 2011 (8).

It is interesting to point out that the *SHOX* genes localization varies across species. This is a good example of genomes evolution between species. The *SHOX* gene is present across different species, but its positioning varies, sometimes being located on a sex chromosome and other times on an autosomal chromosome. In fact, regarding sex chromosomes evolution, they probably evolved from a pair of autosomal chromosomes, after acquiring a sex-determining factor and undergoing additional changes (20). The mechanism of sex determination can differ among species, being determined by genotypic traits like chromosomes or environmental factors such as temperature (21). Mobile genetic elements can also play a role in the different localization of the *SHOX* gene among species by being known to have an important impact on genome evolution (22). Additionally, chromosome rearrangement, such as translocation, is also known to be an important element of genetic variation, suggesting why genes can be preserved between species but not found at the same location on chromosomes (23). An illustrative instance is the study of Down syndrome in mice. In human it is caused by the trisomy of chromosome 21, while in mice, several models have been developed trying to reproduce the syndrome like the one involving 3 segments on 3 distinct chromosomes (chromosome 10, 16 and 17) (24). Another example is been provided by comparative genome analysis, it has revealed a high degree of homology between human X chromosome and chicken chromosomes 1 and 4, localization of *SHOX* gene in this specie (25).

Overall, the presence of *SHOX* genes in vertebrate genome suggests their implication in the development of the skeleton as well as its related structure (4). Despite their various localization in the genomes of different species, showing how genomes evolved differently across species, their function is quite conserved. In addition, *SHOX* expression is more limited compared to *SHOX2*, and their overlapping expression pattern often shows a certain complementarity (26). This suggesting that both genes have an important role in embryonic development because their functions are not totally redundant in humans, as well as in species that also have both genes (10).

3. THE REGULATION OF *SHOX* AND *SHOX2*

<u>*SHOX* regulation</u>

In humans, the *SHOX* gene is located approximately 500kb away from the telomeres of both sex chromosomes, withing PAR1, spanning around 40kb (Fig. 7) (3). In our species, sex determination relies on sex chromosome, with the X chromosome containing many genes (~900-1500) compared to the Y chromosome which contains only a few (~70) (27). The presence of two copies of the X chromosome in females and only one in males can lead to a dosage imbalance in gene expression so to mitigate this imbalance between XX and XY individuals, one of the X chromosomes undergoes a random inactivation during female embryonic development (7). This process, known as X chromosome inactivation (XCI),

is orchestrated by chromatin and DNA modifications, with long-non coding RNAs (lncRNAs) playing a crucial role in its regulation. The *Xist* lncRNA coats the X chromosome by recruiting proteins that modify its structure and putting epigenetic marks to inactivate it. Additionally, epigenetic modifications characteristic of inactivated chromatin, such as H3K9me3, H4K20me3, H3K27me3 and macroH2A, are observed on the inactivated X chromosome, along with a depletion of active marks such as H3K4me3 (28). But not the whole X chromosome is inactivated, PAR regions contain escaper genes. In the human genome, the *SHOX* gene is located on PAR1 region and is an escaper gene because both copies are necessary for normal development. So it is present in two functional copies, one on each sex chromosome (29). These escaper genes have evolved by developing mechanisms to evade silencing so they can continue to be expressed from both the active and inactive X chromosome (28). In human, approximately 15% of genes on the inactivated X chromosome are escaper genes. These genes lack epigenetic marks associated with inactivated genes but exhibit active histone marks such as H3K4me2, H3K4me3, H3K9ac, H3K27ac, and H3K9me1, along with a depletion of inactive marks (28).

Containing 7 exons, the *SHOX* gene generates 2 transcripts, *SHOXa* and *SHOXb*, through alternative splicing. Both transcripts include exons 1-5 but differ in their 3' end due to the inclusion of different final exons, *SHOXa* lacks exon 7 while *SHOXb* is lacking exon 6 (Fig. 7) (29). Although both transcripts feature a homeodomain, *SHOXb* lacks the OAR (otp, aristaless, and rax) transactivation domain, meaning it is not capable of functioning as a transcriptional activator (3). The expression of the *SHOX* gene is under the control of two alternative promoters, P1 and P2. The presence of two promoters remains uncertain, however, it is suggested that one or the other is preferentially used based on the physiological context (3).

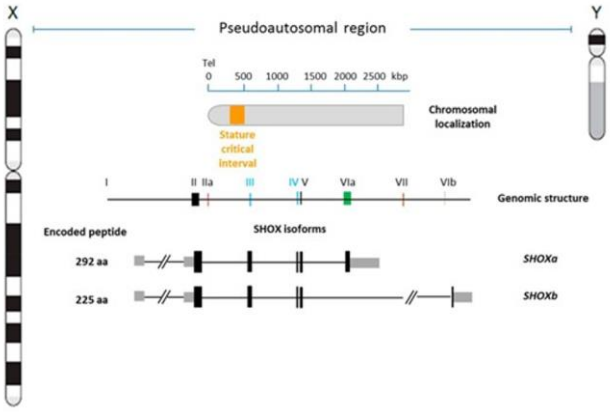

Figure 7: SHOX gene localization and encode transcripts. The SHOX gene is located on both X and Y chromosomes, about 500kb away from telomeres on PAR1 region. The SHOX gene encodes 2 isoforms SHOXa and SHOXb. Modified from Marchini A et al, 2016 (3).

Within the mammalian genome, approximately 95% contains non-coding regions, where around 40% is believed to have a regulatory potential (30). Usually, for developmental genes, such as *SHOX*, promoters are not sufficient to precisely regulate gene transcription and rely additionally on cis-regulatory elements that allow a correct spatiotemporal expression of the genes (31) (32). These regulatory elements, which mainly function as enhancers or as repressors, play a vital role in gene expression by regulating their expression across various tissues and can be located proximal or distal from their target gene (29). In the case of *SHOX* gene, cis-regulatory elements spanning a 1Mb region of PAR1 control its expression (33). Conserved non-coding elements (CNEs) primarily function as enhancers during embryonic development, coordinating spatial and temporal gene expression patterns and usually cluster around developmental genes (34). CNEs are distributed in the genome at nonrandom locations, mostly in regions with low density genes (34). Within PAR1, 11 CNEs have been identified, with 3 evolutionary conserved regions located upstream (CNE-5, CNE-3, CNE-2) and 8 located downstream (CNE2, CNE3,

CNE4, CNE5, CNE6, CNE7, CNE9) of the *SHOX* gene (Fig. 8.A) (29). In chickens limbs buds, chromatin interactions between several CNE and *Shox* promoter have been shown, as well as in human U20S cells (31). More generally, CNEs are very well conserved across species where *SHOX* is present with over 70% sequence conservation, where the level of conservation of CNEs was found to decrease as the evolutionary distance increased (34) (36). In zebrafish, CNE5 predominantly induces expression in the central nervous system, while CNE9 exhibited more frequent activity in the eyes, with less frequent expression observed in the fins (37). In chickens, it has been revealed that only CNE4, CNE5 and CNE9 exhibit enhancer activity during limb development (36). Among all the CNEs, it is suggested that CNE9 is particularly important to regulate SHOX limb expression since it is situated in a region frequently deleted in Leri-Weill dyschondrosteosis (LWD) (37). Moreover, in mice, the removal of CNE5 located downstream *Shox2* leads to the absence of the limb expression domain (36). So it is suggested that the conservation of CNEs implies a similar involvement across species (3).

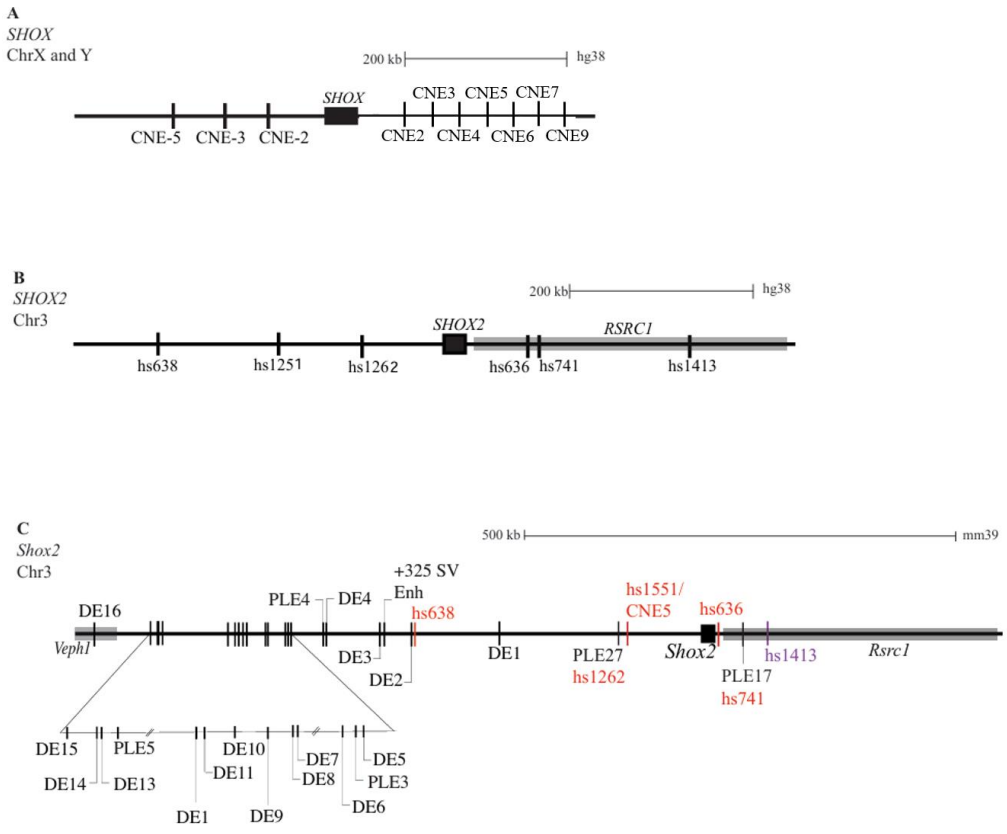

Figure 8: Summary figure of SHOX genes enhancers. (A) Human SHOX evolutionary conserved regions: CNE-5, CNE-3 and CNE-2 are located upstream of the gene, while CNE2, CNE3, CNE4, CNE5, CNE6, CNE7, CNE8 and CNE9 are located downstream. (B) Human SHOX2 enhancers. (C) Mouse Shox2 enhancers. In red, human enhancers where their function is also conserved in mouse. In purple, enhancers that have not been tested in mouse embryo but show a sequence conservation.

<u>*SHOX2* regulation</u>

In humans, *SHOX2* is located on chromosome 3, 3q25-q26, and was first discovered because of its homology with *SHOX* and mouse *Shox2* (9). As *SHOX*, *SHOX2* encodes two isoforms, *SHOX2a* the longer isoform and *SHOX2b* the shortest isoform (9). As for SHOX, these two isoforms are produced by two different promoters (19). Although the precise functions of the isoforms remain unclear, it is known that mouse *Shox2a* and *Shox2b* isoforms serve as tissue-specific transcription factors during embryogenesis (14) (38).

That location of the mouse *Shox2* is known to be around a gene desert, a chromosomal region spanning at least 500kb without protein-coding gene (39). Gene deserts are regions that play a crucial role in embryonic development and organogenesis due to their frequent proximity to developmental genes and their abundance of regulatory elements (39). It is believed that 25% of the human genome are gene deserts (39). The mouse *Shox2* gene desert has been suggested to be similar to the human *SHOX* gene desert (39) Therefore, it is been suggested that human *SHOX* gene desert could share a functional homology with the mouse *Shox2* and be required during embryonic development by regulating *SHOX* transcription (39). In mouse, *Shox2* is positioned approximately 675kb downstream from the centromeric region of the third chromosome and is adjacent to a stable gene desert (Fig. 9.A) (39). The *Shox2* gene desert contains distal enhancers, where their interaction with *Shox2* promoter is facilitated within a topological domain (39). As previously mentioned, *Shox2* is under the regulation of several enhancers, that are located within or in the border region of a topologically-associated domain (TAD) (39). TADs are kilo- to mega- base segments that are part of the genome architecture and organization (40). They are a 3D structure element that play a crucial role in regulating gene expression by confining interactions between cis-regulatory elements and their target. The TAD itself that is the self-associating domain, loop-like that contains cis-regulatory elements and their target gene. Positioned in between TADs specific proteins such as CTCF act as TAD boundaries, that function as insulating elements to limit interactions, in fact, removing them leads to ectopic gene expression (41). TADs are found in the genome of various species, from human to Caenorhabditis elegans, showing degrees of variation. TADs are highly conserved in mammals, where their average size is about 880kb (42). TADs size varies a lot within a specie and across species (42).

The mouse *Shox2* gene is located within a 1Mb TAD, with another gene, *Rsrc1*, located upstream *Shox2* (Fig. 9.A) (39). The sub-structure delineates into upstream (U-dom) and downstream (D-dom) domains flanking *Shox2*, with the D-dom extending across nearly the entire gene desert (39). Within *Shox2*-TAD at least 20 elements showing enhancers activity where identified in mouse embryos (39). Each of these enhancers are involved in distinct roles regulating *Shox2* transcription in specific tissues and participating in various embryonic developmental processes, underscoring the pleiotropic nature of *Shox2* (39). In fact, *Shox2* gene desert primarily contributes to proximal limb morphogenesis, craniofacial patterning, as well as defining the sinus venosus and the sinoatrial node (SAN) by encoding tissue specific enhancers (Fig. 9.B) (39). In addition, *SHOX2* gene desert is essential for embryonic viability (39). In human, *SHOX2* is also under the regulation of various cis-regulatory elements such as hs741, hs636, hs1262, hs1251, and hs638. Moreover, the core of the mouse +325SV enhancer is also conserved in humans (Fig. 8.B) (39).

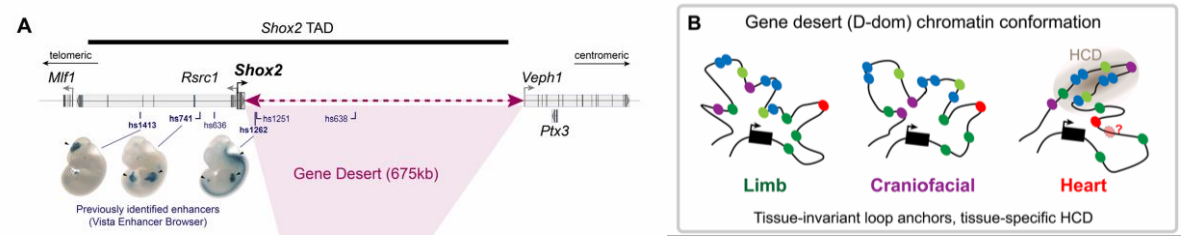

Figure 9 : Mouse Shox2 regulatory landscape. (A) Mouse Shox2 TAD and gene desert localization. (B) Chromatin conformation of the Shox2 gene desert with tissue-specific enhancers involved in regulation of enhancers activities. Modified from Abassah-Oppong S. et al, 2023 (39).

Human *SHOX* and *SHOX2*, and mouse *Shox2* regulatory landscape conservation

Overall, the regulatory landscape of *SHOX* genes shares similarities between species. The regulatory landscape of the human *SHOX* and *SHOX2* can be studied by using the mouse as a model even if they have lost their *SHOX* gene during their evolution because the transcription control of the mouse *Shox2* exhibits a certain level of conservation (37). In fact, mouse *Shox2* and human *SHOX* genes share a similar-sized gene desert, as well as an identical DNA-interacting homeodomain (39). In addition, human *SHOX2* shares a conserved locus architecture with mouse *Shox2* (39). The human *SHOX2* also has a similar downstream gene desert to the mouse one, that contains neural enhancer (hs1251, hs1262), SAN formation enhancer (+325SV enhancer) and stylopod enhancer (hs741, hs1262 and hs638). But regarding mouse *Shox2* regulatory landscape, the loss of stylopod enhancer does not give a complete loss of *Shox2* limb expression, suggesting a higher complexity and redundancy within enhancers landscape in mouse (39).

Mouse *Shox2* expression is under the regulation of various enhancers. Some of these enhancers are related to the human enhancers. In fact, some of the human enhancers have been tested in transgenic mouse embryo and show a conservation of their regulatory potential. In fact, hs741, hs636, hs1262, hs1251, hs638, and CNE5 keep their enhancer activity and expression pattern in mouse embryo (Fig. 8.C) (43) (44) (37) (VISTA). But all enhancers have not been tested yet, so the study of the sequences homology can be used as an indicator of the conservation of these enhancers between humans and mouse. By using BLAST, out of the CNEs, only CNE5 did show a 100 bases sequence conservation of about 85% between those species (BLAST). As well as hs1413, it has not been tested but it shows about 92% of conservation in the mouse genome (Fig. 8.C) (BLAST). So, these enhancers do show a sequence conservation but have to be tested on mouse embryo to be able to say that they do have a conservation of function. Overall, even though mouse have lost their *SHOX* gene during evolution, their regulatory landscape seems to contain regulatory sequences that combine *human SHOX* and *SHOX2* enhancers.

4. CLINICAL IMPLICATION OF *SHOX* AND *SHOX2*

As development is a process finely regulated, mutation in homeobox genes has long been known to be the cause of several human disorders, as well as the mutation of their enhancers (11) (37). The *SHOX* gene was first described in 1997 because of its implication in short stature in Turner Syndrome (TS) patients but now, its implication in other disorders such as LWD or Langer Mesomelic Dysplasia (LMD) or Idiopathic Short Stature (ISS) was also highlighted (45). Each of these disorders is the result of a distinct aberration of the *SHOX* gene (29). On the other hand, for a long time it was thought that *SHOX2* was not linked with any diseases, but more recently, *SHOX2* disruption has been associated with atrial fibrillation disease (14).

Turner Syndrome

The TS is a rare disorder, affecting 1/2500 female's world-wide, manifesting primarily as short stature, yet encompassing a spectrum of associated medical conditions (10). This includes cardiovascular disorders (e.g. bicuspid aortic valve and aortic dilatation) endocrine and metabolic disorders (e.g. hypogonadotropic hypogonadism, glucose intolerance,) gastrointestinal and hepatic disorders (indicated by elevated levels of hepatic enzymes), among others, with varying prevalence rates (46). TS is associated with complete loss (45,X0 karyotype), partial loss of the X chromosome, and other chromosomal abnormalities (10). It was revealed that a deletion in the terminal region of the short arm of the sex chromosome, currently known as the PAR1 region, leads to a short stature, thereby highlighting the involvement of the *SHOX* gene in growth process (7). In fact, *SHOX* plays a critical role in growth by modulating various genes involved in chondrocyte development and being involved in bone morphology (33) (10). Nevertheless, *SHOX* effect is beyond its implication in height where *SHOX* pattern expression is linked with TS clinical expression (10). Indeed, *SHOX* expression in first

and second pharyngeal arches allows the understanding of common features observed in TS patients, such as high-arched palate, micrognathia, sensorineural deafness or middle ear infection. However, it is unlikely that *SHOX* is involved in the development of all non-skeletal somatic features often seen in TS patients (10).

Leri-Weill dyschondrosteosis

LWD is a disorder characterized by short stature, Madelung deformity of the wrist, and mesomelia. LWD also present other variable features such as scoliosis, increased BMI, high-arched palate, among others (47). *SHOX* deficiency has been identified as an underlying factor of LWD, in approximately 70-90% of LWD, the condition is attributed to a defect in cis-regulatory element or a heterozygous deletion. Where in around 15-40% of the time, it is associated with a deletion of enhancer elements, which is the highest rate of enhancer disruption in any disease (3). The CNE9 represents the most frequent deleted element in LWD patients, but CNE4 and CNE5 seem to be as much important (36). These deletions are probably linked with *SHOX* localization on PAR1, which goes 17 times more under recombination that the rest of the genome during male meiosis (35). Even if LWD is often inherited, the deletion of the enhancers typically arises as a de novo mutation (48). In addition to that, LWD can also be caused by a haploinsufficiency of the *SHOX* gene (49).

Langer mesomelic dysplasia

While LWD can be caused by the loss of one copy of the *SHOX* gene, LMD is caused by the loss of both copy of the gene, as a result being considered the homozygous manifestation of LWD (49). LMD is a rare disorder, characterized by an extreme phenotype of skeletal dysplasia (50). Its clinical expression is mainly characterized by a short stature, mesomelic and rhizomelic dysgenesis of the limbs involving hypoplasia/aplasia of the ulna and fibula (49). Other manifestations can be found, such as a bilateral conductive hearing loss, but is part of the rare manifestation of the gene deficiency (50). In addition to the biallelic loss of the gene, the implication of *SHOX* regulatory regions has been identified in about 75% of cases in LMD, either a mutation or a deletion (50).

Atrial fibrillation

The heart functions as a pump due to its electrical activity mediated by pacemakers that regulate heartbeats (38). The primary pacemaker of the heart is the SAN, located on the wall of the right atrium. During heart development, all cardiomyocytes possess pacemaker abilities, but later differentiate to form the working myocardium, while a subset contributes to the SAN and conductive tissues (38). During embryo development, *Shox2* is expressed in the developing heart where it plays a crucial role in the development and function of the SAN by controlling the expression of Nkx2.5, the primary transcription factor involved in heart development, acting downstream of *Shox2* (38) (51). In fact, it is the balance in the expression patterns of *Shox2* and *Nkx2.5* that is essential for the differentiation of cardiac progenitor cells (14). While for a long time no diseases associated with *SHOX2* in humans were previously reported, it was known that the disruption of *SHOX2* orthologues genes lead to cardiovascular defects in other species (37) (52). Thus, *SHOX2* was suggested to be involved in atrial fibrillation in human, a condition that affects the electrical activity of the heart, resulting in a fast and irregular heartbeats due to abnormal electrical activity, touching 1-2% of the population (52). In fact, a study has identified a positive correlation between a mutation in the *SHOX2* gene and early-onset atrial fibrillation. This mutation involves a substitution of a histidine with a glutamine and is located on the 6th exon of the gene. It adversely affects the transactivation activity of *SHOX2*, no longer allowing it to regulate its targets (52). Additionally, the measure of *SHOX2* expression in patients heart tissues showed that its expression is significantly reduced in atrial fibrillation patients compared to healthy patients (52).

5. CONCLUSION AND PERSPECTIVES

SHOX genes are homeobox genes important for embryonic development. Both *SHOX* and *SHOX2* are expressed during embryogenesis in different species, with an exception in a few species such as rodents that have lost *SHOX* during their evolution, making chicken and zebrafish privileged model organisms. These genes are well conserved across species and show a similar expression in various tissues. The *SHOX* gene is mainly involved in limb and pharyngeal arches development while *SHOX2* in heart and nervous system development, but both show a pleiotropic expression, and sometimes a partially overlapping but complementary expression. By being genes involved in development, *SHOX* genes are highly regulated genes, under the regulation of many cis-regulatory elements. In addition, the regulatory landscape of these genes is conserved between species, showing similarities between the human *SHOX* and *SHOX2*, and the mouse *Shox2* even though the loss of the *SHOX* gene during evolution of rodents. The complex regulation of these genes makes it easy to disrupt, leading to various disorders such as TS, LWD, LMD, or AF. More globally, a better understanding of the regulation of *SHOX* genes would allow a better understanding of human embryogenesis, as well as human syndromes and congenital abnormalities. Leading to better patient care and the possibility of offering innovative treatments.

REFERENCES

1. Holland PWH. Evolution of homeobox genes. WIREs Developmental Biology. 2013;2(1):31-45.

2. Bürglin TR, Affolter M. Homeodomain proteins: an update. Chromosoma. 1 juin 2016;125(3):497-521.

3. Marchini A, Ogata T, Rappold GA. A Track Record on SHOX: From Basic Research to Complex Models and Therapy. Endocrine Reviews. 1 août 2016;37(4):417-48.

4. Liu H, Espinoza-Lewis RA, Chen C, Hu X, Zhang Y, Chen Y. The Role of Shox2 in SAN Development and Function. Pediatr Cardiol. 1 août 2012;33(6):882-9.

5. Hubert KA, Wellik DM. Hox genes in development and beyond. Development. 16 janv 2023;150(1):dev192476.

6. Zhong Y fu, Holland PW. The dynamics of vertebrate homeobox gene evolution: gain and loss of genes in mouse and human lineages. BMC Evolutionary Biology. 16 juin 2011;11(1):169.

7. Lyon MF. Gene Action in the X-chromosome of the Mouse (Mus musculus L.). Nature. avr 1961;190(4773):372-3.

8. Kenyon EJ, McEwen GK, Callaway H, Elgar G. Functional Analysis of Conserved Non-Coding Regions Around the Short Stature hox Gene (shox) in Whole Zebrafish Embryos. PLOS ONE. 24 juin 2011;6(6):e21498.

9. Blaschke RJ, Monaghan AP, Schiller S, Schechinger B, Rao E, Padilla-Nash H, et al. SHOT, a SHOX-related homeobox gene, is implicated in craniofacial, brain, heart, and limb development. Proceedings of the National Academy of Sciences. 3 mars 1998;95(5):2406-11.

10. Clement-Jones M, Schiller S, Rao E, Blaschke RJ, Zuniga A, Zeller R, et al. The short stature homeobox gene SHOX is involved in skeletal abnormalities in Turner syndrome. Human Molecular Genetics. 22 mars 2000;9(5):695-702.

11. Semina EV, Reiter RS, Murray JC. A New Human Homeobox Gene OGI2X is a Member of the Most Conserved Homeobox Gene Family and is Expressed During Heart Development in Mouse. Human Molecular Genetics. 1 mars 1998;7(3):415-22.

12. Durand C, Roeth R, Dweep H, Vlatkovic I, Decker E, Schneider KU, et al. Alternative Splicing and Nonsense-Mediated RNA Decay Contribute to the Regulation of SHOX Expression. PLOS ONE. 23 mars 2011;6(3):e18115.

13. De Baere E, Speleman F, van Roy N, De Paepe A, Messiaen L. Assignment of SHOX2 (alias OG12X and SHOT) to human chromosome bands 3q25→q26.1 by in situ hybridization. Cytogenetics and Cell Genetics. 16 déc 1998;82(3-4):228-9.

14. Hu W, Xin Y, Zhao Y, Hu J. Shox2: The Role in Differentiation and Development of Cardiac Conduction System. The Tohoku Journal of Experimental Medicine. 2018;244(3):177-86.

15. Tiecke E, Bangs F, Blaschke R, Farrell ER, Rappold G, Tickle C. Expression of the short stature homeobox gene Shox is restricted by proximal and distal signals in chick limb buds and affects the length of skeletal elements. Developmental Biology. 15 oct 2006;298(2):585-96.

16. Sawada R, Kamei H, Hakuno F, Takahashi SI, Shimizu T. In vivo loss of function study reveals the short stature homeobox-containing (shox) gene plays indispensable roles in early embryonic growth and bone formation in zebrafish. Developmental Dynamics. 2015;244(2):146-56.

17. Laureano AS, Flaherty K, Hinman AM, Jadali A, Nakamura T, Higashijima S ichi, et al. shox2 is required for vestibular statoacoustic neuron development. Biology Open. 21 déc 2022;11(12):bio059599.

18. Hoffmann S, Roeth R, Diebold S, Gogel J, Hassel D, Just S, et al. Identification and Tissue-Specific Characterization of Novel SHOX-Regulated Genes in Zebrafish Highlights SOX Family Members Among Other Genes. Frontiers in Genetics [Internet]. 2021 [cité 2 févr 2024];12. Disponible sur: https://www.frontiersin.org/articles/10.3389/fgene.2021.688808

19. Blaschke RJ, Hahurij ND, Kuijper S, Just S, Wisse LJ, Deissler K, et al. Targeted Mutation Reveals Essential Functions of the Homeodomain Transcription Factor Shox2 in Sinoatrial and Pacemaking Development. Circulation. 10 avr 2007;115(14):1830-8.

20. Muller HJ. GENETIC VARIABILITY, TWIN HYBRIDS AND CONSTANT HYBRIDS, IN A CASE OF BALANCED LETHAL FACTORS. Genetics. 1 sept 1918;3(5):422-99.

21. Kratochvíl L, Stöck M, Rovatsos M, Bullejos M, Herpin A, Jeffries DL, et al. Expanding the classical paradigm: what we have learnt from vertebrates about sex chromosome evolution. Philosophical Transactions of the Royal Society B: Biological Sciences. 26 juill 2021;376(1833):20200097.

22. Platt RN, Vandewege MW, Ray DA. Mammalian transposable elements and their impacts on genome evolution. Chromosome Res. 1 mars 2018;26(1):25-43.

23. Damas J, Corbo M, Lewin HA. Vertebrate Chromosome Evolution. Annual Review of Animal Biosciences. 15 févr 2021;9(Volume 9, 2021):1-27.

24. Gupta M, Dhanasekaran AR, Gardiner KJ. Mouse models of Down syndrome: gene content and consequences. Mamm Genome. 1 déc 2016;27(11):538-55.

25. Graves JAM, Wakefield MJ, Toder R. The Origin and Evolution of the Pseudoautosomal Regions of Human Sex Chromosomes. Human Molecular Genetics. 1 déc 1998;7(13):1991-6.

26. Yu L, Liu H, Yan M, Yang J, Long F, Muneoka K, et al. *Shox2* is required for chondrocyte proliferation and maturation in proximal limb skeleton. Developmental Biology. 15 juin 2007;306(2):549-59.

27. Fang H, Disteche CM, Berletch JB. X Inactivation and Escape: Epigenetic and Structural Features. Frontiers in Cell and Developmental Biology [Internet]. 2019 [cité 4 févr 2024];7. Disponible sur: https://www.frontiersin.org/articles/10.3389/fcell.2019.00219

28. Balaton BP, Brown CJ. Escape Artists of the X Chromosome. Trends in Genetics. 1 juin 2016;32(6):348-59.

29. Spurna Z, Capkova P, Srovnal J, Duchoslavova J, Punova L, Aleksijevic D, et al. Clinical impact of variants in non-coding regions of SHOX – Current knowledge. Gene. 15 avr 2022;818:146238.

30. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. Nature. oct 2013;502(7472):499-506.

31. Verdin H, Fernández-Miñán A, Benito-Sanz S, Janssens S, Callewaert B, De Waele K, et al. Profiling of conserved non-coding elements upstream of SHOX and functional characterisation of the SHOX cis-regulatory landscape. Sci Rep. 3 déc 2015;5(1):17667.

32. Kleinjan DA, van Heyningen V. Long-Range Control of Gene Expression: Emerging Mechanisms and Disruption in Disease. Am J Hum Genet. janv 2005;76(1):8-32.

33. Fukami M, Seki A, Ogata T. SHOX Haploinsufficiency as a Cause of Syndromic and Nonsyndromic Short Stature. Molecular Syndromology. 12 avr 2016;7(1):3-11.

34. Polychronopoulos D, King JWD, Nash AJ, Tan G, Lenhard B. Conserved non-coding elements: developmental gene regulation meets genome organization. Nucleic Acids Research. 15 déc 2017;45(22):12611-24.

35. Fanelli A, Vannelli S, Babu D, Mellone S, Cucci A, Monzani A, et al. Copy number variations residing outside the SHOX enhancer region are involved in Short Stature and Léri-Weill dyschondrosteosis. Molecular Genetics & Genomic Medicine. 2022;10(1):e1793.

36. Sabherwal N, Bangs F, Röth R, Weiss B, Jantz K, Tiecke E, et al. Long-range conserved non-coding SHOX sequences regulate expression in developing chicken limb and are associated with short stature phenotypes in human patients. Human Molecular Genetics. 15 janv 2007;16(2):210-22.

37. Rosin JM, Abassah-Oppong S, Cobb J. Comparative transgenic analysis of enhancers from the human SHOX and mouse Shox2 genomic regions. Human Molecular Genetics. 1 août 2013;22(15):3063-76.

38. Liu H, Chen CH, Espinoza-Lewis RA, Jiao Z, Sheu I, Hu X, et al. Functional Redundancy between Human SHOX and Mouse Shox2 Genes in the Regulation of Sinoatrial Node Formation and Pacemaking Function *. Journal of Biological Chemistry. 13 mai 2011;286(19):17029-38.

39. Abassah-Oppong S, Mannion BJ, Zoia M, Rouco R, Tissieres V, Spurrell CH, et al. A gene desert required for regulatory control of pleiotropic Shox2 expression and embryonic survival [Internet]. bioRxiv; 2023 [cité 3 févr 2024]. p. 2020.11.22.393173. Disponible sur: https://www.biorxiv.org/content/10.1101/2020.11.22.393173v2

40. Okhovat M, VanCampen J, Nevonen KA, Harshman L, Li W, Layman CE, et al. TAD evolutionary and functional characterization reveals diversity in mammalian TAD boundary properties and function. Nat Commun. 7 déc 2023;14(1):8111.

41. McArthur E, Capra JA. Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. Am J Hum Genet. 4 févr 2021;108(2):269-83.

42. Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. Science Advances. 10 avr 2019;5(4):eaaw1668.

43. Ye W, Song Y, Huang Z, Osterwalder M, Ljubojevic A, Xu J, et al. A unique stylopod patterning mechanism by Shox2-controlled osteogenesis. Development. 15 juill 2016;143(14):2548-60.

44. Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, et al. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature. févr 2018;554(7691):239-43.

45. Oliveira CS, Alves C. The role of the *SHOX* gene in the pathophysiology of Turner syndrome. Endocrinología y Nutrición. 1 oct 2011;58(8):433-42.

46. Gravholt CH, Viuff MH, Brun S, Stochholm K, Andersen NH. Turner syndrome: mechanisms and management. Nat Rev Endocrinol. oct 2019;15(10):601-14.

47. Vodopiutz J, Steurer LM, Haufler F, Laccone F, Garczarczyk-Asim D, Hilkenmeier M, et al. Leri–Weill Dyschondrosteosis Caused by a Leaky Homozygous SHOX Splice-Site Variant. Genes. avr 2023;14(4):877.

48. Barroso E, Benito-Sanz S, Belinchón A, Yuste-Checa P, Gracia R, Aragones Á, et al. Identification of the first *de novo* PAR1 deletion downstream of *SHOX in* an individual diagnosed with Léri–Weill dyschondrosteosis (LWD). European Journal of Medical Genetics. 1 juill 2010;53(4):204-7.

49. Thomas NS, Maloney V, Bass P, Mulik V, Wellesley D, Castle B. SHOX mutations in a family and a fetus with Langer mesomelic dwarfism. American Journal of Medical Genetics Part A. 2004;128A(2):179-84.

50. Gürsoy S, Hazan F, Aykut A, Nalbantoğlu Ö, Korkmaz HA, Demir K, et al. Detection of SHOX Gene Variations in Patients with Skeletal Abnormalities with or without Short Stature. J Clin Res Pediatr Endocrinol. 25 nov 2020;12(4):358-65.

51. Cao C, Li L, Zhang Q, Li H, Wang Z, Wang A, et al. Nkx2.5: a crucial regulator of cardiac development, regeneration and diseases. Frontiers in Cardiovascular Medicine [Internet]. 2023 [cité 19 févr 2024];10. Disponible sur: https://www.frontiersin.org/articles/10.3389/fcvm.2023.1270951

52. Hoffmann S, Clauss S, Berger IM, Weiß B, Montalbano A, Röth R, et al. Coding and non-coding variants in the SHOX2 gene in patients with early-onset atrial fibrillation. Basic Res Cardiol. 30 avr 2016;111(3):36.

ChatGPT was used to help rephrase some sentences.